Fundamental limits of inference: A statistical physics approach.

Léo Miolane

PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l' École Normale Supérieure

Fundamental limits of inference: A statistical physics approach.

Limites fondamentales de l'inférence statistique: Une approche par la physique statistique.

Soutenue par

Léo Miolane

Le 6 juin 2019

École doctorale n°386

Sciences Mathématiques de Paris Centre

Spécialité

Mathématiques

Composition du jury :

Gérard Ben Arous New York University	*Président*
Nike Sun Massachusetts Institute of Technology	*Rapporteur*
Dmitry Panchenko University of Toronto	*Rapporteur*
Élisabeth Gassiat Université Paris-Sud	*Examinateur*
Stéphane Boucheron Université Paris-Diderot	*Examinateur*
Marc Lelarge École Normale Supérieure & Inria	*Directeur de thèse*

ENS | PSL

Inria
inventeurs du monde numérique

Contents

Remerciements

Mes premiers remerciements sont bien évidemment pour Marc pour tout ce qu'il a fait pour moi au cours de cette thèse. Toujours de bonne humeur, disponible, encourageant, il m'a très bien guidé et a su rendre ces trois années très enrichissantes et agréables. Je pense que j'aurais difficilement pu espérer avoir un encadrant aussi sympa, on a bien rigolé.

Ce fut également un bonheur de travailler avec Lenka Zdeborová et Florent Krzakala que je remercie profondément pour m'avoir fait voyager, mais surtout pour toutes ces discussions qui m'ont appris à parler physicien. Ces trois années n'auraient pas été aussi intéressantes si je n'avais pas croisé leur route.

Je m'estime aussi extrêmement chanceux d'avoir croisé celle d'Andrea Montanari, qui m'a ensuite gentiment invité à Stanford. J'ai énormément apprécié travailler avec quelqu'un d'aussi motivant et passionné. Je lui suis très reconnaissant.

Je voudrais également remercier Nicolas Macris et Jean Barbier pour m'avoir accueilli à Lausanne, point de départ de l'aventure "stoInt_GAMP" qui fut riche en émotions. Je suis aussi très heureux d'avoir pu collaborer avec Francesco Caltagirone, Thibault Lesieur, Aukosh Jagannath et Patrick Lopatto.

Je remercie l'Inria et toute l'équipe Dyogene pour sa bonne ambiance. Je tiens aussi à remercier le groupe du "Golosino seminar" pour les bons moments passés autour de délicieuses pizzas.

Je remercie chaleureusement Nike Sun et Dmitry Panchenko pour avoir accepté d'être les rapporteurs de cette thèse. Un grand merci également à Gérard Ben Arous, Stéphane Boucheron et Élisabeth Gassiat pour en constituer le jury.

Je tiens à remercier mes professeurs d'Orsay pour m'avoir donné envie de faire une thèse, avec une pensée particulière pour Christophe Giraud dont les bons conseils m'avaient amené vers Marc.

Pour tout le reste, je voudrais remercier mes amis, mes parents, ma sœur et mon frère.

Foreword

This thesis is about the fundamental limits of statistical inference. Suppose that we are given data (in the form of a graph, a matrix, a list of measurements...) that contains some underlying information/structure corrupted by noise. Our main question is "how well is it possible to recover this information", by any means? In other words, given some data what is the best possible result I can hope for?

We shall consider very basic – and therefore fundamental – estimations tasks: finding "communities" in a large graph, supervised and unsupervised clustering in high dimensions, recovering hidden structures in matrices and tensors. We will not work on real dataset but rather on random data: we will always consider a probabilistic model. If this scenario is less realistic, it provides a coherent mathematical framework and allows to derive precise expressions that could lead to practical insights.

We will look at these problems in large dimension, when we observe a large amount of data and when the signal we aim at estimating has a lot of parameters. This setting is particularly relevant for modern applications (as "real" datasets always get larger) and contains a lot of interesting theoretical challenges. While estimating a single parameter is well understood within the classical statistical theory, estimating a number of parameters that goes to infinity with the number of observations unveils a number of surprising and new phenomena.

Indeed, motivated by deep insights from statistical physics, it has been conjectured that in high dimensions many statistical problems (such as community detection on graphs, Principal Component Analysis (PCA), sparse PCA, Gaussian mixture clustering, synchronization on groups, low-rank tensor estimation, linear and generalized linear models...) may encounter phase transitions. More precisely, there exists a critical value of the noise intensity above which it is impossible to recover (even partially) the signal. This means there exists fundamental limits on the noise level in order to make non-trivial estimation possible: this threshold is known as the *information-theoretic threshold*. However, even in the regime where estimation is theoretically possible, there are many cases where no efficient algorithm is known to recover the signal. The noise level has to be below a second critical value called the *algorithmic threshold* in order for polynomial-time methods to work. In the case where the algorithmic and the I.T. thresholds do not coincide, we say that there exists a computational gap, meaning that there exists a regime where non-trivial estimation is theoretically possible (using exponential-time algorithms) but where no polynomial-time algorithm will perform better than a random guess.

This thesis aims to precisely characterize the information-theoretic threshold for a number of classical statistical problems. Comparing this threshold to known algorithms allows

us to understand when computational gaps appears. The manuscript is organised as follows.

- Chapter 1 introduces the basics concepts of Bayes-optimal inference that we will used repeatedly throughout this manuscript. I am grateful to Andrea Montanari for interesting discussions regarding Section 1.2.

- Chapter 2 presents a useful "decoupling principle" that will simplify the study of the more complex models investigated in this thesis.

- Chapter 3 and Chapter 4 focus on low-rank matrix estimation and establish the information-theoretic limits for this problems. They are based on the paper [132] in collaboration with Marc Lelarge and the work [145].

- Chapter 5 studies the community detection problem in the Stochastic Block Model, when there is two communities of unequal sizes. It is excerpt from the joint work [40] together with Francesco Caltagirone and Marc Lelarge.

- Chapter 6 pursue the directions of Chapters 3-4. The first part studies the statistical limits of low-rank tensor estimation is based on the paper [136] with Thibault Lesieur, Marc Lelarge, Florent Krzakala and Lenka Zdeborová. The second part analyzes maximum likelihood estimations, following the paper [110] with Aukosh Jagannath and Patrick Lopatto.

- Chapter 7 focuses on the statistical and computational limits of estimation in Generalized Linear Models. It is based on the work [19] with Jean Barbier, Florent Krzakala, Nicolas Macris and Lenka Zdeborová.

- Chapters 8 and 9 concern the Lasso estimator. They establish uniform control of the distribution of the Lasso and study methods to tune the penalization parameter. This work is a collaboration with Andrea Montanari [146].

All the inference models that we consider in this manuscript have an equivalent within the statistical physics literature. The only difference is that the problems we study here contain a *planted solution* (representing some signal) whereas their equivalent in physics can be seen as "pure noise" models. More precisely, the Spiked Wigner model from Chapter 3 corresponds to the Sherrington-Kirkpatrick model [191] while the Spiked Wishart model from Chapter 4 is the analog to the bipartite SK model [22] or the Hopfield model [105]. The spiked tensor model of Chapter 6 is linked to the p-spin model and the Generalized Linear Models from Chapter 7 can be seen as perceptrons [88, 87, 89] with various threshold functions. Finally, the Lasso estimator studied in Chapter 8 and 9 is linked to the "Shcherbina and Tirozzi" model [190].

Some notations

$\mathbb{R}_{\geq 0}$	non-negative numbers: $[0, +\infty)$
$\mathbb{R}_{>0}$	positive numbers: $(0, +\infty)$
$\mathbb{E}$	expectation with respect to all random variables
$\mathbb{E}_X$	expectation with respect to the random variable X only
W_2	Wasserstein distance of order 2
D_{TV}	total variation distance
D_{KL}	Kullback-Leibler divergence
$\xrightarrow{(\mathrm{d})}$	convergence in distribution
$\mathbb{S}^{n-1}$	unit sphere of $\mathbb{R}^n$
$(x; y)$ or $\langle x, y \rangle$	dot product between x and y
$x \cdot y$	normalized dot product: $x \cdot y = \frac{1}{n} \sum_{i=1}^{n} x_i y_i$
$\|x\|_0$	ℓ_0 norm of x: $\|x\|_0 = \#\{i \| x_i \neq 0\}$
$\|x\|$	ℓ_1 norm of x: $\|x\| = \sum \|x_i\|$
$\|x\|$	ℓ_2 norm of x: $\|x\| = \sqrt{\sum x_i^2}$
$\phi(x)$	standard Gaussian density function $\phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$
$\Phi(x)$	standard Gaussian distribution function $\Phi(x) = \int_{-\infty}^{x} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$
$\mathrm{MMSE}(X\|Y)$	Minimal Mean Square Error: $\mathrm{MMSE}(X\|Y) = \mathbb{E}\|X - \mathbb{E}[X\|Y]\|^2$

Chapter 1

Bayes-optimal inference

We introduce in this chapter some general properties of Bayes-optimal inference, that will be used repeatedly in the sequel. Let us first define what we mean by *Bayes-optimal inference*.

We consider a statistical problem where we would like to recover a signal vector $\boldsymbol{X} \in \mathbb{R}^n$ from some observations $\boldsymbol{Y} \in \mathbb{R}^m$. We assume that $(\boldsymbol{X}, \boldsymbol{Y})$ is drawn from some probability distribution μ over $\mathbb{R}^n \times \mathbb{R}^m$. Given a performance criterion, a Bayes-optimal estimator (or simply Bayes estimator) is an estimator of $\boldsymbol{X}$ given $\boldsymbol{Y}$ that achieves the best performance for this criterion. For instance if we measure the performance of an estimator $\widehat{\boldsymbol{x}}$ by its mean square error $\mathrm{MSE}(\widehat{\boldsymbol{x}}) = \mathbb{E}\|\boldsymbol{X} - \widehat{\boldsymbol{x}}(\boldsymbol{Y})\|^2$, then the Bayes-optimal estimator is simply the posterior mean $\widehat{\boldsymbol{x}}^{\mathrm{Bayes}}(\boldsymbol{Y}) = \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}]$.

The goal of this chapter is to present some general properties of Bayes-optimal estimators. In Section 1.1 we introduce what we will call (according to the statistical physics terminology) the "Nishimori identity" which is nothing more than a rewriting of Bayes rule. In Section 1.2 we will study the links between various natural performance metrics for estimators and show that they are in some sense equivalent. In Sections 1.3 we analyse the special case where $\boldsymbol{Y} = \sqrt{\lambda}\boldsymbol{X} + \boldsymbol{Z}$, where $\lambda \geq 0$ and $\boldsymbol{Z}$ is some Gaussian noise. This is the starting point of the study of the "spiked" matrix and tensor models. Finally we consider in Section 1.4 a simple example to illustrate the tools of this chapter.

1.1 The Nishimori identity

In order to analyze Bayes-optimal estimators, we will need to examine the posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$. To do so we will often consider i.i.d. samples $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k)}$ from the posterior distribution $P(\cdot \,|\, \boldsymbol{Y})$, independently of everything else. Such samples are called replicas. The (obvious) identity below (which is simply Bayes rule) is named after the works of Nishimori [163, 164] on "gauge-symmetric" spin glasses. It states that the planted solution $\boldsymbol{X}$ behaves like a replica.

Proposition 1.1.1 (*Nishimori identity*)

Let $(\boldsymbol{X}, \boldsymbol{Y})$ be a couple of random variables on a polish space. Let $k \geq 1$ and let $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k)}$ be k i.i.d. samples (given $\boldsymbol{Y}$) from the distribution $\mathbb{P}(\boldsymbol{X} = \cdot \,|\, \boldsymbol{Y})$, independently of every other random variables. Let us denote $\langle \cdot \rangle$ the expectation with respect to $\mathbb{P}(\boldsymbol{X} = \cdot \,|\, \boldsymbol{Y})$ and $\mathbb{E}$ the expectation with respect to $(\boldsymbol{X}, \boldsymbol{Y})$. Then, for all continuous bounded function f

$$\mathbb{E}\Big\langle f(\boldsymbol{Y}, \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k)}) \Big\rangle = \mathbb{E}\Big\langle f(\boldsymbol{Y}, \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k-1)}, \boldsymbol{X}) \Big\rangle .$$

Proof. It is equivalent to sample the couple $(\boldsymbol{X}, \boldsymbol{Y})$ according to its joint distribution or to sample first $\boldsymbol{Y}$ according to its marginal distribution and then to sample $\boldsymbol{X}$ conditionally to $\boldsymbol{Y}$ from its conditional distribution $\mathbb{P}(\boldsymbol{X} = \cdot \,|\, \boldsymbol{Y})$. Thus the $(k+1)$-tuple $(\boldsymbol{Y}, \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k)})$ is equal in law to $(\boldsymbol{Y}, \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k-1)}, \boldsymbol{X})$. □

1.2 Performance measure and optimal estimators

We consider two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ that live respectively in $\mathbb{R}^n$ and $\mathbb{R}^m$. We assume (for simplicity) that $\|\boldsymbol{X}\| = 1$ almost surely. As explained above, given the observations $\boldsymbol{Y}$, our goal is to estimate $\boldsymbol{X}$ with an estimator $\widehat{\boldsymbol{x}}(\boldsymbol{Y})$. In order to evaluate the performance of such an estimator, what criterion should we take?

The probably most natural way to characterize the performance of $\widehat{\boldsymbol{x}}$ is by its mean-squared error:

$$\mathrm{MSE}(\widehat{\boldsymbol{x}}) = \mathbb{E}\|\boldsymbol{X} - \widehat{\boldsymbol{x}}(\boldsymbol{Y})\|^2.$$

By Pythagorean theorem, we know that the optimal estimator which respect to this metric is the posterior mean $\widehat{\boldsymbol{x}}(\boldsymbol{Y}) = \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}]$ which achieves the minimal mean square error:

$$\mathrm{MMSE}(\boldsymbol{X}|\boldsymbol{Y}) \stackrel{\text{def}}{=} \mathbb{E}\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}]\|^2. \tag{1.2.1}$$

However, the MSE is not always an appropriate criterion. Indeed in many cases it is only possible to recover $\boldsymbol{X}$ up to its sign: think for instance of the Spiked Wigner Model $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{X}^\intercal + \text{Noise}$ with $\boldsymbol{X} \sim \mathrm{Unif}(\mathbb{S}^{n-1})$. In such case, $\mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}] = 0$: the best estimator in term of MSE does not even depend on the observations $\boldsymbol{Y}$!

For this kind of problems one should rather consider the correlation (also known as cosine similarity) *in absolute value* between $\widehat{\boldsymbol{x}}$ and the signal $\boldsymbol{X}$:

$$\sup_{\|\widehat{\boldsymbol{x}}\|=1} \mathbb{E}\Big[|(\widehat{\boldsymbol{x}}(\boldsymbol{Y}); \boldsymbol{X})|\Big] \qquad \text{or} \qquad \sup_{\|\widehat{\boldsymbol{x}}\|=1} \mathbb{E}\Big[(\widehat{\boldsymbol{x}}(\boldsymbol{Y}); \boldsymbol{X})^2\Big], \tag{1.2.2}$$

where $(\cdot;\cdot)$ denotes the Euclidean inner product and where the suprema are taken over all estimators $\widehat{\boldsymbol{x}} : \mathbb{R}^m \to \mathbb{S}^{n-1}$.

Let us introduce some notations. We will use the Gibbs Bracket $\langle \cdot \rangle$ to write expectations with respect to the posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$:

$$\langle f(\boldsymbol{x}) \rangle = \mathbb{E}[f(\boldsymbol{X})|\boldsymbol{Y}],$$

for all measurable function f such that $f(\boldsymbol{X})$ is integrable. In particular, we will be interested by the $n \times n$ positive semi-definite (random) matrix:

$$\boldsymbol{M} \stackrel{\text{def}}{=} \langle \boldsymbol{x}\boldsymbol{x}^\mathsf{T} \rangle = \mathbb{E}[\boldsymbol{X}\boldsymbol{X}^\mathsf{T}|\boldsymbol{Y}]. \tag{1.2.3}$$

$\boldsymbol{M}$ is the Bayes-optimal estimator (in terms of mean square error) for estimating the matrix $\boldsymbol{X}\boldsymbol{X}^\mathsf{T}$:

$$\text{MMSE}(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}|\boldsymbol{Y}) = \mathbb{E}\|\boldsymbol{X}\boldsymbol{X}^\mathsf{T} - \mathbb{E}[\boldsymbol{X}\boldsymbol{X}^\mathsf{T}|\boldsymbol{Y}]\|^2. \tag{1.2.4}$$

An easy computation gives $\text{MMSE}(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}|\boldsymbol{Y}) = 1 - \mathbb{E}[\text{Tr}(\boldsymbol{M}^2)]$. The matrix $\boldsymbol{M}$ is also related to the second quantity of (1.2.2) through its largest eigenvalue $\lambda_{\max}(\boldsymbol{M})$:

Lemma 1.2.1

$$\sup_{\|\widehat{\boldsymbol{x}}\|=1} \mathbb{E}\Big[(\widehat{\boldsymbol{x}}(\boldsymbol{Y}); \boldsymbol{X})^2\Big] = \mathbb{E}\Big[\lambda_{\max}(\boldsymbol{M})\Big]$$

and the optimal estimator for this metric is a unit eigenvector of $\boldsymbol{M}$ associated to its largest eigenvalue $\lambda_{\max}(\boldsymbol{M})$.

Proof. Let $\widehat{\boldsymbol{x}}$ be an estimator of $\boldsymbol{X}$. By the Nishimori identity (Proposition 1.1.1), we have

$$\mathbb{E}[(\widehat{\boldsymbol{x}}(\boldsymbol{Y}); \boldsymbol{X})^2] = \mathbb{E}[\widehat{\boldsymbol{x}}(\boldsymbol{Y})^\mathsf{T}\boldsymbol{X}\boldsymbol{X}^\mathsf{T}\widehat{\boldsymbol{x}}(\boldsymbol{Y})] = \mathbb{E}[\langle \widehat{\boldsymbol{x}}(\boldsymbol{Y})^\mathsf{T}\boldsymbol{x}\boldsymbol{x}^\mathsf{T}\widehat{\boldsymbol{x}}(\boldsymbol{Y})\rangle] = \mathbb{E}[\widehat{\boldsymbol{x}}(\boldsymbol{Y})^\mathsf{T}\boldsymbol{M}\widehat{\boldsymbol{x}}(\boldsymbol{Y})],$$

the lemma follows. □

Lemma 1.2.1 tells us that the top unit eigenvector $\widehat{\boldsymbol{v}}$ of $\boldsymbol{M}$ maximizes $\mathbb{E}\,[(\widehat{\boldsymbol{x}}(\boldsymbol{Y}); \boldsymbol{X})^2]$. In the following, we will show that under a simple condition (that will hold for the models we consider in this manuscript), the estimator $\widehat{\boldsymbol{v}}$ is "asymptotically optimal" (in the limit of large dimension) for the two metrics (1.2.2) and $\lambda_{\max}\widehat{\boldsymbol{v}}\widehat{\boldsymbol{v}}^\mathsf{T}$ is optimal for the estimation of $\boldsymbol{X}\boldsymbol{X}^\mathsf{T}$ in terms of mean square error.

To introduce this condition and the asymptotic limit, we need to consider to a sequence of inference problems. We assume that for all $n \geq 1$ we have two random vectors $\boldsymbol{X}_{[n]}$ and $\boldsymbol{Y}_{[n]}$ respectively in $\mathbb{S}^{n-1}$ and $\mathbb{R}^{m_n}$, for some sequence $(m_n)_{n\geq 1}$. Our goal is again to estimate $\boldsymbol{X}_{[n]}$ from the observation of $\boldsymbol{Y}_{[n]}$ when n is very large: we would like for instance to compute the limits of (1.2.2) and (1.2.4) as $n \to \infty$. Moreover, we would like to know which estimators are "asymptotically optimal", i.e. whose performance reach in the $n \to \infty$ limit the optimal one. In the following, in order to simplify the notations, we will write $\boldsymbol{X}$ and $\boldsymbol{Y}$ instead of $\boldsymbol{X}_{[n]}$ and $\boldsymbol{Y}_{[n]}$.

Proposition 1.2.1

Let us denote by G_n the posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$. Notice that G_n is a random probability distributions on $\mathbb{S}^{n-1}$. Assume that there exists $q \in [0,1]$ such that for $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \overset{\text{i.i.d.}}{\sim} G_n$ we have

$$\left|\left(\boldsymbol{x}^{(1)}; \boldsymbol{x}^{(2)}\right)\right| \xrightarrow[n\to\infty]{\text{(d)}} q. \tag{1.2.5}$$

Then $\text{Tr}(\boldsymbol{M}^2) \xrightarrow[n\to\infty]{\text{(d)}} q^2$ and

$$\lambda_{\max}(\boldsymbol{M}) \xrightarrow[n\to\infty]{\text{(d)}} q.$$

Proof. Let us compute $\mathrm{Tr}(\boldsymbol{M}^2) = \mathrm{Tr}(\langle \boldsymbol{x}\boldsymbol{x}^\mathsf{T}\rangle\langle \boldsymbol{x}\boldsymbol{x}^\mathsf{T}\rangle) = \langle (\boldsymbol{x}^{(1)};\boldsymbol{x}^{(2)})^2\rangle \xrightarrow[n\to\infty]{} q^2$, by assumption. If $q = 0$, then the result is obvious since $\lambda_{\max}(\boldsymbol{M})^2 \le \mathrm{Tr}(\boldsymbol{M}^2)$.

Notice that $\mathrm{Tr}(\boldsymbol{M}^3) \le \lambda_{\max}(\boldsymbol{M})\mathrm{Tr}(\boldsymbol{M}^2)$, so it suffices to show that

$$\mathrm{Tr}(\boldsymbol{M}^3) = \langle (\boldsymbol{x}^{(1)};\boldsymbol{x}^{(2)})(\boldsymbol{x}^{(2)};\boldsymbol{x}^{(3)})(\boldsymbol{x}^{(3)};\boldsymbol{x}^{(1)})\rangle \xrightarrow[n\to\infty]{} q^3,$$

for $\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\boldsymbol{x}^{(3)} \overset{\text{i.i.d.}}{\sim} G_n$. This follows from Lemma 1.2.2 below. □

Lemma 1.2.2

Under the assumptions of Proposition 1.2.1, we have for $\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\boldsymbol{x}^{(3)} \overset{\text{i.i.d.}}{\sim} G_n$,

$$(\boldsymbol{x}^{(1)};\boldsymbol{x}^{(2)})(\boldsymbol{x}^{(2)};\boldsymbol{x}^{(3)})(\boldsymbol{x}^{(3)};\boldsymbol{x}^{(1)}) \xrightarrow[n\to\infty]{\text{(d)}} q^3.$$

Lemma 1.2.2 will be proved in Appendix A. From Proposition 1.2.1 we deduce the main result of this section:

Proposition 1.2.2

Let $\widehat{\boldsymbol{v}}$ *be a leading unit eigenvector of* $\boldsymbol{M}$ *(which is defined by (1.2.3)). Under the assumptions of Proposition 1.2.1, we have*

$$|(\widehat{\boldsymbol{v}};\boldsymbol{X})| \xrightarrow[n\to\infty]{\text{(d)}} \sqrt{q}. \tag{1.2.6}$$

Further $\lim_{n\to\infty} \mathrm{MMSE}(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}|\boldsymbol{Y}) = 1 - q^2$,

$$\lim_{n\to\infty} \sup_{\|\widehat{\boldsymbol{x}}\|=1} \mathbb{E}\Big[|(\widehat{\boldsymbol{x}}(\boldsymbol{Y});\boldsymbol{X})|\Big] = \sqrt{q}, \qquad \textit{and} \qquad \lim_{n\to\infty} \sup_{\|\widehat{\boldsymbol{x}}\|=1} \mathbb{E}\Big[(\widehat{\boldsymbol{x}}(\boldsymbol{Y});\boldsymbol{X})^2\Big] = q. \tag{1.2.7}$$

Proof. Let us abbreviate $\lambda_{\max} \overset{\text{def}}{=} \lambda_{\max}(\boldsymbol{M})$. By Lemma 1.2.1 and Proposition 1.2.1 we have

$$\mathbb{E}[(\widehat{\boldsymbol{v}};\boldsymbol{X})^2] = \mathbb{E}[\lambda_{\max}] \xrightarrow[n\to\infty]{} q. \tag{1.2.8}$$

Hence if $q = 0$, the Proposition follows easily. Assume now that $q > 0$. Using Pythagorean Theorem and the Nishimori identity (Proposition 1.1.1) we get

$$\begin{aligned}
\mathbb{E}\|\boldsymbol{X}^{\otimes 4} - \lambda_{\max}^2 \widehat{\boldsymbol{v}}^{\otimes 4}\|^2 &\ge \mathbb{E}\|\boldsymbol{X}^{\otimes 4} - \langle \boldsymbol{x}^{\otimes 4}\rangle\|^2 \\
&= 1 + \mathbb{E}[(\langle \boldsymbol{x}^{\otimes 4}\rangle;\langle \boldsymbol{x}^{\otimes 4}\rangle)] - 2\mathbb{E}[(\boldsymbol{X}^{\otimes 4};\langle \boldsymbol{x}^{\otimes 4}\rangle)] \\
&= 1 - \mathbb{E}\left[\left\langle (\boldsymbol{x}^{(1)};\boldsymbol{x}^{(2)})^4 \right\rangle\right] \xrightarrow[n\to\infty]{} 1 - q^4,
\end{aligned}$$

where the last limit follows from the assumption (1.2.5). Since

$$\mathbb{E}\|\boldsymbol{X}^{\otimes 4} - \lambda_{\max}^2 \widehat{\boldsymbol{v}}^{\otimes 4}\|^2 = 1 + \mathbb{E}[\lambda_{\max}^4] - 2\mathbb{E}[\lambda_{\max}^2 (\boldsymbol{X};\widehat{\boldsymbol{v}})^4],$$

using Proposition 1.2.1, we deduce (recall that we assumed $q > 0$) that $\limsup_{n\to\infty} \mathbb{E}[(\widehat{\boldsymbol{v}};\boldsymbol{X})^4] \le q^2$. Together with (1.2.8) this gives that $|(\widehat{\boldsymbol{v}};\boldsymbol{X})| \xrightarrow[n\to\infty]{} \sqrt{q}$.

The next point is a consequence of Proposition 1.2.1 because $\mathrm{MMSE}(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}|\boldsymbol{Y}) = 1 - \mathbb{E}[\mathrm{Tr}(\boldsymbol{M}^2)]$. To prove (1.2.7) simply notice that

$$\mathbb{E}[|(\widehat{\boldsymbol{v}};\boldsymbol{X})|]^2 \le \sup_{\|\widehat{\boldsymbol{x}}\|=1} \mathbb{E}[|(\widehat{\boldsymbol{x}}(\boldsymbol{Y});\boldsymbol{X})|]^2 \le \sup_{\|\widehat{\boldsymbol{x}}\|=1} \mathbb{E}[(\widehat{\boldsymbol{x}}(\boldsymbol{Y});\boldsymbol{X})^2] = \mathbb{E}[\lambda_{\max}],$$

which proves (1.2.7) using (1.2.6) and Proposition 1.2.1. □

From Proposition 1.2.2, we deduce that the estimator $\widehat{\boldsymbol{A}} \stackrel{\text{def}}{=} \lambda_{\max}(\boldsymbol{M})\, \widehat{\boldsymbol{v}}\widehat{\boldsymbol{v}}^{\mathsf{T}}$ achieves asymptotically the minimal mean square error for the estimation of $\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}}$:

$$\lim_{n\to\infty} \mathbb{E}\|\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}} - \widehat{\boldsymbol{A}}\|^2 = 1 - q^2.$$

Remark 1.2.1. *For simplicity we assumed in this section that $\|\boldsymbol{X}\|^2 = 1$ almost surely. However we will need to work in the next chapters under a slightly weaker condition, namely $\|\boldsymbol{X}\|^2 \xrightarrow[n\to\infty]{} 1$. It is not difficult to modify the proofs of this section to see that Lemma 1.2.1, Proposition 1.2.1, Lemma 1.2.2 and Proposition 1.2.2 still hold, provided that $\|\boldsymbol{X}\|^2 \xrightarrow[n\to\infty]{} 1$ for the Wasserstein distance of order 4 (i.e. $\|\boldsymbol{X}\|^2 \xrightarrow[n\to\infty]{} 1$ in distribution and $\mathbb{E}\|\boldsymbol{X}\|^8 \xrightarrow[n\to\infty]{} 1$).*

1.3 Bayesian inference with Gaussian noise

We will now focus on the following model:

$$\boldsymbol{Y} = \sqrt{\lambda}\, \boldsymbol{X} + \boldsymbol{Z}\,, \tag{1.3.1}$$

where the signal $\boldsymbol{X}$ is sampled according to some probability distribution P_X over $\mathbb{R}^n$, and where the noise $\boldsymbol{Z} = (Z_1, \dots, Z_n) \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ is independent from $\boldsymbol{X}$. In Chapters 3 and 4, $\boldsymbol{X}$ will typically be a low-rank matrix. The parameter $\lambda \geq 0$ plays the role of a signal-to-noise ratio. We assume that P_X admits a finite second moment: $\mathbb{E}\|\boldsymbol{X}\|^2 < \infty$.

Given the observation channel (1.3.1), the goal of the statistician is to estimate $\boldsymbol{X}$ given the observations $\boldsymbol{Y}$. Again, we assume to be in the "Bayes-optimal" setting, where the statistician knows all the parameters of the inference model, that is the prior distribution P_X and the signal-to-noise ratio λ. We measure the performance of an estimator $\widehat{\theta}$ (i.e. a measurable function of the observations $\boldsymbol{Y}$) by its Mean Squared Error $\mathrm{MSE}(\widehat{\theta}) = \mathbb{E}\|\boldsymbol{X} - \widehat{\theta}(\boldsymbol{Y})\|^2$. One of our main quantity of interest will be the Minimum Mean Squared Error

$$\mathrm{MMSE}(\lambda) \stackrel{\text{def}}{=} \min_{\widehat{\theta}} \mathrm{MSE}(\widehat{\theta}) = \mathbb{E}\left[\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}]\|^2\right],$$

where the minimum is taken over all measurable function $\widehat{\theta}$ of the observations $\boldsymbol{Y}$. Since the optimal estimator (in term of Mean Squared Error) is the posterior mean of $\boldsymbol{X}$ given $\boldsymbol{Y}$, a natural object to study is the posterior distribution of $\boldsymbol{X}$. By Bayes rule, the posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$ is

$$dP(\boldsymbol{x}\,|\,\boldsymbol{Y}) = \frac{1}{\mathcal{Z}(\lambda, \boldsymbol{Y})} e^{H_{\lambda,\boldsymbol{Y}}(\boldsymbol{x})} dP_X(\boldsymbol{x})\,, \tag{1.3.2}$$

where

$$H_{\lambda,\boldsymbol{Y}}(\boldsymbol{x}) = \sqrt{\lambda}\, \boldsymbol{x}^{\mathsf{T}}\boldsymbol{Y} - \frac{\lambda}{2}\|\boldsymbol{x}\|^2 = \sqrt{\lambda}\, \boldsymbol{x}^{\mathsf{T}}\boldsymbol{Z} + \lambda\, \boldsymbol{x}^{\mathsf{T}}\boldsymbol{X} - \frac{\lambda}{2}\|\boldsymbol{x}\|^2\,.$$

Definition 1.3.1

$H_{\lambda,\boldsymbol{Y}}$ is called the Hamiltonian[1] *and the normalizing constant*

$$\mathcal{Z}(\lambda, \boldsymbol{Y}) = \int dP_X(\boldsymbol{x}) e^{H_{\lambda,\boldsymbol{Y}}(\boldsymbol{x})}$$

is called the partition function.

Expectations with respect the posterior distribution (1.3.2) will be denoted by the Gibbs brackets $\langle \cdot \rangle_\lambda$:

$$\left\langle f(\boldsymbol{x}) \right\rangle_\lambda = \mathbb{E}\left[f(\boldsymbol{X}) | \boldsymbol{Y} \right] = \frac{1}{\mathcal{Z}(\lambda, \boldsymbol{Y})} \int dP_X(\boldsymbol{x}) f(\boldsymbol{x}) e^{H_{\lambda,\boldsymbol{Y}}(\boldsymbol{x})} ,$$

for any measurable function f such that $f(\boldsymbol{X})$ is integrable.

Definition 1.3.2

$F(\lambda) = \mathbb{E} \log \mathcal{Z}(\lambda, \boldsymbol{Y})$ is called the free energy[2]*. It is related to the mutual information between $\boldsymbol{X}$ and $\boldsymbol{Y}$ by*

$$F(\lambda) = \frac{\lambda}{2} \mathbb{E}\|\boldsymbol{X}\|^2 - I(\boldsymbol{X}; \boldsymbol{Y}) . \tag{1.3.3}$$

Proof. The mutual information $I(\boldsymbol{X}; \boldsymbol{Y})$ is defined as the Kullback-Leibler divergence between $P_{(X,Y)}$, the joint distribution of $(\boldsymbol{X}, \boldsymbol{Y})$ and $P_X \otimes P_Y$ the product of the marginal distributions of $\boldsymbol{X}$ and $\boldsymbol{Y}$. $P_{(X,Y)}$ is absolutely continuous with respect to $P_X \otimes P_Y$ with Radon-Nikodym derivative:

$$\frac{dP_{(X,Y)}}{dP_X \otimes P_Y}(\boldsymbol{X}, \boldsymbol{Y}) = \frac{\exp\left(-\frac{1}{2}\|\boldsymbol{Y} - \sqrt{\lambda}\boldsymbol{X}\|^2\right)}{\int \exp\left(-\frac{1}{2}\|\boldsymbol{Y} - \sqrt{\lambda}\boldsymbol{x}\|^2\right) dP_X(\boldsymbol{x})} .$$

Therefore

$$\begin{aligned} I(\boldsymbol{X}; \boldsymbol{Y}) &= \mathbb{E} \log \left(\frac{dP_{(X,Y)}}{dP_X \otimes P_Y}(\boldsymbol{X}, \boldsymbol{Y}) \right) = -\mathbb{E} \log \int dP_X(\boldsymbol{x}) \exp\left(\sqrt{\lambda} \boldsymbol{x}^\intercal \boldsymbol{Y} - \sqrt{\lambda} \boldsymbol{X}^\intercal \boldsymbol{Y} - \frac{\lambda}{2}\|\boldsymbol{x}\|^2 + \frac{\lambda}{2}\|\boldsymbol{X}\|^2 \right) \\ &= -F(\lambda) + \frac{\lambda}{2}\mathbb{E}\|\boldsymbol{X}\|^2 . \end{aligned}$$

□

We state now two basic properties of the MMSE. A more detailed analysis can be found in [97, 216].

Proposition 1.3.1

$\lambda \mapsto \mathrm{MMSE}(\lambda)$ is non-increasing over $\mathbb{R}_{\geq 0}$. Moreover

- $\mathrm{MMSE}(0) = \mathbb{E}\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]\|^2$,
- $\mathrm{MMSE}(\lambda) \xrightarrow[\lambda \to +\infty]{} 0$.

Proposition 1.3.2

$\lambda \mapsto \mathrm{MMSE}(\lambda)$ is continuous over $\mathbb{R}_{\geq 0}$.

[1] According to the physics convention, this should be minus the Hamiltonian, since a physical system tries to minimize its energy. However, we chose here to remove it for simplicity.

[2] This is in fact minus the free energy, but we chose to remove the minus sign for simplicity.

The proofs of Proposition 1.3.1 and 1.3.2 can respectively be found in Appendix B.1 and B.2.

We present now the very useful "I-MMSE" relation from [96]. This relation was previously known (under a different formulation) as "de Brujin identity" see [193, Equation 2.12].

Proposition 1.3.3

For all $\lambda \geq 0$,

$$\frac{\partial}{\partial\lambda} I(\boldsymbol{X};\boldsymbol{Y}) = \frac{1}{2}\mathrm{MMSE}(\lambda) \qquad \text{and} \qquad F'(\lambda) = \frac{1}{2}\mathbb{E}\langle \boldsymbol{x}^\intercal \boldsymbol{X}\rangle_\lambda = \frac{1}{2}\Big(\mathbb{E}\|\boldsymbol{X}\|^2 - \mathrm{MMSE}(\lambda)\Big)\,. \tag{1.3.4}$$

F thus is a convex, differentiable, non-decreasing, and $\frac{1}{2}\mathbb{E}\|\boldsymbol{X}\|^2$-Lipschitz function over $\mathbb{R}_{\geq 0}$. If P_X is not a Dirac mass, then F is strictly convex.

Proposition 1.3.3 is proved in Appendix B.3. Proposition 1.3.3 reduces the computation of the MMSE to the computation of the free energy. This will be particularly useful because the free energy F is much easier to handle than the MMSE.

We end this section with the simplest model of the form (1.3.1), namely the additive Gaussian scalar channel:

$$Y = \sqrt{\lambda}X + Z\,, \tag{1.3.5}$$

where $Z \sim \mathcal{N}(0,1)$ and X is sampled from a distribution P_0 over $\mathbb{R}$, independently of Z. The corresponding free energy and the MMSE are respectively

$$\psi_{P_0}(\lambda) = \mathbb{E}\log\int dP_0(x) e^{\sqrt{\lambda}Yx - \lambda x^2/2} \quad \text{and} \quad \mathrm{MMSE}_{P_0}(\lambda) = \mathbb{E}\Big[\Big(X - \mathbb{E}[X|Y]\Big)^2\Big]\,. \tag{1.3.6}$$

The study of this simple inference channel will be very useful in the following, because we will see that the inference problems that we are going to study enjoy asymptotically a "decoupling principle" that reduces them to scalar channels like (1.3.5).

Let us compute the mutual information and the MMSE for particular choices of prior distributions:

Example 1.3.1 (Gaussian prior: $P_0 = \mathcal{N}(0,1)$)**.** *In that case $\mathbb{E}[X|Y]$ is simply the orthogonal projection of X on Y:*

$$\mathbb{E}[X|Y] = \frac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]}Y = \frac{\sqrt{\lambda}}{1+\lambda}Y.$$

One deduces $\mathrm{MMSE}_{P_0}(\lambda) = \frac{1}{1+\lambda}$. Using (1.3.4), we get $I(X;Y) = \frac{1}{2}\log(1+\lambda)$ and $\psi_{P_0}(\lambda) = \frac{1}{2}\Big(\lambda - \log(1+\lambda)\Big)$.

Remark 1.3.1 (Worst-case prior)**.** *Let P_0 be a probability distribution on $\mathbb{R}$ with unit second moment $\mathbb{E}_{P_0}[X^2] = 1$. By considering the estimator $\widehat{x} = \frac{\sqrt{\lambda}}{1+\lambda}Y$, one obtain $\mathrm{MMSE}_{P_0}(\lambda) \leq \frac{1}{1+\lambda}$. We conclude:*

$$\sup_{P_0} \mathrm{MMSE}_{P_0}(\lambda) = \frac{1}{1+\lambda} \qquad \text{and} \qquad \inf_{P_0} \psi_{P_0}(\lambda) = \frac{1}{2}(\lambda - \log(1+\lambda)),$$

where the supremum and infimum are both over the probability distributions that have unit second moment. The standard normal distribution $P_0 = \mathcal{N}(0,1)$ achieves both extrema.

Example 1.3.2 (Rademacher prior: $P_0 = \frac{1}{2}\delta_{+1}+\frac{1}{2}\delta_{-1}$). *We compute* $\psi_{P_0}(\lambda) = \mathbb{E}\log\cosh(\sqrt{\lambda}Z+\lambda) - \frac{\lambda}{2}$ *and* $I(X;Y) = \lambda - \mathbb{E}\log\cosh(\sqrt{\lambda}Z+\lambda)$. *The I-MMSE relation gives*

$$\begin{aligned}\frac{1}{2}\mathrm{MMSE}(\lambda) = \frac{\partial}{\partial\lambda}I(X;Y) &= 1 - \mathbb{E}\Big[\Big(\frac{1}{2\sqrt{\lambda}}Z+1\Big)\tanh\Big(\sqrt{\lambda}Z+\lambda\Big)\Big] \\ &= 1 - \mathbb{E}\tanh(\sqrt{\lambda}Z+\lambda) - \frac{1}{2}\mathbb{E}\tanh'(\sqrt{\lambda}Z+\lambda) \\ &= \frac{1}{2} - \mathbb{E}\tanh(\sqrt{\lambda}Z+\lambda) + \frac{1}{2}\mathbb{E}\tanh^2(\sqrt{\lambda}Z+\lambda)\end{aligned}$$

where we used Gaussian integration by parts. Since by the Nishimori property $\mathbb{E}\langle xX\rangle_\lambda = \mathbb{E}\langle x\rangle_\lambda^2$, *one has* $\mathbb{E}\tanh(\sqrt{\lambda}Z+\lambda) = \mathbb{E}\tanh^2(\sqrt{\lambda}Z+\lambda)$ *and therefore* $\mathrm{MMSE}(\lambda) = 1 - \mathbb{E}\tanh(\sqrt{\lambda}Z+\lambda)$.

1.4 A warm-up: the "needle in a haystack" problem

In order to illustrate the results seen in the previous sections, we study now a very simple inference model. Let $(e_1,\dots,e_{2^n})$ be the canonical basis of $\mathbb{R}^{2^n}$. Let $\sigma_0 \sim \mathrm{Unif}(\{1,\dots,2^n\})$ and define $\boldsymbol{X} = e_{\sigma_0}$ (i.e. $\boldsymbol{X}$ is chosen uniformly over the canonical basis of $\mathbb{R}^{2^n}$). Suppose here that we observe:

$$\boldsymbol{Y} = \sqrt{\lambda n}\boldsymbol{X} + \boldsymbol{Z}\,,$$

where $\boldsymbol{Z} = (Z_1,\dots,Z_{2^n}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, independently from σ_0. The goal here is to estimate $\boldsymbol{X}$ or equivalently to find σ_0. The posterior distribution reads:

$$\begin{aligned}\mathbb{P}(\sigma_0 = \sigma|\boldsymbol{Y}) = \mathbb{P}(\boldsymbol{X} = e_\sigma|\boldsymbol{Y}) &= \frac{1}{\mathcal{Z}_n(\lambda)}2^{-n}\exp\Big(\sqrt{\lambda n}e_\sigma^\intercal\boldsymbol{Y} - \frac{\lambda n}{2}\|e_\sigma\|^2\Big) \\ &= \frac{1}{\mathcal{Z}_n(\lambda)}2^{-n}\exp\Big(\sqrt{\lambda n}Z_\sigma + \lambda n\mathbb{1}(\sigma=\sigma_0) - \frac{\lambda n}{2}\Big),\end{aligned}$$

where $\mathcal{Z}_n(\lambda)$ is the partition function

$$\mathcal{Z}_n(\lambda) = \frac{1}{2^n}\sum_{\sigma=1}^{2^n}\exp\Big(\sqrt{\lambda n}Z_\sigma + \lambda n\mathbb{1}(\sigma=\sigma_0) - \frac{\lambda n}{2}\Big)\,.$$

We will be interested in computing the free energy $F_n(\lambda) = \frac{1}{n}\mathbb{E}\log\mathcal{Z}_n(\lambda)$ in order to deduce then the minimal mean squared error using the I-MMSE relation (1.3.4) presented in the previous section.

Although its simplicity, this model is interesting for many reasons. First, it is one of the simplest statistical model for which one observes a phase transition. Second it is the "planted" analog of the random energy model (REM) introduced in statistical physics by Derrida [57, 58], for which the free energy reads $\frac{1}{n}\mathbb{E}\log\sum_\sigma\frac{1}{2^n}\exp\Big(\sqrt{\lambda n}Z_\sigma\Big)$. Third, as we will see in Section 6.1.1, this model correspond to the "large order limit" of a rank-one tensor estimation model.

We start by computing the limiting free energy:

Theorem 1.4.1

$$\lim_{n\to\infty} F_n(\lambda) = \begin{cases} 0 & \text{if} \quad \lambda \le 2\log 2\,, \\ \frac{\lambda}{2} - \log(2) & \text{if} \quad \lambda \ge 2\log 2\,. \end{cases}$$

Proof. Using Jensen's inequality

$$\begin{aligned} F_n(\lambda) &\le \frac{1}{n}\mathbb{E}\log\mathbb{E}\left[\mathcal{Z}_n(\lambda)|\sigma_0, Z_{\sigma_0}\right] = \frac{1}{n}\mathbb{E}\log\left(1 - \frac{1}{2^n} + e^{\sqrt{\lambda n}Z_{\sigma_0} + \frac{\lambda n}{2} - \log(2)n}\right) \\ &\le \frac{1}{n}\mathbb{E}\log\left(1 + e^{\frac{\lambda n}{2} - \log(2)n}\right) + \sqrt{\frac{\lambda}{n}} \xrightarrow[n\to\infty]{} \begin{cases} 0 & \text{if} \quad \lambda \le 2\log(2)\,, \\ \frac{\lambda}{2} - \log(2) & \text{if} \quad \lambda \ge 2\log(2)\,. \end{cases} \end{aligned}$$

F_n is non-negative since $F_n(0) = 0$ and F_n is non-decreasing. We have therefore $F_n(\lambda) \xrightarrow[n\to\infty]{} 0$ for all $\lambda \in [0, 2\log(2)]$. We have also, by only considering the term $\sigma = \sigma_0$:

$$F_n(\lambda) \ge \frac{1}{n}\mathbb{E}\log\left(\frac{e^{\sqrt{\lambda n}Z_{\sigma_0} + \frac{\lambda n}{2}}}{2^n}\right) = \frac{\lambda}{2} - \log(2)\,.$$

We obtain therefore that $F_n(\lambda) \xrightarrow[n\to\infty]{} \frac{\lambda}{2} - \log(2)$ for $\lambda \ge 2\log(2)$. □

Using the I-MMSE relation (1.3.4), we deduce the limit of the minimum mean squared error $\mathrm{MMSE}_n(\lambda) = \min_{\widehat{\theta}} \mathbb{E}\|\boldsymbol{X} - \widehat{\theta}(\boldsymbol{Y})\|^2$:

$$\mathrm{MMSE}_n(\lambda) = \mathbb{E}\|\boldsymbol{X}\|^2 - 2F_n'(\lambda) = 1 - 2F_n'(\lambda)\,.$$

F_n is a convex function of λ, thus (see Proposition C.1) its derivative converges to the derivative of its limit at each λ at which the limit is differentiable, i.e. for all $\lambda \in (0, +\infty) \setminus \{2\log(2)\}$. We obtain therefore that for all $\lambda > 0$,

- if $\lambda < 2\log(2)$, then $\mathrm{MMSE}_n(\lambda) \xrightarrow[n\to\infty]{} 1$: one can not recover $\boldsymbol{X}$ better than a random guess.
- if $\lambda > 2\log(2)$, then $\mathrm{MMSE}_n(\lambda) \xrightarrow[n\to\infty]{} 0$: one can recover $\boldsymbol{X}$ perfectly.

Of course, the result we obtain here is (almost) trivial since the maximum likelihood estimator

$$\widehat{\sigma}(\boldsymbol{Y}) = \operatorname*{arg\,max}_{1\le\sigma\le 2^n} Y_\sigma$$

of σ_0 is easy to analyze. Indeed, $\max_\sigma Z_\sigma \simeq \sqrt{2\log(2)n}$ with high probability so that the maximum likelihood estimator recovers perfectly the signal for $\lambda > 2\log(2)$ with high probability.

Chapter 2

A decoupling principle

We present in this section a general "decoupling principle" that will be particularly useful in the study of planted models. We consider here the setting where $\boldsymbol{X} = (X_1, \dots, X_n) \overset{\text{i.i.d.}}{\sim} P_0$ for some probability distribution P_0 over $\mathbb{R}$ with support S. Let $\boldsymbol{Y} \in \mathbb{R}^m$ be another random variable that accounts for noisy observation of $\boldsymbol{X}$. The goal is again to recover the planted vector $\boldsymbol{X}$ from the observations $\boldsymbol{Y}$. We suppose that the distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$ takes the following form

$$\mathbb{P}(\boldsymbol{X} \in A \mid \boldsymbol{Y}) = \frac{1}{\mathcal{Z}_n(\boldsymbol{Y})} \int_{\boldsymbol{x} \in A} dP_0^{\otimes n}(\boldsymbol{x}) e^{H_n(\boldsymbol{x}, \boldsymbol{Y})}, \quad \text{for all Borel set } A \subset \mathbb{R}^n, \tag{2.0.1}$$

where H_n is a measurable function on $\mathbb{R}^n \times \mathbb{R}^m$ that can be equal to $-\infty$ (in which case, we use the convention $\exp(-\infty) = 0$) and $\mathcal{Z}_n(\boldsymbol{Y}) = \int dP_0^{\otimes n}(\boldsymbol{x}) e^{H_n(\boldsymbol{x}, \boldsymbol{Y})}$ is the appropriate normalization. We assume that $\mathbb{E}|\log \mathcal{Z}_n(\boldsymbol{Y})| < \infty$ in order to define the free energy

$$F_n = \frac{1}{n} \mathbb{E} \log \mathcal{Z}_n(\boldsymbol{Y}) = \frac{1}{n} \mathbb{E} \log \left(\int dP_0^{\otimes n}(\boldsymbol{x}) e^{H_n(\boldsymbol{x}, \boldsymbol{Y})} \right).$$

In the following, we are going to drop the dependency in $\boldsymbol{Y}$ of $H_n(\boldsymbol{x}, \boldsymbol{Y})$ and simply write $H_n(\boldsymbol{x})$.

We introduce now an important notation: the overlap between to vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$. This is simply the normalized scalar product:

$$\boldsymbol{u} \cdot \boldsymbol{v} = \frac{1}{n} \sum_{i=1}^{n} u_i v_i \,.$$

One should see $\boldsymbol{x}$ as a system of n spins $(x_1, \dots, x_n)$ interacting through the (random) Hamiltonian H_n. Our inference problem should be understood as the study of this spin glass model. A central quantity of interest in spin glass theory is the overlaps $\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)}$ between two replicas, i.e. the normalized scalar product between two independent samples $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ from (2.0.1). Understanding this quantity is fundamental because it allows to deduce the distance between two typical configurations of the system and thus encodes the "geometry" of the "Gibbs measure" (2.0.1).

In our statistical inference setting we have $\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} \cdot \boldsymbol{X}$ in law, by the Nishimori identity (Proposition 1.1.1). Thus the overlap $\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)}$ corresponds to the correlation between a typical configuration and the planted configuration. Moreover it is linked to the Minimum Mean Squared Error by

$$\text{MMSE} = \frac{1}{n} \mathbb{E} \left[\|\boldsymbol{X} - \langle \boldsymbol{x} \rangle\|^2 \right] = \mathbb{E}_{P_0}[X^2] - \mathbb{E} \langle \boldsymbol{x} \cdot \boldsymbol{X} \rangle \,,$$

where $\langle \cdot \rangle$ denotes the expectation with respect to $\boldsymbol{x}$ which is sampled from the posterior $\mathbb{P}(\boldsymbol{X} = \cdot \,|\, \boldsymbol{Y})$ (defined by Equation 2.0.1), independently of everything else.

In this section we will see a general principle that states that under a small perturbation of the Gibbs distribution (2.0.1), the overlap $\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)}$ between two replicas concentrates around its mean. Such behavior is called "Replica-Symmetric" in statistical physics. It remains to define what "a small perturbation of the Gibbs distribution" is. In spin glass theory, such perturbations are usually obtained by adding small extra terms to the Hamiltonian. In our context of Bayesian inference a small perturbation will correspond to a small amount of side-information given to the statistician. This extra information will lead to a new posterior distribution. In the following, we will consider two different kind of side-information and we show that the overlaps under the induced posterior concentrate around their mean.

2.1 The pinning Lemma

We suppose here that the support S of P_0 is finite. We make this assumption in order to be able to work with the discrete entropy.

In this section, we give extra information to the statistician by revealing a (small) fraction of the coordinates of $\boldsymbol{X}$. Let us fix $\epsilon \in [0,1]$, and suppose that we have access to the additional observations

$$Y_i' = \begin{cases} X_i & \text{if } L_i = 1\,, \\ * & \text{if } L_i = 0\,, \end{cases} \qquad \text{for} \quad 1 \le i \le n,$$

where $L_i \overset{\text{i.i.d.}}{\sim} \mathrm{Ber}(\epsilon)$ and $*$ is a value that does not belong to S. The posterior distribution of $\boldsymbol{X}$ is now

$$\mathbb{P}(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{Y}, \boldsymbol{Y}') = \frac{1}{\mathcal{Z}_{n,\epsilon}} \left(\prod_{i | L_i = 1} \mathbb{1}(x_i = Y_i') \right) \left(\prod_{i | L_i = 0} P_0(x_i) \right) e^{H_n(\boldsymbol{x})}\,, \tag{2.1.1}$$

where $\mathcal{Z}_{n,\epsilon}$ is the appropriate normalization constant. For $\boldsymbol{x} \in S^n$ we will write

$$\bar{\boldsymbol{x}} = (\bar{x}_1, \dots, \bar{x}_n) = (L_1 X_1 + (1 - L_1) x_1, \dots, L_n X_n + (1 - L_n) x_n)\,. \tag{2.1.2}$$

$\bar{\boldsymbol{x}}$ is thus obtained by replacing the coordinates of $\boldsymbol{x}$ that are revealed by $\boldsymbol{Y}'$ by their revealed values. The notation $\bar{\boldsymbol{x}}$ allows us to obtain a convenient expression for the free energy of the perturbed model:

$$F_{n,\epsilon} = \frac{1}{n} \mathbb{E} \log \mathcal{Z}_{n,\epsilon} = \frac{1}{n} \mathbb{E} \Big[\log \sum_{\boldsymbol{x} \in S^n} P_0(\boldsymbol{x}) \exp(H_n(\bar{\boldsymbol{x}})) \Big]\,.$$

Proposition 2.1.1

For all $n \ge 1$ and all $\epsilon \in [0,1]$, we have

$$|F_{n,\epsilon} - F_n| \le H(P_0)\epsilon\,.$$

Proof. Let us compute

$$\begin{aligned} P(\boldsymbol{Y}' \,|\, \boldsymbol{Y}, \boldsymbol{L}) &= \int \mathbb{1}(x_i = Y_i' \text{ for all } i \text{ such that } L_i = 1) dP(\boldsymbol{x} \,|\, \boldsymbol{Y}) \\ &= \frac{1}{\mathcal{Z}_n} \sum_{\boldsymbol{x} \in S^n} \mathbb{1}(x_i = Y_i' \text{ for all } i \text{ such that } L_i = 1) e^{H_n(\boldsymbol{x})} \prod_{i=1}^n P_0(x_i) \\ &= \frac{\mathcal{Z}_{n,\epsilon}}{\mathcal{Z}_n} \prod_{i | L_i = 1} P_0(Y_i') = \frac{\mathcal{Z}_{n,\epsilon}}{\mathcal{Z}_n} P(\boldsymbol{Y}' \,|\, \boldsymbol{L}) \,. \end{aligned}$$

Therefore, $nF_{n,\epsilon} - nF_n = H(\boldsymbol{Y}' \,|\, \boldsymbol{L}) - H(\boldsymbol{Y}' \,|\, \boldsymbol{Y}, \boldsymbol{L})$ and the proposition follows from the fact that $0 \le H(\boldsymbol{Y}'|\boldsymbol{Y}, \boldsymbol{L}) \le H(\boldsymbol{Y}'|\boldsymbol{L}) = n\epsilon H(P_0)$. □

From now we suppose $\epsilon_0 \in (0, 1]$ to be fixed and consider $\epsilon \in [0, \epsilon_0]$. The following lemma comes from [148] and is sometimes known as the "pinning lemma". It shows that the extra information $\boldsymbol{Y}'$ forces the correlations between the spins under the posterior (2.1.1) to vanish.

Lemma 2.1.1 (*Lemma 3.1 from [148]*)

For all $\epsilon_0 \in [0, 1]$, we have

$$\int_0^{\epsilon_0} d\epsilon \left(\frac{1}{n^2} \sum_{1 \le i,j \le n} I(X_i; X_j | \boldsymbol{Y}, \boldsymbol{Y}') \right) \le \frac{2}{n} H(P_0) \,.$$

Let $\langle \cdot \rangle_{n,\epsilon}$ denote the expectation with respect to two independent samples $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}$ from the posterior (2.1.1). Lemma 2.1.1 implies that the overlap between these two replicas concentrates:

Proposition 2.1.2

There exists a constant $C > 0$ that only depends on P_0 such that for all $\epsilon_0 \in [0, 1]$,

$$\int_0^{\epsilon_0} d\epsilon \, \mathbb{E} \left\langle \left(\frac{1}{n} \sum_{i=1}^n x_i^{(1)} x_i^{(2)} - \left\langle \frac{1}{n} \sum_{i=1}^n x_i^{(1)} x_i^{(2)} \right\rangle_{n,\epsilon} \right)^2 \right\rangle_{n,\epsilon} \le C \sqrt{\frac{\epsilon_0}{n}} .$$

Proof.

$$\begin{aligned} \big\langle (\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} - \langle \boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} \rangle_{n,\epsilon})^2 \big\rangle_{n,\epsilon} &= \big\langle (\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)})^2 \big\rangle_{n,\epsilon} - \big\langle \boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} \big\rangle_{n,\epsilon}^2 \\ &= \frac{1}{n^2} \sum_{1 \le i,j \le n} \big\langle x_i^{(1)} x_i^{(2)} x_j^{(1)} x_j^{(2)} \big\rangle_{n,\epsilon} - \big\langle x_i^{(1)} x_i^{(2)} \big\rangle_{n,\epsilon} \big\langle x_j^{(1)} x_j^{(2)} \big\rangle_{n,\epsilon} \\ &= \frac{1}{n^2} \sum_{1 \le i,j \le n} \langle x_i x_j \rangle_{n,\epsilon}^2 - \langle x_i \rangle_{n,\epsilon}^2 \langle x_j \rangle_{n,\epsilon}^2 . \end{aligned}$$

Let now $i, j \in \{1, \dots, n\}$. The support of P_0 is finite and thus included in $[-K, K]$ for some $K > 0$. This gives:

$$\begin{aligned} \langle x_i x_j \rangle_{n,\epsilon}^2 - \langle x_i \rangle_{n,\epsilon}^2 \langle x_j \rangle_{n,\epsilon}^2 &\le 2K^2 |\langle x_i x_j \rangle_{n,\epsilon} - \langle x_i \rangle_{n,\epsilon} \langle x_j \rangle_{n,\epsilon}| \\ &= 2K^2 \Big| \sum_{x_i, x_j} x_i x_j \mathbb{P}(X_i = x_i, X_j = x_j | \boldsymbol{Y}, \boldsymbol{Y}') - x_i x_j \mathbb{P}(X_i = x_i | \boldsymbol{Y}, \boldsymbol{Y}') \mathbb{P}(X_j = x_j | \boldsymbol{Y}, \boldsymbol{Y}') \Big| \\ &\le 4K^2 \mathrm{D}_{\mathrm{TV}} \big(\mathbb{P}(X_i = \cdot, X_j = \cdot | \boldsymbol{Y}, \boldsymbol{Y}'); \mathbb{P}(X_i = \cdot | \boldsymbol{Y}, \boldsymbol{Y}') \otimes \mathbb{P}(X_j = \cdot | \boldsymbol{Y}, \boldsymbol{Y}') \big) \\ &\le 4K^2 \sqrt{\mathrm{D}_{\mathrm{KL}} \big(\mathbb{P}(X_i = \cdot, X_j = \cdot | \boldsymbol{Y}, \boldsymbol{Y}'); \mathbb{P}(X_i = \cdot | \boldsymbol{Y}, \boldsymbol{Y}') \otimes \mathbb{P}(X_j = \cdot | \boldsymbol{Y}, \boldsymbol{Y}') \big)} \end{aligned}$$

by Pinsker's inequality. Since

$$I(X_i;X_j|\boldsymbol{Y},\boldsymbol{Y}') = \mathbb{E}\big[\mathrm{D}_{\mathrm{KL}}\big(\mathbb{P}(X_i=\cdot,X_j=\cdot|\boldsymbol{Y},\boldsymbol{Y}');\mathbb{P}(X_i=\cdot|\boldsymbol{Y},\boldsymbol{Y}')\otimes\mathbb{P}(X_j=\cdot|\boldsymbol{Y},\boldsymbol{Y}')\big)\big],$$

we get using Lemma 2.1.1:

$$\begin{aligned}\int_0^{\epsilon_0} d\epsilon\,\mathbb{E}\left\langle\Big(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}-\langle\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\rangle_{n,\epsilon}\Big)^2\right\rangle_{n,\epsilon} &\le 4K^2\sqrt{\epsilon_0\int_0^{\epsilon_0}d\epsilon\Big(\frac{1}{n^2}\sum_{1\le i,j\le n}I(X_i;X_j|\boldsymbol{Y},\boldsymbol{Y}')\Big)}\\ &\le 4K^2\sqrt{\frac{2\epsilon_0 H(P_0)}{n}}.\end{aligned}$$

□

2.2 Noisy side Gaussian channel

We consider in this section of a different kind of side-information: an observation of the signal $\boldsymbol{X}$ perturbed by some Gaussian noise. It was proved in [125] for CDMA systems that such perturbations forces the overlaps to concentrate around their means. The principle here is in fact more general and holds for any observation system, provided some concentration property of the free energy.

We suppose here that the prior P_0 has a bounded support $S\subset[-K,K]$, for some $K>0$. Let $a>0$ and $(s_n)_n\in(0,1]^{\mathbb{N}}$. Let $(Z_i)_{1\le i\le n}\overset{\text{i.i.d.}}{\sim}\mathcal{N}(0,1)$ independently of everything else. The extra side-information takes now the form

$$Y_i' = a\sqrt{s_n}X_i + Z_i,\quad \text{for } 1\le i\le n. \tag{2.2.1}$$

The posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y},\boldsymbol{Y}'$ is now $P(\boldsymbol{x}\,|\,\boldsymbol{Y},\boldsymbol{Y}') = \frac{1}{\mathcal{Z}_{n,a}^{(\mathrm{pert})}}P_0^{\otimes n}(\boldsymbol{x})\exp\Big(H_{n,a}^{(\mathrm{pert})}(\boldsymbol{x})\Big)$, where $H_{n,a}^{(\mathrm{pert})}(\boldsymbol{x}) = H_n(\boldsymbol{x})+h_{n,a}(\boldsymbol{x})$ and

$$h_{n,a}(\boldsymbol{x}) = \sum_{i=1}^n a\sqrt{s_n}Z_ix_i + a^2s_nx_iX_i - \frac{1}{2}a^2s_nx_i^2\,.$$

$\mathcal{Z}_{n,a}^{(\mathrm{pert})}$ is the appropriate normalization. Let us define

$$\phi: a\mapsto \frac{1}{ns_n}\log\left(\int dP_0^{\otimes n}(\boldsymbol{x})e^{H_{n,a}^{(\mathrm{pert})}(\boldsymbol{x})}\right).$$

We fix now $A\ge 2$. Define also $v_n(s_n) = \sup_{1/2\le a\le A+1}\mathbb{E}|\phi(a)-\mathbb{E}\phi(a)|$. The following result shows that, in the perturbed system (under some conditions on v_n and s_n) the overlap between two replicas concentrates asymptotically around its expected value.

Proposition 2.2.1 (*Overlap concentration*)

Assume that $v_n(s_n)\xrightarrow[n\to\infty]{}0$. Then there exists a constant $C>0$ that only depends on K such that for all $A\ge 2$,

$$\frac{1}{A-1}\int_1^A\mathbb{E}\left\langle\Big(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}-\mathbb{E}\langle\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\rangle_{n,a}\Big)^2\right\rangle_{n,a}da \;\le\; C\Big(\frac{1}{\sqrt{ns_n}}+\sqrt{v_n(s_n)}\Big),$$

where $\langle\cdot\rangle_{n,a}$ denotes the distribution of $\boldsymbol{X}$ given $(\boldsymbol{Y},\boldsymbol{Y}')$. $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ are two independent samples from $\langle\cdot\rangle_{n,a}$, independently of everything else.

Proposition 2.2.1 is the analog of [170, Theorem 3.2] (the Ghirlanda-Guerra identities, see [90]) and is proved analogously is the remaining of the section. Denote for $\boldsymbol{x} \in S^n$

$$U(\boldsymbol{x}) = \frac{1}{ns_n}\frac{\partial}{\partial a} h_{n,a}(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\sqrt{s_n}} Z_i x_i + 2a x_i X_i - a x_i^2 \,.$$

Lemma 2.2.1

Let $\boldsymbol{x}$ be a sample from $\langle\cdot\rangle_{n,a}$, independently of everything else. Under the conditions of Proposition 2.2.1, we have for all $A \geq 2$

$$\frac{1}{A-1}\int_1^A \mathbb{E}\Big\langle \Big| U(\boldsymbol{x}) - \mathbb{E}\langle U(\boldsymbol{x})\rangle_{n,a}\Big|\Big\rangle_{n,a} da \leq C\Big(\frac{1}{\sqrt{ns_n}} + \sqrt{v_n(s_n)}\Big)\,,$$

for some constant $C > 0$ that only depends on K.

Before proving Lemma 2.2.1, let us show how it implies Proposition 2.2.1.

Proof of Proposition 2.2.1. By the bounded support assumption on P_0, the overlap between two replicas is bounded by K^2, thus

$$\Big|\mathbb{E}\Big\langle U(\boldsymbol{x}^{(1)})\, \boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\Big\rangle_{n,a} - \mathbb{E}\Big\langle \boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\Big\rangle_{n,a}\mathbb{E}\Big\langle U(\boldsymbol{x}^{(1)})\Big\rangle_{n,a}\Big| \leq K^2\mathbb{E}\Big\langle \big|U(\boldsymbol{x}) - \mathbb{E}\langle U(\boldsymbol{x})\rangle_{n,a}\big|\Big\rangle_{n,a}\,. \tag{2.2.2}$$

Let us compute the left-hand side of (2.2.2). By Gaussian integration by parts and using the Nishimori identity (Proposition 1.1.1) we get $\mathbb{E}\langle U(\boldsymbol{x}^{(1)})\rangle_{n,a} = 2a\,\mathbb{E}\langle \boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\rangle_{n,a}$. Therefore

$$\mathbb{E}\left\langle \boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\right\rangle_{n,a}\mathbb{E}\left\langle U(\boldsymbol{x}^{(1)})\right\rangle_{n,a} = 2a\left(\mathbb{E}\left\langle \boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\right\rangle_{n,a}\right)^2.$$

Using the same tools, we compute for $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{x}^{(3)}, \boldsymbol{x}^{(4)} \overset{\text{i.i.d.}}{\sim} \langle\cdot\rangle_{n,a}$, independently of everything else:

$$\begin{aligned}
&\mathbb{E}\langle U(\boldsymbol{x}^{(1)})(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})\rangle_{n,a} \\
&= 2a\mathbb{E}\langle(\boldsymbol{x}^{(1)}\cdot\boldsymbol{X})(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})\rangle_{n,a} + \frac{1}{n\sqrt{s_n}}\sum_{i=1}^{n}\mathbb{E}Z_i\langle x_i^{(1)}(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})\rangle_{n,a} - \frac{a}{n}\sum_{i=1}^{n}\mathbb{E}\langle (x_i^{(1)})^2(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})\rangle_{n,a} \\
&= 2a\mathbb{E}\langle(\boldsymbol{x}^{(1)}\cdot\boldsymbol{X})(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})\rangle_{n,a} + a\mathbb{E}\langle(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})^2\rangle_{n,a} - a\mathbb{E}\langle(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(3)} + \boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(4)})(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})\rangle_{n,a} \\
&= 2a\mathbb{E}\langle(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})^2\rangle_{n,a}\,.
\end{aligned}$$

Thus, by (2.2.2) we have for all $a \in [1, A]$

$$\mathbb{E}\Big\langle (\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)} - \mathbb{E}\langle \boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)}\rangle_{n,a})^2\Big\rangle_{n,a} \leq \frac{K^2}{2}\mathbb{E}\left\langle \big|U(\boldsymbol{x}) - \mathbb{E}\,\langle U(\boldsymbol{x})\rangle_{n,a}\big|\right\rangle_{n,a},$$

and we conclude by integrating with respect to a over $[1, A]$ and using Lemma 2.2.1. □

Proof of Lemma 2.2.1. ϕ is twice differentiable on $(0, +\infty)$, and for $a > 0$

$$\phi'(a) = \langle U(\boldsymbol{x})\rangle_{n,a}\,, \tag{2.2.3}$$

$$\phi''(a) = ns_n(\langle U(\boldsymbol{x})^2\rangle_{n,a} - \langle U(\boldsymbol{x})\rangle_{n,a}^2) + \frac{1}{n}\sum_{i=1}^{n}\Big\langle 2x_iX_i - x_i^2\Big\rangle_{n,a}\,. \tag{2.2.4}$$

Thus $\big\langle (U(\boldsymbol{x}) - \langle U(\boldsymbol{x})\rangle_{n,a})^2 \big\rangle_{n,a} \le \frac{1}{ns_n}(\phi''(a) + 2K^2)$ and

$$\int_1^A \mathbb{E}\big\langle (U(\boldsymbol{x}) - \langle U(\boldsymbol{x})\rangle_{n,a})^2 \big\rangle_{n,a} da \le \frac{1}{ns_n}\Big(\mathbb{E}\phi'(A) - \mathbb{E}\phi'(1) + 2K^2(A-1)\Big) \le \frac{CA}{ns_n}\,,$$

for some constant $C > 0$ (that only depend on K), because $\mathbb{E}\phi'(a) = 2a\mathbb{E}\langle \boldsymbol{x}\cdot\boldsymbol{X}\rangle_{n,a}$. It remains to show that $\int_1^A \mathbb{E}|\langle U(\boldsymbol{x})\rangle_{n,a} - \mathbb{E}\langle U(\boldsymbol{x})\rangle_{n,a}|da \le CA\sqrt{v_n(s_n)}$ for some constant $C > 0$ that only depends on K.

We will use the following lemma on convex functions (from [170], Lemma 3.2).

Lemma 2.2.2

If f and g are two differentiable convex functions then, for any $b > 0$

$$|f'(a) - g'(a)| \le g'(a+b) - g'(a-b) + \frac{d}{b}\,,$$

where $d = |f(a+b) - g(a+b)| + |f(a-b) - g(a-b)| + |f(a) - g(a)|$.

We apply this lemma to $\lambda \mapsto \phi(\lambda) + \frac{3}{2}K^2\lambda^2$ and $\lambda \mapsto \mathbb{E}\phi(\lambda) + \frac{3}{2}K^2\lambda^2$ that are convex because of (2.2.4) and the bounded support assumption on P_0. Therefore, for all $a \ge 1$ and $b \in (0, 1/2)$ we have

$$\mathbb{E}|\phi'(a) - \mathbb{E}\phi'(a)| \le \mathbb{E}\phi'(a+b) - \mathbb{E}\phi'(a-b) + 6K^2 b + \frac{3v_n(s_n)}{b}\,. \tag{2.2.5}$$

Notice that for all $a > 0$, $|\mathbb{E}\phi'(a)| = |2a\mathbb{E}\langle \boldsymbol{x}\cdot\boldsymbol{X}\rangle_{n,a}| \le 2aK^2$. Therefore, by the mean value theorem

$$\begin{aligned}\int_1^A (\mathbb{E}\phi'(a+b) - \mathbb{E}\phi'(a-b))da &= (\mathbb{E}\phi(b+A) - \mathbb{E}\phi(b+1)) - (\mathbb{E}\phi(A-b) - \mathbb{E}\phi(1-b)) \\ &= (\mathbb{E}\phi(b+A) - \mathbb{E}\phi(A-b)) + (\mathbb{E}\phi(1-b) - \mathbb{E}\phi(1+b)) \\ &\le 4K^2 b(A+2)\,.\end{aligned}$$

Combining this with equation (2.2.5), we obtain

$$\forall b \in (0,1/2),\ \int_1^A \mathbb{E}|\phi'(a) - \mathbb{E}\phi'(a)|da \le CA\Big(b + \frac{v_n(s_n)}{b}\Big)\,. \tag{2.2.6}$$

for some constant $C > 0$ depending only on K. The minimum of the right-hand side is achieved for $b = \sqrt{v_n(s_n)} < 1/2$ for n large enough. Then, (2.2.6) gives

$$\int_1^A \mathbb{E}|\langle U(\boldsymbol{x})\rangle_{n,a} - \mathbb{E}\langle U(\boldsymbol{x})\rangle_{n,a}|da = \int_1^A \mathbb{E}|\phi'(a) - \mathbb{E}\phi'(a)|da \le 2CA\sqrt{v_n(s_n)}.$$

□

Chapter 3

Low-rank symmetric matrix estimation

3.1 Introduction to the spiked matrix models

Estimating a low-rank object (matrix or tensor) from a noisy observation is a fundamental problem in statistical inference with applications in machine learning, signal processing or information theory. We focus in this chapter (and in Chapter 4) on the so-called "spiked" models where we observe a signal spike perturbed with some additive noise. We should consider here two popular models.

The first one is often denoted as the *spiked Wigner* model. One observes

$$\boldsymbol{Y} = \sqrt{\frac{\lambda}{n}}\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}} + \boldsymbol{Z} \tag{3.1.1}$$

where the "spike" $\boldsymbol{X} = (X_1, \dots, X_n) \overset{\text{i.i.d.}}{\sim} P_0$ is the signal vector and $\boldsymbol{Z}$ is symmetric matrix that account for noise with standard Gaussian entries: $(Z_{i,j})_{i\leq j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. $\lambda \geq 0$ is a signal-to-noise ratio.

The second model that we will consider in Chapter 4 is the non-symmetric version of (3.1.1), sometimes called *spiked Wishart*[1] or spiked covariance model:

$$\boldsymbol{Y} = \sqrt{\frac{\lambda}{n}}\boldsymbol{U}\boldsymbol{V}^{\mathsf{T}} + \boldsymbol{Z} \tag{3.1.2}$$

where $\boldsymbol{U} = (U_1, \dots, U_n) \overset{\text{i.i.d.}}{\sim} P_U$, $\boldsymbol{V} = (V_1, \dots, V_m) \overset{\text{i.i.d.}}{\sim} P_V$ are independent. $\boldsymbol{Z}$ is a noise matrix with standard normal entries: $Z_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. $\lambda > 0$ captures again the strength of the signal. We are here interested in the regime where $n, m \to +\infty$, while $m/n \to \alpha > 0$. In both models (3.1.1)-(3.1.2) the goal of the statistician is to estimate the low-rank signals ($\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}}$ or $\boldsymbol{U}\boldsymbol{V}^{\mathsf{T}}$) from the observation of $\boldsymbol{Y}$. This task is often called Principal Component Analysis (PCA) in the literature.

These spiked models have received a lot of attention since their introduction by [114]. From a statistical point of view, there are two main problems linked to the spiked models (3.1.1)-(3.1.2).

[1] This terminology usually refers to the case where $\boldsymbol{V}$ is a standard Gaussian vector. We consider here a slightly more general case by allowing the entries of $\boldsymbol{V}$ to be taken i.i.d. from any probability distribution.

- The *recovery* problem: how can we recover the planted signal $\boldsymbol{X}$ / $\boldsymbol{U}, \boldsymbol{V}$? Is it possible? Can we do it efficiently?

- The *detection* problem: is it possible to distinguish between the pure noise case ($\lambda = 0$) and the case where a spike is present ($\lambda > 0$)? Is there any efficient test to do this?

We will focus here on the recovery problem. We let the reader refer to [30, 166, 65, 16, 175, 3, 79] and the references therein for a detailed analysis of the detection problem.

The spiked models (3.1.1)-(3.1.2) has been extensively studied in random matrix theory. The seminal work of [13] (for the complex spiked Wishart model, and [14] for the real spiked Wishart) established the existence of a phase transition: there exists a critical value of the signal-to-noise ratio λ above which the largest singular value of $\boldsymbol{Y}/\sqrt{n}$ escapes from the Marchenko-Pastur bulk. The same phenomenon holds for the spiked Wigner model, as observed by Edwards and Jones [77] using the heuristic replica method and then rigorously proved [173, 84, 44]. It turns out that for both models the eigenvector (respectively singular vector) corresponding to the largest eigenvalue (respectively singular value) also undergo a phase transition at the same threshold, see [109, 172, 160, 28, 29].

For the spiked Wigner model (3.1.1), the main result of interest for us is the following. For any probability distribution P_0 such that $\mathbb{E}_{P_0}[X^2] = 1$, we have

- if $\lambda \leq 1$, the top eigenvalue of $\boldsymbol{Y}/\sqrt{n}$ converges a.s. to 2 as $n \to \infty$, and the top eigenvector $\widehat{\boldsymbol{x}}$ (with norm $\|\widehat{\boldsymbol{x}}\|^2 = n$) has asymptotically trivial correlation with $\boldsymbol{X}$: $\frac{1}{n}\langle \widehat{\boldsymbol{x}}, \boldsymbol{X}\rangle \to 0$ a.s.

- if $\lambda > 1$, the top eigenvalue of $\boldsymbol{Y}/\sqrt{n}$ converges a.s. to $\sqrt{\lambda} + 1/\sqrt{\lambda} > 2$ and the top eigenvector $\widehat{\boldsymbol{x}}$ (with norm $\|\widehat{\boldsymbol{x}}\|^2 = n$) has asymptotically nontrivial correlation with $\boldsymbol{X}$: $\left(\frac{1}{n}\langle \widehat{\boldsymbol{x}}, \boldsymbol{X}\rangle\right)^2 \to 1 - 1/\lambda$ a.s.

An analog statement for the spiked Wishart model is proved in [29]. These results give us a precise understanding of the performance of the top eigenvectors (or top singular vectors) for recovering the low-rank signals.

However these naive spectral estimators do not take into account any prior information on the signal. Thus many algorithms have been proposed to exploit additional properties of the signal, such as sparsity [116, 55, 223, 6, 61] or positivity [151].

Another line of works study Approximate Message Passing (AMP) algorithms for the spiked models above, see [177, 60, 134, 152]. Motivated by deep insights from statistical physics, these algorithms are believed (for the models (3.1.1)-(3.1.2), when λ and the priors P_0, P_U, P_V are known by the statistician) to be optimal among all polynomial-time algorithms. A great property of these algorithms is that their performance can be precisely tracked in the high-dimensional limit by a simple recursion called "state evolution", see [26, 112, 25]. For a detailed analysis of message-passing algorithms for the models (3.1.1)-(3.1.2), see [135].

It turns out that fixed points of these AMP algorithms are stationary points of the so-called "TAP[2] free energy" from statistical physics. For the model (3.1.1), the minimization of the TAP free energy was studied in [83], who showed that the minimizer was equal to

the posterior mean of $\boldsymbol{X}$, provided that λ was large enough.

In the following we will not consider any particular estimator but rather try to compute the best performance achievable by any estimator. We will suppose to be in the so-called "Bayes-optimal" setting, where the statistician knows the prior P_0 (or P_U, P_V) and the signal-to-noise ratio λ. In that situation, we will study the posterior distribution of the signal given the observations. As we should see in the sequel, both estimation problems (3.1.1)-(3.1.2) can be seen as mean-field spin glass models similar to the Sherrington-Kirkpatrick model, studied in the ground-breaking book of Mézard, Parisi and Virasoro [143]. Therefore, the methods that we will use here come from the mathematical study of spin glasses, namely from the works of Talagrand [201, 202], Guerra [95] and Panchenko [170].

In order to further motivate the study of the models (3.1.1)-(3.1.2) let us mention some interesting special cases, depending on the choice of the priors P_0 / P_U, P_V.

- **Sparse PCA**. Consider the spiked Wishart model with $P_U = \mathrm{Ber}(\epsilon)$ and $P_V = \mathcal{N}(0,1)$. In that case, one sees that conditionally on $\boldsymbol{U}$ the columns of $\boldsymbol{Y}$ are i.i.d. sampled from $\mathcal{N}\left(\boldsymbol{0}, \mathrm{Id}_n + \lambda/n \boldsymbol{U}\boldsymbol{U}^\intercal\right)$, which is a sparse spiked covariance model. The spiked Wigner model with $P_0 = \mathrm{Ber}(\epsilon)$ has also been used to study sparse PCA.

- **Submatrix localization**. Take $P_0 = \mathrm{Ber}(p)$ in the spiked Wigner model. The goal of submatrix localization is then to extract a submatrix of $\boldsymbol{Y}$ of size $pn \times pn$ with larger mean.

- **Community Detection** in the Stochastic Block Model (SBM). As we should see in Chapter 5 recovering two communities of size pn and $(1-p)n$ in a dense SBM of n vertices is (in some sense) "equivalent" to the spiked Wigner model with prior

$$P_0 = p\,\delta_{\sqrt{\frac{1-p}{p}}} + (1-p)\,\delta_{-\sqrt{\frac{p}{1-p}}}\,.$$

- $\mathbb{Z}/2$ **synchronization**. This corresponds to the spiked Wigner model with Rademacher prior $P_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$.

- **High-dimensional Gaussian mixture clustering**. Consider the multidimensional version of the spiked Wishart model where $\boldsymbol{U} \in \mathbb{R}^{n\times k}$ and $\boldsymbol{V} \in \mathbb{R}^{m\times k}$. If one takes P_V (the distribution of the rows of $\boldsymbol{V}$) to be supported by the canonical basis of $\mathbb{R}^k$, the model is equivalent to the clustering of m points (the columns of $\boldsymbol{Y}$) in n dimensions from a Gaussian mixture model. The centers of the clusters are here the columns of $\boldsymbol{U}$.

[2]named after Thouless, Anderson and Palmer [203] who proposed a variational formula (which was recently rigorously proved by [50]) for the limiting free energy of a mean-field model of spin glasses, the Sherrington-Kirkpatrick model [191].

3.2 Information-theoretic limits in the spiked Wigner model

We consider in this chapter the spiked Wigner model (3.1.1). Let P_0 be a probability distribution on $\mathbb{R}$ that admits a finite second moment and consider the following observations:

$$Y_{i,j} = \sqrt{\frac{\lambda}{n}} X_i X_j + Z_{i,j}, \quad \text{for } 1 \leq i < j \leq n\,, \tag{3.2.1}$$

where $X_i \overset{\text{i.i.d.}}{\sim} P_0$ and $Z_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ are independent random variables. Note that we suppose here to only observe the coefficients of $\sqrt{\lambda/n}\boldsymbol{X}\boldsymbol{X}^\intercal + \boldsymbol{Z}$ that are above the diagonal. The case where all the coefficients are observed can be directly deduced from this case. In the following, $\mathbb{E}$ will denote the expectation with respect to the $\boldsymbol{X}$ and $\boldsymbol{Z}$ random variables.

Our main quantity of interest is the Minimum Mean Squared Error (MMSE) defined as:

$$\begin{aligned}\mathrm{MMSE}_n(\lambda) &= \min_{\widehat{\theta}} \frac{2}{n(n-1)} \sum_{1\leq i<j\leq n} \mathbb{E}\left[\left(X_iX_j - \widehat{\theta}_{i,j}(\boldsymbol{Y})\right)^2\right] \\ &= \frac{2}{n(n-1)} \sum_{1\leq i<j\leq n} \mathbb{E}\left[\left(X_iX_j - \mathbb{E}\left[X_iX_j|\boldsymbol{Y}\right]\right)^2\right],\end{aligned}$$

where the minimum is taken over all estimators $\widehat{\theta}$ (i.e. measurable functions of the observations $\boldsymbol{Y}$). We have the trivial upper-bound

$$\mathrm{MMSE}_n(\lambda) \leq \mathrm{DMSE} \overset{\text{def}}{=} \mathbb{E}_{P_0}[X^2]^2 - \mathbb{E}_{P_0}[X]^4\,,$$

obtained by considering the "dummy" estimator $\widehat{\theta}_{i,j} = \mathbb{E}_{P_0}[X]^2$. One can also compute the Mean Squared Error achieved by naive PCA. Let $\widehat{\boldsymbol{x}}$ be the leading eigenvector of $\boldsymbol{Y}$ with norm $\|\widehat{\boldsymbol{x}}\|^2 = n$. If we take an estimator proportional to $\widehat{x}_i\widehat{x}_j$, i.e. $\widehat{\theta}_{i,j} = \delta\widehat{x}_i\widehat{x}_j$ for $\delta \geq 0$, we can compute explicitly (using the results presented in Section 3.1) the resulting MSE as a function of δ and minimize it. The optimal value for δ depends on λ, more precisely if $\lambda < \mathbb{E}_{P_0}[X^2]^{-2}$, then $\delta = 0$ while for $\lambda \geq \mathbb{E}_{P_0}[X^2]^{-2}$, the optimal of value for δ is $\mathbb{E}_{P_0}[X^2] - \lambda^{-1}\mathbb{E}_{P_0}[X^2]^{-1}$, resulting in the following MSE for naive PCA:

$$\mathrm{MSE}_n^{\mathrm{PCA}}(\lambda) \xrightarrow[n\to\infty]{} \begin{cases} \mathbb{E}_{P_0}[X^2]^2 & \text{if } \lambda \leq \mathbb{E}_{P_0}[X^2]^{-2}, \\ \lambda^{-1}\left(2 - \lambda^{-1}\mathbb{E}_{P_0}[X^2]^{-2}\right) & \text{otherwise.} \end{cases} \tag{3.2.2}$$

We will see in Section 3.3 that in the particular case of $P_0 = \mathcal{N}(0,1)$, PCA is optimal: $\lim_{n\to\infty} \mathrm{MSE}_n^{\mathrm{PCA}} = \lim_{n\to\infty} \mathrm{MMSE}_n$.

Posterior distribution and free energy. In order to formulate our inference problem as a statistical physics problem we introduce the random Hamiltonian

$$H_n(\boldsymbol{x}) = \sum_{i<j} \sqrt{\frac{\lambda}{n}} x_i x_j Z_{i,j} + \frac{\lambda}{n} X_i X_j x_i x_j - \frac{\lambda}{2n} x_i^2 x_j^2\,. \tag{3.2.3}$$

The posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$ takes then the form

$$dP(\boldsymbol{x}\,|\,\boldsymbol{Y}) = \frac{1}{\mathcal{Z}_n(\lambda)} dP_0^{\otimes n}(\boldsymbol{x}) \exp\left(\sum_{i<j} x_i x_j \sqrt{\frac{\lambda}{n}} Y_{i,j} - \frac{\lambda}{2n} x_i^2 x_j^2\right) = \frac{1}{\mathcal{Z}_n(\lambda)} dP_0^{\otimes n}(\boldsymbol{x}) e^{H_n(\boldsymbol{x})}\,, \tag{3.2.4}$$

where $\mathcal{Z}_n(\lambda)$ is the appropriate normalization. The free energy is defined as

$$F_n(\lambda) = \frac{1}{n}\mathbb{E}\Big[\log \int dP_0^{\otimes n}(\boldsymbol{x})\; e^{H_n(\boldsymbol{x})}\Big] = \frac{1}{n}\mathbb{E}\log \mathcal{Z}_n(\lambda).$$

We will first compute the limit of the free energy F_n and then deduce the limit of MMSE_n by an I-MMSE (see Proposition 1.3.3) argument. We express the limit of F_n using the following function

$$\mathcal{F} : (\lambda, q) \mapsto \psi_{P_0}(\lambda q) - \frac{\lambda}{4}q^2 = \mathbb{E}\log\left(\int dP_0(x)\exp\left(\sqrt{\lambda q}Zx + \lambda q x X - \frac{\lambda}{2}qx^2\right)\right) - \frac{\lambda}{4}q^2, \tag{3.2.5}$$

where $Z \sim \mathcal{N}(0,1)$ and $X \sim P_0$ are independent random variables. Recall that ψ_{P_0} denotes the free energy (1.3.6) of the scalar channel (1.3.5). The main result of this section is:

Theorem 3.2.1 (*Replica-Symmetric formula for the spiked Wigner model*)

For all $\lambda > 0$,

$$F_n(\lambda) \xrightarrow[n\to\infty]{} \sup_{q\geq 0} \mathcal{F}(\lambda, q)\,. \tag{3.2.6}$$

Theorem 3.2.1 is proved in Section 3.4. In the case of Rademacher prior ($P_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$), Theorem 3.2.1 was proved in [59]. The expression (3.2.6) for general priors was conjectured by [133]. For discrete priors P_0 for which the map $\mathcal{F}(\lambda, \cdot)$ has not more than 3 stationary points, the statement of Theorem 3.2.1 was obtained in [18]. The full version of Theorem 3.2.1 as well as its multidimensional generalization (where $\boldsymbol{X} \in \mathbb{R}^{n\times k}$, k fixed) was proved in [132].

Theorem 3.2.1 allows us to compute the limit of the mutual information between the signal $\boldsymbol{X}$ and the observations $\boldsymbol{Y}$. Indeed, by using (1.3.3):

Corollary 3.2.1

$$\lim_{n\to+\infty}\frac{1}{n}I(\boldsymbol{X};\boldsymbol{Y}) = \frac{\lambda\mathbb{E}_{P_0}(X^2)^2}{4} - \sup_{q\geq 0}\mathcal{F}(\lambda, q)\,.$$

We will now use Theorem 3.2.1 to obtain the limit of the Minimum Mean Squared Error MMSE_n by the I-MMSE relation of Proposition 1.3.3. Let us define

$$D = \{\lambda > 0 \mid \mathcal{F}(\lambda, \cdot) \text{ has a unique maximizer } q^*(\lambda)\}\,.$$

We start by computing the derivative of $\lim_{n\to\infty} F_n(\lambda)$ with respect to λ.

Proposition 3.2.1

D is equal to $\mathbb{R}_{>0}$ minus some countable set and is precisely the set of $\lambda > 0$ at which the function $f : \lambda \mapsto \sup_{q\geq 0}\mathcal{F}(\lambda, q)$ is differentiable. Moreover, for all $\lambda \in D$

$$f'(\lambda) = \frac{q^*(\lambda)^2}{4}\,.$$

Proof. Let $\lambda > 0$ and compute

$$\frac{\partial}{\partial q}\mathcal{F}(\lambda, q) = \lambda\psi'_{P_0}(\lambda q) - \frac{\lambda q}{2} \leq \frac{\lambda}{2}(\mathbb{E}_{P_0}[X^2] - q)\,,$$

because ψ_{P_0} is $\frac{1}{2}\mathbb{E}_{P_0}[X^2]$-Lipschitz by Proposition 1.3.3. Consequently, the maximum of $\mathcal{F}(\lambda,\cdot)$ is achieved on $[0,\mathbb{E}_{P_0}[X^2]]$. If q^* maximizes $\mathcal{F}(\lambda,\cdot)$, the optimality condition gives $q^* = 2\psi'_{P_0}(\lambda q^*)$. Consequently

$$\frac{\partial}{\partial\lambda}\mathcal{F}(\lambda,q^*) = q^*\psi'_{P_0}(\lambda q^*) - \frac{(q^*)^2}{4} = \frac{(q^*)^2}{4}\,.$$

Now, Proposition D.2 in Appendix D gives that the $\lambda > 0$ at which f is differentiable is exactly the $\lambda > 0$ for which

$$\left\{\frac{\partial}{\partial\lambda}\mathcal{F}(\lambda,q^*) = \frac{1}{4}(q^*)^2 \,\middle|\, q^* \text{ maximizer of } \mathcal{F}(\lambda,\cdot)\right\}$$

is a singleton. These λ are precisely the elements of D. Moreover, Proposition D.2 gives also that for all $\lambda \in D$, $f'(\lambda) = \frac{q^*(\lambda)^2}{4}$, which concludes the proof. □

We deduce then the limit of MMSE_n:

Corollary 3.2.2

For all $\lambda \in D$,

$$\mathrm{MMSE}_n(\lambda) \xrightarrow[n\to\infty]{} (\mathbb{E}_{P_0}X^2)^2 - q^*(\lambda)^2\,. \tag{3.2.7}$$

Proof. By Proposition 1.3.3, $(F_n)_{n\geq 1}$ is a sequence of differentiable convex functions that converges pointwise on $\mathbb{R}_{>0}$ to f. By Proposition C.1, $F_n'(\lambda) \xrightarrow[n\to\infty]{} f'(\lambda)$ for every $\lambda > 0$ at which f is differentiable, that is for all $\lambda \in D$. We conclude using the I-MMSE relation (1.3.4):

$$\frac{n-1}{4n}(\mathbb{E}_{P_0}[X^2]^2 - \mathrm{MMSE}_n(\lambda)) = F_n'(\lambda) \xrightarrow[n\to\infty]{} f'(\lambda) = \frac{q^*(\lambda)^2}{4}\,. \tag{3.2.8}$$

□

Let us now define the information-theoretic threshold

$$\lambda_c = \inf\left\{\lambda \in D \,\middle|\, q^*(\lambda) > (\mathbb{E}_{P_0}X)^2\right\}. \tag{3.2.9}$$

If the above set is empty, we define $\lambda_c = 0$. By Corollary 3.2.2 we obtain that

- if $\lambda > \lambda_c$, then $\lim_{n\to\infty} \mathrm{MMSE}_n < \mathrm{DMSE}$: one can estimate the signal better than a random guess.
- if $\lambda < \lambda_c$, then $\lim_{n\to\infty} \mathrm{MMSE}_n = \mathrm{DMSE}$: one can not estimate the signal better than a random guess.

Thus, there is no hope for reconstructing the signal below λ_c. Interestingly, one can not even detect if the measurements $\boldsymbol{Y}$ contains some signal below λ_c. If one denotes by Q_λ the distribution of $\boldsymbol{Y}$ given by (3.2.1), the work [3] shows that for $\lambda < \lambda_c$ one can not asymptotically distinguish between Q_λ and Q_0: both distributions are contiguous.

3.3 Information-theoretic and algorithmic phase transitions

3.3.1 Approximate Message Passing (AMP) algorithms

Approximate Message Passing (AMP) algorithms, introduced in [71] for compressed sensing, have then be used for various other tasks. Rigorous properties of AMP algorithms

have been established in [26, 112, 25, 31], following the seminal work of Bolthausen [35]. In the context of low-rank matrix estimation an AMP algorithm has been proposed by [177] for the rank-one case and then by [139] for finite-rank matrix estimation. For detailed review and developments about matrix factorization with message-passing algorithms, see [135]. We will only give a brief description of AMP here and we let the reader refer to [177, 60, 133, 152]. In this section, we follow [152] who provides the most advanced results for our problem (3.1.1). For simplicity, we assume here that P_0 has a unit second moment: $\int x^2 dP_0(x) = 1$.

Starting from an initialization $\boldsymbol{x}^0$, the AMP algorithm produces vectors $\boldsymbol{x}^1, \dots, \boldsymbol{x}^t$ according to the following recursion:

$$\boldsymbol{x}^{t+1} = (\boldsymbol{Y}/\sqrt{n}) f_t(\boldsymbol{x}^t) - b_t f_{t-1}(\boldsymbol{x}^{t-1}), \tag{3.3.1}$$

where $b_t = \frac{1}{n}\sum_{i=1}^n f_t'(x_i^t)$ and where the functions f_t act componentwise on vectors. After t iterations of (3.3.1), the AMP estimate of $\boldsymbol{X}$ is defined by $\widehat{\boldsymbol{x}}^t = f_t(\boldsymbol{x}^t)$.

A natural choice for the initialization is to take $\boldsymbol{x}^0$ proportional to $\boldsymbol{\varphi}_1$, the leading unit eigenvector of $\boldsymbol{Y}$:

$$\boldsymbol{x}^0 = \sqrt{n(\lambda^2 - 1)}\boldsymbol{\varphi}_1.$$

We need now to specify the "denoisers" $(f_t)_{t\geq 1}$. Let us consider the following one-dimensional recursion:

$$\begin{cases} q_0 & = (1-\lambda^{-1})_+, \\ q_{t+1} & = 2\psi_{P_0}'(\lambda q_t) = 1 - \mathrm{MMSE}_{P_0}(\lambda q_t). \end{cases} \tag{3.3.2}$$

Recall the additive Gaussian scalar channel from Section 1.3: $Y_0 = \sqrt{\gamma} X_0 + Z_0$. Let us define $g_{P_0}(y, \gamma) = \mathbb{E}[X_0 | \sqrt{\gamma} X_0 + Z_0 = y]$. We define then

$$f_t(x) = g_{P_0}\Big(x/\sqrt{q_t}, \lambda q_t\Big). \tag{3.3.3}$$

The next theorem is a consequence of the more general results of [152], specified to our setting.

Theorem 3.3.1

For all $t \geq 0$,

$$\lim_{n\to\infty} \frac{|\langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}\rangle|}{\|\widehat{\boldsymbol{x}}^t\|\|\boldsymbol{X}\|} = \lim_{n\to\infty} \|\widehat{\boldsymbol{x}}^t\| = \sqrt{q_t}.$$

Consequently,

$$\mathrm{MSE}_t^{\mathrm{AMP}} \overset{\mathrm{def}}{=} \lim_{n\to\infty} \frac{1}{n^2}\mathbb{E}\|\boldsymbol{X}\boldsymbol{X}^\intercal - \widehat{\boldsymbol{x}}^t(\widehat{\boldsymbol{x}}^t)^\intercal\|^2 = 1 - q_t^2. \tag{3.3.4}$$

By Proposition 1.3.3, the function ψ_{P_0}' is increasing and bounded. The sequence $(q_t)_{t\geq 0}$ converges therefore to a point $q_\infty \geq 0$ that verifies $q_\infty = 2\psi_{P_0}'(\lambda q_\infty)$. q_∞ is therefore a critical point of $\mathcal{F}(\lambda, \cdot)$. In the case where q_∞ is the global minimizer of $\mathcal{F}(\lambda, \cdot)$, i.e. $q_\infty = q^*(\lambda)$, we see using Corollary 3.2.2 that $\lim_{t\to\infty} \mathrm{MSE}_t^{\mathrm{AMP}} = \mathrm{MMSE}(\lambda)$: AMP achieves the Bayes-optimal accuracy.

In the case where $q_\infty \neq q^*(\lambda)$, AMP does not reach the information-theoretically optimal performance. However, AMP is conjectured (see for instance [220, 7]) to be optimal among polynomial-time algorithms, i.e. $\lim_{t\to\infty} \mathrm{MSE}_t^{\mathrm{AMP}}$ is conjectured to be the best Mean Squared Error achievable by any polynomial-time algorithm.

3.3.2 Examples of phase transitions

We give here some illustrations and interpretations of the results presented in the previous sections. Let us first study the case where $P_0 = \mathcal{N}(0,1)$ where the formulas (3.2.6) and (3.2.7) can be evaluated explicitly. Indeed, we saw in Example 1.3.1 in Section 1.3 that $\psi_{\mathcal{N}(0,1)}(q) = \frac{1}{2}\big(q - \log(1+q)\big)$. We can then compute $q^*(\lambda) = (1-\lambda^{-1})_+$ which gives

$$\lim_{n\to\infty} \mathrm{MMSE}_n(\lambda) = \begin{cases} 0 & \text{if } \lambda \le 1\,, \\ \frac{1}{\lambda}\big(2 - \frac{1}{\lambda}\big) & \text{if } \lambda \ge 1\,. \end{cases}$$

Comparing the limit above with the performance of (naive) PCA given by (3.2.2) we see that in the case $P_0 = \mathcal{N}(0,1)$, PCA is information-theoretically optimal.

However, as we see on (3.2.2), the MSE of PCA only depends on the second moment of P_0: naive PCA is not able to exploit additional properties of the signal. We compare on Figure 3.1 the asymptotic performance of the naive PCA (3.2.2) and the Approximate Message Passing (AMP) algorithm (3.3.4) to the asymptotic Minimum Mean Squared Error for the prior

$$P_0 = p\,\delta_{\sqrt{\frac{1-p}{p}}} + (1-p)\,\delta_{-\sqrt{\frac{p}{1-p}}}\,, \tag{3.3.5}$$

where $p \in (0,1)$. This is a two-points distribution with zero mean and unit variance. It is of particular interest because it is related with the community detection problem in the (dense) Stochastic Block Model as we will see in Section 5.9. We see on Figure 3.1 that

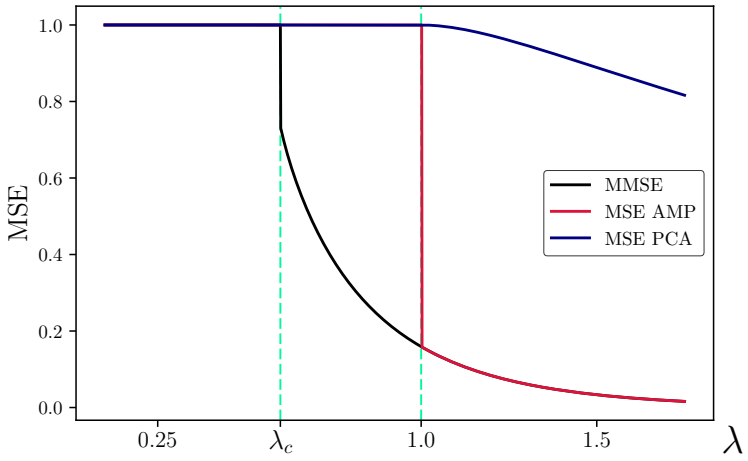
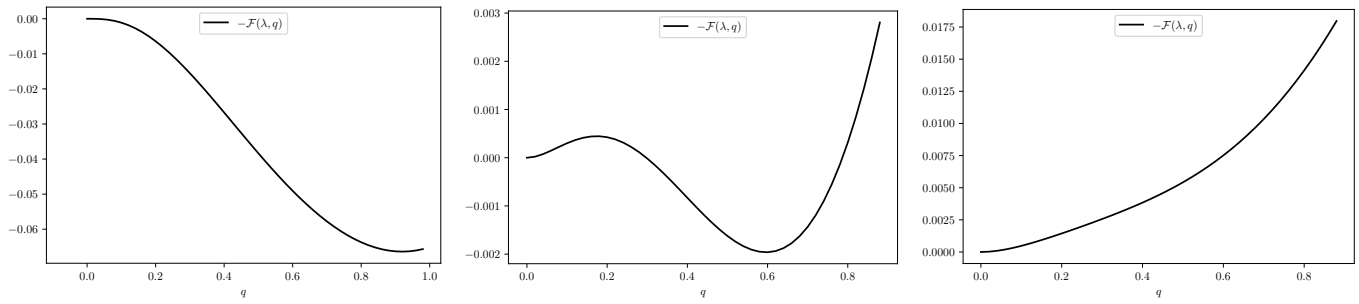


Figure 3.1: Mean Squared Errors for the Spiked Wigner model with prior P_0 given by (3.3.5) with $p = 0.05$.

the MMSE is equal to 1 for λ below the information-theoretic threshold $\lambda_c \simeq 0.6$. One can not asymptotically recover the signal better than a random guess in this region: we call this region the "impossible" phase. For $\lambda > 1$ we see that spectral methods and AMP perform better than random guessing. This region is therefore called the "easy" phase, because non-trivial estimation is here possible using efficient algorithms. Notice also that AMP achieves the Minimum Mean Squared Error for $\lambda > 1$, as proved in [152]. The region $\lambda_c < \lambda < 1$ is more intriguing. It is still possible to build a non-trivial estimator (for instance by computing the posterior mean), but our two polynomial-time algorithms fail. This region is thus denoted as the "hard" phase because it is conjectured that polynomial-time algorithms can only provide trivial estimates (based on the belief that AMP is here

optimal among polynomial-time algorithms).

Quite surprisingly, one can guess in which phase (easy-hard-impossible) we are, simply by plotting the "potential" $q \mapsto -\mathcal{F}(\lambda, q)$. This is done in Figure 3.2. By Corollary 3.2.2



(a) "Easy" phase ($\lambda = 1.01$) (b) "Hard" phase ($\lambda = 0.625$) (c) "Impossible" phase ($\lambda = 0.5$)

Figure 3.2: Plots of $q \mapsto -\mathcal{F}(\lambda, q)$ for different values of λ and P_0 given by (3.3.5) with $p = 0.05$.

we know that the limit of the MMSE is equal to $1 - q^*(\lambda)^2$ where $q^*(\lambda)$ is the minimizer of $-\mathcal{F}(\lambda, \cdot)$. Thus when $-\mathcal{F}(\lambda, \cdot)$ is minimal at $q = 0$, we are in the impossible phase.

When $q^*(\lambda) > 0$, the shape of $-\mathcal{F}(\lambda, \cdot)$ indicates whether we are in the easy or hard phase. If the $q = 0$ is a local maximum, then we are in the easy phase, whereas when it is a local minimum we are in a hard phase. The shape of $-\mathcal{F}(\lambda, \cdot)$ could be interpreted as a simplified "free energy landscape": the hard phase appears when the "informative" minimum $q^*(\lambda) > 0$ is separated from the non-informative critical point $q = 0$ by a "free energy barrier" as in Figure 3.2 (b).

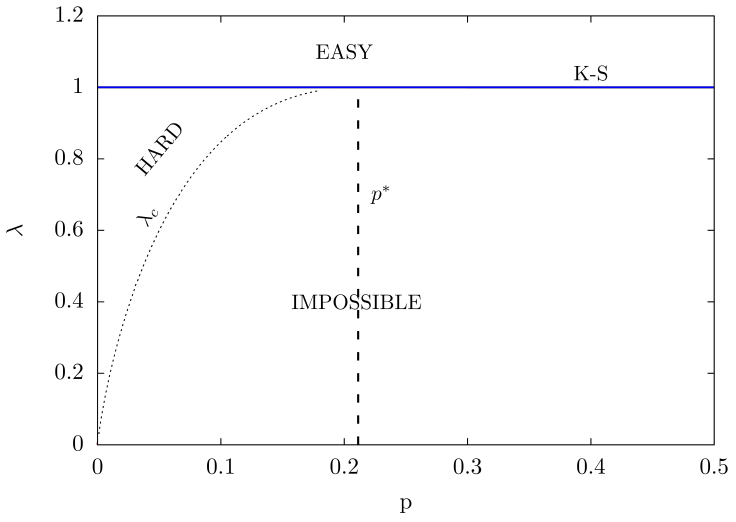


Figure 3.3: Phase diagram for the spiked Wigner model with prior (3.3.5).

The phase diagram from Figure 3.3 displays the three phases on the (p, λ)-plane. One observes that the hard phase only appears when the prior is sufficiently asymmetric, i.e. for $p < p^* = \frac{1}{2} - \frac{1}{2\sqrt{3}}$, as computed in [18, 40]. For a more detailed analysis of the phase transitions in the spiked Wigner model, see [135] where many other priors are considered.

3.4 Proof of the Replica-Symmetric formula (Theorem 3.2.1)

We prove Theorem 3.2.1 in this section, following [132]. We have to mention that other proofs of Theorem 3.2.1 have appeared since then: see [20, 80, 158].

Because of an approximation argument presented in Section 3.4.7 it suffices to prove Theorem 3.2.1 for priors P_0 with finite (and thus bounded) support $S \subset [-K, K]$, for some $K > 0$. From now, we assume to be in that case.

3.4.1 The lower bound: Guerra's interpolation method

The following result comes from [129]. It adapts arguments from the study of the gauge symmetric p-spin glass model of [124] to the inference model (3.2.1). It is based on Guerra's interpolation technique for the Sherrington-Kirkpatrick model, see [95]. We reproduce the proof for completeness.

Proposition 3.4.1

$$\liminf_{n\to\infty} F_n(\lambda) \geq \sup_{q\geq 0} \mathcal{F}(\lambda, q)\,. \tag{3.4.1}$$

Proof. Let $q \geq 0$. For $t \in [0,1]$ we define

$$H_{n,t}(\boldsymbol{x}) = \sum_{i<j} \sqrt{\frac{\lambda t}{n}} Z_{i,j} x_i x_j + \frac{\lambda t}{n} x_i x_j X_i X_j - \frac{\lambda t}{2n} x_i^2 x_j^2 + \sum_{i=1}^n \sqrt{(1-t)\lambda q} Z_i' x_i + (1-t)\lambda q x_i X_i - \frac{(1-t)\lambda q}{2} x_i^2\,.$$

Let $\langle\cdot\rangle_{n,t}$ denote the Gibbs measure associated with the Hamiltonian $H_{n,t}(\boldsymbol{x})$:

$$\langle f(\boldsymbol{x})\rangle_{n,t} = \frac{\sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x}) f(\boldsymbol{x}) e^{H_{n,t}(\boldsymbol{x})}}{\sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x}) e^{H_{n,t}(\boldsymbol{x})}}\,,$$

for any function f on S^n. The Gibbs measure $\langle\cdot\rangle_{n,t}$ corresponds to the distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$ and $\boldsymbol{Y}'$ in the following inference channel:

$$\begin{cases} Y_{i,j} = \sqrt{\frac{\lambda t}{n}} X_i X_j + Z_{i,j} & \text{for } 1 \leq i < j \leq n, \\ Y_i' = \sqrt{(1-t)\lambda q} X_i + Z_i' & \text{for } 1 \leq i \leq n, \end{cases}$$

where $X_i \overset{\text{i.i.d.}}{\sim} P_0$ and $Z_{i,j},\ Z_i' \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ are independent random variables. We will therefore be able to apply the Nishimori property (Proposition 1.1.1) to the Gibbs measure $\langle\cdot\rangle_{n,t}$. Let us define

$$\psi : t \in [0,1] \mapsto \frac{1}{n}\mathbb{E}\log \sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x}) e^{H_{n,t}(\boldsymbol{x})}\,.$$

We have $\psi(1) = F_n(\lambda)$ and

$$\begin{aligned} \psi(0) &= \frac{1}{n}\mathbb{E}\log \sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x}) \exp\left(\sum_{i=1}^n \sqrt{\lambda q} Z_i' x_i + \lambda q x_i X_i - \frac{\lambda q}{2} x_i^2\right) \\ &= \frac{1}{n}\mathbb{E}\log \prod_{i=1}^n \left(\sum_{x_i\in S} P_0(x_i) \exp\left(\sqrt{\lambda q} Z_i' x_i + \lambda q x_i X_i - \frac{\lambda q}{2} x_i^2\right)\right) \\ &= \mathcal{F}(\lambda, q) + \frac{\lambda q^2}{4}\,. \end{aligned}$$

ψ is continuous on $[0,1]$, differentiable on $(0,1)$. For $0 < t < 1$,

$$\psi'(t) = \frac{1}{n}\mathbb{E}\left\langle \sum_{i<j} \frac{\sqrt{\lambda}}{2\sqrt{nt}} Z_{i,j} x_i x_j + \frac{\lambda}{n} x_i x_j X_i X_j - \frac{\lambda}{2n} x_i^2 x_j^2 - \sum_{i=1}^{n} \frac{\sqrt{\lambda q}}{2\sqrt{1-t}} Z_i' x_i - \lambda q x_i X_i + \frac{\lambda q}{2} x_i^2 \right\rangle_{n,t}, \tag{3.4.2}$$

where $\boldsymbol{x}$ is a sample from the Gibbs measure $\langle\cdot\rangle_{n,t}$, independently of everything else. For $1 \le i < j \le n$ we have, by Gaussian integration by parts and by the Nishimori property

$$\begin{aligned}\mathbb{E}\left[Z_{i,j} \left\langle \frac{\sqrt{\lambda}}{2\sqrt{nt}} x_i x_j \right\rangle_{n,t} \right] &= \frac{\lambda}{2n}\Big(\mathbb{E}\langle x_i^2 x_j^2\rangle_{n,t} - \mathbb{E}\langle x_i x_j\rangle_{n,t}^2\Big) = \frac{\lambda}{2n}\Big(\mathbb{E}\langle x_i^2 x_j^2\rangle_{n,t} - \mathbb{E}\langle x_i^{(1)} x_j^{(1)} x_i^{(2)} x_j^{(2)}\rangle_{n,t}\Big) \\ &= \frac{\lambda}{2n}\Big(\mathbb{E}\langle x_i^2 x_j^2\rangle_{n,t} - \mathbb{E}\langle x_i x_j X_i X_j\rangle_{n,t}\Big),\end{aligned}$$

where $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ are two independent samples from the Gibbs measure $\langle\cdot\rangle_{n,t}$, independently of everything else. Similarly, we have for $1 \le i \le n$

$$\mathbb{E}\left\langle \frac{\sqrt{\lambda q}}{2\sqrt{1-t}} Z_i' x_i \right\rangle_{n,t} = \frac{\lambda q}{2}\left(\mathbb{E}\langle x_i^2\rangle_{n,t} - \mathbb{E}\langle x_i X_i\rangle_{n,t}\right).$$

Therefore (3.4.2) simplifies

$$\begin{aligned}\psi'(t) &= \frac{1}{n}\mathbb{E}\Big\langle \sum_{i<j} \frac{\lambda}{2n} x_i x_j X_i X_j - \sum_{i=1}^{n} \frac{\lambda q}{2} x_i X_i \Big\rangle_{n,t} = \frac{\lambda}{4}\mathbb{E}\Big\langle (\boldsymbol{x}\cdot\boldsymbol{X})^2 - 2q\,\boldsymbol{x}\cdot\boldsymbol{X} \Big\rangle_{n,t} + o_n(1) \\ &= \frac{\lambda}{4}\mathbb{E}\Big\langle (\boldsymbol{x}\cdot\boldsymbol{X} - q)^2 \Big\rangle_{n,t} - \frac{\lambda q^2}{4} + o_n(1) \ge -\frac{\lambda q^2}{4} + o_n(1)\,,\end{aligned} \tag{3.4.3}$$

where $o_n(1)$ denotes a quantity that goes to 0 uniformly in $t \in (0,1)$. Then

$$F_n(\lambda) - \mathcal{F}(\lambda, q) - \frac{\lambda}{4}q^2 = \psi(1) - \psi(0) = \int_0^1 \psi'(t)dt \ge -\frac{\lambda}{4}q^2 + o_n(1)\,.$$

Thus $\liminf_{n\to\infty} F_n(\lambda) \ge \mathcal{F}(\lambda, q)$, for all $q \ge 0$. □

3.4.2 Adding a small perturbation

It remains to prove the converse bound of (3.4.1). For this purpose, we need to show that the overlap $\boldsymbol{x}\cdot\boldsymbol{X}$ (where $\boldsymbol{x}$ is a sample from the posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$, independently of everything else) concentrates around its mean. To obtain such a result, we follow the ideas of Section 2.1 that states that giving a small amount of side information to the statistician forces the overlap to concentrate, while keeping the free energy almost unchanged.

Let us fix $\epsilon \in [0,1]$, and suppose that we have access, in addition of $\boldsymbol{Y}$, to the additional information, for $1 \le i \le n$

$$Y_i' = \begin{cases} X_i & \text{if } L_i = 1, \\ * & \text{if } L_i = 0, \end{cases} \tag{3.4.4}$$

where $L_i \overset{\text{i.i.d.}}{\sim} \mathrm{Ber}(\epsilon)$ and $*$ is a value that does not belong to S. Recall the free energy that corresponds to this perturbed inference channel is

$$F_{n,\epsilon} = \frac{1}{n}\mathbb{E}\Big[\log \sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x}) \exp(H_n(\bar{\boldsymbol{x}}))\Big],$$

where

$$\bar{\boldsymbol{x}} = (\bar{x}_1, \dots, \bar{x}_n) = (L_1 X_1 + (1 - L_1)x_1, \dots, L_n X_n + (1 - L_n)x_n)\,. \tag{3.4.5}$$

From now we suppose $\epsilon_0 \in (0, 1]$ to be fixed and consider $\epsilon \in [0, \epsilon_0]$. We will compute the limit of $F_{n,\epsilon}$ as $n \to \infty$ and then let $\epsilon \to 0$ to deduce the limit of F_n, because by Proposition 2.1.1

$$|F_{n,\epsilon} - F_n| \leq H(P_0)\epsilon\,.$$

3.4.3 Aizenman-Sims-Starr scheme

The Aizenman-Sims-Starr scheme was introduced in [2] in the context of the SK model. This is what physicists call a "cavity computation": one compare the system with $n + 1$ variables to the system with n variables and see what happen to the $(n + 1)^{\text{th}}$ variable we add.

With the convention $F_{0,\epsilon} = 0$, we have $F_{n,\epsilon} = \frac{1}{n}\sum_{k=0}^{n-1} A^{(0)}_{k,\epsilon}$ where

$$A^{(0)}_{k,\epsilon} = (k + 1)F_{k+1,\epsilon} - kF_{k,\epsilon} = \mathbb{E}[\log(\mathcal{Z}_{k+1,\epsilon})] - \mathbb{E}[\log(\mathcal{Z}_{k,\epsilon})]\,.$$

We recall that $\mathcal{Z}_{n,\epsilon} = \sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x})e^{H_n(\bar{\boldsymbol{x}})}$ where the notation $\bar{\boldsymbol{x}}$ is defined by equation (3.4.5). Consequently

$$\limsup_{n\to\infty} \int_0^{\epsilon_0} d\epsilon\, F_{n,\epsilon} \leq \limsup_{n\to\infty} \int_0^{\epsilon_0} d\epsilon\, A^{(0)}_{n,\epsilon}\,. \tag{3.4.6}$$

We now compare H_{n+1} with H_n. Let $\boldsymbol{x} \in S^n$ and $\sigma \in S$. σ plays the role of the $(n + 1)^{\text{th}}$ variable. We decompose $H_{n+1}(\boldsymbol{x}, \sigma) = H'_n(\boldsymbol{x}) + \sigma z_0(\boldsymbol{x}) + \sigma^2 s_0(\boldsymbol{x})$, where

$$\begin{aligned}
H'_n(\boldsymbol{x}) &= \sum_{1\leq i<j\leq n} \sqrt{\frac{\lambda}{n+1}} Z_{i,j}x_i x_j + \frac{\lambda}{n+1} X_i X_j x_i x_j - \frac{\lambda}{2(n+1)} x_i^2 x_j^2\,,\\
z_0(\boldsymbol{x}) &= \sum_{i=1}^{n} \sqrt{\frac{\lambda}{n+1}} Z_{i,n+1}x_i + \frac{\lambda}{n+1} X_i X_{n+1} x_i\,,\\
s_0(\boldsymbol{x}) &= -\frac{\lambda}{2(n+1)} \sum_{i=1}^{n} x_i^2\,.
\end{aligned}$$

Let $(\widetilde{Z}_{i,j})_{1\leq i<j\leq n}$ be independent, standard Gaussian random variables, independent of all other random variables. We have then $H_n(\boldsymbol{x}) = H'_n(\boldsymbol{x}) + y_0(\boldsymbol{x})$ in law, where

$$y_0(\boldsymbol{x}) = \sum_{1\leq i<j\leq n} \frac{\sqrt{\lambda}}{\sqrt{n(n+1)}} \widetilde{Z}_{i,j}x_i x_j + \frac{\lambda}{n(n+1)} X_i X_j x_i x_j - \frac{\lambda}{2(n+1)n} x_i^2 x_j^2\,.$$

We define the Gibbs measure $\langle\cdot\rangle_{n,\epsilon}$ by

$$\langle f(\boldsymbol{x})\rangle_{n,\epsilon} = \frac{1}{\mathcal{Z}_{n,\epsilon}} \sum_{\boldsymbol{x}\in S^n} P_0(\boldsymbol{x}) f(\bar{\boldsymbol{x}}) \exp(H'_n(\bar{\boldsymbol{x}}))\,, \tag{3.4.7}$$

for any function f on S^n. The Gibbs measure $\langle\cdot\rangle_{n,\epsilon}$ corresponds to the posterior distribution of $\boldsymbol{X}$ given $(\sqrt{\lambda/(n+1)}X_iX_j + Z_{i,j})_{1\leq i<j\leq n}$ and $\boldsymbol{Y}'$ from (3.4.4). We will therefore be able to apply the Nishimori identity (Proposition 1.1.1) and Proposition 2.1.2

to the Gibbs measure $\langle\cdot\rangle_{n,\epsilon}$. Let us define $\bar\sigma = (1-L_{n+1})\sigma + L_{n+1}X_{n+1}$. We can rewrite $\mathcal{Z}_{n+1,\epsilon} = \sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x}) e^{H'_n(\bar{\boldsymbol{x}})}\Big(\sum_{\sigma\in S} P_0(\sigma)\exp(\bar\sigma z_0(\bar{\boldsymbol{x}}) + \bar\sigma^2 s_0(\bar{\boldsymbol{x}}))\Big)$ and $\mathcal{Z}_{n,\epsilon} = \sum_{\boldsymbol{x}\in S^n} P_0^{\otimes n}(\boldsymbol{x}) e^{H'_n(\bar{\boldsymbol{x}})} e^{y_0(\bar{\boldsymbol{x}})}$. Thus

$$A^{(0)}_{n,\epsilon} = \mathbb{E}\log\Big\langle \sum_{\sigma\in S} P_0(\sigma)\exp\Big(\bar\sigma z_0(\boldsymbol{x}) + \bar\sigma^2 s_0(\boldsymbol{x})\Big)\Big\rangle_{n,\epsilon} - \mathbb{E}\log\Big\langle \exp(y_0(\boldsymbol{x}))\Big\rangle_{n,\epsilon}\,.$$

In the sequel, it will be more convenient to use slightly simplified versions of z_0, s_0 and y_0 in order to obtain nicer expressions in the sequel. We define

$$\begin{aligned}
z(\boldsymbol{x}) &= \sum_{i=1}^n \sqrt{\frac{\lambda}{n}} Z_{i,n+1}x_i + \frac{\lambda}{n}X_iX_{n+1}x_i = \sqrt{\frac{\lambda}{n}}\sum_{i=1}^n x_iZ_{i,n+1} + \lambda(\boldsymbol{x}\cdot\boldsymbol{X})X_{n+1}\,,\\
s(\boldsymbol{x}) &= -\frac{\lambda}{2n}\sum_{i=1}^n x_i^2 = -\frac{\lambda}{2}\boldsymbol{x}\cdot\boldsymbol{x}\,,\\
y(\boldsymbol{x}) &= \frac{\sqrt{\lambda}}{\sqrt{2}n}\sum_{i=1}^n Z''_i x_i^2 + \frac{\lambda}{2n^2}\sum_{i=1}^n\Big(x_i^2X_i^2 - \frac{x_i^4}{2}\Big) + \frac{\sqrt{\lambda}}{n}\sum_{1\le i<j\le n} x_ix_j\Big(\widetilde{Z}_{i,j} + \frac{\sqrt{\lambda}}{n}X_iX_j\Big) - \frac{\lambda}{2n^2}x_i^2x_j^2\\
&= \frac{\sqrt{\lambda}}{\sqrt{2}n}\sum_{i=1}^n Z''_i x_i^2 + \frac{\sqrt{\lambda}}{n}\sum_{1\le i<j\le n} x_ix_j\widetilde{Z}_{i,j} + \frac{\lambda}{2}\Big((\boldsymbol{x}\cdot\boldsymbol{X})^2 - \frac{1}{2}(\boldsymbol{x}\cdot\boldsymbol{x})^2\Big)\,,
\end{aligned}$$

where $Z''_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ independently of any other random variables. Define now

$$A_{n,\epsilon} = \mathbb{E}\log\Big\langle\sum_{\sigma\in S} P_0(\sigma)\exp(\bar\sigma z(\boldsymbol{x}) + \bar\sigma^2 s(\boldsymbol{x}))\Big\rangle_{n,\epsilon} - \mathbb{E}\log\langle\exp(y(\boldsymbol{x}))\rangle_{n,\epsilon}\,.$$

Using Gaussian interpolation techniques, it is not difficult to show that $\int_0^{\epsilon_0} d\epsilon\,(A_{n,\epsilon} - A^{(0)}_{n,\epsilon}) \xrightarrow[n\to\infty]{} 0$ because the modifications made in z_0, s_0 and y_0 are of negligible order. Using (3.4.6) we conclude

$$\limsup_{n\to\infty}\int_0^{\epsilon_0} d\epsilon\, F_{n,\epsilon} \le \limsup_{n\to\infty}\int_0^{\epsilon_0} d\epsilon\, A_{n,\epsilon}\,. \tag{3.4.8}$$

3.4.4 Overlap concentration

Proposition 2.1.2 implies that the overlap between two replicas, i.e. two independent samples $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ from the Gibbs distribution $\langle\cdot\rangle_{n,\epsilon}$, concentrates. Let us define the random variables

$$Q = \Big\langle\frac{1}{n}\sum_{i=1}^n x_i^{(1)}x_i^{(2)}\Big\rangle_{n,\epsilon} \qquad\text{and}\qquad b_i = \langle x_i\rangle_{n,\epsilon}\,.$$

Notice that $Q = \frac{1}{n}\sum_i b_i^2 \ge 0$. By Proposition 2.1.2 we know that

$$\int_0^{\epsilon_0} d\epsilon\mathbb{E}\Big\langle(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)} - Q)^2\Big\rangle_{n,\epsilon} \xrightarrow[n\to\infty]{} 0\,. \tag{3.4.9}$$

Thus, using the Nishimori property (Proposition 1.1.1) we deduce:

$$\int_0^{\epsilon_0} d\epsilon\mathbb{E}\Big\langle(\boldsymbol{x}\cdot\boldsymbol{X} - Q)^2\Big\rangle_{n,\epsilon} \xrightarrow[n\to\infty]{} 0 \qquad\text{and}\qquad \int_0^{\epsilon_0} d\epsilon\mathbb{E}\Big\langle(\boldsymbol{x}\cdot\boldsymbol{b} - Q)^2\Big\rangle_{n,\epsilon} \xrightarrow[n\to\infty]{} 0\,. \tag{3.4.10}$$

3.4.5 The main estimate

Let us denote, for $\epsilon \in [0,1]$,

$$\mathcal{F}_\epsilon : (\lambda, q) \mapsto -\frac{\lambda}{4}q^2 + \epsilon(\mathbb{E}_{P_0}X^2)\frac{\lambda q}{2} + (1-\epsilon)\mathbb{E}\Big[\log \sum_{x\in S} P_0(x)\exp\Big(\sqrt{\lambda q}Zx + \lambda q x X - \frac{\lambda}{2}qx^2\Big)\Big]$$

where the expectation $\mathbb{E}$ is taken with respect to the independent random variables $X \sim P_0$ and $Z \sim \mathcal{N}(0,1)$. The following proposition is one of the key steps of the proof.

Proposition 3.4.2

For all $\epsilon_0 \in [0,1]$,

$$\int_0^{\epsilon_0} d\epsilon\, (A_{n,\epsilon} - \mathbb{E}[\mathcal{F}_\epsilon(\lambda, Q)]) \xrightarrow[n\to\infty]{} 0\,.$$

The proof of Proposition 3.4.2 is deferred to Section 3.4.6. We deduce here Theorem 3.2.1 from Proposition 3.4.2 and the results of the previous sections. Because of Proposition 3.4.1, we only have to show that $\limsup_{n\to\infty} F_n \le \sup_{q\ge 0}\mathcal{F}(\lambda, q)$.

By Proposition 2.1.1 we have

$$\epsilon_0 F_n \le \int_0^{\epsilon_0} d\epsilon F_{n,\epsilon} + \frac{1}{2}H(P_0)\epsilon_0^2\,.$$

Therefore by equation (3.4.8) and Proposition 3.4.2

$$\epsilon_0 \limsup_{n\to\infty} F_n \le \limsup_{n\to\infty}\int_0^{\epsilon_0} d\epsilon A_{n,\epsilon} + \frac{1}{2}H(P_0)\epsilon_0^2 \le \limsup_{n\to\infty}\int_0^{\epsilon_0} d\epsilon\, \mathbb{E}\mathcal{F}_\epsilon(\lambda, Q) + \frac{1}{2}H(P_0)\epsilon_0^2\,. \tag{3.4.11}$$

It remains then to show that $\limsup_{n\to\infty}\int d\epsilon\, \mathbb{E}\mathcal{F}_\epsilon(\lambda, Q) \le \epsilon_0 \sup_{q\ge 0}\mathcal{F}(\lambda, q) + O(\epsilon_0^2)$. We have for $\epsilon \in [0,1]$,

$$\begin{aligned}\sup_{q\in[0,K^2]} |\mathcal{F}_\epsilon(\lambda, q) - \mathcal{F}(\lambda, q)| &\le \epsilon \sup_{q\in[0,K^2]}\Big\{\frac{\lambda q}{2}\mathbb{E}_{P_0}[X^2] + \Big|\mathbb{E}\log\sum_{x\in S}P_0(x)\exp(\sqrt{\lambda q}Zx + \lambda q x X - \frac{\lambda}{2}qx^2)\Big|\Big\}\\ &\le C\epsilon\,,\end{aligned}$$

for some constant C that only depends on λ and P_0. Noticing that $Q \in [0, K^2]$ a.s., we have then $|\mathbb{E}\mathcal{F}_\epsilon(\lambda, Q) - \mathbb{E}\mathcal{F}(\lambda, Q)| \le C\epsilon_0$, for all $\epsilon \in [0, \epsilon_0]$ and therefore

$$\int_0^{\epsilon_0} d\epsilon\, \mathbb{E}\mathcal{F}_\epsilon(\lambda, Q) \le \epsilon_0 \sup_{q\ge 0}\mathcal{F}(\lambda, q) + \frac{1}{2}C\epsilon_0^2\,.$$

Combined with (3.4.11), this implies $\limsup_{n\to\infty} F_n \le \sup_{q\ge 0}\mathcal{F}(\lambda, q) + \frac{1}{2}H(P_0)\epsilon_0 + \frac{1}{2}C\epsilon_0$, for all $\epsilon_0 \in (0,1]$. Theorem 3.2.1 is proved.

3.4.6 Proof of Proposition 3.4.2

In this section, we prove Proposition 3.4.2 which is a consequence of Lemmas 3.4.1 and 3.4.2 below. In order to lighten the formulas, we will use the following notations

$$X' = X_{n+1} \quad \text{and} \quad Z'_i = Z_{i,n+1}.$$

Recall

$$A_{n,\epsilon} = \mathbb{E}\log\Big\langle \sum_{\sigma\in S} P_0(\sigma)\exp(\bar\sigma z(\boldsymbol{x}) + \bar\sigma^2 s(\boldsymbol{x}))\Big\rangle_{n,\epsilon} - \mathbb{E}\log\Big\langle \exp(y(\boldsymbol{x}))\Big\rangle_{n,\epsilon}, \tag{3.4.12}$$

where for $\sigma \in S$, $\bar\sigma = (1-L_{n+1})\sigma + L_{n+1}X'$. We recall that $\langle\cdot\rangle_{n,\epsilon}$ denotes the expectation with respect to $\boldsymbol{x}$ sampled from the Gibbs measure defined by (3.4.7). The computations here are closely related to the cavity computations in the SK model, see for instance [201].

Lemma 3.4.1

$$\int_0^{\epsilon_0} d\epsilon \Big| \mathbb{E}\log\Big\langle \sum_{\sigma\in S} P_0(\sigma)\exp(\bar\sigma z(\boldsymbol{x}) + \bar\sigma^2 s(\boldsymbol{x}))\Big\rangle_{n,\epsilon} - \Big(\epsilon(\mathbb{E}_{P_0}X^2)\mathbb{E}\frac{\lambda Q}{2} + (1-\epsilon)\mathbb{E}\log\sum_{\sigma\in S} P_0(\sigma)\exp\Big(\sqrt{\lambda Q}\sigma Z_0 + \lambda Q\sigma X' - \frac{\lambda\sigma^2}{2}Q\Big)\Big)\Big| \xrightarrow[n\to\infty]{} 0\,,$$

where $Z_0 \sim \mathcal{N}(0,1)$ is independent of all other random variables.

Lemma 3.4.2

$$\int_0^{\epsilon_0} d\epsilon \Big| \mathbb{E}\log\Big\langle \exp(y(\boldsymbol{x}))\Big\rangle_{n,\epsilon} - \frac{\lambda}{4}\mathbb{E}Q^2\Big| \xrightarrow[n\to\infty]{} 0\,.$$

We will only prove Lemma 3.4.1 here since Lemma 3.4.2 follows from the same kind of arguments (the full proof can be found in [132]). The remaining of the section is thus devoted to the proof of Lemma 3.4.1.

Let us write $f(z,s) = \sum_{\sigma\in S} P_0(\sigma)e^{\bar\sigma z + \bar\sigma^2 s}$ and we define:

$$U = \Big\langle f(z(\boldsymbol{x}), s(\boldsymbol{x}))\Big\rangle_{n,\epsilon},$$
$$V = \sum_{\sigma\in S} P_0(\sigma)\exp\left(\bar\sigma\sqrt{\frac{\lambda}{n}}\sum_{i=1}^n b_i Z_i' + \lambda Q X'\bar\sigma - \frac{\lambda Q}{2}\bar\sigma^2\right).$$

Lemma 3.4.3

$$\int_0^{\epsilon_0} d\epsilon\, \mathbb{E}\Big[(U-V)^2\Big] \xrightarrow[n\to\infty]{} 0\,.$$

Proof. It suffices to show that $\int d\epsilon\,|\mathbb{E}U^2 - \mathbb{E}V^2| \xrightarrow[n\to\infty]{} 0$ and $\int d\epsilon\,|\mathbb{E}UV - \mathbb{E}V^2| \xrightarrow[n\to\infty]{} 0$. Let $\mathbb{E}_{\boldsymbol{Z}'}$ denote the expectation with respect to $\boldsymbol{Z}' = (Z_{i,n+1})_{1\le i\le n}$ only. Compute

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{Z}'}V^2 &= \mathbb{E}_{\boldsymbol{Z}'}\sum_{\sigma_1,\sigma_2\in S} P_0(\sigma_1,\sigma_2)\exp\Big((\bar\sigma_1+\bar\sigma_2)\sqrt{\frac{\lambda}{n}}\sum_{i=1}^n b_i Z_i' + \lambda Q X'(\bar\sigma_1+\bar\sigma_2) - \frac{\lambda Q}{2}(\bar\sigma_1^2+\bar\sigma_2^2)\Big)\\
&= \sum_{\sigma_1,\sigma_2\in S} P_0(\sigma_1,\sigma_2)\exp\Big((\bar\sigma_1+\bar\sigma_2)^2\frac{\lambda}{2}Q + \lambda Q X'(\bar\sigma_1+\bar\sigma_2) - \frac{\lambda Q}{2}(\bar\sigma_1^2+\bar\sigma_2^2)\Big)\\
&= \sum_{\sigma_1,\sigma_2\in S} P_0(\sigma_1,\sigma_2)\exp\left(\bar\sigma_1\bar\sigma_2\lambda Q + \lambda Q X'(\bar\sigma_1+\bar\sigma_2)\right)
\end{aligned} \tag{3.4.13}$$

where we write for $i=1,2$, $\bar\sigma_i = (1-L_{n+1})\sigma_i + L_{n+1}X'$, as before.

