A probabilistic approach to convex ($\phi$)-entropy decay for Markov chains

Giovanni Conforti

April 3, 2021

We study the exponential dissipation of entropic functionals along the semigroup generated by a continuous time Markov chain and the associated convex Sobolev inequalities, including MLSI and Beckner inequalities. We propose a method that combines the Bakry Émery approach and coupling arguments, which we use as a probabilistic alternative to the discrete Bochner identities. In particular, the validity of the method is not limited to the perturbative setting and we establish convex entropy decay for interacting random walks beyond the high temperature/weak interaction regime. In this framework, we show that the exponential contraction of the Wasserstein distance implies MLSI. We also revisit classical examples often obtaining new inequalities and sometimes improving on the best known constants. In particular, we analyse the zero range dynamics, hardcore and Bernoulli-Laplace models and the Glauber dynamics for the Curie Weiss and Ising model.

Contents

1 Introduction 1

2 Coupling rates and convex entropy decay 4
   2.1 Setup and main assumptions 4
   2.2 Coupling rates and second derivative of the entropy 5

3 Interacting random walks 9
   3.1 A sufficient non perturbative condition 10
   3.2 Examples 17
   3.3 Contraction of the Wasserstein distance 19

4 Glauber dynamics 22
   4.1 Sufficient condition for Glauber dynamics 22
   4.2 Applications to spin systems
      4.2.1 Curie-Weiss model 24
      4.2.2 Ising model 25

5 More examples 27
   5.1 Bernoulli Laplace 27
   5.2 The hardcore model 28
   5.3 Zero-range dynamics 30

6 Appendix 34

1 Introduction

Functional inequalities are powerful tools to quantify the trend to equilibrium of Markov semigroups and have a wide range important applications to the concentration of measure phenomenon and hypercontractivity. In a seminal work [2] Bakry and Émery showed that a diffusion process on a Riemannian manifold whose generator is of the form $\mathcal{L} = \Delta + \nabla V \cdot \nabla$ satisfies the logarithmic Sobolev inequality if the pointwise bound

$$\text{Ric} + \text{Hess}V \geq \kappa > 0$$  \hspace{1cm} (1)

*Département de Mathématiques Appliquées, École Polytechnique, Route de Saclay, 91128, Palaiseau Cedex, France. giovanni.conforti@polytechnique.edu
holds. Over the last three decades many profound results have been obtained in connection with the Bakry-Émery condition [1] and optimal transport, see the monographs [31, 3]. Because of this success, considerable effort has been put into transferring the ideas and concepts of Bakry-Émery theory to the setting of continuous time Markov chains. For example, Caputo et al. [4, 5, 9, 12] developed a method based on a discrete analogous of Bochner’s identity and obtained estimates on the spectral gap and entropy dissipation estimates for a large class of non local dynamics, whereas the more general problem of defining a notion of discrete Ricci curvature has been tackled in [31, 28, 19]. In particular, the notion of entropic Ricci curvature put forward in [26, 14] has deep implications in terms of functional inequalities. Explicit lower bounds for the entropic Ricci curvature in concrete examples have been recently obtained in [15, 16] and [13]. The exponential decay of general entropic functionals, called \( \phi \)-entropies has been recently investigated in [23], whereas functional inequalities for non linear Markov chains are the object of the preprint [12]. In this article we develop a probabilistic approach to establish convex Sobolev inequalities and quantify the exponential decay of \( \phi \)-entropies for continuous time Markov chains via the Bakry-Émery method. Moreover, we apply this method on specific model examples and obtain explicit lower bounds.

**Discrete convex Sobolev inequalities** In order to introduce \( \phi \)-entropies and discrete convex Sobolev inequalities, we consider a continuous time Markov chain on a countable state space \( \Omega \), whose infinitesimal generator \( \mathcal{L} \) takes the form

\[
\mathcal{L} f(\eta) = \sum_{\sigma \in G} c(\eta, \sigma) \nabla_{\sigma} f(\eta),
\]

where \( G \) is a collection of maps \( \sigma : \Omega \to \Omega \) called moves, \( c : \Omega \times G \to \mathbb{R}_{\geq 0} \) are the transition rates and \( \nabla_{\sigma} f(\eta) \) denotes the discrete gradient \( \nabla_{\sigma} f(\eta) = f(\sigma \eta) - f(\eta) \). We will also assume that the Markov chain is reversible and denote \( m \) the reversible measure. Given a convex function \( \phi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) and a positive function \( f : \Omega \to \mathbb{R}_{\geq 0} \) the \( \phi \)-entropy \( \mathcal{H}^\phi(f|m) \) of \( f \) is defined as follows:

\[
\mathcal{H}^\phi(f|m) = \sum_{\eta \in \Omega} \phi(f(\eta)m(\eta)) - \phi\left( \sum_{\eta \in \Omega} f(\eta)m(\eta) \right).
\]

In this work we are interested in estimating the best constant \( \kappa_\phi \) such that the estimate

\[
\mathcal{H}^\phi(S_t f|m) \leq \exp(-\kappa_\phi t) \mathcal{H}^\phi(f|m)
\]

holds uniformly on \( f > 0 \) and \( t > 0 \). In the above, we denoted by \( S_t \) the Markovian semigroup generated by \( \mathcal{L} \). It is well known that [4] is equivalent to the convex Sobolev inequality

\[
\forall f > 0, \quad \kappa_\phi \mathcal{H}^\phi(f|m) \leq \mathcal{E}(\phi'(f), f),
\]

where \( \mathcal{E}(f, g) \) is the Dirichlet form

\[
\mathcal{E}(f, g) = -\sum_{\eta \in \Omega} f(\eta)(\mathcal{L}g)(\eta)m(\eta).
\]

The family of convex Sobolev inequalities is quite rich. Indeed, defining

\[
\phi_\alpha(a) = \begin{cases} \frac{1}{\alpha} (a^\alpha - a) - a + 1, & \text{if } \alpha \in (1, 2], \\ a \log a - a + 1, & \text{if } \alpha = 1. \end{cases}
\]

we get that [5] is the modified logarithmic Sobolev inequality (henceforth MLSI) for \( \alpha = 1 \),

\[
\forall f > 0, \quad \kappa_1 \mathcal{H}^{\phi_1}(f|m) \leq \mathcal{E}(f, f).
\]

For \( \alpha = 2 \) we recover the Poincaré inequality, whereas for \( \alpha \in (1, 2) \) we find the family of (discrete) Beckner inequalities [4, 5],

\[
\forall f > 0, \quad \kappa_\alpha \mathcal{H}^{\phi_\alpha}(f|m) \leq \frac{\alpha}{\alpha - 1} \mathcal{E}(f^{\alpha - 1}, f).
\]

For diffusions on a Riemannian manifold it is known [1] that [5] holds with \( \kappa_\phi = \kappa \) if the Bakry-Émery condition [1] is satisfied, \( \phi \) is convex and \( \frac{1}{\phi'} \) is concave. Our strategy for establishing [4] and [5] for Markov chains follows the original idea of [2], that is to prove the stronger convexity estimate

\[
\frac{d^2}{dt^2} \mathcal{H}^\phi(S_t f|m) \geq \kappa_\phi \mathcal{E}(\phi'(S_t f), S_t f),
\]
To see why (8) implies (5) we recall that, at least formally we have
\[ \frac{d}{dt} \mathcal{H}^\phi(S_t f | m) = -\mathcal{E}(\phi(S_t f), S_t f). \]

Therefore, we obtain from (6) and Gronwall’s lemma that \( \frac{d}{dt} \mathcal{H}^\phi(S_t f | m) \to 0 \). From this, (5) follows integrating (6) over \([t, \infty)\) provided that \( \mathcal{H}^\phi(S_t f | m) \to 0 \). In the continuous setting (6) is obtained via Bochner’s identity and pointwise comparison between the so called \( \Gamma \) and \( \tilde{\Gamma} \) operators. For Markov chains comparing first and second derivative of the entropy has proven to be quite challenging, and the picture is not fully clear yet.

Probabilistic approach to convex entropy decay In this work, we develop a method for establishing (8) based on the notion coupling rates. In order to define coupling rates, it is convenient to augment the set \( G \) with a null element \( e \) and set \( G^* = G \cup \{ e \} \).

Definition 1.1. Let \( \eta, \bar{\eta} \in \Omega \) and \( \mathcal{L} \) as in (2). We say that \( c_{\text{cpl}}(\eta, \bar{\eta}, \cdot, \cdot) : G^* \times G^* \to \mathbb{R}_{\geq 0} \) are coupling rates for \((\eta, \bar{\eta})\) if
\[
\forall \gamma \in G, \sum_{\gamma' \in G^*} c_{\text{cpl}}(\eta, \bar{\eta}, \gamma, \gamma') = c(\eta, \gamma),
\]
\[
\forall \gamma' \in G, \sum_{\gamma \in G^*} c_{\text{cpl}}(\eta, \bar{\eta}, \gamma, \gamma') = c(\bar{\eta}, \bar{\gamma}).
\]

If coupling rates are available for any pair \((\eta, \bar{\eta})\) then one can define a Markov generator \( \mathcal{L}_{\text{cpl}} \) acting on \( F : \Omega \times \Omega \to \mathbb{R} \) as follows
\[ \mathcal{L}_{\text{cpl}} F(\eta, \bar{\eta}) = \sum_{\gamma, \gamma' \in G^*} c_{\text{cpl}}(\eta, \bar{\eta}, \gamma, \gamma') \nabla_{\gamma, \gamma'} F(\eta, \bar{\eta}), \]
where \( \nabla_{\gamma, \gamma'} F(\eta, \bar{\eta}) := F(\gamma \eta, \gamma' \bar{\eta}) - F(\eta, \bar{\eta}) \). A Markov chain on \( \Omega \times \Omega \) with generator \( \mathcal{L}_{\text{cpl}} \) started at \((\eta, \bar{\eta})\) indeed realizes a coupling of a Markov chain with generator \( \mathcal{L} \) started at \( \eta \) and of a Markov chain with generator \( \mathcal{L} \) started at \( \bar{\eta} \). For diffusions on Riemannian manifolds, it is well known that the fundamental gradient estimate (3, Thm. 3.2.4) can be obtained with a coupling argument, see [36]. Concerning Markov chains, we recall that couplings are a fundamental ingredient in the notion of coarse Ricci curvature [31]. However, this notion is not known to imply neither (3) nor (4). Finally, we remark that Chen has obtained in a series of paper (see for instance [29]) bounds on the spectral gap for birth and death chains by means of coupling arguments and that in the recent preprint [20] couplings are employed to obtain MLSI for inhomogeneous zero range processes using the so called martingale method. We conclude this introductory section summarizing the main contributions of this work.

- We propose a probabilistic alternative to the discrete Bochner identities of [3], upon which a large part of the results about MLSI recently obtained in connection with the Bakry Émery method rely. The notion of “admissible function” [8, Def. 2.3] is replaced by that of coupling rates. Although there is no blackbox for producing an efficient coupling in view of obtaining (5), there are some general guidelines. In particular, as one may expect, it is often convenient to construct the coupling rates in such a way that the associated Markov chain on \( \Omega \times \Omega \) reaches as quickly as possible the diagonal \( \{ (\eta, \eta) : \eta \in \Omega \} \) and if it starts from the set \( \{ \eta, \bar{\eta} : \exists \sigma \in G \text{ s.t. } \eta = \sigma \bar{\eta} \} \), it never leaves it. Therefore, the method is quite robust and could be used to analyse a wider class of models than those studied here.

- A cornerstone result of Bakry Émery theory [2] asserts that strongly log-concave probability measures on \( \mathbb{R}^d \) satisfy the logarithmic Sobolev inequality with a positive constant. This powerful geometric criterion is non perturbative, in the sense that it is satisfied by probability measures that may be far from being product measures. On the contrary, most results for continuous time Markov chains are perturbative in spirit, ensuring positive lower bounds on the MLSI constant only if the interaction is small and \( m \) is almost a product measure. In light of these observations, it is very natural to seek for non perturbative sufficient conditions on the generator of a continuous time Markov chain on \( \mathbb{N}^d \) implying MLSI. To the best of our knowledge, such results have only been obtained for \( d = 1 \), with the exception of some two dimensional examples treated in [9]. It turns out that the use of coupling rates enables to lift the obstacles that have limited non perturbative criteria to the one dimensional setup and we shall propose at Theorem 3.1 below a sufficient non perturbative condition for MLSI and general convex Sobolev inequalities that is valid for any value of \( d \). As a corollary, we obtain that multiplying a multidimensional Poisson distribution by a density of the form \( \exp(-V) \) yields a probability measure satisfying MLSI if a local condition at the origin holds and the (discrete) Hessian of the potential \( V \) has non negative entries, see condition (45).
This creates a curious parallelism with the above mentioned result for probability measures on $\mathbb{R}^d$, where it is the non negativity of $\text{Hess } V$ as a quadratic form that plays an essential role.

- For interacting random walks, we show at Theorem 3.2 that the sufficient condition for MLSI and convex Sobolev inequalities proposed at Theorem 3.1 is equivalent to an exponential contraction estimate for the Wasserstein distance along the semigroup generated by $L$. For diffusions on a Riemannian manifold it is known that the best constant in the logarithmic Sobolev inequality is at least as good as the best constant in the exponential contraction of the Wasserstein distance, see [35] for example. This fundamental result served as an inspiration for the notion of coarse Ricci curvature [31] and it is a natural question to ask whether it admits a counterpart in the setting of continuous time Markov chains. To the best of our knowledge, this question has remained unanswered so far and Theorem 3.2 settles it when the state space is $\mathbb{N}^d$. It is reasonable to expect that the conclusions of Theorem 3.2 hold in a broader setup, for instance in that of section 4.

- The proposed method provides with a unified framework for the study of general convex Sobolev inequalities, MLSI and Beckner inequalities. The literature about convex Sobolev inequalities for Markov chains is not abundant, see [23, 6]. Therefore, in many of the examples we analyse, the lower bounds on $\kappa_\phi$ that we obtain seem to be new. Concerning MLSI, we can sometimes improve on the best known estimates for $\kappa_1$ we are aware of, see sections 4.2.1, 4.2.2 and 5.2.

Organization In section 2 we state the basic assumptions and outline the method in an abstract setup. In section 3 we present a general criterion that applies in particular to interacting random walks. Moreover, we provide an interpretation of the lower bounds in terms of Wasserstein contraction. Another criterion is given in section 4 that covers many classical spin systems. Section 5 deals with some classical models well studied in the literature: Bernoulli-Laplace models, hardcore models and zero range dynamics on the complete graph.

2 Coupling rates and convex entropy decay

In this section we state our main assumptions and give some simple but rather general results on how to use coupling rates to obtain convexity estimates for the evolution of entropic functionals.

2.1 Setup and main assumptions

Given a state space $\Omega$ that is at most countable, a finite set of moves $G$, and non negative transition rates $c(\eta, \sigma)$ we consider the formal generator (2). We make the following basic assumption.

(H0) The set $G$ is finite. $L$ is irreducible and admits an invariant probability measure $m \in \mathcal{P}(\Omega)$ that satisfies

$$\sum_{\eta \in \Omega, \sigma \in G} c(\eta, \sigma) m(\eta) < +\infty.$$ 

It is well known, see for instance [30], that under (H0) the invariant measure is unique and for any initial $\eta \in \Omega$ there exists a continuous time Markov chain $(X_t)_{t \geq 0}$ whose infinitesimal generator is $L$ and such that $X_0 = \eta$. Moreover, $(X_t)_{t \geq 0}$ is non-explosive. Following closely [9] we also assume that $m$ is reversible for $L$ and that each move admits an “inverse”.

(H1) There exists an involution

$$G \rightarrow G$$

$$\sigma \mapsto \sigma^{-1}$$

such that $\sigma^{-1}(\sigma(\eta)) = \eta$ holds whenever $m(\eta)c(\eta, \sigma) > 0$ and

$$\sum_{\eta \in \Omega, \sigma \in G} F(\eta, \sigma)c(\eta, \sigma) m(\eta) = \sum_{\eta \in \Omega, \sigma \in G} F(\sigma^{-1}(\sigma(\eta)), \sigma^{-1})c(\eta, \sigma) m(\eta)$$

holds for all bounded $F : \Omega \times G \rightarrow \mathbb{R}$.
Remark 2.1. The last years have seen a surge of interest around functional inequalities for non reversible Markov dynamics. As we have already said, our method is built around the notion of coupling rates, which is not necessarily related to reversibility. Indeed, there are several non reversible Markov chains for which one can construct contractive couplings that would be natural candidates for making the methods of the present paper work. Moreover, some of our sufficient conditions remain valid even if the underlying Markov chain is not reversible, see [33, 34] for example. However, we often make use of reversibility in the proofs in order to obtain cancellations between certain cross-terms. Therefore additional efforts are needed in order to extend the reach of the techniques we develop here to non reversible Markov chains.

Next, we shall define the functional inequalities that are the main object of interest of this paper. To avoid having to discuss the domain of \( \mathcal{L} \) and of the associated Dirichlet form, we begin by assuming that \( \Omega \) is finite. In this case the Dirichlet form \( \mathcal{E}(f,g) \) can be defined for any pair of real valued functions \( f \) and \( g \) as

\[
\mathcal{E}(f,g) := -\sum_{\eta \in \Omega} g(\eta)(\mathcal{L}f)(\eta)m(\eta). \tag{11}
\]

It is well known (see for instance [9, Eq. (2.12)]) that under (H1) we can rewrite \( \mathcal{E}(f,g) \) using (10) as follows

\[
\mathcal{E}(f,g) = \frac{1}{2} \sum_{\eta \in \Omega \ \sigma \in G} c(\eta, \sigma) \nabla_{\sigma} f(\eta) \nabla_{\sigma} g(\eta) m(\eta). \tag{12}
\]

Therefore, recalling the definition (3) of \( \phi \)-entropy \( H^\phi(.|m) \) we have that for a given convex function \( \phi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \), the convex Sobolev inequality (5) holds with constant \( \kappa_{\phi} \) if and only if

\[
\forall f > 0, \quad H^\phi(f|m) \leq \frac{\kappa_{\phi}}{2} \sum_{\eta \in \Omega \ \sigma \in G} \nabla_{\sigma}(\phi \circ f) \nabla_{\sigma} f(\eta) c(\eta, \sigma)m(\eta). \tag{13}
\]

Note that the convexity of \( \phi \) makes sure that the right hand side of (13) is well defined even when \( \Omega \) is not finite but countable and we shall use (13) as a definition of convex Sobolev inequality for countable state spaces. If \( \Omega \) is finite, the fact that (13) implies the entropy dissipation estimate (4) is an immediate consequence of

\[
\forall t > 0, \quad \frac{d}{dt} H^\phi(S_t f|m) = -\mathcal{E}(\phi'(S_t f), S_t f).
\]

When \( \Omega \) is countable, some extra care has to be used as the above relation may not be valid for all positive \( f \). For MLSI \( (\phi(a) = a \log a - a + 1) \) and under hypothesis (H0)(H1) the validity of (4) is covered by [9, Prop 2.1]. Some minor modifications of the argument therein cover the case of a general convex \( \phi \). For the sake of brevity, we do not provide details here. In this article, we seek for conditions implying the convex Sobolev inequality that can be read directly off the generator \( \mathcal{L} \). In view of (13), it is convenient to introduce the function \( \Phi \) defined by

\[
\Phi : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{>0}, \quad \Phi(a, b) := (\phi'(b) - \phi'(a))(b - a).
\]

A natural assumption for our method to work is the following

(H2) \( \phi \) is convex and the function \( \Phi \) is also convex.

When \( \phi = \phi_a \), the function \( \Phi \) is denoted \( \Phi_a \). We will show at Lemma 2.1 that \( \Phi_a \) satisfies (H2).

2.2 Coupling rates and second derivative of the entropy

The goal of this section is to show how one can use coupling rates to organize the terms originating from differentiating \( \mathcal{E}(\phi'(S_t f), S_t f) \) and find appropriate lower bounds in view of establishing (8). We begin by recording some useful properties of \( \Phi_a \) that we shall use to obtain Beckner inequalities and MLSI. From now on, for a differentiable function \( (a, b) \to \Phi(a, b) \) we denote by \( D\Phi(a, b) \) the Jacobian, i.e. the \( 1 \times 2 \) matrix \([\partial_a \Phi(a, b), \partial_b \Phi(a, b)]\). We also use the notation \( \cdot \) for the standard matrix-vector product.

Lemma 2.1. Let \( \alpha \in [1, 2] \). Then \( \phi_\alpha \) satisfies (H2). Moreover

- If \( a, b, a', b' > 0 \) are such that \( a' = b' \) we have

\[
\Phi_\alpha(a', b') - \Phi_\alpha(a, b) - D\Phi_\alpha(a, b) \cdot \begin{bmatrix} a' - a \\ b' - b \end{bmatrix} \geq (\alpha - 1)\Phi_\alpha(a, b). \tag{14}
\]
For all \(a, b > 0\) we have

\[
\Phi_1(a, a) - \Phi_1(a, b) - \mathbf{D}\Phi_1(a, b) \cdot \begin{bmatrix} 0 \\ a - b \end{bmatrix} + \Phi_1(b, b) - \Phi_1(a, b) - \mathbf{D}\Phi_1(a, b) \cdot \begin{bmatrix} b - a \\ 0 \end{bmatrix} \geq 2\Phi_1(a, b).
\] (15)

We defer the proof of this algebraic lemma to the appendix.

