Electronic Journal of Probability

Electron. J. Probab. **22** (2017), no. 40, 1–23.
ISSN: 1083-6489 DOI: 10.1214/17-EJP49

Ricci curvature bounds for weakly interacting Markov chains

Matthias Erbar* Christopher Henderson† Georg Menz‡
Prasad Tetali§

Abstract

We establish a general perturbative method to prove entropic Ricci curvature bounds for interacting stochastic particle systems. We apply this method to obtain curvature bounds in several examples, namely: Glauber dynamics for a class of spin systems including the Ising and Curie–Weiss models, a class of hard-core models and random walks on groups induced by a conjugacy invariant set of generators.

Keywords: Ricci curvature; modified logarithmic Sobolev inequality; Markov chain; Ising model; hardcore model; random walk; Cayley graph.
AMS MSC 2010: 60K35; 60J22; 60B15.
Submitted to EJP on August 23, 2016, final version accepted on March 14, 2017.

1 Introduction

Bounds on the Ricci curvature are an essential ingredient to control the behavior of diffusion processes on Riemannian manifolds. For instance, the celebrated Bakry–Émery criterion asserts that a bound $\mathrm{Ric} + \mathrm{Hess}\, V \geq \lambda > 0$ guarantees that the drift diffusion process with generator $Lu = \Delta u - \nabla V \cdot \nabla u$ satisfies a logarithmic Sobolev inequality. The latter controls the trend to equilibrium of the associated semigroup through the exponential decay of the entropy. Furthermore, a large number of other geometric and functional inequalities can be derived from curvature bounds.

In view of this wide range of implications, considerable effort has been devoted to developing a notion of (lower bounds for the) Ricci curvature for non-smooth spaces. Bakry and Émery [1] introduced an approach based on algebraic properties of diffusion operators, the so-called Γ-calculus. A different approach based on optimal transport has

*Institute for Applied Mathematics, University of Bonn, Endenicher Allee 60, 53115 Bonn, Germany. E-mail: erbar@iam.uni-bonn.de
†Department of Mathematics, University of Chicago, 5734 S. University Avenue, Chicago, IL 60637, USA. E-mail: henderson@math.uchicago.edu
‡Mathematics Department, UCLA, Los Angeles, CA 90095, USA. E-mail: gmenz@math.ucla.edu
§School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: tetali@math.gatech.edu

been taken by Sturm [22] and Lott and Villani [14] and applies to metric measure spaces. Such a space is said to have Ricci curvature bounded below by κ, provided the relative entropy is κ-convex along geodesics in the Wasserstein space of probability measures. As in the smooth case these notions of curvature bounds entail a large number of functional inequalities.

Unfortunately, this theory does not apply to discrete spaces and Markov chains and many alternative notions of Ricci curvature bounds have been developed in this setting, see e.g. [3, 11, 21]. We will focus on the notion of *entropic Ricci curvature bounds* put forward in [8, 16] which applies to a finite Markov chain and seems particularly well suited to study functional inequalities in the discrete setting. Here the idea is to replace the role of the L^2-Wasserstein distance with a new transportation-type distance in the Lott–Sturm–Villani definition. It has been shown in [8] that, in analogy with the Bakry–Émery criterion, a strictly positive entropic Ricci bound implies a modified logarithmic Sobolev inequality (MLSI). Moreover, it entails a Poincaré inequality and an analogue of Talagrand's transport cost entropy inequality.

In view of these consequences, it is desirable to obtain entropic Ricci bounds in concrete examples of Markov chains. Relatively few results in this direction are available to date: Mielke [19] derived entropic Ricci bounds for one-dimensional birth and death chains and applied these to discretizations of Fokker–Planck equations. Erbar–Maas [8] obtained a tensorization result giving an entropic Ricci curvature bound for the product of two Markov chains in terms of Ricci bounds of the individual chains. In particular, this allows to get sharp bounds for the random walk on the hyper-cube $\{-1,1\}^n$. First results in high dimensions beyond product chains were obtained by Erbar–Maas–Tetali [9], considering the simple exclusion process on the complete graph and the random transposition shuffle models. Fathi–Maas [10] generalized the latter results by considering inhomogeneous jump rates in these models and obtained new results for the zero range process.

In this work, we present a general perturbative criterion to derive entropic Ricci curvature bounds for weakly interacting Markov chains and illustrate this method with a number of examples. Perturbation methods are well-known in the study of functional inequalities, see for instance, the Holley-Stroock criterion for the logarithmic Sobolev inequality (LSI) [12].

To formulate our main results, consider an irreducible and reversible Markov chain on a finite set $\mathcal{X}$ whose generator L can be written in the form

$$L\psi(x) = \sum_{\delta\in G} \Big(\psi(\delta x) - \psi(x)\Big) c(x,\delta) \ ,$$

where G is a collection of bijective maps $\delta : \mathcal{X} \to \mathcal{X}$ and $c : \mathcal{X} \times G \to \mathbb{R}_+$ are the transition rates. Let π denote the unique reversible probability measure on $\mathcal{X}$, i.e. π satisfies the detailed-balance condition $c(x,\delta)\pi(x) = c(\delta x, \delta^{-1})\pi(\delta x)$ for all $x \in \mathcal{X}, \delta \in G$. Then one of our main results is the following (see Theorem 3.9 below).

Theorem 1.1. *Assume that* $\delta\eta x = \eta\delta x$ *for all* $x \in \mathcal{X}, \delta, \eta \in G$ *and that*

$$\lambda \ := \min_{\substack{x\in\mathcal{X},\delta\in G\\ c(x,\delta)>0}} \left[c(x,\delta) - \mathbf{1}_{\delta\neq\delta^{-1}} c(\delta x,\delta) - \sum_{\eta:\eta\neq\delta,\delta^{-1}} \frac{(q-q_*)(\delta x,\delta^{-1},\eta)}{c(x,\delta)\pi(x)} \right] \ \geq \ 0 \ , \qquad (1.1)$$

where we set $q(x,\delta,\eta) = c(x,\delta)c(x,\eta)\pi(x)$ *and* $q_*(x,\delta,\eta) = \min\{q(x,\delta,\eta), q(\delta x,\delta^{-1},\eta),$ $q(\eta x,\delta,\eta^{-1}), q(\delta\eta x,\delta^{-1},\eta^{-1})\}$*. Then, the entropic Ricci curvature of the chain is bounded below by* 2λ*.*

That this is a perturbative criterion can be seen as follows. If the jump rates are homogeneous, in the sense that $c(\delta x,\eta) = c(x,\eta)$, for all x,δ,η, we find $\lambda \geq 0$ and recover the criterion established in [8]. This criterion was used there to prove the tensorization principle for entropic Ricci bounds. Theorem 1.1 is a generalization of this criterion when a *quantitative* bound on the deficit in the homogeneity of the rates is given. As a result, a key advantage of our results is that it gives an explicit condition on the transition rates that can be checked directly on examples.

We apply Theorem 1.1 to derive new entropic Ricci bounds for different statistical mechanics models. In particular, we consider Glauber dynamics for the Ising model on a general weighted graph and a general hard-core model. In the case of the hard-core model, we recover, in particular, the criterion derived in [5] for convex decay of the entropy and the MLSI. In the Ising case, the maps δ correspond to flipping individual spins. We show that (1.1) is satisfied for sufficiently high temperature. For the Ising model on square-lattice and the Curie–Weiss model we obtain a positive bound on the Ricci curvature that is uniform in the size of the system. We note that Ollivier [20, Ex. 17] has obtained a positive bound on his notion of coarse Ricci curvature for this chain under weaker assumptions on the temperature (in fact, down to the single-site Dobrushin condition). However, this notion of curvature is not known to imply the MLSI (2.3), for instance, among other aspects.

Finally, we develop an analogue of Theorem 1.1 for a class of Markov chains based on non-commutative maps. Namely, we consider random walks on Cayley graphs of non-abelian groups generated by a set invariant under conjugation. Prototypical examples are random walks on the symmetric group S_n generated by k-cycles, for $k \geq 2$. Our result also allows to treat *inhomogeneous* jump rates for the random walk. For a precise formulation we refer to Theorem 3.11.

Organization

In Section 2, we recall some basic facts about entropic Ricci curvature bounds for finite Markov chains. In Section 3, we introduce the new perturbative approach to proving Ricci bounds and give the proof of the main results. Finally, we apply this method to different examples in Section 4.

2 Entropic Ricci curvature bounds for Markov chains

Here we briefly recall the definitions of the discrete transport distance $\mathcal{W}$, the entropic Ricci curvature bounds and some of their consequences that we will use in this paper. The discrete transport distance (or its associated Riemannian structure) has been introduced independently in [16, 18]. The notion of entropic Ricci curvature bounds for Markov chains has been introduced and studied in [8].

2.1 Discrete transport distance and Ricci bounds

Let $\mathcal{X}$ be a finite set and let $Q : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ be a collection of transition rates. Then the operator L acting on functions $\psi : \mathcal{X} \to \mathbb{R}$ via

$$L\psi(x) = \sum_{y\in\mathcal{X}} Q(x,y)\big(\psi(y) - \psi(x)\big)$$

is the generator of a continuous time Markov chain on $\mathcal{X}$. We make the convention that $Q(x,x) = 0$ for all $x \in \mathcal{X}$. We shall assume that Q is irreducible, i.e. for all $x, y \in \mathcal{X}$, there exist points $(x_1 = x, x_2, \ldots, x_n = y)$ such that $Q(x_i, x_{i+1}) > 0$ for $i = 1, \ldots, n-1$. This implies that there exists a unique probability measure π on $\mathcal{X}$ that is stationary under the continuous time Markov chain. We shall further assume that Q is reversible w.r.t. π i.e. the detailed-balance condition holds:

$$Q(x,y)\pi(x) = Q(y,x)\pi(y) \quad \forall x, y \in \mathcal{X}\ . \tag{2.1}$$

Since π is strictly positive, we can identify the set of probability measures on $\mathcal{X}$ with the set of probability densities w.r.t. π denoted by

$$\mathcal{P}(\mathcal{X}) = \left\{ \rho \in \mathbb{R}_+^{\mathcal{X}} : \sum_x \rho(x)\pi(x) = 1 \right\} .$$

We consider a distance $\mathcal{W}$ on $\mathcal{P}(\mathcal{X})$ defined for $\rho_0, \rho_1 \in \mathcal{P}(\mathcal{X})$ by

$$\mathcal{W}(\rho_0,\rho_1)^2 := \inf_{\rho,\psi} \left\{ \frac{1}{2} \int_0^1 \sum_{x,y\in\mathcal{X}} (\psi_t(x) - \psi_t(y))^2 \theta\big(\rho_t(x), \rho_t(y)\big) Q(x,y)\pi(x) \, \mathrm{d}t \right\} ,$$

where the infimum runs over all sufficiently regular curves $\rho : [0,1] \to \mathcal{P}(\mathcal{X})$ and $\psi : [0,1] \to \mathbb{R}^{\mathcal{X}}$ satisfying the continuity equation

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\rho_t(x) + \displaystyle\sum_{y\in\mathcal{X}} (\psi_t(y) - \psi_t(x))\theta\big(\rho_t(x), \rho_t(y)\big) Q(x,y) \ = \ 0 \qquad \forall x \in \mathcal{X} \ , \\ \rho|_{t=0} = \rho_0 \ , \qquad \rho|_{t=1} = \rho_1 \ . \end{cases} \tag{2.2}$$