Let us show that $\int d\epsilon \, |\mathbb{E}U^2 - \mathbb{E}V^2| \xrightarrow[n\to\infty]{} 0$.

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{Z}'} U^2 &= \mathbb{E}_{\boldsymbol{Z}'} \big\langle f(z(\boldsymbol{x}), s(\boldsymbol{x})) \big\rangle_{n,\epsilon}^2 \\
&= \mathbb{E}_{\boldsymbol{Z}'} \big\langle f(z(\boldsymbol{x}^{(1)}), s(\boldsymbol{x}^{(1)})) f(z(\boldsymbol{x}^{(2)}), s(\boldsymbol{x}^{(2)})) \big\rangle_{n,\epsilon} \quad (\boldsymbol{x}^{(1)} \text{ and } \boldsymbol{x}^{(2)} \text{ are indep. samples from } \langle\cdot\rangle_{n,\epsilon}) \\
&= \Big\langle \mathbb{E}_{\boldsymbol{Z}'} f(z(\boldsymbol{x}^{(1)}), s(\boldsymbol{x}^{(1)})) f(z(\boldsymbol{x}^{(2)}), s(\boldsymbol{x}^{(2)})) \Big\rangle_{n,\epsilon} \\
&= \left\langle \sum_{\sigma_1,\sigma_2 \in S} P_0(\sigma_1,\sigma_2) \mathbb{E}_{\boldsymbol{Z}'} \exp\left(\bar\sigma_1 z(\boldsymbol{x}^{(1)}) + \bar\sigma_1^2 s(\boldsymbol{x}^{(1)}) + \bar\sigma_2 z(\boldsymbol{x}^{(2)}) + \bar\sigma_2^2 s(\boldsymbol{x}^{(2)}) \right) \right\rangle_{n,\epsilon} .
\end{aligned}$$

The next lemma follows from the simple fact that for $N \sim \mathcal{N}(0,1)$ and $t \in \mathbb{R}$, $\mathbb{E}e^{tN} = \exp(\frac{t^2}{2})$.

Lemma 3.4.4

Let $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S^n$ and $\sigma_1, \sigma_2 \in S$ be fixed. Then

$$\mathbb{E}_{\boldsymbol{Z}'} \exp\left(\sigma_1 \sqrt{\frac{\lambda}{n}} \sum_{i=1}^n x_i^{(1)} Z_i' + \sigma_2 \sqrt{\frac{\lambda}{n}} \sum_{i=1}^n x_i^{(2)} Z_i' \right) = \exp\left(\lambda \sigma_1 \sigma_2 \, \boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} + \frac{\lambda \sigma_1^2}{2n} \|\boldsymbol{x}^{(1)}\|^2 + \frac{\lambda \sigma_2^2}{2n} \|\boldsymbol{x}^{(2)}\|^2 \right).$$

Thus, for all $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S^n$ and $\sigma_1, \sigma_2 \in S$

$$\mathbb{E}_{\boldsymbol{Z}'} e^{\bar\sigma_1 z(\boldsymbol{x}^{(1)}) + \bar\sigma_1^2 s(\boldsymbol{x}^{(1)}) + \bar\sigma_2 z(\boldsymbol{x}^{(2)}) + \bar\sigma_2^2 s(\boldsymbol{x}^{(2)})} = e^{\lambda \bar\sigma_1 \bar\sigma_2 \boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} + \lambda X' (\bar\sigma_1 (\boldsymbol{x}^{(1)} \cdot \boldsymbol{X}) + \bar\sigma_2 (\boldsymbol{x}^{(2)} \cdot \boldsymbol{X}))},$$

where we used the fact that $s(\boldsymbol{x}) = -\frac{\lambda}{2n}\|\boldsymbol{x}\|^2$ for all $\boldsymbol{x} \in S^n$. We have therefore

$$\mathbb{E}_{\boldsymbol{Z}'} U^2 = \left\langle \sum_{\sigma_1,\sigma_2 \in S} P_0(\sigma_1,\sigma_2) \exp\left(\lambda \bar\sigma_1 \bar\sigma_2 \boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} + \lambda X' \left(\bar\sigma_1 (\boldsymbol{x}^{(1)} \cdot \boldsymbol{X}) + \bar\sigma_2 (\boldsymbol{x}^{(2)} \cdot \boldsymbol{X}) \right) \right) \right\rangle_{n,\epsilon} .$$

Define

$$g : (s, r_1, r_2) \in [-K^2, K^2]^3 \mapsto \sum_{\sigma_1,\sigma_2 \in S} P_0(\sigma_1,\sigma_2) \exp\left(\lambda \bar\sigma_1 \bar\sigma_2 s + \lambda X' (\bar\sigma_1 r_1 + \bar\sigma_2 r_2) \right).$$

We have $\mathbb{E}_{\boldsymbol{Z}'} U^2 = \big\langle g(\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)}, \boldsymbol{x}^{(1)} \cdot \boldsymbol{X}, \boldsymbol{x}^{(2)} \cdot \boldsymbol{X}) \big\rangle_{n,\epsilon}$ and by (3.4.13), $\mathbb{E}_{\boldsymbol{Z}'} V^2 = g(Q,Q,Q)$.

Lemma 3.4.5

There exists a constant M that only depends on λ and K, such that g is almost surely M-Lipschitz.

Proof. g is a random function that depends only on the random variables X' and L_{n+1} (because of $\bar\sigma_1$ and $\bar\sigma_2$). g is $\mathcal{C}^1$ on the compact $[-K^2, K^2]^3$. An easy computation show that

$$\forall (s, r_1, r_2) \in [-K^2, K^2]^3, \ \ \|\nabla g(s, r_1, r_2)\| \le 3\lambda K^4 \exp(3\lambda K^4).$$

g is thus M-Lipschitz with $M = 3\lambda K^4 \exp(3\lambda K^4)$. □

Using Lemma 3.4.5 we obtain

$$\begin{aligned}
&\Big\langle |g(\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)}, \boldsymbol{x}^{(1)} \cdot \boldsymbol{X}, \boldsymbol{x}^{(2)} \cdot \boldsymbol{X}) - g(Q,Q,Q)| \Big\rangle_{n,\epsilon} \\
&\qquad \le M \Big\langle \sqrt{(\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} - Q)^2 + (\boldsymbol{x}^{(1)} \cdot \boldsymbol{X} - Q)^2 + (\boldsymbol{x}^{(2)} \cdot \boldsymbol{X} - Q)^2} \Big\rangle_{n,\epsilon} .
\end{aligned}$$

We recall equation (3.4.13) to notice that $g(Q,Q,Q) = \mathbb{E}_{\boldsymbol{Z}'} V^2$. Thus, using (3.4.9) and (3.4.10)

$$\int_0^{\epsilon_0} d\epsilon\, \mathbb{E}|\mathbb{E}_{\boldsymbol{Z}'} U^2 - \mathbb{E}_{\boldsymbol{Z}'} V^2| \leq M \int_0^{\epsilon_0} d\epsilon\, \mathbb{E}\Big\langle \sqrt{(\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)} - Q)^2 + (\boldsymbol{x}^{(1)}\cdot\boldsymbol{X} - Q)^2 + (\boldsymbol{x}^{(2)}\cdot\boldsymbol{X} - Q)^2}\Big\rangle_{n,\epsilon},$$

and the right-hand side goes to 0 by (3.4.9-3.4.10).

Showing that $\int d\epsilon\, |\mathbb{E}UV - \mathbb{E}V^2| \xrightarrow[n\to\infty]{} 0$ goes exactly the same way. We thus omit this part here for the sake of brevity, but the reader can refer to [132] where all details are presented. □

Using the fact that $|\log U - \log V| \leq \max(U^{-1}, V^{-1})|U - V|$ and the Cauchy-Schwarz inequality, we have

$$\mathbb{E}|\log U - \log V| \leq \sqrt{\mathbb{E}U^{-2} + \mathbb{E}V^{-2}}\sqrt{\mathbb{E}(U-V)^2}\,.$$

Lemma 3.4.6

There exists a constant C that depends only on λ and K such that

$$\mathbb{E}U^{-2} + \mathbb{E}V^{-2} \leq C\,.$$

Proof. Using Jensen inequality, we have $U \geq f\big(\langle z(\boldsymbol{x})\rangle_{n,\epsilon}, \langle s(\boldsymbol{x})\rangle_{n,\epsilon}\big)$. Then

$$U^{-2} \leq f\big(\langle z(\boldsymbol{x})\rangle_{n,\epsilon}, \langle s(\boldsymbol{x})\rangle_{n,\epsilon}\big)^{-2} \leq \sum_{\sigma\in S} P_0(\sigma)\exp\big(-2\bar\sigma\langle z(\boldsymbol{x})\rangle_{n,\epsilon} - 2\bar\sigma^2\langle s(\boldsymbol{x})\rangle_{n,\epsilon}\big)\,.$$

It remains to bound $\mathbb{E}\exp(-2\bar\sigma\langle z(\boldsymbol{x})\rangle_{n,\epsilon} - 2\bar\sigma^2\langle s(\boldsymbol{x})\rangle_{n,\epsilon})$. P_0 has a bounded support, therefore

$$\mathbb{E}\exp\big(-2\bar\sigma\langle z(\boldsymbol{x})\rangle_{n,\epsilon} - 2\bar\sigma^2\langle s(\boldsymbol{x})\rangle_{n,\epsilon}\big) \leq C_0\mathbb{E}\exp\left(-2\bar\sigma\sum_{i=1}^n\sqrt{\frac{\lambda}{n}}\langle x_i\rangle_{n,\epsilon}Z_i'\right) = C_0\mathbb{E}\exp(2\lambda Q\bar\sigma^2) \leq C_1\,,$$

for some constant C_0, C_1 depending only on λ and K. Similar arguments show that $\mathbb{E}V^{-2}$ is upper-bounded by a constant. □

Using the previous lemma we obtain $\int d\epsilon\, \mathbb{E}|\log U - \log V| \xrightarrow[n\to\infty]{} 0$. We now compute $\mathbb{E}\log V$ explicitly.

Lemma 3.4.7

$$\mathbb{E}\log V = \epsilon(\mathbb{E}_{P_0}X^2)\mathbb{E}\frac{\lambda Q}{2} + (1-\epsilon)\mathbb{E}\log\sum_{\sigma\in S}P_0(\sigma)\exp\left(\sigma\sqrt{\frac{\lambda}{n}}\sum_{i=1}^n b_iZ_i' + \lambda Q\sigma X' - \frac{\lambda\sigma^2}{2}Q\right).$$

Proof. It suffices to distinguish the cases $L_{n+1}=0$ and $L_{n+1}=1$. If $L_{n+1}=1$ then for all $\sigma\in S$, $\bar\sigma = X'$ and

$$\log V = \log\left(\exp\Big(X'\sqrt{\frac{\lambda}{n}}\sum_{i=1}^n b_iZ_i' + \lambda QX'^2 - \frac{\lambda X'^2}{2}Q\Big)\right) = X'\sqrt{\frac{\lambda}{n}}\sum_{i=1}^n b_iZ_i' + \frac{\lambda X'^2}{2}Q\,.$$

L_{n+1} is independent of all other random variables, thus

$$\mathbb{E}\Big[\mathbb{1}(L_{n+1}=1)\log V\Big] = \epsilon(\mathbb{E}_{P_0}X^2)\frac{\lambda}{2}\mathbb{E}Q\,,$$

because the Z_i' are centered, independent from X' and because X' is independent from Q. The case $L_{n+1} = 0$ is obvious. □

The variables $(b_i)_{1\leq i\leq n}$ and $(Z_i')_{1\leq i\leq n}$ are independent. Recall that $Q = \frac{1}{n}\sum_{i=1}^{n} b_i^2$. Therefore,

$$\left(X', Q, \frac{1}{\sqrt{n}}\sum_{i=1}^{n} b_i Z_i'\right) = \left(X', Q, \sqrt{Q}Z_0\right) \quad \text{in law,}$$

where $Z_0 \sim \mathcal{N}(0,1)$ is independent of Q, X'. The expression of $\mathbb{E}\log V$ from Lemma 3.4.7 simplifies

$$\mathbb{E}\log V = \epsilon(\mathbb{E}_{P_0}X^2)\mathbb{E}\frac{\lambda Q}{2} + (1-\epsilon)\mathbb{E}\log\sum_{\sigma\in S} P_0(\sigma)\exp\left(\sqrt{\lambda Q}\sigma Z_0 + \lambda Q\sigma X' - \frac{\lambda\sigma^2}{2}Q\right),$$

thus

$$\int_0^{\epsilon_0} d\epsilon \left|\mathbb{E}\log U - \left(\epsilon(\mathbb{E}_{P_0}X^2)\mathbb{E}\frac{\lambda Q}{2} + (1-\epsilon)\mathbb{E}\log\sum_{\sigma\in S} P_0(\sigma)\exp\left(\sqrt{\lambda Q}\sigma Z_0 + \lambda Q\sigma X' - \frac{\lambda\sigma^2}{2}Q\right)\right)\right| \xrightarrow[n\to\infty]{} 0\,,$$

which is precisely the statement of Lemma 3.4.1.

3.4.7 Reduction to distribution with finite support

We will show in this section that it suffices to prove Theorems 3.2.1 for input distribution P_0 with finite support.

Suppose the Theorem 3.2.1 holds for all prior distributions over $\mathbb{R}$ with finite support. Let P_0 be a probability distribution that admits a finite second moment: $\mathbb{E}_{P_0}X^2 < \infty$. We are going to approach P_0 with distributions with finite supports.
Let $0 < \epsilon \leq 1$. Let $K > 0$ such that $\mathbb{E}_{P_0}[X^2\mathbb{1}(|X| \geq K)] \leq \epsilon^2$. Let $m \in \mathbb{N}^*$ such that $\frac{K}{m} \leq \epsilon$. For $x \in \mathbb{R}$ we will use the notation

$$\bar{x} = \begin{cases} \frac{K}{m}\left\lfloor \frac{xm}{K}\right\rfloor & \text{if } x \in [-K, K]\,, \\ 0 & \text{otherwise.} \end{cases}$$

Consequently if $x \in [-K, K]$, $\bar{x} \leq x < \bar{x} + \frac{K}{m} \leq \bar{x} + \epsilon$. We define $\bar{P}_0$ the image distribution of P_0 through the application $x \mapsto \bar{x}$. Let $n \geq 1$. We will note $\bar{F}_n$ the free energy corresponding to the distribution $\bar{P}_0$ and $\bar{\mathcal{F}}$ the function $\mathcal{F}$ from (3.2.5) corresponding to the distribution $\bar{P}_0$. $\bar{P}_0$ has a finite support, we have then by assumptions

$$\bar{F}_n(\lambda) \xrightarrow[n\to\infty]{} \sup_{q\geq 0} \bar{\mathcal{F}}(\lambda, q)\,. \tag{3.4.14}$$

By construction we have for all $1 \leq i \leq n$, $\mathbb{E}(X_i - \bar{X}_i)^2 \leq \epsilon^2$. Hence

$$\mathbb{E}\left\|(X_iX_j)_{i<j} - (\bar{X}_i\bar{X}_j)_{i<j}\right\|^2 \leq 2(n-1)n\mathbb{E}_{P_0}[X^2]\epsilon^2\,.$$

Consequently, by "pseudo-Lipschitz" continuity of the free energy with respect to the Wasserstein metric (see Proposition B.1 in Appendix B.4) there exist a constant $C > 0$ depending only on P_0, such that, for all $n \geq 1$ and all $\lambda \geq 0$,

$$|F_n(\lambda) - \bar{F}_n(\lambda)| \leq \lambda C\epsilon\,. \tag{3.4.15}$$

Lemma 3.4.8

There exists a constant $C' > 0$ that depends only on P_0, such that

$$\left| \sup_{q \geq 0} \mathcal{F}(\lambda, q) - \sup_{q \geq 0} \bar{\mathcal{F}}(\lambda, q) \right| \leq \lambda C' \epsilon \,.$$

Proof. First notice that both suprema are achieved over a common compact set $[0, \mathbb{E}_{P_0}[X^2 + \bar{X}^2]]$. Indeed, for $q \geq 0$,

$$\frac{\partial}{\partial q} \mathcal{F}(\lambda, q) = \lambda \psi'_{P_0}(\lambda q) - \frac{\lambda q}{2} \leq \frac{\lambda}{2} (\mathbb{E}_{P_0}[X^2] - q)$$

because ψ_{P_0} is $\frac{1}{2}\mathbb{E}_{P_0}[X^2]$-Lipschitz by Proposition 1.3.3. Consequently, the maximum of $\mathcal{F}(\lambda, \cdot)$ is achieved on $[0, \mathbb{E}_{P_0}[X^2]]$ and similarly the supremum of $\bar{\mathcal{F}}(\lambda, \cdot)$ is achieved over $[0, \mathbb{E}_{P_0}[\bar{X}^2]]$. Using Proposition B.1 in Appendix B.4, we obtain that there exists a constant C' depending only on P_0 such that $\forall q \in [0, \mathbb{E}_{P_0}[X^2 + \bar{X}^2]]$, $|\mathcal{F}(\lambda, q) - \bar{\mathcal{F}}(\lambda, q)| \leq \lambda C' \epsilon$. The lemma follows. □

Combining Equation 3.4.14 and 3.4.15 and Lemma 3.4.8, we obtain that there exists $n_0 \geq 1$ such that for all $n \geq n_0$,

$$|F_n - \sup_{q \geq 0} \mathcal{F}(\lambda, q)| \leq \lambda (C + C' + 1) \epsilon \,,$$

where C and C' are two constants that only depend on P_0. This proves Theorem 3.2.1.

Chapter 4

Non-symmetric low-rank matrix estimation

We consider now the spiked Wishart model (3.1.2). Let P_U and P_V be two probability distributions on $\mathbb{R}$ with finite second moment. We assume that $\mathrm{Var}_{P_U}(U), \mathrm{Var}_{P_V}(V) > 0$. Let $n, m \geq 1$, $\lambda > 0$ and consider $\boldsymbol{U} = (U_1, \dots, U_n) \overset{\text{i.i.d.}}{\sim} P_U$ and $\boldsymbol{V} = (V_1, \dots, V_m) \overset{\text{i.i.d.}}{\sim} P_V$, independent from each other. Suppose that we observe

$$Y_{i,j} = \sqrt{\frac{\lambda}{n}} U_i V_j + Z_{i,j}, \qquad \text{for} \quad 1 \leq i \leq n \quad \text{and} \quad 1 \leq j \leq m\,, \tag{4.0.1}$$

where $(Z_{i,j})_{i,j}$ are i.i.d. standard normal random variables, independent from $\boldsymbol{U}$ and $\boldsymbol{V}$. In the following, $\mathbb{E}$ will denote the expectation with respect to the variables $(\boldsymbol{U}, \boldsymbol{V})$ and $\boldsymbol{Z}$. We define the Minimum Mean Squared Error (MMSE) for the estimation of the matrix $\boldsymbol{U}\boldsymbol{V}^\intercal$ given the observation of the matrix $\boldsymbol{Y}$:

$$\mathrm{MMSE}_n(\lambda) = \min_{\widehat{\theta}} \left\{ \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \mathbb{E}\left[\left(U_i V_j - \widehat{\theta}_{i,j}(\boldsymbol{Y}) \right)^2 \right] \right\} = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \mathbb{E}\left[\left(U_i V_j - \mathbb{E}\left[U_i V_j | \boldsymbol{Y} \right] \right)^2 \right],$$

where the minimum is taken over all estimators $\widehat{\theta}$ (i.e. measurable functions of the observations $\boldsymbol{Y}$). In order to get an upper bound on the MMSE, let us consider the "dummy estimator" $\widehat{\theta}_{i,j} = \mathbb{E}[U_i V_j]$ for all i, j which achieves a "dummy" matrix Mean Squared Error of:

$$\mathrm{DMSE} = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \mathbb{E}\left[\left(U_i V_j - \mathbb{E}[U_i V_j] \right)^2 \right] = \mathbb{E}[U^2]\mathbb{E}[V^2] - (\mathbb{E}U)^2 (\mathbb{E}V)^2\,.$$

4.1 Fundamental limits of estimation

As in Chapter 3, we investigate the posterior distribution of $\boldsymbol{U}, \boldsymbol{V}$ given $\boldsymbol{Y}$. We define the Hamiltonian

$$H_n(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i,j} \sqrt{\frac{\lambda}{n}} u_i v_j Z_{i,j} + \frac{\lambda}{n} u_i U_i v_j V_j - \frac{\lambda}{2n} u_i^2 v_j^2, \quad \text{for } (\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{R}^n \times \mathbb{R}^m. \tag{4.1.1}$$

The posterior distribution of $(\boldsymbol{U}, \boldsymbol{V})$ given $\boldsymbol{Y}$ is then

$$dP\Big(\boldsymbol{u}, \boldsymbol{v} \,\Big|\, \boldsymbol{Y}\Big) = \frac{1}{\mathcal{Z}_n(\lambda)} e^{H_n(\boldsymbol{u}, \boldsymbol{v})} dP_U^{\otimes n}(\boldsymbol{u}) dP_V^{\otimes m}(\boldsymbol{v}), \tag{4.1.2}$$

where $\mathcal{Z}_n(\lambda) = \int e^{H_n(\boldsymbol{u},\boldsymbol{v})} dP_U^{\otimes n}(\boldsymbol{u}) dP_V^{\otimes m}(\boldsymbol{v})$ is the appropriate normalization. The corresponding free energy is

$$F_n(\lambda) = \frac{1}{n}\mathbb{E}\log \mathcal{Z}_n(\lambda) = \frac{1}{n}\mathbb{E}\log\left(\int e^{H_n(\boldsymbol{u},\boldsymbol{v})} dP_U^{\otimes n}(\boldsymbol{u}) dP_V^{\otimes m}(\boldsymbol{v})\right).$$

We consider here the high-dimensional limit where $n, m \to \infty$, while $m/n \to \alpha > 0$. We will be interested in the following fixed point equations, sometimes called "state evolution equations".

Definition 4.1.1

We define the set $\Gamma(\lambda, \alpha)$ *as*

$$\Gamma(\lambda,\alpha) = \left\{(q_u, q_v) \in \mathbb{R}^2_{\geq 0} \;\middle|\; q_u = 2\psi'_{P_U}(\lambda\alpha q_v) \text{ and } q_v = 2\psi'_{P_V}(\lambda q_u)\right\}. \tag{4.1.3}$$

First notice that $\Gamma(\lambda, \alpha)$ is not empty. The function $f : (q_u, q_v) \mapsto (2\psi'_{P_U}(\lambda\alpha q_v), 2\psi'_{P_V}(\lambda q_u))$ is continuous from the convex compact set $[0, \mathbb{E}U^2] \times [0, \mathbb{E}V^2]$ into itself (see Proposition 1.3.3). Brouwer's Theorem gives the existence of a fixed point of f: $\Gamma(\lambda, \alpha) \neq \emptyset$.

We will express the limit of F_n using the following function

$$\mathcal{F} : (\lambda, \alpha, q_u, q_v) \mapsto \psi_{P_U}(\lambda\alpha q_v) + \alpha\psi_{P_V}(\lambda q_u) - \frac{\lambda\alpha}{2} q_u q_v \,. \tag{4.1.4}$$

Recall that ψ_{P_U} and ψ_{P_V}, defined by (1.3.6), are the free energies of additive Gaussian scalar channels (1.3.5) with priors P_U and P_V. The Replica-Symmetric formula states that the free energy F_n converges to the supremum of $\mathcal{F}$ over $\Gamma(\lambda, \alpha)$.

Theorem 4.1.1 (*Replica-Symmetric formula for the spiked Wishart model*)

$$F_n(\lambda) \xrightarrow[n\to\infty]{} \sup_{(q_u,q_v)\in\Gamma(\lambda,\alpha)} \mathcal{F}(\lambda, \alpha, q_u, q_v) = \sup_{q_u\geq 0} \inf_{q_v\geq 0} \mathcal{F}(\lambda, \alpha, q_u, q_v) \,. \tag{4.1.5}$$

Moreover, these extrema are achieved over the same couples $(q_u, q_v) \in \Gamma(\lambda, \alpha)$.

This result proved in [145] was conjectured by [133], in particular $\mathcal{F}$ corresponds to the "Bethe free energy" [133, Equation 47]. Theorem 4.1.1 is proved in Section 4.3. For the rank-k case (where P_U and P_V are probability distributions over $\mathbb{R}^k$), see [145]. As in Chapter 3, the Replica-Symmetric formula (Theorem 3.2.1) allows to compute the limit of the MMSE.

Proposition 4.1.1 (*Limit of the* MMSE)

Let

$$D_\alpha = \left\{\lambda > 0 \;\middle|\; \mathcal{F}(\lambda, \alpha, \cdot, \cdot) \text{ has a unique maximizer } (q_u^*(\lambda,\alpha), q_v^*(\lambda,\alpha)) \text{ over } \Gamma(\lambda,\alpha)\right\}.$$

Then D_α *is equal to* $(0, +\infty)$ *minus a countable set and for all* $\lambda \in D_\alpha$ *(and thus almost every* $\lambda > 0$*)*

$$\mathrm{MMSE}_n(\lambda) \xrightarrow[n\to\infty]{} \mathbb{E}[U^2]\mathbb{E}[V^2] - q_u^*(\lambda,\alpha) q_v^*(\lambda,\alpha) \,. \tag{4.1.6}$$

Again, this was conjectured in [133]: the performance of the Bayes-optimal estimator (i.e. the MMSE) corresponds to the fixed point of the state-evolution equations (4.1.3) which has the greatest Bethe free energy $\mathcal{F}$. Proposition 4.1.1 follows from the same kind of arguments than Corollary 3.2.2 so we omit its proof for the sake of brevity.

Proposition 4.1.1 allows to locate the information-theoretic threshold for our matrix estimation problem. Let us define

$$\lambda_c(\alpha) = \inf\left\{\lambda \in D_\alpha \mid q_u^*(\lambda,\alpha)q_v^*(\lambda,\alpha) > (\mathbb{E}U)^2(\mathbb{E}V)^2\right\}. \tag{4.1.7}$$

If the set of the left-hand side is empty, one defines $\lambda_c(\alpha) = 0$. Proposition 4.1.1 gives that $\lambda_c(\alpha)$ is the information-theoretic threshold for the estimation of $\boldsymbol{U}\boldsymbol{V}^\intercal$ given $\boldsymbol{Y}$:

- If $\lambda < \lambda_c(\alpha)$, then $\mathrm{MMSE}_n(\lambda) \xrightarrow[n\to\infty]{} \mathrm{DMSE}$. It is not possible to reconstruct the signal $\boldsymbol{U}\boldsymbol{V}^\intercal$ better than a "dummy" estimator.
- If $\lambda > \lambda_c(\alpha)$, then $\lim_{n\to\infty} \mathrm{MMSE}_n(\lambda) < \mathrm{DMSE}$. It is possible to reconstruct the signal $\boldsymbol{U}\boldsymbol{V}^\intercal$ better than a "dummy" estimator.

Proposition 4.1.1 gives us the limit of the MMSE for the estimation of the matrix $\boldsymbol{U}\boldsymbol{V}^\intercal$, but does not gives us the minimal error for the estimation of $\boldsymbol{U}$ or $\boldsymbol{V}$ separately. As we will see in the next section with the spiked covariance model, one can be interested in estimating $\boldsymbol{U}\boldsymbol{U}^\intercal$ or $\boldsymbol{V}\boldsymbol{V}^\intercal$, only. Let us define:

$$\mathrm{MMSE}_n^{(u)}(\lambda) = \frac{1}{n^2}\mathbb{E}\Big[\sum_{1\le i,j\le n}\Big(U_iU_j - \mathbb{E}[U_iU_j|\boldsymbol{Y}]\Big)^2\Big],$$
$$\mathrm{MMSE}_n^{(v)}(\lambda) = \frac{1}{m^2}\mathbb{E}\Big[\sum_{1\le i,j\le m}\Big(V_iV_j - \mathbb{E}[V_iV_j|\boldsymbol{Y}]\Big)^2\Big].$$

Theorem 4.1.2

For all $\alpha > 0$ and all $\lambda \in D_\alpha$

$$\mathrm{MMSE}_n^{(u)}(\lambda) \xrightarrow[n\to\infty]{} \mathbb{E}_{P_U}[U^2]^2 - q_u^*(\lambda,\alpha)^2 \quad \textit{and} \quad \mathrm{MMSE}_n^{(v)}(\lambda) \xrightarrow[n\to\infty]{} \mathbb{E}_{P_V}[V^2]^2 - q_v^*(\lambda,\alpha)^2\,.$$

Theorem 4.1.2 is proved in Section 4.4.

4.2 Application to the spiked covariance model

Let us consider now the so-called spiked covariance model. Let $\boldsymbol{U} = (U_1,\dots,U_n) \overset{\text{i.i.d.}}{\sim} P_U$, where P_U is a distribution over $\mathbb{R}$ with finite second moment. Define the "spiked covariance matrix"

$$\boldsymbol{\Sigma} = \mathrm{Id}_n + \frac{\lambda}{n}\boldsymbol{U}\boldsymbol{U}^\intercal, \tag{4.2.1}$$

and suppose that we observe $\boldsymbol{Y}_1,\dots,\boldsymbol{Y}_m \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\boldsymbol{0},\boldsymbol{\Sigma})$, conditionally on $\boldsymbol{\Sigma}$. Given the matrix $\boldsymbol{Y} = (\boldsymbol{Y}_1|\cdots|\boldsymbol{Y}_m)$, one would like to estimate the "spike" $\boldsymbol{U}\boldsymbol{U}^\intercal$. We deduce from Theorem 4.1.2 above the minimal mean squared error for this task, in the asymptotic regime where $n, m \to +\infty$ and $m/n \to \alpha > 0$.

Corollary 4.2.1

For all $\alpha > 0$, the function

$$q \mapsto \left\{ \psi_{P_U}(\lambda\alpha q) + \frac{\alpha}{2}\big(q + \log(1-q)\big) \right\}$$

admits for almost all $\lambda > 0$ a unique maximizer $q^*(\lambda, \alpha)$ on $[0, 1)$ and

$$\mathrm{MMSE}_n^{(u)}(\lambda) = \frac{1}{n^2}\mathbb{E}\Big[\big\|\boldsymbol{U}\boldsymbol{U}^\intercal - \mathbb{E}[\boldsymbol{U}\boldsymbol{U}^\intercal|\boldsymbol{Y}]\big\|^2\Big] \xrightarrow[n\to\infty]{} \mathbb{E}_{P_U}[U^2]^2 - \left(\frac{q^*(\lambda,\alpha)}{\lambda(1-q^*(\lambda,\alpha))}\right)^2 .$$

Proof. There exists independent Gaussian random variables $\boldsymbol{V} = (V_1, \dots, V_m) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and $Z_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, independent from $\boldsymbol{U}$ such that

$$\boldsymbol{Y} = (\boldsymbol{Y}_1|\cdots|\boldsymbol{Y}_m) = \sqrt{\frac{\lambda}{n}}\boldsymbol{U}\boldsymbol{V}^\intercal + \boldsymbol{Z}\,.$$

Therefore, the limit of the MMSE for the estimation of $\boldsymbol{U}\boldsymbol{U}^\intercal$ is given by Theorem 4.1.2 above. It remains only to evaluate the formulas of Theorems 4.1.1 and 4.1.2 in the case $P_V = \mathcal{N}(0,1)$. As computed in Example 1.3.1, $\psi_{\mathcal{N}(0,1)}(q) = \frac{1}{2}(q - \log(1+q))$. Thus, the limit of the free energy (4.1.5) becomes (after evaluation of the supremum in q_u):

$$\sup_{q_v \in [0,1)} \left\{ \psi_{P_U}(\lambda\alpha q_v) + \frac{\alpha}{2}(q_v + \log(1-q_v)) \right\}.$$

By Theorem 4.1.2 for all $\alpha > 0$ and almost all $\lambda > 0$ this supremum admits a unique maximizer $q_v^*(\lambda,\alpha)$ and $\mathrm{MMSE}_n^{(u)}(\lambda) \to \mathbb{E}_{P_U}[U^2]^2 - q_u^*(\lambda,\alpha)^2$ where q_u^* verifies (recall that $(q_u^*, q_v^*) \in \Gamma$):

$$q_v^*(\lambda,\alpha) = 2\psi'_{\mathcal{N}(0,1)}(\lambda q_u^*(\lambda,\alpha)) = \frac{\lambda q_u^*(\lambda,\alpha)}{1+\lambda q_u^*(\lambda,\alpha)}\,.$$

We deduce from the equation above that $q_u^*(\lambda,\alpha) = \frac{q_v^*(\lambda,\alpha)}{\lambda(1-q_v^*(\lambda,\alpha))}$, which concludes the proof. □

We will now compare the MMSE given by Corollary 4.2.1 to the mean squared errors achieved by PCA and Approximate Message Passing (AMP).

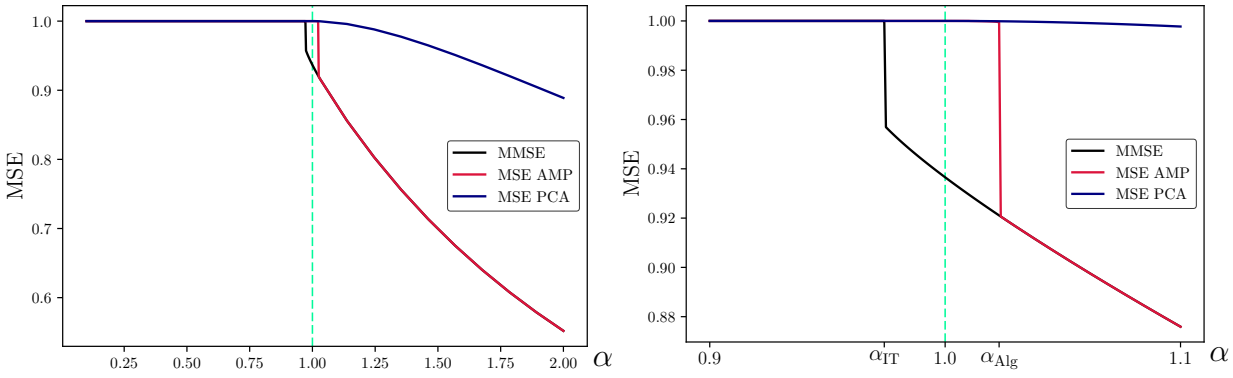


Figure 4.1: Mean Squared Errors for the spiked covariance model, where the spike is generated by (4.2.3) with $s = 0.15$, $\lambda = 1$. The right-hand side panel is a zoom of the left-hand side panel around $\alpha = 1$.

Let $\widehat{\boldsymbol{u}}$ be a singular vector of $\boldsymbol{Y}/\sqrt{n}$ associated with σ_1, the top singular value of $\boldsymbol{Y}/\sqrt{n}$, such that $\|\widehat{\boldsymbol{u}}\| = \sqrt{n}$. Then results from [29, 66] give that almost surely:

$$\lim_{n\to\infty} \left(\widehat{\boldsymbol{u}} \cdot \boldsymbol{U}\right)^2 = \begin{cases} \frac{\lambda^2\alpha - 1}{\lambda(\lambda\alpha+1)} & \text{if } \lambda^2\alpha \geq 1\,, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad \lim_{n\to\infty} \sigma_1 = \begin{cases} \sqrt{\frac{(1+\lambda)(\alpha^{-1}+\lambda)}{\lambda}} & \text{if } \lambda^2\alpha \geq 1\,, \\ 1 + 1/\sqrt{\alpha} & \text{otherwise.} \end{cases}$$

We are then going to estimate $\boldsymbol{U}\boldsymbol{U}^\intercal$ using $\widehat{\theta}^{\mathrm{PCA}} = \delta\widehat{\boldsymbol{u}}\widehat{\boldsymbol{u}}^\intercal$, where δ is chosen in order to minimize the mean squared error. The optimal choice of δ is $\delta^* = \left(\frac{\lambda^2\alpha-1}{\lambda(\lambda\alpha+1)}\right)_+$, which can be estimated using σ_1. We obtain the mean squared error of the spectral estimator $\widehat{\theta}^{\mathrm{PCA}}$:

$$\lim_{n\to\infty} \mathrm{MSE}_n^{\mathrm{PCA}} = \begin{cases} \frac{1+\lambda}{\lambda(\lambda\alpha+1)}\left(2 - \frac{1+\lambda}{\lambda(\lambda\alpha+1)}\right) & \text{if } \lambda^2\alpha \geq 1\,, \\ 1 & \text{otherwise.} \end{cases}$$

As in the symmetric case (see Section 3.3.1) one can define an Approximate Message Passing (AMP) algorithm to estimate $\boldsymbol{U}\boldsymbol{U}^\intercal$. For a precise description of the algorithm, see [177, 61, 134]. The MSE achieved by AMP after t iterations is:

$$\lim_{n\to\infty} \mathrm{MSE}_n^{\mathrm{AMP}} = 1 - \left(q_u^t\right)^2,$$

where q_u^t is given by the recursion:

$$\begin{cases} q_u^t = 2\psi'_{P_U}(\lambda\alpha q_v^t) \\ q_v^{t+1} = 2\psi'_{P_V}(\lambda q_u^t)\,, \end{cases} \tag{4.2.2}$$

with initialization $(q_u^0, q_v^0) = (0,0)$. We know by Proposition 1.3.3 that the functions ψ'_{P_U} and ψ'_{P_V} are both non-decreasing and bounded. This ensures that $(q_u^t, q_v^t)_{t\geq 0}$ converges as $t \to \infty$ to some fixed point $(q_u^{\mathrm{AMP}}, q_v^{\mathrm{AMP}}) \in \Gamma$. If this fixed point turns out to be the one that maximizes $\mathcal{F}(\lambda, \alpha, \cdot, \cdot)$, i.e. that $(q_u^{\mathrm{AMP}}, q_v^{\mathrm{AMP}}) = (q_u^*(\lambda,\alpha), q_v^*(\lambda,\alpha))$, then AMP achieves the minimal mean squared error!

For the plots of Figure 4.1, we consider a case where the signal is sparse:

$$P_U = s\mathcal{N}(0, 1/s) + (1-s)\delta_0\,, \tag{4.2.3}$$

for some $s \in (0,1]$, so that $\mathbb{E}_{P_U}[U^2] = 1$. We plot the different MSE on Figure 4.1. We chose $\lambda = 1$ so the "spectral threshold" (the minimal value of α for which PCA performs better than a random guess) it at $\alpha = 1$ (green dashed line). This threshold corresponds also to the threshold for AMP: $\mathrm{MSE}^{\mathrm{AMP}} = 1$ for $\alpha < 1$ while $\mathrm{MSE}^{\mathrm{AMP}} < 1$ for $\alpha > 1$. The information-theoretic threshold α_{IT} is however strictly less than 1. For $\alpha \in (\alpha_{\mathrm{IT}}, 1)$ inference is "hard": it is information-theoretically possible to achieve a MSE strictly less than 1, but PCA and AMP fail (and it is conjectured that any polynomial-time algorithm will also fail).

However, even for $\alpha > 1$, AMP does not always succeed to reach the MMSE. For $\alpha \in (1, \alpha_{\mathrm{Alg}})$, $\mathrm{MSE}^{\mathrm{AMP}}$ is strictly less than 1 but is still very bad. So, the region $\alpha \in (1, \alpha_{\mathrm{Alg}})$ is also a "hard region" in the sense that achieving the MMSE seems impossible for polynomial-time algorithms (under the conjecture that AMP is optimal among polynomial-time algorithms). The scenario presented on Figure 4.1 is not the only one possible: various cases have been studied in great details in [135]. See in particular Figure 6 from [135] and the phase diagrams of Figure 7 and 8.

4.3 Proof of the Replica-Symmetric formula (Theorem 4.1.1)

4.3.1 Proof ideas

The proof of the Replica formula for the non-symmetric case is a little bit more involved compared to the symmetric case, because one can not use the convexity argument of Proposition 3.4.1 to obtain the lower bound. Indeed, a key step in the proof of Proposition 3.4.1 was the inequality (3.4.3) that was obtained by saying that for every $q \geq 0$

$$\mathbb{E}\Big[(\boldsymbol{x} \cdot \boldsymbol{X} - q)^2\Big] \geq 0\,, \tag{4.3.1}$$

where $\boldsymbol{x}$ is a sample from the posterior distribution of $\boldsymbol{X}$ given some observations (we omit the notation's details here in order to focus on the main ideas).

However, if we apply the strategy of Proposition 3.4.1 to the non-symmetric case, one obtains

$$\mathbb{E}\Big[(\boldsymbol{u} \cdot \boldsymbol{U} - q_u)(\boldsymbol{v} \cdot \boldsymbol{V} - q_v)\Big] \tag{4.3.2}$$

where $(\boldsymbol{u}, \boldsymbol{v})$ is a sample from the posterior distribution of $(\boldsymbol{U}, \boldsymbol{V})$ given some observations, instead of (4.3.1). Now, it not obvious anymore that (4.3.2) is non-negative. In order to prove it, one has to investigate further the distributions of the overlaps $\boldsymbol{u} \cdot \boldsymbol{U}$ and $\boldsymbol{v} \cdot \boldsymbol{V}$. By following the approach used by Talagrand in [201] to prove the TAP equations (discovered by Thouless, Anderson and Palmer in [203]) for the Sherrington-Kirkpatrick model, one can show that the overlaps approximately satisfy (when n and m are large)

$$\begin{cases} \boldsymbol{u} \cdot \boldsymbol{U} \;\simeq\; 2\psi'_{P_U}(\lambda\alpha \boldsymbol{v} \cdot \boldsymbol{V}) \\ \boldsymbol{v} \cdot \boldsymbol{V} \;\simeq\; 2\psi'_{P_V}(\lambda \boldsymbol{u} \cdot \boldsymbol{U})\,. \end{cases}$$

These are precisely the fixed point equations verified by $(q_u, q_v) \in \Gamma(\lambda, \alpha)$. Thus one has

$$\mathbb{E}\Big[(\boldsymbol{u} \cdot \boldsymbol{U} - q_u)(\boldsymbol{v} \cdot \boldsymbol{V} - q_v)\Big] \simeq \mathbb{E}\Big[(2\psi'_{P_U}(\lambda\alpha \boldsymbol{v} \cdot \boldsymbol{V}) - 2\psi'_{P_U}(\lambda\alpha q_v))(\boldsymbol{v} \cdot \boldsymbol{V} - q_v)\Big] \geq 0\,, \tag{4.3.3}$$

because by Proposition 1.3.3, ψ'_{P_U} is non-decreasing. One obtain thus the analog of the lower-bound of Proposition 3.4.1 for the non-symmetric case. The converse upper-bound is proved following the Aizenman-Sims-Starr scheme, as in the symmetric case.

In the following sections we will not, however, follow the proof strategy that we just described. This was done in [145]. We will instead provide a more straightforward proof from [21] that uses an evolution of Guerra's interpolation technique, see [20].

4.3.2 Interpolating inference model

We prove Theorem 4.1.1 in this section. First, notice that is suffices to prove Theorem 4.1.1 for $\lambda = 1$, because the dependency in λ can be "incorporated" in the prior P_U. We will thus consider in this section that $\lambda = 1$ and consequently alleviate the notations by removing the dependencies in λ. Second, it suffices to prove that

$$F_n \xrightarrow[n\to\infty]{} \sup_{q_u \geq 0} \inf_{q_v \geq 0} \mathcal{F}(\alpha, q_u, q_v) \tag{4.3.4}$$

because the equality with $\sup_{(q_u,q_v)\in\Gamma(\lambda,\alpha)} \mathcal{F}(\alpha, q_u, q_v)$ follows then from simple convex analysis arguments (Proposition C.6) presented in Appendix C.

Third, by a straightforward adaptation of the approximation argument of Section 3.4.7 to the non-symmetric case, it suffices to prove (4.3.4) in the case where the priors P_U and P_V have bounded supports included in $[-K, K]$ for some $K > 0$. We suppose now that the above conditions are verified and we will show that (4.3.4) holds.

Let $q_1, q_2 : [0,1] \to \mathbb{R}_{\geq 0}$ be two differentiable functions. For $0 \leq t \leq 1$ we consider the following observation channel

$$\begin{cases} \boldsymbol{Y}_t &= \sqrt{(1-t)/n}\,\boldsymbol{U}\boldsymbol{V}^\intercal &+ \boldsymbol{Z} \\ \boldsymbol{Y}_t^{(u)} &= \sqrt{\alpha q_1(t)}\,\boldsymbol{U} &+ \boldsymbol{Z}^{(u)} \\ \boldsymbol{Y}_t^{(v)} &= \sqrt{q_2(t)}\,\boldsymbol{V} &+ \boldsymbol{Z}^{(v)}\,, \end{cases} \tag{4.3.5}$$

where $Z_i^{(u)}, Z_j^{(v)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, are independent from everything else. The observation channel (4.3.5) interpolates between the initial matrix estimation problem (4.0.1) ($t = 0$, provided that $q_1(0)$ and $q_2(0)$ are small), and two decoupled inference channels on $\boldsymbol{U}$ and $\boldsymbol{V}$ ($t = 1$). For $r_1, r_2 \geq 0$, we define the Hamiltonian:

$$\begin{aligned} H_{n,t}(\boldsymbol{u}, \boldsymbol{v}; r_1, r_2) =& \sum_{i,j} \sqrt{\frac{(1-t)}{n}} u_i v_j Z_{i,j} + \frac{(1-t)}{n} u_i v_j U_i V_j - \frac{(1-t)}{2n} u_i^2 v_j^2 \\ &+ \sum_{i=1}^{n} \sqrt{\alpha r_1}\, u_i Z_i^{(u)} + \alpha r_1 u_i U_i - \frac{\alpha r_1}{2} u_i^2 + \sum_{j=1}^{m} \sqrt{r_2}\, v_j Z_j^{(v)} + r_2\, v_j V_j - \frac{r_2}{2}\, v_j^2\,. \end{aligned}$$

The posterior distribution of $(\boldsymbol{U}, \boldsymbol{V})$ given $(\boldsymbol{Y}_t, \boldsymbol{Y}_t^{(u)}, \boldsymbol{Y}_t^{(v)})$ is then

$$dP\Big(\boldsymbol{u}, \boldsymbol{v} \,\Big|\, \boldsymbol{Y}_t, \boldsymbol{Y}_t^{(u)}, \boldsymbol{Y}_t^{(v)}\Big) = \frac{1}{\mathcal{Z}_{n,t}} e^{H_{n,t}(\boldsymbol{u},\boldsymbol{v};q_1(t),q_2(t))} dP_U^{\otimes n}(\boldsymbol{u}) dP_V^{\otimes m}(\boldsymbol{v})\,, \tag{4.3.6}$$

where $\mathcal{Z}_{n,t}$ is the appropriate normalization. We will often drop the dependencies in $q_1(t), q_2(t)$ and write simply $H_{n,t}(\boldsymbol{u}, \boldsymbol{v})$. The Gibbs bracket $\langle \cdot \rangle_{n,t}$ denotes the expectation with respect to samples $(\boldsymbol{u}, \boldsymbol{v})$ from the posterior (4.3.6):

$$\Big\langle f(\boldsymbol{u}, \boldsymbol{v}) \Big\rangle_{n,t} = \frac{1}{\mathcal{Z}_{n,t}} \int f(\boldsymbol{u}, \boldsymbol{v}) e^{H_{n,t}(\boldsymbol{u},\boldsymbol{v};q_1(t),q_2(t))} dP_U^{\otimes n}(\boldsymbol{u}) dP_V^{\otimes m}(\boldsymbol{v})\,, \tag{4.3.7}$$

for all function f for which the right-hand side is well defined. The corresponding free energy is then

$$f_n(t) = \frac{1}{n}\mathbb{E}\log \mathcal{Z}_{n,t} = \frac{1}{n}\mathbb{E}\log\left(\int e^{H_{n,t}(\boldsymbol{u},\boldsymbol{v})} dP_U^{\otimes n}(\boldsymbol{u}) dP_V^{\otimes m}(\boldsymbol{v})\right). \tag{4.3.8}$$

Notice that

$$\begin{cases} f_n(0) &= F_n + O(q_1(0) + q_2(0)) \\ f_n(1) &= \psi_{P_U}(\alpha q_1(1)) + \frac{m}{n}\psi_{P_V}(q_2(1))\,. \end{cases} \tag{4.3.9}$$

$f_n(1)$ looks similar to the limiting expression $\mathcal{F}$ defined by (4.1.4). We would therefore like to compare $f_n(1)$ and $F_n = f_n(0) + O(q_1(0) + q_2(0))$. We thus compute the derivative of f_n:

Lemma 4.3.1

For all $t \in (0,1)$,

$$f_n'(t) = \frac{\alpha}{2} q_1'(t) q_2'(t) - \frac{1}{2}\mathbb{E}\left\langle \Big(\boldsymbol{u}\cdot\boldsymbol{U} - q_2'(t)\Big)\Big(\frac{m}{n}\boldsymbol{v}\cdot\boldsymbol{V} - \alpha q_1'(t)\Big)\right\rangle_{n,t}. \tag{4.3.10}$$

Proof. Let $t \in (0,1)$. Compute

$$f_n'(t) = \frac{1}{n}\mathbb{E}\left\langle \frac{\partial}{\partial t} H_{n,t}(\boldsymbol{u},\boldsymbol{v})\right\rangle_{n,t}.$$

Using Gaussian integration by parts and the Nishimori property (Proposition 1.1.1) as in the proof of Proposition 1.3.3, one obtains:

$$\frac{1}{n}\mathbb{E}\left\langle \frac{\partial}{\partial t} H_{n,t}(\boldsymbol{u},\boldsymbol{v})\right\rangle_{n,t} = \frac{1}{2}\alpha q_1'(t)\mathbb{E}\left\langle \boldsymbol{u}\cdot\boldsymbol{U}\right\rangle_{n,t} + \frac{1}{2}q_2'(t)\mathbb{E}\left\langle \frac{m}{n}\boldsymbol{v}\cdot\boldsymbol{V}\right\rangle_{n,t} - \frac{1}{2}\mathbb{E}\left\langle (\boldsymbol{u}\cdot\boldsymbol{U})(\frac{m}{n}\boldsymbol{v}\cdot\boldsymbol{V})\right\rangle_{n,t},$$

which leads to (4.3.10). □

Our goal now is to show that the expectation of the Gibbs measure in (4.3.10) vanishes. If this is the case, the relation $F_n \simeq f_n(0) = f_n(1) - \int_0^1 f_n'(t)dt$ would give us almost the formula that we want to prove. The arguments can be summarized as follows:

- First, we show that the overlap $\boldsymbol{u}\cdot\boldsymbol{U}$ concentrates around its mean $\mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}$.
- Then, we chose q_2 to be solution of the differential equation $q_2'(t) = \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}$ in order to cancel the Gibbs average in (4.3.10).

4.3.3 Overlap concentration

Following the ideas of Section 2.2, we show here that the overlap $\boldsymbol{u}\cdot\boldsymbol{U}$ concentrates around its mean, on average over small perturbations of our observation model.

Proposition 4.3.1

Let $R_1, R_2 : [0,1]\times(0,+\infty)^2 \to \mathbb{R}_{\geq 0}$ be two continuous, bounded functions that admits partial derivatives with respect to their second and third arguments, that are continuous and non-negative. Let $s_n = n^{-1/32}$. For $\epsilon \in [1,2]^2$, we let $q_1(\cdot,\epsilon), q_2(\cdot,\epsilon)$ be the unique solution of

$$\begin{cases} q_1(0) = s_n\epsilon_1 \\ q_2(0) = s_n\epsilon_2 \end{cases} \quad \begin{cases} q_1'(t) = R_1(t, q_1(t), q_2(t)) \\ q_2'(t) = R_2(t, q_1(t), q_2(t)). \end{cases} \tag{4.3.11}$$

Then there exists a constant $C > 0$ that only depends on K, α, $\|R_1\|_\infty$ and $\|R_2\|_\infty$, such that for all $t \in [0,1]$,

$$\int_1^2\int_1^2 \mathbb{E}\Big\langle \Big(\boldsymbol{u}\cdot\boldsymbol{U} - \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}\Big)^2\Big\rangle_{n,t} d\epsilon_1 d\epsilon_2 \leq \frac{C}{n^{1/8}},$$

where $\langle\cdot\rangle_{n,t}$ is the Gibbs measure (4.3.7) with $(q_1,q_2) = (q_1(\cdot,\epsilon), q_2(\cdot,\epsilon))$.

Proof. The existence and uniqueness of the solution of the Cauchy problem (4.3.11) comes from the usual Cauchy-Lipschitz theorem (see for instance Theorem 3.1 in Chapter V from [100]). Let us fix $t \in [0,1]$ The flow

$$Q_t : \epsilon \mapsto (q_1(t,\epsilon), q_2(t,\epsilon))$$

of (4.3.11) is a $\mathcal{C}^1$-diffeomorphism. Its Jacobian is given by the Liouville formula (see for instance Corollary 3.1 in Chapter V from [100]):

$$J(\epsilon) \stackrel{\text{def}}{=} \det\Big(\frac{\partial Q_t}{\partial \epsilon}(\epsilon)\Big) = s_n^2 \exp\Big(\int_0^t \frac{\partial R_1}{\partial q_1}(s, Q_s(\epsilon))ds + \int_0^t \frac{\partial R_2}{\partial q_2}(s, Q_s(\epsilon))ds\Big) \geq s_n^2, \tag{4.3.12}$$

because the partial derivatives inside the exponential are both non-negative. The quantity

$$\mathbb{E}\Big\langle \big(\boldsymbol{u}\cdot\boldsymbol{U} - \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}\big)^2\Big\rangle_{n,t}$$

is a function of the signal-to-noise ratios q_1 and q_2, that we denote by V. Let us write $\Omega = Q([1,2]^2)/s_n$ and $M = \max(\|R_1\|_\infty, \|R_2\|_\infty) + 2$. Notice that $\Omega \subset [1, M/s_n]^2$ because q_1, q_2 are by (4.3.11) non-decreasing and $\max(\|R_1\|_\infty, \|R_2\|_\infty)$-Lipschitz. By the change of variable $(r_1, r_2) = Q(\epsilon_1, \epsilon_2)/s_n$ we have

$$\begin{aligned}
\int_1^2\int_1^2 \mathbb{E}\Big\langle \big(\boldsymbol{u}\cdot\boldsymbol{U} - \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}\big)^2\Big\rangle_{n,t} d\epsilon_1 d\epsilon_2 &= \int_1^2\int_1^2 V(q_1(t,\epsilon_1), q_2(t,\epsilon_2))d\epsilon_1 d\epsilon_2 \\
&= \int_\Omega V(s_n r_1, s_n r_2)\frac{s_n^2 dr_1 dr_2}{J(Q_t^{-1}(s_n r))} \\
&\leq \int_1^{M/s_n}\int_1^{M/s_n} V(s_n r_1, s_n r_2)dr_1 dr_2,
\end{aligned}$$

where we used (4.3.12) for the last inequality. By the change of variable $r_1 = a^2$, we have for all $r_2 \geq 0$:

$$\int_1^{M/s_n} V(s_n r_1, s_n r_2)dr_1 = \int_1^{\sqrt{M/s_n}} V(s_n a^2, s_n r_2)2ada \leq 2\sqrt{\frac{M}{s_n}}\int_1^{\sqrt{M/s_n}} V(s_n a^2, s_n r_2)da.$$

By definition of V, the quantity $V(s_n a^2, s_n r_2)$ is the variance of the overlap $\boldsymbol{u}\cdot\boldsymbol{U}$ where $\boldsymbol{u}$ is sampled from the posterior distribution of $\boldsymbol{U}$ given $\boldsymbol{Y}_t$, $a\sqrt{\alpha s_n}\boldsymbol{U} + \boldsymbol{Z}^{(u)}$ and $\sqrt{s_n r_2}\boldsymbol{V} + \boldsymbol{Z}^{(v)}$. By Proposition 2.2.1 we have for all $1 \leq r_2 \leq M/s_n$

$$\frac{1}{\sqrt{M/s_n} - 1}\int_1^{\sqrt{M/s_n}} V(s_n a^2, s_n r_2)da \leq C\Big(\frac{1}{\sqrt{ns_n}} + \sqrt{v_n}\Big)$$

where $C > 0$ is a constant that depends only on K, α,

$$v_n = \sup_{t\in[0,1]}\ \sup_{0\leq r_1, r_2 \leq M/s_n} \mathbb{E}\big|\phi_t(r_1, r_2) - \mathbb{E}\phi_t(r_1, r_2)\big|$$

and

$$\phi_t : (r_1, r_2) \mapsto \frac{1}{ns_n}\log\Big(\int_{\boldsymbol{u},\boldsymbol{v}} dP_U^{\otimes n}(\boldsymbol{u})dP_V^{\otimes m}(\boldsymbol{v})e^{H_{n,t}(\boldsymbol{u},\boldsymbol{v};s_n r_1, s_n r_2)}\Big).$$

Consequently

$$\int_1^2\int_1^2 \mathbb{E}\Big\langle \big(\boldsymbol{u}\cdot\boldsymbol{U} - \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}\big)^2\Big\rangle_{n,t} d\epsilon_1 d\epsilon_2 \leq 2\Big(\frac{M}{s_n}\Big)^2 C\Big(\frac{1}{\sqrt{ns_n}} + \sqrt{v_n}\Big) \leq \frac{C'}{s_n^2}\Big(\frac{1}{\sqrt{ns_n}} + \sqrt{v_n}\Big),$$

for some constant $C' > 0$. We now use the following lemma to control v_n:

Lemma 4.3.2

There exists a constant $C > 0$ (that only depends on K, M and α) such that

$$v_n \leq C n^{-1/2} s_n^{-1}.$$

We delay the proof of Lemma 4.3.2 to Section 4.3.5. We deduce that

$$\int_1^2\int_1^2 \mathbb{E}\Big\langle \big(\boldsymbol{u}\cdot\boldsymbol{U} - \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}\big)^2\Big\rangle_{n,t} d\epsilon_1 d\epsilon_2 \leq \frac{2C'}{n^{1/8}}$$

if we choose $s_n = n^{-1/32}$. □

4.3.4 Lower and upper bounds

From now we write $\mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}$, as a function of $(t, q_1(t), q_2(t))$:

$$\mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t} = Q(t, q_1(t), q_2(t)). \tag{4.3.13}$$

Notice that Q is continuous, non-negative on $[0,1]\times(0,+\infty)^2$, bounded by K^2 and admits partial derivatives with respect to its second and third argument. These derivatives are both continuous. Moreover, notice that

$$\mathbb{E}_{P_U}[U^2] - Q(t, r_1, r_2) = \mathbb{E}\Big\|\boldsymbol{U} - \mathbb{E}\Big[\boldsymbol{U}\Big|\boldsymbol{Y}_t, \sqrt{\alpha r_1}\boldsymbol{U} + \boldsymbol{Z}^{(u)}, \sqrt{r_2}\boldsymbol{V} + \boldsymbol{Z}^{(v)}\Big]\Big\|^2$$

is of course non-increasing with respect to r_1 and r_2. The partial derivatives of Q with respect to its second and third argument are thus non-negative.

For simplicity we will now omit the dependencies on λ and α in $\mathcal{F}$. The proof of (4.3.4) will follow from the two matching lower- and upper-bounds below.

Proposition 4.3.2

In the setting of Proposition 4.3.1, for $\epsilon\in[1,2]^2$ we let $q_1(t,\epsilon), q_2(t,\epsilon)$ be the solution of (4.3.11), with the choice $R_2 = Q$. For this choice of functions q_1, q_2, we have:

$$F_n = \int_{[1,2]^2}\int_0^1 \Big(\psi_{P_U}(\alpha q_1(1,\epsilon)) + \alpha\psi_{P_V}(q_2(1,\epsilon)) - \frac{\alpha}{2}q_1'(t,\epsilon)q_2'(t,\epsilon)\Big)dtd\epsilon + o_n(1).$$

Proof. Let us fix $\epsilon\in[1,2]^2$. With the choice $R_2 = Q$, we have for all $t\in[0,1]$:

$$q_2'(t,\epsilon) = Q(t, q_1(t,\epsilon), q_2(t,\epsilon)) = \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}.$$

The derivative of (4.3.10) becomes then by Proposition 4.3.1:

$$f_n'(t) = \frac{\alpha}{2}q_1'(t)q_2'(t) - \frac{1}{2}\mathbb{E}\left\langle \Big(\boldsymbol{u}\cdot\boldsymbol{U} - \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}\Big)\Big(\frac{m}{n}\boldsymbol{v}\cdot\boldsymbol{V} - \alpha q_1'(t)\Big)\right\rangle_{n,t} = \frac{\alpha}{2}q_1'(t)q_2'(t) + o_n(1)$$

where $o_n(1)$ denotes a quantity that goes to 0 as $n\to\infty$, uniformly in t,ϵ. By Proposition B.1 we have $f_n(0) = F_n + O_n(s_n)$. We have also: $f_n(1) = \psi_{P_U}(\alpha q_1(1,\epsilon)) + \alpha\psi_{P_V}(q_2(1,\epsilon)) + o_n(1)$. We conclude by

$$\begin{aligned} F_n &= \int_{[1,2]^2} f_n(0)d\epsilon + o_n(1) = \int_{[1,2]^2}\Big(f_n(1) - \int_0^1 f_n'(t)dt\Big)d\epsilon + o_n(1) \\ &= \int_{[1,2]^2}\int_0^1 \Big(\psi_{P_U}(\alpha q_1(1,\epsilon)) + \alpha\psi_{P_V}(q_2(1,\epsilon)) - \frac{\alpha}{2}q_1'(t,\epsilon)q_2'(t,\epsilon)\Big)dtd\epsilon + o_n(1). \end{aligned}$$

□

Lower bound

One deduces from Proposition (4.3.2) the following lower bound:

Proposition 4.3.3

$$\liminf_{n\to\infty} F_n \geq \sup_{q_1\geq 0}\inf_{q_2\geq 0}\mathcal{F}(q_2, q_1)\,.$$

Proof. We apply Proposition 4.3.2 with $R_1 = r$, for some $r \geq 0$. We get $q_1(t,\epsilon) = \epsilon_1 s_n + rt$, so that:

$$\begin{aligned} F_n &= \int_{[1,2]^2} \int_0^1 \Big(\psi_{P_U}(\alpha(s_n\epsilon_1 + r)) + \alpha\psi_{P_V}(q_2(1,\epsilon)) - \frac{\alpha}{2} r q_2'(t,\epsilon)\Big) dt d\epsilon + o_n(1). \\ &= \int_{[1,2]^2} \Big(\psi_{P_U}(\alpha r) + \alpha\psi_{P_V}(q_2(1,\epsilon)) - \frac{\alpha}{2} r q_2(1,\epsilon)\Big) d\epsilon + o_n(1). \\ &\geq \inf_{q_2 \geq 0} \mathcal{F}(q_2, r) + o_n(1), \end{aligned}$$

where we used the fact that ψ_{P_U} is $\frac{1}{2}K^2$-Lipschitz, and that $s_n \to 0$. This proves the proposition since the last inequality holds for all $r \geq 0$. □

Upper bound

We will now prove the converse upper bound.

Proposition 4.3.4

$$\limsup_{n\to\infty} F_n \leq \sup_{q_1 \geq 0} \inf_{q_2 \geq 0} \mathcal{F}(q_2, q_1)\,.$$

Proof. We apply Proposition 4.3.2 with $R_1 = 2\alpha\psi'_{P_V} \circ Q$. R_1 verifies the conditions of Proposition 4.3.1 because ψ_{P_V} is a $\mathcal{C}^2$ convex Lipschitz function (Proposition 1.3.3).

For simplicity we omit briefly the dependencies in ϵ of q_1 and q_2. ψ_{P_U} is $K^2/2$-Lipschitz, and $q_1(0) = \epsilon s_n = o_n(1)$ so $\psi_{P_U}(\alpha q_1(1)) = \psi_{P_U}(\alpha(q_1(1) - q_1(0))) + o_n(1)$, where $o_n(1)$ is a quantity that goes to 0 as $n \to \infty$, uniformly in $\epsilon \in [1,2]^2$. Notice that by convexity of the functions ψ_{P_U} and ψ_{P_V}, we get

$$\psi_{P_U}(\alpha q_1(1)) = \psi_{P_U}\Big(\alpha \int_0^1 q_1'(t)\Big) + o_n(1) \leq \int_0^1 \psi_{P_U}(\alpha q_1'(t)) dt + o_n(1).$$

and similarly: $\psi_{P_V}(q_2(1)) \leq \int_0^1 \psi_{P_V}(q_2'(t)) dt + o_n(1)$. We get by Proposition 4.3.2

$$\begin{aligned} F_n &\leq \int_{[1,2]^2} \int_0^1 \Big(\psi_{P_U}(\alpha q_1'(t,\epsilon)) + \alpha\psi_{P_V}(q_2'(t,\epsilon)) - \frac{\alpha}{2} q_1'(t,\epsilon) q_2'(t,\epsilon)\Big) dt d\epsilon + o_n(1). \\ &= \int_{[1,2]^2} \int_0^1 \mathcal{F}(q_2'(t,\epsilon), q_1'(t,\epsilon)) dt d\epsilon + o_n(1). \end{aligned} \tag{4.3.14}$$

Since we chose $R_2 = Q$ and $R_1 = 2\alpha\psi'_{P_V} \circ Q = 2\alpha\psi'_{P_V} \circ R_2$, Equation (4.3.11) gives:

$$\forall \epsilon \in [1,2]^2,\ \forall t \in [0,1], \qquad q_1'(t,\epsilon) = 2\alpha\psi'_{P_V}(q_2'(t,\epsilon)).$$

By convexity of ψ_{P_V}, this gives that for all $\epsilon \in [1,2]^2$ and all $t \in [0,1]$ we have

$$\mathcal{F}(q_2'(t,\epsilon), q_1'(t,\epsilon)) = \inf_{q_2 \geq 0} \mathcal{F}(q_2, q_1'(t,\epsilon)) \leq \sup_{q_1 \geq 0} \inf_{q_2 \geq 0} \mathcal{F}(q_2, q_1).$$

Together with (4.3.14), this concludes the proof. □

4.3.5 Concentration of the free energy: proof of Lemma 4.3.2

In this section, we prove Lemma 4.3.2: we show that the perturbed free energy concentrates around its mean, uniformly in the perturbation. Lemma 4.3.2 will follow from Lemma 4.3.3 and Lemma 4.3.4 below. Let $\mathbb{E}_z$ denote the expectation with respect to the Gaussian random variables $\boldsymbol{Z}, \boldsymbol{Z}^{(u)}, \boldsymbol{Z}^{(v)}$.

Lemma 4.3.3

There exists a constant $C > 0$, that only depends on K, α, such that for all $t \in [0,1]$, $B \geq 0$ and $(r_1, r_2) \in [0, B]^2$,

$$\mathbb{E}\left|\phi_t(r_1, r_2) - \mathbb{E}_z \phi_t(r_1, r_2)\right| \leq C n^{-1/2} s_n^{-1} \sqrt{1 + B s_n}\,.$$

Proof. Let $(r_1, r_2) \in [0, B]^2$ and consider $\boldsymbol{U}$ and $\boldsymbol{V}$ to be fixed (i.e. we first work conditionally on $\boldsymbol{U}, \boldsymbol{V}$). Consider the function

$$f : (\boldsymbol{Z}, \boldsymbol{Z}^{(u)}, \boldsymbol{Z}^{(v)}) \mapsto \phi_t(r_1, r_2)\,.$$

It is not difficult to verify that

$$\|\nabla f\|^2 \leq \frac{C}{n s_n^2}(1 + B s_n)$$

for some constant $C > 0$ that depends only on K and α. The Gaussian Poincaré inequality (see [37] Chapter 3) gives then

$$\mathbb{E}_z\left(\phi_t(r_1, r_2) - \mathbb{E}_z \phi_t(r_1, r_2)\right)^2 \leq \frac{C}{n s_n^2}(1 + B s_n)\,.$$

We obtain the lemma by integration over $\boldsymbol{U}, \boldsymbol{V}$ and Jensen's inequality. □

Lemma 4.3.4

There exists a constant $C > 0$, that only depends on K, α, such that for all $t \in [0,1]$, $B \geq 0$ and $(r_1, r_2) \in [0, B]^2$,

$$\mathbb{E}\left|\mathbb{E}_z \phi_t(r_1, r_2) - \mathbb{E}\phi_t(r_1, r_2)\right| \leq C n^{-1/2} s_n^{-1} \sqrt{1 + B s_n}\,.$$

Proof. It is not difficult to verify that the function

$$g : (\boldsymbol{U}, \boldsymbol{V}) \mapsto \mathbb{E}_z \phi_t(r_1, r_2)$$

verifies a "bounded difference property" (see [37], Section 3.2) because the components of $\boldsymbol{U}$ and $\boldsymbol{V}$ are bounded by a constant $K > 0$. Then Corollary 3.2 from [37] (which is a corollary from the Efron-Stein inequality) implies that for all $t \in [0,1]$ and $r_1, r_2 \in [0, B]$

$$\mathbb{E}\left(\mathbb{E}_z \phi_t(r_1, r_2) - \mathbb{E}\phi_t(r_1, r_2)\right)^2 \leq C n^{-1} s_n^{-2}(1 + s_n B)\,.$$

for some constant C depending only on K and α. We conclude the proof using Jensen's inequality. □

4.4 Proof of Theorem 4.1.2

In order to prove Theorem 4.1.2, we are going to consider the following model with side information to obtain a lower bound on the MMSE. Suppose that we observe for $\gamma \geq 0$

$$\begin{cases} \boldsymbol{Y}_\lambda &= \sqrt{\frac{\lambda}{n}} \boldsymbol{U}\boldsymbol{V}^\intercal + \boldsymbol{Z} \\ \boldsymbol{Y}'_\gamma &= \sqrt{\frac{\gamma}{n}} \boldsymbol{U}\boldsymbol{U}^\intercal + \boldsymbol{Z}' \end{cases} \tag{4.4.1}$$

where $(Z'_{i,j} = Z'_{j,i})_{i\le j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ are independent from everything else. Define the corresponding free energy

$$F_n(\lambda,\gamma) = \frac{1}{n}\mathbb{E}\log\int dP_U^{\otimes n}(\boldsymbol{u})dP_V^{\otimes m}(\boldsymbol{v})\exp\Big(\sum_{1\le i\le j\le n}\sqrt{\frac{\gamma}{n}}Y'_{i,j}u_iu_j - \frac{\gamma u_i^2u_j^2}{2} + \sum_{i,j}\sqrt{\frac{\lambda}{n}}Y_{i,j}u_iv_j - \frac{\lambda u_i^2v_j^2}{2n}\Big).$$

Proposition 4.4.1

Recall that $\psi^*_{P_U}$ (resp. $\psi^*_{P_V}$) denotes the monotone conjugate (Definition C.2 in Appendix C) of ψ_{P_U} (resp. ψ_{P_V}). For all $\lambda,\gamma\ge 0$, we have

$$F_n(\lambda,\gamma) \xrightarrow[n\to\infty]{} f(\lambda,\gamma) \overset{\text{def}}{=} \sup_{q_u,q_v\ge 0}\Big\{\frac{\gamma q_u^2}{4} + \frac{\alpha\lambda q_u q_v}{2} - \psi^*_{P_U}(q_u/2) - \alpha\psi^*_{P_V}(q_v/2)\Big\}. \quad (4.4.2)$$

Proposition 4.4.1 is proved at the end of this section. Before we deduce Theorem 4.1.2 from Proposition 4.4.1, let us just mention that Proposition 4.4.1 allows to precisely derive the information-theoretic limits for the model (4.4.1), by the "I-MMSE" relation (Proposition 1.3.3).

Corollary 4.4.1

For almost all $\gamma > 0$ the supremum of Proposition 4.4.1 is achieved at a unique $q_u^*(\lambda,\gamma,\alpha)$ and

$$\mathrm{MMSE}_n^{(u)}(\lambda,\gamma) \overset{\text{def}}{=} \frac{1}{n^2}\mathbb{E}\Big[\sum_{1\le i,j\le n}\big(U_iU_j - \mathbb{E}[U_iU_j|\boldsymbol{Y}_\lambda,\boldsymbol{Y}'_\gamma]\big)^2\Big] \xrightarrow[n\to\infty]{} \mathbb{E}[U^2]^2 - q_u^*(\lambda,\gamma,\alpha)^2\,.$$

The model (4.4.1) was considered in [62], in the special case $P_U = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$ and $P_V = \mathcal{N}(0,1)$. Theorem 6 from [62] shows that one can estimate $\boldsymbol{U}\boldsymbol{U}^\intercal$ better than a random guess if and only if $\gamma^2+\alpha\lambda^2 > 1$. Corollary 4.4.1 above is more precise and general because it gives the precise expression of the minimum mean squared error for any prior P_U, P_V. In particular the boundary $\gamma^2+\alpha\lambda^2 = 1$ is not expected to be the information-theoretic threshold for sufficiently sparse or unbalanced priors, see the phase diagram of Figure 3.3 for a similar scenario.

Let us now deduce Theorem 4.1.2 from Proposition 4.4.1. By the "I-MMSE" relation of Proposition 1.3.3:

$$\mathrm{MMSE}_n^{(u)}(\lambda) = \mathrm{MMSE}_n^{(u)}(\lambda,\gamma=0) = \mathbb{E}_{P_U}[U^2] - 4\frac{\partial F_n}{\partial\gamma}(\lambda,0^+). \quad (4.4.3)$$

The sequence of convex functions $(F_n(\lambda,\cdot))_{n\ge1}$ converges pointwise to $f(\lambda,\cdot)$ on $\mathbb{R}_{\ge0}$. Thus, by Proposition C.1:

$$\limsup_{n\to\infty}\frac{\partial F_n}{\partial\gamma}(\lambda,0^+) \le \frac{\partial f}{\partial\gamma}(\lambda,0^+). \quad (4.4.4)$$

We need therefore the following lemma:

Lemma 4.4.1

For all $\alpha>0$ and all $\lambda\in D_\alpha$, $\dfrac{\partial f}{\partial\gamma}(\lambda,0^+) = \dfrac{q_u^*(\alpha,\lambda)^2}{4}$.

Proof. Let $\alpha > 0$ and $\lambda \in D_\alpha$. Then the supremum of (4.4.2) is uniquely achieved at $(q_u^*(\lambda,\alpha), q_v^*(\lambda,\alpha))$ because the couples achieving this supremum are by Proposition C.6 precisely the couples achieving the supremum of $\mathcal{F}(\lambda,\alpha,\cdot,\cdot)$ over $\Gamma(\lambda,\alpha)$. The lemma follows then from the "envelope theorem" of Proposition D.2. □

From Lemma 4.4.1 and equations (4.4.3)-(4.4.4) above, we conclude:

$$\liminf_{n\to\infty} \mathrm{MMSE}_n^{(u)}(\lambda) \geq \mathbb{E}_{P_U}[U^2]^2 - q_u^*(\lambda,\alpha)^2.$$

Let $(\boldsymbol{u},\boldsymbol{v})$ sampled from the posterior distribution of $(\boldsymbol{U},\boldsymbol{V})$ given $\boldsymbol{Y}$, independently of everything else. Then $\mathrm{MMSE}_n^{(u)}(\lambda) = \mathbb{E}_{P_U}[U^2]^2 - \mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})^2\Big] + o_n(1)$. This gives (the corresponding result for V is obtained by symmetry):

$$\limsup_{n\to\infty} \mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})^2\Big] \leq q_u^*(\lambda,\alpha)^2 \qquad \text{and} \qquad \limsup_{n\to\infty} \mathbb{E}\Big[(\boldsymbol{v}\cdot\boldsymbol{V})^2\Big] \leq q_v^*(\lambda,\alpha)^2\,. \tag{4.4.5}$$

Now, we know by Proposition 4.1.1 that

$$\mathbb{E}_{P_U}[U^2]\mathbb{E}_{P_V}[V^2] - \mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})(\boldsymbol{v}\cdot\boldsymbol{V})\Big] = \mathrm{MMSE}_n(\lambda) \xrightarrow[n\to\infty]{} \mathbb{E}_{P_U}[U^2]\mathbb{E}_{P_V}[V^2] - q_u^* q_v^*\,,$$

which gives $\mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})(\boldsymbol{v}\cdot\boldsymbol{V})\Big] \to q_u^* q_v^*$. By Cauchy-Schwarz inequality we have

$$\mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})(\boldsymbol{v}\cdot\boldsymbol{V})\Big]^2 \leq \mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})^2\Big]\,\mathbb{E}\Big[(\boldsymbol{v}\cdot\boldsymbol{V})^2\Big]$$

which gives, by taking the liminf:

$$(q_u^* q_v^*)^2 \leq \left(\liminf_{n\to\infty} \mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})^2\Big]\right)\left(\liminf_{n\to\infty} \mathbb{E}\Big[(\boldsymbol{v}\cdot\boldsymbol{V})^2\Big]\right).$$

Combining this with (4.4.5), we get that $\lim \mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})^2\Big] = (q_u^*)^2$ and the relation $\mathrm{MMSE}_n^{(u)}(\lambda) = \mathbb{E}_{P_U}[U^2]^2 - \mathbb{E}\Big[(\boldsymbol{u}\cdot\boldsymbol{U})^2\Big] + o_n(1)$ gives the result.

Proof of Proposition 4.4.1

It suffices to prove the result in the case where P_U and P_V have bounded support, because we can then proceed by approximation as in Section 3.4.7. From now, we suppose to be in that case. Since the dependency in γ can be incorporated in the prior P_U and the one in λ in the prior P_V, we only have to prove Proposition 4.4.1 in the case $\gamma = \lambda = 1$. In the sequel we will therefore remove the dependencies in λ, γ. Define for $r \geq 0$

$$L_n(r) = \frac{1}{n}\mathbb{E}\log \int dP_U^{\otimes n}(\boldsymbol{u}) dP_V^{\otimes m}(\boldsymbol{v}) \exp\left(H_n(\boldsymbol{u},\boldsymbol{v}) + \sum_{i=1}^n \sqrt{r} Z_i'' u_i + r U_i u_i - \frac{r}{2}u_i^2\right),$$

where $Z_i'' \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, independently of everything else and where the Hamiltonian $H_n(\boldsymbol{u},\boldsymbol{v})$ is defined by (4.1.1) (with $\lambda = 1$). L_n is the free energy (expected log-partition function) for observing jointly $\boldsymbol{Y} = \frac{1}{\sqrt{n}}\boldsymbol{U}\boldsymbol{V}^\intercal + \boldsymbol{Z}$ and $\boldsymbol{Y}'' = \sqrt{r}\boldsymbol{U} + \boldsymbol{Z}''$. By an straightforward extension of Theorem 4.1.1 we have for all $r \geq 0$:

$$L_n(r) \xrightarrow[n\to\infty]{} L(r) \tag{4.4.6}$$

where

$$L(r) = \sup_{q_u \geq 0} \inf_{q_v \geq 0} \left\{\psi_{P_U}(\alpha q_v + r) + \alpha \psi_{P_V}(q_u) - \frac{\alpha q_u q_v}{2}\right\}.$$

Lemma 4.4.2

$$F_n(\lambda,\gamma) \xrightarrow[n\to\infty]{} \sup_{r\geq 0}\left\{L(r)-\frac{r^2}{4}\right\}. \tag{4.4.7}$$

Proof. We will follow the same steps than in Section 4.3: we will therefore only present the main steps. Let $r:[0,1]\to\mathbb{R}_{\geq 0}$ be a differentiable function. For $0\leq t\leq 1$ we consider the following observation channel

$$\begin{cases} \boldsymbol{Y} &= \sqrt{1/n}\boldsymbol{U}\boldsymbol{V}^\intercal &+ \boldsymbol{Z} \\ \boldsymbol{Y}_t' &= \sqrt{(1-t)/n}\boldsymbol{U}\boldsymbol{U}^\intercal &+ \boldsymbol{Z}' \\ \boldsymbol{Y}_t'' &= \sqrt{r(t)}\boldsymbol{U} &+ \boldsymbol{Z}'', \end{cases} \tag{4.4.8}$$

We will denote (analogously to (4.3.8)) by $f_n(t)$ the interpolating free energy and by $\langle\cdot\rangle_{n,t}$ (analogously to (4.3.7)) corresponding Gibbs measure. We have the analog of Equation (3.4.3) and Lemma 4.3.1:

$$f_n'(t) = -\frac{1}{4}\mathbb{E}\Big\langle\big(\boldsymbol{u}\cdot\boldsymbol{U}-r'(t)\big)^2\Big\rangle_{n,t}+\frac{r'(t)^2}{4}+o_n(1), \tag{4.4.9}$$

where $o_n(1)\to 0$, uniformly in $t\in[0,1]$. By taking $r(t)=rt$ for all $t\in[0,1]$, we obtain

$$F_n = f_n(0) = f_n(1)-\int_0^1 f_n'(t)dt \geq L_n(r)-\frac{r^2}{4}.$$

Therefore $\liminf F_n \geq \liminf_{n\to\infty}\{L_n(r)-\frac{r^2}{4}\}$ which gives $\liminf F_n \geq L(r)-\frac{r^2}{4}$ for all $r\geq 0$, hence $\liminf F_n \geq \sup_{r\geq 0}\{L(r)-\frac{r^2}{4}\}$.

To prove the converse upper-bound we proceed as in Section 4.3 and chose r to be solution $r(\cdot;\epsilon)$ of the Cauchy problem:

$$\begin{cases} r(0) = \epsilon n^{-1/32} \\ r'(t) = \mathbb{E}\langle \boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t} \end{cases}$$

where $\epsilon\in[1,2]$ is a parameter. The analog of Proposition 4.3.1 holds:

$$\int_1^2 \mathbb{E}\Big\langle\big(\boldsymbol{u}\cdot\boldsymbol{U}-\mathbb{E}\langle\boldsymbol{u}\cdot\boldsymbol{U}\rangle_{n,t}\big)^2\Big\rangle_{n,t}d\epsilon \leq \frac{C}{n^{1/8}}$$

for some constant $C>0$. Using (4.4.9) we get

$$\begin{aligned} F_n = f_n(0)+o_n(1) &= \int_1^2\Big(f_n(1)-\int_0^1 f_n'(t)dt\Big)d\epsilon + o_n(1) \\ &= \int_1^2\Big(L_n(r(1,\epsilon))-\int_0^1\frac{r'(t,\epsilon)^2}{4}dt\Big)d\epsilon+o_n(1) \\ &\leq \int_1^2\int_0^1\Big(L_n(r'(t,\epsilon))-\frac{r'(t,\epsilon)^2}{4}\Big)dtd\epsilon+o_n(1) \leq \sup_{0\leq r\leq\rho_u}\Big\{L_n(r)-\frac{r^2}{4}\Big\}+o_n(1), \end{aligned}$$

where $\rho_u=\mathbb{E}_{P_U}[U^2]$. The free energy L_n is (by the usual arguments, see Section 1.3) convex and non-decreasing and converges to L which is thus convex (therefore continuous) and non-decreasing. By Dini's second theorem we get that the convergence in (4.4.6) is uniform in r over all compact subsets of $\mathbb{R}_{\geq 0}$. We conclude

$$\limsup_{n\to\infty} F_n \leq \sup_{0\leq r\leq\rho_u}\{L(r)-\frac{r^2}{4}\} \leq \sup_{r\geq 0}\{L(r)-\frac{r^2}{4}\}.$$

□

In order to prove Proposition 4.4.1, it remains to show that

$$\sup_{r\geq 0}\left\{L(r)-\frac{r^2}{4}\right\}=\sup_{q_u,q_v\geq 0}\left\{\frac{\gamma q_u^2}{4}+\frac{\alpha\lambda q_u q_v}{2}-\psi^*_{P_U}(q_u/2)-\alpha\psi^*_{P_V}(q_v/2)\right\}.$$

This is a consequence of the following Lemma:

Lemma 4.4.3

Let f, g be two non-decreasing lower semi-continuous convex functions on $\mathbb{R}_{\geq 0}$, such that $f(0)$ and $g(0)$ are finite. Let f^* and g^* denote their monotone conjugate (see Definition C.2 in Appendix C). Then

$$\sup_{r\geq 0}\sup_{q_1\geq 0}\inf_{q_2\geq 0}\left\{f(q_1)+g(r+q_2)-q_1q_2-\frac{r^2}{2}\right\}=\sup_{q_1,q_2\geq 0}\left\{\frac{q_1^2}{2}+q_1q_2-f^*(q_2)-g^*(q_1)\right\}$$

Proof. Let $r\geq 0$. Let us write $g_r : q\mapsto g(q+r)$. By Proposition C.6, we have

$$\begin{aligned}\sup_{q_1\geq 0}\inf_{q_2\geq 0}\left\{f(q_1)+g(r+q_2)-q_1q_2\right\}&=\sup_{q_1,q_2\geq 0}\left\{q_1q_2-f^*(q_2)-g_r^*(q_1)\right\}\\&=\sup_{q_2\geq 0}\left\{g(r+q_2)-f^*(q_2)\right\}\\&=\sup_{q_1,q_2\geq 0}\left\{q_1(q_2+r)-f^*(q_2)-g^*(q_1)\right\},\end{aligned}$$

where we used Proposition C.4 for the two last equalities. Therefore

$$\begin{aligned}\sup_{r\geq 0}\sup_{q_1\geq 0}\inf_{q_2\geq 0}\left\{f(q_1)+g(r+q_2)-q_1q_2-\frac{r^2}{2}\right\}&=\sup_{q_1,q_2\geq 0}\left\{\sup_{r\geq 0}\left\{q_1r-\frac{r^2}{2}\right\}+q_1q_2-f^*(q_2)-g^*(q_1)\right\}\\&=\sup_{q_1,q_2\geq 0}\left\{\frac{q_1^2}{2}+q_1q_2-f^*(q_2)-g^*(q_1)\right\}.\end{aligned}$$

□

Chapter 5

Community detection in the asymmetric stochastic block model

5.1 Introduction

The community detection problem is fundamental problem in statistics. It consists in dividing the vertices of a given graph into groups which are more densely connected than the rest of the graph. The stochastic block model is a popular model of random graphs that exhibit communities: it is generated according to an underlying partition of the vertices. This model has been studied for a long time in statistics (see [103]) computer science (see [76]) and more recently in statistical physics (see [56]).

Definition 5.1.1 (*Stochastic block model (SBM)*)

Let M be a 2×2 symmetric matrix whose entries are in $[0,1]$. Let $n \in \mathbb{N}^*$ and $p \in [0,1]$. We define the stochastic block model with parameters (M,n,p) as the random graph G defined by:

1. *The vertices of G are the integers in $\{1,\dots,n\}$.*
2. *For each vertex $i \in \{1,\dots,n\}$ one draws independently $X_i \in \{1,2\}$ according to $\mathbb{P}(X_i = 1) = p$. X_i will be called the label (or the class, or the community) of the vertex i.*
3. *For each pair of vertices $\{i,j\}$ the unoriented edge $G_{i,j}$ is then drawn conditionally on X_i and X_j according to a Bernoulli distribution with mean M_{X_i,X_j}.*

Our main focus will be on the community detection problem: given the graph G, is it possible to retrieve the labels X better than a random guess?

We investigate this question in the asymptotic of large sparse graphs, when $n \to +\infty$ while the average degree remains fixed. Our quantitative results will then be obtained when the average degree tends to infinity. We define the connectivity matrix M as follows:

$$M = \frac{d}{n}\begin{pmatrix} a & b \\ b & c \end{pmatrix}, \tag{5.1.1}$$

where a, b, c, d remain fixed as $n \to +\infty$. In this case, it is not possible to correctly classify more than a certain fraction of the vertices correctly and we will say that community de-

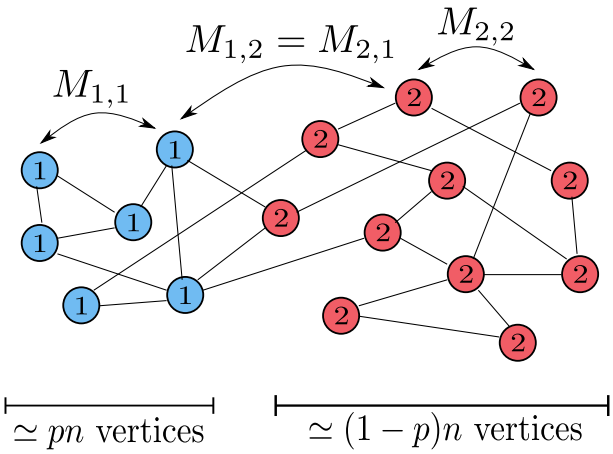


Figure 5.1: The Stochastic Block Model with two communities of different sizes.

tection is solvable if there is some algorithm that recovers the communities more accurately than a random guess would.

The symmetric case where $p = 1/2$ and $a = c$ has been extensively studied starting with [56] and gives rise to an interesting phenomenon: if $d(1-b)^2 < 1$ then community detection is not solvable [155], while if $d(1-b)^2 > 1$, it is solvable (in polynomial time) [138, 154]. Much less is known in the case where (5.1.2) holds and $p < 1/2$. The aim of this work is to investigate this asymmetric case. The main question is: does the asymmetry change the location of the above mentioned threshold?

A simple argument (see below) shows that if $pa+(1-p)b \neq pb+(1-p)c$ then non-trivial information on the community of a vertex can be gained just by looking at its degree and the community detection is then solvable. In this paper, we concentrate on the case:

$$pa + (1-p)b = pb + (1-p)c = 1. \tag{5.1.2}$$

Under condition (5.1.2), the matrix R defined by

$$R = \begin{pmatrix} pa & (1-p)b \\ pb & (1-p)c \end{pmatrix} \tag{5.1.3}$$

is a stochastic matrix with two eigenvalues: $\lambda_1 = 1 > \lambda_2 = 1 - b$. Define

$$\lambda = d\lambda_2^2 = d(1-b)^2 . \tag{5.1.4}$$

If $\lambda > 1$, it is shown in [161] that it is easy to distinguish G from an Erdős-Rényi model with the same average degree. In this regime, the spectral algorithm based on the non-backtracking matrix solves the community detection problem [36, 1]. Here, we prove that if a vanishing fraction of labels is given, then a local algorithm (belief propagation) allows to solve the community detection problem.

The case $\lambda < 1$ (known as below the Kesten-Stigum bound) is more challenging. It is shown in [161], that the community detection problem is still solvable for some values of p but it is expected in [56] that no computationally efficient reconstruction is possible. In [161], some bounds are given on the non-reconstructability region but they are not expected to be tight.

In this work, we are mainly interested in a regime where d tends to infinity while $\lambda = d(1-b)^2$ remains constant (in particular a, b, c tend to one). Note that we first let n tend

to infinity and then let d tend to infinity mainly in order to get explicit formulas. In this regime, we show that for all values of $p \in (p^*, 1/2)$ with $p^* = \frac{1}{2} - \frac{1}{2\sqrt{3}}$, the situation is similar to the balanced case: below the Kesten-Stigum bound, i.e. when $\lambda < 1$, the community detection problem is not solvable. For $p < p^*$, we compute a function $p \mapsto \lambda_{sp}(p) < 1$ such that for $\lambda < \lambda_{sp}(p)$, the community detection problem is not solvable. As shown by [161], there are points in the region $\lambda_{sp}(p) < \lambda < 1$ where the community detection problem is solvable but we do not expect the bound $\lambda_{sp}(p)$ to be tight, i.e. the information theoretic threshold for community detection should be above $\lambda_{sp}(p)$ for $p < p^*$.

There is an important probabilistic interpretation of the matrix R relating to the local structure of the SBM. As explained below, the SBM converges locally toward a labeled Poisson Galton-Watson branching process with mean offspring d: the label of the root is 1 with probability p and 2 with probability $1 - p$ and then conditioned on the parent's label being i, its children's labels are independently chosen to be j with probability R_{ij}. A problem closely related to the detection problem in the SBM is the reconstruction problem on this random tree: given some information about the labels at depth n from the root, is it possible to infer some information about the label of the root when $n \to \infty$? It is known [153] that the Kesten-Stigum bound corresponds to census-solvability (i.e. knowing only the number of labels 1 and 2 at depth n allows to get some information about the label of the root). When $d \to \infty$, we show that $\lambda_{sp}(p)$ corresponds to the solvability threshold for the reconstruction problem on the tree (i.e. knowing the labels at depth n allows to get some information about the label of the root). We also consider the reconstruction problem where the label of each node at depth n is revealed with probability q. Then in the region $\lambda_{sp}(p) < \lambda < 1$, we compute the minimal value of q such that some information about the label of the root can be recovered from the revealed labels. Above the Kesten-Stigum bound, i.e. when $1 < \lambda$, this minimal value is 0.

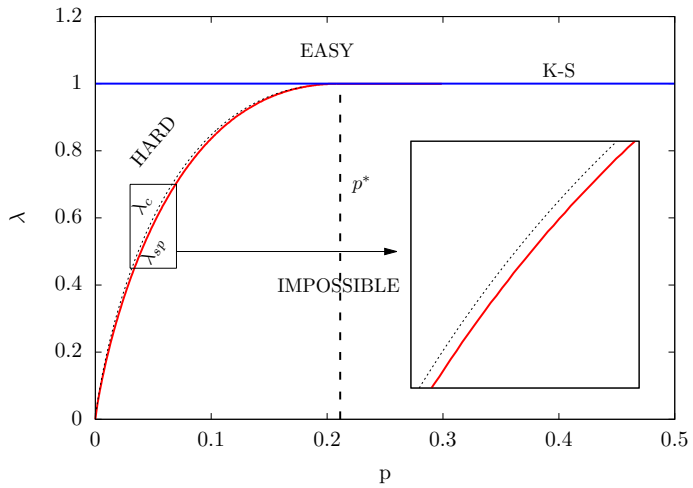


Figure 5.2: Phase diagram for the asymmetric community detection problem. The easy phase follows from [36], the impossible phase below the spinodal curve (red curve) is proved in this paper and the hard phase is a conjecture. The dotted curve corresponding to $\lambda_c(p)$ is the conjectured curve for solvability of the community detection problem (see discussion in Section 5.4).

We summarize our main results on the phase diagram of Figure 5.2:

- Above the Kesten-Stigum bound (blue line), reconstruction is possible by a local algorithm given that an arbitrary small fraction of the labels is revealed. Moreover,

the local algorithms (with this arbitrary small side information) achieve then the best possible performance without side information (see Proposition 5.3.1).

- Between the blue and the red line, we show that local algorithms are efficient for reconstruction when a certain fraction of labels is revealed (see Proposition 5.4.2).
- We show that reconstruction is impossible below the spinodal curve (red line), see Proposition 5.4.1.

In Section 5.2, we define the community detection problem and its variation when some labels are revealed. In Section 5.3, we give our main results about reconstruction above the Kesten-Stigum bound and in Section 5.4, we describe what happens below the Kesten-Stigum bound. Section 5.5 defines the various notions of solvability for the problem of reconstruction on trees and gives our main result for this problem. We first use the cavity method on trees in Section 5.6 and then relate these results to the original problem of community detection in Section 5.7. Finally, we show in Section 5.9 that the community detection problem is in some sense equivalent to the low-rank matrix estimation problem of Chapter 3. This allows to use the results of Chapter 3 to obtain precise reconstruction limits for the stochastic block model.

5.2 The community detection problem

We are interested in inferring the labels X from the graph G. To do so, we aim at constructing an estimator (i.e. a function of the observation G) $T(G) = (T_1(G), \dots, T_n(G)) \in \{1,2\}^n$, such that $T_i(G)$ is 'close' to X_i. We will measure the performance of T using the 'rescaled average success probability' defined as follows

$$P_{\mathrm{suc}}(T) = \frac{1}{n}\sum_{i=1}^{n}\Big(\mathbb{P}(T_i(G)=1|X_i=1) + \mathbb{P}(T_i(G)=2|X_i=2) - 1\Big). \tag{5.2.1}$$

The '-1' is here to rescale the success probability and to ensure that $P_{\mathrm{suc}} = 0$ for 'dummy estimators' (i.e. estimators that do not depend on the observed graph G). The optimal test with respect to this metric is

$$T_i^{\mathrm{opt}}(G) = \begin{cases} 1 & \text{if } \log\frac{\mathbb{P}(X_i=1|G)}{\mathbb{P}(X_i=2|G)} \geq \log(\frac{p}{1-p}), \\ 2 & \text{otherwise.} \end{cases} \tag{5.2.2}$$

Let s_0 be uniformly chosen among the vertices of G, independently of all other random variables. The maximal achievable rescaled success probability reduces then to

$$P_{\mathrm{suc}}(T^{\mathrm{opt}}) = \mathbb{P}(T_{s_0}^{\mathrm{opt}}(G)=1|X_{s_0}=1) + \mathbb{P}(T_{s_0}^{\mathrm{opt}}(G)=2|X_{s_0}=2) - 1.$$

We define a notion of solvability for the community detection problem.

Definition 5.2.1

We say that the community detection problem on a given stochastic block model is solvable if

$$\liminf_{n\to\infty} P_{suc}(T^{opt}) > 0.$$

We have another equivalent characterization for solvability given by the following proposition:

Proposition 5.2.1

We have

$$P_{suc}(T^{opt}) = D_{TV}(P_1, P_2),$$

where P_1 and P_2 denote the conditional distribution of the graph G, conditionally respectively on $X_{s_0} = 1$ and $X_{s_0} = 2$, where s_0 is a uniformly chosen random vertex of G. D_{TV} denotes the total variation distance.

Proof. The set of estimators of X_{s_0} is precisely $\{1 + 1_A | A \text{ measurable set}\}$. Consequently

$$\begin{aligned} P_{\text{suc}}(T^{\text{opt}}) &= \sup_{A \text{ measurable}} \mathbb{P}(G \in A^c | X_{s_0} = 1) + \mathbb{P}(G \in A | X_{s_0} = 2) - 1 \\ &= \sup_{A \text{ measurable}} \mathbb{P}_1(A^c) + \mathbb{P}_2(A) - 1 = \sup_{A \text{ measurable}} \mathbb{P}_2(A) - \mathbb{P}_1(A) \\ &= D_{\text{TV}}(P_1, P_2). \end{aligned}$$

□

In our setting (5.1.1), the asymptotic degree distribution of a given vertex is a Poisson random variable. Let $i \in \{1, \dots, n\}$ be a vertex of G, we will denote by d_i its degree. Then we have

$$\begin{aligned} d_i &\xrightarrow[n\to+\infty]{(d)} \text{Poi}(d(pa + (1-p)b)) \text{ conditionally on } \{X_i = 1\}, \\ d_i &\xrightarrow[n\to+\infty]{(d)} \text{Poi}(d(pb + (1-p)c)) \text{ conditionally on } \{X_i = 2\}. \end{aligned}$$

If the asymptotic average degrees differ from class 1 to class 2, we see easily that the problem is solvable.

Lemma 5.2.1

If $pa + (1-p)b \neq pb + (1-p)c$ then the community detection problem is solvable.

Proof. Using the definition of solvability in terms of the total variation distance, we have:

$$\begin{aligned} \liminf_{n\to\infty} D_{\text{TV}}(P_1, P_2) &= \liminf_{n\to\infty} \sup_{\text{event } A} |P_1(A) - P_2(A)| \\ &\geq \liminf_{n\to\infty} \sup_{B\subset\mathbb{N}} |P_1(\{d_{s_0} \in B\}) - P_2(\{d_{s_0} \in B\})| \\ &= D_{\text{TV}}(\text{Poi}(d(pa + (1-p)b)), \text{Poi}(d(pb + (1-p)c))) \\ &> 0. \end{aligned}$$

□

In the rest of the paper, we will always assume that (5.1.2) is valid, so that the average degree in the graph is d. Most of our results below will be obtained in the limit where first n tends to infinity and then d tends to infinity but the parameter λ will always remain fixed, as well as the parameter $p \in [0, 1/2]$ corresponding to the proportion of nodes in the first community.

We now introduce a variation of the standard community detection problem where a fraction q of the vertices have their labels revealed. This labels correspond to 'side-information' given with the graph. More formally, in this setting the label X_v of each vertex $v \in \{1, \dots, n\}$ is observed with probability $q \in [0, 1]$ independently of everything else and the estimator $T(G, q) = (T_1(G, q), \dots, T_n(G, q))$ is then a function of the observed

graph G and the observed labels. The probability of success is again defined by (5.2.1) where the $T_i(G)$'s are replaced by the $T_i(G,q)$'s and the optimal test $T^{\text{opt}}(G,q)$ is then

$$T_i^{\text{opt}}(G,q) = \begin{cases} 1 & \text{if } \log \frac{\mathbb{P}(X_i=1|G,Y)}{\mathbb{P}(X_i=2|G,Y)} \geq \log(\frac{p}{1-p}) \\ 2 & \text{otherwise,} \end{cases}$$

where Y denotes the (random) set of observed vertices. Note that we have

$$\liminf_{n\to\infty} P_{\text{suc}}(T^{\text{opt}}(G,q)) \geq q \geq 0,$$

so that the notion of solvability of Definition 5.2.1 does not make sense in this case where we have some side-information.

We end this section by some technical definitions. In order to state our main results, we need to define the following function F from $\mathbb{R}_{\geq 0}$ to $\mathbb{R}$:

$$F(\mu) = \frac{\lambda}{(1-p)^2}\mathbb{E}\left[\frac{1}{p+(1-p)\exp(\sqrt{\mu}Z - \mu/2)} - 1\right],$$

where $Z \sim \mathcal{N}(0,1)$. Note that F is also a function of the parameters λ and p which are considered as fixed.

Definition 5.2.2 (*Spinodal curve*)

The spinodal curve is defined as the function

$$\lambda_{sp} : p \mapsto \sup\left\{\lambda \geq 0 \mid 0 \text{ is the unique fixed point of } F\right\}. \tag{5.2.3}$$

Let us define

$$p^* = \frac{1}{2} - \frac{1}{2\sqrt{3}}. \tag{5.2.4}$$

The spinodal curve λ_{sp} and p^* are represented on Figure 5.2. p^* corresponds to the critical value of p below which λ_{sp} goes strictly below the 'Kesten-Stigum' line $\lambda = 1$. As we can see on the phase diagram Figure 5.2, a "hard region" appears when $p < p^*$. The following conjecture shows that λ_{sp} is well defined and summarizes its main properties.

Conjecture 5.2.1

(i) If $\lambda > 1$, then F has two fixed points: 0 and $\alpha > 0$. Moreover, 0 is unstable and α is stable.

(ii) For $p^* \leq p \leq 1/2$, we have $\lambda_{sp}(p) = 1$.

(iii) For $0 \leq p < p^*$, we have $\lambda_{sp}(p) < 1$ and if $\lambda_{sp}(p) < \lambda < 1$, then F has three fixed points: $0 < \beta < \alpha$. Moreover, 0 and α are stable and β is unstable.

The analysis of the function F seems challenging and we were only able to verify Conjecture 5.2.1 numerically.

Proof. The exact value of p^* follows from the following (non-rigorous) argument. a small μ expansion of the function F gives

$$F(\mu) \approx \lambda\mu + \frac{\lambda}{2}(1 - 6p(1-p))\mu^2.$$

Thus, if $1 - 6p(1-p) > 0$ (i.e. if $p < p^*$), then F is convex in a neighborhood of 0. Thus, when $\lambda < 1$, F is likely to have 3 fixed points. □

5.3 Reconstructability above the Kesten-Stigum bound

We first consider the case $\lambda > 1$.

Proposition 5.3.1

If $\lambda > 1$, then we have

$$\limsup_{d\to\infty} \limsup_{n\to\infty} P_{suc}(T^{opt}) \leq 2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1, \tag{5.3.1}$$

where $\alpha > 0$ is the stable fixed point in Conjecture 5.2.1 (i). Moreover, for all $0 < q < 1$, we have

$$\liminf_{d\to\infty} \liminf_{n\to\infty} P_{suc}(T^{opt}(G, q)) \geq 2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1.$$

Proposition 5.3.1 will follow from Corollary 5.7.1 and Corollary 5.7.2 below.

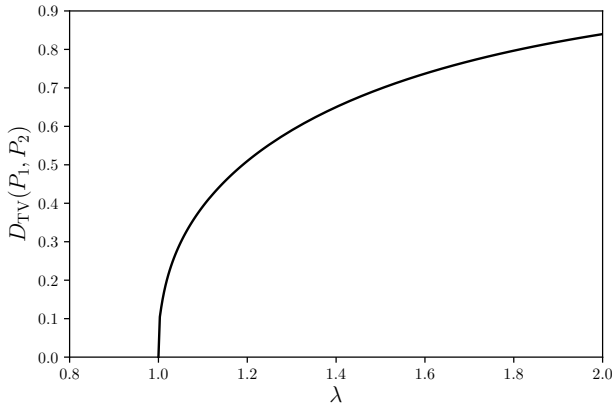


Figure 5.3: Lower bound for the probability to recover the true label of a typical vertex by an optimal local algorithm with side information for $p > p^*$ as a function of λ (i.e. function $2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1$ for $p = 0.25 > p^*$).

In words, we see that if a vanishing fraction of the labels is revealed, then the probability to recover the true label of a typical vertex by the optimal algorithm is $2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1$. Indeed, we believe that (5.3.1) should be an equality. On Figure 5.3, we give a drawing of this curve as a function of λ for $p = 0.25 > p^*$ and on Figure 5.4 for $p = 0.005 < p^*$. Note that at this stage, we only gave an interpretation of the curve for $\lambda > 1$. We deal with the case $\lambda < 1$ in the next section.

Before that, we give a result which shows that if a vanishing fraction of the labels is revealed then the optimal recovery is achieved by a local algorithm. Similar results in the case where (5.1.2) does not hold have been proved in [157]. In the large degree regime, our result improves Proposition 3 in [120] which deals only with the case $p = 0.5$ and λ larger than a large constant C. The fact that local algorithms are very efficient as soon as $q > 0$ (even optimal in the limit $q \to 0$) leads to linear time algorithms for community detection (when some labels are revealed). Indeed from a practical perspective, we believe that our analysis carries over to the labeled stochastic block model [101, 131]. It is then possible to devise new clustering algorithms based on a similarity graph which are shown to be optimal for a wide range of models [187] and also local semi-supervised learning clustering algorithms, see [186] for more details in this direction.

We now define local algorithms. For an integer t, a test $T(G,q) = (T_1(G,q), \dots, T_n(G,q))$ is t-local if each $T_i(G,q)$ is a function of the graph $B_t(G,i)$ induced by the vertices of G whose distance from i is at most t. We denote by Loc_t the set of t-local tests and by $\mathrm{Loc} = \cup_{t\geq 0}\mathrm{Loc}_t$ the set of local tests.

Proposition 5.3.2

If $\lambda > 1$, then we have for all $0 < q \leq 1$,

$$\sup_{T\in Loc} \lim_{d\to\infty} \lim_{n\to\infty} P_{suc}(T(G,q)) \geq 2\mathbb{P}(\mathcal{N}(\alpha/2,\alpha) > 0) - 1,$$

where $\alpha > 0$ is the stable fixed point in Conjecture 5.2.1 (i).

Proposition 5.3.2 will follow from Corollary 5.7.2. Note in particular that as a vanishing fraction of labels is revealed, i.e. $q \to 0$, the best local algorithm performs at least as well as the optimal algorithm with no revealed labels. An explicit description of an optimal local test (obtained using 'belief-propagation') is given in the proof of Corollary 5.7.2.

5.4 Non reconstructability below the spinodal curve

We now state our second main result which states that reconstruction is impossible below the spinodal curve.

Proposition 5.4.1

If $\lambda < \lambda_{sp}(p)$ then

$$\lim_{d\to\infty} \lim_{n\to\infty} P_{suc}(T^{opt}) = 0.$$

Proposition 5.4.1 will follow from Corollary 5.7.1. If

$$\lambda_c(p) \stackrel{\text{def}}{=} \inf\left\{\lambda \geq 0 \,\middle|\, \text{community detection is solvable}\right\}$$

denotes the solvability threshold, then Proposition 5.4.1 implies that $\lambda_{sp}(p) \leq \lambda_c(p)$. Moreover thanks to [36], the Kesten-Stigum threshold is an upper bound on the solvability threshold so that we have $\lambda_c(p) \leq 1$. For $p \geq p^*$ defined in (5.2.4), the spinodal curve is equal to the Kesten-Stigum threshold by Conjecture 5.2.1 (ii), so that we have in this case $\lambda_c(p) = 1$ and moreover as soon as the community detection problem is solvable, it is solvable in polynomial time thanks to the results in [36]. Figure 5.3 is valid for $\lambda < 1$. However for $p < p^*$, there is a gap between the spinodal curve and Kesten-Stigum threshold and we conjecture that $\lambda_{sp}(p) < \lambda_c(p) < 1$, see Figure 5.2. In the case of dense graphs (where the average degree d is of order n), the value of $\lambda_c(p)$ has been computed in the recent works [129] and [18]. We conjecture that their expression (used in Figure 5.2) is still valid for sparse graphs in the large degree regime.

Our next result states that it is possible to reconstruct the communities below the spinodal curve, given a sufficient amount of side-information.

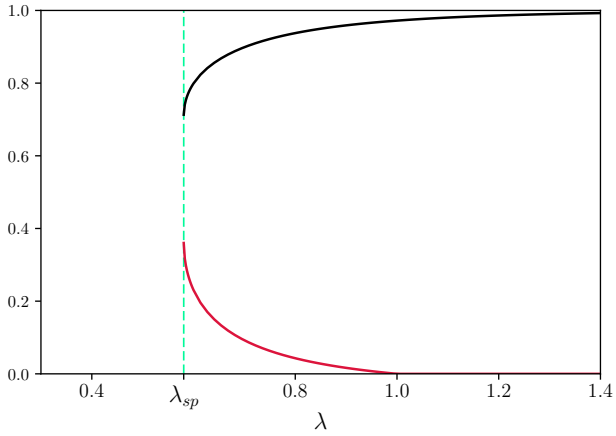


Figure 5.4: Necessary fraction of revealed labels (red) and corresponding lower bound of probability to recover the true label of a typical vertex by an optimal local algorithm (black) for $p < p^*$, i.e. functions $\frac{\beta p(1-p)}{\lambda}$ (in red) and $2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1$ (in black) appearing in Proposition 5.4.2 as a function of λ for $p = 0.05 < p^*$.

Proposition 5.4.2

Consider the case where $p < p^*$ and $\lambda_{sp}(p) < \lambda < 1$. As soon as $q > \frac{\beta p(1-p)}{\lambda}$, we have

$$\begin{aligned}\lim_{d\to\infty} \liminf_{n\to\infty} P_{suc}(T^{opt}(G, q)) &\geq \sup_{T\in Loc} \lim_{d\to\infty} \liminf_{n\to\infty} P_{suc}(T(G, q)) \\ &\geq 2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1,\end{aligned}$$

where $0 < \beta < \alpha$ are the fixed points defined in Conjecture 5.2.1 (iii).

Proposition 5.4.2 will follow from Corollary 5.7.2. In the regime of Proposition 5.4.2 ($\lambda_{sp}(p) < \lambda < 1$ and $q > \frac{\beta p(1-p)}{\lambda}$), we believe that local algorithms are indeed optimal. Figure 5.4 illustrates the case $p < p^*$ with $p = 0.05$ for which we have $\lambda_{sp}(0.05) \approx 0.58$. Also, if the number of revealed entries is sufficiently high (i.e. above the red curve) then local algorithms provide a great improvement in the probability of successfully recovering the label of a typical vertex (the black curve). A description of a local algorithm (belief-propagation) achieving the lower bound in Proposition 5.4.2 is provided in the proof.

5.5 Reconstruction on trees

We will first concentrate on the reconstruction of the labels on trees. The tree structure makes the analysis simpler and allows to deduce results for the stochastic block model, because the SBM is asymptotically locally tree-like. In this section we are going to state the analogous of the well known (see for instance [211]) local convergence of the Erdős-Rényi random graph towards the Galton-Watson branching process, in terms of labeled graphs. The labeled stochastic block model (G, X) will converge locally towards a random labeled tree. We have to introduce first the notion of pointed labeled graphs.

Definition 5.5.1 (*Pointed labeled graphs*)

- *A pointed labeled graph is a triple $G = (g, s_0, x)$ where g is a countable, locally finite and connected graph, s_0 is a distinguished vertex of g called the root of the graph and $x = (x_s)_{s \in V_g} \in \{1, 2\}^{V_g}$ are the labels of the vertices.*
- *Two pointed labeled graphs are equivalent if there exists a graph isomorphism between them, that preserves the root and the labels.*
- *We define, for $r \in \mathbb{N}$, $[G]_r$, the ball of radius r of G, as the pointed labeled graph induced by the root of G and all the vertices at distance at most r from the root.*

The randomly rooted stochastic block model (G, s_0, X) with parameters (M, n, p) is therefore a random pointed labeled graph, that we will denote SBM_n from now. We will also be interested in a second family of random pointed labeled graphs, that will correspond to the local limits of stochastic block models.

Definition 5.5.2 (*Labeled Poisson Galton-Watson branching process*)

Let $A = \begin{pmatrix} \delta & 1-\delta \\ 1-\delta' & \delta' \end{pmatrix}$ (where $\delta, \delta' \in [0, 1]$) be a transition matrix. The labeled Poisson Galton-Watson branching process with parameters (A, p, d) is a random pointed labeled graph (T, s_0, X), where

- *(T, s_0) is a Galton-Watson tree with offspring distribution $Poi(d)$ rooted at s_0.*
- *The labels X of the vertices of T are then chosen as follows:*
  1. *The label of the root $X_{s_0} \in \{1, 2\}$ is chosen accordingly to $\mathbb{P}(X_{s_0} = 1) = p$.*
  2. *Given the label X_p of the parent p of a node s, the probability that $X_s = i \in \{1, 2\}$ is equal to $A_{X_p, i}$ independently from all other random variables.*

In the following, we will denote $\mathrm{GW} = (T, s_0, X)$, the labeled Galton-Watson branching process with parameters (R, d, p) with R defined by (5.1.3). The next well known result states that SBM_n converges locally toward GW.

Theorem 5.5.1

Let f be a (positive or bounded) function of pointed labeled graphs, that depends only on the ball of radius r. Then

$$\mathbb{E}\Big[f(SBM_n)\Big] \xrightarrow[n \to \infty]{} \mathbb{E}\Big[f(GW)\Big].$$

A main ingredient for our proof will therefore be the analysis of this well-studied problem of reconstruction on trees [82, 156, 153]. In the rest of this section, we define the reconstruction problem on trees and give the required results.

We consider here $\mathrm{GW} = (T, s_0, X)$ the labeled Poisson Galton-Watson branching process with parameters (R, p, d). We denote $L_n = \{v \in V_T \mid d(s_0, v) = n\}$, the set of vertices at distance n from the root. We define then $X^{(n)} = (X_s)_{s \in L_n}$ and $c^{(n)} = (c_1^{(n)}, c_2^{(n)}) = (\#\{s \in L_n \mid X_s = i\})_{i=1,2}$. We also define a random subset E_n of the nodes at depth n

as follows: let $q \in [0,1]$ and for $n \in \mathbb{N}$, let E_n be the random subset of L_n obtained by including in E_n each vertex $s \in L_n$ independently with probability q.

We have three kinds of reconstruction problems.

Definition 5.5.3 (*Solvability, q-solvability and census solvability*)

We say that the reconstruction problem is solvable if

$$\liminf_{n\to\infty} D_{TV}(P_1^{(n)}, P_2^{(n)}) > 0,$$

where $P_i^{(n)}$ denotes the conditional distribution of $(T, s_0, X^{(n)})$ given $X_{s_0} = i$. One defines analogously q-solvability (respectively census solvability) by replacing $P_i^{(n)}$ by $P_i^{(n,q)}$ (respectively $\tilde{P}_i^{(n)}$), the conditional distribution of $(T, s_0, (X_s)_{s\in E_n})$ (respectively $(T, s_0, c^{(n)})$) given $X_{s_0} = i$.

Solvability corresponds thus to the special case $q = 1$. Obviously, census solvability and q-solvability imply solvability, but we will see that solvability does not always imply census solvability.

Similarly to the stochastic block model case, this characterization of solvability in terms of total variation can be rewritten in terms of the maximal achievable success probability for the estimation of X_{s_0}, given $X^{(n)}$ (or $c^{(n)}$). We define the rescaled success probability of an estimator T as

$$P_{\text{suc}}(T) = \mathbb{P}(T(X^{(n)}) = 1 | X_{s_0} = 1) + \mathbb{P}(T(X^{(n)}) = 2 | X_{s_0} = 2) - 1.$$

The maximal rescaled success probability is then defined as $\Delta_n = \sup_T P_{\text{suc}}(T)$ where the supremum is taken over all measurable function of $X^{(n)}$. Even though we defined these quantities for the solvability problem, these definitions and the following result can be straightforwardly extended to q-solvability and census-solvability. The following lemma is the analog of Proposition 5.2.1.

Lemma 5.5.1

$$\Delta_n = D_{TV}(P_1^{(n)}, P_2^{(n)}).$$

We recall here the census-solvability criterion for our particular case (which is a straightforward extension of the results presented in [153]).

Theorem 5.5.2

We consider the Poisson Galton-Watson branching process with parameter (R, p, d). If $\lambda > 1$, then the problem is census-solvable and q-solvable for all $0 < q \leq 1$. If $\lambda < 1$, then the problem is not census-solvable.

In the large d limit, we are able to get more quantitative results. We define

$$P_{\text{opt}} = \lim_{d\to\infty} \lim_{n\to\infty} D_{\text{TV}}(P_1^{(n)}, P_2^{(n)}),$$
$$P_{\text{opt}}^{(q)} = \lim_{d\to\infty} \lim_{n\to\infty} D_{\text{TV}}(P_1^{(n,q)}, P_2^{(n,q)}).$$

Proposition 5.5.1

We consider the Poisson Galton-Watson branching process with parameter (R, p, d).

$$P_{opt} = \begin{cases} 2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1 > 0 & \text{if } \lambda > \lambda_{sp}(p), \\ 0 & \text{if } \lambda < \lambda_{sp}(p), \end{cases}$$

where α is the stable fixed point in Conjecture 5.2.1. If $\lambda_{sp} < \lambda < 1$, then we have

$$P_{opt}^{(q)} = \begin{cases} 2\mathbb{P}(\mathcal{N}(\alpha/2, \alpha) > 0) - 1 > 0 & \text{if } q > \frac{\beta p(1-p)}{\lambda}, \\ 0 & \text{if } q < \frac{\beta p(1-p)}{\lambda}, \end{cases}$$

where α and β are the fixed points defined in Conjecture 5.2.1 (iii).

In particular, Proposition 5.5.1 shows that the spinodal curve is the solvability threshold for the reconstruction on trees. Proposition 5.5.1 is proved at the end of the next section.

5.6 Cavity method on trees

To compute the optimal success probability $\Delta_n = D_{\mathrm{TV}}(P_1^{(n)}, P_2^{(n)})$ for the reconstruction problem on trees, we need to study the behaviour of the optimal estimator. This estimator is computed, similarly to (5.2.2), using the marginal distributions of the labels. We aim therefore at computing these marginals.

Our approach here is closely related to the one of [149] which studies the problem of finding one single community. We establish rigorously the 'cavity equations', a recursive method to compute marginals, that originates from statistical physics.

We consider here the labeled branching process $\mathrm{GW} = (T, s_0, X)$ with parameter (R, p, d). In order to obtain quantitative results, we will be interested in the asymptotic of large degrees $d \to \infty$ while λ remains fixed. We also define $\epsilon = 1 - b = \sqrt{\frac{\lambda}{d}}$. We have then

$$a = 1 + \frac{1-p}{p}\epsilon, \quad b = 1 - \epsilon \quad \text{and} \quad c = 1 + \frac{p}{1-p}\epsilon. \tag{5.6.1}$$

5.6.1 The cavity recursions

Let $r \in \mathbb{N}$. For a given vertex s of $\mathrm{GW} = (T, s_0, X)$, we note $\mathcal{T}_s$ the subtree induced by s and its progeny.

With a slight abuse of notation, we write $p(x) = p^{\mathbb{1}_{x=1}}(1-p)^{\mathbb{1}_{x=2}}$. We define also $\psi(x, y) = a^{\mathbb{1}_{y=x=1}} b^{\mathbb{1}_{y\neq x}} c^{\mathbb{1}_{y=x=2}}$. We now introduce the belief-propagation algorithm for label reconstruction. This algorithm computes 'messages', that approximate the marginal distributions of the labels. The message of the vertex $s \in [T]_r$ is defined as the following function from $\{1, 2\} \to \mathbb{R}$:

$$\nu_r^s : x_s \mapsto \begin{cases} p(x_s)\mathbb{1}_{x_s = X_s} & \text{if } s \in E_r, \\ p(x_s) & \text{if } s \in L_r \setminus E_r, \\ p(x_s) \prod_{s \to v} \sum_{x_v} \psi(x_s, x_v)\nu_r^v(x_v) & \text{if } s \notin L_r. \end{cases} \tag{5.6.2}$$

Lemma 5.6.1

For all $s \in [T]_r \setminus L_r$,

$$\nu_r^s(x_s) \propto p(x_s) \sum_{\substack{(x_v)_{v\in\mathcal{T}_s\cap[T]_r\setminus\{s\}} \\ (x_v)_{v\in E_r}=(X_v)}} \prod_{(i\to j)\in[T]_r\cap\mathcal{T}_s} R_{x_i,x_j}, \tag{5.6.3}$$

where $\propto$ means equality up to a multiplicative constant that is independent of x_s.

Proof. We show this lemma by induction on the depth of $d_s \overset{\text{def}}{=} \mathcal{T}_s \cap [T]_r$. Since $s \notin L_r$, we have $d_s \geq 1$. We consider the following recursion hypothesis:

$$\mathcal{H}(d) : \text{for all } s \in [T]_r \text{ such that } d_s \leq d \text{ (5.6.3) holds.}$$

$\mathcal{H}(1)$ follows from the definition of ν_r^s. Suppose now that $\mathcal{H}(d)$ holds for some $d \geq 1$. Let $s \in [T]_r$ such that $d_s = d + 1$. Then, by induction

$$\begin{aligned}
\nu_r^s(x_s) &= p(x_s) \sum_{(x_v)_{s\to v}} \prod_{s\to v} \psi(x_s, x_v)\nu_r^v(x_v) \\
&= p(x_s) \sum_{(x_v)_{s\to v}} \prod_{s\to v} \psi(x_s, x_v)p(x_v) \sum_{\substack{(x_u)_{u\in\mathcal{T}_v\cap[T]_r\setminus\{v\}} \\ (x_u)_{u\in E_r}=(X_u)}} \prod_{(i\to j)\in[T]_r\cap\mathcal{T}_v} R_{x_i,x_j} \\
&= p(x_s) \sum_{(x_v)_{s\to v}} \prod_{s\to v} R_{x_s,x_v} \sum_{\substack{(x_u)_{u\in\mathcal{T}_v\cap[T]_r\setminus\{v\}} \\ (x_u)_{u\in E_r}=(X_u)}} \prod_{(i\to j)\in[T]_r\cap\mathcal{T}_v} R_{x_i,x_j} \\
&= p(x_s) \sum_{\substack{(x_v)_{v\in\mathcal{T}_s\cap[T]_r\setminus\{s\}} \\ (x_v)_{v\in E_r}=(X_v)}} \prod_{(i\to j)\in[T]_r\cap\mathcal{T}_s} R_{x_i,x_j},
\end{aligned}$$

which proves that $\mathcal{H}(d+1)$ holds. □

The next lemma states the well known fact that belief-propagation computes the exact marginals on trees.

Lemma 5.6.2

$$\nu_r^{s_0}(x_{s_0}) = \mathbb{P}(X_{s_0} = x_{s_0} | T, E_r, (X_v)_{v\in E_r}).$$

Proof. The structure of T outside of $[T]_r$ does not provide any information about the labels of the vertices, thus, using Lemma 5.6.1

$$\begin{aligned}
\mathbb{P}(X_{s_0} = x_{s_0} \,|\, T, E_r, (X_s)_{s\in E_r}) &= \mathbb{P}(X_{s_0} = x_{s_0} \,|\, [T]_r, E_r, (X_s)_{s\in E_r}) \\
&= p(x_{s_0}) \sum_{\substack{(x_s)_{s\in[T]_r\setminus\{s_0\}} \\ (x_s)_{s\in E_r}=(X_s)}} \prod_{(i\to j)\in[T]_r} R_{x_i,x_j} \\
&= \nu_r^{s_0}(x_{s_0}).
\end{aligned}$$

□

If we write $\xi_r^s = \log(\frac{\nu_r^s(1)}{\nu_r^s(2)})$, the recursive definition (5.6.2) of the messages gives for $r \geq 1$,

$$\xi_r^{s_0} = h + \sum_{s,s_0\to s} f(\xi_r^s), \tag{5.6.4}$$

where $h = \log(\frac{p}{1-p})$ and $f : x \mapsto \log \frac{ae^x+b}{be^x+c}$.

Definition 5.6.1

We define P_r as the law of $\xi_r^{s_0} = \log \frac{\mathbb{P}(X_{s_0}=1|T,E_r,(X_s)_{s\in E_r})}{\mathbb{P}(X_{s_0}=2|T,E_r,(X_s)_{s\in E_r})}$. We denote $P_r^{(1)}$ and $P_r^{(2)}$ the laws of $\xi_r^{s_0}$ respectively conditionally on $\{X_{s_0}=1\}$ and $\{X_{s_0}=2\}$.

For $r \geq 1$. Let us work conditionally on $\{X_{s_0}=1\}$. By the branching property of GW, conditionally on $\{s_1,\dots,s_L\}$ being the children of s_0, the random variables $\xi_r^{s_1},\dots,\xi_r^{s_L}$ are independent, identically distributed according to

$$paP_{r-1}^{(1)} + (1-p)bP_{r-1}^{(2)}.$$

Indeed, conditionally on $\{X_{s_0}=1\}$, $s_1,\dots,s_L \overset{\text{i.i.d.}}{\sim} pa\delta_1 + (1-p)b\delta_2$. Equation (5.6.4) leads therefore to the following distributional recursions (the second one is obtained by the same arguments):

Proposition 5.6.1 (*Cavity equations*)

For all $r \geq 1$,

$$\xi_r^{(1)} \overset{(d)}{=} h + \sum_{i=1}^{L_{1,1}} f(\xi_{r-1,i}^{(1)}) + \sum_{i=1}^{L_{1,2}} f(\xi_{r-1,i}^{(2)}), \tag{5.6.5}$$

where $\xi_r \sim P_r^{(1)}$, $L_{1,1} \sim Poi(pad)$, $L_{1,2} \sim Poi((1-p)bd)$, $\xi_{r-1,i}^{(1)} \sim P_{r-1}^{(1)}$, $\xi_{r-1,i}^{(2)} \sim P_{r-1}^{(2)}$, and all these variables are independent.

$$\xi_r^{(2)} \overset{(d)}{=} h + \sum_{i=1}^{L_{2,1}} f(\xi_{r-1,i}^{(1)}) + \sum_{i=1}^{L_{2,2}} f(\xi_{r-1,i}^{(2)}), \tag{5.6.6}$$

where $\xi_r \sim P_r^{(2)}$, $L_{2,1} \sim Poi(pbd)$, $L_{2,2} \sim Poi((1-p)cd)$, $\xi_{r-1,i}^{(1)} \sim P_{r-1}^{(1)}$, $\xi_{r-1,i}^{(2)} \sim P_{r-1}^{(2)}$, and all these variables are independent.

5.6.2 Gaussian limit

We are interested in the limit distributions of $P_r^{(1)}$ and $P_r^{(2)}$, because they encode the marginal distributions of the labels and thus allow to derive the optimal reconstruction performances. We study the recursions (5.6.5) and (5.6.6). More precisely, we show that $P_r^{(1)}$ and $P_r^{(2)}$ converge in Wasserstein sense toward Gaussian distributions when $d \to \infty$ and λ remains fixed.

Proposition 5.6.2

For all $r \geq 1$,

$$P_r^{(1)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h + \frac{\mu_r}{2}, \mu_r),$$
$$P_r^{(2)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h - \frac{\mu_r}{2}, \mu_r),$$

where $\xrightarrow{W_2}$ denote the convergence with respect to W_2, the Wasserstein distance of order 2, and $(\mu_r)_{r\geq 1}$ is defined by

$$\begin{cases} \mu_1 &= \frac{q\lambda}{p(1-p)} \\ \mu_{k+1} &= F(\mu_k). \end{cases} \tag{5.6.7}$$

Proposition 5.6.2 is proved in Section 5.8. Proposition 5.5.1 is now a consequence of Conjecture 5.2.1 and the following corollary.

Corollary 5.6.1

For all $0 \leq q \leq 1$,

$$\lim_{r\to\infty} \lim_{d\to\infty} D_{TV}(P_1^{(r,q)}, P_2^{(r,q)}) = 2\mathbb{P}(\mathcal{N}(\mu_\infty/2, \mu_\infty) > 0) - 1,$$

where μ_∞ *is the limit of* (μ_r) *defined in* (5.6.7).

Proof. The optimal test according the performance measure P_{suc} is

$$T_r^{(\text{GW})}(T, E_r, (X_s)_{s\in E_r}) = \begin{cases} 1 & \text{if } \xi_r \geq \log \frac{p}{1-p}, \\ 2 & \text{otherwise.} \end{cases} \tag{5.6.8}$$

Analogously to Lemma 5.5.1, we have

$$\begin{aligned} D_{\text{TV}}(P_1^{(r,q)}, P_2^{(r,q)}) &= P_{\text{suc}}(T_r^{(\text{GW})}) \\ &= \mathbb{P}\left(\xi_r^{(1)} \geq \log \frac{p}{1-p}\right) + \mathbb{P}\left(\xi_r^{(2)} < \log \frac{p}{1-p}\right) - 1 \\ &\xrightarrow[d\to\infty]{} 2\mathbb{P}(\mathcal{N}(\mu_r/2, \mu_r) > 0) - 1 \\ &\xrightarrow[r\to\infty]{} 2\mathbb{P}(\mathcal{N}(\mu_\infty/2, \mu_\infty) > 0) - 1, \end{aligned}$$

where we have used Proposition 5.6.2. □

5.7 Proofs for the stochastic block model

In this section, we apply the results that we obtained for the reconstruction on the branching process to derive bounds for the community detection problem on the stochastic block model.

Consider the case, where a fraction $0 \leq q \leq 1$ is revealed: one observes the graph G and additionally each label X_v with probability q, for $1 \leq v \leq n$, independently of everything else. Let us denote $E_G = \{1 \leq v \leq n \mid X_v \text{ is revealed}\}$.

Let s_0 be uniformly chosen among the vertices of G. For $r \geq 0$ we define (analogously to the case of the branching process) $E_{G,r} = \partial[G]_r \cap E_G$, i.e. the vertices at the boundary of the ball of center s_0 and radius r, whose label has been revealed. Define

$$\xi_{G,r} = \log\left(\frac{\mathbb{P}(X_{s_0} = 1 | [G]_r, (X_s)_{s\in E_{G,r}})}{\mathbb{P}(X_{s_0} = 2 | [G]_r, (X_s)_{s\in E_{G,r}})}\right). \tag{5.7.1}$$

Let us denote $P_{G,r}^{(1)}$ and $P_{G,r}^{(2)}$ the laws of $\xi_{G,r}$ conditionally respectively on $X_{s_0} = 1$ and $X_{s_0} = 2$.

Proposition 5.7.1

$$\text{For } i = 1, 2; \quad P_{G,r}^{(i)} \xrightarrow[n\to\infty]{} P_r^{(i)},$$

where $P_r^{(i)}$ *is defined in Definition* 5.6.1.

Proposition 5.7.1 is a consequence of the local convergence of the stochastic block model towards a branching process (Theorem 5.5.1).

