

HOMOGENISATION OF ONE-DIMENSIONAL DISCRETE OPTIMAL TRANSPORT

PETER GLADBACH, EVA KOPFER, JAN MAAS, AND LORENZO PORTINALE

ABSTRACT. This paper deals with dynamical optimal transport metrics defined by discretisation of the Benamou–Benamou formula for the Kantorovich metric \mathbb{W}_2 . Such metrics appear naturally in discretisations of \mathbb{W}_2 -gradient flow formulations for dissipative PDE. However, it has recently been shown that these metrics do not in general converge to \mathbb{W}_2 , unless strong geometric constraints are imposed on the discrete mesh. In this paper we prove that, in a 1-dimensional periodic setting, discrete transport metrics converge to a limiting transport metric with a non-trivial effective mobility. This mobility depends sensitively on the geometry of the mesh and on the nonlocal mobility at the discrete level. Loosely speaking, the result quantifies to what extent discrete transport can make use of microstructure in the mesh to reduce the cost of transport.

1. INTRODUCTION

In the past decades there has been intense research activity in the area of optimal transport, cf. the monographs [Vil03, Vil08, San15, PC19] for an overview of the subject. In continuous settings, a key result in the field is the *Benamou–Brenier formula* [BB00], which expresses the equivalence of static and dynamic formulations of the optimal transport problem. In discrete settings, the equivalence between static and dynamical optimal transport breaks down, and it turns out that the dynamical formulation (introduced in [Maa11, Mie11]) is essential in applications to evolution equations and discrete Ricci curvature and functional inequalities, see, e.g., [CHLZ12, EM12, Mie13, EM14, EMT15, FM16, EHMT17, EF18].

However, the limit passage from discrete dynamical transport to continuous optimal transport turns out to be nontrivial. In fact, it has been shown in [GKM18] that seemingly natural discretisations of the Benamou–Brenier formula do *not* necessarily converge to the Kantorovich distance \mathbb{W}_2 , even in one-dimensional settings. The main result in [GKM18] asserts that, for a sequence of meshes on a bounded convex domain in \mathbb{R}^d , an isotropy condition on the meshes is required to obtain the convergence of the discrete dynamical transport distances to \mathbb{W}_2 .

It remained an open question to identify the limiting behaviour of the discrete metrics in situations where the isotropy condition fails to hold. The aim of the current paper is to answer this question in the one-dimensional periodic setting.

We start by informally introducing the main objects of study in this paper and present the main result. For more formal definitions we refer to Section 2 below.

Continuous optimal transport. Let $\mathcal{P}(\mathcal{S})$ (resp. $\mathcal{M}(\mathcal{S})$) denote the set of Borel probability measures (resp. signed measures) on a Polish space (\mathcal{S}, d) . We will work on the one-dimensional torus $\mathcal{S}^1 = \mathbb{R}/\mathbb{Z}$ and use the convention that arithmetic operations are understood *modulo* 1.

Date: May 13, 2019.

The *Kantorovich metric* \mathbb{W}_2 (also known as *Wasserstein metric*) on $\mathcal{P}(\mathcal{S})$ is defined by

$$\mathbb{W}_2^2(\mu_0, \mu_1) = \inf_{\gamma \in \Gamma(\mu_0, \mu_1)} \left\{ \int_{\mathcal{S} \times \mathcal{S}} d^2(x, y) d\gamma(x, y) \right\} \quad (1.1)$$

for $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S})$. Here, $\Gamma(\mu_0, \mu_1)$ denotes the set of probability measures on $\mathcal{S} \times \mathcal{S}$ with marginals μ_0 and μ_1 respectively. For $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$ the *Benamou–Brenier formula* yields the equivalent dynamical formulation

$$\mathbb{W}_2^2(\mu_0, \mu_1) = \inf_{\mu, v} \left\{ \int_0^1 \int_{\mathcal{S}^1} \frac{|j|^2}{\mu} : \partial_t \mu + \partial_x j = 0 \right\}, \quad (1.2)$$

where the infimum runs over all curves $\mu : [0, 1] \rightarrow \mathcal{P}(\mathcal{S}^1)$ connecting μ_0 and μ_1 , and all vector fields $j : [0, 1] \times \mathcal{S}^1 \rightarrow \mathbb{R}$ satisfying the stated continuity equation. Here, $\int_0^1 \int_{\mathcal{S}^1} \frac{|j|^2}{\mu}$ is to be understood as $\int_0^1 \int_{\mathcal{S}^1} |v_t(x)|^2 d\mu_t(x) dt$ if $j \ll v$ with $\frac{dj}{d\mu} = v$, and $+\infty$ otherwise.

Discrete dynamical optimal transport. Let \mathcal{X} be a finite set endowed with a reference probability measure $\pi \in \mathcal{P}(\mathcal{X})$. Let $R : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}_+$ denote the transition rates of an irreducible continuous time Markov chain on \mathcal{X} . We assume that the *detailed balance condition* holds, i.e., $\pi(x)R(x, y) = \pi(y)R(y, x)$ for all $x, y \in \mathcal{X}$.

Let $\{\theta_{xy}\}_{x, y \in \mathcal{X}}$ be a collection of admissible means, i.e., each $\theta_{xy} : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is concave, 1-homogeneous, and satisfies $\theta(1, 1) = 1$. We assume that $\theta_{xy}(a, b) = \theta_{yx}(b, a)$ for any $a, b \geq 0$.

The *discrete dynamical transport metric* associated to (\mathcal{X}, R, π) is defined by

$$\mathcal{W}^2(m_0, m_1) = \inf_{m, J} \left\{ \frac{1}{2} \int_0^1 \sum_{x, y \in \mathcal{X}} \frac{J_t^2(x, y)}{\theta_{xy}(m_t(x)R(x, y), m_t(y)R(y, x))} dt \right\}.$$

Here the infimum runs over all curves $m : [0, 1] \rightarrow \mathcal{P}(\mathcal{X})$ connecting m_0 and m_1 , and all discrete vector fields $J : [0, 1] \rightarrow \mathcal{V}(\mathcal{X})$ satisfying the discrete continuity equation

$$\frac{d}{dt} m_t(x) + \sum_{y \in \mathcal{X}} J_t(x, y) = 0 \quad \text{for all } x \in \mathcal{X},$$

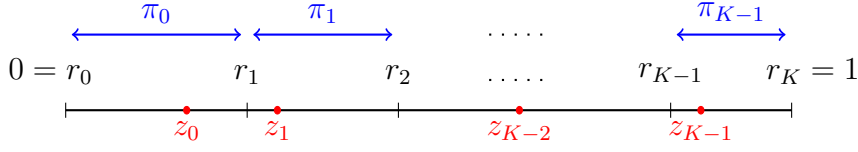
where $\mathcal{V}(\mathcal{X})$ denotes the set of all anti-symmetric functions $V : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$. The definition of \mathcal{W} is a direct analogue of (1.2) with one additional feature: between any pair of points x and y an admissible mean θ_{xy} needs to be chosen to describe the mobility.

Discrete optimal transport on 1-dimensional meshes. In this paper we consider discrete transport metrics induced by a finite volume discretisation of \mathcal{S}^1 .

Fix $0 = r_0 < \dots < r_1 < \dots < r_K = 1$ for some $K \geq 1$. We write $\pi_k := r_{k+1} - r_k$ and $A_k := [r_k, r_{k+1})$, so that $\mathcal{T} := \{A_k\}_{k=0}^{K-1}$ is a partition of \mathcal{S}^1 into disjoint half-open intervals. We also consider a sequence of points $\{z_k\}_{k=0}^{K-1}$ such that each z_k lies in the interior of A_k . The distance between z_k and $z_{k'}$ in \mathcal{S}^1 will be denoted by $d_{kk'}$. Here and below we will often perform calculations *modulo* K .

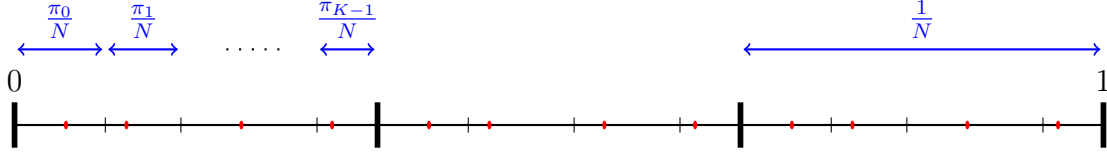
We endow the discrete state space \mathcal{T} with the natural reference measure $\pi \in \mathcal{P}(\mathcal{T})$ given by $\pi(A_k) = \pi_k$. The main object of study in this paper is the transport metric $\mathcal{W}_{\mathcal{T}}$ on $\mathcal{P}(\mathcal{T})$ induced by the Markov transition rates on \mathcal{T} given by

$$R(A_k, A_{k'}) := R_{kk'} := \frac{1}{\pi_k d_{kk'}}$$


 FIGURE 1. The mesh \mathcal{T} on \mathcal{S}^1 .

if $|k - k'| = 1$, and $R_{kk'} = 0$ otherwise. Then we have the detailed balance condition $\pi_k R_{kk'} = \pi_{k'} R_{k'k}$. The rates are chosen to ensure that solutions to the discrete diffusion equation (i.e., the Kolmogorov forward equation associated to the Markov chain given by R) converge to solutions of the diffusion equation $\partial_t \mu = \partial_x^2 \mu$ in the limit of vanishing mesh size [EGH00]. A gradient flow approach in one dimension can be found in [DL15].

The periodic setting. For any mesh \mathcal{T} as above and $N \geq 1$ one can construct a periodic mesh \mathcal{T}_N with NK cells $A_{n,k}$ by concatenating N rescaled copies of \mathcal{T} .


 FIGURE 2. The mesh \mathcal{T}_N on \mathcal{S}^1 .

We then consider the transport metric $\mathcal{W}_N := \mathcal{W}_{\mathcal{T}_N}$ on $\mathcal{P}(\mathcal{T}_N)$ as defined above. Explicitly, we have

$$\mathcal{W}_N^2(m_0, m_1) = \inf_{m, J} \left\{ \frac{1}{N} \int_0^1 \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} d_{k,k+1} \frac{J_t^2(n; k, k+1)}{\theta_{k,k+1} \left(\frac{Nm_t(n;k)}{\pi_k}, \frac{Nm_t(n;k+1)}{\pi_{k+1}} \right)} dt \right\},$$

where the infimum runs over all curves $m : [0, 1] \rightarrow \mathcal{P}(\mathcal{T}_N)$ and $J : [0, 1] \rightarrow \mathcal{V}(\mathcal{T}_N)$ satisfying the discrete continuity equation

$$\frac{d}{dt} m_t(n; k) + J_t(n; k, k+1) - J_t(n; k-1, k) = 0$$

for all $n = 0, \dots, N-1$ and $k = 0, \dots, K-1$. Here we use the shorthand notation $m(n; k) = m(A_{n,k})$ and $J(n; k, k+1) = J(A_{n,k}, A_{n,k+1})$. The main goal of this paper is to analyse the limiting behaviour of \mathcal{W}_N as $N \rightarrow \infty$.

The discrete-to-continuous limit. The first convergence result for discrete dynamical transport metrics (in the sense of Gromov–Hausdorff) was obtained in [GM13]. There it is shown that the discrete transport metric associated to the cubic mesh on the d -dimensional torus converges to \mathbb{W}_2 in the limit of vanishing mesh size.

The limiting behaviour of discrete dynamical transport metrics on more general meshes turns out to be a delicate issue. In fact, it follows from the multi-dimensional results in [GKM18] that the discrete transport metrics \mathcal{W}_N converge to \mathbb{W}_2 if and only if the means $\theta_{kk'}$ are carefully chosen to satisfy an appropriate “balance condition” that reflects the geometry of the mesh \mathcal{T} . In our one-dimensional periodic setting, these results imply

that \mathcal{W}_N converges to \mathbb{W}_2 if and only if there exist constants $\lambda_{k,k+1}, s \in (0, 1)$ such that the following conditions hold for $k = 0, \dots, K-1$:

$$\begin{aligned} r_{k+1} &= \lambda_{k,k+1} z_{k+1} + (1 - \lambda_{k,k+1}) z_k + s, \\ \theta_{k,k+1}(a, b) &\leq \lambda_{k,k+1} a + (1 - \lambda_{k,k+1}) b \quad \text{for any } a, b \geq 0. \end{aligned} \quad (1.3)$$

Thus, to fulfill this condition, the asymmetry of the means $\theta_{k,k+1}$ should reflect the relative location of the points z_k, r_{k+1} , and z_{k+1} . We refer to Section 4 below for a full discussion.

The main contribution of the current paper is the identification of the limiting behaviour of \mathcal{W}_N in the general one-dimensional periodic setting, without assuming (1.3). To state the result, we introduce the canonical projection operator $P_{\mathcal{T}} : \mathcal{P}(\mathcal{S}^1) \rightarrow \mathcal{P}(\mathcal{T})$ defined by

$$(P_{\mathcal{T}}\mu)(\{A\}) = \mu(A) \quad (1.4)$$

for $\mu \in \mathcal{P}(\mathcal{S}^1)$ and $A \in \mathcal{T}$. For brevity we write $P_N := P_{\mathcal{T}_N}$.

The following result asserts that \mathcal{W}_N converges to a Kantorovich metric with an effective mobility determined by the geometry of the mesh and by the choice of the means $\theta_{k,k+1}$.

