ENTROPIC CURVATURE AND CONVERGENCE TO EQUILIBRIUM FOR MEAN-FIELD DYNAMICS ON DISCRETE SPACES

MATTHIAS ERBAR, MAX FATHI, AND ANDRÉ SCHLICHTING

ABSTRACT. We consider non-linear evolution equations arising from mean-field limits of particle systems on discrete spaces. We investigate a notion of curvature bounds for these dynamics based on the convexity of the free energy along interpolations in a discrete transportation distance related to the gradient flow structure of the dynamics. This notion extends the one for linear Markov chain dynamics studied in [22]. We show that positive curvature bounds entail several functional inequalities controlling the convergence to equilibrium of the dynamics. We establish explicit curvature bounds for several examples of mean-field limits of various classical models from statistical mechanics.

1. INTRODUCTION

This work is about longtime behavior for mean-field systems on discrete spaces. Mean-field equations describe the large-scale limit of interacting particle systems where the total force exerted on any given particle is the average of the forces exerted by all other particles on the tagged particle. They are used to describe collective behavior in many areas of sciences. Examples include the modeling of granular flows in physics [2] and collective behavior and self-organization for groups of animals [17, 12]. We refer to [37] for an introduction to the mathematical theory.

One of the important questions in the mathematical analysis of these equations is their longtime behavior. In [10], Carrillo, McCann and Villani obtained quantitative bounds on the rate of convergence to equilibrium for McKean-Vlasov equations in a continuous setting of the form

$$\partial_t \rho = \nabla \cdot [\rho \nabla (S'(\rho) + V + W * \rho)]$$

under strong convexity assumptions on the potentials S, V and W. The core idea underlying their method was the fact that the PDE has a gradient flow structure, i.e. it can be recast as a gradient descent equation $\partial_t \rho = -\nabla F(\rho)$ of the free energy functional $F(\rho) = \int S(\rho) + \int V \, d\rho + \int W * \rho \, d\rho$ in the space of probability measure with respect to the Kantorovitch-Wasserstein distance W_2, which has a formal Riemannian description via Otto calculus [33, 34]. The use of such structures in the study of longtime behavior comes from the fact that as soon as the driving functional satisfies some uniform convexity property (with respect to the particular metric structure), it must decay exponentially fast towards its minimal value along solutions of the evolution equation. Moreover, we can use convexity to derive strong functional inequalities relating the distance, the entropy functional and the entropy dissipation functional [34].

1.1. **Setup and main results.** Our main motivation here is to adapt the approach of [10] to mean-field equations in a discrete setting. We consider discrete mean-field dynamics of the form

$$\dot{\mu}(t) = \mu(t)Q(\mu(t)) \, , \tag{1.1}$$

where μ is a flow of probability measures on a finite set $\mathcal{X}$ and $(Q(\mu)_{xy})_{x,y\in\mathcal{X}}$ is a parametrized collection of Markov kernels. These dynamics naturally arise as scaling limits of interacting particles systems on graphs where the interaction only depends on the normalized empirical measure of the system (which indeed corresponds to mean-field interactions). They generalize linear Markov chains on discrete spaces, which correspond to the case where Q is a constant Markov kernel, independent of μ.

While the Wasserstein gradient flow approach works well on continuous spaces, it fails in the discrete setting, since the Wasserstein L^2-transport distance does not admit any non-trivial absolutely continuous curve. In our previous work [20], we derived a gradient flow structure for (1.1) by replacing the role of the Wasserstein distance with a distance $\mathcal{W}$ constructed via a suitable modification of the Benamou-Brenier formula for optimal transport, extending similar earlier results for linear reversible Markov chains obtained in [30, 32, 14]. Under the condition that the rates Q are Gibbs with a potential $K : \mathcal{P}(\mathcal{X}) \times \mathcal{X} \to \mathbf{R}$ (see Assumption 2.1), i.e. $Q(\mu)$ is reversible with respect to a local Gibbs measure of the form $\pi_x(\mu) = Z(\mu)^{-1}\exp(-H_x(\mu))$, with H_x given in terms of the potential K, we showed that this dynamic is the gradient flow of the free energy functional

$$\mathcal{F}(\mu) = \sum_{x\in\mathcal{X}} \mu_x \log\mu_x + U(\mu), \qquad \text{with} \qquad U(\mu) = \sum_{z\in\mathcal{X}} \mu_z\, K_z(\mu), \tag{1.2}$$

with respect to the distance $\mathcal{W}$ on the simplex of probability measures on $\mathcal{X}$, see Proposition 2.2. This built up on previous works [4, 5] that showed that $\mathcal{F}$ is indeed a Lyapunov functional for the flow. An archetypical example, which we shall discuss in some details later on, is the classical Curie-Weiss model, which corresponds to a mean-field dynamic on a two-point space. Already this easy model exhibits interesting behavior, such as a phase transition at an explicit critical value of a temperature parameter.

In the present work, we exploit this gradient flow structure to analyze the long-term behavior of (1.1) inspired by the approach in [10] by investigating convexity properties of the free energy along discrete optimal transport paths for a non-linear Markov triple $(\mathcal{X}, Q, \pi)$ as above. Following the works of Lott, Villani, and Sturm [29, 36] for metric measure spaces and [22, 32] for linear Markov chains, we make the following

Definition 1.1 (Entropic Ricci curvature lower bound)**.** We say that $(\mathcal{X}, Q, \pi)$ has *Ricci curvature bounded below by* $\kappa \in \mathbf{R}$ (for short $\mathrm{Ric}(\mathcal{X}, Q) \geq \kappa$) if for any $\mathcal{W}$-geodesic $(\mu_t)_{t\in[0,1]}$:

$$\mathcal{F}(\mu_t) \leq (1 - t)\mathcal{F}(\mu_0) + t\mathcal{F}(\mu_1) - \frac{\kappa}{2}t(1 - t)\mathcal{W}(\mu_0, \mu_1)^2 \, .$$

We show, see Theorem 3.7, that Ricci curvature lower bounds can be characterized in terms of a discrete Bochner-type inequality by deriving the Hessian of $\mathcal{F}$ in the Riemannian structure $\mathcal{W}$, as well as in terms of the Evolution Variational Inequality

EVI_κ for the solutions to (1.1):

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\mathcal{W}(\mu_t,\nu)^2 + \frac{\kappa}{2}\mathcal{W}(\mu_t,\nu)^2 \leq \mathcal{F}(\nu) - \mathcal{F}(\mu_t) \ .$$

Further, we show that a positive lower bound on the Ricci curvature entails a number of functional inequalities that control the convergence to equilibrium of the mean-field systems. These involve a discrete Fisher information functional $\mathcal{I} : \mathcal{P}(\mathcal{X}) \to [0,\infty]$ given by

$$\mathcal{I}(\mu) = \frac{1}{2}\sum_{x,y}\Theta(\mu_x Q_{xy}(\mu), \mu_y Q_{yx}(\mu)) \ , \qquad \Theta(a,b) = (a-b)(\log a - \log b) \ ,$$

which arises from the dissipation of $\mathcal{F}$ along solutions to (1.1) as $\frac{d}{dt}\mathcal{F}(\mu_t) = -\mathcal{I}(\mu_t)$. One of our main results is the following theorem which can be seen as a discrete analog of [10, Thm. 2.1].

Theorem 1.2. *Assume that* $\mathrm{Ric}(\mathcal{X},Q,\pi) \geq \lambda$ *for some* $\lambda > 0$. *Then the following hold:*
  (i) *there exists a unique stationary point* π^* *for the evolution* (1.1)*, it is the unique minimizer of* $\mathcal{F}$. *Let* $\mathcal{F}^*(\cdot) := \mathcal{F}(\cdot) - \mathcal{F}(\pi^*)$;
  (ii) *the* modified logarithmic Sobolev inequality *with constant* $\lambda > 0$ *holds, i.e. for all* $\mu \in \mathcal{P}(\mathcal{X})$,

$$\mathcal{F}^*(\mu) \leq \frac{1}{2\lambda}\mathcal{I}(\mu) \ ; \tag{MLSI(λ)}$$

  (iii) *for any solution* $(\mu_t)_{t\geq 0}$ *to* (1.1) *we have exponential decay of the free energy:*

$$\mathcal{F}^*(\mu_t) \leq e^{-2\lambda t}\mathcal{F}^*(\mu_0) \ ;$$

  (iv) *the* entropy-transport inequality *with constant* $\lambda > 0$ *holds, i.e. for all* $\mu \in \mathcal{P}(\mathcal{X})$,

$$\mathcal{W}(\mu,\pi^*) \leq \sqrt{\frac{2}{\lambda}\mathcal{F}^*(\mu)} \ . \tag{ET(λ)}$$

1.2. **Examples.** We establish explicit curvature bounds for several examples of (relatively simple) mean-field dynamics, such as the Curie-Weiss model, zero-range mean-field dynamics, and misanthrope processes. We compute a formula for the second derivative of entropy along geodesics and generalize techniques developed in [23, 21] to the present non-linear situation, to bound this second derivative. The non-linearity of the dynamic gives rise to several extra terms when computing the Hessian of the free energy functional, which complicates the analysis.

In the case of the Curie–Weiss model, we show that a positive lower curvature bound holds down to the critical temperature, see Section 5.1.

Another particular family of dynamic we shall be interested is when the flux of particles from some site x to a site y is a function of the particle density at site y, that is $f(c_y)$. In the situation where f is constant, this would correspond to the scaling limit of independent particles on the complete graph. As in [23, 21], our approach is in some sense perturbative in nature, and we shall consider rates of the form $f(r) = T + g(r)$, and show that if g is not too large in some sense, relative to T, then we can derive a rate of convergence to equilibrium. This is inspired by recent work of Villemonais [40], who proved that the N particle system has a positive Ollivier-Ricci (or coarse Ricci) curvature (another

notion of curvature, corresponding to a contraction rate for the Markovian dynamic) independently of the system size, and hence converges to equilibrium in L^2 distance, via a uniform estimate on the Poincaré constant of the dynamic. Our approach has the advantage of yielding rates of convergence in relative entropy via Theorem 1.2, which is a strictly stronger notion of convergence.

1.3. **Connection to the literature.** The approach of [10] was later extended to other potentials [11, 3]. Other approaches developed later include using uniform convergence estimates for a stochastic particle approximation [13] and coupling arguments [18, 19]. Without convexity, deriving rates of convergence can be quite delicate, since there may be multiple equilibria [39], unlike what happens for linear diffusions.

Our approach developed here builds on earlier work [30, 32, 14] constructing gradient flow structures for linear Markov chain dynamics and [22] studying Ricci curvature and its impact on functional inequalities in this context. It is also related to, but different from the one developed in [27], which uses convexity of the entropy along a different type of paths, the so-called entropic interpolations, rather than geodesic paths, to establish functional inequalities involving relative entropy. In the continuous setting, entropic interpolations are regularizations of geodesics in Wasserstein space, but in the discrete case it seems that the entropic interpolations of [27] are related to a gradient flow structure different from the one of [20] we use here.

Another approach to convergence to equilibrium for discrete quadratic systems was developped in [8], using decompositions of entropy on product spaces similar to, but more general than, the classical Shearer inequality to establish a modified logarithmic Sobolev inequality. The methods in that work are more combinatorial in nature, while the point of view developed here is more analytic.

Organization. The plan of the paper is as follows. Section 2 introduces the mathematical framework we shall work in. Section 3 introduces the notion of curvature bounds in our setting, and contains the computation of the Hessian for general dynamics. Section 4 investigates the consequences of Ricci curvature bounds in terms of functional inequalities and convergence to equilibrium for the non-linear dynamics. Finally, Section 5 investigates curvature bounds for several examples of mean-field dynamics inspired by classical models of statistical physics.

