

GRADIENT FLOW STRUCTURES FOR DISCRETE POROUS MEDIUM EQUATIONS

MATTHIAS ERBAR AND JAN MAAS

ABSTRACT. We consider discrete porous medium equations of the form $\partial_t \rho_t = \Delta \varphi(\rho_t)$, where Δ is the generator of a reversible continuous time Markov chain on a finite set \mathcal{X} , and φ is an increasing function. We show that these equations arise as gradient flows of certain entropy functionals with respect to suitable non-local transportation metrics. This may be seen as a discrete analogue of the Wasserstein gradient flow structure for porous medium equations in \mathbb{R}^n discovered by Otto. We present a one-dimensional counterexample to geodesic convexity and discuss Gromov-Hausdorff convergence to the Wasserstein metric.

1. INTRODUCTION

Recently it has been shown that discretisations of heat equations and Fokker-Planck equations can be formulated as gradient flows of the entropy with respect to a non-local transportation metric \mathcal{W} on the space of probability measures. Results in this direction have been obtained independently in various settings, including Fokker-Planck equations on graphs [5], reversible Markov chains [11], and systems of reaction-diffusion equations [14]. Related gradient flow structures have been found for fractional heat equations [7] and quantum mechanical evolution equations [4, 15].

The above-mentioned results can be regarded as discrete counterparts to the by now classical result of Jordan, Kinderlehrer and Otto [10], who showed that Fokker-Planck equations on \mathbb{R}^n are gradient flows of the entropy with respect to the L^2 -Wasserstein metric on the space of probability measures. In this continuous setting there are many other interesting partial differential equations which admit a formulation as Wasserstein gradient flow. Among them, one of the most prominent examples is the porous medium equation, which has been identified as the Wasserstein gradient flow of the Rényi entropy in Otto's seminal paper [16].

It seems therefore natural to ask whether a similar gradient flow structure exists for discrete versions of porous medium equations. In this paper we show that suitable discretisations indeed admit a gradient flow structure. The associated metrics turn out to be variations on the metric \mathcal{W} , which have already been studied in [11].

1.1. The discrete setting. In this paper we let \mathcal{X} be a finite set. We consider a matrix $Q : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ satisfying $Q(x, y) \geq 0$ for all $x \neq y$ and $\sum_{y \in \mathcal{X}} Q(x, y) = 0$. We assume that Q is irreducible, so that basic Markov chain theory implies the existence of a unique invariant probability measure π on \mathcal{X} . We assume that π is reversible, i.e., the *detailed balance equations*

$$Q(x, y)\pi(x) = Q(y, x)\pi(y) \tag{1.1}$$

Key words and phrases. Non-local transportation metrics, Rényi entropy, porous medium equations.

hold for all $x, y \in \mathcal{X}$. For a function $\psi : \mathcal{X} \rightarrow \mathbb{R}$ we write

$$\Delta\psi(x) := \sum_{y \in \mathcal{X}} Q(x, y)\psi(y) = \sum_{y \in \mathcal{X}} Q(x, y)(\psi(y) - \psi(x)). \quad (1.2)$$

The operator Δ is the generator of a continuous time Markov semigroup $(P_t)_{t \geq 0}$. By the reversibility assumption, this semigroup is selfadjoint on $L^2(\mathcal{X}, \pi)$. Moreover, the set $\mathcal{P}(\mathcal{X})$ of probability densities with respect to π , is invariant under the semigroup $(P_t)_{t \geq 0}$.

We shall consider a class of non-local transportation metrics on the space of probability densities $\mathcal{P}(\mathcal{X})$. The definition of such a metric [11] requires a weight-function $\theta : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ satisfying appropriate conditions (see Definition 2.1 below). For $\rho \in \mathcal{P}(\mathcal{X})$ and $x, y \in \mathcal{X}$ we then define $\hat{\rho}(x, y) = \theta(\rho(x), \rho(y))$. Inspired by the Benamou-Brenier formula [2], we define for $\bar{\rho}_0, \bar{\rho}_1 \in \mathcal{P}(\mathcal{X})$ the distance $\mathcal{W}(\bar{\rho}_0, \bar{\rho}_1)$ by

$$\mathcal{W}(\bar{\rho}_0, \bar{\rho}_1)^2 := \inf_{\rho, \psi} \left\{ \frac{1}{2} \int_0^1 \sum_{x, y \in \mathcal{X}} (\psi(y) - \psi(x))^2 \hat{\rho}(x, y) Q(x, y) \pi(x) dt \right\},$$

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

$$\partial_t \rho_t(x) + \sum_{y \in \mathcal{X}} (\psi_t(y) - \psi_t(x)) \hat{\rho}_t(x, y) Q(x, y) = 0$$

for all $x \in \mathcal{X}$, with boundary conditions $\rho_0 = \bar{\rho}_0$ and $\rho_1 = \bar{\rho}_1$. Actually, \mathcal{W} is the Riemannian distance induced by a Riemannian metric defined on the interior $\mathcal{P}_*(\mathcal{X})$ of $\mathcal{P}(\mathcal{X})$.

Consider now the relative entropy functional $\mathcal{H} : \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ defined by

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

The following result has been shown in [5, 11, 14]: *Solutions to the heat equation $\partial_t \rho_t = \Delta \rho_t$ are gradient flow trajectories of \mathcal{H} with respect to the metric \mathcal{W} , provided that θ is the logarithmic mean defined by*

$$\theta(r, s) := \int_0^1 r^{1-\alpha} s^\alpha d\alpha. \quad (1.3)$$

1.2. A gradient flow structure for discrete porous medium equations. In this paper we shall consider non-linear generalisations of the discrete heat equation. For suitable (see Assumption 3.1) strictly increasing functions $\varphi : [0, \infty) \rightarrow \mathbb{R}$ we consider the *discrete porous medium type equation* given by

$$\partial_t \rho_t = \Delta \varphi(\rho_t), \quad (1.4)$$

where Δ denotes the discrete Laplacian defined in (1.2). This equation can be analysed using classical Hilbertian gradient flow techniques in a suitable “discrete H^{-1} -space” (see Proposition 3.2 below).

Here we are interested in more general gradient flow structures in the spirit of the Wasserstein gradient flow structure for the porous medium equation [16]. For this purpose, for

suitable (see Assumption 3.3) strictly convex functions $f : [0, \infty) \rightarrow \mathbb{R}$, we consider the entropy functional $\mathcal{F} : \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ defined by

$$\mathcal{F}(\rho) := \sum_{x \in \mathcal{X}} f(\rho(x)) \pi(x). \quad (1.5)$$

We shall prove the following result.

Theorem 1.1. *Let φ and f be as above, and let \mathcal{W} denote the non-local transportation metric induced by*

$$\theta(r, s) = \frac{\varphi(r) - \varphi(s)}{f'(r) - f'(s)}. \quad (1.6)$$

Then the gradient flow equation of the entropy functional \mathcal{F} with respect to \mathcal{W} is the discrete porous medium equation $\partial_t \rho = \Delta \varphi(\rho)$.

Remark 1.2. In the special case where $\varphi(r) = r$ and $f(r) = r \log r$ we recover the result from [5, 11, 14]. In this case we have $\theta(r, s) = \frac{r-s}{\log r - \log s}$, which coincides with (1.3). The possibility of allowing more general functions f has already been discussed in [11].

Remark 1.3. Of particular interest is the case where $\varphi(r) = r^m$ and $f(r) = \frac{1}{m-1} r^m$ for some $0 < m \leq 2$. In this case the equation can be considered as a discrete porous medium equation (if $m > 1$) or fast diffusion equation (if $m < 1$). The functional \mathcal{F} becomes a Rényi entropy \mathcal{F}_m , so that Theorem 1.1 can be considered as a discrete analogue of the gradient flow structure obtained by Otto [16]. The expression for θ from (1.6) becomes

$$\theta_m(r, s) = \frac{m-1}{m} \frac{r^m - s^m}{r^{m-1} - s^{m-1}}. \quad (1.7)$$

Several classical means are special cases of this expression. In fact, $\theta_2(r, s) = \frac{r+s}{2}$ is the arithmetic mean, $\lim_{m \rightarrow 1} \theta_m(r, s)$ is the logarithmic mean, and $\theta_{1/2}(r, s) = \sqrt{rs}$ is the geometric mean. Moreover, $\theta_{-1}(r, s) = \frac{2rs}{r+s}$ is the harmonic mean, but we shall not consider negative values of m in the sequel.

Remark 1.4. For integer values of m , the porous medium equation takes the form of a chemical reaction equation, for which a gradient flow structure has been found in [14]. However, there the driving functional is the relative entropy \mathcal{H} , so that the associated weight function can be expressed in terms of logarithmic means. Here we also allow for different driving functionals, such as the Rényi entropy.