Theorem 1.1 (Main result). *Fix a mesh \mathcal{T} on \mathcal{S}^1 , and consider the induced periodic meshes \mathcal{T}_N for $N \geq 1$. For any $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have*

$$\lim_{N \rightarrow \infty} \mathcal{W}_N(P_N \mu_0, P_N \mu_1) = \sqrt{c^*(\theta, \mathcal{T})} \mathbb{W}_2(\mu_0, \mu_1),$$

where

$$c^*(\theta, \mathcal{T}) := \inf \left\{ \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}} \right)} : m \in \mathcal{P}(\mathcal{T}) \right\}. \quad (1.5)$$

Moreover, as $N \rightarrow \infty$ we have Gromov–Hausdorff convergence of metric spaces:

$$(\mathcal{P}(\mathcal{T}_N), \mathcal{W}_N) \rightarrow (\mathcal{P}(\mathcal{S}^1), \sqrt{c^*(\theta, \mathcal{T})} \mathbb{W}_2).$$

Remark 1.2 (Isotropic case). If the compatibility conditions (1.3) are satisfied, it follows that $c^*(\theta, \mathcal{T}) = 1$, and we recover the result of [GKM18]; cf. Section 4.

Remark 1.3 (Convergence of gradient flows). We stress that the limiting behaviour at the level of the transport metrics is in stark contrast with the convergence results of the level of the gradient flow equation. Indeed, consider the discrete transport metric \mathcal{W}_N in the case where each $\theta_{k,k+1}$ is equal to the *logarithmic mean* $\theta_{\log}(a, b) = \int_0^1 a^{1-s} b^s ds$. Then the discrete diffusion equation is the gradient flow equation in $(\mathcal{P}(\mathcal{T}_N), \mathcal{W}_N)$ for the relative entropy with respect to the natural reference measure π_N ; cf. [CHLZ12, Maa11, Mie11]. Similarly, the continuous diffusion equation is the gradient flow in $(\mathcal{P}(\mathcal{S}^1), \mathbb{W}_2)$ for the relative entropy with respect to the Lebesgue measure on \mathcal{S}^1 [JKO98]. Nevertheless, our main result shows that the discrete transport metrics \mathcal{W}_N converge to a limiting metric that is different from \mathbb{W}_2 , unless the mesh is equidistant. For a systematic study of convergence of gradient flow structures we refer to [Mie16b, Mie16a, DFM18].

Remark 1.4 (Convergence on geometric graphs). A convergence result for discrete transport distances on a large class of geometric graphs associated to point clouds on the d -dimensional torus has been obtained in [GT17]. This result applies in particular to iid points sampled from the uniform distribution on the torus. As the results in that paper

apply to sequences of graphs with increasing degree, they do not overlap with the results obtained here.

Heuristics. We briefly sketch a non-rigorous argument that makes Theorem 1.1 plausible. For this purpose we consider a smooth solution to the continuity equation $\partial_t \mu + \partial_x j = 0$, and fix $\alpha \in \mathcal{T}$. Suppressing the time variable, we define a discrete measure m that assigns mass $m(k) := \alpha(k) \mu\left(\left[\frac{n}{N}, \frac{n+1}{N}\right)\right)$ to each cell $A_{n,k}$ in \mathcal{T}_N . This ensures that each interval of the form $\left[\frac{n}{N}, \frac{n+1}{N}\right)$ receives the same mass at the discrete and the continuous level, but within each such interval, the measure α introduces discrete density oscillations.

Let J be the discrete momentum vector field that solves the continuity equation for m . If this vector field is sufficiently regular, we may estimate the discrete energy by

$$\begin{aligned} & \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} d_{k,k+1} \frac{J_t^2(n; k, k+1)}{\theta_{k,k+1} \left(\frac{Nm_t(n;k)}{\pi_k}, \frac{Nm_t(n;k+1)}{\pi_{k+1}} \right)} \\ & \approx \frac{1}{N} \sum_{n=0}^{N-1} \frac{J_t^2(n; 0, 1)}{\mu\left(\left[\frac{n}{N}, \frac{n+1}{N}\right)\right)} \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{\alpha(k)}{\pi_k}, \frac{\alpha(k+1)}{\pi_{k+1}} \right)} \approx c^*(\theta, \mathcal{T}) \int \frac{|j|^2}{\mu}, \end{aligned}$$

after minimisation over $\alpha \in \mathcal{P}(\mathcal{T})$. We thus recover the continuous energy appearing in the Benamou–Brenier formula up to a multiplicative correction, which indeed suggests our main result.

A rigorous argument based on this heuristics clearly requires suitable spatial regularity results for m and J . Indeed, we will show in Section 5 below that any discrete curve can be approximated by a curve of similar energy, which enjoys good Lipschitz bounds for J as well as good Lipschitz bounds for m “up to oscillations within each cell”.

Organisation of the paper. In Section 2 we collect the basic definitions and preliminary results that are used in this paper. In Section 3 we give a simple approach to some of the main convergence results, which only applies in the special case where \mathcal{T} consists of exactly 2 cells. In Section 3 we analyse the formula (1.5) for the effective mobility $c^*(\theta, \mathcal{T})$ and discuss its relation to the geometric conditions from [GKM18].

The bulk of the proof of the main result is contained in Sections 5 and 6, which deal with the lower and upper bounds for \mathcal{W}_N respectively. The key results in these sections are Theorems 5.4 and 6.6. In Section 7 we finish the proof of the main result by proving the Gromov–Hausdorff convergence.

2. PRELIMINARIES

2.1. Continuous optimal transport on \mathcal{S}^1 . For $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$, let $\mathbb{CE}(\mu_0, \mu_1)$ denote the set of all distributional solutions to the continuity equation

$$\partial_t \mu + \partial_x j = 0 \tag{2.1}$$

with boundary conditions $\mu_t|_{t=0} = \mu_0$ and $\mu_t|_{t=1} = \mu_1$. This means that $(\mu_t)_t$ and $(j_t)_t$ are Borel families of measures in $\mathcal{P}(\mathcal{S}^1)$ and $\mathcal{M}(\mathcal{S}^1)$ respectively, such that $\int_0^1 |j_t|(\mathcal{S}^1) dt < \infty$, and (2.1) holds in the sense that

$$\int_0^1 \int_{\mathcal{S}^1} \partial_t \xi(t, x) d\mu_t(x) dt + \int_0^1 \int_{\mathcal{S}^1} \partial_x \xi(t, x) dj_t(x) dt = 0,$$

for any test function $\xi \in C_c^1(\mathcal{S}^1 \times (0, 1))$. For $\mu \in \mathcal{P}(\mathcal{S}^1)$ and $j \in \mathcal{M}(\mathcal{S}^1)$ we set

$$\mathbb{A}(\mu, j) = \int_{\mathcal{S}^1} \left| \frac{dj}{d\mu} \right|^2 d\mu ,$$

if $j \ll \mu$, and $\mathbb{A}(\mu, j) = +\infty$ otherwise. With this notation, the Benamou–Brenier formula [BB00] asserts that

$$\mathbb{W}_2^2(\mu_0, \mu_1) = \inf \left\{ \int_0^1 \mathbb{A}(\mu_t, j_t) dt : (\mu_t, j_t)_t \in \mathbb{CE}(\mu_0, \mu_1) \right\} , \quad (2.2)$$

see, e.g., [AGS08, Lemma 8.1.3] for more details.

2.2. Discrete optimal transport on one-dimensional meshes. As in Section 1, we fix a mesh $\mathcal{T} = \{A_k\}_{k=0}^{K-1}$ on \mathcal{S}^1 , and use the notation $r_k, \pi_k, z_k, d_{kk'}$. The set $\mathcal{V}(\mathcal{T})$ of discrete vector fields is naturally identified with the set of real-valued functions on $\{(k, k+1)\}_{k=0}^{K-1}$.

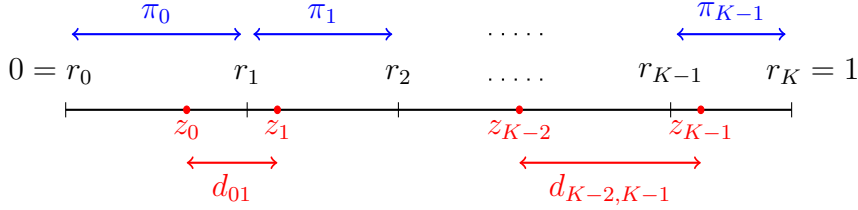


FIGURE 3. The mesh \mathcal{T} on \mathcal{S}^1 .

Definition 2.1 (Discrete continuity equation). *A pair $(m_t, J_t)_{t \in [0,1]}$ is said to satisfy the discrete continuity equation if*

- (i) $m : [0, 1] \rightarrow \mathcal{P}(\mathcal{T})$ is continuous;
- (ii) $J : [0, 1] \rightarrow \mathcal{V}(\mathcal{T})$ is locally integrable;
- (iii) the continuity equation holds in the sense of distributions:

$$\frac{d}{dt} m_t(k) + J_t(k, k+1) - J_t(k-1, k) = 0 \quad \text{for all } k = 0, \dots, K-1 . \quad (2.3)$$

We write $\mathbb{CE}_{\mathcal{T}}(m_0, m_1)$ to denote the collection of pairs $(m_t, J_t)_{t \in [0,1]}$ satisfying $m|_{t=0} = m_0$ and $m|_{t=1} = m_1$.

Definition 2.2 (Admissible mean). *An admissible mean is a function $\theta : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ that is concave, 1-homogeneous, and satisfies $\theta(1, 1) = 1$.*

Note that we do *not* impose that admissible means are symmetric.

Definition 2.3 (Discrete dynamical transport distance). *Let \mathcal{T} be a mesh on \mathcal{S}^1 , and $\{\theta_{k,k+1}\}_{k=0}^{K-1}$ be a family of admissible means.*

- (1) The energy functional $\mathcal{A}_{\mathcal{T}} : \mathcal{P}(\mathcal{T}) \times \mathcal{V}(\mathcal{T}) \rightarrow \mathbb{R} \cup \{+\infty\}$ is given by

$$\mathcal{A}_{\mathcal{T}}(m, J) = \sum_{k=0}^{K-1} d_{k,k+1} f_{k,k+1} \left(\frac{m(k)}{\pi_k}, \frac{m(k+1)}{\pi_{k+1}}, J(k, k+1) \right) ,$$

where $f_{k,k+1}(\rho, \tilde{\rho}, J) = F(\theta_{k,k+1}(\rho, \tilde{\rho}), J)$, and

$$F(\rho, J) = \begin{cases} \frac{J^2}{\rho} & \text{if } \rho > 0, \\ 0 & \text{if } \rho = 0 \text{ and } J = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

(2) The discrete dynamical transportation distance between $m_0, m_1 \in \mathcal{P}(\mathcal{T})$ is given by

$$\mathcal{W}_{\mathcal{T}}(m_0, m_1) = \inf \left\{ \sqrt{\int_0^1 \mathcal{A}_{\mathcal{T}}(m_t, J_t) dt} : (m_t, J_t)_t \in \text{CE}_{\mathcal{T}}(m_0, m_1) \right\}.$$

In the sequel we apply these definitions to the periodic meshes \mathcal{T}_N defined in Section 1. We will then simply write \mathcal{A}_N and \mathcal{W}_N as a shorthand for $\mathcal{A}_{\mathcal{T}_N}$ and $\mathcal{W}_{\mathcal{T}_N}$ respectively.

2.3. A priori bounds. In this section we collect some coarse bounds that will be useful in the sequel. To compare discrete and continuous measures, we consider the canonical embedding $\iota_{\mathcal{T}} : \mathcal{P}(\mathcal{T}) \rightarrow \mathcal{P}(\mathcal{S}^1)$ defined by

$$\iota_{\mathcal{T}} m = \sum_{k=0}^{K-1} m_k \mathcal{U}_{A_k} \quad \text{for } m \in \mathcal{P}(\mathcal{T}),$$

where \mathcal{U}_{A_k} denotes the uniform probability measure on A_k . Note that $\iota_{\mathcal{T}}$ is the right-inverse of the projection map $P_{\mathcal{T}}$ defined by (1.4). We will often write $\iota_N = \iota_{\mathcal{T}_N}$ for brevity.

The following notion of mesh regularity can be found in a multi-dimensional setting in [EGH00, Section 3.1.2].

Definition 2.4 (ζ -regularity). *Let $\zeta \in (0, 1]$. We say that a mesh \mathcal{T} is ζ -regular, if $\zeta < \frac{\min_k \{z_k - r_k, r_{k+1} - z_k\}}{\max_k \pi_k}$ for all $k = 0, \dots, K - 1$.*

A mesh \mathcal{T} is ζ -regular if and only if the ball of radius $\zeta \max_k \pi_k$ around z_k is contained in the interior of the cell A_k for each k . Clearly, any mesh \mathcal{T} on \mathcal{S}^1 is ζ -regular for some $\zeta \in (0, 1]$.

Remark 2.5. If \mathcal{T} is ζ -regular for some $\zeta \in (0, 1]$, then each \mathcal{T}_N is ζ -regular as well.

Let $[\mathcal{T}]$ denote the size of the mesh, i.e., the maximal diameter of its cells:

$$[\mathcal{T}] := \max\{\pi_k : k = 0, \dots, K - 1\}.$$

The following result provide a coarse upper bound for \mathcal{W}_N in terms of \mathbb{W}_2 .

