Some remarks on the effect of the Random Batch Method on phase transition

Arnaud Guillin, Pierre Le Bris, Pierre Monmarché

Some remarks on the effect of the Random Batch Method on phase transition

Arnaud Guillin[*], Pierre Le Bris[†] and Pierre Monmarché[‡]

In memory of Francis Comets

Abstract

In this article, we focus on two toy models : the *Curie-Weiss* model and the system of N particles in linear interactions in a *double well confining potential*. Both models, which have been extensively studied, describe a large system of particles with a mean-field limit that admits a phase transition. We are concerned with the numerical simulation of these particle systems. To deal with the quadratic complexity of the numerical scheme, corresponding to the computation of the $O(N^2)$ interactions per time step, the *Random Batch Method* (RBM) has been suggested. It consists in randomly (and uniformly) dividing the particles into batches of size $p > 1$, and computing the interactions only within each batch, thus reducing the numerical complexity to $O(Np)$ per time step. The convergence of this numerical method has been proved in other works.

This work is motivated by the observation that the RBM, via the random constructions of batches, artificially adds noise to the particle system. The goal of this article is to study the effect of this added noise on the phase transition of the nonlinear limit, and more precisely we study the *effective dynamics* of the two models to show how a phase transition may still be observed with the RBM but at a lower critical temperature.

1 Introduction

1.1 Motivation

Consider a system of N particles $(X^i)_{i\in\{1,...,N\}}$ in interaction

$$dX_t^i = -\nabla U(X_t^i)dt - \frac{1}{N-1}\sum_{j\neq i}\nabla W(X_t^i - X_t^j)dt + \sqrt{2\sigma}dB_t^i, \tag{IPS}$$

where for all $i \in \{1, ..., N\}$ and $t \geq 0$ we have $X_t^i \in \mathbb{R}^d$, U and W are two twice continuously differentiable functions, respectively called *confining potential* and *interaction potential*, $\sigma > 0$ is a *diffusion coefficient* or *temperature*, and $(B^i)_i$ are independent d-dimensional Brownian motions. The name (IPS) refers to *Interacting Particle System*.

It is well known (see [CD22a, CD22b] and references therein) that, under suitable assumptions on U and W, the particle system (IPS) converges as $N \to \infty$ towards its nonlinear mean-field limit, a stochastic differential equation (SDE) of *McKean-Vlasov* type

$$\begin{cases} d\bar{X}_t = -\nabla U(\bar{X}_t)dt - \nabla W * \bar{\rho}_t(\bar{X}_t)dt + \sqrt{2\sigma}dB_t, \\ \bar{\rho}_t = \text{Law}(\bar{X}_t). \end{cases} \tag{NL}$$

[*]Laboratoire de Mathématiques Blaise Pascal - Université Clermont-Auvergne and Institut Universitaire de France. Email : arnaud.guillin[AT]uca.fr

[†]Laboratoire Jacques-Louis Lions - Sorbonne Université. Email : pierre.lebris[AT]sorbonne-universite.fr

[‡]Laboratoire Jacques-Louis Lions - Sorbonne Université. Email : pierre.monmarche[AT]sorbonne-universite.fr

Here, the name (NL) refers to *Nonlinear Limit*, and this equation arises in the modelling of granular media [CGM08].

The quantitative link between of (IPS) and (NL) can be exploited in various ways. On one hand, as it was historically motivated, the study of (way too) large systems of particles cannot be feasible, and boiling it down to the study of the nonlinear limit yields exploitable results. On the other hand, one can see (IPS) as an approximation of (NL), and in particular an approximation that can be numerically simulated. Consider the *Euler-Maruyama* scheme associated to (IPS) with a timestep $\delta > 0$

$$\begin{cases} X_{t+1}^{i,\delta} = X_t^{i,\delta} - \delta \nabla U(X_t^{i,\delta}) - \frac{\delta}{N-1} \sum_{j \neq i} \nabla W(X_t^{i,\delta} - X_t^{j,\delta}) + \sqrt{2\sigma\delta} G_t^i, \\ G_t^i \text{ i.i.d } \sim \mathcal{N}(0,1), \quad t \in \mathbb{N}. \end{cases} \tag{D-IPS}$$

Its name (D-IPS) comes from *Discrete - Interacting Particle System*. The convergence of (D-IPS) towards (NL) has been extensively studied : with bounded Lipschitz coefficients [BT97], with Hölder continuous coefficients [BH19], non-Lipschitz coefficients [DQ21]. The quantitative convergence of the implicit Euler-Maruyama scheme can also be found in [Mal03].

Notice that this numerical scheme requires $O(N^2)$ operations per time step, corresponding to the total number of interactions of pairs $(i,j)_{i,j \in \{1,...,N\}}$. To cope with this possibly limiting complexity, several works have suggested using the *Random Batch Method* (RBM) (see for instance [JLL20]), motivated by the Stochastic Gradient Langevin Dynamics [WT11].

Consider, for a time step $t \in \mathbb{N}$, a partition $\mathcal{P}_t = \left(\mathcal{P}_t^1, ..., \mathcal{P}_t^{N/p}\right)$ of $\{1, ..., N\}$ into N/p subsets of size $p > 1$, assuming for the sake of simplicity that N is a multiple of p, and define

$$\mathcal{C}_t^i = \left\{ j \in \{1, .., N\} \text{ s.t. } \exists l \in \{1, ..., N/p\}, i, j \in \mathcal{P}_t^l \right\}. \tag{1.1}$$

In other words, $\mathcal{C}_t^i$ is the set of indexes that are in the same subset as i at time step t, with the convention $i \in \mathcal{C}_t^i$. We now consider the following numerical scheme

$$\begin{cases} Y_{t+1}^{i,\delta,p} = Y_t^{i,\delta,p} - \delta \nabla U(Y_t^{i,\delta,p}) - \frac{\delta}{p-1} \sum_{j \in \mathcal{C}_t^i \setminus \{i\}} \nabla W(Y_t^{i,\delta,p} - Y_t^{j,\delta,p}) + \sqrt{2\sigma\delta} G_t^i, \\ G_t^i \text{ i.i.d } \sim \mathcal{N}(0,1), \quad i \in \{1, ..., N\}, \quad t \in \mathbb{N}, \end{cases} \tag{D-RB-IPS}$$

where for each time step t the partition $\mathcal{P}_t$ is random and each partition has the same probability of occurring. The name (D-RB-IPS) refers to *Discrete - Random Batch - Interacting Particle System*. The convergence of (D-RB-IPS) towards (NL) can be found in [JLL21, JLYZ22, YZ22].

The idea of using random batches has been shown to be efficient for computing the evolution of large interacting system of quantum particles [GJP21], of particles with Coulomb interactions in molecular dynamics [JLXZ21], but also for Markov Chain Monte Carlo [LXZ20], or for solving PDEs [CJT22, LLT22]. See also references therein.

The starting point of this work is the following observation : the RBM, via the random construction of a partition of $\{1, ..., N\}$, artificially adds noise (or temperature) to a system. We thus ask the following question :

Does the critical temperature of (the mean-field limit of) a system of interacting particles admitting a phase transition decrease when considering a version with random batches ? If so, can we quantify it ?

To partially answer this question, we focus on two specific types of particle systems for which the mean-field limit admits a phase transition : the first one is the *Curie-Weiss* model and the second one is the system (IPS) in dimension 1 with attractive and quadratic interaction potential W and the *double well confining potential* U.

The nonlinear mean-field limits of both models admit, as we will discuss, a phase transition occurring at a certain critical parameter. We consider a version with random batches of size p of each system, consider the limit as $N \to \infty$ (with fixed p) towards a nonlinear model, and then study the phase transition of said limit.

1.2 The Curie-Weiss model

The classical system. The Curie-Weiss model is, and it is the reason we start by studying it, arguably one of the most simple system admitting a phase transition. Consider N spins, given by a configuration $\sigma = (\sigma_1, ..., \sigma_N)$, and $\Omega_N = \{-1, 1\}^N$ the set of possible configurations for the system. On this system we consider the following Hamiltonian

$$\forall \sigma \in \Omega_N, \quad H_N(\sigma) = -\frac{1}{2N} \sum_{i,j} \sigma_i \sigma_j. \tag{1.2}$$

Intuitively, each spin will tend to align with the others. It is a mean field model as H_N only depends in reality on the mean magnetization $m_N(\sigma) := \frac{1}{N} \sum_{i=1}^N \sigma_i$, by

$$H_N(\sigma) = -\frac{N}{2} m_N(\sigma)^2.$$

The evolution for $(\sigma(n))_{n\geq 0}$ in Ω_N is the following : at each discrete time step, a spin is chosen uniformly among the N possible spins. Let us denote i this spin, and $\sigma' = (\sigma'_1, ..., \sigma'_N)$ the configuration such that for all $j \neq i$, $\sigma'_j = \sigma(n)_j$, and $\sigma'_i = -\sigma(n)_i$. We accept σ' as the next step of $\sigma(n)$ with probability $\exp\left(-\beta(H_N(\sigma') - H_N(\sigma))_+\right)$ (i.e if the Hamiltonian decreases then with probability 1, otherwise with a positive probability depending on a parameter β), otherwise the system remains at $\sigma(n)$. Here we use the notation $x_+ = \max(x, 0)$. This parameter β is known as the *inverse temperature*. This yields the following transition probabilities for the Markov chain $(\sigma(n))_{n\geq 0}$:

$$p(\sigma, \sigma') = \begin{cases} \frac{1}{N} \exp\left(-\beta(H_N(\sigma') - H_N(\sigma))_+\right) & \text{if } ||\sigma - \sigma'||_1 = 2 \\ 0 & \text{if } ||\sigma - \sigma'||_1 > 2 \\ 1 - \sum_{\eta \neq \sigma} p(\sigma, \eta) & \text{if } \sigma' = \sigma \end{cases}$$

This dynamics $(\sigma(n))_{n\geq 0}$, which is an irreducible and aperiodic Markov chain on a finite state space Ω_N, is reversible with respect to the Gibbs measure

$$\mu_{\beta,N}(\sigma) = \frac{1}{Z_{\beta,N}} \exp(-\beta H_N(\sigma)), \tag{1.3}$$

where $Z_{\beta,N}$ is a normalizing constant. Instead of studying the dynamics of σ, we look at the mean magnetization $m_N(n) = m_N(\sigma(n))$, which is still a Markov chain. This quantity, at each time step, can only increase or decrease by $\frac{2}{N}$, and the transition probabilities are given by

$$r(m, m') = \begin{cases} \frac{1-m}{2} \exp\left(-\frac{\beta N}{2}(m^2 - m'^2)_+\right) & \text{if } m' = m + \frac{2}{N} \\ \frac{1+m}{2} \exp\left(-\frac{\beta N}{2}(m^2 - m'^2)_+\right) & \text{if } m' = m - \frac{2}{N} \\ 1 - r\left(m, m + \frac{2}{N}\right) - r\left(m, m - \frac{2}{N}\right) & \text{if } m' = m \\ 0 & \text{otherwise.} \end{cases} \tag{1.4}$$

Likewise, this dynamics is reversible with respect to the Gibbs measure

$$\nu_{\beta,N}(m) = \frac{1}{Z_{\beta,N}} \binom{N}{\frac{1+m}{2}N} \exp\left(\frac{\beta N m^2}{2}\right).$$

Many works (see for instance [CK17, EN78, LLP10], the classical reference that is Chapter 4 of [Ell85] or more recently Chapter 2 of [FV17]) have studied Large Deviation Principles for this system, and have shown that there exists a critical inverse temperature $\beta_c = 1$. For the sake of completeness, and because the method will be similar in the case with random batches, we give a proof in Section 2.1 of the phase transition

happening in the following sense : the process $M_t^{(N)} = m_N(\lfloor Nt \rfloor)$ weakly converges to the solution of an ordinary differential equation (ODE). For $\beta > 1$, the limit ODE admits three equilibrium states, and for $\beta \leq 1$ only one. In both cases, 0 is an equilibrium state, and is stable in the case $\beta \leq 1$ and unstable in the case $\beta > 1$. Phase transition for the maximum likelihood estimator of the parameters has also been studied in [CG91].

The Curie-Weiss model with random batches. We then consider the same system, but using the Random Batch Method. At each time step, the chosen spin no longer evolves according to the entire system, but according to a subset of p spins containing the chosen spin.

We thus consider a new evolution for $(\sigma^p(n))_{n\geq 0}$ in Ω_N, where σ^p denotes the new sequence of spin configurations. At each discrete time step, a spin is chosen uniformly among the N possible spins. Let us denote it i, and $\sigma' = (\sigma'_1, ..., \sigma'_N)$ the configuration such that for all $j \neq i$, $\sigma'_j = \sigma^p(n)_j$, and $\sigma'_i = -\sigma^p(n)_i$. We then sample a subset of $\{1, ..., N\}$ of size p containing i, denoted $\mathcal{C}^{i,p}$, uniformly over such subsets, and accept σ' as the next step of $\sigma^p(n)$ with probability $\exp\left(-\beta(H_{N,p}(\sigma', \mathcal{C}^{i,p}) - H_{N,p}(\sigma^p(n), \mathcal{C}^{i,p}))_+\right)$, where

$$H_{N,p}(\sigma, \mathcal{C}^{i,p}) = -\frac{1}{2p} \sum_{j,k \in \mathcal{C}^{i,p}} \sigma_j \sigma_k. \tag{1.5}$$

Likewise, we may study this system in terms of its magnetization, denoted $(m_{N,p}(n))_n$, for which we can explicitly write the transition probabilities (see Lemma 2.1).

This system resembles to some extent the *dilute Curie-Weiss model* [BMP21], in which the spins interact according to an Erdős-Rényi random graph with edge probability $\tilde{p} = \frac{p}{N} \in]0, 1[$, the main difference being that the "graph", in our case, is modified at each time step and there are exactly $p - 1$ spins interacting with a given one.

Studying the Curie-Weiss model with random batches, which is done in Section 2.2, yields the following results.

Theorem 1. *Let $p \in \mathbb{N} \setminus \{0, 1\}$ and $\beta > 0$.*

- *Define*

$$S_1^{p,\beta}(m) = \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^k \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}$$

$$S_2^{p,\beta}(m) = \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^k \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+},$$

$$f_p(\beta, m) = \left(S_1^{p,\beta}(m) - S_2^{p,\beta}(m)\right) - m\left(S_1^{p,\beta}(m) + S_2^{p,\beta}(m)\right).$$

  The process $M_t^{(N,p)} = m_{N,p}(\lfloor Nt \rfloor)$, i.e the magnetization rescaled in time, weakly converges as $N \to \infty$ to the solution of the ODE

$$\frac{d}{dt} m(t) = f_p(\beta, m(t)). \tag{1.6}$$

  For all $\beta > 0$, 0 is an equilibrium state for the solution of (1.6)*.*

- *For $p \in \{2, 3\}$, 0 is the unique equilibrium state, and it is stable.*

- *For $p \geq 4$, there exists $\beta_{c,p}$ such that for all $\beta > \beta_{c,p}$, the equilibrium state 0 is unstable, and for all $\beta \leq \beta_{c,p}$ it is stable. Furthermore, we have the estimate*

$$\beta_{c,p} = 1 + \sqrt{\frac{2}{p\pi}} + o\left(\frac{1}{\sqrt{p}}\right). \tag{1.7}$$

This theorem thus gives a first answer to the main question of the article : the RBM does increase the critical inverse temperature of the system (i.e decreases the critical temperature).

1.3 Numerical scheme and double-well potential

We then go back to the initial motivation concerning numerical schemes for interacting particle systems.

The effective dynamics. Just like we may consider the nonlinear limit of (IPS), we may also consider the limit as $N \to \infty$ of (D-RB-IPS). Define

$$\begin{cases} \bar{Y}_{t+1}^{\delta,p} = \bar{Y}_t^{\delta,p} - \delta\nabla U(\bar{Y}_t^{\delta,p}) - \frac{\delta}{p-1}\sum_{j=1}^{p-1}\nabla W(\bar{Y}_t^{\delta,p} - Y^j) + \sqrt{2\sigma\delta}G_t, \\ G_t \text{ i.i.d } \sim \mathcal{N}(0,1), \quad (Y^j)_j \text{ i.i.d } \sim \bar{\rho}_t^{\delta,p} := \text{Law}(\bar{Y}_t^{\delta,p}). \end{cases} \tag{D-RB-NL}$$

The name (D-RB-NL) stands for *Discrete - Random Batch - Nonlinear Limit*. The convergence of (D-RB-IPS) towards (D-RB-NL) can be found in [JL22]. The proof relies on a coupling method, noticing that, as $N \to \infty$, the probability of constructing batches of fixed size p in (D-RB-IPS) with independent and identically distributed particles goes to 1, thus giving a convergence in total variation distance.

We then, in the spirit of [SS21], construct a continuous process, parameterized by the timestep and the batch size, which is closer to the numerical scheme (D-RB-IPS) than the target (NL). In the dynamics of (D-RB-NL), writing

$$\xi_t = \frac{1}{p-1}\sum_{j=1}^{p-1}\nabla W\left(\bar{Y}_t^{\delta,p} - Y^j\right)$$

notice that

$$\mathbb{E}\left(\xi_t \middle| \bar{Y}_t^{\delta,p}\right) = \nabla W * \bar{\rho}_t^{\delta,p}(\bar{Y}_t^{\delta,p}),$$

and

$$\begin{aligned} \text{Var}\left(\xi_t \middle| \bar{Y}_t^{\delta,p}\right) =& \frac{1}{p-1}\text{Var}_{\bar{\rho}_t^{\delta,p}}\left(\nabla W(\bar{Y}_t^{\delta,p} - \cdot)\middle|\bar{Y}_t^{\delta,p}\right) \\ =& \frac{1}{p-1}\left((\nabla W)^2 * \bar{\rho}_t^{\delta,p}(\bar{Y}_t^{\delta,p}) - (\nabla W * \bar{\rho}_t^{\delta,p}(\bar{Y}_t^{\delta,p}))^2\right), \end{aligned}$$

where the square of a vector has to be understood component-wise. Hence,

$$\bar{Y}_t^{\delta,p} = \bar{Y}_0^{\delta,p} - \delta\sum_{s=0}^{t-1}\nabla U(\bar{Y}_s^{\delta,p}) - \delta\sum_{s=0}^{t-1}\nabla W * \bar{\rho}_s^{\delta,p}(\bar{Y}_s^{\delta,p}) - \delta M_t + \sqrt{2\sigma\delta}\sum_{s=0}^{t-1}G_s\,,$$

where

$$t \mapsto M_t := \sum_{s=0}^{t-1}\left(\xi_s - \nabla W * \bar{\rho}_s^{\delta,p}(\bar{Y}_s^{\delta,p})\right)$$

is a martingale. By martingale CLT, we thus expect the numerical scheme (D-RB-IPS) to be close, for small values of δ, the following non-linear SDE, that we call the *effective dynamics*:

$$\begin{cases} d\bar{X}_t^{e,\delta,p} = -\nabla U(\bar{X}_t^{e,\delta,p})dt - \nabla W * \bar{\rho}_t^{e,\delta,p}(\bar{X}_t^{e,\delta,p})dt + \left(2\sigma + \frac{\delta}{p-1}\Sigma(\bar{X}_t^{e,\delta,p}, \bar{\rho}_t^{e,\delta,p})\right)^{1/2} dB_t, \\ \bar{\rho}_t^{e,\delta,p} = \text{Law}(\bar{X}_t^{e,\delta,p}), \end{cases} \tag{Eff}$$

where we denote $\Sigma(x,\rho) = (\nabla W)^2 * \rho(x) - (\nabla W * \rho(x))^2$. Notice that, although it is a continuous-time process, it depends on the stepsize δ of the numerical schemes.

