The Annals of Applied Probability
2016, Vol. 26, No. 3, 1774–1806
DOI: 10.1214/15-AAP1133

ENTROPIC RICCI CURVATURE BOUNDS FOR DISCRETE INTERACTING SYSTEMS[1]

BY MAX FATHI AND JAN MAAS

Université Paris 6 and Institute of Science and Technology Austria

We develop a new and systematic method for proving entropic Ricci curvature lower bounds for Markov chains on discrete sets. Using different methods, such bounds have recently been obtained in several examples (e.g., 1-dimensional birth and death chains, product chains, Bernoulli–Laplace models, and random transposition models). However, a general method to obtain discrete Ricci bounds had been lacking. Our method covers all of the examples above. In addition, we obtain new Ricci curvature bounds for zero-range processes on the complete graph. The method is inspired by recent work of Caputo, Dai Pra and Posta on discrete functional inequalities.

1. Introduction. Ricci curvature lower bounds for Riemannian manifolds play a crucial role in differential geometry, but also in probability theory and analysis. In particular, such bounds are known to imply important properties of Brownian motion on the manifold. For example, positive curvature bounds imply several useful functional inequalities (logarithmic Sobolev inequality, Poincaré inequality), as well as exponential rates of convergence to equilibrium. Such results were first obtained using the celebrated Bakry–Émery approach developed in [1].

1.1. *Discrete entropic Ricci curvature.* In [18] and [9] a notion of Ricci curvature bounds for Markov chains was introduced, based on convexity properties of the relative entropy along geodesics in the space of probability measures over the state space, for a well-chosen metric structure. It is the discrete analogue of the synthetic notion of Ricci curvature bounds in geodesic metric-measure spaces obtained in the contributions of Lott–Villani [17] and Sturm [24].

The main observation behind the Lott–Villani–Sturm definition of Ricci curvature bounds is that, on a Riemannian manifold $\mathcal{M}$, the Ricci curvature is bounded from below by some constant $\kappa \in \mathbb{R}$ if and only if the Boltzmann–Shannon entropy $\mathcal{H}(\rho) = \int \rho \log \rho \, \mathrm{dvol}$ is κ-convex along geodesics in the L^2-Wasserstein space of probability measures on $\mathcal{M}$. However, this definition is not well adapted to discrete

Received January 2015; revised July 2015.

[1]Supported by the German Research Foundation through the Collaborative Research Center 1060 *The Mathematics of Emergent Effects* and the Hausdorff Center for Mathematics.

MSC2010 subject classifications. 60J10, 60K35.

Key words and phrases. Discrete Ricci curvature, transport metrics, functional inequalities, birth-death processes, zero-range processes, Bernoulli–Laplace model, random transposition model.

spaces, because the L^2-Wasserstein space over a discrete space does not contain any geodesic.

To circumvent this issue, [18] and [20] introduced a different metric $\mathcal{W}$. This metric is built from a Markov kernel in such a way that the heat flow associated to this kernel is the gradient flow of the entropy with respect to the distance $\mathcal{W}$. This is an analogue of the continuous situation, where it has been proven by Jordan, Kinderlehrer and Otto in [15] that the heat flow on $\mathbb{R}^d$ is the gradient flow of the entropy with respect to the Wasserstein metric. A Markov chain is then said to have Ricci curvature bounded from below by some constant κ if the relative entropy (with respect to the invariant probability measure of the chain) is κ-convex along geodesics for the metric $\mathcal{W}$.

It has been shown in [9] that a Ricci curvature lower bound has significant consequences for the associated Markov chain, such as a (modified) logarithmic Sobolev inequality, a Talagrand transportation inequality and a Poincaré inequality. Therefore, it is desirable to obtain sharp Ricci curvature bounds in concrete discrete examples. To this day, there are very few results of this type.

Mielke [20] obtained Ricci curvature bounds for one-dimensional birth and death processes, and applied these bounds to discretizations of 1-dimensional Fokker–Planck equations. Erbar and Maas [9] proved a tensorization principle: the Ricci curvature lower bound for a product system is the minimum of the Ricci curvature bounds of the components. This property is crucial for applications to high-dimensional problems. As a special case, the result yields the sharp Ricci bound for the discrete hypercube $\{0, 1\}^n$. The first high-dimensional results beyond the product-setting have been recently obtained by Erbar, Maas and Tetali [11], who obtained Ricci curvature bounds for the Bernoulli–Laplace model and for the random transposition model on the symmetric group.

In spite of this progress, a systematic method for obtaining discrete Ricci curvature bounds has been lacking. In this paper, we present such a method, which allows us to obtain curvature bounds for several interacting particle systems on the complete graph. The method allows us to recover all of the above-mentioned results, and yields new curvature bounds for zero-range processes on the complete graph.

1.2. *The Bochner-approach.* To explain our method, let us recall that the discrete Ricci curvature bounds considered in this paper correspond to bounds on the second derivative of the entropy along $\mathcal{W}$-geodesics. In the continuous setting, on a Riemannian manifold $\mathcal{M}$, the second derivative of the entropy along a 2-Wasserstein geodesic (ρ_t) is formally given by

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{H}(\rho_t) = \int_{\mathcal{M}} -\langle \nabla\psi_t, \nabla\Delta\psi_t\rangle + \frac{1}{2}\Delta\big(|\nabla\psi_t|^2\big)\,\mathrm{d}\rho_t, \tag{1}$$

where $\psi : [0, 1] \times \mathcal{M} \to \mathbb{R}$ satisfies the continuity equation $\partial_t\rho + \nabla\cdot(\rho\nabla\psi) = 0$ and the Hamilton–Jacobi equation $\partial_t\psi + \frac{1}{2}|\nabla\psi|^2 = 0$. Note that the integrand above can be reformulated using Bochner's formula, which yields

$$-\langle \nabla \psi_t, \nabla \Delta \psi_t \rangle + \frac{1}{2}\Delta\big(|\nabla \psi_t|^2\big) = \frac{1}{2}\big|\mathrm{D}^2\psi\big|^2 + \mathrm{Ric}(\nabla\psi, \nabla\psi). \tag{2}$$

Therefore, it follows at least formally that a lower bound $\mathrm{Ric} \geq \kappa$ implies κ-convexity of the entropy along geodesics.

In the discrete case, the second derivative of the entropy along geodesics is given by an expression $\mathcal{B}(\rho, \psi)$ which somewhat resembles the right-hand side of (1), but the dependence on ρ is more complicated. Unfortunately, there does not seem to be a suitable discrete analogue of Bochner's formula.

In this paper, we present an approach to get around this difficulty. We obtain a convenient lower bound for the second derivative of the entropy, which turns out to be very useful; see Theorem 3.5 below. This expression has the same features as the right-hand side of Bochner's formula (2): one of the terms is nonnegative and contains only second-order (discrete) derivatives, while the other terms contain only first-order derivatives. The expression is remarkably flexible, since it allows us to derive sharp Ricci bounds in the rather diverse examples mentioned above.

The work of Caputo, Dai Pra and Posta [5] which inspired this paper, fits very naturally into this framework. In order to obtain modified logarithmic Sobolev inequalities, these authors followed the Bakry–Émery strategy, which consists of computing the second derivative of the entropy along the "heat equation" $\partial_t \rho = \mathcal{L}\rho$, where $\mathcal{L}$ is the generator of a reversible Markov chain. It can be checked that this second derivative can be expressed as $2\mathcal{B}(\rho, \log\rho)$. Therefore, the bounds obtained in [5] and [8] are a special case of the bounds on $\mathcal{B}(\rho, \psi)$, which we obtain for arbitrary ψ. We refer to Section 2.3 for more details.

1.3. *Organization and notation.* In Section 2, we collect preliminaries on discrete transport metrics, the associated Riemannian structure, and the notion of discrete entropic Ricci curvature. Section 3 contains the general criterion for Ricci curvature bounds. Various examples are studied in Section 4. The Appendix states a few useful properties of the logarithmic mean that are used in the proofs.

Throughout the paper, we use the probabilistic notation

$$\pi[F] = \pi\big[F(\eta)\big] = \sum_{\eta\in\mathcal{X}} F(\eta)\pi(\eta).$$

2. Transport metrics and discrete Ricci curvature. In this section, we briefly recall some basic facts on the discrete transportation metric $\mathcal{W}$, which plays a crucial role in the paper. This metric has been introduced in [18] in the setting of finite Markov chains, and (independently) in [19] for reaction–diffusion systems. Closely related metrics have been considered in [7] in the setting of Fokker–Planck equations on graphs. Our discussion follows [18] with a slightly different normalisation convention (see also [10]).

We work with a discrete, finite space $\mathcal{X}$ and a Markov generator $\mathcal{L}$, acting on functions $\psi : \mathcal{X} \to \mathbb{R}$ by

$$\mathcal{L}\psi(\eta) = \sum_{\tilde{\eta}\in\mathcal{X}} Q(\eta, \tilde{\eta})\big(\psi(\tilde{\eta}) - \psi(\eta)\big).$$

The transition rates $Q(\eta, \tilde{\eta})$ are nonnegative for all distinct $\eta, \tilde{\eta} \in \mathcal{X}$, and we use the convention that $Q(\eta, \eta) = 0$ for all $\eta \in \mathcal{X}$. We assume that Q is *irreducible*, that is, for all $\eta, \tilde{\eta} \in \mathcal{X}$, there exists a sequence $\{\eta_i\}_{i=0}^n \subseteq \mathcal{X}$ such that $\eta_0 = \eta$, $\eta_n = \tilde{\eta}$, and $Q(\eta_i, \eta_{i+1}) > 0$ for all $i = 0, \dots, n-1$. It is a basic result of Markov chain theory that there is a unique stationary probability measure π on $\mathcal{X}$, which means that

$$\sum_{\eta\in\mathcal{X}} \pi(\eta) = 1 \quad \text{and} \quad \pi(\tilde{\eta}) = \sum_{\eta\in\mathcal{X}} Q(\eta, \tilde{\eta})\pi(\eta).$$

We shall always assume that π is *reversible* for Q, that is, the detailed balance equations

$$Q(\eta, \tilde{\eta})\pi(\eta) = Q(\tilde{\eta}, \eta)\pi(\tilde{\eta})$$

hold for all $\eta, \tilde{\eta} \in \mathcal{X}$. We shall refer to $(\mathcal{X}, Q, \pi)$ as a Markov triple.

2.1. *The nonlocal transport metric $\mathcal{W}$.* Let

$$\mathscr{P}(\mathcal{X}) := \left\{\rho : \mathcal{X} \to \mathbb{R}_+ \,\middle|\, \sum_{\eta\in\mathcal{X}} \rho(\eta)\pi(\eta) = 1\right\}$$

be the set of *probability densities* on $\mathcal{X}$ (with respect to the stationary measure π). We consider the nonlocal transport metric $\mathcal{W}$ defined for $\rho_0, \rho_1 \in \mathscr{P}(\mathcal{X})$ by

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

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t(\eta) + \sum_{\tilde{\eta}\in\mathcal{X}} \big(\psi_t(\tilde{\eta}) - \psi_t(\eta)\big)\hat{\rho}_t(\eta, \tilde{\eta}) Q(\eta, \tilde{\eta}) = 0 \qquad \forall\eta \in \mathcal{X}, \tag{3}$$

with boundary conditions $\rho|_{t=0} = \rho_0$ and $\rho|_{t=1} = \rho_1$. It has been shown in [18] that $\mathcal{W}$ defines a distance on the set of probability densities $\mathscr{P}(\mathcal{X})$. Moreover, this distance is induced by a Riemannian structure on its interior

$$\mathscr{P}_*(\mathcal{X}) := \{\rho \in \mathscr{P}(\mathcal{X}) : \rho(x) > 0 \text{ for all } x \in \mathcal{X}\}.$$

This definition is a natural analogue of the Benamou–Brenier formulation of the Wasserstein metric, with one crucial additional feature in the discrete setting: the logarithmic mean

$$\hat{\rho}(\eta, \tilde{\eta}) := \theta\big(\rho(\eta), \rho(\tilde{\eta})\big) \qquad \text{with } \theta(r, s) = \int_0^1 r^{1-p}s^p\,\mathrm{d}p = \frac{r - s}{\log r - \log s} \tag{4}$$

is used to define, loosely speaking, the value of ρ along the edge between η and $\tilde{\eta}$. Note that the integral representation of θ holds for all $r, s \geq 0$, while the alternative expression requires r and s to be strictly positive. The main reason for the appearance of this particular mean is the fact that it allows one to formulate a discrete chain rule $\hat{\rho}\nabla \log \rho = \nabla \rho$, where

$$\nabla \psi(\eta, \tilde{\eta}) := \psi(\tilde{\eta}) - \psi(\eta)$$

denotes the discrete gradient. This identity takes over the role of the usual chain rule $\nabla \log \rho = \nabla \rho / \rho$ that conveniently holds in the continuous setting.

