DISPLACEMENT CONVEXITY FOR THE ENTROPY IN SEMIDISCRETE NONLINEAR FOKKER-PLANCK EQUATIONS

JOSÉ A. CARRILLO, ANSGAR JÜNGEL, AND MATHEUS C. SANTOS

Abstract. The displacement λ-convexity of a nonstandard entropy with respect to a nonlocal transportation metric in finite state spaces is shown using a gradient flow approach. The constant λ is computed explicitly in terms of a priori estimates of the solution to a finite-difference approximation of a nonlinear Fokker-Planck equation. The key idea is to employ a new mean function, which defines the Onsager operator in the gradient flow formulation.

1. Introduction

Displacement convexity, which was introduced by McCann [18], describes the geodesic convexity of functionals on the space of probability measures endowed with a transportation metric. Geodesic convexity has important consequences for the existence and uniqueness of gradient flows in the space of probability measures [1, 5, 20]. It may also provide quantitative contraction estimates between solutions of the gradient flows [4] and exponential decay estimates [1]. Displacement λ-convexity of the entropy is equivalent to a lower bound on the Ricci curvature Ric_M of the Riemannian manifold M, i.e. $\mathrm{Ric}_M \geq \lambda$ [15, 21]. Furthermore, it leads to inequalities in convex geometry and probability theory, such as the Brunn-Minkowski, Talagrand, and log-Sobolev inequalities [23].

We are interested in the question to what extent the concept of displacement convexity can be extended to discrete settings, like numerical discretization schemes of gradient flows. As one step in this direction, we show in this paper that a certain entropy functional, related to the finite-difference approximation of nonlinear Fokker-Planck equations, is displacement convex. Before making this statement more precise, let us review the state of the art of the literature.

The study of discrete gradient flows and related topics is very recent. First results were concerned with Ricci curvature bounds in discrete settings [2]. Markov processes and

Date: May 13, 2017.

2000 *Mathematics Subject Classification.* 60J27, 53C21, 65M20.

Key words and phrases. Entropy, displacement convexity, logarithmic mean, finite differences, fast-diffusion equation.

The first author was partially supported by the Royal Society via a Wolfson Research Merit Award and the EPSRC grant EP/P031587/1. The second author acknowledges partial support from the Austrian Science Fund (FWF), grants P22108, P24304, F65, and W1245. The last author acknowledges the support from the São Paulo Research Foundation (FAPESP), grant 2015/20962-7. The authors warmly thank the Mittag-Leffler Institute for providing a marvellous atmosphere for research while this paper was being finalized.

Fokker-Planck equations on finite graphs were investigated by Chow et al. in [6]. Maas [16] and Mielke [19] introduced nonlocal transportation distances on probability spaces such that continuous-time Markov chains can be formulated as gradient flows of the entropy, and they explored geodesic convexity properties of the functionals. The concept of displacement convexity was used by Gozlan et al. [11] to derive HWI and log-Sobolev inequalities on (complete) graphs. Talagrand's inequality was studied in discrete spaces by Sammer and Tetali [22].

Only few results can be found in the literature on convexity properties of functionals for discretizations of partial differential equations. Exponential decay rates for time-continuous Markov chains were derived by Caputo et al. [3]. This result was also obtained for reversible Markov chains as a consequence of the displacement convexity of the Shannon entropy as first investigated by Mielke [19] and applied to discretizations of one-dimensional linear Fokker-Planck equations (also see the presentation in [13, Section 5.2]). While the proof of Caputo et al. [3] is based on the Bochner-Bakry-Emery method, Mielke [19] employed a gradient flow approach together with matrix estimates. The nonlocal transportation metric, needed for the definition of displacement convexity, is induced by the logarithmic mean,

$$\Lambda(s,t) = \frac{s-t}{\log s - \log t} \quad \text{for } s \neq t, \quad \Lambda(s,s) = s,$$

which has some remarkable properties (proved in [19] and summarized in Lemma 5 below). The approach of [3] (and [10]) was extended to general convex entropy densities $f(s)$ in [14] using the mean function

$$\Lambda^f(s,t) = \frac{s-t}{f'(s) - f'(t)} \quad \text{for } s \neq t, \quad \Lambda^f(s,s) = \frac{1}{f''(s)}, \tag{1}$$

which becomes the logarithmic mean for $f(s) = s(\log s - 1)$.

Concerning nonlinear equations, we are aware only of two results. Erbar and Maas [9] showed that a discrete one-dimensional porous-medium equation is a gradient flow of the Rényi entropy function $f(s) = s^\alpha$ with respect to a suitable nonlocal transportation metric induced by the mean function

$$\Lambda^\alpha(s,t) = \frac{\alpha-1}{\alpha} \frac{s^\alpha - t^\alpha}{s^{\alpha-1} - t^{\alpha-1}} \quad \text{for } s \neq t, \quad \Lambda^\alpha(s,s) = s.$$

However, the Rényi entropy fails to be convex along geodesics with respect to this transportation metric [9]. A weaker notion than geodesic convexity (called convex entropy decay), which is strongly related to the Bakry-Emery method, was introduced by Maas and Matthes [17] to prove exponential decay rates for finite-volume discretizations of the quantum drift-diffusion equation. Its gradient flow formulation is based on the Fisher information and the logarithmic mean.

In this, paper, we propose a new mean function by composing the logarithmic mean with a nonlinear function (coming from the diffusivity), which is suitable for finite-difference discretizations of the nonlinear Fokker-Planck equation

$$\partial_t \rho = \partial_x\big(\partial_x \phi(\rho) + \phi(\rho)\partial_x V\big), \quad x \in (0,1), \ t > 0, \tag{2}$$

supplemented with no-flux boundary conditions and an initial condition. Here, $\phi : [0,\infty) \to [0,\infty)$ is a continuous function and $V(x)$ is a confinement potential. An example is $\phi(\rho) = \rho^\alpha$ with $\alpha > 0$ and $V(x) = \gamma|x|^2/2$ with $\gamma \geq 0$. A computation shows that the entropy

$$\mathcal{F}_c(\rho) = \int_0^1 \big(f(\rho) + \rho V(x)\big)dx, \quad \text{where } f'(s) = \log\phi(s),$$

is nonincreasing along (smooth) solutions to (2). The displacement convexity of equations related to (2) was analyzed in [5]. Our aim is to show that a discrete version of the entropy $\mathcal{F}_c$ is displacement convex along semidiscrete solutions associated to (2).

For the discretization of (2), let $n \in \mathbb{N}$, $h = 1/n > 0$, and $x_i = ih$, $i = 0,\ldots,n$. Let $\rho_i(t)$ approximate the solution $\rho(x_i,t)$ and w_i approximate the function $w(x_i) = e^{-V(x_i)}$. Writing (2) in the form

$$\partial_t\rho = \operatorname{div}\left(\phi(\rho)\nabla\log\frac{\phi(\rho)}{w}\right),$$

a corresponding finite-difference scheme reads as

$$\partial_t\rho_i = \frac{\kappa_i\Lambda_i}{h^2}\left(\log\frac{\phi(\rho_{i+1})}{w_{i+1}} - \log\frac{\phi(\rho_i)}{w_i}\right) - \frac{\kappa_{i-1}\Lambda_{i-1}}{h^2}\left(\log\frac{\phi(\rho_i)}{w_i} - \log\frac{\phi(\rho_{i-1})}{w_{i-1}}\right), \tag{3}$$

where $h > 0$ is the space size and $\kappa_i\Lambda_i$ is an approximation of $\phi(\rho)$ in $[x_i, x_{i+1}]$. Our idea is to employ the *modified* logarithmic mean

$$\Lambda_i = \frac{u_i - u_{i+1}}{\log u_i - \log u_{i+1}} = \frac{\phi(\rho_i)/w_i - \phi(\rho_{i+1})/w_{i+1}}{\log(\phi(\rho_i)/w_i) - \log(\phi(\rho_{i+1})/w_{i+1})}, \tag{4}$$

and to set, as in [19], $\kappa_i = \sqrt{w_i w_{i+1}}$. Since Λ_i approximates $u_i = \phi(\rho_i)/w_i$, it follows that $\kappa_i\Lambda_i$ approximates $\sqrt{w_{i+1}/w_i}\phi(\rho_i)$. Observe that with this choice, the numerical scheme reduces to

$$\partial_t\rho_i = \frac{\kappa_i}{h^2}(u_{i+1} - u_i) - \frac{\kappa_{i-1}}{h^2}(u_i - u_{i-1}), \quad u_i = \frac{\phi(\rho_i)}{w_i},$$

which approximates (2) written in the form $\partial_t\rho = \partial_x(w\partial_x(\phi(\rho)/w))$.