Such dynamics are also known as modified equations in various works considering the backward error analysis of SDEs [Sha06, Zyg11], improving upon a technique that had already provided a better understanding of the numerical methods for ODEs. They have been used in the numerical error analysis of the Stochastic Gradient Langevin Dynamics [VZT16, SS21]. Of course, these references do not consider nonlinear SDEs as we do, and obtaining a formal result in our case is out of the scope of the present work. Let us informally and briefly explain the motivation of the effective dynamics (we refer to [VZT16, SS21] and references within for further details). In the usual stochastic gradient case (which would correspond to (D-RB-NL) where we assume that the law of Y^j is fixed), denote by $\pi_{\delta,p}$, π, π_δ and π_{eff}, respectively, the invariant measures of (D-RB-NL), of the continuous-time limit process (NL), of its Euler scheme (without Random batches, i.e. (D-RB-NL) with $p = \infty$) and of (Eff). For a fixed observable f, from the weak error analysis on the invariant measure (see e.g. [SS21, Proposition 1])), we get that there exists $c_1, c_2 \in \mathbb{R}$ such that $\pi_{\delta,p}(f) \simeq \pi(f) + c_1\delta + c_2\delta/p$, while $\pi_\delta(f) \simeq \pi(f) + c_1\delta$ and $\pi_{eff}(f) \simeq \pi(f) + c_2\delta/p$, where these approximations are all up to a term of order $\delta^2(1 + p^{-3/2})$. In other words, at first order in δ, c_1 and c_2 respectively accounts for the time discretization and stochastic gradient errors. By studying (Eff), at first order, we disregard the error which is purely due to the time discretization and focus on the contribution of the stochastic gradient approximation. Notice that, for the Euler scheme of the overdamped Langevin diffusion (D-RB-NL), except if the variance Σ is very large (which corresponds to the case in [VZT16, SS21] which are not concerned with a mean-field scaling) these two parts of the error are of the same order in δ. However, in practice, second-order schemes for underdamped Langevin or Hamiltonian Monte Carlo are widely used (as in [SS21, GLBMM22]) and in that case the stochastic gradient contribution is the leading term of the bias (see Remark 1.2 below). In any cases, the numerical scheme is closer to the effective dynamics than it is to the continuous-time process (as they only differ, at first order, through the pure discretization error), which motivates in the following the analysis of the effective dynamics (Eff).

Again, we emphasize that providing a quantitative link between the various processes (IPS), (D-IPS), (D-RB-IPS), (D-RB-NL), (Eff), and (NL) would require an entire separate analysis, even though some results are already known. As it would dilute the main message of this work concerning the phase transition of the effective dynamics, we do not address this question here.

The double well confining potential. We now choose in (NL) the dimension to be $d = 1$ and the potentials

$$U(x) = \frac{x^4}{4} - \frac{x^2}{2}, \quad W(x) = L_W \frac{x^2}{2} \quad \text{with } L_W > 0. \tag{1.8}$$

Recall the following result adapted from [Tug14].

Theorem 2 (Theorem 2.1 of [Tug14])**.** *For U and W given by* (1.8)*, there exists $\sigma_c > 0$ such that*

- *For all $\sigma \geq \sigma_c$, there exists a unique stationary distribution $\mu_{\sigma,0}$ for* (NL)*. Furthermore, $\mu_{\sigma,0}$ is symmetric.*
- *For all $\sigma < \sigma_c$, there exist three stationary distributions for* (NL)*. One is symmetric, also denoted $\mu_{\sigma,0}$, and the other two, denoted $\mu_{\sigma,+}$ and $\mu_{\sigma,-}$, satisfy $\pm \int x d\mu_{\sigma,\pm}(dx) > 0$.*

By convention, in the case $\sigma \geq \sigma_c$, we may denote $\mu_\sigma = \mu_{\sigma,\pm} = \mu_{\sigma,0}$.

Our goal is now to study the stationary distribution(s) for the effective dynamics (Eff) in the specific case of the double-well potential (1.8). We wish to understand if, similarly as Theorem 2, there exists a phase transition, and if so compare the critical parameters. We thus prove in Section 3 the following theorem.

Theorem 3. *Let $\sigma_0 \in]0, \sigma_c[$ where σ_c is defined in Theorem 2. For U and W given by* (1.8)*, there exists $c_0 > 0$ such that for all (δ, p) satisfying $\frac{\delta}{p-1} \leq c_0$, denoting*

$$\sigma_c^{eff} = \sigma_c \left(1 - \frac{\delta L_W}{2(p-1)}\right), \tag{1.9}$$

we have the following phase transition for the dynamics (Eff)

- *For all* $\sigma \geq \sigma_c^{eff}$*, there exists a unique stationary distribution* $\mu_{\sigma,0}^{\delta,p}$ *for* (Eff)*. Furthermore,* $\mu_{\sigma,0}^{\delta,p}$ *is symmetric.*
- *For all* $\sigma \in [\sigma_0, \sigma_c^{eff}[$*, there exists exactly three stationary distributions for* (Eff)*. One is symmetric, also denoted* $\mu_{\sigma,0}^{\delta,p}$*, and the other two, denoted* $\mu_{\sigma,+}^{\delta,p}$ *and* $\mu_{\sigma,-}^{\delta,p}$*, satisfy* $\pm \int x d\mu_{\sigma,\pm}^{\delta,p}(x) > 0$*.*

Remark 1.1. *Let us quickly discuss the form of* (1.9)*. In the specific case of* (1.8)*, as discussed in Section 3, one has* $\Sigma(\bar{X}_t^{e,\delta,p}, \bar{\rho}_t^{e,\delta,p}) = L_W^2 Var(\bar{\rho}_t^{e,\delta,p})$*. To insist on the dependence on* σ *rather than* (δ, p)*, let us denote, only in this remark,* $\Sigma_\sigma := \Sigma(\bar{X}_t^{e,\delta,p}, \bar{\rho}_t^{e,\delta,p})$*.*

We will show, but this can be intuitively understood at this stage, that any stationary distribution for (NL) *is a stationary distribution for* (Eff) *although for a smaller value of* σ*. We thus have to study the stationary distribution at the critical value* σ_c*.*

As proved in Lemma 3.2, the variance of the stationary distribution for (NL) *at the critical value is* $Var(\mu_{\sigma_c,0}) = \frac{\sigma_c}{L_W}$*.*

By considering the diffusion term in (Eff) *and ensuring* $2\sigma_c = 2\sigma_c^{eff} + \frac{\delta}{p-1}\Sigma_{\sigma_c}$*, we then obtain* (1.9)*.*

Remark 1.2. *Another consequence of the fact that* $\Sigma(x, \rho)$ *does not depend on* x *is that a stationary solution of* (Eff) *is also the first marginal of a stationary solution of the corresponding effective dynamics for the kinetic Langevin diffusion (with the second marginal, i.e. the distribution of velocities at equilibrium, being the standard Gaussian distribution, and the distribution in the phase space being the product of these two marginals, see e.g. [GM21]). It means that Theorem 3 also applies to the kinetic case. Moreover, in this case, in practice, second-order splitting schemes are used, which means that the discretization error is negligible with respect to the stochastic gradient error, and thus the effective dynamics captures the leading term of the numerical errror, see also [GLBMM22] on this topic.*

Let us sum up the organization of the article.

- The Curie-Weiss model is studied in Section 2. We start by recalling the analysis of the phase transition for the classical Curie-Weiss model in Section 2.1 since the same ideas will be used afterwards. The study of the Curie-Weiss model with random batches and the proof of Theorem 1 are then done in Section 2.2,
- In Section 3 we study the Random Batch Method for interacting particle systems. More specifically we prove Theorem 3 in the specific case of the double-well potential,
- Finally, in Appendix A we gather some technical lemmas, and in Appendix B we prove some results on (NL) for the double-well potential used in Section 3.

Notation

For the Curie-Weiss model, with and without random batches:

- $\Omega_N = \{-1, ..., 1\}^N$: the set of possible configurations,
- $\sigma(n) = (\sigma_1(n), ..., \sigma_N(n))$: the spin configuration at time step n,
- β : the inverse temperature,
- β_c : the critical inverse temperature,
- H_N: the Hamiltonian of the Curie-Weiss model given in (1.2),

- $m_N(n) = \frac{1}{N}\sum_{i=1}^N \sigma_i(n)$: the magnetization at time step n,
- $r(\sigma, \sigma')$: transition probability for the Markov chain $(m_N(n))_n$, given in (1.4).
- $\sigma^p(n) = (\sigma_1^p(n), ..., \sigma_N^p(n))$: the spin configuration of the system with random batches of size p at time step n,
- $H_{N,p}$: the Hamiltonian for the system with random batches of size p, given in (1.5),
- $m_{N,p}(n) = \frac{1}{N}\sum_{i=1}^N \sigma_i^p(n)$: the magnetization at time step n for the system with random batches of size p,
- $r_p(m, m')$: transition probability for the Markov chain $(m_{N,p}(n))_n$, given in Lemma 2.1.
- $\beta_{c,p}$: the critical inverse temperature for the system with random batches of size p.

For the Random Batch Method for interacting particle system :

- U, W : two twice continuously differentiable functions, respectively the confining potential and the interacting potential (see (IPS)),
- $\sigma > 0$: a diffusion coefficient (see (IPS)),
- $(X_t^i)_{i\in\{1,...,N\}}$: the solution at time $t \in \mathbb{R}^+$ of the interacting particle system (IPS),
- $\bar{X}_t$, $\bar{\rho}_t$: the solution at time $t \in \mathbb{R}^+$ of the nonlinear limit (NL) and its law,
- $\delta > 0$: a timestep used in the various numerical schemes,
- $(X_t^{i,\delta})_{i\in\{1,...,N\}}$: the solution at time step $t \in \mathbb{N}$ of the Euler-Maruyama numerical scheme (D-IPS),
- $p \in \mathbb{N} \setminus \{0, 1\}$: the batch size,
- $\mathcal{P}_t$: the partition of $\{1, ..., N\}$ at time step t into subsets of size p,
- $\mathcal{C}_t^i$: the cluster containing index i at time step t (see (1.1)),
- $(Y_t^{i,\delta,p})_{i\in\{1,...,N\}}$: the solution at time step $t \in \mathbb{N}$ of the numerical scheme with random batches (D-RB-IPS),
- $\bar{Y}_t^{\delta,p}$: the solution at time step $t \in \mathbb{N}$ of (D-RB-NL), the nonlinear limit of (D-RB-IPS) as $N \to \infty$,
- $\bar{X}_t^{e,\delta,p}$, $\bar{\rho}_t^{e,\delta,p}$: the effective dynamics (Eff) at time $t \in \mathbb{R}^+$ and its law,
- $\mu_{\sigma,*}$ for $* \in \{0, \pm\}$, σ_c : stationary distributions and critical parameter of (NL) given in Theorem 2,
- $\mu_{\sigma,*}^{\delta,p}$ for $* \in \{0, \pm\}$, σ_c^{eff} : stationary distributions and critical parameter of (Eff) given in Theorem 3.

2 Understanding the problem on the Curie-Weiss model

In order to get a better grasp on the phenomenon we focus on, we begin by studying arguably one of the simplest model admitting a phase transition : the Curie-Weiss model. In Section 2.1, we show how we obtain the value of the critical parameter in the classical case. Then, in Section 2.2, we follow the same steps to compute the new critical inverse temperature in the case with random batches.

2.1 ...without the Random Batch Method

In order to study this critical inverse temperature, we choose to look at the limit of the dynamics with time step $\frac{1}{N}$ as N goes to infinity. $(m_N(n))_n$ is a discrete-time Markov chain with transition operator $U^{(N)}$ given by $U^{(N)} = \left(U^{(N)}_{i,j}\right)_{0\leq i,j\leq N}$ where $U^{(N)}_{i,j} = r\left(-1+\frac{2i}{N}, -1+\frac{2j}{N}\right)$. We denote $A_N = N\left(U^{(N)} - I\right)$. We have, for all continuously differentiable functions f,

$$\begin{aligned} A_N f(m) =& N\frac{1-m}{2}e^{-\beta N\left(\frac{m^2}{2}-\frac{\left(m+\frac{2}{N}\right)^2}{2}\right)_+}\left(f\left(m+\frac{2}{N}\right)-f(m)\right) \\ &+ N\frac{1+m}{2}e^{-\beta N\left(\frac{m^2}{2}-\frac{\left(m-\frac{2}{N}\right)^2}{2}\right)_+}\left(f\left(m-\frac{2}{N}\right)-f(m)\right). \end{aligned}$$

We thus get

$$\begin{aligned} A_N f(m) =& N\frac{1-m}{2}e^{-2\beta\left(-m-\frac{1}{N}\right)_+}\left(f\left(m+\frac{2}{N}\right)-f(m)\right) \\ &+ N\frac{1+m}{2}e^{-2\beta\left(m+\frac{1}{N}\right)_+}\left(f\left(m-\frac{2}{N}\right)-f(m)\right) \\ =& N\frac{1-m}{2}e^{-2\beta\left(-m-\frac{1}{N}\right)_+}\left(\frac{2}{N}f'(m)+O\left(\frac{1}{N^2}\right)\right) \\ &+ N\frac{1+m}{2}e^{-2\beta\left(m+\frac{1}{N}\right)_+}\left(-\frac{2}{N}f'(m)+O\left(\frac{1}{N^2}\right)\right) \\ =& (1-m)e^{-2\beta\left(-m-\frac{1}{N}\right)_+}f'(m) - (1+m)e^{-2\beta\left(m+\frac{1}{N}\right)_+}f'(m) + O\left(\frac{1}{N}\right) \\ &\xrightarrow[N\to\infty]{} f'(m)\left((1-m)e^{-2\beta(-m)_+} - (1+m)e^{-2\beta m_+}\right), \end{aligned}$$

which finally yields

$$A_N f(m) \xrightarrow[N\to\infty]{} 2f'(m)e^{-\beta|m|}\left(\sinh(m\beta) - m\cosh(m\beta)\right).$$

By [Kal97, Theorem 17.28], the process $M^{(N)}_t = m_N(\lfloor Nt\rfloor)$ weakly converges to the solution of

$$\frac{d}{dt}m(t) = 2e^{-\beta|m(t)|}\left(\sinh(\beta m(t)) - m(t)\cosh(\beta m(t))\right).$$

Denote $f(\beta, m) = 2e^{-\beta|m|}\left(\sinh(\beta m) - m\cosh(\beta m)\right)$. We have

$$f(\beta, m) = 0 \quad \Longleftrightarrow \quad \tanh(\beta m) = m.$$

For $\beta > 1$, the equation $f(\beta, m) = 0$ admits three solutions, and for $\beta \leq 1$ only one. Notice that for all $\beta > 0$, $f(\beta, 0) = 0 : 0$ is thus always an equilibrium state for the magnetization. Furthermore

$$\begin{aligned} \forall \beta > 0, \forall m \neq 0, \partial_m f(\beta, m) =& -2\beta \mathrm{sign}(m)e^{-\beta|m|}\left(\sinh(\beta m) - m\cosh(\beta m)\right) \\ &+2e^{-\beta|m|}\left((\beta-1)\cosh(\beta m) - \beta m\sinh(\beta m)\right), \end{aligned}$$

and, extending by continuity, we have $\partial_m f(\beta, 0) = 2(\beta - 1)$. Therefore, for $\beta > 1$, 0 is unstable as $\partial_m f(\beta, 0) > 0$, and for $\beta \leq 1$ it is stable.

Hence a critical inverse temperature $\beta_c = 1$, above which there are two stable equilibrium states, and under which there is only one.

2.2 ...with the Random Batch Method

To follow the same steps in the case with random batches, we need to compute the transition operator before finding its limit.

2.2.1 Transition probabilities

Let us start by giving explicit values for the transitions probabilities for the magnetization using the Random Batch Method. The proof, which relies on combinatorics arguments, is double-checked via numerical simulations in Figure 1.

Lemma 2.1. *In a system of size N, the transition probabilities for the magnetization with random batches of size p are given by*

$$r_p(m,m') = \begin{cases} \frac{1-m}{2}\binom{N-1}{p-1}^{-1} \sum_{k=0}^{p-1} \binom{\frac{1-m}{2}N-1}{k}\binom{\frac{1+m}{2}N}{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} \\ \qquad\qquad \text{if } m' = m + \frac{2}{N} \\ \frac{1+m}{2}\binom{N-1}{p-1}^{-1} \sum_{k=0}^{p-1} \binom{\frac{1-m}{2}N}{k}\binom{\frac{1+m}{2}N-1}{p-1-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+} \\ \qquad\qquad \text{if } m' = m - \frac{2}{N} \\ 1 - r_p\left(m, m+\frac{2}{N}\right) - r_p\left(m, m-\frac{2}{N}\right) \\ \qquad\qquad \text{if } m' = m \\ 0 \qquad\qquad \text{otherwise.} \end{cases} \tag{2.1}$$

Proof. Notice that, for a given m, the number of positive spins is given by $\frac{1+m}{2}N$ and the number of negative spins by $\frac{1-m}{2}N$.