Proposition 2.6 (Coarse upper bound for $\mathcal{W}_{\mathcal{T}}$). *Let $\zeta \in (0, 1]$. There exists a constant $C < \infty$ depending only on ζ such that for any ζ -regular mesh \mathcal{T} of \mathcal{S}^1 and all $m_0, m_1 \in \mathcal{P}(\mathcal{T})$ we have*

$$\mathcal{W}_{\mathcal{T}}(m_0, m_1) \leq C \left(\mathbb{W}_2(\iota_{\mathcal{T}} m_0, \iota_{\mathcal{T}} m_1) + [\mathcal{T}] \right). \quad (2.4)$$

Proof. This result has been proved in [GKM18, Lemma 3.3] for convex domains in \mathbb{R}^d ; the proof on \mathcal{S}^1 proceeds *mutatis mutandi*. \square

The following result provides a coarse bound in the opposite direction.

Proposition 2.7 (Coarse lower bound for $\mathcal{W}_{\mathcal{T}}$). *Fix $\delta, \zeta \in (0, 1)$. There exists a constant $C < \infty$ depending only on δ and ζ , such that for any ζ -regular mesh \mathcal{T} on \mathcal{S}^1 , and for any solution $(m_t, J_t)_t$ to the discrete continuity equation (2.3) satisfying $\delta \leq \frac{m_t(k)}{\pi_k} \leq \delta^{-1}$ for all $t \in [0, 1]$ and $k = 0, \dots, K-1$, we have*

$$\mathbb{W}_2^2(\iota_{\mathcal{T}}m_0, \iota_{\mathcal{T}}m_1) \leq C \int_0^1 \mathcal{A}_{\mathcal{T}}(m_t, J_t) dt . \quad (2.5)$$

Proof. We define

$$\mu_t = \iota_{\mathcal{T}}m_t , \quad j_t(x) = \frac{r_{k+1} - x}{\pi_k} J_t(k-1, k) + \frac{x - r_k}{\pi_k} J_t(k, k+1) ,$$

for $x \in A_k$. It follows that (μ_t, j_t) solves the continuous continuity equation. Moreover,

$$\begin{aligned} \mathbb{A}(\mu_t, j_t) &= \sum_{k=0}^{K-1} \frac{\pi_k}{m_t(k)} \int_{r_k}^{r_{k+1}} \left(\frac{r_{k+1} - x}{\pi_k} J_t(k-1, k) + \frac{x - r_k}{\pi_k} J_t(k, k+1) \right)^2 dx \\ &\leq \frac{1}{2} \sum_{k=0}^{K-1} \frac{\pi_k^2}{m_t(k)} \left(J_t^2(k-1, k) + J_t^2(k, k+1) \right) \\ &= \frac{1}{2} \sum_{k=0}^{K-1} J_t^2(k, k+1) \left(\frac{\pi_k^2}{m_t(k)} + \frac{\pi_{k+1}^2}{m_t(k+1)} \right) . \end{aligned}$$

Write $\rho_t(k) = \frac{m_t(k)}{\pi_k}$. In view of the bounds on $m_t(k)$, we have

$$\begin{aligned} \theta_{k,k+1}(\rho_t(k), \rho_t(k+1)) &\leq \max\{\rho_t(k), \rho_t(k+1)\} \\ &\leq \delta^{-2} \min\{\rho_t(k), \rho_t(k+1)\} \leq 2\delta^{-2} \left(\frac{1}{\rho_t(k)} + \frac{1}{\rho_t(k+1)} \right)^{-1} . \end{aligned}$$

Since $2\zeta[\mathcal{T}] \leq d_{k,k+1}$, we have

$$\frac{\pi_k^2}{m_t(k)} + \frac{\pi_{k+1}^2}{m_t(k+1)} \leq [\mathcal{T}] \left(\frac{\pi_k}{m_t(k)} + \frac{\pi_{k+1}}{m_t(k+1)} \right) \leq \frac{1}{\zeta\delta^2} \frac{d_{k,k+1}}{\theta_{k,k+1}(\rho_t(k), \rho_t(k+1))} .$$

It follows that

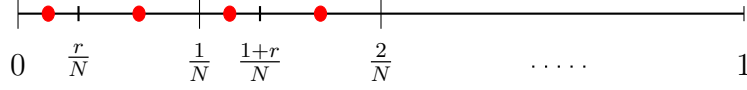
$$\mathbb{A}(\mu_t, j_t) \leq \frac{1}{2\zeta\delta^2} \sum_{k=0}^{K-1} d_{k,k+1} \frac{J_t^2(k, k+1)}{\theta_{k,k+1}(\rho_t(k), \rho_t(k+1))} = \frac{1}{2\zeta\delta^2} \mathcal{A}_{\mathcal{T}}(m_t, J_t) ,$$

which implies the result with $C = \frac{1}{2\zeta\delta^2}$. \square

3. A SIMPLE PROOF OF THE LOWER BOUND IN THE 2-PERIODIC CASE

In this section we focus on the simplest non-trivial periodic setting, which corresponds to taking $K = 2$ in Figure 1. In this setting we present a short proof of the lower bound in Theorem 1.1 by connecting the problem to known results from [GM13, GKM18]. This approach does not appear to generalise to $K \geq 3$.

We fix a parameter $r \in (0, 1)$, and consider the mesh \mathcal{T} with $K = 2$, and $r_0 = 0$, $r_1 = r$, and $r_2 = 1$. The points $z_0 = \frac{r}{2}$ and $z_1 = \frac{r+1}{2}$ are chosen to be the midpoints of the cells, so that $d_{01} = d_{12} = \frac{1}{2}$.

FIGURE 4. A 2-periodic mesh \mathcal{T}_N on \mathcal{S}^1 .

Throughout this section we make the standing assumption that $\theta_{01} = \theta_{21}$, and we simply write $\theta := \theta_{01}$. This implies that the constant $c^*(\theta, r) := c^*(\theta, \mathcal{T})$ is given by

$$c^*(\theta, r) = \inf_{\alpha \in [0,1]} \frac{1}{\theta\left(\frac{\alpha}{r}, \frac{1-\alpha}{1-r}\right)}. \quad (3.1)$$

The notation $m(k) = m(A_k)$ allows us to canonically identify measures on \mathcal{T}_N with measures on the equidistant mesh corresponding to $r = \frac{1}{2}$. We write $\mathcal{T}_{r,N}$ to emphasise the dependence of \mathcal{T}_N on r , and write $\mathcal{A}_{r,N}^\theta$ and $\mathcal{W}_{r,N}^\theta$ to denote the corresponding energy and metric. The cells in $\mathcal{T}_{r,N}$ will be labeled $0, \dots, 2N - 1$.

3.1. Lower bound. The following lemma compares the discrete transport metric on the mesh $\mathcal{T}_{r,N}$ with the corresponding quantity on the equidistant mesh $\mathcal{T}_{\frac{1}{2},N}$.

Lemma 3.1. *Let $r \in (0, 1)$ and $N \geq 1$. For any $m_0, m_1 \in \mathcal{P}(\mathcal{T}_{r,N})$ we have*

$$\mathcal{W}_{r,N}^\theta(m_0, m_1) \geq \sqrt{c^*(\theta, r)} \mathcal{W}_{\frac{1}{2},N}^{\theta_a}(m_0, m_1) \quad (3.2)$$

where θ_a denotes the arithmetic mean.

Proof. Note that

$$\mathcal{A}_{r,N}^\theta(m, J) = \frac{1}{2N^2} \sum_{n=0}^{N-1} \left\{ \frac{(J(2n, 2n+1))^2}{\theta\left(\frac{m(2n)}{r}, \frac{m(2n+1)}{1-r}\right)} + \frac{(J(2n-1, 2n))^2}{\theta\left(\frac{m(2n)}{r}, \frac{m(2n-1)}{1-r}\right)} \right\}.$$

The key observation is that the mean θ of the densities $\frac{m(2n)}{r}$ and $\frac{m(2n\pm 1)}{1-r}$ on the mesh $\mathcal{T}_{r,N}$ can be estimated in terms of the *arithmetic mean* θ^a of the corresponding densities $\frac{m(2n)}{1/2}$ and $\frac{m(2n\pm 1)}{1/2}$ on the symmetric mesh $\mathcal{T}_{\frac{1}{2},N}$. Indeed, as we can write

$$m(2n) = \alpha_n^\pm (m(2n) + m(2n \pm 1)) \quad \text{and} \quad m(2n \pm 1) = (1 - \alpha_n^\pm) (m(2n) + m(2n \pm 1))$$

for some $\alpha_n^\pm \in [0, 1]$, the 1-homogeneity of θ yields

$$\begin{aligned} \theta\left(\frac{m(2n)}{r}, \frac{m(2n \pm 1)}{1-r}\right) &= (m(2n) + m(2n \pm 1)) \theta\left(\frac{\alpha_n^\pm}{r}, \frac{1 - \alpha_n^\pm}{1-r}\right) \\ &\leq \theta^a\left(\frac{m(2n)}{1/2}, \frac{m(2n \pm 1)}{1/2}\right) \frac{1}{c^*(\theta, r)}. \end{aligned}$$

Consequently,

$$\mathcal{A}_{r,N}^\theta(m, J) \geq c^*(\theta, r) \mathcal{A}_{\frac{1}{2},N}^{\theta_a}(m, J).$$

As the continuity equation does not depend on r , this implies the result. \square

The sought lower bound for $\mathcal{W}_{r,N}^\theta$ can now be easily obtained.

Corollary 3.2. *Fix $r \in (0, 1)$. For any $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have*

$$\liminf_{N \rightarrow \infty} \mathcal{W}_{r,N}(P_N \mu_0, P_N \mu_1) \geq \sqrt{c^*(\theta, r)} \mathbb{W}_2(\mu_0, \mu_1).$$

Proof. This follows by applying Lemma 3.1 to the measures $m_i := P_N \mu_i$ and using the known convergence result for symmetric meshes [GM13, GKM18], which asserts that

$$\lim_{N \rightarrow \infty} \mathcal{W}_{\frac{1}{2}, N}^{\theta_a}(P_N \mu_0, P_N \mu_1) = \mathbb{W}_2(\mu_0, \mu_1) .$$

□

For proving the corresponding upper bound

$$\limsup_{N \rightarrow \infty} \mathcal{W}_{r, N}(P_N \mu_0, P_N \mu_1) \leq \sqrt{c^*(\theta, r)} \mathbb{W}_2(\mu_0, \mu_1) ,$$

the 2-periodic setting does not appear to offer conceptual simplifications compared to the general K -periodic setting. Therefore, we will directly treat the K -periodic setting in Section 6.

3.2. Examples. We finish this section by explicitly computing the value of $c^*(\theta, r)$ in a number of cases. We write

$$f_{\theta, r}(\alpha) := \theta\left(\frac{\alpha}{r}, \frac{1-\alpha}{1-r}\right) , \quad \text{so that} \quad c^*(\theta, r) = \inf_{\alpha \in (0,1)} \frac{1}{f_{\theta, r}(\alpha)} .$$

Example 3.3 (θ is r -balanced). Suppose that $\theta(a, b) \leq ra + (1-r)b$ for any $a, b \geq 0$ (i.e., θ is r -balanced in the sense of Definition 4.6 below). Applying this inequality to $a = \frac{\alpha}{r}$ and $b = \frac{1-\alpha}{1-r}$ we immediately obtain $f_{\theta, r}(\alpha) \leq 1$ for all $\alpha \in [0, 1]$. Since $f_{\theta, r}(r) = \theta(1, 1) = 1$ by assumption, it follows that

$$c^*(\theta, r) = \frac{1}{f_{\theta, r}(r)} = 1 .$$

Example 3.4 (Geometric mean). Let $\theta(a, b) = \sqrt{ab}$. Then $f_{\theta, r}(\alpha) = \sqrt{\frac{\alpha(1-\alpha)}{r(1-r)}}$ is uniquely maximised at $\alpha = \frac{1}{2}$, and we obtain

$$c^*(\theta, r) = 2\sqrt{r(1-r)} .$$

Note that the fact that $\alpha^* = \frac{1}{2}$ means that the mass is equally distributed among large and small cells, irrespectively of the value of r . Thus, there will be no oscillations for the optimal discrete measures; however, this means that oscillations at the level of the density do occur.

Example 3.5 (Harmonic mean). Let $\theta(a, b) = \frac{2ab}{a+b}$. In this case we have $g(\alpha) := \frac{1}{f_{\theta, r}(\alpha)} = \frac{1}{2} \left(\frac{1-r}{1-\alpha} + \frac{r}{\alpha} \right)$, and $g'(\alpha) = \frac{1-r}{2(1-\alpha)^2} - \frac{r}{2\alpha^2}$. It follows that g' vanishes at $\alpha^* = \frac{\sqrt{r}}{\sqrt{r} + \sqrt{1-r}}$, which is indeed the unique minimiser of g . Consequently,

$$c^*(\theta, r) = g(\alpha^*) = \frac{1}{2} (\sqrt{r} + \sqrt{1-r})^2 .$$

Example 3.6 (Arithmetic mean). Let $\theta(a, b) = \frac{a+b}{2}$. Then $f_{\theta, r}(\alpha) = \frac{1}{2} \left(\frac{1-\alpha}{1-r} + \frac{\alpha}{r} \right)$ is affine in α . If $r < \frac{1}{2}$ (resp. $r > \frac{1}{2}$), the maximum is attained at $\alpha^* = 1$ (resp. $\alpha^* = 0$). In both cases, this means that all the mass will be assigned to the small cells. It follows that

$$c^*(\theta, r) = 2 \min\{r, 1-r\} .$$

Example 3.7 (Minimum). Let $\theta(a, b) = \min\{a, b\}$. In this case, $f_{\theta,r}(\alpha) = \min\{\frac{1-\alpha}{1-r}, \frac{\alpha}{r}\}$ is uniquely maximised at $\alpha^* = r$. This means that the assigned mass is proportional to the size of the cells, hence there are no oscillations at the level at the density. We find

$$c^*(\theta, r) = \frac{1}{f_{\theta,r}(\alpha^*)} = 1 .$$

4. ANALYSIS OF THE EFFECTIVE MOBILITY

In this section we investigate some basic properties of the effective mobility $c^*(\theta, \mathcal{T})$ defined in (1.5), and relate its value to certain geometric properties of the mesh \mathcal{T} that have been considered in [GKM18]. Recall:

$$c^*(\theta, \mathcal{T}) := \inf \left\{ \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}} \right)} : m \in \mathcal{P}(\mathcal{T}) \right\} . \quad (4.1)$$

We start with a simple observation.

