Euro. Jnl of Applied Mathematics (2019), *vol.* 30, *pp.* 1103–1122 © Cambridge University Press 2018. 1103

doi:10.1017/S0956792517000389

Displacement convexity for the entropy in semi-discrete non-linear Fokker–Planck equations†

JOSÉ A. CARRILLO[1], ANSGAR JÜNGEL[2] and MATHEUS C. SANTOS[3]

[1]*Department of Mathematics, Imperial College London, London SW7 2AZ, UK*
email: carrillo@imperial.ac.uk
[2]*Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstraße 8-10, 1040 Wien, Austria*
email: juengel@tuwien.ac.at
[3]*Departamento de Matemática – IMECC, Universidade Estadual de Campinas, Campinas-SP, Brazil*
email: matheus.santos@ufrgs.com

(*Received 28 May 2017; revised 13 December 2017; accepted 14 December 2017; first published online 10 January 2018*)

The displacement λ-convexity of a non-standard entropy with respect to a non-local 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 non-linear Fokker–Planck equation. The key idea is to employ a new mean function, which defines the Onsager operator in the gradient flow formulation.

Key words: Entropy, displacement convexity, logarithmic mean, finite differences, fast-diffusion equation

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

1 Introduction

Displacement convexity, which was introduced by McCann [20], 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, 6, 22]. It may also provide quantitative contraction estimates between solutions of the gradient flows [5] 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., $\mathrm{Ric}_M \geqslant \lambda$ [17,23]. Furthermore, it leads to inequalities in convex geometry and probability theory, such as the Brunn–Minkowski, Talagrand and log-Sobolev inequalities [25].

† 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.

We are interested in the question to what extent the concept of displacement convexity can be extended to discrete settings, like numerical discretisation 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 non-linear 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 [3]. Markov processes and Fokker–Planck equations on finite graphs were investigated by Chow *et al.* in [7]. Maas [18] and Mielke [21] introduced non-local 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.* [12] to derive HWI inequality (which interpolates the relative entropy H, the Wasserstein distance W and the Fisher information I) and log-Sobolev inequalities on (complete) graphs. Talagrand's inequality was studied in discrete spaces by Sammer and Tetali [24].

Only few results can be found in the literature on convexity properties of functionals for discretisations of partial differential equations. Exponential decay rates for time-continuous Markov chains were derived by Caputo *et al.* [4]. 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 [21] and applied to discretisations of one-dimensional linear Fokker-Planck equations (also see the presentation in [14, Section 5.2]). While the proof of Caputo *et al.* [4] is based on the Bochner–Bakry–Emery method, Mielke [21] employed a gradient flow approach together with matrix estimates. The non-local 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 [21] and summarized in Lemma 7 below). The same mean function has been used for finite volume discretisations of drift-diffusion equations [2, equation (28)]. The approach of [4] (and [11]) was extended to general convex entropy densities $f(s)$ in [15] using the mean function:

$$\Lambda^f(s,t) = \frac{s-t}{f'(s) - f'(t)} \quad \text{for } s \neq t, \quad \Lambda^f(s,s) = \frac{1}{f''(s)}, \tag{1.1}$$

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

Concerning non-linear equations, we are aware only of two results. Erbar and Maas [10] 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 non-local 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 [10]. 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 [19] to prove exponential decay rates for finite-volume discretisations 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 non-linear function (coming from the diffusivity), which is suitable for finite-difference discretisations of the non-linear Fokker–Planck equation

$$\partial_t \rho = \partial_x\big(\partial_x \phi(\rho) + \phi(\rho)\partial_x V\big), \quad x \in (0,1),\ t > 0, \tag{1.2}$$

supplemented with no-flux boundary conditions and an initial condition. Equations with non-linear mobility were already treated under the point of view of optimal transport in [16], and the displacement convexity of equations related to (1.2) was analysed in [6]. Here, $\phi : [0,\infty) \to [0,\infty)$ is a continuous function, e.g., $\phi(\rho) = \rho^\alpha$ with $\alpha > 0$, and $V(x)$ is a quadratic confinement potential $V(x) = \gamma|x|^2/2$ with $\gamma \geqslant 0$. 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 non-increasing along (smooth) solutions to (1.2). Our aim is to analyse the displacement convexity of a discrete version of the entropy $\mathcal{F}_c$ along semi-discrete solutions associated to (1.2).