The main result of the paper is as follows. If ϕ is invertible and $\phi' \circ \phi^{-1}$ is nonincreasing (an example is $\phi(s) = s^\alpha$ with $0 < \alpha < 1$), then the discrete entropy

$$\mathcal{F}(\rho) = \sum_{i=0}^n \big(f(\rho_i) + \rho_i V(x_i)\big), \quad \text{where } f'(s) = \log\phi(s), \tag{5}$$

is displacement λ_h-convex with respect to the nonlocal transportation metric induced by (4), where

$$\lambda_h = \gamma\left(\frac{2}{\gamma h^2}(1 - e^{-\gamma h^2/2})\min_{i=0,\ldots,n}\phi'(\rho_i) - 2\cosh(\gamma h)\max_{i=0,\ldots,n}|\nabla_h\phi'(\rho_i)|\right) \in \mathbb{R},$$

and $\nabla_h\phi'(\rho_i) = h^{-1}(\phi'(\rho_{i+1}) - \phi'(\rho_i))$; see Theorem 3.

Our result is consistent with that one in [19]: If $\phi(s) = s$ is linear (and $V \neq 0$), $\lambda_h \to \gamma$ as $h \to 0$, and the constant is asymptotically sharp. If the minimum of $\phi'(\rho_i)$ is positive and the maximum of $|\nabla_h\phi'(\rho_i)|$ is sufficiently small, then λ_h is positive. We expect that exponential convergence to the steady state holds for sufficiently small $h > 0$, but we are unable to prove these a-priori estimates for our numerical scheme in this whole generality. Such bounds in terms of the initial data can be shown at least in the very restrictive case $V = 0$ and for small initial data depending on the mesh size h; see Corollary 1.

The paper is organized as follows. In Section 2, we introduce the mathematical setting and give the definition of displacement λ-convexity. We show that displacement λ-convexity follows if a certain matrix is positive semidefinite, slightly generalizing Proposition 2.1 in [19]. As a warm-up, we consider in Section 3 the semidiscrete heat equation and prove that the entropy $\mathcal{F}(\rho) = \sum_{i=0}^n f(\rho_i)$ is displacement convex if $f(s) = s(\log s - 1)$ or $f(s) = s^\alpha$ for $1 < \alpha \leq 2$; see Theorem 2. This result is a reformulation of Theorem 5 in [14], but our proof is very simple. Section 4 is concerned with the proof of displacement λ-convexity of (5) and contains our main result. Some properties of mean functions are recalled in Appendix A, and a priori estimates of solutions to (3) with $V = 0$ and small intial data depending on the small size are proved in Appendix B.

2. Displacement convexity

In this section, we specify our setting and give the definition of displacement convexity. Let $n \in \mathbb{N}$ and introduce the finite state space

$$X_n = \left\{ \rho = (\rho_0, \ldots, \rho_n) \in \mathbb{R}^{n+1} : \rho_0, \ldots, \rho_n > 0, \ \sum_{i=0}^n \rho_i = 1 \right\}.$$

This space can be identified with the space of strictly positive probability measures on a $(n+1)$-point set. We will denote by $\langle \cdot, \cdot \rangle$ the Euclidean inner product in $\mathbb{R}^{n+1}$. Let a matrix $Q = (Q_{ij}) \in \mathbb{R}^{(n+1)\times(n+1)}$ be given such that

$$Q_{ij} \geq 0 \text{ for } i \neq j, \quad \sum_{i=0}^n Q_{ij} = 0 \text{ for } j = 1, \ldots, n.$$

The value Q_{ij} is the rate of a particle moving from state j to i. We assume that there exists a unique vector $w \in X_n$ such that the detailed balance condition

$$Q_{ij} w_j = Q_{ji} w_i \quad \text{for all } i, j = 0, \ldots, n$$

is satisfied. Summing this condition for fixed i over $j = 0, \ldots, n$, we see that $Qw = 0$. Note that in Markov chain theory, the detailed balance condition is usually formulated for the transposed matrix $Q^\top$.

Our aim is to show convexity properties of the entropy along solutions $t \mapsto \rho(t)$ to ODE systems of the type

$$\partial_t \rho = Q\phi(\rho), \quad t > 0, \tag{6}$$

where ϕ is some smooth function. This equation can be formulated as a gradient flow. Indeed, given a (smooth) function $f : [0, \infty) \to \mathbb{R}$, we define the entropy $\mathcal{F} : X_n \to \mathbb{R}$,

$$\mathcal{F}(\rho) = \sum_{i=0}^n f_i(\rho_i), \quad \text{where } f_i'(s) = f'\left(\frac{\phi(s)}{w_i}\right), \tag{7}$$

and the Onsager operator $K : X_n \to \mathbb{R}^{(n+1)\times(n+1)}$,

$$K(\rho) = \frac{1}{2}\sum_{i,j=0}^{n} Q_{ij}w_j\Lambda^f\left(\frac{\phi(\rho_i)}{w_i}, \frac{\phi(\rho_j)}{w_j}\right)(e_i - e_j)\otimes(e_i - e_j), \tag{8}$$

where $e_i = (\delta_{i0}, \ldots, \delta_{in})^\top \in \mathbb{R}^{n+1}$ is the ith unit vector and "$\otimes$" is the tensor product. By detailed balance and $Q_{ij}w_j \geq 0$ for $i \neq j$, it follows that $K(\rho)$ is symmetric and positive semidefinite. With these definitions, we can formulate (6) as a gradient system in the sense that it can be rewritten as

$$\partial_t\rho = -K(\rho)D\mathcal{F}(\rho), \tag{9}$$

where $D\mathcal{F}(\rho) = (f_0'(\rho_0), \ldots, f_n'(\rho_n))$.

The space X_n is endowed with the nonlocal transportation distance

$$\mathcal{W}(\rho_0, \rho_1)^2 = \inf_{(\rho,\psi)\in E(\rho_0,\rho_1)} \int_0^1 \big\langle K(\rho(t))\psi(t), \psi(t)\big\rangle dt, \tag{10}$$

where $E(\rho_0, \rho_1)$ is the set of pairs $(\rho(t), \psi(t))$, $t \in [0, 1]$, such that

$$\begin{aligned}&\rho \in C^1([0,1]; X_n),\ \ \psi : [0,1] \to \mathbb{R}^{n+1} \text{ is measurable},\\ &\text{for all } i = 0, \ldots, n,\ t \in [0,1] :\ \partial_t\rho(t) = -K(\rho)\psi(t), \quad \rho(0) = \rho_0,\ \rho(1) = \rho_1.\end{aligned}$$

It is well known that the function $\mathcal{W}$ is a pseudo-metric on $\overline{X}_n$ (the space of probability measures on a $(n+1)$-point set) [16, Theorem 1.1] and the pair $(\overline{X}_n, \mathcal{W})$ defines a geodesic space [9, Prop. 2.3], i.e., for all ρ_0, $\rho_1 \in \overline{X}_n$, there exists at least one curve $\rho : [0, 1] \to \overline{X}_n$, $t \mapsto \rho(t)$, such that $\rho(0) = \rho_0$, $\rho(1) = \rho_1$, and $\mathcal{W}(\rho(s), \rho(t)) = |s - t|\mathcal{W}(\rho_0, \rho_1)$ for all s, $t \in [0, 1]$. Such a curve is called a constant speed geodesics between ρ_0 and ρ_1. By [16, Lemma 3.30], any geodesic can be approximated by curves in X_n. If the pair $(\rho, \psi) \in E(\rho_0, \rho_1)$ attains the infimum in (10), then ρ is a geodesic and satisfies the geodesic equations [9, Prop. 2.5]

$$\begin{cases}\partial_t\rho = K(\rho)\psi\\ \partial_t\psi = -\frac{1}{2}\langle DK(\rho)[\,\cdot\,]\psi, \psi\rangle\end{cases}, \quad t > 0, \tag{11}$$

where the vector $b = \langle DK(\rho)[\,\cdot\,]\psi, \psi\rangle$ is defined by $\langle b, v\rangle = \langle DK(\rho)[v]\psi, \psi\rangle$ for $v \in \mathbb{R}^{n+1}$.

Definition 1 (Displacement convexity). *Let $\lambda \in \mathbb{R}$. We say that a functional $\mathcal{E} : X_n \to \mathbb{R}\cup\{+\infty\}$ is displacement λ-convex on X_n with respect to the metric $\mathcal{W}$ if for any constant speed geodesic curve $\rho : [0, 1] \to X_n$,*

$$\mathcal{E}(\rho(t)) \leq (1-t)\mathcal{E}(\rho(0)) + t\mathcal{E}(\rho(1)) - \frac{\lambda}{2}t(1-t)\mathcal{W}(\rho(0), \rho(1))^2, \quad t \in [0, 1].$$

If $\lambda = 0$, $\mathcal{E}$ is simply called displacement convex. Moreover, if $t \mapsto \mathcal{E}(\rho(t))$ is twice differentiable, $\mathcal{E}$ is displacement λ-convex if and only if

$$\frac{d^2}{dt^2}\mathcal{E}(\rho(t)) \geq \lambda\mathcal{W}(\rho(0), \rho(1))^2, \quad t \in [0, 1].$$

We show that displacement λ-convexity of $\mathcal{F}$ is guaranteed if a certain matrix is positive semidefinite. This result (slightly) generalizes Proposition 2.1 in [19].