In [18, 19], it is shown that the "heat equation" $\partial_t \rho = \mathcal{L}\rho$ is the gradient flow equation for the relative entropy

$$\mathcal{H}(\rho) := \sum_{\eta \in \mathcal{X}} \pi(\eta)\rho(\eta) \log \rho(\eta),$$

with respect to the Riemannian structure associated with $\mathcal{W}$. In this sense, the metric $\mathcal{W}$ is a natural discrete analogue of the L^2-Wasserstein distance. Moreover, at least in some situations [13], it can be shown that the metric $\mathcal{W}$ converges to the L^2-Wasserstein in the continuous limit, in the sense of Gromov–Hausdorff.

Following the Lott–Villani–Sturm approach, [9, 18] gave the following definition of lower Ricci curvature boundedness.

DEFINITION 2.1 (Ricci curvature lower boundedness). We say that a Markov triple $(\mathcal{X}, Q, \pi)$ has Ricci curvature bounded from below by $\kappa \in \mathbb{R}$ if for any constant speed geodesic $(\rho_t)_{t \in [0,1]}$ in $(\mathscr{P}(\mathcal{X}), \mathcal{W})$ we have

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

We shall use the notation $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$ [or briefly $\mathrm{Ric}(Q) \geq \kappa$]. Since $(\mathscr{P}(\mathcal{X}), \mathcal{W})$ is a geodesic space, this definition is nontrivial. This notion of curvature has several interesting properties, obtained in [9]:

- A Markov chain that has $\mathrm{Ric}(Q) \geq \kappa > 0$ satisfies various functional inequalities, such as the modified logarithmic Sobolev inequality

  $$\mathcal{H}(\rho) \leq \frac{1}{\alpha}\mathcal{E}(\rho, \log \rho) \qquad \forall \rho \in \mathscr{P}(\mathcal{X}),$$

  with constant $\alpha = 2\kappa$, as well as a Poincaré inequality

  $$\mathrm{Var}_\pi(\psi) \leq \frac{1}{\lambda}\mathcal{E}(\psi, \psi) \qquad \forall \psi : \mathcal{X} \to \mathbb{R},$$

  with constant $\lambda = \kappa$. Here, $\mathrm{Var}_\pi(\psi) = \pi[\psi^2] - \pi[\psi]^2$, and $\mathcal{E}$ denotes the discrete Dirichlet form given by

  $$\mathcal{E}(\varphi, \psi) := \frac{1}{2}\pi\Big[\sum_{\tilde{\eta}} (\varphi(\eta) - \varphi(\tilde{\eta}))(\psi(\eta) - \psi(\tilde{\eta}))Q(\eta, \tilde{\eta})\Big].$$

These functional inequalities imply the exponential convergence estimates

$$\mathcal{H}(e^{tL}\rho) \le e^{-\alpha t}\mathcal{H}(\rho) \quad \text{and} \quad \|e^{tL}\psi\|_{L^2(\mathcal{X},\pi)} \le e^{-\lambda t}\|\psi\|_{L^2(\mathcal{X},\pi)}$$

along the heat equation $\partial_t \rho = \mathcal{L}\rho$.

- Ricci curvature bounds tensorize: if $\mathrm{Ric}(\mathcal{X}_i, Q_i, \pi_i) \ge \kappa_i$, then the product chain $(\mathcal{X}_1 \times \mathcal{X}_2, Q, \pi_1 \otimes \pi_2)$ defined by

$$Q\big((x_1, x_2), (y_1, y_2)\big) = \begin{cases} Q_1(x_1, y_1), & \text{if } x_2 = y_2, \\ Q_2(x_2, y_2), & \text{if } x_1 = y_1, \\ 0, & \text{otherwise,} \end{cases}$$

satisfies $\mathrm{Ric}(\mathcal{X}_1 \times \mathcal{X}_2, Q, \pi_1 \otimes \pi_2) \ge \min\{\kappa_1, \kappa_2\}$.

REMARK 2.2. Besides the entropic Ricci curvature studied in this paper, several other notions of discrete Ricci curvature have been studied in the literature, including Ollivier's coarse Ricci curvature [22], rough Ricci curvature by Bonciocat and Sturm [3] and Bakry–Émery-type conditions, for example, [16]. We refer to [23] to a survey covering some of these concepts.

2.2. *The Riemannian structure induced by $\mathcal{W}$.* It will be useful to describe the Riemannian structure associated to $\mathcal{W}$ in more detail. Let $\mathcal{E} := \{(x, y) \subseteq \mathcal{X} \times \mathcal{X} : Q(x, y) > 0\}$ be the set of edges in the graph induced by Q, and let $\mathcal{G}$ be the set of discrete gradients, that is, all functions $\Psi : \mathcal{E} \to \mathbb{R}$ of the form $\Psi(\eta, \tilde{\eta}) = \nabla\psi(\eta, \tilde{\eta}) := \psi(\tilde{\eta}) - \psi(\eta)$ for some function $\psi : \mathcal{X} \to \mathbb{R}$.

Note that, at each $\rho \in \mathscr{P}_*(\mathcal{X})$, the tangent space of $\mathscr{P}_*(\mathcal{X})$ is naturally given by $\mathcal{T} := \{\sigma : \mathcal{X} \to \mathbb{R} | \sum_{\eta\in\mathcal{X}} \sigma(\eta)\pi(\eta) = 0\}$. It can be proved (see [18], Section 3) that, for each $\rho \in \mathscr{P}_*(\mathcal{X})$, the mapping

$$\mathcal{K}_\rho : \nabla\psi \mapsto \sum_{\tilde{\eta}\in\mathcal{X}} \big(\psi(\tilde{\eta}) - \psi(\eta)\big)\hat{\rho}(\eta, \tilde{\eta})Q(\eta, \tilde{\eta})$$

defines an bijection between $\mathcal{G}$ and $\mathcal{T}$. At each $\rho \in \mathscr{P}_*(\mathcal{X})$, this map allows us to identify the tangent space with $\mathcal{G}$. In other words, the continuity equation (3) provides an identification between the "vertical tangent vector" $\frac{\mathrm{d}}{\mathrm{d}t}\rho$ and the "horizontal tangent vector" $\nabla\psi$. A Riemannian structure is then defined by the ρ-dependent scalar product $\langle\cdot,\cdot\rangle_\rho : \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ given by

$$\langle\Phi, \Psi\rangle_\rho := \frac{1}{2}\pi\Big[\sum_{\tilde{\eta}\in\mathcal{X}} \Phi(\eta, \tilde{\eta})\Psi(\eta, \tilde{\eta})\hat{\rho}(\eta, \tilde{\eta})Q(\eta, \tilde{\eta})\Big]$$

for $\Phi, \Psi : \mathcal{E} \to \mathbb{R}$. We shall frequently use the notation

$$\mathcal{A}(\rho, \psi) := \|\nabla\psi\|_\rho^2,$$

where $\|\cdot\|_\rho$ denotes the norm induced by the scalar product $\langle\cdot,\cdot\rangle_\rho$.

The geodesic equations for this Riemannian structure are given by

$$\begin{cases} \partial_s \rho_s(\eta) + \sum_{\tilde{\eta} \in \mathcal{X}} \big(\psi_s(\tilde{\eta}) - \psi_s(\eta)\big) \widehat{\rho_s}(\eta, \tilde{\eta}) Q(\eta, \tilde{\eta}) = 0, \\ \partial_s \psi_s(\eta) + \frac{1}{2} \sum_{\tilde{\eta} \in \mathcal{X}} \big(\psi_s(\eta) - \psi_s(\tilde{\eta})\big)^2 (\widehat{\rho_s})_1(\eta, \tilde{\eta}) Q(\eta, \tilde{\eta}) = 0, \end{cases}$$

where, by abuse of notation, $\hat{\rho}_i$ denotes the partial derivative of the logarithmic mean defined in (4), that is, for $i = 1, 2$,

$$\hat{\rho}_i(\eta, \tilde{\eta}) := \partial_i \theta\big(\rho(\eta), \rho(\tilde{\eta})\big).$$

These equations can be regarded as discrete analogues of the continuity equation and the Hamilton–Jacobi equation, respectively. The Hessian of the entropy can then be calculated using the expression

$$\big\langle \operatorname{Hess} \mathcal{H}(\rho) \nabla\psi, \nabla\psi \big\rangle_\rho = \frac{\mathrm{d}^2}{\mathrm{d}s^2}\bigg|_{s=0} \operatorname{Ent}(\rho_s),$$

where (ρ_s) solves the geodesic equations with initial conditions $\rho|_{s=0} = \rho_0$ and $\psi|_{s=0} = \psi_0$. The following result provides an explicit formula for the Hessian which resembles its continuous counterpart (1); we refer to [9] for a proof. With a slight abuse of notation, we write

$$\langle \Phi, \Psi \rangle_\pi = \frac{1}{2} \pi \bigg[\sum_{\tilde{\eta} \in \mathcal{X}} \Phi(\eta, \tilde{\eta}) \Psi(\eta, \tilde{\eta}) Q(\eta, \tilde{\eta}) \bigg].$$

PROPOSITION 2.3 (Identification of the Hessian). *For any $\rho \in \mathscr{P}_*(\mathcal{X})$ and $\psi : \mathcal{X} \to \mathbb{R}$, we have*

$$\begin{aligned} \mathcal{B}(\rho, \psi) &:= \big\langle \operatorname{Hess} \mathcal{H}(\rho) \nabla\psi, \nabla\psi \big\rangle_\rho \\ &= \frac{1}{2} \langle \hat{\mathcal{L}}\rho \cdot \nabla\psi, \nabla\psi \rangle_\pi - \langle \hat{\rho} \cdot \nabla\psi, \nabla\mathcal{L}\psi \rangle_\pi, \end{aligned}$$

where

$$\hat{\mathcal{L}}\rho(\eta, \tilde{\eta}) := \widehat{\rho}_1(\eta, \tilde{\eta}) \mathcal{L}\rho(\eta) + \widehat{\rho}_2(\eta, \tilde{\eta}) \mathcal{L}\rho(\tilde{\eta}).$$

Since the metric structure we consider is Riemannian in $\mathscr{P}_*(\mathcal{X})$, one expects that κ-geodesic convexity of $\mathcal{H}$ is equivalent to

$$\big\langle \operatorname{Hess} \mathcal{H}(\rho) \nabla\psi, \nabla\psi \big\rangle_\rho \geq \kappa \|\nabla\psi\|_\rho^2$$

for any function ψ and any probability density $\rho \in \mathscr{P}_*(\mathcal{X})$. As the Riemannian structure is degenerate on the boundary of $\mathscr{P}(\mathcal{X})$, this is not an immediate result, but this issue has been solved in [9], Theorem 4.4. We thus have the following result.

PROPOSITION 2.4 (Characterization of Ricci boundedness). *Let $\kappa \in \mathbb{R}$. A Markov triple $(\mathcal{X}, Q, \pi)$ satisfies* $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$ *if and only if for all $\rho \in \mathscr{P}_*(\mathcal{X})$ and all functions $\psi : \mathcal{X} \to \mathbb{R}$, we have*

$$\mathcal{B}(\rho, \psi) \geq \kappa \mathcal{A}(\rho, \psi). \tag{5}$$

In this paper, we provide a systematic method to prove such an inequality in concrete situations. A very useful criterion will be given in Section 3.

2.3. *The Bakry–Émery approach to the logarithmic Sobolev inequality.* Let us now compare our method with the Bakry–Émery approach to the (modified) logarithmic Sobolev inequality (MLSI), which has been developed in the discrete setting in [4] and [5]. In order to prove the MLSI $\mathcal{H}(\rho) \leq \frac{1}{\alpha}\mathcal{E}(\rho, \log \rho)$, in the Bakry–Émery method one considers the behaviour of the entropy $h(t) := \mathcal{H}(\rho_t)$ for solutions to the heat equation $\partial_t \rho_t = \mathcal{L}\rho_t$. Since $h'(t) = -\mathcal{E}(\rho_t, \log \rho_t)$, the MLSI asserts that $h(t) \leq -\frac{1}{\alpha}h'(t)$. Rather than approaching this inequality directly, the idea is to investigate the second derivative of the entropy. Suppose that one could prove the "convex entropy decay inequality" $h''(t) \geq -\alpha h'(t)$. Since $h(t), h'(t) \to 0$ as $t \to \infty$, this inequality implies the MLSI after integration. Taking into account that

$$h''(t) = \pi[\mathcal{L}\rho_t \mathcal{L} \log \rho_t] + \pi\left[\frac{(\mathcal{L}\rho_t)^2}{\rho_t}\right],$$

the following lemma (taken from [5]) summarises this discussion.

LEMMA 2.5. *Let $\alpha > 0$ and suppose that the convex entropy decay inequality*

$$\pi[\mathcal{L}\rho \mathcal{L} \log \rho] + \pi\left[\frac{(\mathcal{L}\rho)^2}{\rho}\right] \geq \alpha \mathcal{E}(\rho, \log \rho)$$

holds for all $\rho \in \mathscr{P}_*(\mathcal{X})$. Then the modified logarithmic Sobolev inequality $\mathcal{H}(\rho) \leq \frac{1}{\alpha}\mathcal{E}(\rho, \log \rho)$ holds as well. Moreover, (2.5) *implies the exponential convergence bound*

$$\mathcal{E}\big(e^{t\mathcal{L}}\rho, \log e^{t\mathcal{L}}\rho\big) \leq e^{-\alpha t}\mathcal{E}(\rho, \log \rho)$$

for all $\rho \in \mathscr{P}_*(\mathcal{X})$.