**Remark 2.2.** Assumption \([H2]\) is different from that of \([23]\), where it is assumed that \((a - b)/(\varphi'(a) - \varphi'(b))\) is concave. This assumption implies in particular that \(a \mapsto \frac{1}{\varphi'(a)}\) is concave, which is the classical hypothesis used for diffusions on manifolds. However, in order to go beyond the study of one dimensional birth and death processes, an extra homogeneity assumption has to be added there.

In the next lemma we observe that coupling rates can be used to organize the terms coming from \(\frac{d}{dt}\mathcal{E}(\varphi'(f_t), f_t)|_{t=0}\) and find a first general upper bound. For the next lemma and in all what follows it is convenient to define \(\mathcal{S} \subseteq \Omega \times G\) as

\[
\mathcal{S} = \{(\eta, \sigma) \in \Omega \times G : c(\eta, \sigma) > 0\}.
\]

Moreover, we recall that \(G^*\) is the set of moves augmented with the null-move \(e\), i.e. \(G^* = G \cup \{e\}\) and \(c_\eta = \eta\) for all \(\eta \in \Omega\). To streamline proofs and avoid technicalities we assume that \(\Omega\) is finite, although we believe this assumption not to be essential. Finally, to ease notation we shall write \(f_t\) instead of \(S_t f\).

**Lemma 2.2.** Assume \([H0] \text{ and } [H2]\) and let \(\{c^{\text{cpl}}(\eta, \sigma, \cdot, \cdot)\}_{(\eta, \sigma) \in S}\) be coupling rates. For all \(f > 0\) define \(f^\varphi\) as

\[
f^\varphi : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0}, \quad f^\varphi(\eta, \bar{\eta}) = \Phi(f(\eta), f(\bar{\eta})).
\]

We have:

\[
\frac{d}{dt} \mathcal{E}(\varphi'(f_t), f_t)|_{t=0} = \sum_{(\eta, \sigma) \in S} c(\eta, \sigma) c^{\text{cpl}}(\eta, \sigma, \gamma, \bar{\gamma}) \mathbf{D}\Phi(f(\eta), f(\sigma)) \cdot \left[ \nabla_{\gamma} f(\eta) \right] \mathbf{m}(\eta).
\] (16)

Consequently,

\[
\frac{d}{dt} \mathcal{E}(\varphi'(f_t), f_t)|_{t=0} \leq \sum_{(\eta, \sigma) \in S} c(\eta, \sigma) c^{\text{cpl}}(\eta, \sigma, \gamma, \bar{\gamma}) \mathbf{D}\Phi(f(\eta), f(\sigma)) \cdot \left[ \nabla_{\gamma} f(\eta) \right] \mathbf{m}(\eta) - \sum_{(\eta, \sigma) \in S} c(\eta, \sigma) c^{\text{cpl}}(\eta, \sigma, \gamma, \bar{\gamma}) \mathbf{D}\Phi(f(\eta), f(\sigma)) \cdot \left[ \nabla_{\gamma} f(\eta) \right] \mathbf{m}(\eta).
\] (17)

**Proof.** By definition of \(\Phi\) and \([12]\) we have for all \(f > 0\)

\[
2\mathcal{E}(\varphi'(f), f) = \sum_{(\eta, \sigma) \in S} c(\eta, \sigma) \Phi(f(\eta), f(\sigma)) \mathbf{m}(\eta).
\]

Therefore,

\[
\frac{d}{dt} \mathcal{E}(\varphi'(f_t), f_t)|_{t=0} = \sum_{(\eta, \sigma) \in S} c(\eta, \sigma) \Phi(f_t(\eta), f_t(\sigma)) \cdot \left[ \mathcal{L} f(\eta) \mathbf{m}(\eta) \right] \mathbf{m}(\eta).
\] (18)

From Definition \([11]\) of coupling rates we get that for \((\eta, \sigma) \in S\),

\[
\mathcal{L} f(\eta) = \sum_{\gamma \in G} c(\eta, \gamma) \nabla_{\gamma} f(\eta) = \sum_{\gamma, \bar{\gamma} \in G^*} c^{\text{cpl}}(\eta, \sigma, \gamma, \bar{\gamma}) \nabla_{\gamma} f(\eta).
\]

Rewriting \(\mathcal{L} f(\sigma \eta)\) analogously and plugging the two resulting expression back into \([18]\) we arrive at \([16]\). To derive \([17]\) we can first add and subtract

\[
\sum_{(\eta, \sigma) \in S} c(\eta, \sigma) c^{\text{cpl}}(\eta, \sigma, \gamma, \bar{\gamma}) \mathbf{D}\Phi(f(\eta), f(\sigma)) \cdot \left[ \nabla_{\gamma} f(\eta) \right] \mathbf{m}(\eta)
\]
in \([16]\). We obtain the equivalent expression
\[
\frac{d}{dt} E'(f_t, f_t) |_{t=0} = \sum_{\eta, \sigma \in S} c(\eta, \sigma) e^{cpl}(\eta, \sigma, \gamma, \tilde{\gamma}) \nabla_{\gamma, \tilde{\gamma}} f^{\phi}(\eta, \sigma, \eta) m(\eta) \\
- \sum_{\eta, \sigma \in S} c(\eta, \sigma) e^{cpl}(\eta, \sigma, \gamma, \tilde{\gamma}) \left( \nabla_{\gamma, \tilde{\gamma}} f^{\phi}(\eta, \sigma, \eta) - D\Phi(f(\eta), f(\sigma)) \cdot \left[ \frac{\nabla_{\gamma} f(\eta)}{\nabla_{\gamma} f(\sigma)} \right] \right) m(\eta).
\]
(19)

Since \(\nabla_{\gamma, \tilde{\gamma}} f^{\phi}(\eta, \sigma, \eta) = \Phi(f(\gamma), f(\tilde{\gamma} \sigma \eta)) - \Phi(f(\eta), f(\sigma \eta))\), we deduce from the convexity of \(\Phi\) that for all \((\eta, \sigma) \in S\) and all \(\gamma, \tilde{\gamma} \in G^*\) we have
\[
\nabla_{\gamma, \tilde{\gamma}} f^{\phi}(\eta, \sigma, \eta) - D\Phi(f(\eta), f(\sigma)) \cdot \left[ \frac{\nabla_{\gamma} f(\eta)}{\nabla_{\gamma} f(\sigma)} \right] \geq 0.
\]
Thus, we obtain the upper bound \([17]\) by dropping in \([19]\) all terms such that \(\gamma \eta \neq \tilde{\gamma} \sigma \eta\).

We now present a sufficient condition to obtain convex Sobolev inequalities. In there, we again assume for simplicity that \(\Omega\) is finite although this is probably not necessary. Under this assumption, the proof that \([8]\) implies \([13]\) is straightforward. At Corollary 2.1 we provide a simple sufficient condition that allows to extend the results for finite to countable state spaces via a localization procedure.

**Proposition 2.1.** Let \(\Omega\) be finite, assume \([H0, \ H1, \ H2]\) and let \(\{e^{cpl}(\eta, \sigma, \cdot, \cdot, \cdot)\}_{(\eta, \sigma) \in S}\) be coupling rates. If

- There exists \(\kappa' \geq 0\) such that
  \[
  \frac{1}{2} \sum_{(\eta, \sigma) \in S} c(\eta, \sigma) e^{cpl}(\eta, \sigma, \gamma, \tilde{\gamma}) \nabla_{\gamma, \tilde{\gamma}} f^{\phi}(\eta, \sigma, \eta) m(\eta) \leq -\kappa' E'(\phi'(f), f)
  \]
  holds uniformly on \(f > 0\)

- There exist \(\kappa'', \kappa''' \geq 0\) such that
  \[
  \inf_{(\eta, \sigma) \in S} \min\{e^{cpl}(\eta, \sigma, \sigma, v), e^{cpl}(\eta, \sigma, v, \sigma^{-1})\} \geq \kappa''
  \]
  \[
  \inf_{(\eta, \sigma) \in S} \sum_{\gamma, \tilde{\gamma} \in G^*} e^{cpl}(\eta, \sigma, \gamma, \tilde{\gamma}) \geq \kappa'''
  \]
  hold.

Then

(i) The convex Sobolev inequality \([5]\) holds with \(\kappa_{\phi} = \kappa'\) for all \(\Phi\) satisfying \([H2]\).

(ii) The modified logarithmic Sobolev inequality \([8]\) holds with \(\kappa_1 = \kappa' + 2\kappa''\).

(iii) For \(\alpha \in (1, 2]\), the discrete Beckner inequality \([7]\) holds with \(\kappa_{\alpha} = \kappa' + (\alpha - 1)\kappa'''\).

**Proof.** We begin by proving (i). Let \(f > 0\) and \(\Phi\) satisfy \([H2]\) and consider the bound \([17]\). Using the convexity of \(\Phi\) and \([20]\) we deduce that
\[
\frac{d}{dt} E'(\phi'(f_t), f_t) |_{t=0} \leq -\kappa' E'(\phi'(f), f).
\]
We have therefore established the Bakry-Émery convexity estimate \([8]\) with \(\kappa_{\phi} = \kappa'\), from which the desired conclusion follows. To prove (ii) we observe that the convexity of \(\Phi\) gives
\[
\sum_{(\eta, \sigma) \in S \atop \gamma, \tilde{\gamma} \in G^*} c(\eta, \sigma) e^{\tilde{\gamma}_i (\eta, \sigma, \gamma, \tilde{\gamma})} \left( \nabla_{\gamma, \tilde{\gamma}} f^{\phi_i (\eta, \sigma \eta)} - \Phi_1 (f(\eta), f(\sigma \eta)) \cdot \left[ \nabla_{\gamma} f(\eta) \right] \right) m(\eta) \\
\geq \sum_{(\eta, \sigma) \in S \atop \gamma, \tilde{\gamma} \in G^*} c(\eta, \sigma) e^{\tilde{\gamma}_i (\eta, \sigma, e, \sigma^{-1})} \left( \nabla_{e, \sigma^{-1}} f^{\phi_i (\eta, \sigma \eta)} - \Phi_1 (f(\eta), f(\sigma \eta)) \cdot \left[ -\nabla_{\sigma} f(\eta) \right] \right) m(\eta) \\
+ \sum_{(\eta, \sigma) \in S \atop \gamma, \tilde{\gamma} \in G^*} c(\eta, \sigma) \left( \nabla_{e, \sigma^{-1}} f^{\phi_i (\eta, \sigma \eta)} - \Phi_1 (f(\eta), f(\sigma \eta)) \right) \cdot \left[ 0 \right] m(\eta) \\
\geq c'' \sum_{(\eta, \sigma) \in S \atop \gamma, \tilde{\gamma} \in G^*} c(\eta, \sigma) \left( \nabla_{e, \sigma^{-1}} f^{\phi_i (\eta, \sigma \eta)} - \Phi_1 (f(\eta), f(\sigma \eta)) \right) \cdot \left[ 0 \right] m(\eta) \\
+ \nabla_{e, \sigma^{-1}} f^{\phi_i (\eta, \sigma \eta)} - \Phi_1 (f(\eta), f(\sigma \eta)) \cdot \left[ 0 \right] m(\eta) \\
\geq 4c'' \mathcal{E}(\phi' (f), f) m(\eta),
\]

where to obtain the last inequality we used (15) with \( a = f(\eta), b = f(\sigma \eta) \). Combining this last estimate with (20) in (17) yields (8) with \( \kappa_1 = \kappa' + 2c'' \), from which the desired conclusion follows. The proof of (iii) is analogous. Indeed, (14) with \( a = f(\eta), b = f(\sigma \eta) \) and \( a' = \gamma \eta \) gives

\[
\sum_{(\eta, \sigma) \in S \atop \gamma, \tilde{\gamma} \in G^*} c(\eta, \sigma) e^{\tilde{\gamma}_i (\eta, \sigma, \gamma, \tilde{\gamma})} \left( \nabla_{\gamma, \tilde{\gamma}} f^{\phi_{\alpha} (\eta, \sigma \eta)} - \Phi_\alpha (f(\eta), f(\sigma \eta)) \cdot \left[ \nabla_{\gamma} f(\eta) \right] \right) m(\eta) \\
\geq (\alpha - 1)c'' \sum_{(\eta, \sigma) \in S \atop \gamma, \tilde{\gamma} \in G^*} c(\eta, \sigma) \Phi_\alpha (f(\eta), f(\sigma \eta)) m(\eta) \\
= 2c'' \mathcal{E}(\phi' (f), f)
\]

Using this last estimate and (20) in (17) yields (8) with \( \kappa_\alpha = \kappa' + (\alpha - 1)c'' \), from which the desired conclusion follows. \( \square \)

It is known [9, 23] that the best constant \( \kappa_\alpha \) in (7) is such that \( \kappa_\alpha \leq \kappa_2 \) for all \( \alpha \in [1, 2] \), i.e. Beckner’s inequalities and in particular MLSI are stronger than the Poincaré inequality. When working in a continuous state space, using the fact that the operator \( \Gamma \) is a derivation, more relations between the constants \( \kappa_\alpha \) are known [3, Sec. 2.8]. In particular, a Poincaré inequality implies a Beckner inequality for \( \alpha \in [1, 2] \) with a non-optimal constant. Because of the non locality of generators, this fact that does not carry over in a straightforward way to continuous time Markov chains and no general result in this direction is known to the author. We conclude this section adapting Proposition 2.1 to countable state spaces. As we said above, we achieve this with a localization procedure. To describe it, let \( \mathcal{L} \) be a generator defined on a countable state space \( \Omega \) such that (H0) and (H1) are satisfied and let \( m(\cdot) \) be the associated reversible probability measure. Next, consider an increasing sequence of finite subsets \( (\Omega^n)_{n \in \mathbb{N}} \subseteq \Omega \) such that \( \bigcup_{n \in \mathbb{N}} \Omega^n = \Omega \). For \( \sigma \in G \) define the move \( \sigma^n : \Omega^n \to \Omega^n \) and the rates \( c(\eta, \sigma^n) \) as follows:

\[
\forall \eta \in \Omega^n, \sigma \in G, \quad \sigma^n \eta = \begin{cases} 
\sigma \eta, & \text{if } \sigma \eta \in \Omega^n \\
\eta, & \text{otherwise},
\end{cases} \quad c(\eta, \sigma^n) = c(\eta, \sigma).
\]

Finally, we consider the generator \( \mathcal{L}^n \) on \( \Omega^n \) given by

\[
\mathcal{L}^n f(\eta) = \sum_{\sigma \in G} c(\eta, \sigma^n) \nabla_{\sigma^n} f(\eta).
\]

It is not hard to see that if we denote by \( m^n \in \mathcal{P}(\Omega^n) \) the conditioning of \( m \) to \( \Omega^n \), i.e.

\[
\forall A \subseteq \Omega^n, \quad m^n (A) = \frac{1}{m(\Omega^n)} m(A),
\]

then \( m^n \) is invariant for \( \mathcal{L}^n \) and (H1) is satisfied with \( (\sigma^n)^{-1} = (\sigma^{-1})^n \).
Corollary 2.1. Assume that $\mathcal{L}$ satisfies $(\text{H0})$ and $(\text{H1})$ that $\Omega$ is countable and that $\phi$ satisfies $(\text{H2})$. Consider an increasing sequence of finite sets $(\Omega^n)_{n\in\mathbb{N}}$ such that $\Omega = \bigcup_{n\in\mathbb{N}} \Omega^n$ and define $\mathcal{L}^n$ by $(\text{23})$. Moreover, assume that there exists $\kappa > 0$ such that for all $n \in \mathbb{N}$ the convex Sobolev inequality

$$\forall f > 0, \quad \kappa \mathcal{H}(f|m^n) \leq \frac{1}{2} \sum_{\eta \in \Omega^n} c(\eta, \sigma^n) \Phi(f(\eta), f(\sigma^n)(\eta))$$

holds uniformly in $n$. Then the convex Sobolev inequality $(\text{13})$ holds with $\kappa_\phi = \kappa$.

Proof. Let $f: \Omega \to \mathbb{R}$ be compactly supported and denote $\bar{\Omega}$ the support. If $n$ is large enough so that $\Omega^n \supseteq \{\eta \in \Omega, \sigma \in G\}$ for all $n \geq \bar{n}$. Applying $(\text{25})$ to the restriction of $f$ to $\Omega^n$ and using the definition of the rates $(\text{23})$ yields for all $n \geq \bar{n}$

$$\kappa \left[ \sum_{\eta \in \Omega} \phi(f(\eta)) m(\Omega^n) - \phi \left( \sum_{\eta \in \Omega} f(\eta) m(\Omega^n) \right) \right] \leq \frac{1}{2} \sum_{\eta \in \Omega} c(\eta, \sigma) \Phi(f(\eta), f(\sigma)) m(\Omega^n).$$

Letting $n \to +\infty$ and recalling that $\nabla_{\sigma}(\phi' \circ f)\nabla_{\sigma} f(\eta) = \Phi(f(\eta), f(\sigma))$ gives $(\text{13})$. A standard approximation argument using compactly supported functions and monotone convergence allows to extend the result to a non compactly supported $f > 0$.

In the remainder of this article, we use Proposition 2.1 and Corollary 2.1 to obtain lower bounds for the best constant in $(\text{13})$ in various concrete examples.

3 Interacting random walks

A cornerstone result of Bakry-Émery theory asserts that if we multiply the standard Gaussian distribution on $\mathbb{R}^d$ by a log-concave density, then the resulting probability measure $m$ satisfies the logarithmic Sobolev inequality

$$\kappa \mathcal{H}(f|m) \leq \int |\nabla \log f|^2(x) f m(dx) \quad \forall f > 0.$$  

with a constant $\kappa$ that is at least as large as the optimal constant for the Gaussian distribution. This result is a geometric and non-perturbative sufficient condition implying the logarithmic Sobolev inequality. Indeed, log-concave probability measures are not necessarily close to product measures. On the lattice $\mathbb{N}^d$ the fundamental role played by the Gaussian distribution on $\mathbb{R}^d$ is taken up by the (multidimensional) Poisson distribution $\mu_\lambda \in \mathcal{P}(\mathbb{N}^d)$:

$$\mu_\lambda(\eta_1, \ldots, \eta_d) = \prod_{i=1}^d \exp \left( \frac{-1}{\lambda} \frac{\eta_i}{\eta_i!} \right),$$

where $1/\lambda > 0$ is the intensity parameter. In analogy with Bakry-Émery theory, one is lead to consider the following problem

- Find non perturbative conditions on $V: \mathbb{N}^d \to \mathbb{R}$ implying that $m = \exp(-V)\mu_\lambda$ satisfies the convex Sobolev inequality with a positive constant.

In the above, by non perturbative we mean that the sought conditions do not necessarily imply that $V$ has to be small. In the language of statistical mechanics, this means that we try to go beyond the high temperature/weak interaction regime. Of course, in order to give a meaning to the inequality $(\text{13})$, we need to first choose a generator $\mathcal{L}$ for which $m$ is the reversible measure. Following the classical choice made in $[\text{3, 9, 15}]$, we begin by setting $\Omega = \mathbb{N}^d, G = \{\gamma_i : i = 1, \ldots, d\}$ and recalling the standard notation $(e_i)_{i=1,\ldots,d}$ for the canonical basis of $\mathbb{N}^d$. Next, we define the formal generator $\mathcal{L}$ as

$$\mathcal{L} f(\eta) = \sum_{i=1}^d c(\eta, \gamma_i^+ \eta) \nabla^+_i f(\eta) + c(\eta, \gamma_i^- \eta) \nabla^-_i f(\eta),$$

where for all $\eta \in \mathbb{N}^d$ and $1 \leq i \leq d$

$$\gamma_i^+ \eta = \eta + e_i, \quad \gamma_i^- \eta = \eta - e_i 1_{\eta_i > 0}, \quad \nabla^+_i f(\eta) = \nabla \gamma_i^+ f(\eta),$$

and

$$c(\eta, \gamma_i^+) = \exp(-\nabla^+_i V(\eta)), \quad c(\eta, \gamma_i^-) = \lambda \eta_i.$$  

We shall give in the sequel some natural conditions on $V$ ensuring that $(\text{H0})$ and $(\text{H1})$ hold.
Contraction of the Wasserstein distance and functional inequalities Another fundamental result for diffusions on manifolds is that Wasserstein contraction at rate $\kappa$ implies the logarithmic Sobolev inequality (and others) with the same constant. To illustrate this, consider the semigroup generated by the Kolmogorov diffusion
\begin{equation}
\frac{dX^x_t}{dt} = -\nabla V(X^x_t) + \sqrt{2}dB_t,
\end{equation}

We say that the $p$-Wasserstein distance $W_p(\cdot, \cdot)$ contracts at rate $\kappa$ if for all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with finite $p$-th moment and all $t > 0$ we have
\begin{equation}
W_p(\mu, \nu_t) \leq \exp(-\kappa t)W_p(\mu, \nu)
\end{equation}
where $\mu_t = \mu S_t$ and $\nu_t = \nu S_t$ are the laws of the Kolmogorov diffusion at time $t$ started at $\mu$ and $\nu$ respectively. For diffusion processes on manifolds of the form (29) it is known that for any fixed $V$ potentials respectively. For diffusion processes on manifolds of the form (29) it is known that for any fixed $V$ potentials
\begin{equation}
\frac{d}{dt}W_2(\mathcal{L}^\alpha \mu, \mathcal{L}^\alpha \nu) \leq -\kappa W_2(\mu, \nu),
\end{equation}

As for the contraction properties of $\kappa$-convexity of $V$. Going back to the lattice $\mathbb{N}^d$, the following question arises naturally:

- Is there a relation between the contraction properties of $\mathcal{L}$ as defined in (28) and the best constant in the convex Sobolev inequality (13)?