Proposition 1. *The entropy $\mathcal{F}$, defined in* (7)*, is displacement λ-convex for some $\lambda \in \mathbb{R}$ if for any $\rho \in X_n$,*

$$M(\rho) \geq \lambda K(\rho), \tag{12}$$

i.e. if $M(\rho) - \lambda K(\rho)$ is positive semidefinite, where

$$M(\rho) = \frac{1}{2}\big(DK(\rho)[Q\phi(\rho)] - Q\Phi'(\rho)K(\rho) - K(\rho)\Phi'(\rho)Q^\top\big) \tag{13}$$

and $\Phi'(\rho) = \mathrm{diag}(\phi'(\rho_1), \ldots, \phi'(\rho_n))$.

Proof. Let ρ_0, $\rho_1 \in X_n$ and let $\rho : [0,1] \to X_n$ be a geodesic curve with $(\rho, \psi) \in E(\rho_0, \rho_1)$. Then (ρ, ψ) satisfies the geodesic equations (11), implying that

$$\frac{d}{dt}\mathcal{F}(\rho) = \langle D\mathcal{F}(\rho), \partial_t\rho\rangle = \langle D\mathcal{F}(\rho), K(\rho)\psi\rangle.$$

Differentiating a second time and using the symmetry of $K(\rho)$ and $DK(\rho)[\partial_t\rho]$, we find that

$$\begin{aligned}\frac{d^2}{dt^2}\mathcal{F}(\rho) &= \langle D^2\mathcal{F}(\rho)\partial_t\rho, K(\rho)\psi\rangle + \langle D\mathcal{F}(\rho), DK(\rho)[\partial_t\rho]\psi\rangle + \langle D\mathcal{F}(\rho), K(\rho)\partial_t\psi\rangle \\ &= \langle K(\rho)D^2\mathcal{F}(\rho)\partial_t\rho, \psi\rangle + \langle DK(\rho)[\partial_t\rho]D\mathcal{F}(\rho), \psi\rangle + \langle K(\rho)D\mathcal{F}(\rho), \partial_t\psi\rangle.\end{aligned}$$

Inserting the geodesic equations (11) yields

$$\begin{aligned}\frac{d^2}{dt^2}\mathcal{F}(\rho) &= \big\langle K(\rho)D^2\mathcal{F}(\rho)K(\rho)\psi + DK(\rho)[K(\rho)\psi]D\mathcal{F}(\rho), \psi\big\rangle \\ &\quad - \frac{1}{2}\big\langle DK(\rho)[K(\rho)D\mathcal{F}(\rho)]\psi, \psi\big\rangle.\end{aligned} \tag{14}$$

We differentiate $K(\rho)D\mathcal{F}(\rho) = -Q\phi(\rho)$ with respect to ρ:

$$K(\rho)D^2\mathcal{F}(\rho) + DK(\rho)[\,\cdot\,]D\mathcal{F}(\rho) = -Q\Phi'(\rho).$$

Thus, we can replace the first bracket on the right-hand side of (14) by $-Q\Phi'(\rho)K(\rho)\psi$:

$$\begin{aligned}\frac{d^2}{dt^2}\mathcal{F}(\rho) &= -\langle Q\Phi'(\rho)K(\rho)\psi, \psi\rangle + \frac{1}{2}\big\langle DK(\rho)[Q\phi(\rho)]\psi, \psi\big\rangle \\ &= \frac{1}{2}\big\langle\big(DK(\rho)[Q\phi(\rho)] - Q\Phi'(\rho)K(\rho) - K(\rho)\Phi'(\rho)Q^\top\big)\psi, \psi\big\rangle.\end{aligned} \tag{15}$$

We infer from (12) that

$$\frac{d^2}{dt^2}\mathcal{F}(\rho) \geq \lambda\langle K(\rho)\psi, \psi\rangle$$

for all geodesic curves ρ and vector fields ψ such that $(\rho, \psi) \in E(\rho_0, \rho_1)$. Consequently,

$$\frac{d^2}{dt^2}\mathcal{F}(\rho(t)) \geq \lambda\mathcal{W}(\rho_0, \rho_1)^2, \quad t \in [0,1],$$

and by Definition 1, $\mathcal{F}$ is displacement λ-convex. □

3. Semidiscrete heat equation

As a warm-up, we consider the semidiscrete heat equation

$$\partial_t \rho_i = h^{-2}(\rho_{i-1} - 2\rho_i + \rho_{i+1}), \quad i = 0, \dots, n, \ t > 0, \tag{16}$$

where $n \in \mathbb{N}$ and $h = 1/n > 0$. The no-flux boundary conditions are realized by setting $\rho_{-1} = \rho_0$ and $\rho_{n+1} = \rho_n$. We write $\rho = (\rho_0, \dots, \rho_n)$. Equation (16) can be written as (6) by setting $\phi(s) = s$ and $Q = -G^\top G$ with the discrete gradient $G \in \mathbb{R}^{n\times(n+1)}$, $G_{ij} = h^{-1}(\delta_{ij} - \delta_{i+1,j})$. By slightly abusing the notation, we set $w_i = 1$ for $i = 0, \dots, n$ and note that for a function $f : [0,\infty) \to \mathbb{R}$, the corresponding entropy given in (7) reduces to

$$\mathcal{F}(\rho) = \sum_{i=0}^{n} f(\rho_i). \tag{17}$$

Then, for the respective Onsager operator given in (8) with the mean function Λ^f, we claim that the entropy $\mathcal{F}$ is displacement convex, under suitable conditions on f.

Theorem 2. *Let f be such that Λ^f, defined in* (1)*, is concave in both variables. Then the entropy* (17) *is displacement convex with respect to the metric* (10) *induced by Λ^f.*

If $f(s) = s(\log s - 1)$ or $f(s) = s^\alpha$ for $1 < \alpha \le 2$, Λ is concave in both variables (see Lemma 7), thus fulfilling the assumption of the theorem.

Proof. We formulate $Q\rho = -G^\top G\rho = -G^\top L(\rho)Gf'(\rho)$, where $L(\rho) = \operatorname{diag}(\Lambda^f(\rho_i, \rho_{i+1}))_{i=0}^{n-1}$ and $f'(\rho) = (f'(\rho_i))_{i=0}^n$. Then, setting $K(\rho) = G^\top L(\rho)G$, we can write (16) as the gradient system

$$\partial_t \rho = Q\rho = -K(\rho)D\mathcal{F}(\rho),$$

where we identify $D\mathcal{F}(\rho)$ with $f'(\rho)$. Thus, by Proposition 1, it is sufficient to show that the matrix $M(\rho)$, defined in (13), is positive semidefinite. In fact, because of the special structure of $K(\rho)$, we can simplify this condition. Let $\psi \in \mathbb{R}^{n+1}$. Then, using $DK(\rho)[\,\cdot\,] = G^\top DL(\rho)[\,\cdot\,]G$ and $Q = -G^\top G$,

$$\begin{aligned}
\langle M(\rho)\psi, \psi\rangle &= \frac{1}{2}\Big\langle \big(DK(\rho)[Q\rho] - QK(\rho) - K(\rho)Q^\top\big)\psi, \psi\Big\rangle \\
&= \frac{1}{2}\Big\langle G^\top\big(DL(\rho)[Q\rho]G + GG^\top L(\rho)G + L(\rho)GG^\top G\big)\psi, \psi\Big\rangle \\
&= \frac{1}{2}\Big\langle \big(DL(\rho)[Q\rho] + GG^\top L(\rho) + L(\rho)GG^\top\big)G\psi, G\psi\Big\rangle.
\end{aligned}$$

Hence, it is sufficient to show that

$$\widetilde{M} := -DL(\rho)[G^\top G\rho] + GG^\top L(\rho) + L(\rho)GG^\top$$

is positive semidefinite.

We show this claim by verifying that $\widetilde{M}$ is diagonally dominant. To this end, we observe that $\widetilde{M}$ is a symmetric tridiagonal matrix with entries

$$\widetilde{M} = \frac{1}{h^2}\begin{pmatrix} a_0 & b_0 & 0 & \cdots & 0 \\ b_0 & a_1 & b_1 & \ddots & \vdots \\ 0 & b_1 & \ddots & & 0 \\ \vdots & \ddots & & a_{n-2} & b_{n-2} \\ 0 & \cdots & 0 & b_{n-2} & a_{n-1} \end{pmatrix},$$

where the coefficients are given by

$$\begin{aligned} a_i &= 4\Lambda^f(\rho_i,\rho_{i+1}) - \partial_1\Lambda^f(\rho_i,\rho_{i+1})(2\rho_i - \rho_{i-1} - \rho_{i+1}) \\ &\quad - \partial_2\Lambda^f(\rho_i,\rho_{i+1})(2\rho_{i+1} - \rho_i - \rho_{i+2}), \quad i = 0,\ldots,n-1 \\ b_i &= -\big(\Lambda^f(\rho_i,\rho_{i+1}) + \Lambda^f(\rho_{i+1},\rho_{i+2})\big) \le 0, \quad i = 0,\ldots,n-2. \end{aligned}$$

We also set $b_{-1} = -\Lambda^f(\rho_{-1},\rho_0) - \Lambda^f(\rho_0,\rho_1) \le 0$ and $b_{n-1} = -\Lambda^f(\rho_{n-1},\rho_n) - \Lambda^f(\rho_n,\rho_{n+1}) \le 0$, where by the non-flux boundary conditions, we have $\rho_{-1} = \rho_0$ and $\rho_{n+1} = \rho_n$.