Proof. Let us define

$$\varphi_n : (\widetilde{G}, \widetilde{s}_0, \widetilde{X}, \widetilde{E}_r) \mapsto \log \frac{\mathbb{P}(X_{s_0} = 1 | [G]_r = [\widetilde{G}]_r, (X_s)_{s \in E_{G,r}} = (\widetilde{X}_s)_{s \in \widetilde{E}_r}, E_{G,r} = \widetilde{E}_r)}{\mathbb{P}(X_{s_0} = 2 | [G]_r = [\widetilde{G}]_r, (X_s)_{s \in E_{G,r}} = (\widetilde{X}_s)_{s \in \widetilde{E}_r}, E_{G,r} = \widetilde{E}_r)}$$

and

$$\varphi_\infty : (\widetilde{G}, \widetilde{s}_0, \widetilde{X}, \widetilde{E}_r) \mapsto \begin{cases} \log \frac{\mathbb{P}(X_{s_0}=1 | [T]_r = [\widetilde{G}]_r, (X_s)_{s \in E_r} = (\widetilde{X}_s)_{s \in \widetilde{E}_r}, E_r = \widetilde{E}_r)}{\mathbb{P}(X_{s_0}=2 | [T]_r = [\widetilde{G}]_r, (X_s)_{s \in E_r} = (\widetilde{X}_s)_{s \in \widetilde{E}_r}, E_r = \widetilde{E}_r)} & \text{if } [\widetilde{G}]_r \text{ is a tree,} \\ 0 & \text{otherwise.} \end{cases}$$

Let $(\widetilde{G}, \widetilde{s}_0, \widetilde{X})$ be a fixed pointed labeled graph such that $[\widetilde{G}]_r$ is a tree. Let $\widetilde{E}_r$ be a subset of the vertices in $\partial[\widetilde{G}]_r$. A straightforward extension of Theorem 5.5.1 gives us

$$\varphi_n(\widetilde{G}, \widetilde{s}_0, \widetilde{X}, \widetilde{E}_r) \xrightarrow[n \to \infty]{} \varphi_\infty(\widetilde{G}, \widetilde{s}_0, \widetilde{X}, \widetilde{E}_r). \tag{5.7.2}$$

Another consequence of the local convergence of (SBM_n) toward GW is that one can couple $(\mathrm{SBM}_n, E_{G,r})_n$ and (GW, E_r) on a probability space such that there exists $n_0 \in \mathbb{N}$ such that

$$\forall n \geq n_0, [\mathrm{SBM}_n, E_{G,r}]_r = [\mathrm{GW}, E_r]_r.$$

Let $n \geq n_0$

$$\varphi_n(\mathrm{SBM}_n, E_{G,r}) = \varphi_n(\mathrm{GW}, E_r) \xrightarrow[n \to \infty]{} \varphi_\infty(\mathrm{GW}, E_r).$$

On this probability space $\varphi_n(\mathrm{SBM}_n, E_{G,r})$ converges almost surely to $\varphi_\infty(\mathrm{GW}, E_r)$, hence the convergence of the conditional distributions. □

Define the local test

$$T_r^{\mathrm{loc}}(G, X) = \begin{cases} 1 & \text{if } \xi_{G,r} \geq \log \frac{p}{1-p}, \\ 2 & \text{otherwise.} \end{cases} \tag{5.7.3}$$

T_r^{loc} is the optimal r-local test, with side information $(X_s)_{s \in E_{G,r}}$. Note that $\xi_{G,r}$ (and thus T_r^{loc}) is computed by the belief propagation algorithm. Using the results on the branching process, we are now able to fully characterize the performance of T^{loc}.

Proposition 5.7.2

For all $0 \leq q \leq 1$,

$$\lim_{r \to \infty} \lim_{d \to \infty} \lim_{n \to \infty} P_{suc}(T_r^{loc}) = 2\mathbb{P}(\mathcal{N}(\mu_\infty/2, \mu_\infty) > 0) - 1,$$

where μ_∞ *is the limit of the sequence defined by* (5.6.7).

Proof. Using Propositions 5.7.1 and 5.6.2

$$\begin{cases} P_{G,r}^{(1)} \xrightarrow[n \to \infty]{(d)} P_r^{(1)} \xrightarrow[d \to \infty]{(d)} \mathcal{N}(h + \frac{\mu_r}{2}, \mu_r), \\ P_{G,r}^{(2)} \xrightarrow[n \to \infty]{(d)} P_r^{(2)} \xrightarrow[d \to \infty]{(d)} \mathcal{N}(h - \frac{\mu_r}{2}, \mu_r), \end{cases}$$

where μ_r is defined by the recursion (5.6.7). Therefore (recall that $h = \log \frac{p}{1-p}$)

$$\begin{aligned} P_{\text{suc}}(T_r^{\text{loc}}) &= \mathbb{P}(\xi_{G,r} \geq h | X_{s_0} = 1) + \mathbb{P}(\xi_{G,r} < h | X_{s_0} = 2) - 1 \\ &\xrightarrow[n\to\infty]{} P_{G,r}^{(1)}(\xi_{G,r} \geq h) + P_{G,r}^{(2)}(\xi_{G,r} < h) - 1 \\ &\xrightarrow[d\to\infty]{} 2\mathbb{P}(\mathcal{N}(\mu_r/2, \mu_r) > 0) - 1 \\ &\xrightarrow[r\to\infty]{} 2\mathbb{P}(\mathcal{N}(\mu_\infty/2, \mu_\infty) > 0) - 1. \end{aligned}$$

□

5.7.1 Upper bound

We can now deduce an upper bound on the optimal performance for the community detection problem.

Corollary 5.7.1

We have:

$$\limsup_{d\to\infty} \limsup_{n\to\infty} P_{suc}(T^{opt}) \leq 2\mathbb{P}(\mathcal{N}(\mu_\infty/2, \mu_\infty) > 0) - 1,$$

where μ_∞ is the limit of the sequence defined by (5.6.7) in the case $q = 1$.

Proposition 5.4.1 and the first part of Proposition 5.3.1 are then consequences of this corollary:

- when $\lambda < \lambda_{sp}(p)$, $\mu_\infty = 0$ and consequently $\lim_{d\to\infty} \limsup_{n\to\infty} P_{\text{suc}}(T^{\text{opt}}) = 0$.
- when $\lambda > 1$, then $\mu_\infty = \alpha$, hence the first bound of Proposition 5.3.1.

Proof. Let $r > 0$. Let s_0 be uniformly chosen from the vertices of G. We aim at estimating X_{s_0} from the rooted graph (G, s_0). As seen in Section 5.2, the optimal test in terms of rescaled success probability P_{suc} is

$$T^{\text{opt}}(G) = \begin{cases} 1 & \text{if } \log(\frac{\mathbb{P}(X_{s_0}=1|G)}{\mathbb{P}(X_{s_0}=2|G)}) \geq \log(\frac{p}{1-p}), \\ 2 & \text{otherwise.} \end{cases}$$

We are going to analyze the oracle

$$T_r^*(G, X) = \begin{cases} 1 & \text{if } \log \frac{\mathbb{P}(X_{s_0}=1|G,(X_s)_{s\in\partial[G]_r})}{\mathbb{P}(X_{s_0}=2|G,(X_s)_{s\in\partial[G]_r})} \geq \log \frac{p}{1-p}, \\ 2 & \text{otherwise.} \end{cases}$$

Obviously, $P_{\text{suc}}(T^{\text{opt}}) \leq P_{\text{suc}}(T_r^*)$. The oracle T_r^* uses extra information $(X_s)_{s\in\partial[G]_r}$ but is a local test, i.e. a test that only depends on the ball of radius r:

$$\xi_{G,r}^* \overset{\text{def}}{=} \log \frac{\mathbb{P}(X_{s_0} = 1|G, (X_s)_{s\in\partial[G]_r})}{\mathbb{P}(X_{s_0} = 2|G, (X_s)_{s\in\partial[G]_r})} = \log \frac{\mathbb{P}(X_{s_0} = 1|[G]_r, (X_s)_{s\in\partial[G]_r})}{\mathbb{P}(X_{s_0} = 2|[G]_r, (X_s)_{s\in\partial[G]_r})}.$$

$\xi_{G,r}^*$ is thus equal to $\xi_{G,r}$ from equation (5.7.1), when $q = 1$. We conclude using Proposition 5.7.2 and the fact $P_{\text{suc}}(T^{\text{opt}}) \leq P_{\text{suc}}(T_r^*)$. □

5.7.2 Lower bound

We now establish a lower bound for estimation when a fraction q of the labels is revealed.

Corollary 5.7.2

For all $0 \leq q \leq 1$,

$$\liminf_{d\to\infty} \liminf_{n\to\infty} P_{suc}(T^{opt}(G,q)) \geq 2\mathbb{P}(\mathcal{N}(\mu_\infty/2, \mu_\infty) > 0) - 1, \tag{5.7.4}$$

where μ_∞ *is the limit of the sequence defined by* (5.6.7).

Proof. Here, we are going to bound $P_{\rm suc}(T^{\rm opt}(G,q))$ from below by the performance of the local test $T_r^{\rm loc}$ (that corresponds to the estimator derived from belief-propagation) defined by (5.7.3). Obviously, $P_{\rm suc}(T^{\rm opt}(G,q)) \geq P_{\rm suc}(T_r^{\rm loc})$ and Proposition 5.7.2 gives then the result. □

The second part of Proposition 5.3.1 follows from this corollary. Indeed, when $\lambda > 1$ and $q > 0$, $\mu_\infty = \alpha > 0$.

Corollary 5.7.2 leads also to proposition 5.4.2: when $q > \frac{\beta p(1-p)}{\lambda}$, then $\mu_1 > \beta$ and thus (μ_k) converges to the fixed point $\alpha > \beta$ of G: $\mu_\infty = \alpha > 0$.

We deduce also Proposition 5.3.2 from the proof of Corollary 5.7.2. Indeed, we will see that the lower bound in (5.7.4) is achieved by a local test.

5.8 Proof of Proposition 5.6.2

We prove Proposition 5.6.2 by induction over $r \geq 1$. Proposition 5.6.2 follows from Lemmas 5.8.1 and 5.8.3 below.

5.8.1 Initialization

First of all, we are going to show that $\xi_1^{(1)}$ and $\xi_1^{(2)}$ converge towards Gaussian distributions.

Lemma 5.8.1

$$\xi_1^{(1)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h + \frac{\mu_1}{2}, \mu_1),$$
$$\xi_1^{(2)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h - \frac{\mu_1}{2}, \mu_1).$$

Proof. We will only prove the convergence for $\xi_1^{(1)}$, the convergence for $\xi_1^{(2)}$ can be obtained analogously. We have

$$\xi_0^{(1)} = \begin{cases} +\infty & \text{with probability } q \\ h & \text{with probability } 1-q, \end{cases} \quad \text{and} \quad \xi_0^{(2)} = \begin{cases} -\infty & \text{with probability } q \\ h & \text{with probability } 1-q. \end{cases}$$

Therefore, the recursion (5.6.5) gives

$$\xi_1^{(1)} \overset{(d)}{=} h + L\log\left(\frac{a}{b}\right) + L'\log\left(\frac{b}{c}\right),$$

where $L \sim \mathrm{Poi}(padq)$ and $L' \sim \mathrm{Poi}((1-p)bdq)$ are independent. By isolating the means

$$\xi_1^{(1)} \overset{(d)}{=} h + (L - padq)\log\left(\frac{a}{b}\right) + (L' - (1-q)bdp)\log\left(\frac{b}{c}\right) + dq\left(pa\log\left(\frac{a}{b}\right) + (1-p)b\log\left(\frac{b}{c}\right)\right). \tag{5.8.1}$$

In the $d \to \infty$ limit, $\log(\frac{a}{b}) = \frac{\epsilon}{p} + \frac{1}{2}\epsilon^2(1-(\frac{1-p}{p})^2) + o(\epsilon^2)$ and $\log(\frac{b}{c}) = -\frac{\epsilon}{1-p} - \frac{1}{2}\epsilon^2(1-(\frac{p}{1-p})^2) + o(\epsilon^2)$, so the last term in (5.8.1) becomes

$$\begin{aligned} dq\left(pa\log\frac{a}{b} + (1-p)b\log\frac{b}{c}\right) &= q\lambda\left(\frac{1}{p} + \frac{2p-1}{2p(1-p)}\right) + o(1) \\ &= \frac{q\lambda}{2p(1-p)} + o(1). \end{aligned}$$

Now, to deal with the remaining terms in (5.8.1) we are going to use the following corollary the Central Limit Theorem.

Lemma 5.8.2

Let $(a_n)_n \in (0,+\infty)^{\mathbb{N}}$, such that $a_n \xrightarrow[n\to\infty]{} +\infty$. Let X_n be a sequence of random variable such that $X_n \sim Poi(a_n)$. Then

$$\frac{1}{\sqrt{a_n}}(X_n - a_n) \xrightarrow[n\to\infty]{W_2} \mathcal{N}(0,1).$$

Proof. We define $v_n \overset{\text{def}}{=} \frac{a_n}{\lceil a_n \rceil} \to 1$. Let $(Y_i^{(n)}) \overset{\text{i.i.d.}}{\sim} \mathrm{Poi}(v_n) - v_n$. By the central limit theorem we have

$$\frac{1}{\sqrt{a_n}}(X_n - a_n) \overset{(d)}{=} \frac{1}{\sqrt{a_n}} \sum_{i=1}^{\lceil a_n \rceil} Y_i^{(n)} \xrightarrow[n\to\infty]{W_2} \mathcal{N}(0,1).$$

□

Applying this result to the terms in (5.8.1), we obtain $(L - padq)\log(\frac{a}{b}) \xrightarrow[d\to\infty]{W_2} \mathcal{N}(0, \frac{\lambda q}{p})$ and $(L' - (1-p)bdq)\log(\frac{b}{c}) \xrightarrow[d\to\infty]{W_2} \mathcal{N}(0, \frac{\lambda q}{1-p})$ and finally

$$\xi_1^{(1)} \xrightarrow[d\to\infty]{W_2} \mathcal{N}(h + \frac{\mu_1}{2}, \mu_1).$$

□

5.8.2 Induction

The following lemma (combined with Lemma 5.8.1) concludes the proof of Proposition 5.6.2 by induction.

Lemma 5.8.3

Suppose that, for a fixed $r \geq 1$ we have

$$\xi_r^{(1)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h + \frac{\mu_r}{2}, \mu_r),$$
$$\xi_r^{(2)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h - \frac{\mu_r}{2}, \mu_r).$$

Then

$$\xi_{r+1}^{(1)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h + \frac{\mu_{r+1}}{2}, \mu_{r+1}),$$
$$\xi_{r+1}^{(2)} \xrightarrow[d\to+\infty]{W_2} \mathcal{N}(h - \frac{\mu_{r+1}}{2}, \mu_{r+1}),$$

where $\mu_{r+1} = F(\mu_r)$.

Proof. We first compute the limits of the mean and the variance of $\xi_{r+1}^{(1,2)}$. To do so, we will need the following well known Wald formulas.

Lemma 5.8.4 (*Wald formulas*)

Let $X_1, \dots, X_n$ be i.i.d. integrable real random variables, and T a $\mathbb{N}$ valued integrable random variable, independent from the X_i. Then $\sum_{i=1}^{T} X_i$ is integrable and

$$\mathbb{E}\left[\sum_{i=1}^{T} X_i\right] = \mathbb{E}[T]\mathbb{E}[X].$$

Moreover, if the variables (X_i) are square integrable and centered

$$Var\left(\sum_{i=1}^{T} X_i\right) = \mathbb{E}[T]\mathbb{E}\left[X^2\right].$$

Using Wald formulas above and Proposition 5.6.1, we obtain

$$\mathbb{E}\xi_{r+1}^{(1)} = h + pad\ \mathbb{E}f(\xi_r^{(1)}) + (1-p)bd\ \mathbb{E}f(\xi_r^{(2)}), \tag{5.8.2}$$
$$\mathbb{E}\xi_{r+1}^{(2)} = h + pbd\ \mathbb{E}f(\xi_r^{(1)}) + (1-p)cd\ \mathbb{E}f(\xi_r^{(2)}). \tag{5.8.3}$$

We will now compute an approximation for the function f in the large d limit. $f(\xi) = \log(\frac{\frac{a}{b}e^{\xi}+1}{e^{\xi}+\frac{c}{b}})$ and we have $\frac{a}{b} = 1 + \frac{1}{p}\epsilon + O(\epsilon^2)$ and $\frac{c}{b} = 1 + \frac{1}{1-p}\epsilon + O(\epsilon^2)$. Thus

$$f(\xi) = \log\left(1 + \epsilon\frac{e^{\xi}}{p(1+e^{\xi})} + O(\epsilon^2)\right) - \log\left(1 + \epsilon\frac{1}{(1-p)(1+e^{\xi})} + O(\epsilon^2)\right).$$

The Taylor-Lagrange formula ensures that

$$\forall a \in [0,1], \forall x \geq 0, \quad \left|\log(1+ax) - ax + \frac{a^2x^2}{2}\right| \leq \frac{1}{3}x^3.$$

Therefore

$$f(x) = \epsilon\frac{e^x}{p(1+e^x)} - \frac{\epsilon^2}{2p^2}\left(\frac{e^x}{1+e^x}\right)^2 - \epsilon\frac{1}{(1-p)(1+e^x)} + \frac{\epsilon^2}{2(1-p)^2}\left(\frac{1}{1+e^x}\right)^2 + O(\epsilon^3) \tag{5.8.4}$$

To simplify the computation, we need the following so-called 'Nishimori condition'.

Lemma 5.8.5

For all continuous bounded function g:

$$\mathbb{E}\left[g(\xi_r^{(2)})\right] = \frac{p}{1-p}\mathbb{E}\left[g(\xi_r^{(1)})e^{-\xi_r^{(1)}}\right].$$

Proof. This is a consequence of Bayes rule.

$$\begin{aligned}
\mathbb{P}(\xi_r^2 \in A) = \mathbb{P}(\xi_r \in A | X_{s_0} = 2) &= \frac{\mathbb{P}(\xi_r \in A, X_{s_0} = 2)}{1-p} \\
&= \frac{1}{1-p}\mathbb{E}(1(\xi_r \in A)\mathbb{P}(X_{s_0} = 2|\mathrm{GW}, E_r, (X_s)_{s\in E_r})) \\
&= \frac{1}{1-p}\mathbb{E}(1(\xi_r \in A)\mathbb{P}(X_{s_0} = 1|\mathrm{GW}, E_r, (X_s)_{s\in E_r})e^{-\xi_r}) \\
&= \frac{p}{1-p}\mathbb{E}(1(\xi_r \in A)e^{-\xi_r}|X_{s_0} = 1).
\end{aligned}$$

□

We use Lemma 5.8.5 to deduce the following identities.

Corollary 5.8.1

$$p\mathbb{E}\frac{e^{\xi_r^{(1)}}}{1+e^{\xi_r^{(1)}}} + (1-p)\mathbb{E}\frac{e^{\xi_r^{(2)}}}{1+e^{\xi_r^{(2)}}} = p,$$

$$p\mathbb{E}\left(\frac{e^{\xi_r^{(1)}}}{1+e^{\xi_r^{(1)}}}\right)^2 + (1-p)\mathbb{E}\left(\frac{e^{\xi_r^{(2)}}}{1+e^{\xi_r^{(2)}}}\right)^2 = p\mathbb{E}\frac{e^{\xi_r^{(1)}}}{1+e^{\xi_r^{(1)}}}.$$

Replacing f by its approximation (5.8.4) in equations (5.8.2) and (5.8.3) and applying Corollary 5.8.1, we obtain that $\mathbb{E}\xi_{r+1}^{(1)} \xrightarrow[d\to+\infty]{} h - \frac{1}{2}F(\mu_r)$ and $\mathbb{E}\xi_{r+1}^{(2)} \xrightarrow[d\to+\infty]{} h + \frac{1}{2}F(\mu_r)$. Similar calculations show that $\mathrm{Var}(\xi_{r+1}^{(1)}) \xrightarrow[d\to+\infty]{} F(\mu_r)$ and $\mathrm{Var}(\xi_{r+1}^{(2)}) \xrightarrow[d\to+\infty]{} F(\mu_r)$.

It remains to show that $\xi_{r+1}^{(1)}$ is converging toward a Gaussian distribution in the Wasserstein sense. We will need the following lemma.

Lemma 5.8.6

For $i = 1, 2$, $\sqrt{d}\ \mathbb{E}[f(\xi_r^{(i)})]$ and $d\ \mathbb{E}[f(\xi_r^{(i)})^2]$ are both converging to constants when $d \to +\infty$.

Proof. $\sqrt{d}f(x) = \sqrt{d}\frac{\epsilon}{p}\frac{e^x}{1+e^x} - \sqrt{d}\frac{\epsilon}{1-p}\frac{1}{1+e^x} + o(1)$. So

$$\sqrt{d}\ \mathbb{E}\Big[f(\xi_r^{(i)})\Big] = \sqrt{d}\frac{\epsilon}{p}\mathbb{E}\left[\frac{e^{\xi_r^{(i)}}}{1+e^{\xi_r^{(i)}}}\right] - \sqrt{d}\frac{\epsilon}{1-p}\mathbb{E}\left[\frac{1}{1+e^{\xi_r^{(i)}}}\right] + o(1),$$

and all the terms in this expression are converging by weak convergence of $\xi_r^{(i)}$ and the fact that $\epsilon \sim \sqrt{\frac{\lambda}{d}}$. The other limit is proved analogously. □

Compute now

$$\begin{aligned}
\xi_{r+1}^{(1)} - \mathbb{E}\xi_{r+1}^{(1)} = &\sum_{i=1}^{L_{1,1}} f(\xi_{r,i}^{(1)}) - \mathbb{E}f(\xi_{r,i}^{(1)}) + \sum_{i=1}^{L_{1,2}} f(\xi_{r,i}^{(2)}) - \mathbb{E}f(\xi_{r,i}^{(2)}) \\
&+ (L_{1,1} - \mathbb{E}L_{1,1})\mathbb{E}f(\xi_r^{(1)}) + (L_{1,2} - \mathbb{E}L_{1,2})\mathbb{E}f(\xi_r^{(2)}) \qquad (5.8.5)
\end{aligned}$$

We write $X_i = f(\xi_{r,i}^{(1)}) - \mathbb{E}f(\xi_{r,i}^{(1)})$ and $Y_i = f(\xi_{r,i}^{(2)}) - \mathbb{E}f(\xi_{r,i}^{(2)})$. Let us decompose the first sum:

$$\sum_{i=1}^{L_{1,1}} X_i = \sum_{i=1}^{\mathbb{E}L_{1,1}} X_i + \underbrace{\sum_{i=1}^{L_{1,1}} X_i - \sum_{i=1}^{\mathbb{E}L_{1,1}} X_i}_{S}.$$

We first show that $S \xrightarrow[d\to\infty]{W_2} \delta_0$. Wald identities give us

$$\begin{aligned}\mathrm{Var}(S) &= \mathrm{Var}\left(\sum_{i=1}^{|L_{1,1}-\mathbb{E}L_{1,1}|} X_i\right) = \mathbb{E}\big[|L_{1,1} - \mathbb{E}L_{1,1}|\big]\mathbb{E}[X_1^2] \\ &= \underbrace{\frac{1}{d}\mathbb{E}\big[|L_{1,1} - \mathbb{E}L_{1,1}|\big]}_{\xrightarrow[d\to\infty]{}0} \underbrace{b\mathrm{Var}(f(\xi_r^{(0)}))}_{O(1)} \xrightarrow[d\to\infty]{} 0.\end{aligned}$$

And therefore $S \xrightarrow[d\to\infty]{W_2} \delta_0$ because $\mathbb{E}S = 0$. Next, we apply the Central Limit Theorem to the sum

$$\sum_{i=1}^{\mathbb{E}L_{1,1}} X_i = \frac{1}{\sqrt{(1-p)a}} \sum_{i=1}^{(1-p)a} \underbrace{\sqrt{(1-p)a}X_i}_{\text{i.i.d., bounded random variables}}.$$

We obtain that the sum converges with respect to the Wasserstein metric to a normal distribution.

The two first sums in (5.8.5) are independent and converge to Gaussian distributions in the Wasserstein sense. It remains to show that the last two terms are converging toward Gaussian distributions. This is indeed the case because

$$(L_{1,1} - \mathbb{E}L_{1,1})\mathbb{E}f(\xi_r^{(0)}) = \frac{1}{\sqrt{d}}(L_{1,1} - \mathbb{E}L_{1,1})\ \sqrt{d}\mathbb{E}f(\xi_r^{(0)}).$$

$\frac{1}{\sqrt{d}}(L_{1,1} - \mathbb{E}L_{1,1})$ converges toward a Gaussian distribution and $\sqrt{d}\mathbb{E}f(\xi_r^{(0)})$ converges (to a constant) as $d \to \infty$. The last term is treated the same way.

$\xi_{r+1}^{(1)}$ is therefore converging toward a Gaussian distribution in the Wasserstein sense. The mean and the variance of this Gaussian distribution are necessarily equal the limits of the means and the variance of $\xi_{r+1}^{(1)}$ that we computed. □

5.9 Connection with rank-one matrix estimation

We adopt now a completely different point of view from the previous sections, in order to relate the community detection problem to the rank-one matrix estimation problem of Chapter 3. Rougthly, we will show now that the graph $\boldsymbol{G}$ contains, in some sense, "as much information" about the classes $\boldsymbol{X}$ than the matrix $\boldsymbol{Y}$ given by

$$Y_{i,j} = \sqrt{\frac{\lambda}{n}}\widetilde{X}_i\widetilde{X}_j + Z_{i,j} \qquad \text{for } 1 \le i < j \le n$$

where $Z_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ independently of everything else and $\widetilde{X}_i = \phi_p(X_i)$ where $\phi_p(1) = \sqrt{\frac{1-p}{p}}$ and $\phi_p(2) = -\sqrt{\frac{p}{1-p}}$. In the case of two symmetric communities (i.e. $p = 1/2$) this was done in [59] who showed that the mutual information of the two models where approximately equal: $\frac{1}{n}I(\boldsymbol{X};\boldsymbol{G}) \simeq \frac{1}{n}I(\widetilde{\boldsymbol{X}};\boldsymbol{Y})$. Following the proof strategy of [59] we generalize this to the asymmetric case:

Theorem 5.9.1

There exists a constant $C > 0$ such that, for d large enough

$$\limsup_{n\to\infty} \frac{1}{n}\Big|I(\boldsymbol{X};\boldsymbol{G}) - I(\widetilde{\boldsymbol{X}},\boldsymbol{Y})\Big| \leq Cd^{-1/2}.$$

Theorem 3.2.1 that we will see in Chapter 3 allows us to compute the limit of $\frac{1}{n}I(\widetilde{\boldsymbol{X}},\boldsymbol{Y})$. Define

$$\mathcal{F}(\lambda,q) = -\frac{\lambda q^2}{4} + \mathbb{E}\log\Big[p\exp(\sqrt{\frac{1-p}{p}}(\sqrt{\lambda q}Z_0 + \lambda q\widetilde{X}_0) - \frac{\lambda(1-p)}{2p}q) \\ + (1-p)\exp(-\sqrt{\frac{p}{1-p}}(\sqrt{\lambda q}Z_0 + \lambda q\widetilde{X}_0) - \frac{\lambda p}{2(1-p)}q)\Big] \quad (5.9.1)$$

where the expectation is taken over $Z_0 \sim \mathcal{N}(0,1)$ and $\widetilde{X}_0 \sim p\delta_{\phi_p(1)} + (1-p)\delta_{\phi_p(2)}$ independently from Z_0. Define also

$$\mathcal{I} : \lambda \mapsto \frac{\lambda}{4} - \sup_{q\geq 0} \mathcal{F}_g(\lambda,q). \quad (5.9.2)$$

Corollary 3.2.1 gives then

Corollary 5.9.1

There exists a constant $C > 0$ such that, for d large enough

$$\limsup_{n\to\infty} \Big|\frac{1}{n}I(\boldsymbol{X};\boldsymbol{G}) - \mathcal{I}(\lambda)\Big| \leq Cd^{-1/2}.$$

From Corollary 5.9.1 one can deduce the precise threshold for reconstruction in the Stochastic Block Model (when $d \to \infty$).

Definition 5.9.1 (*Estimator*)

An estimator of the labels $\boldsymbol{X}$ is a function $\widehat{\boldsymbol{x}} : \boldsymbol{G} \mapsto \{1,2\}^n$ that could depend on auxiliary randomness (random variables independent of $\boldsymbol{X}$).

For a labeling $\boldsymbol{x} \in \{1,2\}^n$ and $i \in \{1,2\}$ we define $S_i(\boldsymbol{x}) = \{k \in \{1,\dots,n\} | x_k = i\}$, i.e. the indices of the nodes that have the label i according to $\boldsymbol{x}$. We now recall a popular performance measure for estimators.

Definition 5.9.2 (*Community Overlap*)

For $\boldsymbol{x},\boldsymbol{y} \in \{1,2\}^n$ we define the community overlap of the configuration $\boldsymbol{x}$ and $\boldsymbol{y}$ as

$$\text{Overlap}(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{n}\max_{\sigma}\sum_{i=1,2}\Big(\#S_i(\boldsymbol{x}) \cap S_{\sigma(i)}(\boldsymbol{y}) - \frac{1}{n}\#S_i(\boldsymbol{x})\#S_{\sigma(i)}(\boldsymbol{y})\Big)$$

where the maximum is taken over the permutations of $\{1,2\}$.

Two configurations have thus a positive community overlap if they are correlated, up to a permutation of the classes. We will then say that the community detection problem is solvable, if there exists an estimator (i.e. an algorithm) that achieves a positive overlap with positive probability.

Definition 5.9.3 (*Solvability*)

We say that the community detection problem is solvable (in the limit of large degrees) if there exists an estimator $\widehat{\boldsymbol{x}}(\boldsymbol{G})$ such that

$$\liminf_{d\to\infty}\liminf_{n\to\infty}\mathbb{E}(\mathrm{Overlap}(\widehat{\boldsymbol{x}}(\boldsymbol{G}),\boldsymbol{X}))>0.$$

The next Theorem states that

$$\lambda_c(p)\stackrel{\mathrm{def}}{=}\inf\Big\{\lambda>0\,\Big|\,\mathcal{I}(\lambda)<\frac{\lambda}{4}\Big\}.$$

will be the threshold for solvability in the stochastic block model.

Theorem 5.9.2

- *If $\lambda>\lambda_c(p)$, then the community detection problem is solvable.*
- *If $\lambda<\lambda_c(p)$, then the community detection problem is not solvable.*

5.9.1 The limit of the mutual information: proof of Theorem 5.9.1

We are going to compute $I(\boldsymbol{X};\boldsymbol{G})$ and $I(\widetilde{\boldsymbol{X}},\boldsymbol{Y})$. For $\boldsymbol{x}\in\{1,2\}^n$ we denote $\widetilde{\boldsymbol{x}}=(\phi_p(x_1),\dots,\phi_p(x_n))\in S_p^n$, where $\phi_p(1)=\sqrt{\frac{1-p}{p}}$, $\phi_p(2)=-\sqrt{\frac{p}{1-p}}$ and $S_p=\Big\{-\sqrt{\frac{p}{1-p}},\sqrt{\frac{1-p}{p}}\Big\}$. Recall the notations from (5.6.1):

$$a=1+\frac{1-p}{p}\epsilon,\quad b=1-\epsilon\quad\text{and}\quad c=1+\frac{p}{1-p}\epsilon. \tag{5.9.3}$$

For $x\in\{1,2\}$ we define $P_0(x)=P_0(\widetilde{x})=p$, if $x=1$ and $P_0(x)=P_0(\widetilde{x})=1-p$, if $x=2$. For $\boldsymbol{x}\in\{1,2\}^n$ we will write, with a slight abuse of notation

$$P_0(\boldsymbol{x})=P_0(\widetilde{\boldsymbol{x}})=\prod_{i=1}^n P_0(x_i).$$

By definition of the mutual information, a simple computation gives:

Lemma 5.9.1

$$I(\widetilde{\boldsymbol{X}},\boldsymbol{Y})=-\mathbb{E}\left[\log\sum_{\widetilde{\boldsymbol{x}}\in S_p^n}P_0(\widetilde{\boldsymbol{x}})\exp\left(\sum_{i<j}\widetilde{x}_i\widetilde{x}_j\sqrt{\frac{\lambda}{n}}Z_{i,j}-\frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j-\widetilde{X}_i\widetilde{X}_j)^2\right)\right].$$

Define $V_{i,j}=\epsilon(G_{i,j}-\mathbb{E}(G_{i,j}|X_i,X_j))=\epsilon(G_{i,j}-\frac{d}{n}(1+\widetilde{X}_i\widetilde{X}_j))$.

Lemma 5.9.2

For d large enough,

$$\begin{aligned}&I(\boldsymbol{X};\boldsymbol{G})+O(d^2+n\epsilon)\\&=-\mathbb{E}\log\sum_{\boldsymbol{x}\in\{1,2\}^n}P_0(\boldsymbol{x})\exp\Big(\sum_{i<j}(\widetilde{x}_i\widetilde{x}_j-\widetilde{X}_i\widetilde{X}_j)V_{i,j}-\frac{\epsilon}{2}(\widetilde{x}_i\widetilde{x}_j)^2V_{i,j}-\frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j-\widetilde{X}_i\widetilde{X}_j)^2\Big).\end{aligned}$$

Proof. By definition, $I(\boldsymbol{X};\boldsymbol{G}) = \mathbb{E}\log\frac{\mathbb{P}(\boldsymbol{X},\boldsymbol{G})}{\mathbb{P}(\boldsymbol{X})\mathbb{P}(\boldsymbol{G})} = -\mathbb{E}\log\frac{\mathbb{P}(\boldsymbol{G})}{\mathbb{P}(\boldsymbol{G}|\boldsymbol{X})}$. Thus

$$I(\boldsymbol{X};\boldsymbol{G}) = -\mathbb{E}\log\frac{\sum_{\boldsymbol{x}\in\{1,2\}^n} P_0(\boldsymbol{x})\mathbb{P}(\boldsymbol{G}|\boldsymbol{x})}{\mathbb{P}(\boldsymbol{G}|\boldsymbol{X})}.$$

Recall that

$$\mathbb{P}(\boldsymbol{G}|\boldsymbol{x}) = \prod_{i<j} M_{x_i,x_j}^{G_{i,j}}(1-M_{x_i,x_j})^{1-G_{i,j}} = \exp\Big(\sum_{i<j} G_{i,j}\log M_{x_i,x_j} + (1-G_{i,j})\log(1-M_{x_i,x_j})\Big).$$

This leads to

$$I(\boldsymbol{X};\boldsymbol{G}) = -\mathbb{E}\left[\log\sum_{\boldsymbol{x}\in\{1,2\}^n} P_0(\boldsymbol{x})\exp\left(\sum_{i<j} G_{i,j}\log\left(\frac{M_{x_i,x_j}}{M_{X_i,X_j}}\right) + (1-G_{i,j})\log\left(\frac{1-M_{x_i,x_j}}{1-M_{X_i,X_j}}\right)\right)\right]. \tag{5.9.4}$$

Notice that $M_{x_i,x_j} = \frac{d}{n}(1+\widetilde{x}_i\widetilde{x}_j\epsilon)$. Therefore $\log(\frac{M_{x_i,x_j}}{M_{X_i,X_j}}) = \log(\frac{1+\widetilde{x}_i\widetilde{x}_j\epsilon}{1+\widetilde{X}_i\widetilde{X}_j\epsilon})$. By the Taylor-Lagrange inequality, there exist a constant $C>0$ such that, for ϵ small enough (i.e. for d large enough):

$$\Big|\log\Big(\frac{1+\widetilde{x}_i\widetilde{x}_j\epsilon}{1+\widetilde{X}_i\widetilde{X}_j\epsilon}\Big) - \epsilon(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j) + \frac{1}{2}\epsilon^2((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)\Big| \le C\epsilon^3\,, \tag{5.9.5}$$

$$\Big|\log\Big(\frac{1-M_{x_i,x_j}}{1-M_{X_i,X_j}}\Big) + \frac{d}{n}\epsilon(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\Big| \le C\frac{d^2}{n^2}\,. \tag{5.9.6}$$

By summation and triangle inequality:

$$\begin{aligned}
A_{i,j} &\stackrel{\text{def}}{=} G_{i,j}\log\left(\frac{M_{x_i,x_j}}{M_{X_i,X_j}}\right) + (1-G_{i,j})\log\left(\frac{1-M_{x_i,x_j}}{1-M_{X_i,X_j}}\right)\\
&= \epsilon(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)G_{i,j} - \frac{1}{2}\epsilon^2((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)G_{i,j} - (1-G_{i,j})(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\frac{d}{n}\epsilon + O\Big(G_{i,j}\epsilon^3 + \frac{d^2}{n^2}\Big)\,,
\end{aligned}$$

because of equations (5.9.5) and (5.9.6). Since $\epsilon G_{i,j} - V_{i,j} = \frac{\epsilon d}{n}(1+\epsilon\widetilde{X}_i\widetilde{X}_j) = \frac{\epsilon d}{n} + \frac{\lambda}{n}\widetilde{X}_i\widetilde{X}_j$, we get

$$\begin{aligned}
A_{i,j} =&(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)V_{i,j} - \frac{1}{2}((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)V_{i,j} + \frac{\lambda}{n}\widetilde{X}_i\widetilde{X}_j(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\\
&- \frac{\epsilon}{2}((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)\frac{\epsilon d}{n}(1+\epsilon\widetilde{X}_i\widetilde{X}_j) - G_{i,j}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\frac{d}{n}\epsilon + O\Big(G_{i,j}\epsilon^3 + \frac{d^2}{n^2}\Big)\\
=&(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)V_{i,j} - \frac{1}{2}((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)V_{i,j} + \frac{\lambda}{n}\widetilde{X}_i\widetilde{X}_j(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\\
&- \frac{\lambda}{2n}((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2) - G_{i,j}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\frac{d}{n}\epsilon + O\Big(G_{i,j}\epsilon^3 + \frac{d^2}{n^2} + \frac{\epsilon}{n}\Big)\\
=&(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)V_{i,j} - \frac{1}{2}((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)V_{i,j} - \frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)^2\\
&- G_{i,j}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\frac{d}{n}\epsilon + O\Big(G_{i,j}\epsilon^3 + \frac{d^2}{n^2} + \frac{\epsilon}{n}\Big).
\end{aligned}$$

Notice that $G_{i,j}\in\{0,1\}$, we have therefore, for some constant $C>0$,

$$\Big|\sum_{i<j} G_{i,j}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)\frac{d}{n}\Big| \le C\frac{d}{n}\sum_{i<j} G_{i,j}.$$

Therefore

$$\sum_{i<j} A_{i,j} = (\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)V_{i,j} - \frac{1}{2}((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)V_{i,j} - \frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)^2 + \Delta_n$$

where $|\Delta_n| \leq C((\frac{d}{n}+\epsilon^3)\sum_{i<j} G_{i,j} + d^2 + n\epsilon)$, for some constant $C > 0$. Since $I(\boldsymbol{X};\boldsymbol{G}) = -\mathbb{E}\log\sum_{\boldsymbol{x}\in\{1,2\}^n} P_0(\boldsymbol{x})e^{A_{i,j}}$, we get

$$\mathbb{E}\log\sum_{\boldsymbol{x}\in\{1,2\}^n} P_0(\boldsymbol{x})\exp\Big(\sum_{i<j}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)V_{i,j} - \frac{\epsilon}{2}((\widetilde{x}_i\widetilde{x}_j)^2 - (\widetilde{X}_i\widetilde{X}_j)^2)V_{i,j} - \frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)^2\Big)$$
$$= -I(\boldsymbol{X},\boldsymbol{G}) + O\Big(\Big(\frac{d}{n}+\epsilon^3\Big)\mathbb{E}\Big[\sum_{i<j} G_{i,j}\Big] + d^2 + n\epsilon\Big) = -I(\boldsymbol{X},\boldsymbol{G}) + O(d^2+n\epsilon).$$

□

Lindeberg argument

We recall the following Lindeberg generalization theorem (Theorem 2 from [126]) which is a variant of the generalized "Lindeberg principle" from [45].

Theorem 5.9.3 (*Lindeberg generalization theorem*)

Let $(U_i)_{1\leq i\leq n}$ and $(V_i)_{1\leq i\leq n}$ be two collection of random variables with independent components and $f : \mathbb{R}^n \to \mathbb{R}$ a $\mathcal{C}^3$ function. Denote $a_i = |\mathbb{E}U_i - \mathbb{E}V_i|$ and $b_i = |\mathbb{E}U_i^2 - \mathbb{E}V_i^2|$. Then

$$|\mathbb{E}f(U) - \mathbb{E}f(V)| \leq \sum_{i=1}^n \Big(a_i\mathbb{E}|\partial_i f(U_{1:i-1},0,V_{i+1:n})| + \frac{b_i}{2}\mathbb{E}|\partial_i^2 f(U_{1:i-1},0,V_{i+1:n})|$$
$$+ \frac{1}{2}\mathbb{E}\int_0^{U_i}|\partial_i^3 f(U_{1:i-1},0,V_{i+1:n})|(U_i-s)^2 ds$$
$$+ \frac{1}{2}\mathbb{E}\int_0^{V_i}|\partial_i^3 f(U_{1:i-1},0,V_{i+1:n})|(V_i-s)^2 ds\Big).$$

Define

$$J(\boldsymbol{X},\boldsymbol{Z}) = -\mathbb{E}\log\sum_{\widetilde{\boldsymbol{x}}\in S_p^n} P_0(\widetilde{\boldsymbol{x}})\exp\Big(\sum_{i<j}\sqrt{\frac{\lambda}{n}}Z_{i,j}\Big((\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j) - \frac{\epsilon}{2}(\widetilde{x}_i\widetilde{x}_j)^2\Big) - \frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)^2\Big).$$

We show, using Theorem 5.9.3, that $J(\boldsymbol{X},\boldsymbol{Z})$ is close to $I(\boldsymbol{X};\boldsymbol{G})$.

Lemma 5.9.3

For d large enough we have

$$\frac{1}{n}|I(\boldsymbol{X};\boldsymbol{G}) - J(\boldsymbol{X},\boldsymbol{Z})| = O(\epsilon + d^2/n).$$

Proof. We apply here Theorem 5.9.3 conditionally to $\boldsymbol{X}$ to the function

$$\Phi(u) = -\log\sum_{\widetilde{\boldsymbol{x}}\in S_p^n} P_0(\widetilde{\boldsymbol{x}})\exp\Big(\sum_{i<j}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)u_{i,j} - \frac{\epsilon}{2}(\widetilde{x}_i\widetilde{x}_j)^2 u_{i,j} - \frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)^2\Big).$$

Φ is $\mathcal{C}^3$ with bounded derivatives. Notice that $I(\boldsymbol{X};\boldsymbol{G}) = \mathbb{E}\Phi(\boldsymbol{V})$ and $J(\boldsymbol{X},\boldsymbol{Z}) = \mathbb{E}\Phi(\sqrt{\frac{\lambda}{n}}\boldsymbol{Z})$. Let us compute $V_{i,j}$ moments, conditionally to $\boldsymbol{X}$.

$$\begin{aligned}\mathbb{E}(V_{i,j}|\boldsymbol{X}) &= 0\\ \mathbb{E}(V_{i,j}^2|\boldsymbol{X}) &= \epsilon^2\mathrm{Var}(G_{i,j}|\boldsymbol{X}) = \epsilon^2\frac{d}{n}(1+\widetilde{X}_i\widetilde{X}_j\epsilon)(1+O(\frac{d}{n}))\\ &= \frac{\lambda}{n}+O(\frac{\epsilon}{n}) = \frac{\lambda}{n}\mathbb{E}(Z_{i,j}^2)+O(\frac{\epsilon}{n})\end{aligned}$$

Analogously, $\mathbb{E}(V_{i,j}^3|\boldsymbol{X}) = O(\frac{\epsilon}{n})$. Using the Lindeberg generalization theorem we obtain

$$\left|\mathbb{E}[\Phi(\sqrt{\lambda/n}\boldsymbol{Z})] - \mathbb{E}[\Phi(\boldsymbol{V})]\right| \le \sum_{i<j} O(\epsilon/n) = O(n\epsilon).$$

□

Gaussian interpolation

It remains to show

Lemma 5.9.4

$$I(\boldsymbol{X};\boldsymbol{Y}) = J(\boldsymbol{X},\boldsymbol{Z}) + O(n\epsilon).$$

Proof. We define:

$$\begin{aligned}H(\boldsymbol{x},\boldsymbol{X},\boldsymbol{Z},\epsilon) &= \sum_{i<j}\widetilde{x}_i\widetilde{x}_j\sqrt{\frac{\lambda}{n}}Z_{i,j} - \frac{\lambda}{2n}(\widetilde{x}_i\widetilde{x}_j - \widetilde{X}_i\widetilde{X}_j)^2 - \frac{1}{2}\epsilon(\widetilde{x}_i\widetilde{x}_j)^2\sqrt{\frac{\lambda}{n}}Z_{i,j},\\ F(\epsilon) &= \mathbb{E}\log\sum_{\boldsymbol{x}\in\{1,2\}^n} P_0(\boldsymbol{x})\exp(H(\boldsymbol{x},\boldsymbol{X},\boldsymbol{Z},\epsilon)).\end{aligned}$$

Notice that $F(0) = I(\boldsymbol{X};\boldsymbol{Y})$ and $F(\epsilon) = J(\boldsymbol{X},\boldsymbol{Z})$. We are going to control the derivative of F. We note $\langle\cdot\rangle$ the expectation with respect to the Gibbs measure: $\langle g(\widetilde{\boldsymbol{x}})\rangle \overset{\text{def}}{=} \frac{\sum_{\widetilde{\boldsymbol{x}}} P_0(\widetilde{\boldsymbol{x}})g(\widetilde{\boldsymbol{x}})\exp(H(\boldsymbol{x},\boldsymbol{X},\boldsymbol{Z},\epsilon))}{\sum_{\widetilde{\boldsymbol{x}}} P_0(\widetilde{\boldsymbol{x}})\exp(H(\boldsymbol{x},\boldsymbol{X},\boldsymbol{Z},\epsilon))}$.
The derivative of F reads

$$F'(\epsilon) = -\frac{1}{2}\sqrt{\frac{\lambda}{n}}\sum_{i<j}\mathbb{E}[Z_{i,j}\langle(\widetilde{x}_i\widetilde{x}_j)^2\rangle].$$

Here $\langle(\widetilde{x}_i\widetilde{x}_j)^2\rangle$ is a continuously differentiable function of $Z_{i,j}$ and $\partial_{Z_{i,j}}\langle(\widetilde{x}_i\widetilde{x}_j)^2\rangle = O(n^{-1/2})$. Using Gaussian integration by parts: $F'(\epsilon) = -\frac{\sqrt{\lambda}}{2\sqrt{n}}\sum_{i<j}\mathbb{E}[\partial_{Z_{i,j}}\langle(\widetilde{x}_i\widetilde{x}_j)^2\rangle] = O(1)$. We conclude $|F(0)-F(\epsilon)| = O(n\epsilon)$. □

5.9.2 From mutual information to solvability: proof of Theorem 5.9.2

We are first going to introduce an estimation metric that will allow us to make the link between the minimum mean squared error for matrix estimation, and the overlap for community detection. Define

$$\mathrm{MMSE}_n^G(\lambda) = \min_{\hat{\theta}}\frac{2}{n(n-1)}\sum_{i<j}\mathbb{E}\Big(\widetilde{X}_i\widetilde{X}_j - \hat{\theta}_{i,j}(\boldsymbol{G})\Big)^2 = \frac{2}{n(n-1)}\sum_{i<j}\mathbb{E}\Big(\widetilde{X}_i\widetilde{X}_j - \mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G})\Big)^2 \tag{5.9.7}$$

where the minimum is taken over all function $\hat{\theta}$ of $\boldsymbol{G}$. By considering the trivial estimator $\hat{\theta} = 0$, we see that $\mathrm{MMSE}_n^G(\lambda) \in [0,1]$. This estimation metric correspond (up to a vanishing error term) to the derivative of the mutual information between the graph $\boldsymbol{G}$ and the labels $\boldsymbol{X}$.

Proposition 5.9.1

Let $\lambda_0 > 0$. There exists a constant $C > 0$ such that, for all $\lambda \in (0, \lambda_0]$, $d \geq 1$ and $n \geq 1$

$$\Big|\frac{1}{n}\frac{\partial I(\boldsymbol{X};\boldsymbol{G})}{\partial \lambda} - \frac{1}{4}\mathrm{MMSE}_n^G(\lambda)\Big| \leq C\Big(d^{-1/2} + \frac{d}{n} + \frac{d^{1/2}\lambda^{-1/2}}{n} + d^{3/2}n^{-2}\lambda^{-1/2}\Big) \tag{5.9.8}$$

Proof. We are going to differentiate $H(\boldsymbol{X}|\boldsymbol{G})$ with respect to λ. To do so we will use a differentiation formula from [59] (Lemma 7.1), which was first proved in [141]. Let us recall the setting (taken from [59]) of this Lemma.

For n an integer, denote by $\mathcal{P}_n$ the set of unordered pairs in $[n]$ (in particular $\#\mathcal{P}_n = \binom{n}{2}$)). We will use $e, e_1, e_2, \dots$ to denote elements of $\mathcal{P}_n$. For for each $e = (i,j)$ we are given a one-parameter family of discrete noisy channels indexed by $\theta \in J$ (with $J = (a_1, a_2)$ a non-empty interval), with finite input alphabet $\mathcal{X}_0$ and finite output alphabet $\mathcal{Y}$. Concretely, for any e, we have a transition probability

$$\{p_{e,\theta}(y|x)\}_{x\in\mathcal{X}_0, y\in\mathcal{Y}}\,, \tag{5.9.9}$$

which is differentiable in θ. We shall omit the subscript θ since it will be clear from the context.

We then consider $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$ a random vector in $\mathcal{X}^n$, and $\boldsymbol{Y} = (Y_{ij})_{(i,j)\in\mathcal{P}_n}$ a set of observations in $\mathcal{Y}^{\mathcal{P}_n}$ that are conditionally independent given $\boldsymbol{X}$. Further Y_{ij} is the noisy observation of $X_iX_j \in \mathcal{X}_0$ through the channel $p_{ij}(\,\cdot\,|\,\cdot\,)$. In formulae, the joint probability density function of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is

$$p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y}) = p_{\boldsymbol{X}}(\boldsymbol{x}) \prod_{(i,j)\in\mathcal{P}_n} p_{ij}(y_{ij}|x_ix_j)\,. \tag{5.9.10}$$

This obviously include the two-groups stochastic block model as a special case. In that case $\boldsymbol{Y} = \boldsymbol{G}$ is just the adjacency matrix of the graph. In the following we write $\boldsymbol{Y}_{-e} = (Y_{e'})_{e'\in\mathcal{P}_n\setminus e}$ for the set of observations excluded e, and $X_e = X_iX_j$ for $e = (i,j)$.

Lemma 5.9.5

With the above notation, we have:

$$\frac{\partial H(\boldsymbol{X}|\boldsymbol{Y})}{\partial\theta} = \sum_{e\in\mathcal{P}_n}\sum_{x_e,y_e}\frac{\partial p_e(y_e|x_e)}{\partial\theta}\mathbb{E}\Big\{p_{X_e|\boldsymbol{Y}_{-e}}(x_e|\boldsymbol{Y}_{-e})\log\Big[\sum_{x'_e}\frac{p_e(y_e|x'_e)}{p_e(y_e|x_e)}p_{X_e|\boldsymbol{Y}_{-e}}(x'_e|\boldsymbol{Y}_{-e})\Big]\Big\} \tag{5.9.11}$$

We apply Lemma 5.9.5 to the stochastic block model. Let $\lambda_0 > 0$ and $\lambda \in (0, \lambda_0]$. Instead of having $G_{i,j} \in \{0,1\}$, it will be more convenient to consider $G_{i,j} \in \{-1,1\}$: $G_{i,j} = 1$ if $i \sim j$, $G_{i,j} = -1$ else. Notice that neither the mutual information nor MMSE_n^G are affected by this change. $G_{i,j}$ is, conditionally to $\widetilde{X}_i\widetilde{X}_j$, independent of any other random variable, and distributed as follows

$$\mathbb{P}(G_{i,j} = 1|\widetilde{X}_i\widetilde{X}_j) = \frac{d}{n}(1 + \sqrt{\frac{\lambda}{d}}\widetilde{X}_i\widetilde{X}_j)$$

The transition probability from equation (5.9.9) is then $p_\lambda(g_{i,j}|\widetilde{x}_i\widetilde{x}_j) = \frac{1-g_{i,j}}{2} + g_{i,j}\frac{d}{n}(1+\widetilde{x}_i\widetilde{x}_j\sqrt{\frac{\lambda}{d}})$. Thus

$$\frac{\partial}{\partial\lambda}p_\lambda(g_{i,j}|\widetilde{x}_i\widetilde{x}_j) = \frac{1}{2n}g_{i,j}\widetilde{x}_i\widetilde{x}_j\sqrt{\frac{d}{\lambda}}.$$

Lemma 5.9.5 gives

$$\begin{aligned}
&\frac{\partial H(\boldsymbol{X}|\boldsymbol{G})}{\partial\lambda} \\
&= \frac{1}{2n}\sqrt{\frac{d}{\lambda}}\sum_{i<j}\underbrace{\sum_{\substack{\widetilde{x}_i,\widetilde{x}_j\\ g_{i,j}}} g_{i,j}\widetilde{x}_i\widetilde{x}_j\mathbb{E}\Big[p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}_i,\widetilde{x}_j|\boldsymbol{G}_{-(i,j)})\log\sum_{\widetilde{x}'_i,\widetilde{x}'_j} p_\lambda(g_{i,j}|\widetilde{x}'_i\widetilde{x}'_j)p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}'_i,\widetilde{x}'_j|\boldsymbol{G}_{-(i,j)})\Big]}_{A_{i,j}} \\
&- \frac{1}{2n}\sqrt{\frac{d}{\lambda}}\sum_{i<j}\underbrace{\sum_{\substack{\widetilde{x}_i,\widetilde{x}_j\\ g_{i,j}}} g_{i,j}\widetilde{x}_i\widetilde{x}_j\mathbb{E}\Big[p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}_i,\widetilde{x}_j|\boldsymbol{G}_{-(i,j)})\log\big(p_\lambda(g_{i,j}|\widetilde{x}_i\widetilde{x}_j)\big)\Big]}_{B_{i,j}} \qquad (5.9.12)
\end{aligned}$$

Compute

$$\begin{aligned}
B_{i,j} &= \sum_{\substack{\widetilde{x}_i,\widetilde{x}_j\\ g_{i,j}}} g_{i,j}\widetilde{x}_i\widetilde{x}_j p(\widetilde{x}_i,\widetilde{x}_j)\log\big(p_\lambda(g_{i,j}|\widetilde{x}_i\widetilde{x}_j)\big) \\
&= p^2\frac{1-p}{p}\big(\log(\frac{ad}{n}) - \log(1-\frac{ad}{n})\big) - 2p(1-p)\big(\log(\frac{bd}{n}) - \log(1-\frac{bd}{n})\big) \\
&\quad + (1-p)^2\frac{p}{1-p}\big(\log(\frac{cd}{n}) - \log(1-\frac{cd}{n})\big) \\
&= p(1-p)\log(\frac{ac}{b^2}) + p(1-p)d(2\frac{b}{n} - \frac{a}{n} - \frac{c}{n}) + O(\frac{d^2}{n^2}) \\
&= p(1-p)\epsilon(\frac{1-p}{p} + \frac{p}{1-p} + 2) + O(\epsilon^2) + O(\frac{d^2}{n^2}) \\
&= \epsilon + O(\epsilon^2) + O(\frac{d^2}{n^2})
\end{aligned}$$

and

$$\begin{aligned}
A_{i,j} &= \sum_{g_{i,j}} g_{i,j}\mathbb{E}\Big[\mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G}_{-(i,j)})\log\big(\sum_{\widetilde{x}'_i,\widetilde{x}'_j} p_\lambda(g_{i,j}|\widetilde{x}'_i\widetilde{x}'_j)p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}'_i,\widetilde{x}'_j|\boldsymbol{G}_{-(i,j)})\big)\Big] \\
&= \mathbb{E}\Big[\mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G}_{-(i,j)})\log\frac{\sum_{\widetilde{x}'_i,\widetilde{x}'_j} p_\lambda(1|\widetilde{x}'_i\widetilde{x}'_j)p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}'_i,\widetilde{x}'_j|\boldsymbol{G}_{-(i,j)})}{\sum_{\widetilde{x}'_i,\widetilde{x}'_j} p_\lambda(-1|\widetilde{x}'_i\widetilde{x}'_j)p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}'_i,\widetilde{x}'_j|\boldsymbol{G}_{-(i,j)})}\Big].
\end{aligned}$$

Define $\hat{a}_{i,j} = \mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G}_{-(i,j)})$.

$$\begin{aligned}
A_{i,j} &= \mathbb{E}\Big[\hat{a}_{i,j}\log\frac{\sum_{\widetilde{x}'_i,\widetilde{x}'_j}\frac{d}{n}(1+\widetilde{x}'_i\widetilde{x}'_j\sqrt{\frac{\lambda}{d}})p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}'_i,\widetilde{x}'_j|\boldsymbol{G}_{-(i,j)})}{\sum_{\widetilde{x}'_i,\widetilde{x}'_j}(1-\frac{d}{n}(1+\widetilde{x}'_i\widetilde{x}'_j\sqrt{\frac{\lambda}{d}}))p_{(\widetilde{X}_i,\widetilde{X}_j)|\boldsymbol{G}_{-(i,j)}}(\widetilde{x}'_i,\widetilde{x}'_j|\boldsymbol{G}_{-(i,j)})}\Big] \\
&= \mathbb{E}\Big[\hat{a}_{i,j}\log\frac{\frac{d}{n}(1+\hat{a}_{i,j}\sqrt{\frac{\lambda}{d}})}{1-\frac{d}{n}(1+\hat{a}_{i,j}\sqrt{\frac{\lambda}{d}})}\Big] = \mathbb{E}\Big[\hat{a}_{i,j}\big(\log\frac{d}{n} + \epsilon\hat{a}_{i,j} + O(\epsilon^2) + O(\frac{d}{n})\big)\Big].
\end{aligned}$$

$\mathbb{E}\hat{a}_{i,j} = \mathbb{E}\widetilde{X}_i\widetilde{X}_j = (\mathbb{E}\widetilde{X}_1)^2 = 0$ therefore $A_{i,j} = \epsilon\mathbb{E}\hat{a}_{i,j}^2 + O(\epsilon^2 + \frac{d}{n})$. By replacing $A_{i,j}$ and $B_{i,j}$ in (5.9.12), we have

$$\begin{aligned}\frac{\partial H(\boldsymbol{X}|\boldsymbol{G})}{\partial\lambda} &= \frac{1}{2n}\epsilon^{-1}\sum_{i<j}\Big(\epsilon\mathbb{E}\hat{a}_{i,j}^2 + O(\epsilon^2) + O(\frac{d}{n})\Big) - \frac{1}{2n}\epsilon^{-1}\sum_{i<j}\Big(\epsilon + O(\epsilon^2) + O(\frac{d^2}{n^2})\Big)\\ &= \frac{1}{2n}\sum_{i<j}\Big(\mathbb{E}\hat{a}_{i,j}^2 - 1\Big) + O(n\epsilon) + O(d^{1/2}\lambda^{-1/2} + d^{3/2}n^{-1}\lambda^{-1/2}). \end{aligned} \tag{5.9.13}$$

Define $a_{i,j} = \mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G})$. Using Bayes rule, we have

$$p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}) = \frac{p(G_{i,j}|\widetilde{x}_i\widetilde{x}_j)p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}_{-(i,j)})}{\sum_{\widetilde{x}_i',\widetilde{x}_j'} p(G_{i,j}|\widetilde{x}_i'\widetilde{x}_j')p(\widetilde{x}_i', \widetilde{x}_j'|\boldsymbol{G}_{-(i,j)})}.$$

If $G_{i,j} = 1$, then

$$p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}) = \frac{\frac{d}{n}(1+\widetilde{x}_i\widetilde{x}_j\epsilon)p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}_{-(i,j)})}{\sum_{\widetilde{x}_i',\widetilde{x}_j'}\frac{d}{n}(1+\widetilde{x}_i'\widetilde{x}_j'\epsilon)p(\widetilde{x}_i', \widetilde{x}_j'|\boldsymbol{G}_{-(i,j)})} = \frac{(1+\epsilon\widetilde{x}_i\widetilde{x}_j)p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}_{-(i,j)})}{1+\epsilon\hat{a}_{i,j}}.$$

Thus

$$a_{i,j} = \frac{\hat{a}_{i,j} + \epsilon\mathbb{E}((\widetilde{X}_i\widetilde{X}_j)^2|\boldsymbol{G}_{-(i,j)})}{1+\epsilon\hat{a}_{i,j}} = \hat{a}_{i,j} + O(\epsilon)$$

If $G_{i,j} = -1$,

$$p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}) = \frac{(1-\frac{d}{n}(1+\widetilde{x}_i\widetilde{x}_j\epsilon))p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}_{-(i,j)})}{1-\sum_{\widetilde{x}_i',\widetilde{x}_j'}\frac{d}{n}(1+\widetilde{x}_i'\widetilde{x}_j'\epsilon)p(\widetilde{x}_i', \widetilde{x}_j'|\boldsymbol{G}_{-(i,j)})} = (1-\frac{d}{n}(1+\widetilde{x}_i\widetilde{x}_j\epsilon))p(\widetilde{x}_i, \widetilde{x}_j|\boldsymbol{G}_{-(i,j)}) + O(\frac{d}{n}).$$

Therefore $a_{i,j} = \hat{a}_{i,j} + O(\frac{d}{n})$. Equation (5.9.13) becomes then

$$\begin{aligned}\frac{\partial H(\boldsymbol{X}|\boldsymbol{G})}{\partial\lambda} &= \frac{1}{2n}\sum_{i<j}\Big(\mathbb{E}a_{i,j}^2 - 1\Big) + O(n\epsilon + d) + O(d^{1/2}\lambda^{-1/2} + d^{3/2}n^{-1}\lambda^{-1/2})\\ &= -\frac{1}{2n}\sum_{i<j}\mathbb{E}\Big((\widetilde{X}_i\widetilde{X}_j - \mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G}))^2\Big) + O(n\epsilon + d + d^{1/2}\lambda^{-1/2} + d^{3/2}n^{-1}\lambda^{-1/2}).\end{aligned} \tag{5.9.14}$$

Decomposing $I(\boldsymbol{X};\boldsymbol{G}) = H(\boldsymbol{X}) - H(\boldsymbol{X}|\boldsymbol{G})$ we obtain the desired result. □

Consequently, if one consider a sufficiently large d (in order to apply Corollary 5.9.1) and one integrate equation (5.9.8) from 0 to $\lambda > 0$, and let n tend to infinity,

$$\limsup_{n\to\infty}\left|\int_0^\lambda \mathrm{MMSE}_n^G(\lambda')d\lambda' - \mathcal{I}(\lambda)\right| \le Kd^{-1/2}, \tag{5.9.15}$$

for some constant $K > 0$ depending on λ but not on d.

Proposition 5.9.2

For $\lambda < \lambda_c(p)$

$$\lim_{d\to\infty}\liminf_{n\to\infty}\mathrm{MMSE}_n^G(\lambda) = 1 \tag{5.9.16}$$

For $\lambda > \lambda_c(p)$

$$\limsup_{d\to\infty}\limsup_{n\to\infty}\mathrm{MMSE}_n^G(\lambda) < 1 \tag{5.9.17}$$

Proof. This is a consequence of equation (5.9.15) and the definition of $\lambda_c(p)$. We will show (5.9.16) first. $\mathcal{I}$ is continuous, so by definition of λ_c, $\mathcal{I}(\lambda_c) = \lambda_c/4$. Equation (5.9.15) gives then

$$\limsup_{n\to\infty} \int_0^{\lambda_c} |1 - \mathrm{MMSE}_n^G(\lambda)| d\lambda \leq 4Kd^{-1/2}$$

because $\mathrm{MMSE}_n^G(\lambda) \leq 1$. Equation (5.9.16) follows.

Equation (5.9.17) is proved analogously. If $\limsup_{d\to\infty} \limsup_{n\to\infty} \mathrm{MMSE}_n^G(\lambda_0) = 1$ for some $\lambda_0 > \lambda_c$, then $\limsup_{d\to\infty} \limsup_{n\to\infty} |\int_0^{\lambda_0} \frac{1}{4}\mathrm{MMSE}_n^G(\lambda)d\lambda - \frac{\lambda_0}{4}| = 0$ which implies (by equation (5.9.15)) that $\mathcal{I}(\lambda_0) = 0$ which contradicts the definition of λ_c. □

We would now like to rewrite the result of Proposition 5.9.2 in terms of Overlap instead of MMSE. The following two lemmas will be useful to make the link between the MMSE and the overlap.

Lemma 5.9.6

Let $n \in \mathbb{N}^*$. Let (A_1, A_2) and (B_1, B_2) be two partitions of $\{1, \dots, n\}$. Then

$$\#A_1 \cap B_1 - \frac{1}{n}\#A_1\#B_1 = \#A_2 \cap B_2 - \frac{1}{n}\#A_2\#B_2, \tag{5.9.18}$$

$$\#A_1 \cap B_1 - \frac{1}{n}\#A_1\#B_1 = -\Big(\#A_1 \cap B_2 - \frac{1}{n}\#A_1\#B_2\Big). \tag{5.9.19}$$

Proof. We prove (5.9.18) first. Remark that $\#A_2 \cap B_2 = \#B_2 - (\#A_1 - \#B_1 \cap A_1)$ and $\#A_2\#B_2 = n^2 - n(\#A_1 + \#B_1) + \#A_1\#B_1$. So that

$$\begin{aligned}\#A_1 \cap B_1 - \frac{1}{n}\#A_1\#B_1 &= \#A_2 \cap B_2 - \frac{1}{n}\#A_2\#B_2 + \#B_2 - \#A_1 - n + \#A_1 + \#B_1 \\ &= \#A_2 \cap B_2 - \frac{1}{n}\#A_2\#B_2.\end{aligned}$$

To prove (5.9.19), write $\#A_1 \cap B_1 = \#A_1 - \#A_1 \cap B_2$. Thus

$$\#A_1 \cap B_1 - \frac{1}{n}\#A_1\#B_1 = \#A_1 - \#A_1 \cap B_2 - \frac{1}{n}\#A_1(n - \#B_2) = -\Big(\#A_1 \cap B_2 - \frac{1}{n}\#A_1\#B_2\Big).$$

□

Lemma 5.9.7

Recall that for $\boldsymbol{x} \in \{1,2\}^n$ and $i = 1, 2$ we define $S_i(\boldsymbol{x}) = \{j \,|\, x_j = i\}$. Let $\boldsymbol{x} = (x_1, \dots, x_n) \in \{1,2\}^n$. Recall that $\widetilde{\boldsymbol{x}} = (\phi_p(x_1), \dots, \phi_p(x_n))$ where $\phi_p(1) = \sqrt{\frac{1-p}{p}}$ and $\phi_p(2) = -\sqrt{\frac{p}{1-p}}$. Then

$$\frac{1}{n}\Big|\sum_{i=1}^n \widetilde{X}_i\widetilde{x}_i\Big| = \frac{1}{2p(1-p)}\mathrm{Overlap}(\boldsymbol{x}, \boldsymbol{X}) + O\Big(\Big|\frac{S_1(\boldsymbol{X})}{n} - p\Big| + \Big|\frac{S_2(\boldsymbol{X})}{n} - (1-p)\Big|\Big).$$

Proof.

$$
\begin{aligned}
&\frac{1}{n}\Big|\sum_{i=1}^{n}\widetilde{X}_i\widetilde{x}_i\Big| \\
&= \frac{1}{n}\Big|\frac{1-p}{p}\#S_1(\boldsymbol{x})\cap S_1(\boldsymbol{X}) + \frac{p}{1-p}\#S_2(\boldsymbol{x})\cap S_2(\boldsymbol{X}) - (n - \#S_1(\boldsymbol{x})\cap S_1(\boldsymbol{X}) - \#S_2(\boldsymbol{x})\cap S_2(\boldsymbol{X}))\Big| \\
&= \frac{1}{n}\Big|\frac{1}{p}\#S_1(\boldsymbol{x})\cap S_1(\boldsymbol{X}) + \frac{1}{1-p}\#S_2(\boldsymbol{x})\cap S_2(\boldsymbol{X}) - n\Big| \\
&= \frac{1}{n}\Big|\frac{1}{p}(\#S_1(\boldsymbol{x})\cap S_1(\boldsymbol{X}) - p\#S_1(\boldsymbol{x})) + \frac{1}{1-p}(\#S_2(\boldsymbol{x})\cap S_2(\boldsymbol{X}) - (1-p)\#S_2(\boldsymbol{x}))\Big| \\
&= \frac{1}{n}\Big|\frac{1}{p}(\#S_1(\boldsymbol{x})\cap S_1(\boldsymbol{X}) - \frac{1}{n}\#S_1(\boldsymbol{X})\#S_1(\boldsymbol{x})) + \frac{1}{1-p}(\#S_2(\boldsymbol{x})\cap S_2(\boldsymbol{X}) - \frac{1}{n}\#S_2(\boldsymbol{X})\#S_2(\boldsymbol{x}))\Big| \\
&\quad + O\Big(\frac{S_1(\boldsymbol{x})}{pn}\Big|p - \frac{S_1(\boldsymbol{X})}{n}\Big| + \frac{S_2(\boldsymbol{x})}{(1-p)n}\Big|(1-p) - \frac{S_2(\boldsymbol{X})}{n}\Big|\Big).
\end{aligned}
$$

Using Lemma 5.9.6, one obtains then

$$
\frac{1}{n}\Big|\sum_{i=1}^{n}\widetilde{X}_i\widetilde{x}_i\Big| = \frac{1}{2p(1-p)}\mathrm{Overlap}(\boldsymbol{x},\boldsymbol{X}) + O\left(\Big|\frac{S_1(\boldsymbol{X})}{n} - p\Big| + \Big|\frac{S_2(\boldsymbol{X})}{n} - (1-p)\Big|\right).
$$

□

We are now going to prove Theorem 5.9.2. Remark that

$$
\begin{aligned}
\mathrm{MMSE}_n^G(\lambda) &= \frac{2}{n(n-1)}\sum_{i<j}\mathbb{E}\Big(\widetilde{X}_i\widetilde{X}_j - \mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G})\Big)^2 \\
&= \frac{2}{n(n-1)}\sum_{i<j}\mathbb{E}\Big(1 - (\mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G}))^2\Big)
\end{aligned}
$$

We will denote $\langle\cdot\rangle_G$ the expectation with respect to the posterior distribution $\mathbb{P}(\boldsymbol{X} = \cdot|\boldsymbol{G})$. We have then $\mathbb{E}(\mathbb{E}(\widetilde{X}_i\widetilde{X}_j|\boldsymbol{G})^2) = \mathbb{E}\langle\widetilde{X}_i\widetilde{X}_j\widetilde{x}_i\widetilde{x}_j\rangle_G$, where $\widetilde{x}$ is sampled, conditionally to $\boldsymbol{G}$, from $\mathbb{P}(\cdot|\boldsymbol{G})$. Thus

$$
\mathrm{MMSE}_n^G(\lambda) = 1 - \frac{1}{n^2}\sum_{i,j}\mathbb{E}\langle\widetilde{X}_i\widetilde{X}_j\widetilde{x}_i\widetilde{x}_j\rangle_G + o(1) = 1 - \mathbb{E}\Big\langle\Big(\frac{1}{n}\sum_{i=1}^{n}\widetilde{X}_i\widetilde{x}_i\Big)^2\Big\rangle_G + o(1). \quad (5.9.20)
$$

Suppose $\lambda > \lambda_c$. Then (5.9.20) and Proposition 5.9.2 imply

$$
\liminf_{d\to\infty}\liminf_{n\to\infty}\mathbb{E}\Big\langle\Big|\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\widetilde{X}_i\Big|\Big\rangle_G > 0.
$$

Using Lemma 5.9.7, this gives

$$
\liminf_{d\to\infty}\liminf_{n\to\infty}\mathbb{E}\Big\langle\mathrm{Overlap}(\boldsymbol{x},\boldsymbol{X})\Big\rangle_G > 0,
$$

where $\boldsymbol{x}$ is sampled according to the posterior distribution of $\boldsymbol{X}$. Sampling from the posterior distribution $\mathbb{P}(\boldsymbol{X} = \cdot|\boldsymbol{G})$ provides thus an estimator that achieves a non zero overlap: the community detection problem is solvable.

Suppose now $\lambda < \lambda_c$. Suppose that the community detection problem is solvable. There is therefore an estimator $\boldsymbol{a}$ that achieves a non zero overlap. Lemma 5.9.7 gives then

$$\alpha \stackrel{\text{def}}{=} \liminf_{d\to\infty} \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}|\sum_{i=1}^{n} \widetilde{X}_i\widetilde{a}_i| > 0$$

Compute now for $\delta \in (0,1]$

$$\begin{aligned}\frac{2}{n(n-1)}\sum_{i<j}\mathbb{E}(\widetilde{X}_i\widetilde{X}_j - \delta\widetilde{a}_i\widetilde{a}_j)^2 &= \frac{1}{n^2}\sum_{i,j}\mathbb{E}(\widetilde{X}_i\widetilde{X}_j - \delta\widetilde{a}_i\widetilde{a}_j)^2 + o(1)\\ &= \frac{1}{n^2}\sum_{i\neq j}\mathbb{E}(\widetilde{X}_i^2\widetilde{X}_j^2) + O(\delta^2) - \frac{2\delta}{n^2}\mathbb{E}(\sum_{i=1}^{n}\widetilde{X}_i\widetilde{a}_i)^2 + o(1)\\ &\leq 1 + O(\delta^2) - 2\delta\Big(\frac{1}{n}\mathbb{E}|\sum_{i=1}^{n}\widetilde{X}_i\widetilde{a}_i|\Big)^2 + o(1).\end{aligned}$$

So that

$$\liminf_{d\to\infty}\liminf_{n\to\infty}\mathrm{MMSE}_n^G(\lambda) \leq \liminf_{d\to\infty}\liminf_{n\to\infty}\frac{2}{n(n-1)}\sum_{i<j}\mathbb{E}(\widetilde{X}_i\widetilde{X}_j - \delta\widetilde{a}_i\widetilde{a}_j)^2 = 1 - 2\delta\alpha^2 + O(\delta^2).$$

The right-hand side will be strictly inferior to 1 for δ sufficiently small. This is contradictory with Proposition 5.9.2 (recall that $\lambda < \lambda_c$). The community detection problem is not solvable. Theorem 5.9.2 is proved.

Chapter 6

Statistical limits of rank-one tensor estimation

We consider in this chapter the tensor-analog of the spiked Wigner model from Chapter 3, namely the "spiked tensor model". Let $k \geq 2$ and consider

$$\boldsymbol{Y} = \sqrt{\frac{\lambda}{n^{k-1}}} \boldsymbol{X}^{\otimes k} + \boldsymbol{Z}\,, \tag{6.0.1}$$

where $\boldsymbol{X} \in \mathbb{R}^n$ such that $\|\boldsymbol{X}\| \simeq \sqrt{n}$ and $Z_{i_1,\dots,i_k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. The spiked tensor model (6.0.1) was introduced by Montanari and Richard [180], as a natural extension to tensors of the spiked matrix models studied in the previous chapters. The paper [180] shows that there exists a finite value λ_c such that weak recovery is possible for $\lambda > \lambda_c$. Montanari, Reichman and Zeitouni [150] obtained bounds on λ_c that were then improved by Perry, Wein and Bandeira [174] who obtained tight bounds in the $k \to \infty$ limit. The threshold for hypothesis testing is also of order 1 [150, 174] and precise expressions for the threshold has recently been obtained in [47, 49].

In this chapter we derive in Section 6.1 precise information-theoretic limits for estimation. We then turn our attention to the maximum likelihood estimator in Section 6.3 and compute its asymptotic performance.

6.1 Information-theoretic limits

6.1.1 Large order limit: $k \to \infty$

We start by a simple analysis of the symmetric tensor estimation model (6.1.5) with Rademacher prior $P_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$ in the limit $k \to \infty$. This large k scenario has been studied in [174], where the detection problem was also investigated. We suppose to observe here

$$Y_{i_1,\dots,i_k} = \sqrt{\frac{\lambda}{n^{k-1}}} X_{i_1} \dots X_{i_k} + Z_{i_1,\dots,i_k}\,, \tag{6.1.1}$$

for all $(i_1,\dots,i_k) \in \{1,\dots,n\}^k$. The $Z_{i_1,\dots,i_k}$ are i.i.d. standard Gaussian, independent from $X_1,\dots,X_n \overset{\text{i.i.d.}}{\sim} \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$. We will denote by

$$F_n^{(k)}(\lambda) = \frac{1}{n}\mathbb{E}\log \sum_{\boldsymbol{x}\in\{-1,1\}^n} 2^{-n} \exp\Big(\sum_{i_1,\dots,i_k} \sqrt{\frac{\lambda}{n^{k-1}}} Y_{i_1,\dots,i_k} x_{i_1} \dots x_{i_k} - \frac{\lambda n}{2} \Big)$$

and $\mathrm{MMSE}_n^{(k)}(\lambda) = \frac{1}{n^k}\mathbb{E}\|\boldsymbol{X}^{\otimes k} - \mathbb{E}[\boldsymbol{X}^{\otimes k}|\boldsymbol{Y}]\|^2$ the corresponding free energy and MMSE.

We will see that simple arguments (that does not require the knowledge of the exact formulas of Theorem 6.1.1 and Corollary 6.1.2 that we will see in the next section) give that for large values of k we have

- if $\lambda \leq 2\log(2) - O(\log(k)^2/k^2)$ then $\lim_{n\to\infty} \mathrm{MMSE}_n^{(k)}(\lambda) = 1$,
- if $\lambda \geq 2\log(2)$ then $\lim_{n\to\infty} \mathrm{MMSE}_n^{(k)}(\lambda) = O(\log(k)^2/k^2)$.

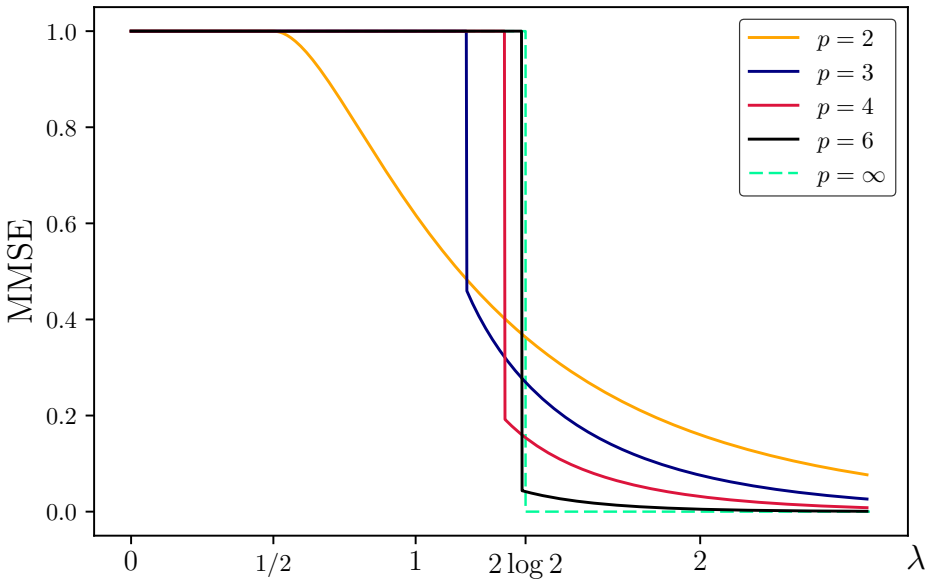


Figure 6.1: Minimal Mean Squared Errors $\mathrm{MMSE}_n^{(k)}$ for tensor estimation (6.1.1) with Rademacher prior, for $k = 2, 3, 4, 6$, as given by Corollary 6.1.2.

We start with the study of the free energy. For $\lambda_0 \in \mathbb{R}$ we define

$$f_{\lambda_0}(\lambda) = \begin{cases} 0 & \text{if} \quad \lambda \leq \lambda_0\,, \\ \frac{1}{2}(\lambda - \lambda_0) & \text{if} \quad \lambda \geq \lambda_0\,. \end{cases}$$

Proposition 6.1.1

There exists an increasing sequence $(\lambda_k^\sim)_{k\geq 2}$ such that $\lambda_2^\sim = 1/2$ and $\lambda_k^\sim = 2\log(2) - O(\log(k)^2/k^2)$ and for all $k, n, \lambda \geq 0$

$$f_{2\log(2)}(\lambda) \leq F_n^{(k)}(\lambda) \leq f_{\lambda_k^\sim}(\lambda) + \sqrt{\frac{\lambda}{n}} + \frac{\log(2n+1)}{n}\,. \tag{6.1.2}$$

Proof. We start with the lower bound. For all $\lambda \geq 0$ we have $F_n^{(k)}(\lambda) \geq 0$ and

$$\begin{aligned} F_n^{(k)}(\lambda) &= \frac{1}{n}\mathbb{E}\log \sum_{\boldsymbol{x}\in\{-1,1\}^n} \frac{1}{2^n} \exp\Big(\sum_{i_1,\dots,i_k} \sqrt{\frac{\lambda}{n^{k-1}}} Y_{i_1,\dots,i_k} x_{i_1}\dots x_{i_k} - \frac{\lambda n}{2}\Big) \\ &\geq \frac{1}{n}\mathbb{E}\log\left(\frac{1}{2^n} \exp\Big(\sum_{i_1,\dots,i_k} \sqrt{\frac{\lambda}{n^{k-1}}} Y_{i_1,\dots,i_k} X_{i_1}\dots X_{i_k} - \frac{\lambda n}{2}\Big)\right) = \frac{\lambda}{2} - \log(2)\,. \end{aligned}$$

Let us now prove the upper-bound. For $\boldsymbol{x} \in \{-1,1\}^n$ let us write $Z(\boldsymbol{x}) = n^{-k/2} \sum_{i_1,\dots,i_k} Z_{i_1,\dots,i_k} x_{i_1} \dots x_{i_k}$. By Jensen's inequality, we have

$$F_n^{(k)}(\lambda) \le \frac{1}{n}\mathbb{E}\log\left(\sum_{\boldsymbol{x}\in\{-1,1\}^n} \frac{1}{2^n}\mathbb{E}\left[\exp\left(\sqrt{\lambda n}Z(\boldsymbol{x}) + \lambda n(\boldsymbol{x}\cdot\boldsymbol{X})^k - \frac{\lambda n}{2}\right)\middle|\boldsymbol{X}, Z(\boldsymbol{X})\right]\right).$$

For $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in \{-1,1\}^n$ fixed, the covariance between $Z(\boldsymbol{x}^{(1)})$ and $Z(\boldsymbol{x}^{(2)})$ is $\mathbb{E}[Z(\boldsymbol{x}^{(1)})Z(\boldsymbol{x}^{(2)})] = (\boldsymbol{x}^{(1)}\cdot\boldsymbol{x}^{(2)})^k$. Consequently, for $\boldsymbol{x} \in \{-1,1\}^n$ the law of $Z(\boldsymbol{x})$ conditionally on $Z(\boldsymbol{X})$ is $\mathcal{N}\big((\boldsymbol{x}\cdot\boldsymbol{X})^k Z(\boldsymbol{X}),\, 1-(\boldsymbol{x}\cdot\boldsymbol{X})^{2k}\big)$ and therefore

$$\mathbb{E}\left[\exp\left(\sqrt{\lambda n}Z(\boldsymbol{x})\right)\middle|\boldsymbol{X}, Z(\boldsymbol{X})\right] = \exp\left(\sqrt{\lambda n}(\boldsymbol{x}\cdot\boldsymbol{X})^k Z(\boldsymbol{X}) + \frac{\lambda n}{2}(1-(\boldsymbol{x}\cdot\boldsymbol{X})^{2k})\right).$$

We obtain thus:

$$\begin{aligned} F_n^{(k)}(\lambda) &\le \frac{1}{n}\mathbb{E}\log\left(\sum_{\boldsymbol{x}\in\{-1,1\}^n} \frac{1}{2^n}\exp\left(\sqrt{\lambda n}|Z(\boldsymbol{X})| + \frac{\lambda n}{2}(2(\boldsymbol{x}\cdot\boldsymbol{X})^k - (\boldsymbol{x}\cdot\boldsymbol{X})^{2k})\right)\right) \\ &\le \frac{1}{n}\mathbb{E}\log\left(\sum_{\boldsymbol{x}\in\{-1,1\}^n} \frac{1}{2^n}\exp\left(\frac{\lambda n}{2}(2(\boldsymbol{x}\cdot\boldsymbol{X})^k - (\boldsymbol{x}\cdot\boldsymbol{X})^{2k})\right)\right) + \sqrt{\frac{\lambda}{n}}. \end{aligned}$$

Now, for $k \in \{-n,\dots,n\}$, we have

$$\#\left\{\boldsymbol{x}\in\{-1,1\}^n \;\middle|\; \sum_{i=1}^n x_i X_i = k\right\} \le 2^n \exp\big(-nh(k/n)\big),$$

where $h(t) = \frac{1+t}{2}\log(1+t) + \frac{1-t}{2}\log(1-t)$. This gives

$$\begin{aligned} F_n^{(k)}(\lambda) &\le \frac{1}{n}\mathbb{E}\log\left(\sum_{k=-n}^{n} \exp\left(\frac{n}{2}(2\lambda(k/n)^k - \lambda(k/n)^{2k} - 2h(k/n))\right)\right) + \sqrt{\frac{\lambda}{n}} \\ &\le \frac{1}{n}\mathbb{E}\log\left((2n+1)\exp\left(\frac{n}{2}\max_{t\in[0,1]}\{2\lambda t^k - \lambda t^{2k} - 2h(t)\}\right)\right) + \sqrt{\frac{\lambda}{n}} \\ &= \frac{1}{2}\max_{t\in[0,1]}\left\{\lambda(2t^k - t^{2k}) - 2h(t)\right\} + \frac{\log(2n+1)}{n} + \sqrt{\frac{\lambda}{n}}. \end{aligned} \tag{6.1.3}$$

The function $\lambda \mapsto \max_{t\in[0,1]}\{\lambda(2t^k - t^{2k}) - 2h(t)\}$ is continuous, 1-Lipschitz and equal to 0 for $\lambda \le \lambda_k^\sim$ where

$$\lambda_k^\sim \stackrel{\text{def}}{=} \inf_{t\in(0,1]} \frac{2h(t)}{2t^k - t^{2k}}.$$

Therefore, for $\lambda \ge \lambda_k^\sim$, $\max_{t\in[0,1]}\{\lambda(2t^k - t^{2k}) - 2h(t)\} \le \lambda - \lambda_k^\sim$, which combined to (6.1.3) proves (6.1.2).

It remains to show $\lambda_2^\sim = 1/2$ and $2\log(2) - O(\log(k)^2/k^2) \le \lambda_k^\sim \le 2\log(2)$. Let us start with the case $k=2$. The maximizer t of $t \mapsto \lambda(2t^2 - t^4) - 2h(t)$ verifies

$$4\lambda(t - t^3) = 2h'(t) = \log\left(\frac{1+t}{1-t}\right) = 2\,\mathrm{artanh}(t).$$

Therefore $t = \tanh(2\lambda(t-t^3))$. For $\lambda \le 1/2$, this equation admits a unique solution $t = 0$, whereas of $\lambda > 1/2$ it admits a second solution $t' > 0$ and $\max_{t\in[0,1]}\lambda(2t^2 - t^4) - 2h(t) > 0$. Thus $\lambda_2^\sim = 1/2$.

Let now $k \geq 3$. Let t_k be the largest minimizer of $h_k : t \mapsto 2h(t)/(2t^k - t^{2k})$ on $(0,1]$. One has $t_k \in (0,1)$. For $t \in (0,1)$, $h_k'(t)$ has the same sign as

$$\frac{t(\log(1+t) - \log(1-t))}{2h(t)} - \frac{2k(1-t^k)}{2-t^k}\,, \tag{6.1.4}$$

which is decreasing in k. This gives that $(t_k)_{k\geq 3}$ is increasing and converges to 1, which is the only possible limit because t_k cancels (6.1.4). One has also $t_k^k \to 1$. Define $u_k = 1 - t_k$. Since $t_k^k \to 1$, one has $ku_k \to 0$. Then $t_k^k = \exp(k\log(1-u_k)) = 1 - ku_k + O(ku_k^2)$. We get

$$-\log(u_k) + o(1) = k\log(2)\frac{ku_k + O(u_k^2 k)}{1+o(1)} = \log(2)k^2 u_k + o(1)\,.$$

Therefore

$$2\log(k) = \log(2)k^2 u_k + \log(k^2 u_k) + o(1) \sim \log(2)k^2 u_k\,.$$

We deduce that $1 - t_k = u_k \sim \frac{2\log(k)}{\log(2)k^2}$. We have then $2t_k^k - t_k^{2k} = 1 - k^2u_k^2 + o(k^2u_k^2)$ and $2h(t_k) = 2\log(2) + O(u_k) + u_k\log(u_k)$. We conclude:

$$\lambda_k^{\sim} = h_k(t_k) = 2\log(2) - O(\log(k)^2/k^2)\,.$$

□

Using the I-MMSE relations (1.3.4) as usual, we deduce:

Corollary 6.1.1

If $\lambda < \lambda_k^{\sim}$ then

$$\mathrm{MMSE}_n^{(k)}(\lambda) \xrightarrow[n\to\infty]{} 1\,,$$

while for $\lambda > 2\log(2)$:

$$\limsup_{n\to\infty} \mathrm{MMSE}_n^{(k)}(\lambda) \leq \frac{2\log(2) - \lambda_k^{\sim}}{\lambda - 2\log(2)} = \frac{O(\log(k)^2/k^2)}{\lambda - 2\log(2)}\,.$$

Proof. By the I-MMSE relation (1.3.4), we have $\mathrm{MMSE}_n^{(k)}(\lambda) = 1 - 2F_n^{(k)\prime}(\lambda)$. If $\lambda < \lambda_k^{\sim}$ then by convexity and the fact that the free energy is non-decreasing:

$$0 \leq F_n^{(k)\prime}(\lambda) \leq \frac{F_n^{(k)}(\lambda_k^{\sim}) - F_n^{(k)}(\lambda)}{\lambda_k^{\sim} - \lambda} \xrightarrow[n\to\infty]{} 0\,,$$

by Proposition 6.1.1. This proves the result for $\lambda < \lambda_k^{\sim}$. For $\lambda > 2\log(2)$ again, by convexity and Proposition 6.1.1 we have

$$\begin{aligned}\liminf_{n\to\infty} F_n^{(k)\prime}(\lambda) &\geq \liminf_{n\to\infty} \frac{F_n^{(k)}(\lambda) - F_n^{(k)}(2\log(2))}{\lambda - 2\log(2)} \\ &\geq \frac{f_{2\log(2)}(\lambda) - f_{\lambda_k^{\sim}}(2\log(2))}{\lambda - 2\log(2)} = \frac{1}{2} - \frac{2\log(2) - \lambda_k^{\sim}}{2(\lambda - 2\log(2))}\,,\end{aligned}$$

which concludes the proof. □

The "abrupt" phase transition at $\lambda = 2\log(2)$ that we see on Figure 6.1 reminds of the phase transition for the "needle in a haystack" problem seen in Section 1.4. This is not surprising, and this has been known for a long time in statistical physics: the Random Energy Model (which is the non-planted analog of the "needle in a haystack" problem) can be seen as the $k \to \infty$ limit of the k-spin model (which corresponds to the spiked tensor model (6.1.1)), see [57].

6.1.2 Information-theoretic limits

We turn now our attention to the precise expression to the limit of the minimal mean square error in the model (6.0.1) for general priors P_0. It will be useful (notably for the proof of Proposition 6.1.2 below) to consider a more general model, where we observe multiple tensors like (6.0.1), of different orders. This is the analog of the "mixed p-spin model" in statistical physics, see for instance [170]. Let $K \geq 1$ and $\lambda_1, \dots \lambda_K \geq 0$. Assume that we observe, for all $k \in \{1, \dots, K\}$:

$$Y_{i_1,\dots,i_k} = \sqrt{\frac{\lambda_k}{n^{k-1}}} X_{i_1} \dots X_{i_k} + Z_{i_1,\dots,i_k} \quad \text{for } 1 \leq i_1, \dots, i_k \leq n, \tag{6.1.5}$$

where $\boldsymbol{X} = (X_1, \dots, X_n) \overset{\text{i.i.d.}}{\sim} P_0$ and $(Z_{i_1,\dots,i_k})_{i_1,\dots,i_k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ are independent. We define the Hamiltonian

$$H_n(\boldsymbol{x}) = \sum_{k=1}^{K} \sum_{i_1,\dots,i_k} \sqrt{\frac{\lambda_k}{n^{k-1}}} Y_{i_1,\dots,i_k} x_{i_1} \dots x_{i_k} - \frac{\lambda_k}{2n^{k-1}} (x_{i_1} \dots x_{i_k})^2 \,, \tag{6.1.6}$$

for $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$. The posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$ reads then:

$$dP(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{Y}) = \frac{1}{\mathcal{Z}_n} dP_0^{\otimes n}(\boldsymbol{x}) e^{H_n(\boldsymbol{x})} \,, \tag{6.1.7}$$

where $\mathcal{Z}_n$ is the appropriate normalizing factor. The free energy is thus $F_n(\boldsymbol{\lambda}) = \frac{1}{n}\mathbb{E} \log \mathcal{Z}_n$. Again, we will express the limit of $F_n(\boldsymbol{\lambda})$ using the following "potential":

$$\mathcal{F} : (\boldsymbol{\lambda}, q) \mapsto \psi_{P_0}\left(\sum_{k=1}^{K} \lambda_k k q^{k-1}\right) - \frac{1}{2}\sum_{k=1}^{K} (k-1)\lambda_k q^k \,, \tag{6.1.8}$$

where ψ_{P_0} is the free energy of the scalar channel (1.3.5), defined by (1.3.6).

Theorem 6.1.1 (*Replica-Symmetric formula for the spiked tensor model*)

Let P_0 be a probability distribution over $\mathbb{R}$, with finite moment of order $2K$. Then, for all $\boldsymbol{\lambda} \in [0, +\infty)^K$

$$F_n(\boldsymbol{\lambda}) \xrightarrow[n\to\infty]{} \sup_{q \geq 0} \mathcal{F}(\boldsymbol{\lambda}, q) \,. \tag{6.1.9}$$

Theorem 6.1.1 was proved in [136] by the same arguments used for Theorem 3.2.1. The paper [20] gave then an alternative proof. In the case of a binary signal $P_0 = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ the work [124] obtained bounds on the limit of F_n. However, these bounds are not tight for all values of λ. By the I-MMSE relation (1.3.4) we deduce from Theorem 6.1.1 the limit of the Minimum Mean Squared Error:

$$\mathrm{MMSE}_n^{(k)}(\boldsymbol{\lambda}) = \inf_{\widehat{\theta}} \left\{ \frac{1}{n^k} \mathbb{E} \left[\sum_{i_1,\dots,i_k} \left(X_{i_1} \dots X_{i_k} - \widehat{\theta}(\boldsymbol{Y})_{i_1 \dots i_k} \right)^2 \right] \right\} ,$$

where the infimum is taken over all measurable functions $\widehat{\theta}$ of the observations $\boldsymbol{Y}$.

Corollary 6.1.2

For all $\boldsymbol{\lambda} \in [0,+\infty)^K$ such that $\mathcal{F}(\boldsymbol{\lambda},\cdot)$ admits a unique maximizer $q^*(\boldsymbol{\lambda})$ over $\mathbb{R}_{\geq 0}$, we have

$$\mathrm{MMSE}_n^{(k)}(\boldsymbol{\lambda}) \xrightarrow[n\to\infty]{} \left(\mathbb{E}_{P_0} X^2\right)^k - q^*(\boldsymbol{\lambda})^k \,. \tag{6.1.10}$$

The proof is exactly the same than for the matrix case, see the proof of Corollary 3.2.2.

Proposition 6.1.2

Assume that P_0 admits a finite moment of order $2K$. Let $\boldsymbol{\lambda} \in [0,+\infty)^K$ such that $\mathcal{F}(\boldsymbol{\lambda},\cdot)$ admits a unique maximizers $q^*(\boldsymbol{\lambda})$. Then

$$\frac{1}{n}\Big|\sum_{i=1}^n x_i X_i\Big| \xrightarrow[n\to\infty]{(d)} q^*(\boldsymbol{\lambda}),$$

where $\boldsymbol{x}$ is a sample from the posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$, independently of everything else. Moreover, if $\lambda_k > 0$ for some odd k, then

$$\frac{1}{n}\sum_{i=1}^n x_i X_i \xrightarrow[n\to\infty]{(d)} q^*(\boldsymbol{\lambda}).$$

From Proposition 6.1.2, we deduce using Proposition 1.2.2 and Remark 1.2.1:

Corollary 6.1.3

Assume that P_0 admits a finite moment of order $\max(2K,8)$. Let $\boldsymbol{\lambda} \in [0,+\infty)^K$ such that $\mathcal{F}(\boldsymbol{\lambda},\cdot)$ admits a unique maximizers $q^*(\boldsymbol{\lambda})$. Then

$$\sup_{\widehat{\boldsymbol{x}}} \mathbb{E}\left[\frac{|\langle \widehat{\boldsymbol{x}}(\boldsymbol{Y}), \boldsymbol{X}^*\rangle|}{\|\widehat{\boldsymbol{x}}(\boldsymbol{Y})\|\|\boldsymbol{X}^*\|}\right] \xrightarrow[n\to\infty]{} \sqrt{\frac{q^*(\boldsymbol{\lambda})}{\rho}}. \tag{6.1.11}$$

Proof of Proposition 6.1.2. We define the random variable $Q_n \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n x_i X_i$. Then for all $k \in \{1,\dots,K\}$, $\mathrm{MMSE}_n^{(k)}(\boldsymbol{\lambda}) = \mathbb{E}_{P_0}[X^2]^k - \mathbb{E}[Q_n^k] + o_n(1)$. Let us fix $\boldsymbol{\lambda} \in [0,+\infty)^K$ such that $\mathcal{F}(\boldsymbol{\lambda},\cdot)$ admits a unique maximizer $q^*(\boldsymbol{\lambda})$. From Corollary 6.1.2 we get that for all k such that $\lambda_k > 0$, we have

$$\mathbb{E}[Q_n^k] \xrightarrow[n\to\infty]{} q^*(\boldsymbol{\lambda})^k. \tag{6.1.12}$$

Let us fix such k. By the "I-MMSE" relation (Proposition 1.3.3) we have

$$\frac{\partial F_n}{\partial \lambda_{2k}}(\boldsymbol{\lambda})\Big|_{\lambda_{2k}^+} = \mathbb{E}[Q_n^{2k}] + o_n(1).$$

Since F_n is convex in λ_{2k}, we have by Proposition C.1:

$$\limsup_{n\to\infty} \mathbb{E}[Q_n^{2k}] = \limsup_{n\to\infty} \frac{\partial F_n}{\partial \lambda_{2k}}(\boldsymbol{\lambda})\Big|_{\lambda_{2k}^+} \leq \frac{\partial}{\partial \lambda_{2k}}\Big[\sup_{q\geq 0} \mathcal{F}(\boldsymbol{\lambda},q)\Big]\Big|_{\lambda_{2k}^+} = q^*(\boldsymbol{\lambda})^{2k}. \tag{6.1.13}$$

From (6.1.12) and (6.1.13) we get that $Q_n^k \xrightarrow[n\to\infty]{(d)} q^*(\boldsymbol{\lambda})^k$, which leads to $|Q_n| \xrightarrow[n\to\infty]{(d)} q^*(\boldsymbol{\lambda})$. If k is odd then $Q_n \xrightarrow[n\to\infty]{(d)} q^*(\boldsymbol{\lambda})$. □

6.2 Hardness of low-rank tensor estimation

The brutal jump of the minimal mean squared error on Figure 6.1 and the fact that tensor estimation is related (for large orders) to the "needle in a haystack" problem of Section 1.4 seems to indicate that the low-rank tensor estimation problem (6.0.1) for $k \geq 3$ is computationally hard. Moreover it is known that computing an eigenvalue, an eigenvector or even the rank of k-tensor is NP-hard when $k \geq 3$, [102]. The study of [180] supports this picture and shows that unless the signal-to-noise ratio λ goes to infinity with n, popular algorithms such as power iteration, tensor unfolding or message passing fail to recover the signal $\boldsymbol{X}$, when $\boldsymbol{X}$ is uniformly distributed on the sphere of radius $\sqrt{n}$.

The spiked tensor model is thus expected to be an extreme example of statistical problems that admit a *statistical-to-algorithmic gap*, when $\boldsymbol{X} \sim \mathrm{Unif}(\sqrt{n}\,\mathbb{S}^{n-1})$. The thresholds for estimation and detection are both order 1 in n; on the other hand, the thresholds for efficient testing and estimation are expected to diverge polynomially in n, $\lambda_{\mathrm{alg}} = O(n^{\alpha})$. Sharp algorithmic thresholds have been shown for semi-definite and spectral relaxations of the maximum likelihood problem [107, 106, 122] as well as optimization of the likelihood itself via Langevin dynamics [8]. Upper bounds have also been obtained for message passing and power iteration [180], as well as gradient descent [8]. All this approaches need at least $\alpha = (k-2)/4$ in order to succeed.

This suggests that we would be in a "hard regime" (where polynomial time algorithms can only achieve trivial performance) for all (finite) values of the signal-to-noise ratio λ. The work [136] provides a more optimistic vision that can be summarized as:

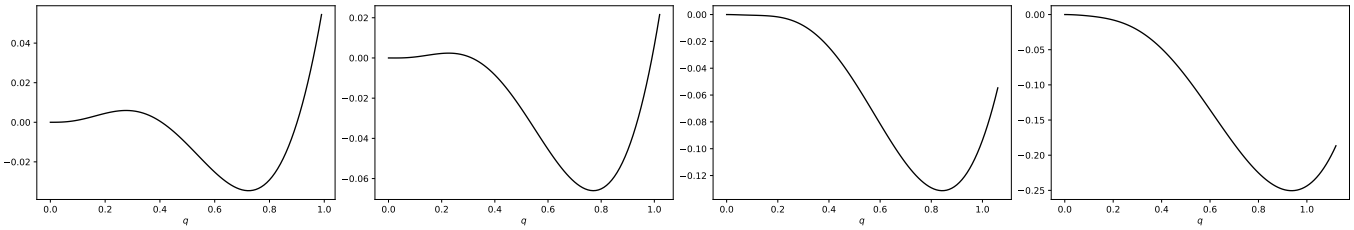
- If the distribution of the signal has zero mean (i.e. $\mathbb{E}_{P_0} X = 0$) we are indeed in a hard phase for all values of λ.
- However, if $\mathbb{E}_{P_0} X \neq 0$ then polynomial-time algorithms (as AMP) can achieve a non-trivial performance and can even be optimal if $\mathbb{E}_{P_0} X$ is not too small.

Let us give some intuition about these points. As we have seen in Section 3.3 the presence of a "hard regime" is characterized (if we believe that AMP are optimal among polynomial-time algorithms) by the fact that $q = 0$ is a local minimum of the potential $q \mapsto -\mathcal{F}(\lambda, q)$. We thus expand around $q = 0$

$$-\mathcal{F}(\lambda, q) = -\frac{\lambda k q^{k-1}}{2}\mathbb{E}_{P_0}[X]^2 + \frac{\lambda(k-1)}{2}q^k + O(q^{2(k-1)})\,,$$

because $\psi_{P_0}(0) = 0$ and $\psi'_{P_0} = \frac{1}{2}\mathbb{E}_{P_0}[X]^2$. Consequently, if $\mathbb{E}_{P_0} X = 0$ and $k \geq 3$, then $q = 0$ is a local minimum of $-\mathcal{F}(\lambda, \cdot)$ and we are in a "hard regime". But if the prior P_0 has a non-zero mean, then $q = 0$ is not a local minimum anymore and it is possible (with AMP for instance) to estimate the signal $\boldsymbol{X}$ with efficient algorithms.

The plots of Figure 6.2 confirm this picture. On the first plot (where $P_0 = \mathcal{N}(\mu = 0, 1)$) we observe that the local minimum $q = 0$ is separated from the global minimum by a barrier, which indicates hardness (see the discussion in Section 3.3.2). Since $\boldsymbol{X} = (X_1, \dots, X_n) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ is up to a normalization uniformly distributed on the sphere of radius $\sqrt{n}$ this is coherent with the results mentioned above. On the second plot, where the prior has a small mean $\mu = 0.15$, the local minimum at $q = 0$ disappears and is replaced by another local



(a) $\mu = 0$ (b) $\mu = 0.15$ (c) $\mu = 0.25$ (d) $\mu = 0.35$

Figure 6.2: Plots of $q \mapsto -\mathcal{F}(\lambda, q)$ for $k = 3$, $\lambda = 10$ and $P_0 = \mathcal{N}(\mu, 1)$, for different values of μ.

minimum at q_0, close to 0. It is possible in this situation to achieve non-trivial performance by efficient algorithms (as AMP), but it is again conjectured that their correlation with the planted solution $\boldsymbol{X}$ will be at most equal to q_0, which is quite small compared to the optimal overlap $q^*(\lambda) \simeq 0.8$. Polynomial-time algorithms can thus have non-trivial performance but are still far from optimal. On the third and fourth plot, we see that for larger means, the local minimum at q_0 disappears completely. It is now possible (using for instance AMP) to achieve the optimal performance in polynomial time.

6.3 Maximum likelihood estimation

6.3.1 Background and related work

We adopt here a slightly different normalization of the spiked tensor model in order to study maximum likelihood estimation. Suppose again that we are given an observation, $\boldsymbol{Y}$, which is a k-tensor of rank 1 in dimension n subject to additive Gaussian noise. That is,

$$\boldsymbol{Y} = \lambda\sqrt{n}\boldsymbol{X}^{\otimes k} + \boldsymbol{Z}, \tag{6.3.1}$$

where $\boldsymbol{X} \in \mathbb{S}^{n-1}$, the unit sphere in $\mathbb{R}^n$ and $\boldsymbol{Z}$ is an i.i.d. Gaussian k-tensor with $Z_{i_1\dots i_k} \sim \mathcal{N}(0,1)$[1]. Throughout this section, we assume that $\boldsymbol{X}$ is drawn from an uninformative prior, namely the uniform distribution on $\mathbb{S}^{n-1}$. Consequently, the typical order of magnitude of the entries of $\boldsymbol{X}$ is $1/\sqrt{n}$: the scaling we take here is equivalent to the one we used in Section 6.1.

It is straightforward to show that maximizing the log-likelihood is equivalent to maximizing $\langle \boldsymbol{Y}, \boldsymbol{x}^{\otimes k}\rangle$ over the sphere, $\boldsymbol{x} \in \mathbb{S}^{n-1}$. The maximum likelihood estimator (MLE) $\widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}$ is then defined as[2]

$$\widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}} = \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathbb{S}^{n-1}} \langle \boldsymbol{Y}, \boldsymbol{x}^{\otimes k}\rangle. \tag{6.3.2}$$

The goal of this section is to study the asymptotic behavior of the maximum likelihood estimator $\widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}$. The function $H_\lambda : \boldsymbol{x} \mapsto \sqrt{n}\langle \boldsymbol{x}^{\otimes k}, \boldsymbol{Y}\rangle$ has been extensively studied in statistical physics and mathematics.

[1] We note here that none of the results of this section are changed if one symmetrizes Z, i.e., if we work with the symmetric Gaussian k-tensor.

[2] As shown in Proposition 6.4.2, $\boldsymbol{x} \mapsto \langle \boldsymbol{x}^{\otimes k}, \boldsymbol{Y}\rangle$ admits almost surely a unique maximizer over $\mathbb{S}^{n-1}$ if k is odd, and two maximizers $\boldsymbol{x}^*$ and $-\boldsymbol{x}^*$ if k is even. In the case of even k, $\widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}$ is simply picked uniformly at random among $\{-\boldsymbol{x}^*, \boldsymbol{x}^*\}$.

In the case $\lambda = 0$, the Hamiltonian H_0 was first studied by Crisanti and Sommers in [54] using the heuristic "replica method". They computed the value of its maximum (the "ground state"), a result that was then rigorously confirmed by [200, 48, 10, 51, 111]. Crisanti and Sommers analyzed in [53] the complexity of the function H_0. In order to characterize how complex the function H_0 can be, one of the main quantity of interest is the number of critical points at a given energy (i.e. likelihood) and with a given correlation with the signal X:

$$\mathrm{Crt}_\star(M, E) \stackrel{\mathrm{def}}{=} \sum_{\boldsymbol{x}:\nabla H_\lambda(\boldsymbol{x})=0} \mathbb{1}\Big\{\langle \boldsymbol{x}, \boldsymbol{X}\rangle \in M\Big\}\, \mathbb{1}\Big\{\frac{1}{n} H_\lambda(\boldsymbol{x}) \in E\Big\} \tag{6.3.3}$$

for $M \subset [-1, 1]$ and $E \subset \mathbb{R}$. One can define similarly $\mathrm{Crt}_0(M, E)$ the number of local maxima with correlation in M and energy in E. The paper [53] shows (by non-rigorous methods) that in the case $\lambda = 0$, $\mathrm{Crt}_0(E) \simeq \exp(nS_0(E))$ provided $E \subset [E_\infty(k), +\infty)$ for some value $E_\infty(k)$. Another derivation of this prediction was done by Fyodorov [86] using tools from random matrix theory.

The work of Auffinger, Ben Arous and Černỳ [11, 10] confirmed rigorously this picture by proving (using the Kac-Rice formula and results from random matrix theory) the conjectured limiting expression for $\frac{1}{n}\log \mathbb{E}[\mathrm{Crt}_\star]$, $\frac{1}{n}\log \mathbb{E}[\mathrm{Crt}_0]$, and for other even more precise quantities. Subag [197] showed then, by a second moment computation, that the number of critical point (at which the Hamiltonian is larger than some value $E_\infty(k)$) is actually concentrated around its mean, giving a precise understanding of the typical number of critical points of the function H_0.

In the case $\lambda > 0$, the model was first considered by Gillin and Sherrington [91] who analyzed the Gibbs measure associated to the Hamiltonian H_λ. The study of the complexity of H_λ was carried out by Ben Arous, Mei, Montanari and Nica [9]. The authors compute the (normalized) logarithm of the expected number of local minima below a certain energy level via the Kac-Rice approach and show that there is a transition at a point λ_s such that for $\lambda < \lambda_s$ this is negative for any strictly positive correlation, and for $\lambda > \lambda_s$ it has a zero with correlation bounded away from zero. In [183], study the (normalized) logarithm of the (random) number of local minima via a novel (but non rigorous) replica theoretic approach. They argued that there are in fact two transitions for the log-likelihood, called λ_s and λ_c. First, for $\lambda < \lambda_s$, all local maxima of the log-likelihood only achieve asymptotically vanishing correlation. For $\lambda_s < \lambda < \lambda_c$, there is a local maximum of the log-likelihood with non-trivial correlation but the maximum likelihood estimator still has vanishing correlation. Finally, for $\lambda_c < \lambda$ the maximum likelihood estimator has strictly positive correlation. In particular, if we let $m(\lambda)$, denote the limiting value of the correlation of the maximum likelihood estimator and X, they predict that $m(\lambda)$ has a jump discontinuity at λ_c. We will verify this with Theorem 6.3.1 in the next section.

6.3.2 Main results

In order to state our main results, we need to introduce some notations. Define for $t \in [0, 1)$,

$$f_\lambda(t) = \lambda^2 t^k + \log(1 - t) + t \tag{6.3.4}$$

and let

$$\lambda_c = \sup\Big\{\lambda \geq 0 \,\Big|\, \sup_{t\in[0,1)} f_\lambda(t) \leq 0\Big\}. \tag{6.3.5}$$

As we will see with Proposition 6.4.1 below, it can be deduced from the results of Section 6.1 that λ_c is the information-theoretic threshold for the model (6.3.1). Our main result is that the preceding transition is also the transition for which maximum likelihood estimation yields an estimator which achieves positive correlation with $\boldsymbol{X}$. Let $q_*(\lambda)$ be defined by

$$q_*(\lambda) = \begin{cases} 0 & \text{if} \quad \lambda < \lambda_c \\ \arg\max_{t \in [0,1)} f_\lambda(t) & \text{if} \quad \lambda > \lambda_c. \end{cases} \tag{6.3.6}$$

As shown in Lemma 6.4.7, the function f_λ admits a unique positive maximizer on $[0, 1)$ when $\lambda > \lambda_c$, so that this is well-defined. Let z_k denote the unique zero on $(0, +\infty)$ of

$$\varphi_k(z) = \frac{1+z}{z^2}\log(1+z) - \frac{1}{z} - \frac{1}{k}. \tag{6.3.7}$$

Finally, let

$$\mathrm{GS}_k = \frac{\sqrt{k}}{\sqrt{1+z_k}}\left(1 + \frac{z_k}{k}\right). \tag{6.3.8}$$

We then have the following.

Theorem 6.3.1

Let $\lambda \geq 0$ and $k \geq 3$. The following limit holds almost surely

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \max_{\boldsymbol{x} \in \mathbb{S}^{n-1}} \left\langle \boldsymbol{x}^{\otimes k}, \boldsymbol{Y} \right\rangle = \begin{cases} \mathrm{GS}_k & \text{if} \quad \lambda \leq \lambda_c \\ \sqrt{k}\dfrac{1+\lambda^2 q_*(\lambda)^{k-1}}{\sqrt{1+\lambda^2 k q_*(\lambda)^{k-1}}} & \text{if} \quad \lambda > \lambda_c. \end{cases} \tag{6.3.9}$$

Furthermore, we have that for $\lambda \neq \lambda_c$

$$\lim_{n\to\infty} \left|\left\langle \widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X} \right\rangle\right| = \sqrt{q_*(\lambda)}. \tag{6.3.10}$$

As a consequence of Corollary 6.4.1, the maximum likelihood estimator achieves maximal correlation. Unlike the case $k = 2$, the transition in $q_*(\lambda)$ is not continuous. See Figure 6.3. Theorem 6.3.1 was proved in [110] where it is also proved that λ_c is the threshold for hypothesis testing, in the case of even k larger than 6. If we denote by Q_λ the law of the observations $\boldsymbol{Y}$ given by (6.3.1), then for $\lambda < \lambda_c$,

$$\mathrm{D}_{\mathrm{TV}}(Q_\lambda, Q_0) \xrightarrow[n\to\infty]{} 0.$$

Below λ_c it is therefore impossible to distinguish the tensor $\boldsymbol{Y}$ from pure noise. As a consequence of Theorem 6.3.1, the threshold λ_c is also the threshold for multiple hypothesis testing: the maximum likelihood is able to distinguish between all of the hypotheses $\lambda > \lambda_c$.

6.3.3 Regarding the spinodal transition

While the regime $\lambda_s < \lambda < \lambda_c$ and the expected transition at λ_s is not relevant for testing and estimation, there is still a natural interpretation from the perspective of the landscape of the maximum likelihood. In [9, 183], this is explained explained in terms of

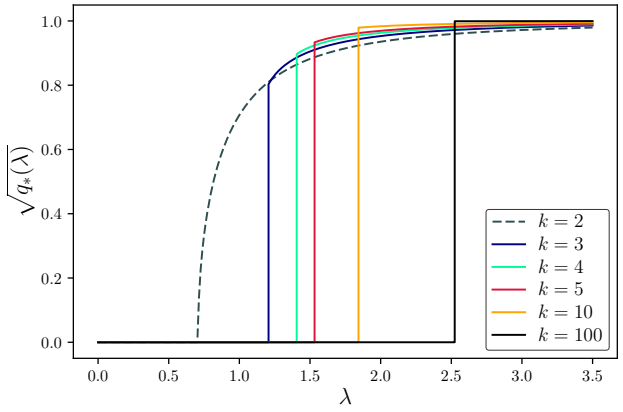


Figure 6.3: Asymptotic correlation $\lim_{n\to\infty} |\langle \hat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X}\rangle| = \sqrt{q_*(\lambda)}$ as a function of the signal-to-noise ratio λ, for different values of k.

the complexity. There is also an explanation in terms of the optimization of the maximum likelihood. We end this section with a brief discussion of this phase. Let λ_s be given by

$$\lambda_s = \sqrt{\frac{(k-1)^{k-1}}{k(k-2)^{k-2}}}. \tag{6.3.11}$$

Consider the constrained maximum likelihood,

$$E_\lambda(m) = \lim_{n\to\infty} \frac{1}{n} \max_{\boldsymbol{x}\in\mathbb{S}^{n-1}, \langle \boldsymbol{X},\boldsymbol{x}\rangle = m} \left\{ \lambda n \langle \boldsymbol{X}, \boldsymbol{x}\rangle^k + \sqrt{n}\langle \boldsymbol{Z}, \boldsymbol{x}^{\otimes k}\rangle \right\}. \tag{6.3.12}$$

This limit exists and is given by an explicit variational problem (see (6.4.5) below). For $\lambda > \lambda_s$, let $\sqrt{q_s(\lambda)}$ be the (unique) positive, strict local maximum of f_λ. By Lemma 6.4.7, this is well-defined and satisfies $q_s(\lambda) = q_*(\lambda)$ for $\lambda > \lambda_c$. In [183], it is argued by the replica method that $E_\lambda(m)$ has a local maximum at $\sqrt{q_s(\lambda)}$ for all $\lambda > \lambda_s$. Establishing this rigorously is a key step in our proof of Theorem 6.3.1. In particular, we prove the following, which is a direct consequence of Lemma 6.4.4 below.

Proposition 6.3.1

For $\lambda > \lambda_s$, the function E_λ has a strict local maximum at $\sqrt{q_s(\lambda)}$.

It is easy to verify (by direct differentiation) that the map $\lambda \mapsto E_\lambda(\sqrt{q_s(\lambda)})$ is strictly increasing on $(\lambda_s, +\infty)$. We have also that $E_{\lambda_c}(\sqrt{q_s(\lambda_c)}) = \mathrm{GS}_k$ by Lemma 6.4.4 and Lemma 6.4.8, so we get that for $\lambda_s < \lambda < \lambda_c$ the strict local maximum at $\sqrt{q_s(\lambda)}$ has $E_\lambda(\sqrt{q_s(\lambda)})$ strictly less than the maximum likelihood. In fact, (6.4.5) can be solved numerically, as it can be shown that one may reduce this variational problem, in the setting we consider here, to a two-parameter family of problems in three real variables. This is discussed in 6.4.2 below. In particular, see Figure 6.4 for an illustration of these two transitions in the case $k = 4$.

6.4 Proof of Theorem 6.3.1

In this section, we prove Theorem 6.3.1. We begin by providing a lower bound for the maximum likelihood for every $\lambda \geq 0$ using results on the ground state of the mixed p-

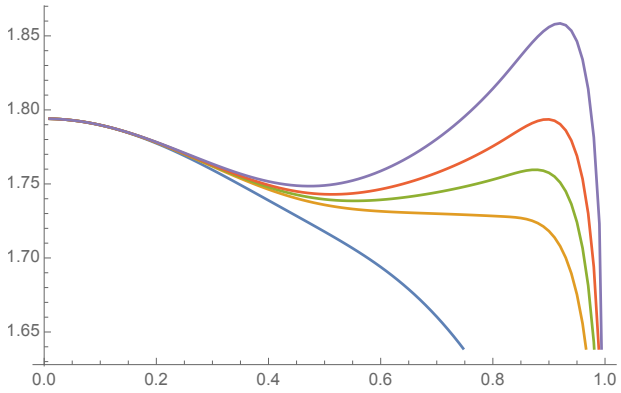


Figure 6.4: Asymptotic constrained maximum likelihood $E_\lambda(m)$ for $k = 4$ with $\lambda = 1, 1.299, 1.35, 1.405, 1.5$. Here $\lambda_s \approx 1.299$ and $\lambda_c \approx 1.405$. For $\lambda < \lambda_s$, the function is (numerically) seen to be monotone. A secondary maximum occurs at the transition $\lambda = \lambda_s$. This local maximum is bounded away from $m = 0$. Finally, at the information theoretic threshold λ_c, the maximum likelihood is now maximized at this second point.

spin model recently proved in [111, 51]. We then use the information-theoretic bound on the maximal correlation achievable by any estimator from [136] to obtain the matching upper bound. We end by proving the desired result for the correlation $(\widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X})$. In the remainder of this section, for ease of notation, we let

$$H_\lambda(\boldsymbol{x}) = H(\boldsymbol{x}) + \lambda n \langle \boldsymbol{x}, \boldsymbol{X} \rangle^k, \tag{6.4.1}$$

where $H(\boldsymbol{x}) = \sqrt{n} \sum_{i_1,\dots,i_k} Z_{i_1,\dots,i_k} x_{i_1} \dots x_{i_k}$.

6.4.1 Variational formula for the ground state of the mixed p-spin model

We begin by recalling the following variational formula for the ground state of the mixed p-spin model. Consider the Gaussian process indexed by $\boldsymbol{x} \in \mathbb{S}^{n-1}$:

$$Y_n(\boldsymbol{x}) = \sqrt{n} \sum_{p \geq 1} a_p \sum_{1 \leq i_1,\dots,i_p \leq n} g_{i_1,\dots,i_p} x_{i_1} \dots x_{i_p},$$

where $g_{i_1,\dots,i_p}$ are i.i.d. standard Gaussian random variables and $\sum_{p \geq 1} 2^p a_p^2 < \infty$. This last condition ensures that the sum above is almost surely finite. The covariance of Y_n is given by

$$\mathbb{E}\Big[Y_n(\boldsymbol{x}) Y_n(\boldsymbol{y})\Big] = n\xi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle),$$

where $\xi(t) = \sum_{p \geq 1} a_p^2 t^p$. Let $\mathcal{C}$ denote the subset of $C([0,1])$ of functions that are positive, non-increasing and concave. For any $h \geq 0$, we let $P_h : \mathcal{C} \to \mathbb{R}$ be

$$P_h(\phi) = \int \xi''(x)\phi(x) + \frac{1}{\phi(x)} dx + (h^2 + \xi'(0))\phi(0).$$

Set

$$\mathcal{G}(\xi, h) = \frac{1}{2} \min_{\phi \in \mathcal{C}} P_h(\phi). \tag{6.4.2}$$

Let us recall the following variational formula. For $\boldsymbol{x} \in \mathbb{S}^{n-1}$, we write $\boldsymbol{x} = (x_1, \dots, x_n)$.

Theorem 6.4.1 (*[51, 111]*)

For all $h \geq 0$,

$$\lim_{n\to\infty} \frac{1}{n} \max_{\boldsymbol{x}\in\mathbb{S}^{n-1}} \Big\{ Y_n(\boldsymbol{x}) + h\sqrt{n} \sum_{i=1}^{n} x_i \Big\} = \mathcal{G}(\xi, h),$$

almost surely and in L^1.

Remark 6.4.1. *While the results of [51, 111] are stated with $\xi'(0) = 0$, they still hold when $\xi'(0) > 0$ by replacing $\xi \mapsto \xi(t) - \xi'(0)t$ and $h^2 \mapsto h^2 + \xi'(0)$. To see this, simply note that the Crisanti-Sommers formula still holds in this setting by the main result of [48]. The reformulation from [111, Eq. (1.0.1)] is then changed by this replacement by simply repeating the integration by parts argument from [111, Lemma 6.1.1]. From here the arguments are unchanged under the above replacement.*

6.4.2 The lower bound

By Borell's inequality, the constrained maximum likelihood (6.3.12) concentrates around its mean with sub-Gaussian tails. In particular, combining this with Borel-Cantelli we see that

$$E_\lambda(m) = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\Big[\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}, \langle \boldsymbol{x}, \boldsymbol{X}\rangle = m} \Big\{ \lambda n \langle \boldsymbol{x}, \boldsymbol{X}\rangle^k + H(\boldsymbol{x}) \Big\} \Big]. \tag{6.4.3}$$

Clearly, $\liminf \frac{1}{n}\mathbb{E}\big[\max_{\mathbb{S}^{n-1}} H_\lambda\big] \geq E_\lambda(m)$ for all $m \in [-1, 1]$. Recall the definition of λ_s from (6.3.11) and $q_s(\lambda)$, see, e.g., Lemma 6.4.7. If we apply this for $\lambda > \lambda_s$ and $m = \sqrt{q_s(\lambda)} > \sqrt{1 - \frac{1}{k-1}}$ (by Lemma 6.4.7), Lemma 6.4.4 below will immediately yield the following lower bound.

Lemma 6.4.1

For all $\lambda > \lambda_s$,

$$\liminf_{n\to\infty} \mathbb{E}\Big[\frac{1}{n} \max_{\boldsymbol{x}\in\mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x}) \Big] \geq \sqrt{k} \frac{1 + \lambda^2 q_s(\lambda)^{k-1}}{\sqrt{1 + \lambda^2 k q_s(\lambda)^{k-1}}}. \tag{6.4.4}$$

We now turn to the proof of Lemma 6.4.4. We begin by observing the following explicit representation for E_λ.

Lemma 6.4.2

For all $m \in [-1, 1]$ the limit in (6.3.12) exists and

$$E_\lambda(m) = \lambda m^k + \mathcal{G}(\xi_m, 0), \tag{6.4.5}$$

where $\xi_m(t) = (m^2 + (1 - m^2)t)^k - m^{2k}$.

Proof. We begin by observing that by rotational invariance,

$$\frac{1}{n}\mathbb{E}\Big[\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}, \langle \boldsymbol{x}, \boldsymbol{X}\rangle = m} \Big\{ \lambda n \langle \boldsymbol{x}, \boldsymbol{X}\rangle^k + H(\boldsymbol{x}) \Big\} \Big] = \lambda m^k + \frac{1}{n}\mathbb{E}\Big[\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}, x_1 = m} H(\boldsymbol{x}) \Big].$$

Let $\boldsymbol{x} \in \mathbb{S}^{n-1}$ such that $x_1 = m$. Then

$$H(\boldsymbol{x}) \overset{(\mathrm{d})}{=} \sqrt{n} m^k g_{1,\dots,1} + \sqrt{n} \sum_{j=0}^{k-1} \binom{k}{j}^{1/2} m^j \sum_{2\leq i_1,\dots,i_{k-j}\leq n} g_{i_1,\dots,i_{k-j}} x_{i_1} \dots x_{i_{k-j}},$$

where $\big((g_{i_1,\dots,i_p})_{1\le i_1,\dots,i_p\le n}\big)_{p\le k}$ are i.i.d. standard Gaussians.

So that $\mathbb{E}\big[\max_{\boldsymbol{x}\in\mathbb{S}^{n-1},\, x_1=m} H(\boldsymbol{x})\big] = \mathbb{E}\big[\max_{\boldsymbol{x}\in\mathbb{S}^{n-2}} H_m(\boldsymbol{x})\big]$, where H_m is given by:

$$H_m(\boldsymbol{x}) = \sqrt{n}\sum_{j=0}^{k-1}\binom{k}{j}^{1/2} m^j(1-m^2)^{(k-j)/2} \sum_{1\le i_1,\dots,i_{k-j}\le n-1} g_{i_1,\dots,i_{k-j}} x_{i_1}\dots x_{i_{k-j}}.$$

The function H_m is a Gaussian process with covariance

$$\mathbb{E}[H_m(\boldsymbol{x})H_m(\boldsymbol{y})] = n\xi_m(\langle \boldsymbol{x},\boldsymbol{y}\rangle),$$

where ξ_m is given by

$$\xi_m(t) = \sum_{j=0}^{k-1}\binom{k}{j} m^{2j}(1-m^2)^{k-j}t^{k-j} = (m^2+(1-m^2)t)^k - m^{2k}. \tag{6.4.6}$$

We conclude using Theorem 6.4.1 to obtain the result. □

We now observe that for m large enough, this formula has a particularly simple form.

Lemma 6.4.3

For all $|m| \ge \sqrt{1-\frac{1}{k-1}}$ we have:

$$E_\lambda(m) = \lambda m^k + \sqrt{k(1-m^2)}. \tag{6.4.7}$$

Proof. In the setting of Theorem 6.4.1 it was also shown in [111, 51] that if $\xi'(1)+h^2 \ge \xi''(1)$ then $\mathcal{G}(\xi,h) = \sqrt{\xi'(1)+h^2}$. Since

$$\begin{aligned}\xi_m'(t) &= k(1-m^2)(m^2+(1-m^2)t)^{k-1}\\ \xi_m''(t) &= k(k-1)(1-m^2)^2(m^2+(1-m^2)t)^{k-2},\end{aligned}$$

the condition $\xi_m'(1) \ge \xi_m''(1)$ corresponds to $(k-1)(1-m^2)\le 1$, i.e. $|m| \ge \sqrt{1-\frac{1}{k-1}}$. When this holds, we get that

$$E_\lambda(m) = \lambda m^k + \mathcal{G}(\xi_m, 0) = \sqrt{\xi_m'(1)} = \lambda m^k + \sqrt{k(1-m^2)}$$

by (6.4.5). □

We end with the desired explicit formula for $E_\lambda(\sqrt{q}_s(\lambda))$.

Lemma 6.4.4

For all $\lambda > \lambda_s$, $\sqrt{q_s(\lambda)}$ is a local maximizer of E_λ and if we write $x(\lambda) = \lambda^2 k q_s^{k-1}(\lambda)$,

$$E_\lambda\Big(\sqrt{q_s(\lambda)}\Big) = \frac{\sqrt{k}}{\sqrt{1+x(\lambda)}}\Big(1+\frac{x(\lambda)}{k}\Big).$$

Proof. Differentiating the expression (6.4.7) for $m \ge \sqrt{1-\frac{1}{k-1}}$ yields

$$E_\lambda'(m) = \lambda k m^{k-1} - \frac{\sqrt{k}m}{\sqrt{1-m^2}} = k\frac{1}{1-m^2}\Big(\lambda k m^{k-1} + \frac{\sqrt{k}m}{\sqrt{1-m^2}}\Big)^{-1}\Big(\lambda^2 k m^{2k-2} - \lambda^2 k m^{2k} - m^2\Big)$$

so that the functions ϕ_λ, f_λ and $m^2 \mapsto E_\lambda(m)$ have precisely the same monotonicity on $[1-\frac{1}{k-1}, 1)$ (recall the expression of the derivatives f'_λ and ϕ'_λ given by (6.4.14)). Lemma 6.4.7 gives that $q_s(\lambda)$ is a local maximum of f_λ and ϕ_λ for $\lambda > \lambda_s$, $\sqrt{q_s(\lambda)}$ is therefore a local maximum of E_λ.

Let us now compute $E_\lambda(\sqrt{q_s(\lambda)})$. By Lemma 6.4.7, $q_s(\lambda) = \frac{x(\lambda)}{1+x(\lambda)}$. Consequently,

$$E_\lambda(q_s(\lambda)^{1/2}) = \lambda q_s(\lambda)^{k/2} + \sqrt{k(1-q_s(\lambda))} = \frac{\sqrt{k}}{\sqrt{1+x(\lambda)}}\Big(1 + \frac{x(\lambda)}{k}\Big).$$

□

6.4.3 The upper bound

We prove here the upper bound.

Lemma 6.4.5

For all $\lambda \geq 0$,

$$\limsup_{n\to\infty} \mathbb{E}\Big[\frac{1}{n}\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x})\Big] \leq \mathrm{GS}_k + \int_0^\lambda q_*(t)^{k/2}\, dt. \tag{6.4.8}$$

We defer the proof of this momentarily to observe the following information-theoretic bounds which will be useful in its proof.

Proposition 6.4.1

Assume that $\boldsymbol{X}$ is uniformly distributed over $\mathbb{S}^{n-1}$, independently of $\boldsymbol{Z}$. Then for all $\lambda \in (0, +\infty) \setminus \{\lambda_c\}$

$$\lim_{n\to\infty} \mathbb{E}\left[\Big\|\boldsymbol{X}^{\otimes k} - \mathbb{E}[\boldsymbol{X}^{\otimes k}|\boldsymbol{Y}]\Big\|^2\right] = 1 - q_*(\lambda)^k.$$

This follows from the results of Section 6.1.2, by approximating the uniform measure on $\mathbb{S}^{n-1}$ by an i.i.d. Gaussian measure. For the completeness, we provide a proof in Section 6.4.6. As a consequence of this, we have the following.

Corollary 6.4.1

Assume that $\boldsymbol{X}$ is uniformly distributed over $\mathbb{S}^{n-1}$, independently from Z. Then for all measurable functions $\widehat{\boldsymbol{x}} : (\mathbb{R}^n)^{\otimes k} \to \mathbb{S}^{n-1}$ and for all $\lambda \neq \lambda_c$ we have

$$\limsup_{n\to\infty} \mathbb{E}\left[\langle \widehat{\boldsymbol{x}}(\boldsymbol{Y}), \boldsymbol{X}\rangle^k\right] \leq q_*(\lambda)^{k/2}.$$

Proof. Compute

$$\begin{aligned}\mathbb{E}\Big[\Big\|\boldsymbol{X}^{\otimes k} - \big(\sqrt{q_*(\lambda)}\widehat{\boldsymbol{x}}(\boldsymbol{Y})\big)^{\otimes k}\Big\|^2\Big] &= \mathbb{E}[\|\boldsymbol{X}^{\otimes k}\|^2] + q_*(\lambda)^k\mathbb{E}[\|\widehat{\boldsymbol{x}}(\boldsymbol{Y})^{\otimes k}\|^2] - 2q_*(\lambda)^{k/2}\mathbb{E}\Big[\langle \widehat{\boldsymbol{x}}(\boldsymbol{Y}), \boldsymbol{X}\rangle^k\Big] \\ &= 1 + q_*(\lambda)^k - 2q_*(\lambda)^{k/2}\mathbb{E}\Big[\langle \widehat{\boldsymbol{x}}(\boldsymbol{Y}), \boldsymbol{X}\rangle^k\Big].\end{aligned}$$

Recall that the posterior mean, $\mathbb{E}[\boldsymbol{X}^{\otimes k}|\boldsymbol{Y}]$, uniquely achieves the minimal mean-square error over all square-integrable tensor-valued estimators, $\widehat{\boldsymbol{T}}(\boldsymbol{Y})$, for $\boldsymbol{X}^{\otimes k}$. The proposition follows then from Proposition 6.4.1 which gives

$$\liminf_{n\to\infty} \mathbb{E}\Big[\Big\|\boldsymbol{X}^{\otimes k} - \big(\sqrt{q_*(\lambda)}\widehat{\boldsymbol{x}}(\boldsymbol{Y})\big)^{\otimes k}\Big\|^2\Big] \geq \liminf_{n\to\infty} \mathbb{E}\Big[\Big\|\boldsymbol{X}^{\otimes k} - \mathbb{E}[\boldsymbol{X}^{\otimes k}|\boldsymbol{Y}]\Big\|^2\Big] = 1 - q_*(\lambda)^k.$$

□

With this in hand we may now prove Lemma 6.4.5.

