Probability Surveys
Vol. 21 (2024) 200–290
ISSN: 1549-5787
https://doi.org/10.1214/24-PS27

Stochastic dynamics and the Polchinski equation: An introduction

Roland Bauerschmidt[1],
Thierry Bodineau[2] **and Benoit Dagallier**[1]

[1] *Courant Institute of Mathematics Sciences, New York University,*
e-mail: bauerschmidt@cims.nyu.edu; bd2543@cims.nyu.edu

[2] *CNRS, IHES,*
e-mail: bodineau@ihes.fr

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.

MSC2020 subject classifications: Primary 60H30; secondary 60K35, 39B62, 82C20, 82C21, 82C27.
Keywords and phrases: Functional inequalities, log-Sobolev inequality, renormalisation group, Polchinski equation, stochastic localisation, transport map, Glauber dynamics, Gibbs measure.

Received July 2023.

Contents

arXiv: 2307.07619

1. Introduction

Functional inequalities have been thoroughly studied in different contexts [66, 9] and one important motivation is to quantify the relaxation of stochastic dynamics by using Poincaré and (possibly modified) log-Sobolev inequalities [72, 55, 91, 40, 25]. 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 [72, 73, 55] 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 [72, 38, 42], 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 [71, 70]. 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 [94], and with its continuous formulation emphasised in particular by Polchinski [83]. 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 [45]. 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 [31] and reviewed in Section 3.7. This estimate applies to the same Polchinki flow, or its equivalent interpretation as stochastic localisation. It originated in the spectral and entropic independence estimates [6, 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 [31] 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 [39] 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 [67], whose origin is in stochastic optimal control theory. This correspondence is at the heart of the Barashkov–Gubinelli variational method [10].

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 [78] 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, Λ 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 $\varepsilon \to 0$ in the case of models defined in the continuum. Sometimes we identify Λ with $[N] = \{1, \dots, N\}$. Spin fields are then random functions $\varphi : \Lambda \to T$ 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 σ instead of φ for a spin configuration.

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

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

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,

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

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

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

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

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

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 φ^4 potential, in which case one usually sets $\beta = 1$ (and r has the role of a temperature),

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

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

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

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

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

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 $\Lambda \to \mathbb{Z}^d$ for the Glauber dynamics of an Ising-type model with continuous spins; or $\Lambda \subset \varepsilon\mathbb{Z}^d$ with $\Lambda \to \mathbb{R}^d$ or $\Lambda \to [0, 1]^d$ in which case a suitably normalised version of φ describes the solution of a singular SPDE in the limit $\varepsilon \to 0$.

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

$$\mathbb{E}_\nu[F(\sigma)] \propto \sum_{\sigma\in\{\pm1\}^\Lambda} e^{-\frac{\beta}{2}(\sigma, A\sigma)} F(\sigma) \tag{2.8}$$

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 σ to σ^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

$$\nu(\sigma) c(\sigma, \sigma^x) = \nu(\sigma^x) c(\sigma^x, \sigma), \tag{2.9}$$

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 $\Lambda \to \mathbb{Z}^d$.

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

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

where

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

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 $F_0 : X \to \mathbb{R}$,

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

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

$$\frac{\partial}{\partial t} F_t = \Delta^H F_t. \tag{2.13}$$

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

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

The right-hand side is the Dirichlet form:

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

In particular, the measure ν is invariant, i.e., if φ_0 is distributed according to ν then φ_t also is:

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

Moreover, we will always impose the following ergodicity assumption:

$$\forall F_0 \in L^2(\nu): \qquad F_t \to \mathbb{E}_\nu[F_0] \qquad \text{in } L^2(\nu). \tag{2.17}$$

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

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

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

Exercise 2.1. Show that $\frac{1}{2}|\nabla H|^2 - \Delta H \to \infty$ as $|\varphi| \to \infty$ 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 $W \to \infty$ has compact resolvent [84, Theorem XIII.67] (a version of Rellich's theorem).

For further general facts on stochastic dynamics in the continuous setting, we refer to [9, 55, 91]. 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

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

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

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

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 $F : \Omega \to \mathbb{R}$, where $\Omega = \{\pm 1\}^\Lambda$ is the finite state space, the generator and Dirichlet form associated with the Glauber dynamics are

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

and

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

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):

$$\forall F_0 : \Omega \to \mathbb{R} : \qquad F_t \to \mathbb{E}_\nu[F_0], \tag{2.23}$$

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., [86].

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

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

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

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

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

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

with the corresponding Dirichlet form

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

The Metropolis jump rates correspond to $\Psi(a,b)=\min\{a,b\}$. For further general discussion of Glauber dynamics in the discrete case, see [72] and [86, 25].

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 Λ finite), one always has the qualitative ergodicity $m_t\to\nu=m_\infty$ 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:

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

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

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

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:

$$\|m_t - \nu\|_{\mathrm{TV}}^2 \leqslant 2\mathbb{H}(m_t|\nu). \tag{2.30}$$

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

$$\frac{\partial}{\partial t}\operatorname{Ent}_\nu(F_t) = -D_\nu(\log F_t, F_t) = -I_\nu(F_t) \leqslant 0, \tag{2.31}$$

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

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

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

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

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

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

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 ν on $X = \mathbb{R}^N$, satisfies the log-Sobolev inequality (LSI) with respect to D_ν if there is a constant $\gamma > 0$ such that the following holds for any smooth, compactly supported function $F : X \to \mathbb{R}_+$:

$$\operatorname{Ent}_\nu(F) \leqslant \frac{2}{\gamma} D_\nu(\sqrt{F}). \tag{2.35}$$

The largest choice of γ in this inequality is the log-Sobolev constant (with respect to D_ν). 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,

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

follows, by Grönwall's lemma, from

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

which, by the de Bruijn 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 motivations 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 [53], see also [80, 50].

Theorem 2.4 (Hypercontractivity [53])**.** *The measure ν satisfies the log-Sobolev inequality (2.35) with constant γ if and only if the associated semigroup $\boldsymbol{T}_t$ is hypercontractive:*

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

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 γ implies the spectral gap inequality (also called Poincaré inequality) with the same constant:*

$$\mathrm{Var}_\nu[F] \leqslant \frac{1}{\gamma}\mathbb{E}_\nu\left[(\nabla F)^2\right]. \tag{2.39}$$

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 ε tend to 0, see, e.g., [86, Lemma 2.2.2].

We refer to [9] for an in-depth account on related functional inequalities and to [66] for 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 Bruijn identity)**.** *Consider the discrete dynamics with invariant measure* ν (2.8) *and Dirichlet form* D_ν *defined in* (2.22)*. Then for* $F_t(\sigma) = \mathbb{E}_{\sigma_0=\sigma}[F(\sigma_t)]$,

$$\frac{\partial}{\partial t}\operatorname{Ent}_\nu(F_t) = -D_\nu(\log F_t, F_t) \leqslant 0. \tag{2.40}$$

The proof is identical to (2.33) replacing the continuous generator by Δ_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 $F : \{\pm 1\}^\Lambda \mapsto \mathbb{R}^+$:

$$\operatorname{Ent}_\nu(F) \leqslant \frac{1}{2\gamma} D_\nu(\log F, F). \tag{2.41}$$

In view of Exercise 2.7, this inequality is weaker than the standard log-Sobolev inequality (2.35), a point discussed in detail in [25].

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

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 [8])**.** *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:*

$$\forall \varphi \in X: \qquad \operatorname{Hess} H(\varphi) \geqslant \lambda \,\mathrm{id}. \tag{2.42}$$

Then the log-Sobolev constant of ν *satisfies* $\gamma \geqslant \lambda$.

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 $\operatorname{Hess} H(\varphi) \geqslant \lambda\,\mathrm{id}$ is to say that H can be written as $H(\varphi) = \frac{1}{2}(\varphi, A\varphi) + V_0(\varphi)$ with a symmetric matrix $A \geqslant \lambda\,\mathrm{id}$ 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 ν. Setting $F_t = P_t(F)$, we note that

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

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

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

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

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

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

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

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

$$\mathrm{Ent}_\nu(F) = \int_0^\infty dt\, I_\nu(F_t) \leqslant \frac{1}{2\lambda} I_\nu(F_0) = \frac{2}{\lambda} D_\nu(\sqrt{F}). \tag{2.47}$$

To prove (2.46), differentiate again:

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

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

$$(\frac{\partial}{\partial t} - \Delta^H)\frac{(\nabla F_t)^2}{F_t} = -2F_t\Big[\underbrace{|\operatorname{Hess}\log F_t|_2^2}_{\geqslant 0} + (\nabla \log F_t, \underbrace{\operatorname{Hess} H(\varphi)}_{\geqslant \lambda\,\mathrm{id}}\nabla \log F_t)\Big]. \tag{2.49}$$

Hence

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

which implies the claim (2.46). □

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 ν is of the form

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

where $\mathbb{E}_1[\cdot] := \mathbb{E}_1[\cdot|\varphi_2]$ is the conditional measure with respect to the variable φ_2. Then the entropy can be split into two parts:

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

Given φ_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 φ_1:

$$\operatorname{Ent}_1(F)(\varphi_2) = \mathbb{E}_1[\Phi(F)|\varphi_2] - \Phi(\mathbb{E}_1\left[F|\varphi_2\right]). \tag{2.53}$$

The strategy is to estimate this term (uniformly in φ_2) by the desired Dirichlet form acting only on φ_1. The second term $\mathbb{E}_2\left[\Phi(\mathbb{E}_1[F])\right]$ 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 ν_1 and ν_2.

In general, the conditional expectations are intertwined and the second term $\operatorname{Ent}_2(\mathbb{E}_1[F])$ 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 $\kappa \operatorname{Ent}_\nu(F)$ 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 ν, the entropy decomposition (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 ν. 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 [19] 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 ν_1 and ν_2 satisfy log-Sobolev inequalities with the constants γ_1 and γ_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 ν_1, ν_2 are probability measures on $\mathbb{R}$ and denote by $\mathbb{E}_1, \mathbb{E}_2$ their expectations so that $\mathbb{E}_\nu\left[F\right] = \mathbb{E}_2\left[\mathbb{E}_1\left[F\right]\right]$ for functions $F(\varphi_1, \varphi_2)$. As discussed in (2.52), the entropy can be decomposed as

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

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

$$2\gamma \operatorname{Ent}(F) \leqslant \mathbb{E}_2[D_1(\sqrt{F})] + D_2(\sqrt{G}). \tag{2.55}$$

It remains to recover the Dirichlet form associated with the product measure ν:

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

The first derivative is easily identified:

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

For the second derivative:

$$\begin{aligned}\partial_{\varphi_2}\sqrt{G(\varphi_2)} &= \frac{1}{2\sqrt{G(\varphi_2)}}\mathbb{E}_1[\partial_{\varphi_2}F(\cdot,\varphi_2)]\\ &= \frac{1}{\sqrt{G(\varphi_2)}}\mathbb{E}_1[\sqrt{F(\cdot,\varphi_2)}\,\partial_{\varphi_2}\sqrt{F(\cdot,\varphi_2)}],\end{aligned} \tag{2.58}$$

so that by Cauchy-Schwarz inequality, we deduce that

$$\begin{aligned}D_2(\sqrt{G}) = \mathbb{E}_2\left[\left(\partial_{\varphi_2}\sqrt{G(\varphi_2)}\right)^2\right] &\leqslant \mathbb{E}_2\left[\mathbb{E}_1\left[\left(\partial_{\varphi_2}\sqrt{F(\cdot,\varphi_2)}\right)^2\right]\right]\\ = \mathbb{E}_\nu\left[(\partial_{\varphi_2}\sqrt{F})^2\right].\end{aligned} \tag{2.59}$$

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

A similar argument applies in the discrete case, see e.g. [86, 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). □

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

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

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

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

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

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

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

with equality if $G = \log F$.

Proof. Since $\mathrm{Ent}_\nu(F) = \mathbb{E}_\nu\left[F \log \frac{F}{\mathbb{E}_\nu[F]}\right]$, 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

$$\forall a \geqslant 0, b \in \mathbb{R}: \qquad ab \leqslant a \log a - a + e^b, \tag{2.63}$$

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

$$\mathbb{E}_\nu[FG] \leqslant \mathrm{Ent}_\nu(F) - 1 + \mathbb{E}_\nu[e^G] \leqslant \mathrm{Ent}_\nu(F). \tag{2.64}$$

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 $b = \log t$:

$$\begin{aligned} \mathrm{Ent}_\nu(F) &= \mathbb{E}_\nu[F \log F] - a \log a \leqslant \mathbb{E}_\nu[F \log F] - ab - a + e^b \\ &= \mathbb{E}_\nu[F \log F - F \log t - F + t], \end{aligned} \tag{2.65}$$

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 $d\nu^F = F\, d\nu$:

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

with equality if $G = \log F$. □

The Holley–Stroock criterion for the log-Sobolev inequality is a simple consequence of (2.61), see e.g., the presentation in [91].

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

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 $\beta \geqslant 0$ and the Gibbs measure (2.1) has a density on $\mathbb{R}^\Lambda$ of the form

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

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 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 β. These phase transitions separate regions of values of β between which the measure ν 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 [52] or [19, 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 Λ, with the same constant as for the single spin $|\Lambda| = 1$ measure (Example 2.9). This tensorisation property has been generalised for β small enough in terms of mixing conditions and for some spin systems up to the critical value β_c, see [72] for a review. Indeed, for β 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 Λ into boxes (of size larger than ξ_β), and using appropriate conditionings the system can be analysed as a renormalised model of weakly interacting spins [95]. In so-called second order phase transitions, when β approaches the critical β_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 β close to β_c. Nevertheless in some cases [72, 70, 71] the strong dynamical mixing properties were derived up to the critical value β_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 β_c with this type of approach.

To study the detailed static features of measures of the form (2.67) close to β_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 [94]. 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:

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

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

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

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 ε:

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

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

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

where the potential V^ε is of the form (2.5). In the example of the φ^4 model, it turns out that such a limit can be constructed if $d < 4$, but in $d \geqslant 2$ the coefficient μ of the potential must be tuned correctly as a function of $\varepsilon \to 0$, 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 ($\varepsilon^d g \to 0$ as $\varepsilon \to 0$). 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 ε 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 ε and depend only on the large scale parameter L. Techniques to control the regularisation parameter ε 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):

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

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 produce $\varepsilon^d \sum_{x\in\Lambda_{\varepsilon,L}} u_x v_x$, i.e., each $W_t(x)$ is a Brownian motion of variance ε^{-d}. On the torus $\Lambda_{\varepsilon,L} = L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$ of mesh size ε and side length L, the eigenvalues of the Laplacian are

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

All Fourier modes of φ evolve independently according to Ornstein–Uhlenbeck processes:

$$p \in \Lambda^*_{\varepsilon,L}: \qquad d\hat{\varphi}(p) = -\lambda(p)\hat{\varphi}(p)\,dt + \sqrt{2}d\hat{W}_t(p), \tag{2.74}$$

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 [94] and Polchinski [83, 27] 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 P_C the corresponding Gaussian measure with covariance C and by E_C its expectation. The measure

P_C is supported on the image of C. In particular, if C is strictly positive definite on $\mathbb{R}^N$,

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

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

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

corresponding to its probabilistic interpretation that if ζ_1 and ζ_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 ζ_1 and ζ_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 Δ_C on $\mathbb{R}^N$:

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

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

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

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_∞. 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_∞. 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 $F : X \to \mathbb{R}$, let $F_t = \mathsf{P}_{C_t} * F$, i.e., $F_t(\varphi) = \mathsf{E}_{C_t}[F(\varphi + \zeta)]$. Then for all t which are not discontinuity points of $\dot{C}_t$,*

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

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

$$\frac{\partial}{\partial t}\mathsf{P}_{C_t} = \frac{1}{2}\Delta_{\dot{C}_t}\mathsf{P}_{C_t}, \tag{3.6}$$

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. ☐

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

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

where $(B_s)_s$ is a Brownian motion taking values in $\mathbb{R}^N$. By construction ζ_t is a Gaussian variable with mean φ 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,

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

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:

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

For a given model from statistical mechanics, it will be important to adjust the decomposition according to the specific spatial structure (of Λ) 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 dependent 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 ν_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 ν_0 on X with expectation given by

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

with a potential $V_0 : X \to \mathbb{R}$, where the Gaussian expectation acts on the variable ζ. 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 ν_0. For this, let us first introduce some notation.

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

- the *renormalised potential* V_t:

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

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

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

- the *renormalised measure* ν_t:

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

where all the Gaussian expectations apply to ζ.

We stress that the renormalised measure ν_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 ν_t. More generally, the function V_∞ is equivalent to the moment generating function of the measure ν_0: changing variables from ζ to $\zeta + C_\infty h$,

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

The renormalised measure ν_t is related to ν_0 by the following identity.

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

$$\mathbb{E}_{\nu_0}[F] = \mathbb{E}_{\nu_t}[\boldsymbol{P}_{0,t}F]. \tag{3.15}$$

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

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

where ζ is integrated with respect to E_{C_t} and φ with respect to $\mathsf{E}_{C_\infty - C_t}$. 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 ν_t is a probability measure. This completes the proof of the identity (3.15). □

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 φ:

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

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

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

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 ν_0 introduced in (3.10), but the covariance of the Gaussian integration is now C_t. By construction $C_t \leqslant C_\infty$, 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 $F : X \to \mathbb{R}$ and all $t > 0$, the identity (3.15) reads

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

where φ denotes the variable of ν_t and ζ the variable of μ_t^φ. 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 ζ encodes the local interactions, which correspond to the fast scales of the dynamics, and φ 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 $\boldsymbol{P}_{0,t}F(\varphi)$ and a measure ν_t (3.13) which is expected to have better properties than the original measure ν_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)

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

Thus $\boldsymbol{P}_{s,t}$ 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 ν_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

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

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

Proof. Let $Z_t(\varphi) = \mathsf{E}_{C_t}[e^{-V_0(\varphi+\zeta)}]$. 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 φ so is Z_t for any $t > 0$, that $Z_t(\varphi) > 0$ for any t and φ, and that for any $t > 0$ such that C_t is differentiable,

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

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

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

□

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 $(\boldsymbol{P}_{s,t})_{s\leqslant t}$ form a time-dependent Markov semigroup with generators $(\boldsymbol{L}_t)$, in the sense that*

$$s \leqslant r \leqslant t: \qquad \boldsymbol{P}_{t,t} = \mathrm{id} \quad \textit{and} \quad \boldsymbol{P}_{r,t}\boldsymbol{P}_{s,r} = \boldsymbol{P}_{s,t}, \tag{3.24}$$

and $\boldsymbol{P}_{s,t}F \geqslant 0$ *if* $F \geqslant 0$ *with* $\boldsymbol{P}_{s,t}1 = 1$.

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

$$s \leqslant t: \qquad \frac{\partial}{\partial t}\boldsymbol{P}_{s,t}F = \boldsymbol{L}_t\boldsymbol{P}_{s,t}F, \qquad -\frac{\partial}{\partial s}\boldsymbol{P}_{s,t}F = \boldsymbol{P}_{s,t}\boldsymbol{L}_sF, \tag{3.25}$$

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

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

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

$$\mathbb{E}_{\nu_t}[\boldsymbol{P}_{s,t}F] = \mathbb{E}_{\nu_s}[F] \quad (s \leqslant t), \qquad -\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}[F] = \mathbb{E}_{\nu_t}[\boldsymbol{L}_tF]. \tag{3.27}$$

The operator $\boldsymbol{L}_t$ in (3.26) is obtained by linearising the Polchinski equation (3.21) and has a structure similar to the generator Δ^H defined in (2.10). According to (3.27), the renormalised measure evolves according to $-\partial_t\nu_t = \boldsymbol{L}_t^*\nu_t$ where $\boldsymbol{L}_t^*$ is the formal adjoint of $\boldsymbol{L}_t$ (with respect to the Lebesgue measure). This is another way to rephrase the Polchinski equation (3.21).

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

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

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

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

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

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

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

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

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 $F : X \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$,

$$\lim_{t\to\infty} \mathbb{E}_{\nu_t}[g(\boldsymbol{P}_{0,t}F)] = g\big(\mathbb{E}_{\nu_0}[F]\big). \tag{3.31}$$

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

Theorem 3.6. *Consider a measure ν_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

$$\forall \varphi \in X, t > 0 : \qquad \dot{C}_t \operatorname{Hess} V_t(\varphi)\dot{C}_t - \frac{1}{2}\ddot{C}_t \geqslant \dot{\lambda}_t \dot{C}_t, \tag{3.32}$$

and define

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

Then ν_0 satisfies the log-Sobolev inequality

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

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 ν_s:

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

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

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

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

$$\begin{aligned}\dot{C}_t^a \,\mathrm{Hess}\, V_t^a \dot{C}_t^a - \frac{1}{2}\ddot{C}_t^a &= \dot{a}(t)^2\Big[\dot{C}_{a(t)}\,\mathrm{Hess}\, V_{a(t)}\dot{C}_{a(t)} - \frac{1}{2}\ddot{C}_{a(t)}\Big] - \frac{1}{2}\ddot{a}(t)\dot{C}_{a(t)} \\ &\geqslant \dot{\lambda}_t^a \dot{C}_t^a\end{aligned} \tag{3.37}$$

with

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

Thus (3.33) becomes

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

and hence

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

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 ∞ 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 $\ddot{C}_t$ can be formulated, see [15, 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 $(\nu_t)_{t\geqslant 0}$ and a corresponding dual time-dependent Markov semigroup $(\boldsymbol{P}_{s,t})$ with generators $(\boldsymbol{L}_t)$ as in Proposition 3.5.

