



**HAL**  
open science

# Stochastic dynamics and the Polchinski equation: an introduction

Roland Bauerschmidt, Thierry Bodineau, Benoît Dagallier

► **To cite this version:**

Roland Bauerschmidt, Thierry Bodineau, Benoît Dagallier. Stochastic dynamics and the Polchinski equation: an introduction. 2023. hal-04163695

**HAL Id: hal-04163695**

**<https://hal.science/hal-04163695>**

Preprint submitted on 17 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Stochastic dynamics and the Polchinski equation: an introduction

Roland Bauerschmidt\*    Thierry Bodineau†    Benoit Dagallier‡

July 13, 2023

## Abstract

This introduction surveys a renormalisation group perspective on log-Sobolev inequalities and related properties of stochastic dynamics. We also explain the relationship of this approach to related recent and less recent developments such as Eldan’s stochastic localisation and the Föllmer process, the Boué–Dupuis variational formula and the Barashkov–Gubinelli approach, the transportation of measure perspective, and the classical analogues of these ideas for Hamilton–Jacobi equations which arise in mean-field limits.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Background on stochastic dynamics</b>	<b>3</b>
2.1	Motivation: Spin models and their stochastic dynamics . . . . .	3
2.2	Generalities on Glauber–Langevin dynamics . . . . .	4
2.3	Log-Sobolev inequality . . . . .	6
2.4	Bakry–Émery theorem . . . . .	8
2.5	Decomposition and properties of the entropy . . . . .	9
2.6	Difficulties arising from statistical physics perspective . . . . .	12
2.7	Difficulties arising from continuum perspective . . . . .	12
<b>3</b>	<b>Gaussian integration and the Polchinski equation</b>	<b>14</b>
3.1	Gaussian integration . . . . .	14
3.2	Renormalised potential and Polchinski equation . . . . .	16
3.3	Log-Sobolev inequality via a multiscale Bakry–Émery method . . . . .	19
3.4	Derivatives of the renormalised potential . . . . .	22
3.5	Example: Convexification along the Polchinski flow for one variable . . . . .	24
3.6	Aside: Geometric perspective on the Polchinski flow . . . . .	26
3.7	Aside: Entropic stability estimate . . . . .	28
<b>4</b>	<b>Pathwise Polchinski flow and stochastic localisation perspective</b>	<b>32</b>
4.1	Pathwise realisation of the Polchinski semigroup . . . . .	32
4.2	Example: Log-Sobolev inequality by coupling . . . . .	34
4.3	Example: Coupling with the Gaussian reference measure . . . . .	34
4.4	Effective potential and martingales . . . . .	35
4.5	Stochastic localisation perspective . . . . .	35

---

\*University of Cambridge, DPMMS. E-mail: [rb812@cam.ac.uk](mailto:rb812@cam.ac.uk).

†IHES, CNRS. E-mail: [bodineau@ihes.fr](mailto:bodineau@ihes.fr).

‡University of Cambridge, DPMMS. E-mail: [bd444@cam.ac.uk](mailto:bd444@cam.ac.uk).

<b>5</b>	<b>Variational and transport perspectives on the Polchinski flow</b>	<b>37</b>
5.1	Föllmer’s problem . . . . .	37
5.2	Variational representation of the renormalised potential . . . . .	39
5.3	Lipschitz transport . . . . .	40
<b>6</b>	<b>Applications</b>	<b>43</b>
6.1	Applications to Euclidean field theory . . . . .	44
6.2	Applications to Ising models . . . . .	47
6.3	Applications to conservative dynamics . . . . .	54
<b>A</b>	<b>Classical renormalised potential and Hamilton–Jacobi equation</b>	<b>55</b>
A.1	Hamilton–Jacobi equation . . . . .	55
A.2	Example: Mean-field Ising model . . . . .	58

## 1 Introduction

Functional inequalities have been thoroughly studied in different contexts [7,62] and one important motivation is to quantify the relaxation of stochastic dynamics by using Poincaré and (possibly modified) log-Sobolev inequalities [22, 37, 51, 68, 86]. Statistical mechanics offers an interesting setting to apply these inequalities and to analyse the information they provide in various physical regimes. Indeed, one would like to describe the relaxation to equilibrium of lattice gas and spin dynamics, which are modelled by stochastic evolutions on high-dimensional state spaces. Their continuum limits, often described by (singular) SPDEs, are also of a lot of interest.

The structure of the equilibrium Gibbs measures is sensitive to the occurrence of phase transitions and the dynamical behaviour will also be strongly influenced by phase transitions. In the uniqueness regime, namely in absence of a phase transition (typically at high temperatures), one expects that the dynamics relax exponentially fast uniformly in the dimension of the state space with a speed of relaxation diverging when the temperature approaches the critical point. We refer to Sections 2.6 and 2.7 for more details. In a phase transition regime (typically corresponding to low temperatures), one expects different types of behaviours depending strongly on the type of boundary conditions and we will not discuss the corresponding phenomena in these notes.

The fast relaxation towards equilibrium in the uniqueness regime (or at least deep in it) is well understood and we refer to [51, 68, 69] for very complete accounts of the corresponding theory. Roughly speaking, it has been shown for a wide range of models that good mixing properties of the equilibrium measure are equivalent to fast relaxation of the dynamics, namely uniform bounds (with respect to the domains and the boundary conditions) on the Poincaré or the log-Sobolev constants. For the Ising model [35, 38, 68], the validity of the mixing properties have been proved in the whole uniqueness regime leading to strong relaxation statements on the dynamics, and more detailed dynamical features are also understood in that regime [66, 67]. For more general systems and in particular continuous spin systems, the picture is much less complete.

The main goal of this survey is to present a different perspective on the derivation of functional inequalities based on the renormalisation group theory, introduced in physics by Wilson [89], and with its continuous formulation emphasised in particular by Polchinski [78]. The renormalisation group was introduced to study the critical behaviour and the existence of continuum limits of equilibrium models of statistical physics and quantum field theory from a unified perspective. The renormalisation group formalism associates with a Gibbs measure a flow of measures defined in terms of a renormalised potential (see Section 3.2). We show how this structure can be used to prove log-Sobolev inequalities under a condition on the renormalised potential which is a multiscale generalisation of the Bakry–Émery criterion (see Section 3.3). The renormalised potential obeys a second-order Hamilton–Jacobi type equation (the Polchinski equation) with characteristics given by a stochastic evolution (see Section 4.1) which coincide with the stochastic

process of Eldan’s *stochastic localisation* method introduced for very different purposes [40]. Section 4.5 provides a dictionary to relate both points of view.

An alternative to the multiscale Bakry–Émery method to derive log-Sobolev inequalities (with much similarity and both advantages and disadvantages) is the entropic stability estimate recently established in [28] and reviewed in Section 3.7. This estimate applies to the same Polchinski flow, or its equivalent interpretation as stochastic localisation. It originated in the spectral and entropic independence estimates [3, 4] which are similar estimates for a different flow that takes the role of the Polchinski flow in another kind of model. This analogy has already been highlighted in [28] to which we refer for a discussion of this relation. Compared to the established approaches to functional inequalities for statistical mechanical models, which typically rely on spatial decompositions, all of the approaches discussed here are more spectral in nature. Spectral quantities are more global and therefore allow to capture for example the near-critical behaviour better. This is illustrated in a series of applications reviewed in Section 6.

The Polchinski renormalisation and the stochastic localisation can be seen as two sides of the same coin, sharing thus very similar structures. In fact, this type of stochastic equations have been considered much earlier by Föllmer [36] as an optimal way to generate a target measure. In Section 5, the Polchinski renormalisation flow is shown to coincide with the optimal stochastic process associated with a suitable *varying* metric. In applications to statistical mechanics models, this metric captures the notion of *scale* so that the Polchinski flow (a continuous renormalisation group flow) provides a canonical way of decomposing the entropy according to scale. Finally, in Section 5.2, the renormalised potential is rewritten as a variational principle using the Polchinski flow, known as the Boué-Dupuis or Borell formula in the generic context, see [63], whose origin is in stochastic optimal control theory. This correspondance is at the heart of the Barashkov–Gubinelli variational method [8].

In Appendix A, these results are compared with the corresponding control theory for classical Hamilton–Jacobi equations, and a simple comparison with the line of research initiated in [74] is also given.

## 2 Background on stochastic dynamics

**2.1. Motivation: Spin models and their stochastic dynamics.** Our goal is to study dynamical (and also some equilibrium) aspects of continuous and discrete spin models of statistical mechanics such as Euclidean field theories or Ising-type spin models. Throughout this article,  $\Lambda$  will be a general finite set (of vertices), but we have  $\Lambda \subset \mathbb{Z}^d$  large in mind, or  $\Lambda \subset \varepsilon\mathbb{Z}^d$  approximating  $\mathbb{R}^d$  (or a subset of it) when  $\varepsilon \rightarrow 0$  in the case of models defined in the continuum. Sometimes we identify  $\Lambda$  with  $[N] = \{1, \dots, N\}$ . Spin fields are then random functions  $\varphi : \Lambda \rightarrow T$  where, for example,  $T = \mathbb{R}$  in the case of continuous scalar spins or  $T = \{\pm 1\}$  in the case of (discrete) Ising spins. For discrete spins, we often write  $\sigma$  instead of  $\varphi$  for a spin configuration.

*Continuous spins.* In the setting of continuous spins, the equilibrium Gibbs measures have expectation of the form

$$(2.1) \quad \mathbb{E}_\nu[F(\varphi)] \propto \int_{\mathbb{R}^\Lambda} e^{-H(\varphi)} F(\varphi) d\varphi$$

where the symbol  $\propto$  denotes the equality of the measures up to a normalisation factor. We will refer to  $H$  as the action or as the Hamiltonian (depending on the context). The main class of  $H$  that we will focus on are of the following form: for spins  $\varphi = (\varphi_x)_{x \in \Lambda}$  taking values in  $\mathbb{R}$  (or vector spins with values in  $\mathbb{R}^n$ ), an interaction matrix  $A$ , and a local potential  $V$ ,

$$(2.2) \quad H(\varphi) = \frac{1}{2}(\varphi, A\varphi) + V_0(\varphi), \quad V_0(\varphi) = \sum_{x \in \Lambda} V(\varphi_x).$$

Defining the discrete Laplace operator on  $\Lambda \subset \mathbb{Z}^d$  by

$$(2.3) \quad \forall x \in \Lambda, \quad (\Delta^\Lambda f)_x := \sum_{y \in \Lambda: y \sim x} (f_y - f_x),$$

a classical choice of interaction is obtained by setting  $A = -\beta \Delta^\Lambda$  for some (inverse temperature) parameter  $\beta > 0$ . In this case, the Hamiltonian reads

$$(2.4) \quad \forall \varphi \in \mathbb{R}^\Lambda, \quad H(\varphi) = \frac{\beta}{4} \sum_{\substack{x, y \in \Lambda, \\ x \sim y}} (\varphi_x - \varphi_y)^2 + V_0(\varphi),$$

where the nearest neighbour interaction is denoted by  $x \sim y$  and the sum counts each pair  $\{x, y\}$  twice. As an example, a typical choice for the single-spin potential  $V$  is the Ginzburg–Landau–Wilson  $\varphi^4$  potential, in which case one usually sets  $\beta = 1$  (and  $\nu$  has the role of a temperature),

$$(2.5) \quad V(\varphi) = \frac{1}{4}g|\varphi|^4 + \frac{1}{2}\nu|\varphi|^2 \quad \text{with } g > 0 \text{ and } \nu < 0.$$

The following Glauber–Langevin dynamics is reversible for the measure  $\nu$  introduced in (2.1):

$$(2.6) \quad d\varphi_t = -\nabla H(\varphi_t) dt + \sqrt{2}dB_t.$$

For the choices (2.4) and (2.5), this stochastic differential equation (SDE) reads

$$(2.7) \quad d\varphi_t = -A\varphi_t dt - \nabla V(\varphi_t) dt + \sqrt{2}dB_t = \Delta^\Lambda \varphi_t dt - g|\varphi_t|^2 \varphi_t dt - \nu \varphi_t dt + \sqrt{2}dB_t.$$

In this survey, we are interested in the long time behaviour of these dynamics when the number of spins is large. We will consider two cases: either  $\Lambda \rightarrow \mathbb{Z}^d$  for the Glauber dynamics of an Ising-type model with continuous spins; or  $\Lambda \subset \varepsilon \mathbb{Z}^d$  with  $\Lambda \rightarrow \mathbb{R}^d$  or  $\Lambda \rightarrow [0, 1]^d$  in which case a suitably normalised version of  $\varphi$  describes the solution of a singular SPDE in the limit  $\varepsilon \rightarrow 0$ .

*Discrete spins.* In the setting of discrete spins, we focus on the Ising model where  $\sigma \in \{\pm 1\}^\Lambda$  and

$$(2.8) \quad \mathbb{E}_\nu[F(\sigma)] \propto \sum_{\sigma \in \{\pm 1\}^\Lambda} e^{-\frac{\beta}{2}(\sigma, A\sigma)} F(\sigma)$$

for some symmetric coupling matrix  $A$ . Its Glauber dynamics is a continuous- or discrete-time Markov process with local transition rates  $c(\sigma, \sigma^x)$  from a configuration  $\sigma$  to  $\sigma^x$  where  $\sigma^x \in \{\pm 1\}^\Lambda$  denotes the configuration obtained from  $\sigma \in \{\pm 1\}^\Lambda$  by flipping the sign of the spin at  $x$ . The transition rates are assumed to satisfy the detailed-balance condition

$$(2.9) \quad \nu(\sigma)c(\sigma, \sigma^x) = \nu(\sigma^x)c(\sigma^x, \sigma),$$

which implies that the measure (2.8) is invariant. Typical choices are described below in the next section. We will be interested in the large time behaviour of the dynamics when  $\Lambda \rightarrow \mathbb{Z}^d$ .

**2.2. Generalities on Glauber–Langevin dynamics.** We now discuss some standard general properties of the stochastic dynamics such as its (finite-dimensional state space) ergodicity.

*Continuous spins.* The Glauber–Langevin dynamics (2.6) is a Markov process with generator

$$(2.10) \quad \Delta^H = \Delta - (\nabla H, \nabla) = e^{+H}(\nabla, e^{-H}\nabla)$$

where

$$(2.11) \quad \Delta = \sum_{x \in \Lambda} \frac{\partial^2}{\partial \varphi_x^2}, \quad (\nabla H, \nabla) = \sum_{x \in \Lambda} \frac{\partial H}{\partial \varphi_x} \frac{\partial}{\partial \varphi_x}.$$

The state space  $\mathbb{R}^\Lambda$  will often be denoted by  $X$ . The distribution of the spin configuration evolves in time along the stochastic dynamics and we denote by  $m_t$  the distribution at time  $t$  starting from an initial measure  $m_0$ : given  $F_0 : X \rightarrow \mathbb{R}$ ,

$$(2.12) \quad \mathbb{E}_{m_t}[F_0] = \mathbb{E}_{m_0}[F_t] \quad \text{with} \quad F_t(\varphi) = \mathbf{T}_t F(\varphi) := \mathbb{E}_{\varphi_0 = \varphi}[F(\varphi_t)],$$

where  $\mathbf{T}_t$  is the semigroup associated with the generator  $\Delta^H$ . In particular,  $F_t = \mathbf{T}_t F$  solves the Kolmogorov backward equation

$$(2.13) \quad \frac{\partial}{\partial t} F_t = \Delta^H F_t.$$

Starting from the SDE, this can be verified using Itô's formula. The measure  $\nu$ , introduced in (2.1), is reversible with respect to this dynamics, and the following integration by parts formula holds for sufficient smooth  $F$ :

$$(2.14) \quad \mathbb{E}_\nu[F(-\Delta^H G)] = \mathbb{E}_\nu[(\nabla F, \nabla G)].$$

The right-hand side is the Dirichlet form:

$$(2.15) \quad D_\nu(F, G) := \mathbb{E}_\nu[(\nabla F, \nabla G)] \quad \text{and} \quad D_\nu(F) := D_\nu(F, F).$$

In particular, the measure  $\nu$  is invariant, i.e., if  $\varphi_0$  is distributed according to  $\nu$  then  $\varphi_t$  also is:

$$(2.16) \quad \frac{\partial}{\partial t} \mathbb{E}_\nu[F_t] = \mathbb{E}_\nu[\Delta^H F_t] = \mathbb{E}_\nu[(\nabla F_t, \nabla 1)] = 0.$$

Moreover, we will always impose the following ergodicity assumption:

$$(2.17) \quad \forall F_0 \in L^2(\nu), \quad F_t \rightarrow \mathbb{E}_\nu[F_0] \quad \text{in } L^2(\nu).$$

In particular, for any bounded smooth functions  $F_0 : X \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$(2.18) \quad \lim_{t \rightarrow \infty} \mathbb{E}_\nu[g(F_t)] = g(\mathbb{E}_\nu[F_0]).$$

As the next exercise shows, the ergodicity assumption is qualitative if  $\Lambda$  is finite and holds in all examples of interest.

**Exercise 2.1.** Show that  $\frac{1}{2}|\nabla H|^2 - \Delta H \rightarrow \infty$  as  $|\varphi| \rightarrow \infty$  implies that  $-\Delta^H$  has discrete spectrum on  $L^2(\nu)$  with unique minimal eigenvalue 0, and deduce (2.17) and (2.18).

For the discreteness of the spectrum, one may observe that the multiplication operator  $U = e^{\frac{1}{2}H}$  is an isometry from  $L^2(\nu)$  onto  $L^2(\mathbb{R}^N)$  that maps  $-\Delta^H$  to the Schrödinger operator  $-\Delta + W$  on  $\mathbb{R}^N$  with  $W = \frac{1}{4}|\nabla H|^2 - \frac{1}{2}\Delta H$ . The result therefore follows from the spectral theorem and the result that a Schrödinger operator with a potential  $W \in L^1_{\text{loc}}(\mathbb{R}^N)$  that is bounded below and satisfies  $W \rightarrow \infty$  has compact resolvent [79, Theorem XIII.67] (a version of Rellich's theorem).

For further general facts on stochastic dynamics in the continuous setting, we refer to [7, 51, 86]. Even though we will not need it, let us also mention that if the distribution of  $m_t$  is written as  $dm_t = G_t d\nu$  where  $\nu = m_\infty$  is the invariant measure and  $G_t = dm_t/d\nu$  is the density of  $m_t$  relative to it, then

$$(2.19) \quad \frac{\partial}{\partial t} G_t = (\Delta^H)^* G_t = (\Delta^H)^* G_t$$

where  $(\Delta^H)^* = \Delta^H$  is the adjoint of  $\Delta^H$  with respect to  $\nu$ . This can also be expressed as an equation for  $m_t$  (interpreted in a weak sense), which is the Fokker–Planck equation:

$$(2.20) \quad \frac{\partial m_t}{\partial t} = \Delta m_t + (\nabla, m_t \nabla H) = (\nabla, m_t \nabla (\log m_t + H)).$$

*Discrete spins.* A similar structure can also be associated with discrete dynamics. In particular, the Glauber dynamics of an Ising model is determined by its local jump rates  $c(\sigma, \sigma^x)$  satisfying the detailed balance condition as in (2.9). For all  $F : \Omega \rightarrow \mathbb{R}$ , where  $\Omega = \{\pm 1\}^\Lambda$  is the finite state space, the generator and Dirichlet form associated with the Glauber dynamics are

$$(2.21) \quad \Delta_c F(\sigma) = \sum_{x \in \Lambda} c(\sigma, \sigma^x) (F(\sigma^x) - F(\sigma))$$

and

$$(2.22) \quad D_\nu(F) = - \sum_{x \in \Lambda} \sum_{\sigma \in \Omega} F(\sigma) \Delta_c F(\sigma) \nu(\sigma) = \frac{1}{2} \sum_{x \in \Lambda} \sum_{\sigma \in \Omega} c(\sigma, \sigma^x) (F(\sigma^x) - F(\sigma))^2 \nu(\sigma),$$

where we used the detailed balance condition for the second equality. We will again write  $D_\nu(F, F)$  for the quadratic form associated with  $D_\nu(F)$  by polarisation. As in the continuous setting, we will always impose an irreducibility assumption which is equivalent to the analogue of (2.17):

$$(2.23) \quad \forall F_0 : \Omega \rightarrow \mathbb{R}, \quad F_t \rightarrow \mathbb{E}_\nu[F_0],$$

where  $F_t(\sigma) = e^{\Delta_c t} F_0(\sigma) = \mathbb{E}_{\sigma_0 = \sigma} [F(\sigma_t)]$ . Indeed, assuming irreducibility, the convergence (2.23) is a consequence of the Perron–Frobenius theorem, see, e.g., [81].

Many choices of jump rates can be considered, but as long as the jump rates are uniformly bounded from above and below the different Dirichlet forms are equivalent and the large time behaviour of the microscopic dynamics will be similar. Often a natural choice of jump rates is that corresponding to the *standard* Dirichlet form. This choice formally corresponds to  $c(\sigma, \sigma^x) = 1$  in (2.22) which however are not the jump rates of the associated Markov process because the constant function 1 does not satisfy the detailed balance condition. However, rewriting (2.22) as

$$(2.24) \quad D_\nu(F) = \frac{1}{2} \sum_{x \in \Lambda} \sum_{\sigma \in \Omega} \frac{1}{2} \left[ c(\sigma, \sigma^x) + c(\sigma^x, \sigma) \frac{\nu(\sigma^x)}{\nu(\sigma)} \right] (F(\sigma^x) - F(\sigma))^2 \nu(\sigma),$$

we see that the standard Dirichlet form corresponds to the jump rates (satisfying detailed balance)

$$(2.25) \quad c(\sigma, \sigma^x) = \frac{1}{2} \left( 1 + \frac{\nu(\sigma^x)}{\nu(\sigma)} \right).$$

Another popular choice are the heat-bath jump rates which are given by

$$(2.26) \quad c^{\text{HB}}(\sigma, \sigma^x) = \frac{\nu(\sigma^x)}{\nu(\sigma) + \nu(\sigma^x)} = \left( 1 + \frac{\nu(\sigma)}{\nu(\sigma^x)} \right)^{-1}$$

with the corresponding Dirichlet form

$$(2.27) \quad D_\nu^{\text{HB}}(F) = \frac{1}{2} \sum_{x \in \Lambda} \sum_{\sigma \in \Omega} \Psi(\nu(\sigma), \nu(\sigma^x)) (F(\sigma) - F(\sigma^x))^2, \quad \Psi(a, b) = \frac{ab}{a+b}.$$

The Metropolis jump rates correspond to  $\Psi(a, b) = \min\{a, b\}$ . For further general discussion of Glauber dynamics in the discrete case, see [68] and [22, 81].

As in (2.12), the distribution at time  $t$  will be denoted by  $m_t$ .

**2.3. Log-Sobolev inequality.** In the above examples (with  $\Lambda$  finite), one always has the qualitative ergodicity  $m_t \rightarrow \nu = m_\infty$  which amounts to an irreducibility condition. One of the main questions we are interested in is how fast this convergence is. A very good measure for the distance between  $m_t$  and  $\nu = m_\infty$  with many further applications is the relative entropy:

$$(2.28) \quad \mathbb{H}(m_t | \nu) = \mathbb{E}_\nu [F_t \log F_t] = \text{Ent}_\nu(F_t), \quad F_t = \frac{dm_t}{d\nu}.$$

More generally, when  $F$  is nonnegative but not necessarily satisfies  $\mathbb{E}_\nu[F] = 1$ , define

$$(2.29) \quad \text{Ent}_\nu(F) = \mathbb{E}_\nu[\Phi(F)] - \Phi(\mathbb{E}_\nu[F]), \quad \Phi(x) = x \log x.$$

The relative entropy is not symmetric and thus not a metric, but it has many very useful properties making it a good quantity, and it controls the total variation distance by Pinsker's inequality:

$$(2.30) \quad \|m_t - \nu\|_{\text{TV}}^2 \leq 2\mathbb{H}(m_t|\nu).$$

One of the most important properties is that the relative entropy decreases under the dynamics. We begin with the continuous case.

**Proposition 2.2** (de Bruin identity). *Consider the (continuous spin) stochastic dynamics (2.6) with invariant measure  $\nu$  and Dirichlet form  $D_\nu$  defined in (2.15). Then for  $F_t(\varphi) = \mathbb{E}_{\varphi_0=\varphi}[F(\varphi_t)]$  as in (2.12),*

$$(2.31) \quad \frac{\partial}{\partial t} \text{Ent}_\nu(F_t) = -D_\nu(\log F_t, F_t) = -I_\nu(F_t) \leq 0,$$

where the Fisher information is defined in terms of the Dirichlet form (2.15):

$$(2.32) \quad I_\nu(F_t) := \mathbb{E}_\nu \left[ \frac{(\nabla F_t)^2}{F_t} \right] = 4D_\nu(\sqrt{F_t}).$$

*Proof.* Since  $\Phi(\mathbb{E}_\nu[F_t]) = \Phi(\mathbb{E}_\nu[F_0])$  is independent of  $t$  and recalling that  $\Delta^H$  is defined in (2.10),

$$(2.33) \quad \begin{aligned} \frac{\partial}{\partial t} \text{Ent}_\nu(F_t) &= \frac{\partial}{\partial t} \mathbb{E}_\nu[\Phi(F_t)] = \mathbb{E}_\nu[\Phi'(F_t)\dot{F}_t] = \mathbb{E}_\nu[\Phi'(F_t)\Delta^H F_t] \\ &= -D_\nu(\Phi'(F_t), F_t) = -D_\nu(\log F_t + 1, F_t) = -D_\nu(\log F_t, F_t). \end{aligned}$$

To complete the identity (2.31), it is enough to notice that

$$(2.34) \quad D_\nu(\log F_t, F_t) = \mathbb{E}_\nu[(\nabla \log F_t, \nabla F_t)] = \mathbb{E}_\nu \left[ \frac{(\nabla F_t)^2}{F_t} \right] = 4\mathbb{E}_\nu[(\nabla \sqrt{F_t})^2]. \quad \square$$

Using the identity (2.31), the decay of the entropy can be quantified in terms of the *log-Sobolev constant* which will be a key quantity we study.

**Definition 2.3.** *A probability measure  $\nu$  on  $X = \mathbb{R}^N$ , satisfies the log-Sobolev inequality (LSI) with respect to  $D_\nu$  if there is a constant  $\gamma > 0$  such that the following holds for any smooth, compactly supported function  $F : X \rightarrow \mathbb{R}_+$ :*

$$(2.35) \quad \text{Ent}_\nu(F) \leq \frac{2}{\gamma} D_\nu(\sqrt{F}).$$

*The largest choice of  $\gamma$  in this inequality is the log-Sobolev constant (with respect to  $D_\nu$ ). The normalisation with the above factor 2 is convenient (see Proposition 2.5).*

The upshot is that the exponential decay of the relative entropy  $\mathbb{H}(m_t|\nu)$  of the distribution  $m_t$  (defined in (2.12)) along the flow of the Glauber-Langevin dynamics,

$$(2.36) \quad \mathbb{H}(m_t|\nu) \leq e^{-2\gamma t} \mathbb{H}(m_0|\nu),$$

follows, by Gronwall's lemma, from

$$(2.37) \quad \frac{\partial}{\partial t} \mathbb{H}(m_t|\nu) \leq -2\gamma \mathbb{H}(m_t|\nu),$$

which, by the de Bruin identity, is a consequence of the log-Sobolev inequality (2.35). Thus the log-Sobolev constant provides a quantitative estimate on the speed of relaxation of the dynamics towards its stationary measure. This is one of the main motivation for deriving the log-Sobolev inequality.

The log-Sobolev inequality (2.35) has also other consequences. Especially, it is equivalent to the hypercontractivity of the associated Markov semigroup. The hypercontractivity was, in fact, its original motivation [49], see also [45, 76].



**Theorem 2.4** (Hypercontractivity [49]). *The measure  $\nu$  satisfies the log-Sobolev inequality (2.35) with constant  $\gamma$  if and only if the associated semigroup  $\mathbf{T}_t$  is hypercontractive:*

$$(2.38) \quad \|\mathbf{T}_t F\|_{L^{q(t)}(\nu)} \leq \|F\|_{L^p(\nu)} \quad \text{with} \quad \frac{q(t) - 1}{p - 1} = e^{2\gamma t}.$$

We note that the hypercontractivity does not follow in this form from the modified log-Sobolev inequality which will be introduced in (2.41) below for dynamics with discrete state spaces.

More generally, the log-Sobolev inequality is part of a larger class of functional inequalities. In particular, it implies the spectral gap inequality (also called Poincaré inequality).

**Proposition 2.5** (Spectral gap inequality). *The log-Sobolev inequality with constant  $\gamma$  implies the spectral gap inequality (also called Poincaré inequality) with the same constant:*

$$(2.39) \quad \text{Var}_\nu[F] \leq \frac{1}{\gamma} \mathbb{E}_\nu [(\nabla F)^2].$$

The same conclusion holds assuming the modified log-Sobolev inequality (2.41) below instead of the log-Sobolev inequality. The proof follows by applying the log-Sobolev inequality to the test function  $1 + \varepsilon F$  and then letting  $\varepsilon$  tend to 0, see, e.g., [81, Lemma 2.2.2].

We refer to [7] for an in-depth account on related functional inequalities and to [62] for the applications of the log-Sobolev inequality to the concentration of measure phenomenon.

For discrete spin models, the counterpart of Proposition 2.2 is the following proposition.

**Proposition 2.6** (de Bruin identity). *Consider the discrete dynamics with invariant measure  $\nu$  (2.8) and Dirichlet form  $D_\nu$  defined in (2.22). Then for  $F_t(\sigma) = \mathbb{E}_{\sigma_0=\sigma} [F(\sigma_t)]$ ,*

$$(2.40) \quad \frac{\partial}{\partial t} \text{Ent}_\nu(F_t) = -D_\nu(\log F_t, F_t) \leq 0.$$

The proof is identical to (2.33) replacing the continuous generator by  $\Delta_c$  defined in (2.21). As the chain rule no longer applies in the discrete setting, the Fisher information cannot be recovered. Nevertheless the exponential decay of the entropy (2.36) can be established under the modified log-Sobolev inequality (mLSI), i.e., if there is  $\gamma > 0$  such that for any function  $F : \{\pm 1\}^\Lambda \mapsto \mathbb{R}^+$ :

$$(2.41) \quad \text{Ent}_\nu(F) \leq \frac{1}{2\gamma} D_\nu(\log F, F).$$

In view of Exercise 2.7, this inequality is weaker than the standard log-Sobolev inequality (2.35), a point discussed in detail in [22].

**Exercise 2.7.** In the discrete case, the different quantities in (2.34) are not equal. Verify the inequality  $4(\sqrt{a} - \sqrt{b})^2 \leq (a - b) \log(a/b)$  for  $a, b > 0$  and hence show  $D_\nu(\log F, F) \geq 4D_\nu(\sqrt{F})$ .

**2.4. Bakry–Émery theorem.** In verifying the log-Sobolev inequality for spins taking values in a continuous space, a very useful criterion is the Bakry–Émery theorem which applies to log-concave probability measures.

**Theorem 2.8** (Bakry–Émery [6]). *Consider a probability measure on  $X = \mathbb{R}^N$  (or a linear subspace) of the form (2.1) and assume that there is  $\lambda > 0$  such that as quadratic forms:*

$$(2.42) \quad \forall \varphi \in X : \quad \text{Hess } H(\varphi) \geq \lambda \text{id}.$$

*Then the log-Sobolev constant of  $\nu$  satisfies  $\gamma \geq \lambda$ .*

For quadratic  $H$ , one can verify (by a simple choice of test function) that in fact the equality  $\gamma = \lambda$  holds. An equivalent way to state the assumption  $\text{Hess } H(\varphi) \geq \lambda \text{id}$  is to say that  $H$  can be written as  $H(\varphi) = \frac{1}{2}(\varphi, A\varphi) + V_0(\varphi)$  with a symmetric matrix  $A \geq \lambda \text{id}$  and  $V_0$  convex.

*Proof of Theorem 2.8.* The entropy (2.29) can be estimated by interpolation along the semigroup (2.12) of the Langevin dynamics associated with  $\nu$ . Setting  $F_t = P_t(F)$ , we note that

$$(2.43) \quad \text{Ent}_\nu(F) = \mathbb{E}_\nu[\Phi(F)] - \Phi(\mathbb{E}_\nu[F]) = \mathbb{E}_\nu[\Phi(F_0) - \Phi(F_\infty)]$$

where we used that the dynamics converges to the invariant measure (2.18) which implies

$$(2.44) \quad \Phi(\mathbb{E}_\nu[F]) = \lim_{t \rightarrow \infty} \mathbb{E}_\nu[\Phi(F_t)].$$

Indeed, we may assume that  $F$  takes values in a compact interval  $I \subset (0, \infty)$ , and by the positivity of the semigroup  $F_t$  then takes values in  $I$  for all  $t \geq 0$ , and we can replace  $\Phi$  by a bounded smooth function  $g$  that coincides with  $\Phi$  on  $I$ . By the de Bruin identity (2.31) (and using that  $\mathbb{E}_\nu[F_t]$  is independent of  $t$ ), therefore

$$(2.45) \quad \text{Ent}_\nu(F) = - \int_0^\infty dt \frac{\partial}{\partial t} \mathbb{E}_\nu[\Phi(F_t)] = - \int_0^\infty dt I_\nu(F_t).$$

Provided that the Fisher information  $I_\nu(F_t)$  introduced in (2.32) satisfies

$$(2.46) \quad I_\nu(F_t) \leq e^{-2\lambda t} I_\nu(F_0),$$

the log-Sobolev inequality follows from (2.43) and (2.45) by integrating in time:

$$(2.47) \quad \text{Ent}_\nu(F) = \int_0^\infty dt I_\nu(F_t) \leq \frac{1}{2\lambda} I_\nu(F_0) = \frac{2}{\lambda} D_\nu(\sqrt{F}).$$

To prove (2.46), differentiate again:

$$(2.48) \quad \frac{\partial}{\partial t} I(\nu_t | \nu) = \frac{\partial}{\partial t} \mathbb{E}_\nu \left[ \frac{(\nabla F_t)^2}{F_t} \right] = \mathbb{E}_\nu \left[ \left( \frac{\partial}{\partial t} - \Delta^H \right) \frac{(\nabla F_t)^2}{F_t} \right],$$

where we used that  $\mathbb{E}_\nu[\Delta^H G] = 0$  for every sufficiently nice function  $G : X \rightarrow \mathbb{R}$ . It is an elementary but somewhat tedious exercise to verify that

$$(2.49) \quad \left( \frac{\partial}{\partial t} - \Delta^H \right) \frac{(\nabla F_t)^2}{F_t} = -2F_t \left[ \underbrace{|\text{Hess } \log F_t|_2^2}_{\geq 0} + \underbrace{(\nabla \log F_t, \text{Hess } H(\varphi) \nabla \log F_t)}_{\geq \lambda \text{id}} \right].$$

Hence

$$(2.50) \quad \frac{\partial}{\partial t} I_\nu(F_t) \leq -2\lambda \mathbb{E}_\nu[F_t (\nabla \log F_t, \nabla \log F_t)] = -2\lambda \mathbb{E}_\nu \left[ \frac{(\nabla F_t)^2}{F_t} \right] = -2\lambda I_\nu(F_t)$$

which implies the claim (2.46).  $\square$

**2.5. Decomposition and properties of the entropy.** The Bakry–Émery criterion (Theorem 2.8) implies the validity of the log-Sobolev inequality for all the Gibbs measures with strictly convex potentials. For more general measures, the log-Sobolev inequality is often derived by decomposing the entropy thanks to successive conditionings. We are going to sketch this procedure below.