The matrix $\widetilde{M}$ is diagonally dominant if

$$a_0 + b_0 \ge 0, \quad a_{n-1} + b_{n-2} \ge 0, \tag{18}$$

$$a_i + b_{i-1} + b_i \ge 0 \quad \text{for } i = 1,\ldots,n-2. \tag{19}$$

We will prove (19) for $i = 0,\ldots,n-1$. The first two conditions (18) follow from (19) for $i = -1$ and $i = n-1$, since $a_0 + b_0 = (a_0 + b_{-1} + b_0) - b_{-1} \ge a_0 + b_{-1} + b_0 \ge 0$ and $a_{n-1} + b_{n-2} = (a_{n-1} + b_{n-2} + b_{n-1}) - b_{n-1} \ge a_{n-1} + b_{n-2} + b_{n-1} \ge 0$. Thus, it remains to prove (19). We compute

$$\begin{aligned} a_i + b_{i-1} + b_i &= 2\Lambda^f(\rho_i,\rho_{i+1}) - \Lambda^f(\rho_{i+1},\rho_{i+2}) - \Lambda^f(\rho_{i-1},\rho_i) \\ &\quad - \partial_1\Lambda^f(\rho_i,\rho_{i+1})\big(2\rho_i - \rho_{i-1} - \rho_{i+1}\big) - \partial_2\Lambda^f(\rho_i,\rho_{i+1})\big(2\rho_{i+1} - \rho_i - \rho_{i+2}\big). \end{aligned}$$

Since Λ^f is assumed to be concave, we may apply Lemma 6, which shows that this expression is nonnegative, and hence, $\widetilde{M}$ is positive semidefinite. □

For nonlinear functions ϕ and nonconstant steady states (w_i), the proof of nonnegativity of $a_i + b_{i-1} + b_i$ is, unfortunately, not as simple as above, and we need more properties of the mean function. It turns out that the logarithmic mean satisfies these properties. Such a situation is considered in the next section.

4. Semidiscrete nonlinear Fokker-Planck equations

We discretize the nonlinear Fokker-Planck equation

$$\partial_t\rho = \partial_x(\partial_x\phi(\rho) + \phi(\rho)\partial_x V) = \partial_x\left(\phi(\rho)\partial_x \log\frac{\phi(\rho)}{w}\right),$$

where $w(x) = e^{-V(x)}$. We choose the quadratic potential $V(x) = \gamma|x|^2/2$ with $\gamma > 0$ but other choices are possible. Let $n \in \mathbb{N}$, $h = 1/n > 0$, and $x_i = ih$. Approximating $\rho(x_i, t)$ by $\rho_i(t)$, $w(x_i)$ by w_i and setting $u_i = \phi(\rho_i)/w_i$, the numerical scheme reads as

$$\partial_t \rho_i = h^{-2}\kappa_i(u_{i+1} - u_i) - h^{-2}\kappa_i(u_i - u_{i-1}), \tag{20}$$

where $\kappa_i = \sqrt{w_i w_{i+1}}$ approximates $w(x_{i+1/2})$. The no-flux boundary conditions are realized by $u_{-1} = u_0$ and $u_{n+1} = u_n$. Setting $Q = G^\top \operatorname{diag}(\kappa_i) G \operatorname{diag}(w_i^{-1})$ and, slightly abusing the notation, $\rho = (\rho_0, \ldots, \rho_n)$, we see that the scheme can be formulated as $\partial_t \rho = Q\phi(\rho)$, and thus, the framework of Section 2 applies. Hence, (20) can be written as the gradient system

$$\partial_t \rho = -K(\rho) \log u, \quad K(\rho) = G^\top L(\rho) G,$$

where $\log u = (\log u_i)_{i=0}^n$,

$$L(\rho) = \operatorname{diag}\left(\kappa_i \Lambda(u_i, u_{i+1})\right)_{i=0}^{n-1}, \quad u_i = \frac{\phi(\rho_i)}{w_i},$$

and Λ is the logarithmic mean. The above system can be written as in (9) by choosing $f(s) = s(\log s - 1)$, and therefore, by (7), the entropy reads as

$$\mathcal{F}(\rho) = \sum_{i=0}^{n} \left(f(\rho_i) + \frac{\gamma}{2} x_i^2 \rho_i \right),$$

since

$$f_i'(s) = f'\left(\frac{\phi(s)}{w_i}\right) = \log \phi(s) - \log w_i = f'(s) + \frac{\gamma}{2} x_i^2, \qquad i = 0, \ldots n.$$

Thus, $D\mathcal{F}(\rho) = \log u$. By Proposition 1, we know that the convexity of $\mathcal{F}$ is related to the matrix $M(\rho)$ defined in (13). Then, if $\mathcal{W}$ is the nonlocal transportation distance defined in (10), we have the following result.

Theorem 3. *Let ϕ be invertible, $\phi' \circ \phi^{-1}$ be nonincreasing, and $\gamma > 0$. Then, for each $\rho \in X_n$, there exist $\lambda_h(\rho) \in \mathbb{R}$ such that*

$$M(\rho) \geq \lambda_h(\rho) K(\rho),$$

where

$$\lambda_h(\rho) = \gamma \left(\frac{2}{\gamma h^2}(1 - e^{-\gamma h^2/2}) \min_{i=0,\ldots,n} \phi'(\rho_i) - 2\cosh(\gamma h) \max_{i=0,\ldots,n} |\nabla_h \phi'(\rho_i)| \right) \in \mathbb{R}.$$

For every $(\rho, \psi) \in E(\rho_0, \rho_1)$, the entropy $\mathcal{F}$ satisfies

$$\frac{d^2}{dt^2} \mathcal{F}(\rho(t)) \geq \lambda_h(\rho(t)) \mathcal{W}(\rho_0, \rho_1)^2, \quad t \in [0, 1].$$

If $\phi(s) = s$, we have $\lambda_h = (2/h^2)(1 - e^{-\gamma h^2/2}) \to \gamma$ as $h \to 0$.

The function $\phi(s) = s^\alpha$ satisfies the assumptions of the theorem if $0 < \alpha \leq 1$. In the linear case $\phi(s) = s$, we recover essentially the result of [19]. Increasing nonlinearities behaving like a power law $\phi(s) = s^\alpha$, $0 < \alpha \leq 1$, near zero and being linear at infinity also satisfy the assumptions of our theorem.

From numerical analysis and the expected large-time asymptotics of the equations involved, we expect that $\min_{i=0,\ldots,n}\phi'(\rho_i)$ and $\max_{i=0,\ldots,n}|\nabla_h\phi'(\rho_i)|$ are independent of h and bounded only by discrete norms of $\rho(0)$ under suitable assumptions on ϕ and V. However, we are unable to obtain these a-priori estimates for our scheme in this generality. In Appendix B, we provide such estimates for the case $V=0$ and very small initial data depending on the mesh size. These estimates only show that λ_h is positive if $\max_{i=0,\ldots,n}|\nabla_h\phi(\rho_i(0))|$ is sufficiently small with respect to h, more precisely, if it is smaller than $Ch^{1/2}$ with sufficiently small constant $C>0$. However, the estimates are not uniform in h.

Since $\lambda_h(\rho)$ depends on ρ, we need to be careful with the definition of displacement convexity. As explained in the previous paragraph, it is expected that $\lambda_h(\rho)$ can be bounded in terms of $\rho(0)=\rho_0$. Thus, if $|\rho_0|\le C$ for some constant $C>0$, $\lambda_h(\rho)$ does not depend on ρ and the standard notion of displacement convexity makes sense. Another issue arises since the space X_n is not complete. However, it is shown in [7] that a geodesically λ-convex gradient system on X_n can be extended to the completion $\overline{X}_n$, which is again a geodesically λ-convex gradient system. We refer to [19, Section 3.3] for a detailed discussion on this issue.