For the discretisation of (1.2), let $n \in \mathbb{N}$, $h = 1/n > 0$, and $x_i = ih$, $i = 0,\dots,n$. 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 (1.2) 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

$$\partial_t \rho_i = \frac{\kappa_i \Lambda_i}{h^2}\left(\log\frac{\phi(\rho_{i+1})}{w_{i+1}} - \log\frac{\phi(\rho_i)}{w_i}\right) - \frac{\kappa_{i-1}\Lambda_{i-1}}{h^2}\left(\log\frac{\phi(\rho_i)}{w_i} - \log\frac{\phi(\rho_{i-1})}{w_{i-1}}\right), \tag{1.3}$$

where $h > 0$ is the space size and $\kappa_i\Lambda_i$ is an approximation of $\phi(\rho)$ in $[x_i, x_{i+1}]$. To simplify the notation, let us write $u_i = \phi(\rho_i)/w_i$, for $i = 0,\dots,n$. Our idea is to employ the *modified* logarithmic mean

$$\Lambda_i = \frac{u_i - u_{i+1}}{\log u_i - \log u_{i+1}}, \tag{1.4}$$

and to set, as in [21], $\kappa_i = \sqrt{w_i w_{i+1}}$. Since Λ_i approximates u_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

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

which approximates (1.2) 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 non-increasing (an example is $\phi(s) = s^\alpha$ with $0 < \alpha \leqslant 1$), then the discrete entropy

$$\mathcal{F}(\rho) = \sum_{i=0}^{n} \left(f(\rho_i) + \gamma \frac{x_i^2}{2} \rho_i \right), \quad \text{where } f'(s) = \log \phi(s), \tag{1.5}$$

is displacement λ_h-convex with respect to the non-local transportation metric induced by (1.4), where

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

and $\nabla_h \phi'(\rho_i) = h^{-1}(\phi'(\rho_{i+1}) - \phi'(\rho_i))$; see Theorem 3. Notice that γ is the convexity constant of the quadratic potential $V(x)$ and that $\lambda_h = 0$ for $\gamma = 0$.

Our result is consistent with that one in [21]: If $\phi(s) = s$ is linear (and $V \neq 0$), $\lambda_h \to \gamma$ as $h \to 0$, and the constant is asymptotically sharp. When $\gamma = 0$, our main result shows that $\mathcal{F}(\rho)$ is displacement convex.

For $\gamma > 0$, if the minimum of $\phi'(\rho_i)$ is positive and the maximum of $|\nabla_h \phi'(\rho_i)|$ is sufficiently small, then $\lambda_h > 0$. 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 for $V = 0$ and for small initial data depending on the mesh size h; see Corollary 1. This is an indication that such *a priori* estimates might hold true for $\gamma > 0$.

The main motivation of this work is to find a family of non-linear drift-diffusion equations for which the strategy developed in [21] can be carried over. As a byproduct, we obtain semi-discrete finite difference schemes for equation (1.2) enjoying mass conservation, positivity preservation and the natural entropy dissipation property, i.e., $\mathcal{F}(\rho)$ non-increasing.

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 semi-definite, slightly generalising Proposition 2.1 in [21]. As a warm-up, we consider in Section 3 the semi-discrete 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 $1 < \alpha \leqslant 2$; see Theorem 2. This result is a reformulation of Theorem 5 in [15], but our proof is very simple. Section 4 is concerned with the proof of displacement λ-convexity of (1.5) and contains our main result. Some properties of mean functions are recalled in Appendix A, and *a priori* estimates of solutions to (1.3) with $V = 0$ and small initial 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:

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

The closure of X_n can be identified with the space of probability measures on a $(n+1)$-point set. We restrict our attention to X_n in order to have the entropy dissipation term $D\mathcal{F}(\rho)$ well defined independently of the assumptions on ϕ at the origin. 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