Assume that the expectation under  $\nu$  is of the form

$$(2.51) \quad \mathbb{E}_\nu[F] := \mathbb{E}_\nu[F(\varphi_1, \varphi_2)] = \mathbb{E}_2 \left[ \mathbb{E}_1[F(\varphi_1, \varphi_2) | \varphi_2] \right] = \mathbb{E}_2[\mathbb{E}_1[F]],$$

where  $\mathbb{E}_1[\cdot] := \mathbb{E}_1[\cdot|\varphi_2]$  is the conditional measure wrt the variable  $\varphi_2$ . Then the entropy can be split into two parts:

$$(2.52) \quad \begin{aligned} \text{Ent}(F) &= \mathbb{E}_\nu[\Phi(F)] - \Phi(\mathbb{E}_\nu[F]) \\ &= \mathbb{E}_2 \left[ \underbrace{\mathbb{E}_1[\Phi(F)] - \Phi(\mathbb{E}_1[F])}_{\text{Ent}_1(F)} \right] + \mathbb{E}_2 \left[ \underbrace{\Phi(\mathbb{E}_1[F]) - \Phi(\mathbb{E}_2[\mathbb{E}_1[F]])}_{\text{Ent}_2(\mathbb{E}_1[F])} \right]. \end{aligned}$$

Given  $\varphi_2$ , the first term involves the relative entropy of a simpler measure  $\mathbb{E}_1(\cdot|\varphi_2)$  as the integration refers only to the coordinate  $\varphi_1$ :

$$(2.53) \quad \text{Ent}_1(F)(\varphi_2) = \mathbb{E}_1[\Phi(F)|\varphi_2] - \Phi(\mathbb{E}_1[F|\varphi_2]).$$

The strategy is to estimate this term (uniformly in  $\varphi_2$ ) by the desired Dirichlet form acting only on  $\varphi_1$ . The second term  $\mathbb{E}_2[\Phi(\mathbb{E}_1[F])]$  is more complicated because the expectation  $\mathbb{E}_1[F(\cdot, \varphi_2)|\varphi_2]$  is inside the relative entropy. For a product measure  $\nu = \nu_1 \otimes \nu_2$ , this term can be estimated easily (as recalled in Example 2.9 below). In this way, the log-Sobolev inequality for  $\nu = \nu_1 \otimes \nu_2$  is reduced to establishing log-Sobolev inequalities for the simpler measures  $\nu_1$  and  $\nu_2$ .

In general, the conditional expectations are intertwined and the second term  $\text{Ent}_2(\mathbb{E}_1[F])$  is much more difficult to estimate. There are two general strategies: either one also bounds this term by the desired Dirichlet form (and thus one somehow has to move the expectation out of the entropy) or one bounds it by  $\kappa \text{Ent}(F)$  with  $\kappa < 1$ . In the latter case, the estimate reduces to the first term, at the expense of an overall factor  $(1 - \kappa)^{-1}$ .

For a given measure  $\nu$ , the decomposition entropy (2.52) can be achieved with different choices of the measures  $\mathbb{E}_1, \mathbb{E}_2$ . The optimal choice depends on the structure of the measure  $\nu$ . In this survey, we focus on Gibbs measures of the form (2.1) which arise naturally in statistical mechanics. The renormalisation group method constitutes a framework to study such Gibbs measures (see [17] for an introduction and references) and provides strong insight on a good entropy decomposition. This is the core of the method presented in Section 3, which is based on the Polchinski equation, a continuous version of the renormalisation group.

**Example 2.9** (Tensorisation). Assume that probability measures  $\nu_1$  and  $\nu_2$  satisfy log-Sobolev inequalities with the constants  $\gamma_1$  and  $\gamma_2$ . Then the product measure  $\nu = \nu_1 \otimes \nu_2$  also satisfies a log-Sobolev inequality with constant  $\gamma = \min\{\gamma_1, \gamma_2\}$  (with the natural Dirichlet form on the product space).

*Proof.* For simplicity, assume that  $\nu_1, \nu_2$  are probability measures on  $\mathbb{R}$  and denote by  $\mathbb{E}_1, \mathbb{E}_2$  their expectations so that  $\mathbb{E}_\nu[F] = \mathbb{E}_2[\mathbb{E}_1[F]]$  for functions  $F(\varphi_1, \varphi_2)$ . As discussed in (2.52), the entropy can be decomposed as

$$(2.54) \quad \text{Ent}(F) = \mathbb{E}_2[\text{Ent}_1(F)] + \text{Ent}_2(G) \quad \text{with} \quad G(\varphi_2) = \mathbb{E}_1[F(\cdot, \varphi_2)].$$

The log-Sobolev inequalities for  $\mathbb{E}_1$  and  $\mathbb{E}_2$  imply that for  $\gamma = \min\{\gamma_1, \gamma_2\}$ :

$$(2.55) \quad 2\gamma \text{Ent}(F) \leq \mathbb{E}_2[D_1(\sqrt{F})] + D_2(\sqrt{G}).$$

It remains to recover the Dirichlet form associated with the product measure  $\nu$ :

$$(2.56) \quad D_\nu(\sqrt{F}) = \mathbb{E}_\nu \left[ (\partial_{\varphi_1} \sqrt{F})^2 + (\partial_{\varphi_2} \sqrt{F})^2 \right].$$

The first derivative is easily identified:

$$(2.57) \quad \mathbb{E}_2[D_1(\sqrt{F})] = \mathbb{E}_2 \left[ \mathbb{E}_1 \left[ (\partial_{\varphi_1} \sqrt{F})^2 \right] \right] = \mathbb{E}_\nu \left[ (\partial_{\varphi_1} \sqrt{F})^2 \right].$$

For the second derivative:

$$(2.58) \quad \partial_{\varphi_2} \sqrt{G(\varphi_2)} = \frac{1}{2\sqrt{G(\varphi_2)}} \mathbb{E}_1[\partial_{\varphi_2} F(\cdot, \varphi_2)] = \frac{1}{\sqrt{G(\varphi_2)}} \mathbb{E}_1[\sqrt{F(\cdot, \varphi_2)} \partial_{\varphi_2} \sqrt{F(\cdot, \varphi_2)}],$$

so that by Cauchy-Schwarz inequality, we deduce that

$$(2.59) \quad D_2(\sqrt{G}) = \mathbb{E}_2 \left[ \left( \partial_{\varphi_2} \sqrt{G(\varphi_2)} \right)^2 \right] \leq \mathbb{E}_2 \left[ \mathbb{E}_1 \left[ \left( \partial_{\varphi_2} \sqrt{F(\cdot, \varphi_2)} \right)^2 \right] \right] = \mathbb{E}_\nu \left[ \left( \partial_{\varphi_2} \sqrt{F} \right)^2 \right].$$

This reconstructs the Dirichlet form (2.56) and completes the proof.

A similar argument applies in the discrete case, see for example [81, Lemma 2.2.11]. More abstractly, the tensorisation of the log-Sobolev constant also follows from the equivalence between the log-Sobolev inequality and hypercontractivity (which tensorises more obviously).  $\square$

We conclude this section by stating useful variational characterisations of the entropy.

**Proposition 2.10** (Entropy inequality). *The entropy of a function  $F \geq 0$  can be rewritten as*

$$(2.60) \quad \text{Ent}_\nu(F) = \sup \left\{ \mathbb{E}_\nu[FG] : \text{Borel functions } G \text{ such that } \mathbb{E}_\nu[e^G] \leq 1 \right\}$$

with equality if  $G = \log\left(\frac{F}{\mathbb{E}_\nu[F]}\right)$ , or as

$$(2.61) \quad \text{Ent}_\nu(F) = \sup \left\{ \mathbb{E}_\nu[F \log F - F \log t - F + t] : t > 0 \right\}$$

with equality if  $t = \mathbb{E}_\nu[F]$ . Finally, one has (also called the entropy inequality):

$$(2.62) \quad \text{Ent}_\nu(F) = \sup \left\{ \mathbb{E}_\nu[FG] - \mathbb{E}_\nu[F] \log \mathbb{E}_\nu[e^G] : \text{Borel functions } G \right\}$$

with equality if  $G = \log F$ .

*Proof.* Since  $\text{Ent}_\nu(F) = \mathbb{E}_\nu \left[ F \log \frac{F}{\mathbb{E}_\nu[F]} \right]$ , to show (2.60), it is enough to consider the case  $\mathbb{E}_\nu[F] = 1$ , by homogeneity of both sides. Applying Young's inequality

$$(2.63) \quad \forall a \geq 0, b \in \mathbb{R}, \quad ab \leq a \log a - a + e^b,$$

with  $\mathbb{E}_\nu[e^G] \leq 1$ , we get

$$(2.64) \quad \mathbb{E}_\nu[FG] \leq \text{Ent}_\nu(F) - 1 + \mathbb{E}_\nu[e^G] \leq \text{Ent}_\nu(F).$$

This implies (2.60) as the converse inequality holds with  $G = \log F$ .

The variational formula (2.61) follows directly from Young's inequality by choosing  $a = \mathbb{E}_\nu[F]$  and  $b = \log t$ :

$$(2.65) \quad \text{Ent}_\nu(F) = \mathbb{E}_\nu[F \log F] - a \log a \leq \mathbb{E}_\nu[F \log F] - ab - a + e^b = \mathbb{E}_\nu[F \log F - F \log t - F + t],$$

where again (2.61) follows since equality holds with  $t = \mathbb{E}_\nu[F]$ .

To show (2.62), we may again assume  $\mathbb{E}_\nu[F] = 1$ . Then apply Jensen's inequality with respect to the probability measure  $d\nu^F = F d\nu$ :

$$(2.66) \quad \log \mathbb{E}_\nu[e^G] = \log \mathbb{E}_{\nu^F}[F^{-1} e^G] \geq \mathbb{E}_{\nu^F}[\log(F^{-1} e^G)] = -\mathbb{E}_\nu[F \log F] + \mathbb{E}_\nu[FG],$$

with equality if  $G = \log F$ .  $\square$

The Holley–Stroock criterion for the log-Sobolev inequality is a simple consequence of (2.61), see e.g., the presentation in [86].

**Exercise 2.11** (Holley–Stroock criterion). Assume a measure  $\nu$  satisfies the log-Sobolev inequality with constant  $\gamma$ . Then the measure  $\nu^F$  with  $d\nu^F/d\nu = F$  satisfies a log-Sobolev inequality with constant  $\gamma^F \geq (\inf F / \sup F)\gamma$ .

**2.6. Difficulties arising from statistical physics perspective.** To explain the difficulties arising in the derivation of log-Sobolev inequalities and to motivate our set-up of renormalisation, we are going to consider lattice spin systems with continuous spins and Hamiltonian of the form (2.4). The strength of the interaction is tuned by the parameter  $\beta \geq 0$  and the Gibbs measure (2.1) has a density on  $\mathbb{R}^\Lambda$  of the form

$$(2.67) \quad \nu(d\varphi) \propto \exp \left[ -\frac{\beta}{4} \sum_{\substack{x,y \in \Lambda, \\ x \sim y}} (\varphi_x - \varphi_y)^2 - \sum_{x \in \Lambda} V(\varphi_x) \right] \prod_{x \in \Lambda} d\varphi_x.$$

Further examples will be detailed in Section 6.1.

We are interested in the behaviour of the measure (2.67) in the limit where the the number of sites  $|\Lambda|$  (and thus the dimension of the configuration space) is large. In this limit, when the potential  $V$  is not convex, the measure can have one or more phase transitions at critical values of the parameter  $\beta$ . These phase transitions separate regions of values of  $\beta$  between which the measure  $\nu$  has a different correlation structure, different concentration properties, and so on. The speed of convergence of the associated Glauber dynamics (2.7) is also affected. See the book [47] or [17, Chapter 1] for background on phase transitions in statistical mechanics.

To analyse the log-Sobolev inequality for the measure (2.67), note that the lack of convexity precludes the use of the Bakry–Émery criterion (Theorem 2.8), and the Holley–Stroock criterion (Exercise 2.11) is not effective due to the large dimension of the configuration space when  $\beta > 0$ .

On the other hand, when  $\beta = 0$ , the Gibbs measure is a product measure and the log-Sobolev inequality holds uniformly in  $\Lambda$ , with the same constant as for the single spin  $|\Lambda| = 1$  measure (Example 2.9). This tensorisation property has been generalised for  $\beta$  small enough in terms of mixing conditions and for some spin systems up to the critical value  $\beta_c$ , see [68] for a review. Indeed, for  $\beta$  small, the interaction between the spins is small, in the sense that one can show that correlations between spins decay exponentially in their distance. At distances larger than a correlation length  $\xi_\beta < \infty$  approximate independence between the spins is then recovered. By splitting the domain  $\Lambda$  into boxes (of size larger than  $\xi_\beta$ ), and using appropriate conditionings the system can be analysed as a renormalised model of weakly interacting spins [90]. In so-called second order phase transitions, when  $\beta$  approaches the critical  $\beta_c$  the correlation length diverges as a function of  $\beta_c - \beta$ , and so does the inverse log-Sobolev constant, i.e., the dynamics slows down (as can usually be verified by simple test functions in the spectral gap or log-Sobolev inequality). Spins are thus more and more correlated for  $\beta$  close to  $\beta_c$ . Nevertheless in some cases [66–68] the strong dynamical mixing properties were derived up to the critical value  $\beta_c$  by using a strategy which however can be seen as a (large) perturbation of the product case with respect to  $\beta > 0$ . For this reason, it seems difficult to extract the precise divergence of the log-Sobolev constant near  $\beta_c$  with this type of approach.

To study the detailed static features of measures of the form (2.67) close to  $\beta_c$ , different types of renormalisation schemes have been devised with an emphasis on the Gaussian structure of the interaction. In many cases one expects that the long range structure at the critical point is well described in terms of a Gaussian free field [89]. Compared with the previously mentioned approaches, the perturbation theory no longer uses the product measure as a reference, but the Gaussian free field. In the following, this structure will serve as a guide to decompose the entropy as alluded to in (2.52). Before describing this procedure in Section 3, the elementary example of the Gaussian free field, which illustrates the difficulty of many length scales equilibrating at different rates, is presented in Example 2.12 below.

**2.7. Difficulties arising from continuum perspective.** The long-distance problem discussed in the previous subsection is closely related to the short-distance problem occurring in the study of continuum limits as they appear in quantum field theory and weak interaction limits, which also arise as invariant measure of (singular) SPDEs. In field theory, one is interested in the physical

behaviour of a measure that is defined not on a lattice, but in the continuum (say on  $L\mathbb{T}^d$  with  $\mathbb{T}^d = [0, 1)^d$  the  $d$ -dimensional torus), formally reading:

$$(2.68) \quad \nu_L(d\varphi) \propto e^{-H_L(\varphi)} \prod_{x \in L\mathbb{T}^d} d\varphi_x.$$

Now  $\varphi$  should be a (generalised) function from  $L\mathbb{T}^d$  to  $\mathbb{R}$ , and a typical example for  $H_L$  would be the continuum  $\varphi^4$  model, defined for  $\lambda > 0$  and  $\mu \in \mathbb{R}$  by:

$$(2.69) \quad H_L(\varphi) = \frac{1}{2} \int_{L\mathbb{T}^d} |\nabla\varphi|^2 dx + \int_{L\mathbb{T}^d} \left[ \frac{\lambda}{4} \varphi(x)^4 + \frac{\mu}{2} \varphi(x)^2 \right] dx.$$

Of course, the formal definition (2.68) does not make sense as it stands, and a standard approach to understand such measures is as a limit of measures defined on lattices  $\Lambda_{\varepsilon,L} = L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$  with a vanishing  $\varepsilon$ :

$$(2.70) \quad \nu_{\varepsilon,L}(d\varphi) \propto e^{-H_{\varepsilon,L}(\varphi)} \prod_{x \in \Lambda_{\varepsilon,L}} d\varphi_x,$$

for a discrete approximation  $H_{\varepsilon,L}$  of  $H_L$  of the form

$$(2.71) \quad H_{\varepsilon,L}(\varphi) = \frac{\varepsilon^{d-2}}{4} \sum_{y \sim x} (\varphi_y - \varphi_x)^2 + \varepsilon^d \sum_{x \in \Lambda_{\varepsilon,L}} V^\varepsilon(\varphi_x),$$

where the potential  $V^\varepsilon$  is of the form (2.5). In the example of the  $\varphi^4$  model, it turns out that such a limit can be constructed if  $d < 4$ , but in  $d \geq 2$  the coefficient  $\mu$  of the potential must be tuned correctly as a function of  $\varepsilon \rightarrow 0$ , see Section 6.1 for details. This tuning is known as addition of “counterterms” in quantum field theory. These are the infamous infinities arising there. The relation to the statistical physics (long-distance) problem is that these limits correspond to statistical physics models near a phase transition, with scaled (weak) interaction strength ( $\varepsilon^d \lambda \rightarrow 0$  as  $\varepsilon \rightarrow 0$ ). In particular, due to the counterterms, the resulting measures are usually again very non-convex microscopically, precluding the use of the Bakry–Émery theory and the Holley–Stroock criteria. Nonetheless the regularisation parameter  $\varepsilon$  is not expected to have any influence on the physics of the model, in the sense that the existence of a phase transition, the speed of the Glauber dynamics, concentration properties, and so on, should all be uniform in the small scale parameter  $\varepsilon$  and depend only on the large scale parameter  $L$ . Techniques to control the regularisation parameter  $\varepsilon$  are often simpler than for the large scale problem near the critical point, but they are also based on renormalisation arguments relying on comparisons with the Gaussian free field (corresponding to a quadratic  $H_{\varepsilon,L}$ ).

Finally, the problem of relaxation at different scales is illustrated next in this simple model.

**Example 2.12** (Free field dynamics). Consider now the Gaussian free field dynamics corresponding to  $V^\varepsilon = 0$  in (2.71):

$$(2.72) \quad d\varphi_t = -A\varphi_t dt + \sqrt{2}dW_t,$$

where  $A$  is the Laplace operator on  $\Lambda_{\varepsilon,L}$  as in (2.71) and the white noise is defined with respect to the inner product  $\varepsilon^d \sum_{x \in \Lambda_{\varepsilon,L}} u_x v_x$ , i.e., each  $W_t(x)$  is a Brownian motion of variance  $\varepsilon^{-d}$ . On the torus  $\Lambda_{\varepsilon,L} = L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$  of mesh size  $\varepsilon$  and side length  $L$ , the eigenvalues of the Laplacian are

$$(2.73) \quad \lambda(p) = \varepsilon^{-2} \sum_{i=1}^d 2(\cos(\varepsilon p_i) - 1) \underset{|p| \lesssim 1}{\approx} -|p|^2, \quad p \in \Lambda_{\varepsilon,L}^* = \left(-\frac{\pi}{\varepsilon}, \frac{\pi}{\varepsilon}\right]^d \cap \frac{2\pi}{L}\mathbb{Z}^d.$$

All Fourier modes of  $\varphi$  evolve independently according to Ornstein–Uhlenbeck processes:

$$(2.74) \quad p \in \Lambda_{\varepsilon, L}^*, \quad d\hat{\varphi}(p) = -\lambda(p)\hat{\varphi}(p) dt + \sqrt{2}d\hat{W}_t(p),$$

where the  $\hat{W}(p) = (\hat{W}_t(p))_t$  are independent standard Brownian motions for  $p \in \Lambda^*$ . In particular, small scales corresponding to  $|p| \gg 1$  converge very quickly to equilibrium, while the large scales  $|p| \ll 1$  are slowest. Thus the main contribution to the log-Sobolev constant comes from the large scales and we expect that a similar structure remains relevant in many interacting systems close to a critical point.

In both the statistical and continuum perspectives, for measures with an interaction  $V_0(\varphi) = \sum_{x \in \Lambda} V(\varphi_x) \neq 0$  on top of the free field interaction, the main difficulties result from the simple fact that the local (in real space) interaction do not interact well with the above Fourier decomposition. The Polchinski flow that we will introduce in the next section can be seen as a replacement for the Fourier decomposition, in which the Fourier variable  $p$  takes the role of scale, by a smoother scale decomposition.

### 3 Gaussian integration and the Polchinski equation

In this section, we first review abstractly a continuous renormalisation procedure, which goes back to Wilson [89] and Polchinski [24, 78] in physics in the context of equilibrium phase transitions and quantum field theory (viewed as an problem of statistical mechanics in the continuum). We then explain how the entropy of a measure can be decomposed by this method in order to derive a log-Sobolev inequality via a multiscale Bakry–Émery criterion.

**3.1. Gaussian integration.** For  $C$  a positive semi-definite matrix on  $\mathbb{R}^N$ , we denote by  $\mathbf{P}_C$  the corresponding Gaussian measure with covariance  $C$  and by  $\mathbf{E}_C$  its expectation. The measure  $\mathbf{P}_C$  is supported on the image of  $C$ . In particular, if  $C$  is strictly positive definite on  $\mathbb{R}^N$ ,

$$(3.1) \quad \mathbf{E}_C[F] \propto \int_{\mathbb{R}^N} e^{-\frac{1}{2}(\zeta, C^{-1}\zeta)} F(\zeta) d\zeta.$$

A fundamental property of the Gaussian measure is its semigroup property: if  $C = C_1 + C_2$  with  $C_1, C_2$  also positive semi-definite then

$$(3.2) \quad \mathbf{E}_C[F(\zeta)] = \mathbf{E}_{C_2}[\mathbf{E}_{C_1}[F(\zeta_1 + \zeta_2)]],$$

corresponding to its probabilistic interpretation that if  $\zeta_1$  and  $\zeta_2$  are independent Gaussian random variables then  $\zeta_1 + \zeta_2$  is also Gaussian and the covariance of  $\zeta_1 + \zeta_2$  is the sum of the covariances of  $\zeta_1$  and  $\zeta_2$ .

As discussed in Section 2.5, recall that our goal is to decompose the entropy of a measure by splitting this measure into simpler parts as in (2.52). The above Gaussian decomposition will be a basic step for this. As we have seen in Example 2.12, the dynamics of a spin or particle system close to a phase transition will depend on a very large number of modes and it will be necessary to iterate the decomposition (3.2) many times in order to decouple all the relevant modes. In fact, it is even convenient to introduce a continuous version of the decomposition (3.2), as follows.

For a covariance matrix  $C$  as above, define an associated Laplace operator  $\Delta_C$  on  $\mathbb{R}^N$ :

$$(3.3) \quad \Delta_C = \sum_{i, j} C_{ij} \frac{\partial^2}{\partial \varphi_i \partial \varphi_j},$$

and write  $(\cdot, \cdot)_C$  for the inner product associated with the covariance  $C$ :

$$(3.4) \quad (u, v)_C = \sum_{i, j} C_{ij} u_i v_j \quad \text{and} \quad |u|_C^2 = (u)_C^2 = (u, u)_C.$$

The standard scalar product is denoted by  $(u, v) = \sum_i u_i v_i$ .

Let  $t \in [0, +\infty] \mapsto C_t$  be a function of positive semidefinite matrices on  $\mathbb{R}^N$  increasing continuously as quadratic forms to a matrix  $C_\infty$ . More precisely, we assume that  $C_t = \int_0^t \dot{C}_s ds$  for all  $t$ , where  $t \mapsto \dot{C}_t$  is a bounded cadlag (right-continuous with left limits) function with values in the space of positive semidefinite matrices that is the derivative of  $C_t$  except at isolated points. We say that  $C_\infty = \int_0^\infty \dot{C}_s ds$  is a *covariance decomposition* and write  $X \subset \mathbb{R}^N$  for the image of  $C_\infty$ . We emphasise that the (closed) interval  $[0, +\infty]$  parametrising the covariances has no special significance and that all constructions will be invariant under appropriate reparametrisation. For example, one can equivalently use  $[0, 1]$ .

**Proposition 3.1.** *For a  $C^2$  function  $F : X \rightarrow \mathbb{R}$ , let  $F_t = P_{C_t} * F$ , i.e.,  $F_t(\varphi) = \mathbb{E}_{C_t} [F(\varphi + \zeta)]$ . Then for all  $t$  which are not discontinuity points of  $\dot{C}_t$ ,*

$$(3.5) \quad \frac{\partial}{\partial t} F_t = \frac{1}{2} \Delta_{\dot{C}_t} F_t, \quad F_0 = F.$$

Thus the Gaussian measures  $P_{C_t}$  satisfy the heat equation

$$(3.6) \quad \frac{\partial}{\partial t} P_{C_t} = \frac{1}{2} \Delta_{\dot{C}_t} P_{C_t},$$

interpreted in a weak sense if  $C_t$  is not strictly positive definite.

*Proof.* In the case that  $C_{t-\varepsilon}$  is strictly positive definite for some  $\varepsilon > 0$  (and by monotonicity then also for all larger times), this is a direct computation from (3.1). For  $t$  that are not discontinuity points of  $\dot{C}_t$ , the image  $X_s$  of  $C_s$  is independent of  $s \in [t-\varepsilon, t+\varepsilon]$  and one has the representation (3.1) on  $X_t$ .  $\square$

Alternatively, one can prove the proposition using Itô's formula. For any  $\varphi \in \mathbb{R}^N$ , define the process

$$(3.7) \quad \forall t \geq 0, \quad \zeta_t = \varphi + \int_0^t \sqrt{\dot{C}_s} dB_s \in \mathbb{R}^N,$$

where  $(B_s)_s$  is a Brownian motion taking values in  $\mathbb{R}^N$ . By construction  $\zeta_t$  is a Gaussian variable with mean  $\varphi$  and variance  $C_t = \int_0^t \dot{C}_s ds$ . In particular  $F_t(\varphi) = \mathbb{E} [F(\zeta_t)]$  and by Itô's formula,

$$(3.8) \quad \forall t \geq 0, \quad \frac{\partial}{\partial t} F_t(\varphi) = \mathbb{E} \left[ \frac{1}{2} \Delta_{\dot{C}_t} F(\zeta_t) \right] = \frac{1}{2} \Delta_{\dot{C}_t} \mathbb{E}_{C_t} [F(\varphi + \zeta)] = \frac{1}{2} \Delta_{\dot{C}_t} F_t(\varphi),$$

with the derivative interpreted as the right-derivative at the discontinuity points of  $\dot{C}_t$ .

Note that the decomposition (3.7) is the natural extension of the discrete decomposition  $\zeta = \zeta_1 + \zeta_2$ . Given a covariance matrix  $C$ , many decompositions are possible, such as:

$$(3.9) \quad C = \int_0^\infty \dot{C}_s ds \quad \text{with} \quad \dot{C}_s = C \mathbf{1}_{s \in [0,1]}.$$

For a given model from statistical mechanics, it will be important to adjust the decomposition according to the specific spatial structure (of  $\Lambda$ ) of this model. In Example 2.12, the Gaussian free field has covariance matrix  $A^{-1}$  with  $A$  the discrete Laplace operator as in (2.3). There are many decompositions of  $A^{-1}$  of the form  $\int_0^\infty \dot{C}_s ds$  and the best choice will depend on the application. Nevertheless a suitable decomposition should capture the mode structure of the decomposition (2.74) in order to separate the different scales in the dynamics. Indeed, the key idea of a renormalisation group approach is to integrate the different scales one after the other in order. This is especially important in strongly correlated systems, in which the different scales do not decouple, and integrating some scales has an important effect on the remaining scales: the interaction potential will get renormalised. We refer to Section 6 for several applications.



**3.2. Renormalised potential and Polchinski equation.** In this section, we define the Polchinski flow and analyse its structure. A simple explicit example is worked out in Example 3.5 below. We will focus on probability measures  $\nu_0$  supported on a linear subspace  $X \subset \mathbb{R}^N$ . By considering the measure  $\nu_0(A - a)$  for  $a \in \mathbb{R}^N$ , this also includes measures supported on an affine subspace which is of interest for conservative dynamics. For generalisations to non-linear spaces, see Section 3.6.

Let  $C_\infty = \int_0^\infty \dot{C}_t dt$  be a covariance decomposition, and consider a probability measure  $\nu_0$  on  $X$  with expectation given by

$$(3.10) \quad \mathbb{E}_{\nu_0}[F] \propto \mathbf{E}_{C_\infty} \left[ e^{-V_0(\zeta)} F(\zeta) \right],$$

with a potential  $V_0 : X \rightarrow \mathbb{R}$ , where the Gaussian expectation acts on the variable  $\zeta$ . To avoid technical problems, we always assume in the following that  $V_0$  is bounded below. We are going to use the Gaussian representation introduced in the previous subsection in order to decompose the measure  $\nu_0$ . For this, let us first introduce some notation.

**Definition 3.2.** For  $t > s > 0$ ,  $F : X \rightarrow \mathbb{R}$  bounded, and  $\varphi \in X$ , define:

- the renormalised potential  $V_t$ :

$$(3.11) \quad V_t(\varphi) = -\log \mathbf{E}_{C_t} \left[ e^{-V_0(\varphi+\zeta)} \right];$$

- the Polchinski semigroup  $\mathbf{P}_{s,t}$ :

$$(3.12) \quad \mathbf{P}_{s,t}F(\varphi) = e^{V_t(\varphi)} \mathbf{E}_{C_t-C_s} \left[ e^{-V_s(\varphi+\zeta)} F(\varphi + \zeta) \right];$$

- the renormalised measure  $\nu_t$ :

$$(3.13) \quad \mathbb{E}_{\nu_t}[F] = \mathbf{P}_{t,\infty}F(0) = e^{V_\infty(0)} \mathbf{E}_{C_\infty-C_t} \left[ e^{-V_t(\zeta)} F(\zeta) \right],$$

where all the Gaussian expectations apply to  $\zeta$ .

We stress that the renormalised measure  $\nu_t$  evolving according to the Polchinski semigroup is different from the measure  $m_t$  in (2.12) evolving along the flow of the Langevin dynamics (which we will not discuss directly in this section).

Note that in (3.13),  $e^{+V_\infty(0)}$  is the normalisation factor of the probability measure  $\nu_t$ . More generally, the function  $V_\infty$  is equivalent to the moment generating function of the measure  $\nu_0$ : changing variables from  $\zeta$  to  $\zeta + C_\infty h$ ,

$$(3.14) \quad \begin{aligned} V_\infty(C_\infty h) &= -\log \mathbf{E}_{C_\infty} [e^{-V_0(C_\infty h + \zeta)}] = \frac{1}{2}(h, C_\infty h) - \log \mathbf{E}_{C_\infty} [e^{-V_0(\zeta)} e^{(h, \zeta)}] \\ &= \frac{1}{2}(h, C_\infty h) - \log \mathbb{E}_{\nu_0} [e^{(h, \zeta)}] + V_\infty(0). \end{aligned}$$

The renormalised measure  $\nu_t$  is related to  $\nu_0$  by the following identity.

**Proposition 3.3.** For  $t \geq 0$  and any  $F : X \mapsto \mathbb{R}$  such that the following quantities make sense,

$$(3.15) \quad \mathbb{E}_{\nu_0} [F] = \mathbb{E}_{\nu_t} [\mathbf{P}_{0,t}F].$$

*Proof.* Starting from (3.10), from the Gaussian decomposition (3.2) we get

$$(3.16) \quad \begin{aligned} \mathbf{E}_{C_\infty} \left[ e^{-V_0(\zeta)} F(\zeta) \right] &= \mathbf{E}_{C_\infty-C_t} \left[ \mathbf{E}_{C_t} \left[ e^{-V_0(\varphi+\zeta)} F(\varphi + \zeta) \right] \right] \\ &= \mathbf{E}_{C_\infty-C_t} \left[ e^{-V_t(\varphi)} e^{+V_t(\varphi)} \mathbf{E}_{C_t} \left[ e^{-V_0(\varphi+\zeta)} F(\varphi + \zeta) \right] \right] \\ &\propto \mathbb{E}_{\nu_t} [\mathbf{P}_{0,t}F], \end{aligned}$$

where  $\zeta$  is integrated with respect to  $\mathbf{E}_{C_t}$  and  $\varphi$  with respect to  $\mathbf{E}_{C_\infty-C_t}$ . We used the definitions (3.12) and (3.13) in the last line, and recall that  $\propto$  is an equality up to a normalising factor so that  $\nu_t$  is a probability measure. This completes the proof of the identity (3.15).  $\square$

Using the definition (3.11) of  $V_t$ , the action of the Polchinski semigroup (3.12) can be interpreted as a conditional expectation with respect to  $\varphi$ :

$$(3.17) \quad \mathbf{P}_{0,t}F(\varphi) = \frac{\mathbb{E}_{C_t} [e^{-V_0(\varphi+\zeta)} F(\varphi + \zeta)]}{\mathbb{E}_{C_t} [e^{-V_0(\varphi+\zeta)}]} =: \mathbb{E}_{\mu_t^\varphi} [F(\zeta)].$$

This defines a probability measure  $\mu_t^\varphi$  called the *fluctuation measure*. Assuming  $C_t$  is invertible and changing variables from  $\varphi + \zeta$  to  $\zeta$ , the fluctuation measure can be written equivalently as

$$(3.18) \quad \mu_t^\varphi(d\zeta) = e^{+V_t(\varphi)} e^{-\frac{1}{2}(\varphi-\zeta, C_t^{-1}(\varphi-\zeta)) - V_0(\zeta)} d\zeta \propto e^{-\frac{1}{2}(\zeta, C_t^{-1}\zeta) + (\zeta, C_t^{-1}\varphi) - V_0(\zeta)} d\zeta.$$

Besides the addition of an external field  $C_t^{-1}\varphi$ , the structure of this new measure is similar to the one of the original measure  $\nu_0$  introduced in (3.10), but the covariance of the Gaussian integration is now  $C_t$ . By construction  $C_t \leq C_\infty$ , so that the Hamiltonian of the conditional measure (3.17) is more convex and will hopefully be easier to handle. The fluctuation measure is central in the *stochastic localisation* framework which will be presented in Section 4.5.

For all bounded function  $F : X \rightarrow \mathbb{R}$  and all  $t > 0$ , the identity (3.15) reads

$$(3.19) \quad \mathbb{E}_{\nu_0}[F] = \mathbb{E}_{\nu_t}[\mathbf{P}_{0,t}F(\varphi)] = \mathbb{E}_{\nu_t}[\mathbb{E}_{\mu_t^\varphi}[F(\zeta)]],$$

where  $\varphi$  denotes the variable of  $\nu_t$  and  $\zeta$  the variable of  $\mu_t^\varphi$ . This is therefore an instance of the measure decomposition (2.51) by successive conditionings. The splitting of the covariance  $C_\infty = C_\infty - C_t + C_t$  will be chosen so that the field  $\zeta$  encodes the local interactions, which correspond to the fast scales of the dynamics, and  $\varphi$  the long range part of the interaction, associated with the slow dynamical modes. Integrating out the short scales boils down to considering a new test function  $\mathbf{P}_{0,t}F(\varphi)$  and a measure  $\nu_t$  (3.13) which is expected to have better properties than the original measure  $\nu_0$ . This is illustrated in a one-dimensional case in Example 3.5.

Models from statistical mechanics often involve a multiscale structure when approaching the phase transition. For this reason, it is not enough to split the measure into two parts as in (3.15). The renormalisation procedure is based on a recursive procedure with successive integrations of the fast scales in order to simplify the measure step by step. As an example, let us describe a two step procedure: for  $s < t$ , splitting the covariance into  $C_s, C_t - C_s, C_\infty - C_t$ , can be achieved by applying twice (3.15)

$$(3.20) \quad \mathbb{E}_{\nu_0}[F] = \mathbb{E}_{\nu_s}[\mathbf{P}_{0,s}F] = \mathbb{E}_{\nu_t}[\mathbf{P}_{s,t}(\mathbf{P}_{0,s}F)] = \mathbb{E}_{\nu_t}[\mathbf{P}_{0,t}F].$$

Thus  $\mathbf{P}_{s,t}$  inherits a semigroup property from the nested integrations. For infinitesimal renormalisation steps, we are going to show in Proposition 3.5 that the Polchinski semigroup is in fact a Markov semigroup with a structure reminiscent of the Langevin semigroup (2.13). To implement this renormalisation procedure, one has also to control the renormalised measure  $\nu_t$ . For infinitesimal renormalisation steps, its potential  $V_t$  evolves according to the following Hamilton–Jacobi–Bellman equation, known as *Polchinski equation*.