Going right. Let us calculate the probability of going from m to $m + \frac{2}{N}$. To do so, the chosen spin, denoted i, must be of value -1, and this will happen with probability $\frac{1-m}{2}$. Then, depending on the cluster $\mathcal{C}$ to which spin i belongs, switching the spin from -1 to $+1$ happens with probability

$$\mathbb{P}(\sigma_i^p(n+1) = 1 | \sigma_i^p(n) = -1, \mathcal{C}) = \exp\left(-\beta\left(-\frac{1}{2p}\sum_{j,l\in\mathcal{C}} \sigma_j'\sigma_l' + \frac{1}{2p}\sum_{j,l\in\mathcal{C}} \sigma_j^p(n)\sigma_l^p(n)\right)_+\right),$$

where σ' denotes the configuration such that for all $j \neq i$, $\sigma_j' = \sigma_j^p(n)$, and $\sigma_i' = -\sigma_i^p(n)$. We have

$$\begin{aligned} -\frac{1}{2p}\sum_{j,l\in\mathcal{C}} \sigma_j'\sigma_l' + \frac{1}{2p}\sum_{j,l\in\mathcal{C}} \sigma_j^p(n)\sigma_l^p(n) & \\ = -\frac{1}{2p}\Bigg(\sum_{j,l\in\mathcal{C},j\neq i,l\neq i} \sigma_j'\sigma_l' - \sum_{j,l\in\mathcal{C},j\neq i,l\neq i} \sigma_j^p(n)\sigma_l^p(n) + 2\sum_{j\in\mathcal{C},j\neq i} \sigma_j'\sigma_i' & \\ -2\sum_{j\in\mathcal{C},j\neq i} \sigma_j^p(n)\sigma_i^p(n) + (\sigma_i')^2 - (\sigma_i^p(n))^2\Bigg) & \\ = -\frac{1}{2p}\left(-2\sigma_i^p(n)\sum_{j\in\mathcal{C},j\neq i} \sigma_j^p(n) - 2\sigma_i^p(n)\sum_{j\in\mathcal{C},j\neq i} \sigma_j^p(n)\right) & \\ = \frac{2}{p}\sigma_i^p(n)\sum_{j\in\mathcal{C},j\neq i} \sigma_j^p(n). & \end{aligned}$$

We classify the possible clusters containing i based on the number of negative spins. The number of clusters containing i and k other negative spins is $\binom{\frac{1-m}{2}N-1}{k}\binom{\frac{1+m}{2}N}{p-1-k}$ (choosing k spins among the $\frac{1-m}{2}N-1$ negative spins that are not i, then the $p-1-k$ spins that remain to construct cluster $\mathcal{C}$ among the positive spins). For k negative spins in cluster $\mathcal{C}$ (without counting i), we have

$$\sum_{j\in\mathcal{C},j\neq i}\sigma_j^p(n)=\sum_{j\in\mathcal{C},j\neq i,\sigma_j^p(n)=1}1-\sum_{j\in\mathcal{C},j\neq i,\sigma_j^p(n)=-1}1=p-1-k-k,$$

and thus, since $\sigma_i^p(n)=-1$

$$-\frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma_j'\sigma_l'+\frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma_j^p(n)\sigma_l^p(n)=2\frac{2k+1-p}{p}$$

The total number of possible choices for $\mathcal{C}$ is $\binom{N-1}{p-1}$ (choosing the $(p-1)$ spins that are not i). Hence

$$r_p\left(m,m+\frac{2}{N}\right)=\frac{1-m}{2}\frac{1}{\binom{N-1}{p-1}}\sum_{k=0}^{p-1}\binom{\left(\frac{1-m}{2}\right)N-1}{k}\binom{\left(\frac{1+m}{2}\right)N}{p-1-k}e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}$$

Going left. Similar calculations yield the probability of going left : the probability of choosing a spin of value $+1$ is $\frac{1+m}{2}$, then we classify the possible clusters containing this spin based on the number of negative spins. □

Remark 2.1. *The values given in* (2.1) *are consistent in the case $p=N$. Observe for instance that the only nonzero term in the sum defining $r_N\left(m,m+\frac{2}{N}\right)$ is obtained for $k=\frac{1-m}{2}N-1$. Thus*

$$r_N\left(m,m+\frac{2}{N}\right)=\frac{1-m}{2}e^{-2\beta\left(-m-\frac{1}{N}\right)_+}=r\left(m,m+\frac{2}{N}\right),$$

where the value of r is given in (1.4)*.*

Remark 2.2. *We observe how the transition probabilities evolve with the parameter p in Figure 2. Furthermore, the values given in* (2.1) *allow us to define, on the state space $\{-1,1+\frac{2}{N},...,1-\frac{2}{N},1\}$, a transition matrix for the magnetization. The latter is an irreducible and aperiodic Markov chain on a finite state space, and thus admits a unique invariant measure. We can numerically obtain it by iterating the transition matrix (see Figure 3)*

2.2.2 Study of the critical parameter

We now wish to show how adding random batches artificially increases the temperature of the system, thus decreasing the critical temperature (or, equivalently, increasing the critical inverse temperature).

Limit ODE. Let us, like previously, find the limit as N goes to infinity of the dynamics of $(m_{N,p}(n))_n$ with time step $\frac{1}{N}$. This discrete-time Markov chain admits a transition operator $U^{(N,p)}$ given by

$$U^{(N,p)}=\left(U_{i,j}^{(N,p)}\right)_{0\leq i,j\leq N}\quad\text{where}\quad U_{i,j}^{(N,p)}=r_p\left(-1+\frac{2i}{N},-1+\frac{2j}{N}\right).$$

We denote $A_N^{(p)}=N(U^{(N,p)}-I)$ and have, for all continuously differentiable functions f,

$$A_N^{(p)}f(m)=Nr_p\left(m,m+\frac{2}{N}\right)\left(f(m+\frac{2}{N})-f(m)\right)+Nr_p\left(m,m-\frac{2}{N}\right)\left(f(m-\frac{2}{N})-f(m)\right)$$

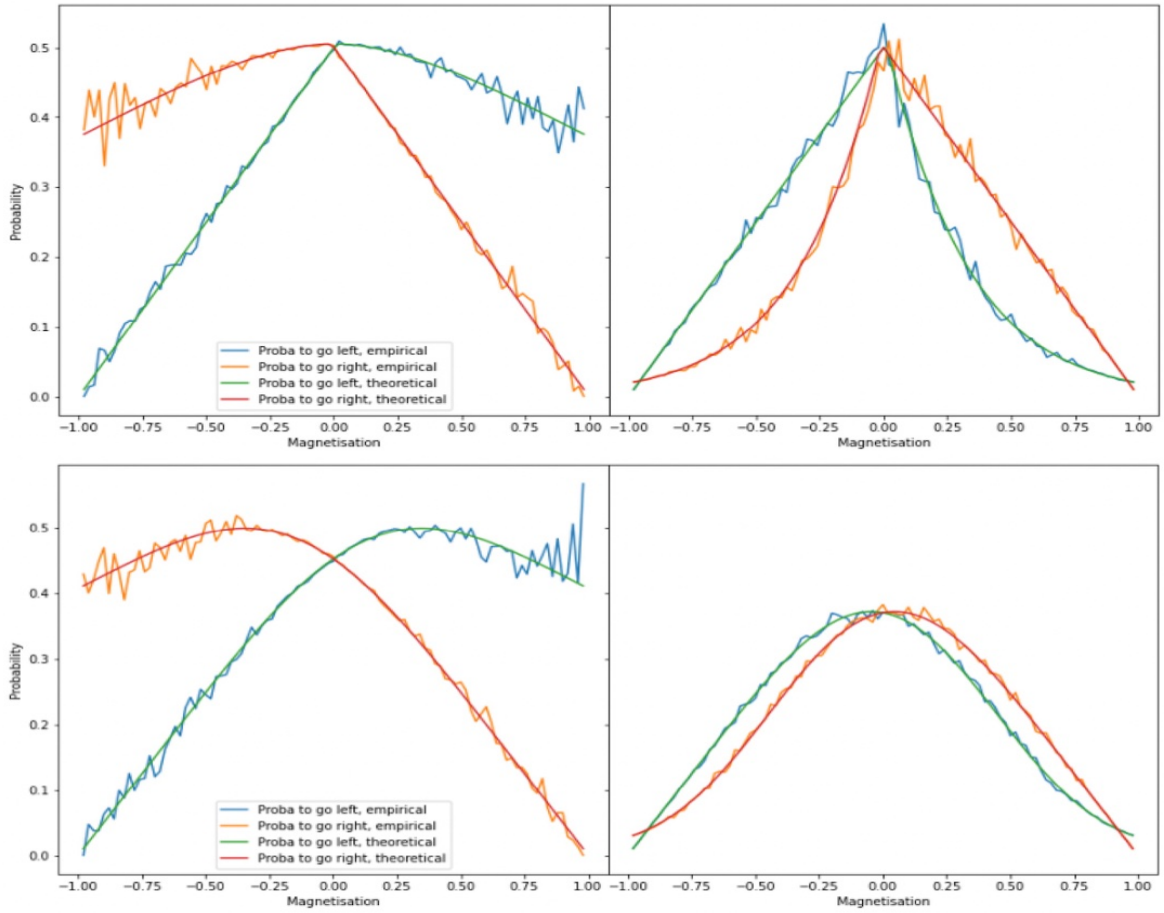

Figure 1: Comparison of theoretical and empirical transition probabilities, for $N = 100$. The theoretical values are those given in Lemma 2.1. To numerically compute the empirical transition probabilities, for each initial magnetization in $\{-1, -1 + \frac{2}{N}, ..., 1 - \frac{2}{N}, 1\}$, 10 processes are simulated during 1000 timesteps, and we consider the proportion of times the processes go left or right. **Top :** without random batches. **Bottom :** with random batches of size $p = 10$. **Left :** for $\beta = 0.5$. **Right :** for $\beta = 2$.

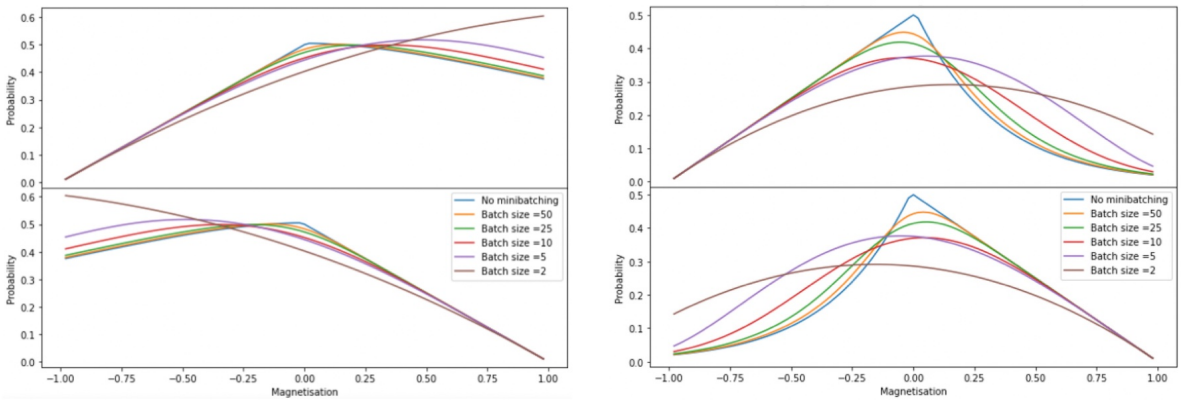

Figure 2: Comparison of transition probabilities depending on batch size, for $N = 100$. The values given are from Lemma 2.1. **Top :** probability of going left. **Bottom :** probability of going right. **Left :** for $\beta = 0.5$. **Right :** for $\beta = 2$.

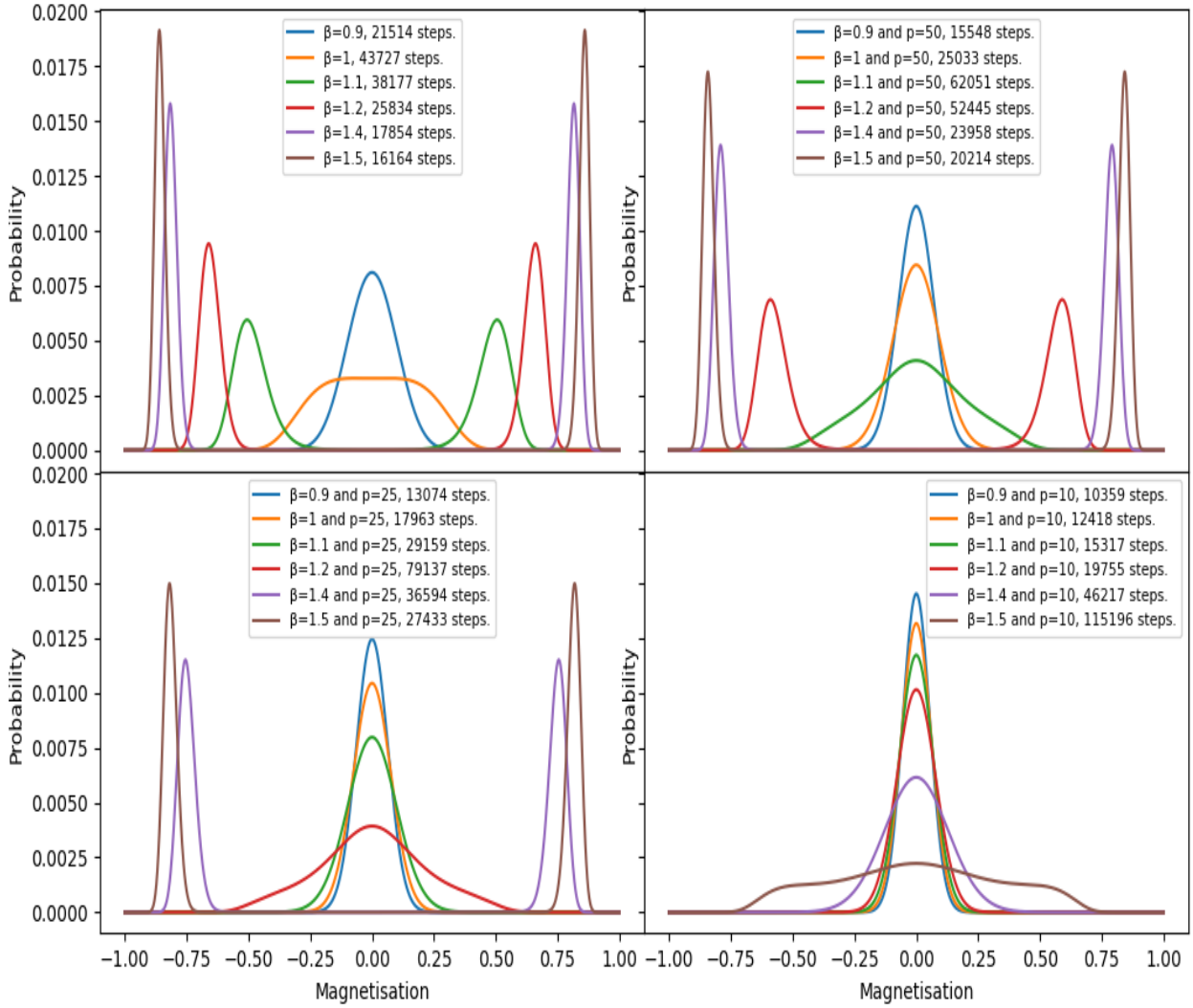

Figure 3: Numerical observation of the invariant distribution for the Curie-Weiss model with N spins. Starting from the uniform distribution for the magnetization, we iterate the transition matrix (given by Lemma 2.1) until the L^1 distance between two consecutive iterations is less than a threshold $N\epsilon$, with $N = 1000$, $\epsilon = 10^{-9}$ and various values for β. We indicate the number of iterations (or steps) needed before convergence. **Top left :** with no random batches. **Top right :** with $p = 50$. **Bottom left :** with $p = 25$. **Bottom right :** with $p = 10$.

$$=r_p\left(m,m+\frac{2}{N}\right)\left(2f'(m)+O\left(\frac{1}{N}\right)\right)-r_p\left(m,m-\frac{2}{N}\right)\left(2f'(m)+O\left(\frac{1}{N}\right)\right).$$

We have, by standard computations

$$\begin{aligned}r_p\left(m,m+\frac{2}{N}\right)&=\frac{1-m}{2}\binom{N-1}{p-1}^{-1}\sum_{k=0}^{p-1}\binom{\left(\frac{1-m}{2}\right)N-1}{k}\binom{\left(\frac{1+m}{2}\right)N}{p-1-k}e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}\\&\xrightarrow[N\to\infty]{}\frac{1-m}{2}\sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^k\left(\frac{1+m}{2}\right)^{p-1-k}e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+},\end{aligned}$$

and likewise

$$r_p\left(m,m-\frac{2}{N}\right)\xrightarrow[N\to\infty]{}\frac{1+m}{2}\sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^k\left(\frac{1+m}{2}\right)^{p-1-k}e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+}.$$

Hence

$$A_N^{(p)}f(m)\xrightarrow[N\to\infty]{}A^{(p)}f(m),$$

where

$$\begin{aligned}A^{(p)}f(m)&=f'(m)\left(S_1^{p,\beta}(m)-S_2^{p,\beta}(m)\right)-mf'(m)\left(S_1^{p,\beta}(m)+S_2^{p,\beta}(m)\right),\\S_1^{p,\beta}(m)&=\sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^k\left(\frac{1+m}{2}\right)^{p-1-k}e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}\\S_2^{p,\beta}(m)&=\sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^k\left(\frac{1+m}{2}\right)^{p-1-k}e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+}.\end{aligned}$$

Remark 2.3. *Notice that*

$$S_1^{p,\beta}(m)=\mathbb{E}\left(e^{-2\beta\left(\frac{2X_{m,p}+1-p}{p}\right)_+}\right),\quad S_2^{p,\beta}(m)=\mathbb{E}\left(e^{-2\beta\left(\frac{p-1-2X_{m,p}}{p}\right)_+}\right),$$

where $X_{m,p}$ is a random variable following a binomial distribution of parameters $p-1$ and $\frac{1-m}{2}$. Intuitively, for an infinite number of spins, the dynamics of the system relies on the construction of a cluster of size p (containing the chosen spin that may change), which is done by independently taking the remaining $p-1$ spins from an infinite pool containing a proportion $\frac{1-m}{2}$ of negative spins.

Denoting $f_p(\beta,m)=\left(S_1^{p,\beta}(m)-S_2^{p,\beta}(m)\right)-m\left(S_1^{p,\beta}(m)+S_2^{p,\beta}(m)\right)$, by [Kal97, Theorem 17.28], the process $M_t^{(N,p)}=m_{N,p}(\lfloor Nt\rfloor)$ weakly converges to the solution of

$$\frac{d}{dt}m(t)=f_p(\beta,m(t)).$$

The cases $p=2$ and $p=3$. We may directly compute

$$\begin{aligned}S_1^{2,\beta}(m)=&\frac{1+m}{2}+\frac{1-m}{2}e^{-\beta},\qquad S_2^{2,\beta}(m)=\frac{1+m}{2}e^{-\beta}+\frac{1-m}{2},\\S_1^{3,\beta}(m)=&\left(\frac{1+m}{2}\right)^2+2\left(\frac{1+m}{2}\right)\left(\frac{1-m}{2}\right)+\left(\frac{1-m}{2}\right)^2e^{-\frac{4\beta}{3}},\end{aligned}$$

$$S_2^{3,\beta}(m) = \left(\frac{1+m}{2}\right)^2 e^{-\frac{4\beta}{3}} + 2\left(\frac{1+m}{2}\right)\left(\frac{1-m}{2}\right) + \left(\frac{1-m}{2}\right)^2,$$

which yield

$$\begin{aligned}
f_2(\beta, m) =& m(1-e^{-\beta}) - m(1+e^{-\beta}) = -2me^{-\beta},\\
f_3(\beta, m) =& \left(\left(\frac{1+m}{2}\right)^2 - \left(\frac{1-m}{2}\right)^2\right)\left(1-e^{-\frac{4\beta}{3}}\right)\\
&- m\left(\left(\left(\frac{1+m}{2}\right)^2 + \left(\frac{1-m}{2}\right)^2\right)\left(1+e^{-\frac{4\beta}{3}}\right) + (1+m)(1-m)\right)\\
=& m\left(1-e^{-\frac{4\beta}{3}}\right) - m\left(\frac{(1+m^2)}{2}\left(1+e^{-\frac{4\beta}{3}}\right) + 1 - m^2\right)\\
=& -\frac{m}{2}\left(1+3e^{-\frac{4\beta}{3}}\right) + \frac{m^3}{2}\left(1-e^{-\frac{4\beta}{3}}\right).
\end{aligned}$$

For $p=2$ we thus have, $f_2(\beta, m) = 0 \iff m = 0$, and furthermore notice that $\partial_m f_2(\beta, 0) < 0$, which means that 0 is the unique equilibrium state, and it is stable. For $p = 3$,

$$f_3(\beta, m) = 0 \quad \iff \quad m = 0 \quad \text{or} \quad m = \pm\sqrt{\frac{1+3e^{-\frac{4\beta}{3}}}{1-e^{-\frac{4\beta}{3}}}}.$$

However, for all $\beta > 0$ we have $\sqrt{\frac{1+3e^{-\frac{4\beta}{3}}}{1-e^{-\frac{4\beta}{3}}}} > 1$, as well as $\partial_m f_3(\beta, 0) = -\frac{1+3e^{-\frac{4\beta}{3}}}{2} < 0$. The point 0 is thus the unique equilibrium state, and it is stable.