Proposition 4.1. *For any mesh \mathcal{T} on \mathcal{S}^1 and any family of means $\theta = \{\theta_{k,k+1}\}_{k=0}^{K-1}$ we have $c^*(\theta, \mathcal{T}) \leq 1$.*

Proof. This follows by using the competitor $m_k = \pi_k$ in (4.1). □

In view of this result, Theorem 1.1 implies the upper bound

$$\limsup_{N \rightarrow \infty} \mathcal{W}_N(P_N \mu_0, P_N \mu_1) \leq \mathbb{W}_2(\mu_0, \mu_1) ,$$

which had already been proved in [GKM18] using different methods.

Proposition 4.2. *The infimum in (4.1) is attained.*

Proof. This follows by lower-semicontinuity and compactness. □

In the remainder of this section we shall investigate under which conditions on θ and \mathcal{T} we have $c^*(\theta, \mathcal{T}) = 1$. For this purpose, we consider two geometric conditions:

Definition 4.3 (Geometric conditions on the mesh). *Fix $\{\lambda_{k,k+1}\}_{k=0}^{K-1} \in [0, 1]^K$, and set $\lambda_{k+1,k} = 1 - \lambda_{k,k+1}$. We say that a mesh $\mathcal{T} = \mathcal{T}_{\pi,z}$ on \mathcal{S}^1 satisfies*

(1) *the center-of-mass condition with parameters $\{\lambda_{k,k+1}\}_{k=0}^{K-1}$ if, for all k ,*

$$r_{k+1} = \lambda_{k+1,k} z_k + \lambda_{k,k+1} z_{k+1} ; \quad (4.2)$$

(2) *the isotropy condition with parameters $\{\lambda_{k,k+1}\}_{k=0}^{K-1}$ if, for all k ,*

$$\pi_k = \lambda_{k,k-1} d_{k-1,k} + \lambda_{k,k+1} d_{k,k+1} . \quad (4.3)$$

Both of these conditions have been studied for meshes on bounded convex domains in \mathbb{R}^d in [GKM18]. The center-of-mass condition asserts that the center of mass of the cell interfaces lie on the line segment connecting the support points of the respective cells. In dimensions $d \geq 2$, this condition poses a strong geometric condition on the mesh. However, in our one-dimensional context, the condition is always satisfied, for a unique choice of the parameters $\{\lambda_{k,k+1}\}_k$. The isotropy condition is weaker than the center-of-mass condition: it holds with the same parameters, but there is an additional degree of freedom, as the following result shows.

Proposition 4.4. *Let $\mathcal{T} = \mathcal{T}_{\pi,z}$ be a mesh on \mathcal{S}^1 and set $\bar{\lambda}_{k,k+1} = \frac{r_{k+1}-z_k}{z_{k+1}-z_k}$ for $k = 0, \dots, K-1$. For $\{\lambda_{k,k+1}\} \subseteq [0, 1]$ the following assertions hold:*

(1) *The center-of-mass condition holds if and only if*

$$\lambda_{k,k+1} = \bar{\lambda}_{k,k+1} .$$

(2) *The isotropy condition holds if and only if*

$$\lambda_{k,k+1} = \bar{\lambda}_{k,k+1} + \frac{s}{d_{k,k+1}}$$

for some $s \in [-\min_k \bar{\lambda}_{k,k+1} d_{k,k+1}, \min_k \bar{\lambda}_{k+1,k} d_{k,k+1}]$.

Proof. This follows immediately by solving the corresponding linear systems. \square

Remark 4.5 (Relation to the asymptotic isotropy condition). Recall from [GKM18, Definition 1.3] that a family of meshes $\{\mathcal{T}\}$ (in any dimension) is said to satisfy the *isotropy condition* with parameters $\{\lambda_{KL}\}$ if, for any $K \in \mathcal{T}$,

$$\sum_{L \in \mathcal{T}} \lambda_{KL} \frac{|K|L|}{d_{KL}} (z_K - z_L) \otimes (z_K - z_L) \leq |K| (I_d + \eta_{\mathcal{T}}(K)) \quad (4.4)$$

where $\sup_{K \in \mathcal{T}} |\eta_{\mathcal{T}}(K)| \rightarrow 0$ as $\max\{\text{diam}(A) : A \in \mathcal{T}\} \rightarrow 0$.

Applying this condition to the family of one-dimensional periodic meshes \mathcal{T}^N constructed from \mathcal{T} , it reduces to

$$\lambda_{k,k-1} d_{k-1,k} + \lambda_{k,k+1} d_{k,k+1} \leq \pi_k (1 + \eta_N(k))$$

for all $N \geq 1$ and $k = 0, \dots, K-1$, where $\eta_N(k) \rightarrow 0$ as $N \rightarrow \infty$. As the left-hand side does not depend on N , this condition in turn simplifies to

$$\lambda_{k,k-1} d_{k-1,k} + \lambda_{k,k+1} d_{k,k+1} \leq \pi_k \quad (4.5)$$

for all $k = 0, \dots, K-1$.

Clearly, (4.3) implies (4.5). To see that both assertions are equivalent, we note that (4.5) can be written as

$$\lambda_{k,k+1} d_{k,k+1} - \lambda_{k-1,k} d_{k-1,k} \leq \pi_k - d_{k-1,k} . \quad (4.6)$$

To obtain a contradiction, suppose that we have strict inequality in (4.6) for some $k = \bar{k}$. Summation over $k = 0, \dots, K-1$ yields

$$0 = \sum_{k=0}^{K-1} (\lambda_{k,k+1} d_{k,k+1} - \lambda_{k-1,k} d_{k-1,k}) < \sum_{k=0}^{K-1} (\pi_k - d_{k-1,k}) = 0 ,$$

which is absurd.

In summary, we conclude that the isotropy condition (4.3) is equivalent to the asymptotic isotropy condition (4.4) for the family of meshes $\{\mathcal{T}^N\}$.

The next definition will be used to connect geometric properties of the mesh to properties of the means in the definition of the transport distance.

Definition 4.6 (Adaptedness). (1) *Let $\lambda \in [0, 1]$. A mean θ is said to be λ -balanced if $\theta(a, b) \leq \lambda a + (1 - \lambda)b$ for any $a, b \geq 0$.*

(2) *A family of means $\{\theta_{k,k+1}\}$ is said to be adapted to the parameters $\{\lambda_{k,k+1}\}$ if $\theta_{k,k+1}$ is $\lambda_{k,k+1}$ -balanced for each k .*

Remark 4.7. Each continuously differentiable mean θ is λ -balanced for exactly one value of $\lambda \in [0, 1]$, namely

$$\lambda = \partial_1 \theta(1, 1) . \quad (4.7)$$

A non-smooth mean θ can be λ -balanced for multiple values of λ , e.g., the mean $(a, b) \mapsto \min\{a, b\}$ is λ -balanced for any $\lambda \in [0, 1]$.

Now we are ready to state the main result of this section. The result is consistent with the main result in [GKM18], which asserts that the asymptotic isotropy condition is necessary (and essentially sufficient) for Gromov–Hausdorff convergence of the discrete transport distance to \mathbb{W}_2 .

Theorem 4.8 (Isotropy is equivalent to $c^*(\theta, \mathcal{T}) = 1$). *Let $\{\theta_{k,k+1}\}$ be a family of means that are adapted to $\{\lambda_{k,k+1}\}$.*

- (1) *If \mathcal{T} satisfies the isotropy condition with parameters $\{\lambda_{k,k+1}\}$, then $c^*(\theta, \mathcal{T}) = 1$.*
- (2) *Assume that each mean $\theta_{k,k+1}$ is continuously differentiable. If $c^*(\theta, \mathcal{T}) = 1$, then \mathcal{T} satisfies the isotropy condition with parameters $\{\lambda_{k,k+1}\}$.*

Remark 4.9 (Minimum mean). In view of Proposition 4.4, every mesh \mathcal{T} satisfies the isotropy condition for suitable choice of $\{\lambda_k\}$. Since the minimum mean $(a, b) \mapsto \min\{a, b\}$ is λ -balanced for any value of $\lambda \in [0, 1]$, it thus follows from Theorem 4.8 that $c^*(\theta, \mathcal{T}) = 1$ if $\theta_{k,k+1} = \min$ for each k .

Proof. To prove (1), take any sequence $\{m_k\}_k$ with $\sum_{k=0}^{K-1} m_k = 1$. Using Jensen’s inequality, the adaptedness, the periodicity, and the isotropy condition, we obtain

$$\begin{aligned} \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}} \right)} &\geq \left(\sum_{k=0}^{K-1} d_{k,k+1} \theta_{k,k+1} \left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}} \right) \right)^{-1} \\ &\geq \left(\sum_{k=0}^{K-1} d_{k,k+1} \left(\lambda_{k,k+1} \frac{m_k}{\pi_k} + \lambda_{k+1,k} \frac{m_{k+1}}{\pi_{k+1}} \right) \right)^{-1} \\ &= \left(\sum_{k=0}^{K-1} \frac{m_k}{\pi_k} (\lambda_{k,k+1} d_{k,k+1} + \lambda_{k,k-1} d_{k-1,k}) \right)^{-1} \\ &= \left(\sum_{k=0}^{K-1} m_k \right)^{-1} = 1 . \end{aligned}$$

Taking the infimum over $\{m_k\}_k$, we obtain $c^*(\theta, \mathcal{T}) \geq 1$. In view of Proposition 4.1 we infer that $c^*(\theta, \mathcal{T}) = 1$.

To prove (2), we consider the probability measures γ_α^k defined by

$$\gamma_\alpha^k = (\pi_0, \dots, \pi_{k-1}, \pi_k + \alpha, \pi_{k+1} - \alpha, \pi_{k+2}, \dots, \pi_{K-1})$$

for $|\alpha|$ sufficiently small. Let us write

$$h_{\theta, \mathcal{T}}(m) = \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{m_k}{\pi_k}, \frac{m_{k+1}}{\pi_{k+1}} \right)} .$$

As $c^*(\theta, \mathcal{T}) = 1$, we have $h_{\theta, \mathcal{T}}(m) \geq 1$ for all m . Thus, since $h_{\theta, \mathcal{T}}(\gamma_0^k) = h_{\theta, \mathcal{T}}(\pi) = 1$, it follows that $\frac{d}{d\alpha}\big|_{\alpha=0} h_{\theta, \mathcal{T}}(\gamma_\alpha^k) = 0$. A direct computation shows that

$$\frac{d}{d\alpha}\bigg|_{\alpha=0} h_{\theta, \mathcal{T}}(\gamma_\alpha^k) = A_{k+1} - A_k \quad \text{where} \quad A_k := \frac{\lambda_{k,k-1}d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1}}{\pi_k}.$$

As this holds for every k , we infer that there exists a constant $\beta > 0$ such that $A_k = \beta$ for every $k = 0, \dots, K-1$. The latter means that

$$\beta\pi_k = \lambda_{k,k-1}d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1}$$

for all $k = 0, \dots, K-1$. Summation over k yields

$$\beta = \beta \sum_{k=0}^{K-1} \pi_k = \sum_{k=0}^{K-1} (1 - \lambda_{k-1,k})d_{k-1,k} + \lambda_{k,k+1}d_{k,k+1} = \sum_{k=0}^{K-1} d_{k-1,k} = 1,$$

which proves the isotropy condition with parameters $\{\lambda_{k,k+1}\}_k$. \square

5. PROOF OF THE LOWER BOUND

The goal of this section is to prove Theorem 5.4, which yields the lower bound in Theorem 1.1. The crucial ingredient is Proposition 5.3, which ensures the existence of approximately optimal curves with good regularity properties.