**Proposition 3.4.** *Let  $(C_t)$  be as above, and let  $V_0 \in C^2$ . Then for every  $t$  such that  $C_t$  is differentiable the renormalised potential  $V_t$  defined in (3.11) satisfies the Polchinski equation*

$$(3.21) \quad \frac{\partial}{\partial t} V_t = \frac{1}{2} \Delta_{\dot{C}_t} V_t - \frac{1}{2} (\nabla V_t)_{\dot{C}_t}^2$$

where  $\Delta_{\dot{C}_t}$  was defined in (3.3) and the scalar product in (3.4).

*Proof.* Let  $Z_t(\varphi) = \mathbb{E}_{C_t}[e^{-V_0(\varphi+\zeta)}]$ . By Proposition 3.1, it follows that the Gaussian convolution acts as the heat semigroup with time-dependent generator  $\frac{1}{2}\Delta_{\dot{C}_t}$ , i.e., if  $Z_0$  is  $C^2$  in  $\varphi$  so is  $Z_t$  for any  $t > 0$ , that  $Z_t(\varphi) > 0$  for any  $t$  and  $\varphi$ , and that for any  $t > 0$  such that  $C_t$  is differentiable,

$$(3.22) \quad \frac{\partial}{\partial t} Z_t = \frac{1}{2} \Delta_{\dot{C}_t} Z_t, \quad Z_0 = e^{-V_0}.$$

Since  $Z_t(\varphi) > 0$  for all  $\varphi$ , its logarithm  $V_t = -\log Z_t$  is well-defined and satisfies the Polchinski equation

$$(3.23) \quad \frac{\partial}{\partial t} V_t = -\frac{\frac{\partial}{\partial t} Z_t}{Z_t} = -\frac{\Delta_{\dot{C}_t} Z_t}{2Z_t} = -\frac{1}{2} e^{V_t} \Delta_{\dot{C}_t} e^{-V_t} = \frac{1}{2} \Delta_{\dot{C}_t} V_t - \frac{1}{2} (\nabla V_t)_{\dot{C}_t}^2. \quad \square$$

The semigroup structure is analysed in the following proposition. As mentioned above, we assume  $V_0$  to be bounded below to avoid technical problems.

**Proposition 3.5.** *The operators  $(\mathbf{P}_{s,t})_{s \leq t}$  form a time-dependent Markov semigroup with generators  $(\mathbf{L}_t)$ , in the sense that*

$$(3.24) \quad \mathbf{P}_{t,t} = \text{id} \quad \text{and} \quad \mathbf{P}_{r,t} \mathbf{P}_{s,r} = \mathbf{P}_{s,t} \quad \text{for all } s \leq r \leq t,$$

and  $\mathbf{P}_{s,t} F \geq 0$  if  $F \geq 0$  with  $\mathbf{P}_{s,t} 1 = 1$ .

Furthermore for all  $t$  at which  $C_t$  is differentiable (respectively  $s$  at which  $C_s$  is differentiable),

$$(3.25) \quad \frac{\partial}{\partial t} \mathbf{P}_{s,t} F = \mathbf{L}_t \mathbf{P}_{s,t} F, \quad -\frac{\partial}{\partial s} \mathbf{P}_{s,t} F = \mathbf{P}_{s,t} \mathbf{L}_s F, \quad (s \leq t),$$

for all smooth functions  $F$ , where  $\mathbf{L}_t$  acts on a smooth function  $F$  by

$$(3.26) \quad \mathbf{L}_t F = \frac{1}{2} \Delta_{\dot{C}_t} F - (\nabla V_t, \nabla F)_{\dot{C}_t}.$$

The measures  $\nu_t$  evolve dual to  $(\mathbf{P}_{s,t})$  in the sense that

$$(3.27) \quad \mathbb{E}_{\nu_t} [\mathbf{P}_{s,t} F] = \mathbb{E}_{\nu_s} [F] \quad (s \leq t), \quad -\frac{\partial}{\partial t} \mathbb{E}_{\nu_t} [F] = \mathbb{E}_{\nu_t} [\mathbf{L}_t F].$$

*Proof.* By assumption,  $V_0$  is bounded below. The weak convergence of the Gaussian measure  $\mathbb{P}_{C_t - C_s}$  to the Dirac measure at 0 when  $t \downarrow s$  thus implies  $\mathbf{P}_{t,t} = \text{id}$ . The semi-group property, i.e.  $\mathbf{P}_{r,t} \mathbf{P}_{s,r} = \mathbf{P}_{s,t}$  for any  $s \leq r \leq t$ , then follows from (3.20). The definition (3.12) also implies continuity since  $\|\mathbf{P}_{s,t} F\|_\infty \leq \|F\|_\infty$  for each bounded  $F$ . Equation (3.12) also implies that  $\mathbf{P}_{s,t} F \geq 0$  if  $F \geq 0$ .

To verify that the generator  $\mathbf{L}_t$  of the Polchinski semigroup is given by (3.26), set for  $s < t$ :

$$(3.28) \quad F_{s,t}(\varphi) = \mathbf{P}_{s,t} F(\varphi) = e^{V_t(\varphi)} \mathbb{E}_{C_t - C_s} [e^{-V_s(\varphi + \zeta)} F(\varphi + \zeta)].$$

Computing the time derivatives using Propositions 3.1 and 3.4, this leads to

$$(3.29) \quad \begin{aligned} \frac{\partial}{\partial t} F_{s,t} &= \left( \frac{\partial}{\partial t} V_t \right) F_{s,t} + e^{V_t} \frac{1}{2} \Delta_{\dot{C}_t} \mathbb{E}_{C_t - C_s} [e^{-V_s(\cdot + \zeta)} F(\cdot + \zeta)] \\ &= \left( \frac{\partial}{\partial t} V_t \right) F_{s,t} + e^{V_t} \frac{1}{2} \Delta_{\dot{C}_t} (e^{-V_t} F_{s,t}) \\ &= \left( \frac{\partial}{\partial t} V_t \right) F_{s,t} - \left( \frac{1}{2} \Delta_{\dot{C}_t} V_t \right) F_{s,t} + \frac{1}{2} (\nabla V_t)_{\dot{C}_t}^2 F_{s,t} + \frac{1}{2} \Delta_{\dot{C}_t} F_{s,t} - (\nabla V_t, \nabla F_{s,t})_{\dot{C}_t} \\ &= \frac{1}{2} \Delta_{\dot{C}_t} F_{s,t} - (\nabla V_t, \nabla F_{s,t})_{\dot{C}_t} \\ &= \mathbf{L}_t F_{s,t}, \end{aligned}$$

which is the first equality in (3.25). The second equality in (3.25) follows analogously.

The first equality in (3.27) holds as in Proposition 3.3. The second identity follows by taking derivatives in  $s$  in the first identity and then using (3.25) so that

$$(3.30) \quad \frac{\partial}{\partial s} \mathbb{E}_{\nu_s} [F] = \frac{\partial}{\partial s} \mathbb{E}_{\nu_t} [\mathbf{P}_{s,t} F] = \mathbb{E}_{\nu_t} \left[ \frac{\partial}{\partial s} \mathbf{P}_{s,t} F \right] = -\mathbb{E}_{\nu_t} [\mathbf{P}_{s,t} \mathbf{L}_s F].$$

For  $t = s$  then  $\mathbf{P}_{s,s} \mathbf{L}_s F = \mathbf{L}_s F$  and the second identity in (3.27) is recovered.  $\square$

The operator  $\mathbf{L}_t$  in (3.26) is obtained by linearising the Polchinski equation (3.21) and has a structure similar to the generator  $\Delta^H$  defined in (2.10).

**3.3. Log-Sobolev inequality via a multiscale Bakry–Émery method.** In this section, the Polchinski renormalisation is used to derive a log-Sobolev inequality under a criterion on the renormalised potentials, which can be interpreted as a multiscale condition generalising the strict convexity of the Hamiltonian in the Bakry–Émery criterion (Theorem 2.8).

We impose the following technical *continuity assumption* analogous to (2.18): for all bounded smooth functions  $F : X \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$(3.31) \quad \lim_{t \rightarrow \infty} \mathbb{E}_{\nu_t} [g(\mathbf{P}_{0,t} F)] = g(\mathbb{E}_{\nu_0} [F]).$$

This can be easily checked in all examples of practical interest.

**Theorem 3.6.** *Consider a measure  $\nu_0$  of the form (3.10) associated with a covariance decomposition  $\dot{C}_t$  differentiable for all  $t$  (see Section 3.2), and assume also (3.31).*

*Suppose there are real numbers  $\dot{\lambda}_t$  (allowed to be negative) such that*

$$(3.32) \quad \forall \varphi \in X, t > 0 : \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t - \frac{1}{2} \ddot{C}_t \geq \dot{\lambda}_t \dot{C}_t,$$

and define

$$(3.33) \quad \lambda_t = \int_0^t \dot{\lambda}_s ds, \quad \frac{1}{\gamma} = \int_0^\infty e^{-2\lambda_t} dt.$$

Then  $\nu_0$  satisfies the log-Sobolev inequality

$$(3.34) \quad \text{Ent}_{\nu_0} [F] \leq \frac{2}{\gamma} \mathbb{E}_{\nu_0} \left[ (\nabla \sqrt{F})_{\dot{C}_0}^2 \right].$$

Contrary to the Bakry–Émery criterion (Theorem 2.8), the initial potential  $V_0$  is not required to be convex. The relevant parameter is an integrated estimate (3.33) on the Hessian of the renormalised potentials  $V_t$ . Thus if one can prove that the renormalisation flow improves the non-convexity of the original potential so that the integral in (3.33) is finite, then the log-Sobolev inequality holds. In the case of convex potential  $V_0$ , the convexity is preserved by the Polchinski equation (see Proposition 3.13) and the Bakry–Émery criterion can be recovered. In general, the analysis of the renormalised potential  $V_t$  is model dependent.

The covariances  $\dot{C}_t$  play the role of an inverse metric on  $X$ . In our examples of interest, this metric becomes increasingly coarse approximately implementing the “block spin renormalisation picture”. See Section 3.6 for further discussion of this.

**Remark 3.7.** With the same proof, the log-Sobolev inequality (3.34) can be generalised to one for each of the renormalised measures  $\nu_s$ :

$$(3.35) \quad \text{Ent}_{\nu_s}(F) \leq \frac{2}{\gamma_s} \mathbb{E}_{\nu_s} \left[ (\nabla \sqrt{F})_{\dot{C}_s}^2 \right], \quad \frac{1}{\gamma_s} = \int_s^\infty e^{-2(\lambda_u - \lambda_s)} du.$$

**Remark 3.8.** The condition (3.32)–(3.34) is invariant under reparametrisation in  $t$ . For example, if  $a : [0, +\infty] \rightarrow [0, +\infty]$  is a smooth reparametrisation, set

$$(3.36) \quad C_t^a = C_{a(t)}, \quad V_t^a = V_{a(t)}.$$

Then  $\dot{C}_t^a = \dot{a}(t) \dot{C}_{a(t)}$  and  $\ddot{C}_t^a = \ddot{a}(t) \dot{C}_{a(t)} + \dot{a}(t)^2 \ddot{C}_{a(t)}$  and therefore (3.32) is equivalent to

$$(3.37) \quad \dot{C}_t^a \text{Hess } V_t^a \dot{C}_t^a - \frac{1}{2} \ddot{C}_t^a = \dot{a}(t)^2 \left[ \dot{C}_{a(t)} \text{Hess } V_{a(t)} \dot{C}_{a(t)} - \frac{1}{2} \ddot{C}_{a(t)} \right] - \frac{1}{2} \ddot{a}(t) \dot{C}_t^a \geq \dot{\lambda}_t^a \dot{C}_t^a$$

with

$$(3.38) \quad \dot{\lambda}_t^a = \dot{a}(t) \dot{\lambda}_{a(t)} - \frac{1}{2} \frac{\ddot{a}(t)}{\dot{a}(t)} = \dot{a}(t) \dot{\lambda}_{a(t)} - \frac{1}{2} \frac{\partial}{\partial t} \log \dot{a}(t).$$

Thus (3.33) becomes

$$(3.39) \quad \lambda_t^a = \lambda_{a(t)} = \int_0^t \dot{\lambda}_s^a ds = \int_0^t \dot{a}(s) \lambda_{a(s)} ds - \frac{1}{2} \log \frac{\dot{a}(t)}{\dot{a}(0)},$$

and hence

$$(3.40) \quad \begin{aligned} \dot{C}_0^a \int_0^\infty e^{-2\lambda_t^a} dt &= \frac{\dot{C}_0^a}{\dot{a}(0)} \int_0^\infty e^{-2 \int_0^t \lambda_{a(s)} \dot{a}(s) ds} \dot{a}(t) dt \\ &= \dot{C}_0 \int_0^\infty e^{-2 \int_0^u \lambda_u du} du = \dot{C}_0 \int_0^\infty e^{-2\lambda_u} du. \end{aligned}$$

Analogously, one can parametrise by  $[0, T]$  instead of  $[0, +\infty]$ , i.e., use a covariance decomposition  $C = \int_0^T \dot{C}_t dt$ , and then obtain the same conclusion with  $T$  instead of  $\infty$  in the estimates.

**Remark 3.9.** For a covariance decomposition such that  $\dot{C}_t$  is not differentiable for all  $t$ , an alternative criterion that does not involve  $\dot{C}_t$  can be formulated, see [13, Theorem 2.6].

*Proof of Theorem 3.6.* The proof follows the strategy of the Bakry–Émery theorem (Theorem 2.8), replacing the Langevin dynamics by the Polchinski flow. We consider a curve of probability measures  $(\nu_t)_{t \geq 0}$  and a corresponding dual time-dependent Markov semigroup  $(\mathbf{P}_{s,t})$  with generators  $(\mathbf{L}_t)$  as in Proposition 3.5.

For  $F : X \rightarrow \mathbb{R}$  a function with values in a compact subset  $I$  of  $(0, \infty)$ , we write  $F_t = \mathbf{P}_{0,t} F \in I$ . Since the function  $\Phi$  is smooth on  $I$ , it can be extended to a bounded smooth function on  $\mathbb{R}$  and we deduce from (3.31) that

$$(3.41) \quad \lim_{t \rightarrow \infty} \mathbb{E}_{\nu_t} [\Phi(\mathbf{P}_{0,t} F)] = \Phi(\mathbb{E}_{\nu_0} [F]).$$

Thus

$$(3.42) \quad \text{Ent}_{\nu_0}(F) = \mathbb{E}_{\nu_0}[\Phi(F)] - \Phi(\mathbb{E}_{\nu_0}[F]) = - \int_0^\infty dt \frac{\partial}{\partial t} \mathbb{E}_{\nu_t}[\Phi(F_t)].$$

It remains to prove the counterpart of the de Bruin Formula (2.31). Denoting  $\dot{F}_t = \frac{\partial}{\partial t} F_t$ , using first (3.27) and then (3.26), it follows that

$$(3.43) \quad \begin{aligned} -\frac{\partial}{\partial t} \mathbb{E}_{\nu_t}[\Phi(F_t)] &= \mathbb{E}_{\nu_t} \left[ \mathbf{L}_t(\Phi(F_t)) - \Phi'(F_t) \dot{F}_t \right] \\ &= \mathbb{E}_{\nu_t} \left[ \Phi'(F_t) \mathbf{L}_t F_t + \Phi''(F_t) \frac{1}{2} (\nabla F_t)_{\dot{C}_t}^2 - \Phi'(F_t) \dot{F}_t \right] \\ &= \frac{1}{2} \mathbb{E}_{\nu_t} \left[ \Phi''(F_t) (\nabla F_t)_{\dot{C}_t}^2 \right] = 2 \mathbb{E}_{\nu_t} \left[ (\nabla \sqrt{F_t})_{\dot{C}_t}^2 \right]. \end{aligned}$$

Integrating this relation using (3.42) gives

$$(3.44) \quad \text{Ent}_{\nu_0}(F) = 2 \int_0^\infty \mathbb{E}_{\nu_t} \left[ (\nabla \sqrt{\mathbf{P}_{0,t} F})_{\dot{C}_t}^2 \right] dt.$$

The above entropy production formula (3.43) is analogous to the de Bruin identity (2.31) and the entropy decomposition to (2.45), but an important difference is that the reference measure  $\nu_t$  here changes as well. In Section 2, we used that  $\mathbb{E}_\nu[\Delta^H F] = 0$  for any  $F$  in the derivation of the de Bruin identity, but more conceptually what we used is that the measure  $\nu$  satisfies

$$(3.45) \quad -\frac{\partial}{\partial t} \mathbb{E}_\nu[\cdot] = \mathbb{E}_\nu[\Delta^H(\cdot)],$$

since both sides are 0 (because the stationary measure  $\nu$  does not depend on  $t$ ). In the computation above, both  $\nu_t$  and  $F_t$  vary with  $t$ , but in a dual way, and the analogue of (3.45) is (3.15).

It remains to derive the counterpart of (2.46) and show that

$$(3.46) \quad \forall t \geq 0, \quad (\nabla \sqrt{\mathbf{P}_{0,t}F})_{\dot{C}_t}^2 \leq e^{-2\lambda t} \mathbf{P}_{0,t} \left[ (\nabla \sqrt{F})_{\dot{C}_0}^2 \right].$$

Plugging this relation in (3.44) and recalling that  $\mathbb{E}_{\nu_t} \left[ \mathbf{P}_{0,t} \left[ (\nabla \sqrt{F})_{\dot{C}_0}^2 \right] \right] = \mathbb{E}_{\nu_0} \left[ (\nabla \sqrt{F})_{\dot{C}_0}^2 \right]$ , the log-Sobolev inequality (3.34) is recovered.

We turn now to the proof of (3.46). The following lemma is essentially the Bakry–Émery argument adapted to the Polchinski flow.

**Lemma 3.10.** *Let  $\mathbf{L}_t$ ,  $\mathbf{P}_{0,t}$ ,  $\dot{C}_t$ ,  $V_t$  be as in Section 3.2. Then the following identity holds for any  $t$ -independent positive definite matrix  $Q$ :*

$$(3.47) \quad (\mathbf{L}_t - \partial_t)(\nabla \sqrt{\mathbf{P}_{0,t}F})_Q^2 = 2(\nabla \sqrt{\mathbf{P}_{0,t}F}, \text{Hess } V_t \dot{C}_t \nabla \sqrt{\mathbf{P}_{0,t}F})_Q + \frac{1}{4}(\mathbf{P}_{0,t}F) |\dot{C}_t^{1/2} (\text{Hess log } \mathbf{P}_{0,t}F) Q^{1/2}|_2^2,$$

where  $|M|_2^2 = \sum_{p,q} |M_{pq}|^2$  denotes the squared Frobenius norm of a matrix  $M = (M_{pq})$ .

The derivation of the lemma is postponed. Applying it with  $Q = \dot{C}_t$  implies

$$(3.48) \quad (\mathbf{L}_s - \partial_s)(\nabla \sqrt{\mathbf{P}_{0,s}F})_{\dot{C}_s}^2 = 2(\nabla \sqrt{\mathbf{P}_{0,s}F}, \text{Hess } V_s \dot{C}_s \nabla \sqrt{\mathbf{P}_{0,s}F})_{\dot{C}_s} - (\nabla \sqrt{\mathbf{P}_{0,s}F})_{\dot{C}_s}^2 + \frac{1}{4}(\mathbf{P}_{0,s}F) |\dot{C}_s^{1/2} (\text{Hess log } \mathbf{P}_{0,s}F) \dot{C}_s^{1/2}|_2^2.$$

By the assumption (3.32) and since the last term is positive, it follows that

$$(3.49) \quad (\mathbf{L}_s - \partial_s)(\nabla \sqrt{\mathbf{P}_{0,s}F})_{\dot{C}_s}^2 \geq 2\dot{\lambda}_s (\nabla \sqrt{\mathbf{P}_{0,s}F})_{\dot{C}_s}^2.$$

Equivalently,  $\psi(s) := e^{-2\lambda t + 2\lambda_s} \mathbf{P}_{s,t} \left[ (\nabla \sqrt{\mathbf{P}_{0,s}F})_{\dot{C}_s}^2 \right]$  satisfies  $\psi'(s) \leq 0$  for  $s < t$ . This implies  $\psi(t) \leq \psi(0)$  so that (3.46) holds.  $\square$

At first sight, the proof may seem mysterious, but the idea is simply to iterate the entropy decomposition (2.52) by using the Polchinski flow to decompose the measure into its scales. To illustrate this, let us consider a discrete decomposition of the entropy using the Polchinski flow. Given  $\delta > 0$  and the sequence  $(t_i = i\delta)_{i \geq 0}$ , one has

$$(3.50) \quad \begin{aligned} \text{Ent}_{\nu_0}(F) &= \mathbb{E}_{\nu_0}[\Phi(F)] - \Phi(\mathbb{E}_{\nu_0}[F]) \\ &= \sum_i \mathbb{E}_{\nu_{t_i}}[\Phi(\mathbf{P}_{0,t_i}(F))] - \mathbb{E}_{\nu_{t_{i+1}}}[\Phi(\mathbf{P}_{0,t_{i+1}}(F))] \\ &= \sum_i \mathbb{E}_{\nu_{t_{i+1}}}[\mathbf{P}_{t_i,t_{i+1}} \Phi(\mathbf{P}_{0,t_i}(F)) - \Phi(\mathbf{P}_{0,t_{i+1}}(F))] \\ &= \sum_i \mathbb{E}_{\nu_{t_{i+1}}}[\text{Ent}_{\mathbf{P}_{t_i,t_{i+1}}}(\mathbf{P}_{0,t_i}(F))]. \end{aligned}$$

The measure  $\mathbf{P}_{t_i,t_{i+1}}$  associated with a small increment satisfies a log-Sobolev inequality as the associated Gaussian covariance  $C_{t_{i+1}} - C_{t_i}$  is tiny for  $\delta$  small (so that the Hamiltonian corresponding to the measure  $\mathbf{P}_{t_i,t_{i+1}}$  is extremely convex). This suggests that for each interval  $[t_i, t_{i+1}]$ , one can reduce to estimating  $\mathbb{E}_{\nu_{t_{i+1}}} [(\nabla \sqrt{\mathbf{P}_{0,t_i}(F)})_{\delta \dot{C}_{t_i}}]$  and the delicate issue is then to interchange  $\nabla$  and  $\mathbf{P}_{0,t_i}$  (note that a similar step already occurred even in the product case (2.59)). Such a discrete decomposition was implemented in [12] to derive a spectral gap for certain models. The proof of Theorem 3.6 relies on the limit  $\delta$  tends to 0 which greatly simplifies the argument as the analytic structure of the Polchinski flow kicks in.

*Proof of Lemma 3.10.* For a more detailed proof, see [13, Lemma 2.8]. One can first verify the so-called ‘Bochner formula’:

$$(3.51) \quad (\mathbf{L}_t - \partial_t)(\nabla \mathbf{P}_{0,t} F)_Q^2 = 2(\nabla \mathbf{P}_{0,t} F, \text{Hess } V_t \dot{C}_t \nabla \mathbf{P}_{0,t} F)_Q + |\dot{C}_t^{1/2} \text{Hess } \mathbf{P}_{0,t} F Q^{1/2}|_2^2.$$

The claim (3.47) then follows: writing  $F$  instead of  $\mathbf{P}_{0,t} F$  for short, dropping other  $t$ -subscripts, (3.52)

$$(3.52) \quad (\mathbf{L}_t - \partial_t)(\nabla \sqrt{F})_Q^2 = \frac{(\mathbf{L}_t - \partial_t)(\nabla F)_Q^2}{4F} - \frac{(\nabla F)_Q^2 (\mathbf{L}_t - \partial_t)F}{4F^2} - \frac{(\nabla(\nabla F)_Q^2, \nabla F)_{\dot{C}}}{4F^2} + \frac{(\nabla F)_Q^2 (\nabla F)_{\dot{C}}^2}{4F^3}.$$

Using  $(\mathbf{L}_t - \partial_t)F = 0$  and (3.51) the right-hand side equals that in (3.47) since

$$(3.53) \quad F |\dot{C}_t^{1/2} \text{Hess } \log F Q^{1/2}|_2^2 = \frac{|\dot{C}_t^{1/2} \text{Hess } F Q^{1/2}|_2^2}{F} - \frac{(\nabla(\nabla F)_Q^2, \nabla F)_{\dot{C}}}{F^2} + \frac{(\nabla F)_Q^2 (\nabla F)_{\dot{C}}^2}{F^3}.$$

To see this, observe that the left-hand side is (with summation convention)

$$(3.54) \quad F \dot{C}_{ij} Q_{kl} (\text{Hess } \log F)_{ik} (\text{Hess } \log F)_{jl} = F \dot{C}_{ij} Q_{kl} \left( \frac{F_{ik}}{F} - \frac{F_i F_k}{F^2} \right) \left( \frac{F_{jl}}{F} - \frac{F_j F_l}{F^2} \right),$$

and the right-hand side is

$$(3.55) \quad \dot{C}_{ij} Q_{kl} \left[ \frac{F_{ik} F_{jl}}{F} - \frac{(F_k F_l)_i F_j}{F^2} + \frac{F_i F_j F_k F_l}{F^3} \right].$$

So both are indeed equal.  $\square$

**3.4. Derivatives of the renormalised potential.** Checking the multiscale assumption in Theorem 3.6 boils down to controlling the Hessian of the renormalised potential  $V_t$ . For a well chosen covariance decomposition, the structure of the potential  $V_t$  is often expected to improve along the flow of the Polchinski equation (3.21). In particular, one may hope that  $V_t$  becomes more convex. This is illustrated in the Example 3.5 below which considers the case of a single variable. However, for a given microscopic model the convexification can be extremely difficult to check. Some examples where it is possible are discussed in Section 6.

Even though the analysis of the derivatives of  $V_t$  is model dependent, we state a few general identities for these derivatives which will be used later.

**Lemma 3.11.** *Let  $U_t = \nabla V_t$  and  $H_t = \text{Hess } V_t$ . Then*

$$(3.56) \quad \partial_t U_t = \mathbf{L}_t U_t, \quad \partial_t H_t = \mathbf{L}_t H_t - H_t \dot{C}_t H_t.$$

Moreover, for all  $f \in X$  and  $t \geq s \geq 0$ ,

$$(3.57) \quad (f, \nabla V_t) = \mathbf{P}_{s,t}(f, \nabla V_s),$$

$$(3.58) \quad (f, \text{Hess } V_t f) = \mathbf{P}_{s,t}(f, \text{Hess } V_s f) - \left[ \mathbf{P}_{s,t}((f, \nabla V_s)^2) - (\mathbf{P}_{s,t}(f, \nabla V_s))^2 \right].$$

*Proof.* (3.56) follows by differentiating (3.21).

To recover (3.57), we recall from (3.11) that  $V_t(\varphi) = -\log \mathbf{E}_{C_t - C_s} [e^{-V_s(\varphi + \zeta)}]$ . Identity (3.57) follows by differentiating and then identifying  $\mathbf{P}_{s,t}$  by (3.12)

$$(3.59) \quad \nabla V_t(\varphi) = \frac{\mathbf{E}_{C_t - C_s} [e^{-V_s(\varphi + \zeta)} \nabla V_s(\varphi + \zeta)]}{\mathbf{E}_{C_t - C_s} [e^{-V_s(\varphi + \zeta)}]} = \mathbf{P}_{s,t}(\nabla V_s)(\varphi).$$

Identity (3.58) can then be obtained by taking an additional derivative in the previous expression.  $\square$

Alternatively, one can rewrite the derivatives of the renormalised potential in terms of the fluctuation measure  $\mu_t^\varphi$  introduced in (3.18).

**Lemma 3.12.** *The first derivative of the renormalised potential is related to an expectation*

$$(3.60) \quad \nabla V_t(\varphi) = \mathbb{E}_{\mu_t^\varphi}[\nabla V_0(\zeta)] = \mathbb{E}_{\mu_t^\varphi}[C_t^{-1}(\varphi - \zeta)].$$

*The second derivative is encoded by a variance under the fluctuation measure*

$$(3.61) \quad \forall f \in X : \quad (f, \text{Hess } V_t(\varphi)f) = \mathbb{E}_{\mu_t^\varphi}[(f, \text{Hess } V_0(\zeta)f)] - \text{Var}_{\mu_t^\varphi}((f, \nabla V_0(\zeta))) \\ = (f, C_t^{-1}f) - \text{Var}_{\mu_t^\varphi}((C_t^{-1}f, \zeta)),$$

where the second equalities hold if  $C_t$  is invertible.

*Proof.* The first part of (3.60) follows from the identity  $\nabla V_t = \mathbf{P}_{0,t}(\nabla V_0)$  obtained in (3.57) and the identification of the fluctuation measure  $\mu_t^\varphi$  with the semigroup  $\mathbf{P}_{0,t}$  in (3.17). The second equality is obtained by an integration by parts using the form (3.18) of the fluctuation measure.

In the same way the first equality in (3.61) is deduced from (3.58) by identifying  $\mathbf{P}_{0,t}$  and  $\mu_t^\varphi$ . The second equality follows by differentiating  $\varphi \mapsto \mathbb{E}_{\mu_t^\varphi}[C_t^{-1}(\varphi - \zeta)]$  and using (3.18).  $\square$

We consider the case of convex potentials and show that they remain convex along the Polchinski flow.

**Proposition 3.13.** *Assume that  $V_0$  is convex. Then  $V_t$  is convex for all  $t \geq 0$ .*

Thus the standard Bakry–Émery criterion can be recovered from Theorem 3.6: if  $\text{Hess } H \geq \lambda \text{id}$  one can choose  $A = \lambda \text{id}$  and  $\text{Hess } V_0 \geq 0$  and Theorem 3.6 guarantees the log-Sobolev inequality with constant

$$(3.62) \quad \frac{1}{\gamma} \leq \int_0^\infty e^{-\lambda t} dt = \frac{1}{\lambda}.$$

This follows from criterion (3.32) applied with  $\dot{C}_t = e^{-\lambda t} \text{id}$  so that  $\ddot{C}_t = -\lambda e^{-\lambda t} \text{id}$  and  $\dot{\lambda}_t = \lambda/2$ . Note that the decomposition  $\dot{C}_t = \text{id}$  on  $[0, T]$  with  $T = \frac{1}{\lambda}$  (see Remark 3.8) could have been used instead.

*Proof 1.* If  $V_0$  is convex, then  $e^{-V_t(\varphi)}$  is the marginal of the measure  $\propto e^{-V_0(\varphi+\zeta)} \mathbf{P}_{C_t}(d\zeta)$ , with density log-concave in  $(\zeta, \varphi)$ . A theorem of Prékopa then implies that  $V_t$  is convex. It is also possible to directly compute the Hessian: the Brascamp–Lieb inequality [23, Theorem 4.1] states that if a probability measure  $\propto e^{-H}$  has strictly convex potential  $H$  then

$$(3.63) \quad \text{Var}(F) \leq \mathbb{E}[(\nabla F, (\text{Hess } H)^{-1} \nabla F)].$$

Thus by the first identity in (3.61) and then applying the Brascamp–Lieb inequality to estimate the variance, we get

$$(3.64) \quad \text{Hess } V_t(\varphi) = \mathbb{E}_{\mu_t^\varphi}[\text{Hess } V_0(\zeta)] - \text{Var}_{\mu_t^\varphi}(\nabla V_0(\zeta)) \\ \geq \mathbb{E}_{\mu_t^\varphi} \left[ \text{Hess } V_0(\zeta) - \text{Hess } V_0(\zeta)(C_t^{-1} + \text{Hess } V_0(\zeta))^{-1} \text{Hess } V_0(\zeta) \right] \\ = \mathbb{E}_{\mu_t^\varphi} \left[ \text{Hess } V_0(\zeta)(C_t^{-1} + \text{Hess } V_0(\zeta))^{-1} C_t^{-1} \right] \\ = \mathbb{E}_{\mu_t^\varphi} \left[ C_t^{-1/2} C_t^{1/2} \text{Hess } V_0(\zeta) C_t^{1/2} (\text{id} + C_t^{1/2} \text{Hess } V_0(\zeta) C_t^{1/2})^{-1} C_t^{-1/2} \right].$$

Therefore, with  $\hat{H}_t = C_t^{1/2} \text{Hess } V_0 C_t^{1/2} \geq 0$ ,

$$(3.65) \quad C_t^{1/2} \text{Hess } V_t(\varphi) C_t^{1/2} \geq \mathbb{E}_{\mu_t^\varphi} \left[ \frac{\hat{H}_t(\zeta)}{\text{id} + \hat{H}_t(\zeta)} \right] \geq 0. \quad \square$$



*Proof 2 from [55, Theorem 9.1], [30, Theorem 3.3].* This alternative approach puts the emphasis on the PDE structure associated with the renormalised potential by application of the maximum principle. We give the gist of the proof and refer to [30, Theorem 3.3, page 129] for a complete argument. Let  $H_t = \text{Hess } V_t$  with  $H_0 > 0$ , and recall (3.56):

$$(3.66) \quad \frac{\partial H_t}{\partial t} = \mathbf{L}_t H_t - H_t \dot{C}_t H_t.$$

Now assume there is a first time  $t_0 > 0$  and  $\varphi_0 \in X$  such that  $H_{t_0}(\varphi_0)$  has a 0 eigenvalue with eigenvector  $v_0$ , i.e.,  $H_{t_0}(\varphi_0)v_0 = 0$ . Define  $f_t(\varphi) = (v_0, H_t(\varphi)v_0)$ . Therefore

$$(3.67) \quad \frac{\partial f_{t_0}(\varphi_0)}{\partial t} = \mathbf{L}_{t_0} f_{t_0}(\varphi_0) - (v_0, H_{t_0}(\varphi_0) \dot{C}_{t_0} H_{t_0}(\varphi_0) v_0) \geq 0,$$

where we used that  $f_{t_0}(\varphi)$  is minimum at  $\varphi_0$  so that  $\mathbf{L}_{t_0} f_{t_0}(\varphi) = \frac{1}{2} \Delta_{\dot{C}_{t_0}} f_{t_0}(\varphi) \geq 0$  (by the maximum principle) and that by construction  $(v_0, H_{t_0}(\varphi_0) \dot{C}_{t_0} H_{t_0}(\varphi_0) v_0) = 0$ . This shows that  $f_t(\varphi_0)$  cannot cross 0 after  $t_0$ . A more careful argument involves regularisation, see [30, Theorem 3.3].  $\square$

We end this section with a rescaling property of the Polchinski equation.