We may observe this phenomenon in Figure 4, in which we compare the cases $p = 2$ and $p = 3$ with $p = 4$.

Existence of a phase transition for $p \geq 4$. First notice that

$$\begin{aligned}
f_p(\beta, 0) =& \sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1}{2}\right)^{p-1} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} - \sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1}{2}\right)^{p-1} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+}\\
=& 0 \quad \text{by change of variables} \quad k' = p-1-k.
\end{aligned}$$

Thus $m = 0$ is for all $\beta > 0$ an equilibrium state. The remaining questions, in order to prove Theorem 1, are

- is there $\beta_{c,p} > 0$ such that for all $\beta < \beta_{c,p}$ we have $\partial_\kappa f(\beta, 0) < 0$ (in which case $m = 0$ is stable) and such that for all $\beta > \beta_{c,p}$ we have $\partial_m f(\beta, 0) > 0$ (in which case $m = 0$ is unstable) ?
- do we have $\beta_{c,p} > 1$ (in which case the critical temperature has indeed decreased when compared to the case without random batches) ?
- can we give an estimate of $\beta_{c,p}$?

To answer the first question, we may calculate

$$\begin{aligned}
S_1'^{p,\beta}(m) =& -\sum_{k=0}^{p-1}\frac{k}{2}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^{k-1}\left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}\\
&+ \sum_{k=0}^{p-1}\frac{p-1-k}{2}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^{k}\left(\frac{1+m}{2}\right)^{p-2-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}
\end{aligned}$$

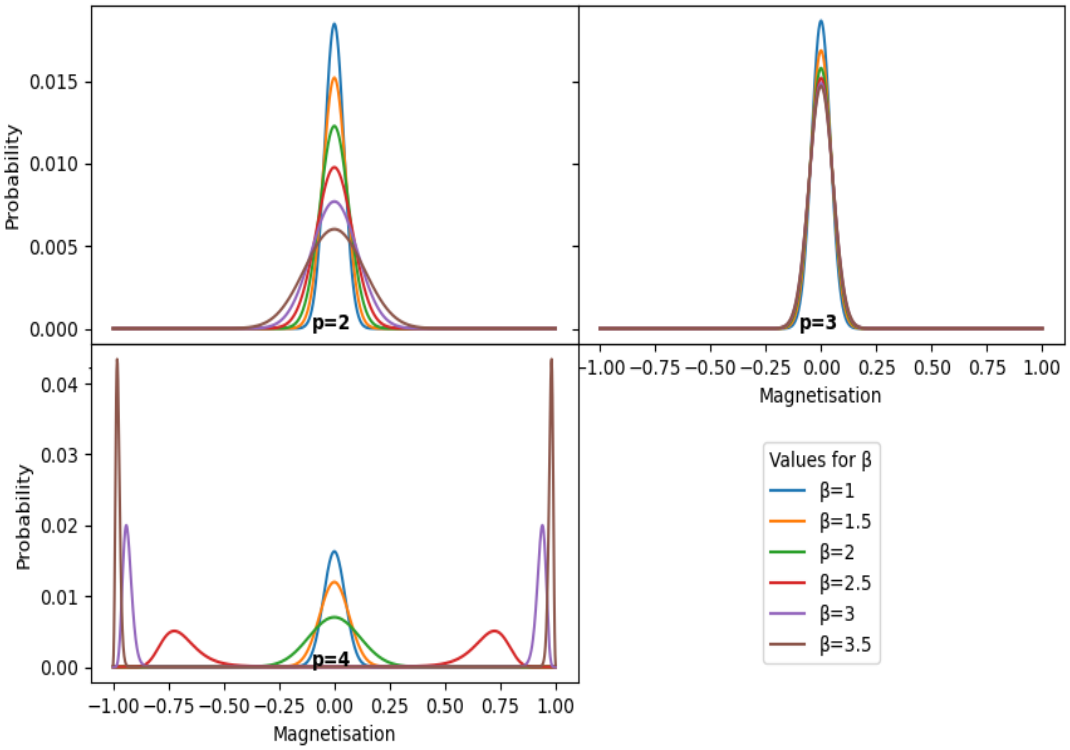

Figure 4: Numerical observation of the invariant distribution for the Curie-Weiss model with N spins in the cases $p = 2$ (Top right), $p = 3$ (Top left) and $p = 4$ (Bottom left). Starting from the uniform distribution for the magnetization, we iterate the transition matrix (given in Lemma 2.1) until the L^1 distance between two consecutive iterations is less than a threshold $N\epsilon$, with $N = 1000$, $\epsilon = 10^{-9}$ and various values for β.

and

$$S_2'^{p,\beta}(m) = -\sum_{k=0}^{p-1} \frac{k}{2}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^{k-1}\left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+} + \sum_{k=0}^{p-1} \frac{p-1-k}{2}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^{k}\left(\frac{1+m}{2}\right)^{p-2-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+},$$

which yields

$$\partial_m f_p(\beta, m) = \left(S_1'^{p,\beta}(m) - S_2'^{p,\beta}(m)\right) - \left(S_1^{p,\beta}(m) + S_2^{p,\beta}(m)\right) - m\left(S_1'^{p,\beta}(m) + S_2'^{p,\beta}(m)\right).$$

We thus have

$$\partial_m f_p(\beta, 0) = 2S_1'^{p,\beta}(0) - 2S_1^{p,\beta}(0) = 2\left(\frac{1}{2}\right)^{p-1}\sum_{k=0}^{p-1}(p-2-2k)\binom{p-1}{k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}.$$

First, notice

$$\partial_\beta(\partial_m f_p(\beta, 0)) = -4\left(\frac{1}{2}\right)^{p-1}\sum_{k=0}^{p-1}(p-2-2k)\left(\frac{2k+1-p}{p}\right)_+\binom{p-1}{k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} > 0.$$

The function $\beta \mapsto \partial_m f_p(\beta, 0)$ is therefore an increasing function, which furthermore satisfies $\partial_m f_p(0, 0) < 0$ and $\lim_{\beta\to\infty} \partial_m f_p(\beta, 0) > 0$, hence a unique critical parameter $\beta_{c,p} > 0$.

Remark 2.4. *We use the assumption $p \geq 4$ in order to prove $\lim_{\beta\to\infty} \partial_m f_p(\beta, 0) > 0$. Indeed*

$$\lim_{\beta\to\infty} \partial_m f_p(\beta, 0) = 2\left(\frac{1}{2}\right)^{p-1}\sum_{k=0}^{p-1}(p-2-2k)\binom{p-1}{k}\mathbb{1}_{k\leq\frac{p-1}{2}}.$$

If p is even, all the terms in the sum are nonnegative, and if $p \geq 4$, at least one term is positive. If p is odd, one term is negative, and if $p \geq 5$ it can easily be shown that it is compensated by the positive terms.

Estimation of the critical parameter. Denoting X_p a random variable following a binomial distribution of parameters $p-1$ and $\frac{1}{2}$, we have

$$\partial_m f_p(\beta, 0) = 2\mathbb{E}\left((p-2-2X_p)e^{-2\beta\left(\frac{2X_p+1-p}{p}\right)_+}\right) := g_p(\beta).$$

We are thus looking for the unique $\beta_{c,p} > 0$ such that $g_p(\beta_{c,p}) = 0$.

Let $Y_p = 2\frac{X_p}{p} - \frac{p-1}{p}$. We have

$$g_p(\beta) = \mathbb{E}\left(2(-pY_p - 1)e^{-2\beta(Y_p)_+}\right). \tag{2.2}$$

Since X_p and $p-1-X_p$ have the same law, Y_p has the same law as $2\frac{p-1-X_p}{p} - \frac{p-1}{p} = \frac{p-1}{p} - 2\frac{X_p}{p} = -Y_p$. Thus

$$\begin{aligned} g_p(\beta) &= -\mathbb{E}\left(pY_p e^{-2\beta(Y_p)_+}\right) - \mathbb{E}\left(p(-Y_p)e^{-2\beta(-Y_p)_+}\right) - \mathbb{E}\left(e^{-2\beta(Y_p)_+}\right) - \mathbb{E}\left(e^{-2\beta(-Y_p)_+}\right) \\ &= -\mathbb{E}\left(pY_p\left(e^{-2\beta(Y_p)_+} - e^{-2\beta(-Y_p)_+}\right)\right) - \mathbb{E}\left(e^{-2\beta(Y_p)_+} + e^{-2\beta(-Y_p)_+}\right) \end{aligned}$$

$$
\begin{aligned}
=&\mathbb{E}\left(2pY_p e^{-\beta|Y_p|}\sinh(\beta Y_p)\right) - \mathbb{E}\left(2e^{-\beta|Y_p|}\cosh(\beta Y_p)\right)\\
=&2\mathbb{E}\left(\cosh(\beta Y_p)e^{-\beta|Y_p|}\left(pY_p\tanh(\beta Y_p)-1\right)\right).
\end{aligned}
$$

As this is an increasing function in β, in order to prove that $\beta_{c,p} > 1$, it is sufficient to prove that $g_p(1) < 0$. The Law of Large Number and the Central Limit Theorem yield

$$
Y_p \xrightarrow[p\to\infty]{a.s} 0 \quad \text{and} \quad \frac{p}{\sqrt{p-1}}Y_p \xrightarrow[p\to\infty]{law} \mathcal{N}(0,1).
$$

We have

$$
\begin{aligned}
g_p(\beta) =&2\mathbb{E}\left(\left(1+\frac{\beta^2Y_p^2}{2}+o(Y_p^2)\right)\left(1-\beta|Y_p|+\frac{\beta^2Y_p^2}{2}+o(Y_p^2)\right)\right.\\
&\left.\times\left(pY_p\left(\beta Y_p-\frac{\beta^3Y_p^3}{3}+o(Y_p^3)\right)-1\right)\right)\\
=&2\mathbb{E}\left(\left(\frac{p-1}{p}\beta\frac{p^2}{p-1}Y_p^2-1\right)-\beta|Y_p|\left(\frac{p-1}{p}\beta\frac{p^2}{p-1}Y_p^2-1\right)+O(Y_p^2)+O(pY_p^4)\right)\\
=&2\left(\beta\frac{p-1}{p}\mathbb{E}\left(\frac{p^2}{p-1}Y_p^2\right)-1-\frac{(p-1)^{3/2}}{p^2}\left(\beta^2\mathbb{E}\left(\left|\frac{p}{\sqrt{p-1}}Y_p\right|^3\right)-\beta\mathbb{E}\left(\left|\frac{p}{\sqrt{p-1}}Y_p\right|\right)\right)\right.\\
&\left.+\frac{1}{p}\mathbb{E}\left(O(pY_p^2)+O(p^2Y_p^4)\right)\right)\\
=&2\left(\beta\frac{p-1}{p}\mathbb{E}\left(Z^2\right)-1-\frac{(p-1)^{3/2}}{p^2}\left(\beta^2\left(\mathbb{E}\left|Z\right|^3+o\left(1\right)\right)-\beta\left(\mathbb{E}\left(|Z|\right)+o(1)\right)\right)+O\left(\frac{1}{p}\right)\right)\\
=&2(\beta-1)-\frac{2}{\sqrt{p}}\sqrt{\frac{2}{\pi}}\left(2\beta^2-\beta\right)+o\left(\frac{1}{\sqrt{p}}\right),
\end{aligned}
$$

where for this last equality, we use Lemma A.1 and the fact that, for $Z\sim\mathcal{N}(0,1)$, $\mathbb{E}|Z| = \sqrt{\frac{2}{\pi}}$ and $\mathbb{E}(|Z|^3) = 2\sqrt{\frac{2}{\pi}}$.

In the end, we obtain, again, the fact that $g_p(1) \xrightarrow[p\to\infty]{} 0$ (hence the correct critical parameter at the limit) and the fact that, at least for p sufficiently large, $g_p(1) < 0$. For smaller values of p, we rely on numerical simulations to verify $g_p(1) < 0$ (See Figure 5). Let us find an approximation of $\beta_{c,p}$ by using the fact that $g_p(\beta_{c,p}) = 0$. We have

$$
2\sqrt{\frac{2}{p\pi}}\beta_{c,p}^2-\left(1+\sqrt{\frac{2}{p\pi}}\right)\beta_{c,p}+1+o\left(\frac{1}{\sqrt{p}}\right)=0,
$$

i.e

$$
\begin{aligned}
\beta_{c,p,\pm} =&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(\left(1+\sqrt{\frac{2}{p\pi}}\right)\pm\left(\left(1+\sqrt{\frac{2}{p\pi}}\right)^2-8\sqrt{\frac{2}{p\pi}}\left(1+o\left(\frac{1}{\sqrt{p}}\right)\right)\right)^{1/2}\right)\\
=&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(1+\sqrt{\frac{2}{p\pi}}\pm\left(1+\frac{2}{p\pi}-6\sqrt{\frac{2}{p\pi}}+o\left(\frac{1}{p}\right)\right)^{1/2}\right)\\
=&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(1+\sqrt{\frac{2}{p\pi}}\pm\left(1+\frac{1}{p\pi}-3\sqrt{\frac{2}{p\pi}}-\frac{9}{p\pi}+o\left(\frac{1}{p}\right)\right)\right),
\end{aligned}
$$

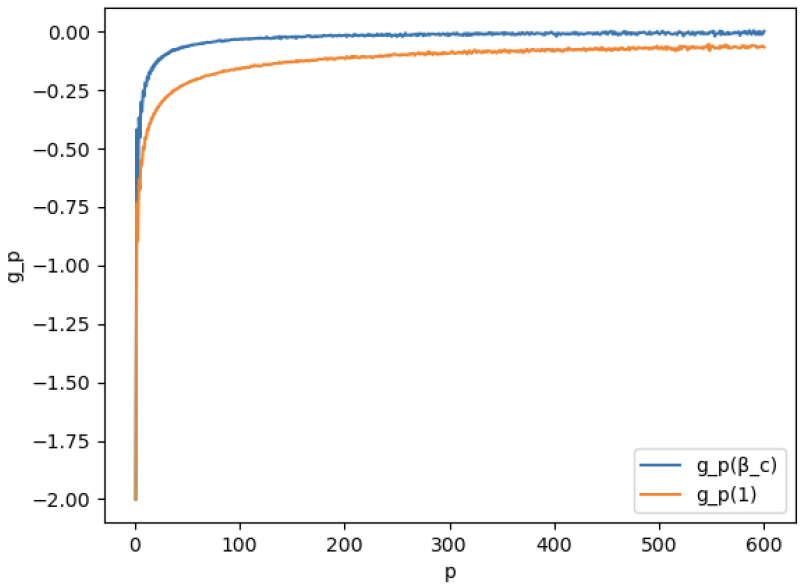

Figure 5: Numerical values for $g_p(1)$ and $g_p\left(\tilde{\beta}_c\right)$, with $\tilde{\beta}_c = 1 + \sqrt{\frac{2}{p\pi}}$. The value of g_p given in (2.2) is computed via Monte-Carlo approximation using $M = 10^8$ samples for Y_p.

thus

$$\begin{aligned}\beta_{c,p} =& \frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(1+\sqrt{\frac{2}{p\pi}}-\left(1-3\sqrt{\frac{2}{p\pi}}-\frac{8}{p\pi}+o\left(\frac{1}{p}\right)\right)\right)\\ =& \frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(4\sqrt{\frac{2}{p\pi}}+\frac{8}{p\pi}+o\left(\frac{1}{p}\right)\right)\\ =& 1+\sqrt{\frac{2}{p\pi}}+o\left(\frac{1}{\sqrt{p}}\right).\end{aligned}$$

We have thus proved Theorem 1.

3 Random Batch Method for interacting particle systems and stationary distribution(s)

We now turn our attention to the study of (D-RB-IPS) for a given batch size $p \in \mathbb{N} \setminus \{0, 1\}$. In the specific case of U and W given in (1.8), we study the phase transition for (Eff) and prove Theorem 3. The stationary distributions of (NL), provided there exists one, are defined by the solutions of

$$\mu_\sigma(dx) = \frac{\exp\left(-\frac{1}{\sigma}\left(U(x) + W * \mu_\sigma(x)\right)\right)}{\int \exp\left(-\frac{1}{\sigma}\left(U(y) + W * \mu_\sigma(y)\right)\right) dy} dx. \tag{3.1}$$

We consider the case of linear interactions in a double well potential in dimension one, i.e U and W given in (1.8), which in particular implies

$$\Sigma(x,\rho) = (\nabla W)^2 * \rho(x) - (\nabla W * \rho(x))^2 = L_W^2\left(\int y^2\rho(dy) - \left(\int y\rho(dy)\right)^2\right) = L_W^2 \mathrm{Var}(\rho).$$

Denote, for a measure μ,

$$\kappa_1(\mu) = \int_{\mathbb{R}} x\mu(dx) \quad \text{and} \quad \kappa_2(\mu) = \int_{\mathbb{R}} (x - \kappa_1(\mu))^2 \mu(dx).$$

The stationary distributions of (Eff), provided there exist one, are thus similarly defined by the solutions of

$$\mu_\sigma^{\delta,p}(dx) = \frac{\exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2\left(\mu_\sigma^{\delta,p}\right)}\left(U(x) + \frac{L_W}{2}\left|x - \kappa_1\left(\mu_\sigma^{\delta,p}\right)\right|^2\right)\right)}{\int \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2\left(\mu_\sigma^{\delta,p}\right)}\left(U(y) + \frac{L_W}{2}\left|y - \kappa_1\left(\mu_\sigma^{\delta,p}\right)\right|^2\right)\right) dy} dx. \tag{3.2}$$

The pair $(\kappa_1(\mu_\sigma^{\delta,p}), \kappa_2(\mu_\sigma^{\delta,p}))$ is therefore a solution of

$$\kappa_1 = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2}\left(U(x) + \frac{L_W}{2}|x - \kappa_1|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2}\left(U(x) + \frac{L_W}{2}|x - \kappa_1|^2\right)\right) dx} \tag{3.3}$$

$$\kappa_2 = \frac{\int_{\mathbb{R}} (x - \kappa_1)^2 \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2}\left(U(x) + \frac{L_W}{2}|x - \kappa_1|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2}\left(U(x) + \frac{L_W}{2}|x - \kappa_1|^2\right)\right) dx}. \tag{3.4}$$

Thanks to (3.2), solving for (κ_1, κ_2) the system of equations (3.3)-(3.4) is equivalent to finding a stationary distribution of (Eff). Define

$$g(x, \sigma, \kappa) = \exp\left(-\frac{1}{\sigma}\left(U(x) + \frac{L_W}{2}|x - \kappa|^2\right)\right), \tag{3.5}$$

$$f_1(\sigma, \kappa) = \frac{\int_{\mathbb{R}} x g(x, \sigma, \kappa) dx}{\int_{\mathbb{R}} g(x, \sigma, \kappa) dx}, \tag{3.6}$$

$$f_2(\sigma, \kappa) = \frac{\int_{\mathbb{R}} (x - \kappa)^2 g(x, \sigma, \kappa) dx}{\int_{\mathbb{R}} g(x, \sigma, \kappa) dx}, \tag{3.7}$$

such that, for the symbol $* \in \{0, \pm\}$ and $\mu_{\sigma,*}$ defined in Theorem 2, $\kappa_1(\mu_{\sigma,*})$ is a solution of $\kappa_1(\mu_{\sigma,*}) = f_1(\sigma, \kappa_1(\mu_{\sigma,*}))$ and $\kappa_2(\mu_{\sigma,*}) = f_2(\sigma, \kappa_1(\mu_{\sigma,*}))$ is the corresponding variance.