1.3. Geodesic convexity of entropy functionals. Of crucial importance in the theory of Wasserstein gradient flows is that the relevant entropy functionals exhibit good convexity properties along Wasserstein geodesics. Recent work [8, 13] shows that analogous properties hold for the relative entropy functional \mathcal{H} in some discrete examples. In particular, \mathcal{H} turns out to be convex along \mathcal{W} -geodesics in one-dimensional discrete Fokker-Planck equations [13] as well as in heat equations on d -dimensional square lattices in arbitrary dimension [8].

It thus seems natural to ask whether similar convexity properties hold for the more general functionals \mathcal{F} along \mathcal{W} -geodesics with the appropriate choice of θ given by (1.6). In the continuous setting, it follows from a fundamental result by McCann [12] that the Rényi entropy $\mathcal{F}_m(\rho) = \frac{1}{m-1} \int_{\mathbb{R}^n} \rho^m(x) dx$ is displacement convex, i.e., convex along L^2 -Wasserstein geodesics in $\mathcal{P}(\mathbb{R}^n)$, for $m \geq 1 - \frac{1}{n}$.

However, in this paper we shall show that the discrete analogue fails in general. We present a counterexample (Proposition 4.5) in the case $m = 2$, which shows that even in one dimension the Rényi entropy fails to be convex along geodesics in the associated non-local transportation metric.

Proposition 1.5. *For $N \geq 6$, let Q be the generator of simple random walk on the discrete circle $\mathbf{T}_N = \mathbb{Z}/N\mathbb{Z}$. Let \mathcal{W} be the non-local transportation metric associated with the arithmetic mean θ_2 . Then the Rényi entropy \mathcal{F}_2 is not convex along \mathcal{W} -geodesics.*

It remains an interesting open question whether the Rényi entropy \mathcal{F}_m is geodesically convex in some non-trivial discrete examples for values of $m \neq 1$.

1.4. Gromov-Hausdorff convergence of discrete transportation metrics. Since the discrete transportation metrics discussed in this paper take over the role of the L^2 -Wasserstein metric in a discrete setting, it seems natural to ask whether they converge to the L^2 -Wasserstein metric by a suitable limiting procedure. First results in this spirit have been obtained in [9], where it was proved that discrete transportation metrics \mathcal{W}_N associated with simple random walk on the discrete torus $(\mathbb{Z}/N\mathbb{Z})^d$ converge to the L^2 -Wasserstein metric W_2 over the torus \mathbb{T}^d . The weight function considered in [9] is the logarithmic mean. In the present paper we observe that the result extends to more general weight functions associated to porous medium equations.

Theorem 1.6. *Let $d \geq 1$, let $0 < m \leq 2$, and let \mathcal{W}_N be the renormalised discrete transportation metric on $\mathcal{P}(\mathbb{T}_N^d)$ associated with simple random walk on \mathbb{T}_N^d and weight function θ_m . Then the metric spaces $(\mathcal{P}(\mathbb{T}_N^d), \mathcal{W}_N)$ converge to $(\mathcal{P}(\mathbb{T}^d), W_2)$ in the sense of Gromov-Hausdorff as $N \rightarrow \infty$.*

Since this result can be obtained by a minor modification of the proof in [9], we do not give a detailed proof here, but merely point out the crucial properties which make the argument go through.

2. PRELIMINARIES ON NON-LOCAL TRANSPORTATION METRICS

In this section we collect some preliminary results on non-local transportation metrics. The presentation here is based on [11].

In this paper, a reversible Markov chain consists of a triple (\mathcal{X}, Q, π) where

- \mathcal{X} is a finite set;
- $Q : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is a Q -matrix, i.e., $Q(x, y) \geq 0$ for all $x \neq y$ and $\sum_{y \in \mathcal{X}} Q(x, y) = 0$ for all $x \in \mathcal{X}$. Throughout this paper we assume that Q is irreducible, i.e., for all $x, y \in \mathcal{X}$ there exist $n \geq 1$ and $x_0, \dots, x_n \in \mathcal{X}$ such that $x_0 = x$, $x_n = y$ and $Q(x_{i-1}, x_i) > 0$ for all $1 \leq i \leq n$;
- π is the associated stationary probability measure on \mathcal{X} , which exists uniquely by basic Markov chain theory. We impose the standing assumption that π is reversible, i.e., we assume that the *detailed balance equations*

$$Q(x, y)\pi(x) = Q(y, x)\pi(y) \tag{2.1}$$

are satisfied for all $x, y \in \mathcal{X}$.

The set of probability densities with respect to π will be denoted by

$$\mathcal{P}(\mathcal{X}) := \left\{ \rho : \mathcal{X} \rightarrow \mathbb{R} \mid \rho(x) \geq 0 \quad \forall x \in \mathcal{X} \ ; \ \sum_{x \in \mathcal{X}} \rho(x)\pi(x) = 1 \right\},$$

and we set

$$\mathcal{P}_*(\mathcal{X}) := \left\{ \rho \in \mathcal{P}(\mathcal{X}) \mid \rho(x) > 0 \quad \forall x \in \mathcal{X} \right\}.$$

Definition 2.1. A weight function is a function $\theta : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ having the following properties:

- (A1) θ is continuous on $[0, \infty) \times [0, \infty)$ and C^∞ on $(0, \infty) \times (0, \infty)$;
- (A2) $\theta(s, t) = \theta(t, s)$ for $s, t \geq 0$;
- (A3) $\theta(s, t) > 0$ for $s, t > 0$;
- (A4) $\theta(r, s) \leq \theta(r, t)$ for all $r \geq 0$ and $0 \leq s \leq t$;
- (A5) $\theta : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ is concave.

Note that (A5) implies in particular the following estimate:

$$\theta(2s, 2t) \leq 2\theta(s, t) \quad \forall s, t \geq 0.$$

Let us fix a reversible Markov chain (\mathcal{X}, Q, π) and the weight function θ throughout the remainder of this section. We recall the definition of the associated non-local transportation metric from [11], as given in an equivalent form in [8]. Actually, the setup in [11, 8] is slightly different from the one considered here, since the starting point there is a Markov kernel instead of a Q -matrix. The transportation metrics defined here are the same as in [11, 8], possibly up to a numerical constant. We refer to Remark 2.7 below for more details.

Definition 2.2 (Non-local transportation metrics). Fix a weight function θ . For $\bar{\rho}_0, \bar{\rho}_1 \in \mathcal{P}(\mathcal{X})$ we define

$$\mathcal{W}(\bar{\rho}_0, \bar{\rho}_1)^2 := \inf \left\{ \int_0^1 \mathcal{A}(\rho_t, \psi_t) dt : (\rho, \psi) \in \mathcal{CE}(\bar{\rho}_0, \bar{\rho}_1) \right\},$$

where $\mathcal{CE}(\bar{\rho}_0, \bar{\rho}_1)$ denotes the collection of pairs (ρ, ψ) satisfying the following conditions:

$$\left\{ \begin{array}{l} (i) \quad \rho : [0, 1] \rightarrow \mathbb{R}^{\mathcal{X}} \text{ is } C^\infty; \\ (ii) \quad \rho_0 = \bar{\rho}_0, \quad \rho_1 = \bar{\rho}_1; \\ (iii) \quad \rho_t \in \mathcal{P}(\mathcal{X}) \text{ for all } t \in [0, 1]; \\ (iv) \quad \psi : [0, 1] \rightarrow \mathbb{R}^{\mathcal{X}} \text{ is measurable}; \\ (v) \quad \text{For all } x \in \mathcal{X} \text{ and all } t \in (0, 1) \text{ we have} \\ \quad \quad \partial_t \rho_t(x) + \sum_{y \in \mathcal{X}} (\psi_t(y) - \psi_t(x)) \hat{\rho}_t(x, y) Q(x, y) = 0, \end{array} \right. \quad (2.2)$$

and for $\rho \in \mathcal{P}(\mathcal{X})$ and $\psi \in \mathbb{R}^{\mathcal{X}}$,

$$\mathcal{A}(\rho, \psi) := \frac{1}{2} \sum_{x, y \in \mathcal{X}} (\psi(y) - \psi(x))^2 \hat{\rho}(x, y) Q(x, y) \pi(x).$$

The equation appearing in (2.2) can be regarded as a “discrete continuity equation”. The analogy with the continuous case becomes more apparent if we introduce some notation. For a function $\psi : \mathcal{X} \rightarrow \mathbb{R}$, we consider the *discrete gradient* $\nabla \psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ given by

$$\nabla \psi(x, y) := \psi(y) - \psi(x).$$

For a function $\Psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ we define its *discrete divergence* $\nabla \cdot \Psi : \mathcal{X} \rightarrow \mathbb{R}$ by

$$(\nabla \cdot \Psi)(x) := \frac{1}{2} \sum_{y \in \mathcal{X}} (\Psi(x, y) - \Psi(y, x)) Q(x, y) \in \mathbb{R}.$$