To formulate this result, we fix a non-negative function $\eta \in C_c^\infty(0, \frac{1}{2})$ with $\int_0^1 \eta(x) dx = 1$. We set $\eta_\lambda(x) = \frac{1}{\lambda}\eta(\frac{x}{\lambda})$ for $x \in [0, 1)$, and consider its periodic extension to \mathcal{S}^1 . For $\lambda \in (0, 1]$ we define a discrete spatial mollifier by

$$\eta_\lambda^N(n) := \frac{N}{\lambda} \int_{\frac{n}{N}}^{\frac{n+1}{N}} \eta\left(\frac{x}{\lambda}\right) dx, \quad n = 0, \dots, N-1,$$

and we extend η_λ^N to \mathbb{Z} periodically modulo N , so that it can be regarded as a function on the discrete torus $\mathbb{T}_N = \mathbb{Z}/N\mathbb{Z}$. It follows that $\frac{1}{N} \sum_{n=0}^{N-1} \eta_\lambda^N(n) = 1$, and the following kernel bounds hold for $n = 0, \dots, N-1$:

$$|\eta_\lambda^N(n)| \leq \frac{\|\eta\|_\infty}{\lambda}, \quad |\eta_\lambda^N(n_1) - \eta_\lambda^N(n_2)| \leq \frac{\|\eta'\|_\infty}{\lambda^2} \frac{|n_1 - n_2|}{N}, \quad (5.1)$$

We consider the convolution operators $\mathbb{M}_\lambda : L^1(\mathcal{S}^1) \rightarrow L^\infty(\mathcal{S}^1)$ given by

$$(\mathbb{M}_\lambda f)(x) = \int_{\mathcal{S}^1} \eta_\lambda(x-y)f(y) dy,$$

as well as the analogous discrete convolution operators $\mathcal{M}_\lambda^N : L^1(\mathbb{T}_N) \rightarrow L^\infty(\mathbb{T}_N)$ defined by

$$(\mathcal{M}_\lambda^N \psi)(n) = \frac{1}{N} \sum_{j=0}^{N-1} \eta_\lambda^N(n-j)\psi(j).$$

The kernel bounds (5.1) imply the following L^1 - L^∞ and L^1 -Lipschitz bounds:

$$\sup_n |\mathcal{M}_\lambda^N \psi(n)| \leq \frac{\|\eta\|_\infty}{\lambda N} \sum_{n=0}^{N-1} |\psi(n)|, \quad (5.2)$$

$$\sup_n |\mathcal{M}_\lambda^N \psi(n_1) - \mathcal{M}_\lambda^N \psi(n_2)| \leq \frac{\|\eta'\|_\infty}{\lambda^2} \frac{|n_1 - n_2|}{N^2} \sum_{n=0}^{N-1} |\psi(n)|. \quad (5.3)$$

The following result collect some basic properties for convolution operators that will be used in the sequel.

Lemma 5.1 (Bounds for convolution operators). *Let $\lambda \in (0, 1]$ and $N \geq 2$. For any $\mu \in \mathcal{P}(\mathcal{S}^1)$ and $m \in \mathcal{P}(\mathbb{T}_N)$ we have*

$$\mathbb{W}_2(\mu, \mathbb{M}_\lambda \mu) \leq C \lambda, \quad (5.4)$$

$$\mathbb{W}_2(\iota_N \mathcal{M}_\lambda^N m, \mathbb{M}_\lambda \iota_N m) \leq \frac{\lambda}{2} + \frac{2}{N}, \quad (5.5)$$

where $C < \infty$ depends only on η .

Proof. The inequality (5.4) follows straightforwardly using the coupling $\gamma(dx, dy) = \eta_\lambda(y - x) d\mu(x) dy$.

To prove (5.5), we note that

$$\mathbb{W}_2^2(\iota_N \mathcal{M}_\lambda^N m, \mathbb{M}_\lambda \iota_N m) \leq \sum_{i=0}^{N-1} m_i \mathbb{W}_2^2(\iota_N \mathcal{M}_\lambda^N \delta_i, \mathbb{M}_\lambda \iota_N \delta_i)$$

by convexity of \mathbb{W}_2^2 . Thus it suffices to prove the lemma for $m = \delta_i$. Since $d(\iota_N \delta_i)(x) = N \mathbf{1}_{\left[\frac{i}{N}, \frac{i+1}{N}\right]}(x) dx$, we have

$$d(\mathbb{M}_\lambda \iota_N \delta_i)(x) = N \left(\eta_\lambda * \mathbf{1}_{\left[\frac{i}{N}, \frac{i+1}{N}\right]} \right)(x) dx. \quad (5.6)$$

On the other hand, we have

$$d(\iota_N \mathcal{M}_\lambda^N \delta_i)(x) = \sum_{n=0}^{N-1} \eta_\lambda^n(n - i) \mathbf{1}_{\left[\frac{n}{N}, \frac{n+1}{N}\right]}(x) dx.$$

Since $\text{supp } \eta_\lambda \subset \left(0, \frac{\lambda}{2}\right)$, we obtain

$$\text{supp } \mathbb{M}_\lambda \iota_N \delta_i \subseteq \left[\frac{i}{N}, \frac{i+1}{N} + \frac{\lambda}{2}\right] \quad \text{and} \quad \text{supp } \iota_N \mathcal{M}_\lambda^N \delta_i \subseteq \left[\frac{i-1}{N}, \frac{i+1}{N} + \frac{\lambda}{2}\right],$$

hence

$$\text{diam} \left(\text{supp}(\iota_N \mathcal{M}_\lambda^N \delta_i) \cup \text{supp}(\mathbb{M}_\lambda \iota_N \delta_i) \right) \leq \frac{\lambda}{2} + \frac{2}{N}.$$

This easily yields the desired result. □

Before stating the crucial regularisation result, we formulate a lemma which asserts that we can decrease the energy at the discrete level by a suitable regularisation. Here it is crucial that the regularisation is performed by averaging the density at spatial locations $nK + k$ and $n'K + k$ that differ by a multiple of the period K . A “naive” regularisation consisting of locally averaging the density, without taking the periodic structure into account, would in general not decrease the energy. We emphasise that the operator \mathcal{M}_λ^N is understood to act on the variable n in the result below.

Lemma 5.2 (Energy bound under periodic smoothing). *Let $\lambda \in (0, 1]$. For any $m \in \mathcal{P}(\mathcal{T}_N)$ and any $J \in \mathcal{V}(\mathcal{T}_N)$ we have*

$$\mathcal{A}_N(\mathcal{M}_\lambda^N m, \mathcal{M}_\lambda^N J) \leq \mathcal{A}_N(m, J).$$

Proof. For brevity we write

$$F_{k,k+1}(m, J, n) = \frac{d_{k,k+1}}{N} f_{k,k+1} \left(N \frac{m(n; k)}{\pi_k}, N \frac{m(n; k+1)}{\pi_{k+1}}, J(n; k, k+1) \right).$$

Applying Jensen's inequality to the jointly convex functions $f_{k,k+1}$ we obtain

$$\begin{aligned} \mathcal{A}_N(\mathcal{M}_\lambda^N m, \mathcal{M}_\lambda^N J) &= \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} F_{k,k+1}(\mathcal{M}_\lambda^N m, \mathcal{M}_\lambda^N J, n) \\ &\leq \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} \frac{1}{N} \sum_{\ell=0}^{N-1} \eta_\lambda^N(n-\ell) F_{k,k+1}(m, J, \ell) \\ &= \sum_{k=0}^{K-1} \sum_{\ell=0}^{N-1} \left(\frac{1}{N} \sum_{n=0}^{N-1} \eta_\lambda^N(n-\ell) \right) F_{k,k+1}(m, J, \ell) \\ &= \sum_{k=0}^{K-1} \sum_{\ell=0}^{N-1} F_{k,k+1}(m, J, \ell) \\ &= \mathcal{A}_N(m, J), \end{aligned}$$

where we used that $\frac{1}{N} \sum_{n=0}^{N-1} \eta_\lambda^N(n) = 1$. □

We are now ready to state the main regularisation result of this section. As we expect that (approximately) optimal densities exhibit oscillations, we cannot expect spatial regularity for such densities. Nevertheless, the lemma above allows us to obtain a restricted form of regularity for such densities, in the sense that good Lipschitz bounds hold if one only compares values of the density at spatial locations $nK + k$ and $n'K + k$ that differ by a multiple of the period K .

Note that the vector field J enjoys better regularity properties: in (5.8e) we even obtain a Lipschitz bound for neighbouring cells.

Proposition 5.3 (Space-time regularisation). *Fix $N \geq 1$, and let $(m_t, J_t)_t$ be a solution to the discrete continuity equation (2.3) in $\mathcal{P}(\mathcal{T}_N)$ satisfying*

$$A := \int_0^1 \mathcal{A}_N(m_t, J_t) dt < \infty.$$

Then, for any $\varepsilon > 0$ there exists a solution $(\tilde{m}_t, \tilde{J}_t)_t$ to (2.3) such that:

- (1) $\mathbb{W}_2(\iota_N \tilde{m}_t, \iota_N m_t) \leq \varepsilon + \frac{C}{N}$ for all $t \in [0, 1]$, where $C < \infty$ depends only on \mathcal{T} ;
- (2) the following action bound holds:

$$\int_0^1 \mathcal{A}_N(\tilde{m}_t, \tilde{J}_t) dt \leq \int_0^1 \mathcal{A}_N(m_t, J_t) dt; \tag{5.7}$$

(3) the following regularity properties hold, for some constants $c_0, \dots, c_5 < \infty$ depending on ε and A , but not on N :

$$c_0^{-1} \leq \min_{n,k} N\tilde{m}_t(n; k) \leq \max_{n,k} N\tilde{m}_t(n; k) \leq c_1, \quad (5.8a)$$

$$\sup_{t \in [0,1]} \max_{n,k} |N\partial_t \tilde{m}_t^N(n; k)| \leq c_2, \quad (5.8b)$$

$$\sup_{t \in [0,1]} \max_{n,k} |N\tilde{m}_t(n; k) - N\tilde{m}_t(n+1; k)| \leq \frac{c_3}{N}, \quad (5.8c)$$

$$\sup_{t \in [0,1]} \max_{n,k} |\tilde{J}_t^N(n; k, k+1)| \leq c_4, \quad (5.8d)$$

$$\sup_{t \in [0,1]} \max_{n,k} |\tilde{J}_t(n; k, k+1) - \tilde{J}_t(n; k-1, k)| \leq \frac{c_5}{N}. \quad (5.8e)$$

Proof. Let $\mathcal{U}_N := P_N \mathcal{L}^1|_{S^1} \in \mathcal{P}(\mathcal{T}_N)$ denote the probability measure that assigns mass $\frac{\pi_k}{N}$ to $A_{n,k}$. Fix a mollifier η as above. For $\lambda, \tau, \delta > 0$ we define a space-time regularisation by

$$\tilde{m}_t(n; k) := \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{M}_\lambda^N \left[(1-\delta)m_u + \delta\mathcal{U}_N \right] (n; k) du, \quad (5.9a)$$

$$\tilde{J}_t(n; k, k+1) := \frac{1-\delta}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{M}_\lambda^N J_u(n; k, k+1) du. \quad (5.9b)$$

In both expressions, the operator \mathcal{M}_λ^N is understood to act on the variable n , i.e., the spatial averaging takes place over cells whose distance is an integer multiple of the period K . Moreover, we use the convention that $m_u = m_0$ and $J_u = 0$ for $u < 0$, and $m_u = m_1$ and $J_u = 0$ for $u > 1$. We claim that this approximation satisfies all the sought properties.

An explicit computation shows that $(\tilde{m}_t, \tilde{V}_t)_t$ solves the discrete continuity equation.

To prove (5.7), we note that by a trifold application of the joint convexity of \mathcal{A}_N ,

$$\begin{aligned} \mathcal{A}_N(\tilde{m}_t, \tilde{J}_t) &\leq \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{A}_N \left(\mathcal{M}_\lambda^N \left[(1-\delta)m_u + \delta\mathcal{U}_N \right], (1-\delta)\mathcal{M}_\lambda^N J_u \right) du \\ &\leq \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{A}_N \left([(1-\delta)m_u + \delta\mathcal{U}_N], (1-\delta)J_u \right) du \\ &\leq \frac{1-\delta}{2\tau} \int_{t-\tau}^{t+\tau} \mathcal{A}_N(m_u, J_u) du. \end{aligned}$$

Here we used the crucial regularisation bound from Lemma 5.2. The desired inequality (5.7) follows.

Moreover, since \mathcal{M}_λ^N preserves positivity, we deduce (5.8a) with $c_0^{-1} = \delta \min_k \pi_k$.

To prove (5.8b), we observe that

$$\partial_t \tilde{m}_t = \frac{1-\delta}{2\tau} \mathcal{M}_\lambda^N [m_{t+\tau} - m_{t-\tau}].$$

Therefore, since m_t is a probability measure, the L^1 - L^∞ -bound in (5.2) yields

$$N|\partial_t \tilde{m}_t(n)| \leq \frac{\|\eta\|_\infty}{\tau\lambda} =: c_2$$

which proves (5.8b).

The inequality (5.8e), with $c_5 = c_2$, follows immediately from (5.8b) and the fact that $(\tilde{m}_t^N, \tilde{J}_t)_t$ solves the continuity equation.

Furthermore, since

$$|\tilde{m}_t(n; k) - \tilde{m}_t(n+1; k)| \leq \sup_t |\mathcal{M}_\lambda^N m_t(n; k) - \mathcal{M}_\lambda^N m_t(n+1; k)| \leq \frac{\|\eta'\|_\infty}{\lambda^2 N^2},$$

we obtain (5.8c) with $c_3 = \frac{\|\eta'\|_\infty}{\lambda^2}$, in view of the Lipschitz bound in (5.3).