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

$$\lim_{t\to\infty} \mathbb{E}_{\nu_t}\left[\Phi(\boldsymbol{P}_{0,t}F)\right] = \Phi\big(\mathbb{E}_{\nu_0}\left[F\right]\big). \tag{3.41}$$

Thus

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

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

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

Integrating this relation using (3.42) gives

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

The above entropy production formula (3.43) is analogous to the de Bruijn identity (2.31) and the entropy decomposition to (2.45), but an important difference is that the reference measure ν_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 Bruijn identity, but more conceptually what we used is that the measure ν satisfies

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

since both sides are 0 (because the stationary measure ν does not depend on t). In the computation above, both ν_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

$$\forall t \geqslant 0: \qquad (\nabla\sqrt{\boldsymbol{P}_{0,t}F})^2_{\dot{C}_t} \leqslant e^{-2\lambda_t}\boldsymbol{P}_{0,t}\left[(\nabla\sqrt{F})^2_{\dot{C}_0}\right]. \tag{3.46}$$

Plugging this relation in (3.44) and recalling that

$$\mathbb{E}_{\nu_t}\left[\boldsymbol{P}_{0,t}\left[(\nabla\sqrt{F})^2_{\dot{C}_0}\right]\right] = \mathbb{E}_{\nu_0}\left[(\nabla\sqrt{F})^2_{\dot{C}_0}\right],$$

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

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

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

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

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

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

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

At first sight, the proof of Theorem 3.6 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 $(t_i = i\delta)_{i\geqslant 0}$, one has

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

The measure $\boldsymbol{P}_{t_i,t_{i+1}}$ 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 δ small (so that the Hamiltonian corresponding to the measure $\boldsymbol{P}_{t_i,t_{i+1}}$ is extremely convex). This suggests that for each interval $[t_i, t_{i+1}]$, one can reduce to estimating $\mathbb{E}_{\nu_{t_{i+1}}}[(\nabla\sqrt{\boldsymbol{P}_{0,t_i}(F)})^2_{\delta\dot{C}_{t_i}}]$ and the delicate issue is then to interchange ∇ and

$\boldsymbol{P}_{0,t_i}$ (note that a similar step already occurred even in the product case (2.59)). Such a discrete decomposition was implemented in [14] to derive a spectral gap for certain models. The proof of Theorem 3.6 relies on the limit where δ 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 [15, Lemma 2.8]. One can first verify the so-called 'Bochner formula':

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

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

$$\begin{aligned}(\boldsymbol{L}_t - \partial_t)(\nabla\sqrt{F})_Q^2 = \frac{(\boldsymbol{L}_t - \partial_t)(\nabla F)_Q^2}{4F} - \frac{(\nabla F)_Q^2(\boldsymbol{L}_t - \partial_t)F}{4F^2}\\ - \frac{(\nabla(\nabla F)_Q^2, \nabla F)_{\dot{C}}}{4F^2} + \frac{(\nabla F)_Q^2(\nabla F)_{\dot{C}}^2}{4F^3}.\end{aligned} \tag{3.52}$$

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

$$\begin{aligned}F|\dot{C}^{1/2} \operatorname{Hess} \log FQ^{1/2}|_2^2 = \frac{|\dot{C}^{1/2} \operatorname{Hess} FQ^{1/2}|_2^2}{F} - \frac{(\nabla(\nabla F)_Q^2, \nabla F)_{\dot{C}}}{F^2}\\ + \frac{(\nabla F)_Q^2(\nabla F)_{\dot{C}}^2}{F^3}.\end{aligned} \tag{3.53}$$

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

$$F\dot{C}_{ij}Q_{kl}(\operatorname{Hess}\log F)_{ik}(\operatorname{Hess}\log F)_{jl} = F\dot{C}_{ij}Q_{kl}\Big(\frac{F_{ik}}{F} - \frac{F_iF_k}{F^2}\Big)\Big(\frac{F_{jl}}{F} - \frac{F_jF_l}{F^2}\Big), \tag{3.54}$$

and the right-hand side is

$$\dot{C}_{ij}Q_{kl}\left[\frac{F_{ik}F_{jl}}{F} - \frac{(F_kF_l)_iF_j}{F^2} + \frac{F_iF_jF_kF_l}{F^3}\right]. \tag{3.55}$$

So both are indeed equal. □

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 $H_t = \operatorname{Hess} V_t$. Then*

$$\partial_t U_t = \boldsymbol{L}_t U_t, \qquad \partial_t H_t = \boldsymbol{L}_t H_t - H_t \dot{C}_t H_t. \tag{3.56}$$

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

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

$$(f, \operatorname{Hess} V_t f) = \boldsymbol{P}_{s,t}(f, \operatorname{Hess} V_s f) - \Big[\boldsymbol{P}_{s,t}((f,\nabla V_s)^2) - (\boldsymbol{P}_{s,t}(f,\nabla V_s))^2\Big]. \tag{3.58}$$

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

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

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

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

Alternatively, one can rewrite the derivatives of the renormalised potential in terms of the fluctuation measure μ_t^φ introduced in (3.18).

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

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

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

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

where the second equalities hold if C_t is invertible.

Proof. The first part of (3.60) follows from the identity $\nabla V_t = \boldsymbol{P}_{0,t}(\nabla V_0)$ obtained in (3.57) and the identification of the fluctuation measure μ_t^φ with the semigroup $\boldsymbol{P}_{0,t}$ 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 $\boldsymbol{P}_{0,t}$ and μ_t^φ. The second equality follows by differentiating $\varphi \mapsto \mathbb{E}_{\mu_t^\varphi}[C_t^{-1}(\varphi - \zeta)]$ and using (3.18). □

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 $t \geqslant 0$.*

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

$$\frac{1}{\gamma} \leqslant \int_0^\infty e^{-\lambda t}\, dt = \frac{1}{\lambda}. \tag{3.62}$$

This follows from criterion (3.32) applied with $\dot{C}_t = e^{-\lambda t} \operatorname{id}$ so that $\ddot{C}_t = -\lambda e^{-\lambda t} \operatorname{id}$ and $\dot{\lambda}_t = \lambda/2$. Note that the decomposition $\dot{C}_t = \operatorname{id}$ 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 probability measure $\propto e^{-V_0(\varphi+\zeta)}\, \mathsf{P}_{C_t}(d\zeta)$, with density log-concave in (ζ, φ). A theorem of Prékopa then implies that V_t is convex. It is also possible to directly compute the Hessian: the Brascamp–Lieb inequality [26, Theorem 4.1] states that if a probability measure $\propto e^{-H}$ has strictly convex potential H then

$$\operatorname{Var}(F) \leqslant \mathbb{E}[(\nabla F, (\operatorname{Hess} H)^{-1} \nabla F)]. \tag{3.63}$$

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

$$\begin{aligned}
\operatorname{Hess} V_t(\varphi) &= \mathbb{E}_{\mu_t^\varphi}[\operatorname{Hess} V_0(\zeta)] - \operatorname{Var}_{\mu_t^\varphi}(\nabla V_0(\zeta)) \\
&\geqslant \mathbb{E}_{\mu_t^\varphi}\Big[\operatorname{Hess} V_0(\zeta) - \operatorname{Hess} V_0(\zeta)(C_t^{-1} + \operatorname{Hess} V_0(\zeta))^{-1} \operatorname{Hess} V_0(\zeta)\Big] \\
&= \mathbb{E}_{\mu_t^\varphi}\Big[\operatorname{Hess} V_0(\zeta)(C_t^{-1} + \operatorname{Hess} V_0(\zeta))^{-1} C_t^{-1}\Big] \\
&= \mathbb{E}_{\mu_t^\varphi}\Big[C_t^{-1/2} C_t^{1/2} \operatorname{Hess} V_0(\zeta) C_t^{1/2} (\operatorname{id} + C_t^{1/2} \operatorname{Hess} V_0(\zeta) C_t^{1/2})^{-1} C_t^{-1/2}\Big].
\end{aligned} \tag{3.64}$$

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

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

Proof 2 from [59, Theorem 9.1], [33, Theorem 3.3]. This alternative approach puts the emphasis on the PDE structure associated with the renormalised potential by application of the maximum princple. We give the gist of the proof and refer to [33, Theorem 3.3, page 129] for a complete argument. Let $H_t = \operatorname{Hess} V_t$ with $H_0 > 0$, and recall (3.56):

$$\frac{\partial H_t}{\partial t} = \boldsymbol{L}_t H_t - H_t \dot{C}_t H_t. \tag{3.66}$$

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

$$\frac{\partial f_{t_0}(\varphi_0)}{\partial t} = \boldsymbol{L}_{t_0} f_{t_0}(\varphi_0) - (v_0, H_{t_0}(\varphi_0)\dot{C}_{t_0} H_{t_0}(\varphi_0) v_0) \geqslant 0, \tag{3.67}$$

where we used that $f_{t_0}(\varphi)$ is minimum at φ_0 so that by the maximum principle $\boldsymbol{L}_{t_0} f_{t_0}(\varphi) = \frac{1}{2}\Delta_{\dot{C}_{t_0}} f_{t_0}(\varphi) \geqslant 0$ 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 [33, Theorem 3.3]. □

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

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

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

where $F_t(h) = V_t(0) - \log \mathsf{E}_{C_t}[e^{-V_0(\zeta)+(h,\zeta)}]$ 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:

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

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

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

and

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

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 $H : \mathbb{R} \to \mathbb{R}$ be a C^2 potential that is strictly convex outside of a segment: $\inf_{|x|\geqslant M} H''(x) \geqslant c > 1$ for some $c, M > 0$, but assume that $\inf_{\mathbb{R}} H'' < 0$, and consider the measure:

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

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, ν_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 ν_t that will be more log-concave than ν_0.

The Polchinski flow is defined in terms of a covariance decomposition, which is supposed to decompose the Gaussian part of ν_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:

$$\forall t \in [0,1]: \qquad C_t := t\,\mathrm{id}, \qquad \dot{C}_t = \mathrm{id}. \tag{3.73}$$

The corresponding renormalised potential reads:

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

and the renormalised measure ν_t defined in (3.13) and fluctuation measure μ_t^φ defined in (3.18) are respectively given by:

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

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

$$\mu_t^\varphi(d\zeta) \propto \exp\Big[-\frac{\zeta^2}{2}\big(\frac{1}{t} - 1\big) + \frac{\zeta\varphi}{t} - H(\zeta)\Big]\, d\zeta. \tag{3.76}$$

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:

$$\partial_t V_t = \frac{1}{2}\partial_\varphi^2 V_t - \frac{1}{2}(\partial_\varphi V_t)^2. \tag{3.77}$$

The motivation for the Polchinski decomposition was that one progressively integrates "small scales" to recover a measure ν_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 ν_t is more log-concave that ν_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., [48, 32, 35, 36] 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:

$$\partial_\varphi^2 V_t(\varphi) = \frac{1}{t} - \frac{1}{t^2}\,\mathrm{Var}_{\mu_t^\varphi}(\zeta). \tag{3.78}$$

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 μ_t^φ satisfies a spectral gap inequality with constant C uniformly in $\varphi \in \mathbb{R}$ and $t > 0$, we deduce:

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

The uniform lower bound (3.79) confirms that ν_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 γ of ν_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

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

the volume form is

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

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

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

In particular,

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

and

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

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

$$\frac{\partial}{\partial t} dm_{g_t} = (\frac{\partial}{\partial t}\log\sqrt{|g_t|})\, dm_{g_t} = \frac{1}{2}\,\mathrm{Tr}(g_t^{-1}\dot{g}_t)\, dm_{g_t}. \tag{3.85}$$

The Ricci curvature tensor associated with the metric g is denoted Ric_g.

Remark 3.16. The notation Δ_{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

$$A^{-1} = \int_0^\infty g_t^{-1}\, dt. \tag{3.86}$$

Then the Polchinski equation reads

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

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

$$\mathrm{Hess}_{g_t}\, V_t + \frac{1}{2}\dot{g}_t \geqslant \dot{\lambda}_t g_t. \tag{3.88}$$

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

$$|f|_{g_t} \leqslant 1 \qquad \Leftrightarrow \qquad |e^{tA} f| \leqslant 1, \tag{3.89}$$

i.e., $f = e^{-tA} g$ for some $|g| \leqslant 1$ 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

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

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

Lemma 3.19.

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

Proof. The Bochner formula with a background metric [9, Theorem C.3.3] implies:

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

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

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

$$\frac{\partial}{\partial t} \mathsf{Z}_t \, dm_{g_t} = (-\frac{1}{2} \Delta_{g_t} \mathsf{Z}_t) \, dm_{g_t}. \tag{3.93}$$

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

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

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

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

$$\mathbb{E}_{\nu_t}[F] \propto \int F e^{-V_t} \mathsf{Z}_t \, dm_{g_t}. \tag{3.96}$$

One can check that

$$\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}[\boldsymbol{P}_{0,t}F] = 0, \tag{3.97}$$

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

$$\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}[F] = -\mathbb{E}_{\nu_t}[\boldsymbol{L}_t F]. \tag{3.98}$$

The analogue of the continuity assumption (3.31) is

$$\mathbb{E}_{\nu_t}[g(\boldsymbol{P}_{0,t}F)] \to g(\mathbb{E}_{\nu_0}[F]), \qquad (t \to T). \tag{3.99}$$

Since the evolution of 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*

$$\forall \varphi \in X, t > 0: \qquad \mathrm{Hess}_{g_t(\varphi)} V_t(\varphi) + \frac{1}{2}\mathrm{Ric}_{g_t(\varphi)} + \frac{1}{2}\dot g_t(\varphi) \geqslant \dot\lambda_t g_t(\varphi), \tag{3.100}$$

and define

$$\lambda_t = \int_0^t \dot\lambda_s \, ds, \qquad \frac{1}{\gamma} = \int_0^T e^{-2\lambda_t}\, dt. \tag{3.101}$$

Then ν_0 *satisfies the log-Sobolev inequality*

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

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 [82] and then [74, 69] and more recently [92, 63, 64, 60].

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 [31] and has its origins in the spectral and entropic independence conditions introduced in [6] and [4]. In [31], 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 [31] 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 μ on X (with all exponential moments) and $h \in X$ write $T_h\mu$ for the tilted probability measure:

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

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

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

$$\forall h \in X : \qquad \dot\Sigma\, \mathrm{Cov}(T_h\mu)\dot\Sigma \leqslant \alpha\dot\Sigma. \tag{3.104}$$

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

$$\frac{1}{2}(\mathbb{E}_\mu[F\zeta] - \mathbb{E}_\mu[\zeta])^2_{\dot\Sigma} \leqslant \alpha\, \mathrm{Ent}_\mu(F). \tag{3.105}$$

*In [31, Definition 29], this inequality is called α-*entropic stability *of the measure μ with respect to the function $\psi(x,y) = \frac{1}{2}(x-y)^2_{\dot\Sigma}$.*

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 (possibly modified) log-Sobolev inequality, see [31, Proposition 39].

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

$$\forall \varphi \in X : \qquad \dot C_t \,\mathrm{Hess}\, V_t(\varphi)\dot C_t - \dot C_t C_t^{-1}\dot C_t \geqslant -\alpha_t \dot C_t, \tag{3.106}$$

or equivalently, with $\dot\Sigma_t = C_t^{-1}\,\dot C_t\, C_t^{-1}$,

$$\forall \varphi \in X : \qquad \dot\Sigma_t\, \mathrm{Cov}(\mu_t^\varphi)\dot\Sigma_t \leqslant \alpha_t\dot\Sigma_t. \tag{3.107}$$

Then the measure μ_t^φ, i.e., $\boldsymbol{P}_{0,t}(\cdot)(\varphi)$, satisfies the following α_t-entropic stability: for all $\varphi \in X$,

$$\begin{aligned}2(\nabla\sqrt{\boldsymbol{P}_{0,t}F})^2_{\dot C_t}(\varphi) = 2(\nabla\sqrt{\mathbb{E}_{\mu_t^\varphi}[F]})^2_{\dot C_t} &\leqslant \alpha_t\, \mathrm{Ent}_{\mu_t^\varphi}(F)\\ &= \alpha_t\Big[\boldsymbol{P}_{0,t}\Phi(F)(\varphi) - \Phi\big(\boldsymbol{P}_{0,t}F(\varphi)\big)\Big].\end{aligned} \tag{3.108}$$

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

$$\mathrm{Ent}_{\nu_0}(F) \leqslant e^{\int_s^\infty \alpha_u\, du}\, \mathbb{E}_{\nu_s}\Big[\mathrm{Ent}_{\mu_s^\varphi}(F)\Big]. \tag{3.109}$$

The condition (3.106) on α_t is very similar to the multiscale Bakry–Émery condition (3.32) on $\dot\lambda_t$, but not identical. We recall that (3.32) reads

$$\dot C_t\, \mathrm{Hess}\, V_t(\varphi)\dot C_t - \frac{1}{2}\ddot C_t \geqslant \dot\lambda_t \dot C_t. \tag{3.110}$$

The log-Sobolev constant and the closely related entropy contraction are estimated by the time integrals of α_t and $\dot\lambda_t$, respectively. In terms of the two strategies for proving log-Sobolev inequalities discussed in Section 2.5, the Bakry–Émery method corresponds more to the first strategy (but also see Remark 3.23 for a rephrasing in terms of the second strategy) while the entropic stability estimate applies the second strategy of entropy contraction.

The multiscale Bakry–Émery condition (3.110) is less singular for t close to 0 as C_t typically vanishes as $t \to 0$ while $\dot C_t$ does not. To see this, consider the simple one-variable case with the covariance function $c_t = \int_0^t \dot c_u \, du \in \mathbb{R}$ (with $\dot c_u > 0$) and assume that $\operatorname{Hess} V_t(\varphi) \geqslant 0$. Then the optimal choices are $-2\dot\lambda_t = \ddot c_t \, \dot c_t^{-1}$ and $\alpha_t = \dot c_t \, c_t^{-1}$ so that

$$-2 \int_s^t \dot\lambda_u \, du = \log \dot c_t - \log \dot c_s \quad \text{and} \quad \int_s^t \alpha_u \, du = \log c_t - \log c_s. \tag{3.111}$$

In particular, one cannot take the limit $s \to 0$ in (3.109) in which the measure μ_s^φ degenerates, and the $\operatorname{Ent}_{\mu_s^\varphi}(F)$ 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 [31]. For example, one can use that since covariance C_s vanishes as $s \to 0$ the measure μ_s^φ becomes uniformly log-concave 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 ν_0.

As a last remark before proving Corollary 3.22, we emphasise that (3.108) is a contraction estimate for the expected entropy of the fluctuation measure μ_t^φ. More precisely, as in the proof of the multiscale Bakry–Émery criterion in (3.43), the expectation of the left-hand side of (3.108) is

$$\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}\big[\operatorname{Ent}_{\mu_t^\varphi}(F)\big] = -\frac{\partial}{\partial t}\operatorname{Ent}_{\nu_t}(\boldsymbol{P}_{0,t}F) = 2\mathbb{E}_{\nu_t}\big[(\nabla\sqrt{\boldsymbol{P}_{0,t}F})^2_{\dot C_t}(\varphi)\big], \tag{3.112}$$

where the first equality follows from the independence of t of the following entropy decomposition:

$$\operatorname{Ent}_{\nu_0}(F) = \operatorname{Ent}_{\nu_t}(\boldsymbol{P}_{0,t}F) + \mathbb{E}_{\nu_t}\big[\operatorname{Ent}_{\mu_t^\varphi}(F)\big], \qquad \boldsymbol{P}_{0,t}F(\varphi) = \mathbb{E}_{\mu_t^\varphi}[F]. \tag{3.113}$$

Thus the entropic stability estimate (3.108) leads to a differential inequality for $\mathbb{E}_{\nu_t}\big[\operatorname{Ent}_{\mu_t^\varphi}(F)\big]$. The multiscale Bakry-Émery criterion also gives an entropy contraction but for ν_t instead of μ_t^φ, i.e., it leads to a differential inequality for $\operatorname{Ent}_{\nu_t}(\boldsymbol{P}_{0,t}F)$ as follows.

Remark 3.23. Assume there are $\dot\lambda_t \in \mathbb{R}$ such that

$$\forall \varphi \in X: \qquad \dot C_t \operatorname{Hess} V_t(\varphi) \dot C_t - \frac{1}{2}\ddot C_t \geqslant \dot\lambda_t \dot C_t. \tag{3.114}$$

Then the measure ν_t satisfies the following entropy contraction: for any $t > 0$, with $\lambda_t = \int_0^t \dot\lambda_s \, ds$,

$$2\mathbb{E}_{\nu_t}\big[(\nabla\sqrt{\boldsymbol{P}_{0,t}F})^2_{\dot C_t}(\varphi)\big] \leqslant \left[\frac{\partial}{\partial t}\log\int_0^t e^{-2\lambda_s}\, ds\right]\operatorname{Ent}_{\nu_t}(\boldsymbol{P}_{0,t}F). \tag{3.115}$$

In particular, one again has an entropy contraction estimate:

$$\mathrm{Ent}_{\nu_0}(F) \leqslant \frac{\int_0^\infty e^{-2\lambda_s}\,ds}{\int_0^t e^{-2\lambda_s}\,ds}\mathbb{E}_{\nu_t}\Big[\mathrm{Ent}_{\mu_t^\varphi}(F)\Big]. \tag{3.116}$$

Note that while the left-hand sides of (3.108) and (3.115) are identical (after taking expectation over ν_t), the right-hand side of (3.108) is expressed in terms of the second term on the right-hand side of (3.113) while (3.115) is expressed in terms of the first term.