**Example 3.14.** Similarly to (3.14), write the renormalised potential as

$$(3.68) \quad V_t(\varphi) = \frac{1}{2}(\varphi, C_t^{-1}\varphi) + F_t(C_t^{-1}\varphi),$$

where  $F_t(h) = V_t(0) - \log \mathbf{E}_{C_t}[e^{-V_0(\zeta) + (h, \zeta)}]$  is the normalised log partition function of the fluctuation measure at external field  $h$ . Then the Polchinski equation for  $V$  is equivalent to a different Polchinski equation for  $F$ :

$$(3.69) \quad \frac{\partial}{\partial t} F_t = \frac{1}{2} \Delta_{\dot{\Sigma}_t} F_t - \frac{1}{2} (\nabla F_t)_{\dot{\Sigma}_t}^2 + \text{Tr}(C_t^{-1} \dot{C}_t), \quad \text{where } \dot{\Sigma}_t = C_t^{-1} \dot{C}_t C_t^{-1}.$$

Note that  $\text{Tr}(C_t^{-1} \dot{C}_t)$  is only a constant. Indeed,  $F_t(h) = V_t(C_t h) - \frac{1}{2}(h, C_t h)$  and thus

$$(3.70) \quad \nabla F_t = C_t \nabla V_t - C_t h, \quad \Delta_{\dot{\Sigma}_t} F_t = \Delta_{\dot{C}_t} V_t - \text{Tr}(C_t^{-1} \dot{C}_t),$$

and

$$(3.71) \quad \begin{aligned} \frac{\partial}{\partial t} F_t &= \frac{1}{2} \Delta_{\dot{C}_t} V_t - \frac{1}{2} (\nabla V_t)_{\dot{C}_t}^2 + (\nabla V_t, \dot{C}_t h) - \frac{1}{2} (h, \dot{C}_t h) \\ &= \frac{1}{2} \Delta_{\dot{C}_t} V_t - \frac{1}{2} (h - \nabla V_t)_{\dot{C}_t}^2 \\ &= \frac{1}{2} \Delta_{\dot{\Sigma}_t} F_t - \frac{1}{2} (\nabla F_t)_{\dot{\Sigma}_t}^2 + \text{Tr}(C_t^{-1} \dot{C}_t). \end{aligned}$$

**3.5. Example: Convexification along the Polchinski flow for one variable.** The aim of this section is to illustrate the claims that the renormalised measure becomes progressively simpler and convex along the Polchinski flow using a simple one variable example. Let  $H : \mathbb{R} \rightarrow \mathbb{R}$  be a  $C^2$  potential that is strictly convex outside of a segment:  $\inf_{|x| \geq M} H''(x) \geq c > 1$  for some  $c, M > 0$ , but assume that  $\inf_{\mathbb{R}} H'' < 0$ , and consider the measure:

$$(3.72) \quad \nu_0(d\varphi) \propto e^{-H(\varphi)} d\varphi \propto \exp \left[ -\frac{\varphi^2}{2} - V_0(\varphi) \right] d\varphi, \quad V_0(\varphi) := H(\varphi) - \frac{\varphi^2}{2}.$$

In (3.72), the Gaussian part  $\frac{1}{2}\varphi^2$  is singled out to define the Polchinski flow, but this is just a convention up to redefining  $V_0$ .

By assumption,  $\nu_0$  is not log-concave, and the Bakry–Émery criterion (Theorem 2.8) does not apply. Let us stress that there are many ways to obtain a log-Sobolev inequality for the above

measure. Our goal, however, is to exemplify that, using the Polchinski flow, how one can still use a convexity-based argument, the multiscale Bakry–Émery criterion of Theorem 3.6, by relying on the convexity of the renormalised measures  $\nu_t$  that will be more log-concave than  $\nu_0$ .

The Polchinski flow is defined in terms of a covariance decomposition, which is supposed to decompose the Gaussian part of  $\nu_0$  into contributions from different scales. In the statistical mechanics examples discussed in Sections 2.6–2.7, the notion of scale was linked with the geometry of the underlying lattice (e.g., small scales corresponding to information pertaining to spins at small lattice distance). In the single variable case, there is no geometry, thus the Gaussian part does not have any structure. The only meaningful decomposition, written here on  $[0, 1]$  instead of  $[0, \infty)$  for convenience, is therefore:

$$(3.73) \quad C_t := t \text{id}, \quad \dot{C}_t = \text{id}, \quad (t \in [0, 1]).$$

The corresponding renormalised potential reads:

$$(3.74) \quad e^{-V_t(\varphi)} = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} \exp \left[ -\frac{\zeta^2}{2t} - V_0(\zeta + \varphi) \right] d\zeta,$$

and the renormalised measure  $\nu_t$  defined in (3.13) and fluctuation measure  $\mu_t^\varphi$  defined in (3.18) are respectively given by:

$$(3.75) \quad \nu_t(d\varphi) \propto \exp \left[ -\frac{\varphi^2}{2(1-t)} - V_t(\varphi) \right] d\varphi, \quad \mu_t^\varphi(d\zeta) \propto \exp \left[ -\frac{\zeta^2}{2t} + \frac{\zeta\varphi}{t} - V_0(\zeta) \right] d\zeta.$$

Note that in terms of the original Hamiltonian  $H(\zeta) = V_0(\zeta) + \frac{\zeta^2}{2}$ , the fluctuation measure  $\mu_t^\varphi$  is more convex than the initial one:

$$(3.76) \quad \mu_t^\varphi(d\zeta) \propto \exp \left[ -\frac{\zeta^2}{2} \left( \frac{1}{t} - 1 \right) + \frac{\zeta\varphi}{t} - H(\zeta) \right] d\zeta.$$

In other words,  $e^{-V_t}$  is the convolution of  $e^{-V_0}$  with the heat kernel on  $\mathbb{R}$  at time  $t$ , and the Polchinski equation becomes the following well-known Hamilton–Jacobi–Bellman equation:

$$(3.77) \quad \partial_t V_t = \frac{1}{2} \partial_\varphi^2 V_t - \frac{1}{2} (\partial_\varphi V_t)^2, \quad (t \in (0, 1)).$$

The motivation for the Polchinski decomposition was that one progressively integrates “small scales” to recover a measure  $\nu_t$  acting on “large scales” that one hopes to be better behaved. In the present case, the only notion of scale refers to the size of fluctuations of the field:

- Even though  $V_0$  may vary a lot on small values of the field, the convolution with the heat kernel at time  $t$  means  $V_t$  is roughly constant on values much smaller than  $\sqrt{t}$ . Thus small details of  $V_0$  are blurred, and  $V_t$  varies more slowly than  $V_0$ . This is the translation to the present case of the general idea that “small scales” (i.e., values below  $\sqrt{t}$ ) have been removed from  $V_t$  and the renormalised potential only sees the “large scales” (values above  $\sqrt{t}$ ).

- Convolution also improves convexity, in the sense that the renormalised measure  $\nu_t$  is more log-concave than  $\nu_0$ . Since  $x \mapsto \frac{1}{2(1-t)} \zeta^2$  becomes increasingly convex as  $t$  approaches 1, proving this statement boils down to proving a lower bound on  $\partial_\varphi^2 V_t$  uniformly on  $t \in [0, 1]$ . Semi-convexity estimates for solutions of the Polchinski equation (3.77) are an active subject of research, connected with optimal transport with entropic regularisation, see, e.g., [29, 32, 33, 43] and references therein and Section 5.1 below. Informally, the convexity of  $V_t$  is given by that of  $V_0$ , plus an “entropic” contribution due to the  $\frac{1}{2t} \zeta^2$  term. In the present simple case, one can directly compute:

$$(3.78) \quad \partial_\varphi^2 V_t(\varphi) = \frac{1}{t} - \frac{1}{t^2} \text{Var}_{\mu_t^\varphi}(\zeta), \quad (\varphi \in \mathbb{R}).$$

This is an instance of the formula of Lemma 3.12 valid for a general covariance decomposition. It is an example of a general feature of the multiscale Bakry–Émery criterion: the log-Sobolev constant, which is not a priori related to spectral properties of the model, can be estimated by lower bounds on  $\partial_\varphi^2 V_t$  which are related to variance bounds, i.e., spectral information.

**Exercise 3.15.** Using the Brascamp-Lieb inequality (3.63) for  $t < t_0$  with  $t_0^{-1} := -\inf_{\mathbb{R}} V_0''$ , and the fact that  $\mu_t^\varphi$  satisfies a spectral gap inequality with constant  $C$  uniformly in  $\varphi \in \mathbb{R}$  and  $t > 0$ , deduce:

$$(3.79) \quad \partial_\varphi^2 V_t \geq \dot{\lambda}_t := \mathbf{1}_{[0, t_0/2]}(t) \left( \frac{-1}{t_0 - t} \right) + \mathbf{1}_{[t_0/2, 1]}(t) \left( \frac{1}{t} - \frac{C}{t^2} \right), \quad (t \in [0, 1]).$$

The uniform lower bound (3.79) confirms that  $\nu_t$  gets more log-concave as  $t$  approaches 1. Injecting the bound (3.79) into the multiscale Bakry–Émery criterion of Theorem 3.6 provides a bound on the log-Sobolev constant  $\gamma$  of  $\nu_0$  in terms on parameter  $t_0$ .

Let us reiterate that one could have obtained a bound on the log-Sobolev constant by standard combination of the usual Bakry–Émery and Holley–Stroock criteria, and here just illustrated that the semi-convexity condition of Theorem 3.6 remains effective in non-convex cases. Theorem 3.6 becomes especially useful in situations with a large state space, where the combination of the Bakry–Émery and Holley–Stroock criteria do not yield dimension-independent bounds on the log-Sobolev constant while effective methods to control the semi-convexity may still exist, see Section 6.

**3.6. Aside: Geometric perspective on the Polchinski flow.** There is a structural resemblance of the renormalisation group flow with geometric flows like the Ricci flow. The matrices  $\dot{C}_t$  take the role of the inverse of a metric  $g_t$  (depending on the flow parameter  $t$ ). We sketch the interpretation of the  $\dot{C}_t$  as a scale-dependent metric on the space of fields, and the natural extension of the above construction in the presence of a non-flat metric.

Suppose  $(X, g)$  is a Riemannian manifold. The metric and its components in coordinates are denoted

$$(3.80) \quad g = (g_{ij}), \quad g^{ij} = (g^{-1})_{ij}, \quad |g| = |\det g|,$$

the volume form is

$$(3.81) \quad m_g(d\varphi) = \sqrt{|g|} d\varphi,$$

and the covariant derivative and Laplace-Beltrami operator are (with summation convention)

$$(3.82) \quad \nabla_g = g^{ij} \partial_j, \quad \operatorname{div}_g U = \frac{1}{\sqrt{|g|}} \partial_i (\sqrt{|g|} U^i), \quad \Delta_g F = \frac{1}{\sqrt{|g|}} \partial_i (\sqrt{|g|} g^{ij} \partial_j F).$$

In particular,

$$(3.83) \quad \operatorname{Hess}_g f = \nabla_g \nabla_g f = g^{ik} \partial_k (g^{jl} \partial_l f),$$

and

$$(3.84) \quad (\nabla_g F)_g^2 = g(\nabla_g F, \nabla_g F) = g^{ij} (\partial_i F) (\partial_j F).$$

For a  $t$ -dependent metric  $g_t$ , the volume form changes according to

$$(3.85) \quad \frac{\partial}{\partial t} dm_{g_t} = \left( \frac{\partial}{\partial t} \log \sqrt{|g_t|} \right) dm_{g_t} = \frac{1}{2} \operatorname{Tr}(g_t^{-1} \dot{g}_t) dm_{g_t}.$$

The Ricci curvature tensor associated with the metric  $g$  is denoted  $\operatorname{Ric}_g$ .

**Remark 3.16.** The notation  $\Delta_{g_t}$  for the Laplace-Beltrami operator is *different* from our previous notation for the covariance-dependent Laplacian  $\Delta_{\dot{C}_t}$  from (3.3). Indeed, the notation differs by an inverse in the index: The Gaussian Laplacian  $\Delta_{\dot{C}_t}$  corresponds to a Laplace-Beltrami operator if  $g_t^{-1} = \dot{C}_t$ . Thus the infinitesimal covariance  $\dot{C}_t$  plays the role of the *inverse* of a metric.

**Example 3.17.** In the above notation, we can reformulate the previous construction as follows: The covariance decomposition is written as

$$(3.86) \quad A^{-1} = \int_0^\infty g_t^{-1} dt.$$

Then the Polchinski equation reads

$$(3.87) \quad \frac{\partial}{\partial t} V_t = \frac{1}{2} \Delta_{g_t} V_t - \frac{1}{2} (\nabla_{g_t} V_t)_{g_t}^2,$$

and the condition (3.32) for the log-Sobolev inequality becomes:

$$(3.88) \quad \text{Hess}_{g_t} V_t + \frac{1}{2} \dot{g}_t \geq \dot{\lambda}_t g_t.$$

**Example 3.18.** Suppose that  $A$  is a Laplace operator on  $\Lambda \subset \mathbb{Z}^d$  and that  $\dot{C}_t = e^{-tA}$  is its heat kernel. Thus the metric  $g_t = e^{+tA}$  is the inverse heat kernel. This means that

$$(3.89) \quad |f|_{g_t} \leq 1 \quad \Leftrightarrow \quad |e^{tA} f| \leq 1,$$

i.e.,  $f = e^{-tA} g$  for some  $|g| \leq 1$  where  $|\cdot|$  denotes the standard Euclidean norm. Therefore the unit ball in the metric  $g_t$  corresponds to elements  $f$  that are obtained by smoothing out elements of the standard unit ball by the heat kernel up to time  $t$ . In this sense, the geometry associated with  $g_t$  implements an approximate block spin picture (in which block averaging has been replaced by convolution with a heat kernel).

We now consider the natural extension to a non-flat metrics. The Laplacian in the presence of a potential  $H$  and metric  $g$  is

$$(3.90) \quad \Delta_g^H = e^H \text{div}_g(e^{-H} \nabla_g F).$$

The analogue of Lemma 3.10 (with  $Q = \dot{C}_t$ ) is as follows.

**Lemma 3.19.**

$$(3.91) \quad (\mathbf{L}_t - \partial_t)(\nabla_g \sqrt{F})_g^2 = (\dot{g} + 2 \text{Hess}_g V_t + \text{Ric}_g)(\nabla_g \sqrt{F}, \nabla_g \sqrt{F}) + \frac{1}{4} |\text{Hess}_g \log F|_g^2.$$

*Proof.* The Bochner formula with a background metric [7, Theorem C.3.3] implies:

$$(3.92) \quad (\mathbf{L}_t - \partial_t)(\nabla_g F)_g^2 = (\dot{g} + 2 \text{Hess}_g V_t + \text{Ric}_g)(\nabla_g F, \nabla_g F) + |\text{Hess}_g F|_g^2.$$

The Bakry-Émery version (with  $\sqrt{F}$  instead of  $F$ ) then follows as in the proof of Lemma 3.10.  $\square$

For  $t \in [0, T)$  with  $T \in (0, +\infty]$ , assume that  $g_t$  is a given  $t$ -dependent metric and that  $Z_t dm_{g_t}$  evolves according to the associated backward heat equation:

$$(3.93) \quad \frac{\partial}{\partial t} Z_t dm_{g_t} = \left(-\frac{1}{2} \Delta_{g_t} Z_t\right) dm_{g_t}.$$

The measure  $Z_t dm_{g_t}$  takes the role of the Gaussian measure with covariance  $C_\infty - C_t$ . Assume:

$$(3.94) \quad \frac{\partial}{\partial t} V_t = \frac{1}{2} \Delta_{g_t} V_t - \frac{1}{2} (\nabla_{g_t} V_t)_{g_t}^2$$

$$(3.95) \quad \frac{\partial}{\partial t} F_t = \frac{1}{2} \Delta_{g_t} F_t - (\nabla_{g_t} V_t, \nabla_{g_t} F_t) = \mathbf{L}_t F_t.$$

The last equation defines the semigroup  $\mathbf{P}_{s,t}F$  with generators  $\mathbf{L}_t$ . The renormalised measure  $\nu_t$  is defined by

$$(3.96) \quad \mathbb{E}_{\nu_t}[F] \propto \int F e^{-V_t} \mathbf{Z}_t dm_{g_t}.$$

One can check that

$$(3.97) \quad \frac{\partial}{\partial t} \mathbb{E}_{\nu_t}[\mathbf{P}_{0,t}F] = 0,$$

and that the renormalised measure  $\nu_t$  again evolves in a dual way to  $\mathbf{P}_{s,t}$ : for  $t < T$ ,

$$(3.98) \quad \frac{\partial}{\partial t} \mathbb{E}_{\nu_t}[F] = -\mathbb{E}_{\nu_t}[\mathbf{L}_t F].$$

The analogue of the continuity assumption (3.31) is

$$(3.99) \quad \mathbb{E}_{\nu_t}[g(\mathbf{P}_{0,t}F)] \rightarrow g(\mathbb{E}_{\nu_0}[F]), \quad (t \rightarrow T).$$

Since the evolution of  $\mathbf{Z}_t$  is in general not explicit in the nonflat case, differently from before, this is now an assumption that seems difficult to verify. The same proof as that of Theorem 3.6 using Lemma 3.19 instead of Lemma 3.10 gives the following condition for the log-Sobolev inequality.

**Theorem 3.20.** *Assume that the continuity assumption (3.99) holds. Suppose there are  $\dot{\lambda}_t$  (allowed to be negative) such that*

$$(3.100) \quad \forall \varphi \in X, t > 0 : \quad \text{Hess}_{g_t(\varphi)} V_t(\varphi) + \frac{1}{2} \text{Ric}_{g_t(\varphi)} + \frac{1}{2} \dot{g}_t(\varphi) \geq \dot{\lambda}_t g_t(\varphi),$$

and define

$$(3.101) \quad \lambda_t = \int_0^t \dot{\lambda}_s ds, \quad \frac{1}{\gamma} = \int_0^T e^{-2\lambda_t} dt.$$

Then  $\nu_0$  satisfies the log-Sobolev inequality

$$(3.102) \quad \text{Ent}_{\nu_0}(F) \leq \frac{2}{\gamma} \mathbb{E}_{\nu_0} \left[ (\nabla \sqrt{F})_{g_0}^2 \right].$$

We currently do not know of any interesting applications of the generalised set-up of Theorem 3.20 over that of Theorem 3.6, but it would be very interesting to find some.

Some references with related constructions (though different motivation) in the context of the Ricci flow appeared in [77] and then [65, 70] and more recently [56, 59, 60, 87].

**3.7. Aside: Entropic stability estimate.** An approach different from the Bakry–Émery method to prove (modified) log-Sobolev inequalities, using the same Polchinski flow, is the entropic stability estimate which underlies [28] and has its origins in the spectral and entropic independence conditions introduced in [4] and [3]. In particular, in [28], this method is applied from the stochastic localisation perspective whose equivalence with the Polchinski flow is discussed in Section 4.5. In this section, we rephrase the entropic stability strategy of [28] with the notations of the Polchinski flow to explain the connection with the Bakry–Émery method.

Let us first introduce some notation. For a probability measure  $\mu$  on  $X$  (with all exponential moments) and  $h \in X$  write  $T_h \mu$  for the tilted probability measure:

$$(3.103) \quad \frac{dT_h \mu}{d\mu}(\zeta) = \frac{e^{(h,\zeta)}}{\mathbb{E}_{\mu}[e^{(h,\zeta)}]},$$

and  $\text{Cov}(\mu)$  for the covariance matrix of  $\mu$ . The key estimate is a proof of the entropic stability from a covariance assumption.

**Lemma 3.21** (Entropic stability estimate [28, Lemmas 31 and 40]). *Let  $\mu$  be a probability measure on  $X$ , let  $\dot{\Sigma}$  be a positive semi-definite matrix, and assume there is  $\alpha > 0$  such that*

$$(3.104) \quad \forall h \in X : \quad \dot{\Sigma} \text{Cov}(T_h \mu) \dot{\Sigma} \leq \alpha \dot{\Sigma}.$$

*Then for all nonnegative  $F$  with  $\mathbb{E}_\mu[F] = 1$ ,*

$$(3.105) \quad \frac{1}{2}(\mathbb{E}_\mu[F\zeta] - \mathbb{E}_\mu[\zeta])_{\dot{\Sigma}}^2 \leq \alpha \text{Ent}_\mu(F).$$

*In [28, Definition 29], this inequality is called  $\alpha$ -entropic stability of the measure  $\mu$  with respect to the function  $\psi(x, y) = \frac{1}{2}(x - y)_{\dot{\Sigma}}^2$ .*

The proof of Lemma 3.21 is postponed to the end of this section and we first show how the entropic stability estimate implies a modified log-Sobolev inequality, see [28, Proposition 39].

**Corollary 3.22.** *Assume there are  $\alpha_t > 0$  such that*

$$(3.106) \quad \forall \varphi \in X : \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t - \dot{C}_t C_t^{-1} \dot{C}_t \geq -\alpha_t \dot{C}_t,$$

*or equivalently, with  $\dot{\Sigma}_t = C_t^{-1} \dot{C}_t C_t^{-1}$ ,*

$$(3.107) \quad \forall \varphi \in X : \quad \dot{\Sigma}_t \text{Cov}(\mu_t^\varphi) \dot{\Sigma}_t \leq \alpha_t \dot{\Sigma}_t.$$

*Then the measure  $\mathbf{P}_{0,t}$  satisfies the following  $\alpha_t$ -entropic stability: for all  $\varphi \in X$ ,*

$$(3.108) \quad 2(\nabla \sqrt{\mathbf{P}_{0,t} F})_{\dot{C}_t}^2(\varphi) \leq \alpha_t \text{Ent}_{\mu_t^\varphi}(F) = \alpha_t \left[ \mathbf{P}_{0,t} \Phi(F)(\varphi) - \Phi(\mathbf{P}_{0,t} F(\varphi)) \right].$$

*This implies the following entropy contraction: for any  $s > 0$ ,*

$$(3.109) \quad \text{Ent}_{\nu_0}(F) \leq e^{\int_s^\infty \alpha_u du} \mathbb{E}_{\nu_s} \left[ \text{Ent}_{\mu_s^\varphi}(F) \right].$$

The condition (3.106) on  $\alpha_t$  is very similar to the multiscale Bakry–Émery condition (3.32) on  $\dot{\lambda}_t$ , but different. We recall that (3.32) reads

$$(3.110) \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t - \frac{1}{2} \ddot{C}_t \geq \dot{\lambda}_t \dot{C}_t.$$

The log-Sobolev constant and the very closely related entropy contraction are estimated by the time integrals of  $\alpha_t$  and  $\dot{\lambda}_t$ , respectively. Note that the condition (3.110) is slightly better behaved for  $t$  close 0 as  $C_t$  vanishes as  $t \rightarrow 0$  and  $\dot{C}_t$  does not. To see this, consider the simple one-variable case with the function  $c_t = \int_0^t \dot{c}_u du \in \mathbb{R}$  (with  $\dot{c}_u > 0$ ) and assume that  $\text{Hess } V_t(\varphi) \geq 0$ . Then the optimal choices are  $\dot{\lambda}_t = -\ddot{c}_t \dot{c}_u^{-1}$  and  $\alpha_t = -\dot{c}_u \dot{c}_u^{-1}$  so that

$$(3.111) \quad \int_s^t \dot{\lambda}_u du = \log \dot{c}_s - \log \dot{c}_t \quad \text{and} \quad \int_s^t \alpha_u du = \log c_s - \log c_t.$$

For this reason, one cannot take the limit  $s \rightarrow 0$  in (3.109), and the  $\text{Ent}_{\mu_s^\varphi}$  term in (3.109) must be treated differently for  $s$  small to recover a log-Sobolev inequality. This last step is called *annealing via a localization scheme* in [28]. For example, one can use that the measure  $\mu_s^\varphi$  is simpler for  $s$  small as the covariance  $C_s$  vanishes. This implies that the Hamiltonian of  $\mu_s^\varphi$  is strictly convex uniformly in  $\varphi$  so that by the standard Bakry–Émery criterion (Theorem 2.8), the measure satisfies a log-Sobolev inequality which can then be plugged into (3.109) to complete the log-Sobolev inequality for  $\nu_0$ .

*Proof.* From (3.17), recall that the Polchinski semigroup  $\mathbf{P}_{0,t}$  coincides with the fluctuation measure  $\mu_t^\varphi$ . Thus for smooth  $F > 0$  and  $\varphi \in X$ , one has

$$(3.112) \quad \begin{aligned} \nabla \log \mathbf{P}_{0,t} F(\varphi) &= \nabla \log \mathbb{E}_{\mu_t^\varphi}[F] = \frac{\mathbb{E}_{\mu_t^\varphi}[F C_t^{-1} \zeta]}{\mathbb{E}_{\mu_t^\varphi}[F]} - \mathbb{E}_{\mu_t^\varphi}[C_t^{-1} \zeta] \\ &= C_t^{-1} \left( \mathbb{E}_{\mu_t^{\varphi, F}}[\zeta] - \mathbb{E}_{\mu_t^\varphi}[\zeta] \right), \end{aligned}$$

where the measure modified by  $F$  is defined by

$$(3.113) \quad \frac{d\mu_t^{\varphi, F}}{d\mu_t^\varphi}(\zeta) = \frac{F(\zeta)}{\mathbb{E}_{\mu_t^\varphi}[F]}.$$

Thus, the estimate (3.108) we are looking for boils down to proving an entropic stability result (3.105) for the measure  $\mu_t^\varphi$ . Indeed, setting  $\dot{\Sigma}_t = C_t^{-1} \dot{C}_t C_t^{-1}$ , then

$$(3.114) \quad 2(\nabla \sqrt{\mathbf{P}_{0,t} F})_{\dot{C}_t}^2(\varphi) = \frac{1}{2}(\nabla \log \mathbf{P}_{0,t} F)_{\dot{C}_t}^2(\mathbf{P}_{0,t} F)(\varphi) = \frac{1}{2}(\mathbb{E}_{\mu_t^{\varphi, F}}[\zeta] - \mathbb{E}_{\mu_t^\varphi}[\zeta])_{\dot{\Sigma}_t}^2(\varphi) \mathbb{E}_{\mu_t^\varphi}[F],$$

and the relative entropy is given by

$$(3.115) \quad \text{Ent}_{\mu_t^\varphi}(F) = \mathbb{H}(\mu_t^{\varphi, F} | \mu_t^\varphi) \mathbb{E}_{\mu_t^\varphi}[F].$$

Assumption (3.107) implies the assumption (3.104) on the covariance of  $\mu_t^\varphi$ ,

$$(3.116) \quad \dot{\Sigma}_t^{1/2} \text{Cov}(T_h \mu_t^\varphi) \dot{\Sigma}_t^{1/2} = \dot{\Sigma}_t^{1/2} \text{Cov}(\mu_t^{\varphi + C_t h}) \dot{\Sigma}_t^{1/2} \leq \alpha_t \text{id} \quad \text{for all } h \in X,$$

so that Lemma 3.21 gives the claim (3.108), i.e.,

$$(3.117) \quad 2(\nabla \sqrt{\mathbf{P}_{0,t} F}(\varphi))_{\dot{C}_t}^2 \leq \alpha_t \text{Ent}_{\mu_t^\varphi}(F).$$

Thus it is enough to show that assumption (3.106) is equivalent to the covariance assumption (3.107). From (3.61), we know that for any  $\varphi \in X$ ,

$$(3.118) \quad \text{Hess } V_t(\varphi) = C_t^{-1} - C_t^{-1} \text{Cov}(\mu_t^\varphi) C_t^{-1},$$

so that

$$(3.119) \quad \begin{aligned} -\dot{C}_t \left( \text{Hess } V_t(\varphi) - C_t^{-1} \right) \dot{C}_t &= \dot{C}_t C_t^{-1} \text{Cov}(\mu_t^\varphi) C_t^{-1} \dot{C}_t \\ &= C_t \dot{\Sigma}_t \text{Cov}(\mu_t^\varphi) \dot{\Sigma}_t C_t \leq \alpha_t C_t \dot{\Sigma}_t C_t = \alpha_t \dot{C}_t, \end{aligned}$$

where the inequality holds if and only if  $\dot{\Sigma} \text{Cov}(\mu_t^\varphi) \dot{\Sigma}_t \leq \alpha_t \dot{\Sigma}_t$ .

We turn now to the second part of the claim and show how the entropic stability estimate (3.108) implies the entropy contraction estimate (3.109). The starting point is the time derivative of the entropy (3.43):

$$(3.120) \quad \frac{\partial}{\partial t} \mathbb{E}_{\nu_t} \left[ \mathbf{P}_{0,t} \Phi(F) - \Phi(\mathbf{P}_{0,t} F) \right] = 2 \mathbb{E}_{\nu_t} \left[ (\nabla \sqrt{\mathbf{P}_{0,t} F})_{\dot{C}_t}^2 \right]$$

which is the counterpart of eq. (27) in [28]. This is bounded thanks to (3.108):

$$(3.121) \quad \frac{\partial}{\partial t} \mathbb{E}_{\nu_t} \left[ \mathbf{P}_{0,t} \Phi(F) - \Phi(\mathbf{P}_{0,t} F) \right] \leq \alpha_t \mathbb{E}_{\nu_t} \left[ \mathbf{P}_{0,t} \Phi(F) - \Phi(\mathbf{P}_{0,t} F) \right],$$

and thus for any  $s < t$ , Gronwall's lemma implies

$$(3.122) \quad \mathbb{E}_{\nu_t} \left[ \mathbf{P}_{0,t} \Phi(F) - \Phi(\mathbf{P}_{0,t} F) \right] \leq e^{\int_s^t \alpha_u du} \mathbb{E}_{\nu_s} \left[ \mathbf{P}_{0,s} \Phi(F) - \Phi(\mathbf{P}_{0,s} F) \right].$$

Taking  $t \rightarrow \infty$ , the entropy is recovered from the left-hand side, as in (3.41), and therefore (3.109) holds for any  $s > 0$ .  $\square$

We finally prove Lemma 3.21 following [28].

*Proof of Lemma 3.21.* It suffices to show that, for any  $h \in X$ ,

$$(3.123) \quad \frac{1}{2}(\mathbb{E}_{T_h\mu}[\zeta] - \mathbb{E}_\mu[\zeta])_\Sigma^2 \leq \alpha \mathbb{H}(T_h\mu|\mu) = \alpha \text{Ent}_\mu \left( \frac{e^{(h,\zeta)}}{\mathbb{E}_\mu[e^{(h,\zeta)}]} \right).$$

Indeed, for any density  $F$  with  $\mathbb{E}_\mu[F] = 1$  and  $\mathbb{E}_\mu[F\zeta] = \mathbb{E}_{T_h\mu}[\zeta]$ , the entropy inequality (2.62) applied with the test function  $G : \zeta \mapsto (h, \zeta)$  implies

$$(3.124) \quad \text{Ent}_\mu(F) = \sup_G \left\{ \mathbb{E}_\mu[FG] - \log \mathbb{E}_\mu[e^G] \right\} \geq \mathbb{E}_{T_h\mu}[(h, \zeta)] - \log \mathbb{E}_\mu[e^{(h,\zeta)}] = \mathbb{H}(T_h\mu|\mu),$$

i.e., the relative entropy over probability measures with given mean is minimised by exponential tilts  $T_h\mu$ . Moreover, if there is no  $h$  such that  $\mathbb{E}_\mu[F\zeta] = \mathbb{E}_{T_h\mu}[\zeta]$  the relative entropy is infinite.

From now on, we may assume that  $\text{Cov}(T_h\mu)$  is strictly positive definite on  $X$  for all  $h \in X$ . Indeed, otherwise consider the largest linear subspace  $X'$  such that  $\text{Cov}(T_h\mu)$  acts and is strictly positive definite on  $X'$ , and note that this subspace is independent of  $h$ . Indeed, let  $f \in \mathbb{R}^N$  be such that  $\text{Var}_\mu((f, \zeta)) = 0$ , and assume without loss of generality that  $\mathbb{E}_\mu[(f, \zeta)] = 0$ . Then under the assumption of exponential moments also  $\mathbb{E}_\mu[(f, \zeta)^4] = 0$  and:

$$(3.125) \quad \begin{aligned} \text{Var}_{T_h\mu}((f, \zeta)) &\propto \frac{1}{2} \mathbb{E}_{\mu \otimes \mu}[(f, \zeta - \zeta')^2 e^{(h, \zeta + \zeta')}] \\ &\leq \frac{1}{2} \mathbb{E}_{\mu \otimes \mu}[(f, \zeta - \zeta')^4]^{1/2} \mathbb{E}_{\mu \otimes \mu}[e^{2(h, \zeta + \zeta')}]^{1/2} = 0. \end{aligned}$$

This implies that  $\mu$  (and thus  $T_h\mu$ ) is supported in an affine subspace of  $X$  which is a translation of  $X'$ , and by recentering one can replace  $X$  by  $X'$  in the following.

The relative entropy of  $T_h\mu$  can be written as:

$$(3.126) \quad \mathbb{H}(T_h\mu|\mu) = \mathbb{E}_{T_h\mu}[(h, \zeta)] - \log(\mathbb{E}_\mu[e^{(h,\zeta)}]) = (h, \theta) - \log(\mathbb{E}_\mu[e^{(h,\zeta)}]), \quad \theta = \mathbb{E}_{T_h\mu}[\zeta].$$

The positive definiteness of  $\text{Cov}(T_h\mu)$  on  $X$  implies that  $X \ni h \mapsto \log \mathbb{E}_\mu[e^{(h,\zeta)}]$  is strictly convex, and hence  $h \mapsto \theta(h) = \mathbb{E}_{T_h\mu}[\zeta]$  is strictly increasing in any direction of  $X$ . Let  $K$  be the image of  $\theta(h)$ , and for  $\theta \in K$ , let  $h(\theta)$  be the inverse function, and then let

$$(3.127) \quad \Gamma(\theta) = \mathbb{H}(T_{h(\theta)}\mu|\mu) = (\theta, h(\theta)) - \log \mathbb{E}_\mu[e^{(\zeta, h(\theta))}].$$

Thus  $\Gamma$  be the Legendre transform of the cumulant generating function of  $\mu$ , and

$$(3.128) \quad \mathbb{H}(T_h\mu|\mu) = \Gamma(\mathbb{E}_{T_h\mu}[\zeta]).$$

In particular,  $h(\mathbb{E}_\mu[\zeta]) = 0$  and properties of Legendre transform imply that, in directions of  $X$ ,

$$(3.129) \quad \nabla \Gamma(\theta) = (h \mapsto \mathbb{E}_{T_h\mu}[\zeta])^{-1}|_{h=h(\theta)}$$

$$(3.130) \quad \text{Hess } \Gamma(\theta) = \left[ \text{Hess } \log \mathbb{E}_\mu[e^{(h,\zeta)}] \Big|_{h=h(\theta)} \right]^{-1} = \text{Cov}(T_{h(\theta)}\mu)^{-1}$$

so that

$$(3.131) \quad \nabla \Gamma(\mathbb{E}_\mu[\zeta]) = 0, \quad \text{Hess } \Gamma(\mathbb{E}_{T_h\mu}[\zeta]) = \text{Cov}(T_h\mu)^{-1}.$$

The assumption  $\dot{\Sigma} \text{Cov}(T_h\mu) \dot{\Sigma} \leq \alpha \dot{\Sigma}$  implies, for  $\theta \in K$ ,

$$(3.132) \quad \alpha \text{Hess } \Gamma(\theta) \geq \dot{\Sigma}.$$

Since  $f(\theta) = \frac{1}{2}(\theta - \mathbb{E}_\mu[\zeta])_\Sigma^2$  satisfies  $\nabla f(\mathbb{E}_\mu[\zeta]) = 0$  and  $\text{Hess } f = \dot{\Sigma}$ , therefore:

$$(3.133) \quad \alpha \mathbb{H}(T_h\mu|\mu) = \alpha \Gamma(\mathbb{E}_{T_h\mu}[\zeta]) \geq \frac{1}{2}(\mathbb{E}_{T_h\mu}[\zeta] - \mathbb{E}_\mu[\zeta])_\Sigma^2 \quad \text{for all } h \in \mathbb{R}^N. \quad \square$$



## 4 Pathwise Polchinski flow and stochastic localisation perspective

**4.1. Pathwise realisation of the Polchinski semigroup.** From Proposition 3.5, we recall that the Polchinski semigroup operates from the right:

$$(4.1) \quad \mathbf{P}_{s,t} = \mathbf{P}_{r,t} \mathbf{P}_{s,r}, \quad (s \leq r \leq t).$$

Thus it acts on probability densities relative to the measure  $\nu_t$ : if  $d\mu_0 = F d\nu_0$  is a probability measure then  $d\mu_t = \mathbf{P}_{0,t} F d\nu_t$  is again a probability measure.