Here θ denotes the logarithmic mean given by

$$\theta(s,t) = \int_0^1 s^\alpha t^{1-\alpha} \, \mathrm{d}\alpha \ .$$

It has been shown in [16] that $\mathcal{W}$ defines a distance on $\mathcal{P}(\mathcal{X})$. It turns out that it is induced by a Riemannian structure on the interior $\mathcal{P}_*(\mathcal{X})$, consisting of all strictly positive probability densities. The distance $\mathcal{W}$ can be seen as a discrete analogue of the Benamou–Brenier formulation [2] of the continuous L^2-transportation cost. The appearance of the logarithmic mean is due to the fact that it allows one to obtain a discrete chain rule for the logarithm, namely $\hat{\rho}\nabla \log \rho = \nabla\rho$, where we write $\nabla\psi(x,y) = \psi(y) - \psi(x)$ and $\hat{\rho}(x,y) = \theta\big(\rho(x), \rho(y)\big)$. This replaces the usual identity $\rho\nabla \log \rho = \nabla\rho$. The distance $\mathcal{W}$ is tailor-made in this way such that the discrete heat equation $\partial_t \rho = L\rho$ is the gradient flow of the relative entropy

$$\mathcal{H}(\rho) = \sum_{x\in\mathcal{X}} \pi(x)\rho(x) \log \rho(x)$$

w.r.t. the Riemannian structure induced by $\mathcal{W}$ [16, 18]. This makes $\mathcal{W}$ a natural replacement of the L^2-Wasserstein distance in the discrete setting. Moreover, it has been proven in [8] that every pair of densities $\rho_0, \rho_1 \in \mathcal{P}(\mathcal{X})$ can be joined by a constant speed $\mathcal{W}$-geodesic $(\rho_s)_{s\in[0,1]}$. Here constant speed geodesic means that $\mathcal{W}(\rho_s, \rho_t) = |s-t|\mathcal{W}(\rho_0, \rho_1)$ for all $s, t \in [0,1]$.

In analogy with the approach of Lott–Sturm–Villani, the following definition of a Ricci curvature lower bound has been given in [8].

Definition 2.1. $(\mathcal{X}, Q, \pi)$ *has* Ricci curvature bounded from below by $\kappa \in \mathbb{R}$*, if for any constant speed geodesic* $\{\rho_t\}_{t\in[0,1]}$ *in* $(\mathcal{P}(\mathcal{X}), \mathcal{W})$*, we have*

$$\mathcal{H}(\rho_t) \le (1-t)\mathcal{H}(\rho_0) + t\mathcal{H}(\rho_1) - \frac{\kappa}{2} t(1-t)\mathcal{W}(\rho_0, \rho_1)^2 \ .$$

In this case, we write $\mathrm{Ric}(\mathcal{X}, Q, \pi) \ge \kappa$*.*

2.2 Equivalent formulation via Bochner-type inequality

Entropic curvature bounds can be expressed more explicitly in terms of an inequality resembling Bochner's inequality in Riemannian geometry. To this end, let us briefly describe the Riemannian structure induced by $\mathcal{W}$.

At each $\rho \in \mathcal{P}_*(\mathcal{X})$ the tangent space to $\mathcal{P}_*(\mathcal{X})$ is given by $\mathcal{T} = \{s \in \mathbb{R}^{\mathcal{X}} : \sum_x s(x)\pi(x) = 0\}$. Given $\psi \in \mathbb{R}^{\mathcal{X}}$ we denote by $\nabla\psi \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ the quantity $\nabla\psi(x,y) = \psi(y) - \psi(x)$, which is the discrete gradient of ψ. Fix $x_0 \in \mathcal{X}$ and let $\mathcal{G} = \{\nabla\psi : \psi \in \mathbb{R}^{\mathcal{X}}, \psi(x_0) = 0\}$ denote the set of all discrete gradient fields modulo constants. It has been shown in [16, Sec. 3] that for each $\rho \in \mathcal{P}_*(\mathcal{X})$, the map

$$K_\rho : \nabla\psi \mapsto \sum_y \nabla\psi(y,x)\hat{\rho}(x,y)Q(x,y)\,,$$

defines a linear bijection between $\mathcal{G}$ and the tangent space $\mathcal{T}$. One can then define a Riemannian metric tensor on $\mathcal{P}_*(\mathcal{X})$ by using this identification and introducing the scalar product $\langle\cdot,\cdot\rangle_\rho$ on $\mathcal{G}$ depending on ρ and given by

$$\langle\nabla\psi,\nabla\phi\rangle_\rho = \frac{1}{2}\sum_{x,y}\nabla\psi(x,y)\nabla\phi(x,y)\hat{\rho}(x,y)Q(x,y)\pi(x)\,.$$

Then $\mathcal{W}$ is the Riemannian distance associated to this Riemannian structure. We will use the notation $\mathcal{A}(\rho,\psi) := \|\nabla\psi\|_\rho^2$.

Entropic Ricci bounds, i.e. convexity of the entropy along $\mathcal{W}$-geodesics, are determined by bounds on the Hessian of the entropy $\mathcal{H}$ in the Riemannian structure defined above. An explicit expression of the Hessian at $\rho \in \mathcal{P}_*(\mathcal{X})$ is given by

$$\operatorname{Hess}\mathcal{H}(\rho)[\nabla\psi] = \frac{1}{2}\sum_{x,y}\left[\frac{1}{2}\hat{L}\rho(x,y)|\nabla\psi(x,y)|^2 - \hat{\rho}(x,y)\nabla\psi(x,y)\nabla L\psi(x,y)\right]Q(x,y)\pi(x)\,,$$

where we have used the notation

$$\hat{L}\rho(x,y) \;:=\; \partial_1\theta\big(\rho(x),\rho(y)\big)L\rho(x) + \partial_2\theta\big(\rho(x),\rho(y)\big)L\rho(y)\,.$$

Setting $\mathcal{B}(\rho,\psi) := \operatorname{Hess}\mathcal{H}(\rho)[\nabla\psi]$ for brevity, we then have the following equivalent characterization of entropic Ricci bounds.

Proposition 2.2 ([8, Thm. 4.4])**.** *A Markov triple $(\mathcal{X},Q,\pi)$ satisfies $\mathrm{Ric}(\mathcal{X},Q,\pi) \geq \kappa$ if and only if for every $\rho \in \mathcal{P}_*(\mathcal{X})$ and every $\psi \in \mathbb{R}^{\mathcal{X}}$ we have*

$$\mathcal{B}(\rho,\psi) \;\geq\; \kappa\mathcal{A}(\rho,\psi)\,.$$

Note that this statement is non-trivial since the Riemannian metric degenerates at the boundary of $\mathcal{P}(\mathcal{X})$. In view of the explicit expressions of $\mathcal{A}$ and $\mathcal{B}$, the criterion above closely resembles (an integrated version of) the classical Bochner inequality or Bakry–Émery Γ_2-criterion. Namely, a Riemannian manifold M satisfies $\mathrm{Ric} \geq \kappa$ if and only if for every pair of smooth functions $\rho,\psi : M \to \mathbb{R}$ we have:

$$\int_M \frac{1}{2}\left[L\rho|\nabla\psi|^2 - \rho\langle\nabla\psi,\nabla L\psi\rangle\right]\,\mathrm{dvol} \;\geq\; \int_M \rho|\nabla\psi|^2\,\mathrm{dvol}\,,$$

where ∇ now denotes the usual gradient and L denotes the Laplace–Beltrami operator. In fact, the left hand side equals the Hessian of the entropy in Otto's formal Riemannian structure on $\mathcal{P}(M)$ associated with the L^2-Wasserstein distance W_2.

2.3 Functional inequalities and trend to equilibrium

Entropic Ricci curvature lower bounds have many consequences in terms of functional inequalities as was shown in [8, Sec. 7]. More precisely, if a Markov triple $(\mathcal{X},Q,\pi)$ satisfies $\mathrm{Ric}(\mathcal{X},Q,\pi) \geq \kappa$ with $\kappa > 0$ then the following hold:

- a modified logarithmic Sobolev inequality MLSI(κ):

$$\mathcal{H}(\rho) \leq \frac{1}{2\kappa}\mathcal{E}(\rho, \log\rho) \quad \forall \rho \in \mathcal{P}_*(\mathcal{X}) \;, \tag{2.3}$$

- a modified Talagrand inequality $\mathrm{T}_{\mathcal{W}}(\kappa)$:

$$\mathcal{W}(\rho, \mathbf{1})^2 \leq \frac{2}{\kappa}\mathcal{H}(\rho) \quad \forall \rho \in \mathcal{P}(\mathcal{X}) \;, \tag{2.4}$$

- a Poincaré inequality P(κ):

$$\mathrm{Var}_\pi(\psi) \leq \frac{1}{\kappa}\mathcal{E}(\psi, \psi) \quad \forall \psi \;, \tag{2.5}$$

where $\mathrm{Var}_\pi(\psi) = \pi[\psi^2] - \pi[\psi]^2$ and $\mathcal{E}$ is a discrete Dirichlet form given as

$$\mathcal{E}(\psi, \phi) = \frac{1}{2}\sum_{x,y\in\mathcal{X}} \big(\psi(y) - \psi(x)\big)\big(\phi(y) - \phi(x)\big)Q(x,y)\pi(x) \;.$$

It is well known that the modified logarithmic Sobolev inequality and the Poincaré inequality govern the trend to equilibrium of the Markov semigroup $P_t = e^{tL}$. Indeed, noting that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(P_t\rho) = -\mathcal{E}(P_t\rho, \log P_t\rho) \;, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Var}(P_t\psi) = -\mathcal{E}(P_t\psi, P_t\psi) \;,$$

the Gronwall lemma together with the inequalities (2.3) and (2.5) yield the exponential convergence estimates

$$\mathcal{H}(P_t\rho) \leq e^{-2\kappa t}\mathcal{H}(\rho) \;. \qquad \mathrm{Var}(P_t\psi) \leq e^{-\kappa t}\mathrm{Var}(\psi) \;.$$

Let us make the connection to the notion of *convex entropy decay* and the Bakry–Émery approach to the MLSI developed in the discrete setting in [4, 5]. This approach is based on the following observation (see [4]):

Lemma 2.3. *Let $\kappa > 0$ and assume that the convex entropy decay inequality*

$$\sum_x \left[L\rho(x)L\log\rho(x) + \frac{(L\rho)^2}{\rho}\right]\pi(x) \geq 2\kappa\mathcal{E}(\rho, \log\rho) \tag{2.6}$$

holds for all $\rho \in \mathcal{P}_*(\mathcal{X})$. Then MLSI($\kappa$) holds.

The idea is to note that

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{H}(P_t\rho) = \sum_x \left[LP_t\rho(x)L\log P_t\rho(x) + \frac{(LP_t\rho)^2}{P_t\rho}\right]\pi(x).$$

Thus, (2.6) asserts that

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{H}(P_t\rho) \geq -2\kappa\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(P_t\rho).$$

After integration, this inequality yields $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(P_t\rho) \leq -2\kappa\mathcal{H}(P_t\rho)$, and thus MLSI($\kappa$).

Now, a direct calculation reveals that

$$\mathcal{A}(\rho, \log\rho) = \mathcal{E}(\rho, \log(\rho)) \;,$$
$$\mathcal{B}(\rho, \log\rho) = \frac{1}{2}\sum_x \left[L\rho(x)L\log\rho(x) + \frac{(L\rho)^2}{\rho}\right]\pi(x) \;.$$

Thus, we obtain that $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$ implies, in particular, the convex entropy decay inequality (2.6).