Proof. The entropy of a test function F decomposes at each scale $t \geqslant 0$ according to (3.113). Under (3.114), the log-Sobolev inequalities for each renormalised measure ν_t provided by Remark 3.7 give, with the same computation as the proof (3.43) of the multiscale Bakry-Émery criterion:

$$\frac{\partial}{\partial t}\,\mathrm{Ent}_{\nu_t}(F_t) = -2\mathbb{E}_{\nu_t}\big[\big(\nabla\sqrt{F_t}\big)^2_{\dot C_t}\big] \leqslant -\gamma_t\,\mathrm{Ent}_{\nu_t}(F_t), \tag{3.117}$$

where $F_t(\varphi) = \boldsymbol{P}_{0,t}F(\varphi) = \mathbb{E}_{\mu_t^\varphi}[F]$ and

$$\gamma_t = \left(\int_t^\infty e^{-2(\lambda_s-\lambda_t)}\,ds\right)^{-1} = -\frac{\partial}{\partial t}\log\int_t^\infty e^{-2\lambda_s}\,ds. \tag{3.118}$$

This implies for each $t \geqslant 0$:

$$\mathrm{Ent}_{\nu_t}(F_t) \leqslant \exp\left[-\int_0^t \gamma_s\,ds\right]\mathrm{Ent}_{\nu_0}(F_0) = \frac{\int_t^\infty e^{-2\lambda_s}\,ds}{\int_0^\infty e^{-2\lambda_s}\,ds}\,\mathrm{Ent}_{\nu_0}(F_0), \tag{3.119}$$

and hence the entropy contraction (3.116) when substituted into (3.113). ☐

Proof of Corollary 3.22. From (3.17), recall that the Polchinski semigroup $\boldsymbol{P}_{0,t}$ coincides with the fluctuation measure μ_t^φ. Thus for smooth $F > 0$ and $\varphi \in X$, one has

$$\begin{aligned}\nabla\log\boldsymbol{P}_{0,t}F(\varphi) = \nabla\log\mathbb{E}_{\mu_t^\varphi}[F] &= \frac{\mathbb{E}_{\mu_t^\varphi}[F\,C_t^{-1}\zeta]}{\mathbb{E}_{\mu_t^\varphi}[F]} - \mathbb{E}_{\mu_t^\varphi}[C_t^{-1}\zeta]\\ &= C_t^{-1}\Big(\mathbb{E}_{\mu_t^{\varphi,F}}[\zeta] - \mathbb{E}_{\mu_t^\varphi}[\zeta]\Big),\end{aligned} \tag{3.120}$$

where the measure modified by F is defined by

$$\frac{d\mu_t^{\varphi,F}}{d\mu_t^\varphi}(\zeta) = \frac{F(\zeta)}{\mathbb{E}_{\mu_t^\varphi}[F]}. \tag{3.121}$$

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

$$\begin{aligned}2(\nabla\sqrt{\boldsymbol{P}_{0,t}F})^2_{\dot C_t}(\varphi) &= \frac{1}{2}(\nabla\log\boldsymbol{P}_{0,t}F)^2_{\dot C_t}(\boldsymbol{P}_{0,t}F)(\varphi)\\ &= \frac{1}{2}(\mathbb{E}_{\mu_t^{\varphi,F}}[\zeta] - \mathbb{E}_{\mu_t^\varphi}[\zeta])^2_{\dot\Sigma_t}(\varphi)\mathbb{E}_{\mu_t^\varphi}[F],\end{aligned} \tag{3.122}$$

and the relative entropy is given by

$$\mathrm{Ent}_{\mu_t^\varphi}(F) = \mathbb{H}(\mu_t^{\varphi,F}|\mu_t^\varphi)\, \mathbb{E}_{\mu_t^\varphi}[F]. \tag{3.123}$$

Assumption (3.107) implies the assumption (3.104) on the covariance of μ_t^φ,

$$\dot\Sigma_t^{1/2}\,\mathrm{Cov}(T_h\mu_t^\varphi)\dot\Sigma_t^{1/2} = \dot\Sigma_t^{1/2}\,\mathrm{Cov}(\mu_t^{\varphi+C_th})\dot\Sigma_t^{1/2} \leqslant \alpha_t\,\mathrm{id} \qquad \text{for all } h\in X, \tag{3.124}$$

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

$$2(\nabla\sqrt{\boldsymbol{P}_{0,t}F}(\varphi))^2_{\dot C_t} \leqslant \alpha_t\,\mathrm{Ent}_{\mu_t^\varphi}(F). \tag{3.125}$$

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 $\varphi\in X$,

$$\mathrm{Hess}\,V_t(\varphi) = C_t^{-1} - C_t^{-1}\,\mathrm{Cov}(\mu_t^\varphi)C_t^{-1}, \tag{3.126}$$

so that

$$\begin{aligned} -\dot C_t\Big(\mathrm{Hess}\,V_t(\varphi) - C_t^{-1}\Big)\dot C_t &= \dot C_t C_t^{-1}\,\mathrm{Cov}(\mu_t^\varphi)C_t^{-1}\dot C_t \\ &= C_t\dot\Sigma_t\,\mathrm{Cov}(\mu_t^\varphi)\dot\Sigma_t C_t \leqslant \alpha_t C_t\dot\Sigma_t C_t = \alpha_t\dot C_t, \end{aligned} \tag{3.127}$$

where the inequality holds if and only if $\dot\Sigma\,\mathrm{Cov}(\mu_t^\varphi)\dot\Sigma_t \leqslant \alpha_t\,\dot\Sigma_t$.

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):

$$\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}\Big[\boldsymbol{P}_{0,t}\Phi(F) - \Phi(\boldsymbol{P}_{0,t}F)\Big] = 2\mathbb{E}_{\nu_t}\Big[(\nabla\sqrt{\boldsymbol{P}_{0,t}F})^2_{\dot C_t}\Big] \tag{3.128}$$

which is the counterpart of eq. (27) in [31]. This is bounded thanks to (3.108):

$$\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}\Big[\boldsymbol{P}_{0,t}\Phi(F) - \Phi(\boldsymbol{P}_{0,t}F)\Big] \leqslant \alpha_t\mathbb{E}_{\nu_t}\Big[\boldsymbol{P}_{0,t}\Phi(F) - \Phi(\boldsymbol{P}_{0,t}F)\Big], \tag{3.129}$$

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

$$\mathbb{E}_{\nu_t}\Big[\boldsymbol{P}_{0,t}\Phi(F) - \Phi(\boldsymbol{P}_{0,t}F)\Big] \leqslant e^{\int_s^t \alpha_u\,du}\,\mathbb{E}_{\nu_s}\Big[\boldsymbol{P}_{0,s}\Phi(F) - \Phi(\boldsymbol{P}_{0,s}F)\Big]. \tag{3.130}$$

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

We finally prove Lemma 3.21 following [31].

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

$$\frac{1}{2}(\mathbb{E}_{T_h\mu}[\zeta] - \mathbb{E}_\mu[\zeta])^2_{\dot\Sigma} \leqslant \alpha \mathbb{H}(T_h\mu|\mu) = \alpha \operatorname{Ent}_\mu\left(\frac{e^{(h,\zeta)}}{\mathbb{E}_\mu[e^{(h,\zeta)}]}\right). \tag{3.131}$$

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

$$\begin{aligned}\operatorname{Ent}_\mu(F) &= \sup_G \Big\{ \mathbb{E}_\mu[FG] - \log \mathbb{E}_\mu[e^G] \Big\} \\ &\geqslant \mathbb{E}_{T_h\mu}[(h,\zeta)] - \log \mathbb{E}_\mu[e^{(h,\zeta)}] = \mathbb{H}(T_h\mu|\mu),\end{aligned} \tag{3.132}$$

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 $\operatorname{Cov}(T_h\mu)$ is strictly positive definite on X for all $h \in X$. Indeed, otherwise consider the largest linear subspace X' such that $\operatorname{Cov}(T_h\mu)$ 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 $\operatorname{Var}_\mu((f,\zeta)) = 0$, 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:

$$\begin{aligned}\operatorname{Var}_{T_h\mu}((f,\zeta)) &\propto \frac{1}{2}\mathbb{E}_{\mu\otimes\mu}[(f,\zeta-\zeta')^2 e^{(h,\zeta+\zeta')}] \\ &\leqslant \frac{1}{2}\mathbb{E}_{\mu\otimes\mu}[(f,\zeta-\zeta')^4]^{1/2}\mathbb{E}_{\mu\otimes\mu}[e^{2(h,\zeta+\zeta')}]^{1/2} = 0.\end{aligned} \tag{3.133}$$

This implies that μ (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.

For $\theta = \mathbb{E}_{T_h\mu}[\zeta]$, the relative entropy of $T_h\mu$ can be written as:

$$\mathbb{H}(T_h\mu|\mu) = \mathbb{E}_{T_h\mu}\left[(h,\zeta)\right] - \log(\mathbb{E}_\mu[e^{(h,\zeta)}]) = (h,\theta) - \log(\mathbb{E}_\mu[e^{(h,\zeta)}]). \tag{3.134}$$

The positive definiteness of $\operatorname{Cov}(T_h\mu)$ 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

$$\Gamma(\theta) = \mathbb{H}(T_{h(\theta)}\mu|\mu) = (\theta, h(\theta)) - \log \mathbb{E}_\mu[e^{(\zeta,h(\theta))}]. \tag{3.135}$$

Thus Γ be the Legendre transform of the cumulant generating function of μ, and

$$\mathbb{H}(T_h\mu|\mu) = \Gamma(\mathbb{E}_{T_h\mu}[\zeta]). \tag{3.136}$$

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

$$\nabla\Gamma(\theta) = (h \mapsto \mathbb{E}_{T_h\mu}[\zeta])^{-1}|_{h=h(\theta)} \tag{3.137}$$

$$\operatorname{Hess}\Gamma(\theta) = \left[\operatorname{Hess}\log\mathbb{E}_\mu[e^{(h,\zeta)}]\Big|_{h=h(\theta)}\right]^{-1} = \operatorname{Cov}(T_{h(\theta)}\mu)^{-1} \tag{3.138}$$

so that

$$\nabla\Gamma(\mathbb{E}_\mu[\zeta]) = 0, \qquad \operatorname{Hess}\Gamma(\mathbb{E}_{T_h\mu}[\zeta]) = \operatorname{Cov}(T_h\mu)^{-1}. \tag{3.139}$$

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

$$\alpha\operatorname{Hess}\Gamma(\theta) \geqslant \dot\Sigma. \tag{3.140}$$

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

$$\alpha\mathbb{H}(T_h\mu|\mu) = \alpha\Gamma(\mathbb{E}_{T_h\mu}[\zeta]) \geqslant \frac{1}{2}(\mathbb{E}_{T_h\mu}[\zeta] - \mathbb{E}_\mu[\zeta])^2_{\dot\Sigma}. \qquad \square \tag{3.141}$$

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:

$$s \leqslant r \leqslant t, \qquad \boldsymbol{P}_{s,t} = \boldsymbol{P}_{r,t}\boldsymbol{P}_{s,r}. \tag{4.1}$$

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

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

$$\mathbb{E}_{\mu_t}[G] = \int G(\boldsymbol{T}_tF)\,d\nu = \int(\boldsymbol{T}_tG)F\,d\nu = \mathbb{E}_{\mu_0}[\boldsymbol{T}_tG]. \tag{4.2}$$

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

Since the Polchinski semigroup is not reversible and time-dependent, this interpretation does not apply to the Polchinski semigroup. Instead, the Polchinski semigroup $\boldsymbol{P}_{s,t}$ 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 $(B_u)_{u\geqslant 0}$ and φ_t, consider the solution to

$$s\leqslant t, \qquad \varphi_s = \varphi_t - \int_s^t \dot C_u\nabla V_u(\varphi_u)\,du + \int_s^t \sqrt{\dot C_u}\,dB_u. \tag{4.3}$$

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 $\varphi_r = \tilde\varphi_{t-r}$ where $\tilde\varphi$ is the solution to the following standard SDE with $\tilde\varphi_0 = \varphi_t$ given and $\tilde B_r = B_t - B_{t-r}$:

$$0 \leqslant r \leqslant t, \qquad d\tilde\varphi_r = -\dot C_{t-r}\, \nabla V_{t-r}(\tilde\varphi_r)dr + \sqrt{\dot C_{t-r}}d\tilde B_r. \tag{4.4}$$

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

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

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

Thus if φ_t is distributed according to the renormalised measure ν_t the backward in time evolution (4.3) ensures that φ_s is distributed according to ν_s for $s < t$.

Our interpretation of this proposition is that, while the renormalised measures ν_t are supported on increasing smooth configurations as t grows, the backward evolution restores the small scale fluctuations of ν_0. Note that for $s = 0$, the identity (4.5) states that the fluctuation measure μ_t^φ introduced in (3.17) is the distribution of the backward process φ_0 conditioned on φ.

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 $\tilde B_r = B_t - B_{t-r}$. Then (4.3) becomes

$$\begin{aligned}\tilde\varphi_r &= \tilde\varphi_0 - \int_{t-r}^t \dot C_u \nabla V_u(\tilde\varphi_{t-u})\, du + \int_{t-r}^t \sqrt{\dot C_u}\, dB_u \\ &= \tilde\varphi_0 - \int_0^r \dot C_{t-u} \nabla V_{t-u}(\tilde\varphi_u)\, du + \int_0^r \sqrt{\dot C_{t-u}}\, d\tilde B_u,\end{aligned} \tag{4.6}$$

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

$$d\tilde f(r,\tilde\varphi_r) = \frac{\partial \tilde f}{\partial r}(r,\tilde\varphi_r) + \boldsymbol{L}_{t-r}\tilde f(r,\tilde\varphi_r) + (\nabla \tilde f(r,\tilde\varphi_r), \sqrt{\dot C_{t-r}}\, d\tilde B_r). \tag{4.7}$$

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

$$df(s,\varphi_s) = \frac{\partial f}{\partial s}(s,\varphi_s) - \boldsymbol{L}_s f(s,\varphi_s) + (\nabla f(s,\varphi_s), \sqrt{\dot C_s} dB_s), \tag{4.8}$$

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

$$\begin{aligned}-d_s f(s,\varphi_s) &= d_r f(t-r,\tilde\varphi_r) \\ &= -\frac{\partial f}{\partial s}(t-r,\tilde\varphi_r) + \boldsymbol{L}_{t-r} f(t-r,\tilde\varphi_r) + (\nabla f(t-r,\tilde\varphi_r), \sqrt{\dot C_{t-r}}\, d\tilde B_r) \\ &= -\frac{\partial f}{\partial s}(s,\varphi_s) + \boldsymbol{L}_s f(s,\varphi_s) - (\nabla f(s,\varphi_s), \sqrt{\dot C_s}\, dB_s).\end{aligned} \tag{4.9}$$

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

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

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 $f(t,\varphi) = \boldsymbol{P}_{s,t}F(\varphi)$ which gives

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

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

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 $a : [0,+\infty] \to [0,+\infty]$. For simplicity, one can choose $a(t) = 1/t$ with $a(0) = +\infty$ and $a(+\infty) = 0$. As in Remark 3.8, set

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

and also

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

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:

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

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 φ (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 $\boldsymbol{P}_{s,t}$ in terms of the above stochastic process, one can alternatively prove Theorem 3.6 using synchronous coupling by adapting the proof from [30] for the Bakry–Émery theorem.

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

$$e^{-2\lambda_t}(\varphi_t-\varphi'_t)^2_{\dot C_t^{-1}} - e^{-2\lambda_s}(\varphi_s-\varphi'_s)^2_{\dot C_s^{-1}}$$
$$= \int_s^t \Big[e^{-2\lambda_u}\Big(-2\dot\lambda_u(\varphi_u-\varphi'_u)^2_{\dot C_u^{-1}} - (\varphi_u-\varphi'_u)^2_{\dot C_u^{-1}\ddot C_u \dot C_u^{-1}}$$
$$+ 2(\dot C_u(\nabla V_u - \nabla V'_u), \dot C_u^{-1}(\varphi_u-\varphi'_u)\Big)\Big]\,du \;\geqslant 0, \tag{4.15}$$

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

$$(\varphi_0-\varphi'_0)^2_{\dot C_0^{-1}} = e^{-2\lambda_0}(\varphi_0-\varphi'_0)^2_{\dot C_0^{-1}} \leqslant e^{-2\lambda_t}(\varphi_t-\varphi'_t)^2_{\dot C_t^{-1}} \tag{4.16}$$

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

$$|\boldsymbol{P}_{0,t}F(\varphi)-\boldsymbol{P}_{0,t}F(\varphi')| = \big|\mathbb{E}\big[(\nabla F(\psi_0),\varphi_0-\varphi'_0)\big]\big|$$
$$= 2\Big|\mathbb{E}\Big[\big(\nabla\sqrt{F(\psi_0)},\sqrt{F(\psi_0)}(\varphi_0-\varphi'_0)\big)\Big]\Big|$$
$$\leqslant 2\mathbb{E}\Big[(\nabla\sqrt{F(\psi_0)})^2_{\dot C_0}\Big]^{1/2}\,\mathbb{E}\Big[F(\psi_0)(\varphi_0-\varphi'_0)^2_{\dot C_0^{-1}}\Big]^{1/2}$$
$$\leqslant 2e^{-\lambda_t}\mathbb{E}\Big[(\nabla\sqrt{F(\psi_0)})^2_{\dot C_0}\Big]^{1/2}\,\mathbb{E}\big[F(\psi_0)\big]^{1/2}\,\sqrt{(\varphi-\varphi')^2_{\dot C_t^{-1}}} \tag{4.17}$$

for some ψ_0 between φ and φ'. Taking $\varphi-\varphi'=\sqrt{\dot C_t}f$ with $|f|_2\to 0$ gives

$$(\nabla\sqrt{\boldsymbol{P}_{0,t}F})^2_{\dot C_t} \leqslant e^{-2\lambda_t}\boldsymbol{P}_{0,t}(\nabla\sqrt{F})^2_{\dot C_0}. \tag{4.18}$$

This is (3.46). □

4.3. Example: coupling with the Gaussian reference measure

Since, by (3.13),

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

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

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

$$\varphi_t = -\int_t^\infty \dot C_u\nabla V_u(\varphi_u)\,du + \int_t^\infty \sqrt{\dot C_u}\,dB_u$$
$$= -\int_t^\infty \dot C_u\nabla V_u(\varphi_u)\,du + \Gamma_t. \tag{4.20}$$

In particular, at $t=0$, this provides a coupling of the full interacting field φ_0 with the Gaussian reference field Γ_0.

As an application of the above coupling, one can relate properties of the Gaussian measure to the interacting one, see for example [22, 12].

4.4. Renormalised 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.

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

and $M_s = V_s(\varphi_s) + \frac{1}{2}\int_s^t (\nabla V_u(\varphi_u))^2_{\dot{C}_u}\, du$ *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),

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

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

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

Proposition 4.4.

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

and $M_s = \nabla V_s(\varphi_s)$ *is a martingale (always with respect to the backwards filtration). Moreover,*

$$\operatorname{Hess} V_t(\varphi) = \mathbb{E}_{\varphi_t=\varphi}\left[\operatorname{Hess} V_0(\varphi_0) - \int_0^t \operatorname{Hess} V_s(\varphi_s)\dot{C}_s \operatorname{Hess} V_s(\varphi_s)\, ds\right] \tag{4.24}$$

and $M_s = \operatorname{Hess} V_s(\varphi_s) - \int_s^t \operatorname{Hess} V_u(\varphi_u)\dot{C}_u \operatorname{Hess} V_u(\varphi_u)\, du$ *is a martingale.*

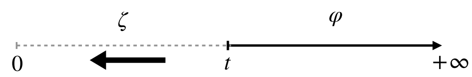

FIG 1. *In the renormalisation group approach, the small scales ζ are averaged out and one considers the projection of the measure to the variables φ encoding the large scales; in the figure above, the Polchinski flow goes from 0 to $+\infty$. In fact, ζ and φ play symmetric roles: in particular for $t = 0$ the original measure is coded by φ, while instead for $t = +\infty$ the original measure is coded by ζ. Stochastic localisation puts the emphasis on the variable ζ and therefore flows in the opposite direction (depicted by the thick arrow).*

Proof. Again, by Itô's formula (4.8) and since $U_s = \nabla V_s$ satisfies $\partial_s U_s = \boldsymbol{L}_s U_s$ by (3.56),

$$dU_s(\varphi_s) = \frac{\partial U_s}{\partial t}(\varphi_s)\,ds - \boldsymbol{L}_s U_s(\varphi_s)\,ds + (\nabla U_s, \sqrt{\dot{C}_s}\,dB_s) = (\nabla U_s, \sqrt{\dot{C}_s}\,dB_s). \tag{4.25}$$

Thus U_t is a martingale, and the expression for its expectation also follows. Similarly, $H_s = \operatorname{Hess} V_s$ satisfies $\partial_s H_s = \boldsymbol{L}_s H_s - H_s \dot{C}_s H_s$ by (3.56), and therefore

$$dH_s(\varphi_s) = -H_s(\varphi_s)\dot{C}_s H_s(\varphi_s)\,ds + (\nabla H_s, \sqrt{\dot{C}_s}\,dB_s) \tag{4.26}$$

so that $H_s(\varphi_s) - \int_s^t \operatorname{Hess} V_u(\varphi_u)\dot{C}_u \operatorname{Hess} V_u(\varphi_u)\,du$ is a martingale. □

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 [45] for a survey on this method and its numerous applications in general, and to [31] for more specific developments on modified log-Sobolev inequalities. The relation between stochastic localisation and a semigroup approach was already pointed out in [62].