This should be compared with the more standard situation of a time-independent semigroup  $\mathbf{T}_{s,t} = \mathbf{T}_{t-s}$  that is reversible with respect to a measure  $\nu$  such as the original Glauber–Langevin semigroup introduced in (2.12). In this case, one has the dual point of view that  $\mathbf{T}$  describes the evolution of an observable: if  $d\mu_0 = F d\nu$  is some initial distribution and  $d\mu_t = (\mathbf{T}_t F) d\nu$  denotes the distribution at time  $t$  then, by reversibility,

$$(4.2) \quad \mathbb{E}_{\mu_t}[G] = \int G(\mathbf{T}_t F) d\nu = \int (\mathbf{T}_t G) F d\nu = \mathbb{E}_{\mu_0}[\mathbf{T}_t G].$$

The dual semigroup can be realised in terms of a Markov process  $(\varphi_t)$  as  $\mathbf{T}_t G(\varphi) = \mathbf{E}_{\varphi_0=\varphi}[G(\varphi_t)]$ .

Since the Polchinski semigroup is not reversible and time-dependent, this interpretation does not apply to the Polchinski semigroup. Instead, the Polchinski semigroup  $\mathbf{P}_{s,t}$  can be realised in terms of an SDE that starts at time  $t$  and runs time in the negative direction from  $t$  to  $s < t$ : Given  $t > 0$ , a standard Brownian motion  $(B_u)_{u \geq 0}$  and  $\varphi_t$ , consider the solution to

$$(4.3) \quad \varphi_s = \varphi_t - \int_s^t \dot{C}_u \nabla V_u(\varphi_u) du + \int_s^t \sqrt{\dot{C}_u} dB_u, \quad (s \leq t).$$

This is the equation for the (stochastic) characteristics of the Polchinski equation, see Appendix A for the classical analogue of a Hamilton–Jacobi equation without viscosity term. By reversing time direction, this backward in time SDE becomes a standard SDE. Indeed, to be concrete, we will *interpret* (4.3) as  $\varphi_t = \tilde{\varphi}_{t-r}$  where  $\tilde{\varphi}$  is the solution to the following standard SDE with  $\tilde{\varphi}_0 = \varphi_t$  given and  $\tilde{B}_r = B_{t-r}$ :

$$(4.4) \quad d\tilde{\varphi}_r = -\dot{C}_{t-r} \nabla V_{t-r}(\tilde{\varphi}_r) dr + \sqrt{\dot{C}_{t-r}} d\tilde{B}_r, \quad (0 \leq r \leq t).$$

Denoting by  $\mathbb{E}_{\varphi_t=\varphi}[\cdot]$  the expectation with respect to the solution  $(\varphi_s)_{s \leq t}$  to (4.3) with  $\varphi_t = \varphi$  given, the Polchinski semigroup can be represented as follows.

**Proposition 4.1.** *For  $s \leq t$  and any bounded  $F : X \rightarrow \mathbb{R}$ ,*

$$(4.5) \quad \mathbf{P}_{s,t} F(\varphi) = \mathbb{E}_{\varphi_t=\varphi}[F(\varphi_s)].$$

*Thus if  $\varphi_t$  is distributed according to the renormalised measure  $\nu_t$  the backward in time evolution (4.3) ensures that  $\varphi_s$  is distributed according to  $\nu_s$  for  $s < t$ .*

Our interpretation of this proposition is that, while the renormalised measures  $\nu_t$  are supported on increasing smooth configurations as  $t$  grows, the backward evolution restores the small scale fluctuations of  $\nu_0$ .

To verify Proposition 4.1, we change time direction so that (4.3) becomes a standard (forward) SDE as follows. Indeed, as discussed above, set  $\tilde{\varphi}_r = \varphi_{t-r}$  and  $\tilde{B}_r = B_{t-r}$ . Then (4.3) becomes

$$(4.6) \quad \begin{aligned} \tilde{\varphi}_r &= \tilde{\varphi}_0 - \int_{t-r}^t \dot{C}_u \nabla V_u(\tilde{\varphi}_{t-u}) du + \int_{t-r}^t \sqrt{\dot{C}_u} dB_u \\ &= \tilde{\varphi}_0 - \int_0^r \dot{C}_{t-u} \nabla V_{t-u}(\tilde{\varphi}_u) du + \int_0^r \sqrt{\dot{C}_{t-u}} d\tilde{B}_u, \end{aligned}$$

i.e.,  $\tilde{\varphi}$  solves the standard SDE (4.4). Itô's formula stated for the forward SDE for  $\tilde{\varphi}$  is

$$(4.7) \quad d\tilde{f}(r, \tilde{\varphi}_r) = \frac{\partial \tilde{f}}{\partial r}(r, \tilde{\varphi}_r) + \mathbf{L}_{t-r}\tilde{f}(r, \tilde{\varphi}_r) + (\nabla \tilde{f}(r, \tilde{\varphi}_r), \sqrt{\dot{C}_{t-r}} d\tilde{B}_r).$$

In terms of  $\varphi$  rather than  $\tilde{\varphi}$  we will state this as

$$(4.8) \quad df(s, \varphi_s) = \frac{\partial f}{\partial s}(s, \varphi_s) - \mathbf{L}_s f(s, \varphi_s) + (\nabla f(s, \varphi_s), \sqrt{\dot{C}_s} dB_s),$$

where the left-hand side is interpreted as follows: with  $s = t - r$ ,

$$(4.9) \quad \begin{aligned} -d_s f(s, \varphi_s) &= d_r f(t - r, \tilde{\varphi}_r) \\ &= -\frac{\partial f}{\partial s}(t - r, \tilde{\varphi}_r) + \mathbf{L}_{t-r} f(t - r, \tilde{\varphi}_r) + (\nabla f(t - r, \tilde{\varphi}_r), \sqrt{\dot{C}_{t-r}} d\tilde{B}_r) \\ &= -\frac{\partial f}{\partial s}(s, \varphi_s) + \mathbf{L}_s f(s, \varphi_s) - (\nabla f(s, \varphi_s), \sqrt{\dot{C}_s} dB_s). \end{aligned}$$

In particular, if  $f$  is smooth and bounded and satisfies  $(\partial_s - \mathbf{L}_s)f = 0$  then

$$(4.10) \quad \mathbb{E}_{\varphi_t = \varphi}[f(s, \varphi_s)] = f(t, \varphi).$$

*Proof of Proposition 4.1.* It is enough to prove the claim for bounded smooth  $F$  and then extend it by density. The claim follows from (4.10) with  $f(t, \varphi) = \mathbf{P}_{s,t}F(\varphi)$  which gives

$$(4.11) \quad \mathbb{E}_{\varphi_t = \varphi}[F(\varphi_s)] = \mathbb{E}_{\varphi_t = \varphi}[\mathbf{P}_{s,t}F(\varphi_s)] = \mathbb{E}_{\varphi_t = \varphi}[f(s, \varphi_s)] = f(t, \varphi) = \mathbf{P}_{s,t}F(\varphi).$$

This proves (4.5). For the last statement, recall from (3.27) that  $\mathbb{E}_{\nu_s}[F] = \mathbb{E}_{\nu_t}[\mathbf{P}_{s,t}F] = \mathbb{E}_{\nu_t}[F(\varphi_s)]$ . This characterises the distribution at time  $s$  of the process (4.3).  $\square$

Finally, we will consider below an analogue of the backward in time SDE (4.3) started at time  $t = +\infty$ , see (4.20). Equation (4.20) can analogously be interpreted by reversing time as follows. Fix any smooth *time-reversing* reparametrisation  $a : [0, +\infty] \rightarrow [0, +\infty]$ . For simplicity, one can choose  $a(t) = 1/t$  with  $a(0) = +\infty$  and  $a(+\infty) = 0$ . As in Remark 3.8, set

$$(4.12) \quad \tilde{C}_t = C_{a(t)}, \quad \tilde{V}_t = V_{a(t)},$$

and also

$$(4.13) \quad \tilde{\varphi}_t = \varphi_{a(t)}, \quad \tilde{B}_t = B_{a(t)}.$$

Analogously to (4.4), the solution (4.20) can then be interpreted as  $\varphi_t = \tilde{\varphi}_{a(t)}$  where  $\tilde{\varphi}$  is the solution to the standard SDE:

$$(4.14) \quad d\tilde{\varphi}_t = \dot{C}_{a(t)} \dot{a}(t) \nabla V_{a(t)}(\tilde{\varphi}_t) dt + \sqrt{\dot{C}_{a(t)}} d\tilde{B}_t, \quad \tilde{\varphi}_0 = 0.$$

More generally, as in Remark 3.8, the SDEs (4.3)–(4.20) are invariant under reparametrisation, thus  $[0, +\infty]$  has no special significance and we could have used  $[0, 1]$  from the beginning instead.

We prefer to consider the backward in time evolution corresponding to  $\varphi$  (rather than the forward SDE for  $\tilde{\varphi}$ ) to comply with the convention that the renormalised potential  $V_t$  evolves forward in time according to the Polchinski equation. From the stochastic analysis point of view, on the other hand, this convention of a stochastic process running backwards in time is less standard and related literature which focuses on the SDE rather than the renormalised potential, as we do, thus uses the opposite convention (see Sections 4.5 and 5.1).

**4.2. Example: Log-Sobolev inequality by coupling.** Using the representation (4.3)-(4.5) of the semigroup  $\mathbf{P}_{s,t}$  in terms of the above stochastic process, one can alternatively prove Theorem 3.6 using synchronous coupling by adapting the proof from [27] for the Bakry–Émery theorem.

*Proof of Theorem 3.6.* Given  $t > 0$ , define  $(\varphi_s)_{s \leq t}$  and  $(\varphi'_s)_{s \leq t}$  as in (4.3) coupled using the same Brownian motions. Then for  $s < t$ ,

$$(4.15) \quad \begin{aligned} e^{-2\lambda t}(\varphi_t - \varphi'_t)_{\dot{C}_t^{-1}}^2 - e^{-2\lambda s}(\varphi_s - \varphi'_s)_{\dot{C}_s^{-1}}^2 \\ = \int_s^t \left[ e^{-2\lambda u} \left( -2\dot{\lambda}_u(\varphi_u - \varphi'_u)_{\dot{C}_u^{-1}}^2 - (\varphi_u - \varphi'_u)_{\dot{C}_u^{-1}}^2 \ddot{C}_u \dot{C}_u^{-1} \right. \right. \\ \left. \left. + 2(\dot{C}_u(\nabla V_u - \nabla V'_u), \dot{C}_u^{-1}(\varphi_u - \varphi'_u)) \right) \right] du \geq 0, \end{aligned}$$

where the inequality follows from the assumption (3.32) and the mean value theorem. Thus

$$(4.16) \quad (\varphi_0 - \varphi'_0)_{\dot{C}_0^{-1}}^2 = e^{-2\lambda_0}(\varphi_0 - \varphi'_0)_{\dot{C}_0^{-1}}^2 \leq e^{-2\lambda t}(\varphi_t - \varphi'_t)_{\dot{C}_t^{-1}}^2$$

with  $\varphi_t = \varphi$  and  $\varphi'_t = \varphi'$  (the dynamics runs backwards) and the mean value theorem gives

$$(4.17) \quad \begin{aligned} |\mathbf{P}_{0,t}F(\varphi) - \mathbf{P}_{0,t}F(\varphi')| &= |\mathbb{E}[(\nabla F(\psi_0), \varphi_0 - \varphi'_0)]| \\ &= 2|\mathbb{E}[(\nabla \sqrt{F(\psi_0)}, \sqrt{F(\psi_0)}(\varphi_0 - \varphi'_0))]| \\ &\leq 2\mathbb{E}[(\nabla \sqrt{F(\psi_0)})_{\dot{C}_0}^2]^{1/2} \mathbb{E}[F(\psi_0)(\varphi_0 - \varphi'_0)_{\dot{C}_0^{-1}}^2]^{1/2} \\ &\leq 2e^{-\lambda t} \mathbb{E}[(\nabla \sqrt{F(\psi_0)})_{\dot{C}_0}^2]^{1/2} \mathbb{E}[F(\psi_0)]^{1/2} \sqrt{(\varphi - \varphi')_{\dot{C}_t^{-1}}^2} \end{aligned}$$

for some  $\psi_0$  between  $\varphi$  and  $\varphi'$ . Taking  $\varphi - \varphi' = \sqrt{\dot{C}_t}f$  with  $\|f\|_2 \rightarrow 0$  gives

$$(4.18) \quad (\nabla \sqrt{\mathbf{P}_{0,t}F})_{\dot{C}_t}^2 \leq e^{-2\lambda t} \mathbf{P}_{0,t}(\nabla \sqrt{F})_{\dot{C}_0}^2.$$

This is (3.46). □

**4.3. Example: Coupling with the Gaussian reference measure.** Since, by (3.13),

$$(4.19) \quad \mathbb{E}_{\nu_t}[F] = \mathbf{P}_{t,\infty}F(0),$$

one can obtain the following coupling of the field distributed under the measure  $\nu_t$  with that of the associated driving Gaussian field from the stochastic realisation of  $\mathbf{P}_{t,\infty}$ .

**Corollary 4.2.** *The distribution of  $\nu_t$  is realised by the solution to the SDE (which we recall can be interpreted as discussed around (4.14)):*

$$(4.20) \quad \begin{aligned} \varphi_t &= - \int_t^\infty \dot{C}_u \nabla V_u(\varphi_u) du + \int_t^\infty \sqrt{\dot{C}_u} dB_u \\ &= - \int_t^\infty \dot{C}_u \nabla V_u(\varphi_u) du + \Gamma_t. \end{aligned}$$

*In particular, at  $t = 0$ , this provides a coupling of the full interacting field  $\varphi_0$  with the Gaussian reference field  $\Gamma_0$ .*

As an application of the above coupling, one can relate properties of the Gaussian measure to the interacting one, see for example [10, 20].

**4.4. Effective potential and martingales.** The stochastic process (4.3) can also be used to obtain a representation of the renormalised potential as follows. These are stochastic interpretations of the formulas in Lemma 3.11.

**Proposition 4.3.**

$$(4.21) \quad \begin{aligned} V_t(\varphi) &= \mathbb{E}_{\varphi_t=\varphi} \left[ V_0(\varphi_0) + \frac{1}{2} \int_0^t (\nabla V_s(\varphi_s))_{\dot{C}_s}^2 ds \right] \\ &= \mathbb{E}_{\varphi_t=\varphi} \left[ V_0 \left( \varphi - \int_0^t \dot{C}_s \nabla V_s(\varphi_s) ds + \int_0^t \sqrt{\dot{C}_s} dB_s \right) + \frac{1}{2} \int_0^t (\nabla V_s(\varphi_s))_{\dot{C}_s}^2 ds \right] \end{aligned}$$

and  $M_s = V_s(\varphi_s) + \frac{1}{2} \int_s^t (\nabla V_u(\varphi_u))_{\dot{C}_u}^2 du$  is a martingale (with respect to the backward filtration).

*Proof.* It suffices to show that  $M_s$  is a martingale. By Itô's formula interpreted as in (4.8),

$$(4.22) \quad \begin{aligned} dM_s &= \left( \frac{\partial V_s}{\partial s} \right)(\varphi_s) - \mathbf{L}_s V_s(\varphi_s) - \frac{1}{2} (\nabla V_s)_{\dot{C}_s}^2 + \text{martingale} \\ &= \left( \frac{\partial V_s}{\partial s} \right)(\varphi_s) - \frac{1}{2} \Delta_{\dot{C}_s} V_s(\varphi_s) + \frac{1}{2} (\nabla V_s)_{\dot{C}_s}^2 + \text{martingale}. \end{aligned}$$

By Polchinski's equation for  $V_s$ , the right-hand side is a martingale.  $\square$

The gradient and Hessian of the renormalised potential have similar representations.

**Proposition 4.4.**

$$(4.23) \quad \nabla V_t(\varphi) = \mathbf{P}_{0,t}[\nabla V_0](\varphi) = \mathbb{E}_{\varphi_t=\varphi}[\nabla V_0(\varphi_0)]$$

and  $U_t(\varphi_t) = \nabla V_t(\varphi_t)$  is a martingale (with respect to the backwards filtration).

*Proof.* Again, by Itô's formula (4.8) and since  $\partial_t U_t = \mathbf{L}_t U_t$  by (3.56),

$$(4.24) \quad dU_t(\varphi_t) = \frac{\partial U_t}{\partial t}(\varphi_t) - \mathbf{L}_t U_t(\varphi_t) + (\nabla U_t, \sqrt{\dot{C}_t} dB_t) = (\nabla U_t, \sqrt{\dot{C}_t} dB_t).$$

Thus  $U_t$  is a martingale, and the expression for its expectation also follows.  $\square$

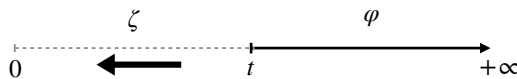


Figure 4.1. In the renormalisation group approach, the small scales  $\zeta$  are averaged out and one considers the projection of the measure to the variables  $\varphi$  encoding the large scales; in the figure above, the Polchinski flow goes from 0 to  $+\infty$ . In fact,  $\zeta$  and  $\varphi$  play symmetric roles: in particular for  $t=0$  the original measure is coded by  $\varphi$ , while instead for  $t=+\infty$  the original measure is coded by  $\zeta$ . Stochastic localisation puts the emphasis on the variable  $\zeta$  and therefore flows in the opposite direction (depicted by the thick arrow).

**4.5. Stochastic localisation perspective.** The stochastic evolution (4.3) has so far been interpreted as the characteristics associated with the Polchinski equation (3.21). In this section, we are going to see that this stochastic process is also, after a suitable change of parametrisation, the flow of the *stochastic localisation*, introduced by Eldan. We refer to [40] for a survey on this method and its numerous applications in general, and to [28] for more specific developments on modified log-Sobolev inequalities. The relation between stochastic localisation and a semigroup approach was already pointed out in [58].

From Lemma 3.12, we recall that the gradient and Hessian of the renormalised potential  $V_t$  can be interpreted as a mean and covariance of the fluctuation measure  $\mu_t^\varphi$  defined in (3.17) by

$$(4.25) \quad \mathbf{P}_{0,t}F(\varphi) = \mathbb{E}_{\mu_t^\varphi}[F].$$

The measure  $\mu_t^\varphi$  is related to  $\mu_t^0$  by the exponential tilt  $e^{(C_t^{-1}\varphi, \zeta)}$ , i.e., by the external field  $C_t^{-1}\varphi$ . In particular, by Lemma 3.12, the gradient of  $V_t$  can be written as

$$(4.26) \quad \nabla V_t(\varphi) = \mathbb{E}_{\mu_t^\varphi}[C_t^{-1}(\varphi - \zeta)] = C_t^{-1}(\varphi - \mathbb{E}_{\mu_t^\varphi}[\zeta])$$

where  $\mathbb{E}_{\mu_t^\varphi}[\zeta] \in X$  is the mean of  $\mu_t^\varphi$ . The stochastic representation (4.3) can therefore be written in terms of the fluctuation measure instead of the renormalised potential. Indeed, let

$$(4.27) \quad h_t = C_t^{-1}\varphi_t, \quad \mu_t = \mu_t^{\varphi_t} = \mu_t^{C_t h_t}, \quad \dot{\Sigma}_t = -\frac{\partial}{\partial t}C_t^{-1} = C_t^{-1}\dot{C}_t C_t^{-1}.$$

Since

$$(4.28) \quad \dot{C}_t \nabla V_t(\varphi_t) = \dot{C}_t C_t^{-1}(\varphi - \mathbb{E}_{\mu_t^\varphi}[\zeta]) = C_t \dot{\Sigma}_t(\varphi - \mathbb{E}_{\mu_t^\varphi}[\zeta]),$$

the external field  $h_t = C_t^{-1}\varphi_t$  satisfies the following SDE equivalent to (4.20): By the Itô formula (4.8) with  $f(t, \varphi_t) = C_t^{-1}\varphi_t$ ,

$$(4.29) \quad \begin{aligned} h_t &= -\int_t^\infty df(u, \varphi_u) \\ &= \int_t^\infty \dot{\Sigma}_u \varphi_u du - \int_t^\infty C_u^{-1} \dot{C}_u \nabla V_u(\varphi_u) du + \int_t^\infty C_u^{-1} \dot{C}_u^{1/2} dB_u \\ &= \int_t^\infty \dot{\Sigma}_u \mathbb{E}_{\mu_u}[\zeta] du + \int_t^\infty C_u^{-1} \dot{C}_u^{1/2} dB_u \\ &= \int_t^\infty \dot{\Sigma}_u \mathbb{E}_{\mu_u}[\zeta] du + \int_t^\infty \dot{\Sigma}_u^{1/2} dB_u, \end{aligned}$$

where the last equality holds in distribution in the case that  $\dot{C}_u$  and  $C_u^{-1}$  do not commute.

What is known as *stochastic localisation* is the process  $(h_t)$  with the direction of time reversed. Thus in the stochastic localisation perspective, the renormalised potential and measure only play implicit roles, and the main object of study is the stochastic process (4.29) and the fluctuation measure (4.25). For this perspective, it is more convenient to assume that time is parameterised by  $[0, T]$  (rather than our previous standard choice  $[0, +\infty]$  — but again everything is reparametrisation invariant, so this is only for notational purposes). The fluctuation measure  $\mu_t = \mu_t^{\varphi_t}$  then “starts” at the final time  $t = T$  as the full measure of interest, and as  $t$  decreases (time runs backwards) its fluctuations get absorbed into the renormalised measure  $\nu_t$  until the fluctuation measure  $\mu_t$  “localises” to a random Dirac measure  $\mu_0 = \delta_{\varphi_0}$  at time  $t = 0$ , with  $\varphi_0$  distributed according to the full measure  $\nu_0 = \mu_T$ . See also Figure 4.1.

Although time runs backwards from  $T$  to 0 in the stochastic localisation perspective written with our time convention, let us change time direction to obtain a forward SDE and connect with the literature on stochastic localisation. Recalling (4.27), the initial measure  $\nu_0 = \mu_T$  coincides with the fluctuation measure at time  $T$  as  $h_T = 0$ . As done previously, we will always use tildes to denote change of time:

$$(4.30) \quad \tilde{\varphi}_t = \varphi_{T-t}, \quad \tilde{C}_t = C_{T-t}, \quad \tilde{V}_t = V_{T-t}, \quad \tilde{\mu}_t = \mu_{T-t}, \quad \tilde{h}_t = h_{T-t}, \quad \tilde{\Sigma}_t = \dot{\Sigma}_{T-t}.$$

Using the notation  $\mathbf{b}(\mu) = \mathbb{E}_\mu[\zeta]$  for the mean of  $\mu$ , the SDE (4.29) for  $\tilde{h}$  can then be written as:

$$(4.31) \quad d\tilde{h}_t = \tilde{\Sigma}_t \mathbf{b}(\tilde{\mu}_t) dt + \tilde{\Sigma}_t^{1/2} d\tilde{B}_t.$$

This equation is the same as the *stochastic localisation* as it appears for example in [28, Fact 14] (after dropping tildes from the notation and with  $y_t$  there corresponding to  $\tilde{h}_t$ ).

The stochastic localisation perspective is different from our renormalisation group perspective in that the object of interest is (again) the fluctuation measure. For example, in the one-variable

case  $|\Lambda| = 1$ , starting from a measure  $\tilde{\mu}_0(dx) = \nu_0(dx) \propto e^{-H(x)}$  (possibly log-concave), the strategy is to make it more convex by considering

$$(4.32) \quad \tilde{\mu}_t(d\zeta) \propto e^{-H(\zeta) - \frac{t}{2}\zeta^2 + \tilde{h}_t\zeta} d\zeta$$

with the choice of the process  $\tilde{h}_t$  such that for any test function

$$(4.33) \quad \forall t \geq 0, \quad \mathbb{E}_{\nu_0}[F] = \mathbb{E}\left[\mathbb{E}_{\tilde{\mu}_t}(F)\right].$$

In this one variable example, the fluctuation measure above is the counterpart of (3.76) for the choice  $C_t = 1/(1+t)$  with  $t$  decreasing from  $+\infty$  to 0 instead of  $C_t = t$  with  $t \in [0, 1]$ . With this reparametrisation, one gets from (4.27) that  $\dot{\Sigma}_t = 1$  so that

$$(4.34) \quad t \geq 0, \quad d\tilde{h}_t = \mathbf{b}(\tilde{\mu}_t) dt + d\tilde{B}_t, \quad \text{with } \tilde{h}_0 = 0.$$

Starting from a general measure  $\tilde{\mu}_0$ , the primary concern in the stochastic localisation perspective is the measure  $\tilde{\mu}_t$  which is now uniformly convex with Hessian at least  $t$  (if say  $H$  is log-concave), thus general concentration inequalities hold for the twisted measure and can be transferred to  $\tilde{\mu}_0$  thanks to (4.33). For example, this is a key tool in current progress on the KLS conjecture, see [40] for a review. The larger  $t$  is, the better in this respect. However, as  $t$  grows the twisted measure  $\tilde{\mu}_t(d\zeta)$  loses the features of the original  $\tilde{\mu}_0$  so there is a trade-off in the choice of  $t$ . Contrary to our renormalisation point of view, in the stochastic localisation point of view, the distribution of  $h_t = C_t^{-1}\varphi_t$  (which is given in terms of  $\tilde{\mu}_t$  in (4.31) but can also be written in terms of our renormalised measure) does not play an important role (see Figure 4.1). The process  $\tilde{h}_t$  is there to twist the measure and sometimes if one adds the correct  $\dot{C}_t$  there are preferred directions to add the convexity.

## 5 Variational and transport perspectives on the Polchinski flow

In this section, we discuss transport-related perspectives on the Polchinski flow. We refer to [34] for additional perspectives such as an interpretation in terms of the Otto calculus that we do not discuss here.

**5.1. Föllmer's problem.** By (4.20), the distribution  $\nu_0$  can be realised as the final time distribution  $\varphi_0$  of the SDE:

$$(5.1) \quad \varphi_t = - \int_t^\infty \dot{C}_u \nabla V_u(\varphi_u) du + \int_t^\infty \sqrt{\dot{C}_u} dB_u,$$

where we recall that the backwards SDE can be interpreted as in (4.14). One can ask whether the distribution  $\nu_0$  can be obtained more efficiently if  $\nabla V_u(\varphi_u)$  is replaced by another drift  $U_u(\varphi_u)$ , i.e., as the distribution of  $\varphi_0^U$  when  $\varphi^U$  is a strong solution of the SDE (again written backward in time):

$$(5.2) \quad \varphi_t^U = - \int_t^\infty \dot{C}_u U_u(\varphi_u^U) du + \int_t^\infty \sqrt{\dot{C}_u} dB_u,$$

where the parameter  $t$  takes values in  $[0, +\infty]$  and  $\varphi_\infty^U = 0$ . As pointed out in Remark 3.8, one could have also considered a parametrisation on a bounded time interval.

Denote by  $\gamma_0 = P_{C_\infty}$  the distribution of the Gaussian reference measure, i.e., of  $\int_0^\infty \sqrt{\dot{C}_u} dB_u$ .

**Theorem 5.1.** *The gradient of the renormalised potential  $V_t$  of the Polchinski flow (3.11) can be interpreted as the optimal drift in (5.2) in the following sense:*

$$(5.3) \quad \mathbb{H}(\nu_0|\gamma_0) = \frac{1}{2}\mathbb{E}\left[\int_0^\infty |\nabla V_t(\varphi_t)|_{\dot{C}_t}^2 dt\right] \leq \frac{1}{2}\mathbb{E}\left[\int_0^\infty |U_t(\varphi_t^U)|_{\dot{C}_t}^2 dt\right],$$

for any drift  $U$  such that (5.2) has a strong solution with  $\varphi_0 \sim \nu_0$ . Recall that  $(\varphi_t)$  follows (5.1).

*Proof.* Let  $U$  be a such that there is a strong solution  $(\varphi_t^U)$  of (5.2) with  $\mathbb{E}[\int_0^\infty |U_t(\varphi_t^U)|_{\dot{C}_t}^2 dt] < \infty$ . By construction  $\varphi_0^U$  has law  $\nu_0$  so that the relative entropy is given by

$$(5.4) \quad \mathbb{H}(\nu_0|\gamma_0) = V_\infty(0) - \mathbb{E}[V_0(\varphi_0^U)] = \int_0^\infty dt \frac{\partial}{\partial t} \mathbb{E}[V_t(\varphi_t^U)],$$

with  $\varphi^U$  evolving according to (5.2), and where we used that  $\nu_0(d\varphi) = e^{+V_\infty(0)} e^{-V_0(\varphi)} \gamma_0(d\varphi)$  with normalisation factor given by  $e^{-V_\infty(0)} = \mathbf{E}_{C_\infty}[e^{-V_0(\zeta)}]$  as in (3.13). The renormalised potential follows the Polchinski equation (3.21):

$$(5.5) \quad \frac{\partial}{\partial t} V_t = \frac{1}{2} \Delta_{\dot{C}_t} V_t - \frac{1}{2} (\nabla V_t)_{\dot{C}_t}^2, \quad (t \in (0, \infty)).$$

Therefore, by Itô's formula,

$$(5.6) \quad \begin{aligned} \frac{\partial}{\partial t} \mathbb{E}[V_t(\varphi_t^U)] &= \mathbb{E} \left[ \frac{\partial}{\partial t} V_t(\varphi_t^U) + (\nabla V_t(\varphi_t^U), U_t(\varphi_t^U))_{\dot{C}_t} - \frac{1}{2} \Delta_{\dot{C}_t} V_t(\varphi_t^U) \right] \\ &= \mathbb{E} \left[ -\frac{1}{2} (\nabla V_t(\varphi_t^U))_{\dot{C}_t}^2 + (\nabla V_t(\varphi_t^U), U_t(\varphi_t^U))_{\dot{C}_t} \right] \\ &= \frac{1}{2} \mathbb{E} \left[ -(\nabla V_t(\varphi_t^U) - U_t(\varphi_t^U))_{\dot{C}_t}^2 \right] + \frac{1}{2} \mathbb{E} \left[ (U_t(\varphi_t^U))_{\dot{C}_t}^2 \right], \end{aligned}$$

where we used the Polchinski equation (5.5) on the second line. Thus

$$(5.7) \quad \frac{1}{2} \mathbb{E} \left[ \int_0^\infty (\nabla U_t)_{\dot{C}_t}^2 dt \right] = \mathbb{H}(\nu_0|\gamma_0) + \frac{1}{2} \mathbb{E} \left[ \int_0^\infty (\nabla V_t - U_t)_{\dot{C}_t}^2 dt \right],$$

and the gradient of the renormalised potential  $V_t$  provides the optimal drift. This completes the proof of Theorem 5.1.  $\square$

It turns out that the right-hand of (5.3) is, in fact, the relative entropy  $\mathbb{H}(\mathbf{Q}|\mathbf{P})$  of the path measure  $\mathbf{Q}$  associated with (5.2) with respect to that of the Gaussian reference process  $\mathbf{P}$ .

**Proposition 5.2.** *The relative entropy of the path measure  $\mathbf{Q}$  associated with a strong solution of (5.9) with respect to the path measure  $\mathbf{P}$  of the Gaussian reference measure is given by*

$$(5.8) \quad \mathbb{H}(\mathbf{Q}|\mathbf{P}) = \frac{1}{2} \mathbb{E} \left[ \int_0^\infty |U_t(\varphi_t^U)|^2 dt \right].$$

*Proof.* This is essentially a consequence of Girsanov's theorem, see [63] for details.  $\square$

Since  $\mathbb{H}(\mathbf{Q}|\mathbf{P}) \geq \mathbb{H}(\nu_0|\gamma_0)$  always holds, by the entropy decomposition (2.52) and the fact that the laws of  $\nu_0$  and  $\gamma_0$  are marginals of the path measures  $\mathbf{Q}$  and  $\mathbf{P}$  respectively, the above shows that the optimal drift  $U_t = \nabla V_t$  in fact achieves equality:  $\mathbb{H}(\mathbf{Q}|\mathbf{P}) = \mathbb{H}(\nu_0|\gamma_0)$ .

The above question was already studied by Föllmer [36], and we refer to [63] for an exposition of this and connections with Gaussian functional inequalities. Föllmer's objective was to find the optimal drift  $b_t$  such that the process  $(X_t)_{t \in [0,1]}$  defined by the following SDE and distributed at time  $t = 1$  according to a given target measure  $\nu$ :

$$(5.9) \quad X_0 = 0, \quad dX_t = b_t(X_t) dt + dB_t \quad \text{and} \quad X_1 \sim \nu,$$

that minimises the dynamical cost

$$(5.10) \quad \frac{1}{2} \mathbb{E} \left[ \int_0^1 |b_t(X_t)|^2 dt \right]$$

over all possible drifts  $b$ . Up to time reversal, parametrisation by  $[0, +\infty]$  instead of  $[0, 1]$ , and introduction of the covariances  $\dot{C}_t$ , this is exactly the set-up of (5.2). For us the introduction of the covariances  $\dot{C}_t$  is a (conceptually and technically) important point, though, with the interpretation that the integral is now an integral over scales measured by the infinitesimal covariances  $\dot{C}_t$  which can also be interpreted as metrics as in Section 3.6.

More generally, one can look for the optimal drift to build a target probability measure of the form  $F(\varphi) \nu_0(d\varphi)$  using now the optimal stochastic flow as a reference process, i.e., we want to determine the drift  $U$  such that for

$$(5.11) \quad t \in [0, +\infty], \quad \psi_t = - \int_t^\infty \dot{C}_s U_s(\psi_s) ds - \int_t^\infty \dot{C}_s \nabla V_s(\psi_s) ds + \int_t^\infty \sqrt{\dot{C}_s} dB_s,$$

the cost  $\frac{1}{2} \mathbb{E} \left[ \int_0^1 |U_t(\psi_t)|_{\dot{C}_t}^2 dt \right]$  is minimised and  $\psi_0$  is distributed according to  $F d\nu_0$ . Proceeding as in the proof of Theorem 5.1, the optimal drift is given in terms of the Polchinski semigroup (3.12) as the gradient of

$$(5.12) \quad W_t(\varphi) = - \log \mathbf{P}_{0,t} F(\varphi) = -V_t(\varphi) - \log \mathbf{E}_{C_t} \left[ F(\varphi + \zeta) e^{-V_0(\varphi + \zeta)} \right].$$

Thus one can check that

$$(5.13) \quad \mathbb{H}(F\nu_0|\nu_0) = \text{Ent}_{\nu_0}(F) = \frac{1}{2} \mathbb{E} \left[ \int_0^\infty |\nabla W_t(\varphi_t)|_{\dot{C}_t}^2 dt \right].$$

In this way, we recover from (5.13) the entropy decomposition (3.44):

$$(5.14) \quad \begin{aligned} 2 \int_0^\infty \mathbb{E}_{\nu_t} \left[ |\nabla \sqrt{\mathbf{P}_{0,t} F}|_{\dot{C}_t}^2 \right] dt &= \frac{1}{2} \int_0^\infty \mathbb{E}_{\nu_t} \left[ |\nabla \log \mathbf{P}_{0,t} F|_{\dot{C}_t}^2 \mathbf{P}_{0,t} F \right] dt \\ &= \frac{1}{2} \mathbb{E} \left[ \int_0^\infty |\nabla W_t(\varphi_t)|_{\dot{C}_t}^2 dt \right], \end{aligned}$$

where we used that the process (5.11) is distributed at time  $t$  with density proportional to

$$(5.15) \quad e^{-W_t(\varphi) - V_t(\varphi)} \mathbf{P}_{C_\infty - C_t}(d\varphi) \propto \mathbf{P}_{0,t} F(\varphi) \nu_t(d\varphi).$$

The above is an instance of the more general version of the Schrödinger problem which is to find the optimal drift so that the stochastic evolution (5.2) interpolates between two probability measures  $\mu$  and  $\nu$ . Here, we discussed only the special case where the process starts from a Dirac measure  $\mu = \delta_0$ , and refer the reader to the survey [64] for a general overview and to [31] for a discussion on the role of the convexity of the potential.

In Section 5.3, we address a related issue, namely that in some cases, the previous flow can be modified in order to achieve an interpolation between the measure of interest and some Gaussian measure.