To prove the upper bound on the density in (5.8a), we take advantage of the L^1 - L^∞ bound (5.2) once more to obtain

$$N\tilde{m}_t(n; k) \leq \frac{\|\eta\|_\infty}{\lambda} =: c_1.$$

Finally, to obtain the L^∞ -bound on the vector field (5.8d), we use (5.2) again to infer

$$\begin{aligned} \sup_{n,k} |\tilde{J}_t(n; k, k+1)| &\leq \frac{1-\delta}{2\tau} \int_{t-\tau}^{t+\tau} \sup_{n,k} |\mathcal{M}_\lambda^N J_u(n; k, k+1)| du \\ &\leq \frac{\|\eta\|_\infty}{\lambda N} \frac{1-\delta}{2\tau} \int_{t-\tau}^{t+\tau} \sup_k \sum_{n=0}^{N-1} |J_u(n; k, k+1)| du. \end{aligned}$$

Writing $\theta_{n,k,k+1} = \theta_{k,k+1} \left(N \frac{m_u(n;k)}{\pi_k}, N \frac{m_u(n;k+1)}{\pi_{k+1}} \right)$ for brevity, we infer that

$$\begin{aligned} \frac{1}{N} \left(\sum_{n,k} |J_u(n; k, k+1)| \right)^2 &\leq \left(\sum_{n,k} \frac{d_{k,k+1}}{N} \frac{J_u^2(n; k, k+1)}{\theta_{n,k,k+1}} \right) \left(\sum_{n,k} \frac{\theta_{n,k,k+1}}{d_{k,k+1}} \right) \\ &= \mathcal{A}_N(m_u, J_u) \sum_{n,k} \frac{\theta_{n,k,k+1}}{d_{k,k+1}}. \end{aligned}$$

Using the bound $\theta_{k,k+1}(a, b) \leq a + b$ we obtain

$$\sum_{n,k} \frac{\theta_{n,k,k+1}}{d_{k,k+1}} \leq \frac{N}{\min_k d_{k,k+1}} \sum_{n,k} \left(\frac{m_u(n; k)}{\pi_k} + \frac{m_u(n; k+1)}{\pi_{k+1}} \right) \leq 2BN,$$

where $B = (\max_k \pi_k^{-1})(\max_k d_{k,k+1}^{-1})$. Combining these bounds, we arrive at

$$\begin{aligned} \sup_{n,k} |\tilde{J}_t(n; k, k+1)| &\leq \frac{\|\eta\|_\infty \sqrt{2B}}{2\tau\lambda} \int_{t-\tau}^{t+\tau} \sqrt{\mathcal{A}_N(m_u, J_u)} du \\ &\leq \frac{\|\eta\|_\infty}{\lambda} \sqrt{\frac{B}{\tau}} \int_0^1 \mathcal{A}_N(m_u, J_u) du, \end{aligned}$$

which yields (5.8d) with $c_4 := \frac{\|\eta\|_\infty}{\lambda} \sqrt{\frac{AB}{\tau}}$. As we will choose $\delta, \lambda, \tau > 0$ depending on ε , the bounds (5.8a)–(5.8e) follow.

It remains to show that $\mathbb{W}_2(\iota_N \tilde{m}_t, \iota_N m_t) \leq \varepsilon + \frac{C}{N}$ for suitable values of δ, λ and τ . We consider the effect of the three different regularisations separately. First we apply the convexity of \mathbb{W}_2^2 to obtain for any $m \in \mathcal{P}(\mathcal{T}_N)$,

$$\mathbb{W}_2^2(\iota_N m, \iota_N[(1-\delta)m + \delta\mathcal{U}_N]) \leq \delta \mathbb{W}_2^2(\iota_N m, \mathcal{L}^1|_{S^1}) \leq \frac{\delta}{4}, \quad (5.10)$$

since the diameter of $(\mathcal{P}(\mathcal{S}^1), \mathbb{W}_2)$ is equal to $\frac{1}{2}$. Moreover, for $m \in \mathcal{P}(\mathcal{T}_N)$, Lemma 5.1 yields

$$\begin{aligned} \mathbb{W}_2(\iota_N m, \iota_N \mathcal{M}_\lambda^N m) &\leq \mathbb{W}_2(\iota_N m, \mathbb{M}_\lambda \iota_N m) + \mathbb{W}_2(\mathbb{M}_\lambda \iota_N m, \iota_N \mathcal{M}_\lambda^N m) \\ &\leq C \left(\lambda + \frac{1}{N} \right), \end{aligned} \quad (5.11)$$

where $C < \infty$ depends only on η . Furthermore, set $\bar{m}_t = \mathcal{M}_\lambda^N((1-\delta)m_t + \delta\mathcal{U}_N)$ and $\bar{J}_t = (1-\delta)\mathcal{M}_\lambda^N J_t$. It then follows that $c_0 \leq N\bar{m}_t \leq c_1$ and $\int_0^1 \mathcal{A}_N(\bar{m}_t, \bar{J}_t) dt \leq A$. Thus, for $s \leq t$, Proposition 2.7 yields a constant $\kappa < \infty$ depending on c_0 and c_1 (hence on δ and λ) such that,

$$\begin{aligned} \mathbb{W}_2^2(\iota_N \bar{m}_s, \iota_N \bar{m}_t) &\leq \kappa \int_0^1 \mathcal{A}(\bar{m}_{(1-a)s+at}, (t-s)\bar{J}_{(1-a)s+at}) da \\ &\leq \kappa(t-s) \int_s^t \mathcal{A}(\bar{m}_u, \bar{J}_u) du \\ &\leq \kappa A(t-s). \end{aligned}$$

By convexity of \mathbb{W}_2^2 , we obtain

$$\mathbb{W}_2^2\left(\iota_N \bar{m}_t, \iota_N \left[\frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \bar{m}_u du \right]\right) \leq \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \mathbb{W}_2^2(\iota_N \bar{m}_t, \iota_N \bar{m}_u) du \leq \frac{\kappa\tau A}{2}. \quad (5.12)$$

Applying the estimates (5.10) with $m = m_t$, (5.11) with $m = (1-\delta)m_t + \delta\mathcal{U}_N$ and (5.12), we arrive at

$$\mathbb{W}_2(\iota_N \tilde{m}_t, \iota_N m_t) \leq C \left(\sqrt{\delta} + \lambda + \frac{1}{N} + \sqrt{\kappa\tau A} \right),$$

for some $C < \infty$ depending only on \mathcal{T} and on η . Thus, choosing first λ and δ sufficiently small, and then τ sufficiently small depending on δ , the result follows. \square

We are now ready to prove the lower bound in Theorem 1.1.

Theorem 5.4 (Lower bound for \mathcal{W}_N). *For any mesh \mathcal{T} and any family of admissible means $\{\theta_{k,k+1}\}_k$ we have*

$$c^*(\theta, \mathcal{T}) \mathbb{W}_2^2(\mu_0, \mu_1) \leq \liminf_{N \rightarrow \infty} \mathcal{W}_N^2(P_N \mu_0, P_N \mu_1),$$

uniformly for all $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$. More precisely, for any $\varepsilon > 0$ there exist $C < \infty$ and $\bar{N} \in \mathbb{N}$ such that for any $N \geq \bar{N}$ and $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have

$$c^*(\theta, \mathcal{T}) \mathbb{W}_2^2(\mu_0, \mu_1) \leq \mathcal{W}_N^2(P_N \mu_0, P_N \mu_1) + \varepsilon + \frac{C}{N}. \quad (5.13)$$

Proof. Fix $\varepsilon > 0$. Applying Proposition 5.3 to an approximate \mathcal{W}_N -geodesic between $P_N \mu_0$ and $P_N \mu_1$, we infer that there exists a curve $(m_t, J_t)_t$ satisfying the bounds

$$\mathbb{W}_2(\iota_N m_i, \iota_N P_N \mu_i) \leq \varepsilon + \frac{C}{N} \quad \text{for } i = 0, 1, \quad (5.14)$$

$$\int_0^1 \mathcal{A}_N(m_t, J_t) dt \leq \mathcal{W}_N^2(P_N \mu_0, P_N \mu_1) + \varepsilon, \quad (5.15)$$

as well as the regularity properties (5.8a)–(5.8e).

For brevity we write

$$\mathcal{A}_N(m, J) = \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{A}_N^n(m, J) ,$$

where

$$\mathcal{A}_N^n(m, J) = \sum_{k=0}^{K-1} d_{k,k+1} f_{k,k+1} \left(N \frac{m(n; k)}{\pi_k}, N \frac{m(n; k+1)}{\pi_{k+1}}, J(n; k, k+1) \right) .$$

We set

$$\widehat{m}_t(n) := \sum_{k=0}^{K-1} m_t(n; k) , \quad \widehat{J}_t(n) := J_t(n; -1, 0) ,$$

and define $\alpha_t : \{0, \dots, N-1\} \times \{0, \dots, K\} \rightarrow \mathbb{R}$ by

$$\alpha_t(n; k) = \frac{m_t(n; \sigma(k))}{\widehat{m}_t(n)} \quad \text{for } k = 0, \dots, K ,$$

where $\sigma(k) = k$ for $k = 0, \dots, K-1$, and $\sigma(K) = 0$. Here it is important to note that $\alpha_t(n; K) \neq \alpha_t(n+1; 0)$. Observe that, for $k = 0, \dots, K-1$,

$$\frac{1}{\theta_{k,k+1} \left(\frac{\alpha_t(n; k)}{\pi_k}, \frac{\alpha_t(n; k+1)}{\pi_{k+1}} \right)} \frac{|\widehat{J}_t(n)|^2}{N \widehat{m}_t(n)} = f_{k,k+1} \left(N \frac{m_t(n; \sigma(k))}{\pi_k}, N \frac{m_t(n; \sigma(k+1))}{\pi_{k+1}}, \widehat{J}_t(n) \right) .$$

Note that, for any n and k ,

$$\frac{c_0^{-1}}{\max_{\ell} \pi_{\ell}} \leq N \frac{m_t(n; k)}{\pi_k} \leq \frac{c_1}{\min_{\ell} \pi_{\ell}} \quad \text{and} \quad |J_t(n; k, k+1)| \leq c_4 .$$

Therefore, since the functions $f_{k,k+1}$ are Lipschitz on the set $[\frac{c_0^{-1}}{\max_{\ell} \pi_{\ell}}, \frac{c_1}{\min_{\ell} \pi_{\ell}}]^2 \times [-c_4, c_4]$, it follows that, for $k = 0, \dots, K-1$,

$$\begin{aligned} & \left| f_{k,k+1} \left(N \frac{m_t(n; k)}{\pi_k}, N \frac{m_t(n; k+1)}{\pi_{k+1}}, J_t(n; k, k+1) \right) - \frac{1}{\theta_{k,k+1} \left(\frac{\alpha_t(n; k)}{\pi_k}, \frac{\alpha_t(n; k+1)}{\pi_{k+1}} \right)} \frac{|\widehat{J}_t(n)|^2}{N \widehat{m}_t(n)} \right| \\ & \leq [f_{k,k+1}]_{\text{Lip}} \left(\frac{N}{\pi_{k+1}} |m_t(n; k+1) - m_t(n; \sigma(k+1))| + |J_t(n; k, k+1) - \widehat{J}_t(n)| \right) \\ & \leq \frac{[f_{k,k+1}]_{\text{Lip}}}{N} \left(\frac{c_3}{\pi_{k+1}} + K c_5 \right) =: \frac{C}{N} , \end{aligned} \tag{5.16}$$

for some $C < \infty$ depending on ε (through c_0, \dots, c_5) and on \mathcal{T} .

Since $\sum_{k=0}^{K-1} \alpha_t(n; k) = 1$, the sequence $\{\alpha_t(n; k)\}_{k=0}^{K-1}$ is, for any n , a competitor for the cell problem (1.5). Taking into account that $\alpha_t(n; 0) = \alpha_t(n; K)$, it follows from (5.16) and the definition (1.5) of $c^*(\theta, \mathcal{T})$ that

$$\begin{aligned} \mathcal{A}_N^n(m_t, J_t) & \geq \frac{|\widehat{J}_t(n)|^2}{N \widehat{m}_t(n)} \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{\alpha_t(n; k)}{\pi_k}, \frac{\alpha_t(n; k+1)}{\pi_{k+1}} \right)} - \frac{C}{N} \\ & \geq c^*(\theta, \mathcal{T}) \frac{|\widehat{J}_t(n)|^2}{N \widehat{m}_t(n)} - \frac{C}{N} . \end{aligned} \tag{5.17}$$

At the continuous level, we define a curve of measures $(\mu_t^N)_t$ with piecewise constant densities, and a vector field j_t^N by piecewise affine interpolation of \widehat{J}_t^N ; more precisely,

$$\begin{aligned}\mu_t^N &= \sum_{n=0}^{N-1} \widehat{m}_t \mathcal{U}_{\widehat{A}_n} , \\ j_t^N(x) &= \sum_{n=0}^{N-1} \chi_{\widehat{A}_n}(x) \left[(n+1 - Nx) \widehat{J}_t(n) + (Nx - n) \widehat{J}_t(n+1) \right] .\end{aligned}$$

As before, $\mathcal{U}_{\widehat{A}_n}$ denotes the normalised Lebesgue measure on $\widehat{A}_n := [\frac{n}{N}, \frac{n+1}{N})$.

We observe that the density ρ_t^N of μ_t^N satisfies

$$\partial_t \rho_t^N(x) = N \sum_{k=0}^{K-1} \partial_t \widehat{m}_t(n; k) = N(\widehat{J}_t(n) - \widehat{J}_t(n+1)) = -\partial_x j_t^N(x)$$

for any $x \in (\frac{n}{N}, \frac{n+1}{N})$, which implies that $(\mu_t^N, j_t^N)_t$ solves the continuity equation.

To estimate the continuous energy, we find

$$\begin{aligned}\mathbb{A}(\mu_t^N, j_t^N) &= \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{\widehat{m}_t(n)} \int_{\frac{n}{N}}^{\frac{n+1}{N}} \left[(n+1 - Nx) \widehat{J}_t(n) + (Nx - n) \widehat{J}_t(n+1) \right]^2 dx \\ &\leq \frac{1}{N} \sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2 + \widehat{J}_t(n+1)^2}{2N \widehat{m}_t(n)} \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2}{\theta_h(N \widehat{m}_t(n), N \widehat{m}_t(n+1))}\end{aligned}$$

where $\theta_h(a, b) = \frac{2ab}{a+b}$ denotes the harmonic mean. Note that (5.8c) implies the Lipschitz bound

$$|N \widehat{m}_t(n) - N \widehat{m}_t(n-1)| \leq \frac{Kc_3}{N} .$$

and (5.8a) yields a lower bound on the density: $N \widehat{m}_t(n) \geq Kc_0^{-1}$. Furthermore, (5.8d) yields the estimate $|\widehat{J}_t(n)| \leq c_4$. Thus, in view of the identity $\frac{1}{\theta_h(a,b)} = \frac{1}{a} + \frac{a-b}{2ab}$ we obtain

$$\mathbb{A}(\mu_t^N, j_t^N) \leq \left(\frac{1}{N} \sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2}{N \widehat{m}_t(n)} \right) + \frac{C}{N} , \tag{5.18}$$

with $C < \infty$ depending on ε (through the c_i 's) and on \mathcal{T} .