The notion of coarse Ricci curvature [31] is based on the contraction of the Wasserstein distance. However, this notion is not known to imply MLSI. In the next two subsections we bring some answers to the questions raised here.

3.1 A sufficient non perturbative condition

The main result of this section is Theorem 3.1, which contains a sufficient condition for [13] to hold. We begin by stating its main assumptions that are general enough to include (28) as a particular case. For two given potentials $V^-, V^+ : \mathbb{N}^d \to \mathbb{R}$ we construct the formal generator $\mathcal{L}$ by
\begin{equation}
\mathcal{L}f(\eta) = \sum_{i=1}^d c(\eta, \gamma^+_i) \nabla^+_i f(\eta) + c(\eta, \gamma^-_i) \nabla^-_i f(\eta),
\end{equation}
where for all $\eta \in \mathbb{N}^d$ and $1 \leq i \leq d$
\begin{equation}
c(\eta, \gamma^+_i) = \exp(-\nabla^+_i V^+ (\eta)), \quad c(\eta, \gamma^-_i) = \exp(-\nabla^-_i V^- (\eta)).
\end{equation}

When $V^+ = V$ and $V^- (\eta) = \sum_{i=1}^d \log(\lambda) \eta_i + \log(\eta_1!)$ we recover (28). We make the following hypothesis, see Figure 1 for an explanation.

(H3.3) For all $\eta \in \Omega, 1 \leq i \leq d$ we have $\kappa^+(\eta, i) \geq 0$ and $\kappa^-(\eta, i) \geq 0$, where
\begin{equation}
\kappa^+(\eta, i) = -\nabla^+_i c(\eta, \gamma^+_i) - \max_{\gamma \neq \gamma^+_i, \gamma^-_i} \{\nabla^+_i c(\eta, \gamma), 0\},
\end{equation}
and
\begin{equation}
\kappa^-(\eta, i) = \nabla^+_i c(\eta, \gamma^-_i) - \max_{\gamma \neq \gamma^+_i, \gamma^-_i} \{-\nabla^+_i c(\eta, \gamma), 0\}.
\end{equation}

Theorem 3.1. Let $V^-, V^+ : \mathbb{N}^d \to \mathbb{R}$ be given and the generator $\mathcal{L}$ be defined by (31). Moreover, assume that (H3.3) holds. Then (H0) (H1) hold with $\mathbf{m} = \frac{1}{Z} \exp(-V^+ - V^-)$, where $Z$ is the normalization constant. If we define
\begin{equation}
\kappa = \inf_{\eta \in \mathbb{N}^d} \inf_{1 \leq i \leq d} \kappa^+(\eta, i) + \kappa^-(\eta, i)
\end{equation}
then the following holds

(i) For any $\phi$ satisfying (H2) the convex Sobolev inequality [13] holds with $\kappa_{\phi} = \kappa$. In particular MLSI holds with $\kappa_1 = \kappa$.

(ii) For $\alpha \in (1, 2]$, the Beckner inequality [7] holds with $\kappa_\alpha = \alpha \kappa$.

To prove this Theorem, we show (20) and (22) with $\kappa' = \kappa'' = \kappa$. In order to do so, we shall construct appropriate coupling rates: we refer to Figure 2 to illustrate some properties of these rates.
Comparison with existing literature  Concerning MLSI and the spectral gap, the results of Theorem 3.1 are well known for $d = 1$. Comparable results can be found e.g. in [16, 28, 8, 7]. For Beckner inequalities and $d = 1$, we refer to [23]. When $d > 1$ much less appear to be known. For MLSI and Poincaré inequalities perturbative sufficient conditions are given in [7, 9]. In [9], a non perturbative two dimensional example is also treated. Erbar et al. gave in [13, Thm 3.9] a general abstract sufficient condition implying positive lower bounds for the entropic Ricci curvature [14, 26]. It can be checked that this criterion, when applied to the setting of [22] provides a bound for $\kappa_1$ that is worse or the same as what Theorem 3.1 gives, and in some cases it may give no positive lower bounds. However, although the authors only apply their result in the weak interaction/high temperature regime, it seems that its validity extends to the non perturbative setup. In particular, it could be used to provide lower bounds for some of the examples we are going to present at section 3.2. Since the hypothesis of [13, Thm 3.9] require some work to be checked and it is not obvious how to come up with a $\mathcal{L}$ satisfying them, it could be interesting to try to use the probabilistic intuition behind coupling rates to construct a candidate $\mathcal{L}$ to which this criterion applies. We conclude this literature overview mentioning the work of Johnson [22], who obtains estimates on the spectral gap and MLSI constant in $d = 1$, as well as some estimates on the spectral gap for (but no estimates on the MLSI constant) for $d \geq 2$. In conclusion, for $d \geq 2$ the results of Theorem 3.1 about Beckner’s inequalities and general convex Sobolev inequalities seem to be new and bring some answers to the questions raised in [22, Sec. 9] about sufficient conditions for MLSI in dimension larger than one.
Figure 1: The condition $\kappa^+(\eta,i) \geq 0$ imposes that the length of the green arrow is at least as much as the total length of the red arrows. The coupling interpretation of this condition is that the random walker starting at $\eta$ can use his/her larger probability to make the $\gamma_i^+$ move in order to run after the walker starting at $\gamma_i^+ \eta$ whenever he/she tries to get at distance two from $\eta$ using the moves $\gamma_j^+ \gamma_j^-$. 
Proof of Theorem 3.1. It is straightforward to check that (H3.3) implies (H0) and (H1) with \((γ^±)^{-1} = (γ^±)\) for all \(1 \leq i \leq d\). The proof strategy consists of concatenating Proposition 2.1 with Corollary 2.1. For this reason we define \(Ω^n = \{η ∈ \mathbb{N}^d : η_i ≤ n \forall i ≤ d\}\) and consider the Markov chain on \(Ω^n\) whose generator \(L^n\) is given by (24) and (23), i.e.

\[
L^n f(η) = \sum_{i=1}^{d} c(η, γ^+_i, n) \nabla^+ f(η) + c(η, γ^-_i, n) \nabla^- f(η)
\]

where for all \(η ∈ \mathbb{N}^d\) and \(i, j ∈ \{1, \ldots, d\}\)

\[
γ^+_i, n η = η + e_i 1_{η_i ≤ n}, \quad γ^-_i η = η - e_i 1_{η_i > 0}
\]

and

\[
c(η, γ^+_i, n) = c(η, γ^-_i, n) = c(η, γ^-_i, n).
\]

For the sake of simplicity, and since there is no ambiguity, we will keep writing \(γ^±\) instead of \(γ^±, n\) and adopt the same convention for discrete gradients. Likewise, we shall write \(c^n(η, γ^±)\) instead of \(c(η, γ^±, n)\). Remark that under the current hypothesis

\[
Σ = \{(η, γ^+_i) : η ∈ Ω^n, 1 ≤ i ≤ d\} ∪ \{(η, γ^-_i) : η ∈ Ω^n, η_i > 0, 1 ≤ i ≤ d\}.
\]

For \((η, γ^+_i) ∈ Σ\) we define (see also figure 2)

\[
c^{pl}(η, γ^+_i, γ^-_i, γ, φ) = \begin{cases}
\min\{c^0(η, γ), c^0(γ^+_i, η, γ)\}, & \text{if } γ = \bar{γ} ∈ G, \\
\max\{\nabla^+ c^0(η, γ, 0)\}, & \text{if } γ = γ^+_i \text{ and } \bar{γ} ∈ G, \gamma \neq γ^+_i, γ^-_i, \\
\max\{-\nabla^+ c^0(η, γ, 0)\}, & \text{if } γ ∈ G, γ \neq γ^+_i, γ^-_i \text{ and } \bar{γ} = γ^-_i, \\
κ^+(η, i), & \text{if } \bar{γ} = γ^+_i, γ = e, \\
κ^−(η, i), & \text{if } \bar{γ} = e, γ = γ^−_i, \\
0, & \text{otherwise.}
\end{cases}
\]

Since \((η, γ^-_i) ∈ Σ \Rightarrow (γ^-_i, η, γ^-_i) ∈ Σ\) we can also define

\[
c^{pl}(η, γ^-_i, γ^-_i, γ, φ) = c^{pl}(γ^-_i, η, γ^+_i, (γ^-_i) η, γ, φ),
\]

uniformly on \(γ, \bar{γ} ∈ G^\ast\). A direct calculation using (H3.3) shows that (36) and (37) define admissible coupling rates. We now move on to prove that (20) holds with \(κ = κ\). We have to show that for all \(f > 0\)

\[
\frac{1}{2} \sum_{(γ^-_i, η) ∈ Σ} \sum_{(η, γ^-_i) ∈ G^\ast} c^0(η, γ^+_i, γ, φ) \nabla f^0(η, γ^+_i η) m^n(η)
+ \frac{1}{2} \sum_{(η, γ^-_i) ∈ Σ} \sum_{(η, γ^-_i) ∈ G^\ast} c^0(η, γ^-_i, γ, φ) \nabla f^0(η, γ^-_i η) m^n(η)
\]

is bounded above by \(-κE(φ′(f), f)\). Using (10) and the fact that \(f^0\) is symmetric in its arguments, we get that the second summand equals the first one in the above expression, which can then be rewritten as \(A + B + C + D\) with

\[
A = \sum_{η ∈ Ω^n} \sum_{γ ∈ G^\ast} c^0(η, γ^+_i) \min\{c^0(η, γ), c^0(γ^+_i, η, γ)\} \nabla f^0(η, γ^+_i η) m^n(η),
\]

\[
B = \sum_{η ∈ Ω^n} \sum_{γ ∈ G^\ast} c^0(η, γ^+_i) \max\{\nabla^+ c^0(η, γ, 0)\} \nabla f^0(η, γ^+_i η) m^n(η),
\]

\[
C = \sum_{η ∈ Ω^n} \sum_{γ ∈ G^\ast} c^0(η, γ^-_i) \max\{-\nabla^+ c^0(η, γ, 0)\} \nabla f^0(η, γ^-_i η) m^n(η),
\]

\[
D = \sum_{η ∈ Ω^n} \sum_{γ ∈ G^\ast} c^0(η, γ^-_i) \min\{\nabla^+ c^0(η, γ, 0)\} \nabla f^0(η, γ^-_i η) m^n(η).
\]

We first rewrite the term \(A\) as \(A.1 + A.2 + A.3\), where
A.1 = \sum_{\eta \in \Omega^n \atop i,j \leq d} c^n(\eta, \gamma_i^+) c^n(\gamma_i^+, \eta) \nabla_{\gamma_i^+, \gamma_i^-} f^\delta(\eta, \gamma_i^+) m^n(\eta)
+ \sum_{\eta \in \Omega^n \atop i,j \leq d} c^n(\eta, \gamma_i^+) c^n(\eta, \gamma_j^-) \nabla_{\gamma_j^- \gamma_j^+} f^\delta(\eta, \gamma_i^+) m^n(\eta).

A.2 = - \sum_{\eta \in \Omega^n \atop i,j \leq d} c^n(\eta, \gamma_i^+) \max\{\nabla_i^+ c^n(\eta, \gamma_j^-), 0\} \nabla_{\gamma_j^+ \gamma_j^-} f^\delta(\eta, \gamma_i^+) m^n(\eta),

A.3 = - \sum_{\eta \in \Omega^n \atop i,j \leq d} c^n(\eta, \gamma_i^+) \max\{-\nabla_i^+ c^n(\eta, \gamma_j^-), 0\} \nabla_{\gamma_j^- \gamma_j^+} f^\delta(\eta, \gamma_i^+) m^n(\eta).

We now consider the terms B and C and observe that B = B.1 + B.2, C = C.1 + C.2 with

\begin{align*}
B.1 &= \sum_{\eta \in \Omega^n \atop i \leq d \atop j \neq i} c^n(\eta, \gamma_i^+) \max\{\nabla_i^+ c^n(\eta, \gamma_j^-), 0\} \nabla_{\gamma_j^+ \gamma_j^-} f^\delta(\eta, \gamma_i^+) m^n(\eta), \\
B.2 &= \sum_{\eta \in \Omega^n \atop i \leq d \atop j \neq i} c^n(\eta, \gamma_i^+) \max\{-\nabla_i^+ c^n(\eta, \gamma_j^-), 0\} \nabla_{\gamma_j^- \gamma_j^+} f^\delta(\eta, \gamma_i^+) m^n(\eta).
\end{align*}

and

\begin{align*}
C.1 &= \sum_{\eta \in \Omega^n \atop j \leq d \atop j \neq i} c^n(\eta, \gamma_j^-) \max\{-\nabla_j^- c^n(\eta, \gamma_j^-), 0\} \nabla_{\gamma_j^- \gamma_j^+} f^\delta(\eta, \gamma_j^-) m^n(\eta), \\
C.2 &= \sum_{\eta \in \Omega^n \atop j \leq d \atop j \neq i} c^n(\eta, \gamma_j^-) \max\{-\nabla_j^- c^n(\eta, \gamma_j^-), 0\} \nabla_{\gamma_j^- \gamma_j^+} f^\delta(\eta, \gamma_j^-) m^n(\eta).
\end{align*}

We claim that A.1 = 0, C.1 + A.3 = B.1 + A.2 = 0, C.2 = B.2 = 0. We break the proof of these identities into three different steps. A fourth step concludes the proof of Theorem 3.1.

- **Step 1**: A.1 = 0. This is done using reversibility \([10]\) on the second summand of A.1 with

\[ F(\eta, \sigma) = 1_{G^-}(\sigma) \sum_{\eta \in \Omega^n \atop i \leq d} c^n(\eta, \gamma_i^+) \nabla_{\sigma, \gamma_i} f^\delta(\eta, \gamma_i^+) \eta, \ G^- = \{\gamma_j^- : 1 \leq j \leq d\}. \]

and observing that \([28]\) implies \(c^n(\eta, \gamma_j^+) c^n(\gamma_j^+, \eta, \gamma_i^+) = c^n(\eta, \gamma_i^+) c^n(\gamma_i^+, \eta, \gamma_j^+)\) for all \(\eta, j, i\).

- **Step 2**: C.2 = B.2 = 0. Using reversibility \([10]\) with

\[ F(\eta, \sigma) = 1_{G^+}(\sigma) \sum_{\gamma \in G^+ \atop \gamma \neq \sigma, \sigma^{-1}} \max\{\nabla_{\sigma} c^n(\eta, \gamma), 0\} \nabla_{\sigma, \gamma} f^\delta(\eta, \sigma) \eta, \ G^+ = \{\gamma_i^+ : 1 \leq i \leq d\} \]

and the symmetry of \(f\) we obtain

\[ B.2 = \sum_{\eta \in \Omega^n \atop i \leq d \atop j \neq i} c^n(\eta, \gamma_j^-) \max\{-\nabla_j^- c^n(\eta, \gamma_j^-), 0\} [f(\eta, \gamma_j^-) - f(\eta, \gamma_j^+)] m^n(\eta). \]

The desired conclusion is obtained rewriting the above expression in a symmetric way by exchanging the labels \(i\) and \(j\) and observing that \([32]\) implies

\[ c^n(\eta, \gamma_i^-) \nabla_i^- c^n(\eta, \gamma_i^-) = c^n(\eta, \gamma_j^-) \nabla_j^- c^n(\eta, \gamma_j^-). \quad (39) \]

The proof that \(C.2 = 0\) is almost identical but simpler as we do not need to invoke \([10]\). Indeed, exchanging the labels \(i\) and \(j\) we arrive at

\[ C.2 = \sum_{\eta \in \Omega^n \atop i \leq d \atop j \neq i} c^n(\eta, \gamma_i^+) \max\{-\nabla_i^+ c^n(\eta, \gamma_j^-), 0\} [f(\eta, \gamma_j^-) - f(\eta, \gamma_j^+)] m^n(\eta). \]
To conclude, we observe that (32) implies
\[ c^n(\eta, \gamma_i^+) \nabla_i^+ c^n(\eta, \gamma_j^+) = c^n(\eta, \gamma_j^+) \nabla_j^+ c^n(\eta, \gamma_i^+). \] (40)

- **Step 3**: We observe that, since \([H3.3]\) implies \( \nabla_i^+ c(\eta, \gamma_i^-) \geq 0 \):

\[ C.1 + A.3 = \sum_{\eta \in \Omega^n} c^n(\eta, \gamma_i^-) \max \{-\nabla_i^+ c^n(\eta, \gamma_j^-), 0\} \left[f(\eta, \gamma_j^- \eta) - f(\gamma_j^- \eta, \gamma_j^- \eta)\right] m^n(\eta), \]

Using reversibility \([10]\) and the symmetry of \( f^\phi \) we obtain the equivalent expression

\[ \sum_{\eta \in \Omega^n} c^n(\eta, \gamma_i^-) \max \{-\nabla_i^- c^n(\eta, \gamma_j^-), 0\} \left[f(\gamma_i^- \eta, \gamma_j^- \eta) - f(\gamma_j^- \eta, \gamma_i^- \eta)\right] m^n(\eta). \]

Arguing as in Step 2, i.e. rewriting the above expression in a symmetric way exchanging the labels \( j \) and \( i \) and using \([39]\) we get the desired result. As in the previous step, the proof that \( B.1 + A.2 = 0 \) is almost identical, the only differences being that we do not need to invoke \([10]\) and \([39]\) is replaced by \([40]\).

- **Step 4**: Combining \([H3.3]\) with the fact that \( \Phi \) vanishes on the diagonal we obtain

\[ D \leq -\kappa \sum_{\eta \in \Omega^n} c^n(\eta, \gamma_i^+) \Phi(f(\eta), f(\gamma_i^+ \eta)) m^n(\eta). \]

Using the reversibility \([10]\) one last time, it follows that the right hand side of the last expression equals

\(-\kappa E(\phi'(f), f)\). Combining this with \( A + B + C = 0 \), which we proved in the former steps, we conclude that \([20]\) holds with \( \kappa' = \kappa \). From the construction of the coupling rates we also have that \([22]\) holds with \( \kappa'' = \kappa \). Applying Proposition 2.1 for any \( n \) and eventually using Corollary 2.1 concludes the proof.
Figure 2: The rates defined at (36) couple the dynamics of two random walkers $X^n_t, X^{\gamma^+\eta}_t$ starting at $\eta, \gamma^+\eta$ respectively in such a way that, on a short time interval $[0, \varepsilon]$ only one of the following movements can be observed:

- The two walkers meet at $\eta$ (resp. $\gamma^+\eta$) with probability $\varepsilon\kappa^-(\eta, i)$ (resp. $\varepsilon\kappa^+(\eta, i)$).
- The two walkers move in parallel making the same move $\gamma$ with probability $\varepsilon\epsilon^{\text{cpl}}(\eta, \gamma^+\eta, \gamma, \gamma)$ (orange lines).
- The walker $X^{\gamma^+\eta}_t$ runs after the walker $X^n_t$ with probability $\varepsilon\epsilon^{\text{cpl}}(\eta, \gamma^+\eta, \gamma, \gamma^-)$ (purple lines).
- The walker $X^n_t$ runs after the walker $X^{\gamma^+\eta}_t$ with probability $\varepsilon\epsilon^{\text{cpl}}(\eta, \gamma^+\eta, \gamma^+, \gamma)$ (blue lines).
3.2 Examples

In this section, we present some examples where explicit lower bounds on $\kappa_\varphi$ can be obtained thanks to Theorem 3.1. In [9, Sec 3.2], the authors managed to establish MLSI for a two dimensional non perturbative example corresponding to $V^+(\eta) = h(\eta_1 + \eta_2), V^-(\eta) = \sum_{i=1}^d \log(\eta_i!)$ with $h$ convex and increasing. There, they raised the question of how to generalize this result to a dimension $d > 2$. Thanks to Theorem 3.1, we can answer this question in the next Corollary. In order to state this result, and in the remainder of this article, for a function $h: \mathbb{R} \to \mathbb{R}$ we use the notation $\nabla^+ h(m)$ to indicate the increment $h(m+1) - h(m)$. Throughout this section the potential $V^-$ is fixed to be

$$V^-(\eta) = \sum_{i=1}^d \log(\lambda \eta_i + \log(\eta_i!))$$

and we will abbreviate $V^+$ with $V$. We are therefore in the setting where $\mathcal{L}$ given by (27), (28).