Proof. We will show that the matrix $M(\rho)-\lambda_h(\rho)K(\rho)$ is positive semidefinite. The derivative of $K(\rho)$ becomes $DK(\rho)[\,\cdot\,]=G^\top DL(\rho)[\,\cdot\,]G$ and

$$(DL(\rho)[\xi])_i=\kappa_i\partial_1\Lambda(u_i,u_{i+1})\frac{\phi'(\rho_i)}{w_i}\xi_i+\kappa_i\partial_2\Lambda(u_i,u_{i+1})\frac{\phi'(\rho_{i+1})}{w_{i+1}}\xi_{i+1}$$

for $i=0,\ldots,n-1$ and $\xi\in\mathbb{R}^{n+1}$. Therefore, for $\psi\in\mathbb{R}^{n+1}$,

$$\begin{aligned}\langle M(\rho)\psi,\psi\rangle&=\frac12\big\langle G^\top\big\{DL(\rho)[Q\phi(\rho)]G+Q\Phi'(\rho)G^\top L(\rho)G+G^\top L(\rho)G\Phi'(\rho)Q^\top\big\}\psi,\psi\big\rangle\\&=\frac12\langle\widetilde{M}G\psi,G\psi\rangle,\end{aligned}$$

where

$$\begin{aligned}\widetilde{M}&=DL(\rho)[Q\phi(\rho)]+\operatorname{diag}(\kappa_i)G\operatorname{diag}(w_i^{-1})\Phi'(\rho)G^\top L(\rho)\\&\quad+L(\rho)G\Phi'(\rho)\operatorname{diag}(w_i^{-1})G^\top\operatorname{diag}(\kappa_i).\end{aligned}$$

This matrix is symmetric and tridiagonal with entries

$$\widetilde{M}=\frac{1}{h^2}\begin{pmatrix}a_0 & b_0 & 0 & \cdots & 0\\ b_0 & a_1 & b_1 & \ddots & \vdots\\ 0 & b_1 & \ddots & & 0\\ \vdots & \ddots & & a_{n-2} & b_{n-2}\\ 0 & \cdots & 0 & b_{n-2} & a_{n-1}\end{pmatrix},$$

where the coefficients are given by

$$a_i=2\kappa_i^2\Lambda_i\bigg(\frac{\phi'(\rho_i)}{w_i}+\frac{\phi'(\rho_{i+1})}{w_i}\bigg)-\kappa_i\frac{\phi'(\rho_i)}{w_i}\partial_1\Lambda_i\big(\kappa_{i-1}(u_i-u_{i-1})+\kappa_i(u_i-u_{i+1})\big)$$

$$- \kappa_i \frac{\phi'(\rho_{i+1})}{w_{i+1}} \partial_2\Lambda_i\big(\kappa_i(u_{i+1}-u_i)+\kappa_{i+1}(u_{i+1}-u_{i+2})\big), \quad i=0,\ldots,n-1$$

$$b_i = -\kappa_i\kappa_{i+1}\frac{\phi'(\rho_{i+1})}{w_{i+1}}(\Lambda_i+\Lambda_{i+1}) \le 0, \quad i=0,\ldots,n-2$$

and we abbreviated

$$\Lambda_i := \Lambda(u_i,u_{i+1}), \quad \partial_j\Lambda_i := \partial_j\Lambda(u_i,u_{i+1}), \quad j=1,2.$$

Using the non-flux boundary conditions, we can also define b_{-1} and b^{n-1} by the same expression as above. We show now that $\widetilde{M}-\lambda_h L(\rho)$ is diagonally dominant for some $\lambda\in\mathbb{R}$. For this, we introduce further abbreviations:

$$\alpha_i = \kappa_i\frac{\phi'(\rho_i)}{w_i}, \quad \beta_i = \kappa_i\frac{\phi'(\rho_{i+1})}{w_{i+1}}.$$

Since $\kappa_i\alpha_{i+1}=\kappa_{i+1}\beta_i$, we compute

$$\begin{aligned}
a_i+b_{i-1}+b_i &= 2\kappa_i\Lambda_i(\alpha_i+\beta_i) - \kappa_i\beta_{i-1}(\Lambda_{i-1}+\Lambda_i) - \kappa_i\alpha_{i+1}(\Lambda_i+\Lambda_{i+1}) \\
&\quad - \kappa_i\alpha_i\partial_1\Lambda_i(u_i-u_{i+1}) - \kappa_i\beta_i\partial_2\Lambda_i(u_{i+1}-u_i) \\
&\quad - \kappa_{i-1}\alpha_i\partial_1\Lambda_i(u_i-u_{i-1}) - \kappa_{i+1}\beta_i\partial_2\Lambda_i(u_{i+1}-u_{i+2}) \\
&= \kappa_i\Lambda_i(2\alpha_i+2\beta_i-\beta_{i-1}-\alpha_{i+1}) \\
&\quad - \kappa_i\alpha_i\partial_1\Lambda_i(u_i-u_{i+1}) - \kappa_i\beta_i\partial_2\Lambda_i(u_{i+1}-u_i) \\
&\quad - \kappa_i\beta_{i-1}(\Lambda_{i-1}-\partial_1\Lambda_i u_{i-1}) - \kappa_i\alpha_{i+1}(\Lambda_{i+1}-\partial_2\Lambda_i u_{i+2}) \\
&\quad - \kappa_{i-1}\alpha_i\partial_1\Lambda_i u_i - \kappa_{i+1}\beta_i\partial_2\Lambda_i u_{i+1} \\
&= I_1+\cdots+I_7. \qquad (21)
\end{aligned}$$

We estimate these expressions term by term. Using property (ii) of Lemma 5, we find that

$$I_2 = -\kappa_i\alpha_i\Lambda_i + \kappa_i\alpha_i\frac{\Lambda_i^2}{u_i}, \quad I_3 = -\kappa_i\beta_i\Lambda_i + \kappa_i\beta_i\frac{\Lambda_i^2}{u_{i+1}}.$$

The first terms on the right-hand sides cancel with some terms in I_1. By property (iv) of Lemma 5, it follows that

$$\begin{aligned}
I_4 &\ge -\kappa_i\beta_{i-1}\max_{r\ge0}\big(\Lambda(r,u_i)-\partial_1\Lambda(u_i,u_{i+1})r\big) = -\kappa_i\beta_{i-1}u_i\partial_2\Lambda(u_i,u_{i+1}) \\
&= -\kappa_i\beta_{i-1}u_i\partial_2\Lambda_i, \\
I_5 &\ge -\kappa_i\alpha_{i+1}\max_{r\ge0}\big(\Lambda(u_{i+1},r)-\partial_2\Lambda(u_i,u_{i+1})r\big) \\
&= -\kappa_i\alpha_{i+1}\max_{r\ge0}(\Lambda(r,u_{i+1})-\partial_1\Lambda(u_{i+1},u_i)r) = -\kappa_i\alpha_{i+1}u_{i+1}\partial_2\Lambda(u_{i+1},u_i) \\
&= -\kappa_i\alpha_{i+1}u_{i+1}\partial_1\Lambda(u_i,u_{i+1}) = -\kappa_i\alpha_{i+1}u_{i+1}\partial_1\Lambda_i.
\end{aligned}$$

Finally, because of $\kappa_i\alpha_{i+1}=\kappa_{i+1}\beta_i$,

$$I_6 = -\kappa_i\beta_{i-1}\partial_1\Lambda_i u_i, \quad I_7 = -\kappa_i\alpha_{i+1}\partial_2\Lambda_i u_{i+1}.$$

Inserting these computations into (21), we arrive at

$$a_i + b_{i-1} + b_i \geq \kappa_i \Lambda_i(\alpha_i + \beta_i - \beta_{i-1} - \alpha_{i+1}) + \kappa_i \Lambda_i^2 \left(\frac{\alpha_i}{u_i} + \frac{\beta_i}{u_{i+1}} \right) - \kappa_i(\beta_{i-1} u_i + \alpha_{i+1} u_{i+1})(\partial_1 \Lambda_i + \partial_2 \Lambda_i).$$

Employing property (iii) of Lemma 5 in the last term, we obtain

$$\begin{aligned} a_i + b_{i-1} + b_i &\geq \kappa_i \Lambda_i(\alpha_i + \beta_i - \beta_{i-1} - \alpha_{i+1}) + \kappa_i \Lambda_i^2 \left(\frac{\alpha_i - \alpha_{i+1}}{u_i} + \frac{\beta_i - \beta_{i-1}}{u_{i+1}} \right) \\ &= J_1 + J_2. \end{aligned} \tag{22}$$

The idea is to replace $\kappa_{i\pm1}$ in β_{i-1} and α_{i+1} by an expression involving only κ_i. By definition of α_i and β_i and since

$$\begin{aligned} \frac{\kappa_{i+1}}{w_{i+1}} &= \frac{\kappa_i}{w_i} \frac{\kappa_{i+1}}{\kappa_i} \frac{w_i}{w_{i+1}} = \frac{\kappa_i}{w_i} \frac{\sqrt{w_i w_{i+2}}}{w_{i+1}} = \frac{\kappa_i}{w_i} e^{-\gamma h^2/2}, \\ \frac{\kappa_{i-1}}{w_i} &= \frac{\kappa_i}{w_{i+1}} \frac{\kappa_{i-1}}{\kappa_i} \frac{w_{i+1}}{w_i} = \frac{\kappa_i}{w_{i+1}} \frac{\sqrt{w_{i-1} w_{i+1}}}{w_i} = \frac{\kappa_i}{w_{i+1}} e^{-\gamma h^2/2}, \end{aligned}$$

we find that

$$\begin{aligned} J_1 &= \kappa_i \Lambda_i \left(\frac{\kappa_i}{w_i} \phi'(\rho_i) - \frac{\kappa_{i+1}}{w_{i+1}} \phi'(\rho_{i+1}) + \frac{\kappa_i}{w_{i+1}} \phi'(\rho_{i+1}) - \frac{\kappa_{i-1}}{w_i} \phi'(\rho_i) \right) \\ &= \frac{\kappa_i^2}{w_i} \Lambda_i \big(\phi'(\rho_i) - e^{-\gamma h^2/2} \phi'(\rho_{i+1}) \big) + \frac{\kappa_i^2}{w_{i+1}} \Lambda_i \big(\phi'(\rho_{i+1}) - e^{-\gamma h^2/2} \phi'(\rho_i) \big). \end{aligned}$$