Putting things together, it follows from (5.17), (5.18) and (5.15) that

$$\begin{aligned}
c^*(\theta, \mathcal{T}) \mathbb{W}_2^2(\mu_0^N, \mu_1^N) &\leq c^*(\theta, \mathcal{T}) \int_0^1 \mathbb{A}(\mu_t^N, j_t^N) dt \\
&\leq c^*(\theta, \mathcal{T}) \int_0^1 \frac{1}{N} \sum_{n=0}^{N-1} \frac{\widehat{J}_t(n)^2}{N \widehat{m}_t(n)} dt + \frac{C}{N} \\
&\leq \int_0^1 \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{A}_N^n(m_t, J_t) dt + \frac{C}{N} \\
&= \int_0^1 \mathcal{A}_N(m_t, J_t) dt + \frac{C}{N} \\
&\leq \mathcal{W}_N^2(P_N \mu_0, P_N \mu_1) + \varepsilon + \frac{C}{N}.
\end{aligned} \tag{5.19}$$

Finally we note that, for $i = 0, 1$, (5.14) yields

$$\begin{aligned}
\mathbb{W}_2(\mu_i, \mu_i^N) &\leq \mathbb{W}_2(\mu_i, \iota_N P_N \mu_i) + \mathbb{W}_2(\iota_N P_N \mu_i, \iota_N m_i) + \mathbb{W}_2(\iota_N m_i, \mu_i^N) \\
&\leq \frac{1}{N} + \left(\varepsilon + \frac{C}{N} \right) + \frac{1}{N} \\
&\leq \varepsilon + \frac{C}{N},
\end{aligned}$$

which implies that

$$\mathbb{W}_2(\mu_0, \mu_1) \leq \mathbb{W}_2(\mu_0^N, \mu_1^N) + 2 \left(\varepsilon + \frac{C}{N} \right). \tag{5.20}$$

Combining (5.19) and (5.20) we obtain the desired result. \square

6. PROOF OF THE UPPER BOUND

In this section we present the proof of the upper bound for \mathcal{W}_N . The idea of the proof of the upper bound is to start from optimal curves of measures at the continuous level, and to introduce the optimal oscillation in their discretations, as determined by the formula for the effective mobility (1.5).

Let $\alpha^* = \{\alpha_k^*\}_{k=0}^{K-1}$ be an optimiser in (1.5), and define $P_N^* : \mathcal{P}(\mathcal{S}^1) \rightarrow \mathcal{P}(\mathcal{T}_N)$ by

$$(P_N^* \mu)(n; k) := \alpha_k^* \mu(\widehat{A}_n), \quad \text{where } \widehat{A}_n := \left[\frac{n}{N}, \frac{n+1}{N} \right), \tag{6.1}$$

as before. Slightly abusing notation, we also define $P_N^* : C(\mathcal{S}^1; \mathbb{R}) \rightarrow \mathcal{V}(\mathcal{T}_N)$ by

$$(P_N^* j)(n; k, k+1) := \left(\sum_{\ell=k+1}^{K-1} \alpha_\ell^* \right) j\left(\frac{n}{N}\right) + \left(\sum_{\ell=0}^k \alpha_\ell^* \right) j\left(\frac{n+1}{N}\right).$$

Since $\sum_{k=0}^{K-1} \alpha_k^* = 1$, the right-hand side is a convex combination of $j\left(\frac{n}{N}\right)$ and $j\left(\frac{n+1}{N}\right)$.

Proposition 6.1 (Discretisation of the continuity equation). *Let $(\mu_t)_{t \in [0,1]}$ be a Borel family of probability measures, and let $(j_t)_{t \in [0,1]}$ be a Borel family of continuous functions satisfying the continuity equation $\partial_t \mu + \partial_x j = 0$ on \mathcal{S}^1 . Then the pair $(m_t, J_t)_{t \in [0,1]}$ defined by*

$$m_t := P_N^* \mu_t, \quad J_t := P_N^* j_t,$$

solves the continuity equation on \mathcal{T}_N .

Proof. As $(\mu_t, j_t)_t$ satisfies the continuity equation, we have

$$\int_0^1 \left(\int_{\mathcal{S}^1} \partial_t \phi_t(x) d\mu_t(x) + \int_{\mathcal{S}^1} \partial_x \phi_t(x) j_t(x) dx \right) dt = \int_{\mathcal{S}^1} \phi_1(x) d\mu_1(x) - \int_{\mathcal{S}^1} \phi_0(x) d\mu_0(x)$$

for any smooth function $\phi : [0, 1] \times \mathcal{S}^1 \rightarrow \mathbb{R}$.

Let $\psi : [0, 1] \rightarrow \mathbb{R}$ be smooth, and define $\eta^\varepsilon : \mathcal{S}^1 \rightarrow \mathbb{R}$ by $\eta^\varepsilon = \chi_{\widehat{A}_n} * \xi^\varepsilon$ for a smooth mollifier ξ^ε supported in an ε -neighbourhood of 0. Set $\phi_t^\varepsilon(x) = \psi(t)\eta^\varepsilon(x)$. Applying the weak formulation of the continuity equation to ϕ^ε , and passing to the limit $\varepsilon \downarrow 0$, we obtain

$$\int_0^1 \psi'(t) \mu_t(\widehat{A}_n) dt + \int_0^1 \psi(t) \left(j_t\left(\frac{n+1}{N}\right) - j_t\left(\frac{n}{N}\right) \right) dt = \psi(1) \mu_1(\widehat{A}_n) - \psi(0) \mu_0(\widehat{A}_n) .$$

Multiplying this identity by α_k^* , and using the fact that

$$\alpha_k^* \left(j_t\left(\frac{n+1}{N}\right) - j_t\left(\frac{n}{N}\right) \right) = J_t(n; k, k+1) - J_t(n; k-1, k) ,$$

we obtain

$$\begin{aligned} & \int_0^1 \psi'(t) m_t(n; k) dt + \int_0^1 \psi(t) \left(J_t(n; k, k+1) - J_t(n; k-1, k) \right) dt \\ & = \psi(1) m_1(n; k) - \psi(0) m_0(n; k) , \end{aligned}$$

which is the distributional form of the discrete continuity equation (2.3). □

Lemma 6.2 (Consistency). *For all $\mu \in \mathcal{P}(\mathcal{S}^1)$ we have*

$$\mathbb{W}_2(\mu, \iota_N P_N^* \mu) \leq \frac{1}{N} .$$

Proof. This readily follows from the definitions; see [GKM18, Lemma 3.2] for a similar result. □

The following proposition is the key result of this section. It proves the required upper bound for the discrete energy under suitable regularity conditions.

For $\delta > 0$, it will be useful to write

$$\mathcal{P}_\delta(\mathcal{S}^1) := \left\{ \mu = \rho dx \in \mathcal{P}(\mathcal{S}^1) : \rho \geq \delta > 0, \quad \text{Lip}(\rho) \leq \frac{1}{\delta} \right\} .$$

Proposition 6.3 (Discrete energy upper bound). *Let $\delta > 0$. There exists $C < \infty$ and $\bar{N} \in \mathbb{N}$ (depending on δ), such that for any $N \geq \bar{N}$, all $\mu \in \mathcal{P}_\delta(\mathcal{S}^1)$, and all vector fields $j : \mathcal{S}^1 \rightarrow \mathbb{R}$ with $\|j\|_{L^\infty} + \text{Lip}(j) \leq \delta^{-1}$, we have*

$$\mathcal{A}_N(P_N^* \mu, P_N^* j) \leq c^*(\theta, \mathcal{T}) \mathbb{A}(\mu, j) + \frac{C}{N} .$$

Proof. Write $m = P_N^* \mu$ and $J = P_N^* j$ for brevity, and set $\bar{\rho}(n) := N\mu(\widehat{A}_n)$. Recall that

$$\mathcal{A}_N(m, J) = \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{A}_N^n(m, J)$$

where

$$\begin{aligned} \mathcal{A}_N^n(m, J) &= \sum_{k=0}^{K-1} d_{k,k+1} f_{k,k+1} \left(N \frac{m(n; k)}{\pi_k}, N \frac{m(n; k+1)}{\pi_{k+1}}, J(n; k, k+1) \right) \\ &= \frac{1}{\bar{\rho}(n)} \sum_{k=0}^{K-1} d_{k,k+1} \frac{(J(n; k, k+1))^2}{\theta_{k,k+1} \left(\frac{\alpha(n; k)}{\pi_k}, \frac{\alpha(n; k+1)}{\pi_{k+1}} \right)}, \end{aligned}$$

with $\alpha(n; k) := \frac{m(n; k)}{\mu(\widehat{A}_n)}$ for $k = 0, \dots, K$. Note that $\alpha(n; k) = \alpha_k^*$ for $k = 0, \dots, K-1$, but

$$\alpha(n; K) = \frac{m(n+1; 0)}{\mu(\widehat{A}_n)} = \alpha_0^* \frac{\mu(\widehat{A}_{n+1})}{\mu(\widehat{A}_n)},$$

which is not necessarily equal to $\alpha_K^* = \alpha_0^*$. Therefore, $\{\alpha(n; k)\}_{k=0}^K$ is not necessarily an admissible competitor in (1.5). Write $d\mu(x) = \rho(x) dx$. We claim that the following estimates hold for sufficiently large N , with $C < \infty$ depending only on θ and \mathcal{T} :

$$|J(n; k, k+1) - j(x)| \leq \frac{\text{Lip}(j)}{N}, \quad (6.2)$$

$$\left| c^*(\theta, \mathcal{T}) - \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{\alpha(n; k)}{\pi_k}, \frac{\alpha(n; k+1)}{\pi_{k+1}} \right)} \right| \leq C \frac{\text{Lip}(\rho)}{\inf \rho} \frac{1}{N}, \quad (6.3)$$

the first one being valid for any $x \in \left[\frac{n}{N}, \frac{n+1}{N} \right]$ and $k = 0, \dots, K-1$. Indeed, writing $\lambda_k = \sum_{\ell=k+1}^{K-1} \alpha_\ell^*$, we obtain

$$|J(n; k, k+1) - j(x)| \leq \lambda_k |j(\frac{n}{N}) - j(x)| + (1 - \lambda_k) |j(\frac{n+1}{N}) - j(x)| \leq \frac{\text{Lip}(j)}{N},$$

which proves (6.2). Furthermore,

$$\begin{aligned} E &:= \left| c^*(\theta, \mathcal{T}) - \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{\alpha(n; k)}{\pi_k}, \frac{\alpha(n; k+1)}{\pi_{k+1}} \right)} \right| \\ &= d_{K-1, K} \left| \frac{1}{\theta_{K-1, K} \left(\frac{\alpha_{K-1}^*}{\pi_{K-1}}, \frac{\alpha_K^*}{\pi_K} \right)} - \frac{1}{\theta_{K-1, K} \left(\frac{\alpha_{K-1}^*}{\pi_{K-1}}, \frac{\alpha(n; K)}{\pi_K} \right)} \right|. \end{aligned}$$

Note that

$$|\alpha_K^* - \alpha(n; K)| = \alpha_K^* \left| 1 - \frac{\mu(\widehat{A}_{n+1})}{\mu(\widehat{A}_n)} \right| = \alpha_K^* \frac{|\mu(\widehat{A}_n) - \mu(\widehat{A}_{n+1})|}{\mu(\widehat{A}_n)} \leq \frac{\alpha_K^* \text{Lip}(\rho)}{N \inf \rho}.$$

If $\alpha_K^* = 0$, we infer that $E = 0$, in which case the claim is proved. If $\alpha_K^* > 0$, we observe that the latter inequality yields

$$\alpha(n; K) \geq \frac{\alpha_K^*}{2} \quad (6.4)$$

for N sufficiently large (depending on δ). Since $\theta_{K-1, K}$ is concave, we have for any $a \geq 0$ and $0 < b \leq y_1 < y_2$,

$$\frac{\theta_{K-1, K}(a, y_2) - \theta_{K-1, K}(a, y_1)}{y_2 - y_1} \leq \frac{\theta_{K-1, K}(a, b) - \theta_{K-1, K}(a, 0)}{b},$$

thus $\theta_{K-1,K}(a, \cdot)$ is Lipschitz on $[b, \infty)$. Let $L < \infty$ denote the Lipschitz constant of $\theta_{K-1,K}\left(\frac{\alpha_{K-1}^*}{\pi_{K-1}}, \cdot\right)$ on $\left[\frac{\alpha_K^*}{2}, \infty\right)$. For N sufficiently large we obtain

$$\begin{aligned} E &\leq \frac{1}{\left(\theta_{K-1,K}\left(\frac{\alpha_{K-1}^*}{\pi_{K-1}}, \frac{\alpha_K^*}{2\pi_K}\right)\right)^2} \left| \theta_{K-1,K}\left(\frac{\alpha_{K-1}^*}{\pi_{K-1}}, \frac{\alpha_K^*}{\pi_K}\right) - \theta_{K-1,K}\left(\frac{\alpha_{K-1}^*}{\pi_{K-1}}, \frac{\alpha(n;K)}{\pi_K}\right) \right| \\ &\leq \frac{L}{\left(\theta_{K-1,K}\left(\frac{\alpha_{K-1}^*}{\pi_{K-1}}, \frac{\alpha_K^*}{2\pi_K}\right)\right)^2} \frac{\alpha_K^* \operatorname{Lip}(\rho)}{\pi_K} \frac{1}{\inf \rho} \frac{1}{N} \end{aligned}$$

which yields our claim (6.3).