**Corollary 3.1.** Let $|\eta| = \sum_{i=1}^d \eta_i$, $h: \mathbb{R} \to \mathbb{R}$ be a convex function and set $V(\eta) = \beta h(|\eta|)$. Consider the generator $\mathcal{L}$ given by (27), (28). If

$$\inf_{m \in \mathbb{N}} \lambda - (d-1)[\exp(-\beta \nabla^+ h(m)) - \exp(-\beta \nabla^+ h(m + 1))] \geq 0,$$

then the conclusion Theorem 3.1 holds with

$$\kappa = \inf_{m \in \mathbb{N}} \lambda - (d-2)[\exp(-\beta \nabla^+ h(m)) - \exp(-\beta \nabla^+ h(m + 1))].$$

In particular, if $h$ is strictly increasing and

$$\beta \geq \frac{\log(d-1) - \log(\lambda)}{h(1) - h(0)},$$

then the conclusion of Theorem 3.1 holds with

$$\kappa = \lambda - (d-2) \exp(-\beta \nabla^+ h(0)).$$

Perturbative criteria typically assert that a probability measure of the form $\mu = \exp(-\beta V)\mu$ where $\mu$ is a product measure satisfy MLSI provided $\beta$ is small enough. This is often called a weak interaction/high temperature condition. On the contrary, (43) looks for a lower bound on $\beta$ and is therefore a non perturbative condition.

**Proof.** Using the convexity of $h$ and the definition of $V$ we obtain that

$$-\nabla^+_i c(\eta, \gamma^+_i) = [\exp(-\beta \nabla^+ h(|\eta|)) - \exp(-\beta \nabla^+ h(|\eta| + 1))]$$

uniformly on $\eta, i, j$. Thus, by convexity of $h$ we obtain that $-\nabla^+_i c(\eta, \gamma^+_i) \geq 0$ and that

$$\kappa^+(\eta, i) = -\nabla^+_i c(\eta, \gamma^+_i) \geq 0.$$

For the same reason

$$\kappa^-(\eta, i) = \lambda - (d-1) \left[ \exp(-\beta \nabla^+ h(|\eta|)) - \exp(-\beta \nabla^+ h(|\eta| + 1)) \right], \quad \forall \eta \in \mathbb{N}^d, i \leq d.$$

Therefore, (41) implies (H3.3) and Theorem 3.1 holds with $\kappa$ given by (42). To prove the last statement, it suffices to observe that by convexity of $h, \kappa^-(\eta, i) \geq \lambda - (d-1) \exp(-\beta \nabla^+ h(0)).$

In the next proposition we show that, combining Theorem 3.1 with a perturbative argument we can establish MLSI with a positive constant for any $\beta > 0$ and relax both the assumption that $h$ is convex increasing and the pointwise condition in (41). The price to pay is that we loose the precise control on the constants (44).

**Proposition 3.1.** Let $h: \mathbb{R} \to \mathbb{R}$ be convex outside a finite interval and set $V(\eta) = \beta h(|\eta|)$. Consider the generator $\mathcal{L}$ given by (27), (28). Then there exists $\kappa_1 > 0$ such that MLSI (6) holds.

**Proof.** Since $h$ is convex outside a finite interval we have that $\exp(-\beta \nabla^+ h(m))$ is decreasing for $m$ large enough and $\lim_{m \to +\infty} \exp(-\beta \nabla^+ h(m))$ exists and is finite, from which it follows that $\lim_{m \to +\infty} \nabla^+ \exp(-\beta \nabla^+ h(m)) = 0$. For $\varepsilon > 0$ small enough consider $M_\varepsilon$ large enough such that $h$ is convex and $\nabla^+ \exp(-\beta \nabla^+ h(m)) \geq -\frac{\varepsilon}{M_\varepsilon}$ outside $[0, M_\varepsilon]$. Define $\tilde{h}: \mathbb{N} \to \mathbb{R}_{\geq 0}$ as follows
\[ \tilde{h}(m) = \begin{cases} h(m), & \text{if } m \geq M_\varepsilon \\ h(M_\varepsilon) - (h(M_\varepsilon + 1) - h(M_\varepsilon))(M_\varepsilon - m), & \text{if } m \leq M_\varepsilon - 1 \end{cases} \]

The function \( \tilde{h} \) satisfies (41) because

\[ \nabla^+ \exp(-\beta \nabla^+ \tilde{h}(m)) = \begin{cases} \nabla^+ \exp(-\beta \nabla^+ h(m)) - \epsilon \frac{\varepsilon}{\sigma}, & \text{if } m \geq M_\varepsilon \\ 0, & \text{if } m \leq M_\varepsilon - 1 \end{cases} \]

If we define \( \tilde{V} = \tilde{h}(|\eta|) \) an application of Corollary 3.1 yields MLSI with constant \( \kappa_1 = \lambda - \varepsilon(d - 2) \). It is easily seen that \( m = \frac{1}{\varepsilon} \exp(-V(\eta)) \mu_\lambda \) and \( \tilde{m} = \frac{1}{\varepsilon} \exp(-\tilde{V}(\eta)) \mu_\lambda \) are equivalent probability measures, i.e., \( 1/K \leq \frac{\tilde{m}}{m} \leq K \) for some finite \( K \). A standard perturbative argument (see e.g. [21]) gives the desired conclusion.

For diffusions on \( \mathbb{R}^d \) [20], it is a lower bound on the spectrum of the Hessian of \( V \) viewed as a quadratic form that implies LSI. In the next Corollary, we show that for interacting random walks a pointwise bound on the entries of the (discrete) Hessian of \( V \) plus a local condition (47) on the behavior of \( V \) at the origin imply the family of convex Sobolev inequalities.

**Corollary 3.2.** Let \( V : \mathbb{N}^d \to \mathbb{R} \) be twice continuously differentiable and such that

\[ \nabla_i \nabla_j^+ V(\eta) \geq 0, \quad \forall \eta \in \mathbb{N}^d, i, j \leq d. \]  

holds uniformly in \( x \in \mathbb{R}^d \). Consider the generator \( \mathcal{L} \) given by (27), (28) and assume that

\[ \inf_{\eta \in \mathbb{N}^d, i = 1, \ldots, d} \lambda - \sum_{j=1}^d \left[ \exp(-\nabla_j^+ V(\eta)) - \exp(-\nabla_j^+ V(\gamma_i^+\eta)) \right] \geq 0. \]  

Then

(i) The conclusion of Theorem 3.1 holds with \( \kappa \) given by

\[ \inf_{\eta \in \mathbb{N}^d, i = 1, \ldots, d} \lambda + \sum_{j=1}^d \left[ \exp(-\nabla_j^+ V(\eta)) - \exp(-\nabla_j^+ V(\gamma_i^+\eta)) \right] - \sum_{j=1}^d \left[ \exp(-\nabla_j^+ V(\eta)) - \exp(-\nabla_j^+ V(\gamma_i^+\eta)) \right] \leq 0. \]

(ii) If

\[ \min_{i = 1, \ldots, d} \lambda - \sum_{j=1}^d \exp(-\nabla_j^+ V(\eta)) \geq 0, \]  

where \( \eta = (0, \ldots, 0) \in \mathbb{N}^d \), then the conclusion of Theorem 3.1 holds with \( \kappa \) given by (47).

**Proof.** Using the definition of the transition rates (28) we obtain from (45) that \( \nabla_i^+ c(\eta, \gamma_i^+) \leq 0 \), which gives

\[ \kappa^+(\eta, i) = -\nabla_i^+ c(\eta, \gamma_i^+) \geq \exp(-V(\eta)) - \exp(-V(\gamma_i^+\eta)) \geq 0, \quad \forall \eta \in \mathbb{N}^d, i \leq d \]

and

\[ \kappa^-(\eta, i) = \lambda - \sum_{j=1}^d \exp(-\nabla_j^+ V(\eta)) - \exp(-\nabla_j^+ V(\gamma_i^+\eta)), \quad \forall \eta \in \mathbb{N}^d, i \leq d, \]

Therefore if (46) holds, so does (H3.3). An application of Theorem 3.1 concludes the proof of (i). From (45) we obtain

\[ \nabla_j^+ V(\eta) \geq \nabla_j^+ V(\eta), \quad \forall \eta \in \mathbb{N}^d, j \leq d, \]

that yields the bound

\[ \forall \eta, i \quad \kappa^-(\eta, i) \geq \lambda - \sum_{j=1}^d \exp(-\nabla_j^+ V(\eta)), \]

This last bound together with (i) prove (ii).

We did not introduce an "inverse temperature" parameter \( \beta \) in Corollary 3.2. If we had done so, we could have seen that, as for Corollary 3.1, the local condition (47) is always satisfied in the low temperature regime \( \beta \to +\infty \).
3.3 Contraction of the Wasserstein distance

Let $p \geq 1$ and $d(\eta, \bar{\eta})$ be the graph distance on $\mathbb{N}^d$:

$$d(\eta, \bar{\eta}) = \sum_{i=1}^{d} |\eta_i - \bar{\eta}_i|.$$

For given $\mu, \nu \in \mathcal{P}_p(\mathbb{N}^d)$ with finite $p$-th moment the Wasserstein distance of order $p$ is defined as

$$W_p(\mu, \nu) = \left( \inf_{\pi \in \Pi(\mu, \nu)} \sum_{\eta', \eta''} d^p(\eta', \eta'') \pi(\eta', \eta'') \right)^{1/p},$$

where $\Pi(\mu, \nu)$ is, as usual, the set of all couplings of $\mu$ and $\nu$. In the next Theorem we show that the hypothesis needed for Theorem 3.1 are equivalent to certain contractivity properties of the Wasserstein distance along the semigroup $(S_t)_{t \geq 0}$ generated by $\mathcal{L}$. In the statement of the next theorem, for $\mu, \nu \in \mathcal{P}_p(\mathbb{N}^d)$ we denote by $\mu_t, \nu_t$ the laws $\mu S_t, \nu S_t$, i.e. $\mu_t$ (resp. $\nu_t$) is the distribution at time $t$ of a Markov chain with generator $\mathcal{L}$ and initial distribution $\mu$ (resp. $\nu$).

**Theorem 3.2.** Let $V^-, V^+ : \mathbb{N}^d \to \mathbb{R}$ be given and $\mathcal{L}$ be defined by (31) (32). Moreover, assume that (H0) (H1) hold and that

$$\inf_{\eta \in \mathbb{N}^d} \nabla_i^+ V^-(\eta) > -\infty, \quad \inf_{\eta \in \mathbb{N}^d} \{ \nabla_i^- V^-(\eta) : \eta_i \leq K \} > -\infty \quad \forall K > 0, i \leq d. \quad (48)$$

The following statements are equivalent for any $\kappa > 0$.

(i) The estimate

$$W_p(\mu_t, \nu_t) \leq \exp \left( -\frac{\kappa}{p} t \right) W_p(\mu, \nu) \quad (49)$$

holds uniformly on $\mu, \nu \in \mathcal{P}_p(\mathbb{N}^d)$, $t > 0$ and $p \geq 1$.

(ii) $V^-, V^+$ satisfy (H3.3) and

$$\inf_{\eta \in \mathbb{N}^d} \sum_{i \leq d} \kappa^+(\eta, i) + \kappa^-(\eta, i) \geq \kappa, \quad (50)$$

where $\kappa^+(\eta, i)$ and $\kappa^-(\eta, i)$ are defined at (33) and (34).

We have the following Corollary.

**Corollary 3.3.** Let $V^-, V^+ : \mathbb{N}^d \to \mathbb{R}$ be given and $\mathcal{L}$ be defined by (31) (32). Assume that (H0) (H1) and (48) hold. If the contraction estimate (49) holds uniformly on $\mu, \nu \in \mathcal{P}_p(\mathbb{N}^d)$, $t > 0$ and $p \geq 1$, then for all $\phi$ satisfying (H2) the convex Sobolev inequality holds with constant $\kappa_\phi = \kappa$ and for any $\alpha \in [1, 2)$ the Beckner inequality holds with $\kappa_0 = \alpha \kappa$.

Note that if $\mathcal{L}$ satisfies (H0) (H1) the only reversible measure is $m = \frac{1}{Z} \exp(-V^+ - V^-)$. The technical assumption (48) is made to simplify the proof and put forward the main ideas but does not play an essential role. It is likely that this assumption can be largely weakened. Moreover, observe that the second condition therein is always satisfied in the setting (28), i.e. when $V^-(\eta) = \sum_{i=1}^{d} \log(\lambda) \eta_i + \log(\eta_i!)$.

**Remark 3.1.** For the Langevin dynamics (29) even more is known [35], indeed, Wasserstein contraction for $p = 1$ implies and is in fact equivalent to the same property for $p \geq 1$, all this properties being equivalent to the $\kappa$ convexity of $V$. In particular, this implies that Wasserstein contraction for $p = 1$ suffices to conclude that the logarithmic Sobolev inequality holds. At the moment of writing, we do not know if the same result holds in $\mathbb{N}^d$. However, it is not hard to see that for continuous Markov chains on $\mathbb{N}^d$, there is no equivalence between Wasserstein contraction for different values of $p$.

In the proof of the Theorem, we will need the following technical lemma, whose proof we defer to the appendix.

**Lemma 3.1.** Under the same hypothesis of Theorem 3.2 consider $\eta \in \mathbb{N}^d$ and set $\mu = \delta_\eta \in \mathcal{P}_p(\mathbb{N}^d)$. Then there exist $C > 0$ such that

$$W_p^p(\mu_t, \bar{\mu}_t) \leq C t^2 \quad (51)$$

holds for $t$ small enough, where

$$\bar{\mu}_t = (1 - t \sum_{\gamma \in G} c(\eta, \gamma)) \delta_\eta + t \sum_{\gamma \in G} c(\eta, \gamma) \delta_{\gamma \eta} \quad (52)$$

holds for $t$ small enough, where

$$\bar{\mu}_t = (1 - t \sum_{\gamma \in G} c(\eta, \gamma)) \delta_\eta + t \sum_{\gamma \in G} c(\eta, \gamma) \delta_{\gamma \eta} \quad (52)$$
Proof of Theorem 3.2. We first prove (i)⇒(ii) in two steps. In the first step we show that (i) implies (H3.3), whereas in the second step we show that it implies (50).

• Step1: (i)⇒(H3.3) If (H3.3) is violated, then there exist \( \eta \in \mathbb{N}^d \) and \( i \leq d \), such that \( \kappa^-(\eta, i) < 0 \) or \( \kappa^+(\eta, i) < 0 \). We only treat the case \( \kappa^-(\eta, i) < 0 \). The proof for the other case is almost identical and we omit the details. We set \( \mu = \delta_{\eta}, \nu = \delta_{\gamma^+_{\eta}} \), and define \( \mu_t, \nu_t \) as in (52). Invoking Lemma 3.1 we have that

\[
W_p^p(\mu, \nu) = W_p^p(\mu_t, \nu_t) + o(t),
\]

(53)

where as usual by \( o(t) \) we denote a function such that \( o(t)/t \to 0 \) as \( t \to 0 \). Consider an arbitrary coupling \( \pi \) of \( \mu_t \) and \( \nu_t \). Then \( \pi \) is supported on the set \( \{ (\gamma\eta, \gamma^+_{\eta}) : \gamma, \bar{\gamma} \in \mathbb{G}^* \} \). We claim that \( \pi(d(\eta', \eta'')) \geq 2 \geq -t\kappa^-(\eta, i) \). To prove this, we observe that

\[
\pi[d(\eta', \eta'') \geq 2] \geq \pi[\eta' = \gamma^+_{\eta}, \eta'' \neq \eta] + \sum_{\gamma \in \mathbb{G}} \pi[\eta' = \gamma\eta, \eta'' \neq \{\eta, \gamma^+_{\eta}\}].
\]

We have

\[
\pi[\eta' = \gamma^+_{\eta}, \eta'' \neq \eta] + \sum_{\gamma \in \mathbb{G}} \pi[\eta' = \gamma\eta, \eta'' \neq \{\eta, \gamma^+_{\eta}\}]
= \pi[\eta' = \gamma^+_{\eta}] - \pi[(\gamma^+_{\eta}, \eta)] + \sum_{\gamma \in \mathbb{G}} \pi[\eta' = \gamma\eta] - \pi[(\gamma\eta, \eta)] - \pi[(\gamma\eta, \gamma^+_{\eta})]
= tc(\eta, \gamma^+_{\eta}) - \sum_{\gamma \in \mathbb{G}} \pi[(\gamma\eta, \eta)] + \sum_{\gamma \in \mathbb{G}} \pi[\eta' = \gamma\eta] - \pi[(\gamma\eta, \gamma^+_{\eta})]
\geq tc(\eta, \gamma^+_{\eta}) - \pi[\eta' = \eta] + \sum_{\gamma \in \mathbb{G}} \max[\pi[\eta' = \gamma\eta] - \pi[\eta' = \gamma^+_{\eta}], 0]
\geq tc(\eta, \gamma^+_{\eta}) - c(\gamma^+_{\eta}, \gamma^+_{\eta}) + t \sum_{\gamma \in \mathbb{G}} \max[-\gamma^+_{\eta} c(\eta, \gamma), 0]
= -t\kappa^-(\eta, i),
\]

which proves the claim. We also observe that \( \pi[d(\eta', \eta'') = 0] = \pi[(\eta\eta, \eta)] + \pi[(\gamma^+_{\eta}, \eta^+_{\eta})] \leq tc(\gamma^+_{\eta}, \gamma^+_{\eta}) + c(\gamma, \gamma^+_{\eta}) \). Therefore, if \( \bar{\pi}_t \) is the optimal coupling for \( W_p^p(\mu_t, \nu_t) \) we have

\[
\frac{W_p^p(\mu_t, \nu_t) - W_p^p(\mu, \nu)}{t} \geq \frac{(2^p - 1)\bar{\pi}_t[d(\eta', \eta'') \geq 2] - \bar{\pi}_t[d(\eta', \eta'') = 0]}{t}
\geq - (2^p - 1)\kappa^-(\eta, i) - [c(\gamma^+_{\eta}, \gamma^+_{\eta}) + c(\gamma, \gamma^+_{\eta})]
\]

Since this quantity is strictly positive for \( p \) large enough we obtain a contradiction with (49) after letting \( t \to 0 \) and recalling that \( W_p^p(\mu_t, \nu_t) = W_p^p(\mu_t, \nu_t) + o(t) \). The proof that (i) implies (H3.3) is now complete.

• Step2: (i)⇒(50). Consider an arbitrary pair \( \eta, i \), set \( \mu = \delta_{\eta}, \nu = \delta_{\gamma^+_{\eta}} \), and let \( \mu_t, \nu_t \) as in (52). To do the proof, we construct explicitly the optimal coupling \( \bar{\pi}_t \) for \( W_p^p(\mu_t, \nu_t) \), which is given by setting \( \bar{\pi}_t[(\eta', \eta'')] = 0 \) outside the set \( \{ (\gamma\eta, \gamma^+_{\eta}) : \gamma, \bar{\gamma} \in \mathbb{G}^* \} \) and defining

\[
\bar{\pi}_t(\gamma\eta, \gamma^+_{\eta}) = \begin{cases} 
 t \min\{c(\gamma\eta, \gamma)\}, & \text{if } \gamma = \gamma \in G, \\
 t \max\{-\nabla^+_{\gamma} c(\gamma\eta, 0)\}, & \text{if } \gamma = \gamma^+ \in G, \gamma \notin \gamma^+_{\eta}, \gamma^-, \\
tc^+ (\eta, i), & \text{if } \gamma \in G, \gamma \notin \gamma^+_{\eta}, \gamma^+, \gamma^-, \\
 c(\gamma, \gamma^+_{\eta}), & \text{if } \gamma = \gamma^+_{\eta}, \gamma^-, \\
 tc^-(\eta, i), & \text{if } \gamma = \gamma^+_{\eta}, \gamma^+, \\
 \bar{\mu}_t(\eta) - tc^-(\eta, i) = \bar{\nu}_t(\eta) - tc^+(\eta, i), & \text{if } \gamma = e, \gamma = \gamma^+_{\eta}, \\
 \end{cases}
\]

(54)

The admissibility and optimality of \( \bar{\pi}_t \) are shown at Lemma 3.2, which we prove separately. By construction, if \((\eta', \eta'')\) is in the support of \( \bar{\pi}_t \) then \( d(\eta', \eta'') \) is worth 0 if \((\eta', \eta'') = (\eta, \eta), (\gamma^+_{\eta}, \gamma^+_{\eta}) \) and 1 otherwise. This gives

\[
W_p^p(\mu_t, \nu_t) = 1 - t[\kappa^-(\eta, i) + \kappa^+(\eta, i)]
\]

(55)
Therefore, invoking (53) and recalling that $W_p(\mu, \nu) = 1$ we obtain
\[
\frac{d}{dt} W_p^t(\mu_t, \nu_t) = -[\kappa^-(\eta, i) + \kappa^+(\eta, i)] \quad \text{and} \quad \frac{d}{dt} W_p(\mu_t, \nu_t) = -\frac{1}{p}[\kappa^-(\eta, i) + \kappa^+(\eta, i)].
\] (56)

Comparing this with the hypothesis (49) yields (50) since the choice of $\eta$ and $i$ was arbitrary.