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

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

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

is satisfied. Summing this condition for fixed i over $j=0,\dots,n$, 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 the system of ordinary differential equations of the type

$$\partial_t\rho=Q\phi(\rho), \quad t>0, \tag{2.1}$$

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 $\mathcal{F}:X_n\to\mathbb{R}$,

$$\mathcal{F}(\rho)=\sum_{i=0}^{n} f_i(\rho_i), \quad \text{where } f_i'(s)=f'\left(\frac{\phi(s)}{w_i}\right), \tag{2.2}$$

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

$$K(\rho)=\frac{1}{2}\sum_{i,j=0}^{n} Q_{ij}w_j\Lambda^f\left(\frac{\phi(\rho_i)}{w_i},\frac{\phi(\rho_j)}{w_j}\right)(e_i-e_j)\otimes(e_i-e_j), \tag{2.3}$$

where $e_i=(\delta_{i0},\dots,\delta_{in})^\top\in\mathbb{R}^{n+1}$ is the ith unit vector and '$\otimes$' is the tensor product. By detailed balance and $Q_{ij}w_j\geqslant 0$ for $i\neq j$, it follows that $K(\rho)$ is symmetric and positive semi-definite. With these definitions, we can formulate (2.1) as a gradient system in the sense that it can be rewritten as

$$\partial_t\rho=-K(\rho)D\mathcal{F}(\rho), \tag{2.4}$$

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

The space X_n is endowed with the non-local transportation distance

$$\mathcal{W}(\rho_0,\rho_1)^2=\inf_{(\rho,\psi)\in E(\rho_0,\rho_1)}\int_0^1\langle K(\rho(t))\psi(t),\psi(t)\rangle dt, \tag{2.5}$$

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

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

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) [18, Theorem 1.1] and the pair $(\overline{X}_n, \mathcal{W})$ defines a geodesic space [10, Proposition 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 [18, 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 (2.5), then ρ is a geodesic and satisfies the geodesic equations [10, Proposition 2.5]

$$\begin{cases} \partial_t \rho = K(\rho)\psi \\ \partial_t \psi = -\frac{1}{2}\langle DK(\rho)[\,\cdot\,]\psi, \psi\rangle \end{cases}, \quad t > 0, \tag{2.6}$$

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 $\rho : [0,1] \to X_n$,*

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

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

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

We show that displacement λ-convexity of $\mathcal{F}$ is guaranteed if a certain matrix is positive semi-definite. This result is an analog of Proposition 2.1 in [21].

Proposition 1 *The entropy $\mathcal{F}$, defined in (2.2), is displacement λ-convex for some $\lambda \in \mathbb{R}$ if and only if, for any $\rho \in X_n$,*

$$M(\rho) \geqslant \lambda K(\rho), \tag{2.7}$$

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

$$M(\rho) = \frac{1}{2}\left(DK(\rho)[Q\phi(\rho)] - Q\Phi'(\rho)K(\rho) - K(\rho)\Phi'(\rho)Q^\top\right) \tag{2.8}$$

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 (2.6), 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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

that

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

Inserting the geodesic equation (2.6) yields

$$\frac{d^2}{dt^2}\mathcal{F}(\rho) = \langle K(\rho)D^2\mathcal{F}(\rho)K(\rho)\psi + DK(\rho)[K(\rho)\psi]D\mathcal{F}(\rho), \psi\rangle$$
$$- \frac{1}{2}\langle DK(\rho)[K(\rho)D\mathcal{F}(\rho)]\psi, \psi\rangle. \tag{2.9}$$

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 (2.9) by $-Q\Phi'(\rho)K(\rho)\psi$:

$$\frac{d^2}{dt^2}\mathcal{F}(\rho) = -\langle Q\Phi'(\rho)K(\rho)\psi, \psi\rangle + \frac{1}{2}\langle DK(\rho)[Q\phi(\rho)]\psi, \psi\rangle$$
$$= \frac{1}{2}\langle\big(DK(\rho)[Q\phi(\rho)] - Q\Phi'(\rho)K(\rho) - K(\rho)\Phi'(\rho)Q^\top\big)\psi, \psi\rangle. \tag{2.10}$$

We infer from (2.7) that

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

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

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

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

For the *only if* part, given $\rho \in X_n$ and $\psi \in \mathbb{R}^n$, we construct a geodesic starting at ρ with initial field ψ and the result follows from (2.10). □