2. Setup

2.1. **Gradient-flow formulation.** The main definitions and results from [20] on which this work builds are collected in this section. We consider a finite space $\mathcal{X}$, and denote by $\mathcal{P}(\mathcal{X})$ (resp. $\mathcal{P}^*(\mathcal{X})$) the space of probability measures on $\mathcal{X}$ (resp. strictly positive probability measures). The gradient flow structure of (1.1) is based on the existence of a suitable potential, which is ensured by the following constraint, which we shall assume to hold throughout the article. We recall that a rate matrix Q of a Markov chain in the continuous-time setting satisfies

$$\forall x \neq y : Q_{xy} \geq 0 \qquad \text{and} \qquad Q_{xx} = -\sum_{y \neq x} Q_{xy} \, .$$

Assumption 2.1. *Let $K : \mathcal{P}(\mathcal{X}) \times \mathcal{X} \to \mathbf{R}$ be such that for each $x \in \mathcal{X}$, $K_x : \mathcal{P}(\mathcal{X}) \to \mathbf{R}$ is a twice continuously differentiable. Let $\{Q(\mu) \in \mathbf{R}^{\mathcal{X} \times \mathcal{X}}\}_{\mu \in \mathcal{P}(\mathcal{X})}$ be a family of rate matrices that is Gibbs with respect to the potential function K, i.e. for each $\mu \in \mathcal{P}(\mathcal{X})$, $Q(\mu)$ is the rate matrix of an irreducible, reversible ergodic Markov chain with respect to the probability measure*

$$\pi_x(\mu) = \frac{1}{Z(\mu)} \exp\bigl(-H_x(\mu)\bigr) \ , \tag{2.1}$$

with

$$H_x(\mu) = \frac{\partial}{\partial \mu_x} U(\mu) \qquad and \qquad U(\mu) = \sum_{x \in \mathcal{X}} \mu_x K_x(\mu) \ . \tag{2.2}$$

In particular $Q(\mu)$ satisfies the detailed balance condition with respect to $\pi(\mu)$, that is for all $x, y \in \mathcal{X}$

$$\pi_x(\mu) Q_{xy}(\mu) = \pi_y(\mu) Q_{yx}(\mu) \tag{2.3}$$

holds. Moreover, we assume that for each $x, y \in \mathcal{X}$ the map $\mu \mapsto Q_{xy}(\mu)$ is Lipschitz continuous over $\mathcal{P}(\mathcal{X})$.

We refer to the triple $(\mathcal{X}, Q, \pi)$ as above for short as a *non-linear Markov triple*.
The specific form of (2.1) with (2.2) emerges from the detailed balance condition of an underlying N-particle system, from which the dynamics we are interested arise in the limit $N \to \infty$ (see [20]). Associated to a non-linear Markov triple $(\mathcal{X}, Q, \pi)$ is the non-linear *master equation*

$$\dot{\mu}(t) = \mu(t) Q(\mu(t)) \ , \tag{2.4}$$

which is the deterministic evolution equation describing the mean-field limit of the underlying particle system. Based on the above assumption a gradient flow formulation of (2.4) is established in [20, Proposition 2.13] as we shall briefly recall.
Consider the *Onsager operator* $\mathcal{K}[\mu] : T_\mu^* \mathcal{P}(\mathcal{X}) \to T_\mu \mathcal{P}(\mathcal{X})$ given by

$$\mathcal{K}[\mu]\psi(x) := -\frac{1}{2} \sum_y \Lambda\bigl(\mu_x Q_{xy}(\mu), \mu_y Q_{yx}(\mu)\bigr)\bigl(\psi(y) - \psi(x)\bigr) \ ,$$

where $\Lambda(a, b) = \int_0^1 a^{1-s} b^s \, \mathrm{d}s = \frac{a-b}{\log a - \log b}$ is the logarithmic mean. Here the space $T_\mu \mathcal{P}(\mathcal{X})$ is the space of signed measures with total mass zero, which should be interpreted as the tangent space at μ, and $T_\mu^* \mathcal{P}(\mathcal{X})$ is the space of functions on $\mathcal{X}$, which shall be interpreted as the cotangent space for the metric we shall define as the domain of the Onsager operator $\mathcal{K}$. Then the master equation can be written in gradient flow form using the functional $\mathcal{F}$ from (1.2):

$$\frac{d\mu_t}{dt} = -\mathcal{K}[\mu_t] D\mathcal{F}(\mu_t) \tag{2.5}$$

where $D\mathcal{F}$ should be understood as the differential of F, given by $D\mathcal{F}(\mu)(x) = \partial_{\mu_x} F(\mu)$. In other words, (2.4) is the gradient flow of $\mathcal{F}$ with respect to the Riemannian structure on $\mathcal{P}(\mathcal{X})$ induced by the metric tensor $\mathcal{K}[\mu]^{-1}$. Since this Riemannian metric degenerates at the boundary of $\mathcal{P}(\mathcal{X})$ we note the following characterization in metric terms. We

consider the distance function on $\mathcal{P}(\mathcal{X})$ that is formally induced by the Riemannian metric $\mathcal{K}[\mu]^{-1}$, i.e. for $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$ we set

$$\mathcal{W}(\mu_0, \mu_1) := \inf_{(\mu, \psi) \in CE} \left(\int_0^1 \mathcal{A}(\mu_t, \psi_t) dt \right)^{1/2}$$

where CE is the set of curves $(\mu_t, \psi_t)_{t \in [0,1]}$ with $t \to \mu_t$ continuous, ψ measurable and integrable in time, and satisfying the continuity equation

$$\dot{\mu}_t = \mathcal{K}[\mu_t] \psi_t \tag{2.6}$$

in distribution sense, and the action functional $\mathcal{A}$ is given by

$$\mathcal{A}(\mu, \psi) = \langle \psi, \mathcal{K}[\mu] \psi \rangle = \frac{1}{2} \sum_{x,y} (\psi(y) - \psi(x))^2 \, \Lambda(\mu_x Q_{xy}(\mu), \mu_y Q_{yx}(\mu)). \tag{2.7}$$

Proposition 2.2 (Gradient flow structure of the mean-field system)**.** *Let $(\mathcal{X}, Q, \pi)$ be a non-linear Markov triple satisfying Assumption 2.1. Then any solution to (2.4) is a gradient flow of $\mathcal{F}$ with respect to the distance $\mathcal{W}$.*

The distance $\mathcal{W}$ and the above gradient flow structure are extensions of the discrete transport distance constructed in [30] and the gradient flow structure of linear Markov chains to the non-linear case. See [20, Section 2.3] for more background on the construction of the distance $\mathcal{W}$.

An immediate consequence of the gradient flow formulation (2.5) is the free energy dissipation relation established in [20, Remark 2.14]:

$$\mathcal{F}(\mu_t) + \int_0^t \mathcal{I}(\mu_s) \, \mathrm{d}s = \mathcal{F}(\mu_0) \qquad \text{for any } t > 0 \, . \tag{2.8}$$

Here, the discrete Fisher information or dissipation $\mathcal{I} : \mathcal{P}(\mathcal{X}) \to [0, \infty]$ is defined by

$$\mathcal{I}(\mu) = \begin{cases} \frac{1}{2} \sum_{(x,y) \in E_\mu} \Theta(\mu_x Q_{xy}(\mu), \mu_y Q_{yx}(\mu)) \, , & \text{for } \mu \in \mathcal{P}^*(\mathcal{X}) \\ +\infty \, , & \text{else} \end{cases}, \tag{2.9}$$

with $\Theta : \mathbf{R}_+ \times \mathbf{R}_+ \to \mathbf{R}_+$ defined by $\Theta(a, b) = (a - b)(\log a - \log b)$ and E_μ the set of pairs (x, y) such that $Q_{xy}(\mu) > 0$. In this framework, the Fisher information can be reinterpreted as the squared modulus of the gradient of the entropy with respect to the discrete transport metric $\mathcal{W}$, i.e. we have

$$\mathcal{I}(\mu) = \langle D\mathcal{F}(\mu), \mathcal{K}[\mu] D\mathcal{F}(\mu) \rangle \, .$$

2.2. **Notation.** We use the following notation throughout the paper.
Given a function $\psi \in \mathbf{R}^{\mathcal{X}}$ we denote by $\nabla \psi \in \mathbf{R}^{\mathcal{X} \times \mathcal{X}}$ its discrete gradient, given by

$$\nabla \psi_{xy} = \psi_y - \psi_x \, .$$

For a function $\Psi \in \mathbf{R}^{\mathcal{X} \times \mathcal{X}}$ we denote by $\nabla \cdot \Psi$ its discrete divergence, given by

$$(\nabla \cdot \Psi)_x = \frac{1}{2} \sum_{y \in \mathcal{X}} \Psi_{xy} - \Psi_{yx} \, .$$

For $\psi, \phi \in \mathbf{R}^{\mathcal{X}}$ and $\Psi, \Phi \in \mathbf{R}^{\mathcal{X} \times \mathcal{X}}$ we denote the Euclidean inner products by

$$\langle \psi, \phi \rangle = \sum_{x \in \mathcal{X}} \psi_x \phi_x \,, \qquad \langle \Psi, \Phi \rangle = \frac{1}{2} \sum_{x,y \in \mathcal{X}} \Psi_{xy} \Phi_{xy} \,.$$

Then we have the integration by parts formula

$$\langle \psi, \nabla \cdot \Phi \rangle = -\langle \nabla \psi, \Phi \rangle \,.$$

Given two functions Φ, Ψ in $\mathbf{R}^{\mathcal{X} \times \mathcal{X}}$, we denote by $\Phi \cdot \Psi$ the component-wise product. Using the shorthand notation $\Lambda(\mu)_{xy} := \Lambda(\mu_x Q(\mu)_{xy}, \mu_y Q(\mu)_{yx})$ we can thus write the continuity equation (2.6) and the action functional (2.7) compactly as

$$\dot{\mu}_t + \nabla \cdot \left(\Lambda(\mu_t) \cdot \nabla \psi_t \right) = 0 \,, \qquad \mathcal{A}(\mu, \psi) = \langle \nabla \psi, \Lambda(\mu) \cdot \nabla \psi \rangle \,.$$

We switch freely between notations for the components of functions $\psi \in \mathbf{R}^{\mathcal{X}}$, $\Psi \in \mathbf{R}^{\mathcal{X} \times \mathcal{X}}$ as ψ_x, Ψ_{xy} or $\psi(x), \Psi(x, y)$ depending on what is more readable in the presence of other indices, e.g. a time parameter t.

2.3. **Equilibria and qualitative longtime behavior.** From the gradient flow formulation, it is straightforward to obtain the following characterization of stationary states, which is completely analog to the McKean-Vlasov equation on $\mathbf{R}^n$ [9, Proposition 2.4 and Corollary 2.5].

Proposition 2.3 (Characterization of stationary points)**.** *Let $(\mathcal{X}, Q, \pi)$ be a non-linear Markov triple satisfying Assumption 2.1. Then, the following statements are equivalent:*

(1) π^* is a stationary solution to (1.1), that is $\pi^* Q(\pi^*) = 0$.
(2) π^* is a fixed point of the map $\mu \mapsto \pi(\mu)$ (2.1), that is $\pi^* = \pi(\pi^*)$.
(3) π^* is a critical point of $\mathcal{F}$ (1.2) on $\mathcal{P}(\mathcal{X})$.
(4) π^* is a global minimizer of $\mathcal{I}$ (2.9), that is $\mathcal{I}(\mu^*) = 0$.

The set of all stationary points π^* is denoted by Π^*.
Moreover, it holds that $\Pi^* \subset \mathcal{P}^*(\mathcal{X})$, i.e. each stationary point has strictly positive density.

Proof. *(1)$\Leftrightarrow$(2):* Let $\pi^* Q(\pi^*) = 0$. The rate matrix $Q^* = Q(\pi^*)$ is by assumption the rate matrix of an irreducible reversible Markov chain with unique reversible measure $\pi(\pi^*)$. In particular, it is also the unique stationary solution to $\pi(\pi^*) Q^* = 0$ and hence $\pi^* = \pi(\pi^*)$. If $\pi^* = \pi(\pi^*)$, we calculate using the local detailed balance condition (2.3) and find

$$\sum_{x \in \mathcal{X}} \pi_x^* Q_{xy}(\pi^*) = \sum_{x \in \mathcal{X}} \pi_x(\pi^*) Q_{xy}(\pi^*) = \sum_{x \in \mathcal{X}} \pi_y(\pi^*) Q_{yx}(\pi^*) = 0 \,,$$

since Q is a rate matrix.

(2)$\Leftrightarrow$(3): Take $\mu \in \mathcal{P}^*(\mathcal{X})$ and any $\nu \in \mathcal{P}(\mathcal{X})$. Let $\mu_s = (1-s)\mu + s\nu$ the standard linear interpolation. Then, it holds

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathcal{F}(\mu_s) \Big|_{s=0} = \sum_{x \in \mathcal{X}} \left(\log \mu_x - 1 + \partial_{\mu_x} U(\mu) \right)(\nu_x - \mu_x) = \sum_{x \in \mathcal{X}} \log \frac{\mu_x}{\pi_x(\mu)}(\nu_x - \mu_x) \,,$$

where we used the relations (2.1) and (2.2). Now, if $\mu = \pi^* = \pi(\pi^*)$ the right hand side is zero and hence π^* a critical point if $\mathcal{F}$. On the other hand, if the right hand side is zero

for all $\nu \in \mathcal{P}(\mathcal{X})$, it follows that $\mu_x = C\pi_x(\mu)$ for a constant C. Since $\mu, \pi(\mu) \in \mathcal{P}^*(\mathcal{X})$, we have that $C = 1$ and hence critical points are fixed points.