Proof of Lemma 6.4.5. By Proposition 6.4.2 and an application of an envelope-type theorem (see, e.g., Proposition D.2), the map $\lambda \mapsto \frac{1}{n}\mathbb{E}\big[\max_{\mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x})\big]$ is differentiable for $\lambda \geq 0$, with derivative

$$\frac{\partial}{\partial\lambda}\mathbb{E}\Big[\frac{1}{n}\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x})\Big] = \mathbb{E}\Big[\langle \widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X}\rangle^k\Big]. \tag{6.4.9}$$

By [111, 51] we know that $\frac{1}{n}\mathbb{E}[\max_{\mathbb{S}^{n-1}} H_0] \to \mathrm{GS}_k$. The reverse Fatou lemma gives then

$$\limsup_{n\to\infty}\mathbb{E}\Big[\frac{1}{n}\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x})\Big] \leq \int_0^\lambda \limsup_{n\to\infty}\mathbb{E}\Big[\langle \widehat{\boldsymbol{x}}_\gamma^{\mathrm{ML}}, \boldsymbol{X}\rangle^k\Big]\,d\gamma + \mathrm{GS}_k \leq \int_0^\lambda q_*(\gamma)^{k/2}\,d\gamma + \mathrm{GS}_k,$$

where the second inequality follows from 6.4.1. □

6.4.4 Proof of first part of Theorem 6.3.1

By an elementary but tedious calculation (see Lemma 6.4.9) the right sides of (6.4.4) and (6.4.8) are equal for $\lambda \geq \lambda_c$ (recall that $q_*(\lambda) = q_s(\lambda)$ for such λ). Thus for all $\lambda > \lambda_c$,

$$\mathbb{E}\Big[\frac{1}{n}\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x})\Big] \xrightarrow[n\to\infty]{} \sqrt{k}\,\frac{1+\lambda^2 q_*(\lambda)^{k-1}}{\sqrt{1+\lambda^2 k q_*(\lambda)^{k-1}}}. \tag{6.4.10}$$

We will now prove that for $\lambda \leq \lambda_c$, $\overline{M}_n(\lambda) \xrightarrow[n\to\infty]{} \mathrm{GS}_k$, where $\overline{M}_n(\lambda)$ is defined by

$$\overline{M}_n(\lambda) = \mathbb{E}\left[\frac{1}{n}\max_{\boldsymbol{x}\in\mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x})\right].$$

Notice that $\overline{M}_n(\lambda)$ is convex as an expectation of a maximum of linear functions. By (6.4.9), it follows that $\overline{M}_n'(0^+) \geq 0$. (When k is odd, we use rotational invariance to see that it is in fact zero.) Consequently, $\overline{M}_n$ is non-decreasing on $[0, +\infty)$.

By [111, 51] (see Theorem 6.4.1), $\lim_{n\to\infty} \overline{M}_n(0) = \mathrm{GS}_k$. By (6.4.10) and Lemma 6.4.9,

$$\lim_{\lambda\to\lambda_c^+}\lim_{n\to\infty} \overline{M}_n(\lambda) = \mathrm{GS}_k.$$

Consequently, we obtain that for all $\lambda \in [0, \lambda_c]$, $\lim_{n\to\infty} \overline{M}_n(\lambda) = \mathrm{GS}_k$.

The almost sure convergence of (6.3.9) follows then from the convergence of the expectation $\overline{M}_n(\lambda)$, combined with Borell's inequality for suprema of Gaussian processes (see for instance [37, Theorem 5.8]) and the Borel-Cantelli Lemma. □

Remark 6.4.2. *By (6.4.5) , E_λ is given by a variational problem over the space $\mathcal{C}$. We first observe that one can easily solve this variational problem numerical due to the following simple reductions. First note that if we let ξ_m be as in (6.4.6), then $(1/\sqrt{\xi_m''})''$ is strictly positive, where the prime denotes differentiation in t. Thus by [111, Theorem 1.2.4], the minimizer ϕ must be of the form $\phi(s) = \int_s^1 d\nu_s$ where $\nu_s = \theta_1\delta_a + \theta_2\delta_b$, where $a, b \in [0,1]$ and $\theta_i \geq 0$. Thus the variational problem (6.4.5) is a variational problem over 4 parameters which can be solved numerically. These observations then rigorously justify the starting point of the discussion in [183, Section 4], namely the "RS" and "1RSB" calculation in [183, Sect. 4.B] in the regime they analyze, called the "$T = 0$" regime there. We refer the reader there for a more in-depth discussion, see [183, Sect. 4.C].*

6.4.5 Proof of second part of Theorem 6.3.1

We now turn to the second part of Theorem 6.3.1, namely (6.3.10). Let $M_n(\lambda)$ denote

$$M_n(\lambda) = \frac{1}{n} \max_{\boldsymbol{x} \in \mathbb{S}^{n-1}} H_\lambda(\boldsymbol{x}).$$

Let $\lambda \in (0, +\infty) \setminus \{\lambda_c\}$. By (6.3.9) and Lemma 6.4.9

$$\lim_{n\to\infty} M_n(\lambda) = \ell(\lambda) \overset{\text{def}}{=} \begin{cases} \mathrm{GS}_k & \text{if} \quad \lambda \le \lambda_c \\ \mathrm{GS}_k + \int_0^\lambda q_*(\gamma)^{k/2}\, d\gamma & \text{if} \quad \lambda > \lambda_c. \end{cases}$$

By Proposition 6.4.2 and the Milgrom-Segal envelope theorem (see Proposition D.2), M_n is differentiable in λ with derivative

$$M_n'(\lambda) = \langle \widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X} \rangle^k,$$

almost surely. As M_n is convex in λ (it is a maximum of linear functions), we see that for any $0 < h < \lambda$,

$$\frac{M_n(\lambda - h) - M_n(\lambda)}{h} \le M_n'(\lambda) \le \frac{M_n(\lambda + h) - M_n(\lambda)}{h}.$$

By taking the $n \to \infty$ limit, we get that almost surely

$$\frac{\ell(\lambda - h) - \ell(\lambda)}{h} \le \liminf_{n\to\infty} \langle \widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X} \rangle^k \le \limsup_{n\to\infty} \langle \widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X} \rangle^k \le \frac{\ell(\lambda + h) - \ell(\lambda)}{h}. \tag{6.4.11}$$

Since ℓ is differentiable for $\lambda \neq \lambda_c$, we may take $h \to 0$ to obtain $\lim_{n\to\infty} \langle \widehat{\boldsymbol{x}}_\lambda^{\mathrm{ML}}, \boldsymbol{X} \rangle^k = q_*(\lambda)^{k/2}$ almost surely, which proves (6.3.10). □

6.4.6 Auxiliary results

Uniqueness of minimizers

This section gathers some basic lemmas that will be useful for the analysis.

Proposition 6.4.2

Recall the definition (6.4.1) of H_λ. We have the following

- *If k is odd, then H_λ has almost surely one unique maximizer over $\mathbb{S}^{n-1}$.*
- *If k is even, then H_λ has almost surely two maximizers over $\mathbb{S}^{n-1}$, $\boldsymbol{x}^*$ and $-\boldsymbol{x}^*$.*

Proof. We note the following basic fact from the theory of Gaussian processes, see, e.g. [123].

Lemma 6.4.6

Let $(Z(t))_{t\in T}$ be a Gaussian process indexed by a compact metric space T such that $t \mapsto Z(t)$ is continuous almost surely. If the intrinsic quasi-metric, $d(s,t)^2 = \mathrm{Var}(Z(s) - Z(t))$, is a metric, i.e., $d(s,t) \neq 0$ for $s \neq t$, then Z admits a unique maximizer on T almost surely.

Observe H_λ is continuous on the compact $\mathbb{S}^{n-1}$. For $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{S}^{n-1}$, we have

$$\mathrm{Var}(H_\lambda(\boldsymbol{x}^1) - H_\lambda(\boldsymbol{x}^2)) = 2n(1 - \langle \boldsymbol{x}^1, \boldsymbol{x}^2 \rangle^k).$$

If k is odd, then the proposition follows directly from the Lemma. If k is even, we apply the Lemma on the quotient space $\mathbb{S}^{n-1}/\sim$ where $\sim$ denotes the equivalence relation defined by $\boldsymbol{x}^1 \sim \boldsymbol{x}^2 \Leftrightarrow (\boldsymbol{x}^1 = \boldsymbol{x}^2 \text{ or } \boldsymbol{x}^1 = -\boldsymbol{x}^2)$. □

Study of the asymptotic equations

Define, for all $q \in [0,1]$

$$\phi_\lambda(q) = \lambda^2 k q^{k-1} - \log(1 + \lambda^2 k q^{k-1}) - \lambda^2 (k-1) q^k. \tag{6.4.12}$$

Lemma 6.4.7

We have for all $\lambda > 0$,

$$\max_{q \in [0,1)} f_\lambda(q) = \max_{q \in [0,1]} \phi_\lambda(q)$$

Furthermore, if we let $\lambda_s = \sqrt{\frac{(k-1)^{k-1}}{k(k-2)^{k-2}}}$:

- *For $\lambda < \lambda_s$, then the functions f_λ and ϕ_λ are decreasing on $[0,1)$.*
- *For $\lambda > \lambda_s$, the functions f_λ and ϕ_λ have a strict local minimum at $q_u(\lambda)$ and a strict local maximum at $q_s(\lambda)$ where $0 < q_u < \frac{k-2}{k-1} < q_s < 1$, and both functions are strictly monotone on the intervals $(0, q_u)$, (q_u, q_s) and $(q_s, 1)$. Moreover, $q_s(\lambda)$ satisfies:*

$$q_s(\lambda) = \frac{\lambda^2 k q_s(\lambda)^{k-1}}{1 + \lambda^2 k q_s(\lambda)^{k-1}}\,. \tag{6.4.13}$$

Finally, for $\lambda > \lambda_c$, $q_*(\lambda) = q_s(\lambda)$ is the unique maximizer of f_λ and ϕ_λ over $[0,1)$.

Proof. We have for $q \in [0,1)$

$$\phi'_\lambda(q) = \frac{k(k-1)\lambda^2 q^{k-1}}{1 + \lambda^2 k q^{k-1}} h(q) \qquad \text{and} \qquad f'_\lambda(q) = \frac{h(q)}{1-q} \tag{6.4.14}$$

where $h(q) = \lambda^2 k q^{k-1} - \lambda^2 k q^k - q$. It suffices therefore to study the variations of f_λ. notice also that

$$\phi_\lambda(q) = f_\lambda(q) + h(q) - \log(1 + h(q)).$$

Since $f'_\lambda(q) = 0$ implies $h(q) = 0$, this implies that

$$\max_{q \in [0,1]} \phi_\lambda(q) = \max_{q \in [0,1)} f_\lambda(q)$$

and that these maxima are achieved at the same points. Let us now study the sign of the polynomial $h(q)$:

$$h(q) = qk\lambda^2 (q^{k-2} - q^{k-1} - \frac{1}{k\lambda^2}). \tag{6.4.15}$$

One verify easily that the polynomial $q^{k-2} - q^{k-1}$ achieves its maximum at $\frac{k-2}{k-1}$ and that the value of this maximum is $\frac{(k-2)^{k-2}}{(k-1)^{k-1}}$. We get that for $\lambda < \lambda_s$, $h'(q) < 0$ for all $q > 0$. For $\lambda > \lambda_s$ we get that h admits exactly 3 zeros on $\mathbb{R}$: $0 < q_u(\lambda) < q_s(\lambda) < 1$. Since the maximum of h is achieved at $\frac{k-2}{k-1}$ we get that $q_u(\lambda) < \frac{k-2}{k-1} < q_s(\lambda)$. This proves the two points of the lemma; (6.4.13) simply follows from the fact that $h(q_s(\lambda)) = 0$. The last statement of Lemma 6.4.7 is then an immediate consequence of the definition of λ_c. □

Recall that z_k is defined as the unique zero of $\varphi_k(z) = \frac{1+z}{z^2}\log(1+z) - \frac{1}{z} - \frac{1}{k}$ on $(0, +\infty)$.

Lemma 6.4.8

The mapping $\lambda \mapsto q_s(\lambda)$ is $\mathcal{C}^\infty$ on $(\lambda_s, +\infty)$. Moreover $\lambda^2 k q_s(\lambda_c)^{k-1} = z_k$.

Proof. The first part follows from a straightforward application of the implicit function theorem.

We get in particular that the mapping $\lambda \mapsto q_s(\lambda)$ is continuous for $\lambda > \lambda_s$. So by definition of λ_c and Lemma 6.4.7, $\phi_{\lambda_c}(q_s(\lambda_c)) = 0$. Let us write $x = \lambda^2 k q_s(\lambda_c)^{k-1}$.

$$0 = \phi_{\lambda_c}(q_s(\lambda_c)) = x - \log(1+x) - \frac{k-1}{k} x q_s(\lambda_c) = x - \log(1+x) - \frac{k-1}{k}\frac{x^2}{1+x}\,,$$

because $q_s(\lambda_c) = \frac{x}{1+x}$ (see (6.4.13)). This gives that $\varphi_k(x) = 0$ and thus $x = z_k$. □

Lemma 6.4.9

Let $\lambda > \lambda_c$ and write $x(\lambda) = \lambda^2 k q_s(\lambda)^{k-1}$. Then we have

$$\frac{\sqrt{k}}{\sqrt{1+x(\lambda)}}\Big(1 + \frac{x(\lambda)}{k}\Big) = \mathrm{GS}_k + \int_{\lambda_c}^{\lambda} q_s(\gamma)^{k/2} d\gamma\,.$$

Proof. Let us write $g(\lambda) = \frac{\sqrt{k}}{\sqrt{1+x(\lambda)}}(1 + \frac{x(\lambda)}{k})$. By Lemma 6.4.4, $\sqrt{q_s(\lambda)}$ is a local maximizer of E_λ and thus a critical point of E_λ. This gives

$$g'(\lambda) = \partial_\lambda\Big[E_\lambda(\sqrt{q_s(\lambda)})\Big] = \partial_\lambda E_\lambda(\sqrt{q_s(\lambda)}) = q_s(\lambda)^{k/2}.$$

The lemma follows then from the fact that $x(\lambda_c) = z_k$ by Lemma 6.4.8 and the definition (6.3.8) of GS_k. □

Proof of 6.4.1

For P_0 a probability distribution over $(\mathbb{R}^n)^{\otimes k}$ with finite second moment, we define the free energy

$$F_{P_0}(\gamma) = \frac{1}{n}\mathbb{E}\log\int P_0(d\boldsymbol{x}) \exp\Big(\sqrt{\gamma n}\langle \boldsymbol{x}, \boldsymbol{Z}\rangle + \gamma n \langle \boldsymbol{x}, \boldsymbol{X}_0\rangle - \frac{1}{2}\gamma n \|\boldsymbol{x}\|^2\Big)$$

where $\boldsymbol{X}_0 \sim P_0$ and $Z_{i_1,\dots,i_k} \sim \mathcal{N}(0,1)$ are independent. Proposition B.1 gives that for two probability distributions P_1, P_2 on $(\mathbb{R}^n)^{\otimes k}$ with finite second moment, we have

$$\big|F_{P_1}(\gamma) - F_{P_2}(\gamma)\big| \le \frac{\gamma}{2}\Big(\sqrt{\mathbb{E}_{P_1}\|\boldsymbol{X}_1\|^2} + \sqrt{\mathbb{E}_{P_2}\|\boldsymbol{X}_2\|^2}\Big) W_2(P_1, P_2),$$

where $W_2(P_1, P_2)$ denotes the Wasserstein distance of order 2 between P_1 and P_2. Let μ_n be the distribution of $\boldsymbol{X}^{\otimes k}$ when $\boldsymbol{X} \sim \mathrm{Unif}(\mathbb{S}^{n-1})$ and let ν_n be the distribution of $\boldsymbol{X}^{\otimes k}$ when $\boldsymbol{X} \sim \mathcal{N}(0, \frac{1}{n}\mathrm{Id}_n)$. Let us compute a bound on $W_2(\mu_n, \nu_n)$. Let X be drawn uniformly over $\mathbb{S}^{n-1}$, and $\boldsymbol{G} \sim \mathcal{N}(0, \mathrm{Id})$, independently from $\boldsymbol{X}$. Then $(\boldsymbol{X}^{\otimes k}, (\|\boldsymbol{G}\|\boldsymbol{X}/\sqrt{n})^{\otimes k})$ is a coupling of μ_n and ν_n, so that, by definition of W_2,

$$W_2(\mu_n, \nu_n)^2 \le \mathbb{E}\Big\|\boldsymbol{X}^{\otimes k} - \big(\boldsymbol{X}\|\boldsymbol{G}\|/\sqrt{n}\big)^{\otimes k}\Big\|^2 = \mathbb{E}\Big[\Big(\Big(\frac{1}{n}\sum_{i=1}^n G_i^2\Big)^{k/2} - 1\Big)^2\Big]$$

where we use that $\mathbb{E}\|\boldsymbol{X}\|^k = 1$. By the law of large numbers, it then follows that

$$\lim_{n\to\infty} |F_{\mu_n}(\gamma) - F_{\nu_n}(\gamma)| \to 0.$$

Recall the definition (6.4.12) of $\phi_\lambda(q)$ and define $L(\gamma) = \frac{1}{2}\max_{q\in[0,1]}\phi_{\sqrt{\gamma}}(q) = \frac{1}{2}\max_{q\in[0,1)} f_{\sqrt{\gamma}}(q)$, where the equality comes from Lemma 6.4.7. By Theorem 6.1.1 we have $\lambda \geq 0$, $F_{\nu_n}(\gamma) \to L(\gamma)$ as $n \to \infty$, which implies $F_{\mu_n}(\gamma) \to L(\gamma)$. The "I-MMSE relation" (Proposition 1.3.3) gives that $\gamma \mapsto F_{\mu_n}(\gamma)$ is convex and differentiable over $[0, +\infty)$ and

$$F'_{\mu_n}(\lambda^2) = \frac{1}{2}\left(1 - \mathbb{E}\Big[\Big\|\boldsymbol{X}^{\otimes k} - \mathbb{E}\big[\boldsymbol{X}^{\otimes k}\big|\boldsymbol{Y}\big]\Big\|^2\Big]\right).$$

By Proposition C.1, F'_{μ_n} converges to L' for each $\lambda > 0$ at which L is differentiable. For $\gamma < \lambda_c^2$, $L(\gamma) = 0$, so L is differentiable on $(0, \lambda_c^2)$ with derivative equal to 0. For $\gamma > \lambda_c^2$, we know by Lemma 6.4.7 that $f_{\sqrt{\gamma}}$ admits a unique maximizer $q_*(\sqrt{\gamma})$ on $[0, 1]$. Proposition D.2 gives that L is differentiable at γ with derivative

$$L'(\gamma) = \frac{1}{2}(\partial_\gamma f_{\sqrt{\gamma}})(q_*(\sqrt{\gamma})) = \frac{1}{2}q_*(\sqrt{\gamma})^k.$$

We conclude that

$$\lim_{n\to\infty}\frac{1}{2}\left(1 - \mathbb{E}\Big[\Big\|\boldsymbol{X}^{\otimes k} - \mathbb{E}[\boldsymbol{X}^{\otimes k}|\boldsymbol{Y}]\Big\|^2\Big]\right) = \lim_{n\to\infty} F'_{\mu_n}(\lambda^2) = \begin{cases} 0 & \text{if } \lambda < \lambda_c \\ \frac{1}{2}q_*(\lambda)^k & \text{if } \lambda > \lambda_c. \end{cases}$$

□

Chapter 7

Phase transitions in Generalized Linear Models

7.1 Introduction: learning a linear classifier

Before presenting the Generalized Linear Models (GLM) in full generalities, we first focus on a simple supervised learning task: learning a linear classification rule. We assume here to have m points $\boldsymbol{g}^1, \dots, \boldsymbol{g}^m \in \mathbb{R}^n$, that are classified between two classes "+1" and "−1", according to the value of

$$Y_\mu = \mathrm{sign}\Big(\langle \boldsymbol{X}^*, \boldsymbol{g}^\mu \rangle\Big),$$

where $\boldsymbol{X}^* \in \mathbb{R}^n$ is a normal vector to the hyperplane that separates the two classes. The supervised learning task is, given the labeled dataset $\Big\{(\boldsymbol{g}^\mu, Y_\mu) \,\Big|\, 1 \leq \mu \leq m\Big\}$, to learn the hyperplane $\mathrm{Span}(\boldsymbol{X}^*)^\perp$ that separates the two classes.

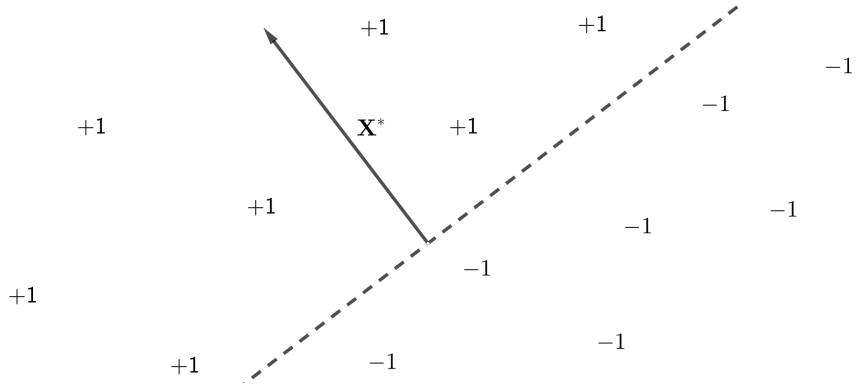


Figure 7.1: Two classes separated by an hyperplane.

7.1.1 Rosenblatt's perceptron algorithm

Rosenblatt [184, 185] proposed an iterative algorithm to find a vector $\widehat{\boldsymbol{x}}$ that separates the data, i.e. such that $\mathrm{sign}(\langle \widehat{\boldsymbol{x}}, \boldsymbol{g}^\mu \rangle) = Y_\mu$ for all μ. The so-called "perceptron" algorithm goes as follows:

- Initialize $\widehat{\boldsymbol{x}}^0 = 0$.

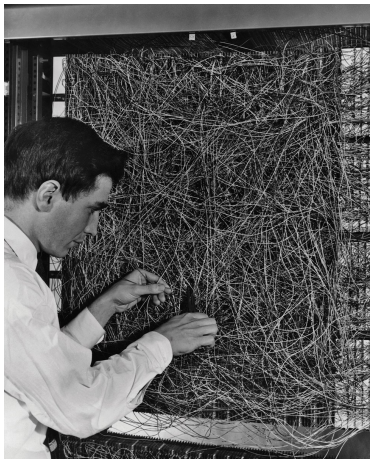



Figure 7.2: Frank Rosenblatt with the "Mark 1 perceptron", a custom-built hardware that implements the perceptron algorithm (circa 1960).

- While $I_t \stackrel{\text{def}}{=} \{\mu \,|\, \text{sign}(\langle \widehat{\boldsymbol{x}}^t, \boldsymbol{g}^\mu \rangle) \neq Y_\mu\} \neq \emptyset$ choose μ_t uniformly at random in I_t and update $\widehat{\boldsymbol{x}}^t$:
$$\widehat{\boldsymbol{x}}^{t+1} = \widehat{\boldsymbol{x}}^t + \text{sign}(Y_{\mu_t})\boldsymbol{g}^{\mu_t}. \tag{7.1.1}$$
- If $I_t = \emptyset$, then return $\widehat{\boldsymbol{x}} = \sqrt{n}\,\widehat{\boldsymbol{x}}^t/\|\widehat{\boldsymbol{x}}^t\|$.

Rosenblatt [185] and then Novikoff [165] proved that the perceptron algorithm terminates after a finite number of iterations:

Proposition 7.1.1

Let $\beta = \max_\mu \|\boldsymbol{g}^\mu\|^2$ and $\gamma = \min_\mu |\langle \boldsymbol{X}^*, \boldsymbol{g}^\mu \rangle|$. Then the perceptron algorithm converges for $t \leq t_{\max}$ where
$$t_{\max} \stackrel{\text{def}}{=} \frac{\beta \|\boldsymbol{X}^*\|^2}{\gamma^2}.$$

Proof. Let t_f be the time at which the algorithm converges, i.e. the first t for which $Y_\mu = \text{sign}(\langle \widehat{\boldsymbol{x}}^t, \boldsymbol{g}^\mu \rangle)$ for all μ. If such a t does not exist, we define $t_f = +\infty$. We have to prove that $t_f \leq t_{\max}$. From the update rule (7.1.1) we have for all $t < t_f$:
$$\langle \widehat{\boldsymbol{x}}^{t+1}, \boldsymbol{X}^* \rangle = \langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^* \rangle + \text{sign}(Y_{\mu_t}) \langle \boldsymbol{g}^{\mu_t}, \boldsymbol{X}^* \rangle \geq \langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^* \rangle + \gamma,$$
so that $\langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^* \rangle \geq t\gamma$ for all $t \leq t_f$. On the other hand, one has
$$\|\widehat{\boldsymbol{x}}^{t+1}\|^2 = \|\widehat{\boldsymbol{x}}^t\|^2 + \|\boldsymbol{g}^{\mu_t}\|^2 + 2\text{sign}(Y_{\mu_t}) \langle \boldsymbol{g}^{\mu_t}, \widehat{\boldsymbol{x}}^t \rangle \leq \|\widehat{\boldsymbol{x}}^t\|^2 + \beta,$$
by definition of β and because $\text{sign}(\langle \boldsymbol{g}^{\mu_t}, \widehat{\boldsymbol{x}}^t \rangle) \neq Y_{\mu_t}$. We get $\|\widehat{\boldsymbol{x}}^t\|^2 \leq t\beta$, for all $t \leq t_f$. From what we have seen above and the Cauchy-Schwarz inequality, we have for all $t \leq t_f$:
$$t^2\gamma^2 \leq \langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^* \rangle^2 \leq \|\widehat{\boldsymbol{x}}^t\|^2 \|\boldsymbol{X}^*\|^2 \leq t\beta \|\boldsymbol{X}^*\|^2.$$
We conclude that $t_f \leq t_{\max}$. □

What is now a typical order of magnitude of $t_{\max}$? We will be interested in the asymptotic regime, where $n, m \to \infty$ such that $m/n \to \alpha > 0$. Let us now suppose that the data points are randomly chosen as $(\boldsymbol{g}^\mu)_{1 \leq \mu \leq m} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \text{Id}_n)$ and assume that $\|\boldsymbol{X}^*\| = \sqrt{n}$.

In that case, the random variables $\langle \boldsymbol{X}^*, \boldsymbol{g}^\mu \rangle$ are i.i.d. $\mathcal{N}(0,n)$ random variables, so there exists constants $c, C > 0$ such that with high probability $\gamma = \min_\mu |\langle \boldsymbol{X}^*, \boldsymbol{g}^\mu \rangle| \geq c\sqrt{n}/m$ and $\max_\mu \|\boldsymbol{g}^\mu\|^2 \leq Cn$. Consequently, with have with high probability

$$t_{\max} = O(n^3).$$

This means that it is possible in this model to find a separating hyperplane in polynomial time.

7.1.2 The generalization problem

We have seen in the previous section an efficient algorithm to find a separating hyperplane of our data. In other words, the perceptron algorithm is able to achieve a "training error" of 0 on the dataset $(\boldsymbol{g}^\mu, Y_\mu)_{\mu \leq m}$. But does this algorithm generalize well? That is, given a new point $\boldsymbol{g}^{\text{new}} \sim \mathcal{N}(0, \mathrm{Id}_n)$ independent of $\boldsymbol{g}$, who has label $Y_{\text{new}} = \text{sign}(\langle \boldsymbol{X}^*, \boldsymbol{g}^{\text{new}} \rangle)$, what is the probability that

$$\text{sign}(\langle \widehat{\boldsymbol{x}}, \boldsymbol{g}^{\text{new}} \rangle) = Y_{\text{new}} \ ?$$

Let us suppose that $\|\boldsymbol{X}^*\| = \sqrt{n}$. Since we assumed $\boldsymbol{g}^{\text{new}}$ to be a standard Gaussian vector independent of $\boldsymbol{g}$ (and therefore also independent of $\widehat{\boldsymbol{x}}$), a simple computation gives

$$\mathbb{P}\Big(\text{sign}(\langle \widehat{\boldsymbol{x}}, \boldsymbol{g}^{\text{new}} \rangle) \neq Y_{\text{new}}\Big) = \frac{\arccos\left(\frac{1}{n}\langle \widehat{\boldsymbol{x}}, \boldsymbol{X}^* \rangle\right)}{\pi}.$$

Not surprisingly, the generalization performance of the perceptron algorithm depends on the correlation between $\widehat{\boldsymbol{x}}$ and the planted solution $\boldsymbol{X}^*$. From now we take $\boldsymbol{X}^* \sim \text{Unif}(\{-1,1\}^n)$, independently from $\boldsymbol{g}$. We plot on Figure 7.3 on page 114 the correlation $\frac{1}{n}\langle \widehat{\boldsymbol{x}}, \boldsymbol{X}^* \rangle$ a simulation with $n = 10000$ and various values of $\alpha = m/n$.

The question is now: how does this compare with the best achievable correlation? Is it possible to generalize perfectly? To study this problem, we have to look at the posterior distribution of $\boldsymbol{X}^*$ given $\boldsymbol{g}$ and $\boldsymbol{Y}$ which is simply the uniform distribution over

$$S_n \stackrel{\text{def}}{=} \left\{ \boldsymbol{x} \in \{-1,1\}^n \,\middle|\, \forall \mu \in \{1,\dots,m\},\ \text{sign}\Big(\langle \boldsymbol{x}, \boldsymbol{g}^\mu \rangle\Big) = Y_\mu \right\}.$$

Obviously $\boldsymbol{X}^* \in S_n$, and $\# S_n$ decreases as m increases. In order to be able to estimate $\boldsymbol{X}^*$ well, one would need that the points of S_n are well correlated with $\boldsymbol{X}^*$. We introduce therefore for $I \subset [-1,1]$,

$$S_n(I) \stackrel{\text{def}}{=} S_n \bigcap \left\{ \boldsymbol{x} \in \{-1,1\}^n \,\middle|\, \frac{1}{n}\langle \boldsymbol{x}, \boldsymbol{X}^* \rangle \in I \right\}.$$

The expected value of $\# S_n(I)$ (or more precisely $\frac{1}{n}\log \mathbb{E}\Big[\# S_n(I)\Big]$ since $\# S_n(I)$ is expected to be of the exponential order) can be computed easily:

Proposition 7.1.2

Let $h(q) = \frac{q+1}{2}\log(1+q) + \frac{1-q}{2}\log(1-q)$ and $\varphi(q) = \arccos(q)/\pi$. Define $f_\alpha(q) = \log(2) - h(q) + \alpha \log \varphi(q)$. Then, for any non-empty interval $I \subset [-1,1]$, we have (recall that $m/n \to \alpha$ as $n, m \to \infty$):

$$\frac{1}{n}\log \mathbb{E}\Big[\# S_n(I)\Big] \xrightarrow[n\to\infty]{} \sup_{q \in I} f_\alpha(q).$$

Proof. For $\boldsymbol{x} \in \{-1,1\}^n$, one writes $q(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^n x_i X_i$. We have, conditionally on $\boldsymbol{X}^*$,

$$\mathbb{E}\#S_n(I) = \sum_{\boldsymbol{x},\, q(\boldsymbol{x})\in I} \mathbb{E}\prod_{\mu=1}^m \mathbb{1}(\mathrm{sign}(\langle \boldsymbol{x}, \boldsymbol{g}^\mu\rangle) = \mathrm{sign}(\langle \boldsymbol{X}^*, \boldsymbol{g}^\mu\rangle)) = \sum_{\boldsymbol{x},\, q(\boldsymbol{x})\in I} \mathbb{P}(\mathrm{sign}(\langle \boldsymbol{x}, \boldsymbol{g}\rangle) = \mathrm{sign}(\langle \boldsymbol{X}^*, \boldsymbol{g}\rangle))^m,$$

because $(\boldsymbol{g}^\mu)_{\mu\geq 1} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathrm{Id}_n)$. An easy computation gives $\mathbb{P}(\mathrm{sign}(\langle \boldsymbol{x}, \boldsymbol{g}\rangle) = \mathrm{sign}(\langle \boldsymbol{X}^*, \boldsymbol{g}\rangle)) = \varphi(q(\boldsymbol{x}))$. For $k \in \mathbb{Z}$ we define $N_{n,k} = \#\{\boldsymbol{x} \in \{-1,1\}^n \,|\, q(\boldsymbol{x}) = k/n\}$. If $n+k$ is odd, then $N_{n,k} = 0$, otherwise $N_{n,k} = \binom{n}{(n+k)/2}$. By Sterling approximation, there exists a constant $c > 0$ such that

$$\frac{c}{\sqrt{n}} 2^n \exp\big(-nh(k/n)\big) \leq N_{n,k} \leq 2^n \exp\big(-nh(k/n)\big),$$

while the upper-bound follows from the standard Chernoff bound for a sum of i.i.d. Rademacher random variables. Since $\mathbb{E}\#S_n(I) = \sum_{k,\, k/n\in I} N_{n,k}\ \varphi(k/n)^m$, we get

$$2^n \max_{k, k/n\in I}\{e^{-nh(k/n)}\varphi(q)^m\} \leq \mathbb{E}\#S_n(I) \leq (2n+1)2^n \max_{q\in I}\{e^{-nh(q)}\varphi(q)^m\}$$

and the result follows from taking the logarithm and dividing by n. □

Let us define

$$\alpha_1 = \inf\left\{\alpha > 0 \,\middle|\, \sup_{q\in[-1,1]} f_\alpha(q) = 0\right\} \simeq 1.45. \tag{7.1.2}$$

Proposition 7.1.2 gives that for $\alpha > \alpha_1$, $\frac{1}{n}\log\mathbb{E}\big[\#S_n\big] \xrightarrow[n\to\infty]{} 0$, which seems to indicate that $\#S_n$ is of sub-exponential order in this regime: it should be possible to estimate $\boldsymbol{X}^*$ almost perfectly:

Corollary 7.1.1

Let $\boldsymbol{x} \sim \mathrm{Unif}(S_n)$, independently of everything else. Then we have for all $\alpha > \alpha_1$,

$$\frac{1}{n}\sum_{i=1}^n X_i x_i \xrightarrow[n\to\infty]{\text{(d)}} 1. \tag{7.1.3}$$

Proof. Let $\epsilon > 0$ and $I_\epsilon = [-1, 1-\epsilon]$. By Proposition 7.1.2 we have

$$\frac{1}{n}\log\mathbb{E}[\#S_n(I_\epsilon)] \xrightarrow[n\to\infty]{} \sup_{q\in I_\epsilon} f_\alpha(q) < 0,$$

because for $\alpha > \alpha_1$, the maximum of f_α is uniquely achieved at $q = 1$. Consequently

$$\mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^n X_i x_i \leq 1-\epsilon\Big) = \mathbb{E}\Big[\frac{\#S_n(I_\epsilon)}{\#S_n}\Big] \leq \mathbb{E}\big[\#S_n(I_\epsilon)\big] \xrightarrow[n\to\infty]{} 0.$$

□

Perfect reconstruction is therefore possible for $\alpha > \alpha_1$. But is α_1 the smallest value of α for which one can estimate $\boldsymbol{X}^*$ perfectly, i.e. the smallest value of α for which (7.1.3) holds? In the physics literature, the current problem was introduced by Gardner and Derrida [89] who predicted (based on numerical simulations) that it was possible to learn $\boldsymbol{X}^*$ perfectly for values of α larger than $\alpha_{\mathrm{IT}} = 1.35 \pm 0.10$. Then Györgyi [98] used the heuristic replica method to derive an exact formula for $\alpha_{\mathrm{IT}} \simeq 1.249$.

Unfortunately, Proposition 7.1.2 does not allow us to obtain the expression of α_{IT}. The issue is that the expected value of $\#S_n$ does not capture the typical order of magnitude

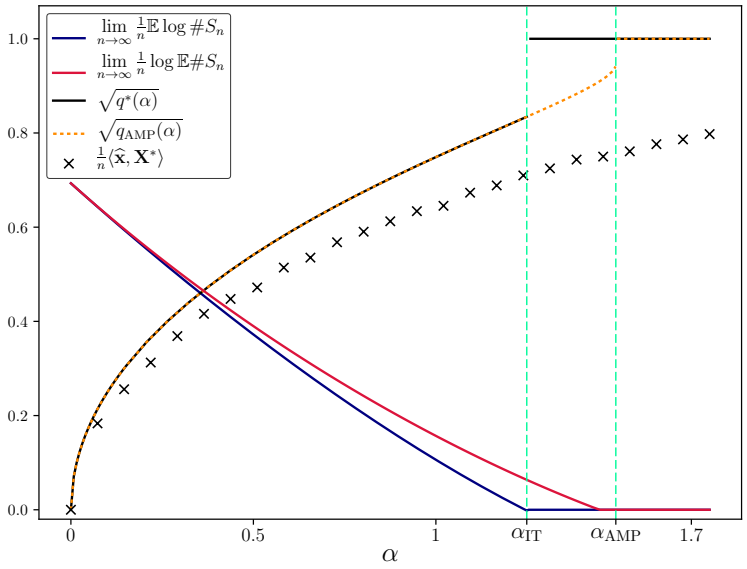


Figure 7.3: Plots of the limits of $\frac{1}{n}\log \mathbb{E}\#S_n$ (given by Proposition 7.1.2), $\frac{1}{n}\mathbb{E}\log \#S_n$ and $\sqrt{q^*(\alpha)}$ (from Proposition 7.1.3), together with the correlation $\sqrt{q_{\mathrm{AMP}}(\alpha)}$ achieved by GAMP and the correlation $\frac{1}{n}\langle \widehat{\boldsymbol{x}}, \boldsymbol{X}^*\rangle$ achieved by the perceptron algorithm (see Section 7.1.1) for $n = 10000$.

of $\#S_n$: the mean of $\#S_n$ is much larger than its median. This comes from exceptional events of extremely small probability on which $\#S_n$ is exceptionally large, which make its first moment too large.

The right quantity to consider is $\frac{1}{n}\mathbb{E}\Big[\log \#S_n\Big]$, which will not be parasitized by these exceptional events. This quantity is, however, harder to compute since the logarithm is now inside the expectation. Computing such quantities will be one of the main goals of this chapter: the result below will follow from an application of Theorems 7.3.1 and 7.3.2 presented in the next section.

Proposition 7.1.3

Let us define for $q \in [0,1]$

$$\mathfrak{f}_\alpha(q) = \log(2) + \inf_{r\geq 0}\left\{\mathbb{E}\Big[\log\cosh(\sqrt{r}Z + r)\Big] + 2\alpha\mathbb{E}\Big[\Phi\Big(\frac{\sqrt{q}\,Z}{\sqrt{1-q}}\Big)\log\Phi\Big(\frac{\sqrt{q}\,Z}{\sqrt{1-q}}\Big)\Big] - \frac{r(q+1)}{2}\right\},$$

where $Z \sim \mathcal{N}(0,1)$. Then we have

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\log \#S_n \xrightarrow[n\to\infty]{} \max_{q\in[0,1]}\mathfrak{f}_\alpha(q).$$

Further, for all values of α for which the maximum in q above is achieved at a unique point $q^*(\alpha)$ (which is the case for all α outside a countable set), we have

$$\frac{1}{n}\sum_{i=1}^{n} X_i x_i \xrightarrow[n\to\infty]{\mathbb{P}} q^*(\alpha). \tag{7.1.4}$$

Notice that (7.1.4) implies that

$$\sup_{\widehat{\boldsymbol{x}}}\mathbb{E}\left[\frac{\langle \widehat{\boldsymbol{x}}(\boldsymbol{Y},\boldsymbol{g}), \boldsymbol{X}^*\rangle}{\|\widehat{\boldsymbol{x}}(\boldsymbol{Y},\boldsymbol{g})\|\|\boldsymbol{X}^*\|}\right] \xrightarrow[n\to\infty]{} \sqrt{q^*(\alpha)},$$

where the supremum is taken over all measurable functions $\widehat{\boldsymbol{x}}$ of $\boldsymbol{Y}$ and $\boldsymbol{g}$. The critical value of α for which one can recover $\boldsymbol{X}^*$ perfectly is then

$$\alpha_{\mathrm{IT}} = \inf\left\{\alpha > 0 \,\middle|\, \max_{q\in[0,1]} \mathfrak{f}_\alpha(q) = 0\right\} = \inf\left\{\alpha > 0 \,\middle|\, q^*(\alpha) = 1\right\} \simeq 1.249.$$

However, as we see on Figure 7.3 on page 114 the perceptron algorithm of Section 7.1.1 does not achieve the optimal correlation $\sqrt{q^*(\alpha)}$. It is a priori unclear if this can be done by an efficient algorithm. We will see in Section 7.4 that there exists a polynomial-time algorithm called GAMP that achieve asymptotically a correlation of $\sqrt{q_{\mathrm{AMP}}(\alpha)}$, where $q_{\mathrm{AMP}}(\alpha)$ is the limit of the sequence $(q^t)_{t\geq 0}$ given by the recursion (7.4.4). A bold conjecture from statistical physics states that it is impossible to do better that GAMP in polynomial time. One sees on Figure 7.3 on page 114 that there is a gap between $\sqrt{q^*}$ and $\sqrt{q_{\mathrm{AMP}}}$ for $\alpha \in (\alpha_{\mathrm{IT}}, \alpha_{\mathrm{AMP}})$. In this regime it is theoretically possible to achieve a "perfect correlation" with the signal, but not with an efficient algorithm.

7.2 Generalized linear estimation: Problem statement

7.2.1 Definition

Let $n, m \in \mathbb{N}^*$. We define a Generalized Linear Model (GLM) as follows. Given a signal vector $\boldsymbol{X}^*$ in $\mathbb{R}^n$ and m "measurement vectors" $\boldsymbol{\Phi}_1, \dots \boldsymbol{\Phi}_m \in \mathbb{R}^n$ we sample $Y_1, \dots, Y_m$ independently (conditionally on $\boldsymbol{X}^*, \boldsymbol{\Phi}$) according to

$$Y_\mu \sim P_{\mathrm{out}}\Big(\,\cdot\,\Big|\frac{1}{\sqrt{n}}\langle \boldsymbol{\Phi}_\mu, \boldsymbol{X}^*\rangle\Big), \qquad 1 \leq \mu \leq m. \tag{7.2.1}$$

Here P_{out} is a transition kernel: for all $x \in \mathbb{R}$, $P_{\mathrm{out}}(\cdot|x)$ is a probability measure on $\mathbb{R}$ such that for all $A \in \mathcal{B}(\mathbb{R})$ the map $x \mapsto P_{\mathrm{out}}(A|x)$ is measurable. The definition (7.2.1) of a Generalized Linear Model is somehow a bit more general than the classical definition of GLMs in statistics. We will discuss this in Section 7.2.2 below.

Given the measurement vectors $(\boldsymbol{\Phi}_\mu)_{1\leq\mu\leq m}$ and the observations $(Y_\mu)_{1\leq\mu\leq m}$, there are two main statistical tasks:

(a) The *estimation* task. The goal is here to recover the signal $\boldsymbol{X}^*$.

(b) The *prediction* task. Given a new point $\boldsymbol{\Phi}_{\mathrm{new}}$ and the dataset $\boldsymbol{\Phi}, \boldsymbol{Y}$, we aim at predicting the corresponding output $Y_{\mathrm{new}} \sim P_{\mathrm{out}}(\cdot \,|\, \frac{1}{\sqrt{n}}\langle \boldsymbol{\Phi}_{\mathrm{new}}, \boldsymbol{X}^*\rangle)$.

Let $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ be the matrix whose rows are $\boldsymbol{\Phi}_1, \dots, \boldsymbol{\Phi}_m$. The GLM (7.2.1) encompasses many statistical models of significant interest, depending on the choice of the kernel P_{out}:

- The linear model: $\boldsymbol{Y} = \frac{1}{\sqrt{n}}\boldsymbol{\Phi}\boldsymbol{X}^* + \sigma\boldsymbol{Z}$, where $\sigma \geq 0$ and Z is some random noise.
- The Poisson regression: $P_{\mathrm{out}}(\cdot|x) = \mathrm{Pois}(e^x)$, where $\mathrm{Pois}(\theta)$ denotes the Poisson distribution of mean θ.
- The (real) phase-retrieval problem or "signless channel": $\boldsymbol{Y} = \frac{1}{\sqrt{n}}|\boldsymbol{\Phi}\boldsymbol{X}^*| + \sigma\boldsymbol{Z}$.

- The "1-bit compressed sensing", that we studied in Section 7.1: $\boldsymbol{Y} = \mathrm{sign}(\boldsymbol{\Phi}\boldsymbol{X}^*)$.
- The logistic model. This is the case when P_{out} is given by

$$P_{\mathrm{out}}(y = 1 \,|\, x) = 1 - P_{\mathrm{out}}(y = -1 \,|\, x) = \frac{1}{1 + e^{-\lambda x}}.$$

In the examples above, the functions absolute value $|\cdot|$ and sign act componentwise on the vector $\boldsymbol{\Phi}\boldsymbol{X}^*$.

7.2.2 Link with the classical definition of Generalized Linear Models in statistics

The definition of Generalized Linear Models in the statistical literature (see for instance [140]) differs slightly from the one given in the previous section. GLM are usually specified by:

- A bijective function $g : \mathbb{R} \to \mathbb{R}$, called "link function".
- A probability distribution P_θ parametrized by $\theta \in \mathbb{R}$ from the exponential family, i.e. whose admits a density f_θ with respect to a reference measure ν given by

$$f_\theta(y) = h(y)e^{\theta y - b(\theta)},$$

  for some functions h and b.

Then, given a "linear predictor" $\eta_\mu = \frac{1}{\sqrt{n}}\langle \boldsymbol{\Phi}_\mu, \boldsymbol{X}^* \rangle$, Y_μ is sampled as $Y_\mu \sim P_{\theta_\mu}$, where θ_μ is given by the relation

$$g^{-1}(\eta_\mu) = b'(\theta_\mu). \tag{7.2.2}$$

This choice of θ_μ ensures that the linear predictor η_μ is related to the mean of Y_μ by $\eta_\mu = g(\mathbb{E}[Y_\mu]) = g(b'(\theta_\mu))$. This definition of a Generalized Linear Model is obviously a particular case of the definition (7.2.1) since the kernel P_{out} can reproduce the generation process that we just described.

Let us now mention some limitations of this "classical" definition, compared to (7.2.1):

- The support of P_θ does not depend on θ (and therefore on $\langle \boldsymbol{\Phi}_\mu, \boldsymbol{X}^* \rangle$). Hence the classical definition does not encompass the clustering problem discussed in Section 7.1.
- The mean of Y_μ is in one-to-one correspondence with η_μ by (7.2.2), whereas with the definition (7.2.1) one could imagine that this mean depends differently on η_μ (think for instance to the "signless channel" where $\mathbb{E}[Y_\mu] = |\eta_\mu|$).

7.2.3 Bayesian framework

We will study the Generalized Linear Model (7.2.1) in a Bayesian framework, where the components $X_1^*, \dots, X_n^*$ of $\boldsymbol{X}^*$ are i.i.d. samples from a probability distribution P_0 over $\mathbb{R}$. We will also assume that the measurement matrix $\boldsymbol{\Phi}$ is independent of $\boldsymbol{X}^*$, with independent entries that have zero mean and unit variance.

We will only consider transition kernels P_{out} that admits a transition density with respect to Lebesgue's measure or the counting measure on $\mathbb{N}$. We will (with a slight abuse

of notation) also use the notation $P_{\text{out}}(\cdot|x)$ to denote this transition density. The posterior distribution of $\boldsymbol{X}^*$ given $\boldsymbol{Y}, \boldsymbol{\Phi}$ takes the form:

$$dP(\boldsymbol{x}|\boldsymbol{Y}, \boldsymbol{\Phi}) = \frac{1}{\mathcal{Z}(\boldsymbol{Y}, \boldsymbol{\Phi})} dP_0^{\otimes n}(\boldsymbol{x}) \prod_{\mu=1}^{m} P_{\text{out}}\Big(Y_\mu \Big| \frac{1}{\sqrt{n}} [\boldsymbol{\Phi}\boldsymbol{x}]_\mu\Big) \tag{7.2.3}$$

$$= \frac{1}{\mathcal{Z}(\boldsymbol{Y}, \boldsymbol{\Phi})} dP_0^{\otimes n}(\boldsymbol{x}) e^{-H(\boldsymbol{x}; \boldsymbol{Y}, \boldsymbol{\Phi})} \tag{7.2.4}$$

where the *Hamiltonian* is defined as

$$H(\boldsymbol{x}; \boldsymbol{Y}, \boldsymbol{\Phi}) \stackrel{\text{def}}{=} -\sum_{\mu=1}^{m} \log P_{\text{out}}\Big(Y_\mu \Big| \frac{1}{\sqrt{n}} [\boldsymbol{\Phi}\boldsymbol{x}]_\mu\Big) \tag{7.2.5}$$

and the *partition function* (the normalization factor) is defined as

$$\mathcal{Z}(\boldsymbol{Y}, \boldsymbol{\Phi}) \stackrel{\text{def}}{=} \int dP_0^{\otimes n}(\boldsymbol{x}) e^{-H(\boldsymbol{x}; \boldsymbol{Y}, \boldsymbol{\Phi})}\,. \tag{7.2.6}$$

The main quantity of interest here is the associated free energy:

$$F_n \stackrel{\text{def}}{=} \frac{1}{n} \mathbb{E} \log \mathcal{Z}(\boldsymbol{Y}, \boldsymbol{\Phi}). \tag{7.2.7}$$

We will compute in Section 7.3.2 the limit of F_n when $n, m \to \infty$ while $m/n \to \alpha > 0$.

Example 7.2.1. *In the case of the "planted perceptron" studied in Section 7.1, we had* $P_{\text{out}}(\cdot|x) = \delta_{\text{sign}(x)}$ *and* $P_0 = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$. *We see that*

$$\mathcal{Z}(\boldsymbol{Y}, \boldsymbol{\Phi}) = \frac{1}{2^n} \#\Big\{\boldsymbol{x} \in \{-1, 1\}^n \,\Big|\, \forall \mu \in \{1, \dots, m\},\ \text{sign}\Big(\langle \boldsymbol{x}, \boldsymbol{\Phi}_\mu \rangle\Big) = Y_\mu\Big\}$$

measures the number of possible values for $\boldsymbol{X}^*$ *given the observations of* $\boldsymbol{Y}, \boldsymbol{\Phi}$.

Random function representation. We introduce now a convenient "random function" representation for the transition kernel P_{out}. Let us consider a function $\varphi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and a probability distribution P_A over $\mathbb{R}$ such that for all $x \in \mathbb{R}$, $P_{\text{out}}(\cdot|x)$ corresponds to the law of

$$Y = \varphi(x, A) + \sqrt{\Delta} Z \tag{7.2.8}$$

where $(A, Z) \sim P_A \otimes \mathcal{N}(0, 1)$ and $\Delta \geq 0$. Notice that any transition kernel can be realized by (7.2.8) by simply taking $\Delta = 0$, $A \sim \text{Unif}([0, 1])$ and $\varphi(x, \cdot)$ to be the generalized inverse of the cumulative distribution function of $P_{\text{out}}(\cdot|x)$, see Proposition 7.2.1 below.

We will however sometimes need to take $\Delta > 0$ in order to "regularize" P_{out}. In that case the transition kernel P_{out} admits a transition density with respect to Lebesgue's measure, given by

$$P_{\text{out}}(y|x) = \frac{1}{\sqrt{2\pi\Delta}} \int dP_A(a) e^{-\frac{1}{2\Delta}(y - \varphi(x,a))^2}\,. \tag{7.2.9}$$

When $\Delta = 0$, we will only consider discrete channels where φ takes values in $\mathbb{N}$[1]. In that case P_{out} admits a transition density with respect the counting measure on $\mathbb{N}$, that we will

[1]Notice that this allows to study any channel whose outputs belong to a countable set S by applying a injection $u : S \to \mathbb{N}$ to the outputs.

write also $P_{\text{out}}(\cdot|x)$.

Let us end this section by looking at the link between the continuity properties of P_{out} and φ. Clearly, if (7.2.8) holds and if for almost all a (with respect to P_A), the function $\varphi(\cdot, a)$ is continuous at some $x_0 \in \mathbb{R}$, then $x \mapsto P_{\text{out}}(\cdot|x)$ is continuous (for the weak convergence) at x_0. The next proposition states that there exists functions φ that verify (7.2.8) and such that the converse is true.

Proposition 7.2.1

There exists a measurable function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ such that for all $x \in \mathbb{R}$, $P_{\text{out}}(\cdot|x)$ is the law of $\varphi(x, A)$ for $A \sim \text{Unif}([0,1])$. Moreover if $x \mapsto P_{\text{out}}(\cdot|x)$ is continuous at x_0 for the weak convergence, i.e.

$$P_{\text{out}}(\cdot|x) \xrightarrow[x \to x_0]{} P_{\text{out}}(\cdot|x_0).$$

then for almost all $a \in [0,1]$ the function $\varphi(\cdot, a)$ is also continuous at x_0.

Proof. We define the cumulative distribution function of $P_{\text{out}}(\cdot|x)$:

$$F(y|x) = \int_{-\infty}^{y} P_{\text{out}}(dt|x),$$

and its generalized inverse:

$$\varphi(x, a) \stackrel{\text{def}}{=} \inf\Big\{ y \in \mathbb{R} \,\Big|\, F(y|x) \geq a \Big\}.$$

For $A \sim \text{Unif}([0,1])$, we know that $\varphi(x, A) \stackrel{(\text{d})}{=} P_{\text{out}}(\cdot|x)$. Let now $x_0 \in \mathbb{R}$ be a point at which $x \mapsto P_{\text{out}}(\cdot|x)$ is continuous. The set $S_{x_0} \stackrel{\text{def}}{=} \{a \in \mathbb{R} \,|\, \exists y_1 \neq y_2,\ a = F(y_1|x_0) = F(y_2|x_0)\}$ is countable. Indeed if $a \in S_{x_0}$ then we can find $q \in \mathbb{Q}$ such that $a = F(q|x_0)$. If we define $\psi(a) = q$ then $\psi : S_{x_0} \to \mathbb{Q}$ is injective hence S_{x_0} is countable.

Let $a \in \mathbb{R} \setminus S_{x_0}$. Since $P_{\text{out}}(\cdot|x) \xrightarrow[x \to x_0]{} P_{\text{out}}(\cdot|x_0)$, we have $F(y|x) \to F(y|x_0)$ for almost all $y \in \mathbb{R}$. Let $y_1 < \varphi(x_0, a) < y_2$ such that $F(y_i|x) \to F(y_i|x_0)$ for $i = 1, 2$. Recall that $a \notin S_{x_0}$ so we have

$$F(y_1|x_0) < a = F(\varphi(x_0, a)|x_0) < F(y_2|x_0).$$

Consequently for x close enough from x_0 we have $F(y_1|x) < a < F(y_2|x)$ which implies that $y_1 < \varphi(x, a) < y_2$. Since y_1 and y_2 can be chosen arbitrarily close to $\varphi(x_0, a)$ we conclude: $\varphi(x, a) \xrightarrow[x \to x_0]{} \varphi(x_0, a)$. □

7.3 Information-theoretic limits

7.3.1 Two scalar inference channels

As for the "spiked models" of the previous chapters, because of the decoupling principle seen in Chapter 2, the limit of the free energy (7.2.7) will be expressed in terms of free energies of simple scalar channels.

The first one is the additive Gaussian channel (1.3.5) that we already studied in Section 1.3:

$$Y_0 = \sqrt{r}\, X_0 + Z_0\,, \tag{7.3.1}$$

where $X_0 \sim P_0$ and $Z_0 \sim \mathcal{N}(0,1)$ are independent and $r \geq 0$. The inference problem consists of retrieving X_0 from the observations Y_0. As we have seen in Section 1.3, the

free energy $\psi_{P_0}(r)$ (defined by (1.3.6)) associated to this problem is related to the mutual information between the signal and observations by:

$$I_{P_0}(r) \stackrel{\text{def}}{=} I(X_0; \sqrt{r}\, X_0 + Z_0) = \frac{r\rho}{2} - \psi_{P_0}(r). \tag{7.3.2}$$

The second inference channel is a non-linear channel, associated to the transition kernel P_{out}. Suppose that $V, W^* \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ where V is known and the inference problem is to recover the unknown W^* from the observation of

$$\widetilde{Y}_0 \sim P_{\text{out}}\Big(\cdot \,|\, \sqrt{q}\, V + \sqrt{\rho - q}\, W^* \Big), \tag{7.3.3}$$

where $\rho > 0$, $q \in [0, \rho]$. Notice that under the representation (7.2.8), the channel (7.3.3) is equivalent to $\widetilde{Y}_0 = \varphi(\sqrt{q}\, V + \sqrt{\rho - q}\, W^*, A) + \sqrt{\Delta} Z$ with $\Delta \geq 0$ and where $(A, Z) \sim P_A \otimes \mathcal{N}(0,1)$, independently of V, W^*. The free energy for this model is again related to the normalization of the posterior $P(W^* | \widetilde{Y}_0, V)$

$$\Psi_{P_{\text{out}}}(q) \stackrel{\text{def}}{=} \mathbb{E} \log \int P_{\text{out}}\Big(\widetilde{Y}_0 | \sqrt{q}\, V + \sqrt{\rho - q}\, w\Big) \frac{e^{-w^2/2}}{\sqrt{2\pi}} dw, \tag{7.3.4}$$

where P_{out} denotes either the transition density with respect to Lebesgue's measure (given by (7.2.9)) in the case where $\Delta > 0$, or the density with respect to the counting measure over $\mathbb{N}$, in the case of discrete observations (φ takes values in $\mathbb{N}$ and $\Delta = 0$). The free energy (7.3.4) is again linked to the mutual information between the observation $\widetilde{Y}_0$ and the signal W^*, given V:

$$\mathcal{I}_{P_{\text{out}}}(q) \stackrel{\text{def}}{=} I(W^*; \widetilde{Y}_0 | V) = \Psi_{P_{\text{out}}}(\rho) - \Psi_{P_{\text{out}}}(q). \tag{7.3.5}$$

We will prove in Section 7.8 that $\Psi_{P_{\text{out}}}$ is convex, differentiable and non-decreasing.

7.3.2 Replica-symmetric formula and mutual information

We present in this section a limiting formula for F_n. The result holds under the following hypotheses.

- (h1) The prior distribution P_0 admits a finite third moment and has at least two points in its support.
- (h2) There exists $\gamma > 0$ such that the sequence $(\mathbb{E}[|\varphi(\frac{1}{\sqrt{n}}[\boldsymbol{\Phi} \boldsymbol{X}^*]_1, A_1)|^{2+\gamma}])_{n \geq 1}$ is bounded.
- (h3) The random variables $(\Phi_{\mu i})$ are independent with zero mean, unit variance and finite third moment that is bounded with n.
- (h4) For almost-all values of $a \in \mathbb{R}$ (w.r.t. P_A), the function $x \mapsto \varphi(x, a)$ is continuous almost everywhere.

We will also assume that one of the two following hypotheses hold:

- (h5.a) $\Delta > 0$.
- (h5.b) $\Delta = 0$ and φ takes values in $\mathbb{N}$.

Remark 7.3.1. *The hypotheses are here stated using the "random function" representation of (7.2.8). In many cases, it can be useful to state them using the transition kernel representation of (7.2.1). The hypotheses (h2) and (h4) are respectively equivalent*[2] *to:*

[2] The implications (h2) $\Leftrightarrow$ (h2') and (h4) $\Rightarrow$ (h4') are easily verified. If (h4') holds, then by Proposition 7.2.1 there exists a function $\varphi : \mathbb{R} \times [0,1] \to \mathbb{R}$ such that (7.2.8) holds for $A_\mu \overset{\text{i.i.d.}}{\sim} P_A = \text{Unif}([0,1])$ and that (h4) is verified.

(h2') *There exists* $\gamma > 0$ *such that* $\mathbb{E}[|Y_1|^{2+\gamma}]$ *remains bounded with* n.
(h4') $x \in \mathbb{R} \mapsto P_{\text{out}}(\cdot|x)$ *is continuous almost everywhere for the weak convergence.*

Under the above hypothesis (h5.a) (respectively (h5.b)), the transition kernel P_{out} admits a density with respect to Lebesgue's measure on $\mathbb{R}$ (resp. the counting measure on $\mathbb{N}$) that will be denoted by $P_{\text{out}}(\cdot|x)$.

Definition 7.3.1

We will say that the kernel P_{out} *is*

- *non-informative if for almost all* $y \in \mathbb{R}$ *(under (h5.a)) – or all* $y \in \mathbb{N}$ *(under (h5.b))– the map* $x \mapsto P_{\text{out}}(y\,|\,x)$ *is almost everywhere equal to a constant. We say that* P_{out} *is informative if it is not non-informative.*
- *symmetric if* $P_{\text{out}}(\cdot|x) = P_{\text{out}}(\cdot|-x)$ *for almost all* $x \in \mathbb{R}$.

If P_{out} is non-informative, it is not difficult to show that estimation is impossible.

Let us define the following *potential* $\mathcal{F}$. Call $\rho \stackrel{\text{def}}{=} \mathbb{E}[X^2]$ where $X \sim P_0$. Then for $(q,r) \in [0,\rho] \times \mathbb{R}_{\geq 0}$ we define

$$\mathcal{F}(q,r) \stackrel{\text{def}}{=} \psi_{P_0}(r) + \alpha \Psi_{P_{\text{out}}}(q) - \frac{rq}{2}\,. \tag{7.3.6}$$

We will also write $\mathcal{F}(\rho,+\infty) = \lim_{r\to\infty} \mathcal{F}(\rho,r)$, which is well-defined in $\mathbb{R} \cup \{-\infty\}$ since $r \mapsto \psi_{P_0}(r) - r\frac{\rho}{2}$ is non-increasing by Proposition 1.3.3. We need also to define the set of the critical points of $\mathcal{F}$:

$$\Gamma \stackrel{\text{def}}{=} \left\{ (q,r) \in [0,\rho] \times (\mathbb{R}_{\geq 0} \cup \{+\infty\}) \;\middle|\; \begin{array}{rcl} q &=& 2\psi'_{P_0}(r) \\ r &=& 2\alpha \Psi'_{P_{\text{out}}}(q;\rho) \end{array} \right\}, \tag{7.3.7}$$

where, with a slight abuse of notation, we extend ψ'_{P_0} and $\Psi'_{P_{\text{out}}}$ by their limits: $\psi'_{P_0}(+\infty) = \lim_{r\to\infty} \psi'_{P_0}(r) = \rho/2$ and $\Psi'_{P_{\text{out}}}(\rho) = \lim_{q\to\rho} \Psi'_{P_{\text{out}}}(q)$. This last limit is well defined in $\mathbb{R} \cup \{+\infty\}$ by convexity of $\Psi_{P_{\text{out}}}$. The elements of Γ are called "fixed points of the state evolution", because – as we shall see in Section 7.4 – they are related to the fixed points of some generalized approximate message passing algorithm.

Theorem 7.3.1

Suppose that hypotheses (h1)-(h2)-(h3)-(h4) hold. Suppose that either hypothesis (h5.a) or (h5.b) holds. Then

$$\lim_{n\to\infty} F_n = \sup_{q\in[0,\rho]} \inf_{r\geq 0} \mathcal{F}(q,r) = \sup_{(q,r)\in\Gamma} \mathcal{F}(q,r)\,. \tag{7.3.8}$$

Moreover, if P_{out} *is informative, then the "*sup inf*" and the supremum over* Γ *in (7.3.8) are achieved over the same couples* (q,r).

An immediate corollary of Theorem 7.3.1 is the limiting expression of the mutual information between the signal and the observations.

Corollary 7.3.1

$$i_\infty \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} I(\boldsymbol{X}^*; \boldsymbol{Y} \,|\boldsymbol{\Phi}) = \inf_{q\in[0,\rho]} \sup_{r\geq 0} i_{\text{RS}}(q,r) = \inf_{(q,r)\in\Gamma} i_{\text{RS}}(q,r)\,, \tag{7.3.9}$$

where

$$i_{\text{RS}}(q,r) \stackrel{\text{def}}{=} I_{P_0}(r) + \alpha \mathcal{I}_{P_{\text{out}}}(q) - \frac{r}{2}(\rho - q)\,. \tag{7.3.10}$$

Proof. This follows from a simple calculation:

$$\begin{aligned}\frac{1}{n} I(\boldsymbol{X}^*; \boldsymbol{Y}|\boldsymbol{\Phi}) &= \frac{1}{n} H(\boldsymbol{Y}|\boldsymbol{\Phi}) - \frac{1}{n} H(\boldsymbol{Y}|\boldsymbol{X}^*, \boldsymbol{\Phi}) = -F_n + \frac{1}{n}\mathbb{E}\log P(\boldsymbol{Y}|\boldsymbol{X}^*, \boldsymbol{\Phi}) \\ &= -F_n + \frac{m}{n}\mathbb{E}\log P_{\text{out}}(Y_1 \,|\, \langle \boldsymbol{\Phi}_1, \boldsymbol{X}^*\rangle/\sqrt{n})\,. \end{aligned} \tag{7.3.11}$$

By the central limit theorem (that we can apply under hypotheses (h1)-(h3)) we have

$$S_n \stackrel{\text{def}}{=} \frac{1}{\sqrt{n}}\langle \boldsymbol{\Phi}_1, \boldsymbol{X}^*\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Phi_{1,i} X_i \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0,\rho)\,.$$

Now, under the hypotheses (h2)-(h4) and either (h5.a) or (h5.b) it is not difficult to verify that

$$\begin{aligned}\mathbb{E}\log P_{\text{out}}(Y_1 \,|\, \langle \boldsymbol{\Phi}_1, \boldsymbol{X}^*\rangle/\sqrt{n}) &= \mathbb{E}\int P_{\text{out}}(dy|S_n)\log P_{\text{out}}(y|S_n) \\ &\xrightarrow[n\to\infty]{} \mathbb{E}\int P_{\text{out}}(dy|\sqrt{\rho}\,V)\log P_{\text{out}}(y|\sqrt{\rho}\,V) = \Psi_{P_{\text{out}}}(\rho)\end{aligned}$$

where $V \sim \mathcal{N}(0,1)$. We conclude, using (7.3.11):

$$\frac{1}{n} I(\boldsymbol{X}^*; \boldsymbol{Y}|\boldsymbol{\Phi}) = -F_n + \alpha \Psi_{P_{\text{out}}}(\rho) + o_n(1). \tag{7.3.12}$$

□

7.3.3 Optimal errors

We compute in this section the optimal errors for both estimation and prediction tasks. Both will be determined by the value of $q \geq 0$ that optimizes (7.3.8) (or equivalently in (7.3.9)).

Proposition 7.3.1

Define

$$D^* \stackrel{\text{def}}{=} \Big\{\alpha > 0 \,\Big|\, \text{(7.3.8) (or equivalently (7.3.9)) admits a unique optimizer } q^*(\alpha)\Big\}\,. \tag{7.3.13}$$

Assume that the assumptions of Theorem 7.3.1 hold and that P_{out} is informative. Then the set D^* is equal to $\mathbb{R}_{>0}$ minus some countable set. Moreover the map $\alpha \mapsto q^*(\alpha)$ is continuous on D^*.

Optimal reconstruction (or estimation) error

We first consider the problem of estimating $\boldsymbol{X}^*$ given $\boldsymbol{Y}$ and $\boldsymbol{\Phi}$. The following theorem states that the optimizer $q^*(\alpha)$ of the replica-symmetric formula (7.3.8) gives the asymptotic correlation between the planted solution $\boldsymbol{X}^*$ and a sample from the posterior distribution $P(\cdot \,|\, \boldsymbol{Y}, \boldsymbol{\Phi})$:

Theorem 7.3.2

Assume that all the moments of P_0 are finite and that P_{out} is informative. Assume that (h1)-(h2)-(h3)-(h4) hold and that either (h5.a) or (h5.b) holds. Then for all $\alpha \in D^*$,

$$\frac{1}{n}\Big|\sum_{i=1}^{n} x_i X_i^*\Big| \xrightarrow[n\to\infty]{\text{(d)}} q^*(\alpha), \tag{7.3.14}$$

where $\boldsymbol{x} = (x_1, \dots, x_n)$ is sampled from the posterior distribution of the signal $P(\cdot \,|\, \boldsymbol{Y}, \boldsymbol{\Phi})$ given by (7.2.3), independently of everything else. Moreover, if P_{out} is not symmetric (see Definition 7.3.1) then:

$$\frac{1}{n}\sum_{i=1}^{n} x_i X_i^* \xrightarrow[n\to\infty]{\text{(d)}} q^*(\alpha). \tag{7.3.15}$$

By Proposition 1.2.2 and Remark 1.2.1 in Section 1.2 we deduce from Theorem 7.3.2:

Corollary 7.3.2

Under the conditions of Theorem 7.3.2 we have for all $\alpha \in D^*$

$$\sup_{\widehat{\boldsymbol{x}}} \mathbb{E}\left[\frac{|\langle \widehat{\boldsymbol{x}}(\boldsymbol{Y}, \boldsymbol{\Phi}), \boldsymbol{X}^*\rangle|}{\|\widehat{\boldsymbol{x}}(\boldsymbol{Y}, \boldsymbol{\Phi})\| \|\boldsymbol{X}^*\|}\right] \xrightarrow[n\to\infty]{} \sqrt{\frac{q^*(\alpha)}{\rho}}, \tag{7.3.16}$$

where the supremum is taken over all measurable functions $\widehat{\boldsymbol{x}} : \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}^n \setminus \{0\}$.

Consequently $\sqrt{q^*(\alpha)/\rho}$ is the best correlation with the signal $\boldsymbol{X}^*$ that one can asymptotically achieve.

Optimal generalization (or prediction) error

We now consider the prediction problem: given the dataset $\boldsymbol{\Phi}, \boldsymbol{Y}$ and a new point $\boldsymbol{\Phi}_{\text{new}}$, we would like to estimate $Y_{\text{new}} \sim P_{\text{out}}(\cdot|\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle/\sqrt{n})$. We define the *generalization error* of an estimator $\widehat{Y}(\boldsymbol{\Phi}_{\text{new}}, \boldsymbol{\Phi}, \boldsymbol{Y})$ by

$$\mathcal{E}_{\text{gen}}(\widehat{Y}) \stackrel{\text{def}}{=} \mathbb{E}\Big[\Big(Y_{\text{new}} - \widehat{Y}(\boldsymbol{\Phi}_{\text{new}}, \boldsymbol{\Phi}, \boldsymbol{Y})\Big)^2\Big]. \tag{7.3.17}$$

The optimal generalization error is then defined as the minimum of $\mathcal{E}_{\text{gen}}$ over all estimators $\widehat{Y}$:

$$\mathcal{E}_{\text{gen}}^{\text{opt}} \stackrel{\text{def}}{=} \min_{\widehat{Y}} \mathcal{E}_{\text{gen}}(\widehat{Y}) = \mathbb{E}\Big[\Big(Y_{\text{new}} - \mathbb{E}[Y_{\text{new}}|\boldsymbol{\Phi}_{\text{new}}, \boldsymbol{\Phi}, \boldsymbol{Y}]\Big)^2\Big]. \tag{7.3.18}$$

In order to express the optimal generalization error we introduce the following function (recall that $\widetilde{Y}_0, V, W^*$ come from the channel (7.3.3)):

$$\mathcal{E}(q) \stackrel{\text{def}}{=} \text{MMSE}(\widetilde{Y}_0|V) = \mathbb{E}\Big[\Big(\widetilde{Y}_0 - \mathbb{E}[\widetilde{Y}_0|V]\Big)^2\Big] \tag{7.3.19}$$

$$= \mathbb{E}\int y^2 P_{\text{out}}(dy|\sqrt{\rho}\, V) - \mathbb{E}\Big[\mathbb{E}_{W^*}\Big[\int y P_{\text{out}}(dy|\sqrt{q}\, V + \sqrt{\rho - q}\, W^*)\Big]^2\Big] \tag{7.3.20}$$

where $\mathbb{E}_{W^*}$ denotes the expectation with respect to W^* only.

Theorem 7.3.3

Under the assumptions of Theorem 7.3.2, we have for all $\alpha \in D^*$

$$\mathcal{E}_{\text{gen}}^{\text{opt}}(\alpha) \xrightarrow[n\to\infty]{} \mathcal{E}(q^*(\alpha)) \tag{7.3.21}$$

where $q^*(\alpha)$ is the optimizer of (7.3.8), see Proposition 7.3.1.

Theorem 7.3.3 follows from a stronger result that we state below:

Theorem 7.3.4

Let $\alpha \in D^*$. Let ν_n be the law of Y_{new} conditionally on $\boldsymbol{\Phi}, \boldsymbol{Y}, \boldsymbol{\Phi}_{\text{new}}$, i.e.

$$\int f(y) d\nu_n(y) = \int f(y) P_{\text{out}}\Big(dy \,\Big|\, \langle \boldsymbol{x}, \boldsymbol{\Phi}_{\text{new}} \rangle / \sqrt{n}\Big) dP(\boldsymbol{x}|\boldsymbol{Y}, \boldsymbol{\Phi}),$$

for all continuous bounded function f. ν_n is therefore a random measure on $\mathbb{R}$. Let ν be the posterior distribution of $\widetilde{Y}_0$ given V in the second scalar channel (7.3.3), i.e.

$$\int f(y) d\nu(y) = \mathbb{E}_{W^*} \int f(y) P_{\text{out}}\Big(dy \,\Big|\, \sqrt{q^*(\alpha)} V + \sqrt{\rho - q^*(\alpha)} W^*\Big)$$

for all continuous bounded function f, where $\mathbb{E}_{W^*}$ denotes expectation with respect to $W^* \sim \mathcal{N}(0,1)$ only. Then, under the hypotheses of Theorem 7.3.2:

$$\nu_n \xrightarrow[n\to\infty]{\text{(wd)}} \nu,$$

where $\xrightarrow[n\to\infty]{\text{(wd)}}$ denotes convergence in distribution of the random variables $(\nu_n)_{n\geq 1}$ to the random variable ν (See [119, Chapter 4] for more details on this mode of convergence).

In other words, the data $\boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}_{\text{new}}$ contains asymptotically as much information on Y_{new} than V on $\widetilde{Y}_0 \sim P_{\text{out}}\Big(\cdot \Big| \sqrt{q^*(\alpha)} V + \sqrt{\rho - q^*(\alpha)} W^*\Big)$. Theorem 7.3.4 is proved in Section 7.7.3.

Theorem 7.3.4 implies Theorem 7.3.3 because:

$$\mathcal{E}_{\text{gen}}^{\text{opt}} = \mathbb{E}\Big[\text{Var}(Y_{\text{new}}|\boldsymbol{\Phi}_{\text{new}}, \boldsymbol{\Phi}, \boldsymbol{Y})\Big] = \mathbb{E}\Big[\text{Var}(\nu_n)\Big] \xrightarrow[n\to\infty]{} \mathbb{E}\Big[\text{Var}(\nu)\Big] = \mathcal{E}(q^*(\alpha)). \tag{7.3.22}$$

To justify the above limit, we need to proceed by truncation since $\mu \mapsto \text{Var}(\mu)$ is not a continuous bounded function over the set of probability measures. For a probability distribution μ on $\mathbb{R}$ and $X \sim \mu$ we define for $R > 0$ the probability measure μ^R as the law of $\text{sign}(X)\min(|X|, R)$, then $\mu \mapsto \text{Var}(\mu^R)$ is a continuous (with respect to the weak topology) bounded function. By Theorem 7.3.4 we get $\mathbb{E}\Big[\text{Var}(\nu_n^R)\Big] \xrightarrow[n\to\infty]{} \mathbb{E}\Big[\text{Var}(\nu^R)\Big]$ and (7.3.22) follows by letting $R \to \infty$, using hypothesis (h2).

Theorem 7.3.4 allows to compute more sophisticated optimal errors. Take for instance a classification problem with $K \in \mathbb{N}^*$ classes (such as the one we studied in Section 7.1) where φ takes values in $\{1, \dots, K\}$. For these kind of problems, a natural error metric to consider for an estimator $\widehat{Y}$ is the probability of misclassification:

$$\mathbb{P}\Big(Y_{\text{new}} \neq \widehat{Y}(\boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}_{\text{new}})\Big).$$

The best estimator for this error metric is known (see for instance [63], Section 2.1) to be the Bayes classifier:

$$\widehat{Y}^{\text{Bayes}} = \underset{k\in\{1,\dots,K\}}{\arg\max}\ \mathbb{P}(Y_{\text{new}} = k \,|\, \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}_{\text{new}}).$$

Notice that the corresponding probability of misclassification is equal to

$$\mathbb{P}\Big(Y_{\text{new}} \neq \widehat{Y}^{\text{Bayes}}\Big) = \mathbb{E}\Big[\sum_{k=1}^{K} \nu_n(\{k\})\mathbb{1}\Big(k \neq \underset{i\in\{1,\dots,K\}}{\arg\max}\, \nu_n(\{i\})\Big)\Big] = \mathbb{E}\Big[1 - \max_{1\leq k\leq K} \nu_n(\{k\})\Big].$$

By Theorem 7.3.4 we deduce:

$$\mathbb{P}\Big(Y_{\text{new}} \neq \widehat{Y}^{\text{Bayes}}\Big) \xrightarrow[n\to\infty]{} = \mathbb{E}\Big[1 - \max_{1\leq k\leq K} \nu(\{k\})\Big] = \mathbb{P}\Big(\widetilde{Y}_0 \neq \widehat{Y}_0^{\text{Bayes}}\Big),$$

where $\widehat{Y}_0^{\text{Bayes}}$ is the Bayes classifier for estimating $\widetilde{Y}_0 \sim P_{\text{out}}\Big(\cdot\,\Big|\sqrt{q^*(\alpha)}V + \sqrt{\rho - q^*(\alpha)}W^*\Big)$ given V:

$$\widehat{Y}_0^{\text{Bayes}} = \underset{k\in\{1,\dots,K\}}{\arg\max}\ \mathbb{P}(\widetilde{Y}_0 = k \,|\, V).$$

7.4 The generalized approximate message-passing algorithm

While the results presented until now are information-theoretic, the next one concerns the performance of a popular algorithm to solve random instances of generalized linear problems, called generalized approximate message-passing (GAMP). This approach has a long history, especially in statistical physics [203, 142, 117, 15], error correcting codes [181], and graphical models [214]. For a modern derivation in the context of linear models, see [71, 127, 212]. The case of generalized linear models was discussed by Rangan in [177], and has been used for classification purpose in [222].

We first need to define two so-called threshold functions that are associated to the two scalar channels (7.3.1) and (7.3.3). The first one is the posterior mean of the signal in channel (7.3.1) with signal-to-noise ratio r:

$$g_{P_0}(y, r) \overset{\text{def}}{=} \mathbb{E}[X_0|Y_0 = y]\,. \tag{7.4.1}$$

The second one is the posterior mean of W^* in channel (7.3.3):

$$g_{P_{\text{out}}}(\widetilde{y}, v, q) \overset{\text{def}}{=} \mathbb{E}[W^*|\widetilde{Y}_0 = \widetilde{y}, \sqrt{q}\,V = v]\,. \tag{7.4.2}$$

These functions act componentwise when applied to vectors.

Given initial estimates $(\widehat{\boldsymbol{x}}^0, \boldsymbol{v}^0)$ – that we take equal to 0 in absence of additional information – for the means and variances of the elements of the signal vector $\boldsymbol{X}^*$, GAMP takes as input the observation vector $\boldsymbol{Y}$ and then iterates the following equations with initialization $g_\mu^0 = 0$ for all $\mu = 1, \dots, m$ (we denote by $\overline{\boldsymbol{u}}$ the average over all the components

of the vector $\boldsymbol{u}$ and $\boldsymbol{\Phi}^\mathsf{T}$ is the transpose of the matrix $\boldsymbol{\Phi}$): From $t = 1$ until convergence,

$$\left\{\begin{array}{lcll} V^t & = & \overline{\boldsymbol{v}^{t-1}} & \\ \boldsymbol{\omega}^t & = & \boldsymbol{\Phi}\widehat{\boldsymbol{x}}^{t-1}/\sqrt{n} - V^t \boldsymbol{g}^{t-1} & \\ g_\mu^t & = & g_{P_{\text{out}}}(Y_\mu, \omega_\mu^t, \rho - V^t) & \forall\, \mu = 1, \dots m \\ \lambda^t & = & \alpha\, \overline{g^2_{P_{\text{out}}}(\boldsymbol{Y}, \boldsymbol{\omega}^t, \rho - V^t)} & \\ \boldsymbol{R}^t & = & \widehat{\boldsymbol{x}}^{t-1} + (\lambda^t)^{-1} \boldsymbol{\Phi}^\mathsf{T} \boldsymbol{g}^t/\sqrt{n} & \\ \widehat{x}_i^t & = & g_{P_0}(R_i^t, \lambda^t) & \forall\, i = 1, \dots n \\ v_i^t & = & (\lambda^t)^{-1}\, \partial_R g_{P_0}(R, \lambda^t)|_{R=R_i^t} & \forall\, i = 1, \dots n. \end{array}\right. \tag{7.4.3}$$

One of the strongest asset of GAMP is that its performance can be tracked rigorously in the limit $n, m \to \infty$ while $m/n \to \alpha$ via a procedure known as state evolution (SE), see [26, 25] for the linear case, and [177, 112] for the generalized one. The state evolution tracks the asymptotic value of the overlap between the true hidden value $\boldsymbol{X}^*$ and its estimate by GAMP $\widehat{\boldsymbol{x}}^t$ defined as $q^t \overset{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n}\langle \boldsymbol{X}^*, \widehat{\boldsymbol{x}}^t\rangle$ via $q^0 = 0$ and

$$\left\{\begin{array}{lcl} q^{t+1} & = & 2\psi'_{P_0}(r^t)\,, \\ r^t & = & 2\alpha \Psi'_{P_{\text{out}}}(q^t)\,. \end{array}\right. \tag{7.4.4}$$

From Theorem 7.3.1 we realize that the fixed points of these equations correspond to the critical points of the asymptotic free energy in (7.3.8). By Proposition 1.3.3 and Proposition 7.8.2 the functions Ψ'_{P_0} and $\Psi'_{P_{\text{out}}}$ are both non-decreasing. This gives that (q^t, r^t) converges as $t \to \infty$ to some couple $(q_{\text{AMP}}, r_{\text{AMP}}) \in \Gamma$.

Perhaps more surprisingly, one can use GAMP for predicting a new output $Y_{\text{new}} \sim P_{\text{out}}(\cdot \,|\, \langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle/\sqrt{n})$ where $\boldsymbol{\Phi}_{\text{new}} \sim \mathcal{N}(0, \mathrm{Id}_n)$ is independent of $\boldsymbol{X}^*, \boldsymbol{\Phi}, \boldsymbol{Y}$. As $\widehat{\mathbf{x}}^t$ is the GAMP estimate of the posterior expectation of $\boldsymbol{X}^*$, the natural heuristic is to consider for the posterior probability distribution (given $\boldsymbol{Y}, \boldsymbol{\Phi}$) of the random variable $\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle/\sqrt{n}$ a Gaussian with mean $\langle \boldsymbol{\Phi}_{\text{new}}, \widehat{\mathbf{x}}^t\rangle/\sqrt{n}$ and variance $\rho - q^t$. This allows to estimate the posterior mean of the output, which leads to the GAMP prediction:

$$\widehat{Y}^{\text{GAMP},t} \overset{\text{def}}{=} \int y\, P_{\text{out}}\Big(dy \,\Big|\, \frac{1}{\sqrt{n}}\langle \boldsymbol{\Phi}_{\text{new}}, \widehat{\mathbf{x}}^{t-1}\rangle + \sqrt{\rho - q^t}\, w\Big) \frac{e^{-w^2/2}}{\sqrt{2\pi}} dw. \tag{7.4.5}$$

7.4.1 Estimation and generalization error of GAMP

The following claim, from [177], gives the precise estimation error of GAMP. It is stated there as a claim because some steps of the proof are missing. The paper [112] affirms in its abstract to prove the claim of [177], but without further details. For these reasons, we believe that the result holds, however we prefer to state it here as a claim (instead of a theorem).

Claim 7.4.1 (*From [177]*)

We have almost surely for all $t \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{1}{n}\langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^*\rangle = \lim_{n\to\infty} \frac{1}{n}\|\widehat{\boldsymbol{x}}^t\|^2 = q^t. \tag{7.4.6}$$

Consequently

$$\lim_{n\to\infty} \frac{\langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^*\rangle}{\|\widehat{\boldsymbol{x}}^t\|\|\boldsymbol{X}^*\|} = \sqrt{\frac{q^t}{\rho}}. \tag{7.4.7}$$

Comparing (7.4.7) with the optimal correlation given by Corollary 7.3.2, we see that if $\lim_{t\to\infty} q^t = q^*(\alpha)$, then GAMP is information-theoretically optimal. Provided that Claim 7.4.1 holds we deduce the generalization error of GAMP:

Proposition 7.4.1

Suppose that hypotheses (h1)-(h2)-(h4) hold. Moreover suppose that either (h5.a) or (h5.b) holds. Assume that $(\Phi_{\mu i}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, *and that* $x \mapsto P_{\text{out}}(\cdot|x)$ *is continuous almost everywhere for the Wasserstein distance of order* 2. *Let* $t \in \mathbb{N}$. *Assume that the limit (7.4.6) holds in probability and that there exists* $\eta > 0$ *such that* $\mathbb{E}[|\widehat{Y}^{\text{GAMP},t}|^{2+\eta}]$ *remains bounded (as* n *grows). Then we have for all* $t \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathcal{E}_{\text{gen}}^{\text{GAMP},t} \overset{\text{def}}{=} \lim_{n\to\infty} \mathbb{E}\Big[\Big(Y_{\text{new}} - \widehat{Y}^{\text{GAMP},t}\Big)^2\Big] = \mathcal{E}(q^t)\,. \tag{7.4.8}$$

Proposition 7.4.1 is proved in Section 7.7.4. We see that this formula matches the one for the Bayes-optimal generalization error, see Theorem 7.3.3, up to the fact that instead of $q^*(\alpha)$ (the optimizer of the replica formula (7.3.8)) appearing in the optimal error formula, here it is q^t which appears. Thus clearly, when q^t converges to $q^*(\alpha)$ (we shall see in Section 7.5 that this is the case in many situations): *GAMP is Bayes-optimal* in a plethora of models (a task often believed to be intractable) and this for large sets of parameters.

7.5 Examples of phase transitions

We illustrate now the results of the previous sections to several models of interest in fields ranging from machine learning to signal processing, and unveil several interesting new phenomena in learning of generalized linear models. For various specific cases of prior P_0 and output P_{out}, we evaluate numerically the free energy potential (7.3.6), its stationary points Γ and identify which of them gives the information-theoretic results, i.e. is the optimizer in (7.3.8). We also identify which of the stationary points corresponds to the result obtained asymptotically by the GAMP algorithm, i.e. the fixed point of the state evolution (7.4.4). Finally we compute the corresponding generalization error (7.4.8). We stress that in this section the results are based on numerical investigation of the resulting formulas: We do not aim at rigor that would involve precise bounds and more detailed analytical control for the corresponding integrals.

7.5.1 Generic observations

Using the functions $g_{P_{\text{out}}}$ and g_{P_0} introduced in Section 7.4 we can rewrite the fixed point equations (7.3.7) as

$$q = 2\psi'_{P_0}(r) = \mathbb{E}[g_{P_0}(Y_0, r)^2]\,, \tag{7.5.1}$$

$$r = 2\alpha\Psi'_{P_{\text{out}}}(q) = \frac{\alpha}{\rho - q}\mathbb{E}[g_{P_{\text{out}}}(\widetilde{Y}_0, \sqrt{q}\,V, q)^2]\,, \tag{7.5.2}$$

where the expectation in (7.5.1) corresponds to the scalar channel (7.3.1) and the expectation in (7.5.2) corresponds to the second scalar channel (7.3.3).

The non-informative fixed point: It is interesting to analyze under what conditions $q^* = 0$ is the optimizer of (7.3.8). Notice that $q^* = 0$ corresponds to the error on the recovery of the signal as large as it would be if we had no observations at our disposition. Theorem 7.3.1 gives that any optimal couple (q^*, r^*) of (7.3.8) should be a fixed point of the state evolution equations (7.5.1)–(7.5.2). A sufficient condition for $(q, r) = (0, 0)$ to be a fixed point of (7.5.1)–(7.5.2) is that:

(a) The transition kernel P_{out} is symmetric (see Definition 7.3.1).

(b) The prior P_0 has zero mean.

In order to see this, notice that if $P_{\text{out}}(y|z)$ is even in z then from the definition (7.4.2) of the function $g_{P_{\text{out}}}$ we have $g_{P_{\text{out}}}(y, 0, 0) = 0$ and consequently from (7.5.2) we have $\Psi'_{P_{\text{out}}}(0) = 0$. For the second point, notice that we have $\psi'_{P_0}(0) = \frac{1}{2}\mathbb{E}_{P_0}[X_0]^2 = 0$.

We assume now that P_{out} is symmetric and that the prior P_0 has zero mean. In order for $q = 0$ to be the global maximizer q^* of (7.3.8) or to be a relevant fixed point of the state evolution (7.4.4) (relevant in the sense that GAMP might indeed converge to it in a practical setting) we need $q = 0$ to be a stable fixed point of the equations (7.5.1)–(7.5.2). We obtain that $q = 0$ is stable if

$$2\alpha\Psi''_{P_{\text{out}}}(0) \times 2\psi''_{P_0}(0) = \alpha \int dy \frac{\left(\int \mathcal{D}z(z^2-1)P_{\text{out}}(y|\sqrt{\rho}z)\right)^2}{\int \mathcal{D}z P_{\text{out}}(y|\sqrt{\rho}z)} < 1\,, \tag{7.5.3}$$

where $\mathcal{D}z$ is the standard Gaussian measure. We conjecture that the condition (7.5.3) delimits precisely the region where polynomial-time algorithms do not perform better than "random guessing" (see the discussion below, where we will make this stability condition explicit for several examples of symmetric output channels). Note that the condition (7.5.3) also appears in a recent work [147] as a barrier for performance of spectral algorithms.

The exact recovery fixed point: Another particular fixed point of (7.5.1)–(7.5.2) is the one corresponding to exact recovery $q = \rho$. A sufficient and necessary condition for this to be a fixed point is that $\lim_{q\to\rho} \Psi'_{P_{\text{out}}}(q) = +\infty$, i.e.

$$\frac{\Psi_{P_{\text{out}}}(\rho) - \Psi_{P_{\text{out}}}(q)}{\rho - q} \xrightarrow[q\to\rho]{} +\infty. \tag{7.5.4}$$

Consider the second scalar channel (7.3.3) and let P_V (respectively P_{V,W^*}) denote the conditional law of $\widetilde{Y}_0$ given V (respectively V, W^*). Then we have $\Psi_{P_{\text{out}}}(\rho) - \Psi_{P_{\text{out}}}(q) = \mathbb{E}[\mathrm{D}_{\mathrm{KL}}(P_V, P_{V,W^*})]$. If the statistical model $(P_{\text{out}}(\cdot|\theta))_{\theta\in\mathbb{R}}$ is regular (see Definitions 1 and 2 in [32]), and if we denote by $J(\theta)$ its Fisher information at θ, then we have for $\theta, \theta_0 \in \mathbb{R}$

$$\mathrm{D}_{\mathrm{KL}}(P_{\text{out}}(\cdot|\theta_0), P_{\text{out}}(\cdot|\theta)) = \frac{1}{2}(\theta - \theta_0)^2 J(\theta_0) + o((\theta - \theta_0)^2).$$

This leads (after some computations) to

$$\Psi_{P_{\text{out}}}(\rho) - \Psi_{P_{\text{out}}}(q) = \mathbb{E}[\mathrm{D}_{\mathrm{KL}}(P_V, P_{V,W^*})] = \frac{1}{2}\mathbb{E}[J(\sqrt{\rho}V)](\rho - q) + o_{q\to\rho}(\rho - q).$$

Consequently (7.5.4) can not hold for such model: perfect recovery is impossible.

In the case of a discrete deterministic channel, where $P_{\text{out}}(\cdot|x) = \delta_{\varphi(x)}$ where $\varphi : \mathbb{R} \to \mathbb{N}$, perfect reconstruction will sometimes be possible. Let us consider the "planted perceptron" problem from Section 7.1 where $\varphi(x) = \text{sign}(x)$ (other functions φ can be treated analogously). In that case $\Psi_{P_{\text{out}}}(\rho) = 0$ and

$$\begin{aligned}\Psi_{P_{\text{out}}}(q) &= 2\mathbb{E}\Big[\Phi\Big(\sqrt{\frac{q}{\rho-q}}Z\Big)\log\Phi\Big(\sqrt{\frac{q}{\rho-q}}Z\Big)\Big] \\ &= 2\sqrt{\frac{\rho-q}{q}}\int_{-\infty}^{+\infty} e^{-\frac{\rho-q}{2q}x^2}\Phi(x)\log\Phi(x)dx \\ &= -\gamma\sqrt{\rho-q} + o_{q\to\rho}\Big(\sqrt{\rho-q}\Big),\end{aligned}$$

where $\gamma = -\frac{2}{\sqrt{\rho}}\int \Phi(x)\log\Phi(x)dx$. Hence $\Psi'_{P_{\text{out}}}(q) \sim \gamma(\rho-q)^{-1/2}$: (7.5.4) holds, $(\rho, +\infty) \in \Gamma$. We have now to verify if $q = \rho$ is a stable fixed point of (7.4.4). Using the change of variables $u = \rho - q$ we have to check whether $u = 0$ is a stable fixed point of the function

$$f(u) = \rho - \psi'_{P_0}(\alpha\Psi'_{P_{\text{out}}}(\rho-u)) = \text{MMSE}_{P_0}(\alpha\Psi'_{P_{\text{out}}}(\rho-u)).$$

The stability of $u = 0$ depends now on the properties of P_0 through the behavior of $\text{MMSE}_{P_0}(r)$ as $r \to \infty$. This "large signal-to-noise ratio" asymptotic has been studied in details in [215] whose conclusions can be roughly summarized as follows:

$$\text{MMSE}_{P_0}(r) = \begin{cases}\Theta(r^{-1}) & \text{if } P_0 \text{ has a "continuous part"} \\ o(r^{-1}) & \text{if } P_0 \text{ is "discrete".}\end{cases}$$

We refer to the paper [215] for a precise definition of "continuous" and "discrete" and consider two examples below:

- $P_0 = \rho\mathcal{N}(0,1) + (1-\rho)\delta_0$, for some $\rho \in (0,1]$. In that case, $\text{MMSE}_{P_0}(r) \sim \frac{\rho}{r}$ and consequently $f(u) \sim \gamma^{-1}\rho\sqrt{u}$ as $u \to 0$. The fixed point $u = 0$ is therefore not stable: perfect reconstruction is not possible in this case. Indeed, as we will see on Figure 7.6 on page 132 (middle plot), the generalization error decreases but never reaches 0.
- $P_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$. In that case, $\text{MMSE}_{P_0}(r) = o(r^{-2})$ and consequently $f(u) = o(u)$ as $u \to 0$. This time, the fixed point $u = 0$ is stable: perfect reconstruction is possible (provided that $(\rho, +\infty)$ maximizes $\mathcal{F}$ over Γ) as we saw it in Section 7.1.

Below we give several examples where exact recovery either is or is not possible, or where there is a phase transition between the two regimes.

7.5.2 Phase diagram of perfect learning

In this section we consider *deterministic (noiseless) output channels* and ask: How many measurements are needed in order to perfectly recover the signal?

As the number of samples (measurements) varies we encounter five different regimes of parameters:

- The *tractable recovery* phase: This is the region in the parameter space where GAMP achieves perfect reconstruction.

- The *non-informative* phase: Region where perfect reconstruction is information-theoretically impossible and moreover even the Bayes-optimal estimator is as bad as a random guess based on the prior information and on the knowledge of the output function.
- The *no recovery* phase: Region where perfect reconstruction is information-theoretically impossible, but an estimator positively correlated with the ground truth exists.
- The *hard* phase: Region where the perfect reconstruction is information-theoretically possible, but where GAMP is unable to achieve it. At the same time, in this region GAMP leads to a better generalization error than the one corresponding to the non-informative fixed point. It remains a challenging open question whether polynomial-time algorithms can achieve perfect reconstruction in this regime.
- The *hard non-informative* phase: This phase corresponds to the region where perfect reconstruction is information-theoretically possible but where GAMP only achieves an error as bad as randomly guessing, given by the non-informative fixed point. In this phase as well, the existence of polynomial-time exact recovery algorithms is an open question.

The linear channel

The case of exact recovery of a sparse signal after it passed trough a noiseless linear channel, i.e. $\varphi(x) = x$, is studied in the literature in great details, especially in the context of compressed sensing [43]. For a signal with a fraction ρ of non-zero entries, as soon as $\alpha > \rho$, perfect reconstruction is theoretically possible, since the linear system $\boldsymbol{Y} = \boldsymbol{\Phi x}$ admits almost surely a unique solution that has ρn non-zero entries, which is the signal $\boldsymbol{X}^*$. However solving $\boldsymbol{Y} = \boldsymbol{\Phi x}$ over the space of ρn-sparse vectors remain (a priori) computationally difficult since it requires to test all the $\binom{n}{\rho n}$ possible positions of the non-zero entries of $\boldsymbol{x}$.

The whole field of compressed sensing builds on the realization that, using the ℓ_1 norm minimization technique:

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x} \,|\, \boldsymbol{Y} = \boldsymbol{\Phi x}}{\arg\min} \|\boldsymbol{x}\|_1,$$

one can efficiently recover the signal for α larger than a critical value $\alpha_{\ell_1}(\rho)$:

$$\alpha_{\ell_1}(\rho) = \min_{\gamma \geq 0} \left\{ \rho(1+\gamma^2) + 2(1-\rho)\Big((1+\gamma^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big) \right\}. \tag{7.5.5}$$

This is the so-called Donoho-Tanner transition [73, 68, 74]. In our Bayesian context, when the empirical distribution of the signal is known, one can fairly easily beat the ℓ_1 transition and reconstruct the signal up to lower values of α using GAMP algorithm [71, 176, 127, 128]. In this case, three different phases are present [127, 128]:

(i) For $\alpha < \rho$, perfect reconstruction is impossible;
(ii) for $\rho < \alpha < \alpha_{\text{AMP}}$ reconstruction is possible, but not with any known polynomial-complexity algorithm;
(iii) for $\alpha > \alpha_{\text{AMP}}$, the so-called *spinodal transition* computed with state evolution, GAMP provides a polynomial-complexity algorithm able to reach perfect reconstruction.

The line $\alpha_{\text{AMP}}(\rho)$ depends on the distribution of the signal. For a Gauss-Bernoulli signal with a fraction ρ of non-zero (Gaussian) values we plot α_{ℓ_1}, α_{AMP} and $\alpha_{\text{IT}}(\rho) = \rho$ on Figure 7.4.

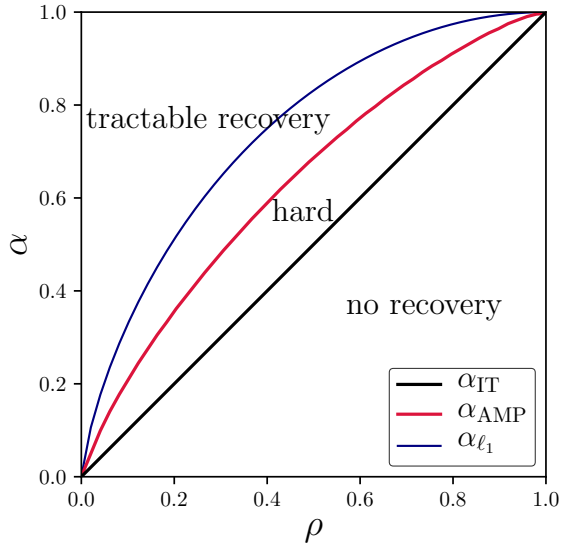


Figure 7.4: Phase diagram of noiseless compressed sensing with Gauss-Bernoulli prior: $P_0 = \rho\mathcal{N}(0,1) + (1-\rho)\delta_0$.

The rectified linear unit (ReLU) channel

Let us start by discussing the case of a generalized linear model with the ReLU output channel, i.e. $\varphi(x) = \max(0, x)$, with a signal coming from a Gauss-Bernoulli distribution $P_0 = \rho\mathcal{N}(0,1) + (1-\rho)\delta_0$, i.e. $\boldsymbol{X}^*$ has a fraction ρ of non-zero (Gaussian) values. We are motivated by the omnipresent use of the ReLU activation function in deep learning, and explore its properties for GLMs that can be seen as a simple single layer neural network.

Our analysis shows that a perfect generalization (and thus a perfect reconstruction of the signal as well) is possible whenever the number of samples per dimension (measurement rate) $\alpha > 2\rho$, and impossible when $\alpha < 2\rho$. This is very intuitive, since half of the measurements (those non-zero) are giving as much information as in the linear case, thus the factor 2.

How hard is it to actually solve the problem with an efficient algorithm? The answer is given by applying the state evolution analysis to GAMP, which tells us that only for larger values of α, beyond the spinodal transition α_{AMP}, does GAMP reach a perfect recovery. Notice, however, that this spinodal transition occurs at a significantly lower measurement rate α than one would reach just keeping the non-zero measurements. This shows that, actually, these zero measurements contain a useful information for the algorithm. The situation is shown in the center panel of Figure 7.5: The zero measurements do not help information-theoretically but they, however, do help algorithmically.

The sign-less channel

We now discuss the sign-less channel where only the absolute value of the linear mixture is observed, i.e. $\varphi(x) = |x|$. This case can be seen as the real-valued analog of the famous phase retrieval problem. We again consider the signal to come from a Gauss-Bernoulli distribution $P_0 = \rho\mathcal{N}(0,1) + (1-\rho)\delta_0$.

Sparse phase retrieval has been well explored in the literature in the regime where the number s of non-zeros is sub-leading in the dimension, $s = o(n)$. This case is known to present a large algorithmic gap. While analogously to compressed sensing exact recovery is information-theoretically possible for a number of measurement $\Omega(s\log(n/s))$, best known

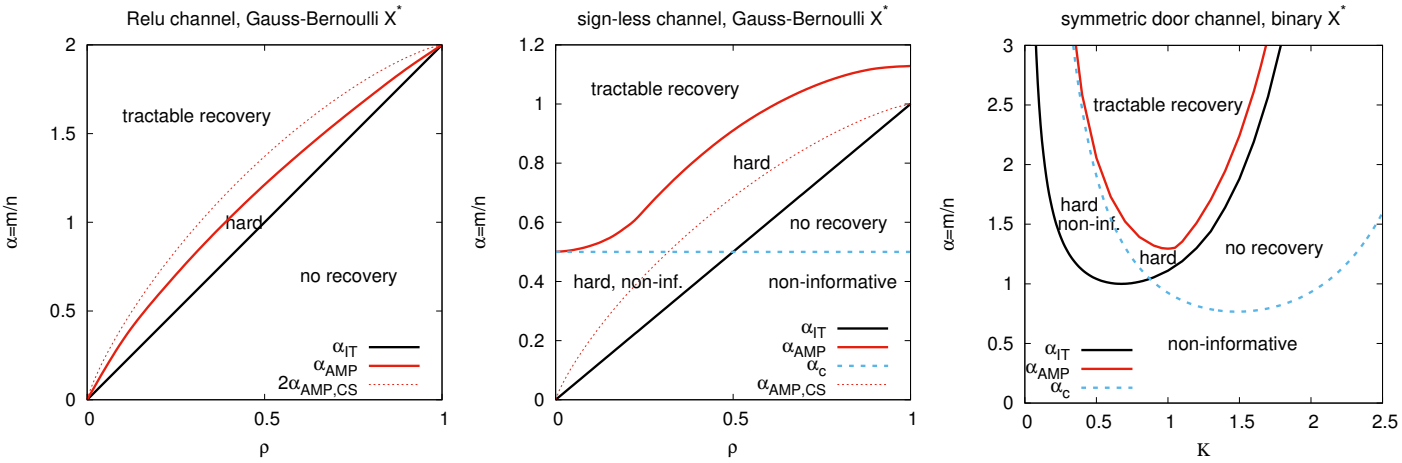


Figure 7.5: Phase diagrams showing boundaries of the region where exact recovery is possible (in absence of noise). **Left:** The case of ReLU activation function, $\varphi(x) = \max(0, x)$ with a Gauss-Bernoulli signal $P_0 = \rho\mathcal{N}(0,1) + (1-\rho)\delta_0$, as a function of the ratio between number of samples/measurements and the dimension $\alpha = m/n$, and the fraction of non-zero components ρ. The dotted red line shows the algorithmic phase transition when using information only about the non-zero observations. **Center:** Analogous to the left panel, for the absolute value function: $\varphi(x) = |x|$. The dotted red line shows for comparison the algorithmic phase transition of the canonical compressed sensing. **Right:** Phase diagram for the symmetric door output function $\varphi(z) = \text{sign}(|z| - K)$ for a Rademacher signal, as a function of α and K.

algorithms achieve it only with $\Omega(s^2/\log n)$ measurements [168], see also [192] and references therein for a good discussion of other related literature. This is sometimes referred to as the "s^2 barrier". We are not aware of a study where, as in our setting, the sparsity is $s = \rho n$ and the number of measurements is αn with α and ρ of order 1.

Perfect reconstruction is information-theoretically possible as soon as $\alpha > \rho$: In other words, the problem is –information-theoretically– as easy, or as hard as the compressed sensing one. When $\alpha > \rho$ one can indeed perfectly reconstruct the signal by the following procedure: Try all 2^m choices of the possible signs for the m outputs, and solve a compressed sensing problem for each of them. Clearly, this should yields a perfect solution only in the case of the actual combination of signs.

Algorithmically, however, the problem is much harder than for the linear output channel. As shown in the left side of Figure 7.5 on page 131, for small ρ one requires a much larger fraction α of measurements in order for GAMP to recover the signal. For the linear channel the algorithmic transition $\alpha_{\text{AMP}}(\rho) \to 0$ as $\rho \to 0$ (see Figure 7.4) while for the sign-less channel we get $\alpha_{\text{AMP}}(\rho) \to 1/2$ as $\rho \to 0$. In other words if one looses the signs one cannot perform recovery in compressed sensing with less than $n/2$ measurements.

What we observe in this example for $\alpha < 1/2$ is in the statistical physics literature on neural networks known as retarded learning [99]. This appears in problems where the $\varphi(x)$ function is symmetric, as seen in Section 7.5.1: There is always a critical point of the mutual information with an overlap value $q = 0$. For this problem, this critical point is actually "stable" (meaning that it is actually a local minimum in q in the mutual information (7.3.9)) for all $\alpha < 1/2$, by (7.5.3), independently of ρ.

This has the two following implications:

(i) In the *non-informative phase*, when $\alpha < 1/2$ and $\alpha < \rho$, the minimum at $q = 0$ is actually the global one. In this case there is no useful information that one can

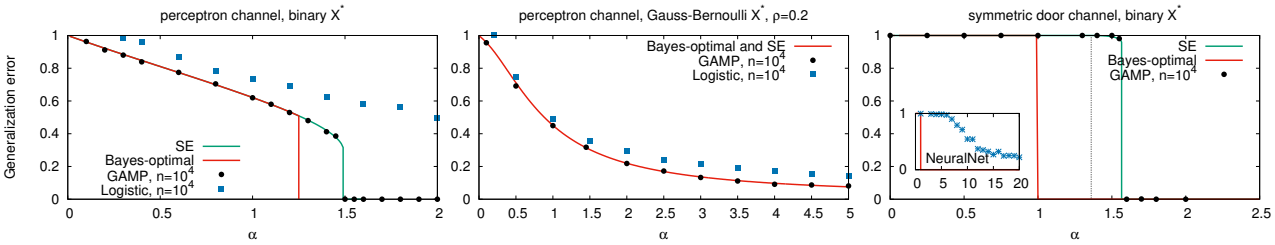


Figure 7.6: Generalization error in three classification problems as a function of the number of data-samples per dimension α. The red line is the Bayes-optimal generalization error, while the green one shows the (asymptotic) performances of GAMP as predicted by the state evolution (SE), when different. For comparison, we also show the result of GAMP (black dots) and, in blue, the performance of a standard out-of-the-box solver, both tested on a single randomly generated instance. **Left:** Perceptron, with $\varphi(x) = \text{sign}(x)$ and a Rademacher (± 1) signal. The results of a logistic regression with fine-tuned ridge penalty are shown for comparison. **Middle:** Perceptron with Gauss-Bernoulli coefficients for the signal. The results of a logistic regression with fine-tuned ℓ_1 sparsity-enhancing penalty are very close to optimal. **Right:** The symmetric door activation rule with parameter $K = K^* \simeq 0.67449$ chosen in order to observe the same number of occurrence of the two classes. Using Keras [52], a neural network with two hidden layers was able to learn approximately the rule, but only for much larger training set sizes (shown in inset).

exploit and no estimator can be better than a random guess.

(ii) In the *hard non-informative phase* when $\rho < \alpha < 1/2$, GAMP initialized at random, i.e. close to the $q = 0$ fixed point, will remain there. This suggests that in this region, even if a perfect reconstruction is information-theoretically possible, it will still be very hard to be better then a random guess with a tractable algorithm.

The symmetric door channel

The third output channel we study in detail is the symmetric door channel, where $\varphi(x) = \text{sign}(|x| - K)$. In case of channels with discrete set of outputs exact recovery is only possible when the prior is also discrete. In the present case we take $\boldsymbol{X}^* \sim \text{Unif}(\{-1, 1\}^n)$, i.e. $P_0 = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}$. This channel was studied previously using the replica method in the context of optimal data compression [108].

This output channel is in the class of symmetric channels for which overlap $q = 0$ is a fixed point. This fixed point is stable for $\alpha < \alpha_c(K)$. Exact recovery is information-theoretically possible above $\alpha_{\text{IT}}(K)$ and tractable with the GAMP algorithm above the spinodal transition $\alpha_{\text{AMP}}(K)$. The values of these three transition lines are depicted in the right panel of Figure 7.5.

We note that $\alpha_{\text{IT}} \geq 1$ is a generic bound on exact recovery for every K, required by a simple counting argument. While a-priori it is not clear whether this bound is saturated for some K, we observe that it is for $K = K^* \simeq 0.67449$ defined by $\mathbb{P}(|Z| \leq K^*) = 1/2$, for $Z \sim \mathcal{N}(0, 1)$. In that case half of the observed measurements are negative and the rest positive. The saturation of the $\alpha_{\text{IT}} \geq 1$ bound was remarked previously in [108]. However, we conjecture that this information-theoretic limit will not be achievable with known efficient algorithms since GAMP recovers the signal for α larger that $\alpha_{\text{AMP}} > 1$.

7.5.3 Examples of generalization errors

In this section we evaluate the Bayes-optimal generalization error (given by Theorem 7.3.3) and the generalization error of GAMP (given by Proposition 7.4.1) for several cases of priors and output functions. We study both regression problems, where the output is real-valued, and classification problems, where the output is discrete.

While in realistic regression and classification problems the matrix $\mathbf{\Phi}$ corresponds to the data, and is thus not i.i.d. random, we view the practical interest of our theory as a benchmark for state-of-the art algorithms. The results of Sections 7.3 and 7.4 provide an exact asymptotic analysis of optimal generalization error and sample complexity for a range of simple rules. The challenge for state-of-the-art multi-purpose algorithms is to try to match as closely as possible the performance that can be obtained with GAMP that is fine-tuned to the specific form of the output and prior.

Threshold output: The perceptron

The example of non-linear output that is the most widely explored in the literature is the threshold output, where the deterministic output function is $\varphi(x) = \text{sign}(x)$. In the setting of the present chapter, this is known as the perceptron problem [89], or equivalently, the one-bit compressed sensing in signal processing [38]. Its solution has been discussed in details within the replica formalism (see for instance [98, 23, 167, 117, 218]) and we confirm all of these heuristic computations within our approach.

In Figure 7.6 (left) we plot the optimal generalization error of the perceptron with a Rademacher signal, the state evolution prediction of the generalization error of the GAMP algorithm, together with the error actually achieved by GAMP on one randomly generated instance of the problem. We also compare these to the performance of a standard logistic regression. As expected from existing literature [89, 98, 189] we confirm that in this case the information-theoretic transition appears at a number of samples per dimension $\alpha_{\text{IT}} = 1.249(1)$, while the algorithmic transition is at $\alpha_{\text{AMP}} = 1.493(1)$. Logistic regression does not seem to be able to match the performance on GAMP in this case.

In Figure 7.6 (center) we plot the generalization error for a Gauss-Bernoulli signal with density $\rho = 0.2$. Cases as this one were studied in detail in the context of one-bit compressed sensing [218] and GAMP was found to match the optimal generalization performance with no phase transitions observed, which is confirmed by our analysis. In this case the logistic regression is rather close to the performance of GAMP.

Symmetric Door

The next classification problem, i.e. discrete output rule, that we study is the symmetric door function $\varphi(x) = \text{sign}(|x| - K)$. In Figure 7.6 (right) we plot the generalization error for $K = K^* \simeq 0,67449$ such that $1/2$ of the outputs are 1 and $1/2$ are -1. The symmetric door output is an example of function for which the optimal generalization error for $\alpha < \alpha_{\text{IT}} = 1$ (for that specific value of K, see phase diagram in the right panel of Figure 7.5) is as bad as if we were guessing randomly. The GAMP algorithm still achieves such a bad generalization until $\alpha_{\text{stab}} \simeq 1.36$ (defined by (7.5.3)), and achieves perfect generalization only for $\alpha > \alpha_{\text{AMP}} = 1.566(1)$.

Interestingly, labels created from this very simple symmetric door rule seem to be very challenging to learn for general purpose algorithms. We tried to optimize parameters of a two-layers neural network and only managed to get the performances shown in the inset

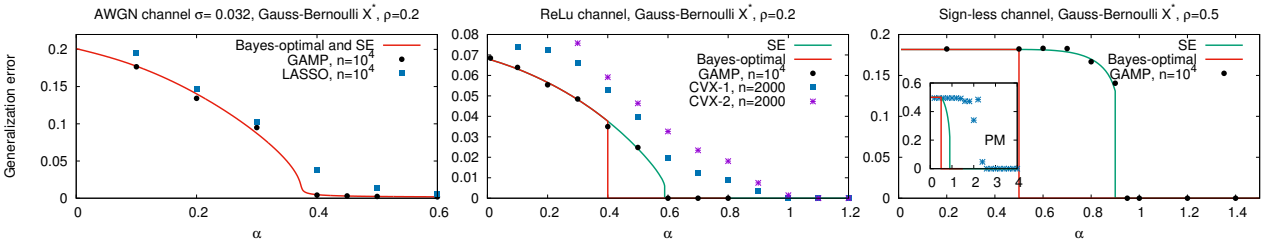


Figure 7.7: The generalization error for three regression problems is plotted as a function of the number of samples per dimension α. The red line is again the Bayes-optimal generalization error, while the green one shows the (asymptotic) performances of GAMP as predicted by the state evolution (SE), when different. Again, we also show the result of GAMP on a particular instance (black dots) and, in blue, the performance of an out-of-the-box solver. **Left:** White Gaussian noise output and a Gauss-Bernoulli signal. For this choice of noise, there is no sharp transition (as opposed to what happens at smaller noises). The results of a Lasso with fine-tuned ℓ_1 sparsity-enhancing penalty are very close to optimal. **Middle:** Here we analyze a ReLU output function $\varphi(x) = \max(0, x)$, still with a Gauss-Bernoulli signal. We show for comparison the results of maximum likelihood estimation performed with CVXPY —a powerful python-embedded language for convex optimization [64]— using two methods that are both amenable to convex optimization: In CVX-1 we use only the non-zero values of $\boldsymbol{Y}$, and perform a minimization of the ℓ_1 norm of $\boldsymbol{x}$ subject to $Y_\mu = \langle \boldsymbol{\Phi}_\mu, \boldsymbol{x} \rangle$ for $\mu \in \{1, \dots, m\}$ such that $Y_\mu \neq 0$, while in CVX-2, we use all the dataset, with the constraint that $Y_\mu = \langle \boldsymbol{\Phi}_\mu, \boldsymbol{x} \rangle$ for $\mu \in \{1, \dots, m\}$ such that $Y_\mu \neq 0$ (as before) and the additional restriction $\langle \boldsymbol{\Phi}_\mu, \boldsymbol{x} \rangle \leq 0$ for $\mu \in \{1, \dots, m\}$ such that $Y_\mu = 0$. In both case, a perfect generalization is obtained only for $\alpha \gtrapprox 1$. **Right:** The sign-less output function $\varphi(x) = |x|$. In inset, we show the performance for the estimation problem using PhaseMax [92].

of Figure 7.6 (right). It is an interesting theoretical challenge whether a deeper neural network can learn this simple rule from fewer samples.

Linear regression

The additive white Gaussian noise (AWGN) channel, or linear regression, is defined by $\varphi(x, A) = x + \sigma A$ with $A \sim \mathcal{N}(0, 1)$. This models the (noisy) linear regression problem, as well as noisy random linear estimation and compressed sensing. In this case (7.3.20) leads to

$$\lim_{n\to\infty} \mathcal{E}_{\rm gen}^{\rm opt} = \rho - q^* + \sigma^2 \,. \tag{7.5.6}$$

This result agrees with the generalization error analyzed heuristically in [189] in the limit $\sigma \to 0$. Figure 7.7 (left) depicts the generalization error for this example. The performance of GAMP in this case is very close to the one of Lasso, that we will study in mode details in the next chapter.

Rectified linear unit (ReLU)

In Figure 7.7 (center) we analyze the generalization error for the ReLU output function, $\varphi(x) = \max(0, x)$. This channel models the behavior of a single neuron with the rectified linear unit activation widely used in multilayer neural networks.

For sparse Gauss-Bernoulli signals in Figure 7.7 (center) we observe again the information-theoretic transition (at $\alpha = 2\rho = 0.4$) to perfect generalization to be distinct from the algorithmic one (at $\alpha_{\rm AMP} = 0.589(1)$). At the same time our test with existing algorithms

were not able to closely match the performance of GAMP. This hence also remains an interesting benchmark.

Sign-less channel

In Figure 7.7 (right) we analyze the generalization error for the sign-less output function where $\varphi(x) = |x|$, that we already discussed in Section 7.5.2. The information-theoretic perfect recovery starts at $\alpha = \rho = 0.5$, but the problem is again harder algorithmically for GAMP that succeeds only above $\alpha_{\text{AMP}} = 0.90(1)$. Again, the problem appears to be hard for other solvers. The state-of-the-art algorithm PhaseMax [92] is for instance able to learn the rule using about four times as many measurements than needed information-theoretically.

7.6 Proof of Theorem 7.3.1

This section is devoted to the proof of Theorem 7.3.1. We will do it under the following hypotheses:

(H1) The support of the prior distribution P_0 is included in $[-S, S]$, for some $S > 0$.
(H2) φ is a bounded $\mathcal{C}^2$ function with bounded first and second derivatives w.r.t. its first argument.
(H3) $(\Phi_{\mu i}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$.
(H4) $\Delta > 0$.

These stronger assumptions can then be relaxed to the weaker assumptions (h1)-(h2)-(h3)-(h4) and (h5.a) or (h5.b). This is done by approximation arguments similar to the ones of Section 3.4.7 for (H1) and (H2). The hypothesis (H3) is relaxed using the "generalized Lindeberg swapping trick" from [45]. Finally the condition $\Delta > 0$ can be replaced by ((h5.b)) by approximation, using Corollary B.2. We refer to [19] for the details of these arguments.

Since the observations (7.2.1) with P_{out} given by (7.2.8), are equivalent to the rescaled observations

$$\Delta^{-1/2} Y_\mu = \Delta^{-1/2} \varphi\Big(\frac{1}{\sqrt{n}} [\boldsymbol{\Phi} \boldsymbol{X}^*]_\mu, A_\mu\Big) + Z_\mu \,, \qquad 1 \le \mu \le m \,, \tag{7.6.1}$$

the variance Δ of the Gaussian noise can be "incorporated" inside the function φ. Thus, it suffices to prove Theorem 7.3.1 for $\Delta = 1$ and we suppose, for the rest of the proof, that we are in this case.

7.6.1 Interpolating estimation problem

We aim at computing the limit of the free energy F_n. To do so, we introduce an estimation problem parametrized by $t \in [0,1]$ that interpolates between the original problem (7.2.1) at $t = 0$ and the two scalar problems described in Section 7.3.1 at $t = 1$ whose free energy is easy to compute. For $t \in (0,1)$ the interpolating estimation problem is a mixture of the original and scalar problems. This interpolation scheme is inspired from the interpolation paths used by Talagrand to study the perceptron, see [201]. There are two major differences between the perceptron studied by Talagrand, and the "planted perceptron" (where there is a "planted solution" $\boldsymbol{X}^*$) that we are investigating:

- In the planted case, the presence of a planted solution forces (under small perturbations) the correlations to vanish for all values of the parameters, see Chapter 2. In the non-planted case, proving such decorrelation is much more involved, and is proved only in a limited region of the parameter space (the high-temperature phase), see [201].

- However, in the planted case, there can be arbitrarily many solutions to the state evolution equations (7.3.7) (see Remark 21 in [217]), whereas in the region studied by [201], there is only one solution.

Let $q, r : [0,1] \to \mathbb{R}_{\geq 0}$ be two continuously differentiable functions such that $q(0) = r(0) = 0$. Define

$$S_{t,\mu} \stackrel{\text{def}}{=} \sqrt{\frac{1-t}{n}}\,[\boldsymbol{\Phi}\boldsymbol{X}^*]_\mu + \sqrt{q(t)}\,V_\mu + \sqrt{\rho t - q(t)}\,W^*_\mu \tag{7.6.2}$$

where $V_\mu, W^*_\mu \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Consider the following observation channels, with two types of observations obtained through

$$\begin{cases} Y_{t,\mu} & \sim & P_{\text{out}}(\,\cdot\,|\,S_{t,\mu})\,, & 1 \leq \mu \leq m, \\ Y'_{t,i} & = & \sqrt{r(t)}\,X^*_i + Z'_i\,, & 1 \leq i \leq n, \end{cases} \tag{7.6.3}$$

where $(Z'_i)_{i=1}^n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. We assume that $\boldsymbol{V} = (V_\mu)_{\mu=1}^m$ is *known*. Then the inference problem is to recover both unknowns $\boldsymbol{W}^* = (W^*_\mu)_{\mu=1}^m$ and $\boldsymbol{X}^* = (X^*_i)_{i=1}^n$ from the knowledge of $\boldsymbol{V}$, $\boldsymbol{\Phi}$ and the observations $\boldsymbol{Y}_t = (Y_{t,\mu})_{\mu=1}^m$ and $\boldsymbol{Y}'_t = (Y'_{t,i})_{i=1}^n$.

Define $u_y(x) \stackrel{\text{def}}{=} \log P_{\text{out}}(y|x)$ and, with a slight abuse of notations, for $(\boldsymbol{x}, \boldsymbol{w}) \in \mathbb{R}^n \times \mathbb{R}^m$

$$s_{t,\mu} = s_{t,\mu}(\boldsymbol{x}, w_\mu) \stackrel{\text{def}}{=} \sqrt{\frac{1-t}{n}}[\boldsymbol{\Phi}\boldsymbol{x}]_\mu + \sqrt{q(t)}\,V_\mu + \sqrt{\rho t - q(t)}\,w_\mu\,. \tag{7.6.4}$$

Notice that $S_{t,\mu} = s_{t,\mu}(\boldsymbol{X}^*, W^*_\mu)$. We introduce the *interpolating Hamiltonian*

$$H_t(\boldsymbol{x}, \boldsymbol{w}; \boldsymbol{Y}_t, \boldsymbol{Y}'_t, \boldsymbol{\Phi}, \boldsymbol{V}) \stackrel{\text{def}}{=} -\sum_{\mu=1}^m \log P_{\text{out}}(Y_{t,\mu}|s_{t,\mu}) + \frac{1}{2}\sum_{i=1}^n \left(Y'_{t,i} - \sqrt{r(t)}\,x_i\right)^2. \tag{7.6.5}$$

The dependence in $\boldsymbol{\Phi}$ and $\boldsymbol{V}$ of the Hamiltonian is through the $(s_{t,\mu})_{\mu=1}^m$. We also introduce the corresponding Gibbs bracket $\langle\cdot\rangle_{n,t}$ which is the expectation w.r.t. the posterior distribution of $(\boldsymbol{X}^*, \boldsymbol{W}^*)$ given $(\boldsymbol{Y}_t, \boldsymbol{Y}'_t, \boldsymbol{\Phi}, \boldsymbol{V})$. It is defined as

$$\langle g(\boldsymbol{x}, \boldsymbol{w})\rangle_{n,t} \stackrel{\text{def}}{=} \frac{1}{\mathcal{Z}_t(\boldsymbol{Y}_t, \boldsymbol{Y}'_t, \boldsymbol{\Phi}, \boldsymbol{V})} \int dP_0^{\otimes n}(\boldsymbol{x})\mathcal{D}\boldsymbol{w}\, g(\boldsymbol{x}, \boldsymbol{w})\, e^{-H_t(\boldsymbol{x},\boldsymbol{w};\boldsymbol{Y}_t,\boldsymbol{Y}'_t,\boldsymbol{\Phi},\boldsymbol{V})}\,, \tag{7.6.6}$$

for every continuous bounded function g on $\mathbb{R}^n \times \mathbb{R}^m$. In (7.6.6) $\mathcal{D}\boldsymbol{w} = (2\pi)^{-m/2}\prod_{\mu=1}^m dw_\mu e^{-w_\mu^2/2}$ is the m-dimensional standard Gaussian distribution and $\mathcal{Z}_t(\boldsymbol{Y}_t, \boldsymbol{Y}'_t, \boldsymbol{\Phi}, \boldsymbol{V})$ is the appropriate normalization (or *partition function*):

$$\mathcal{Z}_t(\boldsymbol{Y}_t, \boldsymbol{Y}'_t, \boldsymbol{\Phi}, \boldsymbol{V}) \stackrel{\text{def}}{=} \int dP_0^{\otimes n}(\boldsymbol{x})\mathcal{D}\boldsymbol{w}\, e^{-H_t(\boldsymbol{x},\boldsymbol{w};\boldsymbol{Y}_t,\boldsymbol{Y}'_t,\boldsymbol{\Phi},\boldsymbol{V})}\,. \tag{7.6.7}$$

Finally the *interpolating free energy* is

$$f_n(t) \stackrel{\text{def}}{=} \frac{1}{n}\mathbb{E}\log \mathcal{Z}_t(\boldsymbol{Y}_t, \boldsymbol{Y}'_t, \boldsymbol{\Phi}, \boldsymbol{V})\,. \tag{7.6.8}$$

Notice that:

$$\left\{ \begin{array}{rcl} f_n(0) & = & F_n \\ f_n(1) & = & \psi_{P_0}(r(1)) - \frac{1}{2}(1+\rho r(1)) + \frac{m}{n}\Psi_{P_{\text{out}}}(q(1)) \end{array} \right. \tag{7.6.9}$$

As discussed above, part of the potential (7.3.6) appears in $f_n(1)$. We would like to relate $F_n = f_n(0)$ to $f_n(1)$. We thus compute the derivative of the free energy along the interpolation path (see Appendix 7.6.4 for the proof):

Proposition 7.6.1

Assume that $\|q'\|_\infty, \|r'\|_\infty = O(1)$ *as* $n \to \infty$. *For all* $t \in (0,1)$

$$f_n'(t) = -\frac{1}{2}\mathbb{E}\Big\langle \Big(\frac{1}{n}\sum_{\mu=1}^{m} u'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu}) - r'(t)\Big)\Big(Q - q'(t)\Big)\Big\rangle_{n,t} + \frac{r'(t)}{2}(q'(t)-\rho) + o_n(1)\,,$$

where $o_n(1)$ *is a quantity that goes to* 0 *as* $n \to \infty$, *uniformly in* $t \in (0,1)$. *Recall that* $s_{t,\mu} = s_{t,\mu}(\boldsymbol{x}, w_\mu)$ *is given by* (7.6.4) *where* $(\boldsymbol{x}, \boldsymbol{w})$ *is a sample from the posterior distribution of* $(\boldsymbol{X}^*, \boldsymbol{W}^*)$ *given* $\boldsymbol{Y}_t, \boldsymbol{Y}'_t, \boldsymbol{\Phi}, \boldsymbol{V}$, *independently of everything else. The expectation with respect to* $(\boldsymbol{x}, \boldsymbol{w})$ *only is denoted by the Gibbs bracket* $\langle\cdot\rangle_{n,t}$, *see* (7.6.6). *Finally the overlap* Q *is*

$$Q \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_i^* x_i. \tag{7.6.10}$$

7.6.2 Overlap concentration

Proposition 7.6.1 is the analog of Lemma 4.3.1 that we have seen for the non-symmetric matrix estimation model in Chapter 4. As in the proof of Theorem 4.1.1 we will need to show that the overlap Q concentrates around its mean, in order to be able to cancel the first term of $f_n'(t)$ (in Proposition 7.6.1) by choosing $q'(t) = \mathbb{E}\langle Q\rangle_{n,t}$.

As seen in Section 2.2 Q concentrates around its mean on average over small perturbations of our model and we need in principle to introduce these small perturbations as in the proof of Theorem 4.1.1. However, in order to make the proof easier to read, we will assume here that for all choice of the functions r, q we have:

$$\int_0^1 \mathbb{E}\Big\langle \Big(Q - \mathbb{E}\langle Q\rangle_{n,t}\Big)^2\Big\rangle_{n,t} dt \xrightarrow[n\to\infty]{} 0. \tag{7.6.11}$$

The reader is invited to refer to Proposition 4.3.2 or to the paper [19] for a precise execution of the perturbation arguments.

Proposition 7.6.2

Assume that (7.6.11) *holds. Assume that we have* $q'(t) = \mathbb{E}\langle Q\rangle_{n,t}$ *for all* $t \in [0,1]$, *and that* $\|r'\|_\infty = O(1)$ *as* $n \to \infty$. *Then*

$$F_n = \psi_{P_0}(r(1)) + \alpha\Psi_{P_{\text{out}}}(q(1)) - \frac{1}{2}\int_0^1 q'(t)r'(t)dt + o_n(1)\,, \tag{7.6.12}$$

where $o_n(1)$ *denotes a quantity that goes to* 0 *as* $n \to \infty$.

Proof. By the Cauchy-Schwarz inequality

$$\Big(\int_0^1 \mathbb{E}\Big\langle\Big(\frac{1}{n}\sum_{\mu=1}^m u'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu}) - r'(t)\Big)(Q-q'(t))\Big\rangle_{n,t}dt\Big)^2 \tag{7.6.13}$$
$$\leq \int_0^1 \mathbb{E}\Big\langle\Big(\frac{1}{n}\sum_{\mu=1}^m u'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu}) - r'(t)\Big)^2\Big\rangle_{n,t}dt \times \int_0^1 \mathbb{E}\big\langle(Q-q'(t))^2\big\rangle_{n,t}dt.$$

Under assumptions (H2)-(H4) the first term of this product is bounded by some constant C that only depend on φ, α and $\|r'\|_\infty$. The second term goes to 0 as $n \to \infty$ by (7.6.11), since we assumed that for all $t \in [0,1]$, $q'(t) = \mathbb{E}\langle Q\rangle_{n,t}$. Consequently (7.6.13) goes to 0 as $n \to \infty$. Therefore from Proposition 7.6.1:

$$f_n(1) - f_n(0) = \int_0^1 f'_n(t)dt = \int_0^1 (q'(t)r'(t) - r'(t)\rho)dt + o_n(1)\,. \tag{7.6.14}$$

When combining (7.6.14) with (7.6.9) we reach the claimed identity. □

7.6.3 Lower and upper matching bounds

We now possess all the necessary tools to prove Theorem 7.3.1 with the following matching lower- and upper-bounds and Propositions C.7 and C.8 in Appendix C.3. Let us start with the lower bound.