With this notation, it follows from the detailed balance equations (2.1) that the integration by parts formula

$$\langle \nabla \psi, \Psi \rangle_\pi = -\langle \psi, \nabla \cdot \Psi \rangle_\pi$$

holds. Here we write, for $\varphi, \psi : \mathcal{X} \rightarrow \mathbb{R}$ and $\Phi, \Psi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$,

$$\begin{aligned} \langle \varphi, \psi \rangle_\pi &= \sum_{x \in \mathcal{X}} \varphi(x) \psi(x) \pi(x), \\ \langle \Phi, \Psi \rangle_\pi &= \frac{1}{2} \sum_{\substack{x, y \in \mathcal{X} \\ x \neq y}} \Phi(x, y) \Psi(x, y) Q(x, y) \pi(x). \end{aligned}$$

The continuity equation above can now be written as

$$\partial_t \rho_t + \nabla \cdot (\hat{\rho}_t \nabla \psi_t) = 0.$$

Moreover, for $\rho \in \mathcal{P}(\mathcal{X})$ we shall write

$$\langle \Phi, \Psi \rangle_\rho := \frac{1}{2} \sum_{\substack{x, y \in \mathcal{X} \\ x \neq y}} \Phi(x, y) \Psi(x, y) \hat{\rho}(x, y) Q(x, y) \pi(x), \quad \|\Psi\|_\rho := \sqrt{\langle \Psi, \Psi \rangle_\rho}.$$

With this notation we have

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

Basic properties of \mathcal{W} have been studied in [11]. Under the current assumptions we have the following result.

Proposition 2.3 (Metric structure). *The function $\mathcal{W} : \mathcal{P}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ defines a metric on $\mathcal{P}(\mathcal{X})$. Moreover, $(\mathcal{P}(\mathcal{X}), \mathcal{W})$ is a geodesic space, i.e., for all $\bar{\rho}_0, \bar{\rho}_1 \in \mathcal{P}(\mathcal{X})$ there exists a curve $\rho : [0, 1] \rightarrow \mathcal{P}(\mathcal{X})$ satisfying $\rho_0 = \bar{\rho}_0$, $\rho_1 = \bar{\rho}_1$, and*

$$\mathcal{W}(\rho_s, \rho_t) = |s - t| \mathcal{W}(\rho_0, \rho_1)$$

for all $s, t \in [0, 1]$.

Proof. Under slightly different assumptions on θ , it has been proved in [11] that \mathcal{W} defines an extended, i.e., (possibly $[0, \infty]$ -valued) metric. To show that \mathcal{W} is finite under the current assumptions, we apply [11, Theorem 3.12, Part(1)], which asserts that it suffices to check that

$$C_\theta := \int_0^1 \frac{1}{\sqrt{\theta(1-r, 1+r)}} dr < \infty.$$

(Let us remark that one of the standing assumptions in [11], denoted by (A5) in that paper, is not in force in the present paper. This doesn't cause problems here, since the proof of [11, Theorem 3.12, Part(1)] is based on [11, Lemma 3.14], in which the additional assumption clearly plays no role). Using the concavity assumption (A5) we infer that

$$\theta(1-r, 1+r) \geq (1-r)\theta(1, 1) + r\theta(0, 2) \geq (1-r)\theta(1, 1)$$

for $r \in [0, 1]$, which implies that $C_\theta < \infty$.

The final assertion has been proved in [8, Theorem 3.2]. \square

Some basic properties of the metric \mathcal{W} are collected in the following result, which asserts that the distance \mathcal{W} is induced by a Riemannian metric on the interior $\mathcal{P}_*(\mathcal{X})$ of $\mathcal{P}(\mathcal{X})$.

Proposition 2.4 (Riemannian structure). *The restriction of \mathcal{W} to $\mathcal{P}_*(\mathcal{X})$ is the Riemannian distance induced by the following Riemannian structure:*

- the tangent space of $\rho \in \mathcal{P}_*(\mathcal{X})$ can be identified with the set of discrete gradients

$$T_\rho := \{\nabla\psi : \psi \in \mathbb{R}^{\mathcal{X}}\}$$

by means of the following identification: given a smooth curve $(-\varepsilon, \varepsilon) \ni t \mapsto \rho_t \in \mathcal{P}_*(\mathcal{X})$ with $\rho_0 = \rho$, there exists a unique element $\nabla\psi_0 \in T_\rho$, such that the continuity equation (2.2)(v) holds at $t = 0$.

- The Riemannian metric on T_ρ is given by the inner product

$$\langle \nabla\varphi, \nabla\psi \rangle_\rho = \frac{1}{2} \sum_{x, y \in \mathcal{X}} (\varphi(x) - \varphi(y))(\psi(x) - \psi(y)) \hat{\rho}(x, y) Q(x, y) \pi(x) .$$

This result allows us to perform Riemannian computations (“Otto-calculus” [16]) in the space of probability measures, even though the underlying space \mathcal{X} is discrete. In particular, if $\mathcal{G} : \mathcal{P}_*(\mathcal{X}) \rightarrow \mathbb{R}$ is a smooth functional and (ρ_t, ψ_t) is a smooth solution to the continuity equation, we have

$$\frac{d}{dt} \mathcal{G}(\rho_t) = \langle \text{grad}_{\mathcal{W}} \mathcal{G}(\rho_t), \nabla\psi_t \rangle_{\rho_t} ,$$

and if (ρ_t, ψ_t) is a geodesic,

$$\frac{d^2}{dt^2} \mathcal{G}(\rho_t) = \langle \text{Hess}_{\mathcal{W}} \mathcal{G}(\rho_t) \nabla\psi_t, \nabla\psi_t \rangle_{\rho_t} .$$

The following result characterises the geodesic equations for \mathcal{W} . The equation for ψ is reminiscent of the Hamilton-Jacobi equation, which describes geodesics in the Wasserstein space.

Proposition 2.5 (Geodesic equations). *Let $\bar{\rho} \in \mathcal{P}_*(\mathcal{X})$ and $\bar{\psi} \in \mathbb{R}^{\mathcal{X}}$. On a sufficiently small time interval around 0, the unique constant speed geodesic with $\rho_0 = \bar{\rho}$ and initial tangent vector $\nabla\psi_0 = \nabla\bar{\psi}$ satisfies the following equations:*

$$\begin{cases} \partial_t \rho_t(x) + \sum_{y \in \mathcal{X}} (\psi_t(y) - \psi_t(x)) \hat{\rho}_t(x, y) Q(x, y) = 0 , \\ \partial_t \psi_t(x) + \frac{1}{2} \sum_{y \in \mathcal{X}} (\psi_t(x) - \psi_t(y))^2 \partial_1 \theta(\rho_t(x), \rho_t(y)) Q(x, y) = 0 . \end{cases} \quad (2.3)$$

Let Δ be the discrete Laplacian associated with Q , i.e., for a function $\psi : \mathcal{X} \rightarrow \mathbb{R}$ we set

$$\Delta\psi(x) := \sum_{y \in \mathcal{X}} Q(x, y) (\psi(y) - \psi(x)) . \quad (2.4)$$

The operator Δ is the generator of the continuous time Markov semigroup associated with Q . Note that we have the usual formula $\Delta = (\nabla \cdot) \nabla$. By the reversibility assumption, Δ is selfadjoint on $L^2(\mathcal{X}, \pi)$.

Let $f \in C([0, \infty); \mathbb{R})$ be a strictly convex function, which is smooth on $(0, \infty)$. We consider the associated entropy functional $\mathcal{F} : \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ defined by

$$\mathcal{F}(\rho) := \sum_{x \in \mathcal{X}} f(\rho(x)) \pi(x) .$$

The following result explains the relevance of the non-local transportation metrics \mathcal{W} .

Proposition 2.6 (Gradient flows). *The heat flow generated by Δ is the gradient flow of the entropy with respect to \mathcal{W} , provided that θ is given by*

$$\theta(r, s) := \frac{r - s}{f'(r) - f'(s)}. \quad (2.5)$$

Let us note that among all weight functions θ of the form (2.5), the logarithmic mean (1.3) is the only one satisfying $\theta(r, r) = r$.

Remark 2.7. We conclude this section by comparing the setup in this paper with the one from [11], and show that both setups are equivalent up to a constant.

Let Q be as in this section, set $q(x) := \sum_{y \neq x} Q(x, y)$ and $q_{\max} := \max_{x \in \mathcal{X}} q(x)$. We may then define a Markov kernel $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ by

$$K(x, y) = \begin{cases} \frac{Q(x, y)}{q_{\max}}, & x \neq y, \\ \frac{q_{\max} - q(x)}{q_{\max}}, & x = y. \end{cases}$$

Clearly, a probability measure π on \mathcal{X} satisfies the detailed balance equations for Q , if and only if the same equations hold for K . It is then immediate to check that

$$\mathcal{W}_K = \sqrt{q_{\max}} \mathcal{W}_Q,$$

where \mathcal{W}_Q denotes the metric defined in this paper and \mathcal{W}_K is the one from [11].