**Proof of (ii)$\Rightarrow$(i)** Let $p \geq 1, \eta, i$ be fixed, $\mu = \delta_\eta, \nu = \delta_\gamma$ and $\mu_t, \nu_t$ be defined as before via (52). The hypothesis (ii) implies that the coupling $\pi_t$ defined at (54) is admissible and from Lemma 3.2 we get that $\pi_t$ is optimal for $W_p(\mu_t, \nu_t)$. Following the proof of Step 2, we obtain the relation (56). Using (50) in there we arrive at
\[
\frac{d}{dt} W_p(\mu_t, \nu_t) \bigg|_{t=0} = -\frac{1}{p} W_p(\mu, \nu)
\] (57)
We can extend (57) to any pair Dirac measures $\mu = \delta_\eta, \nu = \delta_\gamma$ by using the bound (57) along a geodesic connecting $\eta$ and $\gamma$ and the triangular inequality. Next, consider two arbitrarily compactly supported $\mu$ and $\nu$ in $\mathcal{P}_p(\mathbb{N}^d)$ and denote by $\pi_0$ the optimal coupling for $W_p(\mu, \nu)$. For any $(\eta, \eta) \in \text{supp}(\pi_0)$, let $\pi_t^{\eta, \eta}$ the optimal coupling for $W_p(\mu_t^0, \nu_t^0)$, where $\mu_t^0 = \delta_{\eta}, \nu_t^0 = \delta_{\eta}$. By constructing $\pi_t \in \Pi(\mu_t, \nu_t)$ as follows
\[
\pi_t(\eta', \eta'') = \sum_{\eta, \gamma \in \mathbb{N}^d} \pi_0(\eta, \eta) \pi_t^{\eta, \eta}(\eta', \eta'')
\]
it is seen that (57) holds. The extension to non-identical probability measures in $\mathcal{P}_p(\mathbb{R}^d)$ follows a standard approximation argument. The differential inequality (57) extends to an arbitrary value of $t$ by Markovianity. An application of Grönwall’s lemma concludes the proof.

Here we prove the auxiliary Lemma needed for the proof of Theorem 3.2. To follow the proof, it may be helpful to refer to Figure 3.

**Lemma 3.2.** Let the hypothesis of Theorem 3.2 hold and (H3.3) hold as well. For $\mu = \delta_\eta, \nu = \delta_\gamma$ and $t$ small enough, let $\mu_t, \nu_t$ be given by (52). Next, define $\tilde{\pi}_t \in \mathcal{P}(\mathbb{N}^d \times \mathbb{N}^d)$ by setting
\[
\tilde{\pi}_t(\eta, \gamma) = \begin{cases}
\begin{aligned}
t \min\{c(\eta, \gamma), c(\gamma, \eta), c(\gamma, \gamma)\}, & \text{if } \gamma = \gamma \in G, \\
t \max\{\nabla^+ c(\eta, \gamma), 0\}, & \text{if } \gamma = \gamma_i^+, \text{ and } \gamma \in G, \gamma \neq \gamma_i^+, \gamma_i^-, \\
t \max\{-\nabla^+ c(\eta, \gamma), 0\}, & \text{if } \gamma \in G, \gamma \neq \gamma_i^+, \gamma_i^-, \\
t^{\gamma_i^+(\gamma, i)}, & \text{if } \gamma = \gamma_i^+, \gamma = e, \\
t^{\gamma_i^-(\gamma, i)}, & \text{if } \gamma = e, \gamma = e, \\
\bar{\mu}_t(\eta) - t^{\kappa^-(\eta, i)}(\eta', \gamma) = \bar{\nu}_t(\eta) - t^{\kappa^-(\eta, i)}(\eta', \gamma), & \text{if } \gamma \in G, \gamma \neq \gamma_i^+, \gamma_i^-,
\end{aligned}
\end{cases}
\]
and
\[
\pi(\eta', \eta'') = 0, \quad \forall (\eta', \eta'') \notin \{ (\eta, \eta, \gamma) : \gamma \in G^+ \}.
\]

Then $\tilde{\pi}_t \in \Pi(\mu_t, \nu_t)$ and $\tilde{\pi}_t$ is optimal for $W_p(\mu_t, \nu_t)$.

**Proof.** To check that $\tilde{\pi}_t$ is always non-negative, we remark that (H3.3) implies $\kappa^{\pm}(\eta, i) \geq 0$ and that if $t$ is small enough we have $\bar{\mu}_t(\eta) - t^{\kappa^-(\eta, i)}(\eta', \gamma, i) \geq 0$. It can be verified with a direct calculation that the marginals of $\tilde{\pi}_t$ are $\bar{\mu}_t$ and $\bar{\nu}_t$ respectively. To show optimality, we show that the support $\text{supp}(\tilde{\pi}_t)$ is cyclically monotone. That is to say,
\[
(\eta', \eta'', (\xi', \xi'')) \in \text{supp}(\tilde{\pi}_t) \Rightarrow d^p(\eta', \eta'') + d^p(\xi', \xi'') \leq d^p(\eta', \eta'') + d^p(\xi', \xi'').
\] (58)

Note that by construction, $d(\eta', \eta'') \leq 1$ on $\text{supp}(\tilde{\pi}_t)$. Therefore, if (58) is violated we can w.l.o.g. suppose that $\eta' = \xi''$. This can happen only if $\eta' = \xi'' = \eta$ or $\eta' = \xi'' = \gamma_i^+ \eta$. We show that in the first case (58) is always satisfied. The proof for the second case is almost identical and we omit the details. Therefore, from now on we assume $\eta' = \xi'' = \eta$. We observe that
\[
(\eta, \eta') \in \text{supp}(\tilde{\pi}) \Rightarrow \eta'' = \eta \in \{ \eta, \gamma_i^+ \eta \}
\]
\[
(\xi', \eta) \in \text{supp}(\tilde{\pi}) \Rightarrow \xi' \in \{ \eta \} \cup \{ \gamma : \gamma \in G, \gamma \neq \gamma_i^+ \}
\]
We verify case-by-case that (58) holds
- $\eta'' = \eta, \xi' = \eta$ In this case $d(\eta', \eta'') = d(\xi', \xi'') = d(\xi', \eta'') = d(\eta', \xi'') = 0$. 

21
• \( \eta'' = \eta, \xi' = \gamma \eta \) with \( \gamma \in G, \gamma \neq \gamma \). In this case
\[
d^\beta(\eta', \eta''') + d^\beta(\xi', \xi''') = d^\beta(\eta', \eta'') + d^\beta(\xi', \xi'') = d^\beta(\gamma \eta, \eta) = 1.
\]

• \( \eta'' = \gamma \eta, \xi' = \eta \) In this case
\[
d^\beta(\eta', \eta''') + d^\beta(\xi', \xi''') = d^\beta(\eta', \eta'') + d^\beta(\xi', \xi'') = d^\beta(\eta, \gamma \eta) = 1.
\]

• \( \eta'' = \gamma \eta, \xi' = \gamma \eta \) with \( \gamma \in G, \gamma \neq \gamma \). In this case \( d(\xi', \eta') = 2 \) and \( d(\eta', \eta''), d(\xi', \xi'') \leq 1 \) hold since \((\eta, \eta''), (\xi', \xi'') \in \text{supp}(\pi_t)\).

Therefore, (58) is never violated and \( \pi_t \) is optimal for \( W_p(\mu_t, \nu_t) \).

4 Glauber dynamics

In this section, we investigate (5) for the Glauber dynamics. In view of the applications to classical spin systems such as Curie Weiss or the Ising model, we assume that the moves \( \sigma \in G \) are involutions, i.e. \( \sigma^{-1} = \sigma \). However, this is not strictly necessary for our method to work. Given an inverse temperature parameter \( \beta > 0 \) and an Hamiltonian \( H : \Omega \rightarrow \mathbb{R} \) we construct a generator of the form (2) by setting
\[
c(\eta, \sigma) = \exp \left(-\frac{\beta}{2} \nabla_\sigma H(\eta) \right).
\]

If the state space \( \Omega \) is finite, [H0] and [H1] are satisfied and the reversible measure is the Gibbs measure
\[
m(\eta) = \frac{1}{Z_\beta} \exp(-\beta H(\eta)), \quad \forall \eta \in \Omega,
\]
where \( Z_\beta \) is the normalization. Let us now state precisely the assumptions needed for the main result of this section which is Theorem 4.1 below.

(H3.4) The set of moves \( G \) is such that
\[
\sigma^{-1} = \sigma, \quad \forall \sigma \in G,
\]
and the relation
\[
\sigma \gamma \eta = \gamma \sigma \eta
\]
holds uniformly on \( \eta \in \Omega, \sigma, \gamma \in G \).

(H4.4) \( \kappa(\eta, \sigma) \geq 0 \) uniformly on \( \eta \in \Omega, \sigma \in G \), where
\[
\kappa(\eta, \sigma) := c(\eta, \eta) - \sum_{\gamma \in G} \max_{\gamma \neq \sigma} \{-\nabla_\sigma c(\eta, \gamma), 0\}.
\]

4.1 Sufficient condition for Glauber dynamics

Theorem 4.1. Let \( \beta > 0, H : \mathbb{N}^d \rightarrow \mathbb{R} \) be given and the generator \( L \) be defined by (59). Moreover, assume that [H3.4] and [H4.4] hold. If we define
\[
\kappa = \inf_{\eta \in \Omega, \sigma \in G} \kappa(\eta, \sigma), \quad \kappa = \inf_{\eta \in \Omega, \sigma \in G} \kappa(\eta, \sigma)
\]
then the following holds
(i) For any \( \phi \) satisfying [H2] the convex Sobolev inequality [13] holds with with \( \kappa_\phi = \kappa \).
(ii) The modified Logarithmic Sobolev inequality [6] holds with \( \kappa_1 = \kappa + 2 \bar{\kappa} \).
(iii) For \( \alpha \in (1, 2] \), the Beckner inequality [7] holds with \( \kappa_\alpha = \alpha \kappa \).

The assumptions and the proof of Theorems 3.4 and 4.1 bear several resemblances. We could indeed merge them in a single general result. However, for the sake of clarity we prefer to keep the two results distinct.
Proof. The proof is done verifying that the assumptions of Proposition 2.1 hold with \( \kappa' = \kappa''' = \kappa \) and \( \kappa'' = \bar{\kappa} \) for the coupling rates we are going to construct. The proof is then finished applying the same proposition. For any \( \eta \in \Omega \) we define \( \Upsilon^c(\eta), \Upsilon^>(\eta), \Upsilon^= (\eta) \) as follows

\[
\begin{align*}
\Upsilon^c(\eta) &= \{(\sigma, \gamma) \in G \times G : \sigma \neq \gamma, \nabla_\sigma c(\eta, \gamma) < 0\} \\
\Upsilon^>(\eta) &= \{(\sigma, \gamma) \in G \times G : \sigma \neq \gamma, \nabla_\sigma c(\eta, \gamma) > 0\} \\
\Upsilon^= (\eta) &= \{(\sigma, \gamma) \in G \times G : \sigma \neq \gamma, \nabla_\sigma c(\eta, \gamma) = 0\}
\end{align*}
\]

We remark that under the current hypothesis we have \( \Omega \leq \gamma \) and \( \Omega \geq \gamma \). Indeed, (59) implies that

\[
\begin{align*}
\Upsilon^c(\eta) &= \{(\sigma, \gamma) \in G \times G : \sigma \neq \gamma, \nabla_\sigma c(\eta, \gamma) < 0\} \\
\Upsilon^>(\eta) &= \{(\sigma, \gamma) \in G \times G : \sigma \neq \gamma, \nabla_\sigma c(\eta, \gamma) > 0\} \\
\Upsilon^= (\eta) &= \{(\sigma, \gamma) \in G \times G : \sigma \neq \gamma, \nabla_\sigma c(\eta, \gamma) = 0\}
\end{align*}
\]

A direct calculation using (H3.4) and (H4.4) shows that (65) define admissible coupling rates. In particular, (H4.4) ensures that \( c^{pl}(\eta, \sigma, \gamma, \bar{\gamma}) \) and \( c^{pl}(\eta, \sigma, \epsilon, \sigma) \) are non-negative. We now show that (20) holds with \( \kappa' = \kappa \). To this aim observe that for any \( f > 0 \) and \( \phi \) satisfying (H2) the choice (65) give that the left hand side of (20) rewrites as \( \frac{1}{2}(A + B + C + D) \) with

\[
A = \sum_{\eta \in \Omega, \sigma, \gamma \in G \atop \sigma \neq \gamma} c(\eta, \sigma) \min\{c(\eta, \gamma), c(\sigma, \gamma)\} \nabla_{\gamma,\gamma} f^\phi(\eta, \sigma) \mathbf{m}(\eta),
\]

\[
B = -\sum_{\eta \in \Omega} c(\eta, \sigma) \nabla_\sigma c(\eta, \gamma) \nabla_{\gamma,\sigma} f^\phi(\eta, \sigma) \mathbf{m}(\eta),
\]

\[
C = \sum_{\eta \in \Omega} c(\eta, \sigma) \nabla_\sigma c(\eta, \gamma) \nabla_{\gamma,\sigma} f^\phi(\eta, \sigma) \mathbf{m}(\eta),
\]

\[
D = \sum_{\eta \in \Omega} c(\eta, \sigma) \nabla_\sigma c(\eta, \gamma) \nabla_{\gamma,\sigma} f^\phi(\eta, \sigma) \mathbf{m}(\eta).
\]

We now show that \( A = B = C = 0 \). We begin by considering \( B \). Using (H3.4) and (59) we get that for all \( \sigma \neq \gamma \)

\[
(\sigma, \gamma) \in \Upsilon^c(\eta) \Leftrightarrow \nabla_\sigma \nabla_\gamma H(\eta) > 0 \Leftrightarrow (\gamma, \sigma) \in \Upsilon^c(\eta).
\]

Therefore we can rewrite \( B \) exchanging the labels \( \sigma \) and \( \gamma \) as

\[
-\frac{1}{2} \sum_{\eta \in \Omega} c(\eta, \sigma) \nabla_\sigma c(\eta, \gamma) - c(\eta, \gamma) \nabla_\gamma c(\eta, \sigma) [f^\phi(\eta, \gamma) - f^\phi(\eta, \sigma)] \mathbf{m}(\eta)
\]

from which \( B = 0 \) follows. Indeed, (59) implies that \( c(\eta, \sigma) \nabla_\sigma c(\eta, \gamma) = c(\eta, \gamma) \nabla_\gamma c(\eta, \sigma) \) for all \( \eta, \gamma, \sigma \). Using (10) on \( C \) with

\[
F(\eta, \sigma) = \sum_{\gamma \in \Upsilon^c(\eta)} \nabla_\sigma c(\eta, \gamma) \nabla_{\gamma,\sigma} f^\phi(\eta, \sigma)
\]

and (H3.4) yields the equivalent expression

\[
-\sum_{\eta \in \Omega} c(\eta, \sigma) \nabla_\sigma c(\eta, \gamma) [f^\phi(\eta, \gamma) - f^\phi(\eta, \sigma)] \mathbf{m}(\eta).
\]

Using the same argument used to show \( B = 0 \), we conclude that \( C = 0 \). The proof that \( A = 0 \) is done in the auxiliary Lemma 1.1 and follows from reversibility. Finally, recalling that \( f^\phi \) vanishes on the diagonal we easily get that \( D \leq -2\kappa \mathcal{E}(\phi(\gamma), \gamma) \), which completes the proof that (20) holds with \( \kappa' = \kappa \). From the construction of the coupling rates we also have that (21) holds with \( \kappa'' = \bar{\kappa} \) and (22) holds with \( \kappa''' = \kappa \). An application of Proposition 2.1 finishes the proof.

\[\square\]
Lemma 4.1. Under the hypothesis of Theorem 4.1 we have
\[ \sum_{\eta \in \Omega, \sigma, \gamma \in G \atop \sigma \neq \gamma} c(\eta, \sigma) \min\{c(\eta, \gamma), c(\sigma, \gamma)\} \nabla_{\gamma, \gamma} f^\phi(\eta, \sigma \gamma) m(\eta) = 0 \] (66)
holds for all \( f > 0 \).

Proof. Recalling the definition of \( \Upsilon^<(\eta), \Upsilon^=(\eta), \Upsilon^>(\eta) \), we rewrite (66) as the sum of the three terms
\[
\sum_{\eta \in \Omega \atop (\sigma, \gamma) \in \Upsilon^<(\eta)} c(\eta, \sigma) c(\sigma, \gamma) \nabla_{\gamma, \gamma} f^\phi(\eta, \sigma \gamma) m(\eta) \\
\sum_{\eta \in \Omega \atop (\sigma, \gamma) \in \Upsilon^=(\eta)} c(\eta, \sigma) c(\eta, \gamma) \nabla_{\gamma, \gamma} f^\phi(\eta, \sigma \gamma) m(\eta) \\
\frac{1}{2} \sum_{\eta \in \Omega \atop (\sigma, \gamma) \in \Upsilon^>(\eta)} c(\eta, \sigma) [c(\eta, \gamma) + c(\sigma, \gamma)] \nabla_{\gamma, \gamma} f^\phi(\eta, \sigma \gamma) m(\eta). \] (67)

Using the reversibility (10) and ([H3.4]) on the second term with with
\[ F(\eta, \gamma) = \sum_{\sigma, (\sigma, \gamma) \in \Upsilon^>(\eta)} c(\eta, \sigma) \nabla_{\gamma, \gamma} f^\phi(\eta, \sigma \gamma) \]
yields the equivalent form
\[ - \sum_{\eta \in \Omega \atop (\sigma, \gamma) \in \Upsilon^>(\gamma \eta)} c(\eta, \sigma) c(\gamma, \sigma) \nabla_{\gamma, \gamma} f^\phi(\eta, \sigma \gamma) m(\eta) \] (68)
Next observe that, thanks to (59) and ([H3.4]) we obtain
\[ (\sigma, \gamma) \in \Upsilon^>(\gamma \eta) \Leftrightarrow \nabla_{\gamma} \nabla_{\sigma} H(\eta) > 0 \Leftrightarrow \nabla_{\gamma} \nabla_{\sigma} H(\eta) < 0 \Leftrightarrow (\sigma, \gamma) \in \Upsilon^<(\eta). \]
Plugging this back into (68), we obtain that the first two terms in (67) cancel. Arguing as for the first two terms and observing that \( \Upsilon^=(\eta) = \Upsilon^=(\gamma \eta) \) we obtain that the third term is also worth 0. \( \square \)

4.2 Applications to spin systems

4.2.1 Curie-Weiss model

For the Curie-Weiss model we have \( \Omega = \{-1, 1\}^N \) for some \( N > 0 \) and the set of moves is \( G = \{\sigma_i\}_{i=1, \ldots, N} \). \( \sigma_i \) acts on \( \eta \) flipping its \( i \)-th coordinate, i.e.
\[ \sigma_i(\eta)_j = \begin{cases} \eta_j, & \text{if } j \neq i \\ -\eta_j, & \text{if } j = i. \end{cases} \] (69)
The Hamiltonian is given by
\[ H: \{-1, 1\}^N \to \mathbb{R}, \quad H(\eta) = -\frac{1}{2N} \sum_{i,j=1}^N \eta_i \eta_j. \]
For a given \( \beta > 0 \), the transition rates of the Glauber dynamics are then given by
\[ c(\eta, \sigma_i) = \exp \left( -\beta \eta_i m_i(\eta) \right), \quad \text{with } m_i(\eta) = \frac{1}{N} \sum_{j \neq i} \eta_j. \] (70)
To state our result for the Curie-Weiss model, it is convenient to introduce \( f_{CW, \beta, N}: \mathbb{N} \to \mathbb{R} \)
\[ f_{CW, \beta, N}(m) := \exp \left( -\beta \frac{N}{N-1} (N-1-2m) \right) \left[ 1 - (N-1-m) \exp \left( \frac{2\beta}{N} \right) - 1 \right] \\
+ \exp \left( \frac{\beta}{N} (N-1-2m) \right) \left[ 1 - m \exp \left( \frac{2\beta}{N} \right) - 1 \right] \]
Theorem 4.2. Assume that

\[(N - 1)(\exp(2\beta/N) - 1) \leq 1.\] \hspace{1cm} (71)

Then the conclusion of Theorem 4.1 holds with

\[\kappa = f_{\text{CW},\beta,N}([(N - 1)/2]),\]
\[\bar{\kappa} = \exp(-\frac{\beta}{N}(N - 1))\left[1 - (N - 1)(1 - \exp(2\beta/N))\right].\] \hspace{1cm} (72)

In particular, if \(N\) is odd, we have \(f_{\text{CW},\beta,N}([(N - 1)/2]) = 2\left(1 - \frac{N - 1}{N}\right)(\exp(2\beta/N) - 1)\).

Remark 4.1. As \(N \to +\infty\), our condition (71) reads as \(\beta \leq 1/2\), thus improving on [23, Cor 4.5]. Estimates on the best constant for the classical (i.e. non modified) logarithmic Sobolev inequality that are valid up to the critical temperature \(\beta = 1\) have been obtained in [27]. For \(N\) large the MLSI constant we obtain from Theorem 4.1 is \(2(1 - \beta) + 2(1 - 2\beta)\exp(-\beta)\) improving on the value \(4(1 - 2\beta)\exp(2\beta)\exp(-\beta)\) found in [13]. The findings of Theorem 4.1 concerning general convex Sobolev inequalities and Beckner’s inequalities seem to be new.