3 Semi-discrete heat equation

As a warm-up, we consider the semi-discrete heat equation

$$\partial_t\rho_i = h^{-2}(\rho_{i-1} - 2\rho_i + \rho_{i+1}), \quad i = 0, \ldots, n, \ t > 0, \tag{3.1}$$

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 $\rho = (\rho_0, \ldots, \rho_n)$. Equation (3.1) can be written as (2.1) 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 $i = 0, \ldots, n$ and note that for a function $f : [0, \infty) \to \mathbb{R}$, the corresponding entropy given in (2.2) reduces to

$$\mathcal{F}(\rho) = \sum_{i=0}^{n} f(\rho_i). \tag{3.2}$$

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

Then, for the respective Onsager operator given in (2.3) 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* (1.1)*, is concave in both variables. Then the entropy* (3.2) *is displacement convex with respect to the metric* (2.5) *induced by Λ^f.*

If $f(s) = s(\log s - 1)$ or $f(s) = s^\alpha$ for $1 < \alpha \leqslant 2$, Λ is concave in both variables (see Lemma 9), 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 (3.1) 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 (2.8), is positive semi-definite. 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 semi-definite.

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

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

We also set $b_{-1} = -\Lambda^f(\rho_{-1}, \rho_0) - \Lambda^f(\rho_0, \rho_1) \leqslant 0$ and $b_{n-1} = -\Lambda^f(\rho_{n-1}, \rho_n) - \Lambda^f(\rho_n, \rho_{n+1}) \leqslant 0$, 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

$$a_0 + b_0 \geqslant 0, \quad a_{n-1} + b_{n-2} \geqslant 0, \tag{3.3}$$

$$a_i + b_{i-1} + b_i \geqslant 0 \quad \text{for } i = 1, \ldots, n-2. \tag{3.4}$$

We will prove (3.4) for $i = 0, \ldots, n-1$. The first two conditions (3.3) follow from (3.4) for $i = -1$ and $i = n-1$, since $a_0 + b_0 = (a_0 + b_{-1} + b_0) - b_{-1} \geqslant a_0 + b_{-1} + b_0 \geqslant 0$ and $a_{n-1} + b_{n-2} = (a_{n-1} + b_{n-2} + b_{n-1}) - b_{n-1} \geqslant a_{n-1} + b_{n-2} + b_{n-1} \geqslant 0$. Thus, it remains to prove (3.4). 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 8, which shows that this expression is non-negative, and hence, $\widetilde{M}$ is positive semi-definite. □

For non-linear functions ϕ and non-constant steady states (w_i), the proof of non-negativity 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 Semi-discrete non-linear Fokker–Planck equations

We discretise the non-linear 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 $w(x) = e^{-V(x)}$ for $V(x) = \gamma|x|^2/2$ with $\gamma \geqslant 0$. 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

$$\partial_t \rho_i = h^{-2}\kappa_i(u_{i+1} - u_i) - h^{-2}\kappa_i(u_i - u_{i-1}), \tag{4.1}$$

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, (4.1) can be written as the gradient system

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

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

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

and Λ is the logarithmic mean. The above system can be written as in (2.4) by choosing $f(s) = s(\log s - 1)$, and therefore, by (2.2), 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 (2.8). Then, if $\mathcal{W}$ is the non-local transportation distance defined in (2.5), we have the following results.

Theorem 3 *Let ϕ be invertible, $\phi' \circ \phi^{-1}$ be non-increasing, and $\gamma \geqslant 0$.*

(1) *If $\gamma = 0$ then $M(\rho) \geqslant 0$, for all $\rho \in X_n$;*

(2) *If $\gamma > 0$ then for each $\rho \in X_n$, there exist $\lambda_h(\rho) \in \mathbb{R}$ such that*

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

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

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

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 $0 < \alpha \leqslant 1$. In the linear case $\phi(s) = s$, we recover the result of [21]. Increasing non-linearities behaving like a power law $\phi(s) = s^\alpha$, $0 < \alpha \leqslant 1$, near zero and being linear at infinity also satisfy the assumptions of our theorem.