Remark 3.1 (Wellposedness)**.** *Since we are looking for stationary distributions of* (Eff)*, we are in reality considering specific solutions with constant mean and variance. We could thus choose not to worry about wellposedness of* (Eff) *since we could technically forget the nonlinearity.*

However, for the sake of completeness, let us quickly turn our attention to the existence and uniqueness of solutions of (Eff) *in the specific case of U and W given in Theorem 3, for which the process is*

$$\begin{cases} d\bar{X}_t^{e,\delta,p} = u(\bar{X}_t^{e,\delta,p}) + f * \rho_t^{e,\delta,p}(\bar{X}_t^{e,\delta,p})dt + \bar{\sigma}(\rho_t^{e,\delta,p})dB_t, \\ \rho_t^{e,\delta,p} = \mathit{Law}(\bar{X}_t^{e,\delta,p}) \end{cases}$$

in dimension 1, with

$$u(x) = -x^3 + x, \quad f(x) = -L_W x, \quad \textit{and} \quad \bar{\sigma}(\mu) = \left(2\sigma + \frac{\delta L_W^2}{p-1}\mathit{Var}(\mu)\right)^{1/2}.$$

To prove existence and uniqueness, we use [CdS23, Theorem 2.5]. To do so, we notice that there is a constant $C > 0$ (which may change from line to line) such that

$$(x-y)(u(x)-u(y)) \leq C|x-y|^2, \quad |u(x)-u(y)| \leq C(1+x^2+y^2)|x-y|.$$

Another assumption we have to verify, and which is the main technical difficulty, is the following

$$|\bar{\sigma}(\mu) - \bar{\sigma}(\nu)| \leq C\mathcal{W}_2(\mu,\nu).$$

Since the square root function is Lipschitz continuous on $[2\sigma, \infty[$, it is sufficient to prove that there is $C > 0$ such that $|Var(\mu) - Var(\nu)| \leq C\mathcal{W}_2(\mu,\nu)$. Denoting X (resp. Y) a random variable distributed according to μ (resp. ν), we have

$$\begin{aligned}|Var(\mu) - Var(\nu)| \leq& \left|\mathbb{E}(X^2) - \mathbb{E}(Y^2)\right| + \left|\mathbb{E}(X)^2 - \mathbb{E}(Y)^2\right| \\ \leq& |\mathbb{E}((X-Y)(X+Y))| + |\mathbb{E}(X-Y)\mathbb{E}(X+Y)| \\ \leq& 2\sqrt{\mathbb{E}((X-Y)^2)}\sqrt{\mathbb{E}((X+Y)^2)},\end{aligned}$$

where we use for this last line Cauchy-Schwarz inequality and Jensen's inequality. Considering (X,Y) to be distributed according to the optimal coupling for the Wasserstein-2 distance between μ and ν, we thus have

$$|Var(\mu) - Var(\nu)| \leq C\sqrt{\mathbb{E}(X^2) + \mathbb{E}(Y^2)}\mathcal{W}_2(\mu,\nu).$$

We thus should obtain existence and uniqueness for processes with bounded second moments.

Let us now use this idea to prove wellposedness. We consider the following process, for a given $K > 0$

$$\begin{cases} d\bar{X}_t^K = u(\bar{X}_t^K) + f * \rho_t^K(\bar{X}_t^K)dt + \bar{\sigma}_K(\rho_t^K)dB_t, \\ \rho_t^K = Law(\bar{X}_t^K) \end{cases} \tag{3.8}$$

where $\bar{\sigma}_K(\mu) = \left(2\sigma + \frac{\delta L_W^2}{p-1}\left(\int x^2 d\mu(x) \wedge K - \left(\int x d\mu(x)\right)^2 \wedge K\right)\right)^{1/2}$. Let $a, b \in \mathbb{R}$ and $c \geq 0$, we have $\left|a^2 \wedge c - b^2 \wedge c\right| \leq |a-b|\left|\sqrt{a^2 \wedge c} + \sqrt{b^2 \wedge c}\right|$. We obtain similarly as before, for (X,Y) distributed according to the optimal coupling of μ and ν

$$\begin{aligned}\left|\mathbb{E}(X^2) \wedge K\right. &\left. - \mathbb{E}(X)^2 \wedge K - \mathbb{E}(Y^2) \wedge K + \mathbb{E}(Y)^2 \wedge K\right| \\ &\leq \left|\mathbb{E}(X^2) \wedge K - \mathbb{E}(Y^2) \wedge K\right| + \left|\mathbb{E}(X)^2 \wedge K - \mathbb{E}(Y)^2 \wedge K\right| \\ &\leq 2\sqrt{K}\left|\sqrt{\mathbb{E}(X^2)} - \sqrt{\mathbb{E}(Y^2)}\right| + 2\sqrt{K}\left||\mathbb{E}(X)| - |\mathbb{E}(Y)|\right|.\end{aligned}$$

On one hand we have by Cauchy-Schwarz inequality

$$\mathbb{E}((X-Y)^2) - \left|\sqrt{\mathbb{E}(X^2)} - \sqrt{\mathbb{E}(Y^2)}\right|^2 = 2\sqrt{\mathbb{E}X^2}\sqrt{\mathbb{E}Y^2} - 2\mathbb{E}(XY) \geq 0,$$

which implies

$$\left|\sqrt{\mathbb{E}(X^2)} - \sqrt{\mathbb{E}(Y^2)}\right| \leq \sqrt{\mathbb{E}((X-Y)^2)} = \mathcal{W}_2(\mu,\nu).$$

On the other hand

$$||\mathbb{E}(X)| - |\mathbb{E}(Y)|| \leq |\mathbb{E}(X) - \mathbb{E}(Y)| \leq \mathbb{E}|X-Y| \leq \sqrt{\mathbb{E}((X-Y)^2)} = \mathcal{W}_2(\mu,\nu).$$

Hence, $\bar{\sigma}_K$ is Lipschitz continuous.

Thus, [CdS23, Theorem 2.5] applies and we have strong existence and uniqueness for (3.8)*. Furthermore, up until the stopping time*

$$T_K = \inf\left\{ t \geq 0 \text{ s.t. } \int x^2 d\rho_t^K(x) > K \right\},$$

the solution of (3.8) *coincides with the solution of* (Eff)*. Then, applying Itô's formula on the process* (Eff) *for the function $x \mapsto x^2$, we obtain that the second moment of the solution is bounded. Hence, for K large enough depending on the initial condition, $T_K = \infty$. We finally obtain the wellposedness of* (Eff)*.*

Our goal is to compare the stationary distribution(s) for (Eff) to the stationary distribution(s) for (NL), in particular in regards to this critical parameter σ_c. To do so, we begin by showing that any stationary distribution of (NL) is a stationary distribution of (Eff), and conversely.

Lemma 3.1. *Let μ be a probability measure on $\mathbb{R}$.*

- *If μ is a solution of* (3.2) *for a diffusion coefficient σ', then μ is a solution of* (3.1) *for a diffusion coefficient $\sigma = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$,*
- *If μ is a solution of* (3.1) *for a diffusion coefficient σ and $\frac{\delta}{p-1} < \frac{2\sigma}{L_W^2 \kappa_2(\mu)}$, then μ is a solution of* (3.2) *for a diffusion coefficient $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$.*

Proof. Let us prove the two points.
A stationary distribution of (Eff) **is a stationary distribution of** (NL)**.** Assume μ satisfies (3.2) for a given (δ, p, σ'), which in particular is equivalent to the pair $(\kappa_1(\mu), \kappa_2(\mu))$ satisfying

$$\begin{aligned}
\kappa_1(\mu) =& \frac{\int_{\mathbb{R}} x \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx} \\
\kappa_2(\mu) =& \frac{\int_{\mathbb{R}} (x - \kappa_1(\mu))^2 \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx}.
\end{aligned}$$

Denoting $\sigma = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$, we thus have

$$\begin{aligned}
\kappa_1(\mu) =& \frac{\int_{\mathbb{R}} x \exp\left(-\frac{1}{\sigma} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx} \\
\kappa_2(\mu) =& \frac{\int_{\mathbb{R}} (x - \kappa_1(\mu))^2 \exp\left(-\frac{1}{\sigma} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma} \left(U(x) + \frac{L_W}{2}|x - \kappa_1(\mu)|^2 \right) \right) dx}.
\end{aligned}$$

Thus, if μ is a stationary distribution for (Eff) with diffusion coefficient σ' and parameters δ and p, it is also a stationary distribution for (NL) with diffusion coefficient $\sigma = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$.

A stationary distribution of (NL) **is a stationary distribution of** (Eff)**.** Assume μ satisfies (3.2) for a given σ, which in particular is equivalent to the pair $(\kappa_1(\mu), \kappa_2(\mu))$ satisfying

$$\kappa_1(\mu) = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{1}{\sigma}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}$$
$$\kappa_2(\mu) = \frac{\int_{\mathbb{R}} (x-\kappa_1(\mu))^2 \exp\left(-\frac{1}{\sigma}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}.$$

Consider parameters δ and p, and denote $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$. Notice $\kappa_2(\mu)$ is independent of δ and p thus, provided $\frac{\delta}{(p-1)}$ is small enough, we may ensure $\sigma' > 0$ and have

$$\kappa_1(\mu) = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}$$
$$\kappa_2(\mu) = \frac{\int_{\mathbb{R}} (x-\kappa_1(\mu))^2 \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma' + \frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x) + \frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}.$$

Thus, given two parameters δ and p, if μ is a stationary distribution for (NL) with diffusion coefficient σ, it is also a stationary distribution for (Eff) with diffusion coefficient $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$. □

Unfortunately, this lemma does not directly imply the existence of a phase transition for (Eff). Several issues arise :

- the existence of a symmetric stationary distribution $\mu_{\sigma,0}$ for (NL) with diffusion coefficient $\sigma > 0$ only yields the existence of a symmetric stationary distribution for (Eff) for a specific diffusion coefficient $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$. We need to show that, for any $\sigma > 0$, there exists a symmetric stationary distribution for (Eff).
- likewise, the existence of non symmetric stationary distribution for (Eff) is only ensured for specific diffusion coefficients.
- we cannot infer the uniqueness of the symmetric stationary distribution for (Eff) from the uniqueness of the symmetric stationary distribution for (NL). Given $\sigma' > 0$, there may *a priori* be two stationary distributions for (Eff), denoted μ_1 and μ_2, with different variance $\kappa_2(\mu_1) \neq \kappa_2(\mu_2)$, which thus correspond to two different symmetric stationary distributions for (NL) with diffusion coefficients $\sigma_1 = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_1) \neq \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_2) = \sigma_2$.

We therefore dedicate the remainder of this document to the proof of Theorem 3.

3.1 Some results on the stationary distribution(s) of (NL)

The study of the critical parameter of (Eff) relies, by Lemma 3.1, on the study of the one of (NL). We gather here some results concerning the latter. They are numerically illustrated in Figure 6, and the proofs are postponed to Appendix B.

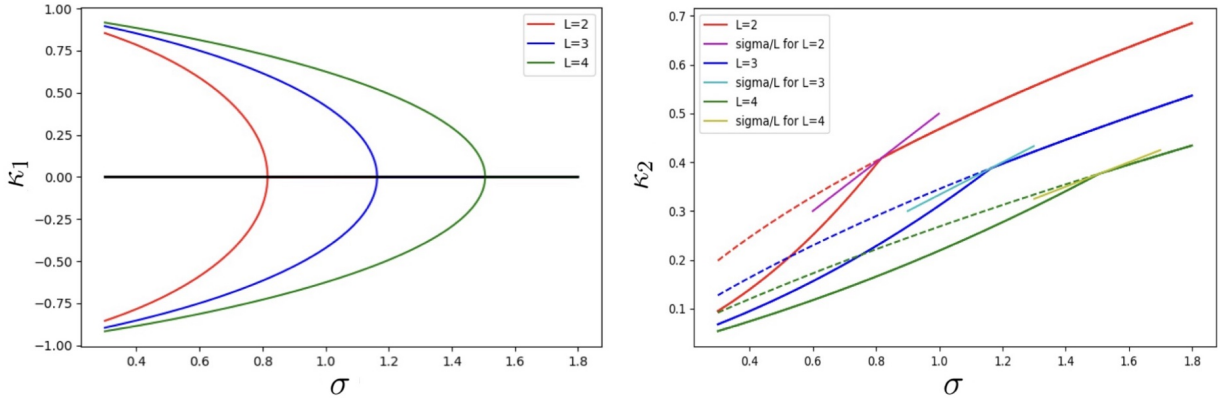

Figure 6: **Left** : The means of $\mu_{\sigma,\pm}$ as a function of σ for different values of L_W, as given in Theorem 2. **Right** : The variances of $\mu_{\sigma,\pm}$ as a function of σ for different values of L_W. In dotted line the variance of $\mu_{\sigma,0}$, and in solid line the variance of $\mu_{\sigma,\pm}$.

Lemma 3.2. *We have the following results concerning the stationary distribution(s) of* (NL).

- ***Symmetry.*** *We have* $\kappa_1(\mu_{\sigma,+}) = -\kappa_1(\mu_{\sigma,-})$ *and* $\kappa_2(\mu_{\sigma,+}) = \kappa_2(\mu_{\sigma,-})$.
- ***Moment bound.*** *Let the symbol* $* \in \{0, \pm\}$. *Consider* $\mu_{\sigma,*}$ *given in Theorem 2, and* $\kappa_1(\mu_{\sigma,*})$ *(resp.* $\kappa_2(\mu_{\sigma,*})$*) the corresponding mean (resp. variance). There exists* $C_{\kappa_1}, C_{\kappa_2} > 0$ *such that for* $\sigma \in [0, \sigma_c]$ *we have*
$$|\kappa_1(\mu_{\sigma,*})| \leq C_{\kappa_1}, \quad \text{and} \quad |\kappa_2(\mu_{\sigma,*})| \leq C_{\kappa_2}. \tag{3.9}$$
- ***Critical variance.*** *We have the equality*
$$\kappa_2\left(\mu_{\sigma_c}\right) = \frac{\sigma_c}{L_W}. \tag{3.10}$$
Furthermore, for $\sigma < \sigma_c$ *we have* $\kappa_2\left(\mu_{\sigma,\pm}\right) < \frac{\sigma}{L_W}$ *and* $\kappa_2\left(\mu_{\sigma,0}\right) > \frac{\sigma}{L_W}$, *and for* $\sigma > \sigma_c$ *we have* $\kappa_2\left(\mu_{\sigma,0}\right) < \frac{\sigma}{L_W}$.
- ***Continuity.*** *The function* $\sigma \mapsto \kappa_1(\mu_{\sigma,+})$, *with the convention* $\mu_{\sigma,+} = \mu_{\sigma,0}$ *for* $\sigma \geq \sigma_c$, *is continuous on* $]0, \infty[$. *In particular, this also yields the continuity of* $\sigma \mapsto \kappa_2(\mu_{\sigma,+}) = f_2(\sigma, \kappa_1(\mu_{\sigma,+}))$.
- ***Lipschitz continuity.*** *Let* $\sigma_0 > 0$. *The functions* $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ *and* $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$ *are Lipschitz continuous, respectively on* $[\sigma_0, \infty[$ *and on* $[\sigma_0, \sigma_c]$. *More precisely, there exists* $C > 0$ *such that for respectively* $\sigma > \sigma_0$ *and* $\sigma \in]\sigma_0, \sigma_c[$ *we have* $\left|\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})\right| \leq C$.

Remark 3.2. *The bound* (3.9)*, combined with the knowledge of the fact that for* $\sigma \geq \sigma_c$ *there only exists a symmetric stationary distribution for* (NL) *as well as Lemma 3.1, shows that we can restrict our study of the stationary distribution for both* (NL) *and* (Eff) *to a compact set of means* $\kappa_1 \in [-C_{\kappa_1}, C_{\kappa_1}]$.

Remark 3.3. *The main technical difficulty lies in the proof of the Lipschitz continuity of* $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ *and* $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$, *since it turns out that the mean* $\sigma \mapsto \kappa_1(\mu_{\sigma,\pm})$ *is not Lipschitz continuous near the critical parameter* σ_c *(See Figure 6). It therefore requires a careful estimation of the mean and variance around* σ_c, *and the proof is a section of its own (See Appendix B.2).*

3.2 Phase transition for the effective dynamics

Let $\sigma_0 > 0$ and define, for $* \in \{0, \pm\}$, the function $g_{eff,*} : \sigma \mapsto \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma,*})$.

From Lemma 3.1, if $\mu_{\sigma,*}$ is a stationary distribution for (NL), then it is a stationary distribution for (Eff) with diffusion coefficient $\sigma' = g_{eff,*}(\sigma)$.

By Lemma 3.2, $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ is a Lipschitz continuous function on $[\sigma_0, \infty[$ and $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$ is a Lipschitz continuous function on $[\sigma_0, \sigma_c]$, and, more precisely, in both cases we obtain that $\left|\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})\right|$ is bounded by some constant $C > 0$.

In this case, the function $\sigma \mapsto g_{eff,*}(\sigma)$ is such that $g'_{eff,*}(\sigma) = 1 - \frac{\delta L_W^2}{2(p-1)}\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})$ and thus $g'_{eff,*}(\sigma) \in \left[1 - \frac{\delta L_W^2}{2(p-1)}C, 1 + \frac{\delta L_W^2}{2(p-1)}C\right]$. In particular, for $\frac{\delta}{p-1}$ sufficiently small, $g_{eff,*}(\sigma)$ is both an increasing continuous function and positive.

Thus, $g_{eff,0}$ and $g_{eff,\pm}$ are two injective functions. In particular, $g_{eff,\pm}$ is a bijection from $[\sigma_0, \sigma_c]$ to $[g_{eff,\pm}(\sigma_0), g_{eff,\pm}(\sigma_c)]$.

Finally, notice that $g_{eff,\pm}(\sigma_c) = g_{eff,0}(\sigma_c) = \sigma_c\left(1 - \frac{\delta L_W}{2(p-1)}\right)$, that $g_{eff,*}(\sigma_0) \leq \sigma_0$, and that, up to the additional assumption $\frac{2(\sigma_c - \sigma_0)}{\sigma_c L_W} > \frac{\delta}{p-1}$, we may assume $g_{eff,*}(\sigma_c) > \sigma_0$.

We may now state the following facts concerning the stationary distribution(s) for (Eff).