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 μ_t^φ defined in (3.17) by

$$\boldsymbol{P}_{0,t} F(\varphi) = \mathbb{E}_{\mu_t^\varphi}[F]. \tag{4.27}$$

The measure μ_t^φ is related to μ_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

$$\nabla V_t(\varphi) = \mathbb{E}_{\mu_t^\varphi}[C_t^{-1}(\varphi - \zeta)] = C_t^{-1}(\varphi - \mathbb{E}_{\mu_t^\varphi}[\zeta]) \tag{4.28}$$

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

$$h_t = C_t^{-1}\varphi_t, \qquad \mu_t = \mu_t^{\varphi_t} = \mu_t^{C_t h_t}, \qquad \dot{\Sigma}_t = -\frac{\partial}{\partial t}C_t^{-1} = C_t^{-1}\dot{C}_t C_t^{-1}. \quad (4.29)$$

Since

$$\dot{C}_t \nabla V_t(\varphi_t) = \dot{C}_t C_t^{-1}(\varphi - \mathbb{E}_{\mu_t^\varphi}[\zeta]) = C_t \dot{\Sigma}_t(\varphi - \mathbb{E}_{\mu_t^\varphi}[\zeta]), \quad (4.30)$$

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 $f(t,\varphi_t) = C_t^{-1}\varphi_t$,

$$\begin{aligned} h_t &= -\int_t^\infty df(u,\varphi_u) \\ &= \int_t^\infty \dot{\Sigma}_u \varphi_u \, du - \int_t^\infty C_u^{-1}\dot{C}_u \nabla V_u(\varphi_u)\, du + \int_t^\infty C_u^{-1}\dot{C}_u^{1/2} dB_u \\ &= \int_t^\infty \dot{\Sigma}_u \mathbb{E}_{\mu_u}[\zeta]\, du + \int_t^\infty C_u^{-1}\dot{C}_u^{1/2} dB_u \\ &= \int_t^\infty \dot{\Sigma}_u \mathbb{E}_{\mu_u}[\zeta]\, du + \int_t^\infty \dot{\Sigma}_u^{1/2} dB_u, \end{aligned} \quad (4.31)$$

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 (4.31) and the fluctuation measure (4.27). 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 ν_t until the fluctuation measure μ_t "localises" to a random Dirac measure $\mu_0 = \delta_{\varphi_0}$ at time $t = 0$, with φ_0 distributed according to the full measure $\nu_0 = \mu_T$. See also Figure 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 (4.29), 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:

$$\begin{aligned} &\tilde{\varphi}_t = \varphi_{T-t}, \qquad \tilde{C}_t = C_{T-t}, \qquad \tilde{V}_t = V_{T-t}, \\ &\tilde{\mu}_t = \mu_{T-t}, \qquad \tilde{h}_t = h_{T-t}, \qquad \dot{\tilde{\Sigma}}_t = \dot{\Sigma}_{T-t}. \end{aligned} \quad (4.32)$$

Using the notation $\mathbf{b}(\mu) = \mathbb{E}_\mu[\zeta]$ for the mean of μ, the SDE (4.31) for $\tilde{h}$ can then be written as:

$$d\tilde{h}_t = \dot{\tilde{\Sigma}}_t \mathbf{b}(\tilde{\mu}_t)\, dt + \dot{\tilde{\Sigma}}_t^{1/2} d\tilde{B}_t. \tag{4.33}$$

This equation is the same as the *stochastic localisation* as it appears for example in [31, 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

$$\tilde{\mu}_t(d\zeta) \propto e^{-H(\zeta) - \frac{t}{2}\zeta^2 + \tilde{h}_t \zeta}\, d\zeta \tag{4.34}$$

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

$$\forall t \geqslant 0: \qquad \mathbb{E}_{\nu_0}[F] = \mathbb{E}\Big[\mathbb{E}_{\tilde{\mu}_t}(F)\Big]. \tag{4.35}$$

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 (4.29) that $\dot{\Sigma}_t = 1$ so that

$$t \geqslant 0: \qquad d\tilde{h}_t = \mathbf{b}(\tilde{\mu}_t)\, dt + d\tilde{B}_t, \qquad \text{with} \quad \tilde{h}_0 = 0. \tag{4.36}$$

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 (4.35). For example, this is a key tool in current progress on the KLS conjecture, see [45] 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 (4.33) but can also be written in terms of our renormalised measure) does not play an important role (see Figure 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 variational and transport-related perspectives on the Polchinski flow. We refer to [37] 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 ν_0 can be realised as the final time distribution φ_0 of the SDE:

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

where we recall that the backwards SDE can be interpreted by reversing time as in (4.14); as pointed out in Remark 3.8, one could have also considered a parametrisation on a bounded time interval. One can ask whether the distribution ν_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 φ_0^U when φ^U is a strong solution of the SDE (again written backward in time):

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

where the parameter t takes values in $[0, +\infty]$ and $\varphi_\infty^U = 0$. More generally, one could consider non-Markovian adapted processes U and the following remains valid, see e.g. [67].

Denote by $\gamma_0 = \mathsf{P}_{C_\infty}$ 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:*

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

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

Proof. Let U be a such that there is a strong solution (φ_t^U) of (5.2) with $\mathbb{E}\big[\int_0^\infty |U_t(\varphi_t^U)|^2_{\dot{C}_t}\, dt\big] < \infty$. By construction φ_0^U has law ν_0 so that the relative entropy is given by

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

with φ^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 $e^{-V_\infty(0)} = \mathsf{E}_{C_\infty}[e^{-V_0(\zeta)}]$ as in (3.13). The renormalised potential follows the Polchinski equation (3.21):

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

Therefore, by Itô's formula,

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

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

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

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

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*

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

Proof. This is essentially a consequence of Girsanov's theorem, see [67] for details. ☐

Since $\mathbb{H}(\mathbf{Q}|\mathbf{P}) \geqslant \mathbb{H}(\nu_0|\gamma_0)$ always holds, by the entropy decomposition (2.52) and the fact that the laws of ν_0 and γ_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 [39], and we refer to [67] 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 ν:

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

minimises the dynamical cost

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

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 an 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 built 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 the process $(\psi_t)_{t\in[0,+\infty]}$ given for $t \geqslant 0$ by

$$\psi_t = -\int_t^\infty \dot{C}_s U_s(\psi_s)\,ds - \int_t^\infty \dot{C}_s \nabla V_s(\psi_s)\,ds + \int_t^\infty \sqrt{\dot{C}_s}\,dB_s, \tag{5.11}$$

the cost $\frac{1}{2}\mathbb{E}\big[\int_0^1 |U_t(\psi_t)|^2_{\dot{C}_t}\,dt\big]$ is minimised and ψ_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

$$W_t(\varphi) = -\log \boldsymbol{P}_{0,t}F(\varphi) = -V_t(\varphi) - \log \mathsf{E}_{C_t}\Big[F(\varphi+\zeta)e^{-V_0(\varphi+\zeta)}\Big]. \tag{5.12}$$

Thus one can check that

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

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

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

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

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

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 μ and ν. 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 [68] for a general overview and to [34] 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 $\nabla V_t = \nabla V_t(\varphi_t)$ and recall that Proposition 4.3 states:

$$V_t(\varphi) = \mathbb{E}_{\varphi_t=\varphi}\left[V_0\Big(\varphi - \int_0^t \dot C_s \nabla V_s\, ds + \int_0^t \sqrt{\dot C_s}\, dB_s\Big) + \frac{1}{2}\int_0^t |\nabla V_s|^2_{\dot C_s}\, ds\right]. \tag{5.16}$$

In particular,

$$V_t(\varphi) \geqslant \inf_U \mathbb{E}\left[V_0\Big(\varphi - \int_0^t \dot C_s U_s\, ds + \int_0^t \sqrt{\dot C_s}\, dB_s\Big) + \frac{1}{2}\int_0^t |U_s|^2_{\dot C_s}\, ds\right], \tag{5.17}$$

where the above infimum is over all adapted processes $U : [0,t] \to X$ (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

$$s \leqslant t: \qquad \varphi_s = \varphi - \int_s^t \dot C_u U_u(\varphi_u)\, du + \int_s^t \sqrt{\dot C_u}\, dB_u. \tag{5.18}$$

The following proposition is a special case of the Boué-Dupuis or Borell formula, see [67], which gives equality in the infimum and is the starting point for the Barashkov–Gubinelli method [10]. An in-depth treatment of stochastic control problems of which this is a special case is given in [51].

Proposition 5.3.

$$V_t(\varphi) = \inf_U \mathbb{E}\left[V_0\Big(\varphi - \int_0^t \dot C_s U_s\, ds + \int_0^t \sqrt{\dot C_s}\, dB_s\Big) + \frac{1}{2}\int_0^t |U_s|^2_{\dot C_s}\, ds\right]. \tag{5.19}$$

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

$$V_t(\varphi) = -\log \mathsf{E}_{C_t}[e^{-V_0(\varphi+\zeta)}] \leqslant \mathrm{Ent}_{\mathsf{P}_{C_t}}(F) + \mathsf{E}_{C_t}\left[F(\zeta)V_0(\varphi+\zeta)\right]. \tag{5.20}$$

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

$$\varphi_0 - \varphi = -\int_0^t \dot C_s U_s\, ds + \int_0^t \sqrt{\dot C_s}\, dB_s. \tag{5.21}$$

Then

$$V_t(\varphi) \leqslant \mathbb{E}\left[\frac{1}{2}\int_0^t |U_s(\varphi_s)|^2_{\dot C_s}\, ds + V_0\Big(\varphi - \int_0^t \dot C_s U_s\, ds + \int_0^t \sqrt{\dot C_s}\, dB_s\Big)\right], \tag{5.22}$$

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. □

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 ν_0. For further details and application to construction of the φ^4_d measures, we refer to [10, 11].

5.3. Lipschitz transport

Instead of a stochastic process, one could also define a map $\hat{S}_t : X \mapsto X$ transporting some measure $\hat{\nu}_t$ to the desired target measure ν_0:

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

Under an assumption on its gradient, such a transport map allows to recover functional inequalities for ν_0 from $\hat{\nu}_t$ as follows. For example, assume that $\hat{\nu}_t$ satisfies a log-Sobolev inequality (with quadratic form $(\cdot,\cdot)_Q$ and denoting the corresponding norm $|f|_Q^2 = (f,f)_Q$ on X):

$$\mathrm{Ent}_{\hat{\nu}_t}[F] \leqslant 2\mathbb{E}_{\hat{\nu}_t}\Big[|\nabla\sqrt{F}|_Q^2\Big], \tag{5.24}$$

and that the transport map $\hat{S}_t$ from $\hat{\nu}_t$ to ν_0 has Jacobian $\nabla\hat{S}_t(\varphi)$ satisfying the uniform bound:

$$\forall \varphi \in X, f \in X: \qquad |{}^t\nabla\hat{S}_t(\varphi)f|_Q^2 \leqslant C^2|f|^2. \tag{5.25}$$

Then ν_0 satisfies also a log-Sobolev inequality:

$$\begin{aligned}
\mathrm{Ent}_{\nu_0}\left[F(\varphi)\right] &= \mathbb{E}_{\hat{\nu}_t}[\Phi(F\circ\hat{S}_t)] - \Phi(\mathbb{E}_{\hat{\nu}_t}[F\circ\hat{S}_t]) \\
&\leqslant 2\mathbb{E}_{\hat{\nu}_t}\left[\Big|\nabla\sqrt{F\circ\hat{S}_t}\Big|_Q^2\right] \\
&= 2\mathbb{E}_{\hat{\nu}_t}\left[\Big|{}^t\nabla\hat{S}_t\,(\nabla\sqrt{F}\circ\hat{S}_t)\Big|_Q^2\right] \\
&\leqslant 2C^2\mathbb{E}_{\hat{\nu}_t}\left[\Big|(\nabla\sqrt{F}\circ\hat{S}_t)\Big|^2\right] \\
&= 2C^2\mathbb{E}_{\nu_0}\left[|\nabla\sqrt{F}|^2\right].
\end{aligned} \tag{5.26}$$

An analogous argument can be applied to more general functional inequalities.

This line of research was first investigated in [28] where it was understood that a convex perturbation of a Gaussian measure leads to a 1-Lipschitz transport map $\hat{S}_t$ to this Gaussian measure, i.e., (5.27) holds with $Q = \mathrm{id}$ and $C = 1$. We refer to [61, 79, 76, 75, 62, 87, 49] 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 [88] which derived a Lipschitz estimate of the form (5.25) from the multiscale Bakry–Émery criterion (3.32) for the covariance decomposition $\dot{C}_t = e^{-tA}$, and then we generalise the result also to other decompositions relevant for applications (see Section 6.3). For this generalisation, it is important that Q in the condition (5.25) is not necessarily equal to the identity. Using the identity $\|M\| = \|{}^tM\|$ ($\|\cdot\|$ denotes the operator norm), this condition can also be equivalently stated in terms of $\nabla\hat{S}_t$ instead of ${}^t\nabla\hat{S}_t$ as

$$|\nabla\hat{S}_t(\varphi)f|^2 \leqslant C^2|f|_{Q^{-1}}. \tag{5.27}$$

Recall that the measure ν_0 gets renormalised to ν_t by the Polchinski flow. By construction $\mathbb{E}_{\nu_t}[\cdot] \propto \mathsf{E}_{C_\infty - C_t}[e^{-V_t}\cdot]$ and ν_t converges to a Dirac mass. As the measure ν_t degenerates, it is more convenient to consider the measure $\hat\nu_t$ obtained by rescaling ν_t by some matrix D_t so that the measures ν_0 and $\hat\nu_t$ are comparable:

$$\mathbb{E}_{\hat\nu_t}\left[F(\varphi)\right] = \mathbb{E}_{\nu_t}\left[F(D_t\varphi)\right]. \tag{5.28}$$

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

$$D_t^{-1}(C_\infty - C_t)^{-1}D_t^{-1} = A \quad \Rightarrow \quad D_t A D_t = (C_\infty - C_t)^{-1}, \tag{5.29}$$

where all inverses are understood to be taken on the range of A. This is implicitly assumed in the rest of the section, with id also denoting the identity matrix on the range of A. Assuming that all the matrices depend smoothly on t and commute, the choice (5.29) implies the following useful relation:

$$2D_t^{-1}\dot D_t = \dot C_t (C_\infty - C_t)^{-1} = \dot C_t D_t A D_t. \tag{5.30}$$

Using again (5.29) we get (recall $C_0 = 0$ and notice $D_0 = \mathrm{id}$):

$$\frac{\partial}{\partial t}(D_t^{-2}) = -A\dot C_t \quad \Rightarrow \quad D_t = (\mathrm{id} - AC_t)^{-1/2}. \tag{5.31}$$

By construction $\lim_{t\to\infty} D_t^{-1}\varphi = 0$ and the renormalised potential satisfies $\lim_{t\to\infty} V_t(D_t^{-1}\varphi) = V_\infty(0)$. This follows from the representation (3.11) of V_t and the standing assumption that V_0 is bounded below. In the same way, we can also show that the following convergence in distribution to a Gaussian measure holds:

$$\lim_{t\to\infty} \mathbb{E}_{\hat\nu_t}\left[F(\varphi)\right] = \mathsf{E}_{C_\infty}\left[F(\varphi)\right] = \mathsf{E}_{A^{-1}}\left[F(\varphi)\right]. \tag{5.32}$$

We are now going to study the transport map $\hat S_t$ between ν_t and $\hat\nu_0$ defined in (5.23).

The properties of $\hat S_t$ are sensitive to the covariance decomposition $\dot C_t$. It was realised in [88] that for the choice $\dot C_t = e^{-tA}$ (often used in applications, see Section 6) the Lipschitz structure associated with $\hat S_t$ is directly related to (a variant of) the multiscale Bakry–Émery criterion (3.32).

Theorem 5.4 ([88])**.** *Let $\dot C_t = e^{-tA}$ for $t \geqslant 0$. Under the assumption*

$$\forall \varphi \in X: \qquad \dot C_t^{1/2}\,\mathrm{Hess}\, V_t(\varphi)\dot C_t^{1/2} \geqslant \dot\mu_t\,\mathrm{id}, \quad \textit{with} \quad \mu_t := \int_0^t ds\,\dot\mu_s, \tag{5.33}$$

the transport map $\hat S_t : X \mapsto X$ introduced in (5.23) is $\exp(-\frac{1}{2}\mu_t)$-Lipschitz, i.e., (5.27) holds with $Q = \mathrm{id}$ and $C = \exp(-\frac{1}{2}\mu_t)$.

From the convergence (5.32) 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 $\mathsf{P}_{C_\infty} = \mathsf{P}_{A^{-1}}$ to ν_0 (see [79, Lemma 2.1]).

Another useful covariance decomposition (see Theorem 6.2) is of the form $\dot{C}_t = (tA + \mathrm{id})^{-2}$. The proof of Theorem 5.4 can be extended to that case as follows.

Theorem 5.5. *Let $\dot{C}_t = (tA + \mathrm{id})^{-2}$ for $t \geqslant 0$. Under the multiscale Bakry-Émery criterion*

$$\forall \varphi \in X : \qquad \dot{C}_t \operatorname{Hess} V_t(\varphi) \dot{C}_t - \frac{1}{2}\ddot{C}_t \geqslant \dot{\lambda}_t \dot{C}_t, \quad \textit{with} \quad \lambda_t := \int_0^t ds\, \dot{\lambda}_s, \quad (5.34)$$

the inverse map $\hat{S}_t = S_t^{-1} : X \mapsto X$ satisfies

$$\forall \varphi \in X, f \in X : \qquad |\nabla \hat{S}_t(\varphi) f|^2 \leqslant e^{-\lambda_t} |\sqrt{1+tA}\, f|^2 = e^{-\lambda_t} |f|^2_{1+tA}. \quad (5.35)$$

Remark 5.6. Contrary to the Lipschitz transport map of Theorem 5.4, the gradient of the map of Theorem 5.5 is bounded with respect to a different input norm that increases with t. In particular, if $A = -\Delta + 1$, then $|\cdot|_{1+tA}$ is a (discrete) Sobolev norm. In our examples (see Section 6.3), the constants λ_t diverge like $\log(1+t)$ (while the constants μ_t in Theorem 5.4 remain bounded) and therefore the combination of $e^{-\lambda_t}(1+tA)$ remains bounded by A uniformly in t. Thus $\nabla \hat{S}_t(\varphi)$ is uniformly bounded from $|\cdot|_A$ to $|\cdot|$ and thus ${}^t\nabla \hat{S}_t(\varphi)$ from $|\cdot|$ to $|\cdot|_{A^{-1}}$. As seen in (5.26), when using transport maps to prove log-Sobolev inequalities, one can use that the Gaussian measure $\hat{\nu}_\infty$ satisfies a log-Sobolev inequality with quadratic form $(\cdot,\cdot)_{A^{-1}}$ to compensate the loss of regularity in the transport map and recover a log-Sobolev inequality with standard quadratic form.