Proof First, note that for any $\gamma \geqslant 0$, we can calculate the derivative of $K(\rho)$ and obtain $DK(\rho)[\,\cdot\,] = G^\top DL(\rho)[\,\cdot\,]G$, where

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

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

$$\begin{aligned}
\langle M(\rho)\psi, \psi \rangle &= \frac{1}{2} \langle \{ G^\top DL(\rho)[Q\phi(\rho)]G + Q\Phi'(\rho)G^\top L(\rho)G + G^\top L(\rho)G\Phi'(\rho)Q^\top \} \psi, \psi \rangle \\
&= \frac{1}{2} \langle \widetilde{M} G\psi, G\psi \rangle,
\end{aligned}$$

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

where

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

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

$$\begin{aligned} a_i &= 2\kappa_i^2\Lambda_i\left(\frac{\phi'(\rho_i)}{w_i} + \frac{\phi'(\rho_{i+1})}{w_{i+1}}\right) - \kappa_i\frac{\phi'(\rho_i)}{w_i}\partial_1\Lambda_i\big(\kappa_{i-1}(u_i - u_{i-1}) + \kappa_i(u_i - u_{i+1})\big) \\ &\quad - \kappa_i\frac{\phi'(\rho_{i+1})}{w_{i+1}}\partial_2\Lambda_i\big(\kappa_i(u_{i+1} - u_i) + \kappa_{i+1}(u_{i+1} - u_{i+2})\big), \quad i = 0, \dots, n-1 \\ b_i &= -\kappa_i\kappa_{i+1}\frac{\phi'(\rho_{i+1})}{w_{i+1}}(\Lambda_i + \Lambda_{i+1}) \leqslant 0, \quad i = 0, \dots, n-2 \end{aligned}$$

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

$$\begin{aligned} a_i + b_{i-1} + b_i &= 2\kappa_i\Lambda_i(\alpha_i + \beta_i) - \kappa_i\beta_{i-1}(\Lambda_{i-1} + \Lambda_i) - \kappa_i\alpha_{i+1}(\Lambda_i + \Lambda_{i+1}) \\ &\quad - \kappa_i\alpha_i\partial_1\Lambda_i(u_i - u_{i+1}) - \kappa_i\beta_i\partial_2\Lambda_i(u_{i+1} - u_i) \\ &\quad - \kappa_{i-1}\alpha_i\partial_1\Lambda_i(u_i - u_{i-1}) - \kappa_{i+1}\beta_i\partial_2\Lambda_i(u_{i+1} - u_{i+2}) \\ &= \kappa_i\Lambda_i(2\alpha_i + 2\beta_i - \beta_{i-1} - \alpha_{i+1}) \\ &\quad - \kappa_i\alpha_i\partial_1\Lambda_i(u_i - u_{i+1}) - \kappa_i\beta_i\partial_2\Lambda_i(u_{i+1} - u_i) \\ &\quad - \kappa_i\beta_{i-1}(\Lambda_{i-1} - \partial_1\Lambda_i u_{i-1}) - \kappa_i\alpha_{i+1}(\Lambda_{i+1} - \partial_2\Lambda_i u_{i+2}) \\ &\quad - \kappa_{i-1}\alpha_i\partial_1\Lambda_i u_i - \kappa_{i+1}\beta_i\partial_2\Lambda_i u_{i+1} \\ &= I_1 + \cdots + I_7. \end{aligned} \tag{4.2}$$

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

We estimate these expressions term by term. Using property (ii) of Lemma 7, 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 7, it follows that

$$\begin{aligned}
I_4 &\geqslant -\kappa_i\beta_{i-1}\max_{r\geqslant 0}\big(\Lambda(r,u_i) - \partial_1\Lambda(u_i,u_{i+1})r\big) = -\kappa_i\beta_{i-1}u_i\partial_2\Lambda(u_i,u_{i+1}) \\
&= -\kappa_i\beta_{i-1}u_i\partial_2\Lambda_i, \\
I_5 &\geqslant -\kappa_i\alpha_{i+1}\max_{r\geqslant 0}\big(\Lambda(u_{i+1},r) - \partial_2\Lambda(u_i,u_{i+1})r\big) \\
&= -\kappa_i\alpha_{i+1}\max_{r\geqslant 0}(\Lambda(r,u_{i+1}) - \partial_1\Lambda(u_{i+1},u_i)r) = -\kappa_i\alpha_{i+1}u_{i+1}\partial_2\Lambda(u_{i+1},u_i) \\
&= -\kappa_i\alpha_{i+1}u_{i+1}\partial_1\Lambda(u_i,u_{i+1}) = -\kappa_i\alpha_{i+1}u_{i+1}\partial_1\Lambda_i.
\end{aligned}$$