- **There exists at least one symmetric stationary distribution.** Since $g'_{eff,0} \geq 1 - \frac{\delta L_W^2}{2(p-1)}C$, $g_{eff,0}$ is an increasing function such that $g_{eff,0}(x) \xrightarrow[x\to\infty]{} \infty$. Thus, $g_{eff,0}$ is a bijection from $[\sigma_0, \infty[$ to $[g_{eff,0}(\sigma_0), \infty[$. Therefore, for all $\sigma \in [g_{eff,0}(\sigma_0), \infty[$, there exists $\tilde{\sigma}$ such that $g_{eff,0}(\tilde{\sigma}) = \sigma$. In other words, $\mu_{\tilde{\sigma},0}$ is also a symmetric stationary distribution for (Eff) with diffusion coefficient σ.

- **There exists at most one symmetric stationary distribution.** Let $\sigma \geq \sigma_0$ and $\kappa_1 = 0$, and assume there are two symmetric stationary distributions of (Eff) with diffusion coefficient σ. This yields two coefficients $\sigma', \sigma'' \geq \sigma > 0$ such that

$$\begin{aligned}\sigma' =& \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right)\\ \sigma'' =& \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right),\end{aligned}$$

  where $\mu_{\sigma,1}^{\delta,p}$ and $\mu_{\sigma,2}^{\delta,p}$ denote the two stationary distributions. We consider $\mu_{\sigma'}$ $(= \mu_{\sigma,1}^{\delta,p})$ and $\mu_{\sigma''}$ $(= \mu_{\sigma,2}^{\delta,p})$ the corresponding (unique) symmetric stationary distributions of (NL). Because σ' and σ'' are greater than σ_0, there exists a constant K, possibly depending on σ_0 and L_W, such that by Lemma 3.2

$$\begin{aligned}&|\kappa_2(\mu_{\sigma''}) - \kappa_2(\mu_{\sigma'})| = \left|f_2(\sigma'', 0) - f_2(\sigma', 0)\right| \leq K|\sigma'' - \sigma'|\\ \text{i.e.}\quad &|\kappa_2(\mu_{\sigma''}) - \kappa_2(\mu_{\sigma'})| \leq \frac{K\delta L_W^2}{2(p-1)}\left|\kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right) - \kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right)\right|. \qquad (3.11)\end{aligned}$$

  Because the stationary distributions of (Eff) are uniquely defined by their mean and variance,

$$\kappa_2(\mu_{\sigma''}) = \kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right) \quad \text{and} \quad \kappa_2(\mu_{\sigma'}) = \kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right),$$

  and we obtain from (3.11) that, for $\frac{\delta}{p-1}$ sufficiently small, $\kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right) = \kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right)$ and thus that $\mu_{\sigma,1}^{\delta,p} = \mu_{\sigma,2}^{\delta,p}$.

- **For $\sigma \in [g_{eff,\pm}(\sigma_0), g_{eff,\pm}(\sigma_c)[$, there exists at least two stationary distributions with nonzero mean.** Because $g_{eff,\pm}$ is a bijection, consider $g^{-1}_{eff,\pm}(\sigma) \in [\sigma_0, \sigma_c[$. There are three stationary distribution for (NL) with diffusion coefficient $g^{-1}_{eff,\pm}(\sigma)$. By Lemma 3.1, $\mu_{g^{-1}_{eff,\pm}(\sigma),+}$ and $\mu_{g^{-1}_{eff,\pm}(\sigma),-}$ are also stationary distributions for (Eff) with diffusion coefficient σ, and they have nonzero mean.

- **For $\sigma \in [g_{eff,\pm}(\sigma_0), g_{eff,\pm}(\sigma_c)[$, there exists at most two stationary distributions with nonzero mean.** By symmetry, it is sufficient to prove that there is at most one stationary distribution with positive mean. Assume there are two such solutions $\mu^{\delta,p}_{\sigma,+,1}$ and $\mu^{\delta,p}_{\sigma,+,2}$.

  Let, for $i \in \{1,2\}$, $\sigma_i = \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu^{\delta,p}_{\sigma,+,i})$. Then, $\mu^{\delta,p}_{\sigma,+,i}$ is a stationary distribution with a positive mean for (NL) with diffusion coefficient σ_i, i.e $\mu^{\delta,p}_{\sigma,+,i} = \mu_{\sigma_i,+}$.

  Thus $\sigma = \sigma_i - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu^{\delta,p}_{\sigma,+,i}) = \sigma_i - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma_i,+}) = g_{eff,+}(\sigma_i)$. Since $g_{eff,+}$ is an injective function, we obtain that $\sigma_1 = \sigma_2$. In particular, $\mu^{\delta,p}_{\sigma,+,1}$ and $\mu^{\delta,p}_{\sigma,+,2}$ are two stationary distribution with a positive mean for (NL) with diffusion coefficient $\sigma_1 = \sigma_2$, thus by uniqueness $\mu^{\delta,p}_{\sigma,+,1} = \mu^{\delta,p}_{\sigma,+,2}$.

- **For $\sigma \geq g_{eff,+}(\sigma_c)$, there does not exists stationary distribution with nonzero mean**. The result is direct if $\sigma \geq \sigma_c$, because if μ is a stationary measure for (Eff) with diffusion coefficient σ, it is is a stationary measure for (NL) with diffusion coefficient $\sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu) \geq \sigma > \sigma_c$, hence it cannot have a nonzero mean.

  Assume $\sigma_c > \sigma \geq g_{eff,+}(\sigma_c) > \sigma_0$ and that there exists such a solution $\mu^{\delta,p}_{\sigma,+}$. Consider

$$\sigma' = \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu^{\delta,p}_{\sigma,+}), \tag{3.12}$$

  such that $\mu^{\delta,p}_{\sigma,+} = \mu_{\sigma',+}$ is a stationary distribution with positive mean for (NL). We thus necessarily have $\sigma' < \sigma_c$. Let

$$\tilde{\sigma} = g_{eff,+}(\sigma') = \sigma' - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma',+}). \tag{3.13}$$

  We obtain $\mu^{\delta,p}_{\sigma,+} = \mu_{\sigma',+} = \mu^{\delta,p}_{\tilde{\sigma},+}$. Since $g_{eff,+}$ is increasing we have $g_{eff,+}(\sigma_c) > g_{eff,+}(\sigma') = \tilde{\sigma}$, we obtain from (3.13)

$$\begin{aligned} g_{eff,+}(\sigma_c) >& \sigma' - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma',+}) = \sigma' - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu^{\delta,p}_{\sigma,+}), \\ \text{i.e} \quad \kappa_2(\mu^{\delta,p}_{\sigma,+}) >& \frac{2(p-1)}{\delta L_W^2}\left(\sigma' - g_{eff,+}(\sigma_c)\right). \end{aligned}$$

  Plugging that back into (3.12), we obtain

$$\sigma' > \sigma + \sigma' - g_{eff,+}(\sigma_c), \quad \text{i.e} \quad \sigma < g_{eff,+}(\sigma_c),$$

  which contradicts the initial assumption.

This concludes the proof of Theorem 3.

A Technical lemmas

We start with this slight extension of the central limit theorem:

Lemma A.1. *Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d random variables in $\mathbb{R}$ such that $\mathbb{E}X_1 = 0$ and $\mathbb{E}(|X_1|^2) = 1$. Assume also that $\mathbb{E}(|X_1|^4) < \infty$. Denote $Z_p = \frac{1}{\sqrt{p}} \sum_{i=1}^p X_i$.*

Then, we have for $Z \sim \mathcal{N}(0,1)$

$$\mathbb{E}\left(|Z_p|^2\right) = \mathbb{E}\left(|Z|^2\right), \quad \textit{and} \quad \mathbb{E}\left(|Z_p|^4\right) = \mathbb{E}\left(|Z|^4\right) + O\left(\frac{1}{p}\right).$$

This in particular also yields $\mathbb{E}\left(|Z_p|\right) \xrightarrow[p\to\infty]{} \mathbb{E}\left(|Z|\right)$ and $\mathbb{E}\left(|Z_p|^3\right) \xrightarrow[p\to\infty]{} \mathbb{E}\left(|Z|^3\right)$

Proof. Direct computations give $\mathbb{E}\left(|Z_p|^2\right) = \mathbb{E}\left(|Z|^2\right)$. Likewise, we may explicitly compute $\mathbb{E}|Z_p|^4$. Keeping only the terms with nonzero expectation, we have

$$\begin{aligned}\mathbb{E}|Z_p|^4 =& \frac{1}{p^2}\sum_{i=1}^p \mathbb{E}|X_i|^4 + \frac{6}{p^2}\sum_{i>j}\mathbb{E}|X_i|^2\mathbb{E}|X_j|^2 \\ =& \frac{\mathbb{E}|X_1|^4}{p} + \frac{6}{p^2}\frac{p(p-1)}{2}.\end{aligned}$$

Noticing that $\mathbb{E}|Z|^4 = 3$ yields the convergence $\mathbb{E}\left(|Z_p|^4\right) = \mathbb{E}\left(|Z|^4\right) + O\left(\frac{1}{p}\right)$. Thus, we have both

- Z_p converges in law to $Z \sim \mathcal{N}(0,1)$,
- and the convergence of the fourth moment $\mathbb{E}\left|Z_p\right|^4 \xrightarrow[p\to\infty]{} \mathbb{E}\left|Z\right|^4$.

By [Vil09, Theorem 6.9], we have the convergence in L^4 Wasserstein distance of the law of Z_p to a law $\mathcal{N}(0,1)$. This implies the convergence in both L^1 Wasserstein distance and L^3 Wasserstein distance, thus the convergence of the first and third moments of Z_p. □

Lemma A.2. *The function $\sigma \in]0,\infty[\mapsto \partial_\kappa^3 f_1(\sigma, 0)$ (with f_1 given by (3.6)) is continuous and satisfies*

$$\forall \sigma > 0, \quad \partial_\kappa^3 f_1(\sigma, 0) < 0.$$

Proof. We have

$$\partial_\kappa^3 f_1(\sigma,0) = \left(\frac{L_W}{\sigma}\right)^3 \left(\frac{\int_{\mathbb{R}} x^4 g}{\int_{\mathbb{R}} g} - 3\left(\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g}\right)^2\right),$$

with g given by (3.5). We wish to prove

$$A(\sigma, L_W) := \frac{\int_{\mathbb{R}} x^4 g \int_{\mathbb{R}} g}{\left(\int_{\mathbb{R}} x^2 g\right)^2} < 3.$$

Remark that $A(\sigma, L_W)$ is by definition the kurtosis of a random variable with probability density $\frac{g}{\int g}$.

Let us rewrite, for $\alpha > 0$

$$\begin{aligned}U(x) + \frac{L_W}{2}x^2 =& \frac{x^4}{4} + \frac{L_W - 1}{2}x^2 \\ =& \frac{x^4}{4} + \frac{L_W - 1 - \alpha}{2}x^2 + \left(\frac{L_W - 1 - \alpha}{2}\right)^2 + \frac{\alpha}{2}x^2 - \left(\frac{L_W - 1 - \alpha}{2}\right)^2\end{aligned}$$

$$=\frac{1}{4}\left(x^2+L_W-1-\alpha\right)^2+\frac{\alpha}{2}x^2-\left(\frac{L_W-1-\alpha}{2}\right)^2,$$

such that

$$g(x,\sigma,0)=\exp\left(-\frac{1}{4\sigma}\left(x^2+L_W-1-\alpha\right)^2\right)\exp\left(\frac{1}{\sigma}\left(\frac{L_W-1-\alpha}{2}\right)^2\right)e^{-\frac{\alpha}{2\sigma}x^2}.$$

This way we can write

$$\begin{aligned}A(\sigma,L_W)=&\frac{\mathbb{E}\left(\exp\left(-\frac{1}{4\sigma}\left(Y^2+L_W-1-\alpha\right)^2\right)\right)\mathbb{E}\left(Y^4\exp\left(-\frac{1}{4\sigma}\left(Y^2+L_W-1-\alpha\right)^2\right)\right)}{\mathbb{E}\left(Y^2\exp\left(-\frac{1}{4\sigma}\left(Y^2+L_W-1-\alpha\right)^2\right)\right)^2},\\ &\qquad\qquad\text{with } Y\sim\mathcal{N}\left(0,\frac{\sigma}{\alpha}\right),\\ =&\frac{\mathbb{E}\left(\exp\left(-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2+L_W-1-\alpha\right)^2\right)\right)\mathbb{E}\left(X^4\exp\left(-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2+L_W-1-\alpha\right)^2\right)\right)}{\mathbb{E}\left(X^2\exp\left(-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2+L_W-1-\alpha\right)^2\right)\right)^2},\\ &\qquad\qquad\text{with } X\sim\mathcal{N}\left(0,1\right).\end{aligned}$$

We have

$$-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2+L_W-1-\alpha\right)^2=-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}\right)^2\left(X^2+\frac{\alpha\left(L_W-1-\alpha\right)}{\sigma}\right)^2.$$

We thus choose $\alpha=\frac{1}{2}\left(L_W-1+\sqrt{(L_W-1)^2+4\sigma}\right)>0$ in order to ensure $\frac{\alpha(L_W-1-\alpha)}{\sigma}=-1$. Finally, denoting

$$\beta(\sigma,L_W)=\left(\frac{L_W-1}{\sqrt{\sigma}}+\sqrt{\left(\frac{L_W-1}{\sqrt{\sigma}}\right)^2+4}\right)^{-2}>0\,,$$

we have

$$\begin{aligned}A(\sigma,L_W)=&\frac{\mathbb{E}\left(\exp\left(-\beta(\sigma,L_W)\left(X^2-1\right)^2\right)\right)\mathbb{E}\left(X^4\exp\left(-\beta(\sigma,L_W)\left(X^2-1\right)^2\right)\right)}{\mathbb{E}\left(X^2\exp\left(-\beta(\sigma,L_W)\left(X^2-1\right)^2\right)\right)^2},\quad X\sim\mathcal{N}\left(0,1\right)\\ :=&A(\beta(\sigma,L_W)).\end{aligned}\tag{A.1}$$

The quantity A given above can be expressed as a function of $\beta(\sigma,L_W)\in]0,\infty[$, that we denote $A(\beta)$. We may then numerically check that $A(\beta)<3$ for all $\beta>0$. (see Figure 7).

Notice that the function $\beta(\sigma,L_W)$, which is in reality a function of the quantity $\frac{L_W-1}{\sqrt{\sigma}}$, is a bijection from $\frac{L_W-1}{\sqrt{\sigma}}\in\mathbb{R}$ to $]0,\infty[$, which satisfies $\beta(\sigma,L_W)\xrightarrow[\frac{L_W-1}{\sqrt{\sigma}}\to\infty]{}0$. And, for $\beta(\sigma,L_W)=0$, direct calculations knowing the moments of the Gaussian law yield $A(0)=3$ and $A'(0)=-24<0$. □

Lemma A.3. *Consider the function $F_1:\sigma\mapsto\partial_\kappa f_1(\sigma,0)$. It is continuously differentiable and, for $\sigma>0$, satisfies $F_1'(\sigma)<0$.*

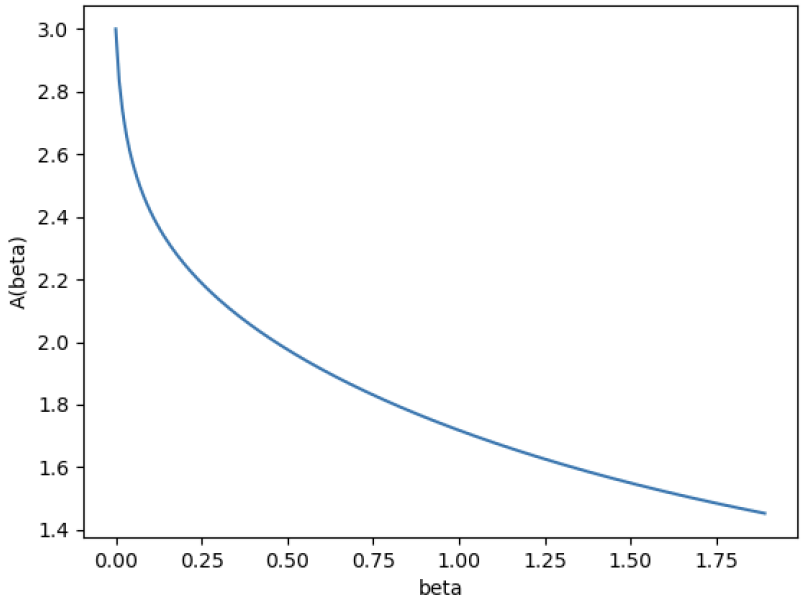

Figure 7: Numerical simulation of the quantity given in (A.1).

Proof. We have

$$F_1(\sigma) = \frac{L_W}{\sigma} \frac{\int x^2 \exp\left(-\frac{x^4}{4\sigma} + \frac{1-L_W}{2\sigma}x^2\right) dx}{\int \exp\left(-\frac{x^4}{4\sigma} + \frac{1-L_W}{2\sigma}x^2\right) dx}.$$

Consider the change of variable $y = \frac{x}{\sqrt{\sigma}}$. We have

$$\begin{aligned} F_1(\sigma) =& \frac{L_W}{\sigma} \frac{\int \sigma y^2 \exp\left(-\sigma^2 \frac{y^4}{4\sigma} + \frac{1-L_W}{2\sigma}\sigma y^2\right) dy}{\int \exp\left(-\sigma^2 \frac{y^4}{4\sigma} + \frac{1-L_W}{2\sigma}\sigma y^2\right) dy} \\ =& L_W \frac{\int y^2 \exp\left(-\frac{\sigma}{4}y^4 + \frac{1-L_W}{2}y^2\right) dy}{\int \exp\left(-\frac{\sigma}{4}y^4 + \frac{1-L_W}{2}y^2\right) dy}, \end{aligned}$$

which then yields

$$F_1'(\sigma) = \frac{L_W}{4}\left(-\mathbb{E}\left(Y^6\right) + \mathbb{E}\left(Y^2\right)\mathbb{E}\left(Y^4\right)\right),$$

where Y is a random variable with probability density (up to renormalization) $\exp\left(-\frac{\sigma}{4}y^4 + \frac{1-L_W}{2}y^2\right) dy$. By Jensen inequality (in a strictly convex case with a non almost surely constant random variable), we have $F_1'(\sigma) < 0$. □

B Proofs of Lemma 3.2

In this section we prove the various results of Lemma 3.2.

B.1 Moment bounds, critical variance and continuity

To prove the first properties stated in Lemma 3.2, we recall the following result, extracted from Theorem 2.1 of [Tug14] and its proof.

Lemma B.1. *The equation (with unknown σ)*

$$\int_{\mathbb{R}^+} \left(x^2 - \frac{1}{2L_W} \right) \exp\left((1 - L_W)x^2 - \sigma x^4 \right) dx = 0, \tag{B.1}$$

admits a unique solution, that is the critical value σ_c.