(2)⇔(4): Let $\pi^* = \pi^*(\pi)$. Since $\mathcal{I}(\mu) \geq 0$ for all $\mu \in \mathcal{P}(\mathcal{X})$, we immediately find from the local detailed balance condition (2.3) that $\mathcal{I}(\pi^*) = 0$. Likewise, any global minimizer π^* satisfies by the definition of $\mathcal{I}$ that $\pi_x^* Q_{xy}(\pi^*) = \pi_y^* Q_{yx}(\pi^*)$, that is the local detailed balance condition (2.3). Since again by assumption $Q(\pi^*)$ has the unique reversible measure $\pi(\pi^*)$, we conclude that $\pi^* = \pi^*(\pi)$.

Finally, the positivity follows from the definition of $\pi(\mu)$ in (2.1) and the assumptions on K implying that H is finite. Hence, $\pi(\mu) \in \mathcal{P}^*(\mathcal{X})$ for all $\mu \in \mathcal{P}(\mathcal{X})$ implies in particular that $\pi(\pi^*) = \pi^* \in \mathcal{P}^*(\mathcal{X})$.   □

Another useful information provided by the gradient flow information is the free energy dissipation relation (2.8), which immediately shows that $\mathcal{F}$ is a Lyapunov function for the evolution (1.1). By standard theory, we can conclude the following qualitative longtime behavior.

Proposition 2.4 (Convergence to stationary points)**.** *Let Q satisfy Assumption 2.1, then* ~~$c(t) \to \pi^*$ for some $\pi^* \in \Pi^*$ as $t \to \infty$~~ $\mathrm{dist}_{\mathcal{P}(\mathcal{X})}(c(t), \Pi^*) \to 0$ as $t \to \infty$.

Proof. The proof follows standard arguments from the theory of dynamical systems (see [38, Section 6]).

By Assumption 2.1, Q is Lipschitz on $\mathcal{P}(\mathcal{X})$, which implies by standard well-posedness for ODEs, that the solutions $(\mu_t)_{t \geq 0}$ to (1.1) are globally defined and generate a semigroup on $\mathcal{P}(\mathcal{X})$. The ω-limit is given by

$$\omega(\mu) = \left\{ \nu \in \mathcal{P}(\mathcal{X}) : \mu_{t_j} \to \nu \text{ for some sequence } t_j \to \infty \right\}.$$

Since $\mathcal{P}(\mathcal{X})$ is compact, each orbit $\mathcal{O}^+(\mu_0) = \bigcup_{t \geq 0} \mu_t$ for any $\mu_0 \in \mathcal{P}(\mathcal{X})$ is also compact in $\mathcal{P}(\mathcal{X})$ and the ω-limit is non-empty and quasi-invariant, that is for $\nu \in \omega(\mu_0)$ it holds $\mathcal{O}^+(\nu) \subseteq \omega(\mu_0)$. Moreover, again thanks to the compactness of $\mathcal{P}(\mathcal{X})$ follows for any $\mu_0 \in \mathcal{P}(\mathcal{X})$ that $\mathrm{dist}_{\mathcal{P}(\mathcal{X})}(\mu_t, \omega(\mu_0)) \to 0$ as $t \to \infty$ (see also [38, Lemma 6.7]).

Since the free energy functional $\mathcal{F}$ is continuous on $\mathcal{P}(\mathcal{X})$ and monotone along the flow, it follows that $\omega(\mu_0)$ consists of complete orbits along which $\mathcal{F}$ has the constant value $\mathcal{F}^\infty = \lim_{t \to \infty} \mathcal{F}(\nu_t)$ with $\nu_0 \in \omega(\mu_0)$. By the free energy dissipation relation (2.8), it follows that for any $\nu_0 \in \omega(\mu_0)$ and any $t > 0$ we have

$$\mathcal{F}^\infty + \int_0^t \mathcal{I}(\nu_s) \, \mathrm{d}s = \mathcal{F}^\infty$$

and hence the non-negativity of $\mathcal{I}$ and continuity of trajectories imply $\mathcal{I}(\nu_s) = 0$ for all $s \in [0, t]$. Hence, $\omega(\mu_0)$ consists of all states ν such that $\mathcal{I}(\nu) = 0$, which by Proposition 2.3 entails $\nu \in \Pi^*$ and moreover also that ν is a stationary solution $\nu Q(\nu) = 0$.   □

Our purpose in this work can be summarized as giving sufficient conditions for which the above statement on convergence to equilibrium can be made quantitative (but which shall automatically enforce that Π^* contains a single element).

3. Curvature for non-linear Markov chains

In this section, we introduce a notion Ricci curvature lower bounds for non-linear Markov chains based on geodesic convexity of the entropy. This generalizes the notion of curvature for linear Markov chains developed in [22] inspired by the approach of Lott, Sturm and Villani [29, 36] to a synthetic notion of lower bounds on Ricci curvature for geodesic metric measure spaces.

Let $(\mathcal{X}, Q, \pi)$ be a non-linear Markov chain according to Assumption 2.1 and let $\mathcal{F}$ be the associated free energy functional (1.2) and $\mathcal{W}$ the associated transport distance.

Definition 3.1 (Entropic Ricci curvature lower bound)**.** We say that $(\mathcal{X}, Q, \pi)$ has *Ricci curvature bounded below by* $\kappa \in \mathbf{R}$ (for short $\mathrm{Ric}(\mathcal{X}, Q) \geq \kappa$) if for any $\mathcal{W}$-geodesic $(\mu_t)_{t \in [0,1]}$:

$$\mathcal{F}(\mu_t) \leq (1-t)\mathcal{F}(\mu_0) + t\mathcal{F}(\mu_1) - \frac{\kappa}{2}t(1-t)\mathcal{W}(\mu_0, \mu_1)^2 \ .$$

We show that a lower bound on the Ricci curvature can be characterized equivalently by a lower bound on the Hessian of the free energy functional $\mathcal{F}$ with respect to the Riemannian structure on $\mathcal{P}^*(\mathcal{X})$ induced by $\mathcal{W}$, or via an Evolution Variational Inequality for the non-linear Markov dynamics.

To this end, we first derive the geodesic equation for the distance $\mathcal{W}$ as well as an expression for the first variation of the free energy.

Lemma 3.2 (Geodesic equation)**.** *Let* $(\mu_t)_{t \in [0,1]}$ *be a constant speed geodesic contained in* $\mathcal{P}^*(\mathcal{X})$*. Then the unique potential* $(\psi_t)_{t \in [0,1]}$ *such that* $(\mu, \psi) \in \mathrm{CE}$ *solves*

$$\dot{\psi}_t(z) + \frac{1}{2}\partial_{\mu(z)}\langle \nabla\psi_t, \Lambda(\mu_t) \cdot \nabla\psi_t \rangle = 0 \ ,$$

or explicitly

$$\dot{\psi}_t(z) + \frac{1}{4}\partial_{\mu(z)}\sum_{x,y}(\psi_t(x) - \psi_t(y))^2 \Lambda\big(\mu_t(x)Q(\mu_t; x, y), \mu_t(y)Q(\mu_t; y, x)\big) = 0 \ , \qquad (3.1)$$

where $\partial_{\mu(z)}$ *is the derivative with respect to* $\mu(z)$*.*

Remark 3.3*.* In the case of a linear Markov chain, where Q is independent of μ, the expression (3.1) simplifies to

$$\dot{\psi}_t(z) + \frac{1}{2}\sum_y(\psi_t(z) - \psi_t(y))^2 \partial_1\Lambda\big(\mu_t(z)Q(z, y), \mu_t(y)Q(y, z)\big)Q(z, y) = 0 \ ,$$

recovering the geodesic equation derived in [22, Prop. 3.4]. For comparison note that $\partial_1\Lambda$ is 0-homogenous and thus $\partial_1\Lambda\big(\mu_t(z)Q(z, y), \mu_t(y)Q(y, z)\big) = \partial_1\Lambda\big(\rho_t(z), \rho_t(y)\big)$ with $\rho_t(x) = \mu_t(x)/\pi(x)$ due to detailed balance.

Proof. Since $\mathcal{P}^*(\mathcal{X})$ is a smooth Riemannian manifold, uniqueness and smoothness of geodesics imply that the curve μ_t is smooth, and that there exists a unique (up to constants) potential ψ_t such that $(\mu, \psi) \in \mathrm{CE}$, which achieves the infimum for the action

$$A(\mu, \psi) = \int_0^1 \mathcal{A}(\mu_t, \psi_t)\,\mathrm{d}t \ , \qquad \mathcal{A}(\mu, \psi) = \langle \nabla\psi, \Lambda(\mu) \cdot \nabla\psi \rangle \ .$$

Moreover, this minimizing ψ is also a smooth curve. We derive (3.1) as the corresponding Euler–Langrange equation. So let $\mu_t^s \in \mathcal{P}^*(\mathcal{X})$ for $s \in [-\varepsilon, \varepsilon]$ be a smooth perturbation of μ such that $\mu_0^s = \mu_0$ and $\mu_1^s = \mu_1$ for all s. Let ψ_t^s be the unique potentials such that $(\mu^s, \psi^s) \in$ CE. Note that ψ_t^s is smooth in s and t. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} A(\mu^s, \psi^s) = 0 \ . \tag{3.2}$$

We compute

$$\frac{\mathrm{d}}{\mathrm{d}s} A(\mu^s, \psi^s) = \int_0^1 2\langle \nabla \psi_t^s, \Lambda(\mu_t^s) \cdot \partial_s \nabla \psi_t^s \rangle + \langle \nabla \psi_t^s, \partial_s \Lambda(\mu_t^s) \cdot \nabla \psi_t^s \rangle \, \mathrm{d}t.$$

From the continuity equation we infer that for any $\phi \in \mathbf{R}^{\mathcal{X}}$

$$\langle \phi, \partial_t \partial_s \mu_t^s \rangle = \langle \phi, \partial_s \partial_t \mu_t^s \rangle = \langle \nabla \phi, \Lambda(\mu_t^s) \cdot \partial_s \nabla \psi_t^s \rangle + \langle \nabla \phi, \partial_s \Lambda(\mu_t^s) \cdot \nabla \psi_t^s \rangle \ .$$

Plugging this into (3.2) for $s = 0$ and integrating by parts in t yields:

$$0 = \int_0^1 2 \left\langle \partial_t \psi_t, \partial_s|_{s=0} \mu_t^s \right\rangle + \langle \nabla \psi_t, \partial_s|_{s=0} \Lambda(\mu_t^s) \cdot \nabla \psi_t \rangle \, \mathrm{d}t \ .$$

The claim then follows by noting that

$$\langle \nabla \psi_t, \partial_s|_{s=0} \Lambda(\mu_t^s) \cdot \nabla \psi_t \rangle$$
$$= \sum_z \partial_s|_{s=0} \mu_t^s(z) \frac{1}{2} \partial_{\mu(z)} \sum_{x,y} (\psi_t(x) - \psi_t(y))^2 \Lambda\big(\mu_t(x)Q(\mu_t; x, y), \mu_t(y)Q(\mu_t; y, x)\big) \ ,$$

and using that the perturbation $\partial_s \mu_t^s$ was arbitrary. $\qquad\square$

In order to give convenient expressions for the first and second variation of the free energy $\mathcal{F}$ along a geodesic, we introduce the following notation.
We set

$$L_\mu \psi(x) := \sum_y Q(\mu; x, y) \psi_y \ , \qquad \hat{L}_\mu \sigma(y) := \sum_x \sigma_x Q(\mu; x, y) \ , \tag{3.3}$$

and note that $\langle L_\mu \psi, \sigma \rangle = \langle \hat{L}_\mu \sigma, \psi \rangle$, so $\hat{L}_\mu$ is the adjoint of L_μ. The master equation (1.1) then reads $\dot{\mu}_t = \hat{L}_{\mu_t} \mu_t$. Note further that we can write

$$\langle \psi, L_\mu \phi \rangle = -\langle \nabla \psi, \big(Q(\mu)\pi(\mu)\big) \cdot \nabla \phi \rangle \ ,$$

where we set $\big(Q(\mu)\pi(\mu)\big)_{xy} = Q(\mu)_{xy} \pi(\mu)_x$, which is symmetric in x, y.

Lemma 3.4 (First variation of the free energy). *Let* $(\mu, \psi) \in$ CE *be a solution to the continuity equation. Then it holds*

$$\frac{d}{dt} \mathcal{F}(\mu_t) = -\langle \mu_t, L_{\mu_t} \psi \rangle \ . \tag{3.4}$$

Note that when the curve is a solution to the gradient flow equation, one can check that the right-hand side is indeed the discrete Fisher information, as given by (2.9).