In the same way, since

$$\kappa_{i+1} \frac{w_i}{w_{i+1}} = \kappa_i \frac{\sqrt{w_i w_{i+2}}}{w_{i+1}} = \kappa_i e^{-\gamma h^2/2}, \quad \kappa_{i-1} \frac{w_{i+1}}{w_i} = \kappa_i \frac{\sqrt{w_{i-1} w_{i+1}}}{w_i} = \kappa_i e^{-\gamma h^2/2},$$

we infer that

$$\begin{aligned} J_2 &= \kappa_i \Lambda_i^2 \left(\kappa_i \frac{\phi'(\rho_i)}{\phi(\rho_i)} - \kappa_{i+1} \frac{w_i}{w_{i+1}} \frac{\phi'(\rho_{i+1})}{\phi(\rho_i)} + \kappa_i \frac{\phi'(\rho_{i+1})}{\phi(\rho_{i+1})} - \kappa_{i-1} \frac{w_{i+1}}{w_i} \frac{\phi'(\rho_i)}{\phi(\rho_{i+1})} \right) \\ &= \kappa_i^2 \Lambda_i^2 \left(\frac{\phi'(\rho_i) - e^{-\gamma h^2/2} \phi'(\rho_{i+1})}{\phi(\rho_i)} + \frac{\phi'(\rho_{i+1}) - e^{-\gamma h^2/2} \phi'(\rho_i)}{\phi(\rho_{i+1})} \right). \end{aligned}$$

Thus, (22) becomes

$$\begin{aligned} a_i + b_{i-1} + b_i &\geq \kappa_i^2 \Lambda_i \left(\frac{\phi'(\rho_i) - e^{-\gamma h^2/2} \phi'(\rho_{i+1})}{w_i} + \frac{\phi'(\rho_{i+1}) - e^{-\gamma h^2/2} \phi'(\rho_i)}{w_{i+1}} \right) \\ &\quad + \kappa_i^2 \Lambda_i^2 \left(\frac{\phi'(\rho_i) - e^{-\gamma h^2/2} \phi'(\rho_{i+1})}{\phi(\rho_i)} + \frac{\phi'(\rho_{i+1}) - e^{-\gamma h^2/2} \phi'(\rho_i)}{\phi(\rho_{i+1})} \right) \\ &= \kappa_i^2 \Lambda_i \big(\phi'(\rho_i) - \phi'(\rho_{i+1}) \big) \left[\Lambda(u_i, u_{i+1}) \left(\frac{1}{\phi(\rho_i)} - \frac{1}{\phi(\rho_{i+1})} \right) + \frac{1}{w_i} - \frac{1}{w_{i+1}} \right] \end{aligned}$$

$$
\begin{aligned}
&+ \kappa_i^2\Lambda_i\big(1-e^{-\gamma h^2/2}\big)\bigg[\frac{\phi'(\rho_i)}{w_{i+1}}+\frac{\phi'(\rho_{i+1})}{w_i} \\
&+ \Lambda(u_i,u_{i+1})\bigg(\frac{\phi'(\rho_i)/w_{i+1}}{u_{i+1}}+\frac{\phi'(\rho_{i+1})/w_i}{u_i}\bigg)\bigg] \\
&= K_1+K_2. \qquad (23)
\end{aligned}
$$

First, we estimate K_2 using property (v) of Lemma 5:

$$
\begin{aligned}
K_2 &\geq 2\kappa_i^2\Lambda_i\big(1-e^{-\gamma h^2/2}\big)\bigg(\frac{\phi'(\rho_i)}{w_{i+1}}+\frac{\phi'(\rho_{i+1})}{w_i}+2\sqrt{\frac{\phi'(\rho_i)\phi'(\rho_{i+1})}{w_iw_{i+1}}}\bigg) \\
&\geq 2\kappa_i\Lambda_i\big(1-e^{-\gamma h^2/2}\big)\sqrt{\phi'(\rho_i)\phi'(\rho_{i+1})} \geq 2\kappa_i\Lambda_i\big(1-e^{-\gamma h^2/2}\big)\min_{i=0,\ldots,n}\phi'(\rho_i).
\end{aligned}
$$

Since $\phi'\circ\phi^{-1}$ is nonincreasing, we have

$$
\big(\phi'(\rho_i)-\phi'(\rho_{i+1})\big)\bigg(\frac{1}{\phi(\rho_i)}-\frac{1}{\phi(\rho_{i+1})}\bigg)\geq 0.
$$

Consequently, since $\Lambda(u_i,u_{i+1})\geq 0$ and $\sinh(s)\leq s\cosh(s)$ for $s\geq 0$,

$$
\begin{aligned}
K_1 &\geq \kappa_i^2\Lambda_i\big(\phi'(\rho_i)-\phi'(\rho_{i+1})\big)\bigg(\frac{1}{w_i}-\frac{1}{w_{i+1}}\bigg) \\
&= \kappa_i\Lambda_i\big(\phi'(\rho_i)-\phi'(\rho_{i+1})\big)\bigg(\sqrt{\frac{w_{i+1}}{w_i}}-\sqrt{\frac{w_i}{w_{i+1}}}\bigg) \\
&= -\kappa_i\Lambda_i\big(\phi'(\rho_i)-\phi'(\rho_{i+1})\big)\big(e^{\gamma(x_{i+1}^2-x_i^2)/4}-e^{-\gamma(x_{i+1}^2-x_i^2)/4}\big) \\
&\geq -2\kappa_i\Lambda_i h\max_{i=0,\ldots,n}|\nabla_h\phi'(\rho_i)|\sinh\bigg(\frac{\gamma}{4}(2i+1)h^2\bigg) \\
&\geq -2\kappa_i\Lambda_i h\max_{i=0,\ldots,n}|\nabla_h\phi'(\rho_i)|\bigg(\frac{\gamma}{4}(2i+1)h^2\bigg)\cosh\bigg(\frac{\gamma}{4}(2i+1)h^2\bigg) \\
&\geq -2\kappa_i\Lambda_i h^2\max_{i=0,\ldots,n}|\nabla_h\phi'(\rho_i)|\gamma\cosh(\gamma h),
\end{aligned}
$$

where we recall that $|\nabla_h\phi'(\rho_i)| := h^{-1}|\phi'(\rho_i)-\phi'(\rho_{i+1})|$ and we used $ih\leq 1$. Then (23) yields

$$
\begin{aligned}
h^{-2}(a_i+b_{i-1}+b_i) &\geq \gamma\kappa_i\Lambda_i\bigg(\frac{2}{\gamma h^2}(1-e^{-\gamma h^2/2})\min_{i=0,\ldots,n}\phi'(\rho_i) \\
&\qquad - 2\cosh(\gamma h)\max_{i=0,\ldots,n}|\nabla_h\phi'(\rho_i)|\bigg) \\
&= \lambda_h\kappa_i\Lambda_i.
\end{aligned}
$$

This proves that $\widetilde{M}-\lambda_h L(\rho)$ is positive semidefinite, finishing the proof. □

If the potential vanishes, we can define $w_i = 1$ for all $i = 0, \dots, n$. Then the entropy

$$\mathcal{F}(\rho) = \sum_{i=0}^{n} f(\rho_i) \quad \text{with } f'(s) = \log \phi(s)$$

is displacement convex with respect to $\mathcal{W}$. The following remark, based on an idea of [9], shows that this result may not hold for other entropies.