The last assertion follows readily from Gronwall's lemma and is actually equivalent to (2.5). A direct calculation shows that

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

hence, in view of Proposition 2.4, the following lemma follows immediately.

LEMMA 2.6. *Let $\kappa \in \mathbb{R}$ and suppose that* $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$. *Then the convex entropy decay inequality* (2.5) *holds with* $\alpha = 2\kappa$.

The condition $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$ is in principle strictly stronger than the convex entropy decay inequality (2.5): we need to check that the inequality $\mathcal{B}(\rho, \psi) \geq \kappa\mathcal{A}(\rho, \psi)$ holds for all functions ψ, and not just for $\psi = \log \rho$. There seems to be no reason why $\psi = \log \rho$ should always be the extremal case in this inequality, but we are unaware of an explicit example which demonstrates this.

In general, the constant we shall obtain here for the lower bound on Ricci curvature is worse (i.e., smaller) than the best known constant for the modified logarithmic Sobolev inequality. This was to be expected, since even in the continuous setting lower bounds on the Ricci curvature have no reason to yield the optimal constant for the logarithmic Sobolev inequality in general.

3. The Bochner method. To explain and apply our method, it will be convenient to use the following point of view. Given an irreducible and reversible Markov triple $(\mathcal{X}, Q, \pi)$, we consider a set G of maps of $\mathcal{X}$ onto itself, that represents the set of allowed moves, and a function $c : \mathcal{X} \times G \to \mathbb{R}_+$, that represents the jump rates.

DEFINITION 3.1. The pair (G, c) is called a mapping representation of $\mathcal{L}$ if the following properties are satisfied:

1. The generator $\mathcal{L}$ can be written as

$$\mathcal{L}\psi(\eta) = \sum_{\delta \in G} \nabla_\delta \psi(\eta) c(\eta, \delta),$$

where $\nabla_\delta \psi(\eta) := \psi(\delta\eta) - \psi(\eta)$.

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

3. For every $F : \mathcal{X} \times G \to \mathbb{R}$, we have

$$\pi\left[\sum_{\delta \in G} F(\eta, \delta) c(\eta, \delta)\right] = \pi\left[\sum_{\delta \in G} F\big(\delta\eta, \delta^{-1}\big) c(\eta, \delta)\right]. \tag{6}$$

Every irreducible reversible Markov chain has a mapping representation. It is always possible to explicitly build one, by considering the set of bijections $t_{\eta,\tilde{\eta}} : \mathcal{X} \to \mathcal{X}$ that exchanges η and $\tilde{\eta}$, and leaves all other points unchanged (see [9]), but in general it is more natural to work with a different, smaller set G, as we shall see in the examples of Section 4. Property 3 expresses the reversibility of the Markov chain.

With this notation, we can rewrite the action functional $\mathcal{A}$ as

$$\mathcal{A}(\rho, \psi) = \frac{1}{2}\pi\left[\sum_{\delta \in G} c(\eta, \delta)\big(\nabla_\delta \psi(\eta)\big)^2 \hat{\rho}(\eta, \delta\eta)\right].$$

As discussed in the Introduction, a key element of the optimal transport approach to curvature in the continuous setting is the Bochner identity $\frac{1}{2}\Delta(|\nabla\psi|^2) - \langle\nabla\psi, \nabla\Delta\psi\rangle = |D^2\psi|^2 + \mathrm{Ric}(\nabla\psi, \nabla\psi)$. This identity allows one to reformulate

$$B(\rho, \psi) := \int_{\mathcal{M}} \frac{1}{2}\Delta\big(|\nabla\psi|^2\big) - \langle\nabla\psi, \nabla\Delta\psi\rangle \,\mathrm{d}\rho,$$

the continuous analogue of $\mathcal{B}$, in the equivalent form

$$B(\rho, \psi) = \int_{\mathcal{M}} \big|D^2\psi\big|^2 + \mathrm{Ric}(\nabla\psi, \nabla\psi)\,\mathrm{d}\rho.$$

Since the first term in this expression is nonnegative, lower bounds on the Ricci curvature immediately translate into lower bounds on $B(\rho, \psi)$.

In the discrete setting, as far as we know, there is no direct analogue of Bochner's identity. To get around this issue (in the context of the Bakry–Émery approach for functional inequalities), it was suggested in [4] and [5] to introduce a function R that satisfies some properties of spatial invariance, in such a way that a one-sided Bochner inequality holds.

ASSUMPTION 3.2. There exists a function $R : \mathcal{X} \times G \times G \to \mathbb{R}_+$ such that

(A1) $R(\eta, \gamma, \delta) = R(\eta, \delta, \gamma)$ for all $\eta \in \mathcal{X}$ and $\gamma, \delta \in G$;
(A2) $\pi[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\psi(\eta, \gamma, \delta)] = \pi[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\psi(\gamma\eta, \gamma^{-1}, \delta)]$ for all bounded functions $\psi : \mathcal{X} \times G \times G \to \mathbb{R}$;
(A3) $\gamma\delta\eta = \delta\gamma\eta$ for all $\eta \in \mathcal{X}$ and $\gamma, \delta \in G$ with $R(\eta, \gamma, \delta) > 0$.

Clearly, (A1) and (A2) imply that for all bounded functions ψ,

$$\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\psi(\eta, \gamma, \delta)\Bigg] = \pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\psi\big(\delta\eta, \gamma, \delta^{-1}\big)\Bigg]. \tag{7}$$

We will show that a function R that satisfies these assumptions automatically satisfies the following identity.

LEMMA 3.3. *Let φ and ψ be two real-valued functions on $\mathcal{X}$, and let $\alpha : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a symmetric function. Then the following identity holds*:

$$\begin{aligned}\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\eta)\Bigg]\\ = \frac{1}{4}\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\gamma\big[\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\big]\nabla_\delta\nabla_\gamma\psi(\eta)\Bigg].\end{aligned}$$

In the case $\alpha \equiv 1$, the right-hand side in this equation can be thought of as a discrete analogue of $\int_{\mathcal{M}} \langle D^2 f, D^2 g\rangle \,\mathrm{d}\rho$. We can therefore think of this identity as a sort of weighted discrete Bochner identity.

REMARK 3.4. This is a generalization of the identity used in [5], where the authors proved this identity in the case $\alpha \equiv 1$. In the applications, we have in mind, we will use this identity with α the logarithmic mean, that is $\alpha(\eta, \delta\eta) = \hat{\rho}(\eta, \delta\eta)$.

PROOF OF LEMMA 3.3. We consider η, δ and γ such that $R(\eta, \gamma, \delta) > 0$. Note that, by assumption (iii), $\delta\gamma\eta = \gamma\delta\eta$. First, we write

$$\begin{aligned}
&\nabla_\gamma\big[\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\big]\nabla_\delta\nabla_\gamma\psi(\eta) \\
&\quad = \alpha(\gamma\eta, \delta\gamma\eta)\nabla_\delta\varphi(\gamma\eta)\nabla_\gamma\psi(\delta\eta) - \alpha(\gamma\eta, \delta\gamma\eta)\nabla_\delta\varphi(\gamma\eta)\nabla_\gamma\psi(\eta) \\
&\qquad - \alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\delta\eta) + \alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\eta).
\end{aligned}$$

We will show that each of the four terms on the right-hand side of this equality, when multiplied by $R(\eta, \gamma, \delta)$, summed over all γ and δ and averaged over π, yields

$$\pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\eta)\bigg].$$

For the fourth term, there is nothing to prove. For the third term, we have

$$\begin{aligned}
&-\pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\delta\eta)\bigg] \\
&\quad = \pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\eta, \delta\eta)\nabla_{\delta^{-1}}\varphi(\delta\eta)\nabla_\gamma\psi(\delta\eta)\bigg] \\
&\quad \overset{(7)}{=} \pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\delta\eta, \eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\eta)\bigg] \\
&\quad = \pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\eta)\bigg].
\end{aligned}$$

The same argument works for the second term. Finally, for the first term, we have

$$\begin{aligned}
&\pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\gamma\eta, \delta\gamma\eta)\nabla_\delta\varphi(\gamma\eta)\nabla_\gamma\psi(\delta\eta)\bigg] \\
&\quad = -\pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\gamma\eta, \delta\gamma\eta)\nabla_\delta\varphi(\gamma\eta)\nabla_{\gamma^{-1}}\psi(\delta\gamma\eta)\bigg] \\
&\quad \overset{(A2)}{=} -\pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\delta\eta)\bigg] \\
&\quad = \pi\bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\alpha(\eta, \delta\eta)\nabla_\delta\varphi(\eta)\nabla_\gamma\psi(\eta)\bigg],
\end{aligned}$$

where the last identity is the one we proved right before for the third term. This completes the proof. □

Following [5], our strategy is to decompose the local weight $c(\eta,\delta)c(\eta,\gamma)$ as $R(\eta,\delta,\gamma)+\Gamma(\eta,\delta,\gamma)$ for a suitable function Γ. Since the contribution of R to the curvature is nonnegative, we will only have to study lower bounds for a functional that depends on Γ.

THEOREM 3.5. *Assume that there exists a function R satisfying Assumption 3.2, and let*

$$\Gamma(\eta,\gamma,\delta) := c(\eta,\gamma)c(\eta,\delta) - R(\eta,\gamma,\delta).$$

Then we have

$$\mathcal{B}(\rho,\psi) \geq (\widetilde{\mathcal{B}}_1 + \widetilde{\mathcal{B}}_2 + \widetilde{\mathcal{B}}_3)(\rho,\psi), \tag{8}$$

where

$$\widetilde{\mathcal{B}}_1(\rho,\psi) := \pi\Big[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\hat{\rho}(\eta,\delta\eta)\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)\Big],$$

$$\widetilde{\mathcal{B}}_2(\rho,\psi) := \frac{1}{2}\pi\Big[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)\Big],$$

$$\widetilde{\mathcal{B}}_3(\rho,\psi) := \frac{1}{4}\pi\Big[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\hat{\rho}(\eta,\delta\eta)\big(\nabla_\gamma\nabla_\delta\psi(\eta)\big)^2\Big]$$

for all $\rho\in\mathscr{P}(\mathcal{X})$ and all $\psi:\mathcal{X}\to\mathbb{R}$.

As a consequence, if $\widetilde{\mathcal{B}}_1+\widetilde{\mathcal{B}}_2+\widetilde{\mathcal{B}}_3 \geq \kappa\mathcal{A}$, then the Ricci curvature of the Markov chain is bounded from below by κ.

Inequality (8) can be viewed as a discrete replacement for Bochner's inequality (2): the quantity $\mathcal{B}(\rho,\psi)$, which is a discrete analogue of the continuous expression $\int_{\mathcal{M}} -\langle\nabla\psi,\nabla\Delta\psi\rangle + \frac{1}{2}\Delta(|\nabla\psi|^2)\,\mathrm{d}\rho$, is estimated in terms of $\widetilde{\mathcal{B}}_3(\rho,\psi)$, a nonnegative expression involving only second-order derivatives, and a number of terms involving first-order derivatives that may be seen as curvature terms.

REMARK 3.6. In most of our applications, we will use the trivial bound $\widetilde{\mathcal{B}}_3 \geq 0$, so that we will only need to obtain a lower bound on the terms $\widetilde{\mathcal{B}}_1, \widetilde{\mathcal{B}}_2$ involving Γ.

REMARK 3.7. This result can be used to recover the criterion of Proposition 5.4 in [9]. Indeed, under their assumptions, we can take $R(\eta,\delta,\gamma) = c(\eta,\delta)c(\eta,\gamma)$ and, therefore, $\Gamma = 0$. The conclusions easily follow. This criterion has been applied in [9] to obtain the sharp Ricci bound for the discrete hypercube, as well as nonnegativity of the Ricci curvature for the discrete cycle $\mathbb{Z}/N\mathbb{Z}$ for $N \geq 2$.