Proposition 7.6.3

$$\liminf_{n\to\infty} F_n \geq \sup_{r\geq 0}\inf_{q\in[0,\rho]} \mathcal{F}(q,r). \tag{7.6.15}$$

Proof. Let us fix $r \geq 0$ and let us choose $r(t) = rt$ and q to be the solution of the following order-1 differential equation ($\mathbb{E}\langle Q\rangle_{n,t}$ is indeed a function of $q(t)$):

$$q(0) = 1 \quad \text{and} \quad \forall t \in [0,1],\ q'(t) = \mathbb{E}\langle Q\rangle_{n,t}.$$

We can then apply Proposition 7.6.2 (since $\|q'\|_\infty \leq \rho$ and $\|r'\|_\infty = r$ are both $O_n(1)$) to get

$$F_n = \psi_{P_0}(r) + \alpha\Psi_{P_{\text{out}}}(q(1)) - \frac{r}{2}q(1) + o_n(1) \geq \inf_{q\in[0,\rho]} \mathcal{F}(q,r) + o_n(1)$$

and thus $\liminf_{n\to\infty} F_n \geq \inf_{q\in[0,\rho]} \mathcal{F}(q,r)$. This is true for all $r \geq 0$ so we get Proposition 7.6.3. □

Proposition 7.6.4

$$\limsup_{n\to\infty} F_n \leq \sup_{r\geq 0}\inf_{q\in[0,\rho]} \mathcal{F}(q,r)\,. \tag{7.6.16}$$

Proof. We now chose (q,r) to be the solution of the following order-1 system of differential equations:

$$\begin{cases} q(0) = 0 \\ r(0) = 0 \end{cases} \quad \text{and} \quad \forall\, t \in [0,1], \quad \begin{cases} q'(t) = \mathbb{E}\langle Q\rangle_{n,t} \\ r'(t) = 2\alpha\Psi'_{P_{\text{out}}}(\mathbb{E}\langle Q\rangle_{n,t}). \end{cases} \tag{7.6.17}$$

As before, $\|q'\|_\infty \le \rho$. By Proposition 7.8.1 the function $\Psi'_{P_{\rm out}}$ is $\mathcal{C}^1$ and bounded, so $\|r'\|_\infty = O(1)$. We apply Proposition 7.6.2 to get

$$\begin{aligned} F_n &= \psi_{P_0}(r(1)) + \alpha\Psi_{P_{\rm out}}(q(1)) - \frac{1}{2}\int_0^1 q'(t)r'(t)dt + o_n(1) \\ &= \psi_{P_0}\Big(\int_0^1 r'(t)dt\Big) + \alpha\Psi_{P_{\rm out}}\Big(\int_0^1 q'(t)dt\Big) - \frac{1}{2}\int_0^1 q'(t)r'(t)dt + o_n(1) \\ &\le \int_0^1 \Big(\psi_{P_0}(r'(t)) + \alpha\Psi_{P_{\rm out}}(q'(t)) - \frac{1}{2}q'(t)r'(t)\Big)dt + o_n(1) \end{aligned} \tag{7.6.18}$$

by Jensen's inequality, because by Propositions 1.3.3 and 7.8.1 the functions ψ_{P_0} and $\Psi_{P_{\rm out}}$ are convex. By definition of r' and q', we have for all $t \in [0,1]$,

$$r'(t) = 2\alpha\Psi'_{P_{\rm out}}(q'(t)). \tag{7.6.19}$$

Therefore, by convexity of $\Psi_{P_{\rm out}}$, we have for all $t \in [0,1]$

$$\alpha\Psi_{P_{\rm out}}(q'(t)) - \frac{1}{2}q'(t)r'(t) = \inf_{q\in[0,\rho]}\{\alpha\Psi_{P_{\rm out}}(q) - \frac{1}{2}q\,r'(t)\}.$$

Plugging this back in (7.6.18), we get:

$$F_n \le \sup_{r\ge 0}\inf_{q\in[0,\rho]}\Big\{\psi_{P_0}(r) + \alpha\Psi_{P_{\rm out}}(q) - \frac{1}{2}q\,r\Big\} + o_n(1),$$

which proves Proposition 7.6.4. □

7.6.4 Derivative of the interpolating free energy: Proof of Proposition 7.6.1

Recall that $u'_y(x)$ is the x-derivative of $u_y(x) = \log P_{\rm out}(y|x)$. Moreover denote $P'_{\rm out}(y|x)$ and $P''_{\rm out}(y|x)$ the first and second x-derivatives of $P_{\rm out}(y|x)$. We will first prove that for all $t \in (0,1)$

$$f_n'(t) = -\frac{1}{2}\mathbb{E}\Big\langle\Big(\frac{1}{n}\sum_{\mu=1}^m u'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu}) - r'(t)\Big)\Big(Q - q'(t)\Big)\Big\rangle_{n,t} + \frac{r'(t)}{2}(q'(t)-\rho) - \frac{A_n}{2}\,, \tag{7.6.20}$$

where recall $Q = \frac{1}{n}\sum_{i=1}^n X_i^* x_i$ and

$$A_n \overset{\rm def}{=} \mathbb{E}\Big[\frac{1}{\sqrt{n}}\sum_{\mu=1}^m \frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n\Big((X_i^*)^2 - \rho\Big)\Big)\frac{1}{n}\log\mathcal{Z}_t\Big]\,. \tag{7.6.21}$$

Once this is done, we will prove that A_n goes to 0 as $n \to \infty$ uniformly in $t \in [0,1]$, in order to obtain Proposition 7.6.1.

Proof of (7.6.20)

Recall definition (7.6.8) which becomes, when written as a function of the interpolating Hamiltonian (7.6.5),

$$f_n(t) = \frac{1}{n}\mathbb{E}\int d\boldsymbol{y}d\boldsymbol{y}'e^{-H_t(\boldsymbol{X}^*,\boldsymbol{W}^*;\boldsymbol{y},\boldsymbol{y}',\boldsymbol{\Phi},\boldsymbol{V})}\log\int dP_0^{\otimes n}(\boldsymbol{x})\mathcal{D}\boldsymbol{w}\,e^{-H_t(\boldsymbol{x},\boldsymbol{w};\boldsymbol{y},\boldsymbol{y}',\boldsymbol{\Phi},\boldsymbol{V})}\,. \tag{7.6.22}$$

We will need the Hamiltonian t-derivative H_t' given by

$$H_t'(\boldsymbol{x},\boldsymbol{w};\boldsymbol{y},\boldsymbol{y}',\boldsymbol{\Phi},\boldsymbol{V}) = -\sum_{\mu=1}^{m}\frac{ds_{t,\mu}}{dt}u'_{y_\mu}(s_{t,\mu}) - \frac{r'(t)}{2\sqrt{r(t)}}\sum_{i=1}^{n}x_i(y_i' - \sqrt{r(t)}x_i)\,. \tag{7.6.23}$$

where we recall (7.6.4):

$$s_{t,\mu} = s_{t,\mu}(\boldsymbol{x},w_\mu) \stackrel{\text{def}}{=} \sqrt{\frac{1-t}{n}}[\boldsymbol{\Phi}\boldsymbol{x}]_\mu + \sqrt{q(t)}\,V_\mu + \sqrt{\rho t - q(t)}\,w_\mu\,. \tag{7.6.24}$$

The derivative of the interpolating free energy thus reads, for $0<t<1$,

$$f_n'(t) = -\underbrace{\frac{1}{n}\mathbb{E}\Big[H_t'(\boldsymbol{X}^*,\boldsymbol{W}^*;\boldsymbol{Y}_t,\boldsymbol{Y}_t',\boldsymbol{\Phi},\boldsymbol{V})\log\mathcal{Z}_t\Big]}_{T_1} - \underbrace{\frac{1}{n}\mathbb{E}\Big\langle H_t'(\boldsymbol{x},\boldsymbol{w};\boldsymbol{Y}_t,\boldsymbol{Y}_t',\boldsymbol{\Phi},\boldsymbol{V})\Big\rangle_{n,t}}_{T_2} \tag{7.6.25}$$

where recall that $\mathcal{Z}_t = \mathcal{Z}_t(\boldsymbol{Y}_t,\boldsymbol{Y}_t',\boldsymbol{\Phi},\boldsymbol{V})$ is given by (7.6.7). Let us compute T_1. Let $1\le\mu\le m$. Let us start with the following term

$$\mathbb{E}\Big[\frac{dS_{t,\mu}}{dt}u'_{Y_{t,\mu}}(S_{t,\mu})\log\mathcal{Z}_t\Big] = \frac{1}{2}\mathbb{E}\Big[\Big(\frac{-[\boldsymbol{\Phi}\boldsymbol{X}^*]_\mu}{\sqrt{n(1-t)}} + \frac{q'(t)}{\sqrt{q(t)}}V_\mu + \frac{\rho-q'(t)}{\sqrt{\rho t-q(t)}}W_\mu^*\Big)u'_{Y_{t,\mu}}(S_{t,\mu})\log\mathcal{Z}_t\Big]\,. \tag{7.6.26}$$

Let us compute the first term of the right-hand side of the last identity. By Gaussian integration by parts w.r.t $\Phi_{\mu i}$ we obtain

$$\begin{aligned}
&\frac{1}{\sqrt{n(1-t)}}\mathbb{E}\Big[[\boldsymbol{\Phi}\boldsymbol{X}^*]_\mu u'_{Y_{t,\mu}}(S_{t,\mu})\log\mathcal{Z}_t\Big]\\
&= \frac{1}{\sqrt{n(1-t)}}\sum_{i=1}^{n}\mathbb{E}\Big[\int d\boldsymbol{y}d\boldsymbol{y}'e^{-H_t(\boldsymbol{X}^*,\boldsymbol{W}^*;\boldsymbol{y},\boldsymbol{y}',\boldsymbol{\Phi},\boldsymbol{V})}\Phi_{\mu i}X_i^*u'_{Y_{t,\mu}}(S_{t,\mu})\log\mathcal{Z}_t(\boldsymbol{y},\boldsymbol{y}',\boldsymbol{\Phi},\boldsymbol{V})\Big]\\
&= \frac{1}{n}\sum_{i=1}^{n}\Big(\mathbb{E}\Big[(X_i^*)^2\Big(u''_{Y_{t,\mu}}(S_{t,\mu}) + u'_{Y_{t,\mu}}(S_{t,\mu})^2\Big)\log\mathcal{Z}_t\Big] + \mathbb{E}\Big\langle X_i^*x_iu'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu})\Big\rangle_{n,t}\Big)\\
&= \mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n}(X_i^*)^2\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\log\mathcal{Z}_t\Big] + \mathbb{E}\Big\langle\frac{1}{n}\sum_{i=1}^{n}X_i^*x_iu'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu})\Big\rangle_{n,t}\,,
\end{aligned} \tag{7.6.27}$$

where we used the identity

$$u''_{Y_{t,\mu}}(x) + u'_{Y_{t,\mu}}(x)^2 = \frac{P''_{\rm out}(Y_{t,\mu}|x)}{P_{\rm out}(Y_{t,\mu}|x)}\,.$$

We now compute the second term of the right hand side of (7.6.26). Using again Gaussian integrations by parts but this time w.r.t $V_\mu, W_\mu^* \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ as well as the previous formula, we obtain similarly

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\frac{q'(t)}{\sqrt{q(t)}}V_\mu + \frac{\rho-q'(t)}{\sqrt{\rho t-q(t)}}W_\mu^*\Big)u'_{Y_{t,\mu}}(S_{t,\mu})\log\mathcal{Z}_t\Big]\\
&= \mathbb{E}\Big[\int d\boldsymbol{y}d\boldsymbol{y}'e^{-H_t(\boldsymbol{X}^*,\boldsymbol{W}^*;\boldsymbol{y},\boldsymbol{y}',\boldsymbol{\Phi},\boldsymbol{V})}\Big(\frac{q'(t)}{\sqrt{q(t)}}V_\mu + \frac{\rho-q'(t)}{\sqrt{\rho t-q(t)}}W_\mu^*\Big)u'_{Y_{t,\mu}}(S_{t,\mu})\log\mathcal{Z}_t\Big]\\
&= \mathbb{E}\Big[\rho\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\log\mathcal{Z}_t\Big] + \mathbb{E}\Big\langle q(t)u'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu})\Big\rangle_{n,t}\,.
\end{aligned} \tag{7.6.28}$$

Combining equations (7.6.26), (7.6.27) and (7.6.28) together, we have

$$
\begin{aligned}
&-\mathbb{E}\Big[\frac{dS_{t,\mu}}{dt}u'_{Y_{t,\mu}}(S_{t,\mu})\log\mathcal{Z}_t\Big]\\
&=\frac{1}{2}\mathbb{E}\Big[\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big(\frac{1}{n}\sum_{i=1}^n(X_i^*)^2-\rho\Big)\log\mathcal{Z}_t\Big]+\frac{1}{2}\mathbb{E}\Big\langle\Big(\frac{1}{n}\sum_{i=1}^n X_i^*x_i-q'(t)\Big)u'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu})\Big\rangle_{n,t}.
\end{aligned}
$$

As seen from (7.6.23), (7.6.25) it remains to compute $\mathbb{E}[X_j^*(Y'_{t,j}-\sqrt{r(t)}X_j^*)\log\mathcal{Z}_t]$. Recalling that for $1\le j\le n$, $Y'_{t,j}-\sqrt{r(t)}X_j^*=Z'_j$ and then using again a Gaussian integration by parts w.r.t $Z'_j\sim\mathcal{N}(0,1)$ we obtain

$$
\begin{aligned}
\mathbb{E}\Big[X_j^*(Y'_{t,j}-\sqrt{r(t)}X_j^*)\log\mathcal{Z}_t\Big]&=\mathbb{E}\Big[X_j^*Z'_j\log\mathcal{Z}_t\Big]=\mathbb{E}\Big[X_j^*\partial_{Z'_j}\log\mathcal{Z}_t\Big]\\
&=-\mathbb{E}\Big[X_j^*\Big\langle\sqrt{r(t)}(X_j^*-x_j)+Z'_j\Big\rangle_{n,t}\Big]\\
&=-\sqrt{r(t)}\Big(\rho-\mathbb{E}\langle X_j^*x_j\rangle_{n,t}\Big).
\end{aligned}
$$

Thus, by taking the sum,

$$
-\frac{r'(t)}{2\sqrt{r(t)}}\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^n X_i^*(Y'_{t,i}-\sqrt{r(t)}X_i^*)\log\mathcal{Z}_t\Big]=\frac{r'(t)\rho}{2}-\frac{r'(t)}{2}\mathbb{E}\Big\langle\frac{1}{n}\sum_{i=1}^n X_i^*x_i\Big\rangle_{n,t}. \tag{7.6.29}
$$

Therefore, for all $t\in(0,1)$,

$$
T_1=\frac{1}{2}\mathbb{E}\Big[\frac{1}{\sqrt{n}}\sum_{\mu=1}^m\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n((X_i^*)^2-\rho)\Big)\frac{1}{n}\log\mathcal{Z}_t\Big]+\frac{r'(t)\rho}{2}-\frac{r'(t)q'(t)}{2} \tag{7.6.30}
$$

$$
+\frac{1}{2}\mathbb{E}\Big\langle\Big(\frac{1}{n}\sum_{\mu=1}^m u'_{Y_{t,\mu}}(S_{t,\mu})u'_{Y_{t,\mu}}(s_{t,\mu})-r'(t)\Big)\Big(\frac{1}{n}\sum_{i=1}^n X_i^*x_i-q'(t)\Big)\Big\rangle_{n,t}. \tag{7.6.31}
$$

To obtain (7.6.20), it remains to show that $T_2=0$. This is a direct consequence of the Nishimori identity (Proposition 1.1.1):

$$
T_2=\frac{1}{n}\mathbb{E}\Big\langle H'_t(\boldsymbol{x},\boldsymbol{w};\boldsymbol{Y}_t,\boldsymbol{Y}'_t,\boldsymbol{\Phi},\boldsymbol{V})\Big\rangle_{n,t}=\frac{1}{n}\mathbb{E}\Big[H'_t(\boldsymbol{X}^*,\boldsymbol{W}^*;\boldsymbol{Y}_t,\boldsymbol{Y}'_t,\boldsymbol{\Phi},\boldsymbol{V})\Big]=0. \tag{7.6.32}
$$

For obtaining the Lemma, it remains to show that A_n goes to 0 uniformly in $t\in[0,1]$.

Proof that A_n vanishes as $n\to\infty$

We now consider the final step, that is showing that A_n given by (7.6.21) vanishes in the $n\to\infty$ limit uniformly in $t\in[0,1]$ under conditions (H1)-(H2)-(H3). First we show that

$$
\mathbb{E}\Big[\frac{1}{\sqrt{n}}\sum_{\mu=1}^m\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n\Big((X_i^*)^2-\rho\Big)\Big)\Big]=0. \tag{7.6.33}
$$

Once this is done, we use the fact that $\frac{1}{n}\log\mathcal{Z}_t$ concentrates around $f_n(t)$ to prove that A_n converges to 0 as $n\to\infty$. We start by noticing the simple fact that for all $s\in\mathbb{R}$, $\int P''_{\rm out}(y|s)dy=0$. Consequently, for $\mu\in\{1,\dots,m\}$,

$$
\mathbb{E}\Big[\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\,\Big|\,\boldsymbol{X}^*,\boldsymbol{S}_t\Big]=\int dyP''_{\rm out}(y|S_{t,\mu})=0. \tag{7.6.34}
$$

Thus, using the "tower property" of the conditional expectation:

$$\mathbb{E}\Big[\sum_{i=1}^{n}((X_i^*)^2-\rho)\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big]=\mathbb{E}\Big[\sum_{i=1}^{n}\Big((X_i^*)^2-\rho\Big)\mathbb{E}\Big[\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\,\Big|\,\boldsymbol{X}^*,\boldsymbol{S}_t\Big]\Big]=0$$

which gives (7.6.33). We now show that A_n goes to 0 uniformly in $t\in[0,1]$ as $n\to\infty$. Using successively (7.6.33) and the Cauchy-Schwarz inequality, we have

$$|A_n|=\Big|\mathbb{E}\Big[\frac{1}{\sqrt{n}}\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big((X_i^*)^2-\rho\Big)\Big(\frac{1}{n}\log\mathcal{Z}_t-f_n(t)\Big)\Big]\Big| \tag{7.6.35}$$

$$\le\ \mathbb{E}\Big[\Big(\frac{1}{\sqrt{n}}\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big)^2\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big((X_i^*)^2-\rho\Big)\Big)^2\Big]^{1/2}\mathbb{E}\Big[\Big(\frac{1}{n}\log\mathcal{Z}_t-f_n(t)\Big)^2\Big]^{1/2}. \tag{7.6.36}$$

Using again the "tower property" of conditional expectations

$$\mathbb{E}\Big[\Big(\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big)^2\Big(\sum_{i=1}^{n}((X_i^*)^2-\rho)\Big)^2\Big] \tag{7.6.37}$$

$$=\mathbb{E}\Big[\Big(\sum_{i=1}^{n}\Big((X_i^*)^2-\rho\Big)\Big)^2\mathbb{E}\Big[\Big(\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big)^2\,\Big|\,\boldsymbol{X}^*,\boldsymbol{S}_t\Big]\Big]. \tag{7.6.38}$$

Now, using the fact that conditionally on $\boldsymbol{S}_t$, the random variables $\Big(\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big)_{1\le\mu\le m}$ are i.i.d. and centered, we have

$$\mathbb{E}\Big[\Big(\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big)^2\,\Big|\,\boldsymbol{X}^*,\boldsymbol{S}_t\Big]=m\mathbb{E}\Big[\Big(\frac{P''_{\rm out}(Y_1|S_{t,1})}{P_{\rm out}(Y_1|S_{t,1})}\Big)^2\,\Big|\,\boldsymbol{S}_t\Big]. \tag{7.6.39}$$

Under condition (H2), it is not difficult to show that there exists a constant $C>0$ such that

$$\mathbb{E}\Big[\Big(\frac{P''_{\rm out}(Y_{t,1}|S_{t,1})}{P_{\rm out}(Y_{t,1}|S_{t,1})}\Big)^2\,\Big|\,\boldsymbol{S}_t\Big]\le C\,. \tag{7.6.40}$$

Combining now (7.6.40), (7.6.39) and (7.6.38) we obtain that

$$\mathbb{E}\Big[\Big(\sum_{\mu=1}^{m}\frac{P''_{\rm out}(Y_{t,\mu}|S_{t,\mu})}{P_{\rm out}(Y_{t,\mu}|S_{t,\mu})}\Big)^2\Big(\sum_{i=1}^{n}\Big((X_i^*)^2-\rho\Big)\Big)^2\Big]\le mC\,\mathbb{E}\Big[\Big(\sum_{i=1}^{n}\Big((X_i^*)^2-\rho\Big)\Big)^2\Big]$$
$$=mnC\,\mathrm{Var}\Big((X_1^*)^2\Big).$$

Going back to (7.6.36), therefore there exists a constant $C'>0$ such that

$$|A_n|\le\ C'\,\mathbb{E}\Big[\Big(\frac{1}{n}\log\mathcal{Z}_t-f_n(t)\Big)^2\Big]^{1/2}. \tag{7.6.41}$$

Using standard concentration arguments (bounded difference property, Gaussian Poincaré inequality) it is not difficult to show that under assumptions (H1), (H2), (H3) and (H4) there exists a constant $C(\varphi,S,\alpha,\Delta,R)\ge 0$ such that for all functions q,r bounded by R we have

$$\mathrm{Var}\Big(\frac{1}{n}\log\mathcal{Z}_t\Big)=\mathbb{E}\Big[\Big(\frac{1}{n}\log\mathcal{Z}_t-\frac{1}{n}\mathbb{E}\log\mathcal{Z}_t\Big)^2\Big]\le\frac{C(\varphi,S,\alpha,\Delta,R)}{n}\,. \tag{7.6.42}$$

Consequently, $\mathbb{E}[(\frac{1}{n}\log\mathcal{Z}_t-f_n(t))^2]\to 0$ as $n\to\infty$ uniformly in $t\in[0,1]$. Thus A_n goes to 0 as $n\to\infty$ uniformly in $t\in[0,1]$. This ends the proof of Proposition 7.6.1.

7.7 Proofs of the limits of optimal errors

7.7.1 Unicity of the optimizer q^* of the replica formula: Proof of Proposition 7.3.1

Since ψ_{P_0} and $\Psi_{P_{\text{out}}}$ are related to I_{P_0} and $\mathcal{I}_{P_{\text{out}}}$ by the relations (7.3.2)-(7.3.5), the optimizers of (7.3.8) and (7.3.9) are the same. We chose here to work with (7.3.9). The function

$$h : \alpha \mapsto \inf_{q\in[0,\rho]} \left\{ \alpha \mathcal{I}_{P_{\text{out}}}(q) + \sup_{r\geq 0} \left\{ I_{P_0}(r) - \frac{r}{2}(\rho - q) \right\} \right\} \tag{7.7.1}$$

is concave (as an infimum of linear functions). Proposition D.2 gives that h is differentiable at α if and only if

$$\left\{ \mathcal{I}_{P_{\text{out}}}(q) \,\middle|\, q \text{ minimizer of (7.7.1)} \right\}$$

is a singleton. We assumed that P_{out} is informative, so Proposition 7.8.4 gives that $\mathcal{I}_{P_{\text{out}}}$ is strictly decreasing. We obtain thus that the set of points at which h is differentiable is exactly D^*. Since h is concave, D^* is equal to $\mathbb{R}_{>0}$ minus a countable set. Proposition D.2 gives also that $h'(\alpha) = \mathcal{I}_{P_{\text{out}}}(q^*(\alpha))$, for all $\alpha \in D^*$. The function h is concave, so its derivative h' is non-increasing. Since $\mathcal{I}_{P_{\text{out}}}$ is strictly decreasing, we obtain that $\alpha \in D^* \mapsto q^*(\alpha)$ is non-decreasing.

Let now $\alpha_0 \in D^*$. By concavity of h, $h'(\alpha) \to h'(\alpha_0)$ when $\alpha \in D^* \to \alpha_0$. Therefore:

$$\mathcal{I}_{P_{\text{out}}}(q^*(\alpha)) \xrightarrow[\alpha\in D\to\alpha_0]{} \mathcal{I}_{P_{\text{out}}}(q^*(\alpha_0))$$

which implies $q^*(\alpha) \to q^*(\alpha_0)$ by strict monotonicity of $\mathcal{I}_{P_{\text{out}}}$. The map $\alpha \mapsto q^*(\alpha)$ is therefore continuous on D^*.

7.7.2 Limit of the overlap: Proof of Theorem 7.3.2

Let Q_n denote the overlap $Q_n \overset{\text{def}}{=} \frac{1}{n}\langle \boldsymbol{X}^*, \boldsymbol{x}\rangle$ between $\boldsymbol{X}^*$ and $\boldsymbol{x} = (x_1, \dots, x_n)$ a sample from the posterior distribution $P(\boldsymbol{X}^* | \boldsymbol{Y}, \boldsymbol{\Phi})$, independently of everything else. In this section we will show that $|Q_n|$ converges in probability to $q^*(\alpha)$, for all $\alpha \in D^*$.

Upper bound on the overlap

Proposition 7.7.1

For all $\alpha \in D^*$ and for all $\epsilon > 0$,

$$\mathbb{P}\Big(|Q_n| \geq q^*(\alpha) + \epsilon \Big) \xrightarrow[n\to\infty]{} 0\,.$$

Let us fix $\alpha \in D^*$ and let $p \geq 1$. In order to obtain an upper bound on the overlap, we consider an observation model with some (small) extra information (that takes the form of a tensor of order $2p$) in addition of the original model (7.2.1), i.e. we observe

$$\begin{cases} \boldsymbol{Y} \sim P_{\text{out}}(\cdot \,|\, \boldsymbol{\Phi}\boldsymbol{X}^*/\sqrt{n})\,, \\ \boldsymbol{Y}' = \sqrt{\frac{\lambda}{n^{2p-1}}}(\boldsymbol{X}^*)^{\otimes 2p} + \boldsymbol{Z}'\,, \end{cases} \tag{7.7.2}$$

where $\lambda \geq 0$, $\boldsymbol{Z}' = (Z'_{i_1 \dots i_{2p}})_{1 \leq i_1, \dots, i_{2p} \leq n} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and $(\boldsymbol{X}^*)^{\otimes 2p} = (X_{i_1} \dots X_{i_{2p}})_{1 \leq i_1, \dots, i_{2p} \leq n}$. The next Proposition gives the limit of the mutual information between the signal and the observations (7.7.2).

Proposition 7.7.2

For all $\lambda \geq 0$,

$$\lim_{n \to \infty} \frac{1}{n} I\Big(\boldsymbol{X}^*; \boldsymbol{Y}, \boldsymbol{Y}' \Big| \boldsymbol{\Phi}\Big) = I(\lambda)\,, \tag{7.7.3}$$

where the right-hand-side is

$$I(\lambda) \overset{\text{def}}{=} \inf_{q \in [0,\rho]} \sup_{r \geq 0} \Big\{ I_{P_0}(r + 2p\lambda q^{2p-1}) + \alpha \mathcal{I}_{P_{\text{out}}}(q) - \frac{r}{2}(\rho - q) + \frac{2p-1}{2}\lambda q^{2p} - \rho p \lambda q^{2p-1} + \frac{\lambda}{2}\rho^{2p} \Big\}. \tag{7.7.4}$$

Proposition 7.7.2 is proved using the same strategy than for Proposition 4.4.1, so we omit its proof for brevity. We are now in position to prove Proposition 7.7.1.

Proof of Proposition 7.7.1. The chain rule for the mutual information gives

$$I(\boldsymbol{X}^*; \boldsymbol{Y}, \boldsymbol{Y}' | \boldsymbol{\Phi}) = I(\boldsymbol{X}^*; \boldsymbol{Y} | \boldsymbol{\Phi}) + I((\boldsymbol{X}^*)^{\otimes 2p}; \boldsymbol{Y}' | \boldsymbol{Y}, \boldsymbol{\Phi}).$$

Thus, using the I-MMSE relation of Proposition 1.3.3,

$$\frac{1}{n}\frac{\partial}{\partial \lambda} I(\boldsymbol{X}^*; \boldsymbol{Y}, \boldsymbol{Y}' | \boldsymbol{\Phi})\Big|_{\lambda = 0^+} = \frac{1}{n}\frac{\partial}{\partial \lambda} I((\boldsymbol{X}^*)^{\otimes 2p}; \boldsymbol{Y}' | \boldsymbol{Y}, \boldsymbol{\Phi})\Big|_{\lambda = 0^+} = \frac{1}{2n^{2p}} \text{MMSE}((\boldsymbol{X}^*)^{\otimes 2p} | \boldsymbol{Y}, \boldsymbol{\Phi})\,.$$

By concavity of $\lambda \mapsto I(\boldsymbol{X}^*; \boldsymbol{Y}, \boldsymbol{Y}' | \boldsymbol{\Phi})$, we have using Proposition 7.7.2 and Proposition C.1:

$$\liminf_{n \to \infty} \frac{1}{2n^{2p}} \text{MMSE}((\boldsymbol{X}^*)^{\otimes 2p} | \boldsymbol{Y}, \boldsymbol{\Phi}) = \liminf_{n \to \infty} \frac{1}{n}\frac{\partial}{\partial \lambda} I(\boldsymbol{X}^*; \boldsymbol{Y}, \boldsymbol{Y}' | \boldsymbol{\Phi})\Big|_{\lambda = 0^+} \geq I'(0^+)\,,$$

For $\alpha \in D^*$ and $\lambda = 0$ the infimum of (7.7.4) is achieved uniquely at $q = q^*(\alpha)$, by definition of D^*. Consequently, by Proposition D.2 we obtain $I(0^+) = \frac{1}{2}(\rho^{2p} - q_*(\lambda)^{2p})$ and thus

$$\liminf_{n \to \infty} \frac{1}{n^{2p}} \text{MMSE}((\boldsymbol{X}^*)^{\otimes 2p} | \boldsymbol{Y}, \boldsymbol{\Phi}) \geq \rho^{2p} - q^*(\alpha)^{2p}\,.$$

One verifies easily that

$$\frac{1}{n^{2p}} \text{MMSE}((\boldsymbol{X}^*)^{\otimes 2p} | \boldsymbol{Y}, \boldsymbol{\Phi}) = \rho^{2p} - \mathbb{E}[Q_n^{2p}] + o_n(1)\,, \tag{7.7.5}$$

so we deduce that $\limsup_{n \to \infty} \mathbb{E}[Q_n^{2p}] \leq q^*(\alpha)^{2p}$. Let $\epsilon > 0$. By Markov's inequality we have

$$\mathbb{P}(|Q_n| \geq q_*(\alpha) + \epsilon) \leq \frac{\mathbb{E}[Q_n^{2p}]}{(q_*(\alpha) + \epsilon)^{2p}}\,.$$

By taking the lim sup in n on both sides we obtain

$$\limsup_{n \to \infty} \mathbb{P}(|Q_n| \geq q_*(\alpha) + \epsilon) \leq \frac{q^*(\alpha)^{2p}}{(q_*(\alpha) + \epsilon)^{2p}}\,,$$

and Proposition 7.7.1 follows by taking the $p \to \infty$ limit in the inequality above. □

Limit of the overlap

Let us fix $\alpha \in D^*$. The sequence of the overlaps $(Q_n)_{n\geq 1}$ is tight (because bounded in L^1). By Prokhorov's Theorem we know that the sequence of the laws of $(Q_n)_{n\geq 1}$ is relatively compact. We can thus consider a subsequence along which it converges in law, to some random variable Q. In order to simplify the notations (and because working with an extraction does not change the proof) we will assume in the sequel that

$$Q_n \xrightarrow[n\to\infty]{(d)} Q\,,$$

for some random variable Q. We aim now at showing that $|Q| = q^*(\alpha)$ almost surely and moreover, if P_{out} is not symmetric, $Q = q^*(\alpha)$ almost surely.

Lemma 7.7.1 (*Upper bound on the overlap*)

$|Q| \leq q^*(\alpha)$ almost surely.

Proof. Let $\epsilon > 0$. The set $[0, q^*(\alpha)+\epsilon]$ is closed, so by Portemanteau's Theorem

$$\mathbb{P}(|Q| \leq q^*(\alpha)+\epsilon) \geq \limsup_{n\to\infty} \mathbb{P}(|Q_n| \leq q^*(\alpha)+\epsilon) = 1\,,$$

by Proposition 7.7.1. So $\mathbb{P}(|Q| \leq q^*(\alpha)+\epsilon) = 1$ for all $\epsilon > 0$ which gives $\mathbb{P}(|Q| \leq q^*(\alpha)) = 1$. □

In order to prove the converse lower bound, we will need to consider the following error quantity, for $f : \mathbb{R} \to \mathbb{R}$ a continuous bounded function:

$$\mathcal{E}_{f,n}(\alpha) \stackrel{\text{def}}{=} \text{MMSE}(f(Y_{\text{new}})|\boldsymbol{\Phi}_{\text{new}}, \boldsymbol{Y}, \boldsymbol{\Phi}) = \mathbb{E}\Big[\Big(f(Y_{\text{new}}) - \mathbb{E}[f(Y_{\text{new}})|\boldsymbol{\Phi}_{\text{new}}, \boldsymbol{Y}, \boldsymbol{\Phi}]\Big)^2\Big], \quad (7.7.6)$$

which is the minimum mean-square error on $f(Y_{\text{new}})$. We define also

$$\mathcal{E}_f(q) \stackrel{\text{def}}{=} \text{MMSE}(f(\widetilde{Y}_0)|V) = \mathbb{E}\Big[\Big(f(\widetilde{Y}_0) - \mathbb{E}[f(\widetilde{Y}_0)|V]\Big)^2\Big] \quad (7.7.7)$$

where $\widetilde{Y}_0 \sim P_{\text{out}}(\cdot\,|\,\sqrt{q}V + \sqrt{\rho - q}W^*)$ is the output of the second scalar channel (7.3.3).

Proposition 7.7.3

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous bounded function. Assume that P_{out} is informative and that (h1)-(h2)-(h3)-(h4) hold and that either (h5.a) or (h5.b) holds. Then for all $\alpha \in D^*$ we have

$$\mathcal{E}_{f,n}(\alpha) \xrightarrow[n\to\infty]{} \mathcal{E}_f(q^*(\alpha)) \quad (7.7.8)$$

where $q^*(\alpha)$ is the optimizer of the replica-symmetric formula (7.3.8), see Proposition 7.3.1.

We first deduce Theorem 7.3.4 from Proposition 7.7.3. Proposition 7.7.3 will then be proved at the end of this Section. Proposition 7.7.3 gives $\mathcal{E}_{f,n}(\alpha) \xrightarrow[n\to\infty]{} \mathcal{E}_f(q^*(\alpha))$. The function $\mathcal{E}_f$ can be written as

$$\mathcal{E}_f(q) = \frac{1}{2}\mathbb{E}\Big[h_f\Big(\sqrt{q}Z_0 + \sqrt{\rho - q}Z_1, \sqrt{q}Z_0 + \sqrt{\rho - q}Z_1'\Big)\Big]$$

where $Z_0, Z_1, Z_1' \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and $h_f : (a,b) \in \mathbb{R}^2 \mapsto \int (f(y_1) - f(y_2))^2 P_{\text{out}}(dy_1|a) P_{\text{out}}(dy_2|b)$.

The next lemma follows from a simple central-limit-type argument (and can for instance be proved using Lindeberg "swapping trick", see [45]).

Lemma 7.7.2

Let $k \geq 1$. Let $(s_n)_{n\geq 1}$ be a sequence of independent random variables that have zero mean, unit variance and finite third moment that is bounded with n. We write $\boldsymbol{s}_n = (s_1, \dots, s_n)$. For $n \geq 1$ let $\boldsymbol{x}_n^{(1)}, \dots, \boldsymbol{x}_n^{(k)} \in \mathbb{R}^n$ be some deterministic vectors. Assume that for all $i \in \{1, \dots, k\}$, $\left(\frac{1}{n}\sum_j (\boldsymbol{x}_{n,j}^{(i)})^3\right)_{n\geq 1}$ is bounded and that for all $i, j \in \{1, \dots, k\}$

$$\frac{1}{n}\langle \boldsymbol{x}_n^{(i)}, \boldsymbol{x}_n^{(j)}\rangle \xrightarrow[n\to\infty]{} R_{i,j},$$

for some matrix $R \in \mathbb{R}^{k\times k}$. Then

$$\left(\frac{1}{\sqrt{n}}\langle \boldsymbol{s}_n, \boldsymbol{x}_n^{(\ell)}\rangle\right)_{1\leq \ell\leq k} \xrightarrow[n\to\infty]{\text{(d)}} \mathcal{N}(0, R).$$

We deduce from Lemma 7.7.2 (that we apply for $k = 2$ conditionally on $\boldsymbol{x}, \boldsymbol{X}^*$):

$$\left(\frac{\langle \boldsymbol{x}, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}, \frac{\langle \boldsymbol{X}^*, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}\right) \xrightarrow[n\to\infty]{(d)} (Z_1, Z_2)\,, \tag{7.7.9}$$

where (Z_1, Z_2) is sampled, conditionally on Q, from $\mathcal{N}\Big(0, \left(\begin{smallmatrix}\rho & Q\\ Q & \rho\end{smallmatrix}\right)\Big)$.

Proposition 7.7.4

We have

$$\mathcal{E}_{f,n}(\alpha) \xrightarrow[n\to\infty]{} \frac{1}{2}\mathbb{E}\Big[h_f(Z_1, Z_2)\Big]\,,$$

where (Z_1, Z_2) is defined in Lemma 7.7.2 above.

Proof. We have

$$\begin{aligned}
\mathcal{E}_{f,n} &= \mathbb{E}\Big[\Big(f(Y_{\text{new}}) - \mathbb{E}[f(Y_{\text{new}})|\boldsymbol{\Phi}_{\text{new}}, \boldsymbol{\Phi}, \boldsymbol{Y}]\Big)^2\Big]\\
&= \frac{1}{2}\mathbb{E}\left[\int (f(y_{\text{new}}) - f(y))^2 P_{\text{out}}\big(dy_{\text{new}}|\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle/\sqrt{n}\big) P_{\text{out}}\big(dy|\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{x}\rangle/\sqrt{n}\big)\right]\\
&= \frac{1}{2}\mathbb{E}\left[h_f\Big(\frac{\langle \boldsymbol{x}, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}, \frac{\langle \boldsymbol{X}^*, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}\Big)\right]\,.
\end{aligned}$$

By (7.7.9) above, we have $(\frac{\langle \boldsymbol{x}, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}, \frac{\langle \boldsymbol{X}^*, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}) \xrightarrow[n\to\infty]{(d)} (Z_1, Z_2)$. Using (h4) (and Remark 7.3.1) we can find a Borel set $S \subset \mathbb{R}$ of full Lebesgue's measure such that h_f is continuous on $S \times S$. The set of discontinuity points of h_f has thus zero measure for the law of (Z_1, Z_2). Indeed if we condition on Q:

- if $|Q| < \rho$, then (Z_1, Z_2) has a density over $\mathbb{R}^2$.
- if $Q = \rho$, then $Z_1 = Z_2$ almost surely, but h_f is continuous on $\{(s, s)\,|\,s \in S\}$ that has full Lebesgue's measure on the diagonal $\{(x, x)\,|\,x \in \mathbb{R}\}$.
- if $Q = -\rho$, then $Z_1 = -Z_2$ almost surely and we use then similar arguments as for the previous point.

We have therefore:

$$h_f\Big(\frac{\langle \boldsymbol{x}, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}, \frac{\langle \boldsymbol{X}^*, \boldsymbol{\Phi}_{\text{new}}\rangle}{\sqrt{n}}\Big) \xrightarrow[n\to\infty]{(d)} h_f(Z_1, Z_2)\,,$$

and Lemma 7.7.4 follows from the fact that h_f is bounded. □

Let us now define:

$$H_f : \left| \begin{array}{ccc} [-\rho,\rho] & \to & \mathbb{R} \\ q & \mapsto & \frac{1}{2}\mathbb{E}\Big[h_f(G^{(q)})\Big] \end{array} \right. \tag{7.7.10}$$

where $G^{(q)} \sim \mathcal{N}\Big(0, \left(\begin{smallmatrix} \rho & q \\ q & \rho \end{smallmatrix}\right)\Big)$. Notice that H_f is equal to the function $\mathcal{E}_f$ on $[0,\rho]$. By Proposition 7.7.4 above and Proposition 7.7.3, we have:

$$H_f(q^*(\alpha)) = \lim_{n\to\infty} \mathcal{E}_{f,n}(\alpha) = \mathbb{E}\Big[H_f(Q)\Big]\,. \tag{7.7.11}$$

Lemma 7.7.3

For all $q \in (0,\rho]$, $H_f(-q) \geq H_f(q)$ with equality if and only if for almost-all $x \in \mathbb{R}$ we have

$$\int f(y) P_{\text{out}}(dy|x) = \int f(y) P_{\text{out}}(dy|-x).$$

Proof. For $x \in \mathbb{R}$ we let $\widetilde{P}_{\text{out}}(\cdot|x)$ denote the law of $f(Y)$ for $Y \sim P_{\text{out}}(\cdot|x)$. Let $q \in (0,\rho]$ and $Z_0, Z_1, Z_1' \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Let Y and Y' be two random variables, that are independent conditionally on Z_0, Z_1, Z_1' and distributed as:

$$Y \sim \widetilde{P}_{\text{out}}(\cdot|\sqrt{q}Z_0 + \sqrt{\rho-q}Z_1) \qquad \text{and} \qquad Y' \sim \widetilde{P}_{\text{out}}(\cdot|-\sqrt{q}Z_0 + \sqrt{\rho-q}Z_1')\,.$$

Y and Y' are independent conditionally on Z_0, so we get

$$\begin{aligned} H_f(-q) &= \frac{1}{2}\mathbb{E}[(Y-Y')^2] \\ &= \frac{1}{2}\mathbb{E}[(Y-\mathbb{E}[Y|Z_0])^2] + \frac{1}{2}\mathbb{E}[(Y'-\mathbb{E}[Y'|Z_0])^2] + \frac{1}{2}\mathbb{E}[(\mathbb{E}[Y|Z_0]-\mathbb{E}[Y'|Z_0])^2] \\ &= \mathbb{E}[(Y-\mathbb{E}[Y|Z_0])^2] + \frac{1}{2}\mathbb{E}[(\mathbb{E}[Y|Z_0]-\mathbb{E}[Y'|Z_0])^2] \\ &= H_f(q) + \frac{1}{2}\mathbb{E}[(\mathbb{E}[Y|Z_0]-\mathbb{E}[Y'|Z_0])^2]. \end{aligned}$$

We get that $H_f(-q) \geq H_f(q)$, with equality if and only if $\mathbb{E}[Y|Z_0] = \mathbb{E}[Y'|Z_0]$ almost surely. This is equivalent to

$$\mathbb{E}\int f(y)P_{\text{out}}(dy|v+\sqrt{\rho-q}Z) = \mathbb{E}\int f(y)P_{\text{out}}(dy|-v+\sqrt{\rho-q}Z), \qquad \text{for a.e. } v \in \mathbb{R},$$

where $Z \sim \mathcal{N}(0,1)$. Thus, if $q = \rho$ then we are done. If $q \in (0,\rho)$, writing

$$F(x) = \int f(y) P_{\text{out}}(dy|\sqrt{\rho-q}x),$$

and $g(x) = F(x) - F(-x)$, we get that $G(v) \overset{\text{def}}{=} \mathbb{E}[g(v+Z)] = 0$ for almost every $v \in \mathbb{R}$. We compute the derivative of G: $G'(v) = \mathbb{E}[Zg(v+Z)] = 0$. By Lemma 7.8.1, we get that g is equal to zero almost-everywhere, which concludes the proof. □

We have now all the tools needed to prove Theorem 7.3.2. Using Lemma 7.7.3 and (7.7.11) above, we get that $\mathbb{E}H_f(|Q|) \leq \mathbb{E}H_f(Q) = H_f(q^*(\alpha))$. Since H_f is equal to $\mathcal{E}_f$ on $[0,\rho]$ this gives

$$\mathbb{E}\Big[\mathcal{E}_f(|Q|)\Big] \leq \mathcal{E}_f(q^*(\alpha))\,. \tag{7.7.12}$$

If $q^*(\alpha) = 0$, then Theorem 7.3.2 follows simply from Proposition 7.7.1. We suppose now that $q^*(\alpha) > 0$ and consider $\epsilon \in (0, q^*(\alpha))$. We define $p(\epsilon) = \mathbb{P}\Big(|Q| \leq q^*(\alpha) - \epsilon\Big)$. We are going to show that $p(\epsilon) = 0$. We assumed that P_{out} is informative, so by Proposition 7.8.6 and Proposition 7.8.7 in Section 7.8, there exists a continuous bounded function $f : \mathbb{R} \mapsto \mathbb{R}$ such that $\mathcal{E}_f$ is strictly decreasing on $[0, \rho]$. In the following, f is assumed to be such a function. We have

$$\begin{aligned}\mathbb{E}\Big[\mathcal{E}_f(|Q|)\Big] &= \mathbb{E}\Big[\mathbb{1}\Big(|Q| \leq q^*(\alpha) - \epsilon\Big)\mathcal{E}_f(|Q|) + \mathbb{1}\Big(|Q| > q^*(\alpha) - \epsilon\Big)\mathcal{E}_f(|Q|)\Big] \\ &\geq p(\epsilon)\mathcal{E}_f(q^*(\alpha) - \epsilon) + (1 - p(\epsilon))\mathcal{E}_f(q^*(\alpha))\,.\end{aligned}$$

because $\mathcal{E}_f$ is non-increasing and because $|Q| \leq q^*(\alpha)$ almost surely (Lemma 7.7.1). Combining this with (7.7.12) leads to

$$p(\epsilon)\mathcal{E}_f(q^*(\alpha)) \geq p(\epsilon)\mathcal{E}_f(q^*(\alpha) - \epsilon)\,.$$

Since $\mathcal{E}_f$ is strictly decreasing: $\mathcal{E}_f(q^*(\alpha)) < \mathcal{E}_f(q^*(\alpha) - \epsilon)$, which implies $p(\epsilon) = 0$. This is true for all $\epsilon > 0$, consequently $|Q| \geq q^*(\alpha)$ almost surely. We get (using Lemma 7.7.1) that

$$|Q| = q^*(\alpha)\,, \quad \text{almost surely.}$$

We conclude that the only possible limit in law of the tight sequence $\Big(|Q_n|\Big)_{n\geq 1}$ is $q^*(\alpha)$. Therefore $|Q_n| \to q^*(\alpha)$.

Let us prove the second part of Theorem 7.3.2. If $q^*(\alpha) = 0$, then this is obvious. Let us suppose now that $q^*(\alpha) > 0$ and let $p = \mathbb{P}(Q = -q^*(\alpha))$. Suppose that $p > 0$. We have seen above that for all continuous bounded function f, $\mathbb{E}H_f(Q) = H_f(q^*(\alpha))$. Hence

$$H_f(q^*(\alpha)) = \mathbb{E}H_f(Q) = (1-p)H_f(q^*(\alpha)) + pH_f(-q^*(\alpha)),$$

and consequently $H_f(q^*(\alpha)) = H_f(-q^*(\alpha))$. Lemma 7.7.3 gives then that

$$\int f(y)P_{\text{out}}(dy|x) = \int f(y)P_{\text{out}}(dy|-x)$$

for almost every $x \in \mathbb{R}$, which then leads to $P_{\text{out}}(\cdot|x) = P_{\text{out}}(\cdot|-x)$ for almost-every $x \in \mathbb{R}$: P_{out} is symmetric. We conclude that if P_{out} is not symmetric, then necessarily $p = 0$ and $Q = q^*(\alpha)$ almost surely.

Proof of Proposition 7.7.3

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous bounded function. Let $\epsilon, \lambda > 0$. In order to get lower-bounds on the generalization error, we will consider the following additional observations

$$U_\mu = \sqrt{\lambda}\, Y'_\mu + Z'_\mu\,, \qquad \text{for} \qquad 1 \leq \mu \leq m' = \epsilon n\,, \tag{7.7.13}$$

where $Z'_\mu \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, and Y'_μ is given by

$$Y'_\mu = f(\widetilde{Y}_\mu)\,, \qquad \widetilde{Y}_\mu \sim P_{\text{out}}\Big(\,\cdot\,\Big|\,\frac{\langle \boldsymbol{\Phi}'_\mu, \boldsymbol{X}^*\rangle}{\sqrt{n}}\Big)\,, \tag{7.7.14}$$

where $\boldsymbol{\Phi}'_\mu \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathrm{Id}_n)$ independently of everything else. We will use the following lemma, proved at the end of this section.

Lemma 7.7.4

For all $\epsilon, \lambda \geq 0$ we have

$$\frac{1}{n} I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') \xrightarrow[n\to\infty]{} \inf_{q\in[0,\rho]} \sup_{r\geq 0} \tilde{i}_{\mathrm{RS}}(q, r, \lambda) - i_\infty \,, \tag{7.7.15}$$

where i_∞ is given by Corollary 7.3.1 and

$$\tilde{i}_{\mathrm{RS}}(q, r, \lambda) \overset{\text{def}}{=} i_{\mathrm{RS}}(q, r) + \epsilon I(f(Y^{(q)}); \sqrt{\lambda}\, f(Y^{(q)}) + Z' | V) \tag{7.7.16}$$

$$= I_{P_0}(r) + \alpha \mathcal{I}_{P_{\mathrm{out}}}(q; \rho) + \epsilon I(f(Y^{(q)}); \sqrt{\lambda}\, f(Y^{(q)}) + Z' | V) - \frac{r}{2}(\rho - q)\,. \tag{7.7.17}$$

Recall that $Y^{(q)}$ is sampled from the "second scalar channel" (7.3.3): $Y^{(q)} \sim P_{\mathrm{out}}(\cdot \,|\, \sqrt{q}\, V + \sqrt{\rho - q}\, W^*)$, where $V, W^* \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$.

Lemma 7.7.5

For all $\alpha, \lambda > 0$ the set

$$D_{\alpha,\lambda} \overset{\text{def}}{=} \left\{ \epsilon \geq 0 \,\middle|\, \text{the infimum in (7.7.15) is achieved at a unique } q^*_{\alpha,\epsilon,\lambda} \right\} \tag{7.7.18}$$

is equal to $[0, +\infty)$ minus some countable set. Moreover, $\epsilon \mapsto q^*_{\alpha,\lambda,\epsilon}$ is continuous on $D_{\alpha,\epsilon}$.

Proof. This follows from the same arguments than in the proof of Proposition 7.3.1. □

From there one can use the I-MMSE relation of Proposition 1.3.3:

$$\begin{aligned} \frac{\partial}{\partial\lambda} \frac{1}{n} I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') &= \frac{1}{2n} \mathrm{MMSE}\Big(\boldsymbol{Y}' \Big| \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}', \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}'\Big) \\ &= \frac{\epsilon}{2} \mathrm{MMSE}(Y_1' | \boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}'). \end{aligned}$$

We will express the limit of $\mathrm{MMSE}(Y_1' | \boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}')$ using the function:

$$M_f : (\lambda, q) \mapsto \mathrm{MMSE}\Big(f(Y^{(q)}) \Big| \sqrt{\lambda}\, f(Y^{(q)}) + Z', V\Big)\,. \tag{7.7.19}$$

Lemma 7.7.6

For all $\alpha, \lambda > 0$, we have for all $\epsilon \in D_{\alpha,\lambda} \setminus \{0\}$

$$\lim_{n\to\infty} \mathrm{MMSE}(Y_1' | \boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') = M_f(\lambda, q^*_{\alpha,\epsilon,\lambda})\,,$$

where $q^*_{\alpha,\epsilon,\lambda}$ is the unique minimizer of (7.7.15).

Proof. Let us fix $\alpha, \epsilon > 0$. Consider the function

$$h_{\alpha,\epsilon} : \lambda \mapsto \inf_{q\in[0,\rho]} \sup_{r\geq 0} \tilde{i}_{\mathrm{RS}}(q, r, \lambda)\,. \tag{7.7.20}$$

Corollary 4 from [144] gives that $h_{\alpha,\epsilon}$ is differentiable at λ if and only if

$$\left\{ \epsilon \frac{\partial}{\partial\lambda} I(f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V) = \frac{\epsilon}{2} M_f(\lambda, q) \,\middle|\, q \text{ minimizer of (7.7.15) (or equivalently of (7.7.20))} \right\}$$

is a singleton (the equality comes from the I-MMSE relation from Proposition 1.3.3). In such case, Corollary 4 from [144] also gives that

$$h'_{\alpha,\epsilon}(\lambda) = \frac{\epsilon}{2} M_f(\lambda, q)\,, \tag{7.7.21}$$

for all q minimizer of (7.7.20). So if now $\epsilon \in D_{\alpha,\lambda} \setminus \{0\}$, then the minimizer is unique and thus $h_{\alpha,\epsilon}$ is differentiable at λ, with derivative $h'_{\alpha,\lambda}(\lambda) = \epsilon M_f(\lambda, q^*_{\alpha,\epsilon,\lambda})/2$. However, by (7.7.15) in Lemma 7.7.4, $h_{\alpha,\epsilon}$ is the pointwise limit on $\mathbb{R}_{\geq 0}$ of the sequence of concave functions

$$(h_n)_{n\geq 1} = \Big(\lambda \mapsto \frac{1}{n} I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') + i_\infty\Big)_{n\geq 1}\,.$$

Consequently, Proposition C.1 gives $h'_n(\lambda) \xrightarrow[n\to\infty]{} h'_{\alpha,\epsilon}(\lambda)$. By the I-MMSE relation (Proposition 1.3.3) we have $h'_n(\lambda) = \epsilon \mathrm{MMSE}(Y'_1 | \boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}')/2$ and we conclude using the fact that $\epsilon \neq 0$. □

Lemma 7.7.7

For all $\alpha \in D^*$ (recall that D^* is defined by (7.3.13)),

$$\lim_{\lambda\to 0} \lim_{\epsilon\to 0} M_f(\lambda, q^*_{\alpha,\epsilon,\lambda}) = \mathcal{E}_f(q^*(\alpha))\,.$$

Proof. Let $\alpha \in D^*$ and $\lambda > 0$. We have by definition of $D_{\alpha,\lambda}$, of D^* and using the link between $\tilde{i}_{\mathrm{RS}}$ and i_{RS} given by (7.7.16), that $0 \in D_{\alpha,\lambda}$. By Lemma 7.7.5 above, we have

$$q^*_{\alpha,\epsilon,\lambda} \xrightarrow[\epsilon\to 0,\ \epsilon\in D_{\alpha,\lambda}]{} q^*_{\alpha,0,\lambda} = q^*(\alpha)\,.$$

Analogously to Proposition 7.8.5, $M_f(\lambda, \cdot)$ is continuous on $[0, \rho]$, thus $\lim_{\epsilon\to 0} M_f(\lambda, q^*_{\alpha,\epsilon,\lambda}) = M_f(\lambda, q^*(\alpha))$. We obtain the result by taking $\lim_{\lambda\to 0} M_f(\lambda, q^*(\alpha)) = \mathcal{E}_f(q^*(\alpha))$, using the continuity of the MMSE at 0, see Proposition 1.3.2. □

In order to simplify the proof, we assume that $m = \alpha n$. By definition of the generalization error (7.7.6) and of the labels $\boldsymbol{Y}'$ given by (7.7.14), we have

$$\mathcal{E}_{f,n}(\alpha) = \mathrm{MMSE}(Y'_1 | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}')\,.$$

Lemma 7.7.8

For all $\alpha \in D^*$,

$$\liminf_{n\to\infty} \mathcal{E}_{f,n}(\alpha) \geq \mathcal{E}_f(q^*(\alpha))\,.$$

Proof. Let $\alpha \in D^*$, $\lambda > 0$ and $\epsilon \in D_{\alpha,\lambda} \setminus \{0\}$. Obviously,

$$\mathcal{E}_{f,n}(\alpha) \geq \mathrm{MMSE}(Y'_1 | \boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') \xrightarrow[n\to\infty]{} M_f(\lambda, q^*_{\alpha,\epsilon,\lambda})\,,$$

where we used Lemma 7.7.6. Consequently $\liminf_{n\to\infty} \mathcal{E}_{f,n}(\alpha) \geq M_f(\lambda, q^*_{\alpha,\epsilon,\lambda})$ and we obtain the lower bound by letting $\epsilon, \lambda \to 0$ and using Lemma 7.7.7. □

Let us now prove the converse upper bound.

Lemma 7.7.9

There exists a constant $C > 0$ (that only depend on f) such that for all $\alpha, \lambda > 0$ and all $\epsilon \in D_{\alpha,\lambda} \setminus \{0\}$

$$\limsup_{n\to\infty} \mathcal{E}_{f,n}(\alpha + \epsilon) \leq M_f(\lambda, q^*_{\alpha,\epsilon,\lambda}) + C\lambda\,.$$

Proof. We will let the signal-to-noise ratio (snr) of the observation of Y'_1 go to zero. Let us denote by λ_1 this snr: $U_1 = \sqrt{\lambda_1}\, Y'_1 + Z'_1$. We will let λ_1 go from λ to 0 while the other snr for the observations of U_μ for $\mu = 2, \dots, \epsilon n$ will remain equal to λ. Using Proposition 9 from [97],

$$\Big|\frac{\partial}{\partial\lambda_1}\mathrm{MMSE}(Y'_1|\boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}')\Big| = \mathbb{E}\big[\mathrm{Var}(Y'_1|\boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}')^2\big] \leq \mathbb{E}\big[(Y'_1)^4\big] \leq \|f\|^4_\infty\,.$$

We define $C \overset{\text{def}}{=} \|f\|^4_\infty$. Consequently, by the mean value theorem,

$$\big|\mathrm{MMSE}(Y'_1|\boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') - \mathrm{MMSE}(Y'_1|\boldsymbol{Y}, (U_\mu)_{\mu=2}^{\epsilon n}, \boldsymbol{\Phi}, \boldsymbol{\Phi}')\big| \leq C\lambda\,. \tag{7.7.22}$$

Since $(U_\mu)_{\mu=2}^{\epsilon n}$ contains less information than $(\widetilde{Y}_\mu)_{\mu=2}^{\epsilon n}$ because of the additional Gaussian noise and the application of the function f, we have

$$\mathrm{MMSE}(Y'_1|\boldsymbol{Y}, (U_\mu)_{\mu=2}^{\epsilon n}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') \geq \mathrm{MMSE}(Y'_1|\boldsymbol{Y}, (\widetilde{Y}_\mu)_{\mu=2}^{\epsilon n}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') = \mathcal{E}_{f,n}(\alpha + \epsilon - 1/n) \geq \mathcal{E}_{f,n}(\alpha + \epsilon)\,.$$

The last identity combined with (7.7.22) leads to

$$\mathrm{MMSE}(Y'_1|\boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') + C\lambda \geq \mathcal{E}_{f,n}(\alpha + \epsilon)\,. \tag{7.7.23}$$

By Lemma 7.7.6 we know that $\lim_{n\to\infty} \mathrm{MMSE}(Y'_1|\boldsymbol{Y}, \boldsymbol{U}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') = M_f(\lambda, q^*_{\alpha,\epsilon,\lambda})$. We conclude by taking the limsup in the inequality above. □

Corollary 7.7.1

For all $\alpha \in D^*$,

$$\limsup_{n\to\infty} \mathcal{E}_{f,n}(\alpha) \leq \mathcal{E}_f(q^*(\alpha))\,.$$

Proof. Let $\alpha \in D^*, \lambda > 0$ and $\epsilon_1 > 0$ such that $\alpha - \epsilon_1 \in D^*$. Since by Lemma 7.7.5 the set $D_{\alpha-\epsilon_1,\lambda}$ is dense in $\mathbb{R}_{\geq 0}$, we can find $\epsilon_2 \in D_{\alpha-\epsilon_1,\lambda}$ such that $0 < \epsilon_2 \leq \epsilon_1$. Using Lemma 7.7.9 above, we have

$$\limsup_{n\to\infty} \mathcal{E}_{f,n}(\alpha - \epsilon_1 + \epsilon_2) \leq M_f(\lambda, q^*_{\alpha-\epsilon_1,\epsilon_2,\lambda}) + C\lambda\,.$$

Now, using the fact that $\epsilon_2 \leq \epsilon_1$ we have

$$\limsup_{n\to\infty} \mathcal{E}_{f,n}(\alpha) \leq \limsup_{n\to\infty} \mathcal{E}_{f,n}(\alpha - \epsilon_1 + \epsilon_2) \leq M_f(\lambda, q^*_{\alpha-\epsilon_1,\epsilon_2,\lambda}) + C\lambda\,.$$

Now, by Lemma 7.7.7 we have

$$\lim_{\lambda\to 0}\lim_{\epsilon_2\to 0} M_f(\lambda, q^*_{\alpha-\epsilon_1,\epsilon_2,\lambda}) + C\lambda = \mathcal{E}_f(q^*(\alpha - \epsilon_1))$$

which leads to $\limsup_{n\to\infty} \mathcal{E}_{f,n}(\alpha) \leq \mathcal{E}_f(q^*(\alpha - \epsilon_1))$. We conclude by letting $\epsilon_1 \to 0$ (recall that by Proposition 7.3.1 D^* is dense in $\mathbb{R}_{\geq 0}$ so it is possible to find $\epsilon_1 > 0$ arbitrary small such that $\alpha - \epsilon_1 \in D^*$), using the continuity of $\mathcal{E}_f$ (by Proposition 7.8.5) and the continuity of q^* (by Proposition 7.3.1). □

Proof of Lemma 7.7.4.

Extending the interpolation method presented in Section 7.6, one can generalize Theorem 7.3.1 to take into account this additional side information. This gives directly

$$\frac{1}{n} I(\boldsymbol{X}^*; \boldsymbol{Y}, \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{\Phi}, \boldsymbol{\Phi}') \xrightarrow[n\to\infty]{} \tilde{I}_\infty(\alpha, \epsilon, \lambda) \stackrel{\text{def}}{=} \inf_{q\in[0,\rho]} \sup_{r\geq 0} \tilde{I}_{\text{RS}}(q, r, \lambda) \tag{7.7.24}$$

where $\tilde{I}_{\text{RS}}(q, r, \lambda)$ is given by

$$\tilde{I}_{\text{RS}}(q, r, \lambda) \stackrel{\text{def}}{=} I_{P_0}(r) + \alpha \mathcal{I}_{P_{\text{out}}}(q; \rho) + \epsilon I(W^*; \sqrt{\lambda} f(Y^{(q)}) + Z' | V) - \frac{r}{2}(\rho - q)\,. \tag{7.7.25}$$

Conditionally on $(V, f(Y^{(q)}))$, the random variables W^* and $\sqrt{\lambda} f(Y^{(q)}) + Z'$ are independent, therefore

$$I(f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V) = I(W^*, f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V)\,.$$

Now, by the chain rule of the mutual information we have

$$I(W^*, f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V) = I(W^*; \sqrt{\lambda} f(Y^{(q)}) + Z' | V) + I(f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V, W^*)\,.$$

We obtain that

$$I(W^*; \sqrt{\lambda} f(Y^{(q)}) + Z' | V) = I(f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V) - I(f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V, W^*)\,. \tag{7.7.26}$$

Notice that the last mutual information in the above equation does not depend on q nor r. Therefore we have:

$$\inf_{q\in[0,\rho]} \sup_{r\geq 0} \tilde{I}_{\text{RS}}(q, r, \lambda) = -\epsilon I(f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V, W^*) + \inf_{q\in[0,\rho]} \sup_{r\geq 0} \tilde{i}_{\text{RS}}(q, r, \lambda)\,. \tag{7.7.27}$$

Now, by the chain rule, we have

$$\frac{1}{n} I(\boldsymbol{X}^*; \boldsymbol{Y}, \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{\Phi}, \boldsymbol{\Phi}') = \frac{1}{n} I(\boldsymbol{X}^*; \boldsymbol{Y} | \boldsymbol{\Phi}) + \frac{1}{n} I(\boldsymbol{X}^*; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}')\,. \tag{7.7.28}$$

The limit of the left-hand side is given by (7.7.24). By Corollary 7.3.1, we have $\lim_{n\to\infty} I(\boldsymbol{X}^*; \boldsymbol{Y} | \boldsymbol{\Phi})/n = i_\infty$. It remains to investigate the last term of the equation above. By the arguments used to prove (7.7.26), we have

$$\begin{aligned} I(\boldsymbol{X}^*; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') &= I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') - I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{X}^*, \boldsymbol{\Phi}') \\ &= I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{\Phi}') - I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{X}^*, \boldsymbol{\Phi}')\,. \end{aligned} \tag{7.7.29}$$

We have $I(\boldsymbol{Y}'; \sqrt{\lambda}\boldsymbol{Y}' + \boldsymbol{Z}' | \boldsymbol{X}^*, \boldsymbol{\Phi}')/n = \epsilon I(Y_1'; \sqrt{\lambda} Y_1' + Z_1' | \boldsymbol{X}^*, \boldsymbol{\Phi}_1')$ and it is not difficult to show, using similar computations as in the proof of Corollary 7.3.1, that

$$I(Y_1'; \sqrt{\lambda} Y_1' + Z_1' | \boldsymbol{X}^*, \boldsymbol{\Phi}_1') \xrightarrow[n\to\infty]{} I(f(Y^{(q)}); \sqrt{\lambda} f(Y^{(q)}) + Z' | V, W^*)\,,$$

(recall that the right-hand side does not depend on q). Combining this with (7.7.29), (7.7.28), (7.7.24), Corollary 7.3.1 and (7.7.27), we obtain the desired result. □

7.7.3 Optimal generalization error: Proof of Theorem 7.3.4

Let $\alpha \in D^*$. We will simply write q^* instead of $q^*(\alpha)$ in order to lighten the notations. For $\boldsymbol{x} \in \mathbb{R}^n$ we write $Z(\boldsymbol{x}) \stackrel{\text{def}}{=} \frac{1}{\sqrt{n}} \langle \boldsymbol{x}, \boldsymbol{\Phi}_{\text{new}} \rangle$. By [119, Theorem 4.11] it suffices to show that for

all continuous bounded function f, $\nu_n f \xrightarrow[n\to\infty]{(\mathrm{d})} \nu f$. Let f be a continuous bounded function and define

$$F(z) = \int f(y) P_{\mathrm{out}}(dy|z).$$

We are going to show that $\nu_n f \xrightarrow[n\to\infty]{(\mathrm{d})} \nu f$ by the moment method. The sequence of the laws of $(\nu_n f)_{n\geq 1}$ is tight because bounded. Let us consider some subsequence $(n_l)_{l\geq 1}$ along which it converges in distribution to some random variable S_f. Let $k \in \mathbb{N}^*$.

$$\mathbb{E}\Big[(\nu_n f)^k\Big] = \mathbb{E}\Big[F(Z(\boldsymbol{x}^{(1)}))\cdots F(Z(\boldsymbol{x}^{(k)}))\Big],$$

where $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k)}$ are i.i.d. samples from the posterior distribution of $\boldsymbol{X}^*$ given $\boldsymbol{Y}, \boldsymbol{\Phi}$, independently of everything else. We define the matrix

$$Q_n = \Big(\langle \boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\rangle\Big)_{1\leq i,j\leq k}.$$

We extract from $(n_l)_{l\geq 1}$ another subsequence $(n'_l)_{l\geq 1}$ along which Q_n converges weakly to some random variable Q. By a central-limit-type argument (Lemma 7.7.2), $(Z(\boldsymbol{x}^{(k)}), \dots, Z(\boldsymbol{x}^{(k)})) \xrightarrow[n'_l\to\infty]{(\mathrm{d})} (Z_1, \dots, Z_k)$ where $\boldsymbol{Z}$ is sampled, conditionally on Q from $\mathcal{N}(0, Q)$. By hypothesis (h4) and Remark 7.3.1, F is continuous almost everywhere. We get:

$$\mathbb{E}\Big[(\nu_n f)^k\Big] \to \mathbb{E}\Big[F(Z_1)\cdots F(Z_k)\Big],$$

where the limit is taken along the subsequence (n'_l) mentioned above. On the other hand

$$\begin{aligned}\mathbb{E}\Big[(\nu f)^k\Big] &= \mathbb{E}\Big[F(\sqrt{q^*}V + \sqrt{\rho - q^*}W_1)\cdots F(\sqrt{q^*}V + \sqrt{\rho - q^*}W_k)\Big]\\ &= \mathbb{E}\Big[F(Z'_1)\cdots F(Z'_k)\Big],\end{aligned}$$

where $V, W_1, \dots, W_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and $\boldsymbol{Z}' \sim \mathcal{N}(0, R)$ where the matrix $R \in \mathbb{R}^{k\times k}$ is defined by

$$R_{i,i} = \rho \qquad \text{and} \qquad R_{i,j} = q^*, \qquad \text{for all } i \neq j.$$

By the Nishimori identity (Proposition 1.1.1) and the law of large numbers we have $Q_{i,i} = \rho$ almost surely for all $i \in \{1, \dots, k\}$. If P_{out} is not symmetric, then by Theorem 7.3.2 we have $Q = R$ almost surely which gives

$$\mathbb{E}\Big[(\nu_n f)^k\Big] \to \mathbb{E}\Big[F(Z_1)\cdots F(Z_k)\Big] = \mathbb{E}\Big[(\nu f)^k\Big].$$

If now P_{out} is symmetric. By Theorem 7.3.2 we have for all $i \neq j$ in $\{1, \dots, k\}$, $Q_{i,j} \in \{-q^*, q^*\}$ almost surely. So if $q^* = 0$, then $Q = R$ almost surely and we are done. Suppose that $q^* > 0$. By Lemma 1.2.2 we have also for all $i, j, \ell \in \{1, \dots, k\}$ distinct, $Q_{i,j}Q_{j,\ell}Q_{\ell,i} = (q^*)^3$, almost surely. The binary relation $\sim$ on $\{1, \dots, k\}$ defined by

$$i \sim j \iff Q_{i,j} > 0$$

is with probability one an equivalence relation. Let $\eta_i = 1$ if $i \sim 1$ and -1 otherwise. By construction we have then

$$\Big(\eta_i Z_i\Big)_{1\leq i\leq k} \overset{(\mathrm{d})}{=} \Big(Z'_i\Big)_{1\leq i\leq k}.$$

We assumed that P_{out} is symmetric, therefore $F(z) = F(|z|)$ for almost all $z \in \mathbb{R}$. Thus $F(\eta_i Z_i) = F(Z_i)$ almost surely. This gives:

$$\begin{aligned}\mathbb{E}\Big[(\nu_n f)^k\Big] &\to \mathbb{E}\Big[F(Z_1)\cdots F(Z_k)\Big] \\ &= \mathbb{E}\Big[F(\eta_1 Z_1)\cdots F(\eta_k Z_k)\Big] \\ &= \mathbb{E}\Big[F(Z_1')\cdots F(Z_k')\Big] = \mathbb{E}\Big[(\nu f)^k\Big].\end{aligned}$$

Recall that $\nu_n f$ converges in distribution to S_f along the subsequence (n_l'). We get $\mathbb{E}S_f^k = \mathbb{E}\Big[(\nu f)^k\Big]$. This last equality holds for all $k \geq 1$ and since νf is a bounded random variable we conclude that $S_f \overset{(\mathrm{d})}{=} \nu f$. νf is therefore the only point of accumulation (for the weak convergence) of the sequence $(\nu_n f)_{n\geq 1}$. This sequence is tight because bounded we conclude that $\nu_n f \xrightarrow[n\to\infty]{(\mathrm{d})} \nu f$ for all continuous bounded function f and therefore $\nu_n \xrightarrow[n\to\infty]{(\mathrm{wd})} \nu$.

7.7.4 Generalization error of GAMP: Proof of Proposition 7.4.1

Let us decompose:

$$\mathcal{E}_{\text{gen}}^{\text{GAMP},t} \overset{\text{def}}{=} \mathbb{E}\Big[\Big(Y_{\text{new}} - \widehat{Y}^{\text{GAMP},t}\Big)^2\Big] = \mathbb{E}\Big[Y_{\text{new}}^2\Big] + \mathbb{E}\Big[\Big(\widehat{Y}^{\text{GAMP},t}\Big)^2\Big] - 2\mathbb{E}\Big[Y_{\text{new}}\widehat{Y}^{\text{GAMP},t}\Big]. \tag{7.7.30}$$

Lemma 7.7.10

Let $V, W \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and let $\mathbb{E}_W$ denotes the expectation with respect to W only. We have

$$\mathbb{E}\Big[Y_{\text{new}}\widehat{Y}^{\text{GAMP},t}\Big] \xrightarrow[n\to\infty]{} \mathbb{E}\Big[\mathbb{E}_W\Big[\int y P_{\text{out}}(dy|\sqrt{q^t}\,V + \sqrt{\rho - q^t}\,W)\Big]^2\Big]. \tag{7.7.31}$$

Proof. Start by writing

$$\mathbb{E}[Y_{\text{new}}\widehat{Y}^{\text{GAMP},t}] = \mathbb{E}\int y\,y'\,P_{\text{out}}\Big(dy\Big|\frac{\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle}{\sqrt{n}}\Big)P_{\text{out}}\Big(dy'\Big|\frac{\langle \boldsymbol{\Phi}_{\text{new}}, \widehat{\boldsymbol{x}}^t\rangle}{\sqrt{n}} + \sqrt{\rho - q^t}\,W\Big)$$

where $W \sim \mathcal{N}(0,1)$ is independent of everything else. $\boldsymbol{\Phi}_{\text{new}} \sim \mathcal{N}(0, \mathrm{Id}_n)$ is independent of $\boldsymbol{X}^*$ and $\widehat{\boldsymbol{x}}^t$, so, conditionally on $\boldsymbol{X}^*, \widehat{\boldsymbol{x}}^t$ we have

$$\Big(\frac{\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle}{\sqrt{n}}, \frac{\langle \boldsymbol{\Phi}_{\text{new}}, \widehat{\boldsymbol{x}}^t\rangle}{\sqrt{n}}\Big) \sim \mathcal{N}\Big(0, \frac{1}{n}\begin{pmatrix} \|\boldsymbol{X}^*\|^2 & \langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^*\rangle \\ \langle \widehat{\boldsymbol{x}}^t, \boldsymbol{X}^*\rangle & \|\widehat{\boldsymbol{x}}^t\|^2 \end{pmatrix}\Big).$$

We assumed that (7.4.6) holds, i.e. $\langle \boldsymbol{X}^*, \widehat{\boldsymbol{x}}^t\rangle/n \to q^t$ and $\|\widehat{\boldsymbol{x}}^t\|^2/n \to q^t$, in probability. By the law of large numbers $\|\boldsymbol{X}^*\|^2/n \to \rho$ in probability. Consequently,

$$\Big(\frac{\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle}{\sqrt{n}}, \frac{\langle \boldsymbol{\Phi}_{\text{new}}, \widehat{\boldsymbol{x}}^t\rangle}{\sqrt{n}}\Big) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\Big(0, \begin{pmatrix} \rho & q^t \\ q^t & q^t \end{pmatrix}\Big).$$

Since $x \mapsto P_{\text{out}}(\cdot|x)$ is continuous almost everywhere for the Wasserstein distance of order 2, the function $h : (a,b) \mapsto \mathbb{E}_W \int y y' P_{\text{out}}(dy|a) P_{\text{out}}(dy'|b + \sqrt{\rho - q^t}\,W)$ with $W \sim \mathcal{N}(0,1)$ is continuous almost everywhere. Therefore

$$H_n \overset{\text{def}}{=} h\Big(\frac{\langle \boldsymbol{\Phi}_{\text{new}}, \boldsymbol{X}^*\rangle}{\sqrt{n}}, \frac{\langle \boldsymbol{\Phi}_{\text{new}}, \widehat{\boldsymbol{x}}^t\rangle}{\sqrt{n}}\Big) \xrightarrow[n\to\infty]{(d)} h(\sqrt{q^t}\,Z_0 + \sqrt{\rho - q^t}\,Z_1, \sqrt{q^t}\,Z_0), \tag{7.7.32}$$

where $Z_0, Z_1 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Let $\eta > 0$. We have by Jensen's inequality

$$\begin{aligned}\mathbb{E}\big[|H_n|^{1+\eta}\big] &\le \mathbb{E}\big[|Y_{\text{new}}\widehat{Y}^{\text{GAMP},t}|^{1+\eta}\big] \le \mathbb{E}\Big[\Big(\frac{1}{2}Y_{\text{new}}^2 + \frac{1}{2}(\widehat{Y}^{\text{GAMP},t})^2\Big)^{1+\eta}\Big] \\ &\le \frac{1}{2}\mathbb{E}|Y_1|^{2+2\eta} + \frac{1}{2}\mathbb{E}|\widehat{Y}^{\text{GAMP},t}|^{2+2\eta}.\end{aligned}$$

By assumption, there exists $\eta > 0$ such that the two last terms above remain bounded with n: H_n is therefore bounded in $L^{1+\eta}$ and is therefore uniformly integrable. From (7.7.32) we thus get

$$\begin{aligned}\mathbb{E}\big[Y_{\text{new}}\widehat{Y}^{\text{GAMP},t}\big] = \mathbb{E}[H_n] &\xrightarrow[n\to\infty]{} \mathbb{E}\big[h(\sqrt{q^t}\,Z_0 + \sqrt{\rho - q^t}\,Z_1, \sqrt{q^t}\,Z_0)\big] \\ &= \mathbb{E}\Big[\mathbb{E}_W\Big[\int y P_{\text{out}}(dy|\sqrt{q^t}\,V + \sqrt{\rho - q^t}\,W)\Big]^2\Big].\end{aligned}$$

□

Following the arguments of Lemma 7.7.10 one can also show that

$$\begin{aligned}\mathbb{E}\Big[\big(\widehat{Y}^{\text{GAMP},t}\big)^2\Big] &\xrightarrow[n\to\infty]{} \mathbb{E}\Big[\mathbb{E}_W\Big[\int y P_{\text{out}}(dy|\sqrt{q^t}\,V + \sqrt{\rho - q^t}\,W)\Big]^2\Big], \\ \mathbb{E}\Big[Y_{\text{new}}^2\Big] &\xrightarrow[n\to\infty]{} \mathbb{E}\int y^2 P_{\text{out}}(dy|\sqrt{\rho}\,V).\end{aligned}$$

This proves (together with (7.7.30) and Lemma 7.7.10) Proposition 7.4.1.

7.8 The non-linear scalar channel

We prove here some properties of the free energy of the second scalar channel (7.3.3), where $V, W^* \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and

$$Y^{(q)} \sim P_{\text{out}}\Big(\cdot\,\Big|\,\sqrt{q}\,V + \sqrt{\rho - q}\,W^*\Big). \tag{7.8.1}$$

In this channel, the statistician observes V and $Y^{(q)}$ and wants to recover W^*.

7.8.1 Study of the scalar free energy

We recall the definition of $\Psi_{P_{\text{out}}}$:

$$\Psi_{P_{\text{out}}}(q) = \mathbb{E}\log\int dw\,\frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}P_{\text{out}}\Big(Y^{(q)}|\sqrt{q}\,V + \sqrt{\rho - q}\,w\Big),$$

where $P_{\text{out}}(\cdot|x)$ denotes the density of $P_{\text{out}}(\cdot|x)$ with respect to the Lebesgue's measure on $\mathbb{R}$ or the counting measure on $\mathbb{N}$ (we will always be in one of these two cases). In this section, we establish the main properties of $\Psi_{P_{\text{out}}}$. Recall that $\mathcal{I}_{P_{\text{out}}}(q) \overset{\text{def}}{=} I(W^*; Y^{(q)}|V) = \Psi_{P_{\text{out}}}(\rho) - \Psi_{P_{\text{out}}}(q)$ so the properties we will prove on $\Psi_{P_{\text{out}}}$ can be directly translated for $\mathcal{I}_{P_{\text{out}}}$, and vice-versa. We will denote the expectation with respect the posterior distribution of W^* given $Y^{(q)}, V$ by the Gibbs bracket $\langle\cdot\rangle_q$:

$$\langle f(w)\rangle_q = \frac{\int f(w)e^{-w^2/2}P_{\text{out}}(Y^{(q)}|\sqrt{q}\,V + \sqrt{\rho - q}\,w)dw}{\int e^{-w^2/2}P_{\text{out}}(Y^{(q)}|\sqrt{q}\,V + \sqrt{\rho - q}\,w)dw}. \tag{7.8.2}$$

Proposition 7.8.1

Suppose that for all $x \in \mathbb{R}$, $P_{\text{out}}(\cdot\,|\,x)$ is the law of $\varphi(x,A) + \sqrt{\Delta}Z$ where $\Delta > 0$, $\varphi : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is a measurable function and $(Z,A) \sim \mathcal{N}(0,1)\otimes P_A$, for some probability distribution P_A over $\mathbb{R}$. In that case P_{out} admits a density given by

$$P_{\text{out}}(y|x) = \frac{1}{\sqrt{2\pi\Delta}} \int dP_A(a) e^{-\frac{1}{2\Delta}(y-\varphi(x,a))^2} .$$

Assume that φ is bounded and $\mathcal{C}^2$ with respect to its first coordinate, with bounded first and second derivatives. Then $q \mapsto \Psi_{P_{\text{out}}}(q)$ is convex, $\mathcal{C}^2$ and non-decreasing on $[0,\rho]$.

Proof. Under the hypotheses we made on φ, we will be able to use continuity and differentiation under the expectation, because all the domination hypotheses are easily verified.

We compute now the first derivative. Recall that $\langle\cdot\rangle_q$, defined in (7.8.2), denotes the posterior distribution of W^* given $Y^{(q)}$. We will use the notation $u_y(x) = \log P_{\text{out}}(y|x)$. For $q \in (0,\rho)$ we have

$$\begin{aligned}\Psi'_{P_{\text{out}}}(q) &= \frac{1}{2}\mathbb{E}\Big\langle u'_{Y^{(q)}}(\sqrt{q}\,V + \sqrt{\rho-q}\,w) u'_{Y^{(q)}}(\sqrt{q}\,V + \sqrt{\rho-q}\,W^*)\Big\rangle_q \\ &= \frac{1}{2}\mathbb{E}\Big\langle u'_{Y^{(q)}}(\sqrt{q}\,V + \sqrt{\rho-q}\,w)\Big\rangle_q^2 \geq 0\,,\end{aligned}$$

where $w \sim \langle\cdot\rangle_q$, independently of everything else. $\Psi_{P_{\text{out}}}$ is therefore non-decreasing. Using the boundedness assumption on φ and its derivatives, it is not difficult to check that $\Psi'_{P_{\text{out}}}$ is indeed bounded.

We will now compute $\Psi''_{P_{\text{out}}}$. To lighten the notations, we write $u'(w)$ for $u'_{Y^{(q)}}(\sqrt{q}\,V + \sqrt{\rho-q}\,w)$. We compute

$$\begin{aligned}\partial_q \mathbb{E}\Big\langle u'(w)u'(W^*)\Big\rangle_q &= \mathbb{E}\Big[\Big(\frac{1}{2\sqrt{q}}\,V - \frac{1}{2\sqrt{\rho-q}}\,W^*\Big)u'(W^*)\Big\langle u'(w)u'(W^*)\Big\rangle_q\Big] && (A)\\ &\quad + 2\mathbb{E}\Big\langle\Big(\frac{1}{2\sqrt{q}}\,V - \frac{1}{2\sqrt{\rho-q}}\,W^*\Big)u''(W^*)u'(w)\Big\rangle_q && (B)\\ &\quad + \mathbb{E}\Big\langle\Big(\frac{1}{2\sqrt{q}}\,V - \frac{1}{2\sqrt{\rho-q}}\,W^*\Big)u'(W^*)^2u'(w)\Big\rangle_q && (C)\\ &\quad - \mathbb{E}\Big\langle u'(W^*)u'(w)\Big\rangle_q\Big\langle\Big(\frac{1}{2\sqrt{q}}\,V - \frac{1}{2\sqrt{\rho-q}}\,w\Big)u'(w)\Big\rangle_q && (D)\end{aligned} \tag{7.8.3}$$

Notice that $(A) = (C)$. Let $w, w^{(1)}, w^{(2)} \overset{\text{i.i.d.}}{\sim} \langle\cdot\rangle_q$, independently of everything else. We compute, using Gaussian integration by parts and the Nishimori identity (Proposition 1.1.1)

$$\begin{aligned}(A) &= \frac{1}{2}\mathbb{E}\Big[u'(W^*)\Big\langle u'(W^*)u''(w)\Big\rangle_q\Big] + \frac{1}{2}\mathbb{E}\Big[u'(W^*)\Big\langle u'(W^*)u'(w)^2\Big\rangle_q\Big] \\ &\quad - \frac{1}{2}\mathbb{E}\Big[u'(W^*)\Big\langle u'(W^*)u'(w)\Big\rangle_q\Big\langle u'(w)\Big\rangle_q\Big]\end{aligned} \tag{7.8.4}$$

$$(B) = \mathbb{E}\Big\langle u''(W^*)u''(w)\Big\rangle_q + \mathbb{E}\Big\langle u''(W^*)u'(w)^2\Big\rangle_q - \mathbb{E}\Big\langle u''(W^*)u'(w)\Big\rangle_q\Big\langle u'(w)\Big\rangle_q \tag{7.8.5}$$

$$\begin{aligned}(D) &= -\mathbb{E}\Big\langle\Big(\frac{1}{2\sqrt{q}}V - \frac{1}{2\sqrt{\rho-q}}W^*\Big)u'(W^*)u'(w^{(1)})u'(w^{(2)})\Big\rangle_q \\ &= -\mathbb{E}\langle u'(W^*)u''(w^{(1)})u'(w^{(2)})\rangle_q - \mathbb{E}\langle u'(W^*)u'(w^{(1)})^2u'(w^{(2)})\rangle_q \\ &\quad + \mathbb{E}\langle u'(W^*)u'(w^{(1)})u'(w^{(2)})\rangle_q\langle u'(w)\rangle_q\end{aligned} \tag{7.8.6}$$

We now replace (7.8.4), (7.8.5) and (7.8.6) in (7.8.3):

$$\begin{aligned}2\Psi''_{P_{\rm out}}(q) &= \mathbb{E}\Big\langle u'(W^*)^2u''(w)\Big\rangle_q + \mathbb{E}\Big\langle u'(W^*)^2u'(w)^2\Big\rangle_q - \mathbb{E}\Big\langle u'(W^*)^2u'(w^{(1)})u'(w^{(2)})\Big\rangle_q\\ &+ \mathbb{E}\Big\langle u''(W^*)u''(w)\Big\rangle_q + \mathbb{E}\Big\langle u''(W^*)u'(w)^2\Big\rangle_q - \mathbb{E}\Big\langle u''(W^*)u'(w^{(1)})u'(w^{(2)})\Big\rangle_q\\ &- \mathbb{E}\langle u'(W^*)u''(w^{(1)})u'(w^{(2)})\rangle_q - \mathbb{E}\langle u'(W^*)u'(w^{(1)})^2u'(w^{(2)})\rangle_q + \mathbb{E}\langle u'(w)\rangle_q^4\,.\end{aligned}$$

Using the identity $u''_Y(x) + u'_Y(x)^2 = \frac{P''_{\rm out}(Y|x)}{P_{\rm out}(Y|x)}$, this factorizes and gives

$$\Psi''_{P_{\rm out}}(q) = \frac{1}{2}\mathbb{E}\Big[\Big(\Big\langle\frac{P''_{\rm out}(Y|\sqrt{q}\,V+\sqrt{\rho-q}\,w)}{P_{\rm out}(Y|\sqrt{q}\,V+\sqrt{\rho-q}w)}\Big\rangle_q - \Big\langle u'_{Y^{(q)}}(\sqrt{q}\,V+\sqrt{\rho-q}\,w)\Big\rangle_q^2\Big)^2\Big] \geq 0\,. \quad (7.8.7)$$

$\Psi_{P_{\rm out}}$ is thus convex on $[0,\rho]$. It is not difficult to verify (by standard arguments of continuity under the integral) that $\Psi''_{P_{\rm out}}$ is continuous on $[0,\rho]$, which gives that $\Psi_{P_{\rm out}}$ is $\mathcal{C}^2$ on its domain. □

Proposition 7.8.2

Suppose that for all $x \in \mathbb{R}$, $P_{\rm out}(\cdot\,|\,x)$ is the law of $\varphi(x,A)+\sqrt{\Delta}Z$ where $\varphi : \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is a measurable function and $(Z,A) \sim \mathcal{N}(0,1)\otimes P_A$, for some probability distribution P_A over $\mathbb{R}$. Assume also that

$$\mathbb{E}[\varphi(\sqrt{\rho}Z,A)^2] < \infty\,, \quad (7.8.8)$$

and that we are in one of the following cases:

(i) $\Delta > 0$.
(ii) $\Delta = 0$ *and* φ *takes values in* $\mathbb{N}$.

Then $q \mapsto \Psi_{P_{\rm out}}(q)$ is continuous, convex and non-decreasing over $[0,\rho]$.

Notice that (7.8.8) is for instance verified under hypotheses (h1)-(h2)-(h3)-(h4). Indeed, by the central limit theorem (that we apply under (h1)-(h3)), $\frac{1}{\sqrt{n}}\langle \boldsymbol{X}^*, \boldsymbol{\Phi}_1\rangle \xrightarrow[n\to\infty]{\rm (d)} \mathcal{N}(0,\rho)$. Then using (h4) we get that

$$\varphi\Big(\frac{1}{\sqrt{n}}\langle \boldsymbol{X}^*, \boldsymbol{\Phi}_1\rangle, A_1\Big) \xrightarrow[n\to\infty]{\rm (d)} \varphi(\sqrt{\rho}Z,A).$$

Finally, by (h2) the sequence above is bounded in L^2, hence is limit has a finite second moment which proves (7.8.8).

Proof. We deduce Proposition 7.8.2 from Proposition 7.8.1 above by an approximation procedure. Since $\Psi_{P_{\rm out}} = \Psi_{P_{\rm out}}(\rho) - \mathcal{I}_{P_{\rm out}}$, we will work with the mutual information $\mathcal{I}_{P_{\rm out}}$. Let us define $U^{(q)} = \varphi(\sqrt{q}\,V + \sqrt{\rho-q}\,W^*, A)$ and $Y^{(q)} = U^{(q)} + \sqrt{\Delta}Z$.

We start by proving Proposition 7.8.2 under the assumption (i). Let $\epsilon > 0$. By density of the $\mathcal{C}^\infty$ functions with compact support in L^2 (see for instance Corollary 4.2.2 from [34]), one can find a $\mathcal{C}^\infty$ function $\widehat{\varphi}$ with compact support, such that

$$\mathbb{E}\Big[\big(\varphi(\sqrt{\rho}\,Z,A) - \widehat{\varphi}(\sqrt{\rho}\,Z,A)\big)^2\Big] \leq \epsilon^2\,.$$

Let us write $\widehat{U}^{(q)} = \widehat{\varphi}(\sqrt{q}\,V + \sqrt{\rho-q}\,W^*, A)$ and $\widehat{Y}^{(q)} = \widehat{U} + \sqrt{\Delta}Z$. We have by the chain rule for the mutual information

$$\begin{aligned}I(U^{(q)};Y^{(q)}|V) &= I(W^*,U^{(q)};Y^{(q)}|V) = I(U^{(q)};Y^{(q)}|V,W^*) + I(W^*;Y^{(q)}|V)\\ &= I(U^{(q)};Y^{(q)}|V,W^*) + \mathcal{I}_{P_{\rm out}}(q)\end{aligned}$$

and similarly, $\mathcal{I}_{\widehat{P}_{\rm out}}(q) = I(\widehat{U}^{(q)};\widehat{Y}^{(q)}|V) - I(\widehat{U}^{(q)};\widehat{Y}^{(q)}|V,W^*)$. By Corollary B.1, there exists a constant $C>0$ such that

$$|I(\widehat{U}^{(q)};\widehat{Y}^{(q)}|V) - I(U^{(q)};Y^{(q)}|V)| \le C\epsilon \quad \text{and} \quad |I(\widehat{U}^{(q)};\widehat{Y}^{(q)}|V,W^*) - I(U^{(q)};Y^{(q)}|V,W^*)| \le C\epsilon\,.$$

We get that for all $q\in[0,\rho]$, $|\mathcal{I}_{P_{\rm out}}(q) - \mathcal{I}_{\widehat{P}_{\rm out}}(q)| \le C\epsilon$. The function $\mathcal{I}_{P_{\rm out}}$ can therefore be uniformly approximated by continuous, concave, non-increasing functions on $[0,\rho]$: $\mathcal{I}_{P_{\rm out}}$ is therefore continuous, concave and non-increasing.

Let us now prove Proposition 7.8.2 under the assumption (ii). Under this assumption we have $\mathcal{I}_{P_{\rm out}}(q) = I(W^*;U^{(q)}|V)$ and by the case (i) we know that the function $i_\Delta(q) = I(W^*;U^{(q)} + \sqrt{\Delta}Z|V)$ is concave and non-increasing for all $\Delta>0$. By Corollary B.2 we obtain that for all $q\in[0,\rho]$ and all $\Delta\in(0,1]$ we have

$$|\mathcal{I}_{P_{\rm out}}(q) - i_\Delta(q)| \le 100 e^{-1/(16\Delta)}\,,$$

which proves (by uniform approximation) that $\mathcal{I}_{P_{\rm out}}$ is continuous, concave and non-increasing. □

Proposition 7.8.3

Under the same hypotheses than Proposition 7.8.2 above, $\Psi_{\rm out}$ is differentiable over $[0,\rho)$ and for all $q\in[0,\rho)$

$$\Psi'_{P_{\rm out}}(q) = \frac{1}{2(\rho-q)}\mathbb{E}\langle w\rangle_q^2\,,$$

where we recall that $\langle\cdot\rangle_q$ is defined by (7.8.2).

Proof. The fact that $\Psi_{P_{\rm out}}$ is differentiable on $(0,\rho)$ follows from differentiation under the expectation sign. In order to see it, we define $X = \sqrt{q}\,V + \sqrt{\rho-q}\,W^*$. Then, for all $q\in[0,\rho)$:

$$\Psi_{P_{\rm out}}(q) = \mathbb{E}\int dX \frac{e^{-\frac{(X-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}} \int P_{\rm out}(dy|X)\log\int dx \frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}} P_{\rm out}(y|x)\,. \tag{7.8.9}$$

We are now in a good setting to differentiate under the expectation sign. We have for all $q\in(0,\rho)$,

$$\frac{\partial}{\partial q}\left[\frac{e^{-\frac{(X-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{\rho-q}}\right] = \frac{e^{-\frac{(X-\sqrt{q}\,V)^2}{2(\rho-q)}}}{2\sqrt{\rho-q}}\Big(\frac{1}{\rho-q} - \frac{(X-\sqrt{q}V)^2}{(\rho-q)^2} + \frac{V(X-\sqrt{q}V)}{\sqrt{q}(\rho-q)}\Big)\,. \tag{7.8.10}$$

Thus

$$\begin{aligned}\Psi'_{P_{\rm out}}(q) = \frac{1}{2}\mathbb{E}\left[\Big(\frac{1}{\rho-q} - \frac{(X-\sqrt{q}\,V)^2}{(\rho-q)^2} + \frac{V(X-\sqrt{q}\,V)}{\sqrt{q}\,(\rho-q)}\Big)\log\int dx\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\rm out}(Y|x)\right]\\ + \frac{1}{2}\mathbb{E}\left\langle \frac{1}{\rho-q} - \frac{(x-\sqrt{q}\,V)^2}{(\rho-q)^2} + \frac{V(x-\sqrt{q}\,V)}{\sqrt{q}\,(\rho-q)}\right\rangle_q\end{aligned}$$

where the Gibbs brackets $\langle\cdot\rangle_q$ denotes the expectation with respect to $x\sim P(X|Y^{(q)},V)$. The second term of the sum above is equal to zero. Indeed by the Nishimori identity (Proposition 1.1.1):

$$\begin{aligned}\mathbb{E}\left\langle \frac{1}{\rho-q} - \frac{(x-\sqrt{q}\,V)^2}{(\rho-q)^2} + \frac{V(x-\sqrt{q}\,V)}{\sqrt{q}\,(\rho-q)}\right\rangle_q &= \mathbb{E}\left[\frac{1}{\rho-q} - \frac{(X-\sqrt{q}\,V)^2}{(\rho-q)^2} + \frac{V(X-\sqrt{q}\,V)}{\sqrt{q}\,(\rho-q)}\right]\\ &= \frac{1}{\rho-q}\mathbb{E}\left[1-(W^*)^2\right] = 0\,.\end{aligned}$$

We now compute, by Gaussian integration by parts with respect to $V \sim \mathcal{N}(0,1)$:

$$\mathbb{E}\left[\frac{V(X-\sqrt{q}\,V)}{\sqrt{q}\,(\rho-q)}\log\int dx\frac{1}{\sqrt{2\pi(\rho-q)}}e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}P_{\text{out}}(Y^{(q)}|x)\right]$$
$$=\mathbb{E}\left[\frac{-1}{\rho-q}\log\int dx\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\text{out}}(Y^{(q)}|x)\right]+\mathbb{E}\left[\frac{(X-\sqrt{q}\,V)^2}{(\rho-q)^2}\log\int dx\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\text{out}}(Y^{(q)}|x)\right]$$
$$+\mathbb{E}\left\langle\frac{(X-\sqrt{q}\,V)(x-\sqrt{q}\,V)}{(\rho-q)^2}\right\rangle_q.$$

Bringing all together, we conclude:

$$\Psi'_{P_{\text{out}}}(q)=\frac{1}{2}\mathbb{E}\left\langle\frac{(X-\sqrt{q}\,V)(x-\sqrt{q}\,V)}{(\rho-q)^2}\right\rangle_q=\frac{1}{2(\rho-q)}\mathbb{E}\langle w\rangle_q^2\,.$$

It remains to show that $\Psi'_{P_{\text{out}}}(0)=\frac{1}{2\rho}\mathbb{E}\langle w\rangle_0^2=0$. This follows from taking the $q\to 0$ limit in the equation above. □

Proposition 7.8.4

Assume that the hypotheses of Proposition 7.8.2 hold and suppose also that the kernel P_{out} is informative. Then $\Psi_{P_{\text{out}}}$ is strictly increasing on $[0,\rho]$.

Proof. By contradiction we suppose that $\Psi_{P_{\text{out}}}$ is not strictly increasing on $[0,\rho]$. There exists thus $q\in(0,\rho)$ such that $\Psi'_{P_{\text{out}}}(q)=0$. By Proposition 7.8.3 this means that $\langle w\rangle_q=0$ almost surely and therefore that

$$\int_{\mathbb{R}}P_{\text{out}}(Y^{(q)}\,|\,\sqrt{q}\,V+\sqrt{\rho-q}\,w)we^{-w^2/2}dw=0$$

almost surely. Let us write $\sigma=\sqrt{\rho-q}$. Consequently,

$$\int_{\mathbb{R}}P_{\text{out}}(y\,|\,v+\sigma w)we^{-w^2/2}dw=0 \tag{7.8.11}$$

for almost all y in $\mathbb{R}$ (if we are under assumption (i)) or all $y\in\mathbb{N}$ (under assumption (ii)) and almost all $v\in\mathbb{R}$. We will now use the following lemma:

Lemma 7.8.1

Let $Z\sim\mathcal{N}(0,1)$ and let $f:\mathbb{R}\to\mathbb{R}$ be a bounded function. Suppose that for almost all $v\in\mathbb{R}$,

$$\mathbb{E}[Zf(v+Z)]=0\,.$$

Then, there exists a constant $C\in\mathbb{R}$ such that $f(v)=C$ for almost every v.

Proof. Let us define the function

$$h:t\mapsto\mathbb{E}[f(Z-t)]=\frac{1}{\sqrt{2\pi}}\int f(x)e^{-(x+t)^2/2}dx\,.$$

We have $h'(t)=\frac{-1}{\sqrt{2\pi}}\int f(x)(x+t)e^{-(x+t)^2/2}dx=-\mathbb{E}[Zf(Z-t)]=0$ and therefore h is equal to some constant $C\in\mathbb{R}$. We are going to show that $f=C$ almost everywhere. Without loss of generality we can assume that $C=0$, otherwise it suffices to consider the function $\tilde{f}=f-C$. Now we have for all $n\geq 0$, $t\in\mathbb{R}$

$$0=h^{(n)}(t)=\frac{1}{\sqrt{2\pi}}\int f(x)\frac{\partial}{\partial t}e^{-(x+t)^2/2}dx=\frac{1}{\sqrt{2\pi}}\int f(x)(-1)^nH_n(x+t)e^{-(x+t)^2/2}dx\,,$$

where H_n is n^{th} Hermite polynomial, defined as $H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}$. Therefore, for all $n \geq 0$,

$$\int f(x) H_n(x) e^{-x^2/2} dx = 0\,,$$

which implies that $f = 0$ almost everywhere since the Hermite functions form an orthonormal basis of $L^2(\mathbb{R})$. □

We apply now Lemma 7.8.1 to (7.8.11) where the function f is given by $f(x) = P_{\text{out}}(y \,|\, \sigma x)$. We thus obtain that for almost every y, $P_{\text{out}}(y \,|\, \cdot)$ is almost everywhere equal to a constant: this contradicts the assumption that P_{out} is informative. □

7.8.2 Study of the generalization function

We turn now our attention to the study of the following "generalization function":

$$\mathcal{E}_f : \left| \begin{array}{ccc} [0,\rho] & \to & \mathbb{R}_{\geq 0} \\ q & \mapsto & \mathbb{E}\Big[\big(f(Y^{(q)}) - \mathbb{E}[f(Y^{(q)})|V]\big)^2\Big] \end{array} \right. \tag{7.8.12}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a continuous bounded function. We will prove that $\mathcal{E}_f$ is continuous (Proposition 7.8.5) and strictly decreasing (Proposition 7.8.6) under the following hypotheses.

(a) For all $x \in \mathbb{R}$, $P_{\text{out}}(\cdot \,|\, x)$ is the law of $\varphi(x, A) + \sqrt{\Delta}\, Z$ where $\varphi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a measurable function and $(Z, A) \sim \mathcal{N}(0,1) \otimes P_A$, for some probability distribution P_A over $\mathbb{R}$.
(b) For almost all $a \in \mathbb{R}$ (w.r.t. P_A), $\varphi(\cdot, a)$ is continuous almost everywhere.

We suppose also that we are in one of the following cases:

(i) $\Delta > 0$.
(ii) $\Delta = 0$ and φ takes values in $\mathbb{N}$.

Proposition 7.8.5

Under the hypotheses presented above, $\mathcal{E}_f$ is continuous on $[0, \rho]$.

Proof. Notice that

$$\mathcal{E}_f(q) = \mathbb{E}[f(\varphi(\sqrt{\rho}\, V, A) + \sqrt{\Delta} Z)^2] - \mathbb{E}_V\big[\mathbb{E}_{W^*,Z,A}\big[f(\varphi(\sqrt{q}\, V + \sqrt{\rho - q}\, W^*, A) + \sqrt{\Delta} Z)\big]^2\big].$$

The first term does not depend on q and the second one is continuous by Lebesgue's convergence theorem. □

Proposition 7.8.6

Assume that the hypotheses of Proposition 7.8.5 hold. Suppose that $x \mapsto \int f(y) P_{\text{out}}(dy \,|\, x)$ is not almost-everywhere equal to a constant. Then $\mathcal{E}_f$ is strictly decreasing on $[0, \rho]$.

Proof. $\mathcal{E}_f(q) = \mathbb{E}[f(Y^{(q)})^2] - \mathbb{E}\big[\mathbb{E}[f(Y^{(q)})|V]^2\big]$. Since the first term does not depend on q, it suffices to show that $H : q \mapsto \mathbb{E}\big[\mathbb{E}[f(Y^{(q)})|V]^2\big]$ is strictly increasing on $[0, \rho]$. We have for $q \in (0, \rho)$:

$$\mathbb{E}[f(Y^{(q)})|V] = \int f(y) \frac{e^{-w^2/2}}{\sqrt{2\pi}} P_{\text{out}}(dy|\sqrt{q}\, V + \sqrt{\rho - q}\, w) dw = \int f(y) \frac{e^{-\frac{(x - \sqrt{q} V)^2}{2(\rho - q)}}}{\sqrt{2\pi(\rho - q)}} P_{\text{out}}(dy|x) dx\,.$$

So we have, using (7.8.10):

$$\frac{\partial}{\partial q}\mathbb{E}[f(Y^{(q)})|V] = \int \frac{f(y)}{2}\Big(\frac{1}{\rho-q} - \frac{(x-\sqrt{q}\,V)^2}{(\rho-q)^2} + \frac{V(x-\sqrt{q}\,V)}{\sqrt{q}(\rho-q)}\Big)\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\mathrm{out}}(dy|x)dw$$
$$= \frac{1}{2(\rho-q)}\mathbb{E}\left[f(Y^{(q)})\Big(1 - W^{*2} + \frac{\sqrt{\rho-q}\,VW^*}{\sqrt{q}}\Big)\Big|V\right].$$

We obtain

$$H'(q) = \frac{1}{\rho-q}\mathbb{E}\left[\mathbb{E}[f(Y^{(q)})|V]\,\mathbb{E}\left[f(Y^{(q)})\Big(1 - W^{*2} + \frac{\sqrt{\rho-q}\,VW^*}{\sqrt{q}}\Big)\Big|V\right]\right]. \tag{7.8.13}$$

We compute by Gaussian integration by parts:

$$\mathbb{E}\left[\mathbb{E}[f(Y^{(q)})|V]\,\mathbb{E}\left[f(Y^{(q)})VW^*\Big|V\right]\right] = \mathbb{E}\left[V\mathbb{E}[f(Y^{(q)})|V]\,\mathbb{E}\left[f(Y^{(q)})W^*\Big|V\right]\right]$$
$$= \mathbb{E}\left[\frac{\partial}{\partial V}\mathbb{E}[f(Y^{(q)})|V]\,\mathbb{E}\left[f(Y^{(q)})W^*\Big|V\right]\right] + \mathbb{E}\left[\mathbb{E}[f(Y^{(q)})|V]\,\frac{\partial}{\partial V}\mathbb{E}\left[f(Y^{(q)})W^*\Big|V\right]\right]. \tag{7.8.14}$$

We compute successively

$$\frac{\partial}{\partial V}\mathbb{E}[f(Y^{(q)})|V] = \frac{\partial}{\partial V}\int f(y)\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\mathrm{out}}(dy|x)dx$$
$$= \int f(y)\frac{\sqrt{q}\,(x-\sqrt{q}\,V)}{\rho-q}\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\mathrm{out}}(dy|x)dx$$
$$= \frac{\sqrt{q}}{\sqrt{\rho-q}}\mathbb{E}\Big[f(Y^{(q)})W^*\Big|V\Big]. \tag{7.8.15}$$

$$\frac{\partial}{\partial V}\mathbb{E}[f(Y^{(q)})W^*|V] = \frac{\partial}{\partial V}\int f(y)\frac{x-\sqrt{q}\,V}{\sqrt{\rho-q}}\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\mathrm{out}}(dy|x)dx$$
$$= \int f(y)\Big(\frac{-\sqrt{q}}{\sqrt{\rho-q}} + \frac{\sqrt{q}(x-\sqrt{q}\,V)^2}{(\rho-q)^{3/2}}\Big)\frac{e^{-\frac{(x-\sqrt{q}\,V)^2}{2(\rho-q)}}}{\sqrt{2\pi(\rho-q)}}P_{\mathrm{out}}(dy|x)dx$$
$$= \frac{\sqrt{q}}{\sqrt{\rho-q}}\mathbb{E}\Big[f(Y^{(q)})(-1 + W^{*2})\Big|V\Big]. \tag{7.8.16}$$

By plugging (7.8.14)-(7.8.15)-(7.8.16) back in (7.8.13) we get:

$$H'(q) = \frac{1}{\rho-q}\mathbb{E}\Big[\mathbb{E}[f(Y^{(q)})W^*|V]^2\Big] \geq 0\,.$$

Let us suppose now that H is not strictly increasing on $[0,\rho]$. This means that we can find $q \in (0,\rho)$ such that $H'(q) = 0$ and therefore $\mathbb{E}[f(Y^{(q)})W^*|V] = 0$ almost surely. This gives that for almost all $v \in \mathbb{R}$,

$$\mathbb{E}\Big[W\int f(y)P_{\mathrm{out}}(dy|\sqrt{q}\,v + \sqrt{\rho-q}\,W)\Big] = 0\,,$$

where $\mathbb{E}$ is the expectation with respect to $W \sim \mathcal{N}(0,1)$. Lemma 7.8.1 gives then that the function $x \mapsto \int f(y)P_{\mathrm{out}}(dy|x)$ is almost everywhere equal to a constant: we obtain a contradiction. We conclude that H is strictly increasing on $[0,\rho]$ and thus $\mathcal{E}_f$ is strictly decreasing on $[0,\rho]$. □

Proposition 7.8.7

Assume that the hypotheses of Proposition 7.8.5 hold. If the channel P_{out} is informative, then there exists a continuous bounded function $f : \mathbb{R} \to \mathbb{R}$ such that $x \mapsto \int f(y)P_{\mathrm{out}}(dy|x)$ is not almost everywhere equal to a constant.

Proof. By contradiction, let us suppose that for all continuous bounded function $f : \mathbb{R} \to \mathbb{R}$ we have

$$\int f(y) P_{\text{out}}(dy|x) = C_f$$

for almost all $x \in \mathbb{R}$, for some constant $C_f \in \mathbb{R}$. Let $X \sim \mathcal{N}(0,1)$ and $Y \sim P_{\text{out}}(\cdot|X)$. We have then $\mathbb{E}[f(Y)|X] = C_f = \mathbb{E}[f(Y)]$ almost surely. Let $g : \mathbb{R} \to \mathbb{R}$ be another continuous bounded function and compute:

$$\mathbb{E}[g(X)f(Y)] = \mathbb{E}\big[g(X)\mathbb{E}[f(Y)|X]\big] = \mathbb{E}[g(X)]\mathbb{E}[f(Y)]\,.$$

It follows that X and Y are independent: The measures $P_{\text{out}}(dy|x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx$ and $\mathbb{E}[P_{\text{out}}(dy|X)]\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx$ are therefore equal. Consequently, for almost every x, y we have

$$P_{\text{out}}(y|x) = \mathbb{E}[P_{\text{out}}(y|X)]\,.$$

This gives that for almost every y, $P_{\text{out}}(y|\cdot)$ is almost everywhere equal to a constant. We conclude that P_{out} is not informative, which is a contradiction. □

Chapter 8

The distribution of the Lasso: Uniform control over sparse balls and adaptive parameter tuning

8.1 Introduction to the Lasso

Given data (x_i, y_i), $1 \le i \le n$, with $x_i \in \mathbb{R}^N$, $y_i \in \mathbb{R}$, the Lasso [206, 46] fits a linear model by minimizing the cost function

$$\begin{aligned} \mathcal{L}_\lambda(\theta) &= \frac{1}{2n} \sum_{i=1}^{n} \Big(y_i - \langle x_i, \theta \rangle \Big)^2 + \frac{\lambda}{n} |\theta| \\ &= \frac{1}{2n} \|y - X\theta\|^2 + \frac{\lambda}{n} |\theta| \,. \end{aligned} \tag{8.1.1}$$

Here $X \in \mathbb{R}^{n \times N}$ is the matrix with rows $x_1, \dots, x_n$, $y = (y_1, \dots, y_n)$, $\|v\|$ denotes the ℓ_2 norm of vector v, and $|v|$ its ℓ_1 norm. To fix normalizations, we will assume that the columns of X have ℓ_2 norm $1 + o(1)$. (Note that this normalization is different from the one that is sometimes adopted in the literature, but the two are completely equivalent.)

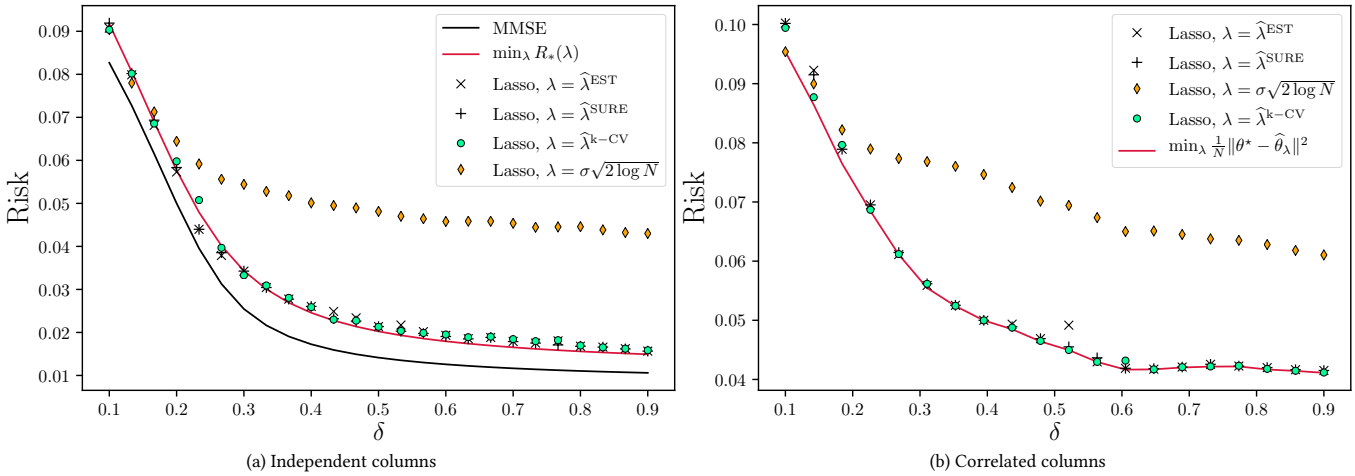
A large body of theoretical work supports the use of ℓ_1 regularization in the high-dimensional regime $n \lesssim N$, when only a small subset of the coefficients θ are expected to be large. Broadly speaking, we can distinguish two types of theoretical approaches. A first line of work makes deterministic assumptions about the design matrix X, such as the restricted isometry property and its generalizations [42, 39]. Under such conditions, minimax optimal estimation rates as well as oracle inequalities have been proved in a remarkable sequence of papers [41, 33, 210, 162, 178]. As an example, assume that that the linear model is correct. Namely,

$$y = X\theta^\star + \sigma z \,, \tag{8.1.2}$$

for $\sigma \ge 0$, $z \sim \mathcal{N}(0, \mathrm{Id}_n)$, and $\theta^\star$ a vector with s_0 non-zero entries. Then, a theorem of Bickel, Ritov and Tsybakov [33] implies that, with high probability,

$$\lambda \ge \sigma \sqrt{c_0 \log N} \quad \Rightarrow \quad \|\widehat{\theta}_\lambda - \theta^\star\|^2 \le C s_0 \lambda^2 \,, \tag{8.1.3}$$

for some constants c_0, C that depend on the specific assumptions on the design. (The normalization of [33] is recovered by setting $\sigma^2 = \sigma_\#^2/n$, where $\sigma_\#^2$ is the noise variance of [33].)



(a) Independent columns

(b) Correlated columns

Figure 8.1: Estimation risk of the Lasso for different choices of λ, as a function of δ. $N = 8000$. In both plots, $\sigma = 0.2$. The true coefficients vector $\theta^\star$ is chosen to be sN-sparse with $s = 0.1$. The entries on the support of $\theta^\star$ are drawn i.i.d. $\mathcal{N}(0,1)$. Cross-validation is carried out using 4 folds. SURE is computed using the estimator $\widehat{\sigma}$ for the plot on the left, and the true value of σ on the right.
Left: A standard random design with $(X_{ij}) \sim_{iid} \mathcal{N}(0, 1/n)$.
Right: The rows of the design matrix X are i.i.d. Gaussian, with correlation structure given by an autoregressive process, see Eq. (8.4.4). Here we used $\phi = 2$.

Unfortunately, this analysis provides limited insight into the choice of the regularization parameter λ which –in practice– can impact significantly the estimation accuracy. As an example, Figure 8.1.3 reports the result of a small simulation in which we compare four different methods of selecting λ. The bound of (8.1.3) suggests to set $\lambda = \sigma\sqrt{c_0 \log N}$. For the standard random design used in the left frame, the optimal constant is expected to be $c_0 = 2$ [69, 72]. We compare this method to three procedures that adapt the choice of λ to the data: cross validation (CV), Stein's Unbiased Risk Estimate (SURE), and a procedure that minimizes an estimate of the risk (EST). We refer to the next sections for further details on these methods. Note that all of these adaptive procedures significantly outperform the 'theory driven' λ: over a broad range of sample sizes n, the resulting estimation error is 2 to 3 times smaller. Further, the error achieved by these methods is quite close to the Bayes optimum.

These empirical observations are not captured by the bound (8.1.3), or by similar results.

An alternative style of analysis postulates an idealized model for the data and derives asymptotically exact results. Throughout this paper we will consider the simplest of such models, by assuming that design matrix to have i.i.d. entries $X_{ij} \sim \mathcal{N}(0, 1/n)$. While this assumption is likely to be violated in practice, it allows to derive useful insights that are mathematically consistent, and susceptible of being generalized to a broader context. This type of analysis was first carried out in the context of the Lasso in [27] and then extended to a number of other problems, see e.g. [121, 205, 67, 204, 81, 198]. As an example, Figure 8.1 reports the predictions of this analysis for the risk of the three adaptive procedure for selecting λ. The agreement with the numerical simulations is excellent.

Unfortunately, the results in [27] (and in follow-up work) do not allow to derive in a mathematically rigorous way curves such as the ones in Figure 8.1. In fact earlier results hold 'pointwise' over λ and hence do not apply to adaptive procedures to select λ. Further they provide asymptotic estimates 'pointwise' over θ, and hence do not allow to compute

–for instance– minimax risk.

In order to clarify these points, it is useful to overview informally the picture emerging from [27, 71]. Fix $\theta \in \mathbb{R}^N$, $\lambda \in \mathbb{R}_{>0}$, and let $\eta(x;b) = (|x| - b)_+ \mathrm{sign}(x)$ be the soft thresholding function. By the KKT conditions the Lasso estimator $\widehat{\theta}_\lambda$ satisfies

$$\widehat{\theta}_\lambda = \eta\Big(\widehat{\theta}^d_\lambda; \alpha\tau\Big)\,, \qquad \widehat{\theta}^d_\lambda = \widehat{\theta}_\lambda + \frac{\alpha\tau}{\lambda} X^{\mathsf{T}}(y - X\widehat{\theta}_\lambda)\,, \tag{8.1.4}$$

where the vector $\widehat{\theta}^d_\lambda$ is also referred to as the 'debiased Lasso' [221, 209, 113]. The above identity holds for arbitrary $\alpha, \tau > 0$. However, [27] predicts that the distribution of the debiased estimator $\widehat{\theta}^d_\lambda$ simplifies dramatically for specific choices of these parameters.

Namely, let Θ be a random variable with distribution given by the empirical distribution of $(\theta_i)_{i\le N}$ (i.e., $\Theta = \theta_i$ with probability $1/N$, for $i \in \{1, \dots, N\}$) and let $Z \sim \mathcal{N}(0,1)$ be independent of Θ. Define α_*, τ_* to be the solution of the following system of equations (we refer to Section 8.3.1 for a discussion of existence and uniqueness):

$$\begin{cases} \tau^2 \;=\; \sigma^2 + \frac{1}{\delta}\mathbb{E}\Big[(\eta(\Theta + \tau Z, \alpha\tau) - \Theta)^2\Big]\,, \\ \lambda \;=\; \alpha\tau\left(1 - \frac{1}{\delta}\mathbb{P}\left(\big|\Theta + \tau Z\big| > \alpha\tau\right)\right). \end{cases} \tag{8.1.5}$$

When α, τ are selected in this way, $\widehat{\theta}^d_\lambda$ is approximately normal with mean $\theta^\star$ (the true parameters vector) and variance τ_*^2: $\widehat{\theta}^d \approx \mathcal{N}(\theta^\star, \tau_*^2 \mathrm{Id})$. More precisely, for any test function $\mathsf{f} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, with $|\mathsf{f}(x) - \mathsf{f}(y)| \le L(1 + \|x\| + \|y\|)\|x - y\|$, almost surely,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \mathsf{f}(\theta^\star_i, \widehat{\theta}^d_{\lambda,i}) = \mathbb{E}\Big[\mathsf{f}(\Theta, \Theta + \tau_* Z)\Big]\,, \tag{8.1.6}$$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \mathsf{f}(\theta^\star_i, \widehat{\theta}_{\lambda,i}) = \mathbb{E}\Big[\mathsf{f}(\Theta, \eta(\Theta + \tau_* Z; \alpha_*\tau_*))\Big]\,. \tag{8.1.7}$$

This is an asymptotic result, which holds along sequences of problems with: (i) Converging aspect ratio $n/N \to \delta \in (0, \infty)$; (ii) Fixed regularization $\lambda \in (0, \infty)$; (iii) Parameter vectors $\theta^\star = \theta^\star(n)$ whose empirical distribution converges (weakly) to a limit law p_Θ. As emphasized above, this does not allow deduce the behavior of the Lasso with adaptive choices of λ (there could be deviations from the above limits for exceptional values of λ), or to compute the minimax risk (there could be deviations for exceptional vectors $\theta^\star$).

The importance of establishing uniform convergence with respect to the regularization parameter λ was recently emphasized by [159]. Among other results, these authors derive a uniform convergence statement for the related approximate message passing (AMP) algorithm. However, in order to establish uniform convergence, they have to construct an ad-hoc smoothing of the quantity of interest, which is roughly equivalent to discretizing the corresponding tuning parameter.

Our goal here is to obtain uniform (in λ) convergence results for the Lasso, hence providing a sound mathematical basis to the comparison of various adaptive procedures, as well as to the study of minimax risk.

The next sections organized as follows. Section 8.2 reviews related work. We state our main theoretical results in Section 8.3. In Section 8.4 we apply these results to two types of statistical questions: estimating the risk and noise level, and selecting λ through adaptive procedures. Further, we illustrate our results in numerical simulations. Finally, Section 8.5 outlines the main proof ideas, with the most technical legwork deferred to Chapter 9.

8.2 Related work

There is –by now– a substantial literature on determining exact asymptotics in high-dimensional statistical models, and a number of mathematical techniques have been developed for this task. We will only provide a few pointers focusing on high-dimensional regression problems.

The original proof of [27] was based on an asymptotically exact analysis of an approximate message passing (AMP) algorithm [26] that was first proposed in [71] to minimize the Lasso cost function. Variants of AMP have been developed in a number of contexts, opening the way to the analysis of various statistical estimation problems. A short list includes generalized linear models [176], phase retrieval [188, 137], robust regression [67], logistic regression [198], generalized compressed sensing [31]. This approach is technically less direct than others, but has the advantage of providing an efficient algorithm, and is and not necessarily limited to convex problems (see [151] for a non-convex example).

As mentioned above, our work was partially motivated by the recent results of [159] that establish a form of uniformity for the AMP estimates –but not for the Lasso solution. It would be interesting to understand whether the approach of [159] could also be used to obtain uniform results for the Lasso or other statistical estimators.

Here we follow a different route that exploits powerful Gaussian comparison inequalities first proved by Gordon [93, 94]. Gordon inequality allows to bound the distribution of a minimax value, i.e. the value of a random variable $G_* = \min_{i\le N}\max_{j\le M} G_{ij}$, where $(G_{ij})_{i\le N,j\le M}$ is a Gaussian process, in terms of a similar quantity for a 'simpler' Gaussian process. The use of Gordon's inequality in this context was pioneered by Stojnic [196] and then developed by a number of authors in the context of regularized regression [205], M-estimation [204], generalized compressed sensing [5], binary compressed sensing [195] and so on. The key idea is to write the optimization problem of interest as a minimax problem, and then apply a suitable version of Gordon's inequality. A matching bound is obtained by convex duality and then a second application of Gordon's inequality. In particular, convexity of the cost function of interest is a crucial ingredient.

While the Gaussian comparison inequality provides direct access to the value of the optimization problem, understanding the properties of the estimator can be more challenging. We identify here a property (that we call *local stability*) that allows to transfer information on the minimum (the Lasso cost) into information about the minimizer (the Lasso estimator). We believe that this strategy can be applied to other examples beyond the Lasso.

Independently, a different approach based on leave-one-out techniques was developed by El Karoui in the context of ridge-regularized robust regression [121, 81].

Finally, a parallel line of research determines exact asymptotics for Bayes optimal estimation, under a model in which the coordinates of θ are i.i.d. with common distribution p_Θ. In particular, the asymptotic Bayes optimal error for linear regression with random designs was recently determined in [17, 179] and is also a Corollary of Theorem 7.3.2 from the previous chapter. Of course –in general– Bayes optimal estimation requires knowledge of the distribution p_Θ, and is not computationally efficient. We will use this Bayes-optimal error as a benchmark of our adaptive procedures.

8.3 Main results

8.3.1 Definitions

As stated above, we consider the standard linear model (8.1.2) where $y = X\theta^\star + \sigma z$, with noise $z \sim \mathcal{N}(0, \mathrm{Id}_n)$, and X a Gaussian design: $(X_{i,j})_{i\le n, j\le N} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/n)$. The Lasso estimator is defined by

$$\widehat{\theta}_\lambda = \arg\min_{\theta\in\mathbb{R}^N} \mathcal{L}_\lambda(\theta)\,. \tag{8.3.1}$$

(The minimizer is almost surely unique since the columns of X are in generic positions.) We set $\delta = n/N$ to be the number of samples per dimension. We are interested in uniform estimation over sparse vectors $\theta^\star$. Following [69, 115] we formalize this notion using ℓ_p-balls (which are convex sets only for $p \ge 1$).

Definition 8.3.1

Define for $p, \xi > 0$ the ℓ_p-ball

$$\mathcal{F}_p(\xi) = \left\{ x \in \mathbb{R}^N \,\middle|\, \frac{1}{N}\sum_{i=1}^N |x_i|^p \le \xi^p \right\},$$

and for $s \in [0,1]$

$$\mathcal{F}_0(s) = \left\{ x \in \mathbb{R}^N \,\middle|\, \|x\|_0 \le sN \right\}.$$

By Jensen's inequality we have for $p \ge p' > 0$, $\mathcal{F}_p(\xi) \subset \mathcal{F}_{p'}(\xi)$.

Let $\phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$ be the standard Gaussian density and $\Phi(x) = \int_{-\infty}^x \phi(t)dt$ be the associated cumulative function. In the case of ℓ_0 balls (sparse vectors), a crucial role is played by the following sparsity level.

Definition 8.3.2

Define the critical sparsity as

$$s_{\max}(\delta) = \delta \max_{\alpha\ge 0} \left\{ \frac{1 - \frac{2}{\delta}\Big((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big)}{1 + \alpha^2 - 2\Big((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big)} \right\}.$$

See Figure 7.4 for a plot of $s_{\max}$. The critical sparsity curve first appears in the seminal work by Donoho and Tanner on compressed sensing [73, 68]. These authors consider the noiseless case ($z = 0$) of model (8.1.2) and reconstruction via ℓ_1 minimization (which corresponds to the $\lambda \to 0$ limit of the Lasso). They prove that ℓ_1 minimization reconstructs exactly $\theta^\star$ with high probability, if $\|\theta^\star\|_0 \le N(s_{\max}(\delta) - \varepsilon)$, and fails with high probability if $\|\theta^\star\|_0 \ge N(s_{\max}(\delta) + \varepsilon)$ (for any $\varepsilon > 0$). A second interpretation of the critical sparsity $s_{\max}(\delta)$ was given in [72, 208, 205]. For $\|\theta^\star\|_0 \le N(s_{\max}(\delta) - \varepsilon)$, the Lasso achieves stable reconstruction. Namely, there exists $M = M(s,\delta) < \infty$ for $s < s_{\max}(\delta)$, such that, if $\|\theta^\star\|_0 \le Ns$, $\|\widehat{\theta}_\lambda - \theta^\star\|_2 \le M(s,\delta)\sigma^2$. Our results provide a third interpretation: uniform limit laws for the Lasso will be obtained on ℓ_0 balls only for $s < s_{\max}(\delta)$.

A crucial role in our results is provided by the following max-min problem:

$$\max_{\beta\geq 0}\min_{\tau\geq\sigma} \quad \psi_\lambda(\beta,\tau)\,, \tag{8.3.2}$$
$$\psi_\lambda(\beta,\tau) \stackrel{\text{def}}{=} \left(\frac{\sigma^2}{\tau}+\tau\right)\frac{\beta}{2}-\frac{1}{2}\beta^2+\frac{1}{\delta}\mathbb{E}\min_{w\in\mathbb{R}}\left\{\frac{w^2}{2\tau}\beta-\beta Z w+\lambda|w+\Theta|-\lambda|\Theta|\right\}.$$

The expectation above is with respect to $(\Theta,Z)\sim\widehat{\mu}_{\theta^\star}\otimes\mathcal{N}(0,1)$, where $\widehat{\mu}_{\theta^\star}$ denotes the empirical distribution of the entries of the vector $\theta^\star$:

$$\widehat{\mu}_{\theta^\star}=\frac{1}{N}\sum_{i=1}^N\delta_{\theta_i^\star}\,.$$

Proposition 8.3.1

The max-min (8.3.2) is achieved at a unique couple $(\beta_*(\lambda),\tau_*(\lambda))$. Moreover, $(\beta_*(\lambda),\tau_*(\lambda))$ is also the unique couple $(\beta,\tau)\in(0,+\infty)^2$ that verify

$$\begin{cases}\tau^2 \;=\; \sigma^2+\frac{1}{\delta}\mathbb{E}\Big[(\eta(\Theta+\tau Z,\tau\frac{\lambda}{\beta})-\Theta)^2\Big]\\ \beta \;=\; \tau\left(1-\frac{1}{\delta}\mathbb{E}\left[\eta'(\Theta+\tau Z,\frac{\tau\lambda}{\beta})\right]\right).\end{cases} \tag{8.3.3}$$

We will also use the notation $\alpha_*(\lambda)=\lambda/\beta_*(\lambda)$ and

$$s_*(\lambda)=\mathbb{E}\Big[\eta'(\Theta+\tau_*(\lambda)Z,\tau_*(\lambda)\alpha_*(\lambda))\Big]=\mathbb{P}\Big(|\Theta+\tau_*(\lambda)Z|\geq\alpha_*(\lambda)\tau_*(\lambda)\Big). \tag{8.3.4}$$

We will sometimes omit the dependency on λ and write simply $\alpha_*,\beta_*,\tau_*,s_*$. The distribution μ_λ^* defined below will correspond (see Theorem 8.3.1 in the next section) to the limit of the empirical distribution of the entries of $(\widehat{\theta}_\lambda,\theta^\star)$.

Definition 8.3.3

Let $(\Theta,Z)\sim\widehat{\mu}_{\theta^\star}\otimes\mathcal{N}(0,1)$. We denote by μ_λ^* the law of the couple

$$\Big(\eta\Big(\Theta+\tau_*(\lambda)Z,\;\alpha_*(\lambda)\tau_*(\lambda)\Big),\;\Theta\Big)\,.$$

8.3.2 Results

We fix from now on $0<\lambda_{\min}\leq\lambda_{\max}$ and $\mathcal{D}\subset\mathbb{R}^N$ that can be either $\mathcal{F}_p(\xi)$ for some $\xi,p>0$, or $\mathcal{F}_0(s)$ for some $s<s_{\max}(\delta)$. Our uniformity domain is defined by $\Omega=\Big(\delta,\sigma,\mathcal{D},\lambda_{\min},\lambda_{\max}\Big)$. Namely, we will control $\widehat{\theta}_\lambda$ uniformly with respect to $\theta^\star\in\mathcal{D}$ and $\lambda\in[\lambda_{\min},\lambda_{\max}]$, with $n/N=\delta$. We will call *constant* any quantity that only depends on Ω. In absence of further specifications, C,c will be constants (that depend only on Ω) that are allowed to change from one line to another.

Our first result shows that the empirical distribution of the entries $\{(\widehat{\theta}_{\lambda,i},\theta_i^\star)\}_{i\leq N}$ is uniformly close to the model μ_λ^*. We quantify deviations using the Wasserstein distance. Recall that, given two probability measures μ,ν on $\mathbb{R}^d$ with finite second moment, their Wasserstein distance of order 2 is

$$W_2(\mu,\nu)=\left(\inf_{\gamma\in\mathcal{C}(\mu,\nu)}\int\|x-y\|_2^2\,\gamma(\mathrm{d}x,\mathrm{d}y)\right)^{1/2}, \tag{8.3.5}$$

where the infimum is taken over all couplings of μ and ν. Note that W_2 metrizes the convergence in Eq. (8.1.7). Namely $\lim_{n\to\infty} W_2(\mu_n, \mu_*) = 0$ if and only if, for any test function $\mathsf{f} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, with $|\mathsf{f}(x) - \mathsf{f}(y)| \le L(1 + \|x\| + \|y\|)\|x - y\|$, we have $\lim_{n\to\infty} \int \mathsf{f}(x)\mu_n(\mathrm{d}x) = \int \mathsf{f}(x)\mu_*(\mathrm{d}x)$ [213]. It provides therefore a natural way to extend earlier results to a non-asymptotic regime.

Theorem 8.3.1

Assume that $\mathcal{D} = \mathcal{F}_p(\xi)$ for some $\xi > 0$ and $p > 0$. Then there exists constants $C, c > 0$ that only depend on Ω, such that for all $\epsilon \in (0, \frac{1}{2}]$

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\left(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} W_2\Big(\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}, \mu^*_\lambda\Big)^2 \ge \epsilon \right) \le C\epsilon^{-\max(1,a)-1} N^{(1/p-1)_+} \exp\left(-cN\epsilon^2\epsilon^a \log(\epsilon)^{-2}\right),$$

where $a = \frac{1}{2} + \frac{1}{p}$.

Theorem 8.3.1 is proved in Section 9.3.2 of Chapter 9.

Remark 8.3.1. *It is worth emphasizing in what sense Theorem 8.3.1 is uniform with respect to $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and to $\theta^\star \in \mathcal{D}$:*

- Uniformity with respect to λ. *We bound (in probability) the maximum (over λ) deviation between the empirical distribution $\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}$ and the predicted distribution μ^*_λ. (The supremum over λ is 'inside' the probability.)*
- Uniformity with respect to $\theta^\star$. *We bound the maximum probability (over $\theta^\star$) of a deviation between $\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}$ and μ^*_λ. (The supremum over $\theta^\star$ is 'outside' the probability.)*

The reader might wonder whether it is possible to strengthen this result and bound the maximum deviation over $\theta^\star$ ('move the supremum over $\theta^\star$ inside'). The answer is negative. In particular, we can choose the support of $\theta^\star$ to coincide with a submatrix of X with atypically small minimum singular value. This will result in larger estimation error $\|\widehat{\theta}_\lambda - \theta^\star\|_2$, and hence in a large Wasserstein distance $W_2(\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}, \mu^*_\lambda)$.

Remark 8.3.2. *Note that Theorem 8.3.1 does not hold for ℓ_0 balls. This is probably a fundamental problem, since controlling W_2 distance uniformly over ℓ_0 balls is impossible even in the simple sequence model (or, equivalently, for orthogonal designs X). Namely, consider the case in which we observe $y_i = \theta^\star_i + z_i$, $i \le N$, where $(z_i)_{i\le N} \overset{i.i.d.}{\sim} \mathcal{N}(0, \tau^2_*)$, and we try to estimate $\theta^\star$ by computing $\widehat{\theta}_{\lambda,i} = \eta(y_i; \lambda)$. Then there are vectors $\theta^\star \in \mathcal{F}_0(s)$ such that the empirical law $\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}$ does not concentrate in Wasserstein distance around its expectation μ^*_λ, i.e. the law of $(\Theta, \eta(\Theta + Z; \lambda))$ for $G \sim \mathcal{N}(0, \tau_*)$.*

In order to see this, it is sufficient to consider the vector

$$\theta^\star = (N, 2N, \dots, kN, 0, \dots, 0)\,.$$

In Section 9.6.1 of Chapter 9, we prove that (for this choice of $\theta^\star$) there exists a constant c_0 such that $W_2(\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}, \mu^*_\lambda) \ge \sqrt{k/N}$ with probability at least $1 - e^{-c_0 k}$ for all N large enough.