Conversely, given a Markov kernel K , we may define $Q : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ by

$$Q(x, y) = \begin{cases} K(x, y), & x \neq y, \\ -\sum_{z \neq x} K(x, z), & x = y. \end{cases}$$

It readily follows from the definitions that $\mathcal{W}_Q = \mathcal{W}_K$.

3. DISCRETE POROUS MEDIUM EQUATIONS AS GRADIENT FLOWS OF THE ENTROPY

In this note we shall be concerned with equations of porous medium-type associated with a reversible Markov chain (\mathcal{X}, Q, π) . The following assumption will be in force throughout this section.

Assumption 3.1. *The function $\varphi : [0, \infty) \rightarrow \mathbb{R}$ is strictly increasing, C^∞ on $(0, \infty)$, and continuous at 0.*

We shall study the *discrete porous medium equation*

$$\partial_t \rho_t = \Delta \varphi(\rho_t), \quad (3.1)$$

where Δ denotes the discrete Laplacian defined in (2.4). Of course, if $\varphi(r) = r$, we recover the usual discrete heat equation associated with Q , which has been studied in [5, 11, 14]. We shall analyse these equations using gradient flow methods, first in a Hilbertian setting of discrete Sobolev spaces, and then in a Riemannian setting of non-local transportation metrics.

3.1. A Hilbertian gradient flow structure. In order to apply Hilbertian gradient flow methods we shall introduce a “discrete H^{-1} -space” \mathcal{H}^{-1} . First we observe that, by the irreducibility assumption, we have $\text{Ker}(\Delta) = \text{lin}\{\mathbf{1}\}$, i.e., the linear span of the function identically equal to 1. Since Δ is selfadjoint on $L^2(\mathcal{X}, \pi)$, it follows that the operator Δ is bijective on

$$\text{Ran}(\Delta) := \left\{ \psi : \mathcal{X} \rightarrow \mathbb{R} \mid \sum_{x \in \mathcal{X}} \psi(x) \pi(x) = 0 \right\}.$$

For $\psi, \psi_1, \psi_2 \in \text{Ran}(\Delta)$ it thus makes sense to define the inner product

$$\langle \psi_1, \psi_2 \rangle_{\mathcal{H}^{-1}} := \langle -\Delta^{-1} \psi_1, \psi_2 \rangle_{\pi},$$

where $\langle \cdot, \cdot \rangle_{\pi}$ denotes the $L^2(\mathcal{X}, \pi)$ -inner product. The associated norm is given by

$$\|\psi\|_{\mathcal{H}^{-1}} := \sqrt{\langle \psi, \psi \rangle_{\mathcal{H}^{-1}}} = \|\nabla \Delta^{-1} \psi\|_{\pi} = \sqrt{\frac{1}{2} \sum_{x, y \in \mathcal{X}} (\Delta^{-1} \psi(y) - \Delta^{-1} \psi(x))^2 Q(x, y) \pi(x)}.$$

For arbitrary functions $\psi : \mathcal{X} \rightarrow \mathbb{R}$ we set $c_{\psi} = \sum_{x \in \mathcal{X}} \psi(x) \pi(x)$ and note that $\psi - c_{\psi} \mathbf{1}$ belongs to $\text{Ran}(\Delta)$. A Hilbertian norm on $\mathbb{R}^{\mathcal{X}}$ can then be defined by

$$\|\psi\|_{\mathcal{H}^{-1}} := \sqrt{c_{\psi}^2 + \|\psi - c_{\psi} \mathbf{1}\|_{\mathcal{H}^{-1}}^2}.$$

The equation $\partial_t \rho_t = \Delta \varphi(\rho_t)$ can now be studied using classical Hilbertian gradient flow arguments:

Proposition 3.2. *Let $\Phi(r) = \int_0^r \varphi(s) \, ds$ and consider the functional $\Phi : \mathbb{R}^{\mathcal{X}} \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by*

$$\Phi(\rho) := \sum_{x \in \mathcal{X}} \Phi(\rho(x)) \pi(x)$$

for $\rho \in \mathcal{P}(\mathcal{X})$, and $\Phi(\rho) = +\infty$ otherwise. Then the functional Φ is strictly convex on $\mathcal{P}(\mathcal{X})$ and attains its unique minimum at $\mathbf{1}$. As a consequence, the following assertions hold for all $\bar{\rho}_0 \in \mathcal{P}(\mathcal{X})$:

- (1) *Among all locally absolutely continuous curve $(\rho_t)_t$ in $\mathcal{P}(\mathcal{X})$ with $\rho_0 = \bar{\rho}_0$, there exists a unique one satisfying the evolution variational inequality*

$$\frac{1}{2} \frac{d}{dt} \|\rho_t - \sigma\|_{\mathcal{H}^{-1}}^2 \leq \Phi(\sigma) - \Phi(\rho_t) \quad (3.2)$$

for all $\sigma \in \mathcal{P}(\mathcal{X})$ and a.e. $t \geq 0$.

- (2) *For all $\bar{\rho}_0 \in \mathcal{P}(\mathcal{X})$ and a.e. $t \geq 0$ the curve (ρ_t) satisfies the ordinary differential equation*

$$\partial_t \rho_t = \Delta \varphi(\rho_t).$$

- (3) *For all $\bar{\rho}_0 \in \mathcal{P}(\mathcal{X})$ we have $\rho_t \rightarrow \mathbf{1}$ as $t \rightarrow \infty$.*

Proof. Strict convexity of Φ follows from the assumption that φ is strictly increasing. Part (1) follows then from the theory of gradient flows of convex functionals in Hilbert spaces [3, Théorème 3.1] or the more general theory in metric spaces [1, Theorem 4.0.4]. To prove (2),

it suffices to compute the \mathcal{H}^{-1} -gradient of Φ . To do this, note that for a smooth curve (ρ_t) we have

$$\frac{d}{dt}\Phi(\rho_t) = \frac{d}{dt}\langle\Phi(\rho_t), \mathbf{1}\rangle_\pi = \langle\varphi(\rho_t), \partial_t\rho_t\rangle_\pi = -\langle\Delta\varphi(\rho_t), \partial_t\rho_t\rangle_{\mathcal{H}^{-1}},$$

which shows that the \mathcal{H}^{-1} -gradient of the functional Φ at $\rho \in \mathcal{P}(\mathcal{X})$ is given by $-\Delta\varphi(\rho)$, as desired.

Since the gradient of the convex functional Φ vanishes at $\mathbf{1}$, it follows that Φ attains its minimum at $\mathbf{1}$. (Alternatively, this follows from Jensen's inequality.) Uniqueness of the minimiser follows from the strict convexity of Φ . Part (3) is then a consequence of [1, Corollary 4.0.6]. \square

3.2. A gradient flow structure for non-local transportation metrics. The gradient flow structure described in Proposition 3.2 can be regarded as a discrete analogue of the H^{-1} -gradient flow structure of the porous medium equation. Let us now focus on different gradient flow structures for the same equation, which are analogous to the Wasserstein-gradient flow structure. We refer to [16, Section 2] for a discussion of different gradient flow structures for the porous medium equation in the continuous setting.

Assumption 3.3. *The function $f : [0, \infty) \rightarrow \mathbb{R}$ is strictly convex, C^∞ on $(0, \infty)$, and continuous at 0. Moreover, the function $\theta_{\varphi, f}$ defined by*

$$\theta_{\varphi, f}(r, s) := \frac{\varphi(r) - \varphi(s)}{f'(r) - f'(s)}$$

for $r \neq s$, extends to a weight function on $[0, \infty) \times [0, \infty)$ in the sense of Definition 2.1.

The associated metric on $\mathcal{P}(\mathcal{X})$ will be denoted by $\mathcal{W}_{\varphi, f}$. Often, if confusion is unlikely to arise, we will simply write \mathcal{W} .

We consider the entropy functional $\mathcal{F} : \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ defined by

$$\mathcal{F}(\rho) := \sum_{x \in \mathcal{X}} f(\rho(x))\pi(x). \quad (3.3)$$

Of course, if $f(r) = r \log r$ we recover the usual relative entropy with respect to π .

Proposition 3.4. *Let \mathcal{W} be any non-local transportation metric. The \mathcal{W} -gradient of the functional \mathcal{F} at $\rho \in \mathcal{P}_*(\mathcal{X})$ is given by $\nabla f'(\rho)$.*

Here we use the identification of the tangent space provided in Proposition 2.4. Note that we do not require that $\mathcal{W} = \mathcal{W}_{\varphi, f}$ in this result.