Remark 4. Erbar and Maas [9] considered the diffusion equation in the form

$$\partial_t \rho = \Delta\phi(\rho) = \operatorname{div}(\rho \nabla U'(\rho)),$$

where U satisfies $sU''(s) = \phi'(s)$. The corresponding numerical scheme becomes

$$\partial_t \rho = -K(\rho)U'(\rho), \quad K(\rho) = G^\top L(\rho) G,$$

where $U'(\rho) = (U'(\rho_0), \dots, U'(\rho_n))$ and the operator $L(\rho)$ is again defined by $L(\rho) = \operatorname{diag}(\Lambda(\rho_i, \rho_{i+1}))$, but with the mean function

$$\Lambda(\rho_i, \rho_{i+1}) = \frac{\phi(\rho_i) - \phi(\rho_{i+1})}{U'(\rho_i) - U'(\rho_{i+1})}. \tag{24}$$

The associated entropy is $\mathcal{F}(\rho) = \sum_{i=0}^n U(\rho_i)$, and if ρ is a geodesic curve on X_n with respect to the nonlinear transportation metric $\mathcal{W}$ induced by (24), then

$$\frac{d^2}{dt^2}\mathcal{F}(\rho) = \frac{1}{2}\langle \widetilde{M}(\rho) G\psi, G\psi \rangle$$

where $\widetilde{M} = DL(\rho)L(\rho)[Q\phi(\rho)] + G\Phi'(\rho)G^\top L(\rho) + L(\rho)G\Phi'(\rho)G^\top$. In fact, $\widetilde{M}$ is the tridiagonal matrix

$$\widetilde{M} = \frac{1}{h^2}\begin{pmatrix} d_0 & c_0 & 0 & \cdots & 0 \\ c_0 & d_1 & c_1 & \ddots & \vdots \\ 0 & c_1 & \ddots & & 0 \\ \vdots & \ddots & & d_{n-2} & c_{n-2} \\ 0 & \cdots & 0 & c_{n-2} & d_{n-1} \end{pmatrix},$$

with the matrix coefficients

$$\begin{aligned} d_i &= 2\Lambda(\rho_i, \rho_{i+1})\big(\phi'(\rho_i) + \phi'(\rho_{i+1})\big) + \partial_1\Lambda(\rho_i, \rho_{i+1})\big(\phi(\rho_{i-1}) - 2\phi(\rho_i) + \phi(\rho_{i+1})\big) \\ &\quad + \partial_2\Lambda(\rho_i, \rho_{i+1})\big(\phi(\rho_i) - 2\phi(\rho_{i+1}) + \phi(\rho_{i+2})\big), \quad i = 1, \dots, n-1, \\ c_i &= -\phi'(\rho_{i+1})\big(\Lambda(\rho_i, \rho_{i+1}) + \Lambda(\rho_{i+1}, \rho_{i+2})\big), \quad i = 1, \dots, n-2. \end{aligned}$$

If $\phi(s) = s^2$, we have $\Lambda(s,t) = (s+t)/2$ and the second principal minor equals

$$\begin{aligned} d_0 d_1 - c_0^2 &= \frac{1}{2}\rho_0^2\rho_1^2 + \frac{3}{2}\rho_0^2\rho_2^2 + 4\rho_0^2\rho_2\rho_3 + \frac{3}{2}\rho_0^2\rho_3^2 + \frac{1}{2}\rho_0^2\rho_4^2 + \rho_0\rho_1^3 + 3\rho_0\rho_1\rho_2^2 \\ &\quad + 8\rho_0\rho_1\rho_2\rho_3 + 3\rho_0\rho_1\rho_3^2 + \rho_0\rho_1\rho_4^2 + \frac{1}{4}\rho_1^4 + 2\rho_1^2\rho_2\rho_3 + \frac{3}{4}\rho_1^2\rho_3^2 + \frac{1}{4}\rho_1^2\rho_4^2 \\ &\quad - 4\rho_1\rho_2^3 - 2\rho_1\rho_2^2\rho_3 - \frac{13}{4}\rho_2^4 - 2\rho_2^3\rho_3 - \frac{1}{4}\rho_2^2\rho_3^2 + \frac{1}{4}\rho_2^2\rho_4^2. \end{aligned}$$

The coefficient 13/4 of the highest power in ρ_2 is negative and therefore, the second principal minor may be negative. According to Sylvester's criterion, $\widetilde{M}$ is not positive semidefinite. For instance, choosing special initial data, the entropy fails to be convex at time $t = 0$. □

Appendix A. Properties of mean functions

We need some properties of the mean function

$$\Lambda^f(s,t) = \frac{s-t}{f'(s)-f'(t)} \quad \text{for } s \neq t, \quad \Lambda^f(s,s) = \frac{1}{f''(s)}, \tag{25}$$

which we recall here. First, we are concerned with the logarithmic mean, i.e. $f'(s) = \log s$, for which we write simply Λ.

Lemma 5 (Properties of the logarithmic mean)**.** *For all s, $t > 0$, we have*

$$\begin{aligned}
&\text{(i)} && \Lambda(s,t) = \Lambda(t,s), \quad \partial_1\Lambda(s,t) = \partial_2\Lambda(t,s),\\
&\text{(ii)} && \partial_1\Lambda(s,t) = \frac{\Lambda(s,t)(s-\Lambda(s,t))}{s(s-t)}, \quad s \neq t,\\
&\text{(iii)} && \partial_1\Lambda(s,t) + \partial_2\Lambda(s,t) = \frac{\Lambda(s,t)^2}{st},\\
&\text{(iv)} && \max_{r\geq 0}\big(\Lambda(r,t) - \partial_1\Lambda(t,s)r\big) = t\partial_1\Lambda(s,t),\\
&\text{(v)} && \Lambda(s,t)\left(\frac{a}{s}+\frac{b}{t}\right) \geq 2\sqrt{ab} \quad \textit{for } a,b>0.
\end{aligned}$$

Proof. Properties (i)-(iii) can be easily verified by a calculation. Properties (iv)-(v) are shown in [19, Appendix A]. □

Lemma 6. *Let $\Lambda \in C^1([0,\infty)^2)$ be any function being concave in both variables, and let $u_0, u_1, u_2, u_3 \geq 0$. Then*

$$\begin{aligned}
&-\Lambda(u_0,u_1) + 2\Lambda(u_1,u_2) - \Lambda(u_2,u_3)\\
&\qquad \geq \partial_1\Lambda(u_1,u_2)(-u_0+2u_1-u_2) + \partial_2\Lambda(u_1,u_2)(-u_1+2u_2-u_3).
\end{aligned} \tag{26}$$

Proof. Since Λ is concave in both variables, we have

$$\begin{aligned}
\Lambda(u_0,u_1) - \Lambda(u_1,u_2) &\leq \partial_1\Lambda(u_1,u_2)(u_0-u_1) + \partial_2\Lambda(u_1,u_2)(u_1-u_2),\\
\Lambda(u_2,u_3) - \Lambda(u_1,u_2) &\leq \partial_1\Lambda(u_1,u_2)(u_2-u_1) + \partial_2\Lambda(u_1,u_2)(u_3-u_2),
\end{aligned}$$

and adding both inequalities gives the conclusion. □

Lemma 7 (Concavity of mean functions)**.** *Let $\Lambda^f : [0,\infty)^2 \to \mathbb{R}$ be given by* (25) *and let either $f(s) = s(\log s - 1)$ or $f(s) = s^\alpha$, where $1 < \alpha \leq 2$. Then Λ^f is concave in both variables.*

Proof. For $f(s) = s(\log s - 1)$, we refer to [8, Section 2]. The statement for $f(s) = s^\alpha$ is proved in [14, Appendix]. □

Appendix B. A priori estimates

Lemma 8 (A priori estimates). *Let ϕ be nondecreasing, $h > 0$ and let $\rho = (\rho_0, \dots, \rho_n) \in C^1([0,T^*];\mathbb{R}^{n+1})$ for some $T^* > 0$ be the solution to*

$$h^2\partial_t\rho_i = \phi(\rho_{i-1}) - 2\phi(\rho_i) + \phi(\rho_{i+1}), \quad i = 0,\dots,n, \tag{27}$$

where $\rho_{-1} = \rho_0$ and $\rho_{n+1} = \rho_n$. Then, for all $i = 0,\dots,n$ and $t > 0$,

$$\min_{i=0,\dots,n} \rho_i(0) \le \rho_i(t) \le \max_{i=0,\dots,n} \rho_i(0), \tag{28}$$

$$\max_{i=0,\dots,n} |\nabla_h\phi(\rho_i(t))| \le h^{-1/2}|\nabla_h\phi(\rho(0))|_2, \tag{29}$$

where $\nabla_h\phi(\rho_i(t)) = h^{-1}(\phi(\rho_{i+1}(t)) - \phi(\rho_i(t)))$ and

$$|\nabla_h\phi(\rho(0))|_2 := \left(\sum_{i=0}^n h\big|\nabla_h\phi(\rho_i(0))\big|^2\right)^{1/2}. \tag{30}$$

Proof. We multiply (27) by $(\rho_i - M)_+ = \max\{0, \rho_i - M\}$ and sum over $i = 0,\dots,n$:

$$\begin{aligned}
&\frac{h^2}{2}\partial_t\sum_{i=0}^n(\rho_i - M)_+^2 \\
&\quad= \sum_{i=0}^n \big(\phi(\rho_{i-1}) - \phi(\rho_i)\big)(\rho_i - M)_+ - \sum_{i=0}^n \big(\phi(\rho_i) - \phi(\rho_{i+1})\big)(\rho_i - M)_+ \\
&\quad= \sum_{j=0}^n \big(\phi(\rho_j) - \phi(\rho_{j+1})\big)(\rho_{j+1} - M)_+ - \sum_{i=0}^n \big(\phi(\rho_i) - \phi(\rho_{i+1})\big)(\rho_i - M)_+ \\
&\quad= -\sum_{i=0}^n \big(\phi(\rho_i) - \phi(\rho_{i+1})\big)\big((\rho_i - M)_+ - (\rho_{i+1} - M)_+\big) \le 0,
\end{aligned}$$

since ϕ is nondecreasing. This shows that

$$\sum_{i=0}^n(\rho_i(t) - M)_+^2 \le \sum_{i=0}^n(\rho_i(0) - M)_+^2.$$

Thus, if $M = \max_{i=0,\dots,n}\rho_i(0)$, the upper bound in (28) follows. The lower bound is proved analogously.