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 (4.2), we arrive at

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

Employing property (iii) of Lemma 7 in the last term, we obtain

$$\begin{aligned}
a_i + b_{i-1} + b_i &\geqslant \kappa_i\Lambda_i(\alpha_i + \beta_i - \beta_{i-1} - \alpha_{i+1}) + \kappa_i\Lambda_i^2\left(\frac{\alpha_i - \alpha_{i+1}}{u_i} + \frac{\beta_i - \beta_{i-1}}{u_{i+1}}\right) \\
&= J_1 + J_2.
\end{aligned} \tag{4.3}$$

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

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

we find that

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

In the same way, since

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

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, (4.3) becomes

$$\begin{aligned} a_i + b_{i-1} + b_i &\geqslant \kappa_i^2\Lambda_i\left(\frac{\phi'(\rho_i) - e^{-\gamma h^2/2}\phi'(\rho_{i+1})}{w_i} + \frac{\phi'(\rho_{i+1}) - e^{-\gamma h^2/2}\phi'(\rho_i)}{w_{i+1}}\right) \\ &\quad + \kappa_i^2\Lambda_i^2\left(\frac{\phi'(\rho_i) - e^{-\gamma h^2/2}\phi'(\rho_{i+1})}{\phi(\rho_i)} + \frac{\phi'(\rho_{i+1}) - e^{-\gamma h^2/2}\phi'(\rho_i)}{\phi(\rho_{i+1})}\right) \\ &= \kappa_i^2\Lambda_i\big(\phi'(\rho_i) - \phi'(\rho_{i+1})\big)\left[\Lambda(u_i,u_{i+1})\left(\frac{1}{\phi(\rho_i)} - \frac{1}{\phi(\rho_{i+1})}\right) + \frac{1}{w_i} - \frac{1}{w_{i+1}}\right] \quad (4.4) \\ &\quad + \kappa_i^2\Lambda_i\big(1 - e^{-\gamma h^2/2}\big)\left[\frac{\phi'(\rho_i)}{w_{i+1}} + \frac{\phi'(\rho_{i+1})}{w_i}\right. \\ &\quad \left. + \Lambda(u_i,u_{i+1})\left(\frac{\phi'(\rho_i)/w_{i+1}}{u_{i+1}} + \frac{\phi'(\rho_{i+1})/w_i}{u_i}\right)\right] \\ &= K_1 + K_2. \quad (4.5) \end{aligned}$$

At this point, let us assume that $\gamma = 0$. Then, $K_2 = 0$. For K_1, note that in this case $w_i = 1$ for all $i = 0, \ldots, n$ and

$$\big(\phi'(\rho_i) - \phi'(\rho_{i+1})\big)\left(\frac{1}{\phi(\rho_i)} - \frac{1}{\phi(\rho_{i+1})}\right) \geqslant 0. \quad (4.6)$$

since $\phi' \circ \phi^{-1}$ is non-increasing. This implies $K_1 \geqslant 0$ and therefore, $M(\rho)$ is diagonally dominant.

Now, let us assume that $\gamma > 0$. We estimate K_2 using property (v) of Lemma 7:

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

By (4.6), since $\Lambda(u_i, u_{i+1}) \geqslant 0$ and $\sinh(s) \leqslant s\cosh(s)$ for $s \geqslant 0$,

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

where we recall that $|\nabla_h \phi'(\rho_i)| := h^{-1}|\phi'(\rho_i) - \phi'(\rho_{i+1})|$ and we used $ih \leqslant 1$. Then, (4.5) yields

$$\begin{aligned}
h^{-2}(a_i + b_{i-1} + b_i) \geqslant \gamma \kappa_i \Lambda_i \Bigg(& \frac{2}{\gamma h^2}(1 - e^{-\gamma h^2/2}) \min_{j=0,\dots,n} \phi'(\rho_j) \\
& - 2\cosh(\gamma h) \max_{j=0,\dots,n} |\nabla_h \phi'(\rho_j)| \Bigg) \\
= \lambda_h \kappa_i \Lambda_i. &
\end{aligned}$$