Proof of Theorem 5.4. Consider S_t a transport map between ν_0 and $\hat{\nu}_t$, so that

$$\mathbb{E}_{\nu_0}\left[F\big(S_t(\varphi)\big)\right] = \mathbb{E}_{\hat{\nu}_t}\left[F(\varphi)\right] = \mathbb{E}_{\nu_t}\left[F(D_t\varphi)\right]. \quad (5.36)$$

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 $S_t : X \to X$ for general general covariance decompositions $\dot{C}_t$ and use the precise form only to conclude the proof. On the one hand,

$$\frac{\partial}{\partial t}\mathbb{E}_{\nu_0}\left[F\big(S_t(\varphi)\big)\right] = \mathbb{E}_{\nu_0}\left[\big(\nabla F\big(S_t(\varphi)\big), \partial_t S_t(\varphi)\big)\right], \quad (5.37)$$

and on the other hand from (3.27):

$$\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}\left[F(D_t\varphi)\right] = \mathbb{E}_{\nu_t}\left[-\boldsymbol{L}_t F(D_t\varphi) + (\dot{D}_t\varphi, \nabla F(D_t\varphi))\right], \quad (5.38)$$

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

$$\begin{aligned}
&\frac{\partial}{\partial t}\mathbb{E}_{\nu_t}\left[F(D_t\varphi)\right]\\
&\quad= \mathbb{E}_{\nu_t}\left[\frac{1}{2}(\dot{C}_t\nabla V_t(\varphi) - \frac{1}{2}\dot{C}_t\,(C_\infty - C_t)^{-1}\varphi + D_t^{-1}\dot{D}_t\varphi, D_t\nabla F(D_t\varphi))\right]\\
&\quad= \mathbb{E}_{\nu_t}\left[\frac{1}{2}(\nabla V_t(\varphi), D_t\nabla F(D_t\varphi))_{\dot{C}_t}\right]\\
&\quad= \mathbb{E}_{\nu_0}\left[\frac{1}{2}(\nabla V_t(D_t^{-1}\cdot), D_t\nabla F)_{\dot{C}_t}(S_t(\varphi))\right],
\end{aligned} \tag{5.39}$$

where we used (5.30). Thus the evolution of S_t is given by

$$\frac{\partial}{\partial t}S_t(\varphi) = \frac{1}{2}D_t\,\dot{C}_t\nabla V_t(D_t^{-1}S_t(\varphi)). \tag{5.40}$$

For $\dot{C}_t = e^{-tA}$, the definition (5.31) implies $D_t = e^{\frac{1}{2}tA} = \dot{C}_t^{-1/2}$ and the evolution (5.40) becomes particularly simple:

$$\frac{\partial}{\partial t}S_t(\varphi) = \frac{1}{2}\dot{C}_t^{1/2}\nabla V_t\big(\dot{C}_t^{1/2}S_t(\varphi)\big). \tag{5.41}$$

The Jacobian evolves according to

$$\frac{\partial}{\partial t}\nabla S_t(\varphi) = \frac{1}{2}\dot{C}_t^{1/2}\operatorname{Hess}V_t\big(\dot{C}_t^{1/2}S_t(\varphi)\big)\dot{C}_t^{1/2}\nabla S_t(\varphi). \tag{5.42}$$

As a consequence of (5.33), the Grönwall inequality implies

$$\forall f\in X: \quad \frac{\partial}{\partial t}|\nabla S_t(\varphi)f|^2 \geqslant \dot{\mu}_t(\nabla S_t(\varphi)f)^2 \quad \Rightarrow \quad |\nabla S_t(\varphi)f|^2 \geqslant \exp(\mu_t)|f|^2. \tag{5.43}$$

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

Proof of Theorem 5.5. For each $t \geqslant 0$, 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 Grönwall estimate for $(B_t\nabla S_t(\varphi)f)^2$ for each $\varphi, f \in X$. Using that $\dot{C}_t = (tA + \mathrm{id})^{-2}$, the definition (5.31) of D_t implies $D_t = \dot{C}_t^{-1/4}$, and in particular D_t commutes with $B_s, C_s, \dot{C}_s, \ddot{C}_s$ for any s. Equation (5.40) gives:

$$\frac{\partial}{\partial t}\nabla S_t(\varphi) = \frac{1}{2}\dot{C}_tD_t\operatorname{Hess}V_t(D_t^{-1}S_t(\varphi))D_t^{-1}\nabla S_t(\varphi). \tag{5.44}$$

Therefore:

$$\begin{aligned}
&\frac{\partial}{\partial t}\big|B_t\nabla S_t(\varphi)f\big|^2\\
&= \Big(B_t\nabla S_t(\varphi)f, \Big[B_t\dot{C}_tD_t\operatorname{Hess}V_t(D_t^{-1}S_t(\varphi))D_t^{-1}B_t^{-1} + 2\dot{B}_tB_t^{-1}\Big]B_t\nabla S_t(\varphi)f\Big).
\end{aligned} \tag{5.45}$$

In order to use the Hessian bound (5.34), we need to choose B_t in such a way that, for each $t \geqslant 0$:

$$B_t \dot{C}_t D_t \operatorname{Hess} V_t(D_t^{-1} S_t(\varphi)) D_t^{-1} B_t^{-1} + 2\dot{B}_t B_t^{-1} \geqslant \dot{C}_t^{1/2} \operatorname{Hess} V_t(D_t^{-1} S_t(\varphi)) \dot{C}_t^{1/2} - \frac{1}{2} \dot{C}_t^{-1/2} \ddot{C}_t \dot{C}_t^{-1/2}. \tag{5.46}$$

With the choice $B_t = D_t^{-1} \dot{C}_t^{-1/2} = \dot{C}_t^{-1/4} = D_t$, we have $\dot{B}_t B_t^{-1} = -\frac{1}{4} \ddot{C}_t \dot{C}_t^{-1}$ and the left- and right-hand sides in (5.46) are in fact equal. As a result,

$$\forall t \geqslant 0: \qquad \frac{\partial}{\partial t} \big| D_t \nabla S_t(\varphi) f \big|^2 \geqslant \dot{\lambda}_t \big| D_t \nabla S_t(\varphi) f \big|^2. \tag{5.47}$$

The Grönwall inequality then yields:

$$\forall t \geqslant 0: \qquad \big| D_t \nabla S_t(\varphi) f \big|^2 \geqslant e^{\lambda_t} |f|^2, \tag{5.48}$$

and the inverse function theorem yields

$$\forall t \geqslant 0: \qquad \big| \nabla \hat{S}_t(\varphi) D_t^{-1} f \big|^2 \leqslant e^{-\lambda_t} |f|^2, \tag{5.49}$$

which is equivalent to the claim. □

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:

$$\forall \varphi \in X, t \geqslant 0: \qquad \dot{C}_t \operatorname{Hess} V_t(\varphi) \dot{C}_t - \frac{1}{2} \ddot{C}_t \geqslant \dot{\lambda}_t \dot{C}_t. \tag{6.1}$$

These lower bounds depend on 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. We expect that the precise choice of the covariance decomposition is technical, as long as it takes into account the important physical features of the model, e.g., the mode structure explained in Example 2.12.

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 $d \geqslant 2$, where either $\varepsilon = 1$ is fixed and $L \to \infty$ (statistical mechanics model) or $\varepsilon \to 0$ is a small regularisation parameter and L is fixed or $L \to \infty$ (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 $\varepsilon \to 0$.

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

1. 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)?
2. Does it give a bound on the log-Sobolev constant independent of the regularisation parameter ε for continuum models?
3. 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) very generally, see, e.g., Theorem 6.11 for the specific case of the Ising model (which extends similarly to a much broader class of models). This regime is also covered by many other criteria (see, e.g., the monographs [85, 55] and [95]), with the notable exception of mean-field spin-glass models, see the discussion in [13], for which the spectral nature of the criterion is important.
- The criterion (6.1) can be sharp enough to reach the critical point, see, e.g., Theorem 6.17 and Example 6.4 below for the Ising and φ^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 $\varepsilon \to 0$ (see Theorems 6.1 and 6.2 for the φ^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. The simplest models are ones with quadratic mean-field interaction such as the Curie-Weiss model, in which case one can answer (iii) in the affirmative. Using this perspective, a detailed analysis above and below the critical temperature was also carried out for mean-field $O(n)$ models in [24]. For certain more general continuum particle systems with mean-field interaction, the behaviour close to the critical point is the subject of ongoing work [16]. For more models with more complicated spatial structure in which computations can still be carried out, such as the hierarchical φ^4_4 model [14], 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 $d \geqslant 5$, the criterion implies polynomial bounds on the log-Sobolev constant at and near the critical temperature [20].

6.1. Applications to Euclidean field theory

In this section, $\Lambda = \Lambda_{\varepsilon,L}$ will be a discrete torus of mesh size ε and side length L (assumed to be a multiple of ε), i.e., $\Lambda_{\varepsilon,L} = L\mathbb{T}^d \cap \varepsilon\mathbb{Z}^d$, and the discrete Laplacian on $\Lambda_{\varepsilon,L}$ is given by

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

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

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

where $d\varphi$ denotes the Lebesgue measure on $\mathbb{R}^{\Lambda_{\varepsilon,L}}$ and the single-site potential V^ε is a real-valued function chosen in such a way that the $\varepsilon \to 0$ limit of the measure exists (in a suitable space of generalised functions) 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 (lattice regularised) SPDE

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

where $dB^{\varepsilon,L}$ is space-time white noise on $\mathbb{R}_+ \times \Lambda_{\varepsilon,L}$, i.e., the $t \mapsto B_{t,x}^{\varepsilon,L}$ are independent Brownian motions with variance ε^{-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 $\varepsilon \to 0$ limit of (6.4) is a singular SPDE. The pathwise (short time) limit theory for such SPDEs is the subject of Hairer's regularity structure theory [58, 56, 57], the paracontrolled method of Gubinelli et al. [54, 29, 97], and the pathwise renormalisation group approach [65, 44]. The log-Sobolev inequality for the associated dynamics takes the form:

$$\mathrm{Ent}_{\nu^{\varepsilon,L}}(F) \leqslant \frac{2}{\gamma} D_{\nu^{\varepsilon,L}}(\sqrt{F}), \tag{6.5}$$

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

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

i.e., $(\nabla^\varepsilon F)_x = (\nabla^\varepsilon_\varphi F)_x = \varepsilon^{-d} \frac{\partial F}{\partial \varphi_x}$. (Thus this gradient acts on functionals of fields $F: \mathbb{R}^{\Lambda_{\varepsilon,L}} \to \mathbb{R}$ 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 single-site potential

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

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

Continuum φ^4 model Let $d = 2$ or $d = 3$. For $g > 0$ and $r \in \mathbb{R}$, the φ^4_d measure is defined by

$$\forall \varphi \in \mathbb{R} : \qquad V^\varepsilon(\varphi) = \frac{g}{4}\varphi^4 + \frac{r + a^\varepsilon(g)}{2}\varphi^2, \tag{6.8}$$

where $a^\varepsilon(g)$ is a divergent counterterm. The notation g, r of the parameters introduced in (2.5) are also used for the continuum setting. Explicitly, for an arbitrary fixed $m^2 > 0$, one can take $a^\varepsilon(g) = a^\varepsilon(g, m^2)$ with

$$a^\varepsilon(g, m^2) := -3g\big(-\Delta^\varepsilon + m^2\big)^{-1}(0,0) + 6g^2\big\|\big(-\Delta^\varepsilon + m^2\big)^{-1}(0,\cdot)\big\|^3_{L^3(\Lambda_{\varepsilon,L})}, \tag{6.9}$$

and the notation $\|f\|^p_{L^p(\Lambda_{\varepsilon,L})} = \varepsilon^d \sum_{x\in\Lambda_{\varepsilon,L}} |f(x)|^p$ 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 φ^4_d 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 $\dot\lambda_t \in \mathbb{R}$,

$$\forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon,L}}, t \geqslant 0 : \qquad \dot C_t \operatorname{Hess} V_t(\varphi)\dot C_t - \frac{1}{2}\ddot C_t \geqslant \dot\lambda_t \dot C_t. \tag{6.10}$$

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

For the sine-Gordon model, this was carried out in [15] by providing an explicit description of the renormalised potential at each scale following [27]. By writing the renormalised potential as the Fourier type series

$$V_t(\varphi) = \sum_{n=0}^\infty \frac{1}{n!}\int_{(\Lambda_{\varepsilon,L}\times\{\pm1\})^n} \tilde V_t(\xi_1,\dots,\xi_n) e^{i\sum_{i=1}^n \sigma_i\varphi_{x_i}}\, d\xi_1\cdots d\xi_n, \tag{6.11}$$

where $\xi_i = (x_i, \sigma_i) \in \Lambda_{\varepsilon,L} \times \{\pm1\}$ and we used the notation

$$\int_{\Lambda_{\varepsilon,L}\times\{\pm1\}} F(\xi)d\xi = \varepsilon^d \sum_{x\in\Lambda_{\varepsilon,L}} \sum_{\sigma\in\{\pm1\}} F(x,\sigma), \tag{6.12}$$

the Polchinski equation for V_t reduces to a triangular system of ODEs for the Fourier coefficients $\tilde{V}_t(\xi_1, \dots, \xi_n)$. For $\beta < 6\pi$ one can obtain the following control on $\operatorname{Hess} V_t$ by estimating these Fourier coefficients. It remains an interesting problem to extend such estimates to the optimal regime $\beta < 8\pi$. (In this regime $\beta < 8\pi$, weaker estimates that are sufficient for the construction of the limiting continuum measure are known [41, 81], see also the discussion in [23]. These estimates are however insufficient for our purposes.)

Theorem 6.1. *For the massive continuum sine-Gordon model with mass $m > 0$, let $A = -\Delta^\varepsilon + m^2 \mathrm{id}$ and $\dot{C}_t = e^{-tA}$ ($t \geqslant 0$). Then, if $\beta < 6\pi$, there is a constant $\mu^* = \mu^*(\beta, z, m, L) > 0$ that does not depend on ε, t, such that with*

$$V_0(\varphi) = \varepsilon^2 \sum_{x \in \Lambda_{\varepsilon,L}} 2\varepsilon^{-\frac{\beta}{4\pi}} z \cos(\sqrt{\beta}\varphi_x) \tag{6.13}$$

the renormalised potential satisfies

$$\forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon,L}}, t \geqslant 0: \qquad \dot{C}_t \operatorname{Hess} V_t(\varphi) \dot{C}_t \geqslant \dot{\mu}_t \dot{C}_t \quad \textit{with} \quad \sup_{t \geqslant 0} \Big| \int_0^t \dot{\mu}_s ds \Big| \leqslant \mu^*. \tag{6.14}$$

Given β, z, m, L, this yields a lower bound, uniform in ε, on the log-Sobolev constant $\inf_\varepsilon \gamma^{\varepsilon,L}(\beta, z, m) > 0$.

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

For the sine-Gordon model, the multiscale Bakry-Émery criterion of Theorem 3.6 is thus seen to provide the optimal independence of ε of the log-Sobolev constant in finite volume, as well as of L under an additional small coupling assumption depending on the external mass. Under the condition (6.14) derived in Theorem 6.1, Shenfeld [88] also used the multiscale Bakry-Émery criterion to construct a Lipschitz transport map between the continuum sine-Gordon model and the free field with the same mass (i.e., the model (6.3) with $V^\varepsilon(\varphi) = \frac{1}{2}m^2\varphi^2$), see Example 6.5.

The log-Sobolev inequality for the φ^4_d model (where $d = 2, 3$) was obtained in [21]. In this case, a sufficiently strong description of the renormalised potential is difficult to obtain directly. To provide some context, we remark that Polchinski's original article [83] assumes a representation of V_t as a *formal* power series analogous to (6.11),

$$V_t(\varphi) = \sum_{m=0}^{\infty} \frac{1}{(2m)!} \int_{(\Lambda_{\varepsilon,L})^{2m}} \tilde{V}_t(x_1, \dots, x_{2m}) \varphi_{x_1} \cdots \varphi_{x_{2m}} \, dx_1 \cdots dx_{2m}. \tag{6.15}$$

In this representation, the Polchinski equation again formally reduces to a system of ODEs for the kernels $\tilde{V}_t$, and this representation is very successful for understanding the renormalised potential as a formal power series. However, the series representation in powers of φ is necessarily divergent as it would imply analyticity of the solution around 0 as a function of the coupling constant g in front of the φ^4 term. It therefore seems difficult to use this naive expansion in powers of φ to obtain nonperturbative analytic information about V_t, as is needed for the construction of the measure and even more so to prove the log-Sobolev inequality.

Nonetheless, the renormalised potential itself is well defined, and it is immediate from the definition (3.11) that its derivatives correspond to correlation functions of the fluctuation measure μ_t^φ defined in (3.18). In particular, by Lemma 3.12,

$$\operatorname{Hess} V_t(\varphi) = C_t^{-1} - C_t^{-1} \operatorname{Cov}(\mu_t^\varphi) C_t^{-1}, \tag{6.16}$$

where $\operatorname{Cov}(\mu_t^\varphi)$ is the covariance matrix of μ_t^φ. Thus understanding the Hessian of the renormalised potential is equivalent to understanding the two-point function of the fluctuation measure. To analyse this two-point function, the following choice of covariance decomposition (known as Pauli–Villars regularisation) is very helpful:

$$C_t = (A + 1/t)^{-1}. \tag{6.17}$$

Indeed, with this choice, the fluctuation measure μ_t^φ has the same structure as the original measure, except for an additional mass term $\frac{1}{2t}\zeta^2$ and an external field term $(\zeta, C_t^{-1}\varphi)$. Therefore various correlation inequalities available for the original φ^4 measure also apply to the fluctuation measure. These include the FKG inequality (for all φ) which shows that the covariance matrix has positive entries, the Griffiths inequalities (when $\varphi = 0$) which shows that the two-point function is monotone in t and μ, and the recent inequality of Ding–Song–Sun (DSS) [42] which shows that the covariance matrix is entrywise maximised at $\varphi = 0$. In this setting, with $\dot C_t = (tA+1)^{-2}$ and $-\frac{1}{2}\ddot C_t = A(tA+1)^{-3}$, the left-hand side of the multiscale Bakry–Émery criterion becomes:

$$\dot C_t \operatorname{Hess} V_t(\varphi) \dot C_t - \frac{1}{2}\ddot C_t = \dot C_t^{1/2}\Big(\frac{1}{t} - \frac{\operatorname{Cov}(\mu_t^\varphi)}{t^2}\Big)\dot C_t^{1/2}, \tag{6.18}$$

and the combination of the FKG and DSS inequalities with the Perron–Frobenius theorem imply that the right-hand side is lower bounded as a quadratic form by the following term involving only $\varphi = 0$ (where $\|M\|$ denotes the operator norm of a matrix M):

$$\Big(\frac{1}{t} - \frac{\|\operatorname{Cov}(\mu_t^{\varphi=0})\|}{t^2}\Big)\dot C_t, \tag{6.19}$$

see [21] for details.

This representation is the starting point for the proof of the following theorem, proved in [21]. For concreteness, we choose the counterterms (6.9) in the definition of the φ^4_d model with $m = 1$ and denote the 0-field susceptibility of the φ^4_d model by

$$\chi^{\varepsilon,L}(g,r) := \varepsilon^d \sum_{x\in\Lambda_{\varepsilon,L}} \big\langle \varphi_0\varphi_x \big\rangle_{g,r}^{\varepsilon,L}, \tag{6.20}$$

where $\langle \cdot \rangle_{g,r}^{\varepsilon,L}$ denotes the expectation of the φ_d^4 measure. The second Griffiths inequality implies that $\chi^{\varepsilon,L}(g,r)$ is monotone in r. We then define the critical point by

$$r_c(g) = \inf \Big\{ r \in \mathbb{R} : \chi^{\varepsilon,L}(g,r) \text{ is uniformly bounded as } \varepsilon \to 0 \text{ and } L \to \infty \Big\}. \tag{6.21}$$

Choosing a linear test function, the Poincaré constant of the φ_d^4 model (and thus the log-Sobolev constant) is not uniformly bounded away from 0 when $r < r_c(g)$. The following theorem shows that it is bounded when $r > r_c(g)$.

Theorem 6.2. *For the continuum φ^4 model in $d = 2, 3$, let $A = -\Delta^\varepsilon + \mathrm{id}$ and $\dot{C}_t = (tA + \mathrm{id})^{-2}$ $(t \geqslant 0)$ and take $m = 1$ in the definition (6.9) of the counterterm. Then with*

$$V_0(\varphi) = \varepsilon^d \sum_{x \in \Lambda_{\varepsilon,L}} \Big(\frac{g}{4}\varphi_x^4 + \frac{r - 1 + a^\varepsilon(g)}{2}\varphi_x^2\Big), \tag{6.22}$$

the renormalised potential satisfies

$$\forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon,L}}, t \geqslant 0 : \qquad \dot{C}_t \operatorname{Hess} V_t(\varphi) \dot{C}_t - \frac{1}{2}\ddot{C}_t \geqslant \Big(\underbrace{\frac{1}{t} - \frac{\chi(g, r + 1/t)}{t^2}}_{\dot{\lambda}_t}\Big)\dot{C}_t, \tag{6.23}$$

and $\lambda_t \geqslant \log(1+t) - C(g,r,L)$ uniformly in $\varepsilon > 0$ for any $g > 0$, $r \in \mathbb{R}$ and any fixed L, and $C(g,r,L)$ is independent of L if $r > r_c(g)$. In particular, the following integral is bounded under the same conditions:

$$\int_0^\infty e^{-2\lambda_t}\, dt, \qquad \lambda_t = \int_0^t \dot{\lambda}_s\, ds. \tag{6.24}$$

The log-Sobolev constant $\gamma^{\varepsilon,L}(g,r)$ thus satisfies $\inf_{\varepsilon,L} \gamma^{\varepsilon,L}(g,r) > 0$ for any $g > 0, r \in \mathbb{R}$ for which the measure does not have a phase transition ($r > r_c(g)$), and $\inf_\varepsilon \gamma^{\varepsilon,L}(g,r) > 0$ for any $g > 0, r \in \mathbb{R}$ and L fixed.

In the φ_d^4 case, the multiscale criterion gives a log-Sobolev constant bounded uniformly in the volume and ε in the entire single-phase region. However, the bound obtained on the log-Sobolev constant has accurate dependence on g, r only far away from the transition, corresponding to values of g, r such that $r > 0$ and $g \ll r$. In $d > 4$ where the φ^4 model does not have a continuum limit, one can obtain a polynomial bound near the critical point on the log-Sobolev constant for the lattice φ_d^4 model. The same applies to the Ising model. The next sections detail these proofs of the log-Sobolev inequality for the lattice φ^4 and Ising models.