Finally, we recall that entropic Ricci bounds also imply exponential contraction in the discrete transport distance $\mathcal{W}$ [8, Prop. 4.7]. More precisely, if $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$, then for all $\rho_0, \rho_1 \in \mathcal{P}(\mathcal{X})$ we have

$$\mathcal{W}(P_t\rho_0, P_t\rho_1) \leq e^{-\kappa t}\mathcal{W}(\rho_0, \rho_1) \;.$$

3 A perturbative approach to Ricci bounds

In this section we present a general method to obtain entropic Ricci bounds for systems of weakly interacting Markov chains. The method starts from Proposition 2.2 and proceeds in two steps to establish the inequality $\mathcal{B} \geq \kappa\mathcal{A}$. The first one consists in reorganizing the $\mathcal{B}$-term, identifying non-negative contributions and giving a first lower bound by neglecting these. A general method for this, the so called Bochner-Bakry-Émery approach, was developed in [4, 5] in the study of spectral gap, MLSI and convex entropy decay and was generalized in [10] to the level of Ricci curvature. We will recall this approach in Section 3.1 and give a short simplified proof. The second step, detailed in Section 3.2, constitutes our main result and gives a final bound on $\mathcal{B}$ using the fact that the interactions are weak.

Before we proceed, we introduce a different representation of the Markov chain that will be convenient in the sequel. Let G be a set of maps from $\mathcal{X}$ to itself (called allowed moves) and consider a function $c : \mathcal{X} \times G \to \mathbb{R}_+$ (called jump rates).

Definition 3.1. *We call the pair (G, c) a* mapping representation *of Q if the following properties hold:*

1. *The generator L can be written in the form*

$$L\psi(x) \;=\; \sum_{\delta\in G} \nabla_\delta\psi(x)c(x,\delta)\;, \quad \text{where } \nabla_\delta\psi(x) = \psi(\delta x) - \psi(x)\;. \tag{3.1}$$

2. *For every $\delta \in G$ there exists a unique $\delta^{-1} \in G$ satisfying $\delta^{-1}(\delta(x)) = x$ for all x with $c(x, \delta) > 0$.*

Note that in particular, we have that $c(x, \delta) = Q(x, \delta x)$, so the detailed-balance condition (2.1) turns into

$$c(x,\delta)\pi(x) = c(\delta x, \delta^{-1})\pi(\delta x) \quad \forall x \in \mathcal{X},\ \delta \in G\;.$$

From this we infer that for every $F : \mathcal{X} \times G \to \mathbb{R}$ we have

$$\sum_{x\in\mathcal{X},\delta\in G} F(x,\delta)c(x,\delta)\pi(x) \;=\; \sum_{x\in\mathcal{X},\delta\in G} F(\delta x,\delta^{-1})c(x,\delta)\pi(x)\;. \tag{3.2}$$

Every irreducible, reversible Markov chain has a mapping representation. In fact, an explicit mapping representation can be obtained as follows. For $x, y \in \mathcal{X}$ consider the bijection $t_{\{x,y\}} : \mathcal{X} \to \mathcal{X}$ that interchanges x and y and keeps all other points fixed. Then let G be the set of all these "transpositions" and set $c(x, t_{\{x,y\}}) = Q(x, y)$ and $c(x, t_{\{y,z\}}) = 0$ for $x \notin \{y, z\}$. Then (G, c) defines a mapping representation. However, in examples it is often more natural to work with a different mapping representation involving a smaller set G.

Using a mapping representation (G, c) of Q, we can write out the quantities $\mathcal{A}$ and $\mathcal{B}$ explicitly. We obtain

$$\mathcal{A}(\rho,\psi) \;=\; \frac{1}{2}\sum_{x\in\mathcal{X},\delta\in G} \big(\nabla_\delta\psi(x)\big)^2\hat{\rho}(x,\delta x)c(x,\delta)\pi(x)\;. \tag{3.3}$$

Moreover, setting for convenience $\hat\rho_i(x,y) := \partial_i\theta(\rho(x),\rho(y))$ for $i=1,2$, we get

$$\begin{aligned}
&\mathcal{B}(\rho,\psi) \;=\\
&\frac14\sum_{x\in\mathcal{X}}\sum_{\delta,\eta\in G}\big(\nabla_\delta\psi(x)\big)^2\Big[\hat\rho_1(x,\delta x)\nabla_\eta\rho(x)c(x,\eta)+\hat\rho_2(x,\delta x)\nabla_\eta\rho(\delta x)c(\delta x,\eta)\Big]c(x,\delta)\pi(x)\\
&\qquad\qquad -2\nabla_\delta\psi(x)\Big[\nabla_\eta\psi(\delta x)c(\delta x,\eta)-\nabla_\eta\psi(x)c(x,\eta)\Big]\hat\rho(x,\delta x)c(x,\delta)\pi(x)\\
&=\;\frac12\sum_{x,\delta,\eta}\Big[|\nabla_\delta\psi|^2(x)\hat\rho_1(x,\delta x)\nabla_\eta\rho(x)+2\nabla_\delta\psi(x)\nabla_\eta\psi(x)\hat\rho(x,\delta x)\Big]c(x,\eta)c(x,\delta)\pi(x)\,. \qquad (3.4)
\end{aligned}$$

Here we have used reversibility and the fact that $\hat\rho_1(x,y)=\hat\rho_2(y,x)$ in the last equality.

Remark 3.2. It will be convenient sometimes to allow more flexibility in the mapping representation by considering a larger space $\mathcal{X}'\supset\mathcal{X}$ and a collection G' of maps from $\mathcal{X}'$ to itself. We trivially extend π by 0 to a probability measure on $\mathcal{X}'$ and similarly the rates Q to $\mathcal{X}'\times\mathcal{X}'$. G' together with a function $c':\mathcal{X}'\times G'\to\mathbb{R}_+$ will still be called a mapping representation if all the properties of Definition 3.1 hold. In particular, we have $c(x,\delta)=0$ if x or δx belongs to $\mathcal{X}'\setminus\mathcal{X}$. Obviously, for any $\rho\in\mathcal{P}(\mathcal{X}'),\psi\in\mathbb{R}^{\mathcal{X}'}$, the expressions in the right hand side of (3.3) and (3.4) calculated with the extended mapping representation (G',c') coincide with the original quantities $\mathcal{A}(\rho|_\mathcal{X},\psi|_\mathcal{X})$ and $\mathcal{B}(\rho|_\mathcal{X},\psi|_\mathcal{X})$.

3.1 The Bochner–Bakry–Émery approach to Ricci bounds

Here we briefly recall the main result of [10], a general method to identify non-negative contributions to the $\mathcal{B}$-term. For convenience, we give a short and simplified proof.

Definition 3.3. *We call a function $R:\mathcal{X}\times G\times G\to\mathbb{R}_+$ admissible for Q if (and only if)*

(i) $\delta\eta x=\eta\delta x$ *for all* x,δ,η *with* $R(x,\delta,\eta)>0$*,*

(ii) $R(x,\delta,\eta)=R(x,\eta,\delta)$ *for all* x,δ,η *with* $c(x,\delta)c(x,\eta)>0$*, and*

(iii) $R(x,\delta,\eta)=R(\delta x,\delta^{-1},\eta)$ *for all* x,δ,η*.*

Proposition 3.4 ([10, Thm. 3.5])**.** *Let R be admissible for Q and define $\Gamma:\mathcal{X}\times G\times G\to\mathbb{R}$ via $\Gamma(x,\delta,\eta)=c(x,\delta)c(x,\eta)\pi(x)-R(x,\delta,\eta)$. Then we have*

$$\mathcal{B}(\rho,\psi)\;\ge\;\sum_{x,\delta,\eta}\Gamma(x,\delta,\eta)\left[\frac12|\nabla_\delta\psi|^2(x)\hat\rho_1(x,\delta x)\nabla_\eta\rho(x)+\nabla_\delta\psi(x)\nabla_\eta\psi(x)\hat\rho(x,\delta x)\right]\,. \qquad (3.5)$$

Proof. The proof works verbatim as [8, Prop. 5.4], using the properties (i)-(iii) of Definition 3.3, instead of the conditions on c given there. Recalling (3.4) it suffices to show that

$$B:=\sum_{x,\delta,\eta}R(x,\delta,\eta)\left[\frac12|\nabla_\delta\psi|^2(x)\hat\rho_1(x,\delta x)\nabla_\eta\rho(x)+\nabla_\delta\psi(x)\nabla_\eta\psi(x)\hat\rho(x,\delta x)\right]\;\ge\;0\,. \qquad (3.6)$$

We first use (iii) to symmetrize in x and δx and obtain

$$\begin{aligned}
B=\frac12\sum_{x,\delta,\eta}R(x,\delta,\eta)\Big[&\frac12|\nabla_\delta\psi|^2(x)\big[\hat\rho_1(x,\delta x)\nabla_\eta\rho(x)+\hat\rho_2(x,\delta x)\nabla_\eta\rho(\delta x)\big]\\
&+\nabla_\delta\psi(x)\big[\nabla_\eta\psi(x)-\nabla_\eta\psi(\delta x)\big]\hat\rho(x,\delta x)\Big]\,.
\end{aligned}$$

In the first term we use the (in-)equalities (3.8) and (3.9), while in the second term we use (i) and the fact that $\nabla_\eta\psi(x) - \nabla_\eta\psi(\delta x) = \nabla_\delta\psi(x) - \nabla_\delta\psi(\eta x)$ provided $\delta\eta x = \eta\delta x$. This yields

$$B \geq \frac{1}{4}\sum_{x,\delta,\eta} R(x,\delta,\eta)\Big[|\nabla_\delta\psi|^2(x)\big[\hat\rho(\eta x,\delta\eta x) + \hat\rho(x,\delta x)\big] - 2\nabla_\delta\psi(x)\nabla_\delta\psi(\eta x)\hat\rho(x,\delta x)\Big]\ .$$

Finally, we use (iii) again to symmetrize in x and ηx, and complete the square to get

$$\begin{aligned} B &= \frac{1}{8}\sum_{x,\delta,\eta} R(x,\delta,\eta)\Big[|\nabla_\delta\psi|^2(x) + |\nabla_\delta\psi|^2(\eta x) - 2\nabla_\delta\psi(x)\nabla_\delta\psi(\eta x)\Big]\big[\hat\rho(\eta x,\delta\eta x) + \hat\rho(x,\delta x)\big] \\ &= \frac{1}{8}\sum_{x,\delta,\eta} R(x,\delta,\eta)\big|\nabla_\delta\psi(x) - \nabla_\delta\psi(\eta x)\big|^2\big[\hat\rho(\eta x,\delta\eta x) + \hat\rho(x,\delta x)\big] \geq 0\ , \end{aligned}$$

which finishes the proof. □

3.2 The perturbative criterion

Here we present our main result: a general entropic Ricci bound for weakly interacting Markov chains (see Theorems 3.9 and 3.11 below).