Proof. Let $\rho \in \mathcal{P}_*(\mathcal{X})$ and pick a smooth curve (ρ_t) satisfying the continuity equation

$$\partial_t\rho_t + \nabla \cdot (\hat{\rho}_t \nabla \psi_t) = 0,$$

where we use the notation from Section 2. It then follows that

$$\frac{d}{dt}\mathcal{F}(\rho_t) = \frac{d}{dt}\langle f(\rho_t), \mathbf{1} \rangle_\pi = -\langle f'(\rho_t), \nabla \cdot (\hat{\rho}_t \nabla \psi_t) \rangle_\pi = \langle \nabla f'(\rho_t), \nabla \psi_t \rangle_{\rho_t},$$

hence the \mathcal{W} -gradient of the functional \mathcal{F} at $\rho \in \mathcal{P}_*(\mathcal{X})$ is given by $\nabla f'(\rho)$. \square

The next result is now an immediate consequence.

Theorem 3.5. *The gradient flow equation of the entropy functional \mathcal{F} with respect to the metric $\mathcal{W}_{\varphi, f}$ is given by the porous medium equation $\partial_t\rho_t = \Delta\varphi(\rho_t)$.*

Proof. The gradient flow equation of the smooth functional \mathcal{F} on $\mathcal{P}_*(\mathcal{X})$ is given by

$$\partial_t \rho_t + \nabla \cdot (\hat{\rho}_t \nabla \psi_t) = 0, \quad \nabla \psi_t = -\text{grad}_{\mathcal{W}} \mathcal{F}(\rho_t).$$

Using Proposition 3.4 and the fact that $\hat{\rho} \nabla f'(\rho) = \nabla \varphi(\rho)$ by the definition of $\theta_{\varphi, f}$, we infer that

$$\nabla \cdot (\hat{\rho}_t \nabla \psi_t) = -\nabla \cdot (\hat{\rho}_t \nabla f'(\rho_t)) = -\Delta \varphi(\rho_t),$$

hence the gradient flow equation reduces to the porous medium equation $\partial_t \rho_t = \Delta \varphi(\rho_t)$, as desired. \square

Clearly, the gradient flow structure for the heat equation given in Proposition 2.6 corresponds to the special case of Theorem 3.5 where $\varphi(r) = r$.

Let us note that also the \mathcal{H}^{-1} -gradient flow structure from Proposition 3.2 is a special case of Theorem 3.5. Indeed, if $f = \Phi$, then $\mathcal{F} = \Phi$. Moreover, the following result asserts that the \mathcal{H}^{-1} -distance on $\mathcal{P}(\mathcal{X})$ coincides with the non-local transportation metric \mathcal{W}_1 , which is given by Definition 2.2 in the special case where $\theta(r, s) = 1$ for all $r, s \geq 0$.

Proposition 3.6. *For $\rho_0, \rho_1 \in \mathcal{P}(\mathcal{X})$ we have*

$$\|\rho_0 - \rho_1\|_{\mathcal{H}^{-1}} = \mathcal{W}_1(\rho_0, \rho_1),$$

where \mathcal{W}_1 denotes the discrete transportation metric with $\theta \equiv 1$.

Proof. Note that the geodesic equations (2.3) for \mathcal{W}_1 reduce to

$$\partial_t \rho_t = -\Delta \psi_t, \quad \partial_t \psi_t = 0.$$

Since ψ does not depend on t , we infer that the \mathcal{W}_1 -geodesic between ρ_0 and ρ_1 in $\mathcal{P}(\mathcal{X})$ is given by linear interpolation, i.e., $\rho_t = (1-t)\rho_0 + \rho_1$. It follows that $\Delta \psi_t = \rho_0 - \rho_1$ for all $t \in [0, 1]$, and since $\|\cdot\|_{\rho_t} = \|\cdot\|_{\pi}$ we obtain

$$\mathcal{W}_1(\rho_0, \rho_1)^2 = \int_0^1 \|\nabla \psi_t\|_{\rho_t}^2 dt = \|\nabla \Delta^{-1}(\rho_0 - \rho_1)\|_{\pi}^2 = \|\rho_0 - \rho_1\|_{\mathcal{H}^{-1}}^2.$$

\square

3.3. Porous medium equations and Rényi entropy. Let us now specialise to a particularly interesting setting, motivated by the Wasserstein gradient flow structure for porous medium equations of the form $\partial_t \rho_t = \Delta \rho_t^m$ in \mathbb{R}^n from [16].

Let $0 < m \leq 2$ and $m \neq 1$, and consider the functions

$$\varphi(r) = \varphi_m(r) = r^m, \quad f(r) = f_m(r) = \frac{1}{m-1} r^m.$$

In this case the porous medium equation and the entropy functional are given by

$$\partial_t \rho_t = \Delta \rho_t^m, \quad \mathcal{F}(\rho) = \mathcal{F}_m(\rho) := \frac{1}{m-1} \sum_{x \in \mathcal{X}} \rho(x)^m \pi(x),$$

thus \mathcal{F}_m is the usual Rényi entropy. The weight function is given by

$$\theta_m(r, s) := \theta_{\varphi_m, f_m}(r, s) := \frac{m-1}{m} \frac{r^m - s^m}{r^{m-1} - s^{m-1}}, \quad (3.4)$$

and we let \mathcal{W}_m denote the associated non-local transportation metric. Note that

$$\lim_{m \rightarrow 1} \theta_m(r, s) = \theta_1(r, s),$$

where θ_1 denotes the logarithmic mean defined in (1.3). The weight function θ_m admits the remarkable integral representation

$$\theta_m(r, s) = \int_0^1 \left((1-\alpha)r^{m-1} + \alpha s^{m-1} \right)^{\frac{1}{m-1}} d\alpha, \quad (3.5)$$

which can be checked by an explicit calculation and naturally generalises (1.3). We collect here some basic properties of the weight functions associated to the Rényi entropy.

Lemma 3.7. *Let $0 < m \leq 2$. Then the function θ_m defines a weight function with the following additional properties:*

- (1) (Homogeneity) $\theta_m(\lambda s, \lambda t) = \lambda \theta_m(s, t)$ for $\lambda > 0$ and $s, t > 0$.
- (2) (Monotonicity in m) For $m < m'$ we have $\theta_m(s, t) \leq \theta_{m'}(s, t)$ for all $s, t > 0$.
- (3) (Zero at the boundary) $\theta_m(0, t) = 0$ for all $t \geq 0$ if and only if $m \leq 1$.

Proof. Properties (A1) – (A4) in Definition 2.1 are obvious from either (3.4) or (3.5).

Concavity of θ_m follows from (3.5), since an explicit computation shows that the Hessian of the function

$$F : (r, s) \mapsto \left((1-\alpha)r^{m-1} + \alpha s^{m-1} \right)^{\frac{1}{m-1}}$$

is given by

$$\text{Hess } F(r, s) = (1-\alpha)\alpha(2-m) \left((1-\alpha)r^{m-1} + \alpha s^{m-1} \right)^{\frac{1}{m-1}-2} r^{m-3} s^{m-3} \begin{pmatrix} -s^2 & rs \\ rs & -r^2 \end{pmatrix},$$

which is negative semi-definite for all $\alpha \in [0, 1]$.

Property (2) follows from the monotonicity in p of L^p -“norms”, applied on a two-point space $\{0, 1\}$ with probability measure $\mu_\alpha := (1-\alpha)\delta_0 + \alpha\delta_1$. This monotonicity follows from Jensen’s inequality and holds for negative p as well.

Finally, (3) follows from an explicit computation and (1) is obvious. \square

Let us note that θ_m is not a weight function for $m > 2$, since in this case Minkowski’s inequality for L^{m-1} -norms implies that θ_m is convex.

Corollary 3.8. *For any reversible Markov chain (\mathcal{X}, Q, π) the distance \mathcal{W}_m is finite and non-increasing in m . More precisely, for any $0 < m \leq m' \leq 2$ and $\rho_0, \rho_1 \in \mathcal{P}(\mathcal{X})$ we have*

$$\mathcal{W}_{m'}(\rho_0, \rho_1) \leq \mathcal{W}_m(\rho_0, \rho_1) < \infty.$$

Proof. This is an immediate consequence of the monotonicity of θ_m stated in Lemma 3.7 combined with the equivalent definition of the distance \mathcal{W} given in [8, Lemma 2.9]. \square

Applied to $\varphi = \varphi_m$ and $f = f_m$, Theorem 3.5 reduces to the following result.

Corollary 3.9. *Let $0 < m \leq 2$. Then the gradient flow equation of the Rényi entropy \mathcal{F}_m with respect to the metric \mathcal{W}_m is the porous medium equation $\partial_t \rho_t = \Delta \rho_t^m$.*

4. GEODESIC κ -CONVEXITY OF ENTROPY FUNCTIONALS

In this section we analyse convexity properties of entropy functionals along geodesics for non-local transportation metrics. We fix a reversible Markov chain (\mathcal{X}, Q, π) , and we fix $\varphi, f : [0, \infty) \rightarrow \mathbb{R}$ satisfying Assumptions 3.1 and 3.3. Throughout this section we set

$$\theta(r, s) = \theta_{\varphi, f}(r, s) = \frac{\varphi(r) - \varphi(s)}{f'(r) - f'(s)}. \quad (4.1)$$

Let \mathcal{W} be the associated metric, and let \mathcal{F} be the entropy functional defined in (3.3).