6.2. Applications to lattice φ^4 models

The lattice φ^4 model corresponds to the case $\varepsilon = 1$ in (6.22), and we denote its single-site potential for $g > 0, r \in \mathbb{R}$ by:

$$V(\varphi) = \frac{g}{4}\varphi^4 + \frac{r}{2}\varphi^2, \tag{6.25}$$

and choose the coupling matrix as $A = -\Delta$ where Δ is the lattice Laplace operator corresponding to $\varepsilon = 1$ in (6.2). Boundary conditions do not matter much, but for concreteness we consider periodic boundary conditions, i.e., the state space is $\Lambda_L = \mathbb{Z}^d/(L\mathbb{Z})^d$ for any $d \geqslant 1$. The associated expectation is denoted by $\langle \cdot \rangle_{g,r}$ and the finite volume susceptibility is defined by

$$\chi^L(g,r) = \sum_{x\in\Lambda_L} \langle \varphi_0\varphi_x \rangle_{g,r}. \tag{6.26}$$

The critical point $r_c(g)$ is again defined as the infimum over all $r \in \mathbb{R}$ such that $\chi^L(g,r)$ is bounded from above uniformly in L. It is the case that $r_c(g) < 0$. As a special case, Theorem 6.2 yields a uniform lower bound on the log-Sobolev constant for all $r > r_c(g)$, but the proof of this statement can be simplified considerably in this case as we now outline.

Example 6.3 ([21, Example 3.1])**.** For the lattice φ^4 model on $\Lambda_L = \mathbb{Z}^d/(L\mathbb{Z})^d$ in any dimension, let $\dot{C}_t = (tA + \mathrm{id})^{-2}$. Then analogous to (6.23):

$$\forall \varphi \in \mathbb{R}^{\Lambda_{\varepsilon,L}}, t \geqslant 0 : \qquad \dot{C}_t \operatorname{Hess} V_t(\varphi)\dot{C}_t - \frac{1}{2}\ddot{C}_t \geqslant \Big(\underbrace{\frac{1}{t} - \frac{\chi(g, r+1/t)}{t^2}}_{\dot{\lambda}_t}\Big)\dot{C}_t, \tag{6.27}$$

and the following integral (and thus the inverse log-Sobolev constant) is bounded for all $r > r_c(g)$:

$$\int_0^\infty e^{-2\lambda_t}\, dt. \tag{6.28}$$

Example 6.4. Assume that the susceptility satisfies the mean-field bound: for some $D > 1/2, \delta \in [0,1]$:

$$\chi^L(g, r+1/t) \leqslant \frac{D}{\delta + 1/t}. \tag{6.29}$$

This bound holds, in particular, for the near-critical φ^4 model on $\mathbb{Z}^d$ when $d \geqslant 5$, for some $D > 1/2$ and $\delta = L^{-d} + r - r_c(g)$ $(r \geqslant r_c(g))$, see [2]. Then

$$\int_0^\infty e^{-2\lambda_t}\, dt \leqslant C(g,D)\delta^{-2D+1} \tag{6.30}$$

is bounded polynomially in $1/\delta$.

Sketch of Examples 6.3–6.4. For $t_0 > 0$ small enough such that $r+1/t_0 > 0$, the measure with coupling constants $(g, r+1/t_0)$ is log-concave and the Brascamp–Lieb inequality (3.63) implies

$$\forall s \leqslant t_0, \quad \chi^L(g, r+1/s) \leqslant \frac{1}{r+1/s} \tag{6.31}$$

and hence

$$\lambda_{t_0} = \int_0^{t_0} \left(\frac{1}{s} - \frac{\chi^L(g, r+1/s)}{s^2} \right) ds \geqslant \int_0^{t_0} \frac{r}{1+rs}\, ds = \log(1+rt_0). \tag{6.32}$$

On the other hand, for any $t > 0$, the second Griffiths inequality implies $\chi^L(g, r+1/t) \leqslant \chi^L(g,r)$ and thus

$$\lambda_t - \lambda_{t_0} \geqslant \int_{t_0}^{t} \left(\frac{1}{s} - \frac{\chi^L(g,r)}{s^2} \right) ds \geqslant \log(\frac{t}{t_0}) - \frac{\chi^L(g,r)}{t_0}. \tag{6.33}$$

Combining both bounds with t_0 such that $r + 1/t_0 > 0$, it follows that the integral (6.28) (and thus the log-Sobolev constant) is bounded below whenever $\chi^L(g,r)$ is bounded above.

Now assume that the susceptibility satisfies the mean-field bound (6.29). Then

$$\lambda_t - \lambda_{t_0} \geqslant \int_{t_0}^{t} \left(\frac{1}{s} - \frac{D}{\delta s^2 + s} \right) ds \geqslant C(D, t_0) + \log t - D \log(\frac{t}{1+\delta t}). \tag{6.34}$$

Combining this bound with $\lambda_{t_0} \geqslant \log(1+rt_0)$ for $t_0 \leqslant 1/|r|$ gives

$$\int_0^\infty e^{-2\lambda_t}\, dt \leqslant C(g,D)\delta^{1-2D} \tag{6.35}$$

as claimed. □

6.3. Applications to transport maps

Another application of the bounds on the Hessian of renormalised potential are to the transport maps of Theorems 5.4– 5.5, which can be used to recover the log-Sobolev inequalities.

Example 6.5. For the continuum sine-Gordon model (under the same assumptions as in Theorem 6.1), the transport map of Theorem 5.4 has Lipschitz constant bounded uniformly in t:

$$\forall \varphi \in X, f \in X: \qquad |\nabla \hat{S}_t(\varphi) f| \leqslant C(\beta, z, L)|f|. \tag{6.36}$$

The constant $C(\beta, z, L)$ is independent of L under the same assumptions as in Theorem 6.1.

Sketch. Using bounds on the Hessian of the renormalised potential from Theorem 6.1, this is a direct consequence of Theorem 5.4. □

Example 6.6. For the continuum φ^4 model (under the same assumptions as in Theorem 6.2) where $A = -\Delta^\varepsilon + \mathrm{id}$, the gradient of the transport map of Theorem 5.5 is uniformly bounded from H^1 to L^2 in the following sense:

$$\forall \varphi \in X, f \in X: \qquad |\nabla \hat{S}_t(\varphi) f| \leqslant C(g,r,L)|f|_A \tag{6.37}$$

where $|f|_A = |\sqrt{A} f|$ is the (discrete) Sobolev norm and $C(g,r,L)$ is independent of L if $r > r_c(g)$.

In particular, in the application (5.26), the reference Gaussian measure $\hat{\nu}_\infty$ satisfies the log-Sobolev inequality with quadratic form $Q = A^{-1}$ and the bound on the transport map (6.37) is exactly the required assumption (5.27) to recover the log-Sobolev inequality for the continuum φ^4 measure.

Sketch. Theorem 5.5 gives the following bound on the gradient of the transport map:

$$|\nabla \hat{S}_t(\varphi) f|^2 \leqslant e^{-\lambda_t} |f|^2_{1+tA}. \tag{6.38}$$

By Theorem 6.2), $\lambda_t \geqslant \log(1+t) - C(g,r)$ and therefore

$$|\nabla \hat{S}_t(\varphi) f|^2 \leqslant \frac{C(g,r)}{1+t} |f|^2_{1+tA} \leqslant C(g,r)|f|^2_A \tag{6.39}$$

where we used that $(1+tA)/(1+t) \leqslant A$ if $A \geqslant \mathrm{id}$. □

Example 6.7. For the lattice φ^4 model, for any $r > r_c(g)$, the transport map of Theorem 5.5 has Lipschitz constant bounded uniformly in L and t. Moreover, if the mean-field bound (6.29) holds then the Lipschitz constant of the transport map is of order $\delta^{-D/2}$:

$$\forall \varphi \in X, f \in X: \qquad |\nabla \hat{S}_t(\varphi) f| \leqslant C(g,D)\delta^{-D/2}|f|. \tag{6.40}$$

Sketch. Using that the operator norm $\|A\|$ is bounded, Theorem 5.5 gives the following bound on the Lipschitz constant of the transport map:

$$\sqrt{t\|A\|+1}\, e^{-\frac{1}{2}\lambda_t} \leqslant C\sqrt{1+t}\, e^{-\frac{1}{2}\lambda_t}. \tag{6.41}$$

On the other hand, by (6.32) and (6.33), for $r > r_c(g)$, we have

$$e^{-\frac{1}{2}\lambda_t} \leqslant C(g,r) \frac{1}{\sqrt{1+t}}. \tag{6.42}$$

Combining both bounds gives the uniform bound on the Lipschitz constant for any $r > r_c(g)$.

Substituting the bound on the Hessian obtained from the mean-field bound on the susceptibility (6.29) into the above bound on the Lipschitz constant of the transport map again yields a polynomial bound in δ. Indeed, by (6.34), the bound obtained is of order

$$\sqrt{t\|A\|+1}\,e^{-\frac{1}{2}\lambda_t} \leqslant C\sqrt{1+t}\,e^{-\frac{1}{2}\lambda_t} \leqslant C(g,D)\left(\frac{t}{1+\delta t}\right)^{D/2} \leqslant C(g,D)\delta^{-D/2} \tag{6.43}$$

which is the claimed bound. □

6.4. 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 [13, 20], while closely related results were obtained using spectral and entropic independence and stability estimates in [46, 4, 31]. Similar ideas apply to $O(n)$ models (see [13]) for which it is however not known in general that the critical point can be reached due to the lack of appropriate correlation inequalities. For concreteness and because the results are most complete in this case, we focus on the situation of Ising models.

6.4.1. Renormalised potential

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

$$\mathbb{E}_\mu[F] = \mathbb{E}_{\mu_{\beta,h}}[F] \propto \sum_{\sigma\in\{\pm1\}^\Lambda} e^{-\frac{1}{2}(\sigma,\beta A\sigma)+(h,\sigma)}F(\sigma). \tag{6.44}$$

Since $\sigma_x^2 = 1$ for each x, the measure is invariant under the change $A \to A+\alpha\,\mathrm{id}$, $\alpha \in \mathbb{R}$. Therefore, without loss of generality, we can 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 β. It is helpful to think of the Ising model as a denerate case of the φ^4 measure in which the Gaussian part has covariance $(\beta A)^{-1}$ and the potential is singular to enforce the values $\{\pm1\}$, i.e., $\int_{\mathbb{R}^\Lambda}(\cdot)e^{-V_0(\varphi)}\,d\varphi$ is replaced by $\sum_{\sigma\in\{\pm1\}^\Lambda}(\cdot)$. It turns out that the Polchinski equation still makes sense and that the renormalised potential indeed becomes smooth immediately. A natural covariance decomposition of $(\beta A)^{-1}$ is

$$\forall 0 \leqslant t \leqslant \beta: \qquad C_t = (tA + (\alpha - t)\,\mathrm{id})^{-1}, \tag{6.45}$$

for a parameter $\alpha > \beta$ 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 β can be written as

$$\mathbb{E}_\mu[F] \propto \sum_{\sigma\in\{\pm1\}^\Lambda} e^{-\frac{1}{2}(\sigma,C_\beta^{-1}\sigma)+(h,\sigma)}F(\sigma). \tag{6.46}$$

Strictly speaking, C_t is not a covariance decomposition since $C_0 = \alpha^{-1}\,\mathrm{id} \neq 0$, 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:

$$V_t(\varphi) = -\log \sum_{\sigma\in\{\pm1\}^\Lambda} e^{-\frac{1}{2}(\sigma-\varphi, C_t^{-1}(\sigma-\varphi))+(h,\sigma)}. \tag{6.47}$$

This leads to a decomposition of the Ising measure (6.44) as:

$$\forall 0 \leqslant t < \beta: \qquad \mathbb{E}_\mu[F] = \mathbb{E}_{\nu_{t,\beta}}\big[\mathbb{E}_{\mu_t^\varphi}[F]\big]. \tag{6.48}$$

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

$$\mu_t^\varphi(\sigma) = \mu_{t,C_t^{-1}\varphi+h} \propto e^{-\frac{1}{2}(\sigma, C_t^{-1}\sigma)+(C_t^{-1}\varphi+h,\sigma)}, \tag{6.49}$$

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

$$\nu_{t,\beta}(d\varphi) \propto e^{-V_t(\varphi)}\,\mathsf{P}_{C_\beta-C_t}(d\varphi) \propto \exp\left[-\frac{1}{2}(\varphi,(C_\beta-C_t)^{-1}\varphi) - V_t(\varphi)\right] d\varphi. \tag{6.50}$$

Even though $\sigma : \Lambda \to \{\pm1\}$ is discrete, the renormalised field $\varphi : \Lambda \to \mathbb{R}$ is continuous as soon as $t > 0$. Convexity-based criterions, such as the Bakry-Émery or multiscale Bakry-Émery criterions 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.

6.4.2. 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 μ be the probability measure on $\{\pm1\}$ with $\mu(+1) = p = 1-q$. The standard Dirichlet form is

$$D_\mu(F) = \frac{1}{2}\mathbb{E}_\mu(F(\sigma^x) - F(\sigma))^2 = \frac{1}{2}(F(+1) - F(-1))^2. \tag{6.51}$$

Proposition 6.8. *Let μ be the probability measure on $\{\pm1\}$ with $\mu(+1) = p = 1-q$. Then*

$$\mathrm{Ent}_\mu(F) \leqslant \frac{pq(\log p - \log q)}{p-q}(\sqrt{F(+1)} - \sqrt{F(-1)})^2. \tag{6.52}$$

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

$$\mathrm{Ent}_\mu(F) \leqslant D(\sqrt{F}) = \frac{2}{\gamma_0} D(\sqrt{F}), \qquad \gamma_0 = 2. \tag{6.53}$$

Proof. The proof can be found in [86] or [7]. □

Ising model, heat bath Dirichlet form With μ as above, the heat-bath Dirichlet form is

$$D^{\mathrm{HB}}(F) = \frac{1}{2}\sum_\sigma \frac{\mu(\sigma)\mu(\sigma^x)}{\mu(\sigma)+\mu(\sigma^x)}(F(\sigma^x)-F(\sigma))^2 = pq(F(+1)-F(-1))^2. \tag{6.54}$$

Proposition 6.9. *Let* μ *be the probability measure on* $\{\pm 1\}$ *with* $\mu(+1) = p = 1-q$*. Then*

$$\mathrm{Ent}_\mu(F) \leqslant pq(\log F(+1) - \log F(-1))(F(+1) - F(-1)). \tag{6.55}$$

Thus the modified log-Sobolev constant with respect to the heat-bath Dirichlet form is at least $1/2$*:*

$$\mathrm{Ent}_\mu F \leqslant D^{\mathrm{HB}}(\log F, F) = \frac{1}{2\gamma_0} D^{\mathrm{HB}}(\log F, F), \qquad \gamma_0 = \frac{1}{2}. \tag{6.56}$$

Proof. See [25, Example 3.8]. □

Similar single-spin inequalities are available for $O(n)$ models [96, 89] and allow to extend for example Theorem 6.11 below for the Ising model to these models with little change [13].

6.4.3. Entropy decomposition

To prove a log-Sobolev inequality (or modified log-Sobolev inequality) for the Ising measure, we start from the decomposition (6.48) with $t = 0$, decomposing $\mu = \mu_\beta^0$ into two parts: an infinite temperature Ising measure μ_0^φ with external field $C_0^{-1}\varphi + h$ and the renormalised measure $\nu_{0,\beta}$. The corresponding entropy decomposition (2.52) is:

$$\mathrm{Ent}_\mu(F) = \mathbb{E}_{\nu_{0,\beta}}[\mathrm{Ent}_{\mu_0^\varphi}(F)] + \mathrm{Ent}_{\nu_{0,\beta}}\left(\mathbb{E}_{\mu_0^\varphi}[F]\right). \tag{6.57}$$

To prove the log-Sobolev inequality (or modified log-Sobolev inequality) with respect to a Dirichlet form D_μ, 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:

$$D_\mu(F) = \frac{1}{2}\sum_{x\in\Lambda} \mathbb{E}_\mu\Big[(F(\sigma) - F(\sigma^x))^2\Big]. \tag{6.58}$$

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 6.4.2 and 6.4.5.

Under the assumption that μ_0^φ satisfies a log-Sobolev inequality with constant γ_0 uniformly in φ, the first term on the right-hand side of (6.57) is bounded by

$$\frac{2}{\gamma_0}\mathbb{E}_{\nu_{0,\beta}}[D_{\mu_0^\varphi}(\sqrt{F})] = \frac{2}{\gamma_0}D_\mu(\sqrt{F}). \tag{6.59}$$

Since μ_0^φ 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 6.4.2. The last equality relies on the specific jump rates in the standard Dirichlet form (6.58). For other Dirichlet forms, one can also get an inequality, see Section 6.4.5.

Bounding the second term on the right-hand side of (6.57) 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

$$\frac{2}{\gamma_{0,\beta}}\mathbb{E}_{\nu_{0,\beta}}\Big[\big|\big(\nabla_\varphi \mathbb{E}_{\mu_0^\varphi}[F]^{1/2}\big)\big|^2\Big] \leqslant \frac{4\alpha^2}{\gamma_0\gamma_{0,\beta}}D_\mu(\sqrt{F}), \tag{6.60}$$

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 6.10. For each $x \in \Lambda$, let $\mu_0^{\varphi,x}$ denote the law of σ_x under the product measure μ_0^φ. Then:

$$\nabla_{\varphi_x}\big(\mathbb{E}_{\mu_0^\varphi}[F]^{1/2}\big) = \frac{\alpha}{2\sqrt{\mathbb{E}_{\mu_0^\varphi}[F]}}\operatorname{Cov}_{\mu_0^\varphi}(F,\sigma_x) = \frac{\alpha}{2\sqrt{\mathbb{E}_{\mu_0^\varphi}[F]}}\mathbb{E}_{\mu_0^\varphi}\Big[\operatorname{Cov}_{\mu_0^{\varphi,x}}(F,\sigma_x)\Big] \tag{6.61}$$

(recall $C_0 = \alpha^{-1}\,\mathrm{id}$) and:

$$\operatorname{Cov}_{\mu_0^{\varphi,x}}(F,\sigma_x)^2 \leqslant 8\mathbb{E}_{\mu_0^{\varphi,x}}[F]\operatorname{Var}_{\mu_0^{\varphi,x}}(\sqrt{F}) \leqslant \frac{8}{\gamma_0}\mathbb{E}_{\mu_0^{\varphi,x}}[F]D_{\mu_0^{\varphi,x}}(\sqrt{F}). \tag{6.62}$$

Sketch. The first equation is a simple computation. The first inequality in (6.62) follows from Cauchy-Schwarz inequality using the following general expression for the covariance of functions G_1, G_2 under a measure m:

$$\operatorname{Cov}_m(G_1,G_2) = \frac{1}{2}\int \big(G_1(x)-G_1(y)\big)\big(G_2(x)-G_2(y)\big)\,dm(x)\,dm(y). \tag{6.63}$$

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 [13]. □

In summary, if μ_0^φ satisfies a log-Sobolev inequality with constant γ_0 and $\nu_{0,\beta}$ satisfies a log-Sobolev inequality with constant $\gamma_{0,\beta}$ then the inverse log-Sobolev constant γ^{-1} of μ satisfies

$$\frac{1}{\gamma} \leqslant \frac{1}{\gamma_0}\left[1+\frac{2\alpha^2}{\gamma_{0,\beta}}\right]. \tag{6.64}$$

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 6.4.4.

Recall that the renormalised measure is $\nu_{0,\beta}(d\varphi) \propto e^{-V_0(\varphi)} \mathsf{P}_{C_\beta - C_0}(d\varphi)$ with the renormalised potential V_0 defined in (6.47). 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 [13] and gives (see Exercise 6.14):

$$\gamma_{0,\beta} \geqslant \alpha - \alpha^2. \tag{6.65}$$

Taking $\alpha \downarrow \beta$, this leads to the following theorem. We recall the convention fixed below (6.44) that the coupling matrix A has spectrum in $[0,1]$. This spectral condition appeared in [13] (and unlike previously existing conditions applies to the Sherrington–Kirkpatrik spin glass). See also [46, 4, 1, 5] for recent results on spin glasses.