We start by introducing the following notation. For any $\psi \in \mathbb{R}^{\mathcal{X}}$ and $\rho \in \mathcal{P}_*(\mathcal{X})$ we write

$$B(\rho,\psi)(x,\delta,\eta) \ := \ \frac{1}{2}|\nabla_\delta\psi|^2(x)\hat\rho_1(x,\delta x)\nabla_\eta\rho(x) + \nabla_\delta\psi(x)\nabla_\eta\psi(x)\hat\rho(x,\delta x)\ . \tag{3.7}$$

We will often suppress the dependence on ρ, ψ, writing simply $B(x,\delta,\eta)$, if no confusion can arise. Note that with this notation

$$\mathcal{B}(\rho,\psi) \ = \ \sum_{x\in\mathcal{X}}\sum_{\delta,\eta\in G} B(\rho,\psi)(x,\delta,\eta)c(x,\delta)c(x,\eta)\pi(x)\ .$$

In this sum, we distinguish between two types of contributions, namely *diagonal* contributions of the form $B(\rho,\psi)(x,\delta,\delta)$ and *off-diagonal* contributions of the form $B(\rho,\psi)(x,\delta,\eta)$ with $\eta \neq \delta$. In the proof of our main result we obtain a lower bound on $\mathcal{B}$ using three ingredients. We will first show in Lemma 3.6 that the diagonal part of $\mathcal{B}$ always gives a positive contribution to curvature. Secondly, provided the interactions are sufficiently weak, expressed through a quantitative assumption on deviation of the jump rates from being homogeneous, we can use the method from the previous section and techniques developed in [9] to discard a large fraction of the off-diagonal contributions. Finally, Lemma 3.7 will allow us to estimate the remaining off-diagonal contributions against the corresponding diagonal contributions.

In the sequel we will use the following properties of the logarithmic mean, see e.g. [8, Lem. 2.2]:

Lemma 3.5. *For any $s,t,u,v > 0$ we have:*

$$u\partial_1\theta(u,v) + v\partial_2\theta(u,v) \ = \ \theta(u,v)\ , \tag{3.8}$$

$$u\partial_1\theta(s,t) + v\partial_2\theta(s,t) \ \geq \ \theta(u,v)\ . \tag{3.9}$$

We have the following bounds on the *on-diagonal* part of $\mathcal{B}$.

Lemma 3.6. *For all $\rho \in \mathcal{P}_*(\mathcal{X})$ and $\psi \in \mathbb{R}^{\mathcal{X}}$, we have that $B(\rho,\psi)(x,\delta,\delta) \geq 0$ for all $x \in \mathcal{X}$ and $\delta \in G$ and it holds:*

$$\sum_{x\in\mathcal{X},\delta\in G} B(\rho,\psi)(x,\delta,\delta)c(x,\delta)\pi(x) \ \geq \ 2\mathcal{A}(\rho,\psi)\ . \tag{3.10}$$

Let H be a subset of G such that $G = H^{-1} \cup H$. Then, we have that

$$\sum_{x\in\mathcal{X},\delta\in H} B(\rho,\psi)(x,\delta,\delta)c(x,\delta)\pi(x) \;\geq\; \frac{1}{2}\mathcal{A}(\rho,\psi)\;. \tag{3.11}$$

Proof. We first use (3.8) and (3.9) to see that

$$\begin{aligned} B(\rho,\psi)(x,\delta,\delta) \;&=\; \frac{1}{2}|\nabla_\delta\psi|^2(x)\left[\hat\rho_1(x,\delta x)\big(\rho(\delta x)-\rho(x)\big)+2\hat\rho(x,\delta x)\right] \\ &=\; \frac{1}{2}|\nabla_\delta\psi|^2(x)\left[\hat\rho_1(x,\delta x)\rho(\delta x)+\hat\rho_2(x,\delta x)\rho(\delta x)+\hat\rho(x,\delta x)\right]\;. \end{aligned} \tag{3.12}$$

Hence $\mathcal{B}(x,\delta,\delta)\geq 0$. This also yields immediately that

$$\begin{aligned} &\sum_{x\in\mathcal{X},\delta\in G} B(\rho,\psi)(x,\delta,\delta)c(x,\delta)\pi(x) \\ &\qquad =\; \mathcal{A}(\rho,\psi)+\sum_{x,\delta}\frac{1}{2}|\nabla_\delta\psi|^2(x)c(x,\delta)\pi(x)\left[\hat\rho_1(x,\delta x)\rho(\delta x)+\hat\rho_2(x,\delta x)\rho(\delta x)\right]\;. \end{aligned}$$

For the second term in the last line we use reversibility, the fact that $\partial_1\theta(s,t)=\partial_2\theta(t,s)$ and (3.9) and obtain

$$\begin{aligned} &\sum_{x,\delta}\frac{1}{2}|\nabla_\delta\psi|^2(x)c(x,\delta)\pi(x)\left[\hat\rho_1(x,\delta x)\rho(\delta x)+\hat\rho_2(x,\delta x)\rho(\delta x)\right] \\ =\;& \sum_{x,\delta}\frac{1}{4}|\nabla_\delta\psi|^2(x)c(x,\delta)\pi(x)\Big[\hat\rho_1(x,\delta x)\big(\rho(\delta x)+\rho(x)\big)+\hat\rho_2(x,\delta x)\big(\rho(\delta x)+\rho(x)\big)\Big] \\ \geq\;& \sum_{x,\delta}\frac{1}{2}|\nabla_\delta\psi|^2(x)c(x,\delta)\pi(x)\hat\rho(x,\delta x) \;=\; \mathcal{A}(\rho,\psi)\;. \end{aligned}$$

Finally, we notice that by symmetrization and reversibility, we have that

$$\mathcal{A}(\rho,\psi)\leq\sum_{\delta\in H}\sum_{x\in\mathcal{X}}|\nabla_\delta\psi(x)|^2\hat\rho(x,\delta x)c(x,\delta)\pi(x)\;,$$

which together with (3.12) immediately yields (3.11). □

We will use the following to estimate the off-diagonal contributions to $\mathcal{B}$. Similar estimates for terms appearing in the study of convex entropy decay can be found in [5, (2.33)].

Lemma 3.7. *For any $\psi\in\mathbb{R}^{\mathcal{X}}$ and $\rho\in\mathcal{P}_*(\mathcal{X})$ and $x\in\mathcal{X},\delta,\eta\in G$ we have*

$$B(x,\delta,\eta)+B(x,\eta,\delta)\;\geq\;-B(\delta x,\delta^{-1},\delta^{-1})-B(\eta x,\eta^{-1},\eta^{-1})\;. \tag{3.13}$$

Proof. Setting $a=\nabla_\delta\psi(x)$, $b=\nabla_\eta\psi(x)$ as well as $s=\rho(x)$, $t=\rho(\delta x)$ and $r=\rho(\eta x)$, it suffices to show that

$$\begin{aligned} &a^2\Big[\partial_1\theta(s,t)(r-s)+\partial_1\theta(t,s)(s-t)+2\theta(t,s)\Big]+2ab\Big[\theta(s,t)+\theta(s,r)\Big] \\ &+b^2\Big[\partial_1\theta(s,r)(t-s)+\partial_1\theta(r,s)(s-r)+2\theta(r,s)\Big]\;\geq\;0\;. \end{aligned}$$

We rewrite this last inequality as $a^2M_{11}+2abM_{12}+b^2M_{22}\geq 0$ with a symmetric 2×2 matrix M. Now, it is readily checked, using the fact that $\partial_1\theta(u,v)=\partial_2\theta(v,u)$ as well as (3.8), (3.9), that M is diagonally dominant and thus non-negative definite. □

3.2.1 Commutative mapping representations

Let $(\mathcal{X}, Q, \pi)$ be a Markov triple and assume that it has a mapping representation (G, c) that is *commutative* in the sense that

$$\delta \circ \eta = \eta \circ \delta \qquad \forall \delta, \eta \in G .$$

A first criterion for entropic Ricci bounds in this setting was given in [8].

Proposition 3.8 ([8, Prop. 5.4])**.** *Assume that*

$$c(\delta x, \eta) = c(x, \eta) \qquad \forall x, \in \mathcal{X}, \delta, \eta \in G . \tag{3.14}$$

Then, we have $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq 0$*. If moreover* $\delta = \delta^{-1}$ *holds for all* $\delta \in G$*, then we have* $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq 2c_*$*, where*

$$c_* \ := \ \min\{c(x, \delta) \ : \ x, \delta \text{ with } c(x, \delta) > 0\} \tag{3.15}$$

denotes the minimal transition rate.

By irreducibility of the Markov chain, Condition (3.14) is in fact equivalent to the transition rates being homogeneous, i.e. $c(x, \eta) = c(y, \eta)$ for all $x, y \in \mathcal{X}, \eta \in G$. Our main result, Theorem 3.9, of this section is a perturbative generalization of this criterion, when an explicit bound on the non-homogeneity of the transition rates is given.

To state the result, we use the following notation. Put $q(x, \delta, \eta) = c(x, \delta)c(x, \eta)\pi(x)$. For $\delta, \eta \in G$ with $\eta \neq \delta, \delta^{-1}$ we define

$$q_*(x, \delta, \eta) \ := \ \min\left\{q(x, \delta, \eta), q(\delta x, \delta^{-1}, \eta), q(\eta x, \delta, \eta^{-1}), q(\delta \eta x, \delta^{-1}, \eta^{-1})\right\} . \tag{3.16}$$

Theorem 3.9. *Assume that*

$$\lambda \ := \ \min_{\substack{x \in \mathcal{X}, \delta \in G \\ c(x,\delta) > 0}} \left[c(x, \delta) - \mathbf{1}_{\delta \neq \delta^{-1}} c(\delta x, \delta) - \sum_{\eta : \eta \neq \delta, \delta^{-1}} \frac{(q - q_*)(\delta x, \delta^{-1}, \eta)}{c(x, \delta)\pi(x)} \right] \ \geq \ 0 \ , \tag{3.17}$$

Then, we have $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq 2\lambda$*.*

Moreover, assume that there are disjoint subsets H_1, H_2 *of* G *such that* $H_1 \cap H_2 = \emptyset$ *and* $H_i \cup H_i^{-1} = G$ *for* $i = 1, 2$*. Set*

$$\lambda_i \ := \ \min_{\substack{x \in \mathcal{X}, \delta \in H_i \\ c(x,\delta) > 0}} \left[c(x, \delta) - \mathbf{1}_{\delta \neq \delta^{-1}} c(\delta x, \delta) - \sum_{\eta : \eta \neq \delta, \delta^{-1}} \frac{(q - q_*)(\delta x, \delta^{-1}, \eta)}{c(x, \delta)\pi(x)} \right] . \tag{3.18}$$

Then, we also have $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \frac{1}{2}(\lambda_1 + \lambda_2)$*.*

Note that we recover Proposition 3.8 as an immediate consequence: In this situation we have $q - q_* \equiv 0$ and hence $\lambda = 0$ or $2c_*$, depending on whether there is δ with $\delta \neq \delta^{-1}$ or not.