4.1. The Hessian of entropy functionals. We start by calculating the Hessian of an entropy functional \mathcal{F} in the Riemannian structure induced by \mathcal{W} . For a probability density $\rho \in \mathcal{P}_*(\mathcal{X})$ we introduce the function $\hat{\Delta}_\varphi \rho : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ defined by

$$\hat{\Delta}_\varphi \rho(x, y) := \partial_1 \theta(\rho(x), \rho(y)) \Delta \varphi(\rho)(x) + \partial_2 \theta(\rho(x), \rho(y)) \Delta \varphi(\rho)(y) .$$

Furthermore, for a function $\psi : \mathcal{X} \rightarrow \mathbb{R}$ we set

$$\begin{aligned} \mathcal{B}(\rho, \psi) &= \frac{1}{2} \langle \hat{\Delta}_\varphi \rho \cdot \nabla \psi, \nabla \psi \rangle_\pi - \langle \hat{\rho} \cdot \nabla \psi, \nabla(\varphi'(\rho) \Delta \psi) \rangle_\pi \\ &= \frac{1}{4} \sum_{x, y, z \in \mathcal{X}} (\psi(x) - \psi(y))^2 Q(x, y) \pi(x) \\ &\quad \times \left(\partial_1 \theta(\rho(x), \rho(y)) (\varphi(\rho(z)) - \varphi(\rho(x))) Q(x, z) \right. \\ &\quad \left. + \partial_2 \theta(\rho(x), \rho(y)) (\varphi(\rho(z)) - \varphi(\rho(y))) Q(y, z) \right) \\ &\quad - \frac{1}{2} \sum_{x, y, z \in \mathcal{X}} (\psi(x) - \psi(y)) \hat{\rho}(x, y) Q(x, y) \pi(x) \\ &\quad \times \left(\varphi'(\rho(x)) Q(x, z) (\psi(z) - \psi(x)) - \varphi'(\rho(y)) Q(y, z) (\psi(z) - \psi(y)) \right) . \end{aligned} \tag{4.2}$$

The proof of the following result is an adaptation of [8, Proposition 4.3].

Proposition 4.1. *For $\rho \in \mathcal{P}_*(\mathcal{X})$ and $\psi \in \mathbb{R}^{\mathcal{X}}$ we have*

$$\langle \text{Hess } \mathcal{F}(\rho) \nabla \psi, \nabla \psi \rangle_\rho = \mathcal{B}(\rho, \psi) .$$

Proof. Take $(\rho_t, \psi_t)_t$ satisfying the geodesic equations (2.3). Using the continuity equation and (4.1) we obtain

$$\begin{aligned} \frac{d}{dt} \mathcal{F}(\rho_t) &= - \langle f'(\rho_t), \nabla \cdot (\hat{\rho}_t \nabla \psi_t) \rangle_\pi \\ &= \langle \nabla f'(\rho_t), \hat{\rho}_t \nabla \psi_t \rangle_\pi \\ &= \langle \nabla \varphi(\rho_t), \nabla \psi_t \rangle_\pi . \end{aligned}$$

Therefore the second derivative of the entropy is given by

$$\begin{aligned} \frac{d^2}{dt^2} \mathcal{F}(\rho_t) &= \langle \nabla \partial_t \varphi(\rho_t), \nabla \psi_t \rangle_\pi + \langle \nabla \varphi(\rho_t), \nabla \partial_t \psi_t \rangle_\pi \\ &= - \langle \partial_t \varphi(\rho_t), \Delta \psi_t \rangle_\pi - \langle \Delta \varphi(\rho_t), \partial_t \psi_t \rangle_\pi . \end{aligned}$$

Using the continuity equation we obtain

$$\begin{aligned} \langle \partial_t \varphi(\rho_t), \Delta \psi_t \rangle_\pi &= \langle \varphi'(\rho_t) \partial_t \rho_t, \Delta \psi_t \rangle_\pi \\ &= - \langle \nabla \cdot (\hat{\rho}_t \nabla \psi_t), \varphi'(\rho_t) \Delta \psi_t \rangle_\pi \\ &= \langle \hat{\rho}_t \nabla \psi_t, \nabla(\varphi'(\rho_t) \Delta \psi_t) \rangle_\pi . \end{aligned}$$

Furthermore, applying the geodesic equations (2.3), the detailed balance equations (2.1), and the fact that $\partial_1\theta(r, s) = \partial_2\theta(s, r)$ by (A2), we infer that

$$\begin{aligned} \langle \Delta\varphi(\rho_t), \partial_t\psi_t \rangle_\pi &= -\frac{1}{2} \sum_{x,y,z \in \mathcal{X}} (\psi_t(x) - \psi_t(y))^2 \partial_1\theta(\rho_t(x), \rho_t(y)) \\ &\quad \times (\varphi(\rho_t(z)) - \varphi(\rho_t(x))) Q(x, y) Q(x, z) \pi(x) \\ &= -\frac{1}{4} \sum_{x,y,z \in \mathcal{X}} (\psi_t(x) - \psi_t(y))^2 \left(\partial_1\theta(\rho_t(x), \rho_t(y)) (\varphi(\rho_t(z)) - \varphi(\rho_t(x))) Q(x, z) \right. \\ &\quad \left. + \partial_2\theta(\rho_t(x), \rho_t(y)) (\varphi(\rho_t(z)) - \varphi(\rho_t(y))) Q(y, z) \right) Q(x, y) \pi(x) \\ &= -\frac{1}{2} \langle \hat{\Delta}_\varphi \rho_t \nabla \psi_t, \nabla \psi_t \rangle_\pi. \end{aligned}$$

Combining the latter three identities, we obtain

$$\begin{aligned} \langle \text{Hess } \mathcal{F}(\rho_t) \nabla \psi_t, \nabla \psi_t \rangle_{\rho_t} &= \frac{d^2}{dt^2} \mathcal{F}(\rho_t) \\ &= -\langle \hat{\rho}_t \nabla \psi_t, \nabla(\varphi'(\rho_t) \Delta \psi_t) \rangle_\pi + \frac{1}{2} \langle \hat{\Delta}_\varphi \rho_t \nabla \psi_t, \nabla \psi_t \rangle_\pi, \end{aligned}$$

which is the desired identity. \square

Since the Riemannian metric does not necessarily extend continuously to the boundary of $\mathcal{P}(\mathcal{X})$, it is not obvious that a lower bound κ on the Hessian in the interior $\mathcal{P}_*(\mathcal{X})$ implies geodesic κ -convexity in the metric space $(\mathcal{P}(\mathcal{X}), \mathcal{W})$. Nevertheless, an Eulerian argument by Daneri and Savaré [6] can be adapted to the current setting, to show that this is indeed the case. We refer to [8, Theorem 4.5] and [13, Proposition 2.1] for the details in the case where θ is the logarithmic mean.

Proposition 4.2. *For $\kappa \in \mathbb{R}$ the following assertions are equivalent:*

(1) *For every constant speed geodesic $(\rho_t)_{t \in [0,1]}$ in $(\mathcal{P}(\mathcal{X}), \mathcal{W})$ we have*

$$\mathcal{F}(\rho_t) \leq (1-t)\mathcal{F}(\rho_0) + t\mathcal{F}(\rho_1) - \frac{\kappa}{2}t(1-t)\mathcal{W}(\rho_0, \rho_1)^2. \quad (4.3)$$

(2) *For all $\bar{\rho}_0 \in \mathcal{P}(\mathcal{X})$, the solution (ρ_t) to the porous medium equation with $\rho_0 = \bar{\rho}_0$ from Proposition 3.2, satisfies the evolution variational inequality*

$$\frac{1}{2} \frac{d}{dt} \mathcal{W}(\rho_t, \sigma)^2 + \frac{\kappa}{2} \mathcal{W}(\rho_t, \sigma)^2 \leq \mathcal{F}(\sigma) - \mathcal{F}(\rho_t) \quad (4.4)$$

for all $\sigma \in \mathcal{P}(\mathcal{X})$ and a.e. $t \geq 0$.

(3) *For all $\rho \in \mathcal{P}_*(\mathcal{X})$ we have*

$$\text{Hess } \mathcal{F}(\rho) \geq \kappa.$$

(4) *For all $\rho \in \mathcal{P}_*(\mathcal{X})$ and $\psi \in \mathbb{R}^{\mathcal{X}}$ we have*

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

The evolution variational inequality (4.4) can be regarded as a definition of a (strong) notion of a gradient flow in the setting of a metric space. This inequality has been extensively studied recently, see for example [1], [?], and [6, Section 3].