Theorem 6.11 (Spectral high temperature condition [13])**.** *For $\beta < 1$ (under the conventions stated below (6.44)), the Ising model μ satisfies the log-Sobolev inequality: for each $F : \{-1,1\}^\Lambda \to \mathbb{R}_+$,*

$$\mathrm{Ent}_\mu(F) \leqslant \Big(1+\frac{2\beta}{1-\beta}\Big) D_\mu(\sqrt{F}). \tag{6.66}$$

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:

$$\frac{1}{\gamma_{0,\beta}} \leqslant |\dot C_0| \int_0^\beta e^{-2\lambda_t\, dt} = \frac{1}{\alpha^2}\int_0^\beta e^{-2\lambda_t}\, dt, \qquad \lambda_t = \int_0^t \dot\lambda_s\, ds, \tag{6.67}$$

and substituting this into (6.64) gives the following log-Sobolev inequality:

$$\mathrm{Ent}_\mu(F) \leqslant \frac{2}{\gamma} D_\mu(\sqrt{F}), \qquad \frac{1}{\gamma} \leqslant \frac{1}{\gamma_0}\left[1 + 2\int_0^\beta e^{-2\lambda_t} dt\right], \qquad \lambda_t = \int_0^t \dot\lambda_s\, ds. \tag{6.68}$$

In the same situation, instead of using the multiscale Bakry–Émery criterion to bound $\gamma_{0,\beta}$ in (6.64), one can prove a log-Sobolev inequality for μ by using the entropic stability estimate discussed in Section 3.7. This was done in [31] (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 (6.45), 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 6.15. 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:

$$\mathrm{Ent}_\mu(F) \leqslant e^{-\lambda_\beta}\mathbb{E}_{\nu_{0,\beta}}[\mathrm{Ent}_{\mu_0^\varphi}(F)], \qquad \lambda_t = \int_0^t \dot\lambda_s\, ds, \tag{6.69}$$

and with the uniform log-Sobolev inequality for μ_0^φ this gives

$$\mathrm{Ent}_\mu(F) \leqslant \frac{2}{\gamma} D_\mu(\sqrt{F}), \qquad \frac{1}{\gamma} \leqslant \frac{1}{\gamma_0} e^{-\lambda_\beta}, \qquad \lambda_t = \int_0^t \dot\lambda_s\, ds. \tag{6.70}$$

To be precise, for the Ising model where $C_0 \neq 0$, the estimate (3.109) holds with the left-hand side there replaced by $\mathrm{Ent}_\mu(F)$ instead of $\mathrm{Ent}_{\nu_0}(F)$. To see this, replace $\boldsymbol{P}_{0,t}$ by $\boldsymbol{P}_t := \mathbb{E}_{\mu_t^\varphi}[\cdot] = \boldsymbol{P}_{0,t}\mathbb{E}_{\mu_0^\varphi}[\cdot]$ in the argument leading to (3.130). Then (3.130) continues to hold and gives the claim.

Estimate (6.70) is very similar to the estimate (6.68) 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 $d \geqslant 5$, see Theorem 6.17 below. The strategies of the two proofs have different advantages. We summarise the results as follows.

Theorem 6.12 (Covariance conditions for Ising models [20, 31])**.** *The log-Sobolev constant of the Ising model at inverse temperature β is bounded by* (6.68) *or* (6.70)*.*

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 6.4.5. As pointed out in [46], other choices are of interest when the jump rates are unbounded.

6.4.4. 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 follow from 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 6.13. Show that

$$\operatorname{Hess} V_t(\varphi) = C_t^{-1} - C_t^{-1}\Sigma_t(C_t^{-1}\varphi + h)C_t^{-1}, \tag{6.71}$$

where $\Sigma_t(g) = (\mathrm{Cov}_{\mu_{t,g}}(\sigma_x, \sigma_y))_{x,y\in\Lambda}$ is the covariance matrix of the Ising model $\mu_{t,g}$ at inverse temperature t and site-dependent magnetic field g (so that (6.49) 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 (6.65) and conclude the proof of Theorem 6.11.

Exercise 6.14. Let $1 \geqslant \alpha > \beta$, and set $C_t = (tA + (\alpha - t)\,\mathrm{id})^{-1}$. Then V_t is convex for all $t \in [0, \alpha]$.

Proof. Since $\mu_{0,g}$ is a product measure and $|\sigma_x| \leqslant 1$,

$$\Sigma_0(g) = \mathrm{diag}(\mathrm{Var}_{\mu_{0,g}}(\sigma_x))_{x\in\Lambda} \leqslant \mathrm{id}. \tag{6.72}$$

Using that $C_0 = \alpha^{-1}\,\mathrm{id}$, we deduce from (6.71) that

$$\operatorname{Hess} V_0(\varphi) = \alpha\,\mathrm{id} - \alpha^2\Sigma_0(\alpha\varphi + h) \geqslant (\alpha - \alpha^2)\,\mathrm{id}. \tag{6.73}$$

Thus if $\alpha \leqslant 1$, it follows that V_0 is convex. By Proposition 3.13, V_t is convex for all $t > 0$. □

For general $\beta > 0$, semi-convexity criteria on the Hessian can equivalently be formulated as covariance estimates that hold uniformly in an external field.

Exercise 6.15. From $C_t = (tA + (\alpha - t)\,\mathrm{id})^{-1}$, one has $\dot{C}_t = (\mathrm{id} - A)C_t^2$, $\ddot{C}_t = 2(\mathrm{id} - A)\dot{C}_tC_t$. The multiscale Bakry–Émery criterion (3.32) and the entropic stability criterion (3.107) thus hold with

$$-\dot{\lambda}_t = \alpha_t = \bar{\chi}_t \tag{6.74}$$

where

$$\bar{\chi}_t = \sup_{g\in\mathbb{R}^\Lambda} \chi_t(g), \qquad \chi_t(g) = \|\Sigma_t(g)\| \tag{6.75}$$

is a uniform upper bound on the spectral radius of the covariance matrix $\Sigma_t(g)$ of an Ising model uniformly in an external field g.

Now a significant simplification occurs for Ising models with ferromagnetic interaction, meaning $A_{xy} \leqslant 0$ for $x \neq y$. This includes the case of the lattice Laplacian Δ acting on configurations according to $(\Delta\sigma)_x = \sum_{y\sim x}[\sigma_y - \sigma_x]$. For ferromagnetic interactions, it turns out that the spectral radius of the covariance matrix is maximal at 0 field:

$$\bar{\chi}_t = \chi_t(0). \tag{6.76}$$

This is a consequence of the FKG inequality (which implies that the covariance matrix has pointwise nonnegative coefficients), the Perron–Frobenious theorem (which therefore implies that the largest eigenvector has nonnegative entries), and the following remarkable correlation inequality due to Ding–Song–Sun [42] which implies that the covariance between any two spins is maximised at 0 field.

Proposition 6.16 (Ding–Song–Sun inequality [42, Corollary 1.3])**.** *Let $\mu = \mu_{\beta,h}$ be the Ising measure* (6.44) *with ferromagnetic interaction A and external field $h \in [-\infty,\infty]^\Lambda$, with values $\pm\infty$ corresponding to boundary conditions. Then:*

$$\forall (x,y) \in \Lambda^2 : \qquad \mathrm{Cov}_{\mu_{\beta,h}}(\sigma_x,\sigma_y) \leqslant \mathrm{Cov}_{\mu_{\beta,0}}(\sigma_x,\sigma_y) = \mathbb{E}_{\mu_{\beta,0}}[\sigma_x\sigma_y]. \tag{6.77}$$

In particular, if the interaction A is ferromagnetic and (for simplicity) in addition translation invariant, i.e., $A_{x,y} = A_{0,x-y}$, then

$$\bar{\chi}_t = \chi_t = \chi_t(0) = \sum_{x\in\Lambda} \mathbb{E}_{\mu_{t,0}}[\sigma_0\sigma_x] \tag{6.78}$$

is the *susceptibility* of the Ising model. It characterises the phase transition of the ferromagnetic Ising model, in the sense that, e.g., for $\Lambda \subset \mathbb{Z}^d$, the critical value β_c of β satisfies:

$$\beta_c := \sup\left\{\beta > 0 : \sup_{\Lambda\uparrow\mathbb{Z}^d} \sum_{x\in\Lambda} \mathbb{E}_{\mu_{\beta,0}}[\sigma_0\sigma_x] < \infty\right\}. \tag{6.79}$$

By combining the Ding–Song–Sun correlation bound with the multiscale Bakry–Émery criterion, and in view of the above characterisation of β_c, the following log-Sobolev inequality up to the critical point for ferromagnetic Ising models on general geometries was proven in [20].

Theorem 6.17 ([20, Theorem 1.1])**.** *The log-Sobolev constant $\gamma_{\beta,h}$ of the Ising measure* (6.44) *with ferromagnetic interaction satisfies:*

$$\frac{1}{\gamma_{\beta,h}} \leqslant \frac{1}{2} + \int_0^\beta e^{2\int_0^t \chi_s\, ds}\, dt, \tag{6.80}$$

with $\chi_\beta = \sup_x \sum_y \mathbb{E}_{\mu_{\beta,0}}[\sigma_x\sigma_y]$ or more generally equal to 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 [42] (which establishes the strong spatial mixing property).

As seen in Example 6.4 for the lattice φ^4 models, (6.80) 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 (6.80) 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 the constant D in Example 6.18 is equal to 1.

Example 6.18. If $\chi_\beta \leqslant D/(\beta_c - \beta)$ then $\gamma_{\beta,h}$ is bounded polynomially in $\beta_c - \beta$, and if $\chi_\beta \leqslant D/(\beta_c - \beta + L^{-\alpha})$ 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 $d \geqslant 5$ [2].

Similarly, one can recover the main result of [77] from (6.80), which improves the high temperature condition of (6.66) 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 (6.66) cannot be improved for general non-ferromagnetic interactions with bounded spectral radius, but for ferromagnetic models the condition established in [77] is $\beta < \operatorname{artanh}(1/(d-1))$, whereas the condition of (6.66) 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 [77]. This application is summarised in the next example.

Example 6.19. 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 [77, Section 5.4, Eq. (39)], we deduce that for any sites x, y in the graph:

$$\mathbb{E}_{\mu_{\beta,0}}[\sigma_x\sigma_y] \leqslant \sum_{w\in\mathcal{S}(y)} \big(\tanh\beta\big)^{\mathrm{dist}(x,w)}, \tag{6.81}$$

where the bound is obtained by comparing the graph with a d-regular tree as in [93] 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

$$\chi_\beta = \sup_x \sum_y \mathbb{E}_{\mu_{\beta,0}}[\sigma_x\sigma_y] \leqslant 1 + C_d \sum_{\ell\geqslant 1} \big((d-1)\tanh\beta\big)^{\ell} \leqslant \frac{C_d}{1-(d-1)\tanh\beta}, \tag{6.82}$$

i.e., one finds a divergence of the susceptibility as in Example 6.18 as $\beta \uparrow \beta_u = \operatorname{artanh}(1/(d-1))$. Moreover, when β approaches β_u, one can use that for a graph with N sites then trivially $\chi_\beta \leqslant N$, so that

$$\chi_\beta \leqslant \frac{C_d}{1-(d-1)\tanh\beta + N^{-1}}. \tag{6.83}$$

By Theorem 6.17, we deduce a polynomial lower bound for the log-Sobolev constant in the size of the graph at $\beta = \beta_u = \operatorname{artanh}(1/(d-1))$. □

6.4.5. Choice of Dirichlet form

As in the original references [13, 20], the above discussion is formulated in terms of the standard Dirichlet form (6.58). 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 [86]. Namely, if c_1, c_2 are families of jump rates reversible with respect to the same measure ν on $\{-1,1\}^\Lambda$ and γ_1, γ_2 denote the associated log-Sobolev constants, then, for $K > 0$:

$$\forall \sigma, \sigma' \in \{-1,1\}^\Lambda : \qquad K^{-1} c_2(\sigma,\sigma') \leqslant c_1(\sigma,\sigma') \leqslant K c_2(\sigma,\sigma') \\ \Rightarrow \quad K\gamma_2 \leqslant \gamma_1 \leqslant \gamma_2/K. \quad (6.84)$$

One can for instance check that the heat-bath- and canonical jump rates associated with an Ising measure with interaction βA and external field $h \in \mathbb{R}^\Lambda$ satisfy such a bound, with a constant K that depends only on $\|h\|_\infty$ and $\beta \max_i \sum_j |A_{ij}|$. The heat-bath Dirichlet form is [72, (3.4)–(3.6)]:

$$D_\mu^{\mathrm{HB}}(F) = \frac{1}{2} \sum_{\sigma \in \{\pm 1\}^\Lambda} \sum_{x \in \Lambda} \Psi(\mu(\sigma), \mu(\sigma^x))(F(\sigma) - F(\sigma^x))^2, \qquad \Psi(a,b) = \frac{ab}{a+b}. \quad (6.85)$$

There are however situations in which the comparison argument (6.84) is not applicable. This observation was made in [46] in the situation of the SK model treated in [13], which is an Ising model with random couplings which can take arbitrarily large values. This makes the constant K in (6.84) 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 [46, 4, 31].

One can however check that the arguments of [13, 20] sketched above are not specific to the standard Dirichlet form. It is in fact straightforward to apply them to other dynamics as explained 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 (6.85), for which both points are satisfied in view of Proposition 6.9 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 6.8, one can instead apply the single-spin modified LSI with the heat-bath Dirichlet form of Proposition 6.9 since it also holds uniformly in the external field. For instance, for $\mathbb{E}_{\nu_{0,\beta}}[\mathrm{Ent}_{\mu_0^\varphi}(F)]$ the bound becomes:

$$\begin{aligned} &\mathbb{E}_{\nu_{0,\beta}}\left[\mathrm{Ent}_{\mu_0^\varphi}(F(\sigma))\right] \\ &\leqslant \frac{1}{2} \sum_\sigma \sum_x \mathbb{E}_{\nu_{0,\beta}}\left[\Psi(\mu_0^\varphi(\sigma), \mu_0^\varphi(\sigma^x))\right](F(\sigma) - F(\sigma^x))(\log F(\sigma) - \log F(\sigma^x)) \\ &\leqslant \frac{1}{2} \sum_\sigma \sum_x \Psi(\mu(\sigma), \mu(\sigma^x))(F(\sigma) - F(\sigma^x))(\log F(\sigma) - \log F(\sigma^x)) \\ &= D_\mu^{\mathrm{HB}}(F, \log F), \end{aligned} \quad (6.86)$$

where the first inequality is the single-spin modified log-Sobolev inequality from Proposition 6.9, and the second inequality is Jensen's inequality, using $\mu(\sigma) = \mathbb{E}_{\nu_{0,\beta}}[\mu_0^\varphi(\sigma)]$ and that Ψ is concave. The bound for the other term in (6.57) works analogously and gives $4D_\mu^{\mathrm{HB}}(\sqrt{F}) \leqslant D_\mu^{\mathrm{HB}}(F, \log F)$ 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):

$$\mathrm{Ent}_\mu(F) \leqslant \frac{1}{2\gamma_0^{\mathrm{HB}}} D_\mu^{\mathrm{HB}}(F, \log F), \qquad \frac{1}{\gamma_0^{\mathrm{HB}}} \leqslant 2 + 4\int_0^\beta e^{-2\lambda_t}\, dt. \tag{6.87}$$

This strategy can be further generalised to Dirichlet forms which are not concave in the measure, see [18].

6.5. 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:

$$\mu_{N,m}(d\varphi) \propto e^{-V_0(\varphi)}\, d\varphi|_{X_{N,m}}, \tag{6.88}$$

with $X_{N,m}$ the hyperplane of spins with magnetisation $m \in \mathbb{R}$:

$$X_{N,m} := \left\{ \varphi \in \mathbb{R}^N : \sum_i \varphi_i = Nm \right\}. \tag{6.89}$$

The infinite temperature case is $V_0 = \sum_{i=1}^N V(\varphi_i)$ and $V : \mathbb{R} \to \mathbb{R}$ a C^2 potential, assumed to be strictly convex outside of a segment for definiteness. The associated conservative dynamics reads:

$$d\varphi_t = -\nabla V_0(\varphi_t)\, dt + \sqrt{2}\, dB_t, \tag{6.90}$$

where (B_t) is a standard Brownian motion on $X_{N,0}$.

See [18] and the forthcoming work [17] for results in discrete and continuous spin settings, and [15] for such results for the continuum sine-Gordon model.

Appendix 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:

$$\varphi_0 \in \mathrm{argmin}_\varphi\, S(\varphi), \qquad S(\varphi) = \frac{1}{2}|\varphi|^2 + V(\varphi). \tag{A.1}$$

These can be related to the Hamilton–Jacobi equation

$$\frac{\partial V_t}{\partial t} = -\frac{1}{2}(\nabla V_t)^2, \qquad V_0(\varphi) = V(\varphi). \tag{A.2}$$

Indeed, its unique viscosity solution is given by the Hopf–Lax formula [47, Section 3.3.2]:

$$V_t(\varphi) = \min_{\zeta}\left(V_0(\zeta) + \frac{t}{2}|\frac{\varphi-\zeta}{t}|^2\right). \tag{A.3}$$

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 [47, 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 [3] and in particular [78, 43] 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. 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

$$S_t(\varphi) = \frac{|\varphi|^2}{2(1-t)} + V_t(\varphi). \tag{A.4}$$

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 $\varphi_t \to 0$ as $t \to 1$, that V is smooth along $(\varphi_t)_{t\in(0,1)}$, and that φ_t is an isolated* *local minimum of S_t for each $t \in [0,1)$. Then for $t \in [0,1)$,*

$$\varphi_t = -\int_t^1 \nabla V_u(\varphi_u)\, du. \tag{A.5}$$

Different from the situation in Corollary 4.2, the choice of φ is not necessarily unique because S may have multiple minimisers and V need not be globally smooth. Indeed, the equation (A.5) for φ is nothing but the equation for the characteristics of Hamilton–Jacobi equation (A.2), see [47, Section 3.3]. These are the curves (φ_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),

$$\frac{\partial U_t}{\partial t} = -(\nabla U_t, U_t), \tag{A.6}$$

provided U is smooth at (t,φ), then

$$\frac{\partial}{\partial t}U_t(\varphi_t) = -(\nabla U_t(\varphi_t), U_t(\varphi_t)) + (\nabla U_t(\varphi_t), \dot\varphi_t) = -(U_t(\varphi_t) - \dot\varphi_t, \nabla U_t(\varphi_t)) = 0. \tag{A.7}$$

Sketch of Proposition A.1. Since $\varphi_t \to 0$ as $t \to 1$, the claim is equivalent to proving that

$$\dot\varphi_t = \nabla V_t(\varphi_t). \tag{A.8}$$

Since φ_t minimises S_t, it satisfies the Euler–Lagrange equation

$$\frac{\varphi_t}{1-t} + \nabla V_t(\varphi_t) = \nabla S_t(\varphi_t) = 0. \tag{A.9}$$

Differentiating this equation in t,

$$\frac{\dot\varphi_t}{1-t} + \frac{\varphi_t}{(1-t)^2} + (\partial_t \nabla V_t)(\varphi_t) + \operatorname{Hess} V_t(\varphi_t)\dot\varphi_t = 0. \tag{A.10}$$

Using the Euler–Lagrange equation $\varphi_t/(1-t) = -\nabla V_t(\varphi_t)$ again and $\partial_t \nabla V_t = -\operatorname{Hess} V_t \nabla V_t$ (which is (A.6)), we obtain

$$\frac{\dot\varphi_t - \nabla V_t(\varphi_t)}{1-t} - \operatorname{Hess} V_t(\varphi_t)\nabla V_t(\varphi_t) + \operatorname{Hess} V_t(\varphi_t)\dot\varphi_t = 0. \tag{A.11}$$

Therefore

$$\operatorname{Hess} S_t(\varphi_t)[-\nabla V_t(\varphi_t) + \dot\varphi_t] = \Big(\frac{1}{1-t} + \operatorname{Hess} V_t(\varphi_t)\Big)[-\nabla V_t(\varphi_t) + \dot\varphi_t] = 0. \tag{A.12}$$

Since φ_t is an isolated local minimum of S_t, the Hessian on the left-hand side is strictly positive definite so that necessarily $\nabla V_t(\varphi_t) - \dot\varphi_t = 0$. □

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 φ^U solves the classical analogue of (5.2): $\varphi^U_t \to 0$ as $t \to 1$ and

$$\varphi^U_t = -\int_t^1 U_u(\varphi^U_u)\,du, \tag{A.13}$$

with φ^U_0 an absolute minimum of S (which we recall is the classical analogue of demanding that φ^U_0 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 φ^U that satisfies (A.13),*

$$\frac{1}{2}|\varphi^U_0|^2 \leqslant \frac{1}{2}\int_0^1 |U_u(\varphi^U_u)|^2\,du, \tag{A.14}$$

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,

$$V_1(0) - V_0(\varphi^U_0) \leqslant V_0(\varphi^U_0) + \frac{1}{2}|\varphi^U_0|^2 - V_0(\varphi^U_0) = \frac{1}{2}|\varphi^U_0|^2, \tag{A.15}$$

with equality if and only if φ^U_0 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

$$V_1(0) - V_0(\varphi^U_0) = \int_0^1 dt\,\frac{\partial}{\partial t}V_t(\varphi^U_t), \tag{A.16}$$

with φ^U evolving according to (A.13), and therefore (whenever the derivative exists classically)

$$\begin{aligned}\frac{\partial}{\partial t}V_t(\varphi^U_t) &= (\frac{\partial}{\partial t}V_t)(\varphi^U_t) + \big(\nabla V_t(\varphi^U_t), U_t(\varphi^U_t)\big)\\ &= -\frac{1}{2}(\nabla V_t(\varphi^U_t))^2 + \big(\nabla V_t(\varphi^U_t), U_t(\varphi^U_t)\big)\\ &= -\frac{1}{2}\big(\nabla V_t(\varphi^U_t) - U_t(\varphi^U_t)\big)^2 + \frac{1}{2}(U_t(\varphi^U_t))^2,\end{aligned} \tag{A.17}$$

where we used the Hamilton–Jacobi equation on the second line. In particular, if φ^U_0 is an absolute minimum of S,

$$\frac{1}{2}\int_0^1 (\nabla U_t(\varphi^U_t))^2\,dt = \frac{1}{2}|\varphi^U_0|^2 + \frac{1}{2}\int_0^1 \big(\nabla V_t(\varphi^U_t) - U_t(\varphi^U_t)\big)^2\,dt \geqslant \frac{1}{2}|\varphi^U_0|^2, \tag{A.18}$$

and the gradient of the renormalised potential V_t provides the optimal drift. □

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 $t \leqslant 1$ and $\varphi_t^U = \varphi$,

$$\varphi_s^U = \varphi - \int_s^t U_u(\varphi_u^U)\,du, \qquad (s \leqslant t). \tag{A.19}$$

Proposition A.3. *Under certain regularity conditions on V_0, and for $t \leqslant 1$ then*

$$V_t(\varphi) = \inf_U \left[V_0\Big(\varphi - \int_0^t U_s(\varphi_s^U)\,ds\Big) + \frac{1}{2}\int_0^t |U_s(\varphi_s^U)|^2\,ds \right]. \tag{A.20}$$

This is discussed in [51, 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 [51, Theorem 10.1]. This minimiser does not have to be unique if the Hamilton–Jacobi equation has a shock, see [51, Theorem 9.1 and 10.2].