**5.2. Variational representation of the renormalised potential.** Let  $U_t = \nabla V_t(\varphi_t)$  and recall that Proposition 4.3 states:

$$(5.16) \quad V_t(\varphi) = \mathbb{E}_{\varphi_t = \varphi} \left[ V_0 \left( \varphi - \int_0^t \dot{C}_s U_s ds + \int_0^t \sqrt{\dot{C}_s} dW_s \right) + \frac{1}{2} \int_0^t |U_s|_{\dot{C}_s}^2 ds \right].$$

In particular,

$$(5.17) \quad V_t(\varphi) \geq \inf_U \mathbb{E} \left[ V_0 \left( \varphi - \int_0^t \dot{C}_s U_s ds + \int_0^t \sqrt{\dot{C}_s} dW_s \right) + \frac{1}{2} \int_0^t |U_s|_{\dot{C}_s}^2 ds \right],$$



where the above infimum is over all adapted processes  $U : [0, t] \rightarrow X$  (where adapted means backwards in time in our convention) called *drifts*. For our current purposes, it suffices to consider  $U_s = U_s(\varphi_s)$  associated with a strong solution to the (backward in time) SDE

$$(5.18) \quad \varphi_s = \varphi - \int_s^t \dot{C}_u U_u(\varphi_u) du + \int_s^t \sqrt{\dot{C}_u} dW_u, \quad (s \leq t).$$

The following proposition is a special case of the Boué-Dupuis or Borell formula, see [63], which gives equality in the infimum and is the starting point for the Barashkov–Gubinelli method [8]. An in-depth treatment of stochastic control problems of which this is a special case is given in [46].

**Proposition 5.3.**

$$(5.19) \quad V_t(\varphi) = \inf_U \mathbb{E} \left[ V_0 \left( \varphi - \int_0^t \dot{C}_s U_s ds + \int_0^t \sqrt{\dot{C}_s} dW_s \right) + \frac{1}{2} \int_0^t |U_s|_{\dot{C}_s}^2 ds \right].$$

*Sketch.* The entropy inequality (2.62) with  $G(\zeta) = -V_0(\varphi + \zeta)$  applied to the Gaussian measure  $\mathbb{P}_{C_t}$  implies that for any density  $F$  with  $\mathbb{E}_{C_t}[F] = 1$ :

$$(5.20) \quad V_t(\varphi) = -\log \mathbb{E}_{C_t}[e^{-V_0(\varphi + \zeta)}] \leq \text{Ent}_{\mathbb{P}_{C_t}}(F) + \mathbb{E}_{C_t}[F(\zeta)V_0(\varphi + \zeta)].$$

Given any drift  $U_s$ , let  $F d\mathbb{P}_{C_t}$  denote the law of  $\varphi_0 - \varphi$  solving (5.18):

$$(5.21) \quad \varphi_0 - \varphi = - \int_0^t \dot{C}_s U_s ds + \int_0^t \sqrt{\dot{C}_s} dW_s.$$

Then

$$(5.22) \quad V_t(\varphi) \leq \mathbb{E} \left[ \frac{1}{2} \int_0^t |U_s(\varphi_s)|_{\dot{C}_s}^2 ds + V_0 \left( \varphi - \int_0^t \dot{C}_s U_s ds + \int_0^t \sqrt{\dot{C}_s} dW_s \right) \right],$$

where we used that the entropy is bounded by the first term, exactly as in Theorem 5.1. As already discussed, the converse direction follows from Proposition 4.3.  $\square$

The point of view is now that by estimating the expectation on the right-hand side above, for a general drift  $U$ , one can obtain estimates on  $V_t(\varphi)$ , and in particular on  $V_\infty(\varphi)$  which we recall from (3.14) is equivalent to the logarithmic moment generating function of the measure  $\nu_0$ . For further details and application to construction of the  $\varphi_d^4$  measures, we refer to [8, 9].

**5.3. Lipschitz transport.** Instead of a stochastic process, one could also define a map  $\hat{S}_t : X \mapsto X$  transporting  $\nu_0$  to some measure  $\hat{\nu}_t$  in order to recover functional inequalities for  $\nu_0$  from  $\hat{\nu}_t$ . For example, assume that  $\hat{\nu}_t$  satisfies a log-Sobolev inequality with constant  $\gamma$  and that

$$(5.23) \quad \mathbb{E}_{\nu_0}[F(\varphi)] = \mathbb{E}_{\hat{\nu}_t}[F(\hat{S}_t(\varphi))].$$

If the Jacobian  $\nabla \hat{S}_t$  is bounded by some constant  $C$  in the following sense

$$(5.24) \quad \forall \omega \in X, \quad (\nabla \hat{S}_t \omega)^2 \leq C^2(\omega)^2,$$

then  $\nu_0$  satisfies also a log-Sobolev inequality

$$(5.25) \quad \begin{aligned} \text{Ent}_{\nu_0}[F(\varphi)] &= \mathbb{E}_{\hat{\nu}_t}[\Phi(F \circ \hat{S}_t)] - \Phi(\mathbb{E}_{\hat{\nu}_t}[F \circ \hat{S}_t]) \\ &\leq \frac{2}{\gamma} \mathbb{E}_{\hat{\nu}_t} \left[ (\nabla \sqrt{F \circ \hat{S}_t})^2 \right] = \frac{2}{\gamma} \mathbb{E}_{\hat{\nu}_t} \left[ ({}^t \nabla \hat{S}_t (\nabla \sqrt{F})(\hat{S}_t))^2 \right] \\ &\leq \frac{2C^2}{\gamma} \mathbb{E}_{\hat{\nu}_t} \left[ ((\nabla \sqrt{F})(\hat{S}_t))^2 \right] = \frac{2C^2}{\gamma} \mathbb{E}_{\nu_0} \left[ (\nabla \sqrt{F})^2 \right], \end{aligned}$$

where we used that  $\nabla \hat{S}_t$  and its transpose  ${}^t\nabla \hat{S}_t$  have the same operator norm. An analogous argument can be applied to more general functional inequalities.

This line of research was first investigated in [25] where it was understood that a convex perturbation of a Gaussian measure leads to a 1-Lipschitz transport map  $\hat{S}_t$ . We refer to [44, 57, 58, 71, 72, 75, 82] for more recent developments as well as other applications of the Lipschitz properties of transport maps to functional inequalities. In this section, we follow the work [83] which derived a Lipschitz estimate of the form (5.24) from the multiscale criterion (3.32) for some covariance decomposition, and generalise the result to other decompositions.

Recall that the measure  $\nu_0$  gets renormalised to  $\nu_t$  by the Polchinski flow. By construction  $\mathbb{E}_{\nu_t}[\cdot] = \mathbb{E}_{C_\infty - C_t}[e^{-V_t} \cdot]$  and  $\nu_t$  converges to a Dirac mass. As the measure  $\nu_t$  degenerates, it is more convenient to consider the measure  $\hat{\nu}_t$  obtained by rescaling  $\nu_t$  by some matrix  $D_t$  so that the measures  $\nu_0$  and  $\hat{\nu}_t$  are comparable:

$$(5.26) \quad \mathbb{E}_{\hat{\nu}_t} [F(\varphi)] = \mathbb{E}_{\nu_t} [F(D_t\varphi)].$$

If  $V_0 = 0$ , a natural choice for  $D_t$  is to preserve the Gaussian measure, i.e.,

$$(5.27) \quad D_t^{-1}(C_\infty - C_t)^{-1}D_t^{-1} = A \quad \Rightarrow \quad D_t A D_t = (C_\infty - C_t)^{-1},$$

where all inverse are understood to be taken on the range of  $A$ . This is implicitly assumed in the rest of the section, with  $\text{id}$  also denoting the identity matrix on the range of  $A$ . Assuming that all the matrices depend smoothly on  $t$  and commute, then the choice (5.27) implies the following useful relation:

$$(5.28) \quad 2D_t^{-1}\dot{D}_t = \dot{C}_t(C_\infty - C_t)^{-1} = \dot{C}_t D_t A D_t.$$

Using again (5.27) we get (recall  $C_0 = 0$  and notice  $D_0 = \text{id}$ ):

$$(5.29) \quad \frac{\partial}{\partial t}(D_t^{-2}) = -A\dot{C}_t \quad \Rightarrow \quad D_t = (\text{id} - AC_t)^{-1/2}.$$

By construction  $\lim_{t \rightarrow \infty} D_t^{-1}\varphi = 0$  and the renormalised potential satisfies  $\lim_{t \rightarrow \infty} V_t(D_t^{-1}\varphi) = V_\infty(0)$  (because  $\text{Hess } V_t \leq C_t^{-1} \leq C_1^{-1}$  for  $t \geq 1$ , recall Lemma 3.12). Thus the following convergence in distribution to a Gaussian measure holds:

$$(5.30) \quad \lim_{t \rightarrow \infty} \mathbb{E}_{\hat{\nu}_t} [F(\varphi)] = \mathbb{E}_{C_\infty} [F(\varphi)] = \mathbb{E}_{A^{-1}} [F(\varphi)].$$

Consider  $S_t$  a transport map between  $\nu_0$  and  $\hat{\nu}_t$ , so that

$$(5.31) \quad \mathbb{E}_{\nu_0} [F(S_t(\varphi))] = \mathbb{E}_{\hat{\nu}_t} [F(\varphi)] = \mathbb{E}_{\nu_t} [F(D_t\varphi)].$$

Note that ultimately we are interested in  $\hat{S}_t$  which is the inverse of  $S_t$ , see (5.23).

We are going to determine an evolution for  $S_t : X \mapsto X$ . On the one hand,

$$(5.32) \quad \frac{\partial}{\partial t} \mathbb{E}_{\nu_0} [F(S_t(\varphi))] = \mathbb{E}_{\nu_0} [(\nabla F(S_t(\varphi)), \partial_t S_t(\varphi))],$$

and on the other hand from (3.27):

$$(5.33) \quad \frac{\partial}{\partial t} \mathbb{E}_{\nu_t} [F(D_t\varphi)] = \mathbb{E}_{\nu_t} \left[ -\mathbf{L}_t F(D_t\varphi) + (\dot{D}_t\varphi, \nabla F(D_t\varphi)) \right],$$

with  $\mathbf{L}_t F = \frac{1}{2}\Delta_{\dot{C}_t} F - (\nabla V_t, \nabla F)_{\dot{C}_t}$ . Integrating by parts, this gives

$$(5.34) \quad \begin{aligned} \frac{\partial}{\partial t} \mathbb{E}_{\nu_t} [F(D_t\varphi)] &= \mathbb{E}_{\nu_t} \left[ \frac{1}{2}(\dot{C}_t \nabla V_t(\varphi) - \frac{1}{2}\dot{C}_t(C_\infty - C_t)^{-1}\varphi + D_t^{-1}\dot{D}_t\varphi, D_t \nabla F(D_t\varphi)) \right] \\ &= \mathbb{E}_{\nu_t} \left[ \frac{1}{2}(\nabla V_t(\varphi), D_t \nabla F(D_t\varphi))_{\dot{C}_t} \right] \\ &= \mathbb{E}_{\nu_0} \left[ \frac{1}{2}(\nabla V_t(D_t^{-1}\cdot), D_t \nabla F)_{\dot{C}_t}(S_t(\varphi)) \right], \end{aligned}$$

where we used (5.28). Thus the evolution of  $S_t$  is given by

$$(5.35) \quad \frac{\partial}{\partial t} S_t(\varphi) = \frac{1}{2} D_t \dot{C}_t \nabla V_t (D_t^{-1} S_t(\varphi)).$$

It was realised in [83] that for the choice  $\dot{C}_t = e^{-tA}$  (often used in applications, see Section 6) the Lipschitz structure associated with  $S_t$  is directly related to (a variant of) the multiscale Bakry–Émery criterion (3.32).

**Theorem 5.4** ([83]). *Under the assumption*

$$(5.36) \quad \forall \varphi \in X : \quad \dot{C}_t^{1/2} \text{Hess } V_t(\varphi) \dot{C}_t^{1/2} \geq \dot{\mu}_t \text{id}, \quad \text{with } \mu_t := \int_0^t ds \dot{\mu}_s,$$

the inverse map  $\hat{S}_t = S_t^{-1} : X \mapsto X$  is  $\exp(-\frac{1}{2}\mu_t)$ -Lipschitz.

From the convergence (5.30) to the Gaussian measure and if  $\mu_\infty := \int_0^\infty ds \dot{\mu}_s < \infty$ , then from the previous theorem, one can extract a  $\exp(-\frac{1}{2}\mu_\infty)$ -Lipschitz map from the Gaussian measure  $\mathbb{P}_{C_\infty}$  to  $\nu_0$  (see [75, Lemma 2.1]).

*Proof.* For  $\dot{C}_t = e^{-tA}$ , the definition (5.29) implies  $D_t = e^{\frac{1}{2}tA} = \dot{C}_t^{-1/2}$  and the evolution (5.35) becomes particularly simple:

$$(5.37) \quad \frac{\partial}{\partial t} S_t(\varphi) = \frac{1}{2} \dot{C}_t^{1/2} \nabla V_t (\dot{C}_t^{1/2} S_t(\varphi)).$$

The Jacobian evolves according to

$$(5.38) \quad \frac{\partial}{\partial t} \nabla S_t(\varphi) = \frac{1}{2} \dot{C}_t^{1/2} \text{Hess } V_t (\dot{C}_t^{1/2} S_t(\varphi)) \dot{C}_t^{1/2} \nabla S_t(\varphi).$$

As a consequence of (5.36), the Gronwall inequality implies

$$(5.39) \quad \forall \omega \in X, \quad \frac{\partial}{\partial t} (\nabla S_t(\varphi) \omega)^2 \geq \dot{\mu}_t (\nabla S_t(\varphi) \omega)^2 \quad \Rightarrow \quad (\nabla S_t(\varphi) \omega)^2 \geq \exp(\mu_t) (\omega)^2.$$

By the inverse function theorem, we deduce that the operator norm of  $\nabla \hat{S}_t$  is less than  $\exp(-\frac{1}{2}\mu_t)$ .  $\square$

Another useful covariance decomposition (see Theorem 6.2) is of the form  $\dot{C}_t = (tA + \text{id})^{-2}$ . The result of Theorem 5.4 can be extended to that case, but with a worse Lipschitz constant.

**Theorem 5.5.** *Let  $\dot{C}_t = (tA + 1)^{-2}$  for  $t \geq 0$ . Under the assumption*

$$(5.40) \quad \forall \varphi \in X : \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t - \frac{1}{2} \ddot{C}_t \geq \dot{\lambda}_t \dot{C}_t, \quad \text{with } \lambda_t := \int_0^t ds \dot{\lambda}_s,$$

the inverse map  $\hat{S}_t = S_t^{-1} : X \mapsto X$  is  $(\sqrt{t\|A\|} + 1)e^{-\lambda_t/2}$ -Lipschitz, with  $\|M\| = \sup\{\|M\omega\|_2 : \omega \in X, \|\omega\|_2 = 1\}$  the operator norm of a matrix  $M$ .

**Remark 5.6.** In concrete examples, the  $\lambda_t$  diverge with  $t$ , contrary to the  $\mu_t$  in Theorem 5.4 that remained bounded. This divergence precisely compensates the factor  $\sqrt{t}$ , giving a uniform Lipschitz bound on  $\hat{S}_t$ . An application of Theorem 5.5 is given in Proposition 6.3.

*Proof of Theorem 5.5.* For each  $t \geq 0$ , we look for a matrix  $B_t$  depending only on  $C_t$  and its derivatives, commuting with them, and such that we can set up a Gronwall estimate for  $(B_t \nabla S_t(\varphi) \omega)^2$  for each  $\varphi, \omega \in X$ . Note that the definition (5.29) of  $D_t$  implies  $D_t = \dot{C}_t^{-1/4}$ . In particular  $D_t$  commutes with  $B_s, C_s, \dot{C}_s, \ddot{C}_s$  for any  $s$ .

Equation (5.35) gives:

$$(5.41) \quad \frac{\partial}{\partial t} \nabla S_t(\varphi) = \frac{1}{2} \dot{C}_t D_t \text{Hess } V_t(D_t^{-1} S_t(\varphi)) D_t^{-1} \nabla S_t(\varphi).$$

Therefore:

$$(5.42) \quad \frac{\partial}{\partial t} (B_t \nabla S_t(\varphi) \omega)^2 = \left( B_t \nabla S_t(\varphi) \omega, \left[ B_t \dot{C}_t D_t \text{Hess } V_t(D_t^{-1} S_t(\varphi)) D_t^{-1} B_t^{-1} + 2 \dot{B}_t B_t^{-1} \right] B_t \nabla S_t(\varphi) \omega \right).$$

In order to use the Hessian bound (5.40), we need to choose  $B_t$  in such a way that, for each  $t \geq 0$ :

$$(5.43) \quad \begin{aligned} & B_t \dot{C}_t D_t \text{Hess } V_t(D_t^{-1} S_t(\varphi)) D_t^{-1} B_t^{-1} + 2 \dot{B}_t B_t^{-1} \\ & \geq \dot{C}_t^{1/2} \text{Hess } V_t(D_t^{-1} S_t(\varphi)) \dot{C}_t^{1/2} - \frac{1}{2} \dot{C}_t^{-1/2} \ddot{C}_t \dot{C}_t^{-1/2}. \end{aligned}$$

The Hessian term gives  $B_t = D_t^{-1} \dot{C}_t^{-1/2} = \dot{C}_t^{-1/4} = D_t$  as a sufficient condition, and for this choice  $\dot{B}_t B_t^{-1} = -\frac{1}{4} \ddot{C}_t \dot{C}_t^{-1}$  so that (5.43) holds. As a result,

$$(5.44) \quad \frac{\partial}{\partial t} (D_t \nabla S_t(\varphi) \omega)^2 \geq \dot{\lambda}_t (D_t \nabla S_t(\varphi) \omega)^2, \quad (t \geq 0).$$

The Gronwall inequality then yields:

$$(5.45) \quad (D_t \nabla S_t(\varphi) \omega)^2 \geq e^{\lambda_t} (\omega)^2, \quad (t \geq 0).$$

It follows that  $\hat{S}_t$  is  $\|\dot{C}_t^{1/4}\| e^{-\lambda_t/2}$ -Lipschitz, i.e.,  $(\sqrt{t\|A\|} + 1) e^{-\lambda_t/2}$ -Lipschitz as claimed.  $\square$

## 6 Applications

In this section, we present concrete examples to which the multiscale Bakry-Émery criterion of Theorem 3.6 can be applied. The criterion gives a bound on the log-Sobolev constant in terms of real numbers  $\dot{\lambda}_t$  ( $t > 0$ ) obtained through convexity lower bounds on the renormalised potential:

$$(6.1) \quad \forall \varphi \in X, t \geq 0: \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t - \frac{1}{2} \ddot{C}_t \geq \dot{\lambda}_t \dot{C}_t.$$

These lower bounds are determined by the choice of the covariance decomposition ( $C_t$ ). While Theorem 3.6 holds for any decomposition, checking (6.1) for concrete models often requires a specific choice of decomposition. This will be illustrated in examples in the following sections.

For now, we discuss the sharpness of the criterion in Theorem 3.6. To fix ideas, suppose we have a model defined on  $\Lambda_{\varepsilon, L} = L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$  for  $d \geq 2$ , where either  $\varepsilon = 1$  is fixed and  $L \rightarrow \infty$  (statistical mechanics model) or  $\varepsilon \rightarrow 0$  is a small regularisation parameter and  $L$  is fixed or  $L \rightarrow \infty$  (continuum field theory model in finite or infinite volume). From the discussion in Sections 2.6–2.7 recall that the speed of convergence of an associated dynamics is often related to the presence of phase transitions in equilibrium. These phase transitions are phenomena arising in the limit of large volumes, i.e., large  $L$ . In this limit, one typically expects that the log-Sobolev constant should be bounded from below independently of  $\Lambda_{\varepsilon, L}$  as long as no phase transition occurs. On the other hand, the additional regularisation parameter  $\varepsilon > 0$  is not expected to affect the dynamics, i.e., the log-Sobolev constant should also be bounded from below as  $\varepsilon \rightarrow 0$ .

Sharpness of the criterion of Theorem 3.6 is therefore evaluated through the following questions:

- (i) In absence of a phase transition, does the criterion provide a lower bound on the log-Sobolev constant uniform in the volume (i.e., in  $L$ )?

- (ii) Does it give a bound on the log-Sobolev constant independent of the regularisation parameter  $\varepsilon$  for continuum models?
- (iii) If the first point holds, can one correctly estimate how the log-Sobolev constant vanishes as a function of the distance to the critical point, or at the critical point as a function of  $L$ ?

The third point is considerably more involved than the first two. To provide a guideline to read the next sections, we collect here the answers to the above three questions obtained by studying the examples presented below.

- For statistical mechanics models ( $\varepsilon = 1$ ) at high temperature, i.e., far away from the critical point, the multiscale Bakry-Émery criterion implies point (i), see, e.g., Theorem 6.7 which extends to a much broader class of models. This regime is also covered by many other criteria (see e.g. the monographs [51, 80] and [90]), with the notable exception of mean-field spin-glass models, see the discussion of [11].
- The criterion (6.1) can be sharp enough to reach the critical point, see, e.g., Theorem 6.13 and Proposition 6.3 below for the Ising and  $\varphi^4$  models. In other words, there are models for which the criterion implies (i) up to the phase transition. We expect that the criterion should imply (i) up to the critical point for a large class of models.
- The criterion can provide a bound on the log-Sobolev constant that is uniform as  $\varepsilon \rightarrow 0$  (see Theorems 6.1 and 6.2 for the  $\varphi^4$  and sine-Gordon models), thus satisfying point (ii).

Point (iii) is in general open and in this generality hopelessly difficult, but some positive results exist. For models with quadratic mean-field interaction such as the Curie-Weiss model, one can answer (iii) in the affirmative. A detailed analysis above and below the critical temperature was also carried out for mean-field  $O(n)$  models in [21]. For certain continuum particle systems with mean-field interaction, the behaviour close to the critical point is the subject of ongoing work [14]. For more complicated models in which computations can be carried out, such as the hierarchical  $\varphi_4^4$  model [12], the criterion provides almost matching upper and lower bounds on how fast the log-Sobolev constant vanishes at the critical point. For the nearest-neighbour Ising model in  $d \geq 5$ , the criterion implies polynomial bounds on the log-Sobolev constant at and near the critical temperature [18].

**6.1. Applications to Euclidean field theory.** In this section, let  $\Lambda = \Lambda_{\varepsilon, L}$  be a discrete torus of mesh size  $\varepsilon$  and side length  $L$  (assumed to be a multiple of  $\varepsilon$ ):  $\Lambda_{\varepsilon, L} = L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$ . The discrete Laplacian on  $\Lambda_{\varepsilon, L}$  is given by

$$(6.2) \quad \forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon, L}}, \quad (\Delta^\varepsilon \varphi)_x = \varepsilon^{-2} \sum_{y \sim x} (\varphi_y - \varphi_x).$$

The (lattice regularised) Euclidean field theory models we consider are of the form

$$(6.3) \quad \nu^{\varepsilon, L}(d\varphi) \propto \exp \left[ -\frac{\varepsilon^d}{2} \sum_{x \in \Lambda_{\varepsilon, L}} \varphi_x (-\Delta^\varepsilon \varphi)_x + \varepsilon^d \sum_{x \in \Lambda_{\varepsilon, L}} V^\varepsilon(\varphi_x) \right] d\varphi,$$

where  $d\varphi$  denotes the Lebesgue measure on  $\mathbb{R}^{\Lambda_{\varepsilon, L}}$  and  $V^\varepsilon$  is a real-valued function chosen in such a way that the  $\varepsilon \rightarrow 0$  limit of the measure (on a suitable space of generalised functions) exists and is non-Gaussian. Writing  $\nabla V^\varepsilon(\varphi) = ((V^\varepsilon)'(\varphi_x))_x$  for  $\varphi \in \mathbb{R}^{\Lambda_{\varepsilon, L}}$ , the dynamics is the singular SPDE

$$(6.4) \quad d\varphi_t = \Delta^\varepsilon \varphi_t dt - \nabla V^\varepsilon(\varphi_t) dt + \sqrt{2} dW_t^{\varepsilon, L}$$

where  $dW_t^{\varepsilon, L}$  is space-time white noise on  $\mathbb{R}_+ \times \Lambda_{\varepsilon, L}$ , i.e., the  $W_x^{\varepsilon, L}$  are independent Brownian motions with variance  $\varepsilon^{-d}$  for  $x \in \Lambda_{\varepsilon, L}$ , or equivalently a standard Brownian motion with respect

to the continuum inner product  $(u, v)_\varepsilon = \varepsilon^d \sum_{x \in \Lambda_{\varepsilon, L}} u_x v_x$ . The pathwise (short time) limit theory for such SPDEs is the subject of Hairer's regularity structure theory [52–54], the paracontrolled method of Gubinelli et al. [26, 50, 92], and the pathwise renormalisation group approach [39, 61]. The log-Sobolev inequality for the associated dynamics takes the form:

$$(6.5) \quad \text{Ent}_{\nu^{\varepsilon, L}}(F) \leq \frac{2}{\gamma} D_{\nu^{\varepsilon, L}}(\sqrt{F}),$$

with the standard Dirichlet form with respect to the gradient  $\nabla_\varepsilon$  corresponding to  $(\cdot, \cdot)_\varepsilon$ :

$$(6.6) \quad D_{\nu^{\varepsilon, L}}(F) = \mathbb{E}_{\nu^{\varepsilon, L}} \left[ (\nabla^\varepsilon F, \nabla^\varepsilon F)_\varepsilon \right] = \frac{1}{\varepsilon^d} \sum_{x \in \Lambda_{\varepsilon, L}} \mathbb{E}_{\nu^{\varepsilon, L}} \left[ \left| \frac{\partial F}{\partial \varphi_x} \right|^2 \right],$$

i.e.,  $(\nabla^\varepsilon F)_x = (\nabla_\varphi^\varepsilon F)_x = \varepsilon^{-d} \frac{\partial F}{\partial \varphi_x}$ . (Thus this gradient acts on functionals of fields  $F : \mathbb{R}^{\Lambda_{\varepsilon, L}} \rightarrow \mathbb{R}$  while the Laplacian (6.2) acts on fields  $\varphi \in \mathbb{R}^{\Lambda_{\varepsilon, L}}$ .)

We now discuss two prototypical models.

*Continuum sine-Gordon model.* Let  $d = 2$ . For  $0 < \beta < 8\pi$  and  $z \in \mathbb{R}$ , the sine-Gordon model is defined by the potential

$$(6.7) \quad V^\varepsilon(\varphi) = 2z\varepsilon^{-\beta/4\pi} \cos(\sqrt{\beta}\varphi), \quad (\varphi \in \mathbb{R}).$$

One can also add a convex quadratic part to the measure: the massive sine-Gordon model with mass  $m > 0$  corresponds to the potential  $V^\varepsilon(\varphi) + \frac{1}{2}m^2\varphi^2$ .

*Continuum  $\varphi^4$  model.* Let  $d = 2$  or  $d = 3$ . For  $\lambda > 0$  and  $\mu \in \mathbb{R}$ , the  $\varphi_d^4$  measure is defined by

$$(6.8) \quad V^\varepsilon(\varphi) = \frac{\lambda}{4}\varphi^4 + \frac{\mu + a^\varepsilon(\lambda)}{2}\varphi^2, \quad (\varphi \in \mathbb{R}),$$

where  $a^\varepsilon(\lambda)$  is a divergent counterterm. Explicitly, for an arbitrary fixed  $m^2 > 0$ , one can take  $a^\varepsilon(\lambda) = a^\varepsilon(\lambda, m^2)$  with

$$(6.9) \quad a^\varepsilon(\lambda, m^2) := -3\lambda(-\Delta^\varepsilon + m^2)^{-1}(0, 0) + 6\lambda^2 \left\| (-\Delta^\varepsilon + m^2)^{-1}(0, \cdot) \right\|_{L^3(\Lambda_{\varepsilon, L})}^3,$$

and the notation  $\|f\|_{L^p(\Lambda_{\varepsilon, L})}^p = \varepsilon^d \sum_{x \in \Lambda_{\varepsilon, L}} |f(x)|^p$  for  $p > 0$ . The counterterms defined in terms of different  $m^2$  differ by additive constants and thus the choice of  $m^2$  corresponds to a normalisation. In the following we take  $m^2 = 1$  for the definition of the counterterm.

The sine-Gordon and  $\varphi_d^4$  models are defined on the discretised torus  $\Lambda_{\varepsilon, L} = L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$ . As explained in Section 2.7, these models should be thought of as discretised versions of limiting models defined on the continuum torus  $L\mathbb{T}^d$ . Recall that the criterion of Theorem 3.6 asks for a lower bound on the Hessian of the renormalised potential uniformly in the field: for some  $\dot{\lambda}_t \in \mathbb{R}$ ,

$$(6.10) \quad \forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon, L}}, t \geq 0: \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t - \frac{1}{2} \ddot{C}_t \geq \dot{\lambda}_t \dot{C}_t.$$

To obtain information on the dynamics in the  $\varepsilon \rightarrow 0$  limit, one is therefore interested in estimates of the renormalised potential that are uniform in  $\varepsilon, \varphi$ .

For the sine-Gordon model, this was carried out in [13] by providing an explicit description of the renormalised potential at each scale.

**Theorem 6.1.** *For the massive continuum sine-Gordon model with mass  $m > 0$ , let  $A = -\Delta^\varepsilon + m^2 \text{id}$  and  $\dot{C}_t = e^{-tA}$  ( $t \geq 0$ ). Then, if  $\beta < 6\pi$ , there is a constant  $\mu^* = \mu^*(\beta, z, m, L) > 0$  that does not depend on  $\varepsilon, t$ , such that with  $V_0(\varphi) = \varepsilon^2 \sum_{x \in \Lambda_{\varepsilon, L}} 2z \cos(\sqrt{\beta}\varphi_x)$ :*

$$(6.11) \quad \forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon, L}}, t \geq 0: \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t \geq \dot{\mu}_t \dot{C}_t \quad \text{with} \quad \sup_{t \geq 0} \left| \int_0^t \dot{\mu}_s ds \right| \leq \mu^*.$$

Given  $\beta, z, m, L$ , this yields a lower bound on the log-Sobolev constant  $\inf_\varepsilon \gamma_{\beta, z, m}^{\varepsilon, L} > 0$ .

Moreover, the log-Sobolev constant is uniform in the large-scale parameter  $L$  under the following condition. If  $L$  satisfies  $m \geq 1/L$  and the coupling constant  $z$  is such that  $|z| \leq \delta_\beta m^{2+\beta/4\pi}$  for a small enough  $\delta_\beta > 0$ , then  $\inf_\varepsilon \gamma_{\beta, z, m}^{\varepsilon, L} > m^2 - O_\beta(m^{\beta/4\pi}|z|)$ , uniformly in  $L$ .

For the sine-Gordon model, the multiscale Bakry-Émery criterion of Theorem 3.6 is thus seen to provide the optimal independence of  $\varepsilon$  of the log-Sobolev constant in finite volume, as well as of  $L$  under an additional small coupling assumption depending on the external mass.

In the case of the  $\varphi_d^4$  model, a description of the renormalised potential at all scales is more difficult to work out directly. The log-Sobolev inequality was obtained in [19] using a correlation inequality due to Ding, Song, and Sun [38], valid for ferromagnetic interactions (see Proposition 6.12 below). This correlation inequality reduces the proof of the log-Sobolev constant to an estimate of certain correlation functions at field  $\varphi = 0$ , for which many tools are available.

**Theorem 6.2.** *For the continuum  $\varphi^4$  model in  $d = 2, 3$ , let  $A = -\Delta^\varepsilon + \text{id}$  and  $\dot{C}_t = (tA + \text{id})^{-2}$  ( $t \geq 0$ ) and take  $m = 1$  in the definition (6.9) of the counterterm. Then with  $V_0(\varphi) = \varepsilon^d \sum_{x \in \Lambda_{\varepsilon, L}} (\frac{\lambda}{4}\varphi^4 + \frac{\mu-1+a^\varepsilon(\lambda)}{2}\varphi^2)$ :*

$$(6.12) \quad \forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon, L}}, t \geq 0: \quad \dot{C}_t \text{Hess } V_t(\varphi) \dot{C}_t - \frac{1}{2} \ddot{C}_t \geq \left( \frac{1}{t} - \frac{\chi_t}{t^2} \right) \dot{C}_t,$$

with  $\chi_t = \chi_{\lambda, \mu+1/t}^{\varepsilon, L}$  the 0-field susceptibility of the  $\varphi_d^4$  model at scale  $t > 0$ : if  $\langle \cdot \rangle_{\lambda, \mu}^{\varepsilon, L}$  denotes the average under the  $\varphi_d^4$  measure,

$$(6.13) \quad \chi_t := \varepsilon^d \sum_{x \in \Lambda_{\varepsilon, L}} \langle \varphi_0 \varphi_x \rangle_{\lambda, \mu+1/t}^{\varepsilon, L}.$$

This implies that the log-Sobolev constant  $\gamma_{\lambda, \mu}^{\varepsilon, L}$  satisfies  $\inf_{\varepsilon, L} \gamma_{\lambda, \mu}^{\varepsilon, L} > 0$  for any  $\lambda, \mu$  for which the measure does not have a phase transition, and  $\inf_\varepsilon \gamma_{\lambda, \mu}^{\varepsilon, L} > 0$  for any  $\lambda, \mu$  and  $L$  fixed.

In the  $\varphi_d^4$  case, the multiscale criterion gives a log-Sobolev constant bounded uniformly in the volume and  $\varepsilon$  in the entire single-phase region. However, the bound obtained on the log-Sobolev constant has accurate dependence on  $\lambda, \mu$  only far away from the transition, corresponding to values of  $\lambda, \mu$  such that  $\mu > 0$  and when  $\lambda \ll \mu$ .

We conclude the section with a last example, the lattice  $\varphi^4$  model, corresponding to the case  $\varepsilon = 1$  above. Thus its single-site potential is given by:

$$(6.14) \quad V(\varphi) = V^{\varepsilon=1}(\varphi) = \frac{g}{4}\varphi^4 + \frac{\nu}{2}\varphi^2,$$

and the coupling matrix is  $A = -\Delta^{\varepsilon=1} + \text{id}$  setting  $m^2 = 1$  without loss of generality. The state space is  $L\mathbb{T}^d \cap \mathbb{Z}^d$  for any  $d \geq 1$ . As a special case, Theorem 6.2 yields the log-Sobolev inequality whenever the infinite volume susceptibility is finite (though the proof of this statement can be simplified considerably in this case, see [19, Example 3.1]). We nonetheless make an independent statement because the Lipschitz constant of the map of Theorem 5.5 can be bounded uniformly in the volume. Theorem 5.5 does not seem to be effective for the discrete model.

**Example 6.3** ([19, Example 3.1]). For the lattice  $\varphi^4$  model in any dimension with associated expectation  $\langle \cdot \rangle_{g, \nu}^L$ , let  $\dot{C}_t = (tA + \text{id})^{-2}$  and define the lattice susceptibility:

$$(6.15) \quad \chi_t^L := \sum_{x \in \Lambda} \langle \varphi_0 \varphi_x \rangle_{g, \nu+1/t}^L.$$

Then the Hessian bound (6.12) holds with  $\chi_t^L$  in the right-hand side instead of  $\chi_t$ . This implies that the log-Sobolev constant  $\gamma_{g, \nu}^L$  satisfies  $\inf_L \gamma_{g, \nu}^L > 0$  if and only if  $\sup_L \chi_\infty^L < \infty$ .

Moreover, if  $\sup_L \chi_\infty^L < \infty$ , then the transport map of Theorem 5.5 has Lipschitz constant is bounded uniformly in  $L$  and  $t$ .