PROOF OF THEOREM 3.5. The quantity $\mathcal{B}(\rho,\psi)$ from Proposition 2.3 can be written as

$$\mathcal{B}(\rho,\psi) = -\langle \hat{\rho}\nabla\psi, \nabla\mathcal{L}\psi\rangle_\pi + \tfrac{1}{2}\langle \hat{\mathcal{L}}\rho\nabla\psi, \nabla\psi\rangle_\pi,$$

where

$$\begin{aligned}
&\langle \hat{\rho}\nabla\psi, \nabla\mathcal{L}\psi\rangle_\pi \\
&\quad = \frac{1}{2}\pi\Big[\sum_{\gamma,\delta}\nabla_\delta\psi(\eta)\big(\nabla_\gamma\psi(\delta\eta)c(\delta\eta,\gamma) - \nabla_\gamma\psi(\eta)c(\eta,\gamma)\big)c(\eta,\delta)\hat{\rho}(\eta,\delta\eta)\Big]
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{2}\langle \hat{\mathcal{L}}\rho\nabla\psi, \nabla\psi\rangle_\pi = \frac{1}{4}\pi\Big[\sum_{\gamma,\delta}&\big(\nabla_\delta\psi(\eta)\big)^2\big(\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)c(\eta,\gamma) \\
&+ \hat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta)c(\delta\eta,\gamma)\big)c(\eta,\delta)\Big].
\end{aligned} \tag{9}$$

Since the Markov chain is reversible, we can use (6) to obtain

$$\begin{aligned}
&\pi\Big[\sum_{\gamma,\delta}\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\delta\eta)c(\delta\eta,\gamma)c(\eta,\delta)\hat{\rho}(\eta,\delta\eta)\Big] \\
&\qquad = -\pi\Big[\sum_{\gamma,\delta}\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)c(\eta,\gamma)c(\eta,\delta)\hat{\rho}(\eta,\delta\eta)\Big],
\end{aligned}$$

so that

$$\begin{aligned}
-\langle \hat{\rho}\nabla\psi, \nabla\mathcal{L}\psi\rangle_\pi &= \pi\Big[\sum_{\gamma,\delta}\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)c(\eta,\gamma)c(\eta,\delta)\hat{\rho}(\eta,\delta\eta)\Big] \\
&= \pi\Big[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)\hat{\rho}(\eta,\delta\eta)\Big] \\
&\quad + \pi\Big[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)\hat{\rho}(\eta,\delta\eta)\Big].
\end{aligned}$$

The first term on the right-hand side of this equality we leave unchanged, and for the second we use Lemma 3.3 with $\alpha = \hat{\rho}$ to get

$$\begin{aligned}
-\langle \hat{\rho}\nabla\psi, \nabla\mathcal{L}\psi\rangle_\pi &= \pi\Big[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)\hat{\rho}(\eta,\delta\eta)\Big] \\
&\quad + \frac{1}{4}\pi\Big[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\nabla_\delta\nabla_\gamma\psi(\eta)\nabla_\gamma\big[\hat{\rho}(\eta,\delta\eta)\nabla_\delta\psi(\eta)\big]\Big].
\end{aligned} \tag{10}$$

We now take a look at the second term in (9). Using the reversibility of the Markov chain and the fact that $\hat{\rho}_2(\delta\eta,\eta)=\hat{\rho}_1(\eta,\delta\eta)$, we obtain

$$\pi\Big[\sum_{\gamma,\delta}\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta)c(\delta\eta,\gamma)c(\eta,\delta)\Big]$$
$$=\pi\Big[\sum_{\gamma,\delta}\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)c(\eta,\gamma)c(\eta,\delta)\Big],$$

and, therefore,

$$\begin{aligned}\frac{1}{2}\langle\hat{\mathcal{L}}\rho\nabla\psi,\nabla\psi\rangle_\pi &= \frac{1}{2}\pi\Big[\sum_{\gamma,\delta}\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)c(\eta,\gamma)c(\eta,\delta)\Big]\\ &= \frac{1}{2}\pi\Big[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)\Big]\\ &\quad+\frac{1}{2}\pi\Big[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)\Big]\\ &=: T_1+T_2.\end{aligned}$$

Again, we leave the first term on the right-hand side of this last equation unchanged. For the second term, we use assumption (A2) to obtain

$$T_2=\frac{1}{2}\pi\Big[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta)\Big],$$

and thus

$$T_2=\frac{1}{4}\pi\Big[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\big(\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)+\hat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta)\big)\Big].$$

From (i) and (ii) of Lemma A.1 we deduce the inequality

$$\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)+\hat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta)\geq\nabla_\gamma\hat{\rho}(\eta,\delta\eta).$$

Therefore, we have

$$\begin{aligned}&\frac{1}{2}\langle\hat{\mathcal{L}}\rho\nabla\psi,\nabla\psi\rangle_\pi\\ &\quad\geq\frac{1}{2}\pi\Big[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)\Big]\\ &\qquad+\frac{1}{4}\pi\Big[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\nabla_\gamma\hat{\rho}(\eta,\delta\eta)\Big].\end{aligned}\tag{11}$$

To deduce our result from (10) and (11), all that is left is to show that

$$
\begin{aligned}
S &:= \pi\Big[\sum_{\gamma,\delta} R(\eta,\gamma,\delta)\big(\big(\nabla_\delta\psi(\eta)\big)^2\nabla_\gamma\hat\rho(\eta,\delta\eta) \\
&\qquad + \nabla_\delta\nabla_\gamma\psi(\eta)\nabla_\gamma\big[\hat\rho(\eta,\delta\eta)\nabla_\delta\psi(\eta)\big]\big)\Big] \qquad (12)\\
&= \pi\Big[\sum_{\gamma,\delta} R(\eta,\gamma,\delta)\hat\rho(\eta,\delta\eta)\big(\nabla_\gamma\nabla_\delta\psi(\eta)\big)^2\Big].
\end{aligned}
$$

We observe that

$$
\begin{aligned}
&\big(\nabla_\delta\psi(\eta)\big)^2\nabla_\gamma\hat\rho(\eta,\delta\eta) + \nabla_\delta\nabla_\gamma\psi(\eta)\nabla_\gamma\big[\hat\rho(\eta,\delta\eta)\nabla_\delta\psi(\eta)\big]\\
&\quad = \hat\rho(\gamma\eta,\delta\gamma\eta)\big(\nabla_\gamma\nabla_\delta\psi(\eta)\big)^2\\
&\qquad - \hat\rho(\eta,\delta\eta)\nabla_\delta\psi(\gamma\eta)\nabla_\delta\psi(\eta) + \hat\rho(\gamma\eta,\delta\gamma\eta)\nabla_\delta\psi(\gamma\eta)\nabla_\delta\psi(\eta),
\end{aligned}
$$

for all $\eta \in \mathcal{X}$ and $\gamma,\delta \in G$ with $\delta\gamma\eta = \gamma\delta\eta$. In view of (A3), we may insert this identity in the expression for S, which yields

$$
\begin{aligned}
S &= \pi\Big[\sum_{\gamma,\delta} R(\eta,\gamma,\delta)\hat\rho(\gamma\eta,\delta\gamma\eta)\big(\nabla_\gamma\nabla_\delta\psi(\eta)\big)^2\Big]\\
&\quad - \pi\Big[\sum_{\gamma,\delta} R(\eta,\gamma,\delta)\hat\rho(\eta,\delta\eta)\nabla_\delta\psi(\gamma\eta)\nabla_\delta\psi(\eta)\Big]\\
&\quad + \pi\Big[\sum_{\gamma,\delta} R(\eta,\gamma,\delta)\hat\rho(\gamma\eta,\delta\gamma\eta)\nabla_\delta\psi(\gamma\eta)\nabla_\delta\psi(\eta)\Big].
\end{aligned}
$$

It follows from (A2) that the second and third term in the latter formula cancel each other. Another application of (A2) shows that the first term equals the right-hand side of (12), which was the last element we needed to complete the proof. □

4. Examples.

4.1. *Birth and death processes.* In this section, we consider the case of birth and death processes on $\mathbb{N} = \{0, 1, \ldots\}$. These are Markov chains with generator

$$
\mathcal{L}\psi(n) = a(n)\big(\psi(n+1) - \psi(n)\big) + b(n)\big(\psi(n-1) - \psi(n)\big),
$$

where a and b are nonnegative functions on $\mathbb{N}$, such that $b(0) = 0$. The set of allowed moves is $G := \{+,-\}$, where $+(n) = n+1$ and $-(n) = (n-1)\mathbb{1}_{\{n>0\}}$. Following the notation of the previous section, we write $\nabla_\pm\psi(n) = \psi(n\pm 1) - \psi(n)$. The generator can therefore be written as

$$
\mathcal{L}\psi(n) = a(n)\nabla_+\psi(n) + b(n)\nabla_-\psi(n),
$$

and, in accordance with our notation, we set $c(n,+) = a(n)$ and $c(n,-) = b(n)$.

We assume that this Markov chain is irreducible, and that there exists a probability measure π on $\mathbb{N}$ satisfying the detailed balance condition

$$a(n)\pi(n) = b(n+1)\pi(n+1).$$

Clearly, the latter condition is satisfied if and only if

$$\sum_{n=1}^{\infty} \frac{a(n-1)\cdots a(0)}{b(n)\cdots b(1)} < \infty.$$

When applying our method to study curvature bounds for this Markov chain, we obtain the following result.

THEOREM 4.1. *Let $\kappa \in \mathbb{R}_+$. Assume that the rate of birth a is nonincreasing, and that the rate of death b is nondecreasing. Assume moreover that*

$$\begin{aligned} &\tfrac{1}{2}\bigl(a(n) - a(n+1) + b(n+1) - b(n)\bigr) \\ &\quad + \tfrac{1}{2}\Theta\bigl(a(n) - a(n+1), b(n+1) - b(n)\bigr) \geq \kappa \end{aligned} \tag{13}$$

for all $n \in \mathbb{N}$, where

$$\Theta(\alpha, \beta) = \inf_{s,t>0} \theta(s,t)\left(\frac{\alpha}{s} + \frac{\beta}{t}\right).$$

Under these assumptions, the birth and death process has Ricci curvature bounded from below by κ.

REMARK 4.2. Strictly speaking, to include this example in our framework we need to assume that the transition rates $a(n), b(n)$ vanish whenever n is sufficiently large, so that the state spaces becomes finite. However, it is reasonable to expect that the result below and its proof remain valid without this assumption. A rigorous inclusion of the countable setting would require us to modify some technical arguments from [9, 18], which is beyond the scope of the present paper.

REMARK 4.3. The same criterion has been obtained by Mielke [20] using a different method based on diagonal dominance of the matrix representing the Hessian of the entropy. Mielke worked on a finite state space $\{0, 1, \ldots, N\}$ rather than on $\mathbb{N}$, but this does not effect the argument.

PROOF OF THEOREM 4.1. First of all, using the reversibility condition, we can rewrite the action functional as

$$\begin{aligned} \mathcal{A}(\rho, \psi) &= \frac{1}{2}\pi\bigl[a(n)\bigl(\nabla_+\psi(n)\bigr)^2\hat{\rho}(n, n+1) + b(n)\bigl(\nabla_-\psi(n)\bigr)^2\hat{\rho}(n, n-1)\bigr] \\ &= \pi\bigl[a(n)\bigl(\nabla_+\psi(n)\bigr)^2\hat{\rho}(n, n+1)\bigr]. \end{aligned}$$

Following [5], we define the function $R : \mathbb{N} \times \{+,-\}^2 \to \mathbb{R}$ by

$$\begin{cases} R(n,+,+) = a(n)a(n+1), \\ R(n,-,-) = b(n)b(n-1), \\ R(n,+,-) = R(n,-,+) = a(n)b(n) \end{cases}$$

so that $\Gamma(n,+,+) = -a(n)\nabla_+ a(n)$, $\Gamma(n,-,-) = -b(n)\nabla_- b(n)$, and $\Gamma(n,+,-) = \Gamma(n,-,+) = 0$. It is not hard to check that Assumption 3.2 is satisfied. In this case, we have

$$\begin{aligned} \widetilde{\mathcal{B}}_1(\rho,\psi) &= \pi\Big[\sum_{\gamma,\delta} \Gamma(\eta,\gamma,\delta)\hat{\rho}(\eta,\delta\eta)\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)\Big] \\ &= -\pi\big[a(n)\nabla_+ a(n)\big(\nabla_+\psi(n)\big)^2\hat{\rho}(n,n+1) \\ &\quad + b(n)\nabla_- b(n)\big(\nabla_-\psi(n)\big)^2\hat{\rho}(n,n-1)\big] \\ &= \pi\big[a(n)\big(\nabla_+ b(n) - \nabla_+ a(n)\big)\big(\nabla_+\psi(n)\big)^2\hat{\rho}(n,n+1)\big]. \end{aligned} \tag{14}$$

Moreover, using reversibility,

$$\begin{aligned} \widetilde{\mathcal{B}}_2(\rho,\psi) &= \frac{1}{2}\pi\Big[\sum_{\gamma,\delta} \Gamma(\eta,\gamma,\delta)\big(\nabla_\delta\psi(\eta)\big)^2\hat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)\Big] \\ &= -\frac{1}{2}\pi\big[a(n)\nabla_+ a(n)\big(\nabla_+\psi(n)\big)^2\hat{\rho}_1(n,n+1)\nabla_+\rho(n) \\ &\quad + b(n)\nabla_- b(n)\big(\nabla_-\psi(n)\big)^2\hat{\rho}_1(n,n-1)\nabla_-\rho(n)\big] \\ &= -\frac{1}{2}\pi\big[a(n)\nabla_+ a(n)\big(\nabla_+\psi(n)\big)^2\hat{\rho}_1(n,n+1)\nabla_+\rho(n) \\ &\quad + a(n)\nabla_+ b(n)\big(\nabla_+\psi(n)\big)^2\hat{\rho}_2(n,n+1)\nabla_+\rho(n)\big]. \end{aligned}$$

A direct computation shows that the partial derivatives $\theta_i(s,t) = \partial_i\theta(s,t)$ satisfy the identities

$$(t-s)\theta_1(s,t) = \frac{\theta(s,t)^2}{s} - \theta(s,t), \qquad (s-t)\theta_2(s,t) = \frac{\theta(s,t)^2}{t} - \theta(s,t),$$

for any $s,t > 0$. Therefore, recalling that $\hat{\rho}(n,n+1) = \theta(\rho(n),\rho(n+1))$, we have

$$\begin{aligned} \widetilde{\mathcal{B}}_2(\rho,\psi) &= -\frac{1}{2}\pi\big[a(n)\big(\nabla_+ b(n) - \nabla_+ a(n)\big)\big(\nabla_+\psi(n)\big)^2\hat{\rho}(n,n+1)\big] \\ &\quad + \frac{1}{2}\pi\Big[a(n)\Big(\frac{\nabla_+ b(n)}{\rho(n+1)} - \frac{\nabla_+ a(n)}{\rho(n)}\Big)\big(\nabla_+\psi(n)\big)^2\hat{\rho}(n,n+1)^2\Big]. \end{aligned} \tag{15}$$

Using the definition of Θ, we obtain the inequality

$$\frac{\nabla_+ b(n)}{\rho(n+1)} - \frac{\nabla_+ a(n)}{\rho(n)} \geq \frac{\Theta(-\nabla_+ a(n), \nabla_+ b(n))}{\hat{\rho}(n,n+1)}.$$

Summing (14) and (15), and substituting the latter inequality, it follows from the assumption (13) that

$$(\widetilde{\mathcal{B}}_1 + \widetilde{\mathcal{B}}_2)(\rho, \psi) \geq \kappa \mathcal{A}(\rho, \psi).$$

The result follows by an application of Theorem 3.5. $\square$

4.2. *Zero-range processes.* We now consider interacting particle systems on the complete graph, whose sites are labeled $1, \dots, L$, where L is the number of sites. Configurations of particles are given by collections of nonnegative integers $\eta_1, \dots, \eta_L$, where η_x denotes the number of particles at site x. The whole configuration will be denoted by $\eta \in S := \mathbb{N}^L$.