Proof. Starting from the expression

$$\mathcal{F}(\mu) = \sum_x \mu(x) \log \mu(x) + U(\mu) \ ,$$

recalling that $\partial_{\mu_x} U(\mu) = H_x(\mu) = -\log \pi_x(\mu) - \log Z(\mu)$, and setting $\rho_t = \mu_t/\pi(\mu_t)$, we obtain from the continuity equation

$$
\begin{aligned}
\frac{d}{dt}\mathcal{F}(\mu_t) &= \sum_x \Big(1 - \log Z(\mu_t) + \log \mu_t(x) - \log \pi_x(\mu)\Big)\partial_t \mu_t(x) = \langle \log \rho_t, \dot{\mu}_t \rangle \\
&= \langle \nabla \log \rho_t , \; \Lambda(\mu_t) \cdot \nabla \psi_t \rangle = -\langle \mu_t , \; L_{\mu_t}\psi_t \rangle \, .
\end{aligned}
$$

Here, we have also used in the last step that

$$
\Lambda(\mu_t)(x,y) = \frac{\nabla \rho_t(x,y)}{\nabla \log \rho_t(x,y)} Q(\mu_t; x,y)\pi(\mu_t)(x) \, ,
$$

and integrated by parts. $\qquad\square$

To give an expression of the second variation of $\mathcal{F}$, we further introduce the following notation.

Let $\partial_{\mu_z} Q(\mu; x,y)$ denote the partial derivative of $Q(\cdot; x,y)$ with respect to μ_z. Then we write

$$
DQ(\mu, \sigma; x,y) := \sum_{z \in \mathcal{X}} \partial_{\mu_z} Q(\mu; x,y)\sigma_z \, . \tag{3.5}
$$

Furthermore, let us write

$$
M(\mu)\nabla\psi(x,y) := \sum_{z,w} M(\mu; z,w,x,y)\nabla\psi(z,w) \, , \text{where}
$$

$$
M(\mu; z,w,x,y) := \mu_x \Lambda(\mu)(z,w)\big[\partial_{\mu_z} Q(\mu; x,y) - \partial_{\mu_w} Q(\mu; x,y)\big] \, . \tag{3.6}
$$

Then, we set

$$
\begin{aligned}
\partial_i \Lambda(\mu)(x,y) &= \partial_i \Lambda\big(\mu_x Q(\mu; x,y), \mu_y Q(\mu; y,x)\big) \, , i = 1, 2 \, , \\
\hat{L}\Lambda(\mu)(x,y) &= \partial_1 \Lambda(\mu)(x,y)\hat{L}_\mu \mu(x) Q(\mu; x,y) + \partial_2 \Lambda(\mu)(x,y)\hat{L}_\mu \mu(y) Q(\mu; y,x) \, , \\
R\Lambda(\mu)(x,y) &= \partial_1 \Lambda(\mu)(x,y)\mu(x) DQ(\mu, \hat{L}_\mu \mu; x,y) + \partial_2 \Lambda(\mu)(x,y)\mu(y) DQ(\mu, \hat{L}_\mu \mu; y,x) \, .
\end{aligned}
$$

Finally, we can define the following quantity:

$$
\begin{aligned}
\mathcal{B}(\mu, \psi) := \; &\frac{1}{2}\langle \nabla\psi, \hat{L}\Lambda(\mu) \cdot \nabla\psi \rangle - \langle \nabla\psi, \Lambda(\mu) \cdot \nabla L_\mu \psi \rangle \\
&+ \frac{1}{2}\langle \nabla\psi, R\Lambda(\mu) \cdot \nabla\psi \rangle + \langle \nabla\psi, M(\mu)\nabla\psi \rangle \, ,
\end{aligned} \tag{3.7}
$$

Remark 3.5. Note that in the case of a linear Markov chain, the last two terms in the definition of $\mathcal{B}$ vanish, and we recover the formula of [22] for the second derivative of the entropy along geodesics.

Lemma 3.6 (Second variation of the free energy)**.** *Let $(\mu_t)_t$ be a $\mathcal{W}$-geodesic contained in $\mathcal{P}^*(\mathcal{X})$ and let (ψ_t) be the unique potential such that $(\mu, \psi) \in \mathrm{CE}$. Then it holds*

$$
\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{F}(\mu_t) = \mathrm{Hess}\,\mathcal{F}(\mu_t)[\nabla\psi_t] = \mathcal{B}(\mu_t, \psi_t) \, .
$$

Proof. From (3.4) we get

$$\frac{d^2}{dt^2}\mathcal{F}(\mu(t)) = -\frac{d}{dt}\langle \hat{L}_{\mu_t}\mu_t, \psi_t\rangle$$
$$= -\langle \hat{L}_{\mu_t}\mu_t, \dot{\psi}_t\rangle - \langle \dot{\mu}_t, L_{\mu_t}\psi_t\rangle - \langle \mu_t, (\partial_t L_{\mu_t})\psi_t\rangle$$
$$=: I_1 + I_2 + I_3 \ .$$

To calculate I_1, first note that

$$\partial_{\mu(z)}\Lambda(\mu)(x,y) = \partial_1\Lambda(\mu)Q(\mu;x,y)\delta_{xz} + \partial_2\Lambda(\mu)Q(\mu;y,x)\delta_{yz}$$
$$+ \partial_1\Lambda(\mu)\mu(x)\partial_{\mu_z}Q(\mu;x,y) + \partial_2\Lambda(\mu)\mu(y)\partial_{\mu_z}Q(\mu;y,x) \ ,$$

where δ_{xz} denotes the Kronecker delta. Hence, we infer from the geodesic equation (3.1) and (3.5) that

$$I_1 = \frac{1}{2}\langle \nabla\psi_t, \hat{L}\Lambda(\mu_t)\cdot\nabla\psi_t\rangle + \frac{1}{2}\langle \nabla\psi_t, R\Lambda(\mu_t)\cdot\nabla\psi_t\rangle \ .$$

The continuity equation $\dot{\mu}_t = -\nabla\cdot(\Lambda(\mu_t)\cdot\nabla\psi_t)$ readily yields that

$$I_2 = -\langle \nabla\psi_t, \Lambda(\mu_t)\cdot\nabla L_{\mu_t}\psi_t\rangle \ .$$

To calculate I_3, first note that for any ϕ we have

$$\partial_t L_{\mu_t}\phi(x) = \sum_{y\neq x}\partial_t Q(\mu_t, x, y)\nabla\phi(x,y) = DQ(\mu_t, \dot{\mu}_t)\phi(x)$$
$$= -DQ\Big(\mu_t, \nabla\cdot(\Lambda(\mu_t)\cdot\nabla\psi_t)\Big)\phi(x)$$

where we used the fact that $Q(\mu_t, x, x) = -\sum_{y\neq x}Q(\mu_t, x, y)$. Moreover, for any μ and ψ we have

$$-\Big\langle \mu, DQ\Big(\mu, \nabla\cdot(\Lambda(\mu)\cdot\nabla\psi)\Big)\psi\Big\rangle$$
$$= \sum_{x,y,z}\mu_x\partial_{\mu_z}Q(\mu;x,y)\big[\nabla\cdot(\Lambda(\mu)\nabla\psi)\big](z)\nabla\psi(x,y)$$
$$= \frac{1}{2}\sum_{x,y,z,w}\mu_x\partial_{\mu_z}Q(\mu;x,y)\Lambda(\mu)(z,w)\big[\nabla\psi(w,z) - \nabla\psi(z,w)\big]\nabla\psi(x,y)$$
$$= -\frac{1}{2}\sum_{x,y,z,w}\nabla\psi(x,y)\nabla\psi(z,w)\mu_x\Lambda(\mu)(z,w)\big[\partial_{\mu_z}Q(\mu;x,y) - \partial_{\mu_w}Q(\mu;x,y)\big]$$
$$= -\langle \nabla\psi, M(\mu)\nabla\psi\rangle \ . \tag{3.8}$$

Thus, we get $I_3 = \langle \nabla\psi_t, M(\mu_t)\nabla\psi_t\rangle$. As $I_1 + I_2 + I_3 = \mathcal{B}(\mu_t, \psi_t)$, this yields the claim. $\square$

We can now state the following equivalent characterizations of lower Ricci bounds:

Theorem 3.7. *Let $\kappa \in \mathbf{R}$. For a non-linear Markov triple $(\mathcal{X}, Q, \pi)$ the following assertions are equivalent:*

(1) $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$;
(2) For all $\mu \in \mathcal{P}^*(\mathcal{X})$ and $\psi \in \mathbf{R}^{\mathcal{X}}$ we have

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

(3) The following Evolution Variational Inequality *EVI_κ holds: for all $\mu, \nu \in \mathcal{P}(\mathcal{X})$ and all $t \geq 0$:*

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\mathcal{W}(\mu_t, \nu)^2 + \frac{\kappa}{2}\mathcal{W}(\mu_t, \nu)^2 \leq \mathcal{F}(\nu) - \mathcal{F}(\mu_t) , \qquad \text{EVI}_\kappa$$

where μ_t denotes the solution to the non-linear Fokker–Planck equation starting from μ, i.e. $\dot{\mu}_t = \hat{L}_{\mu_t}\mu_t = \mu_t Q(\mu_t)$ and $\mu_0 = \mu$;

By Lemma 3.6, (2) corresponds to a lower bound κ on the Hessian of $\mathcal{F}$ in the Riemannian structure on $\mathcal{P}^*(\mathcal{X})$ induced by $\mathcal{W}$. Note that the equivalence of (1) and (2) is a non-trivial assertion, since the Riemannian metric degenerates at the boundary of $\mathcal{P}(\mathcal{X})$.

Proof. The proof is based on an argument of Daneri and Savaré [16] suitably adapted to the discrete setting. We can follow verbatim the proof of [22, Thm. 4.5] where the analog of Thm. 3.7 is proven for linear Markov chains. The core of the argument is a variation of the action along the evolution equation, [22, Lem. 4.6]. To accommodate the additional terms arising from the non-linear structure in the present situation, we have to replace that lemma with Lemma 3.8 below. $\qquad\square$

Lemma 3.8. *Let $\{\mu^s\}_{s \in [0,1]}$ be a smooth curve in $\mathcal{P}^*(\mathcal{X})$. For each $t \geq 0$, let μ_t^s denote the solution of the non-linear Fokker–Planck equation at time st starting from μ^s and let $\{\psi_t^s\}_{s \in [0,1]}$ be a smooth curve in $\mathbf{R}^{\mathcal{X}}$ satisfying the continuity equation*

$$\partial_s \mu_t^s + \nabla \cdot (\Lambda(\mu_t^s) \cdot \nabla \psi_t^s) = 0 , \qquad s \in [0,1] .$$

Then the identity

$$\frac{1}{2}\partial_t \mathcal{A}(\mu_t^s, \psi_t^s) + \partial_s \mathcal{F}(\mu_t^s) = -s\mathcal{B}(\mu_t^s, \psi_t^s)$$

holds for every $s \in [0,1]$ and $t \geq 0$.