Proof. To prove the first statement, we have to show that for any ρ and ψ ,

$$\mathcal{B}(\rho, \psi) \ \geq \ 2\lambda \mathcal{A}(\rho, \psi) . \tag{3.19}$$

Define a function $R : \mathcal{X} \times G \times G \to \mathbb{R}_+$ as follows. For $\delta, \eta \in G$ with $\eta \neq \delta, \delta^{-1}$ set

$$R(x, \delta, \eta) \ = \ q_*(x, \delta, \eta) \ ,$$

and for $\delta \in G$ with $\delta \neq \delta^{-1}$ set

$$\begin{aligned} R(x,\delta,\delta^{-1}) &= q(x,\delta,\delta^{-1}) , \\ R(x,\delta,\delta) &= c(\delta x,\delta)c(\delta x,\delta^{-1})\pi(\delta x) = q(x,\delta,\delta)\frac{c(\delta x,\delta)}{c(x,\delta)} , \end{aligned}$$

and finally for $\delta \in G$ with $\delta = \delta^{-1}$ we set $R(x,\delta,\delta) = 0$. It is readily checked that R is admissible in the sense of Definition 3.3. Note that the assumption on λ guarantees, in particular, that $c(\delta x,\delta) \leq c(x,\delta)$ when $\delta \neq \delta^{-1}$. Thus, we have that $\Gamma(x,\delta,\eta) = q(x,\delta,\eta) - R(x,\delta,\eta) \geq 0$ for all x,δ,η. Note further that in the case $\delta \neq \delta^{-1}$, we have $\Gamma(x,\delta,\delta^{-1}) = 0$. Using Proposition 3.4 and Lemma 3.7 we now obtain

$$\begin{aligned} \mathcal{B}(\rho,\psi) &\geq \sum_{x\in\mathcal{X},\delta,\eta\in G} \Gamma(x,\delta,\eta)B(x,\delta,\eta) \\ &= \sum_{x,\delta} B(x,\delta,\delta)\Gamma(x,\delta,\delta) + \frac{1}{2}\sum_{x,\delta\neq\eta} \big[B(x,\delta,\eta) + B(x,\eta,\delta)\big]\Gamma(x,\delta,\eta) \\ &\geq \sum_{x,\delta} B(x,\delta,\delta)\Gamma(x,\delta,\delta) - \frac{1}{2}\sum_{x,\delta\neq\eta} \big[B(\delta x,\delta^{-1},\delta^{-1}) + B(\eta x,\eta^{-1},\eta^{-1})\big]\Gamma(x,\delta,\eta) \\ &= \sum_{x,\delta} B(x,\delta,\delta)\Gamma(x,\delta,\delta) - \sum_{x,\delta\neq\eta} B(\delta x,\delta^{-1},\delta^{-1})\Gamma(x,\delta,\eta) . \end{aligned}$$

Here we have used that $\Gamma \geq 0$. We can further reorganize this expression to obtain

$$\begin{aligned} &\mathcal{B}(\rho,\psi) \\ &\geq \sum_{x,\delta} B(x,\delta,\delta)\left[\Gamma(x,\delta,\delta) - \sum_{\eta:\eta\neq\delta^{-1}} \Gamma(\delta x,\delta^{-1},\eta)\right] \\ &= \sum_{x,\delta} B(x,\delta,\delta)\left[q(x,\delta,\delta) - \mathbf{1}_{\{\delta\neq\delta^{-1}\}}q(\delta x,\delta^{-1},\delta) - \sum_{\eta:\eta\neq\delta,\delta^{-1}} (q-q_*)(\delta x,\delta^{-1},\eta)\right] \\ &\geq \sum_{x,\delta} B(x,\delta,\delta)c(x,\delta)\pi(x)\left[c(x,\delta) - \mathbf{1}_{\{\delta\neq\delta^{-1}\}}c(\delta x,\delta) - \sum_{\eta:\eta\neq\delta,\delta^{-1}} \frac{(q-q_*)(\delta x,\delta^{-1},\eta)}{c(x,\delta)\pi(x)}\right] . \end{aligned}$$

Here we have also used in the second inequality the fact that $B(x,\delta,\delta) \geq 0$, by Lemma 3.6. Now, invoking (3.17) and (3.10) from Lemma 3.6 finishes the proof of statement i).

To obtain the second statement, we proceed in the same way. In the last step, we note that by (3.17) each summand is non-negative. Thus we obtain the estimate

$$\begin{aligned} &\mathcal{B}(\rho,\psi) \geq \\ &\sum_{x,\delta\in H_1\cup H_2} B(x,\delta,\delta)c(x,\delta)\pi(x)\left[c(x,\delta) - \mathbf{1}_{\{\delta\neq\delta^{-1}\}}c(\delta x,\delta) - \sum_{\eta:\eta\neq\delta,\delta^{-1}} \frac{(q-q_*)(\delta x,\delta^{-1},\eta)}{c(x,\delta)\pi(x)}\right] \end{aligned}$$

and we conclude by invoking (3.18) and (3.11). □

In Section 4, we will apply the first part of Theorem 3.9 to derive lower Ricci bounds for the Glauber dynamics of the Ising model. The second part of Theorem 3.9 is applied to derive lower Ricci bounds for the hard-core model.

The following corollary will illustrate that our method allows to obtain rough entropic Ricci curvature bounds under very explicit and easy-to-check conditions on the transition rates. In practice, however, a direct application of Theorem 3.9 will give sharper results.

Assume for simplicity that $\delta = \delta^{-1}$ for all $\delta \in G$ and set,

$$N := \#\Big\{\{\delta,\eta\} \subset G : c(\delta x,\eta) \neq c(x,\eta) \text{ or } c(\eta x,\delta) \neq c(x,\delta) \text{ for some } x \in \mathcal{X}\Big\} ,$$

$$\alpha := \begin{cases} +\infty , & \text{if there exist } x,\delta,\eta \text{ s.t. } c(x,\delta) = 0, c(\eta x,\delta) > 0 , \\ \max\left\{ \frac{c(\eta x,\delta)}{c(x,\delta)} : x \in \mathcal{X}, \delta,\eta \in G \right. , & \text{else} , \end{cases}$$

$$\beta := \max\left\{ \frac{c(x,\eta)}{c(x,\delta)} : x \in \mathcal{X}, \delta,\eta \in G \text{ with } c(x,\delta) > 0 \right\} .$$

Corollary 3.10. *With the above notation, assume that*

$$\varepsilon := \beta N\Big(\alpha^2 - 1\Big) \leq 1 .$$

Then, we have that $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq (1-\varepsilon)2c_*$*, for* c_* *as in* (3.15).

Proof. The result will follow from Theorem 3.9 by estimating the minimum in (3.17). First, note that if $c(x,\delta) > 0$ we have

$$\frac{q(x,\delta,\eta)}{q(x,\delta,\delta)} = \frac{c(x,\eta)}{c(x,\delta)} \leq \beta .$$

Now, if δ,η are such that $c(\delta x,\eta) = c(x,\eta)$ and $c(\eta x,\delta) = c(x,\delta)$ for all x, then, using the detailed-balance condition, we infer that $q(x,\delta,\eta) = q_*(x,\delta,\eta)$. Otherwise, we have the bound $q(x,\delta,\eta) \leq \alpha^2 q_*(x,\delta,\eta)$. Note also that by construction we have that $q_*(x,\delta,\eta) = q_*(\delta x,\delta^{-1},\eta)$. This implies that

$$\frac{(q-q_*)(\delta x,\delta^{-1},\eta)}{q(x,\delta,\delta)} \leq \beta\frac{(q-q_*)(\delta x,\delta^{-1},\eta)}{q(x,\delta,\eta)} \leq \beta\frac{(q-q_*)(\delta x,\delta^{-1},\eta)}{q_*(x,\delta,\eta)} \leq \beta\Big(\alpha^2 - 1\Big) .$$

From this we obtain that

$$\lambda = \min_{x,\delta} c(x,\delta)\Big[1 - \sum_{\eta\neq\delta}\frac{(q-q_*)(\delta x,\delta,\eta)}{q(x,\delta,\delta)}\Big] \geq c_*\Big[1 - N\beta\Big(\alpha^2-1\Big)\Big] = c_*(1-\varepsilon) . \qquad \square$$

3.2.2 Conjugacy-invariant Cayley graphs

Here we establish entropic Ricci bounds for a class of Markov chains with not necessarily commutative mapping representation. Namely, we consider random walks on weighted conjugacy-invariant Cayley graphs.

Let $\mathcal{G}$ be a finite group and let G be a set of generators for $\mathcal{G}$, i.e. every $g \in \mathcal{G}$ can be written as a word $g = \delta_1\delta_2\cdots\delta_n$ for suitable $\delta_i \in G$. We assume that G is

(i) closed under taking inverse: $\delta^{-1} \in G$ for all $\delta \in G$,

(ii) conjugacy-invariant: $g\delta g^{-1} \in G$ for all $\delta \in G, g \in \mathcal{G}$.

The Cayley graph associated to the generating set G is the (directed) graph with vertex set $\mathcal{G}$ and edge set $E = \{(x,y) : x,y \in \mathcal{G},\ yx^{-1} \in G\}$. We consider a natural irreducible Markov dynamics on the group $\mathcal{G}$ by choosing a function $c : \mathcal{G} \times G \to (0,\infty)$ and considering the mapping representation (G,c). The associated Markov triple $(\mathcal{G},Q,\pi)$ is the natural random walk on the weighted directed graph $(\mathcal{G},E)$, where $c(x,\delta)$ is considered as the weight of the edge $(x,\delta x)$.

We have the following perturbative Ricci bound in the present situation:

Theorem 3.11. *Let us set*

$$\alpha_1 := \max\left\{\frac{c(\delta x,\delta)}{c(x,\delta)} : x\in\mathcal{G},\delta\in G \text{ with } \delta\neq\delta^{-1}\right\}\,,$$
$$\alpha_2 := \max\left\{\frac{c(\delta x,\eta)}{c(x,\eta)},\frac{c(\delta x,\delta\eta\delta^{-1})}{c(x,\eta)} : x\in\mathcal{G},\delta,\eta\in G \text{ with } \delta\neq\eta,\eta^{-1}\right\}\,,$$
$$\beta := \max\left\{\frac{c(x,\eta)}{c(x,\delta)} : x\in\mathcal{G},\delta,\eta\in G\right\}\,,$$

and assume that $\varepsilon := \alpha_1+\beta(|G|-1)\Big(\alpha_2^2-1\Big)\leq 1$*. Then we have that*

$$\mathrm{Ric}(\mathcal{G},Q,\pi)\geq(1-\varepsilon)2c_*\,,$$

where c_* *is the minimal transition rate defined in* (3.15)*. Moreover, if we assume that* $\delta=\delta^{-1}$ *for all* $\delta\in G$ *and that* $\varepsilon' := \beta(|G|-1)\Big(\alpha_2^2-1\Big)\leq 1$*, we then have the improved bound* $\mathrm{Ric}(\mathcal{G},Q,\pi)\geq(1-\varepsilon')2c_*$*.*

In particular, we obtain a Ricci bound for the simple random walk on a conjugacy-invariant Cayley graph. In this case, we have for some constant $c>0$ that

$$c(x,\delta)=c\qquad\forall x\in G,\delta\in G\,.$$

Corollary 3.12. *The simple random walk on a conjugacy-invariant Cayley graph satisfies* $\mathrm{Ric}(\mathcal{G},Q,\pi)\geq 0$*. If* $\delta=\delta^{-1}$ *holds for all* $\delta\in G$*, then we even have that* $\mathrm{Ric}(\mathcal{G},Q,\pi)\geq 2c$*.*

In Section 4.3, we will apply Corollary 3.12 to analyze the curvature of some random walks on the symmetric group.

Since the mapping representation in the present situation is not commutative, the Bochner–Bakry–Émery method developed in [10] (see Prop. 3.4) does not apply immediately. Instead, we will combine it with a technique developed in [9] which consists in partitioning the $\mathcal{B}$-term into contributions coming from square subgraphs of $(\mathcal{G},E)$. We need some notation before we come to the proof of Theorem 3.11.

A *square* in the Cayley graph is a tuple $\square=(x_1,x_2,x_3,x_4)$ such that $x_{i+1}x_i^{-1}\in G$ for all $i=1,\dots,4$ with the convention that $x_5=x_1$. Two squares that differ only by a cyclic permutation will be identified. We write for short $\delta_i := x_{i+1}x_i^{-1}$. Given two maps δ and $\eta\neq\delta,\delta^{-1}$ in G, we obtain for each $x\in\mathcal{G}$ a square

$$\square(x,\delta,\eta)=(x_1=x,x_2=\delta x,x_3=\eta\delta x,x_4=\eta x)\,. \tag{3.20}$$

Indeed, by invariance of G under conjugation, we have that $x_4x_3^{-1}=\eta x(\eta\delta x)^{-1}=\eta\delta^{-1}\eta^{-1}\in G$. The other relations $x_{i+1}x_i^{-1}\in G$ for $i=1,2,4$ hold trivially. The squares obtained in this way fall into two classes depending on whether δ and η commute or not. Let S_1 be the collection of all squares obtained from commuting maps δ,η and let S_2 denote the collection of all squares obtained from non-commuting maps.