Changes in the configuration are made by moving a particle (if any) from a site x to a different site y. Allowed moves are therefore given by maps of the form $\eta \mapsto \eta^{xy}$, where $x \neq y$ are two different sites, and η^{xy} is given by

$$\begin{cases} \eta_z^{xy} = \eta_z, & \text{if } z \notin \{x, y\} \text{ or } \eta_x = 0, \\ \eta_x^{xy} = \eta_x - 1, & \text{if } \eta_x > 0, \\ \eta_y^{xy} = \eta_y + 1, & \text{if } \eta_x > 0. \end{cases}$$

The set of allowed moves is $G := \{xy; x \neq y\}$, where xy denotes the map $\eta \mapsto \eta^{xy}$. We write ∇_{xy} to denote the corresponding discrete gradient, and note that $(xy)^{-1} = (yx)$ in the sense of Definition 3.1.

In this section, we consider zero-range processes on the complete graph with L vertices. In such processes, the jump rate from site x to site y depends only on x and on the number of particles present at x. The rates are thus given by a family of functions $c_x : \mathbb{N} \to [0, \infty)$ such that $c_x(0) = 0$ and $c_x(n) > 0$ for all $n > 0$ and $x \in \{1, \dots, L\}$. Here, $c_x(\eta_x)$ is the rate at which a particle is moved from site x to a site y, with y chosen randomly, with uniform probability on $\{1, \dots, L\}$. The generator of the Markov chain we just described is

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

This dynamic conserves the total number of particles $N := \sum_x \eta_x$. We define a probability measure π_N on configurations with N particles as

$$\pi_N(\eta) := \frac{1}{Z_N} \prod_{x=1}^{L} \prod_{k=1}^{\eta_x} \frac{1}{c_x(k)},$$

with the usual convention that the latter product equals 1 whenever $\eta_x = 0$. The constant Z_N is a normalization constant, which is indeed finite, since there are a finite number of configurations with N particles. It can then be checked that the Markov chain is reversible with respect to π_N. In the sequel, we shall systematically omit the subscript N, which we consider fixed.

When applying our method to study the curvature of this process, we obtain the following result.

THEOREM 4.4. *Assume that there exists* $c > 0$ *and* $\delta \in [0, 2c]$ *such that*

$$c \le c_x(n+1) - c_x(n) \le c + \delta \qquad \text{for all } x \in \{1, \dots, L\} \text{ and } n \ge 0. \tag{16}$$

Then the Ricci curvature is bounded from below by $\frac{c}{2} - \frac{5\delta}{4}$. *In particular, if* $\delta < \frac{2}{5}c$, *then we obtain a positive lower bound on the Ricci curvature.*

Note that the assumption implies that the rates c_x are strictly increasing. Of particular interest is the fact that the bound does not depend on either L or N.

REMARK 4.5. Simply assuming that the rates are strictly increasing is not enough to ensure Ricci curvature is positive. In Section 4.2 of [5], a simple example of a zero-range process for which $\sup c_x(1) - \inf c_x(1)$ is large was shown to exhibit nonconvex decay of the entropy and, therefore, the Ricci curvature of the Markov kernel cannot be positive. So some assumption of the form $\delta < Kc$ is necessary. On the other hand, we do not know if the assumption $\delta < 2c/5$ is optimal. Assuming a uniform positive lower bound on the increase of the rates is known to be necessary to obtain a lower bound on the curvature that is independent of the system size and number of particles. Indeed, it was shown in [21] that for the case of constant rates, the spectral gap is of order $\frac{L^2}{L^2+N^2}$.

REMARK 4.6. For independent random walks on the complete graph (which corresponds to $\delta = 0$ and $c = 1$), we recover the size-independent lower bound on Ricci curvature that was obtained in Example 5.1 of [9].

The estimate on the modified logarithmic Sobolev inequality that can be deduced from the curvature bound has a worse dependence on δ than the one obtained in [5] [which would correspond to a constant $(c-\delta)/2$]. The modified logarithmic Sobolev inequality for this model has also been studied in [6].

PROOF OF THEOREM 4.4. First, we observe that

$$\mathcal{A}(\rho, \psi) = \frac{1}{2L}\pi\Bigg[\sum_{x,y} c_x(\eta_x)\big(\nabla_{xy}\psi(\eta)\big)^2 \hat{\rho}\big(\eta, \eta^{xy}\big)\Bigg].$$

Following [5], we define the function R as

$$R(\eta, xy, uv) = \begin{cases} \dfrac{1}{L^2} c_x(\eta_x) c_u(\eta_u), & \text{for } x \neq u, \\ \dfrac{1}{L^2} c_x(\eta_x) c_x(\eta_x - 1), & \text{otherwise,} \end{cases}$$

with the understanding that $c_x(-1) := 0$, so that $\Gamma(\eta, xy, uv) = 0$ if $x \neq u$, and $\Gamma(\eta, xy, xv) = \frac{1}{L^2} c_x(\eta_x)(c_x(\eta_x) - c_x(\eta_x - 1))$. It follows that

$$\begin{aligned}
\widetilde{\mathcal{B}}_2(\rho, \psi) &= \frac{1}{2}\pi\Big[\sum_{\gamma,\delta} \Gamma(\eta, \gamma, \delta)\big(\nabla_\delta \psi(\eta)\big)^2 \hat{\rho}_1(\eta, \delta\eta)\nabla_\gamma \rho(\eta)\Big] \\
&= \frac{1}{2L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\big(\nabla_{xy}\psi(\eta)\big)^2 \\
&\qquad \times \hat{\rho}_1\big(\eta, \eta^{xy}\big)\nabla_{xv}\rho(\eta)\Big].
\end{aligned} \tag{17}$$

Using the reversibility assumption, this quantity is also equal to

$$\begin{aligned}
&\frac{1}{2L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(c_y(\eta_y + 1) - c_y(\eta_y)\big)\big(\nabla_{xy}\psi(\eta)\big)^2 \hat{\rho}_2\big(\eta, \eta^{xy}\big) \\
&\qquad \times \big(\rho\big(\eta^{xv}\big) - \rho\big(\eta^{xy}\big)\big)\Big].
\end{aligned}$$

Adding this identity to (17), we obtain

$$\begin{aligned}
&\widetilde{\mathcal{B}}_2(\rho, \psi) \\
&\quad = \frac{1}{4L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xy}\psi(\eta)\big)^2 \\
&\qquad \times \big(\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\hat{\rho}_1\big(\eta, \eta^{xy}\big)\nabla_{xv}\rho(\eta) \\
&\qquad + \big(c_y(\eta_y + 1) - c_y(\eta_y)\big)\hat{\rho}_2\big(\eta, \eta^{xy}\big)\big(\rho\big(\eta^{xv}\big) - \rho\big(\eta^{xy}\big)\big)\big)\Big] \\
&\quad = \frac{1}{4L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xy}\psi(\eta)\big)^2 \rho\big(\eta^{xv}\big) \\
&\qquad \times \big(\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\hat{\rho}_1\big(\eta, \eta^{xy}\big) + \big(c_y(\eta_y + 1) - c_y(\eta_y)\big)\hat{\rho}_2\big(\eta, \eta^{xy}\big)\big)\Big] \\
&\qquad - \frac{1}{4L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xy}\psi(\eta)\big)^2\big(\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\hat{\rho}_1\big(\eta, \eta^{xy}\big)\rho(\eta) \\
&\qquad + \big(c_y(\eta_y + 1) - c_y(\eta_y)\big)\hat{\rho}_2\big(\eta, \eta^{xy}\big)\rho\big(\eta^{xy}\big)\big)\Big] \\
&\quad =: I_1 + I_2.
\end{aligned}$$

Taking into account that $\hat\rho_i \geq 0$, we use (16) and the assumption that $c \geq \delta/2$ to obtain the lower bound

$$
\begin{aligned}
I_1 &\geq \frac{c}{4L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xy}\psi(\eta)\big)^2\rho\big(\eta^{xv}\big)\big(\hat\rho_1\big(\eta,\eta^{xy}\big)+\hat\rho_2\big(\eta,\eta^{xy}\big)\big)\Big]\\
&\overset{y\leftrightarrow v}{\geq} \frac{\delta}{8L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xv}\psi(\eta)\big)^2\rho\big(\eta^{xy}\big)\big(\hat\rho_1\big(\eta,\eta^{xv}\big)+\hat\rho_2\big(\eta,\eta^{xv}\big)\big)\Big]\\
&=: I_3.
\end{aligned}
$$

This quantity is nonnegative, but we will need it later to compensate a term that appears when computing $\widetilde{\mathcal{B}}_1(\rho,\psi)$. Using (16) once more and Lemma A.1(ii), we obtain

$$
\begin{aligned}
I_2 &\geq -\frac{c+\delta}{4L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xy}\psi(\eta)\big)^2\big(\hat\rho_1\big(\eta,\eta^{xy}\big)\rho(\eta)+\hat\rho_2\big(\eta,\eta^{xy}\big)\rho\big(\eta^{xy}\big)\big)\Big]\\
&= -\frac{c+\delta}{4L}\pi\Big[\sum_{x,y} c_x(\eta_x)\big(\nabla_{xy}\psi(\eta)\big)^2\hat\rho\big(\eta,\eta^{xy}\big)\Big]\\
&= -\frac{c+\delta}{2}\mathcal{A}(\rho,\psi),
\end{aligned}
$$

so that

$$
\widetilde{\mathcal{B}}_2(\rho,\psi) \geq -\frac{c+\delta}{2}\mathcal{A}(\rho,\psi)+I_3. \tag{18}
$$

Next, we observe that

$$
\begin{aligned}
\widetilde{\mathcal{B}}_1(\rho,\psi) &= \pi\Big[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\hat\rho(\eta,\delta\eta)\nabla_\delta\psi(\eta)\nabla_\gamma\psi(\eta)\Big]\\
&= \frac{1}{L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x)-c_x(\eta_x-1)\big)\nabla_{xy}\psi(\eta)\nabla_{xv}\psi(\eta)\hat\rho\big(\eta,\eta^{xy}\big)\Big].
\end{aligned}
$$

By another application of (16) and reversibility,

$$
\begin{aligned}
\widetilde{\mathcal{B}}_1(\rho,\psi) &= \frac{1}{L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x)-c_x(\eta_x-1)\big)\\
&\qquad\times\nabla_{xy}\psi(\eta)\big(\nabla_{xy}\psi(\eta)+\nabla_{yv}\psi\big(\eta^{xy}\big)\big)\hat\rho\big(\eta,\eta^{xy}\big)\Big]\\
&\geq 2c\mathcal{A}(\rho,\psi)+\frac{1}{L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x)-c_x(\eta_x-1)\big)\nabla_{xy}\psi(\eta)\\
&\qquad\times\nabla_{yv}\psi\big(\eta^{xy}\big)\hat\rho\big(\eta,\eta^{xy}\big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&=: 2c\mathcal{A}(\rho,\psi) - \frac{1}{L^2}\pi\Big[\sum_{x,v,y} c_x(\eta_x)\big(c_y(\eta_y+1)-c_y(\eta_y)\big)\\
&\qquad \times \nabla_{xy}\psi(\eta)\nabla_{xv}\psi(\eta)\hat{\rho}\big(\eta,\eta^{xy}\big)\Big].
\end{aligned}
$$