Sketch. The Hopf–Lax formula implies that

$$V_t(\varphi) = \min_\zeta \left(\frac{t}{2}\Big|\frac{\varphi-\zeta}{t}\Big|^2 + V_0(\zeta) \right) \leqslant \frac{1}{2t}|\varphi - \varphi_0^U|^2 + V_0(\varphi_0^U) \tag{A.21}$$

where given any drift U_s, we let φ_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):

$$\frac{1}{2t}|\varphi - \varphi_0^U|^2 \leqslant \frac{1}{2}\int_0^t |U_s(\varphi_s^U)|^2\,ds. \tag{A.22}$$

This shows

$$V_t(\varphi) \leqslant \frac{1}{2}\int_0^t |U_s(\varphi_s^U)|^2\,ds + V_0\Big(\varphi - \int_0^t U_s(\varphi_s^U)\,ds\Big). \tag{A.23}$$

On the other hand, equality if $U = \nabla V$ follows from the Hamilton–Jacobi equation as in Proposition 4.3: if φ is a solution to (A.5),

$$\frac{\partial}{\partial s}\left[V_s(\varphi_s) + \frac{1}{2}\int_s^t (\nabla V_u(\varphi_u))^2\,ds \right] = \frac{\partial V_s}{\partial s}(\varphi_s) + (\nabla V_s(\varphi_s))^2 - \frac{1}{2}(\nabla V_s(\varphi_s))^2 = 0, \tag{A.24}$$

i.e.,

$$V_t(\varphi) - V_0(\varphi_0) = \frac{1}{2}\int_0^t (\nabla V_u(\varphi_u))^2\,du. \tag{A.25}$$

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

$$\mathbb{E}_\nu[G] = \frac{1}{Z_N(\beta,\mathbf{h})} \sum_{\sigma\in\{\pm1\}^N} e^{-\frac{\beta}{4N}\sum_{i,j}(\sigma_i-\sigma_j)^2 + (\sigma,\mathbf{h})} G(\sigma), \tag{A.26}$$

where the vector $\mathbf{h} \in \mathbb{R}^N$ is a possibly site-dependent external field and $Z_N(\beta, \mathbf{h})$ is a normalisation factor. It is convenient to rewrite it as

$$\mathbb{E}_\nu[G] = \frac{1}{Z_N(\beta, \mathbf{h})} \sum_{\sigma \in \{\pm 1\}^N} e^{-\frac{\beta}{2}(\sigma, P\sigma) + (\sigma, \mathbf{h})} G(\sigma), \tag{A.27}$$

where $P = \mathrm{id} - Q$ and Q is the orthogonal projection onto constants: $Qf = \big(\frac{1}{N}\sum_i f_i\big)\mathbf{1}$ 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 $N \to \infty$ of $F_N(\beta, h)$ where

$$F_N(\beta, h) = -\frac{1}{N} \log Z_N(\beta, h\mathbf{1}). \tag{A.28}$$

(Physically more correctly, the right-hand side should have been divided by β, but it is here more convenient to omit this.) It is well-known and easy to check that

$$\frac{\partial F_N}{\partial \beta} = \frac{1}{2N}\frac{\partial^2 F_N}{\partial h^2} - \frac{1}{2}(\frac{\partial F_N}{\partial h})^2, \qquad F_N(0, h) = -\log\cosh(h), \tag{A.29}$$

and thus that the limiting free energy F is the viscosity solution of the Hamilton–Jacobi equation

$$\frac{\partial F}{\partial \beta} = -\frac{1}{2}(\frac{\partial F}{\partial h})^2, \qquad F(0, h) = -\log\cosh(h), \tag{A.30}$$

see in particular [78]. 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 [19, Chapter 1] or [90]):

$$F(\beta, h) = \min_{g \in \mathbb{R}} \left[\frac{1}{2\beta}(g - h)^2 - \log\cosh(g)\right] = \min_{\varphi \in \mathbb{R}} \left[\frac{\beta}{2}\varphi^2 - \log\cosh(\beta\varphi + h)\right]. \tag{A.31}$$

This Hamilton–Jacobi equation for the free energy can be related, as follows, to the Polchinski equation studied earlier in Section 6.4.1. Recall that $\mathrm{id} = P + Q$ with Q the orthogonal projection onto constant vectors in $\mathbb{R}^N$ and $PQ = 0$. For $\alpha > \beta$, let

$$C_t = (tP + (\alpha - t))^{-1}, \qquad \dot{C}_t = (\alpha - t)^{-2} Q, \tag{A.32}$$

where we used $PQ = 0$ to simplify $\dot{C}_t$. For $\varphi \in \mathbb{R}^N$, the renormalised potential (6.47) then is

$$V_t(\varphi) = -\log \sum_{\sigma \in \{\pm 1\}} e^{-\frac{1}{2}(\sigma - \varphi, (tP + (\alpha - t))(\sigma - \varphi))} + (\text{constant}) \tag{A.33}$$

and satisfies the Polchinski equation (for appropriate otherwise irrelevant choice of the constants):

$$\frac{\partial V_t}{\partial t} = \frac{1}{2}\Delta_{\dot{C}_t} V_t - \frac{1}{2}(\nabla V_t)^2_{\dot{C}_t} = (\alpha - t)^{-2}\left[\frac{1}{2}\Delta_Q V_t - \frac{1}{2}(\nabla V_t)^2_Q\right]. \tag{A.34}$$

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 (6.50) 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 $\tilde V(\tilde\varphi) = \frac{1}{N} V(\tilde\varphi \mathbf{1})$. 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

$$\begin{aligned}\frac{\partial \tilde V_t}{\partial t}(\tilde\varphi) = \frac{1}{N}\frac{\partial V_t}{\partial t}(\tilde\varphi\mathbf{1}) &= (\alpha-t)^{-2}\frac{1}{N}\left[\frac{1}{2}\Delta_Q V_t - \frac{1}{2}(\nabla_Q V_t)^2\right](\tilde\varphi\mathbf{1}) \\ &= (\alpha-t)^{-2}\left[\frac{1}{2}\frac{\partial V_t}{\partial\varphi_1^2} - \frac{1}{2}(\frac{\partial V_t}{\partial\varphi_1})^2\right](\tilde\varphi\mathbf{1}) \\ &= (\alpha-t)^{-2}\left[\frac{1}{2N}\tilde V_t'' - \frac{1}{2}(\tilde V_t')^2\right](\tilde\varphi). \end{aligned} \tag{A.35}$$

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 $\overline{V_t}$ as $N \to \infty$ is the unique viscosity solution of the following (one-variable) Hamilton–Jacobi equation (now dropping tilde from the variable φ):

$$\frac{\partial \overline{V_t}}{\partial t}(\varphi) = -\frac{1}{2}(\alpha-t)^{-2}(\overline{V_t}'(\varphi))^2, \qquad \overline{V_0}(\varphi) = \frac{\alpha}{2}\varphi^2 - \log\cosh(\alpha\varphi), \qquad (\varphi\in\mathbb{R}). \tag{A.36}$$

To conclude this section, we relate the Hamilton–Jacobi equation (A.36) for the renormalised 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))

$$V_t(\varphi\mathbf{1}) = N\left[\frac{\alpha-t}{2}\varphi^2 + F_N(t,(\alpha-t)\varphi)\right]. \tag{A.37}$$

Thus $F(t,h) = -\frac{1}{2}(\alpha-t)^{-1}h^2 + \overline{V_t}((\alpha-t)^{-1}h)$ and the Hamilton–Jacobi equation (A.30) for F follows from the one of $\overline{V}$ 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

$$\frac{\partial}{\partial t}V_t = -\frac{1}{2}(\nabla V_t)^2_{\dot C_t} \qquad \Leftrightarrow \qquad \frac{\partial}{\partial t}F_t = -\frac{1}{2}(\nabla F_t)^2_{\dot\Sigma_t}, \tag{A.38}$$

and in the present example the choice of C_t corresponds to $\dot\Sigma_t = Q$.

Acknowledgments

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.

Funding

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).

References

[1] A. Adhikari, C. Brennecke, C. Xu, and H.-T. Yau. Spectral Gap Estimates for Mixed p-Spin Models at High Temperature. *Probability Theory and Related Fields*, pages 1–29, 2024. MR4771106

[2] M. Aizenman. Geometric analysis of φ^4 fields and Ising models. I, II. *Commun. Math. Phys.*, 86(1):1–48, 1982. MR0678000

[3] 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. MR2310196

[4] 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. *CoRR*, abs/2106.04105, 2021. Preprint, arXiv:2106.04105. MR4490089

[5] N. Anari, V. Jain, F. Koehler, H.T. Pham, and T.-D. Vuong. Universality of Spectral Independence with Applications to Fast Mixing in Spin Glasses. In David P. Woodruff, editor, *Proceedings of the 2024 ACM-SIAM Symposium on Discrete Algorithms, SODA 2024, Alexandria, VA, USA, January 7-10, 2024*, pages 5029–5056. SIAM, 2024. Preprint, arXiv:2307.10466. MR4699426

[6] 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. MR4232133

[7] 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. MR1845806

[8] 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. MR0889476
[9] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham, 2014. MR3155209
[10] N. Barashkov and M. Gubinelli. A variational method for Φ^4_3. *Duke Math. J.*, 169(17):3339–3415, 2020. MR4173157
[11] N. Barashkov and M. Gubinelli. The Φ^4_3 measure via Girsanov's theorem. *Electron. J. Probab.*, 26:Paper No. 81, 29, 2021. MR4269211
[12] N. Barashkov, T.S. Gunaratnam, and M. Hofstetter. Multiscale coupling and the maximum of $\mathcal{P}(\phi)_\in$ models on the torus. *Commun. Math. Phys.*, 404(2):833–882, 2023. MR4665719
[13] 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. MR3926125
[14] R. Bauerschmidt and T. Bodineau. Spectral Gap Critical Exponent for Glauber Dynamics of Hierarchical Spin Models. *Commun. Math. Phys.*, 373(3):1167–1206, 2020. MR4061408
[15] R. Bauerschmidt and T. Bodineau. Log-Sobolev inequality for the continuum sine-Gordon model. *Comm. Pure Appl. Math.*, 74(10):2064–2113, 2021. MR4303014
[16] R. Bauerschmidt, T. Bodineau, and B. Dagallier. In preparation.
[17] R. Bauerschmidt, T. Bodineau, and B. Dagallier. In preparation.
[18] R. Bauerschmidt, T. Bodineau, and B. Dagallier. Kawasaki dynamics beyond the uniqueness threshold. 2023. Preprint, arXiv:2310.04609.
[19] R. Bauerschmidt, D.C. Brydges, and G. Slade. *Introduction to a renormalisation group method*, volume 2242 of *Lecture Notes in Mathematics*. Springer, Singapore, 2019. MR3969983
[20] R. Bauerschmidt and B. Dagallier. Log-Sobolev inequality for near critical Ising models. *Comm. Pure Appl. Math.*, 77(4):2568–2576, 2024. MR4705299
[21] R. Bauerschmidt and B. Dagallier. Log-Sobolev inequality for the φ^4_2 and φ^4_3. *Comm. Pure Appl. Math.*, 77(5):2579–2612, 2024. MR4720217
[22] R. Bauerschmidt and M. Hofstetter. Maximum and coupling of the sine-Gordon field. *Ann. Probab.*, 50(2):455–508, 2022. MR4399156
[23] R. Bauerschmidt and C. Webb. The Coleman correspondence at the free fermion point. *J. Eur. Math. Soc.*, 2023+. MR4767492
[24] S. Becker and A. Menegaki. Spectral gap in mean-field $\mathcal{O}(n)$-model. *Commun. Math. Phys.*, 380(3):1361–1400, 2020. MR4179730
[25] S.G. Bobkov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.*, 19(2):289–336, 2006. MR2283379
[26] 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.
[27] D.C. Brydges and T. Kennedy. Mayer expansions and the Hamilton-Jacobi

equation. *J. Statist. Phys.*, 48(1-2):19–49, 1987. MR0914427

[28] L.A. Caffarelli. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Commun. Math. Phys.*, 214(3):547–563, 2000. MR1800860

[29] R. Catellier and K. Chouk. Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *Ann. Probab.*, 46(5):2621–2679, 2018. MR3846835

[30] 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. MR3330820

[31] Y. Chen and R. Eldan. Localization Schemes: A Framework for Proving Mixing Bounds for Markov Chains. 2022. Preprint, arXiv:2203.04163. MR4537195

[32] S. Chewi and A.-A. Pooladian. An entropic generalization of Caffarelli's contraction theorem via covariance inequalities. *Comptes Rendus. Mathématique*, 361(G9):1471–1482, 2022. MR4683324

[33] B. Chow, P. Lu, and L. Ni. *Hamilton's Ricci flow*, volume 77 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2006. MR2274812

[34] 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. MR3947319

[35] G. Conforti. Weak semiconvexity estimates for Schrödinger potentials and logarithmic Sobolev inequality for Schrödinger bridges. *Probability Theory and Related Fields*, 2024. MR4771110

[36] G. Conforti, A. Durmus, and G. Greco. Quantitative contraction rates for Sinkhorn algorithm: beyond bounded costs and compact marginals, 2023. Preprint, arXiv:2304.04451.

[37] Jordan Cotler and Semon Rezchikov. Renormalization group flow as optimal transport. *Phys. Rev. D*, 108:025003, Jul 2023. MR4643342

[38] 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. MR3844472

[39] 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.

[40] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996. MR1410112

[41] J. Dimock and T.R. Hurd. Sine-Gordon revisited. *Ann. Henri Poincaré*, 1(3):499–541, 2000. MR1777310

[42] 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. MR4586225

[43] T. Dominguez and J.C. Mourrat. Statistical mechanics of mean-field disordered systems: a hamilton-jacobi approach. arXiv:2311.08976, 2023.

MR4758104

[44] P. Duch. Flow equation approach to singular stochastic PDEs, 2022. Preprint, arXiv:2109.11380.

[45] R. Eldan. Analysis of high-dimensional distributions using pathwise methods. *to appear in the 2022 ICM Proceedings*, 2021. MR4680402

[46] 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. MR4408509

[47] L.C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998. MR1625845

[48] 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. MR4098037

[49] M. Fathi, D. Mikulincer, and Y. Shenfeld. Transportation onto log-lipschitz perturbations. *Calculus of Variations and Partial Differential Equations*, 63(3):61, 2024. MR4707029

[50] P. Federbush. Partially alternate derivation of a result of Nelson. *Journal of Mathematical Physics*, 10(1):50–52, 1969.

[51] 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. MR2179357

[52] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems*. Cambridge University Press, Cambridge, 2018. A concrete mathematical introduction. MR3752129

[53] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975. MR0420249

[54] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015. MR3406823

[55] 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. MR1971582

[56] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014. MR3274562

[57] M. Hairer. Regularity structures and the dynamical Φ^4_3 model. In *Current developments in mathematics 2014*, pages 1–49. Int. Press, Somerville, MA, 2016. MR3468250

[58] M. Hairer and K. Matetski. Discretisations of rough stochastic PDEs. *Ann. Probab.*, 46(3):1651–1709, 2018. MR3785597

[59] R.S. Hamilton. Three-manifolds with positive Ricci curvature. *J. Differential Geometry*, 17(2):255–306, 1982. MR0664497

[60] R. Haslhofer and A. Naber. Characterizations of the Ricci flow. *J. Eur. Math. Soc. (JEMS)*, 20(5):1269–1302, 2018. MR3790068

[61] Y.-H. Kim and E. Milman. A generalization of Caffarelli's contraction theorem via (reverse) heat flow. *Math. Ann.*, 354(3):827–862, 2012. MR2983070

[62] B. Klartag and E. Putterman. Spectral monotonicity under Gaussian convolution. 2021. Preprint, arXiv:2107.09496. MR4748461

[63] 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. MR3869036

[64] 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. MR4311115

[65] A. Kupiainen. Renormalization group and stochastic PDEs. *Ann. Henri Poincaré*, 17(3):497–535, 2016. MR3459120

[66] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001. MR1849347

[67] J. Lehec. Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):885–899, 2013. MR3112438

[68] C. Léonard. *A survey of the Schrödinger problem and some of its connections with optimal transport*, volume 34, pages 1533–1574. 2014. MR3121631

[69] J. Lott. Optimal transport and Perelman's reduced volume. *Calc. Var. Partial Differential Equations*, 36(1):49–84, 2009. MR2507614

[70] E. Lubetzky and A. Sly. Cutoff for the Ising model on the lattice. *Invent. Math.*, 191(3):719–755, 2013. MR3020173

[71] E. Lubetzky and A. Sly. Information percolation and cutoff for the stochastic Ising model. *J. Amer. Math. Soc.*, 29(3):729–774, 2016. MR3486171

[72] 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. MR1746301

[73] F. Martinelli. Relaxation times of Markov chains in statistical mechanics and combinatorial structures. In *Probability on discrete structures*, pages 175–262. Berlin: Springer, 2004. MR2023653

[74] R.J. McCann and P.M. Topping. Ricci flow, entropy and optimal transportation. *Amer. J. Math.*, 132(3):711–730, 2010. MR2666905

[75] D. Mikulincer and Y. Shenfeld. On the Lipschitz properties of transportation along heat flows. *GAFA Seminar Notes*, 2022. to appear. MR4651223

[76] Dan Mikulincer and Yair Shenfeld. The Brownian transport map. *Probability Theory and Related Fields*, 2024. Preprint, arXiv:2111.11521.

[77] E. Mossel and A. Sly. Exact thresholds for Ising-Gibbs samplers on general graphs. *Ann. Probab.*, 41(1):294–328, 2013. MR3059200

[78] J.-C. Mourrat. Hamilton-Jacobi equations for mean-field disordered systems. *Ann. H. Lebesgue*, 4:453–484, 2021. MR4275243

[79] J. Neeman. Lipschitz changes of variables via heat flow. 2022. Preprint, arXiv:2201.03403.

[80] 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. MR0210416

[81] F. Nicolò, J. Renn, and A. Steinmann. On the massive sine-Gordon equa-

tion in all regions of collapse. *Commun. Math. Phys.*, 105(2):291–326, 1986. MR0849210

[82] G. Perelman. The entropy formula for the Ricci flow and its geometric applications. Preprint, arXiv:0211159.

[83] J. Polchinski. Renormalization and effective lagrangians. *Nuclear Physics B*, 231(2):269 – 295, 1984.

[84] M. Reed and B. Simon. *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich Publishers], 1978. MR0493421

[85] 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. MR2352327

[86] 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. MR1490046

[87] J. Serres. Behavior of the Poincaré constant along the Polchinski renormalization flow. 2022. Preprint, arXiv:2208.08186. MR4760552

[88] Y. Shenfeld. Exact renormalization groups and transportation of measures. *Annales Henri Poincaré*, 2022. MR4711240

[89] J. Sieber. Formulae for the derivative of the Poincaré constant of Gibbs measures. *Stochastic Process. Appl.*, 140:1–20, 2021. MR4276491

[90] B. Simon. *The statistical mechanics of lattice gases. Vol. I.* Princeton Series in Physics. Princeton University Press, 1993. MR1239893

[91] 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. MR1292280

[92] K.-T. Sturm. Super-Ricci flows for metric measure spaces. *J. Funct. Anal.*, 275(12):3504–3569, 2018. MR3864508

[93] D. Weitz. *Counting independent sets up to the tree threshold*, pages 140–149. New York, NY: ACM Press, 2006. MR2277139

[94] K.G. Wilson and J.B. Kogut. The renormalization group and the ε expansion. *Physics Reports*, 12(2):75–200, 1974.

[95] N. Yoshida. The log-Sobolev inequality for weakly coupled lattice fields. *Probab. Theory Related Fields*, 115(1):1–40, 1999. MR1715549

[96] 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. MR2854732

[97] R. Zhu and X. Zhu. Lattice approximation to the dynamical Φ^4_3 model. *Ann. Probab.*, 46(1):397–455, 2018. MR3758734