We can think of several possibilities to overcome this intrinsic non-uniformity over ℓ_0 balls. One option would be to consider a weaker notion of distance between probability measures. Here we follow a different route, and prove uniform estimates over ℓ_0 balls for several specific quantities of interest. In order to state these results, we introduce the following quantities, which correspond to the risk and the prediction error (and are expressed in terms of the solution (τ_*, β_*) of (8.3.3))

$$R_*(\lambda) = \delta\Big(\tau_*(\lambda)^2 - \sigma^2\Big)\,, \tag{8.3.6}$$

$$P_*(\lambda) = \beta_*(\lambda)^2 + \frac{2\sigma^2}{\delta} s_*(\lambda) - \frac{\sigma^2}{\delta}\,. \tag{8.3.7}$$

Theorem 8.3.2

Assume here that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $0 \le s < s_{\max}(\delta)$ and $\xi > 0, p > 0$. There exists constants $C, c > 0$ that only depend on Ω, such that for all $\epsilon \in (0,1]$

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\bigg(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \Big(\frac{1}{N} \|\widehat{\theta}_\lambda - \theta^\star\|^2 - R_*(\lambda) \Big)^2 \ge \epsilon \bigg) \le \frac{C}{\epsilon^2} N^q e^{-cN\epsilon^2}, \tag{8.3.8}$$

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\bigg(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \Big(\frac{1}{n} \|y - X\widehat{\theta}_\lambda\|^2 - \beta_*(\lambda)^2 \Big)^2 \ge \epsilon \bigg) \le \frac{C}{\epsilon^2} N^q e^{-cN\epsilon^2}, \tag{8.3.9}$$

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\bigg(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \Big(\frac{1}{n} \|X(\theta^\star - \widehat{\theta}_\lambda)\|^2 - P_*(\lambda) \Big)^2 \ge \epsilon \bigg) \le \frac{C}{\epsilon^2} N^q e^{-cN\epsilon^2}, \tag{8.3.10}$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

The statement (8.3.8) is proved in Section 9.3.2, while (8.3.9)-(8.3.10) are proved in Section 9.4 of Chapter 9.

So far we focused on the Lasso estimator $\widehat{\theta}_\lambda$. The *debiased Lasso* estimator is defined as

$$\widehat{\theta}^d_\lambda = \widehat{\theta}_\lambda + \frac{X^\mathsf{T}(y - X\widehat{\theta}_\lambda)}{1 - \frac{1}{n}\|\widehat{\theta}_\lambda\|_0}\,.$$

This estimator plays a crucial role in the construction of confidence intervals and p-values [221, 209, 113, 199], and provide an explicit construction of the 'direct observations' model in the sense that $\widehat{\theta}^d_\lambda$ is approximately distributed as $\mathcal{N}(\theta^\star, \tau_*^2 \mathrm{Id})$. We let $\mu^{(d)}_\lambda$ be the law of the couple $\Big(\Theta + \tau_*(\lambda) Z,\ \Theta\Big)$, where $(\Theta, Z) \sim \hat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$.

Theorem 8.3.3

Let $\widehat{\mu}_{(\widehat{\theta}^d_\lambda, \theta^\star)}$ denote the empirical distribution (on $\mathbb{R}^2$) of the entries of $(\widehat{\theta}^d_\lambda, \theta^\star)$. There exists constants $c, C > 0$ such that for all $\epsilon \in (0,1]$,

$$\sup_{\theta^\star \in \mathcal{F}_4(\xi)} \mathbb{P}\bigg(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} W_2(\widehat{\mu}_{(\widehat{\theta}^d_\lambda, \theta^\star)}, \mu^{(d)}_\lambda) \ge \epsilon \bigg) \le \frac{C}{\epsilon^{11}} e^{-cN\epsilon^{17}}.$$

Theorem 8.3.3 is proved in Section 9.6.5 of Chapter 9.

8.4 Applications

8.4.1 Estimation of the risk and the noise level

In order to select the regularization parameter and to evaluate the quality of the Lasso solution $\widehat{\theta}_\lambda$, it is useful to estimate the risk and noise level. The paper [24] developed a suite of estimators of these quantities based on the asymptotic theory of [27]. The same paper also proposed generalizations of these estimators to correlated designs. Here we revisit these estimators and prove stronger guarantees. First, we obtain quantitative bound on the consistency rate of our estimators. Second, our results are uniform over λ, which justifies using these estimators to select λ.

Let us start with the estimation of $\tau_*(\lambda)$ which plays a crucial role in the asymptotic theory. We define

$$\widehat{\tau}(\lambda) = \sqrt{n}\frac{\|y - X\widehat{\theta}_\lambda\|}{n - \|\widehat{\theta}_\lambda\|_0}\,.$$

We will see with Theorem 9.6.1 presented in Section 9.6.4 of Chapter 9 that

$$\lim_{N,n\to\infty} \frac{1}{N}\|\widehat{\theta}_\lambda\|_0 = \mathbb{P}(|\Theta + \tau_* Z| \geq \tau_*\lambda/\beta_*) \stackrel{\text{def}}{=} s_*(\lambda)\,.$$

Further, by Theorem 8.3.2, we have $\frac{1}{\sqrt{n}}\|y - X\widehat{\theta}_\lambda\| = \beta_*(\lambda) + o_n(1)$. Recall that by (8.3.3) we have $\beta_*(\lambda) = \tau_*(\lambda)\big(1 - \frac{1}{\delta}s_*(\lambda)\big)$. We deduce $\widehat{\tau}(\lambda) = \tau_*(\lambda) + o_n(1)$. More precisely we have the following consistency result.

Corollary 8.4.1

Assume here that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $0 \leq s < s_{\max}(\delta)$ and $\xi > 0, p > 0$. There exists constants $C, c > 0$ that only depend on Ω such that for all $\epsilon \in (0, 1]$

$$\sup_{\theta^\star\in\mathcal{D}} \mathbb{P}\left(\sup_{\lambda\in[\lambda_{\min},\lambda_{\max}]} |\widehat{\tau}(\lambda) - \tau_*(\lambda)| \geq \epsilon\right) \leq C\epsilon^{-6}N^q \exp\left(-cN\epsilon^6\right)\,,$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

We next consider estimating the ℓ_2 error of the Lasso. Following [27], we define

$$\widehat{R}(\lambda) = \widehat{\tau}(\lambda)^2\Big(\frac{2}{N}\|\widehat{\theta}_\lambda\|_0 - 1\Big) + \frac{\big\|X^\mathsf{T}(y - X\widehat{\theta}_\lambda)\big\|^2}{N\big(1 - \frac{1}{n}\|\widehat{\theta}_\lambda\|_0\big)^2}\,.$$

Corollary 8.4.2

Assume here that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $0 \leq s < s_{\max}(\delta)$ and $\xi > 0, p > 0$. There exists constants $C, c > 0$ such that for all $\epsilon \in (0, 1]$,

$$\sup_{\theta^\star\in\mathcal{D}} \mathbb{P}\Big(\sup_{\lambda\in[\lambda_{\min},\lambda_{\max}]} \Big|\widehat{R}(\lambda) - \frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2\Big| \geq \epsilon\Big) \leq \frac{C}{\epsilon^6}N^q e^{-cN\epsilon^6}\,,$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

Corollary 8.4.2 is proved in Section 9.6.6 of Chapter 9. Since by Corollary 8.4.2, Corollary 8.4.1, Theorem 8.3.2 we have with high probability $\widehat{R}(\lambda) \simeq \frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2 \simeq \delta(\tau_*(\lambda)^2 - \sigma^2) \simeq \delta(\widehat{\tau}(\lambda)^2 - \sigma^2)$, the estimator

$$\widehat{\sigma}^2(\lambda) = \widehat{\tau}(\lambda)^2 - \frac{N}{n}\widehat{R}(\lambda) = \widehat{\tau}(\lambda)^2\Big(1 + \frac{N}{n} - \frac{2}{n}\|\widehat{\theta}_\lambda\|_0\Big) - \frac{\big\|X^\mathsf{T}(y - X\widehat{\theta}_\lambda)\big\|^2}{n\Big(1 - \frac{1}{n}\|\widehat{\theta}_\lambda\|_0\Big)^2} \tag{8.4.1}$$

is a consistent estimator of the noise level σ^2.

Corollary 8.4.3

There exists constants $C, c > 0$ that only depend on Ω, such that for all $\epsilon \in (0, 1]$

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\left(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \big|\widehat{\sigma}^2(\lambda) - \sigma^2\big| > \epsilon\right) \leq \frac{C}{\epsilon^6} N^q e^{-cN\epsilon^6}\,.$$

Finally, we consider the prediction error $\|X\theta^\star - X\widehat{\theta}_\lambda\|$. Stein Unbiased Risk Estimator (SURE) provides a general method to estimate the prediction error, see e.g. [194, 78, 207]. In the present case, it takes the form

$$\widehat{P}^{\mathrm{SURE}}(\lambda) = \frac{1}{n}\|y - X\widehat{\theta}_\lambda\|^2 + \frac{2\sigma^2}{n}\|\widehat{\theta}_\lambda\|_0\,. \tag{8.4.2}$$

Tibshirani and Taylor [207] proved that $\widehat{P}^{\mathrm{SURE}}(\lambda)$ is an unbiased estimator of the prediction error, namely

$$\mathbb{E}\big[\widehat{P}^{\mathrm{SURE}}(\lambda)\big] = \frac{1}{n}\|X\theta^\star - X\widehat{\theta}_\lambda\|^2 + \sigma^2\,. \tag{8.4.3}$$

The next result establishes consistency, uniformly over λ and $\theta^\star$, with quantitative concentration estimates.

Corollary 8.4.4

Assume here that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $0 \leq s < s_{\max}(\delta)$ and $\xi > 0, p > 0$. There exists constants $C, c > 0$ that only depend on Ω such that for all $\epsilon \in (0, 1]$

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \Big|\frac{1}{n}\|X\theta^\star - X\widehat{\theta}_\lambda\|^2 + \sigma^2 - \widehat{P}^{\mathrm{SURE}}(\lambda)\Big| \geq \epsilon\Big) \leq \frac{C}{\epsilon^6} N^q e^{-cN\epsilon^6}\,,$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

The same result holds if σ in (8.4.2) is replaced by an estimator of the noise level satisfying the same consistency condition as $\widehat{\sigma}$ defined by (8.4.1) (cf. Corollary 8.4.3).

This corollary follows simply from Theorem 9.6.1 from Chapter 9 and Theorem 8.3.2.

Remark 8.4.1. *Notice that exact unbiasedness of $\widehat{P}^{\mathrm{SURE}}(\lambda)$ only holds if the noise z in the linear model (8.1.2) is Gaussian [207]. In contrast, it is not hard to generalize the proofs in the present paper to include other noise distributions.*

8.4.2 Adaptive selection of λ

As anticipated, we can use our uniform bounds to select λ through an adaptive procedure. We discuss here three such procedures, that have already been illustrated in Figure 8.1:

(i) Selecting λ by minimizing the estimate $\widehat{\tau}(\lambda)$, we denote this by $\widehat{\lambda}^{\mathrm{EST}}$; (ii) Select λ as to minimize Stein's Unbiased Risk Estimate $\widehat{P}^{\mathrm{SURE}}(\lambda)$, $\widehat{\lambda}^{\mathrm{SURE}}$; (iii) Select λ by k-fold cross-validation, $\widehat{\lambda}^{k\text{-CV}}$. We will next describe these procedures in greater detail, and state the corresponding guarantees.

Minimization of $\widehat{\tau}(\lambda)$. Since the ℓ_2 risk of the Lasso is by Theorem 8.3.2 approximately equal to $R_*(\lambda) = \delta(\tau_*(\lambda)^2 - \sigma^2)$ and since by Corollary 8.4.1, $\widehat{\tau}$ is a consistent estimator (uniformly in λ) of τ_*, a natural procedure for selecting λ is to minimize $\widehat{\tau}$. We then define

$$\widehat{\lambda}^{\mathrm{EST}} = \operatorname*{arg\,min}_{\lambda\in[\lambda_{\min},\lambda_{\max}]} \widehat{\tau}(\lambda)\,.$$

The next result is an immediate consequence of Theorem 8.3.2 and Corollary 8.4.1:

Proposition 8.4.1

Assume here that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $0 \le s < s_{\max}(\delta)$ and $\xi > 0, p > 0$. There exists constants $C, c > 0$ that only depend on Ω such that for all $\epsilon \in (0,1]$

$$\inf_{\theta^\star\in\mathcal{D}} \mathbb{P}\left(\frac{1}{N}\|\widehat{\theta}_{\widehat{\lambda}^{\mathrm{EST}}} - \theta^\star\|^2 \le \inf_{\lambda\in[\lambda_{\min},\lambda_{\max}]}\left\{\frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2\right\} + \epsilon\right) \ge 1 - C\frac{N^q}{\epsilon^6}e^{-cN\epsilon^6},$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p-1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

Minimization of SURE. We define

$$\widehat{\lambda}^{\mathrm{SURE}} = \operatorname*{arg\,min}_{\lambda\in[\lambda_{\min},\lambda_{\max}]} \widehat{P}^{\mathrm{SURE}}(\lambda)\,.$$

Here, it is understood that we can use either σ or $\widehat{\sigma}(\lambda)$, cf. Eq. (8.4.1), in the definition of $\widehat{P}^{\mathrm{SURE}}$. We deduce from Corollary 8.4.4:

Proposition 8.4.2

Assume here that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $0 \le s < s_{\max}(\delta)$ and $\xi > 0, p > 0$. There exists constants $C, c > 0$ that only depend on Ω such that for all $\epsilon \in (0,1]$

$$\inf_{\theta^\star\in\mathcal{D}} \mathbb{P}\left(\frac{1}{n}\|X\widehat{\theta}_{\widehat{\lambda}^{\mathrm{SURE}}} - X\theta^\star\|^2 \le \inf_{\lambda\in[\lambda_{\min},\lambda_{\max}]}\left\{\frac{1}{n}\|X\widehat{\theta}_\lambda - X\theta^\star\|^2\right\} + \epsilon\right) \ge 1 - C\epsilon^{-6}N^q\exp\left(-cN\epsilon^6\right),$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p-1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

Cross-validation. We analyze now k-fold Cross Validation. Let $k \ge 2$ and define $n_k = n(k-1)/k$. We partition the rows of X in k groups: we obtain k-submatrices of size $(n/k)\times N$ that we denote $X^{(1)}, \dots, X^{(k)}$. Let us also write for $i \in \{1,\dots,k\}$, $X^{(\text{-}i)}$ for the submatrix of X obtained by removing the rows $X^{(i)}$. We denote by $y^{(i)}$, $z^{(i)}$ and $y^{(\setminus i)}$, $z^{(\setminus i)}$ the corresponding subvectors of y and z.

The estimator $\widehat{R}^{k\text{-CV}}$ of the risk using k-fold cross validation if defined as follows. For $i = 1,\dots,k$ solve the Lasso problem

$$\widehat{\theta}^i_\lambda = \operatorname*{arg\,min}_{\theta\in\mathbb{R}^N}\left\{\frac{1}{2n_k}\left\|y^{(\setminus i)} - X^{(\setminus i)}\theta\right\|^2 + \frac{\lambda}{n}|\theta|\right\},$$

and then compute

$$\widehat{R}^{k\text{-CV}}(\lambda) = \frac{1}{N}\sum_{i=1}^{k}\left\|y^{(i)} - X^{(i)}\widehat{\theta}^{i}_{\lambda}\right\|^2 .$$

Finally, we set λ as follows

$$\widehat{\lambda}^{k\text{-CV}} = \operatorname*{arg\,min}_{\lambda\in[\lambda_{\min},\lambda_{\max}]} \widehat{R}^{k\text{-CV}}(\lambda) .$$

The next Proposition shows that $\widehat{R}^{k\text{-CV}}(\lambda)$ is equal to the true risk (shifted by $\delta\sigma^2$) up to $O(k^{-1/2})$.

Proposition 8.4.3

There exists constants $c, C > 0$ that depend only on Ω, such that for all $k \geq 2$ such that $s_{\max}\Big((k-1)\delta/k\Big) > s$ in the case where $\mathcal{D} = \mathcal{F}_0(s)$, we have

$$\sup_{\theta^\star\in\mathcal{D}}\mathbb{P}\left(\sup_{\lambda\in[\lambda_{\min},\lambda_{\max}]}\left|\widehat{R}^{k\text{-}CV}(\lambda) - \frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2 - \delta\sigma^2\right| \geq \frac{C}{\sqrt{k}}\right) \leq Ck^6N^q e^{-cN/k^6} ,$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p-1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

Proposition 8.4.3 is proved in Section 9.6.7 from Chapter 9. It follows from Proposition 8.4.3 that with high probability,

$$\frac{1}{N}\|\widehat{\theta}_{\widehat{\lambda}^{k\text{- CV}}} - \theta^\star\|^2 \leq \inf_{\lambda\in[\lambda_{\min},\lambda_{\max}]}\frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2 + O(k^{-1/2}).$$

8.4.3 Numerical experiments

In this Section we compare numerically various different choices for the regularization parameter λ, namely $\widehat{\lambda}^{\text{EST}}$, $\widehat{\lambda}^{\text{SURE}}$ and $\widehat{\lambda}^{k\text{-CV}}$, presented in the previous section. For these experiments we take the components $\theta^\star_1, \dots, \theta^\star_N$ to be i.i.d. from

$$P_0 = s\mathcal{N}(0,1) + (1-s)\delta_0 .$$

Within this probabilistic model, we can compare achieved by our various choice of λ to the Bayes optimal error (Minimal Mean Squared Error):

$$\text{MMSE}_N = \min_{\widehat{\theta}}\mathbb{E}\left[\left\|\theta^\star - \widehat{\theta}(y,X)\right\|^2\right] = \mathbb{E}\left[\left\|\theta^\star - \mathbb{E}[\theta^\star|y,X]\right\|^2\right] ,$$

where the minimum is taken over all estimators $\widehat{\theta}$ (i.e. measurable functions of X, y). The limit of the MMSE has been recently computed by [17] and [179] and is given by Theorem 8.4.1 below, which is a direct corollary from Theorem 7.3.2. Recall, that given two random variables U, V, their mutual information is the Kullback-Leibler divergence between their joint distribution and the product of the marginals: $I(U;V) \stackrel{\text{def}}{=} D_{\text{KL}}(p_{U,V}\|p_U \otimes p_V)$.

Theorem 8.4.1

Define the function

$$\Psi_{\delta,\sigma}(m) = I_{P_0}\Big(\frac{\sigma^{-2}}{1+m}\Big) + \frac{\delta}{2}\Big(\log(1+m) - \frac{m}{1+m}\Big)\,,$$

where $I_{P_0}(r) = I(\Theta; \sqrt{r}\Theta + Z)$ *for* $(\Theta, Z) \sim P_0 \otimes \mathcal{N}(0,1)$. *Then, for almost every* $\delta > 0$ *the function* $\Psi_{\delta,\sigma}$ *admits a unique maximizer* $m^*(\delta,\sigma)$ *on* $\mathbb{R}_{\geq 0}$ *and*

$$\mathrm{MMSE}_N \xrightarrow[N\to\infty]{} \delta\sigma^2 m^*(\delta,\sigma)\,.$$

Figure 8.1 reports the risk achieved by the various choices of λ as a function of the number of samples per dimension δ. We also compare the data-driven procedures of the previous section to the theory-driven choice $\lambda = \sigma\sqrt{2\log N}$. In the left frame, we consider uncorrelated random designs: $X_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/n)$. On the right, we consider i.i.d. Gaussian rows with covariance structure determined by an auto-regressive model. Explicitly, the columns $(X_j)_{1\leq j\leq N}$ of X are generated according to:

$$X_1 = u_0, \qquad X_{j+1} = \frac{1}{\sqrt{1+\phi^2}}\big(\phi X_j + u_j\big) \tag{8.4.4}$$

where $u_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathrm{Id}/n)$ and $\phi = 2$. For both types of designs, $\widehat{\lambda}^{\mathrm{EST}}$, $\widehat{\lambda}^{\mathrm{SURE}}$ and $\widehat{\lambda}^{k\text{-CV}}$ perform similarly, and substantially outperform the theoretical choice $\lambda = \sigma\sqrt{2\log N}$.

For uncorrelated designs, the resulting risk is closely tracked by the asymptotic theory, and is surprisingly close to the asymptotic prediction for the Bayes risk MMSE_N.

While our theory does not cover the case of correlated designs, the qualitative behavior is remarkably similar. We also observed that in this case, the risk estimator $\widehat{R}(\lambda)$ is not consistent but its minimum is roughly located at the same value of λ as for uncorrelated designs.

Next we study adaptivity to sparsity. On Figure 8.2, we plot the risk as a function of the sparsity of the signal $\theta^\star$. We compare the three adaptive procedures (namely, $\widehat{\lambda}^{\mathrm{EST}}$, $\widehat{\lambda}^{\mathrm{SURE}}$ and $\widehat{\lambda}^{k\text{-CV}}$), to the following choice

$$\begin{aligned}
\lambda^{\mathrm{MM}}(s_0) &= \alpha_0\sigma\sqrt{1 - \frac{1}{\delta}M_{s_0}(\alpha_0)}\,,\\
M_s(\alpha) &= s(1+\alpha^2) + 2(1-s)\Big((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big)\,,\\
\alpha_0 &= \arg\min_{\alpha\geq 0} M_{s_0}(\alpha)\,,
\end{aligned}$$

where $s_0 < s_{\max}(\delta)$ is a nominal value for the sparsity (for Figure 8.2, we use $s_0 = 0.3$). The value $\lambda^{\mathrm{MM}}(s_0)$ is expected to be asymptotically minimax optimal over $\mathcal{F}_0(s_0)$ [72].

Also in this example, adaptive procedures dramatically outperform the fixed choice $\lambda = \sigma\sqrt{2\log N}$, and also the minimax optimal λ at the nominal sparsity level.

8.5 Proof strategy

As mentioned above, our proofs are based on Gaussian comparison inequalities, and in particular on Gordon's min-max theorem [93, 94]. In this section we review the application

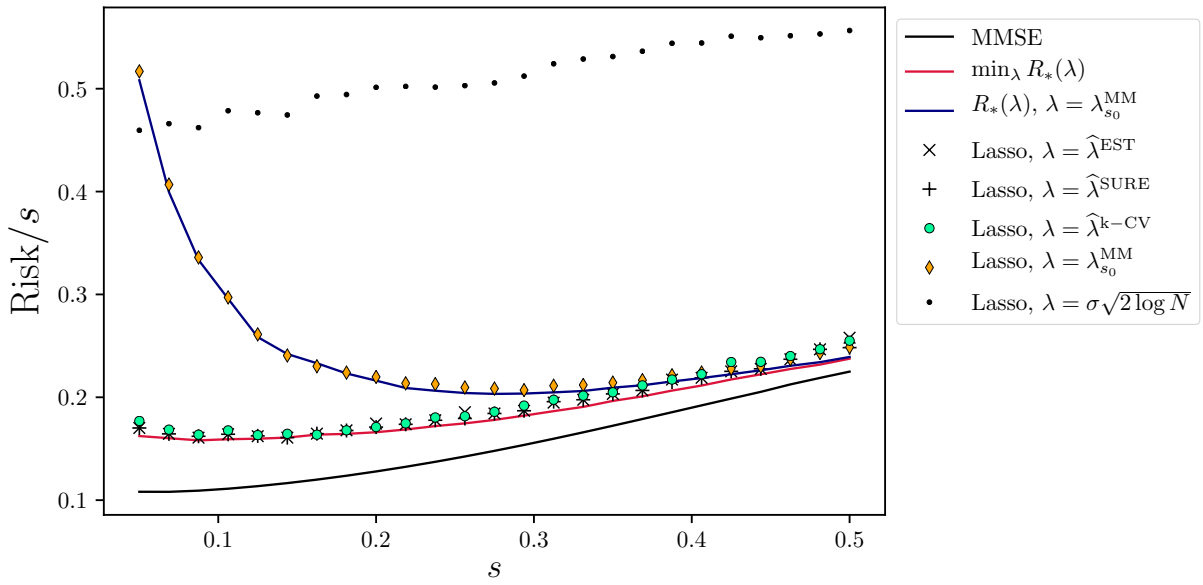


Figure 8.2: Risk of the Lasso for different choices of λ. $N = 10000$, $\sigma = 0.2$, $\delta = 0.8$. Here $\theta^\star$ is chosen to be sN-sparse, and we vary the sparsity level s. The entries on the support of $\theta^\star$ are i.i.d. $\mathcal{N}(0,1)$. Cross-validation is carried out using 4 folds. SURE is computed using the estimator $\widehat{\sigma}$. The minimax regularization $\lambda^{\mathrm{MM}}(s_0)$ is used at the nominal level $s_0 = 0.3$.

of this inequality to the Lasso as developed in [205]. We then discuss the limitations of earlier work, which does not characterize the empirical distribution of the Lasso estimator $\widehat{\theta}_\lambda$ (or need extra sparsity assumptions [169]) nor uniform bounds as in Theorem 8.3.1. A key challenge is related to the fact that the Lasso cost function (8.1.1) is convex but not strongly convex. Hence, a small change in λ could cause *a priori* a large change in the minimizer $\widehat{\theta}_\lambda$.

In order to overcome these problems, we establish a property that we call 'local stability.' Namely, if the empirical distribution of $(\widehat{\theta}_\lambda, \theta^\star)$ deviates from our prediction, then the value of the optimization problem increases significantly. This implies that the empirical distribution is stable with respect to perturbations of the cost (e.g. changes in λ). Gordon's comparison is again crucial to prove this stability property.

Finally, we describe how local stability is used to prove the results of the previous sections. A full description of the proofs is provided in Chapter 9.

8.5.1 Tight Gaussian min-max theorem

It is more convenient (but equivalent) to study $\widehat{w}_\lambda = \widehat{\theta}_\lambda - \theta^\star$ instead of $\widehat{\theta}_\lambda$. The vector $\widehat{w}_\lambda$ is the minimizer of the cost function

$$\mathcal{C}_\lambda(w) = \frac{1}{2n}\|Xw - \sigma z\|^2 + \frac{\lambda}{n}\Big(|w + \theta^\star| - |\theta^\star|\Big)\,. \tag{8.5.1}$$

Following [205], we rewrite the minimization of $\mathcal{C}_\lambda$ as a saddle point problem:

$$\min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) = \min_{w\in\mathbb{R}^N}\max_{u\in\mathbb{R}^n}\left\{\frac{1}{n}u^{\mathsf{T}}\Big(Xw - \sigma z\Big) - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n}\Big(|w+\theta^\star| - |\theta^\star|\Big)\right\}. \tag{8.5.2}$$

We apply the following Theorem from [205] which improves over Gordon's Theorem [94] by exploiting convex duality.

Theorem 8.5.1 (*Theorem 3 from [205]*)

Let $S_w \subset \mathbb{R}^N$ and $S_u \subset \mathbb{R}^n$ be two compact sets and let $Q : S_w \times S_u \to \mathbb{R}$ be a continuous function. Let $G = (G_{i,j}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, $g \sim \mathcal{N}(0, \mathrm{Id}_N)$ and $h \sim \mathcal{N}(0, \mathrm{Id}_n)$ be independent standard Gaussian vectors. Define

$$\begin{cases} \mathcal{C}^*(G) = \min\limits_{w \in S_w} \max\limits_{u \in S_u} u^\mathsf{T} G w + Q(w,u)\,, \\ L^*(g,h) = \min\limits_{w \in S_w} \max\limits_{u \in S_u} \|u\|_2 g^\mathsf{T} w + \|w\|_2 h^\mathsf{T} u + Q(w,u)\,. \end{cases}$$

Then we have:

- *For all $t \in \mathbb{R}$,*
$$\mathbb{P}\Big(\mathcal{C}^*(G) \le t\Big) \le 2\mathbb{P}\Big(L^*(g,h) \le t\Big)\,.$$
- *If S_w and S_u are convex and if Q is convex concave, then for all $t \in \mathbb{R}$*
$$\mathbb{P}\Big(\mathcal{C}^*(G) \ge t\Big) \le 2\mathbb{P}\Big(L^*(g,h) \ge t\Big)\,.$$

For the reader's convenience, we provide in Section 9.7.2 of Chapter 9 a proof of this theorem.

Because of Gordon's Theorem, it suffices now to study (see Corollary 8.5.1 below) for $(g, g', h) \sim \mathcal{N}(0, \mathrm{Id}_N) \otimes \mathcal{N}(0,1) \otimes \mathcal{N}(0, \mathrm{Id}_n)$.

$$L_\lambda(w) = \frac{1}{2}\left(\sqrt{\frac{\|w\|^2}{n} + \sigma^2}\,\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}}\right)_+^2 + \frac{\lambda}{n}|w + \theta^\star| - \frac{\lambda}{n}|\theta^\star|\,. \tag{8.5.3}$$

Corollary 8.5.1

(a) Let $D \subset \mathbb{R}^N$ be a closed set. We have for all $t \in \mathbb{R}$
$$\mathbb{P}\Big(\min_{w \in D} \mathcal{C}_\lambda(w) \le t\Big) \le 2\mathbb{P}\Big(\min_{w \in D} L_\lambda(w) \le t\Big)\,.$$

(b) Let $D \subset \mathbb{R}^N$ be a convex closed set. We have for all $t \in \mathbb{R}$
$$\mathbb{P}\Big(\min_{w \in D} \mathcal{C}_\lambda(w) \ge t\Big) \le 2\mathbb{P}\Big(\min_{w \in D} L_\lambda(w) \ge t\Big)\,.$$

Proof. We will only prove the first point, since the second follows from the same arguments. Define for $(w,u) \in \mathbb{R}^N \times \mathbb{R}^n$

$$c_\lambda(w,u) = \frac{1}{n} u^\mathsf{T} X w - \frac{\sigma}{n} u^\mathsf{T} z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n}(|w + \theta^\star| - |\theta^\star|)\,,$$
$$l_\lambda(w,u) = -\frac{1}{n^{3/2}}\|u\| g^\mathsf{T} w + \frac{1}{n}\|u\| g'\sigma + \sqrt{\frac{\|w\|^2}{n} + \sigma^2}\,\frac{h^\mathsf{T} u}{n} - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n}(|w + \theta^\star| - |\theta^\star|)\,.$$

Notice that for all $w \in \mathbb{R}^N$, $L_\lambda(w) = \max_{u \in \mathbb{R}^n} l_\lambda(w,u)$ and $\mathcal{C}_\lambda(w) = \max_{u \in \mathbb{R}^n} c_\lambda(w,u)$.

Let us suppose that X, z, g, h, g' live on the same probability space and are independent. Let $\epsilon \in (0,1]$. Let $\sigma_{\max}(X)$ denote the largest singular value of the matrix X. By tightness we can find $K > 0$ such that the event

$$\Big\{\sigma_{\max}(X) \leq K, \quad \|z\| \leq K, \quad \|g\| \leq K, \quad \|h\| \leq K, \quad |g'| \leq K\Big\} \tag{8.5.4}$$

has probability at least $1-\epsilon$. Let $D \subset \mathbb{R}^N$ be a (non-empty, otherwise the result is trivial) closed set. Let us fix $w_0 \in D$. On the event (8.5.4) $\mathcal{C}_\lambda(w_0)$ and $L_\lambda(w_0)$ are both upper bounded by some non-random quantity R. Let now $w \in D$ such that $\mathcal{C}_\lambda(w) \leq R$. We have then $\frac{\lambda}{n}|w+\theta^\star| \leq R + \frac{\lambda}{n}|\theta^\star|$, which implies that $\|w\|$ is upper bounded by some non-random quantity R_1. This implies that, on the event (8.5.4), the minimum of $\mathcal{C}_\lambda$ over D is achieved on $D \cap B(0,R_1)$. Similarly on (8.5.4) the minimum of L_λ over D is achieved on $D \cap B(0,R_2)$, for some non-random quantity R_2. Without loss of generalities, one can assume $R_1 = R_2$. On the event (8.5.4) we have

$$\min_{w\in D} \mathcal{C}_\lambda(w) = \min_{w \in D\cap B(0,R_1)} \mathcal{C}_\lambda(w) = \min_{w\in D\cap B(0,R_1)} \max_{u \in B(0,R_3)} c_\lambda(w,u)\,,$$

for some non-random $R_3 > 0$. This gives that for all $t \in \mathbb{R}$, we have

$$\mathbb{P}\Big(\min_{w\in D} \mathcal{C}_\lambda(w) \leq t\Big) \leq \mathbb{P}\Big(\min_{w\in D\cap B(0,R_1)} \max_{u\in B(0,R_3)} c_\lambda(w,u) \leq t\Big) + \epsilon\,,$$

and similarly

$$\mathbb{P}\Big(\min_{w\in D\cap B(0,R_1)} \max_{u\in B(0,R_3)} l_\lambda(w,u) \leq t\Big) \leq \mathbb{P}\Big(\min_{w\in D} L_\lambda(w) \leq t\Big) + \epsilon\,.$$

Since the sets $D \cap B(0,R_1)$ and $B(0,R_3)$ are compact, one can apply Theorem 8.5.1 to c_λ and l_λ and obtain:

$$\mathbb{P}\Big(\min_{w\in D} \mathcal{C}_\lambda(w) \leq t\Big) \leq 2\mathbb{P}\Big(\min_{w\in D} L_\lambda(w) \leq t\Big) + 2\epsilon\,.$$

The Corollary follows then from the fact that one can take ϵ arbitrarily small. □

8.5.2 Local stability

In order to prove that (for instance) $\widehat{w}_\lambda$ verifies with high probability some property, let's say for instance that the empirical distribution of $(\widehat{\theta}_\lambda = \theta^\star + \widehat{w}_\lambda, \theta^\star)$ is close to $\mu_\lambda^\star$, we define a set $D_\epsilon \subset \mathbb{R}^N$ that contains all the vectors that do not verify this property, e.g. $D_\epsilon = \Big\{w \in \mathbb{R}^N \,\Big|\, W_2\Big(\widehat{\mu}_{(\theta^\star+w,\theta^\star)}, \mu_\lambda^*\Big)^2 \geq \epsilon\Big\}$, for some $\epsilon \in (0,1)$. The goal now is to prove that with high probability

$$\min_{w\in D} \mathcal{C}_\lambda(w) \geq \min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) + \epsilon\,,$$

for some $\epsilon > 0$. Using Gordon's min-max Theorem (Corollary 8.5.1) we will be able to show

$$\mathbb{P}\Big(\min_{w\in D_\epsilon} \mathcal{C}_\lambda(w) \leq \min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) + \epsilon\Big) \leq 2\mathbb{P}\Big(\min_{w\in D_\epsilon} L_\lambda(w) \leq \min_{w\in\mathbb{R}^N} L_\lambda(w) + \epsilon\Big) + o_N(1)\,. \tag{8.5.5}$$

Informally, this is a consequence of the following two remarks. First, by applying parts (a) and (b) of Corollary 8.5.1 to the convex domain $\mathbb{R}^N$, we deduce that $\min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) \approx \min_{w\in\mathbb{R}^N} L_\lambda(w)$. Second, by applying part (a) to the closed domain D, we obtain $\min_{w\in D_\epsilon} \mathcal{C}_\lambda(w) \gtrsim \min_{w\in D_\epsilon} L_\lambda(w)$

It remains now to study the cost function L_λ, which is much simpler. This is done in Section 9.2 of Chapter 9. The key step will be to establish the following 'local stability' result (the next statement is an immediate consequence of Proposition 9.2.1 and Theorem 9.2.1 in Chapter 9. We prove in fact that the cost function L_λ is strongly convex on a neighborhood of its minimizer.).

Theorem 8.5.2

The minimizer $w_\lambda^* = \arg\min_w L_\lambda(w)$ exists and is almost surely unique. Further, there exists constants $\gamma, c, C > 0$ that only depend on Ω such that for all $\theta^\star \in \mathcal{D}$, all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and all $\epsilon \in (0, 1]$

$$\mathbb{P}\Big(\exists w \in \mathbb{R}^N, \quad \frac{1}{N}\|w - w_\lambda^*\|^2 > \epsilon \quad \text{and} \quad L_\lambda(w) \leq \min_{v \in \mathbb{R}^N} L_\lambda(v) + \gamma\epsilon\Big) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2} .$$

We do not obtain an equally strong result for the cost function $\mathcal{C}_\lambda(w)$, but we prove the following statement, which is sufficient for obtaining uniform control (for the sake of argument, we focus here on the domain $\mathcal{F}_p(\xi)$ and control of the empirical distribution).

Theorem 8.5.3

Assume that $\mathcal{D} = \mathcal{F}_p(\xi)$ for some $\xi, p > 0$. There exists constants $C, c, \gamma > 0$ that only depend on Ω such that for all $\epsilon \in (0, \frac{1}{2}]$

$$\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\exists \theta \in \mathbb{R}^N, \quad W_2\big(\widehat{\mu}_{(\theta,\theta^\star)}, \mu_\lambda^*\big)^2 \geq \epsilon \quad \text{and} \quad \mathcal{L}_\lambda(\theta) \leq \min \mathcal{L}_\lambda + \gamma\epsilon\Big)$$
$$\leq C\epsilon^{-\max(1,a)} \exp\left(-cN\epsilon^2\epsilon^a \log(\epsilon)^{-2}\right) ,$$

where $a = \frac{1}{2} + \frac{1}{p}$.

Theorem 8.5.3 is proved in Section 9.3.1 of Chapter 9.

8.5.3 Sketch of proof of main results

For the sake of simplicity, we will illustrate the prove strategy by considering the empirical distribution of $\widehat{w}_\lambda = \widehat{\theta}_\lambda - \theta^\star$, as the argument is similar for other quantities. According to Theorem 8.3.1, this should be well approximated by $\overline{\mu}_\lambda$ that is the law of $\widehat{\Theta} - \Theta$, when $(\widehat{\Theta}, \Theta) \sim \mu_\lambda^*$, cf. Definition 8.3.3.

As anticipated, Eq. (8.5.5) and Theorem 8.5.2, allow to control $W_2(\widehat{\mu}_{\widehat{w}_\lambda}, \overline{\mu}_\lambda)$ for a fixed λ ($\widehat{\mu}_{\widehat{w}_\lambda}$ denotes the empirical distribution of the entries of $\widehat{w}_\lambda$). Namely, we can define D_ε to be the set of vectors w such that $W_2(\widehat{\mu}_w, \overline{\mu}_\lambda) \geq \varepsilon > 0$. We then prove that the minimizer w_λ^* of L_λ has empirical distribution close to $\overline{\mu}_\lambda$, and therefore by Theorem 8.5.2, $L_\lambda(w) > L_\lambda(w_\lambda^*) + \gamma\epsilon$ for all $w \in D_\epsilon$, with high probability. This imply that the right-hand side of (8.5.5) is very small and we deduce that, with high probability, all minimizers or near minimizers of $\mathcal{C}_\lambda(w)$ have empirical distribution close to $\overline{\mu}_\lambda$,

We now would like to prove Theorem 8.3.1 and show that with high probability $\widehat{\mu}_{\widehat{w}_\lambda} \approx \overline{\mu}_\lambda$, uniformly in $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. To do so, we apply the above argument for $\lambda = \lambda_1, \ldots, \lambda_k$, where $\lambda_1, \ldots, \lambda_k$ is an ϵ-net of $[\lambda_{\min}, \lambda_{\max}]$. This implies that, with high probability for $\lambda \in \{\lambda_1, \ldots, \lambda_k\}$, $W_2(\widehat{\mu}_{\widehat{w}_{\lambda_i}}, \overline{\mu}_{\lambda_i}) \leq \varepsilon$. Next, for $\lambda \in [\lambda_i, \lambda_{i+1}]$, we show that

$$\mathcal{C}_{\lambda_i}(\widehat{w}_\lambda) = \min_{w \in \mathbb{R}^N} \mathcal{C}_{\lambda_i}(w) + O(|\lambda_{i+1} - \lambda_i|) .$$

Consequently if $|\lambda_{i+1} - \lambda_i| = O(\epsilon)$ (using again Eq. (8.5.5) and Theorem 8.5.2), we obtain that $W_2(\widehat{\mu}_{\widehat{w}_\lambda}, \overline{\mu}_{\lambda_i}) = O(\epsilon)$ and therefore $W_2(\widehat{\mu}_{\widehat{w}_\lambda}, \overline{\mu}_\lambda) = O(\epsilon)$. We conclude that

$W_2(\widehat{\mu}_{\widehat{w}_\lambda}, \overline{\mu}_\lambda) = O(\epsilon)$ for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, with high probability, which is the desired claim.

If the strategy exposed above allows to obtain the risk of the Lasso and the empirical distribution of its coordinates, it is not enough to get its sparsity $\|\widehat{\theta}_\lambda\|_0$ or to obtain the empirical distribution of the debiased lasso

$$\widehat{\theta}^d_\lambda = \widehat{\theta}_\lambda + \frac{X^\mathsf{T}(y - X\widehat{\theta}_\lambda)}{1 - \frac{1}{n}\|\widehat{\theta}_\lambda\|_0}.$$

Therefore, we will need to analyze the vector

$$\widehat{v}_\lambda = \frac{1}{\lambda} X^\mathsf{T}(y - X\widehat{\theta}_\lambda)\,,$$

which is a subgradient of the ℓ_1-norm at $\widehat{\theta}_\lambda$. We are able to study $\widehat{v}_\lambda$ using Gordon's min-max Theorem because $\widehat{v}_\lambda$ is the unique maximizer of

$$v \mapsto \min_{w \in \mathbb{R}^N} \left\{ \frac{1}{2n} \left\| Xw - \sigma z \right\|^2 + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) \right\}.$$

The detailed analysis is done in Section 9.5 from Chapter 9.

Chapter 9

Proofs of the results on the Lasso

This chapter is devoted to the proofs of the results of Chapter 8. It is organised as follows. In Section 9.1, we study the max-min scalar problem (8.3.2). In Section 9.2 we investigate Gordon's optimization problem associated to the Lasso cost and deduce the results about the risk and the empirical law of the Lasso estimator in Section 9.3. We conduct the same kind of analysis for the Lasso residual $\widehat{u}_\lambda \stackrel{\text{def}}{=} X\widehat{\theta}_\lambda - y$ in Section 9.4 and the subgradient $\widehat{v}_\lambda \stackrel{\text{def}}{=} \frac{1}{\lambda}X^\mathsf{T}(y - X\widehat{\theta}_\lambda)$ of the ℓ_1-norm at $\widehat{\theta}_\lambda$ in Section 9.5. Section 9.6 gathers some auxiliary results and proofs such as the study of the sparsity of the Lasso estimator, the law of the debiased Lasso $\widehat{\theta}^{(d)}_\lambda$, or the performance of k-fold Cross Validation. Finally Section 9.7 contains a recap of the notations, a proof of the "tight Gordon min-max" Theorem and some basic concentration results.

9.1 Study of the scalar optimization problem

In this section we study the scalar optimization problem (8.3.2):

$$\max_{\beta\geq 0}\min_{\tau\geq\sigma}\left(\frac{\sigma^2}{\tau}+\tau\right)\frac{\beta}{2}-\frac{1}{2}\beta^2+\frac{1}{\delta}\mathbb{E}\min_{w\in\mathbb{R}}\left\{\frac{w^2}{2\tau}\beta-\beta Z w+\lambda|w+\Theta|-\lambda|\Theta|\right\}, \tag{9.1.1}$$

where $(\Theta, Z)\sim P_0\otimes\mathcal{N}(0,1)$, for some probability distribution P_0 with finite first moment: $\mathbb{E}_{P_0}|\Theta|<\infty$. Of course, we will be mainly interested by the case where $P_0=\widehat{\mu}_{\theta^\star}$, the empirical distribution of the entries of $\theta^\star$. Define

$$\psi_\lambda(\beta,\tau)=\left(\frac{\sigma^2}{\tau}+\tau\right)\frac{\beta}{2}-\frac{1}{2}\beta^2+\frac{1}{\delta}\mathbb{E}\min_{w\in\mathbb{R}}\left\{\frac{w^2}{2\tau}\beta-\beta Z w+\lambda|w+\Theta|-\lambda|\Theta|\right\}.$$

9.1.1 Basic properties of the scalar optimization problem

Lemma 9.1.1 (*From [71]*)

For all $\delta\in(0,1)$, the equation

$$(1+\alpha^2)\Phi(-\alpha)-\alpha\phi(\alpha)=\frac{\delta}{2}$$

admits a unique positive solution $\alpha_{\min}=\alpha_{\min}(\delta)>0$.

Proof. Let $\varphi : \alpha \mapsto (1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)$. φ is continuous on $\mathbb{R}_{\geq 0}$, we have $\varphi(0) = \frac{1}{2}$ and $\varphi(+\infty) = 0$. It remains to show that φ is strictly decreasing on $\mathbb{R}_{\geq 0}$. Compute $\varphi'(\alpha) = 2\alpha\Phi(-\alpha) - 2\phi(\alpha)$ and $\varphi''(\alpha) = 2\Phi(-\alpha) > 0$. Since $\varphi'(+\infty) = 0$, we have that for all $\alpha \geq 0$, $\varphi'(\alpha) < 0$. φ is thus strictly decreasing on $\mathbb{R}_{\geq 0}$. □

Let us define

$$\beta_{\max} = \beta_{\max}(\delta, \lambda) = \frac{\lambda}{\alpha_{\min}(\delta)}\,.$$

More generally, we will always write $\alpha = \lambda/\beta$. We prove in this section the following theorem and some auxiliary results.

Theorem 9.1.1

The max-min (8.3.2) is achieved at a unique couple (β_*, τ_*) and $0 < \beta_* < \beta_{\max}$. Moreover, (τ_*, β_*) is also the unique couple in $(0, +\infty)^2$ that verify

$$\begin{cases} \tau^2 &= \sigma^2 + \frac{1}{\delta}\mathbb{E}\Big[(\eta(\Theta + \tau Z, \tau\frac{\lambda}{\beta}) - \Theta)^2\Big] \\ \beta &= \tau\left(1 - \frac{1}{\delta}\mathbb{E}\left[\eta'(\Theta + \tau Z, \frac{\tau\lambda}{\beta})\right]\right)\,. \end{cases} \tag{9.1.2}$$

Lemma 9.1.2

$$-\frac{\lambda}{\delta}\mathbb{E}|\Theta| \leq \max_{\beta\geq 0}\min_{\tau\geq\sigma}\psi_\lambda(\beta,\tau) \leq \frac{\sigma^2}{2}\,.$$

Proof. We have $\max_\beta \min_\tau \psi_\lambda(\beta,\tau) \geq \min_\tau \psi_\lambda(0,\tau) = -\frac{\lambda}{\delta}\mathbb{E}|\Theta|$. Then, by taking $w = 0$ one get

$$\max_{\beta\geq 0}\min_{\tau\geq\sigma}\psi_\lambda(\beta,\tau) \leq \max_{\beta\geq 0}\min_{\tau\geq\sigma}\left(\frac{\sigma^2}{\tau} + \tau\right)\frac{\beta}{2} - \frac{1}{2}\beta^2 = \frac{\sigma^2}{2}\,.$$

□

Define for $\alpha \geq 0$ and $y \in \mathbb{R}$,

$$\ell_\alpha(y) = \min_{x\in\mathbb{R}}\left\{\frac{1}{2}(y-x)^2 + \alpha|x|\right\}$$

and for $Z \sim \mathcal{N}(0,1), x \in \mathbb{R}$, $\Delta_\alpha(x) = \mathbb{E}\Big[\ell_\alpha(x+Z) - \alpha|x|\Big]$.

Lemma 9.1.3

$$\mathbb{E}\left[\min_{w\in\mathbb{R}}\left\{\frac{w^2}{2\tau}\beta - \beta Z w + \lambda|w+\Theta| - \lambda|\Theta|\right\}\right] = \tau\beta\mathbb{E}\left[\Delta_\alpha\Big(\frac{\Theta}{\tau}\Big)\right] - \frac{\beta\tau}{2}\,. \tag{9.1.3}$$

where $\alpha = \lambda/\beta$.

Proof. Let $\beta > 0$ and compute

$$\begin{aligned} &\mathbb{E}\left[\min_{w\in\mathbb{R}}\left\{\frac{w^2}{2\tau}\beta - \beta Z w + \lambda|w+\Theta| - \lambda|\Theta|\right\}\right] \\ &\qquad = \frac{-\beta\tau}{2} + \frac{\beta}{\tau}\mathbb{E}\min_{w\in\mathbb{R}}\left\{\frac{1}{2}(w-\tau Z)^2 + \frac{\tau\lambda}{\beta}|w+\Theta| - \frac{\tau\lambda}{\beta}|\Theta|\right\} \\ &\qquad = \frac{-\beta\tau}{2} + \beta\tau\mathbb{E}\min_{w\in\mathbb{R}}\left\{\frac{1}{2}(w-Z)^2 + \alpha\Big|w + \frac{\Theta}{\tau}\Big| - \alpha\Big|\frac{\Theta}{\tau}\Big|\right\}\,, \end{aligned}$$

where $\alpha = \lambda/\beta$. Thus

$$\mathbb{E}\left[\min_{w\in\mathbb{R}}\left\{\frac{w^2}{2\tau}\beta - \beta Z w + \lambda|w+\Theta| - \lambda|\Theta|\right\}\right] = \frac{-\beta\tau}{2} + \beta\tau\mathbb{E}\left[\Delta_\alpha\left(\frac{\Theta}{\tau}\right)\right].$$

□

Lemma 9.1.4

- *If* $\beta > \beta_{\max}$, $\quad \psi_\lambda(\beta,\tau) \xrightarrow[\tau\to+\infty]{} -\infty$.
- *If* $\beta = \beta_{\max}$, $\quad \psi_\lambda(\beta,\tau) \xrightarrow[\tau\to+\infty]{} -\frac{\beta^2}{2} - \frac{\lambda}{\delta}\mathbb{E}\big|\Theta\big|$.

Proof. By (9.1.3) and the fact that $\Delta_\alpha(0) = \frac{1}{2} + \alpha\phi(\alpha) - (\alpha^2+1)\Phi(-\alpha)$ by Lemma 9.6.13, we get that for all $\beta,\tau > 0$

$$\psi_\lambda(\beta,\tau) = \frac{\sigma^2\beta}{2\tau} - \frac{\beta^2}{2} + \frac{\tau\beta}{\delta}\left(\frac{\delta}{2} + \alpha\phi(\alpha) - (\alpha^2+1)\Phi(-\alpha)\right) + \frac{\beta}{\delta}\xi_\alpha(\tau),$$

where $\alpha = \lambda/\beta$ and

$$\xi_\alpha(\tau) = \tau\mathbb{E}\left[\left(\Delta_\alpha\left(\frac{\Theta}{\tau}\right) - \Delta_\alpha(0)\right)\right].$$

Using the definition of $\beta_{\max}$: if $\beta > \beta_{\max}$ then $\alpha < \alpha_{\min}$ and therefore $\frac{\delta}{2} + \alpha\phi(\alpha) - (\alpha^2+1)\Phi(-\alpha) < 0$. If $\beta = \beta_{\max}$, $\frac{\delta}{2} + \alpha\phi(\alpha) - (\alpha^2+1)\Phi(-\alpha) = 0$. It remains to compute the limit of $\xi_\alpha(\tau)$ as $\tau\to\infty$.

Using the expression (see Lemma 9.6.13) of the left-and right-derivatives of Δ_α at 0, we have almost surely:

$$\tau\left(\Delta_\alpha\left(\frac{\Theta}{\tau}\right) - \Delta_\alpha(0)\right) \xrightarrow[\tau\to\infty]{} -\alpha|\Theta|.$$

Suppose that $\mathbb{E}|\Theta| < \infty$. By Lemma 9.6.13, Δ_α is α-Lipschitz. Consequently, for all $\tau > 0$:

$$\left|\tau\left(\Delta_\alpha\left(\frac{\Theta}{\tau}\right) - \Delta_\alpha(0)\right)\right| \le \alpha|\Theta|.$$

Since we have assumed that $\mathbb{E}|\Theta| < \infty$ we can apply the dominated convergence theorem to obtain that $\xi_\alpha(\tau) \xrightarrow[\tau\to+\infty]{} -\alpha\mathbb{E}|\Theta|$. □

Define

$$w^*(\alpha,\tau) = \eta\Big(\Theta + \tau Z, \alpha\tau\Big) - \Theta.$$

$w^*(\alpha,\tau)$ is the minimizer of $w \mapsto \frac{w^2}{2\tau}\beta - \beta Z w + \lambda|w+\Theta|$ (recall that we always write $\alpha = \lambda/\beta$).

Lemma 9.1.5

If $\beta \geq \beta_{\max}$ the equation

$$\tau^2 = \sigma^2 + \frac{1}{\delta}\mathbb{E}\Big[w^*(\alpha,\tau)^2\Big] = \sigma^2 + \frac{1}{\delta}\mathbb{E}\Big[(\eta(\Theta + \tau Z, \tau\frac{\lambda}{\beta}) - \Theta)^2\Big]\,. \tag{9.1.4}$$

does not admits any solution on $(0,+\infty)$. For all $\beta \in (0,\beta_{\max})$, the function $\psi_\lambda(\beta,\cdot)$ admits a unique minimizer $\tau_*(\beta)$ on $(0,+\infty)$ that is also the unique solution of (9.1.4). Moreover, $\alpha \mapsto \tau_*(\alpha)$ is $\mathcal{C}^\infty$ on $(\alpha_{\min},+\infty)$ and for all $\alpha > \alpha_{\min}$

$$\left|\frac{\partial \tau_*}{\partial \alpha}(\alpha)\right| \leq (\alpha+1)\frac{\tau_*(\alpha)^3}{\delta\sigma^2}\,.$$

Proof. Most of this lemma was already proved in [71], we however provide a full proof for completeness. We have to study the fixed point equation

$$\tau^2 = \sigma^2 + \frac{1}{\delta}\mathbb{E}\left[w^*(\alpha,\tau)^2\right] = F_\alpha(\tau^2)\,,$$

where $\alpha = \lambda/\beta$. We can compute F_α explicitly:

$$\begin{aligned} F_\alpha(\tau^2) &= \sigma^2 + \frac{\tau^2}{\delta}\Big(1+\alpha^2 \\ &+ \mathbb{E}\left[(x^2-\alpha^2-1)(\Phi(\alpha-x)-\Phi(-\alpha-x))-(x+\alpha)\phi(\alpha-x)+(x-\alpha)\phi(x+\alpha)\right]\Big)\,, \end{aligned}$$

where we used the notation $x = \frac{\Theta}{\tau}$. We can then compute the derivatives:

$$\begin{aligned} F'_\alpha(\tau^2) &= \frac{1}{\delta}(1+\alpha^2)\mathbb{E}[\Phi(x-\alpha)+\Phi(-x-\alpha)] - \frac{1}{\delta}\mathbb{E}[(x+\alpha)\phi(x-\alpha)-(x-\alpha)\phi(x+\alpha)]\,, \\ F''_\alpha(\tau^2) &= \frac{-1}{2\delta\tau^2}\mathbb{E}\left[x^3(\phi(x-\alpha)-\phi(x+\alpha))\right] \leq 0\,. \end{aligned}$$

F_α is therefore concave. By dominated convergence

$$F'_\alpha(\tau^2) \xrightarrow[\tau\to+\infty]{} \frac{2}{\delta}\Big((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big)\begin{cases} <1 & \text{if } \beta < \beta_{\max} \\ \geq 1 & \text{if } \beta \geq \beta_{\max}\,. \end{cases}$$

Since $F_\alpha(0) = \sigma^2 > 0$ and by concavity of F_α, the fixed point equation admits a unique solution $\tau_*(\alpha)$ if and only is $\beta \in (0,\beta_{\max})$. In that case we have also $F'_\alpha(\tau_*(\alpha)) < 1$.

Let us now assume that $\beta \in (0,\beta_{\max})$. We have almost surely

$$\frac{\partial}{\partial\tau}\min_{w\in\mathbb{R}}\left\{\frac{w^2}{2\tau}\beta + \beta Z w + \lambda|w+\Theta|\right\} = -\frac{\beta}{2\tau^2}w^*(\alpha,\tau)^2\,.$$

Since $|w^*(\alpha,\tau)| \leq \alpha\tau + \tau|Z|$, we have by derivation under the expectation

$$\frac{\partial}{\partial\tau}\psi_\lambda(\beta,\tau) = \frac{\beta}{2} - \frac{\beta\sigma^2}{2\tau^2} - \frac{\beta}{2\delta\tau^2}\mathbb{E}\left[w^*(\alpha,\tau)^2\right] = \frac{\beta}{2\tau^2}\Big(\tau^2 - \Big(\sigma^2 + \frac{1}{\delta}\mathbb{E}\Big[w^*(\alpha,\tau)^2\Big]\Big)\Big)\,.$$

Consequently, $\tau_*(\beta)$ is the unique minimizer of $\psi_\lambda(\beta,\cdot)$ over $(0,+\infty)$.

Let us now compute $\frac{\partial\tau_*^2}{\partial\alpha}$. Since F_α is a $\mathcal{C}^\infty$ function of τ^2, one can apply the implicit function theorem to obtain that the mapping $\alpha \mapsto \tau_*(\alpha)^2$ is $\mathcal{C}^\infty$ and moreover:

$$\frac{\partial\tau_*^2}{\partial\alpha}(\alpha) = \frac{\frac{\partial F_\alpha}{\partial\alpha}(\tau_*^2(\alpha))}{1 - F'_\alpha(\tau_*^2(\alpha))}\,. \tag{9.1.5}$$

Compute

$$\frac{\partial F_\alpha}{\partial \alpha}(\tau^2) = \frac{2\tau^2}{\delta}\mathbb{E}\left[\alpha(\Phi(-\alpha+x)+\Phi(-\alpha-x)) - (\phi(\alpha-x)+\phi(\alpha+x))\right] .$$

One verify easily that

$$-1 \le -2\phi(0) \le 2\alpha\Phi(-\alpha) - 2\phi(\alpha) \le \frac{\delta}{2\tau^2}\frac{\partial F_\alpha}{\partial \alpha}(\tau^2) \le \alpha . \tag{9.1.6}$$

By concavity on has that $F'_\alpha(\tau_*^2(\alpha))$ is smaller than the slope of the line between the points of coordinates $(0,\sigma^2)$ and $(\tau_*^2(\alpha),\tau_*^2(\alpha))$:

$$F'_\alpha(\tau_*^2(\alpha)) \le 1 - \frac{\sigma^2}{\tau_*^2(\alpha)} . \tag{9.1.7}$$

From equations (9.1.5-9.1.6-9.1.7) we get

$$\left|\frac{\partial \tau_*^2}{\partial \alpha}(\alpha)\right| \le 2\frac{\tau_*(\alpha)^4}{\sigma^2\delta}(\alpha+1) .$$

The result follows then from the fact that $\frac{\partial \tau_*^2}{\partial \alpha}(\alpha) = 2\tau_*(\alpha)\frac{\partial \tau_*}{\partial \alpha}(\alpha)$. □

Define now

$$\Psi_\lambda : \beta \mapsto \min_{\tau \ge \sigma} \psi_\lambda(\beta,\tau) .$$

Lemma 9.1.6

The function Ψ_λ is differentiable on $(0,\beta_{\max})$ with derivative

$$\Psi'_\lambda(\beta) = \tau_*(\alpha) - \beta - \frac{1}{\delta}\mathbb{E}\left[Zw^*(\alpha,\tau_*(\alpha))\right] \tag{9.1.8}$$

$$= \tau_*(\alpha)\left(1 - \frac{1}{\delta}\mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\alpha)} - \alpha\Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\alpha)} - \alpha\Big)\right]\right) - \beta . \tag{9.1.9}$$

Proof. Ψ_λ is differentiable on $(0,\beta_{\max})$ (because of Lemma 9.1.5) with derivative

$$\begin{aligned}\Psi'_\lambda(\beta) &= \frac{1}{2}\Big(\frac{\sigma^2}{\tau_*(\alpha)} + \tau_*(\alpha)\Big) - \beta + \frac{1}{\delta}\Big(\frac{1}{2\tau_*(\alpha)}\mathbb{E}[w^*(\alpha,\tau_*(\alpha))^2] - \mathbb{E}[Zw^*(\alpha,\tau_*(\alpha))]\Big)\\ &= \tau_*(\alpha) - \beta - \frac{1}{\delta}\mathbb{E}\left[Zw^*(\alpha,\tau_*(\alpha))\right] ,\end{aligned}$$

because of (9.1.4). The second equality follows by Gaussian integration by parts. □

Corollary 9.1.1

The function Ψ_λ achieves its maximum over $\mathbb{R}_{\ge 0}$ at a unique $\beta_* \in (0,\beta_{\max})$.

Proof. Ψ_λ is the minimum of a collection of 1-strongly concave functions: it is therefore 1-strongly concave and admits thus a unique maximizer β_* over $\mathbb{R}_{\ge 0}$. By Lemma 9.1.4 we know that $\beta_* < \beta_{\max}$. Indeed, notice that $\max_\beta \Phi_\lambda(\beta) \ge \Psi_\lambda(0) = -\frac{\lambda}{\delta}\mathbb{E}|\Theta|$. Lemme 9.1.4 gives that $\beta_* \in [0,\beta_{\max})$, because $\Psi_\lambda(\beta_{\max}) \le \Psi_\lambda(0) - \frac{1}{2}\beta_{\max}^2 < \Psi_\lambda(0)$. By dominated convergence:

$$\mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\alpha)} - \alpha\Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\alpha)} - \alpha\Big)\right] \xrightarrow[\beta\to 0^+]{} 0 .$$

Indeed, when $\beta \to 0^+$, $\alpha = \lambda/\beta \to +\infty$ and $|\frac{\Theta}{\tau_*(\alpha)}| \leq \frac{|\Theta|}{\sigma}$. Therefore by Lemma 9.1.6 we obtain

$$\liminf_{\beta \to 0^+} \Psi'_\lambda(\beta) \geq \sigma > 0\,.$$

By concavity, we deduce that $\beta_* \in (0, \beta_{\max})$. □

Proposition 9.1.1

The function $\lambda \mapsto \beta_*(\lambda)$ is $\mathcal{C}^\infty$ and is $2\alpha_{\min}^{-1}$-Lipschitz over $(0, +\infty)$. $\lambda \mapsto \alpha_*(\lambda)$ is $\mathcal{C}^\infty$ over $(0, +\infty)$ and strictly increasing.

Proof. Let us define $\gamma_*(\lambda) = \beta_*(\lambda)/\lambda$. $\gamma_*(\lambda)$ is the unique maximizer of

$$\min_{\tau \geq \sigma} \left(\frac{\sigma^2}{\tau} + \tau\right) \frac{\gamma}{2} - \frac{\lambda}{2}\gamma^2 + \frac{1}{\delta}\mathbb{E} \min_{w \in \mathbb{R}} \left\{ \frac{w^2}{2\tau}\gamma - \gamma Z w + |w + \Theta| - |\Theta| \right\} = h(\gamma) - \frac{\lambda}{2}\gamma^2\,,$$

where h is a concave $\mathcal{C}^\infty$ function on $\mathbb{R}_{>0}$. $\gamma_*(\lambda)$ is thus the unique solution of

$$h'(\gamma) - \lambda\gamma = 0\,,$$

on $\mathbb{R}_{>0}$. $\gamma \mapsto h'(\gamma) - \lambda\gamma$ is $\mathcal{C}^\infty$ with derivative $\gamma \mapsto h''(\gamma) - \lambda < 0$. Consequently, the implicit function theorem gives that the mapping $\lambda \in \mathbb{R}_{>0} \mapsto \gamma_*(\lambda)$ is $\mathcal{C}^\infty$ and that

$$\frac{\partial \gamma_*}{\partial \lambda}(\lambda) = \frac{-\gamma_*(\lambda)}{\lambda - h''(\gamma_*(\lambda))} < 0\,.$$

One deduces that $\lambda \mapsto \alpha_*(\lambda) = \gamma_*(\lambda)^{-1}$ is $\mathcal{C}^\infty$ and strictly increasing and that $\lambda \mapsto \beta_*(\lambda) = \lambda\gamma_*(\lambda)$ is $\mathcal{C}^\infty$. Moreover

$$\left|\frac{\partial \beta_*}{\partial \lambda}(\lambda)\right| = \left|\lambda \frac{\partial \gamma_*}{\partial \lambda}(\lambda) + \gamma_*(\lambda)\right| \leq 2\gamma_*(\lambda) \leq \frac{2}{\alpha_{\min}}\,.$$

□

Proof of Theorem 9.1.1. By Corollary 9.1.1, the maximum in β in (8.3.2) is achieved at a unique $\beta_* \in (0, \beta_{\max})$. To this β_* corresponds a unique $\tau_*(\beta_*)$ that achieves the minimum in (8.3.2), by Lemma 9.1.5. By (9.1.4) and (9.1.9) we obtain that $(\tau_*(\beta_*), \beta_*)$ is solution of the system (9.1.2). Let now $(\tau, \beta) \in (0, +\infty)^2$ be another solution of (9.1.2). τ is therefore solution of (9.1.4) which gives that $\beta \in (0, \beta_{\max})$ and $\tau = \tau_*(\beta)$ by Lemma 9.1.5. The second equality in (9.1.2) gives that $\Psi'_\lambda(\beta) = 0$ and thus that $\beta = \beta_*$ by strong concavity of Ψ_λ. We conclude $(\tau, \beta) = (\tau_*(\beta_*), \beta_*)$. □

9.1.2 Control on β_*, τ_*

The goal of this section is to show that β_* and τ_* remain bounded when $\theta^\star$ varies in $\mathcal{D}$.

Theorem 9.1.2

There exists constants $\beta_{\min}, \tau_{\max} > 0$ that only depend on Ω such that for all $\theta^\star \in \mathcal{D}$ and all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$,

$$\beta_{\min} \leq \beta_*(\lambda) < \beta_{\max} \quad \textit{and} \quad \sigma \leq \tau_*(\lambda) \leq \tau_{\max}\,.$$

To prove Theorem 9.1.2, we separate the case where $\mathcal{D} = \mathcal{F}_p(\xi)$ (where it follows from Lemma 9.1.9 and Corollary 9.1.2 below) from the case where $\mathcal{D} = \mathcal{F}_0(s)$ (where it follows from Lemmas 9.1.11 and 9.1.12).

Technical lemmas

Lemma 9.1.7

We have

$$\max_{\beta\geq 0}\min_{\tau\geq\sigma}\psi_\lambda(\beta,\tau)=\psi_\lambda(\beta_*,\tau_*(\beta_*))=\frac{1}{2}\beta_*^2+\frac{\lambda}{\delta}\mathbb{E}\Big[|w^*(\alpha_*,\tau_*(\beta))+\Theta|-|\Theta|\Big]$$
$$=\frac{1}{2}\beta_*^2+\tau_*(\beta_*)\frac{\lambda}{\delta}\mathbb{E}\left[H_{\alpha_*}\left(\frac{\Theta}{\tau_*(\beta_*)}\right)\right],$$

where

$$H_\alpha(x)=(x-\alpha)\Phi(-\alpha+x)+(-x-\alpha)\Phi(-x-\alpha)+\phi(-x+\alpha)+\phi(x+\alpha)-|x|\,.$$

Proof. Using the optimality condition (9.1.4) of $\tau_*(\beta)$, we have for all $\beta\in(0,\beta_{\max})$

$$\psi_\lambda(\beta,\tau_*(\beta))=-\frac{1}{2}\beta^2+\beta\tau_*(\beta)-\frac{\beta}{\delta}\mathbb{E}\left[Zw^*(\alpha,\tau_*(\beta))\right]+\frac{\lambda}{\delta}\mathbb{E}\left[|w^*(\alpha,\tau_*(\beta))+\Theta|-|\Theta|\right].$$

At β_* the optimality condition (see (9.1.8)) reads $\beta_*=\tau_*(\beta_*)-\frac{1}{\delta}\mathbb{E}[Zw^*(\alpha_*,\tau_*(\beta_*))]$, thus

$$\psi_\lambda(\beta_*,\tau_*(\beta_*))=\frac{1}{2}\beta_*^2+\frac{\lambda}{\delta}\mathbb{E}[|w^*(\alpha_*,\tau_*(\beta))+\Theta|-|\Theta|]\,. \tag{9.1.10}$$

Compute for $\alpha,\tau>0$

$$\mathbb{E}|w^*(\alpha,\tau)+\Theta|=\mathbb{E}|\eta(\Theta+\tau Z,\alpha\tau)|=\tau\mathbb{E}\Big|\eta\Big(\frac{\Theta}{\tau}+Z,\alpha\Big)\Big|\,. \tag{9.1.11}$$

Now, for $x\in\mathbb{R}$,

$$\begin{aligned}\mathbb{E}|\eta(x+Z,\alpha)|&=\int_{\alpha-x}^{+\infty}(x+z-\alpha)\phi(z)dz+\int_{-\infty}^{-\alpha-x}(-x-z-\alpha)\phi(z)dz\\&=(x-\alpha)\Phi(x-\alpha)+(-x-\alpha)\Phi(-x-\alpha)+\phi(\alpha-x)+\phi(\alpha+x)\\&=H_\alpha(x)+|x|\,.\end{aligned}$$

and we obtain the Lemma by putting this together with (9.1.11) and (9.1.10). □

The next Lemma summarizes the main properties of H_α.

Lemma 9.1.8

H_α is a continuous, even function and for $x>0$

$$H'_\alpha(x)=\Phi(x-\alpha)-\Phi(-x-\alpha)-1\in(-1,0)\,.$$

H_α is therefore 1-Lipschitz. H_α admits a maximum at 0 and

$$H_\alpha(0)=2\phi(\alpha)-2\alpha\Phi(-\alpha)>0\,.$$

Moreover $H_\alpha(x)\xrightarrow[x\to+\infty]{}-\alpha$.

On ℓ_p-balls

Lemma 9.1.9

Assume that $\mathbb{E}\Big[|\Theta|^p\Big] \leq \xi^p$ *for some* $\xi, p > 0$. *Then, there exists a constant* $\beta_{\min} = \beta_{\min}(\delta, \lambda_{\min}, \xi, p, \sigma)$ *such that for all* $\lambda \geq \lambda_{\min}$,

$$0 < \beta_{\min} \leq \beta_*(\lambda) < \beta_{\max}\,.$$

Proof. Let $\beta \in (0, \beta_{\max})$. By Lemma 9.1.6 we have

$$\Psi'_\lambda(\beta) = \tau_*(\beta)\left(1 - \frac{1}{\delta}\mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\beta)} - \alpha\Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\beta)} - \alpha\Big)\right]\right) - \beta\,.$$

The function $g_\alpha : x \mapsto \Phi(x-\alpha) + \Phi(-x-\alpha)$ is even, and increasing over $\mathbb{R}_{\geq 0}$. Let $K > 0$ such that $\frac{\xi^p}{K^p\sigma^p} \leq \frac{\delta}{4}$. By Markov's inequality we have

$$\mathbb{P}\left(\left|\frac{\Theta}{\tau_*(\beta)}\right| \geq K\right) \leq \mathbb{P}\left(\left|\frac{\Theta}{\sigma}\right|^p \geq K^p\right) \leq \frac{1}{K^p\sigma^p}\mathbb{E}|\Theta|^p \leq \frac{\delta}{4}\,.$$

Thus

$$\mathbb{E}\left[g_\alpha\left(\frac{\Theta}{\tau_*(\beta)}\right)\right] \leq g_\alpha(K) + \frac{\delta}{4}\,.$$

As $\beta \to 0$, $\alpha \geq \lambda_{\min}/\beta \to +\infty$. Since $g_\alpha(K) \xrightarrow[\alpha\to+\infty]{} 0$ there exists $\beta_0 = \beta_0(K, \lambda_{\min}, \delta) > 0$ such that for all $\beta \in (0, \beta_0)$, $g_\alpha(K) \leq \frac{\delta}{4}$. Thus for all $\beta \in (0, \beta_0)$,

$$\Psi'_\lambda(\beta) \geq \tau_*(\beta)\left(1 - \frac{\frac{\delta}{4} + \frac{\delta}{4}}{\delta}\right) - \beta \geq \frac{\sigma}{2} - \beta\,.$$

Let $\beta_{\min} = \min(\frac{\sigma}{2}, \beta_0)$. We conclude that for all $\beta \in (0, \beta_{\min})$, $\Psi'_\lambda(\beta) > 0$. By concavity we have then that $\beta_* \geq \beta_{\min}$. The other inequality $\beta_* < \beta_{\max}$ was already proved in Corollary 9.1.1.

□

Corollary 9.1.2

Assume that $\mathbb{E}\Big[|\Theta|^p\Big] \leq \xi^p$ *for some* $\xi, p > 0$. *Then there exists a constant* $\tau_{\max} = \tau_{\max}(\xi, p, \delta, s, \lambda_{\min}, \lambda_{\max})$ *such that for all* $\lambda \in [\lambda_{\min}, \lambda_{\max}]$,

$$\sigma \leq \tau_*(\beta_*(\lambda)) \leq \tau_{\max}\,.$$

Proof. Let $t \geq \xi$. By Markov's inequality we have $\mathbb{P}(|\Theta| \geq t) \leq \left(\frac{\xi}{t}\right)^p \leq 1$, since $\theta^\star \in \mathcal{F}_p(\xi)$.

$$\begin{aligned}\mathbb{E}\left[H_{\alpha_*}\left(\frac{\Theta}{\tau_*(\alpha_*)}\right)\right] &= \mathbb{E}\left[\mathbb{1}(|\Theta| < t)H_{\alpha_*}\left(\frac{\Theta}{\tau_*(\alpha_*)}\right)\right] + \mathbb{E}\left[\mathbb{1}(|\Theta| \geq t)H_{\alpha_*}\left(\frac{\Theta}{\tau_*(\alpha_*)}\right)\right] \\ &\geq \left(1 - \left(\frac{\xi}{t}\right)^p\right)\left(H_{\alpha_*}(0) - \frac{t}{\tau_*(\alpha_*)}\right) - \alpha_*\left(\frac{\xi}{t}\right)^p,\end{aligned}$$

because by Lemma 9.1.8, H_{α_*} is 1-Lipschitz and for all $x \in \mathbb{R}$, $-\alpha_* \leq H_{\alpha_*}(x) \leq H_{\alpha_*}(0)$. Replacing $H_{\alpha_*}(0)$ by its expression given by Lemma 9.1.8 we get

$$\mathbb{E}\left[H_{\alpha_*}\left(\frac{\Theta}{\tau_*(\alpha_*)}\right)\right] \geq 2(\phi(\alpha_*) - \alpha_*\Phi(-\alpha_*)) - \left(\frac{\xi}{t}\right)^p(\alpha_* + 2(\phi(\alpha_*) - \alpha_*\Phi(-\alpha_*))) - \frac{t}{\tau_*(\alpha_*)}\,.$$

Since $\alpha_* \leq \lambda_{\max}/\beta_{\min}$ and $\phi(\alpha_*) - \alpha_*\Phi(-\alpha_*) > 0$ (because $\alpha_* > \alpha_{\min}$), we can find a constant $t = t(\delta, \sigma, \lambda_{\min}, \lambda_{\max}, p, \xi) \geq \xi$ such that

$$(\alpha_* + 2(\phi(\alpha_*) - \alpha_*\Phi(-\alpha_*))) \left(\frac{\xi}{t}\right)^p \leq \phi(\alpha_*) - \alpha_*\Phi(-\alpha_*)\,.$$

For this choice of t we have then

$$\frac{\tau_*(\alpha_*)\lambda}{\delta}\mathbb{E}\left[H_{\alpha_*}\left(\frac{\Theta}{\tau_*}\right)\right] \geq \frac{\lambda}{\delta}\tau_*(\alpha_*)(\phi(\alpha_*) - \alpha_*\Phi(-\alpha_*)) - \frac{\lambda t}{\delta}\,.$$

Consequently by Lemma 9.1.2 and Lemma 9.1.7 we have

$$\frac{\beta_*^2}{2} + \frac{\lambda}{\delta}\tau_*(\phi(\alpha_*) - \alpha_*\Phi(-\alpha_*)) - \frac{\lambda t}{\delta} \leq \psi_\lambda(\beta_*, \tau_*(\alpha_*)) \leq \frac{\sigma^2}{2}\,,$$

which finally gives

$$\tau_*(\alpha_*) \leq \frac{\delta\sigma^2\lambda^{-1} + t}{\phi(\alpha_{\max}) - \alpha_{\max}\Phi(-\alpha_{\max})}\,.$$

□

On sparse balls

Define the critical function:

$$M_s : \alpha \mapsto s(1+\alpha^2) + 2(1-s)\Big((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big)\,.$$

M_s corresponds to the worst mean squared error achievable by soft-thresholding with threshold α to estimate a vector $\theta^\star \in \mathcal{F}_0(s)$ from the observations $y = \theta^\star + w$, where $w \sim \mathcal{N}(0, I_N)$, see [70, 69, 115].

Lemma 9.1.10 (*From [71]*)

Assume that

$$s < s_{\max}(\delta) = \delta \max_{\alpha \geq 0} \left\{ \frac{1 - \frac{2}{\delta}\Big((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big)}{1 + \alpha^2 - 2\Big((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)\Big)} \right\}\,.$$

Then there exists $\alpha \geq 0$ such that $M_s(\alpha) < \delta$.

Proof. Let $s < s_{\max}(\delta)$. From the definition of $s_{\max}(\delta)$, we can find $\alpha \in \mathbb{R}$ such that

$$\delta\frac{1 - \frac{2}{\delta}((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha))}{1 + \alpha^2 - 2((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha))} > s\,,$$

which gives $M_s(\alpha) < \delta$. □

We assume in this section that $s < s_{\max}(\delta)$. Let us compute the derivatives

$$\begin{aligned} M_s'(\alpha) &= 2\left(\alpha s + 2(1-s)(\alpha\Phi(-\alpha) - \phi(\alpha))\right)\,, \\ M_s''(\alpha) &= 2\left(s + (1-s)2\Phi(-\alpha)\right) > 0\,. \end{aligned}$$

Notice that $M_s(\alpha) = \frac{1}{2}\Big(\alpha M_s'(\alpha) + M_s''(\alpha)\Big)$. Let α_0 be the unique $\alpha > 0$ such that $M_s'(\alpha) = 0$ and let $\alpha_1 < \alpha_2$ be such that $M_s(\alpha_1) = M_s(\alpha_2) = \delta$. We can then easily plot the variations of M_s:

α	0	α_1	α_0	α_2	$+\infty$
M_s'		$-$	0	$+$	
M_s	1	$\searrow$ δ $\searrow$	$M_s(\alpha_0)$	$\nearrow$ δ $\nearrow$	$+\infty$
$\frac{1}{2}M_s''$	1	$\searrow$	$M_s(\alpha_0)$	$\searrow$	s

Lemma 9.1.11

Let $s < s_{\max}(\delta)$ and assume that $\mathbb{P}(\Theta \neq 0) \leq s$. Then, there exists a constant $\beta_{\min} = \beta_{\min}(\delta, \lambda_{\min}, s, \sigma)$ such that for all $\lambda \geq \lambda_{\min}$

$$0 < \beta_{\min} \leq \beta_*(\lambda) < \beta_{\max}(\lambda_{\max}, \delta) := \lambda_{\max}/\alpha_{\min}(\delta)\,.$$

Proof. We already proved in Corollary 9.1.1 that $\beta_*(\lambda) < \lambda/\alpha_{\min}$. For all $0 < \beta < \lambda/\alpha_{\min}$, we have by Lemma 9.1.6

$$\Psi_\lambda'(\beta) = \tau_*(\beta)\left(1 - \frac{1}{\delta}\mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\beta)} - \alpha\Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\beta)} - \alpha\Big)\right]\right) - \beta\,.$$

The function $g_\alpha : x \mapsto \Phi(x-\alpha) + \Phi(-x-\alpha)$ is even, and increasing over $\mathbb{R}_{\geq 0}$. Therefore

$$\Psi_\lambda'(\beta) \geq \frac{\tau_*(\beta)}{\delta}(\delta - s - (1-s)2\Phi(-\alpha)) - \beta\,.$$

Let $\beta_0 = \beta_0(\lambda_{\min}, \delta, s) > 0$ such that for all $\beta \in (0, \beta_0)$, $2\Phi(-\alpha) \leq \frac{1}{2}(\delta - s)$. For all $\beta \in (0, \beta_0)$ we have then

$$\Psi_\lambda'(\beta) \geq \frac{\sigma(\delta - s)}{2\delta} - \beta\,.$$

Let $\beta_{\min} = \min(\frac{\sigma(\delta-s)}{2\delta}, \beta_0)$: for all $\beta \in (0, \beta_{\min})$, $\Psi_\lambda'(\beta) > 0$. By concavity we conclude that $\beta_* \geq \beta_{\min}$. □

Lemma 9.1.12

Let $s < s_{\max}(\delta)$ and assume that $\mathbb{P}(\Theta \neq 0) \leq s$. Then for all $\beta, \tau, \lambda > 0$ we have

$$\psi_\lambda(\beta, \tau) \geq \frac{\beta\sigma^2}{2\tau} - \frac{\beta^2}{2} + \frac{\tau\beta}{2\delta}\Big(\delta - M_s(\alpha)\Big)\,.$$

Proof. By (9.1.3) we have for all $\beta, \tau > 0$

$$\psi_\lambda(\beta, \tau) = \frac{\beta}{2}\left(\frac{\sigma^2}{\tau} + \tau\right) - \frac{\beta^2}{2} + \frac{\tau\beta}{\delta}\mathbb{E}\left[\Delta_\alpha\Big(\frac{\Theta}{\tau}\Big) - \frac{1}{2}\right]\,.$$

Since by Lemma 9.6.13, Δ_α is even and non-increasing over $\mathbb{R}_{\geq 0}$, we have

$$\begin{aligned}
\mathbb{E}\left[\Delta_\alpha\Big(\frac{\Theta}{\tau}\Big) - \frac{1}{2}\right] &\geq s\Delta_\alpha(+\infty) + (1-s)\Delta_\alpha(0) - \frac{1}{2} \\
&= -s\frac{\alpha^2}{2} + (1-s)\Big(\frac{1}{2} + \alpha\phi(\alpha) - (1+\alpha^2)\Phi(-\alpha)\Big) - \frac{1}{2} \\
&= -\frac{1}{2}M_s(\alpha)\,.
\end{aligned}$$

□

Lemma 9.1.13

Let $s < s_{\max}(\delta)$ and assume that $\mathbb{P}(\Theta \neq 0) \leq s$. Then the following inequalities hold

$$\beta_* \tau_*(\beta_*)(\delta - M_s(\alpha_*)) \leq \delta\Big(\sigma^2 + \beta_*^2\Big)\,, \tag{9.1.12}$$

$$-\tau_*(\beta_*)\lambda M_s'(\alpha_*) \leq \delta\sigma^2\,, \tag{9.1.13}$$

$$\tau_*(\beta_*)(\delta - \frac{1}{2}M_s''(\alpha_*)) \leq \beta_*\,. \tag{9.1.14}$$

Proof. The inequality (9.1.12) simply follows from the previous lemma and from the fact that

$$\psi_\lambda(\beta, \tau_*(\beta)) \leq \max_{\beta \geq 0} \min_{\tau \geq \sigma} \psi_\lambda(\beta, \tau) \leq \frac{\sigma^2}{2}\,,$$

by Lemma 9.1.2. Let us prove (9.1.13). By Lemma 9.1.7, we have

$$\psi_\lambda(\beta_*, \tau_*(\beta_*)) \geq \frac{\lambda \tau_*(\beta_*)}{\delta} \mathbb{E}\left[H_{\alpha_*}\Big(\frac{\Theta}{\tau_*(\beta_*)}\Big)\right]\,.$$

Since by Lemma 9.1.8, H_{α_*} is even, decreasing on $\mathbb{R}_{\geq 0}$, we have

$$\begin{aligned}\mathbb{E}\left[H_{\alpha_*}\Big(\frac{\Theta}{\tau_*(\beta_*)}\Big)\right] &\geq s H_{\alpha_*}(+\infty) + (1-s)H_{\alpha_*}(0) \\ &= -s\alpha_* + 2(1-s)\big(\phi(\alpha_*) - \alpha_*\Phi(-\alpha_*)\big) = -\frac{1}{2}M_s'(\alpha_*)\,,\end{aligned}$$

which proves (9.1.13). To prove (9.1.14) we use the optimality condition at β_*:

$$0 = \Psi_\lambda'(\beta_*) = \tau_*(\alpha_*)\left(1 - \frac{1}{\delta}\mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\alpha_*)} - \alpha_*\Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\alpha_*)} - \alpha_*\Big)\right]\right) - \beta_*\,. \tag{9.1.15}$$

The function $x \mapsto \Phi(x - \alpha_*) + \Phi(-x - \alpha_*)$ is even, increasing on $\mathbb{R}_{\geq 0}$. Therefore

$$\mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\alpha_*)} - \alpha_*\Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\alpha_*)} - \alpha_*\Big)\right] \leq s + 2(1-s)\Phi(-\alpha_*) = \frac{1}{2}M_s''(\alpha_*)\,.$$

Combining this inequality with (9.1.15) leads to (9.1.14). □

Proposition 9.1.2

Let $s < s_{\max}(\delta)$ and assume that $\mathbb{P}(\Theta \neq 0) \leq s$. Then, there exists a constant $\tau_{\max} = \tau_{\max}(\delta, \lambda_{\min}, \lambda_{\max}, s, \sigma)$ such that for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$,

$$\sigma \leq \tau_*(\beta_*(\lambda)) \leq \tau_{\max}\,.$$

Proof. Let (β_*, τ_*) be the unique optimal couple and recall $\alpha_* = \lambda/\beta_*$. We distinguish 3 cases:

Case 1: $\alpha_* \geq \alpha_0$. In that case $\frac{1}{2}M_s''(\alpha_*) \leq M_s(\alpha_0) < \delta$. The inequality (9.1.14) gives

$$\tau_*(\beta_*)(\delta - \frac{1}{2}M_s''(\alpha_*)) \leq \beta_* \leq \beta_{\max}\,,$$

which gives $\tau_*(\beta_*) \leq \frac{\beta_{\max}}{\delta - M_s(\alpha_0)}$.

Case 2: $\alpha_* \in [(\alpha_1+\alpha_0)/2, \alpha_0]$. In that case $\delta - M_s(\alpha_*) \geq c > 0$, for some constant $c = c(\delta, s) > 0$. Now, by (9.1.12)

$$\beta_* \tau_*(\beta_*)(\delta - M_s(\alpha_*)) \leq \delta(\sigma^2 + \beta_*^2)\,.$$

Therefore,

$$\tau_* \leq \frac{\delta}{c\beta_{\min}}(\sigma^2 + \beta_{\max}^2)\,.$$

Case 3: $\alpha_* < (\alpha_1+\alpha_0)/2$. In that case $M_s'(\alpha_*) \leq -c$, for some constant $c = c(\delta, s) > 0$. Consequently by (9.1.13) we get

$$\tau_*(\beta_*) \leq \frac{\sigma^2 \delta}{\lambda c}\,.$$

□

9.1.3 Dependency in λ

Proposition 9.1.3

- *The mapping $\lambda \mapsto \beta_*(\lambda)$ is $\mathcal{C}^\infty$ and $2\alpha_{\min}^{-1}$-Lipschitz on $\mathbb{R}_{>0}$.*
- *The mapping $\lambda \mapsto \tau_*(\lambda)$ is $\mathcal{C}^\infty$ and M-Lipschitz on $[\lambda_{\min}, \lambda_{\max}]$, for some constant $M(\Omega) > 0$.*

Proof. The first point has already been by Proposition 9.1.1. $\lambda \mapsto \tau_*(\lambda)$ is the composition of the mappings $\lambda \mapsto \alpha_*(\lambda)$ and $\alpha \mapsto \tau_*(\alpha)$, that are both $\mathcal{C}^\infty$ by Lemma 9.1.5 and Proposition 9.1.1. Compute the derivative:

$$\frac{\partial \tau_*}{\partial \lambda}(\lambda) = \frac{\partial \alpha_*}{\partial \lambda}(\lambda)\frac{\partial \tau_*}{\partial \alpha}(\alpha_*(\lambda))\,.$$

Recall that $\alpha_*(\lambda) = \lambda/\beta_*(\lambda)$. Thus

$$\left|\frac{\partial \alpha_*}{\partial \lambda}(\lambda)\right| = \left|\frac{1}{\beta_*(\lambda)} - \frac{\partial \beta_*}{\partial \lambda}(\lambda)\frac{\lambda}{\beta_*(\lambda)^2}\right| \leq \beta_{\min}^{-1} + 2\alpha_{\min}^{-1}\lambda_{\max}\beta_{\min}^{-2}\,.$$

By Lemma 9.1.5, we have

$$\left|\frac{\partial \tau_*}{\partial \alpha}(\alpha_*(\lambda))\right| \leq (\alpha_*(\lambda)+1)\frac{\tau_*(\alpha_*)^3}{\delta\sigma^2}\,.$$

Since by Theorem 9.1.2, $\tau_*(\alpha_*) \leq \tau_{\max}(\Omega)$ and $\alpha_* \leq \lambda_{\max}/\beta_{\min}(\Omega)$, the derivative of τ_* with respect to λ is bounded on $[\lambda_{\min}, \lambda_{\max}]$. □

9.2 Study of Gordon's optimization problem for $\widehat{w}_\lambda$

In this section we study L_λ defined by (8.5.3). Define, for $w \in \mathbb{R}^N$ and $\beta \geq 0$

$$\ell_\lambda(w, \beta) = \left(\sqrt{\frac{\|w\|^2}{n} + \sigma^2}\,\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}}\right)\beta - \frac{1}{2}\beta^2 + \frac{\lambda}{n}|w+\theta^\star| - \frac{\lambda}{n}|\theta^\star|\,. \quad (9.2.1)$$

So that $L_\lambda(w) = \max_{\beta \geq 0} \ell_\lambda(w, \beta)$. Let us define the vector $\mathsf{w}_\lambda \in \mathbb{R}^N$ by

$$\mathsf{w}_{\lambda,i} = \eta\left(\theta_i^\star + \tau_*(\lambda) g_i, \frac{\tau_*(\lambda)\lambda}{\beta_*(\lambda)}\right) - \theta_i^\star \,. \tag{9.2.2}$$

The goal of this section is to prove that, with high probability, the minimizer of L_λ is close to w_λ and that L_λ is strongly convex around w_λ.

Proposition 9.2.1

L_λ admits almost surely a unique minimizer w_λ^* on $\mathbb{R}^N$.

Proof. L_λ is a convex function that goes to $+\infty$ at infinity, so it admits minimizers over $\mathbb{R}^N$. We distinguish two cases:

Case 1: there is a minimizer w such that $\sqrt{\frac{\|w\|^2}{n} + \sigma^2}\, \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} > 0$.
In that case, there exist a neighborhood O_w of w such that for all $w' \in O_w$

$$a(w') := \sqrt{\frac{\|w'\|^2}{n} + \sigma^2}\, \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w' + \frac{g'\sigma}{\sqrt{n}} > 0 \,.$$

Thus for all $w' \in O_w$, $L_\lambda(w') = \frac{1}{2} a(w')^2 + \frac{\lambda}{n}|w' + \theta^\star| - \frac{\lambda}{n}|\theta^\star|$. Recall that the composition of a strictly convex function and a strictly increasing function is strictly convex. L_λ is therefore strictly convex on O_w because a is strictly convex and remains strictly positive on O_w and because $x > 0 \mapsto x^2$ is strictly increasing. w is thus the only minimizer of L_λ.

Case 2: for all minimizer w we have $\sqrt{\frac{\|w\|^2}{n} + \sigma^2}\, \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \leq 0$.
Let w be a minimizer of L_λ. The optimality condition gives

$$\frac{-1}{\lambda}\left(\sqrt{\frac{\|w\|^2}{n} + \sigma^2}\, \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}}\right)_+ \left(\frac{w}{\sqrt{\frac{\|w\|^2}{n} + \sigma^2}}\, \frac{\|h\|}{\sqrt{n}} - g\right) \in \partial|\theta^\star + w|.$$

We obtain then $0 \in \partial|\theta^\star + w|$ which implies $w = -\theta^\star$: L_λ has a unique minimizer. □

9.2.1 Local stability of Gordon's optimization

Theorem 9.2.1

There exists constants $\gamma, c, C > 0$ that only depend on Ω such that for all $\theta^\star \in \mathcal{D}$, all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and all $\epsilon \in (0, 1]$

$$\mathbb{P}\Big(\exists w \in \mathbb{R}^N, \quad \frac{1}{N}\|w - \mathsf{w}_\lambda\|^2 > \epsilon \quad \text{and} \quad L_\lambda(w) \leq \min_{v \in \mathbb{R}^N} L_\lambda(v) + \gamma\epsilon\Big) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2} \,.$$

We deduce from Theorem 9.2.1 that for all $\epsilon \in (0, 1]$ with probability at least $1 - C\epsilon^{-1}e^{-cn\epsilon^2}$, $\frac{1}{N}\|w_\lambda^* - \mathsf{w}_\lambda\|^2 \leq \epsilon$. From this we deduce easily that with the same probability $|L_\lambda(w_\lambda^*) - L_\lambda(\mathsf{w}_\lambda)| \leq M\epsilon$, for some constant $M > 0$, which gives by Proposition 9.6.1:

Corollary 9.2.1

Define

$$L_*(\lambda) = \psi_\lambda(\beta_*(\lambda), \tau_*(\lambda))\,. \tag{9.2.3}$$

The exists constants $c, C > 0$ that only depend on Ω such that

$$\mathbb{P}\left(\left|\min_{w\in\mathbb{R}^N} L_\lambda(w) - L_*(\lambda)\right| \geq \epsilon\right) \leq \frac{C}{\epsilon} e^{-nc\epsilon^2}\,.$$

9.2.2 Proof of Theorem 9.2.1

Proposition 9.2.2

For all $R > 0$ there exists constants $c, C > 0$ that only depend on (Ω, R), such that for all $\epsilon \in (0, 1]$,

$$\forall \theta^\star \in \mathcal{D},\ \forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \mathbb{P}\Big(L_\lambda(\mathsf{w}_\lambda) \leq \min_{\|w\|\leq\sqrt{n}R} L_\lambda(w) + \epsilon\Big) \geq 1 - \frac{C}{\epsilon} e^{-cn\epsilon^2}\,.$$

Proof. Notices that it suffices to proves the proposition for ϵ smaller than some constant. Let $\theta^\star \in \mathcal{D}$, $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. Let $R > 0$ and $\epsilon \in (0, \min(1, \sigma^2/2)]$. Define

$$\ell^\circ_\lambda(w, \beta) = \left(\sqrt{\frac{\|w\|^2}{n} + \sigma^2} - \frac{1}{n} g^\mathsf{T} w\right)\beta - \frac{1}{2}\beta^2 + \frac{\lambda}{n}|w + \theta^\star| - \frac{\lambda}{n}|\theta^\star|\,.$$

On the event

$$\left\{\left|\frac{1}{n}\|h\|^2 - 1\right| \leq \epsilon\right\} \bigcap \left\{\left|\frac{g'\sigma}{\sqrt{n}}\right| \leq \epsilon\right\}\,, \tag{9.2.4}$$

which has probability at least $1 - Ce^{-cn\epsilon^2}$, we have, for all $w \in B(0, R\sqrt{n})$ and $\beta \in [0, \beta_{\max}]$:

$$\begin{aligned}
|\ell_\lambda(w, \beta) - \ell^\circ_\lambda(w, \beta)| &= \beta\sqrt{\frac{\|w\|^2}{n} + \sigma^2}\ \left|\frac{\|h\|}{\sqrt{n}} - 1\right| + \beta\left|\frac{g'\sigma}{\sqrt{n}}\right| \\
&\leq \beta_{\max}\sqrt{\sigma^2 + R^2}\left|\frac{1}{n}\|h\|^2 - 1\right| + \beta_{\max}\epsilon \\
&\leq \underbrace{\beta_{\max}(\sqrt{\sigma^2 + R^2} + 1)}_{K}\epsilon\,.
\end{aligned}$$

For simplicity we write $(\beta_*, \tau_*) = (\beta_*(\lambda), \tau_*(\lambda))$. We have on the event (9.2.4):

$$\begin{aligned}
\min_{\|w\|\leq R\sqrt{n}} L_\lambda(w) &= \min_{\|w\|\leq R\sqrt{n}} \max_{\beta\geq 0} \ell_\lambda(w, \beta) \geq \min_{\|w\|\leq R\sqrt{n}} \ell_\lambda(w, \beta_*) \\
&\geq \min_{\|w\|\leq R\sqrt{n}} \ell^\circ_\lambda(w, \beta_*) - K\epsilon\,.
\end{aligned}$$

Using the fact that for $w \in B(0, R\sqrt{n})$

$$\sqrt{\frac{\|w\|^2}{n} + \sigma^2} = \min_{\sigma\leq\tau\leq\sqrt{\sigma^2+R^2}} \left\{\frac{\frac{\|w\|^2}{n} + \sigma^2}{2\tau} + \frac{\tau}{2}\right\}\,,$$

we obtain that

$$\min_{\|w\|\le R\sqrt{n}} \ell_\lambda^\circ(w,\beta_*) = \min_{\sigma\le\tau\le\sqrt{\sigma^2+R^2}} \left\{ \frac{\beta_*}{2}\left(\frac{\sigma^2}{\tau}+\tau\right) - \frac{\beta_*^2}{2} + \frac{1}{n}\min_{\|w\|\le R\sqrt{n}} \left\{ \frac{\beta_*}{2\tau}\|w\|^2 - \beta_* g^\mathsf{T} w + \lambda|w+\theta^\star| - \lambda|\theta^\star| \right\} \right\}.$$

For all $\tau \in [\sigma, \sqrt{\sigma^2+R^2}]$ the function

$$g \mapsto \min_{\|w\|\le R\sqrt{n}} \left\{ \frac{\beta_*}{2\tau}\|w\|^2 - \beta_* g^\mathsf{T} w + \lambda|w+\theta^\star| - \lambda|\theta^\star| \right\}$$

is $\beta_{\max} R\sqrt{n}$-Lipschitz. Therefore

$$\mathsf{F}(\tau,g) = \frac{\beta_*}{2}\left(\frac{\sigma^2}{\tau}+\tau\right) - \frac{\beta_*^2}{2} + \frac{1}{n}\min_{\|w\|\le R\sqrt{n}} \left\{ \frac{\beta_*}{2\tau}\|w\|^2 - \beta_* g^\mathsf{T} w + \lambda|w+\theta^\star| - \lambda|\theta^\star| \right\}$$

is $\beta_{\max}^2 R^2 n^{-1}$-sub-Gaussian. Therefore there exists constants $C, c > 0$ such that for all $\tau \in [\sigma, \sqrt{\sigma^2+R^2}]$, we have

$$\mathbb{P}\left(|\mathsf{F}(\tau,g) - \mathbb{E}\mathsf{F}(\tau,g)| > \epsilon\right) \le Ce^{-cn\epsilon^2}.$$

$\mathsf{F}(\cdot, g)$ is almost surely a $\beta_{\max}(1+\frac{R^2}{\sigma^2})$-Lipschitz function on $[\sigma, \sqrt{\sigma^2+R^2}]$. Therefore, by an ϵ-net argument one can find constants $C, c > 0$ that only depend on (Ω, R), such that for all $\epsilon > 0$ the event

$$\left\{ \sup_{\tau\in[\sigma,\sqrt{\sigma^2+R^2}]} |\mathsf{F}(\tau,g) - \mathbb{E}\mathsf{F}(\tau,g)| \le \epsilon \right\} \tag{9.2.5}$$

has probability at least $1 - \frac{C}{\epsilon}e^{-cn\epsilon^2}$. On the event (9.2.5) we have then $\min_{\tau\in[\sigma,\sqrt{\sigma^2+R^2}]} \mathsf{F}(\tau,g) \ge \min_{\tau\in[\sigma,\sqrt{\sigma^2+R^2}]} \mathbb{E}\left[\mathsf{F}(\tau,g)\right] - \epsilon$. Notice that for all $\tau > 0$ we have

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}\left[\min_{\|w\|\le R\sqrt{n}} \left\{ \frac{\beta_*}{2\tau}\|w\|^2 - \beta_* g^\mathsf{T} w + \lambda|w+\theta^\star| - \lambda|\theta^\star| \right\}\right] \\
&\qquad \ge \frac{1}{n}\sum_{i=1}^N \mathbb{E}\left[\min_{w_i\in\mathbb{R}} \left\{ \frac{\beta_*}{2\tau} w_i^2 - \beta_* g_i w_i + \lambda|w_i+\theta_i^\star| - \lambda|\theta_i^\star| \right\}\right] \\
&\qquad = \frac{1}{\delta}\mathbb{E}\left[\min_{w\in\mathbb{R}} \left\{ \frac{\beta_*}{2\tau} w^2 - \beta_* Z w + \lambda|w+\Theta| - \lambda|\Theta| \right\}\right],
\end{aligned}$$

where the last expectation is with respect $(\Theta, Z) \sim \widehat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$. Consequently on the event (9.2.4) and (9.2.5), we have

$$(1+K)\epsilon + \min_{\|w\|\le R\sqrt{n}} L_\lambda(w) \ge \min_{\sigma\le\tau\le\sqrt{\sigma^2+R^2}} \psi_\lambda(\beta_*,\tau) \ge \min_{\sigma\le\tau} \psi_\lambda(\beta_*,\tau) = \psi_\lambda(\beta_*,\tau_*).$$

By Proposition 9.6.1 we have that

$$\psi_\lambda(\beta_*,\tau_*) \ge L_\lambda(\mathsf{w}_\lambda) - \epsilon,$$

with probability at least $1 - Ce^{-cn\epsilon^2}$. Then, for all $\epsilon \in (0,1)$ we have with probability at least $1 - \frac{C}{\epsilon}e^{-cn\epsilon^2}$

$$\min_{\|w\|\le R\sqrt{n}} L_\lambda(w) + (K+2)\epsilon \ge L_\lambda(\mathsf{w}_\lambda).$$

□

Lemma 9.2.1

Let f be a convex function on $\mathbb{R}^N$. Let $w \in \mathbb{R}^N$ and $r > 0$. Suppose that f is γ-strongly convex on the ball $B(w, r)$, for some $\gamma > 0$. Assume that

$$f(w) \leq \min_{x \in B(w,r)} f(x) + \epsilon\,,$$

for some $\epsilon \leq \frac{r^2\gamma}{8}$. Then f admits a unique minimizer x^* over $\mathbb{R}^N$. We have $x^* \in B(w, r)$ and therefore

$$\|x^* - w\|^2 \leq \frac{2}{\gamma}\epsilon\,.$$

Moreover, for every $x \in \mathbb{R}^N$ such that $f(x) \leq \min f + \epsilon$ we have $\|x - w\|^2 \leq \frac{8}{\gamma}\epsilon$.

Proof. f is convex on $B(w, r)$, it admits therefore a minimizer x^* on $B(w, r)$. By strong convexity we have

$$\|x^* - w\|^2 \leq \frac{2}{\gamma}\epsilon \leq \frac{r^2}{4}\,.$$

Consequently, x^* is in the interior of $B(w, r)$. By strong convexity, x^* is then the unique minimizer of f over $\mathbb{R}^N$. By strong convexity, for any x outside of $B(w, r)$ we have

$$f(x) > f(x^*) + \frac{1}{2}\gamma\left(\frac{r}{2}\right)^2 \geq f(x^*) + \epsilon\,.$$

Consequently, if $f(x) \leq \min f + \epsilon$ then $x \in B(w, r)$ and thus $\|x - x^*\|^2 \leq \frac{2}{\gamma}\epsilon$. □

Proof of Theorem 9.2.1. Let $t = \min(\frac{1}{16}\beta_{\min}, \sigma)$. By Lemma 9.6.1 the event

$$\left\{\left|\frac{\|\mathsf{w}_\lambda\|^2}{n} - \frac{\mathbb{E}\|\mathsf{w}_\lambda\|^2}{n}\right| \leq t^2, \quad \frac{g^\mathsf{T}\mathsf{w}_\lambda}{n} \leq \mathbb{E}\left[\frac{g^\mathsf{T}\mathsf{w}_\lambda}{n}\right] + t, \right. \tag{9.2.6}$$
$$\left. \|g\| \leq 2\sqrt{N}, \quad \left|\frac{\sigma g'}{\sqrt{n}}\right| \leq \frac{\beta_{\min}}{4}, \quad (1 - \frac{\beta_{\min}}{8\tau_{\max}}) \leq \frac{\|h\|}{\sqrt{n}} \leq 2\right\}$$

has probability at least $1 - Ce^{-cn}$, for some constants $C, c > 0$. On the event (9.2.6)

$$\sqrt{\frac{\|\mathsf{w}_\lambda\|^2}{n} + \sigma^2} \geq \sqrt{\frac{\mathbb{E}\|\mathsf{w}_\lambda\|^2}{n} + \sigma^2 - t^2} \geq \tau_* - t\,.$$

Therefore

$$\sqrt{\frac{\|\mathsf{w}_\lambda\|^2}{n} + \sigma^2}\frac{\|h\|}{\sqrt{n}} \geq \frac{\|h\|}{\sqrt{n}}\tau_* - \frac{\|h\|}{\sqrt{n}}t \geq \tau_* - \frac{\beta_{\min}}{8} - 2t \geq \tau_* - \frac{\beta_{\min}}{4}\,.$$

Consequently, on the event (9.2.6) we have

$$\sqrt{\frac{\|\mathsf{w}_\lambda\|^2}{n} + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T}\mathsf{w}_\lambda + \frac{\sigma g'}{\sqrt{n}} \geq \tau_* - \frac{1}{n}\mathbb{E}\left[g^\mathsf{T}\mathsf{w}_\lambda\right] - \frac{3}{4}\beta_{\min} \geq \frac{1}{4}\beta_{\min}\,,$$

because $\tau_* - \frac{1}{n}\mathbb{E}[g^\mathsf{T}\mathsf{w}_\lambda] = \beta_* \geq \beta_{\min}$. Moreover, on the event (9.2.6) the function

$$f : w \mapsto \sqrt{\frac{\|w\|^2}{n} + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T}w + \frac{g'\sigma}{\sqrt{n}}$$

is $\frac{2\sqrt{N}}{n} + \frac{2}{\sqrt{n}}$-Lipschitz. We have seen above that on (9.2.6), $f(\mathsf{w}_\lambda) \geq \frac{1}{4}\beta_{\min}$. Thus we can find a constant $r > 0$ such that on the event (9.2.6) we have for all $w \in B(\mathsf{w}_\lambda, r\sqrt{n})$

$$f(w) > \frac{1}{8}\beta_{\min}\,.$$

By Lemma 9.6.14, the function f is $\frac{a}{n}$-strongly convex on $B(\mathsf{w}_\lambda, r\sqrt{n})$, for some constant $a > 0$. For all $w \in B(\mathsf{w}_\lambda, r\sqrt{n})$ we have

$$L_\lambda(w) = \frac{1}{2}f(w)^2 + \frac{\lambda}{n}\left(|w + \theta^\star| - |\theta^\star|\right)\,.$$

Compute the Hessian for $w \in B(\mathsf{w}_\lambda, r\sqrt{n})$:

$$\nabla^2\Big(\frac{1}{2}f^2\Big)(w) = f(w)\nabla^2 f(w) + \nabla f(w)\nabla f(w)^\mathsf{T} \succeq \frac{a\beta_{\min}}{8n}\mathrm{Id}_N\,,$$

which means that L is $\frac{\gamma}{n}$-strongly convex on $B(\mathsf{w}_\lambda, r\sqrt{n})$, for some constant $\gamma > 0$.

Notice that it suffices to prove Theorem 9.2.1 for $\epsilon \in (0, q]$ for some constant $q > 0$. Let $\epsilon \in (0, \frac{\gamma r^2}{8})$. Let now apply Proposition 9.2.2 with $R = \tau_{\max} + r$: with probability at least $1 - \frac{C}{\epsilon}e^{-cn\epsilon^2}$

$$\Big\{L_\lambda(\mathsf{w}_\lambda) \leq \min_{\|w\| \leq R\sqrt{n}} L_\lambda(w) + \epsilon\Big\}\,. \tag{9.2.7}$$

Notice that on the event (9.2.6),

$$\frac{\|\mathsf{w}_\lambda\|^2}{n} \leq \mathbb{E}\left[\frac{\|\mathsf{w}_\lambda\|^2}{n}\right] + t^2 \leq \mathbb{E}\left[\frac{\|\mathsf{w}_\lambda\|^2}{n}\right] + \sigma^2 = \tau_*^2 \leq \tau_{\max}^2\,.$$

Therefore, on (9.2.6), $B(\mathsf{w}_\lambda, r\sqrt{n}) \subset B(0, R\sqrt{n})$. Using then (9.2.7) we get

$$L_\lambda(\mathsf{w}_\lambda) \leq \min_{w \in B(\mathsf{w}_\lambda, r\sqrt{n})} L_\lambda(w) + \epsilon\,.$$

Consequently, on the events (9.2.6) and (9.2.7) that have probability at least $1 - \frac{C}{\epsilon}e^{-cn\epsilon^2}$, Lemma 9.2.1 gives that for all $w \in \mathbb{R}^N$ such that $L_\lambda(w) \leq \min_{v \in \mathbb{R}^N} L_\lambda(v) + \epsilon$ we have $\|\mathsf{w}_\lambda - w\|^2 \leq \frac{8n}{\gamma}\epsilon$. □

9.3 Empirical distribution and risk of the Lasso

9.3.1 Proofs of local stability of the Lasso cost

Application of Gordon's min-max Theorem

Proposition 9.3.1

There exists constants $c, C > 0$ that only depend on Ω such that for all closed set $D \subset \mathbb{R}^N$ and for all $\epsilon \in (0, 1]$,

$$\mathbb{P}\left(\min_{w \in D} \mathcal{C}_\lambda(w) \leq \min_{w \in \mathbb{R}^N} \mathcal{C}_\lambda(w) + \epsilon\right) \leq 2\mathbb{P}\left(\min_{w \in D} L_\lambda(w) \leq \min_{w \in \mathbb{R}^N} L_\lambda(w) + 3\epsilon\right) + \frac{C}{\epsilon}e^{-cn\epsilon^2}\,.$$

In order to prove this, we start by showing that the optimal Lasso cost concentrates around $L_*(\lambda)$.

Proposition 9.3.2

There exists constants $c, C > 0$ that only depend on Ω such that for all closed set $D \subset \mathbb{R}^N$ and for all $\epsilon \in (0, 1]$,

$$\mathbb{P}\left(\left|\min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) - L_*(\lambda)\right| \geq \epsilon\right) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2}.$$

Proof. By Corollary 8.5.1, we have

$$\mathbb{P}\left(\min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) - L_*(\lambda) \geq \epsilon\right) \leq 2\mathbb{P}\left(\min_{w\in\mathbb{R}^N} L_\lambda(w) - L_*(\lambda) \geq \epsilon\right) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2},$$

where the last inequality comes from Corollary 9.2.1. The bound of the probability of the converse inequality is proved analogously. □

Proof of Proposition 9.3.1. Recall that $L_*(\lambda)$ is defined by (9.2.3). Let $D \subset \mathbb{R}^N$ be a closed set.

$$\begin{aligned}
&\mathbb{P}\left(\min_{w\in D} \mathcal{C}_\lambda(w) \leq \min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) + \epsilon\right) \\
&\leq \mathbb{P}\left(\min_{w\in D} \mathcal{C}_\lambda(w) \leq \min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) + \epsilon \quad \text{and} \quad \min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) \leq L_*(\lambda) + \epsilon\right) \\
&\quad + \mathbb{P}\left(\min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) > L_*(\lambda) + \epsilon\right) \\
&\leq \mathbb{P}\left(\min_{w\in D} \mathcal{C}_\lambda(w) \leq L_*(\lambda) + 2\epsilon\right) + \frac{C}{\epsilon} e^{-cn\epsilon^2},
\end{aligned}$$

where we used Proposition 9.3.2 above. We can now apply the first point of Corollary 8.5.1 to obtain:

$$\mathbb{P}\left(\min_{w\in D} \mathcal{C}_\lambda(w) \leq L_*(\lambda) + 2\epsilon\right) \leq 2\mathbb{P}\left(\min_{w\in D} L_\lambda(w) \leq L_*(\lambda) + 2\epsilon\right). \tag{9.3.1}$$

We thus get

$$\begin{aligned}
&\mathbb{P}\left(\min_{w\in D} \mathcal{C}_\lambda(w) \leq \min_{w\in\mathbb{R}^N} \mathcal{C}_\lambda(w) + \epsilon\right) \leq 2\mathbb{P}\left(\min_{w\in D} L_\lambda(w) \leq L_*(\lambda) + 2\epsilon\right) + \frac{C}{\epsilon} e^{-cn\epsilon^2} \\
&\leq 2\mathbb{P}\left(\min_{w\in D} L_\lambda(w) \leq \min_{w\in\mathbb{R}^N} L_\lambda(w) + 3\epsilon\right) + 2\mathbb{P}\left(\min_{w\in\mathbb{R}^N} L_\lambda(w) < L_*(\lambda) - \epsilon\right) + \frac{C}{\epsilon} e^{-cn\epsilon^2} \\
&\leq 2\mathbb{P}\left(\min_{w\in D} L_\lambda(w) \leq \min_{w\in\mathbb{R}^N} L_\lambda(w) + 3\epsilon\right) + \frac{C}{\epsilon} e^{-cn\epsilon^2},
\end{aligned}$$

for some constants $c, C > 0$, because of Corollary 9.2.1. □

Local stability of the empirical distribution of the Lasso estimator: proof of Theorem 8.5.3

For $w \in \mathbb{R}^N$, let $\mu_0(w)$ be the probability distribution over $\mathbb{R}^2$ defined by

$$\widehat{\mu}_0(w) = \frac{1}{N} \sum_{i=1}^{N} \delta_{(w_i+\theta_i^\star, \theta_i^\star)}$$

Theorem 8.5.3 follows from Proposition 9.3.1 and the following Lemma.

Lemma 9.3.1

Assume that $\mathcal{D} = \mathcal{F}_p(\xi)$ for some $\xi, p > 0$. There exists constants $\gamma, c, C > 0$ that depend only on Ω, such that for all $\epsilon \in (0, \frac{1}{2}]$ we have

$$\mathbb{P}\left(\min_{w \in D_\epsilon} L_\lambda(w) \leq \min_{w \in \mathbb{R}^N} L_\lambda(w) + 3\gamma\epsilon\right) \leq C\epsilon^{-\max(1,a)} \exp\left(-cN\epsilon^2\epsilon^a \log(\epsilon)^{-2}\right),$$

where $D_\epsilon = \left\{w \in \mathbb{R}^N \,\middle|\, W_2(\widehat{\mu}_0(w), \mu_\lambda^*)^2 \geq \epsilon\right\}$ and $a = \frac{1}{2} + \frac{1}{p}$.

Proof. By Theorem 9.2.1 and Proposition 9.6.2 there exists constants $\gamma, c, C > 0$ such that for all $\epsilon \in (0, \frac{1}{2}]$ the event

$$\left\{\forall w \in \mathbb{R}^N,\ L_\lambda(w) \leq \min_{v \in \mathbb{R}^N} L_\lambda(v) + 3\gamma\epsilon \Rightarrow \frac{1}{N}\|w - \mathsf{w}_\lambda\|^2 \leq \frac{\epsilon}{5}\right\} \bigcap \left\{W_2(\mu_\lambda^*, \widehat{\mu}_0(\mathsf{w}_\lambda))^2 \leq \frac{\epsilon}{4}\right\} \quad (9.3.2)$$

has probability at least

$$1 - C\epsilon^{-1}e^{-cn\epsilon^2} - C\epsilon^{-a}\exp\left(-cN\epsilon^2\epsilon^a\log(\epsilon)^{-2}\right) \geq 1 - C\epsilon^{-\max(1,a)}\exp\left(-cN\epsilon^2\epsilon^a\log(\epsilon)^{-2}\right),$$

where $a = \frac{1}{2} + \frac{1}{p}$. On the event (9.3.2), we have for all $w \in D_\epsilon$:

$$\frac{1}{N}\|w - \mathsf{w}_\lambda\|^2 \geq W_2(\widehat{\mu}_0(w), \widehat{\mu}_0(\mathsf{w}_\lambda))^2 \geq \left(W_2(\widehat{\mu}_0(w), \mu_\lambda^*) - W_2(\mu_\lambda^*, \widehat{\mu}_0(\mathsf{w}_\lambda))\right)^2 \geq \frac{\epsilon}{4}.$$

This gives that on the event (9.3.2), for all $w \in D_\epsilon$, $L_\lambda(w) > \min_{v \in \mathbb{R}^N} L_\lambda(v) + 3\gamma\epsilon$. The intersection of (9.3.2) with the event $\left\{\min_{w \in D_\epsilon} L_\lambda(w) \leq \min_{w \in \mathbb{R}^N} L_\lambda(w) + 3\gamma\epsilon\right\}$ is therefore empty: the lemma is proved. □

Local stability of the risk of the Lasso estimator

We prove here the analog of Theorem 8.5.3 for the risk of the Lasso estimator.

Theorem 9.3.1

for the risk of the Lasso estimator. There exists constants $C, c, \gamma > 0$ that only depend on Ω such that for all $\epsilon \in (0, 1]$

$$\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\exists \theta \in \mathbb{R}^N,\ \Big(\frac{1}{N}\|\theta - \theta^\star\|^2 - R_*(\lambda)\Big)^2 \geq \epsilon \quad \text{and} \quad \mathcal{L}_\lambda(\theta) \leq \min \mathcal{L}_\lambda + \gamma\epsilon\Big) \leq \frac{C}{\epsilon}e^{-cN\epsilon^2}.$$

Theorem 9.3.1 follows from Proposition 9.3.1 and the following Lemma.

Lemma 9.3.2

There exists constants $\gamma, c, C > 0$ that only depend on Ω such that for all $\epsilon \in (0, 1]$ we have

$$\mathbb{P}\left(\min_{w \in D_\epsilon} L_\lambda(w) \leq \min_{w \in \mathbb{R}^N} L_\lambda(w) + 3\gamma\epsilon\right) \leq \frac{C}{\epsilon}e^{-cn\epsilon^2},$$

where $D_\epsilon = \left\{w \in \mathbb{R}^N \,\middle|\, \left(\|w\| - \sqrt{NR_*(\lambda)}\right)^2 \geq N\epsilon\right\}$.

Proof. By Theorem 9.2.1 and Lemma 9.6.1 there exists constants $\gamma, c, C > 0$ such that for all $\epsilon \in (0,1]$ the event

$$\Big\{\forall w \in \mathbb{R}^N,\ L_\lambda(w) \le \min_{v\in\mathbb{R}^N} L_\lambda(v) + 3\gamma\epsilon \Rightarrow \frac{1}{N}\|w - \mathsf{w}_\lambda\|^2 \le \frac{\epsilon}{5}\Big\} \bigcap \Big\{ \big(\|\mathsf{w}_\lambda\| - \sqrt{NR_*(\lambda)}\big)^2 \le N\frac{\epsilon}{4}\Big\} \tag{9.3.3}$$

has probability at least $1 - \frac{C}{\epsilon}e^{-cn\epsilon^2}$. On the event (9.3.3), we have for all $w \in D_\epsilon$:

$$\frac{1}{N}\|w - \mathsf{w}_\lambda\|^2 \ge \frac{1}{N}\big(\|w\| - \|\mathsf{w}_\lambda\|\big)^2 \ge \frac{1}{N}\big(\sqrt{N\epsilon} - \frac{1}{2}\sqrt{N\epsilon}\big)^2 \ge \frac{\epsilon}{4}\,.$$

This gives that on the event (9.3.3), for all $w \in D_\epsilon$, $L_\lambda(w) > \min_{v\in\mathbb{R}^N} L_\lambda(v) + 3\gamma\epsilon$. The intersection of (9.3.3) with the event $\big\{ \min_{w\in D_\epsilon} L_\lambda(w) \le \min_{w\in\mathbb{R}^N} L_\lambda(w) + 3\gamma\epsilon\big\}$ is therefore empty: the lemma is proved. □

9.3.2 Uniform control over λ: proofs of Theorems 8.3.1 and 8.3.2-(8.3.8)

Control of the ℓ_1-norm of the Lasso estimator

Proposition 9.3.3

Let $\xi > 0, p > 0$. Define $K = 2\xi + \frac{2\delta\sigma^2}{\lambda_{\min}}$. Then

$$\forall \theta^\star \in \mathcal{F}_p(\xi), \quad \mathbb{P}\left(\forall \lambda \ge \lambda_{\min}, \quad \frac{1}{N}\big|\widehat{w}_\lambda\big| \le K N^{(1/p-1)_+}\right) \ge 1 - e^{-n/2}\,.$$

Proof. Since $\mathcal{F}_{p'}(\xi) \subset \mathcal{F}_p(\xi)$ for $p' \ge p$, it suffices to prove the Proposition for $p \in (0,1]$: we suppose now to be in that case. With probability at least $1 - e^{-n/2}$ we have $\|z\| \le 2\sqrt{n}$ and therefore $\min \mathcal{L}_\lambda \le \mathcal{L}_\lambda(\theta^\star) \le 2\sigma^2 + \frac{\lambda}{n}|\theta^\star|$ for all $\lambda \ge 0$. One has thus with probability at least $1 - e^{-n/2}$,

$$\forall \theta^\star \in \mathcal{F}_1(\xi),\ \forall \lambda > 0, \quad \frac{\lambda}{n}|\widehat{\theta}_\lambda| \le \mathcal{L}_\lambda(\widehat{\theta}_\lambda) \le 2\sigma^2 + \frac{\lambda}{n}|\theta^\star|\,,$$

which implies that $\frac{1}{N}|\widehat{\theta}_\lambda| \le \frac{2\delta\sigma^2}{\lambda} + \xi N^{1/p-1}$ since $\frac{1}{N}|\theta^\star| \le \frac{1}{N}\big(\sum_{i=1}^N |\theta_i^\star|^p\big)^{1/p} \le \xi N^{1/p-1}$. □

Proposition 9.3.4

Assume that $0 < \delta < 1$ and $\sigma > 0$. Let $s < s_{\max}(\delta)$. Then, there exists constants $c, K > 0$ such that

$$\forall \theta^\star \in \mathcal{F}_0(s), \quad \mathbb{P}\left(\forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \frac{1}{n}\Big||\widehat{w}_\lambda + \theta^\star| - |\theta^\star|\Big| \le K\right) \ge 1 - 2e^{-cn}\,.$$

Proposition 9.3.4 follows from the arguments of [208] that we reproduce below.

Lemma 9.3.3

Suppose that $\theta^\star \in \mathcal{F}_0(s)$ for some $s < s_{\max}(\delta)$. There exists constants $c, a > 0$ that only depend on (δ, s) such that with probability at least $1 - e^{-cn}$, for all $w \in \mathbb{R}^N$ such that $|w + \theta^\star| - |\theta^\star| \le 0$ we have

$$\|Xw\|^2 \ge a\|w\|^2\,.$$

Proof. Define

$$\mathcal{K} = \mathsf{D}(|\cdot|, \theta^\star) = \bigcup_{r>0} \left\{ u \in \mathbb{R}^N \;\middle|\; |\theta^\star + ru| \leq |\theta^\star| \right\},$$

the descent cone of the ℓ_1-norm at $\theta^\star$. Define

$$\nu_{\min}(X, \mathcal{K}) = \inf \left\{ \|Xx\| \,\middle|\, x \in \mathcal{K}, \|x\| = 1 \right\},$$

Let $\omega(\mathcal{K})$ be the Gaussian width of $\mathcal{K}$:

$$\omega(\mathcal{K}) = \mathbb{E}\left[\sup_{u \in \mathcal{K}, \|u\|=1} \langle g, u \rangle \right],$$

where the expectation is taken with respect to $g \sim \mathcal{N}(0, \mathrm{Id}_N)$. The following result goes back to Gordon's work, [93], [94]. It can be found in for instance [208] (Proposition 3.3).

Proposition 9.3.5

For all $t \geq 0$,

$$\mathbb{P}\left(\sqrt{n}\, \nu_{\min}(X, \mathcal{K}) \geq \sqrt{n-1} - \omega(\mathcal{K}) - t \right) \geq 1 - e^{-t^2/2}.$$

Recall that $M_s(\alpha) = s(1+\alpha^2) + 2(1-s)((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha))$ is the "critical function" studied in Section 9.1.2.

Lemma 9.3.4

For all $\alpha \geq 0$

$$\omega(\mathcal{K})^2 \leq N M_s(\alpha) = N\big(s(1+\alpha^2) + 2(1-s)((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha))\big).$$

Proof. Let $v \in \partial|\theta^\star|$. By convexity, we have for all $w \in \mathcal{K}$ we have $\langle w, v \rangle \leq |w + \theta^\star| - |\theta^\star| \leq 0$. Now for all $x \in \mathbb{R}^N$ and $\alpha \geq 0$

$$\|x - \alpha v\| = \sup_{\|w\|=1} \langle x - \alpha v, w \rangle \geq \sup_{w \in \mathcal{K},\, \|w\|=1} \left\{ \langle w, x \rangle - \langle w, \alpha v \rangle \right\} \geq \sup_{w \in \mathcal{K},\, \|w\|=1} \langle w, x \rangle. \tag{9.3.4}$$

Let S_0 denote the support of $\theta^\star$. Let $g \sim \mathcal{N}(0, \mathrm{Id}_N)$, $\alpha \geq 0$ and define

$$v_i = \begin{cases} \mathrm{sign}(\theta_i^\star) & \text{if } i \in S_0, \\ \alpha^{-1} g_i \mathbb{1}(|g_i| \leq \alpha) + \mathrm{sign}(g_i)\mathbb{1}(|g_i| > \alpha) & \text{otherwise.} \end{cases}$$

Notice that $v \in \partial|\theta^\star|$, therefore by (9.3.4):

$$\begin{aligned} \omega(\mathcal{K})^2 &\leq \mathbb{E}\left[\Big(\sup_{u \in \mathcal{K}, \|u\|=1} \langle g, u \rangle \Big)^2 \right] \leq \mathbb{E}\left[\|g - \alpha v\|^2 \right] \\ &\leq \mathbb{E}\left[\sum_{i \in S_0} (g_i - \alpha \mathrm{sign}(\theta_i^\star))^2 + \sum_{i \notin S_0} \eta(g_i, \alpha)^2 \right] \\ &\leq Ns(1+\alpha^2) + 2N(1-s)((1+\alpha^2)\Phi(-\alpha) - \alpha\phi(\alpha)). \end{aligned}$$

□

Since $s \leq s_{\max}(\delta)$, there exists (see Lemma 9.1.10) $\alpha \geq 0$ and $t \in (0,1)$ such that $M_s(\alpha) \leq \delta(1-t)^2$. Consequently $\omega(\mathcal{K}) \leq \sqrt{n}(1-t)$. Therefore, there exists some constants $a, c > 0$ that only depends on s and δ such that

$$\mathbb{P}\left(\nu_{\min}(X, \mathcal{K}) \geq a \right) \geq 1 - e^{-cn}.$$

On the above event, for all $w \in \mathcal{K}$, $\|Xw\|^2 \geq a^2\|w\|^2$, which proves the Lemma. □

Proof of Proposition 9.3.4. Let us work on the event

$$\left\{\forall w \in \mathbb{R}^N,\ |w+\theta^\star|-|\theta^\star| \leq 0 \Rightarrow \|Xw\|^2 \geq a\|w\|^2\right\} \bigcap \left\{\|z\| \leq 2\sqrt{n}\right\}, \tag{9.3.5}$$

which has probability at least $1 - 3e^{-cn/2}$. Let $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. Notice that on the event (9.3.5) we have $\min \mathcal{C}_\lambda \leq \mathcal{C}_\lambda(0) \leq 2\sigma^2$ and therefore

$$\frac{\lambda}{n}(|\widehat{w}_\lambda + \theta^\star| - |\theta^\star|) \leq 2\sigma^2 .$$

We distinguish two cases:

Case 1: $|\widehat{w}_\lambda + \theta^\star| - |\theta^\star| \geq 0$. In that case we obtain $\frac{1}{n}||\widehat{w}_\lambda + \theta^\star| - |\theta^\star|| \leq \frac{2\sigma^2}{\lambda_{\min}}$.

Case 2: $|\widehat{w}_\lambda + \theta^\star| - |\theta^\star| \leq 0$. In that case

$$\begin{aligned} 2\sigma^2 \geq \mathcal{C}_\lambda(\widehat{w}_\lambda) \geq \frac{1}{2n}\|X\widehat{w}_\lambda - \sigma z\|^2 - \frac{\lambda}{n}|\widehat{w}_\lambda| &\geq \frac{1}{4n}\|X\widehat{w}_\lambda\|^2 - \frac{\sigma^2}{2n}\|z\|^2 - \frac{\lambda\sqrt{N}}{n}\|\widehat{w}_\lambda\| \\ &\geq \frac{a}{2n}\|\widehat{w}_\lambda\|^2 - 2\sigma^2 - \frac{\lambda}{\sqrt{\delta n}}\|\widehat{w}_\lambda\| . \end{aligned}$$

This implies that there exists a constant $C = C(s, \delta, \sigma) > 0$ such that $\frac{1}{\sqrt{n}}\|\widehat{w}_\lambda\| \leq C(1 + \lambda)$. One conclude

$$-C\delta^{-1/2}(1 + \lambda_{\max}) \leq -\frac{1}{n}|\widehat{w}_\lambda| \leq \frac{1}{n}(|\widehat{w}_\lambda + \theta^\star| - |\theta^\star|) \leq 0 .$$

□

Lipschitz continuity of the limiting risk and empirical distribution

Proposition 9.3.6

The function $\lambda \mapsto \mu^*_\lambda$ is M-Lipschitz on $[\lambda_{\min}, \lambda_{\max}]$ with respect to the Wasserstein distance W_2, for some constant $M = M(\Omega) > 0$.

Proof. Let $\lambda_1, \lambda_2 \in [\lambda_{\min}, \lambda_{\max}]$.

$$\begin{aligned} W_2(\mu^*_{\lambda_1}, \mu^*_{\lambda_2})^2 &\leq \mathbb{E}\left[(\eta(\Theta + \tau_*(\lambda_1)Z, \alpha_*(\lambda_1)\tau_*(\lambda_1)) - \eta(\Theta + \tau_*(\lambda_2)Z, \alpha_*(\lambda_2)\tau_*(\lambda_2)))^2\right] \\ &\leq 2\mathbb{E}\left[(\tau_*(\lambda_1)Z - \tau_*(\lambda_2)Z)^2 + (\alpha_*(\lambda_1)\tau_*(\lambda_1) - \alpha_*(\lambda_2)\tau_*(\lambda_2))^2\right] \\ &\leq 2\left(\tau_*(\lambda_1) - \tau_*(\lambda_2)\right)^2 + 2(\alpha_*(\lambda_1)\tau_*(\lambda_1) - \alpha_*(\lambda_2)\tau_*(\lambda_2))^2 \\ &\leq 2(1 + \alpha^2_{\max})\left(\tau_*(\lambda_1) - \tau_*(\lambda_2)\right)^2 + 2\tau^2_{\max}(\alpha_*(\lambda_1) - \alpha_*(\lambda_2))^2 . \end{aligned}$$

Since by Proposition 9.1.3 the functions $\lambda \mapsto \alpha_*(\lambda)$ and $\lambda \mapsto \tau_*(\lambda)$ are both M-Lipschitz on $[\lambda_{\min}, \lambda_{\max}]$, for some constant $M = M(\Omega) > 0$, we obtain:

$$W_2(\mu^*_{\lambda_1}, \mu^*_{\lambda_2})^2 \leq 2M^2(1 + \alpha^2_{\max} + \tau^2_{\max})(\lambda_1 - \lambda_2)^2 ,$$

which proves the Lemma. □

Proposition 9.3.7

The function $\lambda \mapsto R_*(\lambda) = \delta(\tau_*(\lambda)^2 - \sigma^2)$ is M-Lipschitz on $[\lambda_{\min}, \lambda_{\max}]$, for some constant $M = M(\Omega) > 0$.

Proof. This is a consequence of Proposition 9.1.3. □

Proofs of Theorems 8.3.1 and 8.3.2

Lemma 9.3.5

Assume that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $s < s_{\max}(\delta)$ and $\xi \geq 0, p > 0$. Define

$$q = \begin{cases} (1/p - 1)_+ & \text{if} \quad \mathcal{D} = \mathcal{F}_p(\xi)\,, \\ 0 & \text{if} \quad \mathcal{D} = \mathcal{F}_0(s)\,. \end{cases}$$

Then there exists constants $K, C, c > 0$ that depend only on Ω such that for all $\theta^\star \in \mathcal{D}$

$$\mathbb{P}\Big(\forall \lambda, \lambda' \in [\lambda_{\min}, \lambda_{\max}],\ \mathcal{L}_{\lambda'}(\widehat{\theta}_\lambda) \leq \min_{x \in \mathbb{R}^N} \mathcal{L}_{\lambda'}(x) + KN^q|\lambda - \lambda'|\Big) \geq 1 - Ce^{-cn}. \tag{9.3.6}$$

Proof. $K = K(\Omega) > 0$ be a constant such that for all $\theta^\star \in \mathcal{D}$, the event

$$\Big\{\forall \lambda \in [\lambda_{\min}, \lambda_{\max}],\ \frac{1}{n}\big||\widehat{\theta}_\lambda| - |\theta^\star|\big| \leq KN^q\Big\} \tag{9.3.7}$$

has probability at least $1 - Ce^{-cn}$. Such K exists by Propositions 9.3.3 and 9.3.4. On the event (9.3.7) we have for all $\lambda, \lambda' \in [\lambda_{\min}, \lambda_{\max}]$:

$$\begin{aligned}
\mathcal{L}_{\lambda'}(\widehat{\theta}_\lambda) &= \mathcal{L}_\lambda(\widehat{\theta}_\lambda) + (\lambda' - \lambda)\frac{1}{n}|\widehat{\theta}_\lambda| \\
&\leq \mathcal{L}_\lambda(\widehat{\theta}_{\lambda'}) + (\lambda' - \lambda)\frac{1}{n}|\widehat{\theta}_\lambda| = \mathcal{L}_{\lambda'}(\widehat{\theta}_{\lambda'}) + \frac{1}{n}(\lambda - \lambda')(|\widehat{\theta}_{\lambda'}| - |\widehat{\theta}_\lambda|) \\
&\leq \min_{\theta \in \mathbb{R}^N} \mathcal{L}_{\lambda'}(\theta) + \frac{1}{n}|\lambda' - \lambda|\Big(\big||\widehat{\theta}_{\lambda'}| - |\theta^\star|\big| + \big||\widehat{\theta}_\lambda| - |\theta^\star|\big|\Big) \\
&\leq \min_{\theta \in \mathbb{R}^N} \mathcal{L}_{\lambda'}(\theta) + 2KN^q|\lambda - \lambda'|\,.
\end{aligned}$$

□

Theorem 8.3.1 and Theorem 8.3.2-(8.3.8) are proved the same way.