Averaging both expressions for $\widetilde{\mathcal{B}}_1(\rho,\psi)$, applying (16), and using the inequality $-2|ab| \geq -a^2-b^2$, we obtain

$$
\begin{aligned}
&\widetilde{\mathcal{B}}_1(\rho,\psi)\\
&\quad \geq c\mathcal{A}(\rho,\psi)\\
&\qquad + \frac{1}{2L^2}\Big[\sum_{x,v,y} c_x(\eta_x)\big(c_x(\eta_x)-c_x(\eta_x-1)-\big(c_y(\eta_y+1)-c_y(\eta_y)\big)\big)\\
&\qquad \times \nabla_{xy}\psi(\eta)\nabla_{xv}\psi(\eta)\hat{\rho}\big(\eta,\eta^{xy}\big)\Big]\\
&\quad \geq \Big(c-\frac{\delta}{2}\Big)\mathcal{A}(\rho,\psi) - \frac{\delta}{4L^2}\pi\Big[\sum_{x,v,y} c_x(\eta_x)\big(\nabla_{xv}\psi(\eta)\big)^2\hat{\rho}\big(\eta,\eta^{xy}\big)\Big].
\end{aligned}
\tag{19}
$$

Adding (18) and (19) and applying Theorem 3.5, we obtain

$$
\mathcal{B}(\rho,\psi) \geq \Big(\frac{c}{2}-\delta\Big)\mathcal{A}(\rho,\psi) + I_4, \tag{20}
$$

where

$$
I_4 := I_3 - \frac{\delta}{4L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xv}\psi(\eta)\big)^2\hat{\rho}\big(\eta,\eta^{xy}\big)\Big].
$$

Since reversibility yields

$$
\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xv}\psi(\eta)\big)^2\hat{\rho}\big(\eta,\eta^{xy}\big)\Big] = \pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xv}\psi(\eta)\big)^2\hat{\rho}\big(\eta^{xv},\eta^{xy}\big)\Big],
$$

we obtain after averaging,

$$
\begin{aligned}
I_4 = \frac{\delta}{8L^2}\pi\Big[&\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xv}\psi(\eta)\big)^2\big(\rho\big(\eta^{xy}\big)\big(\hat{\rho}_1\big(\eta,\eta^{xv}\big)\\
&+\hat{\rho}_2\big(\eta,\eta^{xv}\big)\big) - \hat{\rho}\big(\eta,\eta^{xy}\big) - \hat{\rho}\big(\eta^{xv},\eta^{xy}\big)\big)\Big].
\end{aligned}
$$

Applying Lemma A.2 with $r=\rho(\eta^{xy})$, $s=\rho(\eta)$ and $t=\rho(\eta^{xv})$, we infer that

$$
I_4 \geq -\frac{\delta}{8L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)\big(\nabla_{xv}\psi(\eta)\big)^2\hat{\rho}\big(\eta,\eta^{xv}\big)\Big] = -\frac{\delta}{4}\mathcal{A}(\rho,\psi).
$$

In view of (20), we obtain the desired result. $\square$

4.3. *Bernoulli–Laplace models.* We consider the exclusion process on the complete graph with L sites and N particles, where $1 \le N < L$. The moves are of the form $\eta \mapsto \eta^{xy}$, where

$$\begin{cases} \eta_z^{xy} = \eta_z, & \text{if } z \notin \{x, y\}, \\ \eta_x^{xy} = 0, & \text{if } \eta_x(1-\eta_y) = 1, \text{ otherwise } \eta_x^{xy} = \eta_x, \\ \eta_y^{xy} = 1, & \text{if } \eta_x(1-\eta_y) = 1, \text{ otherwise } \eta_y^{xy} = \eta_y. \end{cases}$$

Note that such moves conserve the total number of particles. The state space S consists of all configurations with at most one particle on each of the L sites, that is, $S = \{\eta \in \{0,1\}^L : \sum_x \eta_x = N\}$. The set of moves is given by $G := \{(x, y) \in \{1, \dots, L\}^2 : x \neq y\}$. To simplify notation, we will frequently write xy instead of (x, y). The transition rates $q : S \times G \to \mathbb{R}_+$ are given by

$$q(\eta, xy) = \frac{\lambda_x}{L}\eta_x(1-\eta_y),$$

hence the generator $\mathcal{L}$ takes the form

$$\mathcal{L}\psi(\eta) = \frac{1}{L}\sum_{xy\in G}\lambda_x\eta_x(1-\eta_y)\nabla_{xy}\psi(\eta),$$

where $\nabla_{xy}\psi(\eta) = \psi(\eta^{xy}) - \psi(\eta)$. We shall frequently use reversibility in the form

$$\pi\Big[\sum_{xy\in G} F(\eta, xy)q(\eta, xy)\Big] = \pi\Big[\sum_{xy\in G} F(\eta^{xy}, yx)q(\eta, xy)\Big] \tag{21}$$

for arbitrary functions $F : S \times G \to \mathbb{R}$. We observe that

$$\mathcal{A}(\rho, \psi) = \frac{1}{2L}\pi\Big[\sum_{x,y}\lambda_x(\nabla_{xy}\psi(\eta))^2\hat{\rho}(\eta, \eta^{xy})\Big].$$

Note that we can omit the factor $\eta_x(1-\eta_y)$, since $\nabla_{xy}\psi(\eta) = 0$ whenever $\eta_x(1-\eta_y) = 0$. This remark will allow us to forget this factor in most of the expressions that we shall manipulate.

We obtain the following result on the Ricci curvature for Bernoulli–Laplace models.

THEOREM 4.7. *Let $1 \le N < L-1$, and assume that there exists $c > 0$ and $\delta \in [0, 2c]$ such that*

$$c \le \lambda_x \le c + \delta \qquad \textit{for all } x \in \{1, \dots, L\}. \tag{22}$$

Then the Ricci curvature of the Bernoulli–Laplace model is bounded from below by $\frac{c}{2} - \frac{7\delta}{8}$. In particular, when $\delta < \frac{4}{7}c$, the Ricci curvature is positive.

In the special case where $\delta = 0$, we recover the result obtained in [11] using a different method, with exactly the same constant.

REMARK 4.8. If $\delta < \frac{4}{7}c$, our result yields a modified logarithmic Sobolev inequality with constant $c - \frac{7\delta}{4}$. Once more, this bound is slightly weaker than the one obtained in [5] (which is $c-\delta$ in our notation). Note however that both bounds coincide for the homogeneous model, in which $\delta = 0$. The modified logarithmic Sobolev inequality for the homogeneous model has also been studied in [2, 12, 14].

PROOF OF THEOREM 4.7. As in [5], we take

$$R(\eta, xy, uv) = \begin{cases} \frac{1}{L^2}\lambda_x\lambda_u\eta_x(1-\eta_y)\eta_u(1-\eta_v), & \text{for } \left|\{x,y,u,v\}\right| = 4, \\ 0, & \text{otherwise.} \end{cases}$$

It follows from the definition that $\Gamma(\eta, xy, uv) = 0$ when $|\{x,y,u,v\}| = 4$ and $\Gamma(\eta, xy, uv) = \frac{1}{L^2}\lambda_x\lambda_u\eta_x(1-\eta_y)\eta_u(1-\eta_v)$ otherwise. We also notice that $\nabla_{xy}\varphi\nabla_{yz}\psi = \nabla_{xy}\varphi\nabla_{zx}\psi = 0$ for any choice of x, y and z, and for any φ and ψ. We then have

$$\begin{aligned}
\widetilde{\mathcal{B}}_2(\rho,\psi) &= \frac{1}{2L^2}\pi\Big[\sum_{x,y}\lambda_x^2\big(\nabla_{xy}\psi(\eta)\big)^2\hat{\rho}_1\big(\eta,\eta^{xy}\big)\nabla_{xy}\rho(\eta)\Big] \\
&\quad + \frac{1}{2L^2}\pi\Big[\sum_{|\{x,y,u\}|=3}\lambda_x\lambda_u\big(\nabla_{xy}\psi(\eta)\big)^2\hat{\rho}_1\big(\eta,\eta^{xy}\big)\nabla_{uy}\rho(\eta)\Big] \\
&\quad + \frac{1}{2L^2}\pi\Big[\sum_{|\{x,y,v\}|=3}\lambda_x^2\big(\nabla_{xy}\psi(\eta)\big)^2\hat{\rho}_1\big(\eta,\eta^{xy}\big)\nabla_{xv}\rho(\eta)\Big] \\
&=: J_1 + J_2 + J_3.
\end{aligned}$$

Using reversibility in the form of (21), we obtain, after averaging with the original expression,

$$\begin{aligned}
J_1 &= -\frac{1}{2L^2}\pi\Big[\sum_{x,y}\lambda_x\lambda_y\big(\nabla_{xy}\psi(\eta)\big)^2\hat{\rho}_2\big(\eta,\eta^{xy}\big)\nabla_{xy}\rho(\eta)\Big] \\
&= \frac{1}{4L^2}\pi\Big[\sum_{x,y}\lambda_x\big(\nabla_{xy}\psi(\eta)\big)^2\big(\lambda_x\hat{\rho}_1\big(\eta,\eta^{xy}\big) - \lambda_y\hat{\rho}_2\big(\eta,\eta^{xy}\big)\big)\big(\rho\big(\eta^{xy}\big)-\rho(\eta)\big)\Big] \\
&\geq -\frac{\delta}{4L^2}\pi\Big[\sum_{x,y}\lambda_x\big(\nabla_{xy}\psi(\eta)\big)^2\hat{\rho}\big(\eta,\eta^{xy}\big)\Big] \\
&= -\frac{\delta}{2L}\mathcal{A}(\rho,\psi),
\end{aligned}$$

where the inequality is obtained using (24) and (22).

Another application of reversibility and averaging yields

$$
\begin{aligned}
J_3 &= \frac{1}{2L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x\lambda_y\big(\nabla_{xy}\psi(\eta)\big)^2\hat{\rho}_2\big(\eta,\eta^{xy}\big)\big(\rho\big(\eta^{xv}\big)-\rho\big(\eta^{xy}\big)\big)\Big] \\
&= \frac{1}{4L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x\big(\nabla_{xy}\psi(\eta)\big)^2\big(\lambda_x\hat{\rho}_1\big(\eta,\eta^{xy}\big)\nabla_{xv}\rho(\eta) \\
&\qquad + \lambda_y\hat{\rho}_2\big(\eta,\eta^{xy}\big)\big(\rho\big(\eta^{xv}\big)-\rho\big(\eta^{xy}\big)\big)\big)\Big] \\
&= \frac{1}{4L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \eta_x(1-\eta_y)(1-\eta_v)\lambda_x\big(\nabla_{xy}\psi(\eta)\big)^2\rho\big(\eta^{xv}\big) \\
&\qquad \times\big(\lambda_x\hat{\rho}_1\big(\eta,\eta^{xy}\big)+\lambda_y\hat{\rho}_2\big(\eta,\eta^{xy}\big)\big)\Big] \\
&\qquad - \frac{1}{4L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \eta_x(1-\eta_y)(1-\eta_v)\lambda_x\big(\nabla_{xy}\psi(\eta)\big)^2 \\
&\qquad \times\big(\lambda_x\hat{\rho}_1\big(\eta,\eta^{xy}\big)\rho(\eta)+\lambda_y\hat{\rho}_2\big(\eta,\eta^{xy}\big)\rho\big(\eta^{xy}\big)\big)\Big].
\end{aligned}
$$

Using the inequality $\lambda_x, \lambda_y \geq c \geq \delta/2$ and Lemma A.1(i), we infer that

$$
\lambda_x\hat{\rho}_1\big(\eta,\eta^{xy}\big)+\lambda_y\hat{\rho}_2\big(\eta,\eta^{xy}\big) \geq \frac{\delta}{2}\big(\hat{\rho}_1\big(\eta,\eta^{xy}\big)+\hat{\rho}_2\big(\eta,\eta^{xy}\big)\big),
$$

$$
\lambda_x\hat{\rho}_1\big(\eta,\eta^{xy}\big)\rho(\eta)+\lambda_y\hat{\rho}_2\big(\eta,\eta^{xy}\big)\rho\big(\eta^{xy}\big) \geq c\hat{\rho}\big(\eta,\eta^{xy}\big).
$$

Applying these bounds, we arrive at

$$
\begin{aligned}
J_3 &\geq \frac{\delta}{8L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \eta_x(1-\eta_y)(1-\eta_v)\lambda_x\big(\nabla_{xy}\psi(\eta)\big)^2\rho\big(\eta^{xv}\big) \\
&\qquad \times\big(\hat{\rho}_1\big(\eta,\eta^{xy}\big)+\hat{\rho}_2\big(\eta,\eta^{xy}\big)\big)\Big] \\
&\qquad - \frac{(L-N-1)(c+\delta)}{2L}\mathcal{A}(\rho,\psi),
\end{aligned}
$$

where the $(L-N-1)$ factor appears because, if there is a particle at x and no particle at y, there are exactly $(L-N-1)$ possible sites v different from x and y where there are no particles.