*Proof.* Let us prove the bound on the Lipschitz constant. Theorem 5.5 gives the following bound on the Lipschitz constant of the transport map:

$$(6.16) \quad \sqrt{t\|A\| + 1} \exp \left[ -\frac{1}{2} \int_0^t \left( \frac{1}{s} - \frac{\chi_s^L}{s^2} \right) ds \right].$$

The Brascamp-Lieb inequality (3.63) ensures:

$$(6.17) \quad \inf_L \inf_{s \leq 1} \left( \frac{1}{s} - \frac{\chi_s^L}{s^2} \right) > 0.$$

For  $t \geq 1$ , the second Griffiths inequality [48] implies that  $t \mapsto \chi_t^L$  is increasing, thus is bounded by  $\chi_\infty^L$  so that

$$(6.18) \quad \frac{1}{t} - \frac{\chi_t^L}{t^2} \geq \frac{1}{t} - \frac{\chi_\infty^L}{t^2}, \quad (t \geq 1).$$

Combining the previous inequalities, we deduce that there is a constant  $C = C(\chi_\infty^L) > 0$  such that for all  $t > 0$ :

$$(6.19) \quad \exp \left[ -\frac{1}{2} \int_0^t \left( \frac{1}{s} - \frac{\chi_s^L}{s^2} \right) ds \right] \leq C \min \{t^{-1/2}, 1\},$$

and the observation  $\|A\| \leq \|\Delta^{\varepsilon=1} + \text{id}\| \leq 4d + 1$  implies that the Lipschitz constant (6.16) is uniformly bounded. This concludes the proof.  $\square$

**6.2. Applications to Ising models.** In this section, we explain how to apply the ideas developed in Section 3 to Ising models with discrete spins. Using the Bakry–Émery criterion and its multiscale version, these ideas were developed in [11, 18], while closely related results were obtained using spectral and entropic independence and stability estimates in [3, 28, 41].

*6.2.1. Renormalised potential.* The Ising model with coupling matrix  $A$  at inverse temperature  $\beta > 0$  and (site-dependent) external field  $h$  on a finite set  $\Lambda$  is defined by:

$$(6.20) \quad \mathbb{E}_\mu[F] = \mathbb{E}_{\mu_{\beta,h}}[F] \propto \sum_{\sigma \in \{\pm 1\}^\Lambda} e^{-\frac{1}{2}(\sigma, \beta A \sigma) + (h, \sigma)} F(\sigma).$$

Since  $\sigma_x^2 = 1$  for each  $x$ , the measure is invariant under the change  $A \rightarrow A + \alpha \text{id}$ ,  $\alpha \in \mathbb{R}$ . Without loss of generality, we can thus assume that the coupling matrix  $A$  is positive definite. We also assume that it has spectral radius bounded by 1; this just amounts to a choice of normalisation for  $\beta$ . A natural covariance decomposition is

$$(6.21) \quad C_t = (tA + (\alpha - t) \text{id})^{-1},$$

for a parameter  $\alpha > 0$  which will be unimportant. For  $t < \alpha$  the matrix  $C_t$  is positive definite and for  $\beta < \alpha$ , as explained above, the Ising model at inverse temperature  $\beta$  can be written as

$$(6.22) \quad \mathbb{E}_\mu[F] \propto \sum_{\sigma \in \{\pm 1\}^\Lambda} e^{-\frac{1}{2}(\sigma, C_\beta^{-1} \sigma) + (h, \sigma)} F(\sigma).$$

Strictly speaking,  $C_t$  is not a covariance decomposition since  $C_0 = \alpha^{-1} \text{id} \neq 0$ , different from the assumption in Section 3. However, all results from that section can be applied to the covariance decomposition  $C_t - C_0$ , and we will do this without further emphasis. The renormalised potential can be defined analogously to the continuous setting as:

$$(6.23) \quad V_t(\varphi) = -\log \sum_{\sigma \in \{\pm 1\}^\Lambda} e^{-\frac{1}{2}(\sigma - \varphi, C_t^{-1}(\sigma - \varphi)) + (h, \sigma)}.$$



This leads to a decomposition of the Ising measure (6.20) as:

$$(6.24) \quad \mathbb{E}_\mu[F] = \mathbb{E}_{\nu_{t,\beta}}[\mathbb{E}_{\mu_t^\varphi}[F]] \quad (0 \leq t < \beta).$$

By analogy with (3.18), the fluctuation measure  $\mu_t^\varphi$  used above is the Ising measure with coupling matrix  $C_t^{-1}$  and external field  $C_t^{-1}\varphi + h$ :

$$(6.25) \quad \mu_t^\varphi(\sigma) = \mu_{t, C_t^{-1}\varphi+h} \propto e^{-\frac{1}{2}(\sigma, C_t^{-1}\sigma) + (C_t^{-1}\varphi+h, \sigma)},$$

and the renormalised measure  $\nu_t = \nu_{t,\beta}$  is supported on the image of  $C_\beta - C_t$  in  $\mathbb{R}^\Lambda$ :

$$(6.26) \quad \nu_{t,\beta}(d\varphi) \propto e^{-V_t(\varphi)} \mathbb{P}_{C_\beta - C_t}(d\varphi) \propto \exp\left[-\frac{1}{2}(\varphi, (C_\beta - C_t)^{-1}\varphi) - V_t(\varphi)\right] d\varphi.$$

Even though  $\sigma : \Lambda \rightarrow \{\pm 1\}$  is discrete, the renormalised field  $\varphi : \Lambda \rightarrow \mathbb{R}$  is continuous as soon as  $t > 0$ . Convexity-based criteria, such as the Bakry-Émery or multiscale Bakry-Émery criteria of Theorems 2.8 and 3.6, can therefore be used to derive log-Sobolev inequalities for  $\nu_{t,\beta}$ .

Before discussing these, we summarise results about the infinite temperature (product) Ising model, which serve as input to these arguments.

*6.2.2. Preliminaries: single-spin inequalities.* At infinite temperature  $\beta = 0$  the spin models we consider become product measures. By tensorisation, it thus suffices to know the log-Sobolev constant of a single spin (see Example 2.9). The following summarises known results for these.

*Ising model, standard Dirichlet form.* Let  $\mu$  be the probability measure on  $\{\pm 1\}$  with  $\mu(+1) = p = 1 - q$ . The standard Dirichlet form is

$$(6.27) \quad D_\mu(F) = \frac{1}{2}\mathbb{E}_\mu(F(\sigma^x) - F(\sigma))^2 = \frac{1}{2}(F(+1) - F(-1))^2.$$

**Proposition 6.4.** *Let  $\mu$  be the probability measure on  $\{\pm 1\}$  with  $\mu(+1) = p = 1 - q$ . Then*

$$(6.28) \quad \text{Ent}_\mu(F) \leq \frac{pq(\log p - \log q)}{p - q}(\sqrt{F(+1)} - \sqrt{F(-1)})^2.$$

*Thus the log-Sobolev constant with respect to the standard Dirichlet form on  $\{\pm 1\}$  is at least 2:*

$$(6.29) \quad \text{Ent}_\mu(F) \leq D(\sqrt{F}) = \frac{2}{\gamma_0}D(\sqrt{F}), \quad \gamma_0 = 2.$$

*Proof.* The proof can be found in [81] or [5]. □

*Ising model, heat bath Dirichlet form.* With  $\mu$  as above, the heat-bath Dirichlet form is

$$(6.30) \quad D^{\text{HB}}(F) = \frac{1}{2} \sum_{\sigma} \frac{\mu(\sigma)\mu(\sigma^x)}{\mu(\sigma) + \mu(\sigma^x)} (F(\sigma^x) - F(\sigma))^2 = pq(F(+1) - F(-1))^2$$

**Proposition 6.5.** *Let  $\mu$  be the probability measure on  $\{\pm 1\}$  with  $\mu(+1) = p = 1 - q$ . Then*

$$(6.31) \quad \text{Ent}_\mu(F) \leq pq(\log F(+1) - \log F(-1))(F(+1) - F(-1)).$$

*Thus the modified log-Sobolev constant with respect to the heat-bath Dirichlet form is at least 1/2:*

$$(6.32) \quad \text{Ent}_\mu F \leq D^{\text{HB}}(\log F, F) = \frac{1}{2\gamma_0}D^{\text{HB}}(\log F, F), \quad \gamma_0 = \frac{1}{2}.$$

*Proof.* See [22, Example 3.8]. □

Similar single-spin inequalities are available for  $O(n)$  models [84, 91] and allow to extend for example Theorem 6.7 below for the Ising model to these models with little change [11].

*6.2.3. Entropy decomposition.* To prove a log-Sobolev inequality (or modified log-Sobolev inequality) for the Ising measure, we start from the decomposition (6.24) with  $t = 0$ , decomposing  $\mu = \mu_\beta^0$  into two parts: an infinite temperature Ising measure  $\mu_0^\varphi$  with external field  $C_0^{-1}\varphi + h$  and the renormalised measure  $\nu_{0,\beta}$ . The corresponding entropy decomposition (2.52) is:

$$(6.33) \quad \text{Ent}_\mu(F) = \mathbb{E}_{\nu_{0,\beta}}[\text{Ent}_{\mu_0^\varphi}(F)] + \text{Ent}_{\nu_{0,\beta}}(\mathbb{E}_{\mu_0^\varphi}[F]).$$

To prove the log-Sobolev inequality (or modified log-Sobolev inequality) with respect to a Dirichlet form  $D_\mu$ , we want to bound both terms by a multiple of  $D_\mu(\sqrt{F})$  (or  $D_\mu(F, \log F)$ ). In this discussion, we focus on the log-Sobolev inequality with respect to the standard Dirichlet form:

$$(6.34) \quad D_\mu(F) = \frac{1}{2} \sum_{x \in \Lambda} \mathbb{E}_\mu \left[ (F(\sigma) - F(\sigma^x))^2 \right].$$

However, the same strategy applies with different jump rates and a modified log-Sobolev inequality or a spectral inequality as input (and output), see Sections 6.2.2 and 6.2.5.

Under the assumption that  $\mu_0^\varphi$  satisfies a log-Sobolev inequality with constant  $\gamma_0$  uniformly in  $\varphi$ , the first term on the right-hand side of (6.33) is bounded by

$$(6.35) \quad \frac{2}{\gamma_0} \mathbb{E}_{\nu_{0,\beta}}[D_{\mu_0^\varphi}(\sqrt{F})] = \frac{2}{\gamma_0} D_\mu(\sqrt{F}).$$

Since  $\mu_0^\varphi$  is a product measure, this assumption is well understood and one can take  $\gamma_0 = 2$  for the standard Dirichlet form, as discussed in Section 6.2.2. The last equality relies on the specific jump rates in the standard Dirichlet form (6.34). For other Dirichlet forms, one can also get an inequality, see Section 6.2.5.

Bounding the second term on the right-hand side of (6.33) essentially amounts to estimating the log-Sobolev constant of the renormalised measure  $\nu_{0,\beta}$ . Indeed, if  $\nu_{0,\beta}$  satisfies a log-Sobolev inequality with constant  $\gamma_{0,\beta}$  (and standard Dirichlet form) then this term is bounded by

$$(6.36) \quad \frac{2}{\gamma_{0,\beta}} \mathbb{E}_{\nu_{0,\beta}} \left[ |(\nabla_\varphi \mathbb{E}_{\mu_0^\varphi}[F]^{1/2})|^2 \right] \leq \frac{4\alpha^2}{\gamma_0 \gamma_{0,\beta}} D_\mu(\sqrt{F}),$$

where the second inequality is elementary and follows from the next exercise (which is similar to the argument in the tensorisation proof of Example 2.9) and the Cauchy–Schwarz inequality.

**Exercise 6.6.** For each  $x \in \Lambda$ , let  $\mu_0^{\varphi,x}$  denote the law of  $\sigma_x$  under the product measure  $\mu_0^\varphi$ . Then:

$$(6.37) \quad \nabla_{\varphi,x}(\mathbb{E}_{\mu_0^\varphi}[F]^{1/2}) = \frac{\alpha}{2\sqrt{\mathbb{E}_{\mu_0^\varphi}[F]}} \text{Cov}_{\mu_0^\varphi}(F, \sigma_x) = \frac{\alpha}{2\sqrt{\mathbb{E}_{\mu_0^\varphi}[F]}} \mathbb{E}_{\mu_0^\varphi} \left[ \text{Cov}_{\mu_0^{\varphi,x}}(F, \sigma_x) \right]$$

(recall  $C_0 = \alpha^{-1} \text{id}$ ) and:

$$(6.38) \quad \text{Cov}_{\mu_0^{\varphi,x}}(F, \sigma_x)^2 \leq 8 \mathbb{E}_{\mu_0^{\varphi,x}}[F] \text{Var}_{\mu_0^{\varphi,x}}(\sqrt{F}) \leq \frac{8}{\gamma_0} \mathbb{E}_{\mu_0^{\varphi,x}}[F] D_{\mu_0^{\varphi,x}}(\sqrt{F}).$$

*Sketch.* The first equation is a simple computation. The first inequality in (6.38) follows from Cauchy–Schwarz inequality using the following general expression for the covariance of functions  $G_1, G_2$  under a measure  $m$ :

$$(6.39) \quad \text{Cov}_m(G_1, G_2) = \frac{1}{2} \int (G_1(x) - G_1(y))(G_2(x) - G_2(y)) dm(x) dm(y).$$

The second inequality is the general fact that the spectral gap is always larger than the log-Sobolev constant, see Proposition 2.5. Details can be found in [11].  $\square$

In summary, if  $\mu_0^\varphi$  satisfies a log-Sobolev inequality with constant  $\gamma_0$  and  $\nu_{0,\beta}$  satisfies a log-Sobolev inequality with constant  $\gamma_{0,\beta}$  then the inverse log-Sobolev constant  $\gamma^{-1}$  of  $\mu$  satisfies

$$(6.40) \quad \frac{1}{\gamma} \leq \frac{1}{\gamma_0} \left[ 1 + \frac{2\alpha^2}{\gamma_{0,\beta}} \right].$$

At this point the objective is to bound the log-Sobolev constant  $\gamma_{0,\beta}$  of the renormalised measure  $\nu_{0,\beta}$ . Results are stated next under different conditions, with the corresponding verifications postponed to Section 6.2.4.

Recall that the renormalised measure is  $\nu_{0,\beta}(d\varphi) \propto e^{-V_0(\varphi)} \mathbf{P}_{C_\beta - C_0}(d\varphi)$  with the renormalised potential  $V_0$  defined in (6.23). If the temperature is sufficiently high, namely if  $\beta < \alpha < 1$ , then it turns out that  $V_0$  is strictly convex so that the standard Bakry–Émery criterion is applicable [11] and gives (see Exercise 6.10):

$$(6.41) \quad \gamma_{0,\beta} \geq \alpha - \alpha^2.$$

Taking  $\alpha \downarrow \beta$ , this leads to the following theorem. We recall the convention fixed below (6.20) that the coupling matrix  $A$  has spectrum in  $[0, 1]$ . This spectral condition appeared in [11] (and unlike previously existing conditions applies to the SK spin glass).

**Theorem 6.7** (Spectral high temperature condition [11]). *For  $\beta < 1$  (under the conventions stated below (6.20)), the Ising model  $\mu$  satisfies the log-Sobolev inequality: for each  $F : \{-1, 1\}^\Lambda \rightarrow \mathbb{R}_+$ ,*

$$(6.42) \quad \text{Ent}_\mu(F) \leq \left( 1 + \frac{2\beta}{1-\beta} \right) D_\mu(\sqrt{F}).$$

For  $\beta > 1$ , the renormalised potential  $V_0$  is in general not convex and the log-Sobolev constant  $\gamma_{0,\beta}$  of the renormalised measure  $\nu_{0,\beta}$  cannot be bounded using the Bakry–Émery criterion. However, the argument above can be generalised by using the multiscale Bakry–Émery criterion instead. Assuming  $\dot{\lambda}_t$  are as in the multiscale Bakry–Émery condition (6.1), Theorem 3.6 gives:

$$(6.43) \quad \frac{1}{\gamma_{0,\beta}} \leq |\dot{C}_0| \int_0^\beta e^{-2\lambda_t} dt = \frac{1}{\alpha^2} \int_0^\beta e^{-2\lambda_t} dt, \quad \lambda_t = \int_0^t \dot{\lambda}_s ds,$$

and substituting this into (6.40) gives the following log-Sobolev inequality:

$$(6.44) \quad \text{Ent}_\mu(F) \leq \frac{2}{\gamma} D_\mu(\sqrt{F}), \quad \frac{1}{\gamma} \leq \frac{1}{\gamma_0} \left[ 1 + 2 \int_0^\beta e^{-2\lambda_t} dt \right], \quad \lambda_t = \int_0^t \dot{\lambda}_s ds.$$

In the same situation, instead of using the multiscale Bakry–Émery criterion to bound  $\gamma_{0,\beta}$  in (6.40), one can prove a log-Sobolev inequality for  $\mu$  by using the entropic stability estimate discussed in Section 3.7. This was done in [28] (to prove a modified log-Sobolev inequality, but the argument generalises to a log-Sobolev inequality for the standard Dirichlet form, see below). It turns out that, for the decomposition (6.21), the conditions (3.32) and (3.107) of the multiscale Bakry–Émery criterion and of the entropic stability estimate are *identical* provided  $\dot{\lambda}_t = -\alpha_t$ , see Exercise 6.11. This is not the case for other covariance decompositions, and in particular not for those used for continuous models, see the discussion in Section 3.7. Using the entropic stability estimate (3.109), the entropy is therefore bounded by:

$$(6.45) \quad \text{Ent}_\mu(F) \leq e^{-\lambda_\beta} \mathbb{E}_{\nu_{0,\beta}}[\text{Ent}_{\mu_0^\varphi}(F)], \quad \lambda_t = \int_0^t \dot{\lambda}_s ds,$$

and with the uniform log-Sobolev inequality for  $\mu_0^\varphi$  this gives

$$(6.46) \quad \text{Ent}_\mu(F) \leq \frac{2}{\gamma} D_\mu(\sqrt{F}), \quad \frac{1}{\gamma} \leq \frac{1}{\gamma_0} e^{-\lambda_\beta}, \quad \lambda_t = \int_0^t \dot{\lambda}_s ds.$$

This estimate is very similar to the estimate (6.44) obtained using the multiscale Bakry–Émery criterion, but not exactly identical. Both estimates can be applied up to the critical point in a very general setting for ferromagnetic Ising models and yield a polynomial bound on the log-Sobolev constant under the mean-field bound which holds on  $\Lambda \subset \mathbb{Z}^d$  in  $d \geq 5$ , see Theorem 6.13 below. The strategies of the two proofs have different advantages. We summarise the results as follows.

**Theorem 6.8** (Covariance conditions for Ising models [18, 28]). *The log-Sobolev constant of the Ising model at inverse temperature  $\beta$  is bounded by (6.44) or (6.46).*

As discussed previously, we formulated the results for the log-Sobolev inequality with respect to the standard Dirichlet form. This is a canonical choice (as already explained in Section 2.2), but the argument can be adapted easily to other choices of jump rates with the conclusion of a possibly modified log-Sobolev inequality, see Section 6.2.5. As pointed out in [41], other choices are of interest when the jump rates are unbounded.

*6.2.4. Hessian of the renormalised potential and covariance.* In both strategies, using the multiscale Bakry–Émery criterion or the entropic stability estimate, the estimate of the log-Sobolev constant reduces to estimating the constants  $\dot{\lambda}_t = -\alpha_t$  bounding the Hessian of the renormalised potential from below. From Lemma 3.12, recall that these estimates are equivalent to bounds on the covariance of the fluctuation measure, a point of view that is particularly useful for the Ising model. Indeed, the Hessian of the renormalised potential can be represented as follows.

**Exercise 6.9.** Show that

$$(6.47) \quad \text{Hess } V_t(\varphi) = C_t^{-1} - C_t^{-1} \Sigma_t (C_t^{-1} \varphi + h) C_t^{-1},$$

where  $\Sigma_t(g) = (\text{Cov}_{\mu_{t,g}}(\sigma_x, \sigma_y))_{x,y \in \Lambda}$  is the covariance matrix of the Ising model  $\mu_{t,g}$  at inverse temperature  $t$  and site-dependent magnetic field  $g$  (so that (6.25) reads  $\mu_t^\varphi = \mu_{t, C_t^{-1} \varphi + h}$ ).

For  $\beta < 1$ , one can obtain the following convexity directly from this representation, which allows to apply the standard Bakry–Émery criterion to derive (6.41) and conclude the proof of Theorem 6.7.

**Exercise 6.10.** Let  $1 \geq \alpha > \beta$ , and set  $C_t = (tA + (\alpha - t)\text{id})^{-1}$ . Then  $V_t$  is convex for all  $t \in [0, \alpha]$ .

*Proof.* Since  $\mu_{0,g}$  is a product measure and  $|\sigma_x| \leq 1$ ,

$$(6.48) \quad \Sigma_0(g) = \text{diag}(\text{Var}_{\mu_{0,g}}(\sigma_x))_{x \in \Lambda} \leq \text{id}.$$

Using that  $C_0 = \alpha^{-1} \text{id}$ , we deduce from (6.47) that

$$(6.49) \quad \text{Hess } V_0(\varphi) = (\alpha - \alpha^2 \Sigma_0(\alpha\varphi + h)) \text{id} \geq (\alpha - \alpha^2) \text{id}.$$

Thus if  $\alpha \leq 1$ , it follows that  $V_0$  is convex. By Proposition 3.13,  $V_t$  is convex for all  $t > 0$ .  $\square$

For general  $\beta > 0$ , the semi-convexity criterion of the Hessian can be equivalently formulated as covariance estimate uniformly in an external field.

**Exercise 6.11.** From  $C_t = (tA + (\alpha - t)\text{id})^{-1}$ , one has  $\dot{C}_t = (\text{id} - A)C_t^2$ ,  $\ddot{C}_t = 2(\text{id} - A)\dot{C}_t C_t$  and thus the multiscale Bakry–Émery criterion (3.32) and the entropic stability criterion (3.107) hold with  $-\dot{\lambda}_t = \alpha_t = \bar{\chi}_t$  where  $\bar{\chi}_t$  is a uniform upper bound on the spectral radius of the covariance matrix of an Ising model uniformly in an external field:

$$(6.50) \quad \bar{\chi}_t = \sup_{g \in \mathbb{R}^\Lambda} \chi_t(g), \quad \chi_t(g) = \|\Sigma_t(g)\|.$$

Now a significant simplification occurs for Ising models with ferromagnetic interaction, meaning  $A_{x,y} \leq 0$  for  $x \neq y$ . This includes the case of the lattice Laplacian  $\Delta\sigma_x = \sum_{y \sim x} [\sigma_y - \sigma_x]$ . For ferromagnetic interactions, it turns out that the spectral radius is maximal at 0 field:

$$(6.51) \quad \bar{\chi}_t = \chi_t(0).$$

This is a consequence of the FKG inequality (which implies that the covariance matrix has point-wise nonnegative coefficients), the Perron–Frobenius theorem (which therefore implies that the largest eigenvector has nonnegative entries), and the following remarkable correlation inequality due to Ding–Song–Sun [18] which implies that the covariance between any two spins is maximised at 0 field.

**Proposition 6.12** (Ding–Song–Sun inequality [38, Corollary 1.3]). *Let  $\mu = \mu_{\beta,h}$  be the Ising measure (6.20) with ferromagnetic interaction  $A$ , and external field  $h \in [-\infty, \infty]^\Lambda$ , with values  $\pm\infty$  corresponding to boundary conditions. Then:*

$$(6.52) \quad \text{Cov}_{\mu_{\beta,h}}(\sigma_x, \sigma_y) \leq \text{Cov}_{\mu_{\beta,0}}(\sigma_x, \sigma_y) = \mathbb{E}_{\mu_{\beta,0}}[\sigma_x \sigma_y], \quad (x, y \in \Lambda).$$

In particular, if the interaction  $A$  is ferromagnetic and (for simplicity) in addition translation invariant, i.e.,  $A_{x,y} = A_{x-y}$ , then

$$(6.53) \quad \bar{\chi}_t = \chi_t = \chi_t(0) = \sum_{x \in \Lambda} \mathbb{E}_{\mu_{t,0}}[\sigma_x \sigma_y]$$

is the *susceptibility* of the Ising model. It characterises the phase transition of the ferromagnetic Ising model, in the sense that, e.g., for  $\Lambda \subset \mathbb{Z}^d$  ( $d \geq 2$ ), the critical value  $\beta_c$  of  $\beta$  satisfies:

$$(6.54) \quad \beta_c := \sup \left\{ \beta > 0 : \sup_{\Lambda \uparrow \mathbb{Z}^d} \sum_{x \in \Lambda} \mathbb{E}_{\mu_{\beta,0}}[\sigma_0 \sigma_x] < \infty \right\}.$$

By combining the Ding–Song–Sun correlation bound with the multiscale Bakry–Émery criterion, and in view of the above characterisation of  $\beta_c$ , the following log-Sobolev inequality up to the critical point for ferromagnetic Ising models on general geometries was proven in [18].

**Theorem 6.13** ([18, Theorem 1.1]). *The log-Sobolev constant  $\gamma_{\beta,h}$  of the Ising measure (6.20) with ferromagnetic interaction satisfies:*

$$(6.55) \quad \frac{1}{\gamma_{\beta,h}} \leq \frac{1}{2} + \int_0^\beta e^{2 \int_0^t \chi_s ds} dt,$$

with  $\chi_\beta = \sup_x \sum_y \mathbb{E}_{\mu_{\beta,0}}[\sigma_x \sigma_y]$  or more generally the largest eigenvalue of  $(\mathbb{E}_{\mu_{\beta,0}}[\sigma_x \sigma_y])_{x,y}$ .

The above result implies that  $\gamma_{\beta,h}$  is bounded below uniformly in the size of the lattice as long as  $\beta < \beta_c$ . For special geometries, this could already be argued from [38] (which establishes the strong spatial mixing property).

More interesting is the fact that (6.55) also gives an explicit bound on the log-Sobolev constant. For Ising models with mean-field interactions, the bound implies that  $\gamma_{\beta,h}$  is of order  $\beta_c - \beta$ , which is the correct scaling. More significantly, the bound (6.55) implies a polynomial bound on  $1/\gamma_{\beta,h}$  on  $\Lambda \subset \mathbb{Z}^d$  in dimension five and higher (and more generally under the so-called mean-field bound on the susceptibility, i.e., when the susceptibility diverges linearly), where polynomial means as a function of  $\beta_c - \beta$  when  $\beta < \beta_c$ , and of the lattice size when  $\beta = \beta_c$ . The degree of this polynomial is not expected to be sharp unless  $D = 1$ .

**Example 6.14.** If  $\chi_\beta \leq D/(\beta_c - \beta)$  then  $\gamma_{\beta,h}$  is bounded polynomially in  $\beta_c - \beta$ , and if  $\chi_\beta \leq D/(\beta_c - \beta + L^{-\alpha})$  then  $\gamma_{\beta_c,h}$  is polynomial in  $L$ . These assumptions hold for the ferromagnetic nearest-neighbour Ising model on  $\Lambda \subset \mathbb{Z}^d$  when  $d \geq 5$  [1].

Similarly, one can recover the main result of [73] from (6.55), which improves the high temperature condition of (6.42) in the case of ferromagnetic Ising models on graphs with maximal degree  $d$  (for the mixing time and spectral gap). For the ferromagnetic Ising model,  $A$  is the negative adjacency matrix of the graph which we now (for comparison) *do not* normalise to have spectrum contained in  $[0, 1]$ . The condition of (6.42) cannot be improved for general non-ferromagnetic interactions with bounded spectral radius, but for ferromagnetic models the condition established in [73] is  $\beta < \operatorname{artanh}(1/(d-1))$ , whereas the condition of (6.42) translates in this case to  $\beta < 1/(2d)$ . The value  $\beta_u = \operatorname{artanh}(1/(d-1))$  is the uniqueness threshold for the ferromagnetic Ising model on the infinite  $d$ -regular tree and also the critical point for ferromagnetic Ising models on random  $d$ -regular graphs [73]. This application is summarised in the next example.

**Example 6.15.** Let  $A$  be the negative adjacency matrix of a finite graph of maximal degree  $d$  (not normalised to have spectrum in  $[0, 1]$ ). If  $(d-1) \tanh \beta < 1$  then  $\gamma_{\beta, h}$  is uniformly bounded below (and diverges polynomially in  $\beta \uparrow \beta_u = \operatorname{artanh}(1/(d-1))$  and in the size of the graph if  $\beta = \beta_u$ ).

*Sketch.* From [73, Section 5.4, Eq. (39)], we deduce that for any sites  $x, y$  in the graph:

$$(6.56) \quad \mathbb{E}_{\mu_{\beta, 0}}[\sigma_x \sigma_y] \leq \sum_{w \in \mathcal{S}(y)} (\tanh \beta)^{\operatorname{dist}(x, w)},$$

where the bound is obtained by comparing the graph with a  $d$ -regular tree as in [88] and  $\mathcal{S}(y)$  stands for the set of sites associated with  $y$  in this tree. Summing over the sites  $y$  in the graph boils down to sum over all the sites in  $\cup_y \mathcal{S}(y)$ , i.e., in the  $d$ -regular tree. Thus we get

$$(6.57) \quad \chi_\beta = \sup_x \sum_y \mathbb{E}_{\mu_{\beta, 0}}[\sigma_x \sigma_y] \leq 1 + C_d \sum_{\ell \geq 1} ((d-1) \tanh \beta)^\ell \leq \frac{C_d}{1 - (d-1) \tanh \beta},$$

i.e., one finds a divergence of the susceptibility as in Example 6.14 as  $\beta \uparrow \beta_u = \operatorname{artanh}(1/(d-1))$ . Moreover, when  $\beta$  approaches  $\beta_u$ , one can use that for a graph with  $N$  sites then trivially  $\chi_\beta \leq N$ , so that

$$(6.58) \quad \chi_\beta \leq \frac{C_d}{1 - (d-1) \tanh \beta + N^{-1}}.$$

By Theorem 6.13, we deduce a polynomial lower bound for the log-Sobolev constant in the size of the graph at  $\beta = \beta_u = \operatorname{artanh}(1/(d-1))$ .  $\square$

*6.2.5. Choice of Dirichlet form.* As in the original references [11, 18], the above discussion is formulated in terms of the standard Dirichlet form (6.34). There exist general comparison arguments between Dirichlet forms that ensure that one can transfer the log-Sobolev inequality obtained for a certain dynamics to another one, see, e.g., Chapter 4 in [81]. Namely, if  $c_1, c_2$  are families of jump rates reversible with respect to the same measure  $\nu$  on  $\{-1, 1\}^\Lambda$  and  $\gamma_1, \gamma_2$  denote the associated log-Sobolev constants, then, for  $K > 0$ :

$$(6.59) \quad \forall \sigma, \sigma' \in \{-1, 1\}^\Lambda, \quad K^{-1} c_2(\sigma, \sigma') \leq c_1(\sigma, \sigma') \leq K c_2(\sigma, \sigma') \quad \Rightarrow \quad K \gamma_2 \leq \gamma_1 \leq \gamma_2 / K.$$

One can for instance check that the heat-bath- and canonical jump rates associated with an Ising measure with interaction  $\beta A$  and external field  $h \in \mathbb{R}^\Lambda$  satisfy such a bound, with a constant  $K$  that depends only on  $\|h\|_\infty$  and  $\beta \max_i \sum_j |A_{i,j}|$ . The heat-bath Dirichlet form is [68, (3.4)–(3.6)]:

$$(6.60) \quad D_\mu^{\text{HB}}(F) = \frac{1}{2} \sum_{\sigma \in \{\pm 1\}^\Lambda} \sum_{x \in \Lambda} \Psi(\mu(\sigma), \mu(\sigma^x)) (F(\sigma) - F(\sigma^x))^2, \quad \Psi(a, b) = \frac{ab}{a+b}.$$

There are however situations in which the comparison argument (6.59) is not applicable. This observation was made in [41] in the situation of the SK model treated in [11], which is an Ising

model with random couplings which can take arbitrarily large values. This makes the constant  $K$  in (6.59) arbitrarily small. In such cases, if one is interested in a dynamics different from the one induced by the canonical jump rates, it is desirable to directly obtain the log-Sobolev inequality (or modified log-Sobolev inequality) for this dynamics. This was done for the heat-bath dynamics, for the modified log-Sobolev inequality, in [3, 28, 41].

One can however check that the arguments of [11, 18] sketched above are not specific to the standard Dirichlet form. It is in fact straightforward to apply them to other dynamics as sketched below in the heat-bath case, provided:

- the associated single-spin LSI constant (or modified LSI constant) is uniform in the field;
- the associated Dirichlet form is a concave function of the measure (this can be generalised).

Let us show how this works in the heat-bath case (6.60), for which both points are satisfied in view of Proposition 6.5 and the concavity of  $\Psi(a, b) = ab/(a + b)$ . Instead of using the single-spin log-Sobolev inequality for the standard Dirichlet form of Proposition 6.4, one can instead apply the single-spin modified LSI with the heat-bath Dirichlet form of Proposition 6.5 since it also holds uniformly in the external field. For instance, for  $\mathbb{E}_{\nu_{0,\beta}}[\text{Ent}_{\mu_0^\varphi}(F)]$  the bound becomes:

$$\begin{aligned}
& \mathbb{E}_{\nu_{0,\beta}}[\text{Ent}_{\mu_0}(F(\sigma))] \\
& \leq \frac{1}{2} \sum_{\sigma} \sum_x \mathbb{E}_{\nu_{0,\beta}}[\Psi(\mu_0(\sigma), \mu_0(\sigma^x))](F(\sigma) - F(\sigma^x))(\log F(\sigma) - \log F(\sigma^x)) \\
(6.61) \quad & \leq \frac{1}{2} \sum_{\sigma} \sum_x \Psi(\mu(\sigma), \mu(\sigma^x))(F(\sigma) - F(\sigma^x))(\log F(\sigma) - \log F(\sigma^x)) = D_{\mu}^{\text{HB}}(F, \log F),
\end{aligned}$$

where the first inequality is the single-spin modified log-Sobolev inequality from Proposition 6.5, and the second inequality is Jensen's inequality, using  $\mu(\sigma) = \mathbb{E}_{\nu_{0,\beta}}[\mu_0(\sigma)]$  and that  $\Psi$  is concave. The bound for the other term in (6.33) works analogously and gives  $4D_{\mu}^{\text{HB}}(\sqrt{F}) \leq D_{\mu}^{\text{HB}}(F, \log F)$  instead of  $D_{\mu}(\sqrt{F})$  with the standard Dirichlet form.

The conclusion is that the modified log-Sobolev constant for the heat-bath Dirichlet form satisfies exactly the same bound (up to an overall factor 4 from different normalisations):

$$(6.62) \quad \text{Ent}_{\mu}(F) \leq \frac{1}{2\gamma_{\text{HB}}} D_{\mu}^{\text{HB}}(F, \log F), \quad \frac{1}{\gamma_{\text{HB}}} \leq 2 + 4 \int_0^{\beta} e^{-2\lambda t} dt.$$

This strategy can be further generalised to Dirichlet forms which are not concave in the measure, see [16].

**6.3. Applications to conservative dynamics.** The criterion of Theorem 3.6 in principle also applies to dynamics with a conservation law. This is for instance the case for spin models with constrained magnetisation:

$$(6.63) \quad \mu_{N,m}(d\varphi) \propto e^{-V_0(\varphi)} d\varphi|_{X_{N,m}},$$

with  $X_{N,m}$  the hyperplane of spins with magnetisation  $m \in \mathbb{R}$ :

$$(6.64) \quad X_{N,m} := \left\{ \varphi \in \mathbb{R}^N : \sum_i \varphi_i = Nm \right\}.$$

For instance, an infinite temperature example is given by  $V_0 = \sum_{i=1}^N V(\varphi_i)$  and  $V : \mathbb{R} \rightarrow \mathbb{R}$  a  $C^2$  potential, assumed to be strictly convex outside of a segment for definiteness. The associated conservative dynamics reads:

$$(6.65) \quad d\varphi_t = -\nabla V_0(\varphi_t) dt + \sqrt{2} dB_t,$$

where  $(B_t)$  is a standard Brownian motion on  $X_{N,0}$ .