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

From numerical analysis and the expected large-time asymptotics of the equations involved, we expect that $\min_{i=0,\dots,n} \phi'(\rho_i)$ and $\max_{i=0,\dots,n} |\nabla_h \phi'(\rho_i)|$ are independent of h and bounded only by discrete norms of $\rho(0)$ under suitable assumptions on ϕ. In Appendix B, we provide these *a priori* estimates for the case $\gamma = 0$ and very small initial data, as an indication that they might hold true for the case $\gamma > 0$.

In the case $\gamma > 0$, 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 $|\rho_0| \leqslant C$ 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 [8] 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 [21, Section 3.3] for a detailed discussion on this issue.

Remark 4 *In the case $\gamma = 0$, the previous result implies that the entropy*

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

is convex along solutions of the semi-discrete diffusion equation

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

where $\rho_{-1} = \rho_0$ *and* $\rho_{n+1} = \rho_n$ *and* $\phi(s) = s^\alpha$ *for* $0 < \alpha \leqslant 1$.

The following remark, based on an idea of [10], shows that the standard entropy for the diffusion equation is not displacement convex along the solutions.

Remark 5 *Erbar and Maas [10] considered the diffusion equation in the form*

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

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* $L(\rho) = \mathrm{diag}(\Lambda(\rho_i, \rho_{i+1}))$, *but with the mean function*

$$\Lambda(\rho_i, \rho_{i+1}) = \frac{\phi(\rho_i) - \phi(\rho_{i+1})}{U'(\rho_i) - U'(\rho_{i+1})}. \tag{4.7}$$

The associated entropy is $\mathcal{F}(\rho) = \sum_{i=0}^n U(\rho_i)$. *We show next that* $\mathcal{F}(\rho)$ *is not displacement convex in general. Indeed, if* ρ *is a geodesic curve on* X_n *with respect to the non-linear transportation metric* $\mathcal{W}$ *induced by* (4.7), *then,*

$$\frac{d^2}{dt^2} \mathcal{F}(\rho) = \frac{1}{2} \langle \widetilde{M}(\rho) G\psi, G\psi \rangle, \tag{4.8}$$

where $\widetilde{M} = DL(\rho)[Q\phi(\rho)] + G\Phi'(\rho)G^\top L(\rho) + L(\rho)G\Phi'(\rho)G^\top$. *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

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

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

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

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 semi-definite. By choosing as initial data a vector of positive densities and the direction ψ such that the right-hand side of (4.8) *is negative at $t = 0$, we achieve that $\frac{d^2}{dt^2}\mathcal{F}(\rho) < 0$ for a small time interval. Therefore, the entropy fails to be convex at time $t = 0$.*

Remark 6 *We point out that showing the convergence of the numerical scheme could be done by looking at the convergence of the metric structures from the discrete to the continuum setting. A result in this direction was obtained in [10, Theorem 5.1] using Gromov–Haussdorff convergence arguments. It is an open problem to generalise it to the present setting.*

Acknowledgement

The authors warmly thank the Mittag-Leffler Institute for providing a marvellous atmosphere for research while this paper was being finalised.

References

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

[2] Bessemoulin-Chatard M. (2012) A finite volume scheme for convection-diffusion equations with nonlinear diffusion derived from the Scharfetter-Gummel scheme. *Numer. Math.* **121**, 637–670.

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

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

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

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

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

[8] Daneri, S. & Savaré, G. (2008) Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.* **40**, 1104–1122.

[9] Erbar, M. & Maas, J. (2012) Ricci curvature bounds for finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.* **206**, 997–1038.

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

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

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

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

[14] Jüngel, A. (2016) *Entropy Methods for Diffusive Partial Differential Equations*, BCAM SpringerBriefs, Cham.