4.2. Examples. In this section we will study \mathcal{W} -geodesic convexity of \mathcal{F} in some simple concrete examples. To simplify notation let us write, for $i = 1, 2$,

$$\partial_i \hat{\rho}(x, y) := \partial_i \theta(\rho(x), \rho(x)) .$$

Set

$$\kappa_Q := \sup\{\kappa \in \mathbb{R} : \mathcal{F} \text{ is } \kappa\text{-convex along } \mathcal{W}_{\varphi, f}\text{-geodesics}\} , \quad (4.5)$$

with the understanding that $\sup\{\emptyset\} = -\infty$.

4.2.1. The two-point space. We consider the two-point space $\mathcal{X} = \{a, b\}$ endowed with the Q -matrix Q defined by

$$Q(a, b) = p , \quad Q(b, a) = q ,$$

for $p, q > 0$. In this case the stationary probability measure π is given by

$$\pi(a) = \frac{q}{p+q} , \quad \pi(b) = \frac{p}{p+q} .$$

In this case we have the following characterisation of κ_Q in terms of p and q .

Proposition 4.3 (Geodesic convexity on the two-point space). *Let Q be as above. Then*

$$\kappa_Q = \frac{1}{2} \inf \left\{ p\varphi'(r) + q\varphi'(s) + \theta(r, s)(pf''(r) + qf''(s)) \right\} ,$$

where $r = \frac{p+q}{2p}(1 - \alpha)$, $s = \frac{p+q}{2p}(1 + \alpha)$, and the infimum runs over all $\alpha \in (-1, 1)$.

Proof. An explicit computation shows that

$$\mathcal{A}(\rho, \psi) = \frac{pq}{p+q} \frac{\varphi(\rho(a)) - \varphi(\rho(b))}{f'(\rho(a)) - f'(\rho(b))} (\psi(a) - \psi(b))^2 ,$$

and, taking into account that $\partial_1 \theta(r, s) = \partial_2 \theta(s, r)$,

$$\begin{aligned} \mathcal{B}(\rho, \psi) = \frac{pq}{p+q} \left[\frac{1}{2} \left(q\partial_2 \hat{\rho}(a, b) - p\partial_1 \hat{\rho}(a, b) \right) \left(\varphi(\rho(a)) - \varphi(\rho(b)) \right) \right. \\ \left. + \hat{\rho}(a, b) \left(p\varphi'(\rho(a)) + q\varphi'(\rho(b)) \right) \right] (\psi(a) - \psi(b))^2 . \end{aligned}$$

In view of Proposition 4.2, the result follows from these expressions. \square

The following result provides a simplified expression for κ_Q in the case where π is symmetric and $\theta = \theta_m$. The case $m = 1$ has already been considered in [11, Proposition 2.12].

Corollary 4.4. *If $p = q$ and $\varphi = \varphi_m$, and $f = f_m$, then*

$$\kappa_Q = \begin{cases} \frac{pm}{2} \inf_{\alpha \in (-1, 1)} \left\{ \frac{m-1}{m} \frac{(1-\alpha)^m - (1+\alpha)^m}{(1-\alpha)^{m-1} - (1+\alpha)^{m-1}} \left((1-\alpha)^{m-2} + (1+\alpha)^{m-2} \right) \right. \\ \quad \left. + (1-\alpha)^{m-1} + (1+\alpha)^{m-1} \right\} , & 0 < m \leq 2, m \neq 1 , \\ p \inf_{\alpha \in (-1, 1)} \left\{ \frac{1}{1-\alpha^2} \frac{\alpha}{\operatorname{arctanh}(\alpha)} + 1 \right\} = 2p , & m = 1 . \end{cases}$$

Proof. This follows readily from Proposition 4.3. \square

4.2.2. *The discrete circle.* As announced in Proposition 1.5, we will exhibit an instance of the discrete porous medium equation where convexity fails. For this purpose, we consider the discrete circle of length N , i.e., $\mathcal{X} = \mathbf{T}_N = \mathbb{Z}/N\mathbb{Z}$ for some $N \geq 2$. All computations below are understood modulo N . Let Q denote the discrete Laplacian, normalised so that $Q(x, y) = q$ if $|x - y| = 1$ and $Q(x, y) = 0$ otherwise. In this case π is the uniform probability given by $\pi(x) = \frac{1}{N}$ for all $x \in \mathcal{X}$.

Proposition 4.5. *Let $m = 2$, let f_m , φ_m and θ_m be as in Section 3.3, and let \mathcal{W} be the associated non-local transportation metric. Let Q be the discrete Laplacian on the discrete circle $\mathbf{T}_N = \mathbb{Z}/N\mathbb{Z}$ with $N \geq 6$ defined above. Then the functional \mathcal{F}_m is not convex along \mathcal{W} -geodesics. More precisely,*

$$\kappa_Q \leq -\frac{qN}{2}.$$

Proof. Since $m = 2$ we have the simple expressions

$$\varphi(r) = f(r) = r^2, \quad \theta(r, s) = \frac{r+s}{2}, \quad \partial_1\theta(r, s) = \partial_2\theta(r, s) = \frac{1}{2}. \quad (4.6)$$

Inserting these into (4.2), we calculate the Hessian of the Rényi entropy as

$$\begin{aligned} \mathcal{B}(\rho, \psi) &= \frac{q^2}{4N} \sum_{i=1}^N (\psi_i - \psi_{i+1})^2 \left(\rho_{i-1}^2 + \rho_{i+2}^2 + 3\rho_i^2 + 3\rho_{i+1}^2 + 8\rho_i\rho_{i+1} \right) \\ &\quad - \frac{q^2}{N} \sum_{i=1}^N (\psi_i - \psi_{i+1})(\psi_{i+1} - \psi_{i+2}) \left(2\rho_{i+1}^2 + \rho_{i+1}(\rho_i + \rho_{i+2}) \right). \end{aligned}$$

Choosing in particular

$$\psi = (\psi_1, \dots, \psi_N) = (0, 1, 2, \dots, 2, 0), \quad \rho = (\rho_1, \dots, \rho_N) = (\varepsilon, N - (N-1)\varepsilon, \varepsilon, \dots, \varepsilon)$$

we see that only three terms in the first and one term in the second sum are non-zero, and we find

$$\mathcal{B}(\rho, \psi) = -\frac{q^2N}{2} + O(\varepsilon).$$

Since $\mathcal{A}(\rho, \psi) = q + O(\varepsilon)$, the result follows. \square

Let us remark that in the case where $q = N^2$, the discrete Laplacian converges to the continuous Laplacian on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. In this limit space, it is well known that the Rényi entropy \mathcal{F}_2 is convex along 2-Wasserstein geodesics in $\mathcal{P}(\mathbb{T})$. However, the previous result shows that the behaviour at the discrete level is very different, since $\kappa_Q \leq -N^3/2$.

4.3. Consequences of geodesic convexity. In this subsection we will show that convexity of the entropy functional along \mathcal{W} -geodesics implies a contraction property for solutions to the porous medium equation as well as a number of functional inequalities for the invariant measure of the Markov chain. These results are in the spirit of the work of Otto–Villani [17]. Similar results for the heat flow associated to a finite Markov chain have been obtained in [8].

As a first result we single out the following κ -contractivity property, which is a direct consequence of the evolution variational inequality (4.4).

Proposition 4.6 (κ -Contractivity of the PME). *Suppose that $\kappa_Q \in \mathbb{R}$ and let $(\rho_t)_t, (\sigma_t)_t$ be two solutions of the discrete porous medium equation as given by Proposition 3.2. Then we have for all $t \geq 0$:*

$$\mathcal{W}(\rho_t, \sigma_t) \leq e^{-\kappa t} \mathcal{W}(\rho_0, \sigma_0) .$$

Proof. This follows from Proposition 4.2 by applying [6, Proposition 3.1] to the functional \mathcal{F} on the metric space $(\mathcal{P}(\mathcal{X}), \mathcal{W})$. \square

We introduce the following functional, which we regard as an analogue of the Fisher information:

$$\mathcal{I}(\rho) := \frac{1}{2} \sum_{x, y \in \mathcal{X}} [f'(\rho(y)) - f'(\rho(x))] [\varphi(\rho(y)) - \varphi(\rho(x))] Q(x, y) \pi(x) .$$

If f is not differentiable at 0, we use the convention that $\mathcal{I}(\rho) = +\infty$ for $\rho \notin \mathcal{P}_*(\mathcal{X})$. Using Proposition 3.4, we see that for $\rho \in \mathcal{P}_*(\mathcal{X})$ we have

$$\mathcal{I}(\rho) = \|\text{grad}_{\mathcal{W}} \mathcal{F}(\rho)\|_{\rho}^2 .$$

The significance of this functional is due to the fact that it gives the change, or dissipation, of the entropy functional along solutions of the equation $\partial_t \rho_t = \Delta \varphi(\rho_t)$, namely:

$$\frac{d}{dt} \mathcal{F}(\rho_t) = -\mathcal{I}(\rho_t) .$$

Note that in the setting of Section 3.3, the dissipation functional \mathcal{I}_m associated to φ_m and f_m is given by

$$\mathcal{I}_m(\rho) := \frac{1}{2} \frac{m}{m-1} \sum_{x, y \in \mathcal{X}} [\rho(y)^{m-1} - \rho(x)^{m-1}] [\rho(y)^m - \rho(x)^m] Q(x, y) \pi(x) .$$

As before, let $\mathbf{1} \in \mathcal{P}(\mathcal{X})$ denotes the density of the stationary distribution, which is everywhere equal to 1. It follows from the definition that $\mathcal{F}(\mathbf{1}) = f(1)$.