Proof. First of all, setting $\rho_t^s = \frac{\mu_t^s}{\pi(\mu_t^s)}$ we compute as in Lemma 3.4 that

$$\partial_s \mathcal{F}(\mu_t^s) = \langle \log \rho_t^s , \partial_s \mu_t^s \rangle = -\langle \hat{L}_{\mu_t^s} \mu_t^s , \psi_t^s \rangle .$$

Furthermore,

$$\frac{1}{2}\partial_t \mathcal{A}(\mu_t^s, \psi_t^s) = \langle \partial_t \nabla \psi_t^s , \Lambda(\mu_t^s) \nabla \psi_t^s \rangle + \frac{1}{2}\langle \nabla \psi_t^s , \partial_t \Lambda(\mu_t^s) \cdot \nabla \psi_t^s \rangle$$

$$=: \bar{I}_1 + \bar{I}_2 .$$

In order to further manipulate $\bar{I}_1$ we first note that

$$\partial_t \mu_t^s = s \cdot \hat{L}_{\mu_t^s} \mu_t^s$$

where the factor s comes from the scaling of time in the definition of μ_t^s. Further, we observe that for any $\phi \in \mathbf{R}^{\mathcal{X}}$

$$\langle \nabla \phi, \Lambda(\mu_t^s) \cdot \partial_t \nabla \psi_t^s \rangle + \langle \nabla \phi, \partial_t \Lambda(\mu_t^s) \cdot \nabla \psi_t^s \rangle$$

$$= \langle \mu_t^s, L_{\mu_t^s}\phi \rangle + s\langle \partial_s \mu_t^s, L_{\mu_t^s}\phi \rangle + s\langle \mu_t^s, \partial_s L_{\mu_t^s}\phi \rangle . \qquad (3.9)$$

To show (3.9), note that the left-hand side equals $\partial_t \partial_s \langle \mu_t^s, \phi \rangle$, while the right-hand side equals $\partial_s \partial_t \langle \mu_t^s, \phi \rangle$. Integrating by parts repeatedly and using (3.9) we obtain

$$\bar{I}_1 = -\langle \nabla \psi_t^s \, , \, \partial_t \Lambda(\mu_t^s) \cdot \nabla \psi_t^s \rangle + \langle \mu_t^s, L_{\mu_t^s} \psi_t^s \rangle + s \langle \partial_s \mu_t^s, L_{\mu_t^s} \psi_t^s \rangle + s \langle \mu_t^s, (\partial_s L_{\mu_t^s}) \psi_t^s \rangle$$

$$= -2\bar{I}_2 - \partial_s \mathcal{F}(\mu_t^s) + s \langle \nabla \psi_t^s \, , \, \Lambda(\mu_t^s) \cdot \nabla L_{\mu_t^s} \psi_t^s \rangle + s \langle \mu_t^s, (\partial_s L_{\mu_t^s}) \psi_t^s \rangle$$

Thus, we arrive at

$$\frac{1}{2} \partial_t \mathcal{A}(\mu_t^s, \psi_t^s) + \partial_s \mathcal{F}(\mu_t^s)$$

$$= -\frac{1}{2} \langle \nabla \psi_t^s \, , \, \partial_t \Lambda(\mu_t^s) \cdot \nabla \psi_t^s \rangle + s \langle \nabla \psi_t^s \, , \, \Lambda(\mu_t^s) \cdot \nabla L_{\mu_t^s} \psi_t^s \rangle + s \langle \mu_t^s, (\partial_s L_{\mu_t^s}) \psi_t^s \rangle \ .$$

To conclude, it suffices to note that

$$\partial_t \Lambda(\mu_t^s) = s \cdot \hat{L} \Lambda(\mu_t^s) + s \cdot R \Lambda(\mu_t^s) \ ,$$

further remark that for any ϕ we have

$$\partial_s L_{\mu_t^s} \phi = DQ(\mu_t^s, \partial_s \mu_t^s) \phi = -DQ\Big(\mu_t^s, \nabla \cdot \big(\Lambda(\mu_t^s) \cdot \nabla \psi_t^s\big)\Big) \phi \ ,$$

and then use again (3.8). $\qquad \square$

To end this section, we use Theorem 3.7 to give an expression of the optimal lower Ricci bound on the two-point space.

Lemma 3.9 (Two-point space). *Let* $(\{0,1\}, Q, \pi)$ *be a non-linear Markov triple on the base space* $\mathcal{X} = \{0,1\}$ *and let* $p(\mu) := Q(\mu; 0, 1)$ *and* $q(\mu) := Q(\mu; 1, 0)$ *as well as* $p'(\mu) = [\partial_{\mu_0} - \partial_{\mu_1}] p(\mu)$ *and* $q'(\mu) = [\partial_{\mu_1} - \partial_{\mu_0}] q(\mu)$. *Then, the optimal constant* κ *such that* $\mathrm{Ric}(\{0,1\}, Q, \pi) \geq \kappa$ *is given by*

$$\kappa_{opt} = \inf_{\mu \in \mathcal{P}(\mathcal{X})} \left(\frac{p(\mu) + q(\mu)}{2} + \frac{\mu(0) p'(\mu) + \mu(1) q'(\mu)}{2} \right. \tag{3.10}$$
$$\left. + \frac{\Lambda(\mu)(0,1)}{2} \left(\frac{1}{\mu(0)\,\mu(1)} + \frac{p'(\mu)}{p(\mu)} + \frac{q'(\mu)}{q(\mu)} \right) \right)$$

Remark 3.10. Note that in the case of a linear Markov chain, where p and q are independent of μ, and in particular $p' \equiv 0 \equiv q'$, we recover the formula in [30, Remark 2.11].

Proof. First, we compute from (3.7) for any $\mu \in \mathcal{P}^*(\{0,1\})$ and non-constant ψ:

$$\frac{\mathcal{B}(\mu, \psi)}{(\psi(0) - \psi(1))^2} = \frac{1}{2} \Big(\partial_1 \Lambda(\mu)(0,1) p(\mu) \hat{L}_\mu \mu(0) + \partial_2 \Lambda(\mu)(0,1) q(\mu) \hat{L}_\mu \mu(1) \Big)$$

$$+ \Lambda(\mu)(0,1)(p(\mu) + q(\mu))$$

$$+ \frac{1}{2} \partial_1 \Lambda(\mu)(0,1) \mu(0) \Big(\partial_{\mu_0} p(\mu) \hat{L}_\mu \mu(0) + \partial_{\mu_1} p(\mu) \hat{L}_\mu \mu(1) \Big)$$

$$+ \frac{1}{2} \partial_2 \Lambda(\mu)(0,1) \mu(1) \Big(\partial_{\mu_0} q(\mu) \hat{L}_\mu \mu(0) + \partial_{\mu_1} q(\mu) \hat{L}_\mu \mu(1) \Big)$$

$$+ \Lambda(\mu)(0,1) \big[\mu(0) (\partial_{\mu_0} p(\mu) - \partial_{\mu_1} p(\mu)) - \mu(1) (\partial_{\mu_0} q(\mu) - \partial_{\mu_1} q(\mu)) \big] \ .$$

Now, note that $\hat{L}_\mu \mu(0) = -\hat{L}_\mu \mu(1) = \mu(1)q(\mu) - \mu(0)p(\mu)$, yielding

$$\frac{\mathcal{B}(\mu, \psi)}{(\psi(0) - \psi(1))^2}$$
$$= \Lambda(\mu)(0,1)\big[(p(\mu) + q(\mu) + \mu(0)p'(\mu) + \mu(1)q'(\mu)\big]$$
$$+ \frac{1}{2}\big[\partial_1\Lambda(\mu)(0,1)p(\mu) - \partial_2\Lambda(\mu)(0,1)q(\mu)\big](\mu(1)q(\mu) - \mu(0)p(\mu))$$
$$+ \frac{1}{2}\big[\partial_1\Lambda(\mu)(0,1)\mu(0)p'(\mu) - \partial_2\Lambda(\mu)(0,1)\mu(1)q'(\mu)\big](\mu(1)q(\mu) - \mu(0)p(\mu))$$

Furthermore, $\mathcal{A}(\mu, \psi) = \Lambda(\mu)(0,1)(\psi(1) - \psi(0))^2$. Thus by Theorem 3.7 we get the optimal curvature bound κ_{opt} by dividing the above identity by $\Lambda(\mu)(0,1)$ and minimize in μ. Now, we use the identities

$$\Lambda_1(a,b)(a-b) = \Lambda(a,b) - \frac{\Lambda(a,b)^2}{a} \quad \text{and} \quad \Lambda_2(a,b)(a-b) = -\Lambda(a,b) + \frac{\Lambda(a,b)^2}{b}$$

to get rid of the partial derivatives and obtain after some further simplifications the result (3.10). $\qquad\square$

4. Consequences of Ricci bounds

In this section, we derive consequences of Ricci curvature lower bounds for non-linear Markov chains in terms of functional inequalities and the trend to equilibrium for the dynamics. Throughout this section, let $(\mathcal{X}, Q, \pi)$ be a non-linear Markov triple satisfying Assumption 2.1.

We first note the following expansion bound for the transport distance between solutions to the non-linear Markov dynamics.

Proposition 4.1. *Assume that* $\mathrm{Ric}(\mathcal{X}, Q) \geq \kappa$ *for some* $\kappa \in \mathbf{R}$. *Then for any two solutions* $(\mu_t^i)_{t \geq 0}$ *to the non-linear evolution equation* $\dot\mu_t^i = \mu_t^i Q(\mu_t^i)$, $i = 1, 2$ *we have*

$$\mathcal{W}(\mu_t^1, \mu_t^2) \leq e^{-\kappa t}\mathcal{W}(\mu_0^1, \mu_0^2) .$$

In particular, when $\kappa > 0$, solutions with different initial data get closer at an exponential speed.

Proof. This is a consequence of the EVI$_\kappa$. It follows from [16, Prop. 3.1] applied to the functional $\mathcal{F}$ on the metric space $(\mathcal{P}(\mathcal{X}), \mathcal{W})$. $\qquad\square$

Next, we prove some consequences of Ricci bounds in terms of different functional inequalities. These results can be seen as non-linear discrete analogs of classical results of Bakry and Émery [1] and of Otto and Villani [34]. They extend results that have been obtained in [22] for linear Markov chains, and are reminiscent of results of Carrillo, McCann and Villani [10] obtained for McKean–Vlasov equations in a continuous setting. Let $\mathcal{F}$ be the free energy functional associated with $(\mathcal{X}, Q, \pi)$ given by

$$\mathcal{F}(\mu) = \sum_{x \in \mathcal{X}} \mu_x \log \mu_x + U(\mu), \qquad \text{with} \qquad U(\mu) = \sum_{z \in \mathcal{X}} \mu_z K_z(\mu),$$

and recall that $\mathcal{F}$ attains its minimum on $\mathcal{P}(\mathcal{X})$. We set

$$\mathcal{F}^*(\mu) := \mathcal{F}(\mu) - \min_{\nu \in \mathcal{P}(\mathcal{X})} \mathcal{F}(\nu) .$$

so that $\min \mathcal{F}^* = 0$. Recall that $\mathcal{I}$ is the discrete Fisher information, given by

$$\mathcal{I}(\mu) = \frac{1}{2} \sum_{x,y \in \mathcal{X}} \Theta(\mu_x Q_{xy}(\mu), \mu_y Q_{yx}(\mu)) , \qquad \Theta(a,b) = (a-b)(\log a - \log b) ,$$

provided $\mu \in \mathcal{P}^*(\mathcal{X})$ and $\mathcal{I}(\mu) = +\infty$ else. Recall that $\mathcal{I}$ gives the dissipation of $\mathcal{F}$ along a solution (μ_t) to the non-linear Fokker–Planck equation $\dot{\mu}_t = \mu_t Q(\mu_t)$. More precisely, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{F}(\mu_t) = -\mathcal{I}(\mu_t) .$$

Note further that with $\rho = \mu/\pi(\mu)$ we have the expression $\mathcal{I}(\mu) = \mathcal{A}(\mu, -\log \rho)$. The next result relates $\mathcal{F}$, $\mathcal{I}$ and the transport distance $\mathcal{W}$ under a Ricci bound.

Theorem 4.2. *Assume that* $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$ *for some* $\kappa \in \mathbf{R}$. *Then the* $\mathcal{FWI}$ *inequality holds with constant* $\kappa \in \mathbf{R}$, *i.e. for all* $\mu, \nu \in \mathcal{P}(\mathcal{X})$,

$$\mathcal{F}(\mu) \leq \mathcal{F}(\nu) + \mathcal{W}(\mu, \nu)\sqrt{\mathcal{I}(\mu)} - \frac{\kappa}{2}\mathcal{W}(\mu, \nu)^2 . \qquad\qquad \mathcal{FWI}(\kappa)$$

Proof. Fix $\mu, \nu \in \mathcal{P}(\mathcal{X})$ and assume without restriction that $\mu \in \mathcal{P}^*(\mathcal{X})$ since otherwise there is nothing to prove. Denote by μ_t the solution to $\dot{\mu}_t = \mu_t Q(\mu_t)$ with $\mu_0 = \mu$ and set $\rho_t = \mu_t/\pi(\mu_t)$. Theorem 3.7 yields that EVI_κ holds, so in particular for $t = 0$:

$$\mathcal{F}(\mu) \leq \mathcal{F}(\nu) - \frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} \mathcal{W}(\mu_t, \nu)^2 - \frac{\kappa}{2}\mathcal{W}(\mu, \nu)^2 .$$

From the triangle inequality and the fact that $t \mapsto \mu_t$ is continuous with respect to $\mathcal{W}$ we obtain

$$-\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} \mathcal{W}(\mu_t, \nu)^2 = \liminf_{s \searrow 0} \frac{1}{2s}\left(\mathcal{W}(\mu, \nu)^2 - \mathcal{W}(\mu_s, \nu)^2\right)$$

$$\leq \limsup_{s \searrow 0} \frac{1}{s}\mathcal{W}(\mu_s, \mu) \cdot \mathcal{W}(\mu, \nu) .$$