Proof. We obtain from (70) that for all \(\eta, i, j \neq i:\)

\[c(\sigma_\eta, \sigma_j) = \exp\left(\frac{2\beta_\eta_\eta_j}{N}\right)\]

Moreover, if \(|\{j \neq i : \eta_i = 1\}| = m\) we have

\[c(\sigma_\eta, \sigma_i) = \exp\left(-\frac{\beta}{N}(N - 1 - 2m)\right)\]

Therefore

\[\kappa(\eta, \sigma_i) = \exp\left(-\frac{\beta}{N}(N - 1 - 2m)\right) - \sum_{j : \eta_i, \eta_j = -1} (\exp\left(\frac{2\beta}{N}\right) - 1)c(\sigma_\eta, \sigma_j)\]

Next, we observe that if \(\eta_i = -1\), we have that \(c(\sigma_\eta, \sigma_j) = c(\sigma_\eta, \sigma_i)\). Since there are \(N - 1 - m\) spins of this type, we obtain

\[\kappa(\eta, \sigma_i) = \exp\left(-\frac{\beta}{N}(N - 1 - 2m)\right)[1 - (N - 1 - m)(\exp\left(\frac{2\beta}{N}\right) - 1)]\] \hspace{1cm} (73)

In particular, we obtain that (71) implies (63) and we can apply Theorem 4.1. From (73) we also obtain that \(\bar{\kappa}\) therein is given by (72). To finish the proof, observe that

\[\kappa(\eta, \sigma_i) + \kappa(\sigma_\eta, \sigma_i) = \exp\left(-\frac{\beta}{N}(N - 1 - 2m)\right)[1 - (N - 1 - m)(\exp\left(\frac{2\beta}{N}\right) - 1)] + \exp\left(\frac{\beta}{N}(N - 1 - 2m)\right)[1 - m(\exp\left(\frac{2\beta}{N}\right) - 1)] = f_{\text{CW},\beta,N}(m).\]

The right hand side being a convex function \(m \in [0, N - 1]\) and symmetric around \(m = (N - 1)/2\), it achieves its minimum on \([0, \ldots, N - 1]\) at \(m = [(N - 1)/2]\). \(\square\)

4.2.2 Ising model

Let \(\Lambda \subseteq \mathbb{Z}^d\) be a connected subset, which we endow with the natural graph structure \(\sim\) inherited from \(\mathbb{Z}^d\). The state space is \(\Omega = \{-1, 1\}^\Lambda\) and the Hamiltonian is

\[H : \{-1, 1\}^\Lambda \longrightarrow \mathbb{R}, \quad H(\eta) = \frac{1}{2} \sum_{x \sim y} \eta_x \eta_y.\]

where \(x \sim y\) means that \(x\) and \(y\) are neighbors in \(\Lambda\). The set of moves is \(G = \{\sigma_x\}_{x \in \Lambda}\), where \(\sigma_x\) is the flip of the spin at site \(x\), see [60]. Therefore, for \(\beta > 0\) the transition rates [50] for the Glauber dynamics are

\[c(\eta, \sigma_x) = \exp\left(-\beta \eta_x \sum_{y \sim x} \eta_y\right)\]
Theorem 4.3. Assume that

\[ 2d(1 - \exp(-2\beta)) \exp(4d\beta) \leq 1 \]  

(74)

Then the conclusion of Theorem 4.1 holds with

\[ \kappa = 2 - 2d(1 - \exp(-2\beta)) \exp(2\beta d) \]

(75)

and

\[ \bar{\kappa} = \exp(-2\beta d) - 2d(1 - \exp(-2\beta)) \exp(2\beta d). \]

(76)

Remark 4.2. In the article [27, Cor 4.4] the authors establish entropic Ricci curvature bounds, and in particular MLSI, under the condition \( \varepsilon(\beta) \leq 1 \), where

\[ \varepsilon(\beta) = (2d - 1)(1 - \exp(-2\beta)) \exp(4d\beta) \]

The condition (74) of Theorem 4.3 is therefore more demanding. However, when both results apply, the bound on the MLSI constant \( \kappa + 2\bar{\kappa} \) provided by Theorem 4.3 is better than the bound \( 4(1 - \varepsilon(\beta)) \exp(-2\beta d) \) found there, at least for \( d \geq 2 \). Indeed, after some calculations, one can find that the difference between the two bounds is

\[ 2(1 - \exp(-2\beta d)) + (4 - 3\frac{2d}{2d - 1}) \varepsilon(\beta) \exp(-2\beta d). \]

As for the Curie Weiss, the findings of Theorem 4.3 concerning Beckner inequalities and general convex Sobolev inequalities seem to be new and estimates on the best constant for the non modified logarithmic Sobolev inequality have been obtained in [27] under a condition that appears to be weaker than (74).

Proof. The proof is done verifying that the hypothesis of Theorem 4.1 hold with \( \kappa, \bar{\kappa} \) as in (75) and (76). We first observe that for all \( x \neq y \) we have

\[ \frac{c(\sigma_x \eta, \sigma_y)}{c(\eta, \sigma_y)} = \begin{cases} \exp(2\beta \eta_x \eta_y) & \text{if } x \sim y, \\ 1 & \text{otherwise,} \end{cases} \]

and that if \( |\{y \sim x : \eta_x \eta_y = 1\}| = m \in \{0, \ldots, 2d\} \) then we have

\[ c(\eta, \sigma_x) = \exp(2\beta(d - m)), \quad c(\sigma_x \eta, \sigma_x) = \exp(2\beta(m - d)). \]

Therefore, Recalling the definition (63) of \( \kappa(\eta, \sigma_x) \) we have

\[
\kappa(\eta, \sigma_x) = c(\sigma_x \eta, \sigma_x) + \sum_{y \sim x \atop \eta_y \eta_x = -1} c(\sigma_x \eta, \eta_y) - c(\eta, \sigma_y)
\]

\[ = c(\sigma_x \eta, \sigma_x) + (1 - \exp(2\beta)) \sum_{y \sim x \atop \eta_y \eta_x = -1} c(\sigma_x \eta, \eta_y) \]

If \( y \sim x \) with \( \eta_x \eta_y = -1 \) then

\[ c(\sigma_x \eta, \eta_y) = \exp(-\beta \eta_y \sum_{z \sim y (z \neq x)} (\sigma_x \eta_z)) \]

\[ = \exp \left( -\beta \eta_y \sum_{z \sim y (z \neq x)} \eta_z + \beta \eta_y \eta_x \right) \leq \exp(2\beta(d - 1)). \]

and therefore

\[ \kappa(\eta, \sigma_x) \geq \exp(2\beta(m - d)) - (1 - \exp(-2\beta))(2d - m) \exp(2\beta d). \]

(77)

In particular, if (74) holds, then (H4.4) is satisfied and we can apply Theorem 4.1. It remains to compute \( \kappa \) and \( \bar{\kappa} \). From (77) we immediately get that \( \bar{\kappa} \) can be taken as in (76). Using the same argument that led to (74) one gets

\[ \kappa(\sigma_x \eta, \sigma_x) \geq \exp(2\beta(d - m)) - (1 - \exp(-2\beta))m \exp(2\beta d) \]

If \( \kappa \) is given by (75), observing that \( a + 1/a \geq 2 \) we get that \( \kappa(\eta, \sigma_x) + \kappa(\sigma_x \eta, \sigma_x) \geq \kappa \) holds uniformly in \( \eta \in \Omega, x \in \Lambda \). The conclusion follows by Theorem 4.1.
5 More examples

5.1 Bernoulli Laplace

The Bernoulli Laplace model is the simple exclusion process on the complete graph. Given $L > N \in \mathbb{N}$, where $L$ represents the number of sites and $N$ the number of particles we consider the state space

$$\Omega = \{\eta : \{1, \ldots, L\} \to \{0, 1\} : \sum_{i=1}^{L} \eta_i = N\},$$

where $\eta_i = 1$ means that a particle is present at site $i$. For any $1 \leq i \leq L$ we define $\delta_i \in \Omega$ by

$$(\delta_i)_k = \begin{cases} 1, & \text{if } i = k, \\
0, & \text{otherwise.} \end{cases}$$

The set of moves is $G = \{\sigma_{ij}, i, j \in \{1, \ldots, L\}\}$, where

$$\sigma_{ij}(\eta) = \begin{cases} \eta - \delta_i + \delta_j, & \text{if } \eta_i(1 - \eta_j) = 0, \\
\eta, & \text{otherwise.} \end{cases}$$

The map $\sigma_{ij}$ moves a particle from site $i$ to site $j$, when this is possible. The jump rates for the Bernoulli-Laplace model are defined by

$$c(\eta, \sigma_{ij}) = \eta_i(1 - \eta_j).$$

and therefore the reversible measure $m$ is the uniform measure on $\Omega$. Assumptions $(H0)$ and $(H1)$ are clearly satisfied. In particular, $(\sigma_{ij})^{-1} = \sigma_{ji}$.

**Theorem 5.1.** For the Bernoulli Laplace model the following hold

(i) If $\phi$ satisfies $(H2)$, then the convex Sobolev inequality $(13)$ holds with $\kappa_0 = L$.

(ii) The modified log Sobolev inequality $(6)$ holds with $\kappa_1 = L + 2$.

(iii) For $\alpha \in (1, 2]$, the Beckner inequality $(7)$ holds with $\kappa_\alpha = \alpha L$.

**Remark 5.1.** Poincaré inequalities and MLSI for the Bernoulli Laplace model have been extensively studied, see [17, 18, 6]. The estimate ons $\kappa_1$ and $\kappa_2$ given by Theorem 5.1 match the best known results [8, 15, 10]. Beckner inequalities have been studied in [6] and [23]. Our constant agrees with the one found in [23]. In there, the more general case of non-homogeneous rates is treated as well. Arguably, our method also works in this case but we leave it to future work to verify this. For general functions $\phi$ satisfying $(H2)$ the convex Sobolev inequality obtained at Theorem 5.1 seems to be new.

**Proof.** Let $(\eta, \sigma_{ij}) \in S$, i.e. $\eta_i = 1, \eta_j = 0$. We define

$$e^{\phi}(\eta, \sigma_{ij} \eta, \gamma, \tilde{\gamma}) = \begin{cases} \min\{c(\eta, \gamma), c(\sigma_{ij} \eta, \gamma)\}, & \text{if } \gamma = \tilde{\gamma} \in G, \\
1, & \text{if } \gamma = \sigma_{ij}, \tilde{\gamma} = e \text{ or } e = \gamma, \tilde{\gamma} = \sigma_{ji}, \\
(1 - \eta_i), & \text{if } \gamma = \sigma_{ij}, \tilde{\gamma} = \sigma_{ji}, l \notin \{i, j\}, \\
\eta_i, & \text{if } \gamma = \sigma_{kj}, \tilde{\gamma} = \sigma_{kl}, k \notin \{i, j\}, \\
0, & \text{otherwise.} \end{cases}$$ (78)

It can be verified with a direct calculation that $(78)$ defines admissible coupling rates. Let $\phi$ satisfy $(H2)$ and $f > 0$. In view of $(20)$, the left hand side of $(20)$ can be written as $\frac{1}{2}(A + B + C + D)$ with

$$A = \sum_{\eta_{(i,j)}, k, l} c(\eta, \sigma_{ij}) \min\{c(\eta, \sigma_{kl}), c(\sigma_{ij} \eta, \sigma_{kl})\} \nabla_{\sigma_{ij}, \sigma_{kl}} f^\phi(\eta, \sigma_{ij} \eta) m(\eta),$$

$$B = \sum_{\eta_{(i,j)}} c(\eta, \sigma_{ij}) \left[ \nabla_{\sigma_{ij}} e f^\phi(\eta, \sigma_{ij} \eta) + \nabla_{e, \sigma_{ij}} f^\phi(\eta, \sigma_{ij} \eta) \right] m(\eta),$$

$$C = \sum_{\eta_{(i,j)}} c(\eta, \sigma_{ij}) \left[ \sum_{k \neq i, j} (1 - \eta_k) \nabla_{\sigma_{ij}, \sigma_{kl}} f^\phi(\eta, \sigma_{ij} \eta) \right] m(\eta),$$

$$D = \sum_{\eta_{(i,j)}} c(\eta, \sigma_{ij}) \left[ \sum_{k \neq i, j} \eta_k \nabla_{\sigma_{ij}, \sigma_{kl}} f^\phi(\eta, \sigma_{ij} \eta) \right] m(\eta).$$
We show that $A = 0$. To do this, we first observe that
\[
c(\eta, \sigma_{ij}) \min\{c(\eta, \sigma_{kl}), c(\sigma_{ij} \eta, \sigma_{kl})\} = \begin{cases} 1, & \text{if } i \neq k, \eta_k = \eta_h = 1, \eta_j = \eta_l = 0, j \neq l, \\ 0, & \text{otherwise.} \end{cases}
\]
Therefore, $c(\eta, \sigma_{ij}) \min\{c(\eta, \sigma_{kl}), c(\sigma_{ij} \eta, \sigma_{kl})\} = c(\eta, \sigma_{kl}) \min\{c(\eta, \sigma_{ij}), c(\sigma_{kl} \eta, \sigma_{ij})\}$ and we can rewrite $A$ as
\[
\sum_{\eta, (k,l)} c(\eta, \sigma_{kl}) F(\eta, \sigma_{kl}) \mathbf{m}(\eta)
\]
with
\[
F(\eta, \sigma_{kl}) = \sum_{(i,j)} \min\{c(\eta, \sigma_{ij}), c(\sigma_{kl} \eta, \sigma_{ij})\} \nabla_{\sigma_{kl}, \sigma_{ij}} f^\phi(\eta, \gamma_{ij} \eta).
\]
Using (10) on this last expression we then obtain $A = -A$, whence $A = 0$. Using that $f^\phi$ vanishes on the diagonal and the fact that any $\eta \in \Omega$ has $N$ particles occupying $L$ sites and that $(\eta, \sigma_{ij}) \in S$ implies $\eta_i = 1, \eta_j = 0$:
\[
\begin{align*}
B &= -4\mathcal{E}(\phi'(f), f), \\
C &= -2(L - N - 1)\mathcal{E}(\phi'(f), f), \\
D &= -2(N - 1)\mathcal{E}(\phi'(f), f).
\end{align*}
\]
Therefore (20) holds with $\kappa' = L$. Moreover, from the construction of coupling rates we get that (21) holds with $\kappa'' = 1$ and (22) holds with $\kappa''' = L$. The conclusion follows from Proposition 2.1.

5.2 The hardcore model

Consider a finite undirected graph $(V, E)$ that is also simple ($(x, x) \notin E$) and connected. As usual, if $(x, y) \in E$ we write $x \sim y$ and say that $x, y$ are neighbors. The state space of the classical hardcore model is
\[
\Omega = \{\eta : V \to \{0, 1\} \text{ s.t. } \eta_x \eta_y = 0, \forall x \sim y\}.
\]
For $x \in V$ we define its neighborhood as $N_x = \{y \neq x : y \sim x\}$ and we set $\bar{N}_x = N_x \cup \{x\}$. A configuration $\eta \in \Omega$ is such that if a site $x$ is occupied, then all sites in its neighborhood are empty. For any $x \in V$ we define $\delta_x \in \Omega$ as
\[
(\delta_x)_y = \begin{cases} 1, & \text{if } y = x \\
0, & \text{otherwise.} \end{cases}
\]
The set of moves is $G = \{\gamma^+_x, \gamma^-_x : x \in V\}$, where
\[
\gamma^+_x(\eta) = \begin{cases} \eta + \delta_x, & \text{if } \eta + \delta_x \in \Omega \\
\eta, & \text{otherwise} \end{cases}, \quad \gamma^-_x(\eta) = \begin{cases} \eta - \delta_x, & \text{if } \eta + \delta_x \in \Omega \\
\eta, & \text{otherwise} \end{cases}.
\]
The generator is given by
\[
\mathcal{L} f(\eta) = \sum_{x \in V} c(\eta, \gamma^-_x) \nabla^-_x f(\eta) + c(\eta, \gamma^+_x) \nabla^+_x f(\eta) \tag{79}
\]
where
\[
c(\eta, \gamma^+_x) = \eta_x, \quad c(\eta, \gamma^-_x) = \rho \prod_{y \in \bar{N}_x} (1 - \eta_y).
\]
for some constant $\rho > 0$. The above means that a new particle arrives at rate $\rho$ on an empty site $x \in V$ the neighborhood of $x$ is empty. Each occupied site $x \in V$ is emptied at rate 1. It is clear that (H0) and (H1) hold. In particular, for all $x \in V$ we have $(\gamma^-_x)^{-1} = \gamma^+_x$ and the reversible measure for the hardcore model is known to be (see [9] for example)
\[
\pi(\eta) = \frac{1}{Z} \mathbf{1}_{\eta \in \Omega} \prod_{x \in V} \rho^{\eta_x},
\]
where $Z$ is the normalization.
Theorem 5.2. Let $\Delta$ be the maximum degree of $(V, E)$, $\Delta = \sup_{x \in V} |N_x|$. Assume that

\[ \rho \Delta \leq 1 \]  

and set

\[ \kappa = 1 - \rho (\Delta - 1), \quad \bar{\kappa} = \min\{\rho, (1 - \rho \Delta)\} \]  

Then the following hold

(i) If $\phi$ satisfies (H2), then the convex Sobolev inequality (13) holds with $\kappa_\phi = \kappa$.

(ii) The modified log Sobolev inequality (6) holds with $\kappa_1 = \kappa + 2 \bar{\kappa}$.

(iii) For $\alpha \in (1, 2]$, the Beckner inequality (7) holds with $\kappa_\alpha = \alpha \kappa$.

The hardcore model and its generalizations have been intensively studied, see the discussion in [24, Sec. 22.4] of the book by Levin, Peres and Wilmer. Mixing times have been studied in [25], [11], [33] among others.

Remark 5.2. The best estimates for the MLSI of the hardcore model have been obtained in [26, 13]. For instance, in [13, Cor 4.8] MLSI is shown to hold with constant $1 - \rho (\Delta - 1)$ under assumption (80). Therefore (5.2) improves on this result. We are not aware of previously known results about Beckner inequalities or general convex Sobolev inequalities for the hardcore model. Note that in the two above mentioned references a more general version of the hardcore model is considered. We leave it to future work to see whether the methods of this paper yield interesting results for the general model.

Proof. The proof is done constructing coupling rates such that the assumptions of Proposition 2.1 hold with $\kappa' = \kappa'' = \kappa$, $\kappa'' = \bar{\kappa}$. Consider a pair $(\eta, \gamma_x) \in S$. This means that $x$ and all sites in the neighborhood of $x$ are empty, i.e. $\eta |_{N_x} \equiv 0$. We then define

\[ c^{pl}(\eta, \gamma_x, \eta, \gamma) = \begin{cases} \min\{c(\eta, \gamma), c(\gamma_x^\eta, \gamma)\}, & \text{if } \gamma = \bar{\gamma} \in G, \\ \rho, & \text{if } \gamma = \gamma_x^\eta, \bar{\gamma} = \gamma_x^\eta \text{ with } y \sim x, \eta |_{N_y} \equiv 0, \\ \rho, & \text{if } \gamma = \gamma_x^\eta, \bar{\gamma} = e, \\ 1 - \rho \{y : y \sim x, \eta |_{N_y} \equiv 0\}, & \text{if } \gamma = e, \bar{\gamma} = \gamma_x^\eta, \\ 0, & \text{otherwise.} \end{cases} \]  

(82)

If $(\eta, \gamma_x) \in S$, then $(\gamma_x^\eta, \gamma_x^{-\eta}) \in S$ as well and using the former definition we set

\[ \forall \gamma, \bar{\gamma} \in G^*, \quad c^{pl}(\eta, \gamma_x^\eta, \eta, \bar{\gamma}) = c^{pl}(\gamma_x^{-\eta}, \gamma_x^\eta, \gamma_x^{-\eta} \eta, \bar{\gamma}, \gamma) = c^{pl}(\gamma_x^{-\eta} \eta, \eta, \bar{\gamma}, \gamma). \]  

(83)

Thanks to (80) we have that $c^{pl}(\eta, \gamma_x^\eta, \eta, \bar{\gamma})$ is always non negative. It can verified with a direct calculation that (82), (83) define admissible coupling rates. In order to do so, observe that $c(\eta, \gamma_x^\eta) = c(\gamma_x^\eta, \eta)$ as soon as $y \not\in N_x$ and that $c(\eta, \gamma_x^{-\eta}) = c(\gamma_x^{-\eta}, \eta)$ for all $y \in V$ except for $y = x$ where $c(\eta, \gamma_x^{-\eta}) = 0$ and $c(\gamma_x^{-\eta}, \gamma_x^{-\eta}) = 1$.