We introduce the following functional inequalities.

Definition 4.7. *The Markov chain (\mathcal{X}, Q, π) satisfies*

(1) *an FWI inequality with constant $\kappa \in \mathbb{R}$ if for all $\rho \in \mathcal{P}(\mathcal{X})$,*

$$\mathcal{F}(\rho) - \mathcal{F}(\mathbf{1}) \leq \mathcal{W}(\rho, \mathbf{1}) \sqrt{\mathcal{I}(\rho)} - \frac{\kappa}{2} \mathcal{W}(\rho, \mathbf{1})^2 . \quad (\text{FWI}(\kappa))$$

(2) *an entropy-dissipation inequality with constant $\lambda > 0$ if for all $\rho \in \mathcal{P}(\mathcal{X})$,*

$$\mathcal{F}(\rho) - \mathcal{F}(\mathbf{1}) \leq \frac{1}{2\lambda} \mathcal{I}(\rho) . \quad (\text{EDI}(\lambda))$$

The following result relates these inequalities to \mathcal{W} -geodesic convexity of the entropy functional \mathcal{F} . Recall the definition of κ_Q from (4.5).

Theorem 4.8. *The following assertions hold.*

- (1) *If $\kappa_Q \in \mathbb{R}$, then Q satisfies $\text{FWI}(\kappa_Q)$.*
- (2) *If $\kappa_Q > 0$, then Q satisfies $\text{EDI}(\kappa_Q)$.*

Proof. The proof follows from similar arguments as the corresponding results in [8] and uses the heuristics developed in the continuous case in [17]. However, for the convenience of the reader we give a self-contained proof here.

To prove (1), assume that $\kappa_Q \in \mathbb{R}$. It is sufficient to prove $\mathcal{FWI}(\kappa_Q)$ for $\rho \in \mathcal{P}_*(\mathcal{X})$. The inequality for general $\rho \in \mathcal{P}(\mathcal{X})$ then follows from an easy approximation argument. So fix $\rho \in \mathcal{P}_*(\mathcal{X})$ and let $(\sigma_s)_{s \in [0,1]}$ be a constant speed geodesic connecting $\rho = \sigma_0$ to $\mathbf{1} = \sigma_1$. Using Proposition 4.2 and rearranging terms in (4.3), we find

$$\mathcal{F}(\rho) - \mathcal{F}(\mathbf{1}) \leq -\frac{1}{s} \left(\mathcal{F}(\sigma_s) - \mathcal{F}(\sigma_0) \right) - \frac{\kappa_Q}{2} (1-s) \mathcal{W}(\rho, \mathbf{1})^2. \quad (4.7)$$

Taking into account the smoothness of (σ_t) near $t = 0$, we obtain the bound

$$\lim_{s \rightarrow 0} \frac{|\mathcal{F}(\sigma_s) - \mathcal{F}(\sigma_0)|}{s} \leq \|\text{grad}_{\mathcal{W}} \mathcal{F}(\sigma_0)\|_{\sigma_0} \lim_{s \rightarrow 0} \frac{\mathcal{W}(\sigma_s, \sigma_0)}{s} = \sqrt{\mathcal{I}(\rho)} \mathcal{W}(\rho, \mathbf{1}).$$

Hence, the first assertion of the theorem follows by passing to the limit $s \rightarrow 0$ in (4.7).

Let us now prove (2). Assume that $\kappa_Q > 0$. From part (1) we know that Q satisfies $\mathcal{FWI}(\kappa_Q)$. From this we derive $\text{EDI}(\kappa_Q)$ by an application of Young's inequality:

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

in which we set $x = \mathcal{W}(\rho, \mathbf{1})$, $y = \sqrt{\mathcal{I}(\rho)}$ and $c = \frac{\kappa_Q}{2}$. \square

5. GROMOV-HAUSDORFF CONVERGENCE

In this final section we discuss the convergence of discrete transportation metrics to the Wasserstein metric in a simple setting.

For $N \geq 2$ let $\mathbb{T}_N^d = (\mathbb{Z}/N\mathbb{Z})^d$ be the discrete torus in dimension $d \geq 1$, which we regard as an approximation of the continuous torus $\mathbb{T}^d = [0, 1]^d$. We consider the matrix $Q : \mathbb{T}_N^d \times \mathbb{T}_N^d \rightarrow \mathbb{R}$ defined by

$$Q(i, j) = \begin{cases} -2dN^2, & i = j, \\ N^2, & |i - j| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Note that the entries are scaled in such a way that the associated discrete Laplacian Δ_N defined by (1.2) approximates the Laplacian Δ on \mathbb{T}^d . Let \mathcal{W}_N denote the discrete transportation metric corresponding to a weight function θ_m .

The following result has already been announced in the introduction.

Theorem 5.1. *Let $d \geq 1$ and $0 < m \leq 2$. Then the metric spaces $(\mathcal{P}(\mathbb{T}_N^d), \mathcal{W}_N)$ converge to $(\mathcal{P}(\mathbb{T}^d), W_2)$ in the sense of Gromov-Hausdorff as $N \rightarrow \infty$.*

Since this result can be proved by following the argument in [9], we shall not give a complete proof here. Instead, let us point out the three crucial properties of $\theta = \theta_m$ which allow us to reapply the argument from [9]:

- (1) $\theta(t, t) = t$ for all $t \geq 0$.
- (2) θ is concave.
- (3) For all $a, b > 0$ we have

$$\frac{1}{\overline{\theta}(a, b)} - \frac{1}{\theta(a, b)} \leq \frac{(b-a)^2}{ab} \frac{1}{\overline{\theta}(a, b)},$$

where $\tilde{\theta}$ denotes the harmonic mean defined by $\tilde{\theta}(a, b) := \frac{2ab}{a+b}$.

Assumption (1) is clearly necessary, since it is already checked at the formal level that the metrics may converge to a different limit if (1) is violated. The concavity (2) is used in the proof in [9] to ensure that the discrete heat semigroup $(P_N(t))_{t \geq 0}$ contracts the distance:

$$\mathcal{W}_N(P_N(t)\rho_0, P_N(t)\rho_1) \leq \mathcal{W}_N(\rho_0, \rho_1).$$

This estimate is used in regularisation arguments. Finally, the estimate (3) enters the argument, since it turns out to be easier to compare the Wasserstein metric to the discrete transportation metric $\tilde{\mathcal{W}}_N$, which is defined using the harmonic mean. The estimate (3) can then be combined with a regularisation argument, which shows that the difference between \mathcal{W}_N and $\tilde{\mathcal{W}}_N$ becomes negligible as N gets large, provided that the weight function satisfies (3).

Let us note that these properties are satisfied for θ_m with $0 < m \leq 2$. Indeed, (1) is obvious, and concavity of θ for $0 < m \leq 2$ has been proved in Lemma 3.7. Moreover, it follows from the monotonicity of θ_m in m (proved in Lemma 3.7) that it suffices to check (3) for $m = 2$. In this case an elementary computation shows that (3) indeed holds.

Acknowledgement. The authors are grateful to the anonymous referee for his careful reading of the paper and for giving useful suggestions.

REFERENCES

- [1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [2] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.
- [3] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).
- [4] E. Carlen and J. Maas. An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker-Planck equation is gradient flow for the entropy. *arXiv:1203.5377*, 2012.
- [5] S.-N. Chow, W. Huang, Y. Li, and H. Zhou. Fokker-Planck equations for a free energy functional or Markov process on a graph. *Arch. Ration. Mech. Anal.*, 203(3):969–1008, 2012.
- [6] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.
- [7] M. Erbar. Gradient flows of the entropy for jump processes. *arXiv:1204.2190*, 2012.
- [8] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.*, 206(3):997–1038, 2012.
- [9] N. Gigli and J. Maas. Gromov-Hausdorff convergence of discrete transportation metrics. *arXiv:1207.6501*, 2011.
- [10] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [11] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.
- [12] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
- [13] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *To appear in Calc. Var. Partial Differ. Equ.*
- [14] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.
- [15] A. Mielke. Dissipative quantum mechanics using GENERIC. *To appear in Proc. of the conference “Recent Trends in Dynamical Systems”*, 2012.
- [16] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

- [17] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

UNIVERSITY OF BONN, INSTITUTE FOR APPLIED MATHEMATICS, ENDENICHER ALLEE 60, 53115 BONN, GERMANY

E-mail address: `erbar@iam.uni-bonn.de`

E-mail address: `maas@uni-bonn.de`