Now, note that since $(\mu_t, -\log \rho_t) \in \mathrm{CE}$ we can estimate

$$\mathcal{W}(\mu_s, \mu) \leq \int_0^s \sqrt{\mathcal{A}(\mu_r, \log \rho_r)} \, dr = \int_0^s \sqrt{\mathcal{I}(\mu_r)} \, dr .$$

Since $t \mapsto \mathcal{I}(\mu_t)$ is a continuous function, we obtain

$$\limsup_{s \searrow 0} \frac{1}{s}\mathcal{W}(\mu_s, \mu) \leq \sqrt{\mathcal{I}(\mu)} , \qquad\qquad (4.1)$$

which yields the claim. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Theorem 4.3. *Assume that* $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \lambda$ *for some* $\lambda > 0$. *Then the following hold:*

(i) *there exists a unique stationary point* π^*, *it is the unique minimizer of* $\mathcal{F}$;

(ii) *the* modified logarithmic Sobolev inequality *with constant* $\lambda > 0$ *holds, i.e. for all* $\mu \in \mathcal{P}(\mathcal{X})$,

$$\mathcal{F}^*(\mu) \leq \frac{1}{2\lambda}\mathcal{I}(\mu) ; \qquad\qquad\qquad \mathrm{MLSI}(\lambda)$$

(iii) *for any solution $(\mu_t)_{t \geq 0}$ to $\dot{\mu}_t = \mu_t Q(\mu_t)$ we have exponential decay of the free energy:*

$$\mathcal{F}^*(\mu_t) \leq e^{-2\lambda t} \mathcal{F}^*(\mu_0) ; \tag{4.2}$$

(iv) *the* transport-entropy inequality *with constant $\lambda > 0$ holds, i.e. for all $\mu \in \mathcal{P}(\mathcal{X})$,*

$$\mathcal{W}(\mu, \pi^*) \leq \sqrt{\frac{2}{\lambda} \mathcal{F}^*(\mu)} . \tag{ET(λ)}$$

Proof. **(i)** From Proposition 2.3 we know that the set Π^* of stationary points is non-empty and that it coincides with the set of local minimizers of $\mathcal{F}$. Assume by contradiction that F has two distinct local minima at points μ_0 and μ_1, with $F(\mu_0) \leq F(\mu_1)$ and let $(\mu_s)_{s \in [0,1]}$ be a constant speed geodesic connecting μ_0 and μ_1. Then we infer from $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq \lambda$ that

$$\mathcal{F}(\mu_s) \leq (1-s)\mathcal{F}(\mu_0) + s\mathcal{F}(\mu_1) - \frac{\lambda}{2} s(1-s) \mathcal{W}(\mu_0, \mu_1)^2 .$$

Since μ_1 is a local minimum, there is an $\epsilon > 0$ such that $\mathcal{F}(\mu_{1-\epsilon}) \geq \mathcal{F}(\mu_1)$. This leads to

$$\mathcal{F}(\mu_1) \leq \mathcal{F}(\mu_{1-\epsilon}) \leq \epsilon \mathcal{F}(\mu_0) + (1-\epsilon)\mathcal{F}(\mu_1) - \frac{\lambda}{2}\epsilon(1-\epsilon) < \mathcal{F}(\mu_1) ,$$

a contradiction. Hence, $\Pi^* = \{\pi^*\}$ is a singleton and π^* is the unique global minimizer of $\mathcal{F}$.

(ii) By Theorem 4.2, we have that $\mathcal{FWI}(\kappa)$ holds. Applying $\mathcal{FWI}(\kappa)$ with $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu = \pi^*$, noting that $\mathcal{F}^*(\pi^*) = 0$, and using Young's inequality

$$xy \leq cx^2 + \frac{1}{4c}y^2 \qquad \forall x, y \in \mathbf{R}, \ c > 0,$$

with $x = \mathcal{W}(\mu, \pi^*)$, $y = \sqrt{\mathcal{I}(\mu)}$ and $c = \lambda/2$ yields the claim.

(iii) From $\mathrm{MLSI}(\lambda)$ we infer that for a solution $(\mu_t)_t$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{F}^*(\mu_t) = -\mathcal{I}(\mu_t) \leq -2\lambda \mathcal{F}^*(\mu_t) ,$$

and we obtain (4.2) as a consequence of Gronwall's lemma.

(iv) It suffices to establish $\mathrm{ET}(\lambda)$ for any $\mu \in \mathcal{P}^*(\mathcal{X})$. The inequality for general μ can then be obtained by approximation, taking into account the continuity of $\mathcal{W}$ with respect to the Euclidean metric on $\mathcal{P}(\mathcal{X})$. So fix $\mu \in \mathcal{P}^*(\mathcal{X})$, and let μ_t be the solution to the non-linear Fokker–Planck equation starting from μ. From Proposition 2.4 we have that $\mu_t \to \pi^*$ as $t \to \infty$ and that

$$\mathcal{F}^*(\mu_t) \to 0 \quad \text{and} \quad \mathcal{W}(\mu, \mu_t) \to \mathcal{W}(\mu, \pi^*) . \tag{4.3}$$

The last property follows from the continuity of $\mathcal{W}$ with respect to the Euclidean distance. We now define the function $G : \mathbf{R}_+ \to \mathbf{R}_+$ by

$$G(t) := \mathcal{W}(\mu, \mu_t) + \sqrt{\frac{2}{\lambda} \mathcal{F}^*(\mu_t)} .$$

Obviously we have $G(0) = \sqrt{\frac{2}{\lambda}\mathcal{F}^*(\mu)}$ and by (4.3) we have that $G(t) \to \mathcal{W}(\mu, \pi^*)$ as $t \to \infty$. Hence it is sufficient to show that G is non-increasing. To this end we show that its upper right derivative is non-positive. If $\mu_t \neq \pi^*$ we deduce from (4.1) that

$$\frac{\mathrm{d}^+}{\mathrm{d}t} G(t) \;\leq\; \sqrt{\mathcal{I}(\mu_t)} - \frac{\mathcal{I}(\mu_t)}{\sqrt{2\lambda\mathcal{F}^*(\mu_t)}} \;\leq\; 0\,,$$

where we used MLSI(λ) in the last inequality. If $\mu_t = \pi^*$, then the relation also holds true, since this implies that $\mu_r = \pi^*$ for all $r \geq t$. $\qquad\square$

5. Some examples of curvature bounds

We shall now compute lower bounds on the curvature for several examples of mean-field dynamics, inspired by classical models of statistical physics.

5.1. **Curie-Weiss model.** Let us consider the following example also mentioned in [5, Example 4.2], which is the infinite particle limit of the classical Curie-Weiss model, one of the simplest examples of Markovian dynamic exhibiting a phase transition. Let us take $\mathcal{X} = \{0, 1\}$ and define K for $\beta > 0$ by

$$K_x(\nu) = V_x + \beta \sum_{y \in \mathcal{X}} W_{xy}\, \nu_y, \qquad (x, \nu) \in \mathcal{X} \times \mathcal{P}(\mathcal{X}),$$

with $V \equiv 0$, $W(0, 0) = 0 = W(1, 1)$, and $W(0, 1) = 1 = W(1, 0)$. Hence, we have

$$U(\nu) = 2\beta\nu_0\nu_1\,.$$

The free energy $\mathcal{F}(\mu)$ for the Curie-Weiss model is given by

$$\sum_{x \in \mathcal{X}} (\log \mu(x) + K(\mu; x))\mu(x) = \sum_{x \in \mathcal{X}} \left(\log \mu(x) + \beta \sum_{y \in \mathcal{X}} W(x, y)\mu(y)\mu(x)\right)$$
$$= \mu(0) \log \mu(0) + \mu(1) \log \mu(1) + 2\beta\mu(0)\mu(1).$$

Since $\mu(0) + \mu(1) = 1$, we have that the free energy is essentially given by the function $f_\beta : [0, 1] \to \mathbf{R}$

$$f_\beta(u) = u \log u + (1 - u) \log(1 - u) + 2\beta u(1 - u) \quad \text{and} \quad f_\beta''(u) = \frac{1}{u(1 - u)} - 4\beta,$$

Hence, f_β is convex on $[0, 1]$ for $\beta \in [0, 1]$ and non-convex for $\beta > 1$.
The local detailed balance state $\pi(\mu)$ (2.1) is given by

$$\pi_x(\mu) = \frac{1}{Z(\mu)} \exp(-H_x(\mu)) = \frac{1}{Z(\mu)} \exp\left(-2\beta \sum_{y \in \mathcal{X}} W(x, y)\mu(y)\right)\,.$$

Therefore, it holds

$$\pi_0(\mu) = \frac{\exp(-2\beta\mu(1))}{\exp(-2\beta\mu(0)) + \exp(-2\beta\mu(1))}\,,$$
$$\pi_1(\mu) = \frac{\exp(-2\beta\mu(0))}{\exp(-2\beta\mu(0)) + \exp(-2\beta\mu(1))}\,.$$

We use Glauber rates and set

$$p(\mu) = Q(\mu; 0, 1) = \sqrt{\frac{\pi_1(\mu)}{\pi_0(\mu)}} = \exp\bigl(-\beta(\mu(0) - \mu(1))\bigr) \,,$$

$$q(\mu) = Q(\mu; 1, 0) = \sqrt{\frac{\pi_0(\mu)}{\pi_1(\mu)}} = \exp\bigl(\beta(\mu(0) - \mu(1))\bigr) = \frac{1}{p(\mu)} \,.$$

With this choice, we can estimate the Ricci curvature of the limit with the help of Lemma 3.9.

Proposition 5.1 (λ-Convexity of Curie-Weiss model with Glauber rates)**.** *It holds for* $\beta \in [0, 1]$

$$\kappa_{\text{Glauber}} = 2(1 - \beta). \tag{5.1}$$

As a consequence of this curvature bound, one can derive the modified logarithmic Sobolev inequality for the non-linear dynamic. This inequality could also be derived from a logarithmic Sobolev inequality for the particle Gibbs sampler of [31] and passing to the limit in the number of particles. In [27], the mLSI was also derived via convexity of the entropy, but along a different family of interpolations of probability measures. At a technical level, the proof of [27] requires differentiating the entropy three times rather than two, which involves more technical estimates (this is not much of an issue for a two-point space system like Curie-Weiss, but gets much more complicated for more involved systems, like the ones we shall see later in this section).

Proof. We set $\mu(0) = u$ and $\mu(1) = 1 - u$, for which the rates become $p(\mu) = \exp(-\beta(2u - 1)) = 1/q(u)$. First, we note that with the notation of Lemma 3.9, we have $p'(\mu) = -2\beta p(\mu)$ and $q'(\mu) = -2\beta q(\mu)$.

The expression in the infimum of (3.10) to optimize becomes

$$\kappa(u) = \frac{p + q}{2} - \beta(up + (1 - u)q) + \frac{\Lambda(up, (1 - u)q)}{2}\left(\frac{1}{u(1 - u)} - 4\beta\right).$$

It be convenient to do the variable substitution $x = 2u - 1$. For obtaining the expression in a compact manner, we introduce two auxiliary functions

$$g_1(x) := \cosh(\beta x) - x \sinh(\beta x) \qquad \text{and} \qquad g_2(x) := \sinh(\beta x) - x \cosh(\beta x).$$

We then obtain, using the identities $up + (1 - u)q = g_1(x)$, $up - (1 - u)q = -g_2(x)$ and $\operatorname{arctanh}(x) = \frac{1}{2}\log\frac{1+x}{1-x}$ and after some rewriting,

$$\kappa(x) = \cosh(\beta x) - \beta g_1(x) + \frac{g_2(x)}{\beta x - \operatorname{arctanh}(x)}\left(\frac{1}{1 - x^2} - \beta\right). \tag{5.2}$$

A simple evaluation yields $\kappa(0) = 2(1 - \beta)$, where we note $\frac{g_2(x)}{\beta x - \operatorname{arctanh}(x)} \to 1$ as $x \to 0$. For the lower bound, we proceed in several steps, we first observe that

$$\cosh(\beta x) - \beta g_1(x) = (1 - \beta)\cosh(\beta x) + \beta x \sinh(\beta x) \geq 1 - \beta \,.$$

Now, the claim follows once we have shown

$$\frac{g_2(x)}{\beta x - \operatorname{arctanh}(x)} \geq 1 - x^2 \,. \tag{5.3}$$

Indeed, the last term in (5.2), combined with the above estimate, is bounded from below by $1 - (1-x^2)\beta \geq 1 - \beta$, which proves (5.1). To prove (5.3), we do another substitution and set $x = \tanh(y)$. Therefore, the function g_2 becomes after transformations by hyperbolic trigonometric identities

$$g_2(\tanh y) = \frac{\sinh(\beta \tanh(y) - y)}{\cosh y}.$$

and we can estimate the left hand side of (5.3) by

$$\frac{g_2(x)}{\beta x - \operatorname{arctanh}(x)} = \frac{1}{\cosh(y)} \frac{\sinh(\beta \tanh(y) - y)}{\beta \tanh(y) - y} \geq \frac{1}{\cosh(y)} \ ,$$

where we used the bound $\sinh(z)/z \geq 1$ for all $z \in \mathbf{R}$. This can be further estimated by observing that $\cosh(y) \geq 1$ for all $y \in \mathbf{R}$ and by using the identity $1 - \tanh^2 y = 1/\cosh^2 y$, to obtain

$$\frac{1}{\cosh(y)} \geq \frac{1}{\cosh^2(y)} = 1 - \tanh^2(y) = 1 - x^2 \ ,$$

by the substitution $x = \tanh(y)$, which proves (5.3). $\square$

5.2. **General zero-range/misanthrope processes.** In this section, we consider mean-field limits of particle systems with rate matrix of the form

$$Q(\mu; x, y) = p(x, y)\, c(\mu_x, \mu_y)$$

These systems generalize usual linear Markov chains encoded in $p(x, y)$ by an additional dependency of the jump rate on the population density of the departure and arrival site of the jump. This model, first introduced in [15], incorporates many examples, such as for instance the zero range process, for which $c(\mu_x, \mu_y) = b(\mu_x)$, but also interacting agent/voter models [40], for which $c(\mu_x, \mu_y) = a(\mu_y)$.