Proof of Theorem 8.3.1. Let $\gamma > 0$ as given by Theorem 8.5.3 and let $K = K(\Omega) > 0$ as given by Lemma 9.3.5. Let $M = M(\Omega) > 0$ such that $\lambda \mapsto \mu^*_\lambda$ is M-Lipschitz with respect to the Wasserstein distance W_2 on $[\lambda_{\min}, \lambda_{\max}]$, as given by Proposition 9.3.6.

Let $\epsilon \in (0, 1]$ and define $\epsilon' = \min\Big(\frac{\gamma\epsilon}{2KN^q}, \frac{\epsilon}{M+1}\Big)$. Let $k = \lceil(\lambda_{\max} - \lambda_{\min})/\epsilon'\rceil$. Define, for $i = 0, \dots, k$:

$$\lambda_i = \lambda_{\min} + i\epsilon'\,.$$

By Theorem 8.5.3, the event

$$\Big\{\forall i \in \{1 \dots k\},\ \forall \theta \in \mathbb{R}^N,\ \mathcal{L}_{\lambda_i}(\theta) \leq \min_{x \in \mathbb{R}^N} \mathcal{L}_{\lambda_i}(x) + \gamma\epsilon \ \Rightarrow\ W_2(\widehat{\mu}_{(\theta,\theta^\star)}, \mu^*_{\lambda_i})^2 \leq \epsilon\Big\} \tag{9.3.8}$$

has probability at least $1 - CN^q\epsilon^{-\max(1,a)-1}e^{-cN\epsilon^2\epsilon^a\log(\epsilon)^{-2}}$. Therefore, on the intersection of the event in (9.3.6) and the event (9.3.8) we have for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$

$$\mathcal{L}_{\lambda_i}(\widehat{\theta}_\lambda) \leq \min_{x \in \mathbb{R}^N} \mathcal{L}_{\lambda_i}(x) + 2KN^q|\lambda - \lambda_i| \leq \min_{x \in \mathbb{R}^N} \mathcal{L}_{\lambda_i}(x) + \gamma\epsilon\,,$$

where $1 \leq i \leq k$ is such that $\lambda \in [\lambda_{i-1}, \lambda_i]$. This implies (since we are on the event (9.3.8)) that $W_2\big(\widehat{\mu}_{(\widehat{\theta}_\lambda,\theta^\star)}, \mu^*_{\lambda_i}\big)^2 \leq \epsilon$. We conclude by

$$W_2\big(\widehat{\mu}_{(\widehat{\theta}_\lambda,\theta^\star)}, \mu^*_\lambda\big)^2 \leq 2W_2\big(\widehat{\mu}_{(\widehat{\theta}_\lambda,\theta^\star)}, \mu^*_{\lambda_i}\big)^2 + 2W_2(\mu^*_\lambda, \mu^*_{\lambda_i})^2 \leq 2\epsilon + 2M^2(\lambda - \lambda_i)^2 \leq 4\epsilon\,.$$

This proves the Theorem. □

Proof of Theorem 8.3.2-(8.3.8). Let $\gamma > 0$ as given by Theorem 9.3.1 and let $K = K(\Omega) > 0$ as given by Lemma 9.3.5. Let $M = M(\Omega) > 0$ such that $\lambda \mapsto R_*(\lambda)$ is M-Lipschitz on $[\lambda_{\min}, \lambda_{\max}]$, as given by Proposition 9.3.7.

Let $\epsilon \in (0,1]$ and define $\epsilon' = \min\left(\frac{\gamma\epsilon}{2KN^q}, \frac{\epsilon}{M+1}\right)$. Let $k = \lceil(\lambda_{\max} - \lambda_{\min})/\epsilon'\rceil$. Define, for $i = 0, \dots, k$:

$$\lambda_i = \lambda_{\min} + i\epsilon' .$$

By Theorem 9.3.1, the event

$$\left\{\forall i \in \{1 \dots k\},\ \forall \theta \in \mathbb{R}^N,\ \mathcal{L}_{\lambda_i}(\theta) \le \min_{x\in\mathbb{R}^N} \mathcal{L}_{\lambda_i}(x) + \gamma\epsilon \ \Rightarrow\ \left(\frac{1}{N}\|\theta - \theta^\star\|^2 - R_*(\lambda_i)\right)^2 \le \epsilon\right\} \tag{9.3.9}$$

has probability at least $1 - kC\epsilon^{-1}e^{-cN\epsilon^2} \ge 1 - CN^q\epsilon^{-2}e^{-cN\epsilon^2}$. Therefore, on the intersection of the event in (9.3.6) and the event (9.3.9) we have for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$

$$\mathcal{L}_{\lambda_i}(\widehat{\theta}_\lambda) \le \min_{\theta\in\mathbb{R}^N} \mathcal{L}_{\lambda_i}(\theta) + 2KN^q|\lambda - \lambda_i| \le \min_{\theta\in\mathbb{R}^N} \mathcal{L}_{\lambda_i}(\theta) + \gamma\epsilon ,$$

where $1 \le i \le k$ is such that $\lambda \in [\lambda_{i-1}, \lambda_i]$. This implies (since we are on the event (9.3.9)) that $\left(\frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2 - R_*(\lambda_i)\right)^2 \le \epsilon$. We conclude by

$$\begin{aligned}\left(\frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2 - R_*(\lambda)\right)^2 &\le 2\left(\frac{1}{N}\|\widehat{\theta}_\lambda - \theta^\star\|^2 - R_*(\lambda_i)\right)^2 + 2\left(R_*(\lambda_i) - R_*(\lambda)\right)^2 \\ &\le 2\epsilon + 2M^2(\epsilon')^2 \le 4\epsilon .\end{aligned}$$

This proves (8.3.8). □

9.4 Study of the Lasso residual: proof of (8.3.9)-(8.3.10)

This Section is devoted to the proof of (8.3.9)-(8.3.10) from Theorem 8.3.2. Let us define

$$\widehat{u}_\lambda = X\widehat{w}_\lambda - \sigma z = X\widehat{\theta}_\lambda - y .$$

$\widehat{u}_\lambda$ is the unique maximizer of

$$u \mapsto \min_{w\in\mathbb{R}^N}\left\{u^\mathsf{T}Xw - \sigma u^\mathsf{T}z - \frac{1}{2}\|u\|^2 + \lambda(|\theta^\star + w| - |\theta^\star|)\right\}.$$

In Section 9.4.2 below, we prove the following Theorem:

Theorem 9.4.1

There exists constants $c, C > 0$ such that for all $\epsilon \in (0,1]$, all $\theta^\star \in \mathcal{D}$ and all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$

$$\mathbb{P}\left(\left(\frac{1}{n}\|\widehat{u}_\lambda\|^2 - \beta_*(\lambda)^2\right)^2 \ge \epsilon\right) \le \frac{C}{\epsilon}e^{-cn\epsilon^2} ,$$

and

$$\mathbb{P}\left(\left(\frac{1}{n}\|\widehat{u}_\lambda + \sigma z\|^2 - P_*(\lambda)\right)^2 \ge \epsilon\right) \le \frac{C}{\epsilon}e^{-cn\epsilon^2} .$$

Theorem 8.3.2-(8.3.9)-(8.3.10) will then be deduced from Theorem 9.4.1 in Section 9.4.2.

9.4.1 Study of Gordon's optimization problem

Recall that $g \sim \mathcal{N}(0, I_N)$ and $h \sim \mathcal{N}(0, I_n)$ are independent standard Gaussian vectors. Define for $(w, u) \in \mathbb{R}^N \times \mathbb{R}^n$,

$$m_\lambda(w,u) = -\frac{1}{n^{3/2}}\|u\| g^\mathsf{T} w + \frac{1}{n^{3/2}}\|w\| h^\mathsf{T} u - \frac{\sigma}{n} u^\mathsf{T} z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n}\Big(|w+\theta^\star| - |\theta^\star|\Big)\,,$$

and $U_\lambda(u) = \min_{w\in\mathbb{R}^N} m_\lambda(w,u)$, $\widetilde{U}_\lambda(u) = m_\lambda(\mathsf{w}_\lambda, u)$, where w_λ is defined by (9.2.2). Obviously we have $U_\lambda(u) \le \widetilde{U}_\lambda(u)$. We write also

$$\mathsf{u}_\lambda = \frac{\beta_*(\lambda)}{\tau_*(\lambda)}\left(\sqrt{\tau_*(\lambda)^2 - \sigma^2}\,\frac{h}{\sqrt{n}} - \frac{\sigma}{\sqrt{n}} z\right). \tag{9.4.1}$$

Lemma 9.4.1

There exists constants $C, c > 0$ such that for all $\epsilon \in (0,1]$ and any $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ we have with probability at least $1 - Ce^{-cn\epsilon^2}$

- *$\widetilde{U}_\lambda$ is $1/n$-strongly concave on $\mathbb{R}^n$ and admits therefore a unique maximizer u^*_λ over $\mathbb{R}^n$.*
- *$|\max_{u\in\mathbb{R}^n} \widetilde{U}_\lambda(u) - L_*(\lambda)| \le \epsilon$, where $L_*(\lambda)$ is defined by Corollary 9.2.1.*
- *$\frac{1}{n}\|u^*_\lambda - \mathsf{u}_\lambda\|^2 \le \epsilon$.*

Proof. By Lemma 9.6.1 and Lemma 9.6.2, $\frac{1}{n}\mathsf{w}_\lambda^\mathsf{T} g$ concentrates around $\frac{s_*(\lambda)}{\delta}$ which is greater than some constant $\gamma > 0$. Indeed

$$s_*(\lambda) = \mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*} - \alpha_*\Big) + \Phi\Big(-\frac{\Theta}{\tau_*} - \alpha_*\Big)\right]$$

remains greater than some strictly positive constant while $\theta^\star$ vary in $\mathcal{D}$ and λ vary in $[\lambda_{\min}, \lambda_{\max}]$. By Lemma 9.6.1 we have then that with probability at least $1 - Ce^{-cn}$, $\mathsf{w}_\lambda^\mathsf{T} g \ge 0$ which implies that $\widetilde{U}_\lambda$ is $1/n$-strongly concave. Let us compute

$$\begin{aligned}\max_{u\in\mathbb{R}^n} \widetilde{U}_\lambda(u) &= \max_{\beta\ge 0}\left\{\Big(\Big\|\frac{1}{n}\|\mathsf{w}_\lambda\| h - \frac{\sigma}{\sqrt{n}} z\Big\| - \frac{1}{n} g^\mathsf{T}\mathsf{w}_\lambda\Big)\beta - \frac{1}{2}\beta^2 + \frac{\lambda}{n}\big(|\mathsf{w}_\lambda + \theta^\star| - |\theta^\star|\big)\right\} \\ &= \frac{1}{2}\left(\Big\|\frac{1}{n}\|\mathsf{w}_\lambda\| h - \frac{\sigma}{\sqrt{n}} z\Big\| - \frac{1}{n} g^\mathsf{T}\mathsf{w}_\lambda\right)_+^2 + \frac{\lambda}{n}\big(|\mathsf{w}_\lambda + \theta^\star| - |\theta^\star|\big)\,.\end{aligned}$$

By the concentration properties of w_λ (see Section 9.6.2), have that with probability at least $1 - Ce^{-cn\epsilon^2}$, $|\max_{u\in\mathbb{R}^n} \widetilde{U}_\lambda(u) - L_*(\lambda)| \le \epsilon$. One verify analogously that $\widetilde{U}_\lambda(\mathsf{u}_\lambda) \ge L_*(\lambda) - \epsilon$ with the same probability, which implies the third point by strong concavity. □

9.4.2 Proof of Theorem 9.4.1

Let us only prove the second point since the first one follows from the same arguments. Let $\epsilon \in (0,1]$ and define

$$D_\epsilon = \left\{u \in \mathbb{R}^n \,\middle|\, \Big|\frac{1}{\sqrt{n}}\|u + \sigma z\| - \sqrt{P_*(\lambda)}\Big| \ge 6\epsilon^{1/2}\right\}.$$

Let us define for $(w,u) \in \mathbb{R}^N \times \mathbb{R}^n$:

$$c_\lambda(w,u) = \frac{1}{n} u^\mathsf{T} X w - \frac{\sigma}{n} u^\mathsf{T} z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n}\Big(|w+\theta^\star| - |\theta^\star|\Big) \tag{9.4.2}$$

Lemma 9.4.2

We have almost surely

$$\min_{w\in\mathbb{R}^N}\max_{u\in\mathbb{R}^n} c_\lambda(w,u) = \max_{u\in\mathbb{R}^n}\min_{w\in\mathbb{R}^N} c_\lambda(w,u) = c_\lambda(\widehat{w}_\lambda,\widehat{u}_\lambda)\,.$$

Proof. By definition of $\widehat{w}_\lambda$ and $\widehat{u}_\lambda$ we have

$$c_\lambda(\widehat{w}_\lambda,\widehat{u}_\lambda) = \min_{w\in\mathbb{R}^N}\max_{u\in\mathbb{R}^n} c_\lambda(w,u) \geq \max_{u\in\mathbb{R}^n}\min_{w\in\mathbb{R}^N} c_\lambda(w,u)\,.$$

Let us prove the converse inequality. The optimality condition of $\widehat{w}_\lambda$ gives that there exists $v\in\partial|\theta^\star+\widehat{w}_\lambda|$ such that

$$X^\mathsf{T}\widehat{u}_\lambda + \lambda v = X^\mathsf{T}(X\widehat{w}_\lambda - \sigma z) + \lambda v = 0\,.$$

The function $w\mapsto c_\lambda(w,\widehat{u}_\lambda)$ is convex and

$$\frac{1}{n}X^\mathsf{T}\widehat{u}_\lambda + \frac{\lambda}{n}v = 0$$

is a subgradient at $\widehat{w}_\lambda$. Therefore $\min_{w\in\mathbb{R}^N} c_\lambda(w,\widehat{u}_\lambda) = c_\lambda(\widehat{w}_\lambda,\widehat{u}_\lambda)$, which proves the lemma. □

We compute now

$$\begin{aligned}\mathbb{P}\Big(\widehat{u}_\lambda\in D_\epsilon\Big) &= \mathbb{P}\Big(\max_{u\in D_\epsilon}\min_{w\in\mathbb{R}^N} c_\lambda(w,u) \geq \max_{u\in\mathbb{R}^n}\min_{w\in\mathbb{R}^N} c_\lambda(w,u)\Big)\\ &\leq \mathbb{P}\Big(\max_{u\in D_\epsilon}\min_{w\in\mathbb{R}^N} c_\lambda(w,u) \geq L_*(\lambda)-\epsilon\Big) + \mathbb{P}\Big(\max_{u\in\mathbb{R}^n}\min_{w\in\mathbb{R}^N} c_\lambda(w,u) \leq L_*(\lambda)-\epsilon\Big)\,.\end{aligned}$$

By Lemma 9.4.2 and Proposition 9.3.2 we can bound

$$\mathbb{P}\Big(\max_{u\in\mathbb{R}^n}\min_{w\in\mathbb{R}^N} c_\lambda(w,u) \leq L_*(\lambda)-\epsilon\Big) = \mathbb{P}\Big(\min_{w\in\mathbb{R}^N}\mathcal{C}_\lambda(w) \leq L_*(\lambda)-\epsilon\Big) \leq \frac{C}{\epsilon}e^{-cn\epsilon^2}\,.$$

Now by the same reasoning than Corollary 8.5.1 (we omit here the details for the sake of brevity) we have

$$\begin{aligned}\mathbb{P}\Big(\max_{u\in D_\epsilon}\min_{w\in\mathbb{R}^N} c_\lambda(w,u) \geq L_*(\lambda)-\epsilon\Big) &\leq 2\mathbb{P}\Big(\max_{u\in D_\epsilon}\min_{w\in\mathbb{R}^N} m_\lambda(w,u) \geq L_*(\lambda)-\epsilon\Big)\\ &= 2\mathbb{P}\Big(\max_{u\in D_\epsilon} U_\lambda(u) \geq L_*(\lambda)-\epsilon\Big)\,.\end{aligned}$$

Since $U_\lambda\leq\widetilde{U}_\lambda$ we obtain

$$\mathbb{P}\Big(\max_{u\in D_\epsilon}\min_{w\in\mathbb{R}^N} c_\lambda(w,u) \geq L_*(\lambda)-\epsilon\Big) \leq 2\mathbb{P}\Big(\max_{u\in D_\epsilon}\widetilde{U}_\lambda(u) \geq L_*(\lambda)-\epsilon\Big)\,.$$

Let E be the event of Lemma 9.4.1 above and let us work on the event

$$E\bigcap\Big\{\Big|\frac{1}{\sqrt{n}}\|\mathsf{u}_\lambda+\sigma z\| - \sqrt{P_*(\lambda)}\Big| \leq \epsilon^{1/2}\Big\}\,, \tag{9.4.3}$$

which has probability at least $1-\frac{C}{\epsilon}e^{-cn\epsilon^2}$ (the fact that the second event in the intersection has this probability follows from standard concentration arguments as in Section 9.6.2). Let now $u\in D_\epsilon$, by the definition of D_ϵ and the event above we have $\frac{1}{\sqrt{n}}\|u-\mathsf{u}_\lambda\| \geq 5\epsilon^{1/2}$ and thus $\frac{1}{\sqrt{n}}\|u-u^*_\lambda\| \geq 4\epsilon^{1/2}$. By $1/n$-strong concavity of $\widetilde{U}_\lambda$ we get

$$\widetilde{U}_\lambda(u) \leq \max_{u'\in\mathbb{R}^n}\widetilde{U}_\lambda(u') - 8\epsilon \leq L_*(\lambda) - 7\epsilon\,.$$

Consequently $\mathbb{P}\Big(\max_{u\in D_\epsilon}\widetilde{U}_\lambda(u) \geq L_*(\lambda)-\epsilon\Big) \leq \frac{C}{\epsilon}e^{-cn\epsilon^2}$, which proves the result.

9.4.3 Uniform control over λ: proof of Theorem 8.3.2-(8.3.9)-(8.3.10)

Let $\mathcal{D}$ be either $\mathcal{F}_0(s)$ for some $s < s_{\max}(\delta)$ or $\mathcal{F}_p(\xi)$ for some $\xi \geq 0$, $p > 0$. Let $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$. By Propositions 9.3.3 and 9.3.4 there exists a constant $K = K(\Omega)$ such that the event

$$\left\{\forall \lambda \in [\lambda_{\min}, \lambda_{\max}],\ \frac{1}{n}\Big||\widehat{w}_\lambda + \theta^\star| - |\theta^\star|\Big| \leq KN^q\right\} \tag{9.4.4}$$

has probability at least $1 - Ce^{-cn}$. Let us fix this constant K and let us write

$$D_K = \left\{w \in \mathbb{R}^N \,\middle|\, \frac{1}{n}\Big||w + \theta^\star| - |\theta^\star|\Big| \leq KN^q\right\}.$$

We define also

$$\mathcal{U}_\lambda(u) = \min_{w \in D_K}\left\{\frac{1}{n}u^\mathsf{T}Xw - \frac{\sigma}{n}u^\mathsf{T}z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n}\Big(|w + \theta^\star| - |\theta^\star|\Big)\right\}.$$

Lemma 9.4.3

The function $\mathcal{U}_\lambda$ is $1/n$-strongly concave. On the event (9.4.4), $\widehat{u}_\lambda$ is the (unique) maximizer of $\mathcal{U}_\lambda$.

Proof. Let us work on the event (9.4.4) and let $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. We have, by permutation of max and min:

$$\max_{u \in \mathbb{R}^n} \mathcal{U}_\lambda(u) \leq \min_{w \in D_K} \mathcal{C}_\lambda(w) = \mathcal{C}_\lambda(\widehat{w}_\lambda),$$

because on the event (9.4.4), $\widehat{w}_\lambda$ (the minimizer of $\mathcal{C}_\lambda$) is in D_K. By the optimality condition of $\widehat{w}_\lambda$, one verify easily that $\mathcal{U}_\lambda(\widehat{u}_\lambda) = \mathcal{C}_\lambda(\widehat{w}_\lambda)$ which proves the lemma. □

Theorem 8.3.2-(8.3.9)-(8.3.10) follow then easily from Theorem 9.4.1 (by an ϵ-net argument as in the proof of Theorems 8.3.1 and 8.3.2-(8.3.8), see Section 9.3.2) and the following Proposition:

Proposition 9.4.1

Let $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$. There exists constants $C, c, \kappa > 0$ such that for all $\theta^\star \in \mathcal{D}$ the following event

$$\left\{\forall \lambda, \lambda' \in [\lambda_{\min}, \lambda_{\max}],\ \frac{1}{n}\|\widehat{u}_\lambda - \widehat{u}_{\lambda'}\|^2 \leq \kappa N^q|\lambda - \lambda'|\right\}$$

has probability at least $1 - Ce^{-cn}$.

Proof. Let us work on the event (9.4.4), which has probability at least $1 - Ce^{-cn}$. Let $\lambda, \lambda' \in [\lambda_{\min}, \lambda_{\max}]$. We have

$$\sup_{u \in \mathbb{R}^n} |\mathcal{U}_\lambda(u) - \mathcal{U}_{\lambda'}(u)| \leq \sup_{w \in D_K}\left|\frac{\lambda - \lambda'}{n}(|w + \theta^\star| - |\theta^\star|)\right| \leq KN^q|\lambda - \lambda'|.$$

Therefore

$$\begin{aligned}\mathcal{U}_{\lambda'}(\widehat{u}_\lambda) &\geq \mathcal{U}_\lambda(\widehat{u}_\lambda) - N^qK|\lambda - \lambda'| \geq \mathcal{U}_\lambda(\widehat{u}_{\lambda'}) - N^qK|\lambda - \lambda'| \\ &\geq \mathcal{U}_{\lambda'}(\widehat{u}_{\lambda'}) - 2KN^q|\lambda - \lambda'|,\end{aligned}$$

which gives that $\frac{1}{n}\|\widehat{u}_\lambda - \widehat{u}_{\lambda'}\|^2 \leq 4KN^q|\lambda - \lambda'|$ by $1/n$-strong concavity. □

9.5 Study of the subgradient $\widehat{v}_\lambda$

The goal of this section is to analyze the vector

$$\widehat{v}_\lambda = \frac{1}{\lambda} X^\mathsf{T}(y - X\widehat{\theta}_\lambda)\,,$$

which is a subgradient of the ℓ_1-norm at $\widehat{\theta}_\lambda$. Let us define

$$B_\infty(0,1) = \left\{ v \in \mathbb{R}^N \,\middle|\, \|v\|_\infty \le 1 \right\}.$$

9.5.1 Main results

Let $B = \left\{ w \in \mathbb{R}^N \,|\, |w| \le 2|\theta^\star| + 5\sigma^2 \lambda_{\min}^{-1} n + K \right\}$, where $K > 0$ is some constant (depending only on Ω) that will be fixed later in the analysis (in fact K is the constant given by Lemma 9.5.5). Notice that $\widehat{w}_\lambda \in B$, with probability at least $1 - e^{-n/2}$. Define

$$\mathcal{V}_\lambda(v) = \min_{w \in B} \left\{ \frac{1}{2n}\|Xw - \sigma z\|^2 + \frac{\lambda}{n} v^\mathsf{T}(\theta^\star + w) - \frac{\lambda}{n}|\theta^\star| \right\}.$$

Lemma 9.5.1

With probability at least $1 - e^{-n/2}$ we have for all $\lambda \ge \lambda_{\min}$

$$\min_{w \in \mathbb{R}^N} \mathcal{C}_\lambda(w) = \max_{\|v\|_\infty \le 1} \mathcal{V}_\lambda(v)$$

and $\widehat{v}_\lambda = -\lambda^{-1} X^\mathsf{T}(X\widehat{w}_\lambda - \sigma z)$ is a maximizer of $\mathcal{V}_\lambda$.

Proof. Let us work on the event $\{\|z\| \le 2\sqrt{n}\}$ which has probability at least $1 - e^{-n/2}$. On this event we have $\widehat{w}_\lambda \in B$ and therefore

$$\min_{w \in \mathbb{R}^N} \mathcal{C}_\lambda(w) = \min_{w \in B} \mathcal{C}_\lambda(w) = \max_{\|v\|_\infty \le 1} \mathcal{V}_\lambda(v)\,,$$

where the permutation of the min-max is authorized by Proposition C.2. The optimality condition of $\widehat{w}_\lambda$ gives that

$$\widehat{v}_\lambda = -\lambda^{-1} X^\mathsf{T}(X\widehat{w}_\lambda - \sigma z) \in \partial|\theta^\star + \widehat{w}_\lambda|\,.$$

Therefore $\widehat{v}_\lambda^\mathsf{T}(\widehat{w}_\lambda + \theta^\star) = |\widehat{w}_\lambda + \theta^\star|$. Using the optimality condition again we obtain

$$\begin{aligned}\mathcal{V}_\lambda(\widehat{v}_\lambda) = \min_{w \in B} \left\{ \frac{1}{2}\|Xw - \sigma z\|^2 + \lambda \widehat{v}_\lambda^\mathsf{T}(\theta^\star + w) \right\} &= \frac{1}{2}\|X\widehat{w}_\lambda - \sigma z\|^2 + \lambda \widehat{v}_\lambda^\mathsf{T}(\theta^\star + \widehat{w}_\lambda) \\ &= \frac{1}{2}\|X\widehat{w}_\lambda - \sigma z\|^2 + \lambda|\theta^\star + \widehat{w}_\lambda|\,.\end{aligned}$$

Therefore $\widehat{v}_\lambda$ achieves the optimal value. □

The empirical law of the subgradient

Let ν_λ^* be the law of the couple

$$\left(-\frac{1}{\alpha_*(\lambda)\tau_*(\lambda)} \Big(\eta\big(\Theta + \tau_*(\lambda)Z, \alpha_*(\lambda)\tau_*(\lambda)\big) - \Theta - \tau_*(\lambda)Z \Big),\ \Theta \right), \tag{9.5.1}$$

where $(\Theta, Z) \sim \widehat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$. For $v \in \mathbb{R}^N$ we define

$$\widehat{\mu}_{(v,\theta^\star)} = \frac{1}{N}\sum_{i=1}^{N} \delta_{(v_i,\theta_i^\star)} .$$

Theorem 9.5.1

Assume that $\mathcal{D} = \mathcal{F}_p(\xi)$ for some $\xi, p > 0$. There exists constants $C, c > 0$ that only depend on Ω such that for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and all $\epsilon \in (0, \frac{1}{2}]$,

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(W_2(\widehat{\mu}_{(\widehat{v}_\lambda,\theta^\star)}, \nu_\lambda^*)^2 \geq \epsilon \Big) \leq C\epsilon^{-\max(1,a)} \exp\left(-cN\epsilon^2\epsilon^a \log(\epsilon)^{-2}\right) ,$$

where $a = \frac{1}{2} + \frac{1}{p}$.

Theorem 9.5.1 is proved in Section 9.5.3.

Theorem 9.5.2

Let $\mathcal{D}$ be $\mathcal{F}_p(\xi)$ for some $\xi > 0$ and $p > 0$. For all $\epsilon \in (0, \frac{1}{2}]$,

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} W_2(\widehat{\mu}_{(\widehat{v}_\lambda,\theta^\star)}, \nu_\lambda^*)^2 \geq \epsilon \Big) \leq C\epsilon^{-\max(1,a)-1} N^{(1/p-1)_+} e^{-cN\epsilon^2\epsilon^a \log(\epsilon)^{-2}},$$

where $a = \frac{1}{2} + \frac{1}{p}$.

Theorem 9.5.2 is deduced from Theorem 9.5.1 in Section 9.5.3.

The norm of the subgradient

Let us define

$$\kappa_*(\lambda) = \frac{\beta_*(\lambda)^2}{\lambda^2}\left(1 + \delta - 2s_*(\lambda) - \delta\frac{\sigma^2}{\tau_*(\lambda)^2}\right). \tag{9.5.2}$$

Theorem 9.5.3

There exists a constant $C, c > 0$ such that for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and all $\epsilon \in (0, 1]$,

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\left(\left(\frac{1}{N}\|\widehat{v}_\lambda\|^2 - \kappa_*(\lambda) \right)^2 \geq \epsilon \right) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2} .$$

Theorem 9.5.3 is proved in Section 9.5.3. We deduce as before:

Theorem 9.5.4

Let $\mathcal{D}$ be either $\mathcal{F}_0(s)$ for $s < s_{\max}(\delta)$ or $\mathcal{F}_p(\xi)$ for some $\xi > 0$ and $p > 0$. There exists constants $C, c > 0$ such that for all $\epsilon \in (0, 1]$,

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \left(\frac{1}{N}\|\widehat{v}_\lambda\|^2 - \kappa_*(\lambda) \right)^2 \geq \epsilon \Big) \leq \frac{C}{\epsilon^2} N^q e^{-cn\epsilon^2} ,$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

Theorem 9.5.4 is deduced from Theorem 9.5.3 in Section 9.5.3.

Upper bound on the sparsity of the Lasso estimator

Studying $\widehat{v}_\lambda$ allows to get an upper-bound on the ℓ_0 norm of $\widehat{\theta}_\lambda$. Indeed if $\widehat{\theta}_{\lambda,i} \neq 0$ then $|\widehat{v}_{\lambda,i}| = 1$: therefore $\|\widehat{\theta}_\lambda\|_0 \leq \#\Big\{i \,\Big|\, |\widehat{v}_{\lambda,i}| = 1\Big\}$. For this reason, the following results will be used to prove Theorem 9.6.1 in Section 9.6.4.

Theorem 9.5.5

There exists constants $C, c > 0$ such that for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and all $\epsilon \in (0, 1]$,

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\frac{1}{N}\#\Big\{i \,\Big|\, |\widehat{v}_{\lambda,i}| \geq 1 - \epsilon\Big\} \geq s_*(\lambda) + 2(1 + \alpha_{\max})\epsilon\Big) \leq \frac{C}{\epsilon^3} e^{-cn\epsilon^6} .$$

Theorem 9.5.5 is proved in Section 9.5.3.

Theorem 9.5.6

Let $\mathcal{D}$ be either $\mathcal{F}_0(s)$ for $s < s_{\max}(\delta)$ or $\mathcal{F}_p(\xi)$ for some $\xi > 0$ and $p > 0$. We have for all $\epsilon \in (0, 1]$,

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\exists \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \frac{1}{N}\#\Big\{i \,\Big|\, |\widehat{v}_{\lambda,i}| = 1\Big\} \geq s_*(\lambda) + \epsilon\Big) \leq \frac{C}{\epsilon^6} N^q e^{-cn\epsilon^6} ,$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

Theorem 9.5.6 is deduced from Theorem 9.5.5 in Section 9.5.3.

9.5.2 Gordon's strategy for the subgradient

Application of Gordon's Theorem

Let $g \sim \mathcal{N}(0, \mathrm{Id}_N)$ and $h \sim \mathcal{N}(0, \mathrm{Id}_n)$ be independent standard Gaussian vectors. We define:

$$V_\lambda(v) = \min_{w \in B} \left\{ \frac{1}{2} \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2} \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right)_+^2 + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star| \right\} .$$

The following Proposition is the analog of Corollary 8.5.1.

Proposition 9.5.1

Let $D \subset \left\{ v \in \mathbb{R}^N \,\middle|\, \|v\|_\infty \leq 1 \right\}$ be a closed set.

- *We have for all $t \in \mathbb{R}$*

$$\mathbb{P}\Big(\max_{v \in D} \mathcal{V}_\lambda(v) \geq t \Big) \leq 2\mathbb{P}\Big(\max_{v \in D} V_\lambda(v) \geq t \Big) .$$

- *If D is convex, then we have for all $t \in \mathbb{R}$*

$$\mathbb{P}\Big(\max_{v \in D} \mathcal{V}_\lambda(v) \leq t \Big) \leq 2\mathbb{P}\Big(\max_{v \in D} V_\lambda(v) \leq t \Big) .$$

Proof. Let $v \in \mathbb{R}^N$. By Proposition C.2 one can permute the min-max and obtain:

$$\begin{aligned} \mathcal{V}_\lambda(v) &= \min_{w \in B} \max_{u \in \mathbb{R}^n} \left\{ \frac{1}{n} u^\mathsf{T} X w - \frac{\sigma}{n} u^\mathsf{T} z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n} v^\mathsf{T}(\theta^\star + w) - \frac{\lambda}{n}|\theta^\star| \right\} \\ &= \max_{u \in \mathbb{R}^n} \min_{w \in B} \left\{ \frac{1}{n} u^\mathsf{T} X w - \frac{\sigma}{n} u^\mathsf{T} z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n} v^\mathsf{T}(\theta^\star + w) - \frac{\lambda}{n}|\theta^\star| \right\} . \end{aligned}$$

Let $D \subset \left\{ v \in \mathbb{R}^N \,\middle|\, \|v\|_\infty \leq 1 \right\}$ be a closed set. We can the apply Gordon's Theorem (Corollary 9.7.1) in order to compare

$$\max_{(v,u) \in D \times \mathbb{R}^n} \min_{w \in B} \left\{ \frac{1}{n} u^\mathsf{T} X w - \frac{\sigma}{n} u^\mathsf{T} z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n} v^\mathsf{T}(\theta^\star + w) - \frac{\lambda}{n}|\theta^\star| \right\} , \tag{9.5.3}$$

with

$$\begin{aligned} &\max_{(v,u) \in D \times \mathbb{R}^n} \min_{w \in B} \left\{ \sqrt{\frac{\|w\|^2}{n} + \sigma^2} \frac{h^\mathsf{T} u}{n} - \frac{1}{n^{3/2}}\|u\| g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n} v^\mathsf{T}(\theta^\star + w) - \frac{\lambda}{n}|\theta^\star| \right\} \qquad (9.5.4) \\ &= \max_{v \in D} \min_{w \in B} \max_{u \in \mathbb{R}^n} \left\{ \sqrt{\frac{\|w\|^2}{n} + \sigma^2} \frac{h^\mathsf{T} u}{n} - \frac{1}{n^{3/2}}\|u\| g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n} v^\mathsf{T}(\theta^\star + w) - \frac{\lambda}{n}|\theta^\star| \right\} , \end{aligned}$$

which is equal to $\max_{v \in D} V_\lambda(v)$. Note that the maximums in (9.5.3) and (9.5.4) are not defined on compact sets (since $D \times \mathbb{R}^n$ is not bounded). One has therefore to follow the same procedure than for Corollary 8.5.1, and show that there exists a compact set $\mathcal{K} \subset \mathbb{R}^n$ such that with high probability, the maximum over $u \in \mathbb{R}^n$ is achieved in $\mathcal{K}$. For the sake of brevity we do not provide a complete execution of this argument and refer to the proof of Corollary 8.5.1. □

Study of Gordon's optimization problem

In this section we study the optimization problem $\max_{\|v\|_\infty \leq 1} V_\lambda(v)$. Let us define

$$\mathsf{v}_\lambda = -\alpha_*(\lambda)^{-1}\tau_*(\lambda)^{-1}\Big(\eta\Big(\theta^\star + \tau_*(\lambda) g, \alpha_*(\lambda)\tau_*(\lambda)\Big) - \theta^\star - \tau_*(\lambda) g \Big) .$$

The goal of this section is to prove:

Theorem 9.5.7

There exists constants $\gamma, c, C > 0$ that only depend on Ω such that for all $\theta^\star \in \mathcal{D}$, all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and all $\epsilon \in (0, 1]$

$$\mathbb{P}\Big(\exists v \in B_\infty(0,1), \quad \frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 \geq \epsilon \quad \text{and} \quad V_\lambda(v) \geq \max_{v \in \mathbb{R}^N} V_\lambda(v) - \gamma\epsilon \Big) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2} .$$

Recall that w_λ^* is by Lemma 9.2.1 the unique minimizer of L_λ over $\mathbb{R}^N$.

Lemma 9.5.2

With probability at least $1-2e^{-n/2}$ we have

$$\min_{w\in\mathbb{R}^N} L_\lambda(w) = \max_{\|v\|_\infty\le 1} V_\lambda(v)$$

and the vector

$$v_\lambda^* = -\lambda^{-1}\left(\sqrt{\frac{\|w_\lambda^*\|^2}{n}+\sigma^2}\,\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T} w_\lambda^* + \frac{g'\sigma}{\sqrt{n}}\right)_+ \left(\frac{\|h\|}{\sqrt{n}}\frac{w_\lambda^*}{\sqrt{\|w_\lambda^*\|^2/n+\sigma^2}} - g\right) \tag{9.5.5}$$

verifies $\|v_\lambda^*\|_\infty \le 1$ and is a maximizer of V_λ.

Proof. By Proposition C.2, one can switch the min-max:

$$\begin{aligned}\max_{\|v\|_\infty\le 1} V_\lambda(v) &= \min_{w\in B}\max_{\|v\|_\infty\le 1}\left\{\frac{1}{2}\left(\sqrt{\frac{\|w\|^2}{n}+\sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}}\right)_+^2 + \frac{\lambda}{n}v^\mathsf{T}(w+\theta^\star) - \frac{\lambda}{n}|\theta^\star|\right\}\\ &= \min_{w\in B} L_\lambda(w)\,.\end{aligned}$$

Let us work on the event $\{\|h\|\le 2\sqrt{n}\}\cap\{g'\le\sqrt{n}\}$ which has probability at least $1-2e^{-n/2}$. We have $\frac{\lambda}{n}(|w_\lambda^*+\theta^\star|-|\theta^\star|)\le L_\lambda(w_\lambda^*)\le L_\lambda(0)\le 5\sigma^2$. This gives $w_\lambda^*\in B$ and thus $\max_{\|v\|_\infty\le 1}V_\lambda(v) = \min_{w\in B}L_\lambda(w) = \min_{w\in\mathbb{R}^N}L_\lambda(w)$.

The optimality condition of w_λ^* gives that

$$v_\lambda^* = \frac{-1}{\lambda}\left(\sqrt{\frac{\|w_\lambda^*\|^2}{n}+\sigma^2}\,\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T} w^* + \frac{g'\sigma}{\sqrt{n}}\right)_+ \left(\frac{\|h\|}{\sqrt{n}}\frac{w_\lambda^*}{\sqrt{\|w_\lambda^*\|^2/n+\sigma^2}} - g\right) \in \partial|\theta^\star + w_\lambda^*|\,.$$

Therefore $v_\lambda^{*\mathsf{T}}(w_\lambda^*+\theta^\star) = |w_\lambda^*+\theta^\star|$. Using the optimality condition again we obtain

$$\begin{aligned}V_\lambda(v_\lambda^*) &= \min_{w\in B}\left\{\frac{1}{2}\left(\sqrt{\frac{1}{n}\|w\|^2+\sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}}\right)_+^2 + \frac{\lambda}{n}v_\lambda^{*\mathsf{T}}(w+\theta^\star) - \frac{\lambda}{n}|\theta^\star|\right\}\\ &= \frac{1}{2}\left(\sqrt{\frac{1}{n}\|w_\lambda^*\|^2+\sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n}g^\mathsf{T} w_\lambda^* + \frac{g'\sigma}{\sqrt{n}}\right)_+^2 + \frac{\lambda}{n}|w_\lambda^*+\theta^\star| - \frac{\lambda}{n}|\theta^\star|\\ &= \min_{w\in\mathbb{R}^N} L_\lambda(w) = \max_{\|v\|_\infty\le 1} V_\lambda(v)\,.\end{aligned}$$

Therefore v_λ^* achieves the optimal value. □

Proposition 9.5.2

For all $\theta^\star\in\mathcal{D}$ and all $\lambda\in[\lambda_{\min},\lambda_{\max}]$ we have for all $\epsilon\in(0,1]$

$$\mathbb{P}\left(\frac{1}{N}\|v_\lambda^* - \mathsf{v}_\lambda\|^2 \ge \epsilon\right) \le \frac{C}{\epsilon}e^{-cn\epsilon}\,.$$

Proof. By Theorem 9.2.1 we have for all $\epsilon \in (0,1]$

$$\mathbb{P}\Big(\frac{1}{N}\|w_\lambda^* - \mathsf{w}_\lambda\|^2 \geq \epsilon\Big) \leq \frac{C}{\epsilon} e^{-cn\epsilon},$$

so we deduce the result from the expression (9.5.5) of v_λ^* and the concentration properties of w_λ (see Section 9.6.2). □

By the same arguments used for proving Lemma 9.5.2 it is not difficult to prove:

Lemma 9.5.3

The function

$$\beta \geq 0 \mapsto \min_{w \in B} \ell_\lambda(w, \beta)$$

(recall that ℓ_λ is defined by Equation 9.2.1) admits a unique maximizer b_λ^* over $\mathbb{R}_{\geq 0}$ and

$$b_\lambda^* = \left(\sqrt{\frac{\|w_\lambda^*\|^2}{n} + \sigma^2}\,\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w_\lambda^* + \frac{g'\sigma}{\sqrt{n}}\right)_+ .$$

Moreover, for all $\epsilon \in (0,1]$ we have $\mathbb{P}\Big(|b_\lambda^* - \beta_*(\lambda)| > \epsilon\Big) \leq \frac{C}{\epsilon} e^{-cn\epsilon}$.

Proof of Theorem 9.5.7

Let $v \in \mathbb{R}^N$ such that $\|v\|_\infty \leq 1$. We have by Proposition C.2

$$\begin{aligned}
V_\lambda(v) &= \min_{w \in B} \max_{\beta \geq 0} \left\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) - \frac{\beta^2}{2} + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star| \right\} \\
&= \max_{\beta \geq 0} \min_{w \in B} \left\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) - \frac{\beta^2}{2} + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star| \right\} \\
&= \max_{\beta \geq 0} \min_{0 \leq r \leq R} \left\{ \beta\sqrt{r^2 + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{r}{\sqrt{n}}\|\beta g - \lambda v\| + \beta\frac{g'\sigma}{\sqrt{n}} - \frac{\beta^2}{2} + \frac{\lambda}{n} v^\mathsf{T}\theta^\star - \frac{\lambda}{n}|\theta^\star| \right\},
\end{aligned}$$

because the minimization over the direction of w is easy to perform. Let us define for $\kappa > 0$

$$D_\kappa = \left\{ v \in B_\infty(0,1), \quad \frac{1}{N}\|v - v_\lambda^*\|^2 \leq \kappa^2 \quad \text{and} \quad V_\lambda(v) \geq \max_{\|v'\|_\infty \leq 1} V_\lambda(v') - \frac{1}{8}\kappa^2 \right\}.$$

By concavity of V_λ, D_κ is convex.

Proposition 9.5.3

There exists a constant $\kappa > 0$ such that with probability at least $1 - Ce^{-cn}$ we have $\forall v \in D_\kappa, V_\lambda(v) = \widetilde{V}_\lambda(v)$ where

$$\widetilde{V}_\lambda(v) = \min_{w \in \mathbb{R}^N} \left\{ \max_{\beta \in [\beta_* - \kappa, \beta_* + \kappa]} \left\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) - \frac{\beta^2}{2} \right\} + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star| \right\}.$$

In order to prove Proposition 9.5.3, we start with a Lemma:

Lemma 9.5.4

For all $v \in D_\kappa$, the function

$$f_v : \beta \mapsto \min_{w \in B} \left\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2} \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) - \frac{\beta^2}{2} + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star| \right\}$$

admits a unique maximizer $b_\lambda(v)$ on $[0, +\infty)$ and one has $|b_\lambda(v) - b_\lambda^*| \leq \kappa/2$.

Proof. Let $v \in D_\kappa$. f_v is 1-strongly concave so it admits a unique maximizer $b_\lambda(v)$ on $\mathbb{R}_{\geq 0}$. We have

$$f_v(b_\lambda(v)) = \max_{\beta \geq 0} f_v(\beta) = V_\lambda(v) \geq \max_{\|v'\|_\infty \leq 1} V_\lambda(v') - \frac{1}{8}\kappa^2 \,,$$

because $v \in D_\kappa$. Notice now that $f_v(b_\lambda(v)) \leq \min_{w \in B} \ell_\lambda(w, b_\lambda(v))$ because $v^\mathsf{T}(w + \theta^\star) \leq |w + \theta^\star|$. Permuting the min-max (using Proposition C.2), we have $\max_{\|v'\|_\infty \leq 1} V_\lambda(v') = \max_{\beta \geq 0} \min_{w \in B} \ell_\lambda(w, \beta)$, where we recall that ℓ_λ is defined by (9.2.1). We get

$$\min_{w \in B} \ell_\lambda(w, b_\lambda(v)) \geq \max_{\beta \geq 0} \min_{w \in B} \ell_\lambda(w, \beta) - \frac{1}{8}\kappa^2 \,.$$

The function $\beta \mapsto \min_{w \in B} \ell_\lambda(w, \beta)$ is 1-strongly concave and maximized (by Lemma 9.5.3 above) at b_λ^*, hence $|b_\lambda(v) - b_\lambda^*| \leq \kappa/2$. □

Lemma 9.5.5

There exist constants $K, \kappa > 0$ such that with probability at least $1 - Ce^{-cn}$ the following happens. For all $\beta \geq 0$, $v \in \mathbb{R}^N$ such that $|\beta - b_\lambda^*| \leq 2\kappa$ and $\|v - v_\lambda^*\| \leq \sqrt{N}\kappa$ the minimum over $\mathbb{R}^N$ of

$$w \mapsto \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2} \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) - \frac{\beta^2}{2} + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star|$$

is achieved on $B(0, \sqrt{N}K)$.

Proof. The minimization with respect to the direction of w is easy to perform: w has to be a non-negative multiple of $\beta g - \lambda v$. It remains thus to minimizes with respect to the norm of w. We have to show that under the conditions of the lemma, the minimum of

$$r \geq 0 \mapsto \beta\sqrt{r^2 + \sigma^2} - \frac{r}{\|h\|}\|\beta g - \lambda v\| \tag{9.5.6}$$

is achieved for r smaller than some constant. By Theorem 9.2.1 and Lemma 9.5.3 there exists a constant $R > 0$ (for instance $R = \tau_{\max} + 1$) such that the event

$$\left\{ \|w_\lambda^*\|^2 \leq nR^2 \right\} \bigcap \left\{ b_\lambda^* \geq \beta_{\min}/2 \right\} \bigcap \left\{ \|h\| \geq \sqrt{n}/2 \right\} \tag{9.5.7}$$

has probability at least $1 - Ce^{-cn}$. Let us define the constants $a = \frac{R}{\sqrt{R^2 + \sigma^2}} < 1$ and

$$\kappa = \min \left(\frac{\sqrt{\delta}\beta_{\min}^2(1-a)}{256\lambda_{\max}}, \quad \beta_{\min}/8, \quad \frac{(1-a)\sqrt{\delta}\beta_{\min}}{16\lambda_{\max}} \right).$$

Let us now work on the event (9.5.7). Let $v \in \mathbb{R}^N$ and $\beta \geq 0$ such that $\|v - v_\lambda^*\| \leq \sqrt{N}\kappa$ and $|\beta - b_\lambda^*| \leq 2\kappa$. We have

$$\|g - \frac{\lambda}{\beta} v\| \leq \|g - \frac{\lambda}{b_\lambda^*} v_\lambda^*\| + \|\frac{\lambda}{\beta} v - \frac{\lambda}{b_\lambda^*} v_\lambda^*\| \,.$$

Compute

$$\|g - \frac{\lambda}{b^*_\lambda} v^*_\lambda\| = \|h\| \frac{\|w^*_\lambda\|/\sqrt{n}}{\sqrt{\frac{\|w^*_\lambda\|^2}{n} + \sigma^2}} \leq \|h\| a\,,$$

with probability at least $1 - Ce^{-cn}$. Now

$$\begin{aligned}\|\frac{1}{\beta} v - \frac{1}{b^*_\lambda} v^*_\lambda\| &\leq \frac{1}{\beta^*}\|v - v^*_\lambda\| + \|\frac{1}{\beta} v - \frac{1}{b^*_\lambda} v\| \leq \frac{2}{\beta_{\min}}\|v - v^*_\lambda\| + \frac{\sqrt{N}}{\min(\beta, b^*_\lambda)^2}|\beta - b^*_\lambda| \\ &\leq \frac{2}{\beta_{\min}}\|v - v^*_\lambda\| + \frac{16\sqrt{N}}{\beta^2_{\min}}|\beta - b^*_\lambda| \leq \frac{1-a}{4}\sqrt{n}\,.\end{aligned}$$

Putting all together:

$$\frac{1}{\|h\|}\|g - \frac{\lambda}{\beta} v\| \leq a + \frac{1-a}{2} = \frac{1+a}{2} < 1\,.$$

This gives that the minimum of (9.5.6) is achieved for $r \leq \frac{\sigma}{\sqrt{1-((1+a)/2)^2}}$. One can thus chose $K = \frac{\delta\sigma}{\sqrt{1-((1+a)/2)^2}}$. □

Proof of Proposition 9.5.3. Let us now fix a constant $\kappa \in (0, \beta_{\min}/2)$ that verify the statement of Lemma 9.5.5. Let us work on the intersection of the event $\{|b^*_\lambda - \beta_*(\lambda)| \leq \kappa/2\}$ with the event of Lemma 9.5.5. This intersection has by Lemma 9.5.3 and Lemma 9.5.5 probability at least $1 - Ce^{-cn}$.

Let $v \in D_\kappa$. By Lemma 9.5.4 the unique maximizer $b_\lambda(v)$ of f_v verify $|b_\lambda(v) - b^*_\lambda| \leq \kappa/2$ and therefore $|b_\lambda(v) - \beta_*| \leq \kappa$. Consequently

$$\begin{aligned}V_\lambda(v) = \max_{\beta\in[\beta_*-\kappa,\beta_*+\kappa]} \min_{w\in B} \Bigg\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) \\ - \frac{\beta^2}{2} + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star| \Bigg\}\,.\end{aligned}$$

Now, for $\beta \in [\beta_* - \kappa, \beta_* + \kappa]$, we have $|\beta - b^*_\lambda| \leq 2\kappa$. Since we are working on the event of Lemma 9.5.5, we obtain

$$\begin{aligned}V_\lambda(v) = \max_{\beta\in[\beta_*-\kappa,\beta_*+\kappa]} \min_{w\in\mathbb{R}^N} \Bigg\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) \\ - \frac{\beta^2}{2} + \frac{\lambda}{n} v^\mathsf{T}(w + \theta^\star) - \frac{\lambda}{n}|\theta^\star| \Bigg\}\,,\end{aligned}$$

and Proposition 9.5.3 follows from the permutation of the min – max using Proposition C.2. □

Lemma 9.5.6

There exists a constant $C, c, \gamma > 0$ such that $\widetilde{V}_\lambda$ is γ/N-strongly concave, with probability at least $1 - Ce^{-cn}$.

Proof. The function $f^* : \mathbb{R}^N \to \mathbb{R}$ defined by

$$f^*(v) = \max_{w\in\mathbb{R}^N} \left\{ v^\mathsf{T} w - \frac{n}{\lambda} \max_{\beta\in[\beta_*-\kappa,\beta_*+\kappa]} \left\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2}\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) - \frac{\beta^2}{2} \right\} \right\}$$

is the convex conjugate of the convex function

$$f : w \in \mathbb{R}^N \mapsto \frac{n}{\lambda} \max_{\beta \in [\beta_* - \kappa, \beta_* + \kappa]} \left\{ \beta \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2} \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) - \frac{\beta^2}{2} \right\}$$
$$= \frac{n}{\lambda} \varphi \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2} \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right),$$

where φ is the $\mathcal{C}^1$ function

$$\varphi(x) = \begin{cases} \frac{1}{2}x^2 & \text{if} \quad x \in [\beta_* - \kappa, \beta_* + \kappa], \\ (\beta_* - \kappa)x - \frac{1}{2}(\beta_* - \kappa)^2 & \text{if} \quad x \le \beta_* - \kappa, \\ (\beta_* + \kappa)x - \frac{1}{2}(\beta_* + \kappa)^2 & \text{if} \quad x \ge \beta_* + \kappa. \end{cases}$$

f is a proper closed convex function (because f is convex and its domain is $\mathbb{R}^N$), therefore its convex conjugate f^* is also a proper closed convex function. The Fenchel-Moreau Theorem gives then that $f = f^{**}$. Let us compute the gradient of f for $w \in \mathbb{R}^N$

$$\nabla f(w) = \frac{n}{\lambda} \varphi' \left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2} \frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right) \left(\frac{\|h\|}{\sqrt{n}} \frac{w/n}{\sqrt{\frac{\|w\|^2}{n} + \sigma^2}} - \frac{g}{n} \right).$$

It is not difficult to verify that there exists a constant L such that ∇f is L-Lipschitz on $\mathbb{R}^N$, with probability at least $1 - Ce^{-cn}$. $f = f^{**}$ is therefore $1/L$-strongly smooth (see Definition C.1). By Proposition C.3 this gives that f^* is $1/L$ strongly convex. One deduces then that $\widetilde{V}_\lambda$ is γ/N-strongly concave with $\gamma = \lambda/(L\delta)$. □

Let $0 < \gamma < 1/2$ be a constant that verify the statement of Lemma 9.5.6 and let $\kappa > 0$ be a constant given by Proposition 9.5.3. Notice that it suffices to prove Theorem 9.5.7 for ϵ small enough and let $\epsilon \in (0, \kappa^2)$.

$$\begin{aligned} &\mathbb{P}\Big(\exists v \in B_\infty(0,1),\ \frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 > \epsilon \quad \text{and} \quad V_\lambda(v) \ge \max_{\|v'\|_\infty \le 1} V_\lambda(v') - \frac{1}{4}\gamma\epsilon\Big) \\ &\le \mathbb{P}\Big(\exists v \in B_\infty(0,1),\ \frac{1}{N}\|v - v_\lambda^*\|^2 > \frac{\epsilon}{2} \quad \text{and} \quad V_\lambda(v) \ge \max_{\|v'\|_\infty \le 1} V_\lambda(v') - \frac{1}{4}\gamma\epsilon\Big) + \frac{C}{\epsilon} e^{-cn\epsilon^2} \\ &\le \mathbb{P}\Big(\exists v \in D_\kappa,\ \frac{1}{N}\|v - v_\lambda^*\|^2 > \frac{\epsilon}{2} \quad \text{and} \quad V_\lambda(v) \ge \max_{v' \in D_\kappa} V_\lambda(v') - \frac{1}{4}\gamma\epsilon\Big) + \frac{C}{\epsilon} e^{-cn\epsilon^2}, \end{aligned} \tag{9.5.8}$$

because, if there exists $v \in B_\infty(0,1)$ such that $\frac{1}{N}\|v - v_\lambda^*\|^2 > \frac{\epsilon}{2}$ and $V_\lambda(v) \ge \max_{\|v'\|_\infty \le 1} V_\lambda(v') - \frac{1}{4}\gamma\epsilon$, we can construct $\widetilde{v} \in D_\kappa$ that verifies the same conditions. Indeed:

- if $\frac{1}{N}\|v - v_\lambda^*\|^2 \le \kappa^2$, one simply take $\widetilde{v} = v$.
- otherwise, $\widetilde{v} = v_\lambda^* + \kappa(v - v_\lambda^*)/\|v - v_\lambda^*\|$ is in D_κ and by concavity $V_\lambda(\widetilde{v}) \ge V_\lambda(v)$.

Since with probability at least $1 - Ce^{-cn}$ we have $V_\lambda(v) = \widetilde{V}_\lambda(v)$ for all $v \in D_\kappa$ and $\widetilde{V}_\lambda$ is γ/N-strongly concave, the probability in (9.5.8) above is less that Ce^{-cn}.

9.5.3 Proofs of the main results about the subgradient

Let us start with the analog of Proposition 9.3.1 for the costs functions $\mathcal{V}_\lambda$ and V_λ:

Proposition 9.5.4

There exists constants $c, C > 0$ that only depend on Ω such that for all closed set $D \subset \mathbb{R}^N$ and for all $\epsilon \in (0,1]$,

$$\mathbb{P}\left(\max_{v\in D} \mathcal{V}_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} \mathcal{V}_\lambda(v) - \epsilon\right) \leq 2\mathbb{P}\left(\max_{v\in D} V_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\epsilon\right) + \frac{C}{\epsilon} e^{-cn\epsilon^2}.$$

The proof of Proposition 9.5.4 is omitted for the sake of brevity, and because it follows from the exact same arguments than Proposition 9.3.1.

The norm of $\widehat{v}_\lambda$: proof of Theorem 9.5.3

Lemma 9.5.7

There exists constants $\gamma, c, C > 0$ that depend only on Ω, such that for all $\epsilon \in (0,1]$ we have

$$\mathbb{P}\left(\max_{v\in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\gamma\epsilon\right) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2},$$

where $D_\epsilon = \left\{ v \in B_\infty(0,1) \,\middle|\, \left(\|v\| - \sqrt{N\kappa_*(\lambda)}\right)^2 \geq \epsilon \right\}$ *and $\kappa_*(\lambda)$ is defined by (9.5.2).*

Proof. Similarly to Proposition 9.6.1 it is not difficult to prove that for all $\epsilon \in (0,1]$,

$$\mathbb{P}\Big(\Big|\frac{1}{N}\|\mathsf{v}_\lambda\|^2 - \kappa_*(\lambda)^2\Big| > \epsilon\Big) \leq Ce^{-cN\epsilon^2},$$

for some constants $c, C > 0$. By Theorem 9.5.7 there exists constants $\gamma, c, C > 0$ such that for all $\epsilon \in (0,1]$ the event

$$\left\{\forall v \in B_\infty(0,1),\ V_\lambda(v) \geq \max_{B_\infty(0,1)} V_\lambda - 3\gamma\epsilon \Rightarrow \frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 \leq \frac{\epsilon}{5}\right\} \bigcap \left\{ \left(\|\mathsf{v}_\lambda\| - \sqrt{N\kappa_*(\lambda)}\right)^2 \leq N\frac{\epsilon}{4} \right\} \tag{9.5.9}$$

has probability at least $\frac{C}{\epsilon}e^{-cn\epsilon^2}$. On the event (9.5.9), we have for all $v \in D_\epsilon$:

$$\frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 \geq \frac{1}{N}\left(\|v\| - \|\mathsf{v}_\lambda\|\right)^2 \geq \frac{1}{N}\left(\sqrt{N\epsilon} - \frac{1}{2}\sqrt{N\epsilon}\right)^2 \geq \frac{\epsilon}{4}.$$

This gives that on the event (9.5.9), for all $v \in D_\epsilon$, $V_\lambda(v) < \max_{\|v'\|_\infty \leq 1} V_\lambda(v') - 3\gamma\epsilon$. The intersection of (9.5.9) with the event $\left\{\max_{v\in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \geq 1} V_\lambda(v) - 3\gamma\epsilon\right\}$ is therefore empty: the lemma is proved. □

Proof of Theorem 9.5.3. Let $\gamma > 0$ be a constant that verify the statement of Lemma 9.5.7. Let $\epsilon \in (0,1]$ and define

$$D_\epsilon = \left\{ v \in B_\infty(0,1) \,\middle|\, \left(\|v\| - \sqrt{N\kappa_*(\lambda)}\right)^2 \geq N\epsilon \right\}.$$

D_ϵ is a closed set.

$$\begin{aligned}
&\mathbb{P}\Big(\exists v \in B_\infty(0,1),\ \Big|\frac{1}{N}\|v\|^2 - \kappa_*(\lambda)\Big| \geq \epsilon \text{ and } \mathcal{V}_\lambda(v) \geq \max \mathcal{V}_\lambda - \gamma\epsilon\Big) \\
&\qquad = \mathbb{P}\Big(\max_{v\in D_\epsilon} \mathcal{V}_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} \mathcal{V}_\lambda(v) - \gamma\epsilon\Big) \\
&\qquad \leq 2\mathbb{P}\Big(\max_{v\in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\gamma\epsilon\Big) + Ce^{-cn\epsilon} \leq \frac{C}{\epsilon}e^{-cn\epsilon^2},
\end{aligned}$$

where we used successively Proposition 9.5.4 and Lemma 9.5.7. □

The empirical law of $\widehat{v}_\lambda$: proof of Theorem 9.5.1

Theorem 9.5.1 follows now from Proposition 9.5.4 and the following Lemma.

Lemma 9.5.8

There exists constants $\gamma, c, C > 0$ that depend only on Ω, such that for all $\epsilon \in (0, \frac{1}{2}]$ we have

$$\mathbb{P}\left(\max_{v\in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\gamma\epsilon\right) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2 \log(\epsilon)^{-2}},$$

where $D_\epsilon = \left\{ v \in B_\infty(0,1) \,\middle|\, W_2(\widehat{\mu}_{(v,\theta^\star)}, \nu_\lambda^*)^2 \geq \epsilon \right\}$.

Proof. By Theorem 9.5.7 and Proposition 9.6.2 there exists constants $\gamma, c, C > 0$ such that for all $\epsilon \in (0, \frac{1}{2}]$ the event

$$\left\{ \forall v \in B_\infty(0,1),\ V_\lambda(v) \geq \max_{B_\infty(0,1)} V_\lambda - 3\gamma\epsilon \Rightarrow \frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 \leq \frac{\epsilon}{5} \right\} \bigcap \left\{ W_2\left(\nu_\lambda^*, \widehat{\mu}_{(\mathsf{v}_\lambda,\theta^\star)}\right)^2 \leq \frac{\epsilon}{4} \right\} \quad (9.5.10)$$

has probability at least

$$1 - \frac{C}{\epsilon} e^{-cn\epsilon^2} - C\epsilon^{-a} \exp\left(-cN\epsilon^2\epsilon^a \log(\epsilon)^{-2}\right) \geq 1 - C\epsilon^{-\max(1,a)} \exp\left(-cN\epsilon^2\epsilon^a \log(\epsilon)^{-2}\right).$$

On the event (9.5.10), we have for all $v \in D_\epsilon$:

$$\frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 \geq W_2(\widehat{\mu}_{(v,\theta^\star)}, \widehat{\mu}_{(\mathsf{v}_\lambda,\theta^\star)})^2 \geq \left(W_2(\widehat{\mu}_{(v,\theta^\star)}, \nu_\lambda^*) - W_2(\nu_\lambda^*, \widehat{\mu}_{(\mathsf{v}_\lambda,\theta^\star)})\right)^2 \geq \frac{\epsilon}{4}.$$

This gives that on the event (9.5.10), for all $v \in D_\epsilon$, $V_\lambda(v) < \max_{\|v'\|_\infty \leq 1} V_\lambda(v') - 3\gamma\epsilon$. The intersection of (9.5.10) with the event $\{\max_{v\in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \geq 1} V_\lambda(v) - 3\gamma\epsilon\}$ is therefore empty: the lemma is proved. □

Proof of Theorem 9.5.1. Let $\gamma > 0$ be a constant that verify the statement of Lemma 9.5.8. Let $\epsilon \in (0, \frac{1}{2}]$ and define

$$D_\epsilon = \left\{ v \in B_\infty(0,1) \,\middle|\, W_2(\widehat{\mu}_v, \nu_\lambda^*)^2 \geq \epsilon \right\}$$

D_ϵ is a closed set.

$$\begin{aligned}
&\mathbb{P}\Big(\exists v \in B_\infty(0,1),\ W_2(\widehat{\mu}_v, \nu_\lambda^*)^2 \geq \epsilon \text{ and } \mathcal{V}_\lambda(v) \geq \max \mathcal{V}_\lambda - \gamma\epsilon\Big) \\
&\quad = \mathbb{P}\Big(\max_{v\in D_\epsilon} \mathcal{V}_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} \mathcal{V}_\lambda(v) - \gamma\epsilon\Big) \\
&\quad \leq 2\mathbb{P}\Big(\max_{v\in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\gamma\epsilon\Big) + Ce^{-cn\epsilon} \leq \frac{C}{\epsilon} e^{-cn\epsilon^2 \log(\epsilon)^{-2}},
\end{aligned}$$

where we used successively Proposition 9.5.4 and Lemma 9.5.8. □

Proof of Theorem 9.5.5

Lemma 9.5.9

There exists constants $\gamma, c, C > 0$ that depend only on Ω, such that for all $\epsilon \in (0, 1]$ we have

$$\mathbb{P}\left(\max_{v\in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\gamma\epsilon^3\right) \leq \frac{C}{\epsilon^3} e^{-cn\epsilon^6},$$

where $D_\epsilon = \left\{ v \in B_\infty(0,1) \,\middle|\, \frac{1}{N}\#\left\{ i \,\middle|\, |v_i| \geq 1 - \epsilon \right\} > s_*(\lambda) + 2(1 + \alpha_{\max})\epsilon \right\}$.

Proof. Let $\epsilon \in (0,1]$ and define

$$s_\epsilon = \frac{1}{N}\#\Big\{ i \in \{1,\dots,N\} \,\Big|\, |\mathsf{v}_{\lambda,i}| \geq 1-2\epsilon \Big\} .$$

s_ϵ is the mean of independent Bernoulli random variables. By Hoeffding's inequality we have

$$\mathbb{P}\Big(s_\epsilon \leq \mathbb{P}(|\tau_*^{-1}\Theta + Z| \geq \alpha_* - 2\alpha_*\epsilon) + \epsilon \Big) \geq 1 - e^{-2N\epsilon^2} .$$

Compute

$$\begin{aligned}
&\mathbb{P}(|\tau_*^{-1}\Theta + Z| \geq \alpha_* - 2\alpha_*\epsilon) \\
&\quad = \mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\lambda)} - \alpha_*(\lambda) + 2\alpha_*(\lambda)\epsilon \Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\lambda)} - \alpha_*(\lambda) + 2\alpha_*(\lambda)\epsilon \Big) \right] \\
&\quad \leq s_*(\lambda) + 2\alpha_{\max}\epsilon .
\end{aligned}$$

We obtain $\mathbb{P}\Big(s_\epsilon \leq s_*(\lambda) + 2\alpha_{\max}\epsilon + \epsilon \Big) \geq 1 - e^{-2N\epsilon^2}$. By Theorem 9.5.7 there exists a constant $\gamma > 0$ such that the event

$$\Big\{ \forall v \in B_\infty(0,1),\ V_\lambda(v) \geq \max_{B_\infty(0,1)} V_\lambda - 3\gamma\epsilon^3 \Rightarrow \frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 < \epsilon^3 \Big\} \bigcap \Big\{ s_\epsilon \leq s_*(\lambda) + \alpha_{\max}\epsilon + \epsilon \Big\}$$

has probability at least $1 - \frac{C}{\epsilon^3} e^{-cn\epsilon^6}$. We have on this event, for all $v \in D_\epsilon$, $\frac{1}{N}\|v - \mathsf{v}_\lambda\|^2 \geq \epsilon^3$. Therefore, on the above event we have $\max_{v \in D_\epsilon} V_\lambda(v) < \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\gamma\epsilon^3$, which concludes the proof. □

Proof of Theorem 9.5.5. Let $\gamma > 0$ be a constant that verify the statement of Lemma 9.5.9. Let $\epsilon \in (0,1]$ and define

$$D_\epsilon = \left\{ v \in B_\infty(0,1) \,\middle|\, \frac{1}{N}\#\{ i \,|\, |v_i| \geq 1-\epsilon \} \geq s_*(\lambda) + 2(1+\alpha_{\max})\epsilon \right\} .$$

D_ϵ is a closed set.

$$\begin{aligned}
&\mathbb{P}\Big(\frac{1}{N}\#\{ i \,|\, |\widehat{v}_{\lambda,i}| \geq 1-\epsilon \} \geq s_*(\lambda) + 2(1+\alpha_{\max})\epsilon \Big) \\
&\qquad \leq \mathbb{P}\Big(\max_{v \in D_\epsilon} \mathcal{V}_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} \mathcal{V}_\lambda(v) - \gamma\epsilon \Big) \\
&\qquad \leq 2\mathbb{P}\Big(\max_{v \in D_\epsilon} V_\lambda(v) \geq \max_{\|v\|_\infty \leq 1} V_\lambda(v) - 3\gamma\epsilon \Big) + \frac{C}{\epsilon} e^{-cn\epsilon} \leq \frac{C}{\epsilon} e^{-cn\epsilon^2} ,
\end{aligned}$$

where we used successively Proposition 9.5.4 and Lemma 9.5.9. □

Uniform control over λ: proof of Theorems 9.5.2, 9.5.4 and 9.5.6

Theorems 9.5.2, 9.5.4 and 9.5.6 are deduced from Theorems 9.5.1, 9.5.3 and 9.5.5 by an ϵ-net argument, as we did to deduce Theorems 8.3.1 and 8.3.2 from Theorems 8.5.3 and 9.3.1. Since the ideas are the same, we only present here the key argument:

Proposition 9.5.5

Assume that $\mathcal{D}$ is $\mathcal{F}_0(s)$ or $\mathcal{F}_1(\xi)$ for some $s < s_{\max}(\delta)$ and $\xi \geq 0, p > 0$. Let $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$. Then there exists constants $K, C, c > 0$ that depend only on Ω such that for all $\theta^\star \in \mathcal{D}$

$$\mathbb{P}\Big(\forall \lambda, \lambda' \in [\lambda_{\min}, \lambda_{\max}],\ \frac{1}{N}\|\widehat{v}_\lambda - \widehat{v}_{\lambda'}\|^2 \leq K N^q |\lambda - \lambda'| \Big) \geq 1 - Ce^{-cn} . \tag{9.5.11}$$

Proof. By Proposition 9.4.1, there exists a constant K such that with probability at least $1 - Ce^{-cn}$ we have

$$\forall \lambda, \lambda' \in [\lambda_{\min}, \lambda_{\max}], \quad \frac{1}{n}\|\widehat{u}_\lambda - \widehat{u}_{\lambda'}\|^2 \le KN^q|\lambda - \lambda'|\,.$$

Notice now that $\widehat{v}_\lambda = -\frac{1}{\lambda}X^\mathsf{T}\widehat{u}_\lambda$ and that with probability at least $1-2e^{-n/4}$, $\sigma_{\max}(X) \le \delta^{-1/2}+2$ (by Proposition 9.7.4) which combined with the above inequality, prove the Proposition. □

9.6 Some auxiliary results and proofs

9.6.1 Proof of Remark 8.3.2

Let $k \le N$ and define the vector $\theta^\star = (N, 2N, \dots, kN, 0, \dots, 0)$. With the definitions given in Remark 8.3.2, we claim that $W_2(\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}, \mu^*_\lambda) \ge \sqrt{k/N}$ with probability at least $1 - e^{-ck}$, for some constant $c > 0$. Indeed, consider the case $\lambda = 0$, $\tau_* = 1$, and let $\mathbb{P}$, $\mathbb{E}$ denote probability and expectation with respect to the coupling that achieves the Wasserstein distance. This is a coupling for a triple of random variables (I, Θ, Z), with $I \sim \mathrm{Unif}\big(\{1, \dots, N\}\big)$, $(\Theta, Z) \sim \widehat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$, with

$$W_2(\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}, \mu^*_\lambda)^2 = \mathbb{E}\Big\{(\theta_I - \Theta)^2\Big\} + \mathbb{E}\Big\{(\theta_I + z_I - \Theta - Z)^2\Big\} \equiv A + B\,. \tag{9.6.1}$$

We will proceed to bound separately the two terms above. Define $\delta_i \equiv \mathbb{P}(\Theta \ne \theta^\star_i | I = i)$, and $\delta_{\max} \equiv \max_{i \le k} \delta_i$. Since $\Theta \in \{0, N, \dots, kN\}$ with probability one, we have

$$A \ge \frac{1}{N}\sum_{i=1}^k \mathbb{E}\Big\{(\theta^\star_I - \Theta)^2 \Big| I = i\Big\}\,\delta_i \ge N\sum_{i=1}^k \delta_i \ge N\delta_{\max}\,. \tag{9.6.2}$$

For the second term, we have

$$B \ge \sum_{i=1}^k \mathbb{E}\Big\{(\theta^\star_I + z_I - \Theta - Z)^2 \mathbf{1}_{I=i}\mathbf{1}_{\Theta=\theta^\star_i}\Big\} = \sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^2 \mathbf{1}_{I=i}\mathbf{1}_{\Theta=\theta^\star_i}\Big\} \tag{9.6.3}$$

$$= \sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^2 \mathbf{1}_{\Theta=\theta^\star_i}\Big\} - \sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^2 \mathbf{1}_{I\ne i}\mathbf{1}_{\Theta=\theta^\star_i}\Big\}\,. \tag{9.6.4}$$

Note that, by the coupling definition, $\mathbb{P}\Big(I = i \Big| \Theta = \theta^\star_i\Big) = N\mathbb{P}\Big(I = i\,; \Theta = \theta^\star_i\Big) = 1 - \delta_i$. Using the fact that Θ and Z are independent random variables, together with Cauchy-Schwartz inequality,we get

$$B \ge \sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^2\Big\}\,\mathbb{P}(\Theta = \theta^\star_i) - \sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^4 \mathbf{1}_{\Theta=\theta^\star_i}\Big\}^{1/2}\mathbb{P}\Big(I \ne i; \Theta = \theta^\star_i\Big)^{1/2} \tag{9.6.5}$$

$$\ge \frac{1}{N}\sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^2\Big\} - \frac{1}{N}\sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^4\Big\}^{1/2}\mathbb{P}\Big(I \ne i \Big| \Theta = \theta^\star_i\Big)^{1/2} \tag{9.6.6}$$

$$\ge \frac{1}{N}\sum_{i=1}^k \mathbb{E}\Big\{(z_i - Z)^2\Big\} - \frac{1}{N}\sum_{i=1}^k \delta_i^{1/2}\mathbb{E}\Big\{(z_i - Z)^4\Big\}^{1/2} \tag{9.6.7}$$

$$\ge \frac{1}{N}\sum_{i=1}^k (1 + z_i^2) - \frac{1}{N}\sum_{i=1}^k \delta_i^{1/2}\Big(3 + 6z_i^2 + z_i^4\Big)^{1/2}\,. \tag{9.6.8}$$

where in the last step we used the fact that $Z \sim \mathcal{N}(0,1)$. Using $(3+6x^2+x^4) \le 4(1+x^2)^2$, we thus conclude

$$B \ge \frac{1-2\delta_{\max}^{1/2}}{N} \sum_{i=1}^{k} (1+z_i^2)\,. \tag{9.6.9}$$

By concentration properties of chi-squared random variables, for any $\varepsilon > 0$, there exists $c(\varepsilon) > 0$ such that, with probability at least $1 - e^{-ck}$ we have $\frac{1}{k}\sum_{i=1}^{k} z_i^2 \ge 1 - 2\varepsilon$. Hence, with the same probability

$$W_2(\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}, \mu_\lambda^*)^2 \ge N\delta_{\max} + \frac{2k}{N}(1-2\delta_{\max}^{1/2})(1-\varepsilon) \tag{9.6.10}$$

$$\ge \frac{k}{N}\,. \tag{9.6.11}$$

The last inequality follows by lower bounding the first term for $\delta_{\max} > 1/N$, and the second for $\delta_{\max} \le 1/N$, and fixing ε a sufficiently small constant.

9.6.2 Concentration properties of w_λ

We prove in this section concentrations of the norms and some scalar product of w_λ.

Lemma 9.6.1

There exists constants $c, C > 0$ that only depend on Ω such that for all $t \ge 0$ the event

$$\left\{ \left| \frac{1}{n} g^{\mathsf{T}} \mathsf{w}_\lambda - \mathbb{E}\left[\frac{1}{n} g^{\mathsf{T}} \mathsf{w}_\lambda\right] \right| \le t\,, \quad \left| \frac{\|\mathsf{w}_\lambda\|^2}{n} - \mathbb{E}\left[\frac{\|\mathsf{w}_\lambda\|^2}{n}\right] \right| \le t \quad \text{and} \quad \left| \frac{|\mathsf{w}_\lambda + \theta^\star|}{n} - \mathbb{E}\left[\frac{|\mathsf{w}_\lambda + \theta^\star|}{n}\right] \right| \le t \right\}$$

has probability at least $1 - Ce^{-ct^2 n} - Ce^{-ctn}$.

Proof. The function

$$g \mapsto \mathsf{w}_\lambda = (\eta(\theta_i^\star + \tau_*(\lambda) g_i, \alpha_*(\lambda)\tau_*(\lambda)) - \theta_i^\star)_{1 \le i \le N}$$

is $\tau_{\max}$-Lipschitz. Consequently:

- $g \mapsto \frac{|\mathsf{w}_\lambda + \theta^\star|}{n}$ is $\delta^{-1/2} n^{-1/2} \tau_{\max}$-Lipschitz. Therefore $\frac{|\mathsf{w}_\lambda + \theta^\star|}{n}$ is $\tau_{\max}^2 \delta^{-1} n^{-1}$ sub-Gaussian: for all $t \ge 0$,
$$\mathbb{P}\left(\left| \frac{|\mathsf{w}_\lambda + \theta^\star|}{n} - \mathbb{E}\left[\frac{|\mathsf{w}_\lambda + \theta^\star|}{n}\right] \right| > t \right) \le 2e^{-nt^2\delta/\tau_{\max}^2}\,.$$
- $g \mapsto \frac{\|\mathsf{w}_\lambda\|}{\sqrt{n}}$ is $n^{-1/2}\tau_{\max}$-Lipschitz. Therefore $\frac{\|\mathsf{w}_\lambda\|}{\sqrt{n}}$ is $\tau_{\max}^2 n^{-1}$ sub-Gaussian. Its expectation is bounded by $\mathbb{E}\frac{\|\mathsf{w}_\lambda\|}{\sqrt{n}} \le (\mathbb{E}\frac{\|\mathsf{w}_\lambda\|^2}{n})^{1/2} = \tau_* \le \tau_{\max}$. By Proposition 9.7.3, we obtain that $\frac{\|\mathsf{w}_\lambda\|^2}{n}$ is (Cn^{-1}, Cn^{-1})-sub-Gamma for some constant C and therefore for all $t \ge 0$,
$$\mathbb{P}\left(\left| \frac{\|\mathsf{w}_\lambda\|^2}{n} - \mathbb{E}\left[\frac{\|\mathsf{w}_\lambda\|^2}{n}\right] \right| > t \right) \le 2e^{-cnt^2} + 2e^{-cnt}\,.$$

Now for $i \in \{1, \dots, N\}$,

$$g_i \mathsf{w}_{\lambda,i} = \tau_* g_i^2 + g_i(\mathsf{w}_{\lambda,i} - \tau_* g_i)\,.$$

$|\mathsf{w}_{\lambda,i} - \tau_* g_i| \le \alpha_* \tau_*$ and g_i is 1-sub-Gaussian and $\mathbb{E}[|g_i|] = \sqrt{\frac{2}{\pi}} \le 1$. Consequently, Lemma 9.7.1 gives that $g_i(\mathsf{w}_{\lambda,i} - \tau_* g_i)$ is $48\tau_*^2\alpha_*^2$-sub-Gaussian. This gives that $\frac{1}{n} g^{\mathsf{T}}(\mathsf{w}_\lambda - \tau_* g)$ concentrates exponentially fast around its mean. So does $\frac{1}{n}\|g\|^2$. □

Lemma 9.6.2

$$\frac{1}{n}\mathbb{E}\|\mathsf{w}_\lambda\|^2 + \sigma^2 = \tau_*^2(\lambda) \qquad \textit{and} \qquad \frac{1}{n}\mathbb{E}\Big[g^\mathsf{T}\mathsf{w}_\lambda\Big] = \tau_*(\lambda) - \beta_*(\lambda) = \frac{1}{\delta}s_*(\lambda)\,.$$

Proof. The first equality comes from Lemma 9.1.5: since $(g_i) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, we have $\mathbb{E}\|\mathsf{w}_\lambda\|^2 = N\mathbb{E}\big[w^*(\alpha_*,\tau_*)^2\big]$. The second equality comes from the optimality condition of β_*, see Lemma 9.1.6, and the definition (8.3.4) of $s_*(\lambda)$. □

The next proposition simply follows from Lemma 9.6.4 and standard concentration arguments, so we omit its proof.

Proposition 9.6.1

There exists constant $C, c > 0$ that only depend on Ω such that for all $\epsilon \in [0,1]$,

$$\mathbb{P}\Big(\Big|L_\lambda(\mathsf{w}_\lambda) - \psi_\lambda(\beta_*(\lambda),\tau_*(\lambda))\Big| > \epsilon\Big) \le Ce^{-cn\epsilon^2}\,.$$

9.6.3 Concentration of the empirical distribution

Proposition 9.6.2

Let $\theta^\star \in \mathcal{F}_p(\xi)$, where $p, \xi > 0$. Let $\mu = \widehat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$ and let $\widehat{\mu}$ be the empirical distribution of the entries of $\big(\theta_i^\star, g_i\big)_{1\le i\le N}$, where $g_1, \dots, g_N \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Then there exists constants $C, c > 0$ that only depends on ξ^p, such that for all $\epsilon \in (0, \frac{1}{2}]$,

$$\mathbb{P}\left(W_2(\widehat{\mu},\mu)^2 > \epsilon\right) \le C\epsilon^{-a}\exp\left(-cN\epsilon^2\epsilon^a\log(\epsilon)^{-2}\right)\,,$$

where $a = \frac{1}{2} + \frac{1}{p}$.

Before proving Proposition 9.6.2, we will need two simple lemmas. For $r \ge 0$ and $x \in \mathbb{R}$ we use the notation

$$x_{|r} = \begin{cases} x & \text{if} \quad -r \le x \le r\,, \\ r & \text{if} \quad x \ge r\,, \\ -r & \text{if} \quad x \le -r\,. \end{cases}$$

Let $\mu_{|r}$ be the law of $\big(\Theta, Z_{|r}\big)$ where $(\Theta, Z) \sim \widehat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$.

Lemma 9.6.3

$$W_2(\mu,\mu_{|r})^2 \le e^{-r^2/2}\,.$$

Proof. We have $W_2(\mu,\mu_{|r})^2 \le \mathbb{E}\left[(Z - Z_{|r})^2\right] \le \dfrac{2}{\sqrt{2\pi}}\displaystyle\int_r^{+\infty}(z-r)^2e^{-z^2/2}dz \le e^{-r^2/2}$. □

Let $\widehat{\mu}_{|r}$ be the empirical distribution of the entries of $\big(\theta_i^\star,\, g_{i|r}\big)_{1\le i\le N}$.

Lemma 9.6.4

With probability at least $1 - e^{-\frac{1}{128}N\epsilon^2}$, *we have*

$$W_2(\widehat{\mu}, \widehat{\mu}_{|r})^2 \leq \epsilon + e^{-r^2/2}\,.$$

Proof. Obviously $W_2(\widehat{\mu}, \widehat{\mu}_{|r})^2 \leq \frac{1}{N}\sum_{i=1}^N (g_i - g_{i|r})^2$. The function $x \mapsto x - x_{|r}$ is 1-Lipschitz, so the variables $(g_i - g_{i|r})^2$ are i.i.d. $(16,4)$-sub-Gamma. Therefore for all $\epsilon \in [0,1]$,

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N \left(g_i - g_{i|r}\right)^2 > \mathbb{E}\left(Z - Z_{|r}\right)^2 + \epsilon\right) \leq e^{-\frac{1}{128}N\epsilon^2}\,.$$

And we conclude using $\mathbb{E}(Z - Z_{|r})^2 \leq e^{-r^2/2}$, which we proved in the lemma above. □

We need now some concentration results for empirical measures, in Wasserstein distance. The next proposition follows from a direct application of Theorem 2 from [85] to distributions with bounded support. Notice that the results from [85] are much more general than this.

Proposition 9.6.3

Let $A_1, \dots A_m \overset{i.i.d.}{\sim} \nu$ *be a collection of i.i.d. random variables, bounded by some constant* $r > 0$. *Let*

$$\widehat{\nu}_m = \frac{1}{m}\sum_{i=1}^m \delta_{A_i}$$

be the empirical distribution of $A_1, \dots, A_m$. *Then there exists two absolute constants* $c, C > 0$ *such that for all* $t \geq 0$

$$\mathbb{P}\left(W_2(\nu, \widehat{\nu}_m)^2 \geq r^2 t\right) \leq C\exp(-cmt^2)\,.$$

Proof of Proposition 9.6.2. We are now going to couple $\mu_{|r}$ with $\widehat{\mu}_{|r}$. Let $R > 0$. Let $k \geq 1$ and let $\delta = 2R/k$. Define

$$B_l = [-R + (l-1)\delta, -R + l\delta)\,,$$

for $l = 1, \dots, k$. We define also $B_0 = (-\infty, R) \cup [R, +\infty)$. For $l = 0, \dots k$ we write

$$I_l = \{i \,|\, \theta_i^\star \in B_l\} \quad \text{and} \quad N_l = \#I_l\,.$$

Let $t > 0$. Let $l \in \{1, \dots, k\}$. The random variables $(g_{i|r})_{i \in I_l}$ are i.i.d. and bounded by r. By the proposition above, one can couple $i_l \sim \mathrm{Unif}(I_l)$ with $Z_l \sim \mathcal{N}(0,1)$ such that we have with probability at least $1 - Ce^{-ct^2N}$.

$$\mathsf{E}\left[(Z_{l|r} - g_{i_l|r})^2\right] \leq tr^2\sqrt{\frac{N}{N_l}}\,,$$

where E denotes the expectation with respect to i_l and Z_l. Let $j_l \sim \mathrm{Unif}(I_l)$ independently of everything else.

For $l = 0$, we define $(i_0, Z_0) \sim \mathrm{Unif}(I_0) \otimes \mathcal{N}(0,1)$, independently of everything else. We have with probability at least $1 - Ce^{-ct^2N}$:

$$\mathsf{E}\left[(Z_{0|r} - g_{i_0|r})^2\right] = \mathsf{E}[Z_{0|r}^2] + \mathsf{E}[g_{i_0|r}^2] \leq 2 + tr^2\sqrt{\frac{N}{N_0}}\,,$$

where E denotes the expectation with respect to Z_0 and i_0. Indeed, $\mathbb{E}[g_{i_0|r}^2] = \frac{1}{N_0}\sum_{i\in I_0} g_{i|r}^2 \leq 1 + tr^2\sqrt{\frac{N}{N_0}}$ with probability at least $1 - Ce^{-ct^2N}$. The equality comes from the fact that Z_0 and i_0 are independent. Finally, we define $j_0 = i_0$.

Let us now define the random variable L whose law is given by $\mathbb{P}(L = l) = \frac{N_l}{N}$, independently of everything else. Define

$$\begin{cases} Y_1 = (\theta^\star_{j_L},\, Z_{L|r})\,, \\ Y_2 = (\theta^\star_{i_L},\, g_{i_L|r})\,. \end{cases}$$

(Y_1, Y_2) is a coupling of $(\mu_{|r}, \widehat{\mu}_{|r})$. Let E denote the expectation with respect to $(i_l, Z_l)_{0\leq l\leq k}$ and L. Then

$$\begin{aligned} \mathsf{E}\,\|Y_1 - Y_2\|^2 &= \sum_{l=0}^{k} \frac{N_l}{N} \mathsf{E}\left[\, \left(\theta^\star_{i_l} - \theta^\star_{j_l}\right)^2 + \left(Z_{l|r} - g_{i_l|r}\right)^2\right] \\ &\leq \sum_{l=1}^{k} \frac{N_l}{N}\left(\sqrt{\frac{N}{N_l}}tr^2 + \delta^2\right) + \frac{N_0}{N}\left(2 + tr^2\sqrt{\frac{N}{N_0}}\right) \\ &\leq \delta^2 + \sqrt{k}tr^2 + 2\frac{N_0}{N} \leq \delta^2 + \sqrt{k}tr^2 + 2\frac{\xi^p}{R^p}\,, \end{aligned}$$

with probability at least $1 - C(k+1)e^{-ct^2N}$, where the last inequality comes from Markov's inequality, since $\theta^\star \in \mathcal{F}_p(\xi)$.

Let now $\epsilon \in (0, \frac{1}{2}]$. Let us chose

$$r = \sqrt{-2\log(\epsilon)}, \qquad R = \epsilon^{-1/p} \qquad \text{and} \qquad k = \lceil \epsilon^{-1/2-1/p} \rceil \leq 2\epsilon^{-1/2-1/p},$$

so that $\delta = 2R/k \leq 2\sqrt{\epsilon}$. Consequently

$$\mathbb{E}\|Y_1 - Y_2\|^2 \leq (4 + 2\xi^p)\epsilon + 2\sqrt{2}\epsilon^{-1/4-1/(2p)}|\log(\epsilon)|t\,.$$

So if we chose $t = |\log(\epsilon)|^{-1}\epsilon^{\frac{5}{4}+\frac{1}{2p}}$ we obtain

$$\mathbb{P}\Big(W_2(\mu_{|r}, \widehat{\mu}_{|r})^2 \leq (4 + 2\xi^p + 2\sqrt{2})\epsilon\Big) \geq 1 - C\epsilon^{-1/p-1/2}\exp(-cN\epsilon^2\epsilon^{1/2+1/p}/\log(\epsilon)^2).$$

Combining this with Lemmas 9.6.3 and 9.6.4 proves the proposition. □

9.6.4 Sparsity of the Lasso estimator

The goal of this section is to prove:

Theorem 9.6.1

Assume here that $\mathcal{D}$ is either $\mathcal{F}_0(s)$ or $\mathcal{F}_p(\xi)$ for some $0 \leq s < s_{\max}(\delta)$ and $\xi > 0, p > 0$. There exists constants $C, c > 0$ that only depend on Ω, such that for all $\epsilon \in (0,1)$

$$\sup_{\theta^\star\in\mathcal{D}} \mathbb{P}\left(\sup_{\lambda\in[\lambda_{\min},\lambda_{\max}]} \left|\frac{1}{N}\|\widehat{\theta}_\lambda\|_0 - s_*(\lambda)\right| \geq \epsilon\right) \leq \frac{C}{\epsilon^6}N^q e^{-cN\epsilon^6}\,,$$

where $q = 0$ if $\mathcal{D} = \mathcal{F}_0(s)$ and $q = (1/p - 1)_+$ if $\mathcal{D} = \mathcal{F}_p(\xi)$.

Since $\#\{i \,|\, |\widehat{v}_{\lambda,i}| = 1\} \geq \|\widehat{\theta}_\lambda\|_0$, Theorem 9.5.6 gives that

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\left(\exists \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \frac{1}{N}\|\widehat{\theta}_\lambda\|_0 \geq s_*(\lambda) + \epsilon\right) \leq \frac{C}{\epsilon^6} N^q e^{-cN\epsilon^6}. \tag{9.6.12}$$

It remains to prove the converse lower bound in order to get Theorem 9.6.1. We start with the following 'local stability' property of the Lasso cost:

Proposition 9.6.4

There exists constants $C, c, \gamma > 0$ that only depend on Ω such that for all $\epsilon \in (0, 1]$

$$\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\exists \theta \in \mathbb{R}^N, \ \frac{1}{N}\|\theta\|_0 < s_*(\lambda) - \epsilon \quad \text{and} \quad \mathcal{L}_\lambda(\theta) \leq \min \mathcal{L}_\lambda + \gamma\epsilon^3\Big) \leq \frac{C}{\epsilon^3} e^{-cN\epsilon^6}.$$

Proposition 9.6.4 is a consequence of Proposition 9.3.1 and Lemma 9.6.5 below.

Lemma 9.6.5

There exists constants $\gamma, c, C > 0$ that only depend on Ω such that for all $\epsilon \in (0, 1]$ we have

$$\mathbb{P}\left(\min_{w \in D_\epsilon} L_\lambda(w) \leq \min_{w \in \mathbb{R}^N} L_\lambda(w) + 3\gamma\epsilon^3\right) \leq \frac{C}{\epsilon^3} e^{-cn\epsilon^6},$$

where $D_\epsilon = \Big\{w \in \mathbb{R}^N \,\Big|\, \frac{1}{N}\|w + \theta^\star\|_0 < s_*(\lambda) - \epsilon\Big\}$.

Proof. Define $\mathsf{x}_\lambda = \mathsf{w}_\lambda + \theta^\star = \big(\eta(\theta_i^\star + \tau_* g_i, \alpha_*\tau_*)\big)_{1 \leq i \leq N}$, and for $r > 0$

$$s_r = \frac{1}{N}\#\Big\{i \in \{1, \dots, N\} \,\Big|\, |\mathsf{x}_{\lambda,i}| \geq r\Big\}.$$

s_r is a mean of independent Bernoulli random variables, by Hoeffding's inequality we have:

$$\mathbb{P}\Big(s_r \geq \mathbb{P}(|\Theta + \tau_* Z| \geq \alpha_*\tau_* + r) - \frac{\epsilon}{4}\Big) \geq 1 - e^{-N\epsilon^2/8}.$$

Compute

$$\begin{aligned}\mathbb{P}(|\Theta + \tau_* Z| \geq \alpha_*\tau_* + r) &= \mathbb{E}\left[\Phi\Big(\frac{\Theta}{\tau_*(\lambda)} - \alpha_*(\lambda) - \frac{r}{\tau_*(\lambda)}\Big) + \Phi\Big(-\frac{\Theta}{\tau_*(\lambda)} - \alpha_*(\lambda) - \frac{r}{\tau_*(\lambda)}\Big)\right] \\ &\geq s_*(\lambda) - \frac{r}{\sigma}.\end{aligned}$$

Let us chose $r = \sigma\epsilon/4$. We have then $\mathbb{P}\Big(s_r \geq s_*(\lambda) - \frac{\epsilon}{2}\Big) \geq 1 - e^{-N\epsilon^2/8}$. By Theorem 9.2.1 there exists a constant $\gamma > 0$ such that the event

$$\Big\{\forall w \in \mathbb{R}^N, \ L_\lambda(w) \leq \min_{v \in \mathbb{R}^N} L_\lambda(v) + 3\gamma\epsilon^3 \Rightarrow \frac{1}{N}\|w - \mathsf{w}_\lambda\|^2 < \frac{\sigma^2\epsilon^3}{32}\Big\} \bigcap \Big\{s_r \geq s_*(\lambda) - \frac{\epsilon}{2}\Big\} \tag{9.6.13}$$

has probability at least $1 - \frac{C}{\epsilon^3} e^{-cn\epsilon^6}$. We have on this event, for all $w \in D_\epsilon$

$$\frac{1}{N}\|w - \mathsf{w}_\lambda\|^2 = \frac{1}{N}\|w + \theta^\star - \mathsf{x}_\lambda\|^2 \geq \frac{\epsilon}{2} r^2 = \frac{\sigma^2\epsilon^3}{32}.$$

Therefore, on the event (9.5.10) we have $\min_{w \in D_\epsilon} L_\lambda(w) > \min_{w \in \mathbb{R}^N} L_\lambda(w) + 3\gamma\epsilon^3$. We conclude

$$\mathbb{P}\left(\min_{w \in D_\epsilon} L_\lambda(w) \leq \min_{w \in \mathbb{R}^N} L_\lambda(w) + \gamma\epsilon^3\right) \leq \frac{C}{\epsilon^3} e^{-cn\epsilon^6}.$$

□

Using the same arguments that we use to deduce Theorems 8.3.1 and 8.3.2 (8.3.8) from Theorem 8.5.3 and Theorem 9.3.1 in Section 9.3.2, we deduce from Proposition 9.6.4 that for all $\epsilon \in (0,1]$

$$\sup_{\theta^\star \in \mathcal{D}} \mathbb{P}\Big(\exists \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \frac{1}{N}\|\widehat{\theta}_\lambda\|_0 < s_*(\lambda) - \epsilon \Big) \leq \frac{C}{\epsilon^6} N^q e^{-cn\epsilon^6} .$$

This proves, together with (9.6.12), Theorem 9.6.1.

9.6.5 Proof of Theorem 8.3.3

Recall that the distributions μ^*_λ and ν^*_λ are respectively defined by Definition 8.3.3 and (9.5.1). Let $\epsilon \in (0,1]$. From now, we will work on the event

$$\begin{aligned}\mathcal{E} =& \Big\{ \forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad W_2(\widehat{\mu}_{(\widehat{\theta}_\lambda, \theta^\star)}, \mu^*_\lambda)^2 + W_2(\widehat{\mu}_{(\widehat{v}_\lambda, \theta^\star)}, \nu^*_\lambda)^2 \leq \epsilon^6 \Big\} \\ & \bigcap \Big\{ \forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \Big| \frac{1}{N}\|\widehat{\theta}_\lambda\|_0 - s_*(\lambda) \Big| + \Big| \frac{1}{N} \#\{ i \,|\, |\widehat{v}_{\lambda,i}| = 1 \} - s_*(\lambda) \Big| \leq \epsilon^2 \Big\} ,\end{aligned}$$

which has probability at least $1 - C\epsilon^{-12} e^{-cN\epsilon^{17}}$ from what we have just seen, and Theorems 8.3.1, 9.5.2, 9.6.1 and 9.5.6. From now, E and P will denote the probability with respect to the empirical distributions of the entries of the vectors we study, and the variables that we couple with them. Let $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. On the event $\mathcal{E}$ one can couple $(\Theta^x, Z^x) \sim \hat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$ and $(\Theta^v, Z^v) \sim \hat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$ with $(\Theta, \widehat{\Theta}_\lambda, \widehat{V}_\lambda, \widehat{\Theta}^d_\lambda)$ which is sampled from the empirical distribution of the entries of $(\theta^\star, \widehat{\theta}_\lambda, \widehat{v}_\lambda, \widehat{\theta}^d_\lambda)$, such that

$$\begin{aligned}&\mathsf{E}\Big[\Big(\widehat{\Theta}_\lambda - \eta(\Theta^x + \tau_* Z^x, \alpha_* \tau_*) \Big)^2 + \Big(\Theta - \Theta^x \Big)^2 \Big] \leq \epsilon^6 , \\ &\mathsf{E}\Big[\Big(\widehat{V}_\lambda + \frac{1}{\alpha_* \tau_*} \Big(\eta(\Theta^v + \tau_* Z^v, \alpha_* \tau_*) - \Theta^v - \tau_* Z^v_0 \Big) \Big)^2 + \Big(\Theta - \Theta^v \Big)^2 \Big] \leq \epsilon^6 .\end{aligned}$$

Let

$$E_1 = \Big\{ |\widehat{\Theta}_\lambda - \eta(\Theta^x + \tau_* Z^x, \alpha_* \tau_*)| \leq \epsilon^2, \quad |\widehat{V}_\lambda + \frac{1}{\alpha_* \tau_*} (\eta(\Theta^v + \tau_* Z^v, \alpha_* \tau_*) - \Theta^v - \tau_* Z^v)| \leq \epsilon^2 \Big\} .$$

By Chebychev's inequality, $\mathsf{P}(E_1) \geq 1 - C\epsilon^2$, for some constant $C > 0$. Let us also define the event

$$E_2 = \Big\{ \Theta^x + \tau_* Z^x \neq \alpha_* \tau_* \quad \text{and} \quad \Theta^v + \tau_* Z^v \neq \alpha_* \tau_* \Big\} .$$

$\Theta^x + \tau_* Z^x$ and $\Theta^v + \tau_* Z^v$ admit a density with respect to Lebesgue's measure. Therefore $\mathsf{P}(E_2) = 1$.

Lemma 9.6.6

The event

$$E_3 = \Big\{ |\widehat{\Theta}_\lambda| \notin (0, \epsilon^2] \quad \textit{and} \quad |\widehat{V}_\lambda| \notin [1 - \epsilon^2, 1) \Big\}$$

has probability at least $1 - C\epsilon^2$.

Proof. We denote here by $O(\epsilon^2)$ quantities that are bounded by $C\epsilon^2$, from some constant C. Since $\Theta^x + \tau_* Z^x$ admits a density with respect to Lebesgue's measure we have

$$\mathsf{P}\Big(|\eta(\Theta^x + \tau_* Z^x, \alpha_* \tau_*)| \notin (0, 2\epsilon^2] \Big) = 1 - O(\epsilon^2) .$$

Consequently, since the events E_1 has probability at least $1 - O(\epsilon^2)$, we have

$$\begin{aligned}
&\mathsf{P}(|\widehat{\Theta}_\lambda| \in [0, \epsilon^2]) \\
&= \mathsf{P}\Big(|\widehat{\Theta}_\lambda| \in [0, \epsilon^2] \quad \text{and} \quad |\eta(\Theta^x + \tau_* Z^x, \alpha_* \tau_*)| \notin (0, 2\epsilon^2] \\
&\qquad\qquad \text{and} \quad |\widehat{\Theta}_\lambda - \eta(\Theta^x + \tau_* Z^x, \alpha_* \tau_*)| \leq \epsilon^2\Big) + O(\epsilon^2) \\
&= \mathsf{P}\Big(\eta(\Theta^x + \tau_* Z^x, \alpha_* \tau_*) = 0\Big) + O(\epsilon^2) = s_*(\lambda) + O(\epsilon^2)\,.
\end{aligned}$$

Since $\mathsf{P}(\widehat{\Theta}_\lambda = 0) = s_*(\lambda) + O(\epsilon^2)$ because we are working on $\mathcal{E}$, we conclude that $\mathsf{P}(|\widehat{\Theta}_\lambda| \in (0, \epsilon^2]) = O(\epsilon^2)$. One can prove the same way that $\mathsf{P}(|\widehat{V}_\lambda| \in [1 - \epsilon^2, 1)) = O(\epsilon^2)$, which gives the desired result. □

Lemma 9.6.7

The event

$$E_4 = \Big\{\widehat{\Theta}_\lambda \neq 0 \iff \widehat{V}_\lambda = \operatorname{sign}(\widehat{\Theta}_\lambda)\Big\}$$

has probability at least $1 - C\epsilon^2$, for some constant $C > 0$.

Proof. Since $\widehat{v}_\lambda \in \partial|\widehat{\theta}_\lambda|$, $\widehat{\theta}_{\lambda,i} > 0$ implies that $\widehat{v}_{\lambda,i} = \operatorname{sign}(\widehat{\theta}_{\lambda,i})$. Thus $\mathsf{P}(\widehat{\Theta}_\lambda \neq 0 \implies \widehat{V}_\lambda = \operatorname{sign}(\widehat{\theta}_{\lambda,i})) = 1$. We have thus

$$\{\widehat{\Theta}_\lambda \neq 0\} \subset \{|\widehat{V}_\lambda| = 1\}\,. \tag{9.6.14}$$

On the event $\mathcal{E}$ we have $\Big|\frac{1}{N}\|\widehat{\theta}_\lambda\|_0 - s_*(\lambda)\Big| + \Big|\frac{1}{N}\#\{i \,|\, |\widehat{v}_{\lambda,i}| = 1\} - s_*(\lambda)\Big| \leq \epsilon^2$ which gives

$$\mathsf{P}(\widehat{\Theta}_\lambda \neq 0) = s_*(\lambda) + O(\epsilon^2) \qquad \text{and} \qquad \mathsf{P}(|\widehat{V}_\lambda| = 1) = s_*(\lambda) + O(\epsilon^2).$$

We deduce then from (9.6.14) that $\mathsf{P}(|\widehat{V}_\lambda| = 1 \text{ and } \widehat{\Theta}_\lambda = 0) = O(\epsilon^2)$ and finally $\mathsf{P}(\widehat{V}_\lambda = \operatorname{sign}(\widehat{\Theta}_\lambda) \implies \widehat{\Theta}_\lambda \neq 0) \geq 1 - C\epsilon^2$. □

Lemma 9.6.8

Let $E = E_1 \cap E_2 \cap E_3 \cap E_4$. The event E has probability at least $1 - C\epsilon^2$ and on E we have

$$\Theta^v + \tau_* Z^v \geq \alpha_* \tau_* \iff \Theta^x + \tau_* Z^x \geq \alpha_* \tau_*\,,$$

and

$$\Theta^v + \tau_* Z^v \leq -\alpha_* \tau_* \iff \Theta^x + \tau_* Z^x \leq -\alpha_* \tau_*\,.$$

Proof. Since E_1, E_2, E_3 and E_4 have all a probability greater than $1 - O(\epsilon^2)$, the event $E = E_1 \cap E_2 \cap E_3 \cap E_4$ has probability at least $1 - O(\epsilon^2)$. On E we have

$$\begin{aligned}
\Theta^v + \tau_* Z^v \geq \alpha_* \tau_* &\iff \Theta^v + \tau_* Z_0^v - \eta(\Theta^v + \tau_* Z^v, \alpha_* \tau_*) = \alpha_* \tau_* \\
&\iff \widehat{V}_\lambda \geq 1 - \epsilon^2 \qquad \text{(because we are on the event } E_1\text{)} \\
&\iff \widehat{V}_\lambda = 1 \qquad \text{(because we are on the event } E_3\text{)} \\
&\iff \widehat{\Theta}_\lambda > 0 \qquad \text{(because we are on the event } E_4\text{)} \\
&\iff \widehat{\Theta}_\lambda > \epsilon^2 \qquad \text{(because we are on the event } E_3\text{)} \\
&\iff \eta(\Theta^x + \tau_* Z^x, \alpha_* \tau_*) > 0 \quad \text{(because we are on the event } E_1\text{)} \\
&\iff \Theta^x + \tau_* Z^x \geq \alpha_* \tau_* \qquad \text{(because we are on the event } E_2\text{)}\,.
\end{aligned}$$

The second equivalence is proved exactly the same way. □

Let us define

$$X^d = \eta(\Theta^x + \tau_* Z^x, \alpha_*\tau_*) + \Theta^v + \tau_* Z_0^v - \eta(\Theta^v + \tau_* Z^v, \alpha_*\tau_*)\,.$$

We have

$$\begin{aligned}\mathsf{E}(\widehat{\Theta}_\lambda^d - X^d)^2 &= 2\mathsf{E}\Big[\Big(\widehat{\Theta}_\lambda - \eta(\Theta^x + \tau_* Z^x, \alpha_*\tau_*)\Big)^2\Big] \\ &\quad + 2\mathsf{E}\Big[\Big(\frac{\lambda}{1 - \frac{1}{n}\|\widehat{\theta}_\lambda\|_0}\widehat{V}_\lambda - \Theta^v - \tau_* Z^v + \eta(\Theta^v + \tau_* Z^v, \alpha_*\tau_*)\Big)^2\Big] \\ &\le C\epsilon^4\,,\end{aligned}$$

for some constant $C > 0$, because on the event $\mathcal{E}$, $\frac{1}{N}\Big|\|\widehat{\theta}_\lambda\|_0 - s_*(\lambda)\Big| \le \epsilon^2$, so

$$\frac{\lambda}{1 - \frac{1}{n}\|\widehat{\theta}_\lambda\|_0} = \frac{\lambda}{1 - \frac{1}{\delta}s_*(\lambda)} + O(\epsilon^2) = \alpha_*\tau_* + O(\epsilon^2)\,.$$

By Lemma 9.6.8 above, we have on the event E,

$$X^d = \begin{cases}\Theta^x + \tau_* Z^x & \text{if} \quad \Theta^x + \tau_* Z^x \ge \alpha_*\tau_* \quad \text{or} \quad \Theta^x + \tau_* Z^x \le -\alpha_*\tau_*\,, \\ \Theta^v + \tau_* Z^v & \text{otherwise.}\end{cases}$$

Let us denote $T^x = (\Theta^x + \tau_* Z^x, \Theta^x)$ and $T^v = (\Theta^v + \tau_* Z^v, \Theta^v)$.

Since $\Theta^x + \tau_*^x Z$ and $\Theta^v + \tau_* Z^v$ have the same law and $\mathsf{P}\Big(\Theta^x + \tau_* Z^x \ge \alpha_*\tau_*\Big|E\Big) = \mathsf{P}\Big(\Theta^v + \tau_* Z^v \ge \alpha_*\tau_*\Big|E\Big)$ (by Lemma 9.6.8), we have $\mathsf{P}\Big(\Theta^x + \tau_* Z^x \ge \alpha_*\tau_*\Big|E^c\Big) = \mathsf{P}\Big(\Theta^v + \tau_* Z^v \ge \alpha_*\tau_*\Big|E^c\Big)$. Similarly we have $\mathsf{P}\Big(\Theta^x + \tau_* Z^x \le -\alpha_*\tau_*\Big|E^c\Big) = \mathsf{P}\Big(\Theta^v + \tau_* Z^v \le -\alpha_*\tau_*\Big|E^c\Big)$.

One can therefore define two random variables $\widetilde{T}^x = (\widetilde{\Theta}^x + \tau_*\widetilde{Z}^x,\, \widetilde{\Theta}^x)$ and $\widetilde{T}^v = (\widetilde{\Theta}^v + \tau_*\widetilde{Z}^v,\, \widetilde{\Theta}^v)$ such that

- conditionally on E^c, $\widetilde{T}^x$ (respectively $\widetilde{T}^v$) and T^x (respectively T^v) have the same law.
- On the event E^c, $\widetilde{\Theta}^x + \tau_*\widetilde{Z}^x \ge \alpha_*\tau_* \iff \widetilde{\Theta}^v + \tau_*\widetilde{Z}^v \ge \alpha_*\tau_*$ and $\widetilde{\Theta}^x + \tau_*\widetilde{Z}^x \le -\alpha_*\tau_* \iff \widetilde{\Theta}^v + \tau_*\widetilde{Z}^v \le -\alpha_*\tau_*$.

We define then

$$\Big(\widetilde{X}^d, \widetilde{\Theta}\Big) = \begin{cases}T^x & \text{on the event } E \text{ provided that} \quad |\Theta^x + \tau_* Z^x| \ge \alpha_*\tau_*\,, \\ T^v & \text{on the event } E \text{ provided that} \quad |\Theta^x + \tau_* Z^x| < \alpha_*\tau_*\,, \\ \widetilde{T}^x & \text{on the event } E^c \text{ provided that} \quad |\widetilde{\Theta}^x + \tau_*\widetilde{Z}^x| \ge \alpha_*\tau_*\,, \\ \widetilde{T}^v & \text{on the event } E^c \text{ provided that} \quad |\widetilde{\Theta}^x + \tau_*\widetilde{Z}^x| < \alpha_*\tau_*\,.\end{cases}$$

$(\widetilde{X}^d, \widetilde{\Theta}) \sim \mu_\lambda^d$ which is the law of $(\Theta + \tau_* Z, \Theta)$ where $(\Theta, Z) \sim \widehat{\mu}_{\theta^\star} \otimes \mathcal{N}(0,1)$. Indeed, for

every continuous bounded function f we have

$$\begin{aligned}
\mathsf{E}[f(\widetilde{X}^d,\widetilde{\Theta})] &= \mathsf{E}\Big[\mathbb{1}_E\mathbb{1}_{|\Theta^x+\tau_*Z^x|\geq\alpha_*\tau_*}f(T^x)\Big]+\mathsf{E}\Big[\mathbb{1}_E\mathbb{1}_{|\Theta^v+\tau_*Z^v|<\alpha_*\tau_*}f(T^v)\Big]\\
&\quad+\mathsf{E}\Big[\mathbb{1}_{E^c}\mathbb{1}_{|\widetilde{\Theta}^x+\tau_*\widetilde{Z}^x|\geq\alpha_*\tau_*}f(\widetilde{T}^x)\Big]+\mathsf{E}\Big[\mathbb{1}_{E^c}\mathbb{1}_{|\widetilde{\Theta}^v+\tau_*\widetilde{Z}^v|<\alpha_*\tau_*}f(\widetilde{T}^v)\Big]\\
&= \mathsf{E}\Big[\mathbb{1}_E\mathbb{1}_{|\Theta^x+\tau_*Z^x|\geq\alpha_*\tau_*}f(T^x)\Big]+\mathsf{E}\Big[\mathbb{1}_E\mathbb{1}_{|\Theta^v+\tau_*Z^v|<\alpha_*\tau_*}f(T^v)\Big]\\
&\quad+\mathsf{E}\Big[\mathbb{1}_{E^c}\mathbb{1}_{|\Theta^x+\tau_*Z^x|\geq\alpha_*\tau_*}f(T^x)\Big]+\mathsf{E}\Big[\mathbb{1}_{E^c}\mathbb{1}_{|\Theta^v+\tau_*Z^v|<\alpha_*\tau_*}f(T^v)\Big]\\
&= \mathsf{E}\Big[\mathbb{1}_{|\Theta^x+\tau_*Z^x|\geq\alpha_*\tau_*}f(T^x)\Big]+\mathsf{E}\Big[\mathbb{1}_{|\Theta^v+\tau_*Z^v|<\alpha_*\tau_*}f(T^v)\Big]\\
&= \mathsf{E}\Big[\mathbb{1}_{|\Theta^x+\tau_*Z^x|\geq\alpha_*\tau_*}f(T^x)\Big]+\mathsf{E}\Big[\mathbb{1}_{|\Theta^x+\tau_*Z^x|<\alpha_*\tau_*}f(T^x)\Big]=\mathsf{E}\Big[f(T^x)\Big].
\end{aligned}$$

Let us now compute

$$\mathsf{E}\Big[\big(X^d-\widetilde{X}^d\big)^2\Big]=\mathsf{E}\Big[\mathbb{1}_{E^c}\big(X^d-\widetilde{X}^d\big)^2\Big]\leq C\sqrt{\mathsf{P}(E^c)}\leq C\epsilon\,,$$

and

$$\begin{aligned}
\mathsf{E}\Big[\big(\widetilde{\Theta}-\Theta\big)^2\Big] &\leq \mathsf{E}\Big[\mathbb{1}_E\big(\Theta^x-\Theta\big)^2\Big]+\mathsf{E}\Big[\mathbb{1}_E\big(\Theta^v-\Theta\big)^2\Big]+\mathsf{E}\Big[\mathbb{1}_{E^c}\big(\widetilde{\Theta}^x-\Theta\big)^2\Big]+\mathsf{E}\Big[\mathbb{1}_{E^c}\big(\widetilde{\Theta}^v-\Theta\big)^2\Big]\\
&\leq 2\epsilon^6+2C\sqrt{\mathsf{P}(E^c)}\leq C\epsilon\,.
\end{aligned}$$

Therefore $\mathsf{E}\big\|(\widehat{\Theta}^d_\lambda,\Theta)-(\widetilde{X}^d,\widetilde{\Theta})\big\|^2\leq C\epsilon$ and consequently $W_2(\widehat{\mu}_{(\widehat{\theta}^d_\lambda,\theta^\star)},\mu^d_\lambda)^2\leq C\epsilon$, on the event $\mathcal{E}$ which has probability at least $1-C\epsilon^{-12}e^{-cN\epsilon^{17}}$.

9.6.6 Proof of Corollary 8.4.2

Let $\epsilon\in(0,1]$. Let us work on the intersection the events of Theorem 9.6.1,Corollary 8.4.1 and 9.5.3, which as probability at least $1-\frac{C}{\epsilon^6}e^{-cN\epsilon^6}$. Let $\lambda\in[\lambda_{\min},\lambda_{\max}]$.

$$\frac{1}{N}\big\|X^\mathsf{T}(y-X\widehat{\theta}_\lambda)\|^2=\frac{\lambda^2}{N}\|\widehat{v}_\lambda\|^2=\lambda^2\kappa_*(\lambda)+O(\epsilon)\,.$$

We have also $1-\frac{1}{n}\|\widehat{\theta}_\lambda\|_0=1-\frac{1}{\delta}s_*(\lambda)+O(\epsilon)=\beta_*(\lambda)/\tau_*(\lambda)+O(\epsilon)$. Therefore

$$\frac{\big\|X^\mathsf{T}(y-X\widehat{\theta}_\lambda)\big\|^2}{N(1-\frac{1}{n}\|\widehat{\theta}_\lambda\|_0)^2}=\Big(\frac{\lambda\tau_*(\lambda)}{\beta_*(\lambda)}\Big)^2\kappa_*(\lambda)+O(\epsilon)=\tau_*(\lambda)^2\Big(1+\delta-2s_*(\lambda)\Big)-\delta\sigma^2+O(\epsilon)\,.$$

Now we have $\widehat{\tau}(\lambda)=\tau_*(\lambda)+O(\epsilon)$ and $\frac{1}{N}\|\widehat{\theta}_\lambda\|_0=s_*(\lambda)+O(\epsilon)$. Consequently

$$\widehat{\tau}(\lambda)^2\Big(\frac{2}{N}\|\widehat{\theta}_\lambda\|_0-1\Big)=\tau_*(\lambda)^2(2s_*(\lambda)-1)+O(\epsilon)\,.$$

Putting all together we obtain $\widehat{R}(\lambda)=\delta\tau_*(\lambda)^2-\delta\sigma^2+O(\epsilon)=R_*(\lambda)+O(\epsilon)$, and we conclude using Theorem 8.3.2.

9.6.7 Proof of Proposition 8.4.3

Let $n' \in \{1, \dots, n\}$. We consider a random $n' \times N$ matrix X' and a random vector $z' = (z'_1, \dots, z'_{n'})$ such that $X'_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/n)$ and $z'_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ are independent and independent of everything else.

Lemma 9.6.9

There exists constants $\gamma, c, C > 0$ that only depend on Ω such that for all $\theta^\star$ in $\mathcal{D}$ and all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ such that for all $\epsilon \in (0,1]$,

$$\mathbb{P}\Big(\exists w \in \mathbb{R}^N,\ \Big|\frac{1}{n'}\|X'w - \sigma z'\|^2 - \frac{1}{n}\|w\|^2 - \sigma^2\Big| \geq \sqrt{\epsilon}\frac{n}{n'} \quad \text{and} \quad L_\lambda(w) \leq \min_{v \in \mathbb{R}^N} L_\lambda(v) + \gamma\epsilon\Big) \leq \frac{C}{\epsilon} e^{-cn\epsilon^2}.$$

Proof. The vector w_λ is independent from X', z'. Hence

$$\|X'\mathsf{w}_\lambda - \sigma z'\|^2 = \Big(\frac{1}{n}\|\mathsf{w}_\lambda\|^2 + \sigma^2\Big)\chi \tag{9.6.15}$$

where χ is independent from w_λ and follows a χ-squared distribution with n' degrees of freedom. We have therefore for all $t \geq 0$

$$\mathbb{P}\Big(|\chi - n'| \geq tn'\Big) \leq Ce^{-cn't} + Ce^{-cn't^2}, \tag{9.6.16}$$

for some constants $c, C > 0$. We know by Lemma 9.6.1 and Lemma 9.6.2 that $\frac{1}{n}\|\mathsf{w}_\lambda\|^2$ concentrates exponentially fast around $\tau_*(\lambda)^2 - \sigma^2$, which is (Theorem 9.1.2) bounded by some constant. There exists therefore constants $C, c > 0$ such that

$$\mathbb{P}\Big(\frac{1}{n}\|\mathsf{w}_\lambda\|^2 + \sigma^2 > C\Big) \leq Ce^{-cn}. \tag{9.6.17}$$

From (9.6.15)-(9.6.16) and (9.6.17) above, we deduce that for all $t \geq 0$

$$\mathbb{P}\Big(|\frac{1}{n'}\|X'\mathsf{w}_\lambda - \sigma z'\|^2 - \frac{1}{n}\|\mathsf{w}_\lambda\|^2 - \sigma^2| > t\Big) \leq Ce^{-cn't} + Ce^{-cn't^2} + Ce^{-cn}, \tag{9.6.18}$$

for some constants $c, C > 0$. By Proposition 9.7.4, we know that $\mathbb{P}(\sigma_{\max}(X') > \delta^{-1/2} + \sqrt{n'/n} + 1) \leq e^{-n/2}$. Let $\epsilon \in (0,1]$. Let $w \in \mathbb{R}^N$ such that $\|w - \mathsf{w}_\lambda\|^2 \leq \epsilon N$.

$$\begin{aligned}\frac{1}{\sqrt{n'}}\|X'\mathsf{w}_\lambda - \sigma z'\| + \frac{1}{\sqrt{n'}}\|X'w - \sigma z'\| &\leq \frac{2\sigma}{\sqrt{n'}}\|z'\| + \frac{\sigma_{\max}(X')}{\sqrt{n'}}(\|\mathsf{w}_\lambda\| + \|w\|) \\ &\leq C\sqrt{\frac{n}{n'}}\end{aligned}$$

for some constant $C > 0$, with probability at least $1 - Ce^{-cn}$. Consequently

$$\begin{aligned}\Big|\frac{1}{n'}\|X'\mathsf{w}_\lambda - \sigma z'\|^2 - \frac{1}{n'}\|X'w - \sigma z'\|^2\Big| &\leq \frac{C\sqrt{n}}{n'}\Big|\|X'\mathsf{w}_\lambda - \sigma z'\| - \|X'w - \sigma z'\|\Big| \\ &\leq \frac{C\sqrt{n}}{n'}\|X'(\mathsf{w}_\lambda - w)\| \\ &\leq C\sigma_{\max}(X')\sqrt{\delta^{-1}\epsilon}\frac{n}{n'} \leq C\sqrt{\epsilon}\frac{n}{n'}\end{aligned}$$

with probability at least $1 - Ce^{-cn}$ for some constant $C > 0$. Similarly, we have with probability at least $1 - Ce^{-cn}$,

$$\Big|\frac{1}{n}\|\mathsf{w}_\lambda\|^2 - \frac{1}{n}\|w\|^2\Big| \leq C\sqrt{\epsilon}.$$

We conclude that with probability at least $1-Ce^{-cn}$ we have for all $w \in \mathbb{R}^N$ such that $\|w-\mathsf{w}_\lambda\|^2 \leq N\epsilon$,

$$\Big|\frac{1}{n'}\|X'\mathsf{w}_\lambda - \sigma z'\|^2 - \frac{1}{n}\|\mathsf{w}_\lambda\|^2 - \frac{1}{n'}\|X'w + \sigma z'\|^2 + \frac{1}{n}\|w\|^2\Big| \leq C\sqrt{\epsilon}(1 + \frac{n}{n'}) \leq 2C\sqrt{\epsilon}\frac{n}{n'}.$$

Combining this with (9.6.18), we get that for all $\epsilon \in (0, 1]$,

$$\mathbb{P}\Big(\exists w \in \mathbb{R}^N,\ \|w - \mathsf{w}_\lambda\|^2 \leq N\epsilon \quad \text{and} \quad \Big|\frac{1}{n'}\|X'w + \sigma z'\|^2 - \frac{1}{n}\|w\|^2 - \sigma^2\Big| > 2C\sqrt{\epsilon}\frac{n}{n'}\Big) \leq Ce^{-cn\epsilon}.$$

We conclude using Theorem 9.2.1 that

$$\mathbb{P}\Big(\exists w \in \mathbb{R}^N,\ \Big|\frac{1}{n'}\|X'w-\sigma z'\|^2-\frac{1}{n}\|w\|^2-\sigma^2\Big| \geq \sqrt{\epsilon}\frac{n}{n'} \quad \text{and} \quad L_\lambda(w) \leq \min_{v\in\mathbb{R}^N} L_\lambda(v)+\gamma\epsilon\Big) \leq \frac{C}{\epsilon}e^{-cn\epsilon^2}$$

for some constants $c, C, \gamma > 0$. □

Using Proposition 9.3.1, we deduce

Lemma 9.6.10

There exists constants $\gamma, c, C > 0$ that only depend on Ω such that for all $\theta^\star$ in $\mathcal{D}$ and all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ such that for all $\epsilon \in (0, 1]$,

$$\mathbb{P}\Big(\exists \theta \in \mathbb{R}^N,\ \Big|\frac{1}{n'}\|X'\theta^\star+\sigma z'-X'\theta\|^2-\frac{1}{n}\|\theta-\theta^\star\|^2-\sigma^2\Big| \geq \sqrt{\epsilon}\frac{n}{n'} \quad \text{and} \quad \mathcal{L}_\lambda(\theta) \leq \min \mathcal{L}_\lambda+\gamma\epsilon\Big) \leq \frac{C}{\epsilon}e^{-cn\epsilon^2}.$$

We have

$$\begin{aligned}
\widehat{\theta}^i_\lambda &= \operatorname*{arg\,min}_{\theta\in\mathbb{R}^N}\left\{\frac{1}{2n_k}\left\|y^{(\setminus i)} - X^{(\setminus i)}\theta\right\|^2 + \frac{\lambda}{n}|\theta|\right\} \\
&= \operatorname*{arg\,min}_{\theta\in\mathbb{R}^N}\left\{\frac{1}{2n_k}\left\|X^{(\setminus i)}\theta^\star + \sigma z^{(\setminus i)} - X^{(\setminus i)}\theta\right\|^2 + \frac{\lambda}{n}|\theta|\right\} \\
&= \operatorname*{arg\,min}_{\theta\in\mathbb{R}^N}\left\{\frac{1}{2n}\left\|\sqrt{\frac{k}{k-1}}X^{(\setminus i)}\theta^\star + \sqrt{\frac{k}{k-1}}\sigma z^{(\setminus i)} - \sqrt{\frac{k}{k-1}}X^{(\setminus i)}\theta\right\|^2 + \frac{\lambda}{n}|\theta|\right\}.
\end{aligned}$$

$\widehat{\theta}^i_\lambda$ is thus the minimizer of the Lasso cost (8.3.1) for $\delta^{(k)} = \frac{k-1}{k}\delta$ and $\sigma^{(k)} = \sqrt{k/(k-1)}\sigma$. Let $\tau_*^{(k)}(\lambda)$ be the τ_* defined by Theorem 8.3.1, but with $\delta^{(k)}$ instead of δ and $\sigma^{(k)}$ instead of σ. Define the corresponding 'risk':

$$R_*^{(k)}(\lambda) = \delta^{(k)}\Big(\tau_*^{(k)}(\lambda)^2 - (\sigma^{(k)})^2\Big).$$

It is not difficult to verify that the bounds on τ_*, β_* of Section 9.1.2 are uniform with respect to δ and σ. More precisely

$$\sup_{\delta\in[\delta_{\min},\delta_{\max}]}\ \sup_{\sigma\in[\sigma_{\min},\sigma_{\max}]}\ \sup_{\lambda\in[\lambda_{\min},\lambda_{\max}]}\ \sup_{\theta^\star\in\mathcal{D}}\Big\{\tau_*(\lambda,\delta,\sigma) + \beta_*(\lambda,\delta,\sigma)\Big\} < +\infty,$$

where $\delta_{\max}, \delta_{\min}, \sigma_{\max}, \sigma_{\min} > 0$ such that $s_{\max}(\delta_{\min}) > s$ if we are in the case $\mathcal{D} = \mathcal{F}_0(s)$. This gives that under the assumptions of Proposition 8.4.3, $\tau_*^{(k)}$ and $R_*^{(k)}$ are bounded for all $k \geq 2$ (that verify $s_{\max}(\delta(k-1)/k) > s$ in the case $\mathcal{D} = \mathcal{F}_0(s)$) by some constant that depends only on Ω.

Lemma 9.6.11

There exists constants $C, c > 0$ that only depend on Ω such that for all $\theta^\star \in \mathcal{D}$, for all $i \in \{1, \dots, k\}$ and for all $\epsilon \in (0, 1]$,

$$\mathbb{P}\Big(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \Big| \frac{k}{n} \|y^{(i)} - X^{(i)} \widehat{\theta}^i_\lambda\|^2 - \frac{1}{n} \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - \sigma^2 \Big| \geq k\epsilon \Big) \leq C N^q \epsilon^{-4} e^{-cn\epsilon^4}.$$

Proof. Let $i \in \{1, \dots, k\}$. Let us define for $\theta \in \mathbb{R}^N$,

$$\mathcal{L}^{(\setminus i)}_\lambda(\theta) = \frac{1}{2n_k} \|y^{(\setminus i)} - X^{(\setminus i)}\theta\|^2 + \frac{\lambda}{n}|\theta|.$$

Let $\epsilon \in (0, 1]$. Let $\eta = \frac{\gamma\epsilon}{KN^q}$ and $M = \lceil (\lambda_{\max} - \lambda_{\min})/\eta \rceil$. Define for $j \in \{0, \dots, M\}$, define $\lambda_j = \min\left(\lambda_{\min} + j\eta, \lambda_{\max}\right)$. We apply Lemma 9.6.10 with $n' = n/k$, $X' = X^{(i)}$ and $z' = z^{(i)}$ to obtain that the event

$$E_1 = \Big\{ \forall j \in \{1, \dots, M\}, \quad \forall \theta \in \mathbb{R}^N, \quad \mathcal{L}^{(\setminus i)}_{\lambda_j}(\theta) \leq \min \mathcal{L}^{(\setminus i)}_{\lambda_j} + \gamma\epsilon \\ \Rightarrow \Big| \frac{k}{n} \|y^{(i)} - X^{(i)}\theta\|^2 - \frac{1}{n}\|\theta - \theta^\star\|^2 - \sigma^2 \Big| < \sqrt{\epsilon} k \Big\}$$

has probability at least $1 - M\frac{C}{\epsilon} e^{-c\epsilon^2 n}$. By Lemma 9.3.5 the event

$$E_2 = \Big\{ \forall \lambda, \lambda' \in [\lambda_{\min}, \lambda_{\max}], \ \mathcal{L}^{(\setminus i)}_{\lambda'}(\widehat{\theta}^i_\lambda) \leq \min_{x \in \mathbb{R}^N} \mathcal{L}^{(\setminus i)}_{\lambda'}(x) + KN^q |\lambda - \lambda'| \Big\} \tag{9.6.19}$$

has probability at least $1 - Ce^{-cn}$. On the event E_2, we have for all $j \in \{1, \dots, k\}$ and all $\lambda \in [\lambda_{j-1}, \lambda_j]$

$$\mathcal{L}^{(\setminus i)}_{\lambda_j}(\widehat{\theta}^i_\lambda) \leq \min_{x \in \mathbb{R}^N} \mathcal{L}^{(\setminus i)}_{\lambda_j}(x) + KN^q\eta \leq \min_{x \in \mathbb{R}^N} \mathcal{L}^{(\setminus i)}_{\lambda_j}(x) + \gamma\epsilon.$$

We obtain that on $E_1 \cap E_2$, which has probability at least $1 - CN^q\epsilon^{-2}e^{-cn\epsilon^2}$

$$\forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \Big| \frac{k}{n} \|y^{(i)} - X^{(i)}\widehat{\theta}^i_\lambda\|^2 - \frac{1}{n}\|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - \sigma^2 \Big| < \sqrt{\epsilon} k.$$

□

Proposition 9.6.5

There exists constants $c, C > 0$ that only depend on Ω, such that for all $\theta^\star \in \mathcal{D}$ and for all $i \in \{1, \dots, k\}$

$$\mathbb{P}\Big(\sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \Big| \frac{1}{N} \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - R_*(\lambda) \Big| \leq \frac{C}{\sqrt{k}} \Big) \leq CN^q k^4 e^{-cN/k^4}.$$

Proof. Let us fix $i \in \{1, \dots, k\}$. By Proposition 9.3.7, $\lambda \mapsto R_*(\lambda)$ is K_1-Lipschitz on $[\lambda_{\min}, \lambda_{\max}]$, for some constant $K_1 > 0$. By Propositions 9.3.3 and 9.3.4 there exists a constant $K_2 > 0$ such that the event

$$E_1 = \Big\{ \forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \frac{1}{N} \big| |\widehat{\theta}^i_\lambda| - |\theta^\star| \big| \leq K_2 N^q \quad \text{and} \quad \frac{1}{N} \big| |\widehat{\theta}_\lambda| - |\theta^\star| \big| \leq K_2 N^q \Big\} \tag{9.6.20}$$

has probability at least $1 - Ce^{-cn}$. Let us define $\eta = \min\left(\frac{\delta}{2N^q k K_2}, \frac{1}{K_1\sqrt{k}} \right)$ and $M = \lceil (\lambda_{\max} - \lambda_{\min})/\eta \rceil$. For all $j \in \{0, \dots, M\}$, we write $\lambda_j = \min\left(\lambda_{\min} + j\eta, \lambda_{\max}\right)$.