See the forthcoming works [15] and [16] for results in continuous and discrete spin settings, and [13] for such results for the continuum sine-Gordon model.

## A Classical renormalised potential and Hamilton–Jacobi equation

**A.1. Hamilton–Jacobi equation.** In classical field theory, fields are typically minimisers of an action functional  $S$ :

$$(A.1) \quad \varphi_0 \in \operatorname{argmin}_\varphi S(\varphi), \quad S(\varphi) = \frac{1}{2}|\varphi|^2 + V(\varphi).$$

These can be related to the Hamilton–Jacobi equation

$$(A.2) \quad \frac{\partial V_t}{\partial t} = -\frac{1}{2}(\nabla V_t)^2, \quad V_0(\varphi) = V(\varphi).$$

Indeed, its unique viscosity solution is given by the Hopf–Lax formula [42, Section 3.3.2]:

$$(A.3) \quad V_t(\varphi) = \min_\zeta \left( V_0(\zeta) + \frac{t}{2} \left| \frac{\varphi - \zeta}{t} \right|^2 \right).$$

In particular, the minimum of the action  $S$  is given by  $V_1(0)$ . Note the analogy with the renormalised potential from (3.11). The constructions of Sections 4 and 5 have classical analogues in which the role of the Polchinski equation is replaced by that of the Hamilton–Jacobi equation, as we will see in this appendix.

Unlike solutions of the Polchinski equation, which are smooth at least for  $\dot{C}_t$  nondegenerate and a finite number of variables (as in our discussion), the Hamilton–Jacobi equation can develop shocks and the appropriate weak solutions are not necessarily smooth. However, we can assume that  $V$  is locally Lipschitz continuous and that (A.2) holds almost everywhere. We refer to [42, Chapters 3 and 10] for an introduction. In statistical physics, Hamilton–Jacobi equations are well-known to arise in mean-field *limits* of statistical mechanical models, see [2] and in particular [74] and references for recent work in the context of disordered models, as well as Section A.2 below. Shocks of the Hamilton–Jacobi equations are related to phase transitions. Note that the Polchinski equation in a finite number of variables describes finite systems (rather than limits) and thus has smooth solutions. It therefore provides a complete description of the models (no information is lost) while the Hamilton–Jacobi equations describing mean-field systems are effective equations describing macroscopic information. In the thermodynamic limit, where the number of variables tends to infinity, shocks can also form in the Polchinski equation and then likewise correspond to phase transitions in the statistical mechanical models. We illustrate the above in the simple example of the mean-field Ising model in Section A.2 below.

We now discuss the ‘classical’ analogues of the constructions of Sections 4–5 for the Hamilton–Jacobi equation. Our goal is to emphasise the analogy and to provide a different intuition also for the stochastic constructions, and we will therefore impose convenient regularity assumptions in all statements.

We begin with a classical analogue of Corollary 4.2. Define the classical renormalised action by

$$(A.4) \quad S_t(\varphi) = \frac{|\varphi|^2}{2(1-t)} + V_t(\varphi).$$

Minimisers of  $S_t$  will take the role of the renormalised measure introduced in (3.13).

**Proposition A.1.** *Assume that  $(\varphi_t)_{t \in [0,1]}$  is differentiable in  $t \in (0,1)$ , that  $\varphi_t \rightarrow 0$  as  $t \rightarrow 1$ , that  $V$  is smooth along  $(\varphi_t)_{t \in (0,1)}$ , and that  $\varphi_t$  is an isolated local minimum of  $S_t$  for each  $t \in [0,1)$ . Then for  $t \in [0,1)$ ,*

$$(A.5) \quad \varphi_t = - \int_t^1 \nabla V_u(\varphi_u) du.$$



Different from the situation in Corollary 4.2, the choice of  $\varphi$  is not necessarily unique because  $S$  may have multiple minimisers and  $V$  need not be globally smooth. Indeed, the equation (A.5) for  $\varphi$  is nothing but the equation for the characteristics of Hamilton–Jacobi equation (A.2), see [42, Section 3.3]. These are the curves  $(\varphi_t)$  such that  $\frac{1}{2}U_t(\varphi_t)^2$  is constant in  $t$ , where  $U_t = \nabla V_t$ . Since by the Hamilton–Jacobi equation (A.2),

$$(A.6) \quad \frac{\partial U_t}{\partial t} = -(\nabla U_t, U_t),$$

provided  $U$  is smooth at  $(t, \varphi)$ , then

$$(A.7) \quad \frac{\partial}{\partial t} U_t(\varphi_t) = -(\nabla U_t(\varphi_t), U_t(\varphi_t)) + (\nabla U_t(\varphi_t), \dot{\varphi}_t) = -(U_t(\varphi_t) - \dot{\varphi}_t, \nabla U_t(\varphi_t)) = 0.$$

*Sketch of Proposition A.1.* Since  $\varphi_t \rightarrow 0$  as  $t \rightarrow 1$ , the claim is equivalent to proving that

$$(A.8) \quad \dot{\varphi}_t = \nabla V_t(\varphi_t).$$

Since  $\varphi_t$  minimises  $S_t$ , it satisfies the Euler–Lagrange equation

$$(A.9) \quad \frac{\varphi_t}{1-t} + \nabla V_t(\varphi_t) = \nabla S_t(\varphi_t) = 0.$$

Differentiating this equation in  $t$ ,

$$(A.10) \quad \frac{\dot{\varphi}_t}{1-t} + \frac{\varphi_t}{(1-t)^2} + (\partial_t \nabla V_t)(\varphi_t) + \text{Hess } V_t(\varphi_t) \dot{\varphi}_t = 0.$$

Using the Euler–Lagrange equation  $\varphi_t/(1-t) = -\nabla V_t(\varphi_t)$  again and  $\partial_t \nabla V_t = -\text{Hess } V_t \nabla V_t$  (which follows from (A.2)), we obtain

$$(A.11) \quad \frac{\dot{\varphi}_t - \nabla V_t(\varphi_t)}{1-t} - \text{Hess } V_t(\varphi_t) \nabla V_t(\varphi_t) + \text{Hess } V_t(\varphi_t) \dot{\varphi}_t = 0.$$

Therefore

$$(A.12) \quad \text{Hess } S_t(\varphi_t)[- \nabla V_t(\varphi_t) + \dot{\varphi}_t] = \left( \frac{1}{1-t} + \text{Hess } V_t(\varphi_t) \right) [- \nabla V_t(\varphi_t) + \dot{\varphi}_t] = 0.$$

Since  $\varphi_t$  is an isolated local minimum of  $S_t$ , the Hessian on the left-hand side is strictly positive definite so that necessarily  $\nabla V_t(\varphi_t) - \dot{\varphi}_t = 0$ .  $\square$

The classical version of Föllmer’s problem is the following control problem. We continue to use the convention that the ODE is backwards in time so that the Hamilton–Jacobi equation has initial rather than terminal condition. Suppose that  $U = (U_u(\cdot))_{u \in [0,1]}$  are given smooth functions and that  $\varphi^U$  solves the classical analogue of (5.2):  $\varphi_t^U \rightarrow 0$  as  $t \rightarrow 1$  and

$$(A.13) \quad \varphi_t^U = - \int_t^1 U_u(\varphi_u^U) du,$$

with  $\varphi_0^U$  an absolute minimum of  $S$  (which we recall is the classical analogue of demanding that  $\varphi_0^U$  is a random sample from a desired target measure  $\nu_0 \propto e^{-\beta S}$ ). The classical analogues of Theorem 5.1 and Proposition 5.3 are as follows.

**Proposition A.2.** *The optimal drift is given by  $U = \nabla V$  in the following sense. For any smooth  $U$  and associated trajectory  $\varphi^U$  that satisfies (A.13),*

$$(A.14) \quad \frac{1}{2} |\varphi_0^U|^2 \leq \frac{1}{2} \int_0^1 |U_u(\varphi_u^U)|^2 du,$$

with equality if  $U = \nabla V$ .

Comparing with Theorem 5.1, the classical analogue of  $\mathbb{H}(\nu_0|\gamma_0)$  is simply  $\frac{1}{2}|\varphi_0|^2$  and the classical analogue of the path space entropy  $\mathbb{H}(\mathbf{Q}|\mathbf{P})$  is the cost  $\frac{1}{2}\int_0^1 |U_u(\varphi_u^U)|^2 du = \frac{1}{2}\int_0^1 |\dot{\varphi}_u^U|^2 du$ .

*Sketch.* The proof is analogous to that of Theorem 5.1. Indeed, by the Hopf–Lax formula,

$$(A.15) \quad V_1(0) - V_0(\varphi_0^U) \leq V_0(\varphi_0^U) + \frac{1}{2}|\varphi_0^U|^2 - V_0(\varphi_0^U) = \frac{1}{2}|\varphi_0^U|^2,$$

with equality if and only if  $\varphi_0^U$  is an absolute minimum of  $S$ . On the other hand, assuming  $V$  is locally Lipschitz continuous (so that the fundamental theorem of calculus hold), one has

$$(A.16) \quad V_1(0) - V_0(\varphi_0^U) = \int_0^1 dt \frac{\partial}{\partial t} V_t(\varphi_t^U),$$

with  $\varphi^U$  evolving according to (A.13), and therefore (whenever the derivative exists classically)

$$(A.17) \quad \begin{aligned} \frac{\partial}{\partial t} V_t(\varphi_t^U) &= \left(\frac{\partial}{\partial t} V_t\right)(\varphi_t^U) + (\nabla V_t(\varphi_t^U), U_t(\varphi_t^U)) \\ &= -\frac{1}{2}(\nabla V_t(\varphi_t^U))^2 + (\nabla V_t(\varphi_t^U), U_t(\varphi_t^U)) \\ &= -\frac{1}{2}(\nabla V_t(\varphi_t^U) - U_t(\varphi_t^U))^2 + \frac{1}{2}(U_t(\varphi_t^U))^2, \end{aligned}$$

where we used the Hamilton–Jacobi equation on the second line. In particular, if  $\varphi_0^U$  is an absolute minimum of  $S$ ,

$$(A.18) \quad \frac{1}{2}\int_0^1 (\nabla U_t(\varphi_t^U))^2 dt = \frac{1}{2}|\varphi_0^U|^2 + \frac{1}{2}\int_0^1 (\nabla V_t(\varphi_t^U) - U_t(\varphi_t^U))^2 dt \geq \frac{1}{2}|\varphi_0^U|^2,$$

and the gradient of the renormalised potential  $V_t$  provides the optimal drift.  $\square$

As in (4.3), instead of (A.13), one can also consider the equations for the characteristics in a reduced time interval  $[0, t]$  with  $t \leq 1$  and  $\varphi_t^U = \varphi$ ,

$$(A.19) \quad \varphi_s^U = \varphi - \int_s^t U_u(\varphi_u^U) du, \quad (s \leq t).$$

**Proposition A.3.** *Under certain regularity conditions on  $V_0$ , and for  $t \leq 1$  then*

$$(A.20) \quad V_t(\varphi) = \inf_U \left[ V_0\left(\varphi - \int_0^t U_s(\varphi_s^U) ds\right) + \frac{1}{2}\int_0^t |U_s(\varphi_s^U)|^2 ds \right].$$

This is discussed in [46, Section I.9 and I.10]. Namely, with  $L(t, x, v) = \frac{1}{2}v^2$  and  $\psi = V_0$  (and opposite time direction), the statement follows from [46, Theorem 10.1]. This minimiser does not have to be unique if the Hamilton–Jacobi equation has a shock, see [46, Theorem 9.1 and 10.2].

*Sketch.* The Hopf–Lax formula implies that

$$(A.21) \quad V_t(\varphi) = \min_{\zeta} \left( \frac{t}{2} \left| \frac{\varphi - \zeta}{t} \right|^2 + V_0(\zeta) \right) \leq \frac{1}{2t} |\varphi - \varphi_0^U|^2 + V_0(\varphi_0^U)$$

where given any drift  $U_s$ , we let  $\varphi_0^U$  be the final condition of (A.19). The first term on the right-hand side of (A.21) is bounded as in (A.14):

$$(A.22) \quad \frac{1}{2t} |\varphi - \varphi_0^U|^2 \leq \frac{1}{2} \int_0^t |U_s(\varphi_s^U)|^2 ds.$$

This shows

$$(A.23) \quad V_t(\varphi) \leq \frac{1}{2} \int_0^t |U_s(\varphi_s^U)|^2 ds + V_0\left(\varphi - \int_0^t U_s(\varphi_s^U) ds\right).$$

On the other hand, equality if  $U = \nabla V$  follows from the Hamilton–Jacobi equation as in Proposition 4.3: if  $\varphi$  is a solution to (A.5),

$$(A.24) \quad \frac{\partial}{\partial s} \left[ V_s(\varphi_s) + \frac{1}{2} \int_s^t (\nabla V_u(\varphi_u))^2 ds \right] = \frac{\partial V_s}{\partial s}(\varphi_s) + (\nabla V_s(\varphi_s))^2 - \frac{1}{2} (\nabla V_s(\varphi_s))^2 = 0,$$

i.e.,

$$(A.25) \quad V_t(\varphi) - V_0(\varphi_0) = \frac{1}{2} \int_0^t (\nabla V_u(\varphi_u))^2 du.$$

This solution need not be unique. □

**A.2. Example: Mean-field Ising model.** We conclude this section with the example of the mean-field Ising model, which can be described both in terms of a Polchinski equation and, in the limit, by a Hamilton–Jacobi equation. The mean-field Ising model is given by the measure

$$(A.26) \quad \mathbb{E}_\nu[G] = \frac{1}{Z_N(\beta, \mathbf{h})} \sum_{\sigma \in \{\pm 1\}^N} e^{-\frac{\beta}{4N} \sum_{i,j} (\sigma_i - \sigma_j)^2 + (\sigma, \mathbf{h})} G(\sigma),$$

where the vector  $\mathbf{h} \in \mathbb{R}^N$  is a possibly site-dependent external field. It is convenient to rewrite it as

$$(A.27) \quad \mathbb{E}_\nu[G] = \frac{1}{Z_N(\beta, \mathbf{h})} \sum_{\sigma \in \{\pm 1\}^N} e^{-\frac{\beta}{2} (\sigma, P\sigma) + (\sigma, \mathbf{h})} G(\sigma),$$

where  $P = \text{id} - Q$  and  $Q$  is the orthogonal projection onto constants:  $Qf = \left(\frac{1}{N} \sum_i f_i\right) \mathbf{1}$  with  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^N$ . For  $\mathbf{h} = h\mathbf{1}$  with  $h \in \mathbb{R}$ , the free energy  $F(\beta, h)$  is the limit  $N \rightarrow \infty$  of  $F_N(\beta, h)$  where

$$(A.28) \quad F_N(\beta, h) = -\frac{1}{N} \log Z_N(\beta, h\mathbf{1}).$$

(Physically more correctly, the right-hand side should have been divided by  $\beta$ , but it is here more convenient to omit this.) It is well-known and easy to check that

$$(A.29) \quad \frac{\partial F_N}{\partial \beta} = \frac{1}{2N} \frac{\partial^2 F_N}{\partial h^2} - \frac{1}{2} \left( \frac{\partial F_N}{\partial h} \right)^2, \quad F_N(0, h) = -\log \cosh(h),$$

and thus that the limiting free energy  $F$  is the viscosity solution of the Hamilton–Jacobi equation

$$(A.30) \quad \frac{\partial F}{\partial \beta} = -\frac{1}{2} \left( \frac{\partial F}{\partial h} \right)^2, \quad F(0, h) = -\log \cosh(h),$$

see in particular [74]. Equivalently,  $F$  is given by the Hopf–Lax formula which coincides with the well-known variational formula for the free energy alternatively obtained from Laplace’s Principle (see for example [17, Chapter 1] or [85]):

$$(A.31) \quad F(\beta, h) = \min_{g \in \mathbb{R}} \left[ \frac{1}{2\beta} (g - h)^2 - \log \cosh(g) \right] = \min_{\varphi \in \mathbb{R}} \left[ \frac{\beta}{2} \varphi^2 - \log \cosh(\beta\varphi + h) \right].$$

This Hamilton–Jacobi equation for the free energy can be related, as follows, to the Polchinski equation studied earlier in Section 6.2.1. Recall that  $\text{id} = P + Q$  with  $Q$  the orthogonal projection onto constant vectors in  $\mathbb{R}^N$  and  $PQ = 0$ . For  $\alpha > \beta$ , let

$$(A.32) \quad C_t = (tP + (\alpha - t))^{-1}, \quad \dot{C}_t = (\alpha - t)^{-2}Q,$$

where we used  $PQ = 0$  to simplify  $\dot{C}_t$ . For  $\varphi \in \mathbb{R}^N$ , the renormalised potential (6.23) then is

$$(A.33) \quad V_t(\varphi) = -\log \sum_{\sigma \in \{\pm 1\}} e^{-\frac{1}{2}(\sigma - \varphi, (tP + (\alpha - t))(\sigma - \varphi))} + (\text{constant})$$

and satisfies the Polchinski equation (for appropriate otherwise irrelevant choice of the constants):

$$(A.34) \quad \frac{\partial V_t}{\partial t} = \frac{1}{2}\Delta_{\dot{C}_t} V_t - \frac{1}{2}(\nabla V_t)_{\dot{C}_t}^2 = (\alpha - t)^{-2} \left[ \frac{1}{2}\Delta_Q V_t - \frac{1}{2}(\nabla V_t)_Q^2 \right].$$

Note that the right-hand side only depends on derivatives of  $V_t$  in constant directions (i.e., in the image of  $Q$ ). By (A.32), the covariance  $C_\beta - C_t$  of the renormalised measure  $\nu_t = \nu_{t,\beta}$  defined in (6.26) is also proportional to  $Q$  and therefore supported on constant fields. Thus one can restrict the renormalised potential  $V_t$  to constant fields and the restriction satisfies a closed equation. Explicitly, for  $\tilde{\varphi} \in \mathbb{R}$ , define  $\tilde{V}_t(\tilde{\varphi}) = \frac{1}{N}V(\tilde{\varphi}\mathbf{1})$ . In other words,  $V_t(\varphi) = N\tilde{V}_t(Q\varphi) = N\tilde{V}_t(\frac{1}{N}\sum_i \varphi_i)$  holds for constant fields  $\varphi = Q\varphi$  and

$$(A.35) \quad \begin{aligned} \frac{\partial \tilde{V}_t}{\partial t}(\tilde{\varphi}) &= \frac{1}{N} \frac{\partial V_t}{\partial t}(\tilde{\varphi}\mathbf{1}) = (\alpha - t)^{-2} \frac{1}{N} \left[ \frac{1}{2}\Delta_Q V_t - \frac{1}{2}(\nabla_Q V_t)^2 \right](\tilde{\varphi}\mathbf{1}) \\ &= (\alpha - t)^{-2} \left[ \frac{1}{2} \frac{\partial V_t}{\partial \varphi_1^2} - \frac{1}{2} \left( \frac{\partial V_t}{\partial \varphi_1} \right)^2 \right](\tilde{\varphi}\mathbf{1}) \\ &= (\alpha - t)^{-2} \left[ \frac{1}{2N} \tilde{V}_t'' - \frac{1}{2}(\tilde{V}_t')^2 \right](\tilde{\varphi}). \end{aligned}$$

Thus the reduced (one-variable) Polchinski equation that  $\tilde{V}_t$  satisfies has a prefactor  $1/N$  in front of the Laplacian term, and its limit  $\bar{V}_t$  as  $N \rightarrow \infty$  is the unique viscosity solution of the following (one-variable) Hamilton–Jacobi equation (now dropping tilde from the variable  $\varphi$ ):

$$(A.36) \quad \frac{\partial \bar{V}_t}{\partial t}(\varphi) = -\frac{1}{2}(\alpha - t)^{-2}(\bar{V}_t'(\varphi))^2, \quad \bar{V}_0(\varphi) = \frac{\alpha}{2}\varphi^2 - \log \cosh(\alpha\varphi), \quad (\varphi \in \mathbb{R}).$$

To conclude this section, we relate the Hamilton–Jacobi equation (A.36) for the effective potential to the one satisfied by the free energy (A.30). We follow the argument in Example 3.14 and first note that for a constant field  $\varphi\mathbf{1}$  with  $\varphi \in \mathbb{R}$ , the renormalised potential can be written as (see (3.68))

$$(A.37) \quad V_t(\varphi\mathbf{1}) = N \left[ \frac{\alpha - t}{2}\varphi^2 + F_N(t, (\alpha - t)\varphi) \right].$$

Thus  $F(t, h) = -\frac{1}{2}(\alpha - t)^{-1}h^2 + \bar{V}_t((\alpha - t)^{-1}h)$  and the Hamilton–Jacobi equation (A.30) for  $F$  follows from the one of  $\bar{V}$  exactly as in Example 3.14. Indeed, in the setting of that example with the relation (3.68) between  $V_t$  and  $F_t$  one has in general that

$$(A.38) \quad \frac{\partial}{\partial t} V_t = -\frac{1}{2}(\nabla V_t)_{\dot{C}_t}^2 \quad \Leftrightarrow \quad \frac{\partial}{\partial t} F_t = -\frac{1}{2}(\nabla F_t)_{\dot{\Sigma}_t}^2,$$

and in the present example the choice of  $C_t$  corresponds to  $\dot{\Sigma}_t = Q$ .

## Acknowledgements

This work was supported by the European Research Council under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 851682 SPINRG).

We thank D. Chafai, R. Eldan, N. Gozlan, J. Lehec, Y. Shenfeld, and H.-T. Yau for various discussions related to the material of this introduction, encouragement, and for pointing out several additional references.

We thank the organisers of the following summer schools at which some of the material was presented: the One World Probability Summer School on “PDE and Randomness” in Bath/Zoom organised by Hendrik Weber and Andris Gerasimovics; the “Summer School on SPDE and Related Fields” in Beijing/Zoom organised by Hao Shen, Scott Smith, Rongchan Zhu, and Xiangchan Zhu; and the Summer School on “PDE and Randomness” at the Max Planck Institute for Mathematics in the Sciences organised by Rishabh Gvalani, Francesco Mattesini, Felix Otto, and Markus Tempelmayr. In particular, we also thank Jiwoon Park for leading the exercise classes at the last summer school.

## References

- [1] M. Aizenman. Geometric analysis of  $\varphi^4$  fields and Ising models. I, II. *Commun. Math. Phys.*, 86(1):1–48, 1982.
- [2] M. Aizenman. Perspectives in statistical mechanics. In *Spectral theory and mathematical physics: a Festschrift in honor of Barry Simon’s 60th birthday*, volume 76 of *Proc. Sympos. Pure Math.*, pages 3–24. Amer. Math. Soc., Providence, RI, 2007.
- [3] N. Anari, V. Jain, F. Koehler, H.T. Pham, and T.-D. Vuong. Entropic Independence I: Modified Log-Sobolev Inequalities for Fractionally Log-Concave Distributions and High-Temperature Ising Models. 2021. Preprint, arXiv:2106.04105.
- [4] N. Anari, K. Liu, and S. Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science*, pages 1319–1330. IEEE Computer Soc., Los Alamitos, CA, 2020.
- [5] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses*. Société Mathématique de France, Paris, 2000.
- [6] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.
- [7] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham, 2014.
- [8] N. Barashkov and M. Gubinelli. A variational method for  $\Phi_3^4$ . *Duke Math. J.*, 169(17):3339–3415, 2020.
- [9] N. Barashkov and M. Gubinelli. The  $\Phi_3^4$  measure via Girsanov’s theorem. *Electron. J. Probab.*, 26:Paper No. 81, 29, 2021.
- [10] N. Barashkov, T.S. Gunaratnam, and M. Hofstetter. Multiscale coupling and the maximum of  $\mathcal{P}(\varphi)_2$  models on the torus. 2022.
- [11] R. Bauerschmidt and T. Bodineau. A very simple proof of the LSI for high temperature spin systems. *J. Funct. Anal.*, 276(8):2582–2588, 2019.

- [12] R. Bauerschmidt and T. Bodineau. Spectral Gap Critical Exponent for Glauber Dynamics of Hierarchical Spin Models. *Commun. Math. Phys.*, 373(3):1167–1206, 2020.
- [13] R. Bauerschmidt and T. Bodineau. Log-Sobolev inequality for the continuum sine-Gordon model. *Comm. Pure Appl. Math.*, 74(10):2064–2113, 2021.
- [14] R. Bauerschmidt, T. Bodineau, and B. Dagallier. In preparation.
- [15] R. Bauerschmidt, T. Bodineau, and B. Dagallier. In preparation.
- [16] R. Bauerschmidt, T. Bodineau, and B. Dagallier. Kawasaki dynamics beyond the uniqueness threshold. In preparation.
- [17] R. Bauerschmidt, D.C. Brydges, and G. Slade. *Introduction to a renormalisation group method*, volume 2242 of *Lecture Notes in Mathematics*. Springer, Singapore, 2019.
- [18] R. Bauerschmidt and B. Dagallier. Log-Sobolev inequality for near critical Ising models. *Comm. Pure Appl. Math.* to appear.
- [19] R. Bauerschmidt and B. Dagallier. Log-Sobolev inequality for the  $\varphi_2^4$  and  $\varphi_3^4$  measures. *Comm. Pure Appl. Math.* to appear.
- [20] R. Bauerschmidt and M. Hofstetter. Maximum and coupling of the sine-Gordon field. *Ann. Probab.*, 50(2):455–508, 2022.
- [21] S. Becker and A. Menegaki. Spectral gap in mean-field  $\mathcal{O}(n)$ -model. *Commun. Math. Phys.*, 380(3):1361–1400, 2020.
- [22] S.G. Bobkov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.*, 19(2):289–336, 2006.
- [23] H.J. Brascamp and E.H. Lieb. *On Extensions of the Brunn-Minkowski and Prékopa-Leindler Theorems, Including Inequalities for Log Concave Functions, and with an Application to the Diffusion Equation*, pages 441–464. Springer Berlin Heidelberg, 2002.
- [24] D.C. Brydges and T. Kennedy. Mayer expansions and the Hamilton-Jacobi equation. *J. Statist. Phys.*, 48(1-2):19–49, 1987.
- [25] L.A. Caffarelli. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Commun. Math. Phys.*, 214(3):547–563, 2000.
- [26] R. Catellier and K. Chouk. Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *Ann. Probab.*, 46(5):2621–2679, 2018.
- [27] P. Cattiaux and A. Guillin. Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Math.*, pages 231–292. Springer, Cham, 2014.
- [28] Y. Chen and R. Eldan. Localization Schemes: A Framework for Proving Mixing Bounds for Markov Chains. 2022. Preprint, arXiv:2203.04163.
- [29] Sinho Chewi and Aram-Alexandre Pooladian. An entropic generalization of Caffarelli’s contraction theorem via covariance inequalities, 2022. Preprint, arXiv:2203.04954.
- [30] B. Chow, P. Lu, and L. Ni. *Hamilton’s Ricci flow*, volume 77 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2006.
- [31] G. Conforti. *A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost*, volume 174, pages 1–47. 2019.

- [32] G. Conforti. Weak semiconvexity estimates for Schrödinger potentials and logarithmic Sobolev inequality for Schrödinger bridges, 2022. Preprint, arXiv:2301.00083.
- [33] G. Conforti, A. Durmus, and G. Greco. Quantitative contraction rates for Sinkhorn algorithm: beyond bounded costs and compact marginals, 2023. Preprint, arXiv:2304.04451.
- [34] J. Cotler and S. Rezchikov. Renormalization group flow as optimal transport. February 2022. Preprint, arXiv:2202.11737.
- [35] N. Crawford and W. De Roeck. Stability of the uniqueness regime for ferromagnetic Glauber dynamics under non-reversible perturbations. *Ann. Henri Poincaré*, 19(9):2651–2671, 2018.
- [36] P. Diaconis, E. Nelson, D. Elworthy, G. Papanicolaou, H. Föllmer, and S.R.S. Varadhan. *École d’été de probabilités de Saint-Flour XV-XVII, 1985-87 (2-19 Juil. 1985, 17 Août - 3 Sept. 1986, 1-18 Juil. 1987)*, volume 1362 of *Lect. Notes Math.* Berlin etc.: Springer-Verlag, 1988.
- [37] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.
- [38] J. Ding, J. Song, and R. Sun. A new correlation inequality for Ising models with external fields. *Probab. Theory Related Fields*, 186(1-2):477–492, 2023.
- [39] P. Duch. Flow equation approach to singular stochastic PDEs, 2022. Preprint, arXiv:2109.11380.
- [40] R. Eldan. Analysis of high-dimensional distributions using pathwise methods. *to appear in the 2022 ICM Proceedings*, 2021.
- [41] R. Eldan, F. Koehler, and O. Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature Ising models. *Probab. Theory Related Fields*, 182(3-4):1035–1051, 2022.
- [42] L.C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998.
- [43] M. Fathi, N. Gozlan, and M. Prod’homme. A proof of the Caffarelli contraction theorem via entropic regularization. *Calc. Var. Partial Differential Equations*, 59(3):Paper No. 96, 18, 2020.
- [44] M. Fathi, D. Mikulincer, and Y. Shenfeld. Transportation onto log-lipschitz perturbations. 2023. Preprint, arXiv:2305.03786.
- [45] P. Federbush. Partially alternate derivation of a result of Nelson. *Journal of Mathematical Physics*, 10(1):50–52, 1969.
- [46] W.H. Fleming and H.M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.
- [47] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems*. Cambridge University Press, Cambridge, 2018. A concrete mathematical introduction.
- [48] R.B. Griffiths. Rigorous results for Ising ferromagnets of arbitrary spin. *J. Mathematical Phys.*, 10:1559–1565, 1969.
- [49] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.
- [50] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

- [51] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003.
- [52] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.
- [53] M. Hairer. Regularity structures and the dynamical  $\Phi_3^4$  model. In *Current developments in mathematics 2014*, pages 1–49. Int. Press, Somerville, MA, 2016.
- [54] M. Hairer and K. Matetski. Discretisations of rough stochastic PDEs. *Ann. Probab.*, 46(3):1651–1709, 2018.
- [55] R.S. Hamilton. Three-manifolds with positive Ricci curvature. *J. Differential Geometry*, 17(2):255–306, 1982.
- [56] R. Haslhofer and A. Naber. Characterizations of the Ricci flow. *J. Eur. Math. Soc. (JEMS)*, 20(5):1269–1302, 2018.
- [57] Y.-H. Kim and E. Milman. A generalization of Caffarelli’s contraction theorem via (reverse) heat flow. *Math. Ann.*, 354(3):827–862, 2012.
- [58] B. Klartag and E. Putterman. Spectral monotonicity under Gaussian convolution. 2021. Preprint, arXiv:2107.09496.
- [59] E. Kopfer and K.-T. Sturm. Heat flow on time-dependent metric measure spaces and super-Ricci flows. *Comm. Pure Appl. Math.*, 71(12):2500–2608, 2018.
- [60] E. Kopfer and K.-T. Sturm. Functional inequalities for the heat flow on time-dependent metric measure spaces. *J. Lond. Math. Soc. (2)*, 104(2):926–955, 2021.
- [61] A. Kupiainen. Renormalization group and stochastic PDEs. *Ann. Henri Poincaré*, 17(3):497–535, 2016.
- [62] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.
- [63] J. Lehec. Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):885–899, 2013.
- [64] C. Léonard. *A survey of the Schrödinger problem and some of its connections with optimal transport*, volume 34, pages 1533–1574. 2014.
- [65] J. Lott. Optimal transport and Perelman’s reduced volume. *Calc. Var. Partial Differential Equations*, 36(1):49–84, 2009.
- [66] E. Lubetzky and A. Sly. Cutoff for the Ising model on the lattice. *Invent. Math.*, 191(3):719–755, 2013.
- [67] E. Lubetzky and A. Sly. Information percolation and cutoff for the stochastic Ising model. *J. Amer. Math. Soc.*, 29(3):729–774, 2016.
- [68] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. In *Lectures on probability theory and statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Math.*, pages 93–191. Springer, Berlin, 1999.
- [69] F. Martinelli. Relaxation times of Markov chains in statistical mechanics and combinatorial structures. In *Probability on discrete structures*, pages 175–262. Berlin: Springer, 2004.



- [70] R.J. McCann and P.M. Topping. Ricci flow, entropy and optimal transportation. *Amer. J. Math.*, 132(3):711–730, 2010.
- [71] D. Mikulincer and Y. Shenfeld. The Brownian transport map. 2021. Preprint, arXiv:2111.11521.
- [72] D. Mikulincer and Y. Shenfeld. On the Lipschitz properties of transportation along heat flows. *GAFSA Seminar Notes*, 2022. to appear.
- [73] E. Mossel and A. Sly. Exact thresholds for Ising-Gibbs samplers on general graphs. *Ann. Probab.*, 41(1):294–328, 2013.
- [74] J.-C. Mourrat. Hamilton-Jacobi equations for mean-field disordered systems. *Ann. H. Lebesgue*, 4:453–484, 2021.
- [75] J. Neeman. Lipschitz changes of variables via heat flow. 2022. Preprint, arXiv:2201.03403.
- [76] E. Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles (Proc. Conf., Dedham, Mass., 1965)*, pages 69–73. M.I.T. Press, Cambridge, Mass., 1966.
- [77] G. Perelman. The entropy formula for the Ricci flow and its geometric applications. Preprint, arXiv:0211159.
- [78] J. Polchinski. Renormalization and effective lagrangians. *Nuclear Physics B*, 231(2):269 – 295, 1984.
- [79] M. Reed and B. Simon. *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich Publishers], 1978.
- [80] G. Royer. *An initiation to logarithmic Sobolev inequalities*, volume 14 of *SMF/AMS Texts and Monographs*. American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2007. Translated from the 1999 French original by Donald Babbitt.
- [81] L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, 1997.
- [82] J. Serres. Behavior of the Poincaré constant along the Polchinski renormalization flow. 2022. Preprint, arXiv:2208.08186.
- [83] Y. Shenfeld. Exact renormalization groups and transportation of measures. May 2022. Preprint, arXiv:2205.01642.
- [84] J. Sieber. Formulae for the derivative of the Poincaré constant of Gibbs measures. *Stochastic Process. Appl.*, 140:1–20, 2021.
- [85] B. Simon. *The statistical mechanics of lattice gases. Vol. I*. Princeton Series in Physics. Princeton University Press, 1993.
- [86] D.W. Stroock. Logarithmic Sobolev inequalities for Gibbs states. In *Dirichlet forms (Varenna, 1992)*, volume 1563 of *Lecture Notes in Math.*, pages 194–228. Springer, Berlin, 1993.
- [87] K.-T. Sturm. Super-Ricci flows for metric measure spaces. *J. Funct. Anal.*, 275(12):3504–3569, 2018.
- [88] D. Weitz. *Counting independent sets up to the tree threshold*, pages 140–149. New York, NY: ACM Press, 2006.

- [89] K.G. Wilson and J.B. Kogut. The renormalization group and the  $\varepsilon$  expansion. *Physics Reports*, 12(2):75–200, 1974.
- [90] N. Yoshida. The log-Sobolev inequality for weakly coupled lattice fields. *Probab. Theory Related Fields*, 115(1):1–40, 1999.
- [91] Z. Zhang, B. Qian, and Y. Ma. Uniform logarithmic Sobolev inequality for Boltzmann measures with exterior magnetic field over spheres. *Acta Appl. Math.*, 116(3):305–315, 2011.
- [92] R. Zhu and X. Zhu. Lattice approximation to the dynamical  $\Phi_3^4$  model. *Ann. Probab.*, 46(1):397–455, 2018.