The next step is to prove that (29) holds with $\kappa' = \kappa$. We have to show that for all $f > 0$ and $\phi$ satisfying (H2)

\[ \frac{1}{2} \sum_{(\eta, \gamma_x) \in S} \sum_{(\gamma, \gamma_x) \in G^*} c(\eta, \gamma_x^\eta) c^{pl}(\eta, \gamma_x^\eta, \eta, \bar{\gamma}) \nabla_{\gamma, \bar{\gamma}} f^\phi(\eta, \gamma_x^\eta) m(\eta) \]

\[ + \frac{1}{2} \sum_{(\eta, \gamma_x) \in S} \sum_{(\gamma, \gamma_x) \in G^*} c(\eta, \gamma_x^{-\eta}) c^{pl}(\gamma_x^{-\eta}, \eta, \bar{\gamma}, \gamma) \nabla_{\gamma, \bar{\gamma}} f^\phi(\eta, \gamma_x^{-\eta}) m(\eta) \]  

(84)

is bounded above by $-\kappa \mathcal{E}(\phi'(f), f)$. Using reversibility (10) and the fact that $f^\phi$ is symmetric in its arguments, we get that the second summand in (84) equals the first one. Combining this with

\[ \forall \eta, x, y, \quad c(\eta, \gamma_x^\eta) \leq c(\gamma_x^\eta, \eta, \gamma_x^\eta), \quad c(\eta, \gamma_x^{-\eta}) \geq c(\gamma_x^{-\eta}, \eta, \gamma_x^{-\eta}), \]  

we can rewrite (84) as $A + B + C$, where
For zero range dynamics on the complete graph with $5.3$ vertices, the state space is $L$, and that the first term in $A$ rewrites as

$$A = \sum_{\eta \in \Omega} c(\eta, \gamma_x^+) c(\eta, \gamma_y^-) \nabla_{\gamma_x^-, \gamma_y^+} f^0(\eta, \gamma_x^+ \eta)m(\eta)$$

$$+ \sum_{\eta \in \Omega} c(\eta, \gamma_x^+ \eta, \gamma_y^+) \nabla_{\gamma_x^+, \gamma_y^+} f^0(\eta, \gamma_x^+ \eta)m(\eta)$$

$$B = \sum_{\eta, x, y \in V} c(\eta, \gamma_x^+) \rho \nabla_{\gamma_x^+, \gamma_y^-} f^0(\eta, \gamma_x^+ \eta)m(\eta)$$

$$C = \sum_{\eta \in \Omega} c(\eta, \gamma_x^+) \left[ (1 - \rho) \left\{ y : y \sim x, \eta |_{\mathcal{N}_y} \equiv 0 \right\} \right] \nabla_{\gamma_x^+, \gamma_y^+} f^0(\eta, \gamma_x^+ \eta) \Phi(\eta, \gamma_x^+ \eta) m(\eta).$$

Define the set $G^- = \{ \gamma_y^- : y \in V \}$ and observe that the first term in $A$ rewrites as

$$\sum_{\eta \in \Omega} c(\eta, \gamma) F(\eta, \gamma)m(\eta), \quad F(\eta, \gamma) = 1_{\{\gamma \in G^-\}} \sum_{x \in V} c(\eta, \gamma_x^+) \nabla_{\gamma_x^+, \gamma_y^+} f^0(\eta, \gamma_x^+ \eta)m(\eta).$$

Therefore, reversibility \[10\] yields the equivalent from

$$\sum_{\eta \in \Omega} c(\eta, \gamma_x^+) c(\gamma_y^+ \eta, \gamma_x^+ \eta) \nabla_{\gamma_y^- \gamma_x^+ \eta} f^0(\gamma_y^+ \eta, \gamma_y^+ \eta)m(\eta).$$

(85)

Observing that

$$c(\eta, \gamma_y^+) c(\gamma_y^+ \eta, \gamma_x^+) = c(\eta, \gamma_x^+ \eta, \gamma_y^-) = \begin{cases} \rho^2, & \text{if } x \sim y, \eta |_{\mathcal{N}_x \cup \mathcal{N}_y} \equiv 0 \\ 0, & \text{otherwise}, \end{cases}$$

and that

$$x \sim y, \eta |_{\mathcal{N}_x \cup \mathcal{N}_y} \equiv 0 \Rightarrow \gamma_y^+ \gamma_x^+ \eta = \gamma_x^+ \gamma_y^+ \eta = \eta + \delta_x + \delta_y,$$

we obtain that \[85\] equals

$$- \sum_{\eta \in \Omega} c(\eta, \gamma_x^+) c(\gamma_y^+ \eta, \gamma_x^+ \eta) \nabla_{\gamma_y^- \gamma_x^+ \eta} f^0(\eta, \gamma_x^+ \eta)m(\eta),$$

from which $A = 0$ follows. Next, we consider $B$. Recalling that $(\eta, \gamma_x^+) \in S$ if and only if $\eta |_{\mathcal{N}_x} \equiv 0$ we can rewrite $B$ in a symmetric way exchanging the labels $x$ and $y$

$$2B = \rho \sum_{\eta \in \Omega} c(\eta, \gamma_x^+) [c(\eta, \gamma_x^+) - c(\eta, \gamma_y^+)] [f^0(\eta, \gamma_y^+ \eta) - f^0(\eta, \gamma_x^+ \eta)].$$

This last expression is worth $0$ since $\eta |_{\mathcal{N}_x \cup \mathcal{N}_y} \equiv 0$ implies $c(\eta, \gamma_x^+) = c(\eta, \gamma_y^+) = \rho$. Using the fact that $f^0$ vanishes on the diagonal and the definition of $\Delta$ we get that

$$C \leq - (1 - \rho (\Delta - 1)) \sum_{\eta \in \Omega} c(\eta, \gamma_x^+) \Phi(\eta, \gamma_x^+ \eta)m(\eta)$$

$$= -(1 - \rho (\Delta - 1)) \mathcal{E}(\phi(f), f),$$

where to get the last equality we used the reversibility property \[10\]. Thus, we have proven that \[20\] holds with $\kappa' = \kappa$. From the choice of coupling rates we made, it is easily seen that \[21\] holds with $\kappa'' = \kappa$ and \[22\] holds with $\kappa''' = \kappa$. Proposition \[2.1\] gives the desired conclusion.

### 5.3 Zero-range dynamics

For zero range dynamics on the complete graph with $L$ vertices, the state space is

$$\Omega = \left\{ \eta : \{1, \ldots, L\} \to \text{N.s.t. } \sum_{x=1}^L \eta_x = N \right\},$$
where $N$ is the number of particles $\eta_x = k$ means that there are $k$ particles at site $x$. As before, for $x \in \{1, \ldots, L\}$ we define $\delta_x \in \Omega$ as

$$(\delta_x)_y = \begin{cases} 1, & \text{if } y = x \\ 0, & \text{otherwise.} \end{cases}$$

The set of moves $G$ is $\{x y, x, y \in \{1, \ldots, L\}\}$, the move $\gamma_{x y}$ being defined by

$$\gamma_{x y} \eta = \begin{cases} \eta + \delta_y - \delta_x, & \text{if } \eta_x > 0 \\ \eta, & \text{otherwise.} \end{cases}$$

The move $\gamma_{x y}$ transfers a particle from site $x$ to site $y$. For any site $x$ we consider a non negative function $c_x : \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that $c_x(0) = 0$. $c_x(k)$ is the rate at which particles are expelled from site $x$ when $k$ particles are on site. Thus, the transition rates for the zero range dynamics rates are given by

$$
\forall \eta \in \Omega, 1 \leq x, y \leq L, \quad c(\eta, \gamma_{x y}) = \frac{1}{L} c_x(\eta_x).
$$

and the generator is

$$
\mathcal{L} f(\eta) = \frac{1}{L} \sum_{x, y \leq L} c_x(\eta_x) \nabla_{\gamma_{x y}} f(\eta).
$$

Assumptions (H0) (H1) are easily seen to be satisfied. In particular, $(\gamma_{x y})^{-1} = \gamma_{y x}$ and the reversible measure is given by (see [8])

$$
m(\eta) = \frac{1}{Z} \prod_{x, \eta_x > 0} \prod_{k=1}^{\eta_x} \frac{1}{c_x(k)} , \quad \eta \in \Omega.
$$

where $Z$ is the normalization.

**Theorem 5.3.** Assume that there exist $c, \delta \geq 0$ such that

$$\delta \leq c, \quad c \leq c_x(k+1) - c_x(k) \leq c + \delta, \quad \forall k \in \mathbb{N}, x \in V.
$$

Then for the zero range dynamics the following holds

(i) If $\phi$ satisfies (H2) then the convex Sobolev inequality (5) holds with $\kappa_0 = c - \delta$. In particular, MLSI holds with $\kappa_1 = c - \delta$.

(ii) For $a \in (1, 2]$, the Beckner inequality (7) holds with $\kappa_a = ac - \delta$.

To prove this result we modify slightly the pattern followed so far: instead of using Corollary 2.1 we take one step back and use the identity [16] from Lemma 2.2 as a starting point.

**Remark 5.3.** MLSI has been obtained in [8] with the same estimate for $\kappa_1$ as the one given by Theorem 5.3. In the recent preprint [20] MLSI is established removing the assumption that $c \leq \delta$ and only assuming $\delta < +\infty$. This was done using the so called martingale method. As it is remarked in [8], the entropy may fail to be convex if no relation is imposed between $c$ and $\delta$ and the estimates $\kappa_1 \geq c - \delta$ seems to be the best one obtained so far using the Bakry-Emery approach. Lower bounds for the entropic Ricci curvature of the zero range dynamics have been obtained in [7]. Concerning Beckner’s inequalities with $a \in (1, 2)$, the only results we are aware of are those in [25] on which Theorem 5.3 improves. Indeed the constant obtained there is $ac - (3 + 2^{a-2} - a)\delta$. Finally, Theorem 5.3 establishes the exponential decay of $\phi$-entropies for of arbitrary functions $\phi$ satisfying (H2), whereas the results of [25] cover $\phi = \phi_\alpha$. Concerning the spectral gap ($\alpha = 2$) better estimates are known. In particular under the same assumptions of Theorem 5.3 the bound $\kappa_2 \geq 2\Delta$ is found in [27]. A slight modification of the proof of Theorem 5.3 can be used to obtain the same lower bound.

**Proof.** We first prove (i). Let $(\eta, \gamma_{x y}) \in S$, i.e. $\eta_x > 0$. Define

$$
\mathcal{E}^{\eta}(\eta, \gamma_{x y} \eta, \gamma, \gamma) = \begin{cases} \frac{1}{L} \min \{ c_z(\eta_z), c_z(\gamma_{x y} \eta) \}, & \text{if } \gamma = \gamma_x = \gamma_{xw}, \\
\frac{1}{L} [ c_z(\eta_x) - c_x(\eta_x - 1) - c], & \text{if } \gamma = \gamma_{xw}, \gamma = e, \\
\frac{1}{L} [ c_y(\eta_y) - c_y(\eta_y - 1) - c], & \text{if } \gamma = e, \gamma = \gamma_{yw}, \\
\frac{1}{L} c, & \text{if } \gamma = \gamma_{xw}, \gamma = \gamma_{yw}, \\
0, & \text{otherwise.} \end{cases}
$$

(88)

It can be verified with a direct calculation that (36) defines admissible coupling rates. In order to do this, we observe that since $c_z(\cdot)$ is an increasing function.
\[
\min\{c_x(\eta_x), c_z(\gamma_x \eta_x)\} = \begin{cases} 
c_z(\eta_x - 1), & \text{if } \eta_x > 0 \text{ and } x = z, y \neq z 
c_x(\eta_x), & \text{otherwise.}
\end{cases}
\]

Consider now \(\phi\) satisfying (H2) \(f > 0\) and the expression for \(\frac{d}{dt} 2\mathcal{E}(\phi'(f_t), f_t)\) provided by (16). Using the convexity of \(\Phi\) we get

\[
\sum_{(\eta, \gamma_x \eta_x, \gamma_z \gamma_z) \in S} c(\eta, \gamma_x \eta_x) \mathcal{C}(\eta, \gamma_x \eta_x, \gamma_z \gamma_z, \gamma_z \gamma_z) \partial \Phi(f(\eta), f(\gamma_x \eta_x)) \cdot \begin{bmatrix}
\nabla_{\gamma_x \gamma_z} f(\eta) \\
\nabla_{\gamma_x \gamma_z} f(\gamma_x \eta_x)
\end{bmatrix} m(\eta)
\]

\[
\leq \sum_{(\eta, \gamma_x \eta_x) \in S} c(\eta, \gamma_x \eta_x) \mathcal{C}(\eta, \gamma_x \eta_x, \gamma_z \gamma_z, \gamma_z \gamma_z) \partial \Phi(f(\eta), f(\gamma_x \eta_x)) \cdot \begin{bmatrix}
\nabla_{\gamma_x \gamma_z} f(\eta) \\
\nabla_{\gamma_x \gamma_z} f(\gamma_x \eta_x)
\end{bmatrix} m(\eta) := A
\]

Using again the convexity of \(\Phi\) and the fact that \(\Phi\) vanishes on the diagonal we obtain that

\[
\sum_{(\eta, \gamma_x \eta_x) \in S, \gamma_z \leq L} c(\eta, \gamma_x \eta_x) \mathcal{C}(\eta, \gamma_x \eta_x, \gamma_z \gamma_z, \gamma_z \gamma_z) \partial \Phi(f(\eta), f(\gamma_x \eta_x)) \cdot \begin{bmatrix}
\nabla_{\gamma_x \gamma_z} f(\eta) \\
\nabla_{\gamma_x \gamma_z} f(\gamma_x \eta_x)
\end{bmatrix} m(\eta)
\]

\[
\leq -c \sum_{(\eta, \gamma_x \eta_x) \in S} c(\eta, \gamma_x \eta_x) \mathcal{C}(\eta, \gamma_x \eta_x, \gamma_z \gamma_z, \gamma_z \gamma_z) \partial \Phi(f(\eta), f(\gamma_x \eta_x)) \cdot \begin{bmatrix}
\nabla_{\gamma_x \gamma_z} f(\eta) \\
\nabla_{\gamma_x \gamma_z} f(\gamma_x \eta_x)
\end{bmatrix} m(\eta) = -2c\mathcal{E}(\phi'(f), f).
\]

Plugging (89) and (90) back into (16) we obtain, in view of our choice of coupling rates (88),

\[
\frac{d}{dt} 2\mathcal{E}(\phi'(f_t), f_t) \bigg|_{t=0} \leq -2c\mathcal{E}(\phi'(f), f) + A + B + C,
\]

where

\[
A = \frac{1}{L^2} \sum_{\eta, \gamma_x \eta_x, \gamma_z \gamma_z} c_x(\eta_x) \min\{c_x(\eta_x), c_z(\gamma_x \eta_x)\} \nabla_{\gamma_x \gamma_z, \gamma_z \gamma_z} f(\eta, \gamma_x \eta_x) m(\eta)
\]

\[
B = \frac{1}{L^2} \sum_{\eta, \gamma_x \eta_x, \gamma_z \gamma_z} c_x(\eta_x) (c_x(\eta_x) - c_x(\eta_x - 1) - c) \partial \Phi(f(\eta), f(\gamma_x \eta_x)) \nabla_{\gamma_x \gamma_z, \gamma_z \gamma_z} f(\eta) m(\eta)
\]

\[
C = \frac{1}{L^2} \sum_{\eta, \gamma_x \eta_x, \gamma_z \gamma_z} c_x(\eta_x) (c_x(\eta_x + 1) - c_x(\eta_x) - c) \partial \Phi(f(\eta), f(\gamma_x \eta_x)) \nabla_{\gamma_x \gamma_z, \gamma_z \gamma_z} f(\eta) m(\eta)
\]

We now move on to show that \(A = 0, B = C\) and \(2B \leq 2\mathcal{E}(\phi'(f), f)\). Once this is done, by using these relations in (16), we obtain the proof of (i). The proof that \(A = 0\) is done separately at Lemma 5.1. The other relations are proven following very closely of Lemma 4.1 in [8] which deals with the case \(\phi = \phi_1\). This proof adapts in a straightforward way to the current more general setup and we refer to it whenever the use of the reversibility relation (10) is not explained in full details in what follows. Exchanging the labels \(x, y\) and using that \(m(\eta)c_y(\eta_y) = m(\gamma_x \eta_x)c_\gamma(\gamma_x \eta_x)\) for all \(\eta\) such that \(\eta_y > 0\) we get that

\[
C = \frac{1}{L^2} \sum_{\eta, \gamma_x \eta_x, \gamma_x \gamma_x} c_x(\eta_x) (c_x(\eta_x) - c_x(\eta_x - 1) - c) \partial \Phi(f(\eta), f(\gamma_x \eta_x)) \nabla_{\gamma_x \gamma_z, \gamma_z \gamma_z} f(\gamma_x \eta_x) m(\gamma_x \eta_x)
\]

\[
= \frac{1}{L^2} \sum_{\eta, \gamma_x \eta_x, \gamma_x \gamma_x} c_x(\eta_x) (c_x(\eta_x) - c_x(\eta_x - 1) - c) \partial \Phi(f(\gamma_x \eta_x), f(\eta)) \nabla_{\gamma_x \gamma_z, \gamma_z \gamma_z} f(\eta) m(\eta).
\]

Since \(\Phi\) is symmetric in its arguments, we have \(\partial \Phi(f(\gamma_x \eta_x), f(\eta)) = \partial \Phi(f(\eta), f(\gamma_x \eta_x))\) and therefore \(C = B\). Let us now focus on \(B\). Recalling the definition of \(\Phi\) and exchanging the labels \(y, z\) and \(w, x\) we get

\[
-2B = \sum_{\eta, \gamma_x \eta_x, \gamma_x \gamma_x} c_x(\eta_x) (c_x(\eta_x) - c_x(\eta_x - 1) - c) \phi''(f(\eta)) \nabla_{\gamma_x \gamma_x, \gamma_x \gamma_x} f(\eta) m(\eta)
\]

\[
+ \sum_{\eta, \gamma_x \eta_x, \gamma_x \gamma_x} c_x(\eta_x) (c_x(\eta_x) - c_x(\eta_x - 1) - c) \phi'(f(\gamma_x \eta_x)) - \phi'(f(\eta)) \nabla_{\gamma_x \gamma_x, \gamma_x \gamma_x} f(\eta) m(\eta)
\]

32
Using the convexity of \( \phi \), we obtain that the first summand in the above expression is non negative. Concerning the second summand, rewriting \( \phi'((f(\gamma_{zw}\eta})) - \phi'(f(\eta)) = \phi'(f(\gamma_{zw}\eta)) - \phi'(f(\gamma_{zy}\eta)) + \phi'(f(\gamma_{zy}\eta)) - \phi'(f(\eta)) \) allows to isolate another non-negative term. Therefore,

\[
-L^2B \geq \sum_{\eta, z, w} c_z(\eta_z)(c_z(\eta_z) - c_z(\eta_z) - 1) - c)[\phi'(f(\gamma_{zw}\eta)) - \phi'(f(\gamma_{zy}\eta))] \nabla_{\gamma_{zw}} f(\eta) m(\eta) \\
= \sum_{\eta, z, w} c_z(\eta_z)(c_z(\eta_z) - c_z(\eta_z) - 1) - c)[\phi'(f(\gamma_{zw}\eta)) - \phi'(f(\gamma_{zy}\eta))] f(\gamma_{zy}\eta) m(\eta) := B.1.
\]

where to obtain the last inequality we used that \( \sum_y f(\eta)[\phi'(f(\gamma_{zw}\eta)) - \phi'(f(\gamma_{zy}\eta))] = 0 \). Next, we rewrite \( B.1 \) twice using reversibility. For the first one we use \( m(\eta)c_z(\eta_z) = m(\gamma_{zy}\eta)c_z(\gamma_{zy}\eta) \) to obtain

\[
B.1 = \sum_{\eta, z, w} c_z(\gamma_{zy}\eta_z)(c_z(\gamma_{zy}\eta_z) - c_z(\gamma_{zy}\eta_z) - 1) - c)[\phi'(f(\gamma_{yw}\gamma_{zw}\eta)) - \phi'(f(\gamma_{zw}\eta))] f(\gamma_{zy}\eta) m(\gamma_{zy}\eta) \\
+ \sum_{\eta, z, w} c_z(\eta_z)(c_z(\eta_z) - c_z(\eta_z) - 1) - c)[\phi'(f(\gamma_{yw}\eta)) - \phi'(f(\eta))] f(\gamma_{zy}\eta) m(\eta) \\
= \sum_{\eta, z, w} c_z(\eta_z)(c_z(\eta_z) - c_z(\eta_z) - 1) - c)[\phi'(f(\gamma_{yw}\eta)) - \phi'(f(\eta))] f(\gamma_{zy}\eta) m(\eta) \\
+ \sum_{\eta, z, w} c_z(\eta_z)(c_z(\eta_z) - c_z(\eta_z) - 1) - c)[\phi'(f(\gamma_{yw}\eta)) - \phi'(f(\eta))] f(\gamma_{zy}\eta) m(\eta). \tag{92}
\]

If we use \( m(\eta)c_z(\eta_z) = m(\gamma_{zw}\eta_w)c_w(\gamma_{zw}\eta_w) \), then we get

\[
B.1 = \sum_{\eta, z, w} c_w(\gamma_{zw}\eta_w)(c_w(\gamma_{zw}\eta_w) - c_w(\gamma_{zw}\eta_w) - 1) - c)[\phi'(f(\gamma_{yw}\gamma_{zw}\eta)) - \phi'(f(\gamma_{yw}\gamma_{zw}\eta))] f(\gamma_{zw}\eta) m(\gamma_{zw}\eta) \\
+ \sum_{\eta, z, w} c_w(\eta_w)(c_w(\eta_w) - c_w(\eta_w) - 1) - c)[\phi'(f(\eta)) - \phi'(f(\gamma_{yw}\eta))] f(\gamma_{zw}\eta) m(\eta) \\
= \sum_{\eta, z, w} c_w(\eta_w)(c_w(\eta_w) - c_w(\eta_w) - 1) - c)[\phi'(f(\eta)) - \phi'(f(\gamma_{yw}\eta))] f(\gamma_{zw}\eta) m(\eta) \\
+ \sum_{\eta, z, w} c_w(\eta_w)(c_w(\eta_w) - c_w(\eta_w) - 1) - c)[\phi'(f(\gamma_{yw}\eta)) - \phi'(f(\gamma_{zw}\eta))] f(\gamma_{zw}\eta) m(\eta). \tag{93}
\]