Finally, consider the function

$$\xi(\sigma, \kappa) = \int_{\mathbb{R}} (x - \kappa) \exp\left(-\frac{1}{\sigma} \left(U(x) + \frac{L_W}{2} x^2 - L_W x \kappa \right) \right) dx. \tag{B.2}$$

We have the following properties on ξ :

- *The function $\sigma \mapsto \partial_\kappa \xi(\sigma, 0)$ is decreasing : for $\sigma < \sigma_c$ we have $\partial_\kappa \xi(\sigma, 0) > 0$, for $\sigma > \sigma_c$ we have $\partial_\kappa \xi(\sigma, 0) < 0$, and finally $\partial_\kappa \xi(\sigma_c, 0) = 0$.*
- *For $\sigma \geq \sigma_c$ and $\kappa \geq 0$, the function $\kappa \mapsto \xi(\sigma, \kappa)$ is decreasing (which, in fact, ensures uniqueness of the stationary solution for* (NL)*).*
- *For $\sigma < \sigma_c$ and $\kappa \geq 0$, the function $\kappa \mapsto \xi(\sigma, \kappa)$ is increasing and then decreasing (which,in fact, ensures the thirdness of the stationary solution for* (NL)*).*

Moment bound : Consider $(\bar{X}_t)_t$ the solution of (NL). Itô's formula yields

$$dU(\bar{X}_t) = A_t dt + dM_t,$$

where M_t is a continuous local martingale and

$$A_t = -U'\left(\bar{X}_t\right)^2 - L_W \left(\bar{X}_t - \mathbb{E}\left(\bar{X}_t\right) \right) U'\left(\bar{X}_t\right) + \sigma U''\left(\bar{X}_t\right).$$

There exists $\lambda > 0$ and $C > 0$, both independent of $\sigma \in [0, \sigma_c]$, such that for all $x \in \mathbb{R}$

$$\sigma U''(x) + 2\lambda U(x) \leq \frac{U'(x)^2}{2} + C, \qquad \text{and} \qquad 2L_W^2 x^2 \leq \lambda U(x) + C.$$

Consider for instance $\lambda = 1$ and $C = \max\left(2\sigma + (2\sigma)^{3/2} - \frac{1}{2}, \frac{(1+4L_W^2)^2}{4} \right)$ for $U(x) = \frac{x^4}{4} - \frac{x^2}{2}$. Thus

$$A_t \leq C - \lambda U\left(\bar{X}_t\right) + \left(L_W^2 \left(\bar{X}_t^2 + \mathbb{E}\left(\bar{X}_t\right)^2 \right) - \lambda U\left(\bar{X}_t\right) \right),$$

and, using Fatou's lemma to deal with the local martingale, finally we obtain thanks to Gronwall's lemma

$$\mathbb{E}U\left(\bar{X}_t\right) \leq e^{-\lambda t} \mathbb{E}U\left(\bar{X}_0\right) + \frac{2C}{\lambda}.$$

Since $\mathbb{E}\left(\bar{X}_t\right)^2 \leq \mathbb{E}\left(\bar{X}_t^2\right) \leq \frac{1}{2L_W^2}\left(\lambda \mathbb{E}U\left(\bar{X}_t\right) + C \right)$, and considering $\bar{X}_0$ distributed according to a stationary distribution, we may conclude.

Value of $\kappa_2(\mu_{\sigma_c})$: We rewrite Equation (B.1) defining σ_c, first by using the symmetry in x to obtain

$$\int_{\mathbb{R}} \left(x^2 - \frac{1}{2L_W} \right) \exp\left((1 - L_W)x^2 - \sigma x^4 \right) dx = 0,$$

and then, by a change of variable $x = \frac{y}{\sqrt{2\sigma}}$, this is equivalent to

$$\int_{\mathbb{R}} \left(y^2 - \frac{\sigma}{L_W} \right) \exp\left(-\frac{1}{\sigma} \left(\frac{y^4}{4} - \frac{1 - L_W}{2} y^2 \right) \right) dy = 0.$$

Finally, this amounts to having

$$\frac{\sigma}{L_W} = \frac{\int_{\mathbb{R}} y^2 \exp\left(-\frac{1}{\sigma} \left(\frac{y^4}{4} - \frac{1-L_W}{2} y^2 \right) \right) dy}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma} \left(\frac{y^4}{4} - \frac{1-L_W}{2} y^2 \right) \right) dy},$$

which, since $\kappa_1(\mu_{\sigma_c}) = 0$, yields the value of $\kappa_2(\mu_{\sigma_c})$.

Consider then the function ξ given (B.2). We have $\xi(\sigma, \kappa) = (f_1(\sigma, \kappa) - \kappa) \int g(x, \sigma, \kappa) dx$, and thus

$$\partial_\kappa \xi(\sigma, \kappa) = (\partial_\kappa f_1(\sigma, \kappa) - 1) \int g(x, \sigma, \kappa) dx + (f_1(\sigma, \kappa) - \kappa) \int \partial_\kappa g(x, \sigma, \kappa) dx.$$

Considering the equation above for $\kappa = \kappa_1(\mu_{\sigma,*})$, we obtain

$$\partial_\kappa \xi(\sigma, \kappa_1(\mu_{\sigma,*})) = (\partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,*})) - 1) \int g(x, \sigma, \kappa_1(\mu_{\sigma,*})) dx.$$

We may compute the derivatives of f_1 (see (B.5) below), and obtain

$$\partial_\kappa \xi(\sigma, \kappa_1(\mu_{\sigma,*})) = \left(\frac{L_W}{\sigma} \kappa_2(\mu_{\sigma,*}) - 1 \right) \int g(x, \sigma, \kappa_1(\mu_{\sigma,*})) dx.$$

The values of $\partial_\kappa \xi(\sigma, \kappa)$ for $\kappa = 0$ and $\kappa = \kappa_1(\mu_{\sigma,+})$ depending on σ, as given in Lemma B.1, yields the result.

Continuity of the moments : Notice that f_1 given in (3.6) is continuous on $(\sigma, \kappa) \in \mathbb{R}^{+,*} \times \mathbb{R}^+$. We start by proving the continuity of $\sigma \mapsto \kappa_1(\mu_{\sigma,+})$, with the convention $\mu_{\sigma,+} = \mu_{\sigma,0}$ for $\sigma > \sigma_c$. In this latter case, the function $\sigma \mapsto \kappa_1(\mu_{\sigma,+})$ is trivially continuous as $\kappa_1(\mu_{\sigma,+}) = 0$.

Let us show the continuity at the point σ_c. Let $(\sigma_n)_{n \in \mathbb{N}}$ be a sequence of positive real numbers such that $\sigma_n \xrightarrow[n \to \infty]{} \sigma_c$, and consider the (bounded) sequence $\left(\kappa_1(\mu_{\sigma_n,+})\right)_n$. Up to extraction, we can assume $\kappa_1(\mu_{\sigma_n,+}) \xrightarrow[n \to \infty]{} \kappa_1 \geq 0$. We have, by definition, $\kappa_1(\mu_{\sigma_n,+}) = f_1(\sigma_n, \kappa_1(\mu_{\sigma_n,+}))$ and, by considering the limit $n \to \infty$, thanks to the continuity of f_1, we obtain $\kappa_1 = f_1(\sigma_c, \kappa_1)$. Uniqueness of the fixed point for σ_c then ensures $\kappa_1 = 0 = \kappa_1(\mu_{\sigma_c,+})$. Hence we obtain the desired continuity.

We now consider $\sigma < \sigma_c$. Assume there exists $\epsilon > 0$ and a sequence $(\sigma_n)_{n \in \mathbb{N}}$ such that $\sigma_n \xrightarrow[n \to \infty]{} \sigma$ and $|\kappa_1(\mu_{\sigma_n,+}) - \kappa_1(\mu_{\sigma,+})| > \epsilon$. Again, up to extraction, we have $\kappa_1(\mu_{\sigma_n,+}) \xrightarrow[n \to \infty]{} \kappa_1 \geq 0$ and, since κ_1 is a fixed point that cannot be $\kappa_1(\mu_{\sigma,+})$, we have $\kappa_1 = 0$. Consider the (at least) twice continuously differentiable function ξ given in (B.2). On one hand, we have by continuity $\partial_\kappa \xi(\sigma_n, \kappa_1(\mu_{\sigma_n,+})) \xrightarrow[n \to \infty]{} \partial_\kappa \xi(\sigma, 0) > 0$. On the other hand, by the properties of ξ given in Theorem B.1 , we have $\partial_\kappa \xi(\sigma_n, \kappa_1(\mu_{\sigma_n,+})) < 0$. Hence a contradiction, and $\kappa_1(\mu_{\sigma_n,+}) \xrightarrow[n \to \infty]{} \kappa_1(\mu_{\sigma,+})$ for any sequence $\sigma_n \xrightarrow[n \to \infty]{} \sigma$. We thus obtain the continuity.

B.2 On the variance of the stationary distribution(s) of (NL)

Let $\sigma_0 > 0$, and let us show that the functions $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ and $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$ are Lipschitz continuous. This is useful, as can be seen in Section 3.2, in proving that there exists a phase transition for the effective dynamics (Eff).

Throughout this section, the constant C holds no importance and may change from one line to the next.

We start by showing that $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ is Lipschitz continuous.

Lemma B.2. *Let $\sigma_0 > 0$ and $\kappa_1 \in [-C_{\kappa_1}, C_{\kappa_1}]$ (where C_{κ_1} is given in (3.9)). The function $\sigma \mapsto f_2(\sigma, \kappa_1)$ is Lipschitz continuous on $[\sigma_0, \infty[$ uniformly in $\kappa_1 \in [-C_{\kappa_1}, C_{\kappa_1}]$.*

Applying this lemma for $\kappa_1 = 0$ yields the desired Lipschitz continuity for $\kappa_2(\mu_{\sigma,0})$.

Proof of Lemma B.2. Recall g defined in (3.5). and consider $C = \frac{1+2L_W\kappa_1^2}{4}$, a constant such that, for U given by (1.8), we ensure $U(x) + \frac{L_W}{2}x^2 - L_W x\kappa_1 + C \geq 0$. We have

$$f_2(\sigma, \kappa_1) = \frac{\int_{\mathbb{R}}(x-\kappa_1)^2 g(x,\sigma,\kappa_1)e^{-\frac{C}{\sigma}}dx}{\int_{\mathbb{R}} g(x,\sigma,\kappa_1)e^{-\frac{C}{\sigma}}dx},$$

and thus

$$\begin{aligned}
|\partial_\sigma f_2(\sigma,\kappa_1)| =& \frac{1}{\sigma^2}\left|\frac{\int_{\mathbb{R}}(x-\kappa_1)^2\left(U(x)+\frac{L_W}{2}x^2 - L_W x\kappa_1 + C\right)g(x,\sigma,\kappa_1)dx \int_{\mathbb{R}} g(x,\sigma,\kappa_1)dx}{\left(\int_{\mathbb{R}} g(x,\sigma,\kappa_1)dx\right)^2}\right. \\
&\left. - \frac{\int_{\mathbb{R}}(x-\kappa_1)^2 g(x,\sigma,\kappa_1)dx \int_{\mathbb{R}}\left(U(x)+\frac{L_W}{2}x^2 - L_W x\kappa_1 + C\right)g(x,\sigma,\kappa_1)dx}{\left(\int_{\mathbb{R}} g(x,\sigma,\kappa_1)dx\right)^2}\right| \\
\leq& \frac{1}{\sigma^2}\frac{\int_{\mathbb{R}}(x-\kappa_1)^2\left(U(x)+\frac{L_W}{2}x^2 - L_W x\kappa_1 + C\right)g(x,\sigma,\kappa_1)dx}{\int_{\mathbb{R}} g(x,\sigma,\kappa_1)dx} \qquad \text{(B.3)} \\
&+ \frac{1}{\sigma^2}\frac{\int_{\mathbb{R}}(x-\kappa_1)^2 g(x,\sigma,\kappa_1)dx}{\int_{\mathbb{R}} g(x,\sigma,\kappa_1)dx}\frac{\int_{\mathbb{R}}\left(U(x)+\frac{L_W}{2}x^2 - L_W x\kappa_1 + C\right)g(x,\sigma,\kappa_1)dx}{\int_{\mathbb{R}} g(x,\sigma,\kappa_1)dx}. \qquad \text{(B.4)}
\end{aligned}$$

First

$$\left(\int_{\mathbb{R}} g(x,\sigma,\kappa_1)dx\right)^{-1} \leq \left(\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma_0}\left(U(x)+\frac{L_W}{2}x^2 - L_W x\kappa_1 + C\right)\right)dx\right)^{-1}.$$

Then, for all $x \in \mathbb{R}$ and all $\alpha \geq 0$, we have

$$U(x) + \frac{L_W}{2}x^2 - L_W x\kappa_1 + C = \frac{x^4}{4} - \frac{x^2}{2} + \frac{1}{4} + \frac{L_W}{2}|x-\kappa_1|^2 \geq \alpha x^2 - \beta_\alpha,$$

with

$$\beta_\alpha = \frac{(2\alpha+1)^2}{4} - \frac{1}{4} = \alpha^2 + \alpha.$$

Thus, for all integers $k \geq 0$, we have

$$\int_{\mathbb{R}} x^{2k}\exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}x^2 - L_W x\kappa_1 + C\right)\right)dx \leq e^{\frac{\alpha^2+\alpha}{\sigma}}\int_{\mathbb{R}} x^{2k}\exp\left(-\frac{\alpha x^2}{\sigma}\right)dx$$

$$= e^{\frac{\alpha^2+\alpha}{\sigma}} \sqrt{2\pi} \frac{(2k)!}{2^k k!} \left(\frac{\sigma}{2\alpha}\right)^{k+\frac{1}{2}}.$$

Choosing $\alpha = \frac{\sqrt{\sigma}}{2}$, we obtain

$$\int_{\mathbb{R}} x^{2k} \exp\left(-\frac{1}{\sigma}\left(U(x) + \frac{L_W}{2}x^2 - L_W x \kappa_1 + C\right)\right) dx \leq e^{\frac{1}{4}+\frac{1}{2\sqrt{\sigma}}} \sqrt{2\pi} \frac{(2k)!}{2^k k!} \sigma^{\frac{k}{2}+\frac{1}{4}}.$$

Hence there exists C, independent of σ, such that for $k \leq \frac{7}{2}$, $\frac{1}{\sigma^2}\int_{\mathbb{R}} x^{2k} g(\sigma, x) dx \leq C$ (which allows us to deal with (B.3)) and for $k \leq 2$, $\frac{1}{\sigma^{5/4}}\int_{\mathbb{R}} x^{2k} g(\sigma, x) dx \leq C$ and for $k \leq 1$, $\frac{1}{\sigma^{3/4}}\int_{\mathbb{R}} x^{2k} g(\sigma, x) dx \leq C$ (both allow us to deal with (B.4)). Thus, for $\sigma \geq \sigma_0$, $|\partial_\sigma f_2(\sigma, \kappa_1)|$ is bounded uniformly in κ_1, which yields the result. □

We now show that $\sigma \mapsto \kappa_2(\mu_{\sigma,+})$ is Lipschitz continuous. Let $\sigma_c > \sigma_0 > 0$. We have already proved, in Lemma B.2, that $\sigma \mapsto f_2(\sigma, \kappa_1)$ is Lipschitz continuous uniformly in $\kappa_1 \in [-C_{\kappa_1}, C_{\kappa_1}]$. However, the difficulty lies in the fact that, for $\kappa_2(\mu_{\sigma,+})$ given by $\kappa_2(\mu_{\sigma,+}) = f_2(\sigma, \kappa_1(\mu_{\sigma,+}))$, the mean $\sigma \mapsto \kappa_1(\mu_{\sigma,+})$ is not Lipschitz continuous around σ_c. We will work our way around this fact (and, doing so, also prove it) in the rest of the subsection, but in the meantime this can be numerically observed in Figure 6.

Let us compute the various derivatives of f_1 and f_2 given in (3.6) and (3.7).

$$\begin{aligned}
\partial_\sigma g(x, \sigma, \kappa) =& \frac{\left(U(x) + \frac{L_W}{2}|x-\kappa|^2\right)}{\sigma^2} g(x, \sigma, \kappa), \\
\partial_\kappa g(x, \sigma, \kappa) =& \frac{L_W}{\sigma}(x - \kappa) g(x, \sigma, \kappa), \\
\partial_\sigma f_1(\sigma, \kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\int_{\mathbb{R}} g \int_{\mathbb{R}} x \partial_\sigma g - \int_{\mathbb{R}} \partial_\sigma g \int_{\mathbb{R}} x g\right) \\
=& \frac{1}{\sigma^2}\left(\frac{\int_{\mathbb{R}} x\left(U(x) + \frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}}\left(U(x) + \frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} x g}{\int_{\mathbb{R}} g}\right), \\
\partial_\kappa f_1(\sigma, \kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\int_{\mathbb{R}} g \int_{\mathbb{R}} x \partial_\kappa g - \int_{\mathbb{R}} \partial_\kappa g \int_{\mathbb{R}} x g\right) \\
=& \frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}} x(x-\kappa) g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}}(x-\kappa) g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} x g}{\int_{\mathbb{R}} g}\right) \\
=& \frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g} - \left(\frac{\int_{\mathbb{R}} x g}{\int_{\mathbb{R}} g}\right)^2\right), \\
\partial_\sigma f_2(\sigma, \kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\int_{\mathbb{R}} g \int_{\mathbb{R}} (x-\kappa)^2 \partial_\sigma g - \int_{\mathbb{R}} \partial_\sigma g \int_{\mathbb{R}} (x-\kappa)^2 g\right) \\
=& \frac{1}{\sigma^2}\left(\frac{\int_{\mathbb{R}} (x-\kappa)^2\left(U(x) + \frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}}\left(U(x) + \frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} (x-\kappa)^2 g}{\int_{\mathbb{R}} g}\right), \\
\partial_\kappa f_2(\sigma, \kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\left(2\int_{\mathbb{R}} (\kappa - x) g + \int_{\mathbb{R}} (x-\kappa)^2 \partial_\kappa g\right)\int_{\mathbb{R}} g - \int_{\mathbb{R}} \partial_\kappa g \int_{\mathbb{R}} (x-\kappa)^2 g\right) \\
=& \frac{2\int_{\mathbb{R}} (\kappa - x) g}{\int_{\mathbb{R}} g} + \frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}} (x-\kappa)^3 g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}} (x-\kappa)^2 g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} (x-\kappa) g}{\int_{\mathbb{R}} g}\right).
\end{aligned}$$

In particular, notice that

$$\partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,*})) = \frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*}) \quad \text{and} \quad \partial_\kappa f_2(\sigma, \kappa_1(\mu_{\sigma,*})) = \frac{L_W}{\sigma}\frac{\int_{\mathbb{R}}(x-\kappa_1(\mu_{\sigma,*}))^3 g}{\int_{\mathbb{R}} g}. \tag{B.5}$$

We have

$$\begin{aligned}\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}) =& \partial_\sigma f_1(\sigma, \kappa_1(\mu_{\sigma,*})) + \partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,*}))\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}),\\ \frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*}) =& \partial_\sigma f_2(\sigma, \kappa_1(\mu_{\sigma,*})) + \partial_\kappa f_2(\sigma, \kappa_1(\mu_{\sigma,*}))\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}).\end{aligned}$$

Thus

$$\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}) = \frac{\partial_\sigma f_1(\sigma, \kappa_1(\mu_{\sigma,*}))}{1-\partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,*}))} = \frac{\partial_\sigma f_1(\sigma, \kappa_1(\mu_{\sigma,*}))}{1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})}. \tag{B.6}$$

By the results on the critical variance in Lemma 3.2, $\kappa_1(\mu_{\sigma,*})$ is continuously differentiable on $]\sigma_0, \sigma_c[$. Likewise

$$\begin{aligned}&\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})\left(1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})\right)\\ &\qquad = \left(1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})\right)\partial_\sigma f_2(\sigma, \kappa_1(\mu_{\sigma,*})) + \partial_\kappa f_2(\sigma, \kappa_1(\mu_{\sigma,*}))\partial_\sigma f_1(\sigma, \kappa_1(\mu_{\sigma,*})).\end{aligned} \tag{B.7}$$

The fact that $1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})$ goes to 0 as $\sigma \to \sigma_c$ is what prevents us from giving an upper bound on $\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*})$. By the result on the critical variance in Lemma 3.2 and by continuity, there may be a problem in the limit $\sigma \to \sigma_c^-$, but we wan already say that $\sigma \mapsto \kappa_2(\mu_{\sigma,+})$ is Lipschitz continuous on any interval of the form $[\sigma_1, \sigma_2]$ with $0 < \sigma_1 < \sigma_2 < \sigma_c$. The following lemma gives a more precise speed of convergence of the mean to 0 around the critical parameter.