By Theorem 8.3.2 the event

$$E_2 = \Big\{\forall \lambda \in [\lambda_{\min}, \lambda_{\max}],\ \Big|\frac{1}{N}\|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - R^{(k)}_*(\lambda)\Big| \le 1 \text{ and } \Big|\frac{1}{n}\|y - X\widehat{\theta}_\lambda\|^2 - \beta_*(\lambda)\Big| \le 1\Big\} \quad (9.6.21)$$

has probability at least $1 - CN^q e^{-cN}$. By Lemma 9.6.11, applied with $\epsilon = k^{-1}$,

$$E_3 = \Big\{ \sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \Big|\frac{k}{n}\|y^{(i)} - X^{(i)}\widehat{\theta}^i_\lambda\|^2 - \frac{1}{n}\|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - \sigma^2\Big| \le 1\Big\}$$

has probability at least $1 - Ck^4 N^q e^{-cn/k^4}$. On the event $E_2 \cap E_3$, we have, for all $\lambda \in [\lambda_{\min}, \lambda_{\max}]$,

$$\begin{aligned}
\mathcal{L}_\lambda(\widehat{\theta}^i_\lambda) &= \frac{1}{2n}\|y^{(i)} - X^{(i)}\widehat{\theta}^i_\lambda\|^2 + \frac{1}{2n}\|y^{(-i)} - X^{(-i)}\widehat{\theta}^i_\lambda\|^2 + \frac{\lambda}{n}|\widehat{\theta}^i_\lambda| \\
&\le \frac{1}{k}(\sigma^2 + \frac{1}{n}\|\theta^\star - \widehat{\theta}^i_\lambda\|^2 + 1) + \frac{1}{2n_k}\|y^{(-i)} - X^{(-i)}\widehat{\theta}^i_\lambda\|^2 + \frac{\lambda}{n}|\widehat{\theta}^i_\lambda| \\
&\le \frac{1}{k}(1 + \sigma^2 + \delta^{-1}(R^{(k)}_*(\lambda) + 1)) + \frac{1}{2n_k}\|y^{(-i)} - X^{(-i)}\widehat{\theta}_\lambda\|^2 + \frac{\lambda}{n}|\widehat{\theta}_\lambda| \\
&\le \frac{1}{k}(1 + \sigma^2 + \delta^{-1}(R^{(k)}_*(\lambda) + 1)) + \mathcal{L}_\lambda(\widehat{\theta}_\lambda) + \Big(\frac{1}{2n_k} - \frac{1}{2n}\Big)\|y - X\widehat{\theta}_\lambda\|^2 \\
&\le \mathcal{L}_\lambda(\widehat{\theta}_\lambda) + \frac{C}{k}
\end{aligned}$$

for some constant $C > 0$. Let $j \in \{1, \dots, M\}$. We have $\mathcal{L}_\lambda(\widehat{\theta}^i_\lambda) = \mathcal{L}_{\lambda_j}(\widehat{\theta}^i_\lambda) - \frac{\lambda_j - \lambda}{n}|\widehat{\theta}^i_\lambda|$ and

$$\mathcal{L}_\lambda(\widehat{\theta}_\lambda) \le \mathcal{L}_\lambda(\widehat{\theta}_{\lambda_j}) = \mathcal{L}_{\lambda_j}(\widehat{\theta}_{\lambda_j}) + \frac{\lambda - \lambda_j}{n}|\widehat{\theta}_{\lambda_j}| = \min_{\theta \in \mathbb{R}^N} \mathcal{L}_{\lambda_j}(\theta) + \frac{\lambda - \lambda_j}{n}|\widehat{\theta}_{\lambda_j}|.$$

So we get that on the event $E_1 \cap E_2 \cap E_3$, for all $j \in \{1, \dots, M\}$ and all $\lambda \in [\lambda_{j-1}, \lambda_j]$,

$$\begin{aligned}
\mathcal{L}_{\lambda_j}(\widehat{\theta}^i_\lambda) &\le \min_{\theta \in \mathbb{R}^N} \mathcal{L}_{\lambda_j}(\theta) + \frac{\lambda_j - \lambda}{n}(|\widehat{\theta}^i_{\lambda_j}| - |\widehat{\theta}_{\lambda_j}|) + \frac{C}{k} \\
&\le \min_{\theta \in \mathbb{R}^N} \mathcal{L}_{\lambda_j}(\theta) + \frac{\lambda_j - \lambda}{\delta} 2K_2 N^q + \frac{C}{k} \\
&\le \min_{\theta \in \mathbb{R}^N} \mathcal{L}_{\lambda_j}(\theta) + \frac{\eta}{\delta} 2K_2 N^q + \frac{C}{k} \le \min_{\theta \in \mathbb{R}^N} \mathcal{L}_{\lambda_j}(\theta) + \frac{C_0}{k},
\end{aligned}$$

for some constant $C_0 > 0$, because on E_1 we have $\forall \lambda \in [\lambda_{\min}, \lambda_{\max}]$, $\frac{1}{N}||\widehat{\theta}^i_\lambda| - |\widehat{\theta}_\lambda|| \le 2K_2 N^q$. By Theorem 9.3.1, there exists constants $C, c, \gamma > 0$ such that for all $\epsilon \in (0, 1]$ the event

$$E_4 = \Big\{\forall j \in \{1 \dots M\},\ \forall \theta \in \mathbb{R}^N,\ \mathcal{L}_{\lambda_j}(\theta) \le \min \mathcal{L}_{\lambda_j} + \gamma\epsilon \Big|\ \frac{1}{N}\|\theta - \theta^\star\|^2 - R_*(\lambda_j)\Big| \le \sqrt{\epsilon}\Big\}$$

has probability at least $1 - CM\epsilon^{-1}e^{-cN\epsilon^2}$. Consider the constant $\kappa = \frac{C_0}{\gamma}$. If $k \ge \kappa$, then $\epsilon = \frac{C_0}{\gamma k} \le 1$ and the event E_4 has probability at least $1 - CMke^{-cN/k^2}$. So we obtain that on the event $E_1 \cap E_2 \cap E_3 \cap E_4$, which has probability $1 - CN^q k^4 e^{-cn/k^4}$,

$$\forall j \in \{1, \dots, M\}, \quad \forall \lambda \in [\lambda_{j-1}, \lambda_j], \quad \Big|\frac{1}{N}\|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - R_*(\lambda_j)\Big| \le \frac{C}{\sqrt{k}},$$

for some constant $C > 0$. If now $k < \kappa$. Then on the event E_2 we have $\forall j \in \{1, \dots, M\}, \quad \forall \lambda \in [\lambda_{j-1}, \lambda_j]$,

$$\begin{aligned}
\Big|\frac{1}{N}\|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - R_*(\lambda_j)\Big| &\le \sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} R^{(k)}_*(\lambda) + \sup_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} R_*(\lambda) + 1 \\
&\le C \le \frac{C\sqrt{\kappa}}{\sqrt{k}},
\end{aligned}$$

where C is a constant. We conclude that (in both cases) there exists a constant $C > 0$ such that

$$\forall j \in \{1, \dots, M\}, \quad \forall \lambda \in [\lambda_{j-1}, \lambda_j], \quad \left| \frac{1}{N} \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - R_*(\lambda_j) \right| \leq \frac{C}{\sqrt{k}},$$

holds with probability at least $1 - CN^q k^4 e^{-cN/k^4}$ Proposition 9.6.5 follows from the fact that for all $\lambda \in [\lambda_{j-1}, \lambda_j]$, $|R_*(\lambda) - R_*(\lambda_j)| \leq K_1 |\lambda - \lambda_j| \leq \frac{1}{\sqrt{k}}$. □

Proof of Proposition 8.4.3. We apply Lemma 9.6.11 with $\epsilon = k^{-3/2}$ to obtain that with probability at least $1 - Ck^6 N^q e^{-cn/k^6}$ we have

$$\forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \ \forall i \in \{1 \dots k\}, \ \left| \frac{k}{n} \|y^{(i)} - X^{(i)} \widehat{\theta}^i_\lambda\|^2 - \frac{1}{n} \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - \sigma^2 \right| \leq \frac{1}{\sqrt{k}}.$$

By summing these inequalities for $i = 1 \dots k$ and using the triangular inequality, we get

$$\left| \frac{k}{n} \sum_{i=1}^k \|y^{(i)} - X^{(i)} \widehat{\theta}^i_\lambda\|^2 - \frac{1}{n} \sum_{i=1}^k \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - k\sigma^2 \right| \leq \frac{k}{\sqrt{k}}.$$

and then

$$\left| \frac{1}{N} \sum_{i=1}^k \|y^{(i)} - X^{(i)} \widehat{\theta}^i_\lambda\|^2 - \frac{1}{k} \sum_{i=1}^k \frac{1}{N} \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - \delta\sigma^2 \right| \leq \frac{\delta}{\sqrt{k}}. \tag{9.6.22}$$

By Proposition 9.6.5, we have with probability at least $1 - CN^q k^4 e^{-cN/k^3}$,

$$\forall \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad \forall i \in \{1, \dots, k\}, \quad \left| \frac{1}{N} \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - R_*(\lambda) \right| \leq \frac{C}{\sqrt{k}}.$$

This implies (again by summing and using the triangular inequality) that

$$\left| \frac{1}{k} \sum_{i=1}^k \frac{1}{N} \|\widehat{\theta}^i_\lambda - \theta^\star\|^2 - R_*(\lambda) \right| \leq \frac{C}{\sqrt{k}},$$

which, combined with (9.6.22) proves Proposition 8.4.3. □

9.6.8 The scalar lasso

In this section we study

$$\ell_\alpha(y) = \min_{x \in \mathbb{R}} \left\{ \frac{1}{2}(y - x)^2 + \alpha |x| \right\}. \tag{9.6.23}$$

Lemma 9.6.12

The minimum (9.6.23) is achieved at an unique point $x^* = \eta(y, \alpha)$ and

$$\ell_\alpha(y) = \begin{cases} \frac{1}{2} y^2 & \text{if} \quad -\alpha \leq y \leq \alpha \\ \alpha y - \frac{1}{2}\alpha^2 & \text{if} \quad y \geq \alpha \\ -\alpha y - \frac{1}{2}\alpha^2 & \text{if} \quad y \leq -\alpha \end{cases}$$

Suppose now that

$$y = x + Z\,,$$

for some $x \in \mathbb{R}$ and $Z \sim \mathcal{N}(0, 1)$.

Lemma 9.6.13

Define

$$\Delta_\alpha(x) = \mathbb{E}\Big[\ell_\alpha(x+Z) - \alpha|x|\Big]\,.$$

The function Δ_α is continuous, even, decreasing on $\mathbb{R}_{\geq 0}$, α-Lipschitz. Moreover

$$\begin{cases} \Delta_\alpha(0) = \frac{1}{2} + \alpha\phi(\alpha) - (1+\alpha^2)\Phi(-\alpha) \\ \lim_{x\to\pm\infty} \Delta_\alpha(x) = -\frac{\alpha^2}{2} \end{cases}$$

and $\Delta'_\alpha(0^+) = -\Delta'_\alpha(0^-) = -\alpha$.

Proof. Since Z and $-Z$ have the same law, one verify easily that Δ_α is an even function. We have for all $x > 0$

$$\Delta'_\alpha(x) = \mathbb{E}\left[\ell'_\alpha(x+Z) - \alpha\right] = \mathbb{E}\left[\mathbb{1}(x+Z \in [-\alpha,\alpha])(x+Z-\alpha)\right] \leq 0\,.$$

ℓ_α is convex, therefore $x \mapsto \mathbb{E}[\ell'_\alpha(x+Z)]$ is non-decreasing. $\mathbb{E}[\ell'_\alpha(Z)] = 0$ because ℓ'_α is an odd function. Consequently, for all $x > 0$

$$-\alpha \leq \mathbb{E}[\ell'_\alpha(x+Z) - \alpha] = \Delta'_\alpha(x)\,.$$

This gives (recall that Δ_α is even and continuous over $\mathbb{R}$) that Δ_α is α-Lipschitz. From what we have seen above, we have also $\Delta'_\alpha(0^+) = -\Delta'_\alpha(0^-) = -\alpha$. Compute now, using the fact that ℓ_α is even:

$$\Delta_\alpha(0) = \mathbb{E}[\ell_\alpha(Z)] = \int_0^\alpha z^2\phi(z)dz + \int_\alpha^{+\infty}(2\alpha z - \alpha^2)\phi(z)dz\,.$$

By integration by parts

$$\int_0^\alpha z^2\phi(z)dz = \Big[-z\phi(z)\Big]_0^\alpha + \int_0^\alpha \phi(z)dz = -\alpha\phi(\alpha) + \frac{1}{2} - \Phi(-\alpha)$$

$$\int_\alpha^{+\infty}(2\alpha z - \alpha^2)\phi(z)dz = -\alpha^2\Phi(-\alpha) + 2\alpha\phi(\alpha)\,.$$

Therefore $\Delta_\alpha(0) = \frac{1}{2} + \alpha\phi(\alpha) - (1+\alpha^2)\Phi(-\alpha)$. We have almost surely

$$\ell_\alpha(x+Z) - \alpha|x| \xrightarrow[x\to\pm\infty]{} -\frac{\alpha^2}{2}\,.$$

Thus, by dominated convergence $\lim_{x\to\pm\infty}\Delta_\alpha(x) = -\frac{\alpha^2}{2}$. □

9.6.9 A convexity lemma

Lemma 9.6.14

The function

$$f : x \in \mathbb{R}^N \mapsto \sqrt{\frac{\|x\|^2}{n} + \sigma^2}$$

is $\frac{\sigma^2}{n(R^2+\sigma^2)^{3/2}}$-strongly convex on $B(0,\sqrt{n}R)$.

Proof. Let $x, y \in B(0, \sqrt{n}R)$ and define for $t \in [0,1]$, $g(t) = f(z_t)$, where $z_t = (tx + (1-t)y)$. Compute

$$g'(t) = \frac{\frac{1}{n}(x-y)^\mathsf{T} z_t}{\sqrt{\frac{\|z_t\|^2}{n} + \sigma^2}}\,,$$

and

$$\begin{aligned} g''(t) &= \frac{\frac{1}{n}\|x-y\|^2}{\sqrt{\frac{\|z_t\|^2}{n} + \sigma^2}} - \frac{\left(\frac{1}{n}(x-y)^\mathsf{T} z_t\right)^2}{\left(\frac{\|z_t\|^2}{n} + \sigma^2\right)^{3/2}} \\ &= \frac{1}{\left(\frac{\|z_t\|^2}{n} + \sigma^2\right)^{3/2}} \left(\frac{1}{n}\|x-y\|^2 \Big(\frac{\|z_t\|^2}{n} + \sigma^2 \Big) - \Big(\frac{1}{n}(x-y)^\mathsf{T} z_t \Big)^2 \right) \\ &\geq \frac{\sigma^2}{\left(\frac{\|z_t\|^2}{n} + \sigma^2\right)^{3/2}} \left(\frac{1}{n}\|x-y\|^2 \right) \geq \frac{1}{n}\|x-y\|^2 \frac{\sigma^2}{(R^2+\sigma^2)^{3/2}}\,. \end{aligned}$$

Consequently

$$\begin{aligned} tf(x) + (1-t)f(y) &= tg(1) + (1-t)g(0) \\ &\geq g(t) + \frac{1}{2}t(1-t)\frac{1}{n}\|x-y\|^2 \frac{\sigma^2}{(R^2+\sigma^2)^{3/2}} \\ &= f(tx + (1-t)y) + \frac{1}{2}t(1-t)\frac{1}{n}\|x-y\|^2 \frac{\sigma^2}{(R^2+\sigma^2)^{3/2}}\,. \end{aligned}$$

□

9.7 Toolbox

9.7.1 Notations recap

Recall that X is a $n \times N$ random matrix with entries $X_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/n)$. The random vectors $z \in \mathbb{R}^n$, $g \in \mathbb{R}^N$ and $h \in \mathbb{R}^n$ are standard Gaussian random vectors. The following table displays the main cost (or objective) functions used in this paper and their corresponding optimizers.

Definition	Optimizer
$\mathcal{L}_\lambda(\theta) = \frac{1}{2n}\left\|X\theta - y\right\|^2 + \frac{\lambda}{n}\lvert\theta\rvert$	$\widehat{\theta}_\lambda$
$\mathcal{C}_\lambda(w) = \frac{1}{2n}\|Xw - \sigma z\|^2 + \frac{\lambda}{n}(\lvert w+\theta^\star\rvert - \lvert\theta^\star\rvert)$	$\widehat{w}_\lambda$
$\mathcal{U}_\lambda(u) = \min_{w\in\mathbb{R}^N} \left\{ u^\mathsf{T} X w - \sigma u^\mathsf{T} z - \frac{1}{2}\|u\|^2 + \lambda(\lvert\theta^\star + w\rvert - \lvert\theta^\star\rvert) \right\}$	$\widehat{u}_\lambda$
$\mathcal{V}_\lambda(v) = \min_{w\in B} \left\{ \frac{1}{2n}\|Xw - \sigma z\|^2 + \frac{\lambda}{n} v^\mathsf{T}(\theta^\star + w) - \frac{\lambda}{n}\lvert\theta^\star\rvert \right\}$	$\widehat{v}_\lambda$
$L_\lambda(w) = \frac{1}{2}\left(\sqrt{\frac{\|w\|^2}{n} + \sigma^2}\,\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right)_+^2 + \frac{\lambda}{n}\lvert w+\theta^\star\rvert - \frac{\lambda}{n}\lvert\theta^\star\rvert$	w^*_λ
$U_\lambda(u) = \min_{w\in\mathbb{R}^N} \left\{ \frac{-1}{n^{3/2}}\|u\| g^\mathsf{T} w + \frac{1}{n^{3/2}}\|w\| h^\mathsf{T} u - \frac{\sigma}{n} u^\mathsf{T} z - \frac{1}{2n}\|u\|^2 + \frac{\lambda}{n}\Big(\lvert w+\theta^\star\rvert - \lvert\theta^\star\rvert \Big) \right\}$	u^*_λ
$V_\lambda(v) = \min_{w\in B} \left\{ \frac{1}{2}\left(\sqrt{\frac{1}{n}\|w\|^2 + \sigma^2}\,\frac{\|h\|}{\sqrt{n}} - \frac{1}{n} g^\mathsf{T} w + \frac{g'\sigma}{\sqrt{n}} \right)_+^2 + \frac{\lambda}{n} v^\mathsf{T}(w+\theta^\star) - \frac{\lambda}{n}\lvert\theta^\star\rvert \right\}$	v^*_λ

Table 9.1: Main cost/objective functions

In the definition of V_λ above, $B = \Big\{w \in \mathbb{R}^N \,|\, |w| \leq 2|\theta^\star| + 5\sigma^2\lambda_{\min}^{-1}n + K\Big\}$, where $K > 0$ is the constant given by Lemma 9.5.5. The functions L_λ, U_λ and V_λ are the "corresponding cost/objective functions" to $\mathcal{C}_\lambda$, $\mathcal{U}_\lambda$ and $\mathcal{V}_\lambda$. A main part of the analysis is to show that w_λ^*, u_λ^* and v_λ^* are approximately equal to w_λ, u_λ and v_λ given by:

$$\begin{aligned}
\mathsf{w}_\lambda &= \eta\Big(\theta^\star + \tau_*(\lambda)g, \alpha_*(\lambda)\tau_*(\lambda)\Big) - \theta^\star \\
\mathsf{u}_\lambda &= \tfrac{\beta_*(\lambda)}{\tau_*(\lambda)}\left(\sqrt{\tau_*(\lambda)^2 - \sigma^2}\,\tfrac{h}{\sqrt{n}} - \tfrac{\sigma}{\sqrt{n}}z\right) \\
\mathsf{v}_\lambda &= -\alpha_*(\lambda)^{-1}\tau_*(\lambda)^{-1}\Big(\eta\Big(\theta^\star + \tau_*(\lambda)g, \alpha_*(\lambda)\tau_*(\lambda)\Big) - \theta^\star - \tau_*(\lambda)g\Big)
\end{aligned}$$

Table 9.2: "Asymptotic optimizers"

9.7.2 Gaussian min-max Theorem

In this section, we reproduce the proof of the tight Gaussian min-max comparison theorem from [205] for completeness, but also because we need a slightly more general version of this result.

We recall the classical Gordon's min-max Theorem from [93] (see also Corollary 3.13 from [130]):

Theorem 9.7.1

Let $X_{i,j}$ and $(Y_{i,j})$, $1 \leq i \leq n$, $1 \leq j \leq m$ be two (centered) Gaussian random vectors such that

$$\begin{cases}
\mathbb{E}X_{i,j}^2 = \mathbb{E}Y_{i,j}^2 & \text{for all } i, j\,, \\
\mathbb{E}X_{i,j}X_{i,k} \geq \mathbb{E}Y_{i,j}Y_{i,k} & \text{for all } i, j, k\,, \\
\mathbb{E}X_{i,j}X_{l,k} \leq \mathbb{E}Y_{i,j}Y_{l,k} & \text{for all } i \neq l \text{ and } j, k\,.
\end{cases}$$

Then, for all real numbers $\lambda_{i,j}$:

$$\mathbb{P}\left(\bigcap_{i=1}^{n}\bigcup_{j=1}^{m}\Big\{X_{i,j} > \lambda_{i,j}\Big\}\right) \leq \mathbb{P}\left(\bigcap_{i=1}^{n}\bigcup_{j=1}^{m}\Big\{Y_{i,j} > \lambda_{i,j}\Big\}\right)\,.$$

Theorem 9.7.2

Let $D_u \subset \mathbb{R}^n$ and $D_v \subset \mathbb{R}^m$ be two compact sets. Let $Q : D_u \times D_v \to \mathbb{R}$ be a continuous function. Let $\big(X(u,v)\big)_{(u,v)\in D_u\times D_v}$ and $\big(Y(u,v)\big)_{(u,v)\in D_u\times D_v}$ be two centered Gaussian processes. Suppose that the functions

$$(u,v) \mapsto X(u,v) \qquad \text{and} \qquad (u,v) \mapsto Y(u,v)$$

are continuous on $D_u \times D_v$ almost surely. Assume that

$$\begin{array}{lll} \mathbb{E}\Big[X(u,v)^2\Big] = \mathbb{E}\Big[Y(u,v)^2\Big] & \text{for all} & (u,v) \in D_u \times D_v\,, \\ \mathbb{E}\Big[X(u,v)X(u,v')\Big] \geq \mathbb{E}\Big[Y(u,v)Y(u,v')\Big] & \text{for all} & u \in D_u, v, v' \in D_v\,, \\ \mathbb{E}\Big[X(u,v)X(u',v')\Big] \leq \mathbb{E}\Big[Y(u,v)Y(u',v')\Big] & \text{for all} & u, u' \in D_u, v, v' \in D_v \\ & \text{such that} & u \neq u'\,. \end{array}$$

Then for all $t \in \mathbb{R}$

$$\mathbb{P}\Big(\min_{u\in D_u}\max_{v\in D_v} Y(u,v) + Q(u,v) \leq t\Big) \leq \mathbb{P}\Big(\min_{u\in D_u}\max_{v\in D_v} X(u,v) + Q(u,v) \leq t\Big)\,.$$

Proof. Define the random variable

$$d_0 = \sup\Big\{d \in \mathbb{Q}_+ \,\Big|\, \forall (z,z') \in (D_u \times D_v)^2, \\ \|z - z'\| \leq d \Rightarrow \big(|X(z) - X(z')| \leq \epsilon \text{ and } |Y(z) - Y(z')| \leq \epsilon\big)\Big\}\,.$$

X and Y are continuous on the compact set $D_u \times D_v$ and are therefore uniformly continuous on this set: $d_0 > 0$ almost surely. Let $\epsilon > 0$. By tightness there exists a constant $d > 0$ such that

$$\mathbb{P}(d_0 \geq d) \geq 1 - \epsilon\,.$$

Q is continuous and thus uniformly continuous on $D_u \times D_v$: there exists $\delta \in (0, d]$ such that for all $z, z' \in D_u \times D_v$, $\|z - z'\| \leq \delta \implies |Q(z) - Q(z')| \leq \epsilon$.

Let D_u^δ (respectively D_v^δ) be a $\delta/\sqrt{2}$-net of D_u (respectively D_v). $D_u^\delta \times D_v^\delta$ is thus a δ-net of $D_u \times D_v$. By Theorem 9.7.1 we have for all $t \in \mathbb{R}$

$$\mathbb{P}\Big(\min_{u\in D_u^\delta}\max_{v\in D_v^\delta} X(u,v) + Q(u,v) > t\Big) \leq \mathbb{P}\Big(\min_{u\in D_u^\delta}\max_{v\in D_v^\delta} Y(u,v) + Q(u,v) > t\Big)\,,$$

which gives by taking the complementary:

$$\mathbb{P}\Big(\min_{u\in D_u^\delta}\max_{v\in D_v^\delta} Y(u,v) + Q(u,v) \leq t\Big) \leq \mathbb{P}\Big(\min_{u\in D_u^\delta}\max_{v\in D_v^\delta} X(u,v) + Q(u,v) \leq t\Big)\,.$$

By construction of δ we have with probability at least $1 - \epsilon$

$$\Big|\min_{u\in D_u^\delta}\max_{v\in D_v^\delta} X(u,v) + Q(u,v) - \min_{u\in D_u}\max_{v\in D_v} X(u,v) + Q(u,v)\Big| \leq 2\epsilon\,,$$

and similarly for Y. We have therefore, for all $t \in \mathbb{R}$

$$\mathbb{P}\Big(\min_{u\in D_u}\max_{v\in D_v} Y(u,v) + Q(u,v) \leq t - 2\epsilon\Big) - \epsilon \leq \mathbb{P}\Big(\min_{u\in D_u}\max_{v\in D_v} X(u,v) + Q(u,v) \leq t + 2\epsilon\Big) + \epsilon\,,$$

and thus

$$\mathbb{P}\Big(\min_{u\in D_u}\max_{v\in D_v} Y(u,v) + Q(u,v) \leq t\Big) \leq \mathbb{P}\Big(\min_{u\in D_u}\max_{v\in D_v} X(u,v) + Q(u,v) \leq t + 4\epsilon\Big) + 2\epsilon\,,$$

which proves the theorem by taking $\epsilon \to 0$.

□

Corollary 9.7.1

Let $D_u \subset \mathbb{R}^{n_1+n_2}$ and $D_v \subset \mathbb{R}^{m_1+m_2}$ be compact sets and let $Q : D_u \times D_v \to \mathbb{R}$ be a continuous function. Let $G = (G_{i,j}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, $g \sim \mathcal{N}(0, \mathrm{Id}_{n_1})$ and $h \sim \mathcal{N}(0, \mathrm{Id}_{m_1})$ be independent standard Gaussian vectors. For $u \in \mathbb{R}^{n_1+n_2}$ and $v \in \mathbb{R}^{m_1+m_2}$ we define $\tilde{u} = (u_1, \dots, u_{n_1})$ and $\tilde{v} = (v_1, \dots, v_{m_1})$. Define

$$\begin{cases} C^*(G) = \min\limits_{u \in D_u} \max\limits_{v \in D_v} \tilde{v}^\mathsf{T} G \tilde{u} + Q(u,v)\,, \\ L^*(g,h) = \min\limits_{u \in D_u} \max\limits_{v \in D_v} \|\tilde{v}\| g^\mathsf{T} \tilde{u} + \|\tilde{u}\| h^\mathsf{T} \tilde{v} + Q(u,v)\,. \end{cases}$$

Then we have:

- *For all $t \in \mathbb{R}$,*
$$\mathbb{P}\Big(C^*(G) \le t\Big) \le 2\mathbb{P}\Big(L^*(g,h) \le t\Big)\,.$$
- *If D_u and D_v are convex and if Q is convex concave, then for all $t \in \mathbb{R}$*
$$\mathbb{P}\Big(C^*(G) \ge t\Big) \le 2\mathbb{P}\Big(L^*(g,h) \ge t\Big)\,.$$

Proof. Let us consider the Gaussian processes:

$$\begin{cases} X(u,v) = \|\tilde{v}\| g^\mathsf{T} \tilde{u} + \|\tilde{u}\| h^\mathsf{T} \tilde{v}\,, \\ Y(u,v) = \tilde{v}^\mathsf{T} G \tilde{u} + \|\tilde{u}\| \|\tilde{v}\| z\,, \end{cases}$$

where $z \sim \mathcal{N}(0,1)$ is independent from G. Let $(u,v), (u',v') \in D_u \times D_v$ and compute

$$\begin{aligned} \mathbb{E}[Y(u,v)Y(u',v')] &- \mathbb{E}[X(u,v)X(u',v')] \\ &= \|\tilde{u}\|\|\tilde{v}\|\|\tilde{u}'\|\|\tilde{v}'\| + (\tilde{u}^\mathsf{T}\tilde{u}')(\tilde{v}^\mathsf{T}\tilde{v}') - \|\tilde{v}\|\|\tilde{v}'\|(\tilde{u}^\mathsf{T}\tilde{u}') - \|\tilde{u}\|\|\tilde{u}'\|(\tilde{v}^\mathsf{T}\tilde{v}') \\ &= \big(\|\tilde{u}\|\|\tilde{u}'\| - (\tilde{u}^\mathsf{T}\tilde{u}')\big)\big(\|\tilde{v}\|\|\tilde{v}'\| - (\tilde{v}^\mathsf{T}\tilde{v}')\big) \ge 0\,. \end{aligned}$$

Therefore X and Y verify the covariance inequalities of Theorem 9.7.2: one can apply Theorem 9.7.2:

$$\mathbb{P}\Big(\min_{u \in D_u} \max_{v \in D_v} Y(u,v) + Q(u,v) \le t\Big) \le \mathbb{P}\Big(\min_{u \in D_u} \max_{v \in D_v} Y(u,v) + Q(u,v) \le t\Big)\,,$$

We have then

$$\begin{aligned} \mathbb{P}\Big(\min_{u \in D_u} \max_{v \in D_v} Y(u,v) + Q(u,v) \le t\Big) &\ge \frac{1}{2}\mathbb{P}\Big(\min_{u \in D_u} \max_{v \in D_v} Y(u,v) + Q(u,v) \le t \,\Big|\, z \le 0\Big) \\ &\ge \frac{1}{2}\mathbb{P}\Big(\min_{u \in D_u} \max_{v \in D_v} \tilde{v}^\mathsf{T} G \tilde{u} + Q(u,v) \le t \,\Big|\, z \le 0\Big) \\ &= \frac{1}{2}\mathbb{P}\Big(C^*(G) \le t\Big)\,, \end{aligned}$$

which proves that

$$\mathbb{P}\Big(\min_{u \in D_u} \max_{v \in D_v} \tilde{v}^\mathsf{T} G \tilde{u} + Q(u,v) \le t\Big) \le 2\mathbb{P}\Big(\min_{u \in D_u} \max_{v \in D_v} \|\tilde{v}\| g^\mathsf{T} \tilde{u} + \|\tilde{u}\| h^\mathsf{T} \tilde{v} + Q(u,v) \le t\Big)\,.$$

Let us suppose now that D_u and D_v are convex and that G is convex-concave. We now apply the inequality we just proved, but with the role of u and v being switched (and $-Q$ and $-t$ instead of Q and t):

$$\mathbb{P}\Big(\min_{v \in D_v} \max_{u \in D_u} \tilde{v}^\mathsf{T} G \tilde{u} - Q(u,v) \le -t\Big) \le 2\mathbb{P}\Big(\min_{v \in D_v} \max_{u \in D_u} \|\tilde{v}\| g^\mathsf{T} \tilde{u} + \|\tilde{u}\| h^\mathsf{T} \tilde{v} - Q(u,v) \le -t\Big)\,,$$

which gives (using the fact that (G, g, h) and $(-G, -g, -h)$ have the same law):

$$\mathbb{P}\Big(\max_{v\in D_v}\min_{u\in D_u}\tilde{v}^\mathsf{T}G\tilde{u}+Q(u,v)\geq t\Big)\leq 2\mathbb{P}\Big(\max_{v\in D_v}\min_{u\in D_u}\|\tilde{v}\|g^\mathsf{T}\tilde{u}+\|\tilde{u}\|h^\mathsf{T}\tilde{v}+Q(u,v)\geq t\Big)\,.$$

By Proposition C.2, one can switch the min-max of the left-hand side, because Q is convex-concave and we are working on convex sets D_u and D_v. For the right-hand side, we simply use the fact that:

$$\max_{v\in D_v}\min_{u\in D_u}\|\tilde{v}\|g^\mathsf{T}\tilde{u}+\|\tilde{u}\|h^\mathsf{T}\tilde{v}+Q(u,v)\leq\min_{u\in D_u}\max_{v\in D_v}\|\tilde{v}\|g^\mathsf{T}\tilde{u}+\|\tilde{u}\|h^\mathsf{T}\tilde{v}+Q(u,v)\,,$$

to conclude the proof. □

9.7.3 Basic concentration results

We recall in this section some elementary concentration results, see Chapter 2 from [37] for a more detailed presentation of these facts.

Definition 9.7.1

A real random variable X is said to be

- *σ^2-sub-Gaussian if for every $s\in\mathbb{R}$, $\log\mathbb{E}e^{s(X-\mathbb{E}[X])}\leq\dfrac{s^2\sigma^2}{2}$,*
- *(v,c)-sub-Gamma if for every $s\in(-1/c,1/c)$, $\log\mathbb{E}e^{s(X-\mathbb{E}[X])}\leq\dfrac{s^2v}{2(1-c|s|)}$.*

One deduces immediately from the above definition:

Proposition 9.7.1

Let $(X_1,\dots,X_n)$ be independent real random variables. Define $S=\sum_{i=1}^n X_i$.

- *Suppose that for all $i\in\{1,\dots,n\}$, X_i is σ_i^2-sub-Gaussian. Then S is $\sum_{i=1}^n\sigma_i^2$-sub-Gaussian.*
- *Suppose that for all $i\in\{1,\dots,n\}$, X_i is (v_i,c_i)-sub-Gamma. Then S is $\Big(\sum_{i=1}^n v_i,\max c_i\Big)$-sub-Gamma.*

Proposition 9.7.2

Let X be a real random variable.

- *if X is σ^2-sub-Gaussian, then for all $t>0$*

$$\mathbb{P}(X-\mathbb{E}[X]\geq t)\ \vee\ \mathbb{P}(X-\mathbb{E}[X]\leq -t)\leq e^{-\frac{t^2}{2\sigma^2}}\,,$$

- *if X is (v,c)-sub-Gamma, then for all $t>0$*

$$\mathbb{P}(X-\mathbb{E}[X]\geq\sqrt{2ct}+vt)\ \vee\ \mathbb{P}(X-\mathbb{E}[X]\leq-(\sqrt{2ct}+vt))\leq e^{-t}\,.$$

Remark 9.7.1. *The bound* $\mathbb{P}(X > \sqrt{2vt} + ct) \leq e^{-t}$ *implies that*

$$\mathbb{P}(X > t) \leq \begin{cases} \exp\left(-\frac{t^2}{8v}\right) & \text{for} \quad 0 < t \leq \frac{2v}{c^2}\,, \\ \exp\left(-\frac{t}{2c}\right) & \text{for} \quad t \geq \frac{2v}{c^2}\,. \end{cases}$$

Proposition 9.7.3

If X *is* σ^2*-sub-Gaussian and has mean* μ*, then* X^2 *is a sub-Gamma random variable with parameters*

$$\begin{cases} v = 16\sigma^2 + 4\mu^2\sigma^2\,, \\ c = 4\sigma^2\,. \end{cases}$$

Proof. Let $\mu = \mathbb{E}[X]$ and $Y = X - \mu$. $X^2 = Y^2 + 2\mu Y + \mu^2$.

$$\mathbb{E}\left[(Y^2)^2\right] = \mathbb{E}[(X-\mu)^4] \leq 16\sigma^4\,,$$

$$\mathbb{E}\left[(Y^2)^q\right] = \mathbb{E}\left[(X-\mu)^{2q}\right] \leq 2q!(2\sigma^2)^q = \frac{1}{2}q!(16\sigma^2)(2\sigma^2)^{q-2}\,.$$

By Bernstein's inequality (see for instance Theorem 2.10 in [37])

$$\log \mathbb{E}e^{s(Y^2 - \mathbb{E}[Y^2])} \leq \frac{16\sigma^2 s^2}{2(1 - 2\sigma^2|s|)}\,.$$

$2\mu Y$ is $4\mu^2\sigma^2$-sub-Gaussian, therefore $\log \mathbb{E}e^{2\mu s Y} \leq 2\mu^2\sigma^2 s^2$ and

$$\begin{aligned} \log \mathbb{E}e^{s(X^2 - \mathbb{E}[X^2])} = \log \mathbb{E}e^{s(Y^2 - \mathbb{E}[Y^2]) + 2\mu s Y} &\leq \frac{1}{2}\log \mathbb{E}e^{2s(Y^2 - \mathbb{E}[Y^2])} + \frac{1}{2}\log \mathbb{E}e^{4\mu s Y} \\ &\leq \frac{16\sigma^2 s^2}{1 - 4\sigma^2 s} + 4\mu^2\sigma^2 s^2 \leq \frac{(16\sigma^2 + 4\mu^2\sigma^2)s^2}{1 - 4\sigma^2|s|}\,. \end{aligned}$$

X^2 is therefore a Sub-Gamma random variable with variance factor $v = 16\sigma^2 + 4\mu^2\sigma^2$ and scale parameter $c = 4\sigma^2$. □

Lemma 9.7.1

Let X *be a* σ^2*-sub-Gaussian random variable. Define* $m = \mathbb{E}[|X|]$*. Let* Y *be a random variable bounded by* 1*. Then* XY *is* $16(m^2 + 2\sigma^2)$*-sub-Gaussian.*

Proof. We have $|\mathbb{E}[XY]| \leq m$, therefore

$$\begin{aligned} \mathbb{E}\left[(XY - \mathbb{E}[XY])^{2q}\right] &\leq 2^{2q-1}\mathbb{E}[X^{2q}] + 2^{2q-1}m^{2q} \leq q!(8\sigma^2)^q + q!(4m^2)^q \\ &\leq q!(8\sigma^2 + 4m^2)^q\,. \end{aligned}$$

□

9.7.4 Largest singular value of a Gaussian matrix

The largest singular value of a $n \times N$ matrix A is defined as

$$\sigma_{\max}(A) = \max_{\|x\| \leq 1} \|Ax\|\,.$$

The next classical result is a simple consequence of Slepian's Lemma (see for instance [130], Section 3.3) and the classical Gaussian concentration inequality (see for instance [37], Theorem 5.6).

Proposition 9.7.4

Let G be a $n \times N$ random matrix, whose entries are i.i.d. $\mathcal{N}(0,1)$. For all $t \geq 0$ we have

$$\mathbb{P}(\sigma_{\max}(G) > \sqrt{N} + \sqrt{n} + t) \leq e^{-t^2/2}\,.$$

Appendix

A Proof of Lemma 1.2.2

In order to prove Lemma 1.2.2, we first need to introduce some definitions and results about exchangeable infinite arrays.

Definition A.1

An infinite symmetric random array R is a collection of random variables $(R_{k,k'})_{k,k'\geq 1}$ such that for all $k, k' \geq 1$, $R_{k,k'} = R_{k',k}$ almost surely. We say that R is

- *weakly exchangeable if for all $n \geq 1$ and for all permutation σ of $\{1, \dots, n\}$ we have $(R_{k,k'})_{k,k'\geq 1} \overset{(\mathrm{d})}{=} (R_{\sigma(k),\sigma(k')})_{k,k'\geq 1}$.*
- *positive semi-definite if for all $n \geq 1$ the matrix $R_{|n} \overset{\text{def}}{=} (R_{i,j})_{1\leq i,j\leq n}$ is positive semi-definite with probability one, i.e. $\forall x \in \mathbb{R}^n, x^\intercal R_{|n} x \geq 0$.*

On the one hand, the Aldous-Hoover Theorem [4, 104] states that an infinite weakly exchangeable array is equal in distribution to $(f(w, u_k, v_{k'}, x_{k,k'}))_{k,k'\geq 1}$, for some function f and $w, (u_k), (v_{k'}), (x_{k,k'}) \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}([0,1])$.

On the other hand, if $(R_{k,k'})_{k,k'\geq 1}$ is a deterministic positive semi-definite array, then there exists a separable Hilbert space H, with scalar product $(\cdot;\cdot)$ and $(h_k)_{k\geq 1} \in H^{\mathbb{N}}$ such that for all $k, k' \geq 1$, $R_{k,k'} = (h_k; h_{k'})$.

The Dovbysh-Sudakov Theorem [75] combines somehow these two results:

Theorem A.1 (*Dovbysh-Sudakov*)

Let R be an infinite symmetric random array R which is weakly exchangeable and positive semi-definite. Then there exists a separable Hilbert space H (whose scalar product will be denoted by $(\cdot\,;\,\cdot)$) and a random probability distribution η on $H \times \mathbb{R}_{\geq 0}$ such that R is equal in distribution to

$$\Big((h_k; h_{k'}) + a_k \delta_{k,k'}\Big)_{k,k'\geq 1},$$

where, conditionaly on η, $(h_k, a_k)_{k\geq 1}$ is a sequence of i.i.d. random variables with distribution η. $\delta_{k,k'}$ denotes here the Kronecker delta.

We refer to [171] and [12] for a proof. From the Dovbysh-Sudakov Theorem, one deduces easily:

Proposition A.1

Let $q \in [0,1]$. Let $(R_{k,k'})_{k,k'\geq 1}$ be a random symmetric array, which is weakly exchangeable and positive semi-definite. Assume that we have almost surely

$$R_{1,1} = 1 \qquad \text{and} \qquad |R_{1,2}| = q.$$

Then we have almost surely $R_{1,2}R_{2,3}R_{3,1} = q^3$.

Proof. We apply the Dovbysh-Sudakov Theorem to R. We obtain the existence of a random probability measure η on $H \times \mathbb{R}_{\geq 0}$ such that R is equal in distribution to

$$\big((h_k; h_{k'}) + a_k \delta_{k,k'}\big)_{k,k'\geq 1},$$

where conditionally on η, $(h_k, a_k) \overset{\text{i.i.d.}}{\sim} \eta$. Let us work conditionally on η and let ν denotes the first marginal of η. From our hypothesis we get that $|(h_1; h_2)| = q$, ν-almost surely which implies that $\|h_1\|^2 = q$, ν-almost surely.

Indeed, if $\nu(\|h_1\|^2 \neq q) > 0$ then there exists $h \in H$ such that $\|h\|^2 \neq q$ such that for all $\epsilon > 0$, $\nu(B_\epsilon(h) > 0$. We can now take $\epsilon > 0$ small enough such that for all $a, b \in B_\epsilon(h)$ we have $(a; b) \neq q$. We obtain that $\nu\big(|(h_1; h_2)| \neq q\big) \geq \nu(B_\epsilon(h))^2 > 0$ which leads to a contradiction.

Since $\|h_1\|^2 = \|h_2\|^2 = |(h_1; h_2)| = q$ a.s. the measure ν has at most two points in its support. Consequently

$$(h_1; h_2)(h_2; h_3)(h_3; h_1) = q^3, \qquad \nu\text{-almost surely},$$

which implies that $R_{1,2}R_{2,3}R_{3,1} = q^3$ almost surely. □

We have now all the tools needed to prove Lemma 1.2.2:

Proof of Lemma 1.2.2. The sequence $\big((\boldsymbol{x}^{(1)}; \boldsymbol{x}^{(2)})(\boldsymbol{x}^{(2)}; \boldsymbol{x}^{(3)})(\boldsymbol{x}^{(3)}; \boldsymbol{x}^{(1)})\big)_{n\geq 1}$ is tight because bounded. Let us consider a subsequence $(n_\ell)_{\ell\geq 1}$ along which is converges in law, to some random variable Q.

Let us sample $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{k}, \ldots$ i.i.d. from the posterior distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$ and consider the following random symmetric weakly exchangeable positive semi-definite array

$$R^n = \big((\boldsymbol{x}^{(k)}; \boldsymbol{x}^{(k')})\big)_{k,k'\geq 1}.$$

The scalar products $(\boldsymbol{x}^{(k)}; \boldsymbol{x}^{(k')})$ are bounded, so the sequence of the laws of R^n is tight. We can extract from $(n_\ell)_{\ell\geq 1}$ another subsequence $(n'_\ell)_{\ell\geq 1}$ along which R converges in law to an array R. R is also a random, symmetric, weakly exchangeable, semi-definite positive array. Since $R^n_{1,1} = 1$ and $R^n_{1,2} = q$ almost surely, we have $R_{1,1} = 1$ and $R_{1,2} = q$ almost surely. Proposition A.1 above gives then that $R_{1,2}R_{2,3}R_{3,1} = q^3$ almost surely. The weak convergence $R^n \xrightarrow[n'_\ell\to\infty]{\text{(d)}} R$ implies in particular

$$(\boldsymbol{x}^{(1)}; \boldsymbol{x}^{(2)})(\boldsymbol{x}^{(2)}; \boldsymbol{x}^{(3)})(\boldsymbol{x}^{(3)}; \boldsymbol{x}^{(1)}) = R^n_{1,2}R^n_{2,3}R^n_{3,1} \xrightarrow[n'_\ell\to\infty]{\text{(d)}} R_{1,2}R_{2,3}R_{3,1} = q^3.$$

Hence the only accumulation point of $\big((\boldsymbol{x}^{(1)}; \boldsymbol{x}^{(2)})(\boldsymbol{x}^{(2)}; \boldsymbol{x}^{(3)})(\boldsymbol{x}^{(3)}; \boldsymbol{x}^{(1)})\big)_{n\geq 1}$ is q^3: we conclude that $(\boldsymbol{x}^{(1)}; \boldsymbol{x}^{(2)})(\boldsymbol{x}^{(2)}; \boldsymbol{x}^{(3)})(\boldsymbol{x}^{(3)}; \boldsymbol{x}^{(1)}) \xrightarrow[n\to\infty]{\text{(d)}} q^3$. □

B Proofs of some basic properties of the MMSE and the free energy

B.1 Proof of Proposition 1.3.1

Let $0 < \lambda_2 < \lambda_1$. Define $\Delta_1 = \lambda_1^{-1}$, $\Delta_2 = \lambda_2^{-1}$ and

$$\begin{cases} \boldsymbol{Y}_1 = \boldsymbol{X} + \sqrt{\Delta_1}\boldsymbol{Z}_1 \\ \boldsymbol{Y}_2 = \boldsymbol{X} + \sqrt{\Delta_1}\boldsymbol{Z}_1 + \sqrt{\Delta_2 - \Delta_1}\boldsymbol{Z}_2\,, \end{cases}$$

where $\boldsymbol{X} \sim P_X$ is independent from $\boldsymbol{Z}_1, \boldsymbol{Z}_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathrm{Id}_n)$. Now, by independence between $(\boldsymbol{X}, \boldsymbol{Y}_1)$ and $\boldsymbol{Z}_2$ we have

$$\begin{aligned} \mathrm{MMSE}(\lambda_1) &= \mathbb{E}\left\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}_1]\right\|^2 = \mathbb{E}\left\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}_1, \boldsymbol{Z}_2]\right\|^2 = \mathbb{E}\left\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}_1, \boldsymbol{Y}_2]\right\|^2 \\ &\le \mathbb{E}\left\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}_2]\right\|^2 = \mathrm{MMSE}(\lambda_2)\,. \end{aligned}$$

Next, notice that

$$\mathrm{MMSE}(\lambda_1) = \mathbb{E}\left\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}_1]\right\|^2 \le \mathbb{E}\left\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]\right\|^2 = \mathrm{MMSE}(0)\,. \tag{B.1}$$

This shows that the MMSE is non-increasing on $\mathbb{R}_{\ge 0}$. The first point is obvious while the second follows from:

$$0 \le \mathrm{MMSE}(\lambda) = \mathbb{E}\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}]\|^2 \le \mathbb{E}\|\boldsymbol{X} - \frac{1}{\sqrt{\lambda}}\boldsymbol{Y}\|^2 = \frac{n}{\lambda} \xrightarrow[\lambda\to+\infty]{} 0\,.$$

B.2 Proof of Proposition 1.3.2

We start by proving that MMSE is continuous at $\lambda = 0$. Let $\lambda \ge 0$ and consider $\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}$ as given by (1.3.1). By dominated convergence one has almost surely that

$$\mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}] = \frac{\int dP_X(\boldsymbol{x})\boldsymbol{x}e^{-\frac{1}{2}\|\sqrt{\lambda}\boldsymbol{x}-\boldsymbol{Y}\|^2}}{\int dP_X(\boldsymbol{x})e^{-\frac{1}{2}\|\sqrt{\lambda}\boldsymbol{x}-\boldsymbol{Y}\|^2}} \xrightarrow[\lambda\to 0]{} \mathbb{E}[\boldsymbol{X}]\,.$$

Then by Fatou's Lemma we get

$$\liminf_{\lambda\to 0} \mathrm{MMSE}(\lambda) \ge \mathbb{E}\Big[\liminf_{\lambda\to 0}\Big\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Y}]\Big\|^2\Big] = \mathbb{E}\Big\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]\Big\|^2\,.$$

Combining this with the bound $\mathrm{MMSE}(\lambda) \le \mathbb{E}\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]\|^2$ gives $\mathrm{MMSE}(\lambda) \xrightarrow[\lambda\to 0]{} \mathbb{E}\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]\|^2$. This proves that the MMSE is continuous at $\lambda = 0$.

Let us now prove that the MMSE is continuous on $\mathbb{R}^*_{\ge 0}$. We need here a technical lemma:

Lemma B.1

For all $\lambda > 0$, $p \ge 1$

$$\mathbb{E}\|\boldsymbol{X} - \langle \boldsymbol{x}\rangle_\lambda\|^{2p} \le \frac{2^p(2p!)}{\lambda^p p!}n^{p+1}\,.$$

Proof. We reproduce here the proof from [97], Proposition 5. We start with the equality

$$\sqrt{\lambda}\,(\boldsymbol{X} - \langle \boldsymbol{x} \rangle_\lambda) = \sqrt{\lambda}\boldsymbol{X} - \mathbb{E}[\sqrt{\lambda}\boldsymbol{X}|\boldsymbol{Y}] = \boldsymbol{Y} - \boldsymbol{Z} - \mathbb{E}[\boldsymbol{Y} - \boldsymbol{Z}|\boldsymbol{Y}] = \mathbb{E}[\boldsymbol{Z}|\boldsymbol{Y}] - \boldsymbol{Z}\,.$$

We have therefore

$$\mathbb{E}\|\boldsymbol{X} - \langle \boldsymbol{x} \rangle_\lambda\|^{2p} = \frac{1}{\lambda^p}\mathbb{E}\|\mathbb{E}[\boldsymbol{Z}|\boldsymbol{Y}] - \boldsymbol{Z}\|^{2p} \leq \frac{2^{2p-1}}{\lambda^p}\mathbb{E}[\|\mathbb{E}[\boldsymbol{Z}|\boldsymbol{Y}]\|^{2p} + \|\boldsymbol{Z}\|^{2p}] \leq \frac{2^{2p}}{\lambda^p}\mathbb{E}\|\boldsymbol{Z}\|^{2p}\,.$$

It remains to bound

$$\mathbb{E}\|\boldsymbol{Z}\|^{2p} \leq n^p \mathbb{E}\left[\sum_{i=1}^n Z_i^{2p}\right] = n^{p+1}\frac{(2p)!}{2^p p!}\,.$$

□

Let $\lambda_0 > 0$. The family of random variables $\left(\|\boldsymbol{X} - \langle \boldsymbol{x} \rangle_\lambda\|^2\right)_{\lambda \geq \lambda_0}$ is bounded in L^2 by Lemma B.1 and is therefore uniformly integrable. The function $\lambda \mapsto \|\boldsymbol{X} - \langle \boldsymbol{x} \rangle_\lambda\|^2$ is continuous on $[\lambda_0, +\infty)$, the uniform integrability ensures then that MMSE : $\lambda \mapsto \mathbb{E}\|\boldsymbol{X} - \langle \boldsymbol{x} \rangle_\lambda\|^2$ is continuous over $[\lambda_0, +\infty)$. This is valid for all $\lambda_0 > 0$: we conclude that MMSE is continuous over $(0, +\infty)$.

B.3 Proof of the I-MMSE relation: Proposition 1.3.3

$$\mathrm{MMSE}(\lambda) = \mathbb{E}\|\boldsymbol{X} - \langle \boldsymbol{x} \rangle_\lambda\|^2 = \mathbb{E}\|\boldsymbol{X}\|^2 + \mathbb{E}\|\langle \boldsymbol{x} \rangle_\lambda\|^2 - 2\mathbb{E}\langle \boldsymbol{x}^\intercal \boldsymbol{X} \rangle_\lambda$$

Now, by the Nishimori property $\mathbb{E}\|\langle \boldsymbol{x} \rangle_\lambda\|^2 = \mathbb{E}\left\langle (\boldsymbol{x}^{(1)})^\intercal \boldsymbol{x}^{(2)} \right\rangle_\lambda = \mathbb{E}\langle \boldsymbol{x}^\intercal \boldsymbol{X} \rangle_\lambda$. Thus

$$\mathrm{MMSE}(\lambda) = \mathbb{E}\|\boldsymbol{X}\|^2 - \mathbb{E}\langle \boldsymbol{x}^\intercal \boldsymbol{X} \rangle_\lambda\,. \tag{B.2}$$

By (B.2) and (1.3.3), it suffices now to prove the second equality in (1.3.4). This will follow from the lemmas below.

Lemma B.2

The free energy F is continuous at $\lambda = 0$.

Proof. For all $\lambda \geq 0$,

$$F(\lambda) = \mathbb{E}\log\int dP_X(\boldsymbol{x}) e^{-\frac{1}{2}\|\boldsymbol{Y} - \sqrt{\lambda}\boldsymbol{x}\|^2 + \frac{1}{2}\|\boldsymbol{Y}\|^2} = \mathbb{E}\log\int dP_X(\boldsymbol{x}) e^{-\frac{1}{2}\|\sqrt{\lambda}\boldsymbol{X} - \sqrt{\lambda}\boldsymbol{x} + \boldsymbol{Z}\|^2 + \lambda\mathbb{E}\|\boldsymbol{X}\|^2 + n}\,.$$

By dominated convergence $\int dP_X(\boldsymbol{x}) e^{-\frac{1}{2}\|\sqrt{\lambda}\boldsymbol{X} - \sqrt{\lambda}\boldsymbol{x} + \boldsymbol{Z}\|^2} \xrightarrow[\lambda \to 0]{} e^{-\frac{1}{2}\|\boldsymbol{Z}\|^2}$. Jensen's inequality gives

$$\begin{aligned}\left|\log\int dP_X(\boldsymbol{x}) e^{-\frac{1}{2}\|\sqrt{\lambda}\boldsymbol{X} - \sqrt{\lambda}\boldsymbol{x} + \boldsymbol{Z}\|^2}\right| &= -\log\int dP_X(\boldsymbol{x}) e^{-\frac{1}{2}\|\sqrt{\lambda}\boldsymbol{X} - \sqrt{\lambda}\boldsymbol{x} + \boldsymbol{Z}\|^2} \\ &\leq \frac{1}{2}\int dP_X(\boldsymbol{x})\|\sqrt{\lambda}\boldsymbol{X} - \sqrt{\lambda}\boldsymbol{x} + \boldsymbol{Z}\|^2 \leq \frac{3}{2}(\|\boldsymbol{X}\|^2 + \mathbb{E}\|\boldsymbol{X}\|^2 + \|\boldsymbol{Z}\|^2),\end{aligned}$$

for all $\lambda \in [0, 1]$. One can thus apply the dominated convergence theorem again to obtain that F is continuous at $\lambda = 0$. □

Lemma B.3

For all $\lambda \geq 0$,

$$F(\lambda) - F(0) = \frac{1}{2}\int_0^\lambda \mathbb{E}\langle \boldsymbol{x}^\intercal \boldsymbol{X}\rangle_\gamma d\gamma\,.$$

Proof. Compute for $\lambda > 0$

$$\frac{\partial}{\partial\lambda}\log \mathcal{Z}(\lambda, \boldsymbol{Y}) = \left\langle \frac{1}{2\sqrt{\lambda}}\boldsymbol{x}^\intercal \boldsymbol{Z} + \boldsymbol{x}^\intercal \boldsymbol{X} - \frac{1}{2}\|\boldsymbol{x}\|^2 \right\rangle_\lambda\,.$$

Since $\mathbb{E}\|\boldsymbol{X}\|^2 < \infty$, the right-hand side is integrable and one can apply Fubini's theorem to obtain

$$F(\lambda_2) - F(\lambda_1) = \int_{\lambda_1}^{\lambda_2} \mathbb{E}\left\langle \frac{1}{2\sqrt{\lambda}}\boldsymbol{x}^\intercal \boldsymbol{Z} + \boldsymbol{x}^\intercal \boldsymbol{X} - \frac{1}{2}\|\boldsymbol{x}\|^2 \right\rangle_\lambda d\lambda\,.$$

By Gaussian integration by parts, we have for all $i \in \{1, \dots, n\}$ and $\lambda > 0$

$$\mathbb{E}Z_i\langle x_i\rangle_\lambda = \mathbb{E}\frac{\partial}{\partial Z_i}\langle x_i\rangle_\lambda = \mathbb{E}\left[\langle\sqrt{\lambda}x_i^2\rangle_\lambda - \sqrt{\lambda}\langle x_i\rangle_\lambda^2\right] = \sqrt{\lambda}\mathbb{E}\left[\langle x_i^2\rangle_\lambda - \langle x_i X_i\rangle_\lambda\right]\,,$$

where the last equality comes from the Nishimori property (Proposition 1.1.1). We have therefore

$$F(\lambda_2) - F(\lambda_1) = \frac{1}{2}\int_{\lambda_1}^{\lambda_2} \mathbb{E}\,\langle \boldsymbol{x}^\intercal \boldsymbol{X}\rangle_\lambda\, d\lambda\,.$$

By Lemma B.2, F is continuous at 0 so we can take the limit $\lambda_1 \to 0$ to obtain the result. □

By Proposition 1.3.2, the function $\lambda \mapsto \mathrm{MMSE}(\lambda)$ is continuous over $\mathbb{R}_{\geq 0}$. By (B.2) we deduce that $\lambda \mapsto \mathbb{E}\langle \boldsymbol{x}^\intercal \boldsymbol{X}\rangle_\lambda$ is continuous over $\mathbb{R}_{\geq 0}$ and therefore Lemma B.3 proves (1.3.4).

It remains only to show that F is strictly convex when P_X differs from a Dirac mass. We proceed by truncation. For $N \in \mathbb{N}$ and $x \in \mathbb{R}$ we write $x^{(N)} = x\,\mathbb{1}(-N \leq x \leq N)$. We extend this notation to vectors $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ by $\boldsymbol{x}^{(N)} = (x_1^{(N)}, \dots, x_n^{(N)})$.

For $\boldsymbol{X} \sim P_X$ we define $P_X^{(N)}$ as the distribution of $\boldsymbol{X}^{(N)}$. $F^{(N)}$, $\mathrm{MMSE}^{(N)}$ and $\langle\cdot\rangle_{\lambda,N}$ will denote respectively the corresponding free energy, MMSE and posterior distribution. One can compute the second derivative (since $\boldsymbol{X}^{(N)}$ is bounded, one can easily differentiate under the integral sign) and again, using Gaussian integration by parts and the Nishimori identity one obtains:

$$F^{(N)\prime\prime}(\lambda) = \frac{1}{2}\mathbb{E}\left[\mathrm{Tr}\left(\left(\langle \boldsymbol{x}\boldsymbol{x}^\intercal\rangle_{\lambda,N} - \langle\boldsymbol{x}\rangle_{\lambda,N}\langle\boldsymbol{x}\rangle_{\lambda,N}^\intercal\right)^2\right)\right]\,. \tag{B.3}$$

By Cauchy-Schwarz inequality, we have for all positive, semi-definite matrix $\boldsymbol{M} \in \mathbb{R}^{n\times n}$, $\mathrm{Tr}(\boldsymbol{M})^2 \leq n\mathrm{Tr}(\boldsymbol{M}^2)$. Hence

$$\begin{aligned} F^{(N)\prime\prime}(\lambda) &\geq \frac{1}{2n}\mathbb{E}\left[\mathrm{Tr}\left(\langle \boldsymbol{x}\boldsymbol{x}^\intercal\rangle_{\lambda,N} - \langle\boldsymbol{x}\rangle_{\lambda,N}\langle\boldsymbol{x}\rangle_{\lambda,N}^\intercal\right)^2\right] \\ &\geq \frac{1}{2n}\mathbb{E}\left[\mathrm{Tr}\left(\langle \boldsymbol{x}\boldsymbol{x}^\intercal\rangle_{\lambda,N} - \langle\boldsymbol{x}\rangle_{\lambda,N}\langle\boldsymbol{x}\rangle_{\lambda,N}^\intercal\right)\right]^2 = \frac{1}{2n}\mathrm{MMSE}^{(N)}(\lambda)^2, \end{aligned}$$

by Jensen's inequality. Let now $0 < s < t$. By integrating (B.3) we get

$$F^{(N)\prime}(t) - F^{(N)\prime}(s) \geq \frac{1}{2n}\int_s^t \mathrm{MMSE}^{(N)}(\lambda)^2 d\lambda\,. \tag{B.4}$$

The sequence of convex functions $(F^{(N)})_N$ converges (by Proposition B.1) to F which is differentiable. Proposition C.1 gives that the derivatives $(F^{(N)\prime})_N$ converge to F' and therefore $\mathrm{MMSE}^{(N)}$ converges to MMSE. Therefore, equation (B.4) gives

$$F'(t) - F'(s) \geq \frac{1}{2n}\int_s^t \mathrm{MMSE}(\lambda)^2 d\lambda \geq \frac{1}{2n}(t-s)\mathrm{MMSE}(t)^2\,.$$

If P_0 is not a Dirac measure, then the last term is strictly positive: this concludes the proof.

B.4 Pseudo-Lipschitz continuity of the free energy with respect to the Wasserstein distance

Let P_1 and P_2 be two probability distributions on $\mathbb{R}^n$, that admits a finite second moment. We denote by $W_2(P_1, P_2)$ the Wasserstein distance of order 2 between P_1 and P_2. For $i = 1, 2$ the free energy is defined as

$$F_{P_i}(\lambda) = \mathbb{E}\log\int dP_i(\boldsymbol{x})\exp\left(\sqrt{\lambda}\boldsymbol{x}^\intercal\boldsymbol{Z} + \lambda\boldsymbol{x}^\intercal\boldsymbol{X} - \frac{\lambda}{2}\|\boldsymbol{x}\|^2\right),$$

where the expectation is with respect to $(\boldsymbol{X}, \boldsymbol{Z}) \sim P_i \otimes \mathcal{N}(0, \mathrm{Id}_n)$.

Proposition B.1

For all $\lambda \geq 0$,

$$\left|F_{P_1}(\lambda) - F_{P_2}(\lambda)\right| \leq \frac{\lambda}{2}\left(\sqrt{\mathbb{E}_{P_1}\|\boldsymbol{X}\|^2} + \sqrt{\mathbb{E}_{P_2}\|\boldsymbol{X}\|^2}\right)W_2(P_1, P_2)\,.$$

A similar result was proved in [216] but with a weaker bound for the W_2 distance. By Proposition 1.3.3 we have for $i = 1, 2$, $F_{P_i}(\lambda) = \frac{\lambda}{2}\mathbb{E}\|\boldsymbol{X}_i\|^2 - I(\boldsymbol{X}_i; \sqrt{\lambda}\boldsymbol{X}_i + \boldsymbol{Z})$ where $(\boldsymbol{X}_i, \boldsymbol{Z}) \sim P_i \otimes \mathcal{N}(0, \mathrm{Id}_n)$. We deduce immediately:

Corollary B.1

For all $\lambda \geq 0$,

$$\left|I(\boldsymbol{X}_1; \sqrt{\lambda}\boldsymbol{X}_1 + \boldsymbol{Z}) - I(\boldsymbol{X}_2; \sqrt{\lambda}\boldsymbol{X}_2 + \boldsymbol{Z})\right| \leq \lambda\left(\sqrt{\mathbb{E}_{P_1}\|\boldsymbol{X}\|^2} + \sqrt{\mathbb{E}_{P_2}\|\boldsymbol{X}\|^2}\right)W_2(P_1, P_2)\,.$$

Proof of Proposition B.1. Let $\epsilon > 0$. Let us fix a coupling Q of $\boldsymbol{X}_1 \sim P_1$ and $\boldsymbol{X}_2 \sim P_2$ such that

$$\left(\mathbb{E}\|\boldsymbol{X}_1 - \boldsymbol{X}_2\|^2\right)^{1/2} \leq W_2(P_1, P_2) + \epsilon\,.$$

Let us consider for $t \in [0, 1]$ the observation model

$$\begin{cases} \boldsymbol{Y}_1^{(t)} &= \sqrt{\lambda t}\boldsymbol{X}_1 + \boldsymbol{Z}_1\,, \\ \boldsymbol{Y}_2^{(t)} &= \sqrt{\lambda(1-t)}\boldsymbol{X}_2 + \boldsymbol{Z}_2\,, \end{cases}$$

where $\boldsymbol{Z}_1, \boldsymbol{Z}_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathrm{Id}_n)$ are independent from $(\boldsymbol{X}_1, \boldsymbol{X}_2) \sim Q$. Define

$$f(t) = \mathbb{E}\log\int dQ(\boldsymbol{x}_1, \boldsymbol{x}_2)\exp\left(\sqrt{\lambda t}\boldsymbol{x}_1^\intercal\boldsymbol{Y}_1^{(t)} - \frac{\lambda t}{2}\|\boldsymbol{x}_1\|^2 + \sqrt{\lambda(1-t)}\boldsymbol{x}_2^\intercal\boldsymbol{Y}_2^{(t)} - \frac{\lambda(1-t)}{2}\|\boldsymbol{x}_2\|^2\right).$$

We have $f(0) = F_{P_2}(\lambda)$ and $f(1) = F_{P_1}(\lambda)$. By an easy extension of the I-MMSE relation (1.3.4) we have for all $t \in [0,1]$:

$$f'(t) = \frac{\lambda}{2}\mathbb{E}\Big\langle \boldsymbol{X}_1^\intercal \boldsymbol{x}_1 - \boldsymbol{X}_2^\intercal \boldsymbol{x}_2 \Big\rangle_t ,$$

where $\langle\cdot\rangle_t$ denotes the expectation with respect to $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ sampled from the posterior distribution of $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ given $\boldsymbol{Y}_1^{(t)}, \boldsymbol{Y}_2^{(t)}$, independently of everything else. We have then

$$\begin{aligned}
|\frac{2}{\lambda}f'(t)| &= \Big|\mathbb{E}\Big\langle \boldsymbol{X}_1^\intercal(\boldsymbol{x}_1 - \boldsymbol{x}_2) - (\boldsymbol{X}_2 - \boldsymbol{X}_1)^\intercal \boldsymbol{x}_2 \Big\rangle_t\Big| \\
&\le \Big(\mathbb{E}\|\boldsymbol{X}_1\|^2 \mathbb{E}\langle \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2\rangle_t\Big)^{1/2} + \Big(\mathbb{E}\langle\|\boldsymbol{x}_2\|^2\rangle_t \mathbb{E}\|\boldsymbol{X}_2 - \boldsymbol{X}_1\|^2\Big)^{1/2} \\
&= \Big(\mathbb{E}\|\boldsymbol{X}_1\|^2 \mathbb{E}\|\boldsymbol{X}_1 - \boldsymbol{X}_2\|^2\Big)^{1/2} + \Big(\mathbb{E}\|\boldsymbol{X}_2\|^2 \mathbb{E}\|\boldsymbol{X}_2 - \boldsymbol{X}_1\|^2\Big)^{1/2} \\
&\le \Big((\mathbb{E}\|\boldsymbol{X}_1\|^2)^{1/2} + (\mathbb{E}\|\boldsymbol{X}_2\|^2)^{1/2}\Big)(W_2(P_1, P_2) + \epsilon) ,
\end{aligned}$$

where we used successively the Cauchy-Schwarz inequality and the Nishimori property (Proposition 1.1.1). We then let $\epsilon \to 0$ to obtain the result. □

B.5 Zero-noise limit of the mutual information

The goal of this section is to prove Proposition B.2 and Corollary B.2.

Proposition B.2

Let P_U be a probability distribution over $\mathbb{N}^m$ that admits a finite second moment. Let $\boldsymbol{U} \sim P_U$ and $\boldsymbol{Z} \sim \mathcal{N}(0, \mathrm{Id}_m)$ be two independent random variables. Then $H(\boldsymbol{U}) = -\sum_{\boldsymbol{n}\in\mathbb{N}^m} P_U(\boldsymbol{n}) \log P_U(\boldsymbol{n})$ is finite and for all $\Delta \in (0,1]$,

$$\Big|I(\boldsymbol{U}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) - H(\boldsymbol{U})\Big| \le 48 m e^{-1/(16\Delta)} .$$

Proof. Let us define for $\Delta > 0$, $h(\Delta) = I(\boldsymbol{U}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) = I_{P_U}(\Delta^{-1})$. By Proposition 1.3.3 we have for all $\Delta > 0$,

$$h'(\Delta) = -\frac{1}{2\Delta^2}\mathrm{MMSE}(\boldsymbol{U} \,|\, \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) . \tag{B.5}$$

We are now going to upper bound $\mathrm{MMSE}(\boldsymbol{U} \,|\, \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z})$ by considering the following estimator:

$$\widehat{\theta}_i = \operatorname*{arg\,min}_{u\in\mathbb{N}} |u - U_i + \sqrt{\Delta} Z_i|,$$

for all $i \in \{1, \dots, m\}$. Note that $\widehat{\theta}_i$ is well-defined almost surely since there is a.s. a unique minimizer above. We have

$$\mathbb{P}(\widehat{\theta}_i \neq U_i) \le \mathbb{P}(\sqrt{\Delta}|Z_i| \ge 1/2) = 2\mathbb{P}\Big(\mathcal{N}(0,1) \ge \frac{1}{2\sqrt{\Delta}}\Big) \le 2\frac{2\sqrt{\Delta}}{\sqrt{2\pi}} e^{-1/(8\Delta)} \le 2\sqrt{\Delta} e^{-1/(8\Delta)} ,$$

by usual bounds on the Gaussian cumulative distribution function. We have then

$$\begin{aligned}
\mathrm{MMSE}(\boldsymbol{U}\,|\,\boldsymbol{U}+\sqrt{\Delta}\boldsymbol{Z}) &\le \mathbb{E}\|\boldsymbol{U}-\widehat{\theta}\|^2 = \sum_{i=1}^m \mathbb{E}(U_i-\widehat{\theta}_i)^2 = \sum_{i=1}^m \mathbb{E}\Big[\mathbb{1}(\widehat{\theta}_i\neq U_i)(U_i-\widehat{\theta}_i)^2\Big] \\
&\le \sum_{i=1}^m 2\mathbb{E}\Big[\mathbb{1}(\widehat{\theta}_i\neq U_i)(U_i-(U_i+\sqrt{\Delta}Z_i))^2\Big] + 2\mathbb{E}\Big[\mathbb{1}(\widehat{\theta}_i\neq U_i)(U_i+\sqrt{\Delta}Z_i-\widehat{\theta}_i)^2\Big] \\
&\le \sum_{i=1}^m 2\mathbb{E}\Big[\mathbb{1}(\widehat{\theta}_i\neq U_i)\Delta Z_i^2\Big] + \frac{1}{2}\mathbb{E}\Big[\mathbb{1}(\widehat{\theta}_i\neq U_i)\Big] \\
&\le \sum_{i=1}^m 2\Delta\mathbb{P}(\widehat{\theta}_i\neq U_i)^{1/2}\mathbb{E}[Z_i^4]^{1/2} + \frac{1}{2}\mathbb{P}(\widehat{\theta}_i\neq U_i) \\
&\le me^{-1/(16\Delta)}\Big(2\sqrt{6}\Delta^{5/4}+\sqrt{\Delta}\Big) \le 6me^{-1/(16\Delta)}
\end{aligned}$$

for $\Delta\le 1$. Plugging this inequality in (B.5), we obtain for all $\Delta\in(0,1]$,

$$|h'(\Delta)| \le \frac{3m}{\Delta^2}e^{-1/(16\Delta)}\,. \tag{B.6}$$

Since $h(1)$ is finite and $\int_0^1 \frac{e^{-1/(16\Delta)}}{\Delta^2}d\Delta < +\infty$ we obtain that

$$\sup_{\Delta\in(0,1]} |h(\Delta)| < +\infty\,. \tag{B.7}$$

By definition of h:

$$h(\Delta) = I(\boldsymbol{U};\boldsymbol{U}+\sqrt{\Delta}\boldsymbol{Z}) = -\frac{m}{2} - \mathbb{E}\log\sum_{\boldsymbol{U}\in\mathbb{N}^m} P_U(\boldsymbol{U})\exp\Big(-\frac{1}{2\Delta}\|\boldsymbol{U}+\sqrt{\Delta}\boldsymbol{Z}-\boldsymbol{U}\|^2\Big)\,. \tag{B.8}$$

By the previous equality and (B.7), the family of (non-negative) random variables

$$\left(-\log\sum_{\boldsymbol{U}\in\mathbb{N}^m} P_U(\boldsymbol{U})\exp\Big(-\frac{1}{2\Delta}\|\boldsymbol{U}+\sqrt{\Delta}\boldsymbol{Z}-\boldsymbol{U}\|^2\Big)\right)_{\Delta\in(0,1]}$$

is bounded in L^1. Notice that (by dominated convergence)

$$-\log\sum_{\boldsymbol{U}\in\mathbb{N}^m} P_U(\boldsymbol{U})\exp\Big(-\frac{1}{2\Delta}\|\boldsymbol{U}+\sqrt{\Delta}\boldsymbol{Z}-\boldsymbol{U}\|^2\Big) \xrightarrow[\Delta\to 0]{} -\log\Big(P_U(\boldsymbol{U})e^{-\frac{1}{2}\|\boldsymbol{Z}\|^2}\Big) = \frac{1}{2}\|\boldsymbol{Z}\|^2 - \log P_U(\boldsymbol{U})$$

almost surely. This gives (by Fatou's Lemma) that this almost-sure limit is integrable and thus that $H(\boldsymbol{U}) = -\mathbb{E}\log P_U(\boldsymbol{U})$ is finite. Let us now show that $h(\Delta)\xrightarrow[\Delta\to 0]{} H(\boldsymbol{U})$. We have almost surely

$$\log\Big(P_U(\boldsymbol{U})e^{-\frac{1}{2}\|\boldsymbol{Z}\|^2}\Big) \le \log\sum_{\boldsymbol{U}\in\mathbb{N}^m} P_U(\boldsymbol{U})\exp\Big(-\frac{1}{2\Delta}\|\boldsymbol{U}+\sqrt{\Delta}\boldsymbol{Z}-\boldsymbol{U}\|^2\Big) \le 0\,.$$

Since we now know that the left-hand side is integrable (because $H(\boldsymbol{U})$ is finite), we can apply the dominated convergence theorem to obtain that

$$\mathbb{E}\log\sum_{\boldsymbol{U}\in\mathbb{N}^m} P_U(\boldsymbol{U})\exp\Big(-\frac{1}{2\Delta}\|\boldsymbol{U}+\sqrt{\Delta}\boldsymbol{Z}-\boldsymbol{U}\|^2\Big) \xrightarrow[\Delta\to 0]{} \mathbb{E}\log\Big(P_U(\boldsymbol{U})e^{-\frac{1}{2}\|\boldsymbol{Z}\|^2}\Big) = H(\boldsymbol{U}) - \frac{m}{2}\,,$$

which combined with (B.8) gives $h(\Delta)\xrightarrow[\Delta\to 0]{} H(\boldsymbol{U})$. Now, using the bound on the derivative of h (B.6) we conclude that for all $\Delta\in(0,1]$,

$$|h(\Delta)-H(\boldsymbol{U})| \le 3m\int_0^\Delta \frac{e^{-1/(16t)}}{t^2}dt = 3m\Big[16e^{-1/(16t)}\Big]_0^\Delta = 48me^{-1/(16\Delta)}\,.$$

□

Corollary B.2

Let $\boldsymbol{U}$ be a random variable over $\mathbb{N}^m$ with finite second moment, let $\boldsymbol{X}$ be a random variable over $\mathbb{R}^n$ and let $\boldsymbol{Z} \sim \mathcal{N}(0, \mathrm{Id}_m)$. We assume $(\boldsymbol{U}, \boldsymbol{X})$ to be independent from $\boldsymbol{Z}$. Then, for all $\Delta \in (0, 1]$,

$$\left| I(\boldsymbol{X}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) - I(\boldsymbol{X}; \boldsymbol{U}) \right| \leq 100 m e^{-1/(16\Delta)} .$$

Proof. We have by the chain rule of the mutual information:

$$\begin{aligned} I(\boldsymbol{U}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) &= I(\boldsymbol{U}, \boldsymbol{X}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) \\ &= I(\boldsymbol{X}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) + I(\boldsymbol{U}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z} | \boldsymbol{X}) . \end{aligned}$$

By applying Proposition B.2 twice, we get

$$|I(\boldsymbol{U}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z}) - H(\boldsymbol{U})|, \ |I(\boldsymbol{U}; \boldsymbol{U} + \sqrt{\Delta}\boldsymbol{Z} | \boldsymbol{X}) - H(\boldsymbol{U}|\boldsymbol{X})| \leq 48 m e^{-1/(16\Delta)} .$$

Since $I(\boldsymbol{X}; \boldsymbol{U}) = H(\boldsymbol{U}) - H(\boldsymbol{U}|\boldsymbol{X})$ we obtain the desired inequality. □

C Some results about convex functions

C.1 Convex analysis lemmas

Proposition C.1

Let $I \subset \mathbb{R}$ be an interval, and let $(f_n)_{n \geq 0}$ be a sequence of convex functions on I that converges pointwise to a function f. Then for all $t \in I$ for which these inequalities have a sense

$$f'(t^-) \leq \liminf_{n \to \infty} f_n'(t^-) \leq \limsup_{n \to \infty} f_n'(t^+) \leq f'(t^+).$$

Proof. Let $t \in I$ and $h > 0$. By convexity

$$f_n'(t^+) \leq \frac{f_n(t+h) - f_n(t)}{h} \xrightarrow[n \to \infty]{} \frac{f(t+h) - f(t)}{h} \xrightarrow[h \to 0]{} f'(t^+).$$

The first inequality follows from the same arguments. □

Proposition C.2 (*Corollary 37.3.2 from [182]*)

Let C and D be non-empty closed convex sets in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, and let f be a continuous finite concave-convex function on $C \times D$. If either C or D is bounded, one has

$$\inf_{v \in D} \sup_{u \in C} f(u, v) = \sup_{u \in C} \inf_{v \in D} f(u, v) .$$

Definition C.1

A convex function f over $\mathbb{R}^n$ is said to be

- *γ-strongly convex if $x \mapsto f(x) - \frac{\gamma}{2}\|x\|^2$ is convex.*
- *L-strongly smooth is f is differentiable everywhere and for all $x, y \in \mathbb{R}^n$ we have*

$$f(y) \leq f(x) + (y-x)^\mathsf{T}\nabla f(x) + \frac{L}{2}\|x-y\|^2 .$$

Remark C.1. *If f is convex, differentiable over $\mathbb{R}^n$, and ∇f is L-Lipschitz, then f is L-strongly smooth. Indeed, if we take $x, y \in \mathbb{R}^n$ and if we define $h(t) = f((1-t)x+ty)$ we have*

$$\begin{aligned} f(y) - f(x) = h(1) - h(0) &= \int_0^1 h'(t)dt = \int_0^1 (y-x)^\mathsf{T}\nabla f((1-t)x+ty)dt \\ &= (y-x)^\mathsf{T} f(x) + \int_0^1 (y-x)^\mathsf{T}\Big(\nabla f((1-t)x+ty) - \nabla f(x)\Big)dt \\ &\leq (y-x)^\mathsf{T} f(x) + \int_0^1 tL\|x-y\|^2 dt \leq (y-x)^\mathsf{T} f(x) + \frac{L}{2}\|x-y\|^2 . \end{aligned}$$

Proposition C.3

Let f be a closed convex function over $\mathbb{R}^n$. Then f is γ-strongly convex if and only if f^* is $\frac{1}{\gamma}$-strongly smooth.

This result can be found in the book [219], see Corollary 3.5.11 on page 217 and the Remark 3.5.3 below. A more accessible presentation of this result can be found in [118].

C.2 The monotone conjugate

Definition C.2

We define the monotone conjugate (see [182] p.110) of a non-decreasing convex function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ by:

$$f^*(x) = \sup_{y \geq 0} \{xy - f(y)\}. \tag{C.1}$$

The most fundamental result on the monotone conjugate is the analog of the Fenchel-Moreau theorem:

Proposition C.4 (*[182] Theorem 12.4*)

Let f be a non-decreasing lower semi-continuous convex function on $\mathbb{R}_{\geq 0}$ such that $f(0)$ is finite. Then f^* is another such function and $(f^*)^* = f$.

Proposition C.5

Let f be a non-decreasing lower semi-continuous convex function on $\mathbb{R}_{\geq 0}$ such that $f(0)$ is finite. Then for all $x, y \geq 0$:

$$x \in \partial f^*(y) \iff f(x) + f^*(y) = xy \iff y \in \partial f(x).$$

Proof. Let $x \in \partial f^*(y)$. We get that $(f^*)^*(x) = xy - f^*(y)$ and therefore that $f^*(y) = xy - f(x)$, by Proposition C.4. This gives that x maximizes $s \mapsto sy - f(s)$ over $\mathbb{R}_{\geq 0}$ and thus $y \in \partial f(x)$. It remains to show that $y \in \partial f(x) \implies x \in \partial f^*(y)$. This follows from Proposition C.4 and the implication $x \in \partial f^*(y) \implies y \in \partial f(x)$ that we just showed. □

C.3 Some supremum formulas

This section gathers some tools to deal with "sup-inf" expressions that we encounter in this manuscript.

Proposition C.6

Let f, g be two convex Lipschitz functions on $\mathbb{R}_{\geq 0}$. For $(q_1, q_2) \in \mathbb{R}^2_{\geq 0}$ we define $\varphi(q_1, q_2) = f(q_1) + g(q_2) - q_1 q_2$ and $\psi(q_1, q_2) = q_1 q_2 - f^*(q_2) - g^*(q_1)$. Then the set $\Gamma = \left\{(q_1, q_2) \in \mathbb{R}^2_{\geq 0} \,\middle|\, q_1 \in \partial g(q_2),\ q_2 \in \partial f(q_1)\right\}$ is non-empty and:

$$\sup_{(q_1,q_2)\in\Gamma} \varphi(q_1, q_2) \;=\; \sup_{q_1,q_2\geq 0} \psi(q_1, q_2) \;= \sup_{q_1\geq 0} \inf_{q_2\geq 0} \varphi(q_1, q_2), \tag{C.2}$$

and the two first suprema above are achieved and precisely at the same couples (q_1, q_2).

If moreover f and g are both differentiable and strictly convex, then the same result holds for Γ replaced by

$$\widetilde{\Gamma} = \left\{(q_1, q_2) \in \mathbb{R}^2_{\geq 0} \,\middle|\, q_2 = f'(q_1) \text{ and } q_1 = g'(q_2)\right\}. \tag{C.3}$$

Proof. Let L_f (resp. L_g) be the Lipschitz constant of f (resp. g). For $x > L_f$, $f^*(x) = +\infty$ and (since f^* is lower semi-continuous by Proposition C.5) $f^*(x) \to +\infty$ as $x \to L_f$. Analogously, $g(x) \to +\infty$ as $x \to L_g$. The function ψ is therefore continuous on $[0, L_g) \times [0, L_f)$ and goes to $-\infty$ on the border $(\{L_g\} \times [0, L_f]) \cup ([0, L_g] \times \{L_f\})$.

The functions ψ achieves therefore its maximum at some $(q_1, q_2) \in [0, L_g) \times [0, L_f)$. (q_1, q_2) verifies then $q_2 \in \partial g^*(q_1)$ and $q_1 \in \partial f^*(q_2)$ which gives $(q_1, q_2) \in \Gamma$ by Proposition C.5. The set Γ is therefore non-empty and

$$\sup_{q_1,q_2\geq 0} \psi(q_1, q_2) \leq \sup_{(q_1,q_2)\in\Gamma} \varphi(q_1, q_2).$$

By definition of the conjugates f^* and g^* we have for all $q_1, q_2 \geq 0$

$$\begin{cases} f(q_1) + f^*(q_2) \geq q_1 q_2 \\ g(q_2) + g^*(q_1) \geq q_1 q_2. \end{cases}$$

We get that $\varphi(q_1, q_2) \geq \psi(q_1, q_2)$ with equality if and only if $(q_1, q_2) \in \Gamma$, by Proposition C.5. This gives in particular that

$$\sup_{q_1,q_2\geq 0} \psi(q_1, q_2) \geq \sup_{(q_1,q_2)\in\Gamma} \varphi(q_1, q_2).$$

Hence, both supremum are equal and are achieved over the same couples because we have seen that all couple (q_1, q_2) that achieves the supremum of ψ is in Γ.

We consider now the second equality. Using the definition of the monotone conjugate (C.1) and Proposition C.4:

$$\sup_{q_1\geq 0} \inf_{q_2\geq 0} \varphi(q_1, q_2) = \sup_{q_1\geq 0} \{f(q_1) - g^*(q_1)\} = \sup_{q_1\geq 0} \sup_{q_2\geq 0} \{q_1 q_2 - f^*(q_2) - g^*(q_1)\}.$$

Let us now prove the second part of the Proposition: we now assume that f and g are differentiable, strictly convex. Let $(q_1, q_2) \in \Gamma$ be a couple that achieves the maximum of φ over Γ. It suffices to show that $(q_1, q_2) \in \widetilde{\Gamma}$. If $q_1 > 0$ and $q_2 > 0$, then this is trivial because f and g are differentiable.

Suppose now that $q_1 = 0$ (the case $q_2 = 0$ follows by symmetry). Since $0 = q_1 \in \partial g(q_2)$ and g is strictly increasing, we get that $q_2 = 0$, so that $\sup \psi = \varphi(q_1, q_2) = f(0) + g(0)$. Notice that $f^*(f'(0)) = -f(0)$ and $g^*(g'(0)) = -g(0)$ so

$$f(0) + g(0) = \sup \psi \geq \psi(g'(0), f'(0)) = f(0) + g(0) + f'(0)g'(0) \geq f(0) + g(0). \tag{C.4}$$

We get that $(g'(0), f'(0))$ achieves the supremum of ψ, which implies that $(g'(0), f'(0)) \in \Gamma$. From (C.4) we get also that $f'(0) = 0$ or $g'(0) = 0$. Assume that $f'(0) = 0$ (the case $g'(0) = 0$ follows by symmetry), then $0 \in \partial f(g'(0))$ because $(g'(0), f'(0)) \in \Gamma$. Since f is strictly increasing, so we have $g'(0) = 0$. We conclude that $f'(0) = g'(0) = 0$ which proves that $(q_1, q_2) = (0, 0) \in \widetilde{\Gamma}$. □

We will need the following variants of Proposition C.6 in Chapter 7:

Proposition C.7

Let $f : [0, \rho] \to \mathbb{R}$ be a continuous convex non-decreasing function. Let $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a convex, Lipschitz, non-decreasing function. Define $\rho = \sup_{x \geq 0} g'(x^+)$ the Lipschitz constant of g. Then

$$\sup_{q_2 \geq 0} \inf_{q_1 \in [0,\rho]} \Big\{ f(q_1) + g(q_2) - q_1 q_2 \Big\} = \sup_{q_1 \in [0,\rho]} \inf_{q_2 \geq 0} \Big\{ f(q_1) + g(q_2) - q_1 q_2 \Big\}.$$

Proof. We extend f on $\mathbb{R}_{\geq 0}$ by setting $f(x) = +\infty$ for all $x > \rho$. The extended function f is then a convex, lower semi-continuous function on $\mathbb{R}_{\geq 0}$ and $f(0)$ is finite. We have then

$$\begin{aligned} \sup_{q_2 \geq 0} \inf_{q_1 \in [0,\rho]} \{ f(q_1) + g(q_2) - q_1 q_2 \} &= \sup_{q_2 \geq 0} \inf_{q_1 \geq 0} \{ f(q_1) + g(q_2) - q_1 q_2 \} \\ &= \sup_{q_2 \geq 0} \{ g(q_2) - f^*(q_2) \}, \end{aligned}$$

by Proposition C.4. g is ρ-Lipschitz so $g^*(x) = +\infty$ for all $x > \rho$. Consequently, by Proposition C.4, $g(q_2) = \sup_{q_1 \geq 0} \{ q_1 q_2 - g^*(q_1) \} = \sup_{q_1 \in [0,\rho]} \{ q_1 q_2 - g^*(q_1) \}$. We get

$$\begin{aligned} \sup_{q_2 \geq 0} \inf_{q_1 \in [0,\rho]} \{ f(q_1) + g(q_2) - q_1 q_2 \} &= \sup_{q_2 \geq 0} \sup_{q_1 \in [0,\rho]} \{ q_1 q_2 - g^*(q_1) - f^*(q_2) \} \\ &= \sup_{q_1 \in [0,\rho]} \{ f(q_1) - g^*(q_1) \} \\ &= \sup_{q_1 \in [0,\rho]} \inf_{q_2 \geq 0} \{ f(q_1) + g(q_2) - q_1 q_2 \}. \end{aligned}$$

□

Proposition C.8

Let g be a strictly convex, differentiable, Lipschitz non-decreasing function on $\mathbb{R}_+$. Define $\rho = \sup_{x\geq 0} g'(x)$. Let f be a convex, continuous, strictly increasing function on $[0,\rho]$, differentiable on $[0,\rho)$. For $(q_1,q_2) \in [0,\rho]\times\mathbb{R}_+$ we define $\varphi(q_1,q_2) = f(q_1) + g(q_2) - q_1q_2$. Then

$$\sup_{q_1\in[0,\rho]}\inf_{q_2\geq 0}\varphi(q_1,q_2) = \sup_{(q_1,q_2)\in\Gamma}\varphi(q_1,q_2)\,, \tag{C.5}$$

where

$$\Gamma = \left\{(q_1,q_2)\in[0,\rho]\times(\mathbb{R}_+\cup\{+\infty\}) \;\middle|\; \begin{array}{rcl} q_1 &=& g'(q_2) \\ q_2 &=& f'(q_1)\end{array}\right\},$$

where all the function are extended by there limits at the points at which they may not be defined (for instance $g'(+\infty) = \lim_{q\to\infty} g'(q)$, $f'(\rho) = \lim_{q\to\rho} f'(q)$). Moreover, the above extremas are achieved precisely on the same couples.

Proof. Let q_1^* be a maximizer of $f-g^*$ over $[0,\rho]$. q_1^* is well defined because f is continuous and g^* is continuous over $[0,\rho)$ and is either continuous or goes to $+\infty$ at ρ (because g^* is a lower semi-continuous convex function, see Proposition C.4). We distinguish 3 cases:

Case 1: $0 < q_1^* < \rho$. By strict convexity of g, $\varphi(q_1,\cdot)$ admits a unique minimizer q_2^* and $(g^*)'(q_1^*) = q_2^*$ by Proposition C.5. Thus, the optimality condition at q_1^* gives

$$0 = f'(q_1^*) - (g^*)'(q_1^*) = f'(q_1^*) - q_2^*\,.$$

The optimality of q_2^* gives then $q_1^* \leq g'(q_2^*)$. Suppose that $q_1^* < g'(q_2^*)$. This is only possible when $q_2^* = 0$. Define $q_1' = g'(q_2^*) = g'(0)$. Remark that $g^*(q_1') = -g(0) = g^*(q_1^*)$. We supposed that $q_1' > q_1^*$ thus, by strict monotonicity of f, $f(q_1') - g^*(q_1') > f(q_1^*) - g^*(q_1^*)$ which contradict the optimality of q_1^*. We obtain therefore that $q_1^* = g'(q_2^*)$.

Case 2: $q_1^* = 0$. The optimality condition gives now

$$0 \leq f'(q_1^* = 0) \leq q_2^*\,, \tag{C.6}$$

where q_2^* is again the unique minimizer of $\varphi(q_1^* = 0,\cdot) = f(0) + g$. g is strictly increasing, so $q_2^* = 0$. Therefore $q_2^* = 0 = f'(q_1^* = 0)$, by (C.6). As before we have necessarily, by optimality of q_2^* that $q_1^* = g'(q_2^*)$.

Case 3: $q_1^* = \rho$. In that case $\arg\min_{q_2\geq 0}\{g(q_2) - q_1^*q_2\} = \emptyset$ because g is strictly convex and ρ-Lipschitz. Proposition C.5 gives then that $\partial g^*(\rho) = \emptyset$ which implies (see [182, Theorem 23.3]) that $(g^*)'(\rho^-) = +\infty$. Since $q_1^* = \rho$ maximizes $f - g^*$, we necessarily have then $f'(\rho^-) = +\infty$.

Using the slight abuse of notation explained in the Proposition, we have $f'(q_1^*) = +\infty = q_2^*$, where $q_2^* = +\infty$ is the unique "minimizer" of $\varphi(q_1^*,\cdot)$, by strict convexity of g. By definition of ρ we have also $g'(q_2^*) = g'(+\infty) = \rho = q_1^*$.

We conclude from the tree cases above that the "sup-inf" in (C.5) is achieved, and that all the couples (q_1^*,q_2^*) that achieve this "sup-inf" belong to Γ. Thus

$$\sup_{q_1\in[0,\rho]}\inf_{q_2\geq 0}\varphi(q_1,q_2) \leq \sup_{(q_1,q_2)\in\Gamma}\varphi(q_1,q_2)\,.$$

Let now be $(q_1,q_2)\in\Gamma$. By convexity of g we see easily that $\varphi(q_1,q_2) = \inf_{q_2'}\varphi(q_1,q_2')$. Thus, $\varphi(q_1,q_2) \leq \sup_{q_1'}\inf_{q_2'}\varphi(q_1',q_2')$. Therefore

$$\sup_{(q_1,q_2)\in\Gamma}\varphi(q_1,q_2) \leq \sup_{q_1\in[0,\rho]}\inf_{q_2\geq 0}\varphi(q_1,q_2)\,.$$

This concludes the proof of (C.5). It remains to see that a couple $(q_1^*, q_2^*) \in \Gamma$ that achieves the supremum in (C.5) also achieves the "sup-inf". This simply follows from the fact that $\varphi(q_1^*, q_2^*) = \inf_{q_2} \varphi(q_1^*, q_2)$ and (C.5). □

D Differentiation of a supremum of functions

We recall in this section two results about the differentiation of a supremum of functions from Milgrom and Segal [144]. Let X be a set of parameters and consider a function $f : X \times [0,1] \to \mathbb{R}$. Define, for $t \in [0,1]$

$$V(t) = \sup_{x \in X} f(x,t)\,,$$
$$X^*(t) = \left\{ x \in X \,\middle|\, f(x,t) = V(t) \right\}.$$

Proposition D.1 (*Theorem 1 from [144]*)

Let $t \in [0,1]$ such that $X^*(t) \neq \emptyset$. Let $x^* \in X^*(t)$ and suppose that $f(x^*, \cdot)$ is differentiable at t, with derivative $f_t(x^*, t)$.

- *If $t > 0$ and if V is left-hand differentiable at t, then $V'(t^-) \leq f_t(x^*, t)$.*
- *If $t < 0$ and if V is right-hand differentiable at t, then $V'(t^+) \geq f_t(x^*, t)$.*
- *If $t \in (0,1)$ and if V is differentiable at t, then $V'(t) = f_t(x^*, t)$.*

Proposition D.2 (*Corollary 4 from [144]*)

Suppose that X is nonempty and compact. Suppose that for all $t \in [0,1]$, $f(\cdot, t)$ is continuous. Suppose also that f admits a partial derivative f_t with respect to t that is continuous in (x,t) over $X \times [0,1]$. Then

- *$V'(t^+) = \max_{x^* \in X^*(t)} f_t(x^*, t)$ for all $t \in [0,1)$ and $V'(t^-) = \min_{x^* \in X^*(t)} f_t(x^*, t)$ for all $t \in (0,1]$.*
- *V is differentiable at $t \in (0,1)$ is and only if $\left\{ f_t(x^*, t) \,\middle|\, x^* \in X^*(t) \right\}$ is a singleton. In that case $V'(t) = f_t(x^*, t)$ for all $x^* \in X^*(t)$.*

Bibliography

[1] Emmanuel Abbe and Colin Sandon. Detection in the stochastic block model with multiple clusters: proof of the achievability conjectures, acyclic bp, and the information-computation gap. *arXiv preprint arXiv:1512.09080*, 2015.

[2] Michael Aizenman, Robert Sims, and Shannon L Starr. Extended variational principle for the sherrington-kirkpatrick spin-glass model. *Physical Review B*, 68(21):214403, 2003.

[3] Ahmed El Alaoui, Florent Krzakala, and Michael I Jordan. Finite size corrections and likelihood ratio fluctuations in the spiked wigner model. *arXiv preprint arXiv:1710.02903*, 2017.

[4] David J Aldous. Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598, 1981.

[5] Dennis Amelunxen, Martin Lotz, Michael B McCoy, and Joel A Tropp. Living on the edge: Phase transitions in convex programs with random data. *Information and Inference: A Journal of the IMA*, 3(3):224–294, 2014.

[6] Arash A Amini and Martin J Wainwright. High-dimensional analysis of semidefinite relaxations for sparse principal components. In *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, pages 2454–2458. IEEE, 2008.

[7] Fabrizio Antenucci, Silvio Franz, Pierfrancesco Urbani, and Lenka Zdeborová. Glassy nature of the hard phase in inference problems. *Physical Review X*, 9(1):011020, 2019.

[8] Gerard Ben Arous, Reza Gheissari, and Aukosh Jagannath. Algorithmic thresholds for tensor pca. *arXiv preprint arXiv:1808.00921*, 2018.

[9] Gerard Ben Arous, Song Mei, Andrea Montanari, and Mihai Nica. The landscape of the spiked tensor model. *arXiv preprint arXiv:1711.05424*, 2017.

[10] Antonio Auffinger, Gérard Ben Arous, and Jiří Černỳ. Random matrices and complexity of spin glasses. *Communications on Pure and Applied Mathematics*, 66(2):165–201, 2013.

[11] Antonio Auffinger, Gerard Ben Arous, et al. Complexity of random smooth functions on the high-dimensional sphere. *The Annals of Probability*, 41(6):4214–4247, 2013.

[12] Tim Austin. Exchangeable random measures. In *Annales de l'IHP Probabilités et statistiques*, volume 51, pages 842–861, 2015.

[13] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Annals of Probability*, pages 1643–1697, 2005.

[14] Jinho Baik and Jack W Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis*, 97(6):1382–1408, 2006.

[15] Carlo Baldassi, Alfredo Braunstein, Nicolas Brunel, and Riccardo Zecchina. Efficient supervised learning in networks with binary synapses. *BMC neuroscience*, 8(2):S13, 2007.

[16] Jess Banks, Cristopher Moore, Roman Vershynin, Nicolas Verzelen, and Jiaming Xu. Information-theoretic bounds and phase transitions in clustering, sparse pca, and submatrix localization. In *Information Theory (ISIT), 2017 IEEE International Symposium on*, pages 1137–1141. IEEE, 2017.

[17] Jean Barbier, Mohamad Dia, Nicolas Macris, and Florent Krzakala. The mutual information in random linear estimation. In *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 625–632. IEEE, 2016.

[18] Jean Barbier, Mohamad Dia, Nicolas Macris, Florent Krzakala, Thibault Lesieur, and Lenka Zdeborová. Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula. In *Advances in Neural Information Processing Systems*, pages 424–432, 2016.

[19] Jean Barbier, Florent Krzakala, Nicolas Macris, Léo Miolane, and Lenka Zdeborová. Optimal errors and phase transitions in high-dimensional generalized linear models. *Proceedings of the National Academy of Sciences*, 2019.

[20] Jean Barbier and Nicolas Macris. The stochastic interpolation method: A simple scheme to prove replica formulas in bayesian inference. *arXiv preprint arXiv:1705.02780*, 2017.

[21] Jean Barbier, Nicolas Macris, and Léo Miolane. The layered structure of tensor estimation and its mutual information. *arXiv preprint arXiv:1709.10368*, 2017.

[22] Adriano Barra, Giuseppe Genovese, and Francesco Guerra. Equilibrium statistical mechanics of bipartite spin systems. *Journal of Physics A: Mathematical and Theoretical*, 44(24):245002, 2011.

[23] Eric B Baum and Yuh-Dauh Lyuu. The transition to perfect generalization in perceptrons. *Neural computation*, 3(3):386–401, 1991.

[24] Mohsen Bayati, Murat A Erdogdu, and Andrea Montanari. Estimating lasso risk and noise level. In *Advances in Neural Information Processing Systems*, pages 944–952, 2013.

[25] Mohsen Bayati, Marc Lelarge, Andrea Montanari, et al. Universality in polytope phase transitions and message passing algorithms. *The Annals of Applied Probability*, 25(2):753–822, 2015.

[26] Mohsen Bayati and Andrea Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Transactions on Information Theory*, 57(2):764–785, 2011.

[27] Mohsen Bayati and Andrea Montanari. The lasso risk for gaussian matrices. *IEEE Transactions on Information Theory*, 58(4):1997–2017, 2012.

[28] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, 2011.

[29] Florent Benaych-Georges and Raj Rao Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111:120–135, 2012.

[30] Quentin Berthet, Philippe Rigollet, et al. Optimal detection of sparse principal components in high dimension. *The Annals of Statistics*, 41(4):1780–1815, 2013.

[31] Raphaël Berthier, Andrea Montanari, and Phan-Minh Nguyen. State evolution for approximate message passing with non-separable functions. *Information and Inference: A Journal of the IMA*, 2017.

[32] Peter J Bickel, Chris AJ Klaassen, Peter J Bickel, Y Ritov, J Klaassen, Jon A Wellner, and YA'Acov Ritov. *Efficient and adaptive estimation for semiparametric models*, volume 4. Johns Hopkins University Press Baltimore, 1993.

[33] Peter J Bickel, Ya'acov Ritov, and Alexandre B Tsybakov. Simultaneous analysis of Lasso and Dantzig selector. *The Annals of Statistics*, 37(4):1705–1732, 2009.

[34] Vladimir I Bogachev. *Measure theory*, volume 1. Springer Science & Business Media, 2007.

[35] Erwin Bolthausen. An iterative construction of solutions of the tap equations for the sherrington–kirkpatrick model. *Communications in Mathematical Physics*, 325(1):333–366, 2014.

[36] Charles Bordenave, Marc Lelarge, and Laurent Massoulié. Non-backtracking spectrum of random graphs: community detection and non-regular ramanujan graphs. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 1347–1357. IEEE, 2015.

[37] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[38] Petros T Boufounos and Richard G Baraniuk. 1-bit compressive sensing. In *42nd Annual Conference on Information Sciences and Systems (CISS)*, pages 16–21. IEEE, 2008.

[39] Peter Bühlmann and Sara Van De Geer. *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media, 2011.

[40] Francesco Caltagirone, Marc Lelarge, and Léo Miolane. Recovering asymmetric communities in the stochastic block model. *IEEE Transactions on Network Science and Engineering*, 2017.

[41] Emmanuel Candes and Terence Tao. The Dantzig selector: Statistical estimation when p is much larger than n. *The Annals of Statistics*, 35(6):2313–2351, 2007.

[42] Emmanuel J Candes and Terence Tao. Decoding by linear programming. *IEEE transactions on information theory*, 51(12):4203–4215, 2005.

[43] Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

[44] Mireille Capitaine, Catherine Donati-Martin, and Delphine Féral. The largest eigenvalues of finite rank deformation of large wigner matrices: convergence and nonuniversality of the fluctuations. *The Annals of Probability*, pages 1–47, 2009.

[45] Sourav Chatterjee et al. A generalization of the lindeberg principle. *The Annals of Probability*, 34(6):2061–2076, 2006.

[46] Scott Chen and David L Donoho. Examples of basis pursuit. In *Wavelet Applications in Signal and Image Processing III*, volume 2569, pages 564–575. International Society for Optics and Photonics, 1995.

[47] Wei-Kuo Chen. Phase transition in the spiked random tensor with rademacher prior. *arXiv preprint arXiv:1712.01777*, 2017.

[48] Wei-Kuo Chen et al. The aizenman-sims-starr scheme and parisi formula for mixed p-spin spherical models. *Electron. J. Probab*, 18(94):1–14, 2013.

[49] Wei-Kuo Chen, Madeline Handschy, and Gilad Lerman. Phase transition in random tensors with multiple spikes. *arXiv preprint arXiv:1809.06790*, 2018.

[50] Wei-Kuo Chen and Dmitry Panchenko. On the tap free energy in the mixed p-spin models. *Communications in Mathematical Physics*, 362(1):219–252, 2018.

[51] Wei-Kuo Chen and Arnab Sen. Parisi formula, disorder chaos and fluctuation for the ground state energy in the spherical mixed p-spin models. *Communications in Mathematical Physics*, 350(1):129–173, 2017.

[52] François Chollet. keras. https://github.com/fchollet/keras, 2015.

[53] A Crisanti and H-J Sommers. Thouless-anderson-palmer approach to the spherical p-spin spin glass model. *Journal de Physique I*, 5(7):805–813, 1995.

[54] Andrea Crisanti and H-J Sommers. The sphericalp-spin interaction spin glass model: the statics. *Zeitschrift für Physik B Condensed Matter*, 87(3):341–354, 1992.

[55] Alexandre d'Aspremont, Laurent E Ghaoui, Michael I Jordan, and Gert R Lanckriet. A direct formulation for sparse pca using semidefinite programming. In *Advances in neural information processing systems*, pages 41–48, 2005.

[56] Aurelien Decelle, Florent Krzakala, Cristopher Moore, and Lenka Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Physical Review E*, 84(6):066106, 2011.

[57] Bernard Derrida. Random-energy model: Limit of a family of disordered models. *Physical Review Letters*, 45(2):79, 1980.

[58] Bernard Derrida. Random-energy model: An exactly solvable model of disordered systems. *Physical Review B*, 24(5):2613, 1981.

[59] Yash Deshpande, Emmanuel Abbe, and Andrea Montanari. Asymptotic mutual information for the balanced binary stochastic block model. *Information and Inference: A Journal of the IMA*, 6(2):125–170, 2016.

[60] Yash Deshpande and Andrea Montanari. Information-theoretically optimal sparse pca. In *2014 IEEE International Symposium on Information Theory*, pages 2197–2201. IEEE, 2014.

[61] Yash Deshpande and Andrea Montanari. Sparse pca via covariance thresholding. In *Advances in Neural Information Processing Systems*, pages 334–342, 2014.

[62] Yash Deshpande, Subhabrata Sen, Andrea Montanari, and Elchanan Mossel. Contextual stochastic block models. In *Advances in Neural Information Processing Systems*, pages 8590–8602, 2018.

[63] Luc Devroye, László Györfi, and Gábor Lugosi. *A probabilistic theory of pattern recognition*, volume 31. Springer Science & Business Media, 2013.

[64] Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 2016. To appear.

[65] Edgar Dobriban et al. Sharp detection in pca under correlations: all eigenvalues matter. *The Annals of Statistics*, 45(4):1810–1833, 2017.

[66] Edgar Dobriban, William Leeb, and Amit Singer. Pca from noisy, linearly reduced data: the diagonal case. *arXiv preprint arXiv:1611.10333*, 2016.

[67] David Donoho and Andrea Montanari. High dimensional robust m-estimation: Asymptotic variance via approximate message passing. *Probability Theory and Related Fields*, 166(3-4):935–969, 2016.

[68] David L Donoho. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. *Discrete & Computational Geometry*, 35(4):617–652, 2006.

[69] David L Donoho and Iain M Johnstone. Minimax risk over ℓ_p-balls for ℓ_q-error. *Probability Theory and Related Fields*, 99(2):277–303, 1994.

[70] David L Donoho and Jain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *biometrika*, 81(3):425–455, 1994.

[71] David L Donoho, Arian Maleki, and Andrea Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.

[72] David L Donoho, Arian Maleki, and Andrea Montanari. The noise-sensitivity phase transition in compressed sensing. *IEEE Transactions on Information Theory*, 57(10):6920–6941, 2011.

[73] David L Donoho and Jared Tanner. Neighborliness of randomly projected simplices in high dimensions. *Proceedings of the National Academy of Sciences*, 102(27):9452–9457, 2005.

[74] David L Donoho and Jared Tanner. Sparse nonnegative solution of underdetermined linear equations by linear programming. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9446–9451, 2005.

[75] LN Dovbysh and Vladimir Nikolaevich Sudakov. On the gram–de finetti matrices. *Zapiski Nauchnykh Seminarov POMI*, 119:77–86, 1982.

[76] Martin E. Dyer and Alan M. Frieze. The solution of some random np-hard problems in polynomial expected time. *Journal of Algorithms*, 10(4):451–489, 1989.

[77] SF Edwards and Raymund C Jones. The eigenvalue spectrum of a large symmetric random matrix. *Journal of Physics A: Mathematical and General*, 9(10):1595, 1976.

[78] Bradley Efron. The estimation of prediction error: covariance penalties and cross-validation. *Journal of the American Statistical Association*, 99(467):619–632, 2004.

[79] Ahmed El Alaoui and Michael I Jordan. Detection limits in the high-dimensional spiked rectangular model. In *Conference On Learning Theory*, pages 410–438, 2018.

[80] Ahmed El Alaoui and Florent Krzakala. Estimation in the spiked wigner model: A short proof of the replica formula. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 1874–1878. IEEE, 2018.

[81] Noureddine El Karoui. On the impact of predictor geometry on the performance on high-dimensional ridge-regularized generalized robust regression estimators. *Probability Theory and Related Fields*, 170(1-2):95–175, 2018.

[82] William Evans, Claire Kenyon, Yuval Peres, and Leonard J Schulman. Broadcasting on trees and the ising model. *Annals of Applied Probability*, pages 410–433, 2000.

[83] Zhou Fan, Song Mei, and Andrea Montanari. Tap free energy, spin glasses, and variational inference. *arXiv preprint arXiv:1808.07890*, 2018.

[84] Delphine Féral and Sandrine Péché. The largest eigenvalue of rank one deformation of large wigner matrices. *Communications in mathematical physics*, 272(1):185–228, 2007.

[85] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015.

[86] Yan V Fyodorov. Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices. *Physical review letters*, 92(24):240601, 2004.

[87] E Gardner and B Derrida. Optimal storage properties of neural network models. *Journal of Physics A: Mathematical and general*, 21(1):271, 1988.

[88] Elizabeth Gardner. The space of interactions in neural network models. *Journal of physics A: Mathematical and general*, 21(1):257, 1988.

[89] Elizabeth Gardner and Bernard Derrida. Three unfinished works on the optimal storage capacity of networks. *Journal of Physics A: Mathematical and General*, 22(12):1983, 1989.

[90] Stefano Ghirlanda and Francesco Guerra. General properties of overlap probability distributions in disordered spin systems. towards parisi ultrametricity. *Journal of Physics A: Mathematical and General*, 31(46):9149, 1998.

[91] Peter Gillin and David Sherrington. $p > 2$ spin glasses with first-order ferromagnetic transitions. *Journal of Physics A: Mathematical and General*, 33(16):3081, 2000.

[92] Tom Goldstein and Christoph Studer. Phasemax: Convex phase retrieval via basis pursuit. *IEEE Transactions on Information Theory*, 2018.

[93] Yehoram Gordon. Some inequalities for gaussian processes and applications. *Israel Journal of Mathematics*, 50(4):265–289, 1985.

[94] Yehoram Gordon. On milman's inequality and random subspaces which escape through a mesh in rn. In *Geometric Aspects of Functional Analysis*, pages 84–106. Springer, 1988.

[95] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Communications in mathematical physics*, 233(1):1–12, 2003.

[96] Dongning Guo, Shlomo Shamai, and Sergio Verdú. Mutual information and minimum mean-square error in gaussian channels. *IEEE Transactions on Information Theory*, 51(4):1261–1282, 2005.

[97] Dongning Guo, Yihong Wu, Shlomo S Shitz, and Sergio Verdú. Estimation in gaussian noise: Properties of the minimum mean-square error. *IEEE Transactions on Information Theory*, 57(4):2371–2385, 2011.

[98] Géza Györgyi. First-order transition to perfect generalization in a neural network with binary synapses. *Physical Review A*, 41(12):7097, 1990.

[99] D Hansel, G Mato, and C Meunier. Memorization without generalization in a multilayered neural network. *EPL (Europhysics Letters)*, 20(5):471, 1992.

[100] Philip Hartman. *Ordinary Differential Equations*. Society for Industrial and Applied Mathematics, 2002, 1964.

[101] Simon Heimlicher, Marc Lelarge, and Laurent Massoulié. Community detection in the labelled stochastic block model. *arXiv preprint arXiv:1209.2910*, 2012.

[102] Christopher J Hillar and Lek-Heng Lim. Most tensor problems are np-hard. *Journal of the ACM (JACM)*, 60(6):45, 2013.

[103] Paul W Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic blockmodels: First steps. *Social networks*, 5(2):109–137, 1983.

[104] Douglas N Hoover. Relations on probability spaces and arrays of random variables. *Preprint, Institute for Advanced Study, Princeton, NJ*, 2, 1979.

[105] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

[106] Samuel B Hopkins, Tselil Schramm, Jonathan Shi, and David Steurer. Fast spectral algorithms from sum-of-squares proofs: tensor decomposition and planted sparse vectors. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 178–191. ACM, 2016.

[107] Samuel B Hopkins, Jonathan Shi, and David Steurer. Tensor principal component analysis via sum-of-square proofs. In *Conference on Learning Theory*, pages 956–1006, 2015.

[108] Tadaaki Hosaka, Yoshiyuki Kabashima, and Hidetoshi Nishimori. Statistical mechanics of lossy data compression using a nonmonotonic perceptron. *Physical Review E*, 66(6):066126, 2002.

[109] David C Hoyle and Magnus Rattray. Statistical mechanics of learning multiple orthogonal signals: asymptotic theory and fluctuation effects. *Physical review E*, 75(1):016101, 2007.

[110] Aukosh Jagannath, Patrick Lopatto, and Léo Miolane. Statistical thresholds for tensor pca. *arXiv preprint arXiv:1812.03403*, 2018.

[111] Aukosh Jagannath and Ian Tobasco. Low temperature asymptotics of spherical mean field spin glasses. *Communications in Mathematical Physics*, 352(3):979–1017, 2017.

[112] Adel Javanmard and Andrea Montanari. State evolution for general approximate message passing algorithms, with applications to spatial coupling. *Information and Inference*, page iat004, 2013.

[113] Adel Javanmard and Andrea Montanari. Confidence intervals and hypothesis testing for high-dimensional regression. *The Journal of Machine Learning Research*, 15(1):2869–2909, 2014.

[114] Iain M Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Annals of statistics*, pages 295–327, 2001.

[115] Iain M Johnstone. Function estimation and gaussian sequence models. *Unpublished manuscript*, 2(5.3):2, 2002.

[116] Iain M Johnstone and Arthur Yu Lu. Sparse principal components analysis. *Unpublished manuscript*, 7, 2004.

[117] Y Kabashima. Inference from correlated patterns: a unified theory for perceptron learning and linear vector channels. *Journal of Physics: Conference Series*, 95(1):012001, 2008.

[118] Sham M Kakade, Shai Shalev-Shwartz, and Ambuj Tewari. Regularization techniques for learning with matrices. *Journal of Machine Learning Research*, 13(Jun):1865–1890, 2012.

[119] Olav Kallenberg. *Random measures, theory and applications*. Springer, 2017.

[120] Varun Kanade, Elchanan Mossel, and Tselil Schramm. Global and local information in clustering labeled block models. *IEEE Transactions on Information Theory*, 62(10):5906–5917, 2016.

[121] Noureddine El Karoui. Asymptotic behavior of unregularized and ridge-regularized high-dimensional robust regression estimators: rigorous results. arXiv:1311.2445, 2013.

[122] Chiheon Kim, Afonso S Bandeira, and Michel X Goemans. Community detection in hypergraphs, spiked tensor models, and sum-of-squares. In *2017 International Conference on Sampling Theory and Applications (SampTA)*, pages 124–128. IEEE, 2017.

[123] Jeankyung Kim, David Pollard, et al. Cube root asymptotics. *The Annals of Statistics*, 18(1):191–219, 1990.

[124] Satish Babu Korada and Nicolas Macris. Exact solution of the gauge symmetric p-spin glass model on a complete graph. *Journal of Statistical Physics*, 136(2):205–230, 2009.

[125] Satish Babu Korada and Nicolas Macris. Tight bounds on the capacity of binary input random cdma systems. *IEEE Transactions on Information Theory*, 56(11):5590–5613, 2010.

[126] Satish Babu Korada and Andrea Montanari. Applications of the lindeberg principle in communications and statistical learning. *IEEE Transactions on Information Theory*, 57(4):2440–2450, 2011.

[127] Florent Krzakala, Marc Mézard, François Sausset, YF Sun, and Lenka Zdeborová. Statistical-physics-based reconstruction in compressed sensing. *Physical Review X*, 2(2):021005, 2012.

[128] Florent Krzakala, Marc Mézard, Francois Sausset, Yifan Sun, and Lenka Zdeborová. Probabilistic reconstruction in compressed sensing: algorithms, phase diagrams, and threshold achieving matrices. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(08):P08009, 2012.

[129] Florent Krzakala, Jiaming Xu, and Lenka Zdeborová. Mutual information in rank-one matrix estimation. In *Information Theory Workshop (ITW), 2016 IEEE*, pages 71–75. IEEE, 2016.

[130] Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces: isoperimetry and processes*. Springer Science & Business Media, 2013.

[131] Marc Lelarge, Laurent Massoulié, and Jiaming Xu. Reconstruction in the labelled stochastic block model. *IEEE Transactions on Network Science and Engineering*, 2(4):152–163, 2015.

[132] Marc Lelarge and Léo Miolane. Fundamental limits of symmetric low-rank matrix estimation. *Probability Theory and Related Fields*, pages 1–71, 2016.

[133] Thibault Lesieur, Florent Krzakala, and Lenka Zdeborová. MMSE of probabilistic low-rank matrix estimation: Universality with respect to the output channel. In *53rd Annual Allerton Conference on Communication, Control, and Computing, Allerton 2015, Allerton Park & Retreat Center, Monticello, IL, USA, September 29 - October 2, 2015*, pages 680–687, 2015.

[134] Thibault Lesieur, Florent Krzakala, and Lenka Zdeborová. Phase transitions in sparse PCA. In *IEEE International Symposium on Information Theory, ISIT 2015, Hong Kong, China, June 14-19, 2015*, pages 1635–1639, 2015.

[135] Thibault Lesieur, Florent Krzakala, and Lenka Zdeborová. Constrained low-rank matrix estimation: phase transitions, approximate message passing and applications. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(7):073403, 2017.

[136] Thibault Lesieur, Léo Miolane, Marc Lelarge, Florent Krzakala, and Lenka Zdeborová. Statistical and computational phase transitions in spiked tensor estimation. *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 511–515, 2017.

[137] Junjie Ma, Ji Xu, and Arian Maleki. Optimization-based amp for phase retrieval: The impact of initialization and ℓ_2-regularization. *IEEE Transactions on Information Theory*, 2019.

[138] Laurent Massoulié. Community detection thresholds and the weak ramanujan property. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 694–703. ACM, 2014.

[139] Ryosuke Matsushita and Toshiyuki Tanaka. Low-rank matrix reconstruction and clustering via approximate message passing. In *Advances in Neural Information Processing Systems*, pages 917–925, 2013.

[140] Peter McCullagh. *Generalized linear models*. Routledge, 2018.

[141] Cyril Méasson, Andrea Montanari, Thomas J Richardson, and Rüdiger Urbanke. The generalized area theorem and some of its consequences. *IEEE Transactions on Information Theory*, 55(11):4793–4821, 2009.

[142] Marc Mezard. The space of interactions in neural networks: Gardner's computation with the cavity method. *Journal of Physics A: Mathematical and General*, 22(12):2181, 1989.

[143] Marc Mézard, Giorgio Parisi, and Miguel Virasoro. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, volume 9. World Scientific Publishing Co Inc, 1987.

[144] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.

[145] Léo Miolane. Fundamental limits of low-rank matrix estimation: the non-symmetric case. *arXiv preprint arXiv:1702.00473*, 2017.

[146] Léo Miolane and Andrea Montanari. The distribution of the lasso: Uniform control over sparse balls and adaptive parameter tuning. *arXiv preprint arXiv:1811.01212*, 2018.

[147] Marco Mondelli and Andrea Montanari. Fundamental limits of weak recovery with applications to phase retrieval. *Foundations of Computational Mathematics*, pages 1–71, 2017.

[148] Andrea Montanari. Estimating random variables from random sparse observations. *European Transactions on Telecommunications*, 19(4):385–403, 2008.

[149] Andrea Montanari. Finding one community in a sparse graph. *Journal of Statistical Physics*, 161(2):273–299, 2015.

[150] Andrea Montanari, Daniel Reichman, and Ofer Zeitouni. On the limitation of spectral methods: From the gaussian hidden clique problem to rank-one perturbations of gaussian tensors. In *Advances in Neural Information Processing Systems*, pages 217–225, 2015.

[151] Andrea Montanari and Emile Richard. Non-negative principal component analysis: Message passing algorithms and sharp asymptotics. *IEEE Transactions on Information Theory*, 62(3):1458–1484, 2016.

[152] Andrea Montanari and Ramji Venkataramanan. Estimation of low-rank matrices via approximate message passing. *arXiv preprint arXiv:1711.01682*, 2017.

[153] Elchanan Mossel. Survey-information flow on trees. *DIMACS series in discrete mathematics and theoretical computer science*, 63:155–170, 2004.

[154] Elchanan Mossel, Joe Neeman, and Allan Sly. A proof of the block model threshold conjecture. *arXiv preprint arXiv:1311.4115*, 2013.

[155] Elchanan Mossel, Joe Neeman, and Allan Sly. Reconstruction and estimation in the planted partition model. *Probability Theory and Related Fields*, 162(3-4):431–461, 2015.

[156] Elchanan Mossel, Yuval Peres, et al. Information flow on trees. *The Annals of Applied Probability*, 13(3):817–844, 2003.

[157] Elchanan Mossel and Jiaming Xu. Density evolution in the degree-correlated stochastic block model. In *29th Annual Conference on Learning Theory*, pages 1319–1356, 2016.

[158] Jean-Christophe Mourrat. Hamilton-jacobi equations for mean-field disordered systems. *arXiv preprint arXiv:1811.01432*, 2018.

[159] Ali Mousavi, Arian Maleki, and Richard G Baraniuk. Consistent parameter estimation for lasso and approximate message passing. *The Annals of Statistics*, 45(6):2427–2454, 2017.

[160] Boaz Nadler. Finite sample approximation results for principal component analysis: A matrix perturbation approach. *The Annals of Statistics*, pages 2791–2817, 2008.

[161] Joe Neeman and Praneeth Netrapalli. Non-reconstructability in the stochastic block model. *arXiv preprint arXiv:1404.6304*, 2014.

[162] Sahand N Negahban, Pradeep Ravikumar, Martin J Wainwright, Bin Yu, et al. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. *Statistical Science*, 27(4):538–557, 2012.

[163] Hidetoshi Nishimori. Exact results and critical properties of the ising model with competing interactions. *Journal of Physics C: Solid State Physics*, 13(21):4071, 1980.

[164] Hidetoshi Nishimori. Internal energy, specific heat and correlation function of the bond-random ising model. *Progress of Theoretical Physics*, 66(4):1169–1181, 1981.

[165] Albert B Novikoff. On convergence proofs for perceptrons. Technical report, STANFORD RESEARCH INST MENLO PARK CA, 1963.

[166] Alexei Onatski, Marcelo J Moreira, Marc Hallin, et al. Asymptotic power of sphericity tests for high-dimensional data. *The Annals of Statistics*, 41(3):1204–1231, 2013.

[167] Manfred Opper and David Haussler. Generalization performance of bayes optimal classification algorithm for learning a perceptron. *Physical Review Letters*, 66(20):2677, 1991.

[168] Samet Oymak, Amin Jalali, Maryam Fazel, Yonina C Eldar, and Babak Hassibi. Simultaneously structured models with application to sparse and low-rank matrices. *IEEE Transactions on Information Theory*, 61(5):2886–2908, 2015.

[169] Ashkan Panahi and Babak Hassibi. A universal analysis of large-scale regularized least squares solutions. In *Advances in Neural Information Processing Systems*, pages 3381–3390, 2017.

[170] Dmitry Panchenko. *The Sherrington-Kirkpatrick model*. Springer Science & Business Media, 2013.

[171] Dmitry Panchenko et al. On the dovbysh-sudakov representation result. *Electronic Communications in Probability*, 15:330–338, 2010.

[172] Debashis Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica*, pages 1617–1642, 2007.

[173] Sandrine Péché. The largest eigenvalue of small rank perturbations of hermitian random matrices. *Probability Theory and Related Fields*, 134(1):127–173, 2006.

[174] Amelia Perry, Alexander S Wein, and Afonso S Bandeira. Statistical limits of spiked tensor models. *arXiv preprint arXiv:1612.07728*, 2016.

[175] Amelia Perry, Alexander S Wein, Afonso S Bandeira, Ankur Moitra, et al. Optimality and sub-optimality of pca i: Spiked random matrix models. *The Annals of Statistics*, 46(5):2416–2451, 2018.

[176] Sundeep Rangan. Generalized approximate message passing for estimation with random linear mixing. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pages 2168–2172. IEEE, 2011.

[177] Sundeep Rangan and Alyson K Fletcher. Iterative estimation of constrained rank-one matrices in noise. In *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pages 1246–1250. IEEE, 2012.

[178] Garvesh Raskutti, Martin J Wainwright, and Bin Yu. Minimax rates of estimation for high-dimensional linear regression over ℓ_q-balls. *IEEE Transactions on Information Theory*, 57(10):6976–6994, 2011.

[179] Galen Reeves and Henry D Pfister. The replica-symmetric prediction for compressed sensing with gaussian matrices is exact. In *Information Theory (ISIT), 2016 IEEE International Symposium on*, pages 665–669. IEEE, 2016.

[180] Emile Richard and Andrea Montanari. A statistical model for tensor pca. In *Advances in Neural Information Processing Systems*, pages 2897–2905, 2014.

[181] Tom Richardson and Ruediger Urbanke. *Modern coding theory*. Cambridge university press, 2008.

[182] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.

[183] Valentina Ros, Gerard Ben Arous, Giulio Biroli, and Chiara Cammarota. Complex energy landscapes in spiked-tensor and simple glassy models: Ruggedness, arrangements of local minima, and phase transitions. *Physical Review X*, 9(1):011003, 2019.

[184] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

[185] Frank Rosenblatt. Principles of neurodynamics. *CORNELL AERONAUTICAL LAB INC BUFFALO*, 1962.

[186] Alaa Saade, Florent Krzakala, Marc Lelarge, and Lenka Zdeborová. Fast randomized semi-supervised clustering. In *Journal of Physics: Conference Series*, volume 1036, page 012015. IOP Publishing, 2018.

[187] Alaa Saade, Marc Lelarge, Florent Krzakala, and Lenka Zdeborová. Clustering from sparse pairwise measurements. In *Information Theory (ISIT), 2016 IEEE International Symposium on*, pages 780–784. IEEE, 2016.

[188] Philip Schniter and Sundeep Rangan. Compressive phase retrieval via generalized approximate message passing. *IEEE Transactions on Signal Processing*, 63(4):1043–1055, 2015.

[189] HS Seung, Haim Sompolinsky, and Naftali Tishby. Statistical mechanics of learning from examples. *Physical review A*, 45(8):6056, 1992.

[190] Mariya Shcherbina and Brunello Tirozzi. Rigorous solution of the gardner problem. *Communications in mathematical physics*, 234(3):383–422, 2003.

[191] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.

[192] Mahdi Soltanolkotabi. Structured signal recovery from quadratic measurements: Breaking sample complexity barriers via nonconvex optimization. *arXiv preprint arXiv:1702.06175*, 2017.

[193] Aart J Stam. Some inequalities satisfied by the quantities of information of fisher and shannon. *Information and Control*, 2(2):101–112, 1959.

[194] Charles M Stein. Estimation of the mean of a multivariate normal distribution. *The annals of Statistics*, pages 1135–1151, 1981.

[195] Mihailo Stojnic. Recovery thresholds for ℓ_1 optimization in binary compressed sensing. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 1593–1597. IEEE, 2010.

[196] Mihailo Stojnic. A framework to characterize performance of lasso algorithms. *arXiv preprint arXiv:1303.7291*, 2013.

[197] Eliran Subag et al. The complexity of spherical p-spin models—a second moment approach. *The Annals of Probability*, 45(5):3385–3450, 2017.

[198] Pragya Sur and Emmanuel J Candès. A modern maximum-likelihood theory for high-dimensional logistic regression. arXiv:1803.06964, 2018.

[199] Takashi Takahashi and Yoshiyuki Kabashima. A statistical mechanics approach to de-biasing and uncertainty estimation in lasso for random measurements. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(7):073405, 2018.

[200] Michel Talagrand. Free energy of the spherical mean field model. *Probability theory and related fields*, 134(3):339–382, 2006.

[201] Michel Talagrand. *Mean field models for spin glasses: Volume I: Basic examples*, volume 54. Springer Science & Business Media, 2010.

[202] Michel Talagrand. *Mean field models for spin glasses: Volume II: Advanced Replica-Symmetry and Low Temperature*, volume 55. Springer Science & Business Media, 2011.

[203] David J Thouless, Philip W Anderson, and Robert G Palmer. Solution of'solvable model of a spin glass'. *Philosophical Magazine*, 35(3):593–601, 1977.

[204] Christos Thrampoulidis, Ehsan Abbasi, and Babak Hassibi. Precise error analysis of regularized m-estimators in high-dimensions. *IEEE Transactions on Information Theory*, 2018.

[205] Christos Thrampoulidis, Samet Oymak, and Babak Hassibi. Regularized linear regression: A precise analysis of the estimation error. In *Conference on Learning Theory*, pages 1683–1709, 2015.

[206] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[207] Ryan J Tibshirani, Jonathan Taylor, et al. Degrees of freedom in lasso problems. *The Annals of Statistics*, 40(2):1198–1232, 2012.

[208] Joel A Tropp. Convex recovery of a structured signal from independent random linear measurements. In *Sampling Theory, a Renaissance*, pages 67–101. Springer, 2015.

[209] Sara Van de Geer, Peter Bühlmann, Ya'acov Ritov, Ruben Dezeure, et al. On asymptotically optimal confidence regions and tests for high-dimensional models. *The Annals of Statistics*, 42(3):1166–1202, 2014.

[210] Sara A Van De Geer, Peter Bühlmann, et al. On the conditions used to prove oracle results for the lasso. *Electronic Journal of Statistics*, 3:1360–1392, 2009.

[211] Remco Van Der Hofstad. *Random graphs and complex networks*. Cambridge University Press, 2016.

[212] Jeremy P Vila and Philip Schniter. Expectation-maximization gaussian-mixture approximate message passing. *IEEE Transactions on Signal Processing*, 61(19):4658–4672, 2013.

[213] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[214] Martin J Wainwright, Michael I Jordan, et al. Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning*, 1(1–2):1–305, 2008.

[215] Yihong Wu and Sergio Verdú. Mmse dimension. *IEEE Transactions on Information Theory*, 57(8):4857–4879, 2011.

[216] Yihong Wu and Sergio Verdú. Functional properties of minimum mean-square error and mutual information. *IEEE Transactions on Information Theory*, 58(3):1289–1301, 2012.

[217] Yihong Wu and Sergio Verdú. Optimal phase transitions in compressed sensing. *IEEE Transactions on Information Theory*, 58(10):6241–6263, 2012.

[218] Yingying Xu, Yoshiyuki Kabashima, and Lenka Zdeborová. Bayesian signal reconstruction for 1-bit compressed sensing. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(11):P11015, 2014.

[219] Constantin Zalinescu. *Convex analysis in general vector spaces*. World scientific, 2002.

[220] Lenka Zdeborová and Florent Krzakala. Statistical physics of inference: Thresholds and algorithms. *Advances in Physics*, 65(5):453–552, 2016.

[221] Cun-Hui Zhang and Stephanie S Zhang. Confidence intervals for low dimensional parameters in high dimensional linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):217–242, 2014.

[222] Justin Ziniel, Philip Schniter, and Per Sederberg. Binary linear classification and feature selection via generalized approximate message passing. In *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, pages 1–6. IEEE, 2014.

[223] Hui Zou, Trevor Hastie, and Robert Tibshirani. Sparse principal component analysis. *Journal of computational and graphical statistics*, 15(2):265–286, 2006.

RÉSUMÉ

Nous étudions des problèmes statistiques classiques, tels que la détection de communautés dans un graphe, l'analyse en composantes principales, les modèles de mélanges Gaussiens, les modèles linéaires (généralisés ou non), dans un cadre Bayésien. Nous calculons pour ces problèmes le "risque de Bayes" qui est la plus petite erreur atteignable par une méthode statistique, dans la limite de grande dimension. Nous observons alors un phénomène surprenant: dans de nombreux cas il existe une valeur critique de l'intensité du bruit au-delà de laquelle il n'est plus possible d'extraire de l'information des données. En dessous de ce seuil, nous comparons la performance d'algorithmes polynomiaux à celle optimale. Dans de nombreuses situations nous observons un écart: bien qu'il soit possible – en théorie – d'estimer le signal, aucune méthode algorithmiquement efficace ne parvient à être optimale.

Dans ce manuscrit, nous adoptons une approche issue de la physique statistique qui explique ces phénomènes en termes de transitions de phase. Les méthodes et outils que nous utilisons proviennent donc de la physique, plus précisément de l'étude mathématique des verres de spins.

MOTS CLÉS

inférence statistique, théorie de l'information, physique statistique, verres de spin, détection de communauté dans des graphs, estimation de structure de faible rang, modèle linéaire généralisé, Lasso

ABSTRACT

We study classical statistical problems such as as community detection on graphs, Principal Component Analysis (PCA), sparse PCA, Gaussian mixture clustering, linear and generalized linear models, in a Bayesian framework. We compute the best estimation performance (often denoted as "Bayes Risk") achievable by any statistical method in the high dimensional regime. This allows to observe surprising phenomena: for many problems, there exists a critical noise level above which it is impossible to estimate better than random guessing. Below this threshold, we compare the performance of existing polynomial-time algorithms to the optimal one and observe a gap in many situations: even if non-trivial estimation is theoretically possible, computationally efficient methods do not manage to achieve optimality.

From a statistical physics point of view that we adopt throughout this manuscript, these phenomena can be explained by phase transitions. The tools and methods of this thesis are therefore mainly issued from statistical physics, more precisely from the mathematical study of spin glasses.

KEYWORDS

statistical inference, information theory, statistical physics, spin glasses, community detection on graphs, low-rank estimation, generalized linear models, Lasso