For the proof of (29), we compute

$$\begin{aligned}
\frac{h^2}{2}\partial_t\sum_{i=0}^n \big(\phi(\rho_{i+1}) - \phi(\rho_i)\big)^2 &= h^2\sum_{i=0}^n(\phi(\rho_{i+1}) - \phi(\rho_i))\big(\phi'(\rho_{i+1})\partial_t\rho_{i+1} - \phi'(\rho_i)\partial_t\rho_i\big) \\
&= \sum_{i=0}^n(\phi(\rho_{i+1}) - \phi(\rho_i))\phi'(\rho_{i+1})\big(\phi(\rho_i) - 2\phi(\rho_{i+1}) + \phi(\rho_{i+2})\big) \\
&\quad - \sum_{i=0}^n \big(\phi(\rho_{i+1}) - \phi(\rho_i)\big)\phi'(\rho_i)\big(\phi(\rho_{i-1}) - 2\phi(\rho_i) + \phi(\rho_{i+1})\big).
\end{aligned}$$

Making the change of variables $i \mapsto i-1$ in the first sum and rearranging the terms, we find that

$$\frac{h^2}{2}\partial_t \sum_{i=0}^{n} \big(\phi(\rho_{i+1}) - \phi(\rho_i)\big)^2 = -\sum_{i=0}^{n} \phi'(\rho_i)\big(\phi(\rho_{i-1}) - 2\phi(\rho_i) + \phi(\rho_{i+1})\big)^2 \le 0.$$

Consequently, for any $j = 0, \dots, n-1$ and $t > 0$,

$$\begin{aligned}\big(\phi(\rho_{j+1}(t)) - \phi(\rho_j(t))\big)^2 &\le \sum_{i=0}^{n} \big(\phi(\rho_{i+1}(t)) - \phi(\rho_i(t))\big)^2 \\ &\le \sum_{i=0}^{n} \big(\phi(\rho_{i+1}(0)) - \phi(\rho_i(0))\big)^2 = h|\nabla_h \phi(\rho(0))|_2^2.\end{aligned}$$

Taking the maximum over $j = 0, \dots, n-1$ shows (29). □

Corollary 9. *Let ϕ be nondecreasing and invertible, $h > 0$, and let $\rho = (\rho_0, \dots, \rho_n)$ be the solution to* (27). *We assume that $m := \min_{i=0,\dots,n} \rho_i(0) > 0$ and set $M := \max_{i=0,\dots,n} \rho_i(0)$. Then*

$$\max_{i=0,\dots,n} |\nabla_h \phi'(\rho_i)| \le h^{-1/2} \max_{s \in [\phi^{-1}(m), \phi^{-1}(M)]} \left|\frac{\phi''(s)}{\phi'(s)}\right| |\nabla_h \phi(\rho(0))|_2, \tag{31}$$

where $|\nabla_h \phi(\rho(0))|_2$ is defined in (30).

Proof. First, note that $m \le \rho_i(t) \le M$ for all $i = 0, \dots, n$ and $t > 0$, by Lemma 8. Then the result follows from the mean value theorem. Indeed, we have for some ξ between ρ_{i+1} and ρ_i,

$$\begin{aligned} h^{-1}\big|\phi'(\rho_{i+1}) - \phi'(\rho_i)\big| &= \frac{1}{h}\big|(\phi' \circ \phi^{-1})(\phi(\rho_{i+1})) - (\phi' \circ \phi^{-1})(\phi(\rho_i))\big| \\ &= \frac{1}{h}\left|\frac{\phi''(\phi^{-1}(\xi))}{\phi'(\phi^{-1}(\xi))}\right| \big|\phi(\rho_{i+1}) - \phi(\rho_i)\big| \\ &\le \frac{1}{h} \max_{s \in [\phi^{-1}(m), \phi^{-1}(M)]} \left|\frac{\phi''(s)}{\phi'(s)}\right| \max_{j=0,\dots,n} \big|\phi(\rho_{j+1}) - \phi(\rho_j)\big|, \end{aligned}$$

and we conclude after applying (29). □

Example 1. Let $\phi(s) = s^\alpha$ for $\alpha \in (0,1)$, $h > 0$ and let $\rho = (\rho_0, \dots, \rho_n)$ be the solution to (27) with $m := \min_{i=0,\dots,n} \rho_i(0) > 0$ and $M := \max_{i=0,\dots,n} \rho_i(0)$. We claim that

$$\min_{i=0,\dots,n} \phi'(\rho_i) \le \alpha M^{\alpha-1}, \quad \max_{i=0,\dots,n} |\nabla_h \phi'(\rho_i)| \le (1-\alpha) m^{-1/\alpha} h^{-1/2} |\nabla_h \phi(\rho(0))|_2,$$

where $|\nabla_h \phi(\rho(0))|_2$ is defined in (30). Indeed, the first statement follows from $\alpha < 1$ and (28):

$$\min_{i=0,\dots,n} \phi'(\rho_i) = \alpha\Big(\max_{i=0,\dots,n} \rho_i\Big)^{\alpha-1} \le \alpha M^{\alpha-1},$$

and the second statement is a consequence of Corollary 9 evaluating the right-hand side of (31).

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures.* Lectures in Mathematics, Birkhäuser, Basel, 2005.

[2] A.-I. Bonciocat and K.-T. Sturm. Mass transportation and rough curvature bounds for discrete spaces. *J. Funct. Anal.* 256 (2009), 2944-2966.

[3] P. Caputo, P. Dai Pra, and G. Posta. Convex entropy decay via the Bochner-Bakry-Emery approach. *Ann. Inst. H. Poincaré Prob. Stat.* 45 (2009), 734-753.

[4] J. A. Carrillo, R. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Rational Mech. Anal.* 179 (2006), 217-263.

[5] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepčev. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.* 258 (2010), 1273-1309.

[6] S. Chow, W. Huang, Y. Li, and H. Zhou. Fokker-Planck equations for a free energy functional or Markov process on a graph. *Arch. Rational Mech. Anal.* 203 (2012), 969-1008.

[7] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in theWasserstein distance. *SIAM J. Math. Anal.* 40 (2008), 1104-1122.

[8] M. Erbar and J. Maas. Ricci curvature bounds for finite Markov chains via convexity of the entropy. *Arch. Rational Mech. Anal.* 206 (2012), 997-1038.

[9] M. Erbar and J. Maas. Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Sys.* 34 (2014), 1355-1374.

[10] M. Fathi and J. Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Prob.* 26 (2016), 1774-1806.

[11] N. Gozlan, C. Roberto, Cyril, P.-M. Samson, and P. Tetali. Displacement convexity of entropy and related inequalities on graphs. *Prob. Theory Related Fields* 160 (2014), 47-94.

[12] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* 29 (1998), 1-17.

[13] A. Jüngel. *Entropy Methods for Diffusive Partial Differential Equations.* BCAM Springer Briefs, Springer, Cham, 2016.

[14] A. Jüngel and W. Yue. Discrete Beckner inequalities via the Bochner-Bakry-Emery approach for Markov chains. To appear in *Ann. Appl. Prob.*, 2017. arXiv:1511.06250.

[15] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math.* 169 (2009), 903-991.

[16] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* 261 (2011), 2250-2292.

[17] J. Maas and D. Matthes. Long-time behavior of a finite volume discretization for a fourth order diffusion equation. *Nonlinearity* 29 (2016), 1992-2023.

[18] R. McCann. A convexity principle for interacting gases. *Adv. Math.* 128 (1997), 153-179.

[19] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Diff. Eqs.* 48 (2013), 1-31.

[20] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Diff. Eqs.* 26 (2001), 101-174.

[21] M.-K. von Renesse and K.-Th. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Commun. Pure Appl. Math.* 58 (2005), 923-940.

[22] M. Sammer and P. Tetali. Concentration on the discrete torus using transportation. *Combin. Prob. Comput.* 18 (2009), 835-860.

[23] C. Villani. *Optimal Transport. Old and New.* Springer, Berlin, 2009.

J.A.C.: Department of Mathematics, Imperial College London, London SW7 2AZ, United Kingdom
E-mail address: carrillo@imperial.ac.uk

A.J.: Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8-10, 1040 Wien, Austria
E-mail address: juengel@tuwien.ac.at

M.C.S.: Departamento de Matemática – IMECC, Universidade Estadual de Campinas, 13083-859, Campinas-SP, Brazil
E-mail address: matheus.santos@ufrgs.com