In this section, our method is perturbative, for which reason we restrict to the complete graph, that is $p(x, y) = 1$ for all $x \neq y$. In this case, the mean-field limit from the N-particle system was derived in [24], and the limit equation was investigated in [35]. Since positive curvature is know in the case of independent particles on the complete graph [22], we expect that for $c(\mu_x, \mu_y) = T + \tilde{c}(\mu_x, \mu_y)$ with bounded $\tilde{c} : \mathcal{P}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \to [0, \infty)$, we should also obtain positive entropic curvature for the non-linear models when T is sufficiently large.

To have a gradient flow formulation, the chain has to satisfy the local detailed balance condition (2.3)

$$\pi_x(\mu)\, c(\mu_x, \mu_y) = \pi_y(\mu)\, c(\mu_y, \mu_x) \ . \tag{5.4}$$

For the further analysis, we focus on the separable case, where for some $a, b : [0, 1] \to \mathbf{R}_+$ holds

$$c(\mu_x, \mu_y) = b(\mu_x)\, a(\mu_y) \ .$$

It is easy to verify that (5.4) is satisfied for

$$\pi_x(\mu) = \frac{1}{Z(\mu)} \frac{a(\mu_x)}{b(\mu_x)} \quad \text{and} \quad Z(\mu) = \sum_{x \in \mathcal{X}} \frac{a(\mu_x)}{b(\mu_x)} \ .$$

This is of the form (2.2) for a potential U given e.g. by

$$U(\mu) = \sum_{x \in \mathcal{X}} u(\mu_x) \qquad \text{with} \qquad u(r) = \int_r^1 \log\left(\frac{a(s)}{b(s)}\right) ds \ ,$$

i.e. for K given by $K_x(\mu) = u(\mu_x)/\mu_x$.

Example 5.2. There are two subclasses of models of particular interest:

$$Q_{xy}(\mu) = a(\mu_y) \qquad \text{and} \qquad Q_{xy}(\mu) = b(\mu_x) \ .$$

Both models satisfy the local detailed balance condition (2.3) for

$$\pi_x^a(\mu) = \frac{a(\mu_x)}{Z^a(\mu)} \qquad \text{with} \qquad Z^a(\mu) = \sum_x a(\mu_x)$$

and

$$\pi_x^b(\mu) = \frac{1}{Z^b(\mu)b(\mu_x)} \qquad \text{with} \qquad Z^b(\mu) = \sum_x \frac{1}{b(\mu_x)}$$

For the first, the interacting agent model from [40] is recovered by setting $a(\mu_y) = T/d + f(\mu_y)$, where $d = |\mathcal{X}|$ is the (constant) degree of the complete graph. For these models, [40] proves a spectral gap via another notion of discrete curvature, but which is not strong enough to derive the mLSI. In the second case, this dynamic corresponds to (a scaling limit of) a zero range-process. This type of particle system is commonly used in statistical physics as a toy model for understanding various large-scale features of interacting systems (scaling limits, longtime behavior, phase transitions). We refer to [28] for an overview. Long-time behavior of the N-particle system in various situations was studied for example in [7, 6, 23, 25]. Recently, Hermon and Salez [26] significantly improved on the state of the art using a combination of the Lu-Yau martingale method and a monotone coupling argument, establishing a modified logarithmic Sobolev inequality independent of the number of particles for mean-field zero-range processes in a non-perturbative setting, even in some inhomogeneous situations where the curvature approach *cannot* work.

In this separable case, we can prove the following statement.

Theorem 5.3 (Curvature for separable kernels)**.** *Assume the rates are separable, given by*

$$Q(\mu; x, y) = b(\mu_x)a(\mu_y) \ .$$

Suppose that

$$0 < \underline{a} \le a(\cdot) \le \bar{a} \qquad \text{and} \qquad 0 < \underline{b} \le b(\cdot) \le \bar{b} \ .$$

Moreover, assume that

$$\lambda := \frac{2\,\bar{a}\operatorname{Lip}b + \bar{a}\operatorname{Lip}a}{2\,\underline{a}\,\underline{b}} \le 1 \ . \tag{5.5}$$

Then $\operatorname{Ric} \ge \kappa$ *in the sense of Theorem* 3.7 *with* κ *given by*

$$\kappa := d\left[\underline{a}\,\underline{b} - (1+\lambda)\frac{\bar{a}\,\bar{b}}{2} - \frac{\bar{a}}{2}\left(2 + \frac{\bar{b}}{\underline{b}}\right)\operatorname{Lip}b - \frac{\bar{b}}{2}\left(1 + \frac{\bar{a}}{\underline{a}}\right)\operatorname{Lip}a\right] \tag{5.6}$$

Especially, in the regime $\frac{\max\{\operatorname{Lip}a, \operatorname{Lip}b\}}{\min\{\underline{a}, \underline{b}\}} =: \eta \ll 1$ *it holds*

$$\kappa \ge d\,\underline{a}\,\underline{b}\,(1 + O(\eta)) \ . \tag{5.7}$$

Proof. First, we evaluate some of the quantities occurring in the derivation of the curvature estimate. Let us start with (3.3), for which we have

$$L_\mu \psi(x) = \sum_y (\psi_y - \psi_x) b(\mu_x) a(\mu_y) \tag{5.8}$$

and

$$\hat{L}_\mu \sigma(y) = \sum_x (\sigma_x b(\mu_x) a(\mu_y) - \sigma_y b(\mu_y) a(\mu_x)) .$$

The next quantity (3.5) becomes

$$DQ(\mu, \sigma; x, y) = \sum_z \partial_{\mu_z} b(\mu_x) a(\mu_y) \sigma_z = b'(\mu_x) a(\mu_y) \sigma_x + b(\mu_x) a'(\mu_y) \sigma_y . \tag{5.9}$$

The last quantity is (3.6)

$$M(\mu; z, w, x, y) = \frac{1}{2} \mu_x \Lambda(\mu)(z, w) \Big(\delta_{z,x} b'(\mu_x) a(\mu_y) + \delta_{z,y} b(\mu_x) a'(\mu_y)$$
$$- \delta_{w,x} b'(\mu_x) a(\mu_y) - \delta_{w,y} b(\mu_x) a'(\mu_y) \Big)$$

from which after symmetrization, we obtain the identity

$$M(\mu) \nabla \psi(x, y) = \mu_x \sum_z (\nabla \psi(x, z) \Lambda(\mu)(x, z) b'(\mu_x) a(\mu_y) + \nabla \psi(y, z) \Lambda(y, z) b(\mu_x) a'(\mu_y))$$
$$\tag{5.10}$$

We use the following identity for the logarithmic mean

$$s \partial_1 \Lambda(s, t) + t \partial_2 \Lambda(s, t) = \Lambda(s, t) . \tag{5.11}$$

To compensate off-diagonal terms, we need the following estimate for the logarithmic mean [23, Lemma A.2]

$$r(\partial_1 \Lambda(s, t) + \partial_2 \Lambda(s, t)) + \Lambda(s, t) \geq \Lambda(r, s) + \Lambda(r, t) . \tag{5.12}$$

The above basic identities shall be used to estimate the four terms in (3.7), which we denote by I, II, III and IV in this order of occurrence.

First term I*:* Let us start estimating the first term in (3.7) and use the identity (5.11)

$$\mathrm{I} = \frac{1}{4} \sum_{x,y} |\nabla \psi(x,y)|^2 \Big(\partial_1 \Lambda(\mu)(x,y) \sum_z (\mu_z b(\mu_z) a(\mu_x) - \mu_x b(\mu_x) a(\mu_z)) b(\mu_x) a(\mu_y)$$
$$+ \partial_2 \Lambda(\mu)(x,y) \sum_z (\mu_z b(\mu_z) a(\mu_y) - \mu_y b(\mu_y) a(\mu_z)) b(\mu_y) a(\mu_x) \Big)$$

$$= \frac{1}{4} \sum_{x,y,z} |\nabla \psi(x,y)|^2 \mu_z b(\mu_z) a(\mu_x) a(\mu_y) \Big(\partial_1 \Lambda(\mu)(x,y) b(\mu_x) + \partial_2 \Lambda(\mu)(x,y) b(\mu_y) \Big)$$

$$- \frac{1}{4} \sum_{x,y} |\nabla \psi(x,y)|^2 A(\mu) \Big(\partial_1 \Lambda(\mu)(x,y) \mu_x b(\mu_x)^2 a(\mu_y) + \partial_2 \Lambda(\mu)(x,y) \mu_y b(\mu_y)^2 a(\mu_x) \Big)$$

$$\geq \mathrm{I}_1 - \frac{\sup_{\mu,x}\{b(\mu_x) A(\mu)\}}{2} \mathcal{A}(\mu, \psi) \geq \mathrm{I}_1 - \frac{d \, \bar{a} \, \bar{b}}{2} \mathcal{A}(\mu, \psi) \tag{5.13}$$

where we introduced $A(\mu) = \sum_z a(\mu_z)$. Although I_1 is non-negative, we keep it to compensate for terms from II and IV. To do so, we compactify notation further by introducing the tilted measure

$$\bar{\mu}_x^{b,a} = \mu_x \frac{b(\mu_x)}{a(\mu_x)} \ .$$

With this definition and with the one-homogeneity of Λ, we can rewrite

$$\Lambda(\mu)(x,y) = \Lambda(\mu_x b(\mu_x) a(\mu_y), \mu_y b(\mu_y) a(\mu_x)) = \Lambda\left(\bar{\mu}_x^{b,a}, \bar{\mu}_y^{b,a}\right) a(\mu_x) a(\mu_y)$$

and likewise for the zero-homogeneous derivatives $(i = 1, 2)$

$$\partial_i \Lambda(\mu)(x,y) = \partial_i \Lambda(\mu_x b(\mu_x) a(\mu_y), \mu_y b(\mu_y) a(\mu_x)) = \partial_i \Lambda\left(\bar{\mu}_x^{b,a}, \bar{\mu}_y^{b,a}\right) \ .$$

With this notation we want to employ the estimate (5.12) in the form

$$\bar{\mu}_z^{b,a}\left(\partial_1 \Lambda\left(\bar{\mu}_x^{b,a}, \bar{\mu}_y^{b,a}\right) + \partial_2 \Lambda\left(\bar{\mu}_x^{b,a}, \bar{\mu}_y^{b,a}\right)\right) \geq \Lambda\left(\bar{\mu}_z^{b,a}, \bar{\mu}_x^{b,a}\right) + \Lambda\left(\bar{\mu}_z^{b,a}, \bar{\mu}_y^{b,a}\right) - \Lambda\left(\bar{\mu}_x^{b,a}, \bar{\mu}_y^{b,a}\right)$$
$$= \frac{\Lambda(\mu)(z,x)}{a(\mu_x)a(\mu_z)} + \frac{\Lambda(\mu)(z,y)}{a(\mu_y)a(\mu_z)} - \frac{\Lambda(\mu)(x,y)}{a(\mu_x)a(\mu_y)} \ .$$