Lemma B.3. *There exists $C > 0$ such that*

$$\frac{\kappa_1(\mu_{\sigma,+})}{\sqrt{\sigma_c - \sigma}} \xrightarrow[\sigma\to\sigma_c^-]{} C.$$

Proof. We restrict the study to $\sigma \in [\sigma_0, \sigma_c[$ for some arbitrary $\sigma_0 > 0$. We compute

$$\begin{aligned}\partial_\kappa f_1(\sigma,\kappa) =& \frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g} - \left(\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g}\right)^2\right),\\ \partial_\kappa^2 f_1(\sigma,\kappa) =& \left(\frac{L_W}{\sigma}\right)^2\left(\frac{\int_{\mathbb{R}} x^3 g}{\int_{\mathbb{R}} g} - 3\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g} + 2\left(\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g}\right)^3\right),\\ \partial_\kappa^3 f_1(\sigma,\kappa) =& \left(\frac{L_W}{\sigma}\right)^3\left(\frac{\int_{\mathbb{R}} x^4 g}{\int_{\mathbb{R}} g} - 4\frac{\int_{\mathbb{R}} x^3 g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g} + 12\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g}\left(\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g}\right)^2 - 6\left(\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g}\right)^4 - 3\left(\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g}\right)^2\right).\end{aligned}$$

In particular

$$\begin{aligned}f_1(\sigma, 0) =& 0,\\ \partial_\kappa f_1(\sigma, 0) =& \frac{L_W}{\sigma}\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g} > 0,\\ \partial_\kappa^2 f_1(\sigma, 0) =& 0,\end{aligned}$$

and, by Lemma A.2

$$\partial_\kappa^3 f_1(\sigma,0) = \left(\frac{L_W}{\sigma}\right)^3 \left(\frac{\int_{\mathbb{R}} x^4 g}{\int_{\mathbb{R}} g} - 3\left(\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g}\right)^2\right) \xrightarrow{\sigma\to\sigma_c^-} \partial_\kappa^3 f_1(\sigma_c,0) < 0. \tag{B.8}$$

Let us start by proving that there exists $C>0$ such that, in the limit $\sigma\to\sigma_c$ (or equivalently $\kappa_1(\mu_{\sigma,+})\to 0$), we have

$$\frac{\sigma_c-\sigma}{\kappa_1(\mu_{\sigma,+})^2} < C + o(1). \tag{B.9}$$

We compute

$$\partial_\kappa f_1(\sigma,\kappa) = \partial_\kappa f_1(\sigma,0) + \kappa\partial_\kappa^2 f_1(\sigma,0) + \frac{\kappa^2}{2}\partial_\kappa^3 f_1(\sigma,0) + o(\kappa^2).$$

By Lemma A.3, there exists $C>0$ such that

$$\partial_\kappa f_1(\sigma,0) \geq \partial_\kappa f_1(\sigma_c,0) + C(\sigma_c-\sigma) = 1 + C(\sigma_c-\sigma).$$

Since $\partial_\kappa f_1(\sigma_c,0) = \partial_\kappa f_1(\sigma_c,\kappa_1(\mu_{\sigma_c})) = \frac{L_W}{\sigma_c}\kappa_2(\mu_{\sigma_c}) = 1$ by (3.10)

$$\partial_\kappa f_1(\sigma,\kappa) \geq 1 + C(\sigma_c-\sigma) + \frac{\kappa^2}{2}\left(\partial_\kappa^3 f_1(\sigma_c,0) + o_{\sigma\to\sigma_c}(1)\right) + o(\kappa^2),$$

where the $o(\kappa^2)$ is uniform in $\sigma\in[\sigma_0,\sigma_c]$. Considering $\kappa=\kappa_1(\mu_{\sigma,+})$, which goes to 0 as $\sigma\to\sigma_c$, in the equation above yields

$$\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,+}) \geq 1 + C(\sigma_c-\sigma) + \frac{\kappa_1(\mu_{\sigma,+})^2}{2}\left(\partial_\kappa^3 f_1(\sigma_c,0) + o_{\sigma\to\sigma_c}(1)\right) + o(\kappa_1(\mu_{\sigma,+})^2).$$

By the results of Lemma 3.2 concerning the critical variance, $\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,+}) \leq 1$, which gives

$$0 \geq \frac{C(\sigma_c-\sigma)}{\kappa_1(\mu_{\sigma,+})^2} + \frac{\partial_\kappa^3 f_1(\sigma_c,0)}{2} + o_{\sigma\to\sigma_c}(1).$$

This gives (B.9) and this in turns allows us to state that $\sigma_c-\sigma = O(\kappa_1(\mu_{\sigma,+})^2)$.

We then have

$$\begin{aligned} f_1(\sigma,\kappa) =& f_1(\sigma,0) + \kappa\partial_\kappa f_1(\sigma,0) + \frac{\kappa^2}{2}\partial_\kappa^2 f_1(\sigma,0) + \frac{\kappa^3}{6}\partial_\kappa^3 f_1(\sigma,0) + o(\kappa^3) \\ =& \kappa\partial_\kappa f_1(\sigma,0) + \frac{\kappa^3}{6}\partial_\kappa^3 f_1(\sigma,0) + o(\kappa^3). \end{aligned}$$

Furthermore, defining F_1 as in Lemma A.3, we obtain

$$\begin{aligned} \partial_\kappa f_1(\sigma,0) = F_1(\sigma) =& F_1(\sigma_c) - (\sigma_c-\sigma)F_1'(\sigma_c) + o(\sigma_c-\sigma) \\ =& 1 - (\sigma_c-\sigma)F_1'(\sigma_c) + o(\sigma_c-\sigma), \end{aligned}$$

which then yields

$$f_1(\sigma,\kappa) = \kappa\left(1 - (\sigma_c-\sigma)F_1'(\sigma_c) + o(\sigma_c-\sigma)\right) + \frac{\kappa^3}{6}\left(\partial_\kappa^3 f_1(\sigma_c,0) + o(1)\right) + o(\kappa^3).$$

Since $\kappa_1(\mu_{\sigma,+}) = f_1(\sigma, \kappa_1(\mu_{\sigma,+}))$, we thus get in the limit $\sigma \to \sigma_c^-$

$$\begin{aligned} 0 =& -\kappa_1(\mu_{\sigma,+})(\sigma_c - \sigma)F_1'(\sigma_c) + \frac{\kappa_1(\mu_{\sigma,*})^3}{6}\partial_\kappa^3 f_1(\sigma_c, 0) \\ &+ o(\kappa_1(\mu_{\sigma,+})^3) + \kappa_1(\mu_{\sigma,+})o(\sigma_c - \sigma) \end{aligned} \tag{B.10}$$

$$0 = -(\sigma_c - \sigma)F_1'(\sigma_c) + \frac{\kappa_1(\mu_{\sigma,+})^2}{6}\partial_\kappa^3 f_1(\sigma_c, 0) + o(\kappa_1(\mu_{\sigma,+})^2). \tag{B.11}$$

Thus, thanks to (B.8) and Lemma A.3, there exists $C > 0$ such that

$$\frac{\sigma_c - \sigma}{\kappa_1(\mu_{\sigma,+})^2} \xrightarrow{\sigma \to \sigma_c^-} C,$$

which yields the final result. □

Lemma B.4. *Let $\sigma_0 \in]0, \sigma_c[$. Then $\sigma \mapsto \kappa_2(\mu_{\sigma,+})$ is Lispchitz continuous on $[\sigma_0, \sigma_c]$.*

Proof. Let us write for all $(\sigma, \kappa) \in [\sigma_0, \sigma_c] \times [-C_{\kappa_1}, C_{\kappa_1}]$

$$\begin{aligned} \partial_\kappa f_1(\sigma, \kappa) =& \partial_\kappa f_1(\sigma, 0) + \kappa\partial_\kappa^2 f_1(\sigma, 0) + \frac{\kappa^2}{2}\partial_\kappa^3 f_1(\sigma, 0) + o(\kappa^3) \\ =& \big(F_1(\sigma_c) - (\sigma_c - \sigma)F_1'(\sigma_c) + o(\sigma_c - \sigma)\big) + \frac{\kappa^2}{2}\big(\partial_\kappa^3 f_1(\sigma_c, 0) + O(\sigma_c - \sigma)\big) + o(\kappa^3), \end{aligned}$$

where we used the notation F_1 from Lemma A.3, the fact that $\partial_\kappa^2 f_1(\sigma, 0) = 0$, and where all notation $o(\cdot)$ and $O(\cdot)$ are uniform in $(\sigma, \kappa) \in [\sigma_0, \sigma_c] \times [-C_{\kappa_1}, C_{\kappa_1}]$ by continuity. Since $F_1(\sigma_c) = 1$, we obtain

$$1 - \partial_\kappa f_1(\sigma, \kappa) = (\sigma_c - \sigma)F_1'(\sigma_c) - \frac{\kappa^2}{2}\partial_\kappa^3 f_1(\sigma_c, 0) + o(\sigma_c - \sigma) + o(\kappa^3) + \kappa^2 O(\sigma_c - \sigma).$$

Applying this for $\kappa = \kappa_1(\mu_{\sigma,+})$, and using (B.11) and Lemma B.3, we obtain

$$\begin{aligned} 1 - \partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,+})) =& \frac{\kappa_1(\mu_{\sigma,+})^2}{6}\partial_\kappa^3 f_1(\sigma_c, 0) - \frac{\kappa_1(\mu_{\sigma,+})^2}{2}\partial_\kappa^3 f_1(\sigma_c, 0) + o(\kappa_1(\mu_{\sigma,+})^2) \\ =& -\frac{\kappa_1(\mu_{\sigma,+})^2}{3}\partial_\kappa^3 f_1(\sigma_c, 0) + o(\kappa_1(\mu_{\sigma,+})^2). \end{aligned}$$

Hence, using (B.6), for all $\sigma \in [\sigma_0, \sigma_c[$

$$\left|\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,+})\right| = \frac{|\partial_\sigma f_1(\sigma, \kappa_1(\mu_{\sigma,+}))|}{|1 - \partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,+}))|} \lesssim \frac{C\kappa_1(\mu_{\sigma,+})}{\kappa_1(\mu_{\sigma,+})^2 + o(\kappa_1(\mu_{\sigma,+})^2)} \lesssim \frac{1}{\kappa_1(\mu_{\sigma,+})} + o(\kappa_1(\mu_{\sigma,+})),$$

where we used that $\partial_\sigma f_1(\sigma, 0) = 0$ and that $\partial_\kappa \partial_\sigma f_1$ is bounded over $[\sigma_0, \sigma_c] \times [-C_{\kappa_1}, C_{\kappa_1}]$ to bound the numerator, and Lemma A.2 for the denominator. Besides, since $\partial_\kappa f_2(\sigma, 0) = 0$ (by (B.5) and symmetry),

$$|\partial_\kappa f_2(\sigma, \kappa_1(\mu_{\sigma,+})| = |\partial_\kappa f_2(\sigma, \kappa_1(\mu_{\sigma,+})) - \partial_\kappa f_2(\sigma, 0)| \lesssim |\kappa_1(\mu_{\sigma,+})| + o(\kappa_1(\mu_{\sigma,+}))$$

and thus we bound for all $\sigma \in [\sigma_0, \sigma_c[$

$$\begin{aligned} \left|\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,+})\right| \leq& |\partial_\sigma f_2(\sigma, \kappa_1(\mu_{\sigma,+}))| + \left|\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,+})\right| |\partial_\kappa f_2(\sigma, \kappa_1(\mu_{\sigma,+}))| + o(\kappa_1(\mu_{\sigma,+})) \\ \lesssim& 1 + \frac{|\partial_\kappa f_2(\sigma, \kappa_1(\mu_{\sigma,+})|}{\kappa_1(\mu_{\sigma,+})} + o(1) \\ \lesssim& 1 + o(1). \end{aligned}$$

By continuity (recall that $1 - \partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,+})) = 1 - \frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,+}) > 0$ for $\sigma < \sigma_c$) , this proves that $\sigma \mapsto \kappa_2(\mu_{\sigma,+})$ is Lipschitz on $[\sigma_0, \sigma_c]$. □

Acknowledgements

This work has been (partially) supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency. P. Monmarché's research is supported by the Project SWIDIMS ANR-20-CE40-0022 of the French National Research Agency.

References

[BH19] Jianhai Bao and Xing Huang. Approximations of Mckean-Vlasov SDEs with Irregular Coefficients. *arXiv e-prints*, page arXiv:1905.08522, May 2019.

[BMP21] Anton Bovier, Saeda Marello, and Elena Pulvirenti. Metastability for the dilute Curie-Weiss model with Glauber dynamics. *Electron. J. Probab.*, 26:Paper No. 47, 38, 2021.

[BT97] Mireille Bossy and Denis Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.*, 66(217):157–192, 1997.

[CD22a] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models*, 15(6):895–1015, 2022.

[CD22b] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. II. Applications. *Kinet. Relat. Models*, 15(6):1017–1173, 2022.

[CdS23] Xingyuan Chen, Goncalo dos Reis, and Wolfgang Stockinger. Wellposedness, exponential ergodicity and numerical approximation of fully super-linear McKean–Vlasov SDEs and associated particle systems. *arXiv e-prints*, page arXiv:2302.05133, February 2023.

[CG91] Francis Comets and Basilis Gidas. Asymptotics of maximum likelihood estimators for the Curie-Weiss model. *Ann. Statist.*, 19(2):557–578, 1991.

[CGM08] Patrick Cattiaux, Arnaud Guillin, and Florent Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2):19–40, 2008.

[CJT22] José Antonio Carrillo, Shi Jin, and Yijia Tang. Random batch particle methods for the homogeneous Landau equation. *Commun. Comput. Phys.*, 31(4):997–1019, 2022.

[CK17] Francesca Collet and Richard C. Kraaij. Dynamical moderate deviations for the Curie-Weiss model. *Stochastic Process. Appl.*, 127(9):2900–2925, 2017.

[DQ21] Xiaojie Ding and Huijie Qiao. Euler-Maruyama approximations for stochastic McKean-Vlasov equations with non-Lipschitz coefficients. *J. Theoret. Probab.*, 34(3):1408–1425, 2021.

[Ell85] Richard S. Ellis. *Entropy, large deviations, and statistical mechanics*, volume 271 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.

[EN78] Richard S. Ellis and Charles M. Newman. The statistics of Curie-Weiss models. *J. Statist. Phys.*, 19(2):149–161, 1978.

[FV17] Sacha Friedli and Yvan Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, 2017.

[GJP21] François Golse, Shi Jin, and Thierry Paul. The random batch method for N-body quantum dynamics. *J. Comput. Math.*, 39(6):897–922, 2021.

[GLBMM22] Nicolaï Gouraud, Pierre Le Bris, Adrien Majka, and Pierre Monmarché. HMC and underdamped langevin united in the unadjusted convex smooth case. *arXiv preprint arXiv:2202.00977*, 2022.

[GM21] Arnaud Guillin and Pierre Monmarché. Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes. *Journal of Statistical Physics*, 185:1–20, 2021.

[JL22] Shi Jin and Lei Li. On the mean field limit of the random batch method for interacting particle systems. *Sci. China Math.*, 65(1):169–202, 2022.

[JLL20] Shi Jin, Lei Li, and Jian-Guo Liu. Random batch methods (RBM) for interacting particle systems. *J. Comput. Phys.*, 400:108877, 30, 2020.

[JLL21] Shi Jin, Lei Li, and Jian-Guo Liu. Convergence of the random batch method for interacting particles with disparate species and weights. *SIAM J. Numer. Anal.*, 59(2):746–768, 2021.

[JLXZ21] Shi Jin, Lei Li, Zhenli Xu, and Yue Zhao. A random batch Ewald method for particle systems with Coulomb interactions. *SIAM J. Sci. Comput.*, 43(4):B937–B960, 2021.

[JLYZ22] Shi Jin, Lei Li, Xuda Ye, and Zhennan Zhou. Ergodicity and long-time behavior of the Random Batch Method for interacting particle systems. *arXiv e-prints*, page arXiv:2202.04952, February 2022.

[Kal97] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, 1997.

[LLP10] David A. Levin, Malwina J. Luczak, and Yuval Peres. Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probab. Theory Related Fields*, 146(1-2):223–265, 2010.

[LLT22] Lei Li, Jian-Guo Liu, and Yijia Tang. Some random batch particle methods for the Poisson-Nernst-Planck and Poisson-Boltzmann equations. *Commun. Comput. Phys.*, 32(1):41–82, 2022.

[LXZ20] Lei Li, Zhenli Xu, and Yue Zhao. A random-batch Monte Carlo method for many-body systems with singular kernels. *SIAM J. Sci. Comput.*, 42(3):A1486–A1509, 2020.

[Mal03] Florent Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2):540–560, 2003.

[Sha06] Tony Shardlow. Modified equations for stochastic differential equations. *BIT*, 46(1):111–125, 2006.

[SS21] Inass Sekkat and Gabriel Stoltz. Removing the mini-batching error in Bayesian inference using Adaptive Langevin dynamics. *arXiv e-prints*, page arXiv:2105.10347, May 2021.

[Tug14] Julian Tugaut. Phase transitions of McKean-Vlasov processes in double-wells landscape. *Stochastics*, 86(2):257–284, 2014.

[Vil09] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[VZT16] Sebastian J. Vollmer, Konstantinos C. Zygalakis, and Yee Whye Teh. Exploration of the (non-)asymptotic bias and variance of stochastic gradient Langevin dynamics. *J. Mach. Learn. Res.*, 17:Paper No. 159, 45, 2016.

[WT11] Max Welling and Yee Whye Teh. Bayesian learning via stochastic gradient langevin dynamics. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, page 681–688, Madison, WI, USA, 2011. Omnipress.

[YZ22] Xuda Ye and Zhennan Zhou. Error Analysis of Time-Discrete Random Batch Method for Interacting Particle Systems and Associated Mean-Field Limits. *arXiv e-prints*, page arXiv:2206.02166, June 2022.

[Zyg11] Konstantinos C. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM J. Sci. Comput.*, 33(1):102–130, 2011.