Exchanging the labels \( y, w \) in \( \tag{93} \), summing the result with \( \tag{92} \) and using \( \tag{87} \) yields

\[
\frac{1}{L^2} B.1 \geq -\frac{\delta}{2L} \sum_{\eta, y, w} c_y(\eta_y)[\phi'(f(\gamma_{yw}\eta)) - \phi'(f(\eta))] f(\gamma_{yw}\eta) - f(\eta) m(\eta),
\]

from which it follows that \( 2B \leq 2\delta \mathcal{E}(\phi'(f), f) \). The proof of (i) is now complete. The proof of (ii) is almost identical, the only difference being that we can replace \( \tag{90} \) with a better bound. Indeed, if instead of the convexity of \( \Phi_\alpha \), we use \( \tag{14} \), then we obtain

\[
\sum_{(\eta, \gamma_{xy}) \in S} c(\eta, \gamma_{xy}) c^{\text{pl}}(\eta, \gamma_{xy}, \gamma_{zw}, \gamma_{yw}) \cdots \left[ \left. \left. \nabla_{\gamma_{zw}} f(\gamma_{zy}\eta) \right| \right] m(\eta) \\
\leq -c_\alpha \sum_{(\eta, \gamma_{xy}) \in S} c(\eta, \gamma_{xy}) f(\phi(\eta, \gamma_{xy}\eta)) m(\eta) = -c_\alpha \mathcal{E}(\phi'(f), f).
\]

\( \square \)

**Lemma 5.1.** For any \( f > 0 \) and \( \phi : \mathbb{R} \to \mathbb{R} \) we have \( A = 0 \), where

\[
A = \frac{1}{L^2} \sum_{\eta, \gamma_{xy}, \gamma_{zw}} c_z(\eta_z) \min \{ c_z(\eta_z), c_z(\gamma_{zy}\eta_z) \} \nabla_{\gamma_{zw}, \gamma_{zy}} f(\eta, \gamma_{xy}\eta) m(\eta).
\]
Proof. We begin by observing if $x = y$ or $z = w$, then $\nabla_{(\gamma, \gamma, \gamma)} f^\alpha(\eta, \gamma x y \eta) = 0$. Therefore we can assume \(w.l.o.g.\) that $x \neq y$ and $z \neq w$ in the summation that defines $A$. Subtracting from $A$ the null term

$$\sum_{x \neq y \neq w} Lg^\alpha(\eta)\mathbf{m}(\eta), \text{ where } g^\alpha(\eta) = c_x(\eta_x)f^\alpha(\eta, \gamma x y \eta)$$

and using that for $x \neq y, z \neq w$ we have

$$c_x(\eta_x) \min\left\{c_z(\eta_z), c_z(\gamma x y \eta_z)\right\} - c_x(\eta_x)c_x(\gamma x w \eta_x) = \begin{cases} c_x(\eta_x)[c_x(\eta_x) - c_x(\eta_x + 1)], & \text{if } x = w, \eta_x > 0 \\ 0, & \text{otherwise}, \\ c_x(\eta_x)[c_x(\eta_x) - c_x(\eta_x)], & \text{if } x = z, \eta_x > 0 \\ 0, & \text{otherwise}. \end{cases}$$

we obtain that $L^2 A = A.1 - A.2$ with

$$A.1 = \sum_{\eta \neq \eta_x > 0 \atop x \neq y, z \neq w} c_x(\eta_x)[c_x(\eta_x) - c_x(\eta_x + 1)]f^\alpha(\gamma x z \eta, \gamma y z \eta)\mathbf{m}(\eta),$$

$$A.2 = \sum_{\eta \neq \eta_x > 0 \atop x \neq y, w \neq x} c_x(\eta_x)[c_x(\eta_x) - c_x(\eta_x)]f^\alpha(\eta, \gamma x z \eta)\mathbf{m}(\eta).$$

By reversibility we have that if $\eta_x > 0$ then $\mathbf{m}(\eta) c_x(\eta_x) = \mathbf{m}(\gamma x z \eta) c_x(\gamma x z \eta).$ Therefore

$$A.1 = \sum_{x \neq y \neq w \neq z \neq x} c_x(\gamma x z \eta_x)[c_x(\gamma x z \eta_x) - c_x(\gamma x z \eta_x)]f^\alpha(\gamma x z \eta, \gamma y z \eta)\mathbf{m}(\gamma x z \eta).$$

Since the map $\gamma x z$ induces a bijection between the sets $\{\eta : \eta_x > 0\}$ and $\{\eta : \eta_x > 0\}$ we get that $A.1 = A.2$, and finally that $A = 0$. \hfill \Box

6 Appendix

Proof of Lemma 2.1

Proof. We begin by showing the convexity of $\Phi_\alpha$. The case $\alpha = 1$ is well known. We do the proof for the sake of completeness. In this case $\Phi_1(a, b) = (b - a)(\log b - \log a)$ and for all $a, b > 0$ we get

$$\text{Hess}\Phi_\alpha(a, b) = \begin{pmatrix} \frac{1}{a} & \frac{b}{a^2} \\ -\frac{1}{b} & \frac{1}{b^2} \end{pmatrix},$$

which is a positive semidefinite matrix since the trace is positive and the determinant is 0.

For $\alpha \in (1, 2)$ we get from the definition of $\Phi_\alpha$ that $\Phi_\alpha(a, b) = \frac{a}{a-1} (a - b)(a^{a-1} - b^{a-1}).$ Therefore for all $a, b > 0$:

$$\text{Hess}\Phi_\alpha(a, b) = \begin{pmatrix} a^2a^{a-2} + a(2 - a)a^{a-3}b & -a(a^{a-2} + b^{a-2}) \\ -\alpha(a^{a-2} + b^{a-2}) & a^2\alpha^{-2} + a(2 - \alpha)b^{a-3}a \end{pmatrix}$$

It is easily seen that the trace of this matrix is always non negative. To conclude, we verify that so is its determinant which, after some calculations turns out to be

$$\alpha^2(-a^{a-2} - b^{a-2})^2 + a(2 - a)(ab)^{a-3}(a - b)^2$$

To show that the above expression is always non negative we assume \(w.l.o.g.\) that $b \geq a$, divide by the positive constant $\alpha^2a^{2\alpha-4}$ and set $z = a/b$. We obtain the desired conclusion if we can show that

$$\inf_{z \geq 1} -(z^{a-2} - 1)^2 + (2 - a)az^{a-3}(z - 1) \geq 0$$

Set $g(z) = -(z^{a-2} - 1)^2 + (2 - a)az^{a-3}(z - 1)^2$. A standard calculation yields

$$g'(z) = 2(2 - a)(z^{a-2} - 1)z^{a-3} + (2 - a)a(a - 3)z^{a-4}(z - 1) + 2(2 - a)az^{a-3}(z - 1)$$

Since $2 - a = 2 - a - 2z^{a-4} \geq 0$ we have $g'(z) \geq 0$ if and only if

$$2(z^{a-2} - 1)z + a(a - 3)(z - 1)^2 + 2a(z - 1)z \geq 0$$

$$\Leftrightarrow 2(z^{a-2} - 1)z + (z - 1)(a(a - 1)z + a(3 - a)) \geq 0$$
Now observe that, uniformly on \( z \geq 1 \) we have \( z^{\alpha - 2} \geq z^{-1} \). Therefore
\[
2(z^{\alpha - 2} - 1)z + (z - 1)[\alpha(z - 1)z + 3(3 - \alpha - 2)] \geq (z - 1)[\alpha(z - 1)z + (3 - \alpha - 2) - 2] \geq 0
\]
since \( (z - 1)z - 2 \geq 0 \) in \( [1, 2] \), from which we get that \( g(z) \) is increasing on \( [1, +\infty) \). Since \( \lim_{z \to 1} g(z) = 0 \), the desired conclusion follows. Let us now turn to the proof of (14). Let \( \alpha \in (1, 2] \), if \( \alpha' = b' \) we obtain after some calculations that
\[
\begin{align*}
\Phi_\alpha(a', b') - \Phi_\alpha(a, b) - D\Phi_\alpha(a, b) \cdot \left[ \frac{a'}{b'} - \frac{a}{b} \right] \\
= \alpha a^{\alpha - 2}(b - a)(a' - a) + \alpha b^{\alpha - 2}(a - b)(b' - b) \\
= \alpha a' \left( a^{\alpha - 2} - b^{\alpha - 2} \right)(a - b) + (\alpha - 1) \Phi_\alpha(a, b).
\end{align*}
\]
Using the concavity of \( x \mapsto x^{-1} \) to bound \( a^{\alpha - 2}(b - a) \) and \( b^{\alpha - 2}(a - b) \) yields (14). It remains to prove (15). A direct calculation allows gives that the right hand side of (15) is worth \((a - b)/(a + b)\). Therefore (15) is equivalent to ask that for all \( a, b > 0 \)
\[
(1/a + 1/b) \geq \frac{2(\log b - \log a)}{b - a}
\]
Assume w.l.o.g. that \( a < b \). We have
\[
2\frac{\log b - \log a}{b - a} = \frac{1}{(b - a)^2} \int_a^b \frac{1}{s} \, ds
\]
Using the convexity of \( s \mapsto s^{-1} \) we have
\[
\frac{1}{2} \int_a^b \frac{1}{s} \, ds \leq \frac{1}{2} \int_a^b \frac{1}{s} \, ds = \frac{1}{a} + \frac{1}{b}
\]
\[
\square
\]
Proof of Lemma 3.1
\[
\text{Proof. Step 1: Localization via stopping times} \quad \text{In the proof,} \quad A \text{ denotes a generic positive constant whose value may change from one expression to another. Let} \quad (X^n_t)_{t \geq 0} \text{ be a continuous time random walk with generator} \quad L \quad \text{and initial distribution} \quad \delta_0. \quad \text{We denote by} \quad T_1 \quad \text{the first jump time of the walk and define} \quad \tilde{\mu}_t \quad \text{as the law of} \quad X^n_{t \land T_1}. \quad \text{A straightforward calculation gives that}
\]
\[
\tilde{\mu}_t = \exp(-C(\eta)t)\delta_a + \sum_{\gamma \in G} \frac{1 - \exp(-C(\eta)t)}{C(\eta)}c(\eta, \gamma)\delta_{\gamma \eta}.
\]
where \( C(\eta) = \sum_{\gamma \in G} c(\eta, \gamma) \). It is easily seen that that \( W^p_p(\tilde{\mu}_t, \tilde{\mu}_t) \leq At^2 \). Therefore it suffices to show (51) replacing \( \mu_t \) with \( \tilde{\mu}_t \). Indeed, in this case we would have
\[
W^p_p(\mu_t, \tilde{\mu}_t) \leq (W^p_p(\mu_t, \tilde{\mu}_t) + W_p_p(\tilde{\mu}_t, \tilde{\mu}_t))^p \leq 2^{p-1}(W^p_p(\mu_t, \tilde{\mu}_t) + W_p_p(\tilde{\mu}_t, \tilde{\mu}_t)) \leq At^2.
\]
To this aim, we can exploit the fact that
\[
W^p_p(\mu_t, \tilde{\mu}_t) \leq \mathbb{E}[d(X^n_t, X^n_{t \land T_1})^p].
\]
(\( X^n_t \)) \( \geq 0 \) waits an exponential random time \( T_1 \) and then jumps to the state \( \gamma \eta \) with a probability proportional to \( c(\eta, \gamma) \). Therefore, using the strong Markov property we obtain
\[
\mathbb{P}[d(X^n_t, X^n_{t \land T_1}) \geq k] = \sum_{\gamma \in G} c(\eta, \gamma) \int_0^t \exp(-C(\eta)s)\mathbb{P}[d(X^n_{s}, \gamma \eta) \geq k] \, ds,
\]
where \( (X^n_s)_{s \geq 0} \) is a continuous time Markov chain with generator \( L \) started at \( \gamma \eta \).
Step 2: Bound on \( \mathbb{P}[d(X^n_{t \land T_1}, \gamma \eta) \geq k] \). Fix \( r > 0 \) and observe that
\[
\{d(X^n_{t \land T_1}, \gamma \eta) \geq k\} = \bigcup_{k_1 + \ldots + k_d = k, 1 \leq k_i \leq d} \{(X^n_{t \land T_1})_i = (\gamma \eta)_i \geq k_i, i \in I^+ \} \cap \{(\gamma \eta)_i - (X^n_{t \land T_1})_i = k_i, i \in I^- \}.
\]
For \( i \leq d \) consider the counting process \( (N^i_t)_{t \geq 0} \) defined by
\[
N^i_t - N^i_s = \begin{cases} 1, & \text{if} \quad i \in I^+, X^n_t = \gamma_i^+ X^n_{r \eta} \\ 1, & \text{if} \quad i \in I^-, X^n_t = \gamma_i^- X^n_{r \eta} \quad \text{and} \quad (X^n_{r \eta})_i \leq \gamma_i, \\ 0, & \text{otherwise} \end{cases}
\]

35
We remark that, because of (48) we have $A' < +\infty$, where

$$A' = \max\{\sup\{c(\eta, \gamma_+^i) : \eta \in \mathbb{N}^d\}, \sup\{c(\eta, \gamma_-^i) : \eta \leq \eta_0 + 1\}\}.$$ 

Moreover, observe that for any $i \in I^+$ and conditionally on $X^{\eta_0}_t$, the process $N^i$ jumps up at rate $c(X^{\eta_0}_t, \gamma_+^i) \leq A'$. If $i \in I^-$, then $N^i$ jumps up at rate $c(\eta, \gamma_-^i) \leq A'$ when $X^{\eta_0}_t \leq (\eta \gamma)$, and at rate 0 when $X^{\eta_0}_t > (\eta \gamma)$. Therefore, for any fixed $r > 0$ the random vector $(N^1, \ldots, N^d)$ is stochastically dominated by a random vector $(M^1, \ldots, M^d)$ whose components are independent Poisson random variables of parameter $rA'$. Using this, we obtain

$$\mathbb{P}[(X^{\eta}_t)_i - \gamma \eta_i \geq k, i \in I^+] \leq \mathbb{P}[N^i_{r} \geq k_i, i \in I^+] \leq \prod_{i=1}^{d} \mathbb{P}[M^i_{r} \geq k_i] \leq (\exp(-A'r) + 1)^d \prod_{i=1}^{d} \frac{(A'r)^{k_i}}{k_i!},$$

where to establish the last inequality we used some classical properties of the tails of the Poisson distribution. Using (96) and summing over all possible choices of $k_1, \ldots, k_d$ and $I^+ \subseteq \{1, \ldots, d\}$ we obtain that for all $r > 0$ and $k > 0$

$$\mathbb{P}[d(X^{\eta}_t, \gamma \eta) \geq k] \leq 2^d(1 + \exp(A'))^d \exp(-A'r)\frac{(A'r)^k}{k!}. \quad (97)$$

Step 3: conclusion Combining (97) and (95) we obtain that there exists $A > 0$ such that for all $r > 0, k > 0$

$$\mathbb{P}[d(X^{\eta}_t, X^{\eta}_{t \wedge t}) = k] \leq A^k \frac{t^{k+1}}{(k+1)!},$$

holds. Using this expression to bound side the right of (95) we obtain that, uniformly in $t \leq 1$ and $k$,

$$W^p_{\tilde{\mu}_t}(\mu_t) \leq \sum_{k \geq 1} k^p A^k \frac{t^{k+1}}{(k+1)!} \leq t^2 \sum_{k \geq 1} k^p A^k \frac{t^{k-1}}{(k+1)!} \leq t^2 A.
\square$$

Acknowledgements The author wishes to thank Paolo Dai Pra and Matthias Erbar for useful discussions.

References

[1] Anton Arnold, Peter Markowich, Giuseppe Toscani, and Andreas Unterreiter. On convex sobolev inequalities and the rate of convergence to equilibrium for fokker-planck type equations. Communications in Partial Differential Equations, 26(1-2):43–100, 2001.

[2] D. Bakry and M. Émery. Diffusions hypercontractives. In Séminaire de Probabilités XIX 1983/84, pages 177–206. Springer, 1985.

[3] Dominique Bakry, Ivan Gentil, and Michel Ledoux. Analysis and geometry of Markov diffusion operators, volume 348. Springer Science & Business Media, 2013.

[4] William Beckner. Inequalities in fourier analysis. Annals of Mathematics, pages 159–182, 1975.

[5] William Beckner. A generalized poincaré inequality for gaussian measures. Proceedings of the American Mathematical Society, pages 397–400, 1989.

[6] Sergey G Bobkov and Prasad Tetali. Modified logarithmic sobolev inequalities in discrete settings. Journal of Theoretical Probability, 19(2):289–336, 2006.

[7] Anne-Severine Boudou, Pietro Caputo, Paolo Dai Pra, and Gustavo Posta. Spectral gap estimates for interacting particle systems via a bochner-type identity. Journal of Functional Analysis, 232(1):222–258, 2006.
[8] Pietro Caputo, Paolo Dai Pra, and Gustavo Posta. Convex entropy decay via the bochner–bakry–émery approach. In *Annales de l’IHP Probabilités et statistiques*, volume 45, pages 734–753, 2009.

[9] Paolo Dai Pra and Gustavo Posta. Entropy decay for interacting systems via the bochner-bakry-émery approach. *Electron. J. Probab*, 18(52):1–21, 2013.

[10] Persi Diaconis and Mehrdad Shahshahani. Time to reach stationarity in the bernoulli–laplace diffusion model. *SIAM Journal on Mathematical Analysis*, 18(1):208–218, 1987.

[11] Martin Dyer and Catherine Greenhill. On markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.

[12] Matthias Erbar, Max Fathi, and André Schlichting. Entropic curvature and convergence to equilibrium for mean-field dynamics on discrete spaces. *arXiv preprint arXiv:1908.03397*, 2019.

[13] Matthias Erbar, Christopher Henderson, Georg Menz, and Prasad Tetali. Ricci curvature bounds for weakly interacting markov chains. *Electronic Journal of Probability*, 22, 2017.

[14] Matthias Erbar and Jan Maas. Ricci curvature of finite markov chains via convexity of the entropy. *Archive for Rational Mechanics and Analysis*, 206(3):997–1038, 2012.

[15] Matthias Erbar, Jan Maas, and Prasad Tetali. Discrete ricci curvature bounds for bernoulli-laplace and random transposition models. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 24, pages 781–800, 2015.

[16] Max Fathi, Jan Maas, et al. Entropic ricci curvature bounds for discrete interacting systems. *The Annals of Applied Probability*, 26(3):1774–1806, 2016.

[17] Fuqing Gao, Jeremy Quastel, et al. Exponential decay of entropy in the random transposition and bernoulli-laplace models. *The Annals of Applied Probability*, 13(4):1591–1600, 2003.

[18] Sharad Goel. Modified logarithmic sobolev inequalities for some models of random walk. *Stochastic processes and their applications*, 114(1):51–79, 2004.

[19] Nathael Gozlan, Cyril Roberto, Paul-Marie Samson, and Prasad Tetali. Displacement convexity of entropy and related inequalities on graphs. *Probability Theory and Related Fields*, 160(1-2):47–94, 2014.

[20] Jonathan Hermon and Justin Salez. Entropy dissipation estimates for inhomogeneous zero-range processes. *arXiv preprint arXiv:1903.01410*, 2019.

[21] Richard Holley and Daniel W Stroock. Logarithmic sobolev inequalities and stochastic ising models. 1986.

[22] Oliver Johnson et al. A discrete log-sobolev inequality under a bakry–émery type condition. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 53(4):1952–1970, 2017.

[23] Ansgar Jüngel, Wen Yue, et al. Discrete beckner inequalities via the bochner–bakry–émery approach for markov chains. *The Annals of Applied Probability*, 27(4):2238–2269, 2017.

[24] David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

[25] Michael Luby and Eric Vigoda. Fast convergence of the glauber dynamics for sampling independent sets. *Random Structures & Algorithms*, 15(3-4):229–241, 1999.

[26] Jan Maas. Gradient flows of the entropy for finite markov chains. *Journal of Functional Analysis*, 261(8):2250–2292, 2011.

[27] Katalin Marton. Logarithmic sobolev inequalities in discrete product spaces: a proof by a transportation cost distance. *arXiv preprint arXiv:1507.02803*, 2015.

[28] Alexander Mielke. Geodesic convexity of the relative entropy in reversible markov chains. *Calculus of Variations and Partial Differential Equations*, 48(1-2):1–31, 2013.

[29] Chen Mufa. Optimal markovian couplings and applications. *Acta Mathematica Sinica*, 10(3):260–275, 1994.

[30] J.R. Norris. *Markov Chains*. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, 1999.
[31] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810–864, 2009.

[32] Paul-Marie Samson. Entropic curvature on graphs along schrödinger bridges at zero temperature. *arXiv preprint arXiv:2003.05179*, 2020.

[33] Eric Vigoda. A note on the glauber dynamics for sampling independent sets. *the electronic journal of combinatorics*, pages R8–R8, 2001.

[34] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[35] Max-K von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy and ricci curvature. *Communications on pure and applied mathematics*, 58(7):923–940, 2005.

[36] Feng-Yu Wang. On estimation of the logarithmic sobolev constant and gradient estimates of heat semigroups. *Probability theory and related fields*, 108(1):87–101, 1997.