Now, we can bound I_1 from below by

$$\mathrm{I}_1 \geq \underline{b}\, \frac{1}{4} \sum_{x,y,z} |\nabla \psi(x,y)|^2 a(\mu_x) a(\mu_y) a(\mu_z) \bar{\mu}_z^{b,a}\left(\partial_1 \Lambda\left(\bar{\mu}_x^{b,a}, \bar{\mu}_y^{b,a}\right) + \partial_2 \Lambda\left(\bar{\mu}_x^{b,a}, \bar{\mu}_y^{b,a}\right)\right)$$

$$\geq \underline{a}\,\underline{b}\, \frac{1}{4} \sum_{x,y,z} |\nabla \psi(x,y)|^2 (\Lambda(\mu)(z,x) + \Lambda(\mu)(z,y)) - \frac{d\,\bar{a}\,\underline{b}}{2} \mathcal{A}(\mu, \psi)$$

$$\geq \underline{a}\,\underline{b}\, \mathrm{I}_2 - \frac{d\,\bar{a}\,\bar{b}}{2} \mathcal{A}(\mu, \psi) \ . \tag{5.14}$$

Second term II*:* Let us continue with the second term in (3.7) for which we use (5.8), symmetrize the sum and obtain

$$
\begin{aligned}
\mathrm{II} &= -\langle \nabla\psi, \Lambda(\mu)\nabla L_\mu\psi\rangle \\
&= \sum_{x,y,z} \nabla\psi(x,y)\nabla\psi(x,z)\Lambda(\mu)(x,y)b(\mu_x)a(\mu_z) \\
&= \frac{1}{2}\sum_{x,y,z} a(\mu_z)\Lambda(\mu)(x,y)\nabla\psi(x,y)\Big(\underbrace{\nabla\psi(x,z)}_{=\nabla\psi(x,y)+\nabla\psi(y,z)}b(\mu_x) - \nabla\psi(y,z)b(\mu_y)\Big) \\
&= \frac{1}{2}\sum_{x,y}|\nabla\psi(x,y)|^2\Lambda(\mu)(x,y)b(\mu_x)\sum_z a(\mu_z) \\
&\quad + \frac{1}{2}\sum_{x,y,z}\nabla\psi(x,y)\nabla\psi(x,z)\Lambda(\mu)(x,y)a(\mu_z)\big(b(\mu_x)-b(\mu_y)\big) \\
&\geq \left(d\,\underline{b}\,\underline{a} - \frac{d\,\bar{a}\,\operatorname{Lip} b}{2}\right)\mathcal{A}(\mu,\psi) \\
&\quad - \frac{\bar{a}\,\operatorname{Lip} b}{8}\sum_{x,y,z}|\nabla\psi(x,y)|^2\big(\Lambda(\mu)(x,z)+\Lambda(\mu)(y,z)\big) \\
&\geq \left(d\,\underline{b}\,\underline{a} - \frac{d\,\bar{a}\,\operatorname{Lip} b}{2}\right)\mathcal{A}(\mu,\psi) - \frac{\bar{a}\,\operatorname{Lip} b}{2}I_2\ . \tag{5.15}
\end{aligned}
$$

where we used the Young inequality $uv \leq u^2/2 + v^2/2$.

Third term III*:* For estimating the third term in (3.7), denoted by III, we use (5.9), do a crude estimate to again apply (5.11)

$$
\begin{aligned}
\mathrm{III} = \frac{1}{4}\sum_{x,y}|\nabla\psi(x,y)|^2\Big(&\partial_1\Lambda(\mu)(x,y)\mu_x\big(b'(\mu_x)a(\mu_y)L_\mu\mu(x) + b(\mu_x)a'(\mu_y)L_\mu\mu(y)\big) \\
&+ \partial_2\Lambda(\mu)(x,y)\mu_y\big(b'(\mu_y)a(\mu_x)L_\mu\mu(y) + b(\mu_y)a'(\mu_x)L_\mu\mu(x)\big)\Big) \\
\geq \inf_{\mu,x,y}\left\{\frac{b'(\mu_x)L_\mu\mu(x)}{2b(\mu_x)} + \frac{a'(\mu_y)L_\mu\mu(y)}{2a(\mu_y)}\right\}&\mathcal{A}(\mu,\psi)\ .
\end{aligned}
$$

To bound the infimum, we observe that

$$
|L_\mu\mu(x)| \leq d\,\bar{b}\,\bar{a}\ .
$$

Hence, in total we obtain

$$
\mathrm{III} \geq -\frac{d\,\bar{b}\,\bar{a}}{2}\left(\frac{\operatorname{Lip} b}{\underline{b}} + \frac{\operatorname{Lip} a}{\underline{a}}\right)\mathcal{A}(\mu,\psi)\ . \tag{5.16}
$$

Fourth term IV*:* For estimating the fourth term in (3.7), denoted by IV, we use (5.10) and compensate it partly by I_2 from (5.14)

$$\text{IV} = \frac{1}{2}\sum_{x,y,z}\nabla\psi(x,y)\nabla\psi(x,z)\Lambda(\mu)(x,z)\big(\mu_x b'(\mu_x)a(\mu_y) - \mu_y b(\mu_y)a'(\mu_x)\big)$$

$$\geq -\frac{\bar{a}\,\text{Lip}\,b + \bar{b}\,\text{Lip}\,a}{2}\left[d\,\mathcal{A}(\mu,\psi) + \frac{1}{4}\sum_{x,y,z}|\nabla\psi(x,y)|^2\big(\Lambda(x,z) + \Lambda(y,z)\big)\right].\ (5.17)$$

Conclusion: We combine all the estimates of the individual terms in (3.7) from the rewriting $\mathcal{B} = \text{I} + \text{II} + \text{III} + \text{IV}$. There is one small catch. After having applied the first bound (5.13) to I, we split for $\lambda \in (0,1)$ the non-negative part I_1 into $(1-\lambda)\,I_1$ and $\lambda\,I_1$, where only to the second term $\lambda\,I_1$ the bound (5.14) is applied. The other three estimates (5.15), (5.16) and (5.17) are applied in a straightforward manner to II, III and IV, respectively, to arrive at the lower bound

$$\mathcal{B}(\mu,\psi) \geq (1-\lambda)\,I_1 + \left(\lambda\,\underline{a}\,\underline{b} - \frac{2\,\bar{a}\,\text{Lip}\,b + \bar{b}\,\text{Lip}\,a}{2}\right)I_2$$

$$+ d\left[\underline{a}\,\underline{b} - (1+\lambda)\frac{\bar{a}\,\bar{b}}{2} - \frac{\bar{a}}{2}\left(2 + \frac{\bar{b}}{\underline{b}}\right)\text{Lip}\,b - \frac{\bar{b}}{2}\left(1 + \frac{\bar{a}}{\underline{a}}\right)\text{Lip}\,a\right]\mathcal{A}(\mu,\psi)\ .$$

If λ is chosen according to (5.5) and by $I_1 \geq 0$, we arrive at the bound $\mathcal{B}(\mu,\psi) \geq \kappa\mathcal{A}(\mu,\psi)$ with κ given in (5.6). The final statement (5.7) follows by simple calculus from the bound $\bar{a} \leq \underline{a} + \text{Lip}\,a$, similar for $\bar{b} \leq \underline{b} + \text{Lip}\,b$ and observing that $\lambda = O(\eta)$ in this case.   □

References

[1] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[2] Dario Benedetto, Emanuele Caglioti, and Mario Pulvirenti. A kinetic equation for granular media. *RAIRO Modél. Math. Anal. Numér.*, 31(5):615–641, 1997.

[3] François Bolley, Ivan Gentil, and Arnaud Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2):429–445, 2013.

[4] Amarjit Budhiraja, Paul Dupuis, Markus Fischer, and Kavita Ramanan. Limits of relative entropies associated with weakly interacting particle systems. *Electron. J. Probab.*, 20:no. 80, 22, 2015.

[5] Amarjit Budhiraja, Paul Dupuis, Markus Fischer, and Kavita Ramanan. Local stability of Kolmogorov forward equations for finite state nonlinear Markov processes. *Electron. J. Probab.*, 20:no. 81, 30, 2015.

[6] Pietro Caputo, Paolo Dai Pra, and Gustavo Posta. Convex entropy decay via the Bochner-Bakry-Emery approach. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):734–753, 2009.

[7] Pietro Caputo and Gustavo Posta. Entropy dissipation estimates in a zero-range dynamics. *Probab. Theory Related Fields*, 139(1-2):65–87, 2007.

[8] Pietro Caputo and Alistair Sinclair. Entropy production in nonlinear recombination models. *Bernoulli*, 24(4B):3246–3282, 2018.

[9] J. A. Carrillo, R. S. Gvalani, G. A. Pavliotis, and A. Schlichting. Long-time behaviour and phase transitions for the McKean–Vlasov equation on the torus. *Archive for Rational Mechanics and Analysis*, Jul 2019.

[10] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

[11] José A. Carrillo, Robert J. McCann, and Cédric Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.

[12] Patrick Cattiaux, Fanny Delebecque, and Laure Pédèches. Stochastic Cucker–Smale models: Old and new. *Ann. Appl. Probab.*, 28(5):3239–3286, 2018.

[13] Patrick Cattiaux, Arnaud Guillin, and Florent Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2):19–40, 2008.

[14] Shui-Nee Chow, Wen Huang, Yao Li, and Haomin Zhou. Fokker-Planck equations for a free energy functional or Markov process on a graph. *Arch. Ration. Mech. Anal.*, 203(3):969–1008, 2012.

[15] Christiane Cocozza-Thivent. Processus des misanthropes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 70(4):509–523, 1985.

[16] Sara Daneri and Giuseppe Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.

[17] Pierre Degond, Amic Frouvelle, Jian-Guo Liu, Sebastien Motsch, and Laurent Navoret. Macroscopic models of collective motion and self-organization. In *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Année 2012–2013*, Sémin. Équ. Dériv. Partielles, pages Exp. No. I, 27. École Polytech., Palaiseau, 2014.

[18] Andreas Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, 166(3-4):851–886, 2016.

[19] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. Quantitative Harris type theorems for diffusions and McKean-Vlasov processes. *To appear in Transactions AMS*, 2016.

[20] Matthias Erbar, Max Fathi, Vaios Laschos, and André Schlichting. Gradient flow structure for McKean-Vlasov equations on discrete spaces. *Discrete Contin. Dyn. Syst.*, 36(12):6799–6833, 2016.

[21] Matthias Erbar, Christopher Henderson, Georg Menz, and Prasad Tetali. Ricci curvature bounds for weakly interacting Markov chains. *Electron. J. Probab.*, 22:Paper No. 40, 23, 2017.

[22] Matthias Erbar and Jan Maas. Ricci Curvature of Finite Markov Chains via Convexity of the Entropy. *Archive for Rational Mechanics and Analysis*, 206(3):997–1038, August 2012.

[23] Max Fathi and Jan Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.*, 26(3):1774–1806, 2016.

[24] Stefan Grosskinsky and Watthanan Jatuviriyapornchai. Derivation of mean-field equations for stochastic particle systems. *Stoch. Process. their Appl.*, 129(4):1455–1475, 2019.

[25] Jonathan Hermon and Justin Salez. Cutoff for the mean-field zero-range process with bounded monotone rates. *to appear in Ann. Probab.*, 2018.

[26] Jonathan Hermon and Justin Salez. Entropy dissipation estimates for inhomogeneous zero-range processes. *arxiv preprint*, 2019.

[27] Richard Kraaij. A large deviation perspective on exponential decay of entropy and lower bounds on the Ricci-curvature. *arXiv:1609.09637*, 2016.

[28] Thomas M. Liggett. *Interacting Particle Systems*. Springer, 2004.

[29] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

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

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

[32] Alexander Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations*, 48(1-2):1–31, 2013.

[33] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101–174, January 2001.

[34] Felix Otto and Cédric Villani. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *Journal of Functional Analysis*, 173(2):361–400, June 2000.

[35] André Schlichting. The exchange-driven growth model: basic properties and longtime behavior. *arXiv:1811.05954*, 2018.

[36] Karl-Theodor Sturm. On the geometry of metric measure spaces. I and II. *Acta Math.*, 196(1):65–131, 2006.

[37] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[38] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. Graduate Studies in Mathematics. American Mathematical Soc.

[39] Julian Tugaut. Convergence to the equilibria for self-stabilizing processes in double-well landscape. *Ann. Probab.*, 41(3A):1427–1460, 2013.

[40] Denis Villemonais. Lower bound for the coarse Ricci curvature of continuous-time pure jump processes. *arXiv:1705.06642*, 2017.

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany
Email address: `erbar@math.uni-bielefeld.de`

CNRS & Institut de Mathématiques de Toulouse, Université de Toulouse, 118 route de Narbonne, 31062 Toulouse, France
Email address: `max.fathi@math.univ-toulouse.fr`

Institut für Angewande Mathematik, Endenicher Allee 60, Universität Bonn
Email address: `schlichting@iam.uni-bonn.de`