Given such a square $\square$ and two functions $\rho\in\mathcal{P}_*(\mathcal{G}),\psi\in\mathbb{R}^{\mathcal{G}}$ we set

$$\mathcal{B}^{\mathrm{diag}}_\square(\rho,\psi)=\sum_{i=1}^4 B(x_i,\delta_i,\delta_i)q(x_i,\delta_i,\delta_i)+B(x_i,\delta_{i-1}^{-1},\delta_{i-1}^{-1})q(x_i,\delta_{i-1}^{-1},\delta_{i-1}^{-1})\,,$$
$$\mathcal{B}^{\mathrm{off}}_\square(\rho,\psi)=\sum_{i=1}^4\Big[B(x_i,\delta_i,\delta_{i-1}^{-1})+B(x_i,\delta_{i-1}^{-1},\delta_i)\Big]q(x_i,\delta_i,\delta_{i-1}^{-1})\,,$$

as well as $\mathcal{B}_\square(\rho,\psi)=\mathcal{B}^{\mathrm{diag}}_\square(\rho,\psi)+\mathcal{B}^{\mathrm{off}}_\square(\rho,\psi)$. Note that $\mathcal{B}_\square(\rho,\psi)$ is the quantity $\mathcal{B}$ calculated in the square graph $\square$ with the restrictions of ρ,ψ to $\square$. We will proceed by rearranging the $\mathcal{B}$-term of the full Cayley graph into contributions from squares and applying the techniques of the previous section separately in each square.

Proof of Thm. 3.11. We have to show that

$$\mathcal{B}(\rho,\psi) \geq \Big[1-\alpha_1-\beta(|G|-1)\Big(\alpha_2^2-1\Big)\Big]\, 2c_*\mathcal{A}(\rho,\psi)$$

holds for any $\rho\in\mathcal{P}_*(\mathcal{G}),\psi\in\mathbb{R}^{\mathcal{G}}$. We drop ρ,ψ from the notation for the rest of the proof. We distinguish on- and off-diagonal contributions to $\mathcal{B}$ by writing $\mathcal{B}=\mathcal{B}^{\mathrm{diag}}+\mathcal{B}^{\mathrm{off},1}+\mathcal{B}^{\mathrm{off},2}$ with

$$\begin{aligned}
\mathcal{B}^{\mathrm{diag}} &= \sum_{x\in\mathcal{G},\delta\in G} B(x,\delta,\delta)q(x,\delta,\delta)\ ,\\
\mathcal{B}^{\mathrm{off},1} &= \sum_{x\in\mathcal{G},\delta\in G:\delta\neq\delta^{-1}} B(x,\delta,\delta^{-1})q(x,\delta,\delta^{-1})\ ,\\
\mathcal{B}^{\mathrm{off},2} &= \sum_{x\in\mathcal{G},\delta,\eta\in G:\eta\neq\delta,\delta^{-1}} B(x,\delta,\eta)q(x,\delta,\eta)\ .
\end{aligned}$$

We first estimate $\mathcal{B}^{\mathrm{off},1}$. Symmetrizing in δ,δ^{-1} and using Lemma 3.7, we obtain

$$\mathcal{B}^{\mathrm{off},1}\geq -\frac{1}{2}\sum_{x\in\mathcal{G},\delta\in G:\delta\neq\delta^{-1}} B(x,\delta,\delta)\big[q(\delta x,\delta,\delta^{-1})+q(\delta^{-1}x,\delta^{-1},\delta)\big]\geq -\alpha_1\mathcal{B}^{\mathrm{diag}}\ . \tag{3.21}$$

Now, we claim that

$$\mathcal{B}^{\mathrm{off},2} = \sum_{\square\in S_1}\mathcal{B}^{\mathrm{off}}_{\square}+\frac{1}{2}\sum_{\square\in S_2}\mathcal{B}^{\mathrm{off}}_{\square}\ , \tag{3.22}$$

$$\mathcal{B}^{\mathrm{diag}} = \frac{1}{|G|-1}\left[\sum_{\square\in S_1}\mathcal{B}^{\mathrm{diag}}_{\square}+\frac{1}{2}\sum_{\square\in S_2}\mathcal{B}^{\mathrm{diag}}_{\square}\right]\ . \tag{3.23}$$

Indeed, each term $B(x,\delta,\eta)q(x,\delta,\eta)$ appears in exactly one square from S_1, namely the square $\square(x,\delta,\eta)$ defined in (3.20), if δ and η commute. If they do not commute, then the term $B(x,\delta,\eta)q(x,\delta,\eta)$ appears in exactly two squares from S_2, namely $\square(x,\delta,\eta)$ and $\square(x,\eta,\delta)$. Moreover, each term $B(x,\delta,\delta)q(x,\delta,\delta)$ appears in exactly N_1 squares in S_1 and in exactly $2N_2$ squares in S_2 with $N_1=\#\{\eta\in G:\eta\neq\delta,\delta^{-1},\delta\eta=\eta\delta\}$ and $N_2=\#\{\eta\in\mathcal{G}:\delta\eta\neq\eta\delta\}$. Obviously, $N_1+N_2\leq|G|-1$.

To calculate the $\mathcal{B}$-terms in each square $\square$, we apply the techniques of the previous section by choosing a new mapping representation consisting of two maps $\delta,\tilde{\delta}$ that are involutive and commutative. For instance, set $\delta x_i=x_2,x_1,x_4,x_3$, $\tilde{\delta}x_i=x_4,x_3,x_2,x_1$ for $i=1,2,3,4$. Thus, following the proofs of Theorem 3.9 and Corollary 3.10, we find

$$\mathcal{B}^{\mathrm{off}}_{\square}\geq -\beta\Big(\alpha_2^2-1\Big)\mathcal{B}^{\mathrm{diag}}_{\square}\ .$$

Combing this with (3.21), (3.22), (3.23) and using Lemma 3.6 yields

$$\begin{aligned}
\mathcal{B}=\mathcal{B}^{\mathrm{diag}}+\mathcal{B}^{\mathrm{off},1}+\mathcal{B}^{\mathrm{off},2} &\geq \mathcal{B}^{\mathrm{diag}}\Big[1-\alpha_1-(|G|-1)\beta\Big(\alpha_2^2-1\Big)\Big]\\
&\geq \Big[1-\alpha_1-(|G|-1)\beta\Big(\alpha_2^2-1\Big)\Big]\,2c_*\mathcal{A}\ ,
\end{aligned}$$

which finishes the proof of the first statement. To obtain the second statement, we simply note that $\mathcal{B}^{\mathrm{off},1}=0$ if $\delta=\delta^{-1}$ for all $\delta\in G$. □

4 Examples

In this section we apply our perturbation method to derive Ricci bounds in concrete models. First, we consider a general Ising model in the high temperature regime. Then, we specialize this result to obtain bounds for the Ising model on a finite sub-lattice of $\mathbb{Z}^d$ and the Curie–Weiss model. Moreover, we consider a general hard-core model put forward in [5] and extend the results on convex entropy decay obtained there to the level of Ricci curvature.

4.1 Bounds for a general Ising model

Let $n \in \mathbb{N}$ and introduce the state space $\mathcal{X} = \{-1,1\}^n$. Let $k \in \mathbb{R}^{n\times n}$ be a symmetric matrix modeling the interaction strength between the sites. We set $k_{ii} = 0$ for all i. Then we introduce the Hamiltonian $H : \mathcal{X} \to \mathbb{R}$ via

$$H(x) \;=\; -\sum_{i,j=1}^{n} k_{ij} x_i x_j \;.$$

Note that we make no assumption on the sign of k. We consider the probability measure

$$\pi_\beta(x) \;=\; \frac{1}{Z_\beta} \exp\big(-\beta H(x)\big)\;,$$

where Z_β is a normalizing constant and $\beta \in [0,\infty)$ denotes the inverse temperature. We consider the associated Glauber dynamics, the continuous time Markov chain given by the q-matrix

$$Q_\beta(x,y) \;=\; \begin{cases} \sqrt{\frac{\pi_\beta(y)}{\pi_\beta(x)}}\;, & \text{if } \|x-y\|_{l^1} = 2\;,\\ 0\;, & \text{else.}\end{cases}$$

A natural mapping representation is given as follows. Let $G = \{\delta_i,\ i = 1,\dots,n\}$, where $\delta_i : \mathcal{X} \to \mathcal{X}$ is the map flipping the i-th coordinate, i.e. $\big(\delta_i(x)\big)_i = -x_i$ and $\big(\delta_i(x)\big)_j = x_j$ for all $j \neq i$. Then we put

$$c(x,\delta_i) \;=\; \sqrt{\frac{\pi_\beta(\delta_i x)}{\pi_\beta(x)}} \;=\; e^{-\frac{\beta}{2}\nabla_i H(x)}\;.$$

where we write for short $\nabla_i H(x) = \nabla_{\delta_i} H(x) = H(\delta_i x) - H(x)$. Note that this mapping representation is commutative and involutive, i.e. $\delta_i^{-1} = \delta_i$. We have the following Ricci bound for the Glauber dynamics of the general Ising model.

Theorem 4.1. *Assume that*

$$\varepsilon(\beta) \;:=\; \max_i \sum_{j,j\neq i} \exp\left(2\beta \sum_{m\neq i,j} |k_{im}| + |k_{jm}|\right)\left(e^{4\beta|k_{ij}|} - 1\right) \;\le\; 1\;. \tag{4.1}$$

Then the Glauber dynamics satisfies

$$\mathrm{Ric}(\mathcal{X}, Q_\beta, \pi_\beta) \;\ge\; \big(1-\varepsilon(\beta)\big)2c_*\;,$$

where $c_* = \min\{c(x,\delta) : x, \delta\}$ *denotes the minimal transition rate.*

Proof. The claim is a consequence of the first part of Theorem 3.9 once we have established the following estimate. Let q_* be defined as in (3.16). Then, for all $x \in \mathcal{X}$ and all $i,j = 1,\dots,n$ we have:

$$\frac{q(\delta_i x,\delta_i,\delta_j) - q_*(\delta_i x,\delta_i,\delta_j)}{q(x,\delta_i,\delta_i)} \;\le\; \exp\left(2\beta\sum_{m\neq i,j}|k_{im}| + |k_{jm}|\right)\left(e^{4\beta|k_{ij}|}-1\right)\;. \tag{4.2}$$

Indeed, we first note that

$$q(x,\delta_i,\delta_j) \;=\; \exp\left(-\frac{\beta}{2}\big(H(\delta_i x)+H(\delta_j x)\big)\right) .$$

Note further that for $i\neq j$ we have

$$\begin{aligned} H(x) \;&=\; -\sum_{l,m\neq i,j} k_{lm}x_l x_m - 2\sum_{m\neq i,j} k_{im}x_i x_m - 2\sum_{m\neq i,j} k_{jm}x_j x_m - 2k_{ij}x_i x_j \;, \\ H(\delta_i x) \;&=\; -\sum_{l,m\neq i,j} k_{lm}x_l x_m + 2\sum_{m\neq i,j} k_{im}x_i x_m - 2\sum_{m\neq i,j} k_{jm}x_j x_m + 2k_{ij}x_i x_j \;, \end{aligned}$$

which yields

$$H(\delta_i x)+H(\delta_j x) \;=\; -2\sum_{l,m\neq i,j} k_{lm}x_l x_m + 4k_{ij}x_i x_j \;.$$