Similarly, we have

$$
\begin{aligned}
J_2 &= \frac{1}{2L^2}\pi\Big[\sum_{|\{x,y,u\}|=3} \eta_u\lambda_x\lambda_u\big(\nabla_{xy}\psi(\eta)\big)^2\hat{\rho}_1\big(\eta,\eta^{xy}\big)\nabla_{uy}\rho(\eta)\Big] \\
&= \frac{1}{4L^2}\pi\Big[\sum_{|\{x,y,u\}|=3} \eta_u\lambda_x\lambda_u\big(\nabla_{xy}\psi(\eta)\big)^2
\end{aligned}
$$

$$\times \big(\hat{\rho}_1(\eta,\eta^{xy})\nabla_{uy}\rho(\eta)+\hat{\rho}_2(\eta,\eta^{xy})\big(\rho(\eta^{uy})-\rho(\eta^{xy})\big)\big)\bigg]$$

$$\geq \frac{1}{4L^2}\pi\bigg[\sum_{|\{x,y,u\}|=3}\eta_u\lambda_x\lambda_u\big(\nabla_{xy}\psi(\eta)\big)^2\big(\hat{\rho}_1(\eta,\eta^{xy})+\hat{\rho}_2(\eta,\eta^{xy})\big)\rho(\eta^{uy})\bigg]$$

$$-\frac{(N-1)(c+\delta)}{2L}\mathcal{A}(\rho,\psi)$$

$$\geq -\frac{(N-1)(c+\delta)}{2L}\mathcal{A}(\rho,\psi).$$

We now turn to $\widetilde{\mathcal{B}}_1$. To improve readability, we shall often suppress the variable η in our notation. We have

$$\begin{aligned}\widetilde{\mathcal{B}}_1(\rho,\psi) &= \frac{1}{L^2}\pi\bigg[\sum_{x,y}\lambda_x^2(\nabla_{xy}\psi)^2\hat{\rho}(\eta,\eta^{xy})\bigg]\\ &\quad+\frac{1}{L^2}\pi\bigg[\sum_{|\{x,y,u\}|=3}\lambda_x\lambda_u\nabla_{xy}\psi\nabla_{uy}\psi\hat{\rho}(\eta,\eta^{xy})\bigg]\\ &\quad+\frac{1}{L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x^2\nabla_{xy}\psi\nabla_{xv}\psi\hat{\rho}(\eta,\eta^{xy})\bigg]\\ &=: J_4+J_5+J_6.\end{aligned}$$

We have the immediate bound

$$J_4\geq\frac{2c}{L}\mathcal{A}(\rho,\psi).$$

Another application of (21) yields

$$\begin{aligned}J_6&=\frac{1}{L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x^2\eta_x(1-\eta_y)(1-\eta_v)\nabla_{xy}\psi\big(\psi(\eta^{xv})-\psi(\eta^{xy})\big)\hat{\rho}(\eta,\eta^{xy})\bigg]\\ &\quad+\frac{1}{L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x^2\eta_x(1-\eta_y)(1-\eta_v)(\nabla_{xy}\psi)^2\hat{\rho}(\eta,\eta^{xy})\bigg]\\ &=-\frac{1}{L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x\lambda_y\eta_x(1-\eta_y)(1-\eta_v)\nabla_{xy}\psi\nabla_{xv}\psi\hat{\rho}(\eta,\eta^{xy})\bigg]\\ &\quad+\frac{L-N-1}{L^2}\pi\bigg[\sum_{x,y}\lambda_x^2\eta_x(1-\eta_y)(\nabla_{xy}\psi)^2\hat{\rho}(\eta,\eta^{xy})\bigg].\end{aligned}$$

Averaging the latter expression with the defining formula for J_6, we obtain

$$\begin{aligned}J_6&=\frac{L-N-1}{2L^2}\pi\bigg[\sum_{x,y}\lambda_x^2\eta_x(1-\eta_y)(\nabla_{xy}\psi)^2\hat{\rho}(\eta,\eta^{xy})\bigg]\\ &\quad+\frac{1}{2L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x(\lambda_x-\lambda_y)\eta_x(1-\eta_y)(1-\eta_v)\nabla_{xy}\psi\nabla_{xv}\psi\hat{\rho}(\eta,\eta^{xy})\bigg].\end{aligned}$$

Writing

$$J_7 := \frac{\delta}{4L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x \eta_x (1-\eta_y)(1-\eta_v)(\nabla_{xv}\psi)^2 \hat{\rho}(\eta,\eta^{xy})\Big]$$

for brevity, we obtain using the inequality $-2|ab| \geq -a^2 - b^2$,

$$\begin{aligned} J_6 &\geq \frac{c(L-N-1)}{L}\mathcal{A}(\rho,\psi) \\ &\quad - \frac{\delta}{4L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x \eta_x (1-\eta_y)(1-\eta_v)(\nabla_{xy}\psi)^2 \hat{\rho}(\eta,\eta^{xy})\Big] - J_7 \\ &= \frac{(2c-\delta)(L-N-1)}{2L}\mathcal{A}(\rho,\psi) - J_7. \end{aligned}$$

Using reversibility and averaging, we can show that

$$\begin{aligned} J_7 &\overset{(21)}{=} \frac{\delta}{4L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x \eta_x (1-\eta_y)(1-\eta_v)(\nabla_{xv}\psi)^2 \hat{\rho}(\eta^{xv},\eta^{xy})\Big] \\ &= \frac{\delta}{8L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x \eta_x (1-\eta_y)(1-\eta_v)(\nabla_{xv}\psi)^2 \big(\hat{\rho}(\eta^{xv},\eta^{xy}) + \hat{\rho}(\eta,\eta^{xy})\big)\Big] \\ &\overset{y\leftrightarrow v}{=} \frac{\delta}{8L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x \eta_x (1-\eta_y)(1-\eta_v)(\nabla_{xy}\psi)^2 \big(\hat{\rho}(\eta^{xv},\eta^{xy}) + \hat{\rho}(\eta,\eta^{xv})\big)\Big]. \end{aligned}$$

Therefore, we have

$$\begin{aligned} J_3 + J_6 &\geq \frac{(c-2\delta)(L-N-1)}{2L}\mathcal{A}(\rho,\psi) \\ &\quad + \frac{\delta}{8L^2}\pi\Big[\sum_{|\{x,y,v\}|=3} \lambda_x \eta_x (1-\eta_y)(1-\eta_v)(\nabla_{xy}\psi)^2 \\ &\quad \times \big(\rho(\eta^{xv})\big(\hat{\rho}_1(\eta,\eta^{xy}) + \hat{\rho}_2(\eta,\eta^{xy})\big) - \hat{\rho}(\eta^{xv},\eta^{xy}) - \hat{\rho}(\eta,\eta^{xv})\big)\Big] \\ &\geq \frac{(2c-5\delta)(L-N-1)}{4L}\mathcal{A}(\rho,\psi), \end{aligned}$$

where the last inequality has been obtained through an application of Lemma A.2.

Similarly, we have

$$\begin{aligned} J_5 &\overset{(21)}{=} \frac{1}{L^2}\pi\Big[\sum_{|\{x,y,u\}|=3} \lambda_x \lambda_u \eta_x (1-\eta_y)\eta_u \nabla_{xy}\psi\big(\psi(\eta) - \psi(\eta^{uy})\big)\hat{\rho}(\eta,\eta^{xy})\Big] \\ &\quad + \frac{1}{L^2}\pi\Big[\sum_{|\{x,y,u\}|=3} \lambda_x \lambda_u \eta_x (1-\eta_y)\eta_u (\nabla_{xy}\psi)^2 \hat{\rho}(\eta,\eta^{xy})\Big] \end{aligned}$$

$$\geq -J_5 + \frac{2c(N-1)}{L}\mathcal{A}(\rho, \psi),$$

and, therefore,

$$J_5 \geq \frac{c(N-1)}{L}\mathcal{A}(\rho, \psi).$$

When we sum up, we get

$$\begin{aligned}\widetilde{\mathcal{B}}(\rho, \psi) &\geq \left(\frac{(L-N-1)(2c-5\delta)}{4L} + \frac{(N-1)(c-\delta)}{2L} + \frac{4c-\delta}{2L}\right)\mathcal{A}(\rho, \psi) \\ &\geq \left(\frac{c}{2} - \frac{5(L-1)-3N}{4L}\delta\right)\mathcal{A}(\rho, \psi).\end{aligned} \tag{23}$$

At this point, what we get is a size-independent lower bound on the Ricci curvature of $\frac{c}{2} - \frac{5\delta}{4}$. However, we can improve this bound using the following duality argument for the Bernoulli–Laplace process: if we consider the system with N particles, and then remove all particles while simultaneously adding particles at every empty site, we get a Bernoulli–Laplace model with $L - N$ particles. Since this is just a change in labeling (empty and full sites play symmetric roles), the properties of the dynamic are invariant by this transform. Therefore, without any loss of generality, we can assume that $N \geq L/2$. Under this extra assumption, the bound (23) leads to a lower bound on the Ricci curvature of $\frac{c}{2} - \frac{7\delta}{8}$, which is what we were seeking to prove. $\square$

4.4. *The random transposition model.* We now consider the random transposition model, which is a random walk on the group of permutations $\mathcal{S}_n$ with $n \geq 2$. If we denote by $\mathcal{T}_n$ the set of all transpositions in $\mathcal{S}_n$, we can write the generator of the dynamics as

$$\mathcal{L}f(\sigma) := \frac{2}{n(n-1)} \sum_{\tau \in \mathcal{T}_n} \nabla_\tau f(\sigma),$$

where $\nabla_\tau f(\sigma) = f(\tau \circ \sigma) - f(\sigma)$ and π is the uniform measure on $\mathcal{S}_n$, that is, $\pi(\sigma) = (n!)^{-1}$ for all $\sigma \in \mathcal{S}_n$. We write $\tau = (i, j)$ for the transposition that swaps i and j. We also write (i, j, k) for the mapping that cyclically permutes i, j and k. To simplify notation, we will use the shorthand notation

$$\sigma_{ij} := (i, j) \circ \sigma, \qquad \sigma_{ijk} := (i, j, k) \circ \sigma, \qquad \nabla_{ij} := \nabla_{(i,j)}.$$

We obtain the following result.

THEOREM 4.9. *For $n \geq 2$, the Ricci curvature of the random transposition model is bounded from below by $\frac{4}{n(n-1)}$.*

This bound was already obtained using a different method in [11]. It should be noted that the modified logarithmic Sobolev constant implied by this Ricci curvature bound is significantly worse (by a factor $1/n$) than its known optimal behaviour. We do not know whether the Ricci curvature's true behaviour should match the MLSI-constant.

PROOF OF THEOREM 4.9. It follows from the definition that the action functional can be written as

$$\mathcal{A}(\rho,\psi)=\frac{1}{2n(n-1)}\pi\biggl[\sum_{i\neq j}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\hat{\rho}(\sigma,\sigma_{ij})\biggr].$$

Note that a factor $\frac{1}{2}$ appears, since every transposition (i,j) is counted twice. We define R as

$$R\bigl(\sigma,(i,j),(k,\ell)\bigr)=\begin{cases}\dfrac{4}{n^2(n-1)^2}, & \text{if } \bigl|\{i,j,k,\ell\}\bigr|=4,\\ 0, & \text{otherwise.}\end{cases}$$

Using reversibility and the fact that $(i,j)^{-1}=(i,j)$, it follows that

$$\begin{aligned}\widetilde{\mathcal{B}}_2(\rho,\psi)&=\frac{2}{n^2(n-1)^2}\pi\biggl[\sum_{|\{i,j,k\}|=3}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\hat{\rho}_1(\sigma,\sigma_{ij})\nabla_{ik}\rho(\sigma)\biggr]\\&\quad+\frac{2}{n^2(n-1)^2}\pi\biggl[\sum_{i\neq j}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\hat{\rho}_1(\sigma,\sigma_{ij})\nabla_{ij}\rho(\sigma)\biggr]\\&=\frac{2}{n^2(n-1)^2}\pi\biggl[\sum_{|\{i,j,k\}|=3}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\hat{\rho}_2(\sigma,\sigma_{ij})\bigl(\rho(\sigma_{ijk})-\rho(\sigma_{ij})\bigr)\biggr]\\&\quad-\frac{1}{n^2(n-1)^2}\pi\biggl[\sum_{i\neq j}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\hat{\rho}_2(\sigma,\sigma_{ij})\nabla_{ij}\rho(\sigma)\biggr].\end{aligned}$$

Averaging the latter two expressions, we obtain

$$\begin{aligned}\widetilde{\mathcal{B}}_2(\rho,\psi)&=\frac{1}{n^2(n-1)^2}\pi\biggl[\sum_{|\{i,j,k\}|=3}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\\&\quad\times\bigl(\rho(\sigma_{ik})\hat{\rho}_1(\sigma,\sigma_{ij})+\rho(\sigma_{ijk})\hat{\rho}_2(\sigma,\sigma_{ij})\bigr)\biggr]\\&\quad-\frac{1}{n^2(n-1)^2}\pi\biggl[\sum_{|\{i,j,k\}|=3}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\\&\quad\times\bigl(\rho(\sigma)\hat{\rho}_1(\sigma,\sigma_{ij})+\rho(\sigma_{ij})\hat{\rho}_2(\sigma,\sigma_{ij})\bigr)\biggr]\\&\quad+\frac{1}{2n^2(n-1)^2}\pi\biggl[\sum_{i\neq j}\bigl(\nabla_{ij}\psi(\sigma)\bigr)^2\bigl(\hat{\rho}_1(\sigma,\sigma_{ij})-\hat{\rho}_2(\sigma,\sigma_{ij})\bigr)\nabla_{ij}\rho(\sigma)\biggr].\end{aligned}$$