[15] Jüngel, A. & Yue, W. (2017) Discrete Beckner inequalities via the Bochner–Bakry–Emery approach for Markov chains. *Ann. Appl. Probab.* **27**, 2238–2269.

[16] Lisini, S., Matthes, D. &, Savar, G. (2012) Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Diff. Eq.* **253**(2), 814–850.

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

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

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

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

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

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

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

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

[25] Villani, C. (2009) *Optimal Transport, Old and New,* Springer, Berlin.

Appendix A Properties of mean functions

We need some properties of the mean function

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

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

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

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

Proof Properties (i)–(iii) can be easily verified by a calculation. Properties (iv)–(v) are shown in [21, Appendix A]. □

Lemma 8 *Let* $\Lambda \in C^1([0,\infty)^2)$ *be any function being concave in both variables, and let* $u_0, u_1, u_2, u_3 \geqslant 0$. *Then,*

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

Proof Since Λ is concave in both variables, we have

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

and adding both inequalities gives the conclusion. □

Lemma 9 (Concavity of mean functions) *Let* $\Lambda^f : [0,\infty)^2 \to \mathbb{R}$ *be given by* (A 1) *and let either* $f(s) = s(\log s - 1)$ *or* $f(s) = s^\alpha$, *where* $1 < \alpha \leqslant 2$. *Then* Λ^f *is concave in both variables.*

Proof For $f(s) = s(\log s - 1)$, we refer to [9, Section 2]. The statement for $f(s) = s^\alpha$ is proved in [15, Appendix]. □

Appendix B *a priori* estimates

Lemma 10 (*a priori* estimates) *Let* ϕ *be non-decreasing,* $h > 0$ *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*

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

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

$$\min_{i=0,\dots,n} \rho_i(0) \leqslant \rho_i(t) \leqslant \max_{i=0,\dots,n} \rho_i(0), \tag{B 2}$$

$$\max_{i=0,\dots,n} |\nabla_h\phi(\rho_i(t))| \leqslant h^{-1/2}|\nabla_h\phi(\rho(0))|_2, \tag{B 3}$$

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

$$|\nabla_h\phi(\rho(0))|_2 := \left(\sum_{i=0}^n h|\nabla_h\phi(\rho_i(0))|^2\right)^{1/2}. \tag{B 4}$$

https://doi.org/10.1017/S0956792517000389 Published online by Cambridge University Press

Proof We multiply (B 1) by $(\rho_i - M)_+ = \max\{0, \rho_i - M\}$ and sum over $i = 0, \ldots, n$:

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

since ϕ is non-decreasing. This shows that

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

Thus, if $M = \max_{i=0,\ldots,n} \rho_i(0)$, the upper bound in (B 2) follows. The lower bound is proved analogously.

For the proof of (B 3), we compute

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

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

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

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

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

Taking the maximum over $j = 0, \ldots, n-1$ shows (B 3). □

Corollary 11 *Let ϕ be non-decreasing and invertible, $h > 0$, and let $\rho = (\rho_0, \ldots, \rho_n)$ be the solution to* (B 1). *We assume that $m := \min_{i=0,\ldots,n} \rho_i(0) > 0$ and set $M := \max_{i=0,\ldots,n} \rho_i(0)$. Then,*

$$\max_{i=0,\ldots,n} |\nabla_h \phi'(\rho_i)| \leqslant h^{-1/2} \max_{s\in[\phi^{-1}(m),\phi^{-1}(M)]} \left|\frac{\phi''(s)}{\phi'(s)}\right| |\nabla_h \phi(\rho(0))|_2, \tag{B 5}$$

where $|\nabla_h \phi(\rho(0))|_2$ is defined in (B 4).

Proof First, note that $m \leqslant \rho_i(t) \leqslant M$ for all $i = 0, \ldots, n$ and $t > 0$, by Lemma 10. Then the result follows from the mean value theorem. Indeed, we have for some ξ between ρ_{i+1} and ρ_i,

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

and we conclude after applying (B 3). □

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

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

where $|\nabla_h \phi(\rho(0))|_2$ is defined in (B 4)*. Indeed, the first statement follows from $\alpha < 1$ and* (B 2)*:*

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

and the second statement is a consequence of Corollary 11 evaluating the right-hand side of (B 5).