Since the first term on the right hand side does not depend on the coordinates i,j, we get, for $y\in\{x,\delta_i x,\delta_j x,\delta_i\delta_j x\}$:

$$q(y,\delta_i,\delta_j) \;=\; \exp\left(\beta\sum_{l,m\neq i,j} k_{lm}x_l x_m\right)\exp\big(-2\beta k_{ij}y_i y_j\big) \;,$$

and we conclude that

$$q(\delta_i x,\delta_i,\delta_j)-q_*(\delta_i x,\delta_i,\delta_j) \;\le\; \exp\left(\beta\sum_{l,m\neq i,j} k_{lm}x_l x_m\right)\left(e^{2\beta|k_{ij}|}-e^{-2\beta|k_{ij}|}\right) .$$

Similarly, noting that $q(x,\delta_i,\delta_i)=\exp\big(-\beta H(\delta_i x)\big)$, we obtain the estimate

$$\begin{aligned} q(x,\delta_i,\delta_i) \;&=\; \exp\left(\beta\sum_{l,m\neq i,j} k_{lm}x_l x_m\right)\exp\left(-2\beta\left[\sum_{m\neq i,j}(k_{im}x_i-k_{jm}x_j)x_m + k_{ij}x_i x_j\right]\right) \\ &\ge\; \exp\left(\beta\sum_{l,m\neq i,j} k_{lm}x_l x_m\right)\exp\left(-2\beta\left(|k_{ij}|+\sum_{m\neq i,j}|k_{im}|+|k_{jm}|\right)\right) \;, \end{aligned} \tag{4.3}$$

which yields the claim (4.2). Thus, by (4.1) we find that

$$\lambda=\min_{x,i} c(x,\delta_i)\Big[1-\sum_{j\neq i}\frac{(q-q_*)(\delta_i x,\delta_i,\delta_j)}{q(x,\delta_i,\delta_i)}\Big] \ge c_*\Big(1-\varepsilon(\beta)\Big)\ge 0 \;.$$

Hence, the assumption (3.17) of Theorem 3.9 is satisfied and the thesis follows. □

Remark 4.2. In (4.2) we have given a worst-case estimate in terms of the absolute value of the interaction. This estimate seems rather sharp if the interaction is ferromagnetic, i.e. $k_{ij}\ge 0$ for all i,j. However, in models where the interaction matrix changes sign, a finer estimate making use of frustration effects should be possible. This concerns the second line of (4.3).

Remark 4.3. Note that the perturbative curvature bound in Theorem 4.1 becomes sharp in the limit of infinite temperature, $\beta\to 0$. Indeed, for $\beta=0$, $(\mathcal{X},Q_0,\pi_0)$ is the hypercube equipped with the uniform measure and the Markov chain jumping at rate 1 to any nearest neighbor in the Hamming disance. Since $\varepsilon(0)=0$ and $c_*|_{\beta=0}=1$, we find $\mathrm{Ric}(\mathcal{X},Q_0,\pi_0)\ge 2$. This bound was already obtained in [8] (using a different normalization of Q) and observed to be optimal since it implies the modified logarithmic Sobolev inequality with the optimal constant.

Let us now specialize our result to the d-dimensional Ising model and the Curie–Weiss model.

4.1.1 The d-dimensional Ising model

Let Λ be a finite connected subset of $\mathbb{Z}^d$ endowed with the natural graph structure. Put $n = |\Lambda|$. We consider the Hamiltonian $H : \{-1,1\}^\Lambda \to \mathbb{R}$ given by

$$H(x) \;=\; -\frac{1}{2}\sum_{i\sim j} x_i x_j \;,$$

where $i \sim j$ means that i and j are adjacent. Note that this corresponds to choosing the matrix $k \in \mathbb{R}^{n\times n}$ as

$$k_{ij} \;=\; \begin{cases} \frac{1}{2}\;, & i\sim j\;,\\ 0\;, & \text{else.}\end{cases}$$

Noting that each site has at most $2d$ neighbors we have the following bound

$$\varepsilon(\beta) \;\le\; (2d-1)e^{2\beta(2d-1)}\big(e^{2\beta}-1\big)\;. \tag{4.4}$$

Moreover, the minimal transition rate for the Glauber dynamics becomes $c_* = e^{-2\beta d}$.

Corollary 4.4. *Assume that $\varepsilon(\beta) \le 1$. Then the Glauber dynamics for the d-dimensional Ising model satisfies*

$$\mathrm{Ric}(Q_{Gl}) \;\ge\; \big(1-\varepsilon(\beta)\big)2e^{-2\beta d}\;,$$

where $\varepsilon(\beta)$ is given by (4.4)*.*

In particular, for $d = 2$ we see using the bound (4.4) that the condition $\varepsilon(\beta) \le 1$ is satisfied if

$$3e^{6\beta}\big(e^{2\beta}-1\big) \;\le\; 1\;, \quad \text{approximately} \quad \beta \;\le\; 0.089\;.$$

4.1.2 The Curie–Weiss model

We consider the Hamiltonian $H : \{-1,1\}^n \to \mathbb{R}$ given by

$$H(x) \;=\; -\frac{1}{2n}\sum_{i,j=1}^{n} x_i x_j \;,$$

Note that this corresponds to choosing the matrix $k \in \mathbb{R}^{n\times n}$ as

$$k_{ij} \;=\; \begin{cases} \frac{1}{2n}\;, & i\neq j\;,\\ 0\;, & \text{else.}\end{cases}$$

Thus we see that (4.1) turns into

$$\varepsilon(\beta) \;=\; (n-1)e^{2\beta\frac{n-2}{n}}\big(e^{2\beta\frac{1}{n}}-1\big)\;. \tag{4.5}$$

Moreover, the minimal transition rate for the Glauber dynamics becomes $c_* = e^{-\beta\frac{n-1}{n}}$.

Corollary 4.5. *Assume that $\varepsilon(\beta) \le 1$. Then the Glauber dynamics for the Curie–Weiss model satisfies*

$$\mathrm{Ric}(Q_{Gl}) \;\ge\; \big(1-\varepsilon(\beta)\big)2e^{-\beta\frac{n-1}{n}}\;,$$

where $\varepsilon(\beta)$ is given by (4.5)*.*

Note that if we disregard corrections of the order $O(\frac{1}{n})$, the condition $\varepsilon(\beta) \leq 1$ corresponds to

$$2\beta e^{2\beta} \leq 1\,, \quad \text{approximately} \quad \beta \leq 0.284\,.$$

Recall from Section 2 that an entropic Ricci bound $\mathrm{Ric} \geq \kappa > 0$ implies the modified logarithmic Sobolev inequality (2.3). Thus, we obtain in particular that the Glauber dynamics for the Curie–Weiss model satisfies MLSI up to the inverse temperature $\beta \approx 0.284$. In a recent preprint, Marton [17] showed that the MLSI holds up to the critical inverse temperature $\beta = 1$. It remains an open question whether a strictly positive lower Ricci bound holds up to the critical temperature for the Curie–Weiss model.

4.2 Bounds for a general hard-core model

In this section we derive Ricci bounds for a general hard-core model put forward in [5].

Let T be a finite set and consider the configuration space $S := \{x : T \to \mathbb{N} \cup \{0\}\}$. A set $A \subset S$ is called *decreasing* if for all $x, y \in S$, it holds

$$x \in A\,,\ y(i) \leq x(i),\ \ \forall i \in T \quad \Rightarrow \quad y \in A\,.$$

We fix a finite decreasing set A and call it the set of *allowed configurations*. We fix a function $\nu : T \to (0, \infty)$, called the *intensity*, and define a probability measure π on $\mathcal{X} := A$ by

$$\pi(x) = \frac{1}{Z}\prod_{i\in T}\frac{\nu(i)^{x(i)}}{x(i)!},$$

where Z is a normalization constant. A Markov dynamics on $\mathcal{X}$ is given by the rate matrix

$$Q(x,y) := \begin{cases} \nu(i)\mathbf{1}_{\{x+\mathbf{1}_i\in A\}}\,, & \text{if } y = x + \mathbf{1}_i\,,\\ x(i)\mathbf{1}_{\{x-\mathbf{1}_i\in A\}}\,, & \text{if } y = x - \mathbf{1}_i\,,\\ 0\,, & \text{else}\,. \end{cases}$$

Note that this dynamics is reversible w.r.t. π. A natural mapping representation for this model is given on the extended state space $\mathcal{X}' = S$ (c.f. Remark 3.2) as follows. Let

$$G = \{\gamma_i^+, \gamma_i^- : i \in T\}\,,$$

where $\gamma_i^+, \gamma_i^- : S \to S$ are the creation and annihilation maps defined by

$$\gamma_i^+(x) = x + \mathbf{1}_i\,, \qquad \gamma_i^-(x) = \begin{cases} x - \mathbf{1}_i\,, & \text{if } x(i) > 0\,,\\ x\,, & \text{else}\,. \end{cases}$$

We then may define the transition rates $c : \mathcal{X}' \times G \to \mathbb{R}_+$ by

$$c(x,\gamma_i^+) = \nu(i)\mathbf{1}_{\{x+\mathbf{1}_i\in A\}}\,, \qquad c(x,\gamma_i^-) = x(i)\mathbf{1}_{\{x(i)>0\}}\,.$$

Define

$$\epsilon_0 = \max_{x\in A, i\in T: x(i)>0}\sum_{j\neq i}\nu(j)\mathbf{1}_{\{x+\mathbf{1}_j-\mathbf{1}_i\in A\}}\mathbf{1}_{\{x+\mathbf{1}_j\notin A\}}\,, \tag{4.6}$$

$$\epsilon_1 = \min_{x\in A, i\in T: x(i)>0}\nu(i)\mathbf{1}_{\{x+\mathbf{1}_i\notin A\}}\,. \tag{4.7}$$

We have the following entropic Ricci curvature bound for the general hard-core model.