Using (i) and (ii) of Lemma A.1 and (24), we infer that

$$\widetilde{\mathcal{B}}_2(\rho,\psi) \geq \frac{1}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\big(\nabla_{ij}\psi(\sigma)\big)^2\hat{\rho}(\sigma_{ik},\sigma_{ijk})\Big] - \frac{2(n-2)}{n(n-1)}\mathcal{A}(\rho,\psi).$$

The term $\widetilde{\mathcal{B}}_1(\rho,\psi)$ can be written as

$$\begin{aligned}
\widetilde{\mathcal{B}}_1(\rho,\psi) &= \frac{2}{n^2(n-1)^2}\pi\Big[\sum_{i\neq j}\big(\nabla_{ij}\psi(\sigma)\big)^2\hat{\rho}(\sigma,\sigma_{ij})\Big] \\
&\quad + \frac{4}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\nabla_{ij}\psi(\sigma)\nabla_{ik}\psi(\sigma)\hat{\rho}(\sigma,\sigma_{ij})\Big] \\
&= \frac{4}{n(n-1)}\mathcal{A}(\rho,\psi) + \frac{4}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\big(\nabla_{ij}\psi(\sigma)\big)^2\hat{\rho}(\sigma,\sigma_{ij})\Big] \\
&\quad + \frac{4}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\nabla_{ij}\psi(\sigma)\big(\psi(\sigma_{ik})-\psi(\sigma_{ij})\big)\hat{\rho}(\sigma,\sigma_{ij})\Big].
\end{aligned}$$

Using reversibility and averaging, the latter term can be reformulated as

$$\begin{aligned}
&\frac{4}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\nabla_{ij}\psi(\sigma)\big(\psi(\sigma_{ik})-\psi(\sigma_{ij})\big)\hat{\rho}(\sigma,\sigma_{ij})\Big] \\
&\quad = \frac{4}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\nabla_{ij}\psi(\sigma)\big(\psi(\sigma)-\psi(\sigma_{ijk})\big)\hat{\rho}(\sigma,\sigma_{ij})\Big] \\
&\quad = \frac{2}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\nabla_{ij}\psi(\sigma)\big(\psi(\sigma_{ik})-\psi(\sigma_{ij})+\psi(\sigma)-\psi(\sigma_{ijk})\big) \\
&\qquad \times\hat{\rho}(\sigma,\sigma_{ij})\Big] \\
&\quad = -\frac{4(n-2)}{n(n-1)}\mathcal{A}(\rho,\psi) \\
&\qquad + \frac{2}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\nabla_{ij}\psi(\sigma)\big(\psi(\sigma_{ik})-\psi(\sigma_{ijk})\big)\hat{\rho}(\sigma,\sigma_{ij})\Big],
\end{aligned}$$

which yields

$$\begin{aligned}
\widetilde{\mathcal{B}}_1(\rho,\psi) &= \Big(\frac{4}{n(n-1)} + \frac{4(n-2)}{n(n-1)}\Big)\mathcal{A}(\rho,\psi) \\
&\quad + \frac{2}{n^2(n-1)^2}\pi\Big[\sum_{|\{i,j,k\}|=3}\nabla_{ij}\psi(\sigma)\big(\psi(\sigma_{ik})-\psi(\sigma_{ijk})\big)\hat{\rho}(\sigma,\sigma_{ij})\Big].
\end{aligned}$$

Young's inequality and reversibility imply that

$$\begin{aligned}
\widetilde{\mathcal{B}}_1(\rho,\psi) &\geq \left(\frac{4}{n} - \frac{2(n-2)}{n(n-1)}\right)\mathcal{A}(\rho,\psi) \\
&\quad - \frac{1}{n^2(n-1)^2}\pi\Bigg[\sum_{|\{i,j,k\}|=3} \big(\psi(\sigma_{ik}) - \psi(\sigma_{ijk})\big)^2 \hat{\rho}(\sigma,\sigma_{ij})\Bigg] \\
&= \frac{2}{n-1}\mathcal{A}(\rho,\psi) - \frac{1}{n^2(n-1)^2}\pi\Bigg[\sum_{|\{i,j,k\}|=3} \big(\nabla_{jk}\psi(\sigma)\big)^2 \hat{\rho}(\sigma_{ik},\sigma_{ikj})\Bigg] \\
&\underset{k\leftrightarrow i}{=} \frac{2}{n-1}\mathcal{A}(\rho,\psi) - \frac{1}{n^2(n-1)^2}\pi\Bigg[\sum_{|\{i,j,k\}|=3} \big(\nabla_{ij}\psi(\sigma)\big)^2 \hat{\rho}(\sigma_{ik},\sigma_{ijk})\Bigg].
\end{aligned}$$

When we sum $\widetilde{\mathcal{B}}_1$ and $\widetilde{\mathcal{B}}_2$, we get

$$\mathcal{B}(\rho,\psi) \geq \frac{4}{n(n-1)}\mathcal{A}(\rho,\psi),$$

which is the desired result. □

APPENDIX: PROPERTIES OF THE LOGARITHMIC MEAN

In this paper, we make use of some basic properties of the logarithmic mean, given for $a, b \geq 0$ (resp., $a, b > 0$) by

$$\theta(a,b) := \int_0^1 a^{1-p} b^p \,\mathrm{d}p = \frac{a-b}{\log a - \log b}.$$

The following properties are taken from [9]; see Lemma 5.4. We write $\theta_i(s,t) = \partial_i \theta(s,t)$ for $i = 1, 2$.

LEMMA A.1. *The following assertions hold*:

(i) $u\theta_1(s,t) + v\theta_2(s,t) \geq \theta(u,v)$ *for all* $s, t, u, v > 0$;
(ii) $s\theta_1(s,t) + t\theta_2(s,t) = \theta(s,t)$ *for all* $s, t > 0$;
(iii) θ *is symmetric*, *concave and increasing in both variables*;
(iv) θ *is positively* 1-*homogeneous*, *that is*, *for all* $\lambda, s, t \geq 0$, *we have* $\theta(\lambda s, \lambda t) = \lambda\theta(s,t)$.

It follows directly from this lemma that for every $\lambda_1, \lambda_2 > 0$ and $s, t \geq 0$,

$$\big(\lambda_1\theta_1(s,t) - \lambda_2\theta_2(s,t)\big)(s-t) \leq \big(\max\{\lambda_1,\lambda_2\} - \min\{\lambda_1,\lambda_2\}\big)\theta(s,t). \tag{24}$$

The following inequality plays a crucial role in the proof of the curvature bound for the Bernoulli–Laplace model and for zero-range processes.

LEMMA A.2. *For any $r \geq 0$ and $s, t \geq 0$, we have*

$$r\big(\theta_1(s,t) + \theta_2(s,t)\big) - \big(\theta(r,s) + \theta(r,t)\big) \geq -\theta(s,t).$$

PROOF. If $r = 0$, the inequality is trivially true, so without loss of generality we can assume that $r > 0$. Let $u = s/r$ and $v = t/r$. Using the fact that θ is 1-homogeneous, and θ_1 and θ_2 are 0-homogeneous, the inequality we wish to prove is equivalent to

$$\theta_1(u,v) + \theta_2(u,v) - \theta(1,u) - \theta(1,v) \geq -\theta(u,v). \tag{25}$$

Since θ is concave, we have

$$\begin{aligned}\theta(1,u) + \theta(1,v) &= 2 \times \frac{1}{2}\big(\theta(u,1) + \theta(1,v)\big) \\ &\leq 2\theta\left(\frac{u+1}{2}, \frac{v+1}{2}\right) \\ &= \theta(u+1, v+1).\end{aligned}$$

Using the "curve below tangent" formulation of concavity applied to the function $x \mapsto \theta(u+x, v+x)$, we have

$$\theta(u+1, v+1) \leq \theta(u,v) + \theta_1(u,v) + \theta_2(u,v),$$

and (25) immediately follows, which completes the proof. $\square$

Acknowledgements. Part of this work has been done while M. Fathi visited J. Maas at the University of Bonn in July 2014. We would like to thank the referees for their careful reading of the manuscript.

REFERENCES

[1] BAKRY, D. and ÉMERY, M. (1985). Diffusions hypercontractives. In *Séminaire de Probabilités, XIX*, 1983/84. *Lecture Notes in Math*. **1123** 177–206. Springer, Berlin. MR0889476

[2] BOBKOV, S. G. and TETALI, P. (2006). Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab*. **19** 289–336. MR2283379

[3] BONCIOCAT, A.-I. and STURM, K.-T. (2009). Mass transportation and rough curvature bounds for discrete spaces. *J. Funct. Anal*. **256** 2944–2966. MR2502429

[4] BOUDOU, A.-S., CAPUTO, P., DAI PRA, P. and POSTA, G. (2006). Spectral gap estimates for interacting particle systems via a Bochner-type identity. *J. Funct. Anal*. **232** 222–258. MR2200172

[5] CAPUTO, P., DAI PRA, P. and POSTA, G. (2009). Convex entropy decay via the Bochner–Bakry–Emery approach. *Ann. Inst. Henri Poincaré Probab. Stat*. **45** 734–753. MR2548501

[6] CAPUTO, P. and POSTA, G. (2007). Entropy dissipation estimates in a zero-range dynamics. *Probab. Theory Related Fields* **139** 65–87. MR2322692

[7] CHOW, S.-N., HUANG, W., LI, Y. and ZHOU, H. (2012). Fokker–Planck equations for a free energy functional or Markov process on a graph. *Arch. Ration. Mech. Anal*. **203** 969–1008. MR2928139

[8] DAI PRA, P. and POSTA, G. (2013). Entropy decay for interacting systems via the Bochner–Bakry–Émery approach. *Electron. J. Probab.* **18** 52. MR3065862

[9] ERBAR, M. and MAAS, J. (2012). Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.* **206** 997–1038. MR2989449

[10] ERBAR, M. and MAAS, J. (2014). Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Syst.* **34** 1355–1374. MR3117845

[11] ERBAR, M., MAAS, J. and TETALI, P. (2015). Discrete Ricci curvature bounds for Bernoulli–Laplace and random transposition models. *Annales Fac. Sci. Toulouse* (6) **24** 781–800. MR3434256

[12] GAO, F. and QUASTEL, J. (2003). Exponential decay of entropy in the random transposition and Bernoulli–Laplace models. *Ann. Appl. Probab.* **13** 1591–1600. MR2023890

[13] GIGLI, N. and MAAS, J. (2013). Gromov–Hausdorff convergence of discrete transportation metrics. *SIAM J. Math. Anal.* **45** 879–899. MR3045651

[14] GOEL, S. (2004). Modified logarithmic Sobolev inequalities for some models of random walk. *Stochastic Process. Appl.* **114** 51–79. MR2094147

[15] JORDAN, R., KINDERLEHRER, D. and OTTO, F. (1998). The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29** 1–17. MR1617171

[16] LIN, Y. and YAU, S.-T. (2010). Ricci curvature and eigenvalue estimate on locally finite graphs. *Math. Res. Lett.* **17** 343–356. MR2644381

[17] LOTT, J. and VILLANI, C. (2009). Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math.* (2) **169** 903–991. MR2480619

[18] MAAS, J. (2011). Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* **261** 2250–2292. MR2824578

[19] MIELKE, A. (2011). A gradient structure for reaction–diffusion systems and for energy-drift-diffusion systems. *Nonlinearity* **24** 1329–1346. MR2776123

[20] MIELKE, A. (2013). Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations* **48** 1–31. MR3090532

[21] MORRIS, B. (2006). Spectral gap for the zero range process with constant rate. *Ann. Probab.* **34** 1645–1664. MR2271475

[22] OLLIVIER, Y. (2009). Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.* **256** 810–864. MR2484937

[23] OLLIVIER, Y. (2013). A visual introduction to Riemannian curvatures and some discrete generalizations. In *Analysis and Geometry of Metric Measure Spaces. CRM Proc. Lecture Notes* **56** 197–220. Amer. Math. Soc., Providence, RI. MR3060504

[24] STURM, K.-T. (2006). On the geometry of metric measure spaces. I. *Acta Math.* **196** 65–131. MR2237206

LABORATOIRE DE PROBABILITÉS
ET MODÈLES ALÁTOIRES
UNIVERSITÉ PARIS 6
4, PLACE JUSSIEU
75005 PARIS
FRANCE
E-MAIL: max.fathi@etu.upmc.fr

INSTITUTE OF SCIENCE
AND TECHNOLOGY AUSTRIA (IST AUSTRIA)
AM CAMPUS 1
3400 KLOSTERNEUBURG
AUSTRIA
E-MAIL: jan.maas@ist.ac.at