Theorem 4.6. *Assume that $\epsilon_0 \le 1$. Then, we have that*

$$\mathrm{Ric}(\mathcal{X}, Q, \pi) \ge \frac{1}{2}(1 - \epsilon_0 + \epsilon_1) \ .$$

Proof. The result will be a consequence of the second part of Theorem 3.9. We let $H_1 = \{\gamma_i^+ : i \in T\}$ and $H_2 = \{\gamma_i^- : i \in T\} = H_1^{-1}$. One readily checks that for all $x \in S$ and all $i \neq j$:

$$0 = (q - q_*)(x, \gamma_i^-, \gamma_j^-) = (q - q_*)(x, \gamma_i^-, \gamma_j^+) = (q - q_*)(x, \gamma_i^+, \gamma_j^-) \ .$$

Moreover, we have

$$(q - q_*)(x, \gamma_i^+, \gamma_j^+) = \mathbf{1}_{\{x+\mathbf{1}_i \in A\}} \mathbf{1}_{\{x+\mathbf{1}_j \in A\}} \mathbf{1}_{\{x+\mathbf{1}_i+\mathbf{1}_j \notin A\}} \nu(i)\nu(j)\pi(x) \ .$$

This yields

$$\frac{(q - q_*)(\delta x, \delta^{-1}, \eta)}{c(x, \delta)\pi(x)} = \begin{cases} 0 \ , & \delta = \gamma_i^+, \eta = \gamma_j^- \ , \\ 0 \ , & \delta = \gamma_i^+, \eta = \gamma_j^+ \ , \\ 0 \ , & \delta = \gamma_i^-, \eta = \gamma_j^- \ , \\ \mathbf{1}_{\{x(i)>0\}} \mathbf{1}_{\{x-\mathbf{1}_i+\mathbf{1}_j \in A\}} \mathbf{1}_{\{x+\mathbf{1}_j \notin A\}} \nu(j) \ , & \delta = \gamma_i^-, \eta = \gamma_j^+ \ . \end{cases}$$

In the notation of Theorem 3.9 we thus obtain

$$\begin{aligned} \lambda_1 &= \min_{x \in A, i \in T} c(x, \gamma_i^+) - c(\gamma_i^+ x, \gamma_i^+) = \min_{x \in A, i \in T} \nu(i) \mathbf{1}_{\{x+\mathbf{1}_i \in A\}} \mathbf{1}_{\{x+2\cdot\mathbf{1}_i \notin A\}} \\ &= \min_{x \in A, i \in T : x(i)>0} \nu(i) \mathbf{1}_{\{x+\mathbf{1}_i \notin A\}} = \varepsilon_1 \ . \end{aligned}$$

Moreover, we get

$$\begin{aligned} \lambda_2 &= \min_{x \in A, i \in T} c(x, \gamma_i^-) - c(\gamma_i^- x, \gamma_i^-) - \sum_{j \neq i} \nu(j) \mathbf{1}_{\{x(i)>0\}} \mathbf{1}_{\{x-\mathbf{1}_i+\mathbf{1}_j \in A\}} \mathbf{1}_{\{x+\mathbf{1}_j \notin A\}} \\ &= \min_{x \in A, i \in T : x(i)>0} x(i) - \big(x(i) - 1\big) - \sum_{j \neq i} \nu(j) \mathbf{1}_{\{x-\mathbf{1}_i+\mathbf{1}_j \in A\}} \mathbf{1}_{\{x+\mathbf{1}_j \notin A\}} = 1 - \varepsilon_0 \ . \end{aligned}$$

Applying the second part of Theorem 3.9 yields the claim. □

Remark 4.7. Under the same assumptions as in Theorem 4.6, Dai Pra and Posta established in [5] the convex entropy decay inequality (2.6) with $2\kappa = 1 - \varepsilon_0 + \varepsilon_1$. Recall from Section 2 that $\mathrm{Ric} \ge \kappa$ implies (2.6). Thus, by the previous theorem we recover, in particular, the result in [5].

Let us specialize our result to the standard hard-core model and a model for long hard rods.

4.2.1 The hard-core model

Let $G = (V, E)$ be a finite, connected graph without self-loops. Using our notation above, we let $T := V$, $\nu(i) \equiv \rho$ for some constant $1 > \rho > 0$, and

$$A := \{x \in S : x(i) \in \{0, 1\} \text{ for all } i \in T, x(i)x(j) = 0 \text{ for all } \{i, j\} \in E\}.$$

We define Δ to be the maximum degree of any vertex in the graph. It is easy to see that in this case (4.6) and (4.7) become $\epsilon_0 = \rho\Delta$ and $\varepsilon_1 = \rho$. Thus, we obtain the following corollary.

Corollary 4.8. *If $\rho \le 1/\Delta$, then we have* $\mathrm{Ric}(\mathcal{X}, Q, \pi) \ge \frac{1}{2}\big(1 - \rho(\Delta - 1)\big)$.

This model has been widely studied in the literature. We refer the interested reader to the book of Levin, Peres and Wilmer [13] and the works of Luby and Vigoda [15], Dyer and Greenhill [7], and Vigoda [23], concerning several versions of fast mixing results for the hard-core model.

4.2.2 Long hard rods

Fix two natural numbers L and k in the regime where $L \gg k$. Define the space T_- of horizontal rods of length k to be the collection of all sequence of $(k+1)$ adjacent vertices in $\{0,1,\dots,L\}^2$ of the form

$$\{(u_1,u_2),(u_1+1,u_2),\dots,(u_1+k,u_2)\}.$$

Similarly, define the space T_+ of vertical rods of length k to be the collection of all sequence of adjacent vertices in $\{0,1,\dots,L\}^2$ of the form

$$\{(u_1,u_2),(u_1,u_2+1),\dots,(u_1,u_2+k)\}.$$

We then define T, the set of all rods, as the union of T_- and T_+. The admissible set is defined to be

$$A = \{x \in S : x(i) \in \{0,1\} \text{ for all } i \in T, x(i)x(j) = 0 \text{ if } i \neq j \text{ and } i \cap j \neq \emptyset\}.$$

In other words, we wish to only allow rods which do not overlap. Further, we let $\nu(i) \equiv \rho$ for some constant $\rho > 0$. It is easy to check that in this case $\epsilon_0 = \rho(k^2+4k+1)$ and $\epsilon_1 = \rho$. Thus we obtain the following corollary.

Corollary 4.9. *If* $\rho \le 1/(k^2+4k+1)$, *we have* $\mathrm{Ric}(\mathcal{X},Q,\pi) \ge \frac{1}{2}\Big(1-\rho(k^2+4k)\Big)$.

As pointed out by Disertori and Giuliani [6], for k sufficiently large, there is a phase transition as L tends to infinity at some critical value ρ_c, which is expected to be of the order k^{-2}. Just as for the convex entropy decay considered in [5], our work above yields a uniform (in L) curvature bound in the asymptotically correct regime.

4.3 Random walks on the symmetric group

Let us briefly highlight a class of examples where Corollary 3.12 applies.

Consider the symmetric group S_n of all permutations on n letters. A conjugacy-invariant set of generators is given for instance by the set $G_{n,k}$ of all k-cycles in S_n for $1 < k < n$. Here a k-cycle is a cyclic permutation of length k. The simple random walk $Q_{n,k}$ on the associated Cayley graph is given by $c_{n,k}(x,\delta) \equiv |G_{n,k}|^{-1} = \frac{k(n-k)!}{n!}$. It is reversible w.r.t. the uniform probability measure π_n on S_n. Note that in the case $k=2$ of transpositions we have that the mapping representation is involutive, i.e. it holds $\delta = \delta^{-1}$ for all δ.

Corollary 4.10. *The simple random walk on* S_n *generated by* k*-cycles satisfies*

$$\mathrm{Ric}(S_n, Q_{n,k}, \pi_n) \ \ge \ \begin{cases} \frac{4}{n(n-1)}\ , & k=2\ ,\\ 0\ , & k>2\ . \end{cases}$$

In the case of 2-cycles or transpositions, we recover the result obtained in [9, Thm. 1.2]. In this case, the optimal constant κ in the MLSI (2.3) is known to satisfy the bounds $1/2(n-1) \le \kappa \le 2/(n-1)$. Thus the Ricci bound that we obtain differs roughly by a factor n. It is an open question to determine the correct order for the Ricci bound in the case $k=2$ and whether a strictly positive Ricci bound holds for general k.

References

[1] D. Bakry and M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206. MR-889476

[2] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375–393. MR-1738163

[3] A.-I. Bonciocat and K.-Th. Sturm, *Mass transportation and rough curvature bounds for discrete spaces*, J. Funct. Anal. **256** (2009), no. 9, 2944–2966. MR-2502429

[4] P. Caputo, P. Dai Pra, and G. Posta, *Convex entropy decay via the Bochner-Bakry-Emery approach*, Ann. Inst. Henri Poincaré Probab. Stat. **45** (2009), no. 3, 734–753. MR-2548501

[5] P. Dai Pra and G. Posta, *Entropy decay for interacting systems via the Bochner-Bakry-Émery approach*, Electron. J. Probab. **18** (2013), no. 52, 21. MR-3065862

[6] M. Disertori and A. Giuliani, *The nematic phase of a system of long hard rods*, Comm. Math. Phys. **323** (2013), no. 1, 143–175. MR-3085664

[7] M. Dyer and C. Greenhill, *On Markov chains for independent sets*, J. Algorithms, **35** (2000), 17–49. MR-1747721

[8] M. Erbar and J. Maas, *Ricci curvature of finite Markov chains via convexity of the entropy*, Arch. Ration. Mech. Anal. **206** (2012), no. 3, 997–1038. MR-2989449

[9] M. Erbar, J. Maas, and P. Tetali, *Discrete Ricci curvature bounds for Bernoulli-Laplace and random transposition models*, Ann. Fac. Sci. Toulouse Math. (6) **24** (2015), no. 4, 781–800. MR-3434256

[10] M. Fathi and J. Maas, *Entropic Ricci curvature bounds for discrete interacting systems*, Ann. Appl. Probab. **26** (2016), no. 3, 1774–1806. MR-3513606

[11] N. Gozlan, C. Roberto, P.-M. Samson, and P. Tetali, *Displacement convexity of entropy and related inequalities on graphs*, Probab. Theory Related Fields **160** (2014), no. 1–2, 47–94. MR-3256809

[12] R. Holley and D. Stroock, *Logarithmic Sobolev inequalities and stochastic Ising models*, J. Statist. Phys. **46** (1987), no. 5–6, 1159–1194. MR-893137

[13] D. Levin, Y. Peres, and E. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009, With a chapter by James G. Propp and David B. Wilson. MR-2466937

[14] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. Math. (2) **169** (2009), no. 3, 903–991. MR-2480619

[15] M. Luby and E. Vigoda, *Fast convergence of the Glauber dynamics for sampling independent sets*, Random Structures Algorithms **15** (1999), no. 3–4, 229–241, Statistical physics methods in discrete probability, combinatorics, and theoretical computer science (Princeton, NJ, 1997). MR-1716763

[16] J. Maas, *Gradient flows of the entropy for finite Markov chains*, J. Funct. Anal. **261** (2011), no. 8, 2250–2292. MR-2824578

[17] K. Marton, *Logarithmic Sobolev inequalities in discrete product spaces: a proof by a transportation cost distance*, arXiv:1507.02803 (2015).

[18] A. Mielke, *A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems*, Nonlinearity **24** (2011), no. 4, 1329–1346. MR-2776123

[19] A. Mielke, *Geodesic convexity of the relative entropy in reversible Markov chains*, Calc. Var. Partial Differential Equations **48** (2013), no. 1–2, 1–31. MR-3090532

[20] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256** (2009), no. 3, 810–864. MR-2484937

[21] Y. Ollivier, *A survey of Ricci curvature for metric spaces and Markov chains*, Probabilistic approach to geometry, Adv. Stud. Pure Math., vol. 57, Math. Soc. Japan, Tokyo, 2010, pp. 343–381. MR-2648269

[22] K.-Th. Sturm, *On the geometry of metric measure spaces. I and II*, Acta Math. **196** (2006), no. 1, 65–177. MR-2237207

[23] E. Vigoda, *A note on the Glauber dynamics for sampling independent sets*, Electron. J. Combin. **8** (2001), no. 1, Research Paper 8, 8 pp. (electronic). MR-1814515

Acknowledgments. The authors wish to thank Max Fathi, Jan Maas and André Schlichting for stimulating discussions on this work and related topics. This work originated in discussions that took place at the SQuaRE meetings, *Displacement convexity for* *interacting Markov chains*, at the American Institute for Mathematics (AIM), California. The authors wish to thank AIM for the inspiring atmosphere making this collaboration possible. M.E. gratefully acknowledges support by the German Research Foundation through the Collaborative Research Center 1060 *The Mathematics of Emergent Effects* and the Hausdorff Center for Mathematics. Part of this work was performed within the framework of the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program *Investissements d'Avenir* (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). G.M. gratefully acknowledges support by the National Science Foundation under Grant No. DMS-1440140, while he was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 2015 semester. P.T. gratefully acknowledges support by the NSF grant DMS-1407657.