Taking into account that $\bar{\rho}(n) \geq \delta$ and $\|j\|_\infty \leq \delta^{-1}$, it follows from (6.2) and a twofold application of (6.3) that

$$\begin{aligned} \left| \mathcal{A}_N^n(m, J) - c^*(\theta, \mathcal{T}) \frac{j^2\left(\frac{n}{N}\right)}{\bar{\rho}(n)} \right| &\leq \frac{j^2\left(\frac{n}{N}\right)}{\bar{\rho}(n)} \left| c^*(\theta, \mathcal{T}) - \sum_{k=0}^{K-1} \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\right)} \right| \\ &\quad + \frac{1}{\bar{\rho}(n)} \sum_{k=0}^{K-1} d_{k,k+1} \frac{\left| J^2(n; k, k+1) - j^2\left(\frac{n}{N}\right) \right|}{\theta_{k,k+1} \left(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\right)} \\ &\leq \frac{C}{N} + \frac{C}{N} \sum_k \frac{d_{k,k+1}}{\theta_{k,k+1} \left(\frac{\alpha(n;k)}{\pi_k}, \frac{\alpha(n;k+1)}{\pi_{k+1}}\right)} \\ &\leq \frac{C}{N}, \end{aligned}$$

where $C < \infty$ depends on \mathcal{T} , θ , and δ . Consequently,

$$\mathcal{A}_N(m, J) \leq \frac{c^*(\theta, \mathcal{T})}{N} \sum_{n=0}^{N-1} \frac{j^2\left(\frac{n}{N}\right)}{\bar{\rho}(n)} + \frac{C}{N}.$$

By the arithmetic-harmonic mean inequality,

$$\frac{\left| j\left(\frac{n}{N}\right) \right|^2}{\bar{\rho}(n)} \leq N \left| j\left(\frac{n}{N}\right) \right|^2 \int_{\hat{\mathcal{A}}_n} \frac{1}{\rho(x)} dx \leq N \int_{\hat{\mathcal{A}}_n} \frac{|j(x)|^2}{\rho(x)} dx + \frac{C}{N}.$$

We infer that

$$\mathcal{A}_N(m, J) \leq c^*(\theta, \mathcal{T}) \mathbb{A}(\mu, j) + \frac{C}{N},$$

which completes the proof. □

The previous result shows that the sought upper bound can be achieved once we assume some regularity at the level of the solution of the continuity equation. Therefore in order to conclude the proof of Theorem 1.1 we seek once again for a regularization procedure.

The following result collects some well-known properties of the heat flow on \mathcal{S}^1 .

Lemma 6.4 (Regularisation by heat flow). *Let $s > 0$. There exists a constant $\delta > 0$ such that for any $\mu \in \mathcal{P}(\mathcal{S}^1)$ we have $H_s \mu \in \mathcal{P}_\delta(\mathcal{S}^1)$. Moreover, $\mathbb{W}_2(\mu, H_s \mu) \leq \sqrt{2s}$.*

Proof. See, e.g., [GM13, Proposition 2.9] for a proof of these well-known facts. □

We continue with a well-known regularisation result. For the convenience of the reader we include a simple proof.

Lemma 6.5 (Smooth approximate action minimisers). *Let $\delta > 0$ and let $\varepsilon > 0$. Then there exists $\tilde{\delta} \in (0, \delta)$, such that the following assertion holds: for any $\mu_0, \mu_1 \in \mathcal{P}_\delta(\mathcal{S}^1)$ there exists a curve $(\mu_t, j_t) \in \mathbb{C}\mathbb{E}(\mu_0, \mu_1)$ with $\mu_t \in \mathcal{P}_{\tilde{\delta}}(\mathcal{S}^1)$ and $\|j_t\|_{L^\infty} + \text{Lip}(j_t) \leq \tilde{\delta}^{-1}$ for any $t \in (0, 1)$, such that*

$$\int_0^1 \mathbb{A}(\mu_t, j_t) dt \leq \mathbb{W}_2^2(\mu_0, \mu_1) + \varepsilon .$$

Proof. Let $(\mu_t)_{t \in [0,1]}$ be a \mathbb{W}_2 -geodesic connecting μ_0 and μ_1 , and let $(j_t)_{t \in [0,1]}$ be a vector field such that

$$\int_0^1 \mathbb{A}(\mu_t, j_t) dt = \mathbb{W}_2^2(\mu_0, \mu_1) .$$

The idea of the proof is to regularise μ_0 and μ_1 by applying the heat flow for a short time $s > 0$, and then to connect the regularised measures $H_s\mu_0$ and $H_s\mu_1$ using the natural candidate $(H_s\mu_t)_{t \in [0,1]}$.

Firstly, for $i = 0, 1$ and $s > 0$, set $\gamma_t^{i,s} = H_{st}\mu_i$ for $t \in [0, 1]$, and let $\rho_t^{i,s}$ be the density of $\gamma_t^{i,s}$ with respect to the Haar measure. Then: $\partial_t \gamma_t^{i,s} = s \partial_x^2 \gamma_t^{i,s}$, thus the continuity equation $\partial_t \rho_t^{i,s} + \partial_x k_t^{i,s} = 0$ holds with $k_t^{i,s} = -s \partial_x \rho_t^{i,s}$. Using the contractivity of the Fisher information under the heat flow, and the fact that $\mu_i \in \mathcal{P}_\delta(\mathcal{S}^1)$, we obtain

$$\int_0^1 \mathbb{A}(\gamma_t^{i,s}, k_t^{i,s}) dt = s^2 \int_0^1 \int_{\mathcal{S}^1} \frac{|\partial_x \rho_t^{i,s}(x)|^2}{\rho_t^s(x)} dx dt \leq s^2 \int_{\mathcal{S}^1} \frac{|\partial_x \rho_0^{i,s}(x)|^2}{\rho_0^s(x)} dx \leq \frac{s^2}{\delta^3} .$$

Secondly, for any $s > 0$, we note that $(H_s\mu_t, H_s j_t)_{t \in [0,1]}$ solves the continuity equation, and, by the joint convexity of \mathbb{A} and the fact that H_s is given by a convolution kernel,

$$\mathbb{A}(H_s\mu_t, H_s j_t) \leq \mathbb{A}(\mu_t, j_t) .$$

Fix $\tau \in (0, \frac{1}{2})$, and consider now the curve $(\tilde{\mu}_t, \tilde{j}_t)_{t \in [0,1]} \in \mathbb{C}\mathbb{E}(\mu_0, \mu_1)$ defined by

$$\tilde{\mu}_t := \begin{cases} H_{ts/\tau} \mu_0 & t \in (0, \tau) \\ H_s \mu_{(t-\tau)/(1-2\tau)} & t \in (\tau, 1-\tau) \\ H_{(1-t)s/\tau} \mu_1 & t \in (1-\tau, 1) \end{cases} , \quad \tilde{j}_t := \begin{cases} \frac{1}{\tau} k_{t/\tau}^{0,s} & t \in (0, \tau) \\ \frac{1}{1-2\tau} H_s j_{(t-\tau)/(1-2\tau)} & t \in (\tau, 1-\tau) \\ -\frac{1}{\tau} k_{1-t/\tau}^{1,s} & t \in (1-\tau, 1) \end{cases} ,$$

It follows from the bounds above, using the fact that $\frac{1}{1-2\tau} \leq 1 + 4\tau$ and $\mathbb{W}_2^2 \leq \frac{1}{4}$, that

$$\begin{aligned} \int_0^1 \mathbb{A}(\tilde{\mu}_t, \tilde{j}_t) dt &= \int_0^1 \frac{\mathbb{A}(\gamma_t^{0,s}, k_t^{0,s})}{\tau} + \frac{\mathbb{A}(H_s\mu_t, H_s j_t)}{1-2\tau} + \frac{\mathbb{A}(\gamma_t^{1,s}, k_t^{1,s})}{\tau} dt \\ &\leq \frac{s^2}{\delta^3 \tau} + \frac{\mathbb{W}_2^2(\mu_0, \mu_1)}{1-2\tau} + \frac{s^2}{\delta^3 \tau} \\ &\leq \frac{s^2}{\delta^3 \tau} + (\mathbb{W}_2^2(\mu_0, \mu_1) + \tau) + \frac{s^2}{\delta^3 \tau} . \end{aligned}$$

Let $\varepsilon > 0$, and choose $\tau = \varepsilon/2$, and $s^2 = \delta^3 \tau \varepsilon/4$. Then: $\int_0^1 \mathbb{A}(\tilde{\mu}_t, \tilde{j}_t) dt \leq \mathbb{W}_2^2(\mu_0, \mu_1) + \varepsilon$.

Moreover, by Lemma 6.4, $\tilde{\mu}_t$ belongs to $\mathcal{P}_{\tilde{\delta}}(\mathcal{S}^1)$ for some $\tilde{\delta} > 0$ depending on δ and s . Furthermore,

$$\|H_s j_t\|_{L^\infty(\mathcal{S}^1)} + \|\partial_x H_s j_t\|_{L^\infty(\mathcal{S}^1)} \leq C(s) \|j_t\|_{L^1(\mathcal{S}^1)} \leq C(s) \sqrt{\mathbb{A}(\mu_t, j_t)} = C(s) \mathbb{W}_2(\mu_0, \mu_1) ,$$

where the last inequality follows from the Cauchy-Schwarz inequality. \square

We are now ready to prove the upper bound in Theorem 1.1.

Theorem 6.6 (Upper bound for \mathcal{W}_N). *For any mesh \mathcal{T} and any family of admissible means $\{\theta_{k,k+1}\}_k$ we have*

$$\limsup_{N \rightarrow \infty} \mathcal{W}_N^2(P_N \mu_0, P_N \mu_1) \leq c^*(\theta, \mathcal{T}) \mathbb{W}_2^2(\mu_0, \mu_1) ,$$

uniformly for all $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$. More precisely, for any $\varepsilon > 0$ there exist $C < \infty$ and $\bar{N} \in \mathbb{N}$ such that for any $N \geq \bar{N}$ and $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$, we have

$$\mathcal{W}_N^2(P_N \mu_0, P_N \mu_1) \leq c^*(\theta, \mathcal{T}) \mathbb{W}_2^2(\mu_0, \mu_1) + \varepsilon + \frac{C}{N} . \quad (6.5)$$

Proof. Let $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$ and $\varepsilon \in (0, 1]$. By Lemma 6.4 there exist $s \geq 0$ and $\delta > 0$ such that $\tilde{\mu}_i := H_s \mu_i$ belongs to $\mathcal{P}_\delta(\mathcal{S}^1)$, and

$$\mathbb{W}_2(\mu_i, \tilde{\mu}_i) \leq \varepsilon \quad \text{for } i = 0, 1 .$$

Using that $\mathbb{W}_2 \leq \frac{1}{2}$, it follows that

$$\mathbb{W}_2^2(\tilde{\mu}_0, \tilde{\mu}_1) \leq \mathbb{W}_2^2(\mu_0, \mu_1) + 2\varepsilon . \quad (6.6)$$

Lemma 6.5 yields $\tilde{\delta} \in (0, \delta)$ and a curve $(\tilde{\mu}_t, \tilde{j}_t)_t \in \mathbb{C}\mathbb{E}_{\tilde{\delta}}(\tilde{\mu}_0, \tilde{\mu}_1)$ such that $\tilde{\mu}_t \in \mathcal{P}_{\tilde{\delta}}(\mathcal{S}^1)$ and $\|\tilde{j}_t\|_{L^\infty} + \text{Lip}(\tilde{j}_t) \leq \tilde{\delta}^{-1}$ for any $t \in (0, 1)$, and

$$\int_0^1 \mathbb{A}(\tilde{\mu}_t, \tilde{j}_t) dt \leq \mathbb{W}_2^2(\tilde{\mu}_0, \tilde{\mu}_1) + \varepsilon . \quad (6.7)$$

Set $\tilde{m}_t^N := P_N^* \tilde{\mu}_t$ and $\tilde{J}_t^N = P_N^* \tilde{j}_t$. By Proposition 6.3 there exists $C_1 < \infty$ depending on ε (through $\tilde{\delta}$) such that

$$\mathcal{W}_N^2(\tilde{m}_0^N, \tilde{m}_1^N) \leq \int_0^1 \mathcal{A}(\tilde{m}_t^N, \tilde{J}_t^N) dt \leq c^*(\theta, r) \int_0^1 \mathbb{A}(\tilde{\mu}_t, \tilde{j}_t) dt + \frac{C_1}{N} . \quad (6.8)$$

Set $m_i^N := P_N^* \mu_i$ for $i = 0, 1$. By Proposition 2.6, Lemma 6.2, and Lemma 6.4, there exists $C_2 < \infty$ depending only on \mathcal{T} (possibly varying from line to line) such that

$$\begin{aligned} \mathcal{W}_N(m_i^N, \tilde{m}_i^N) &= \mathcal{W}_N(P_N^* \mu_i, P_N^* H_s \mu_i) \\ &\leq C_2 \left(\mathbb{W}_2(\iota_N P_N^* \mu_i, \iota_N P_N^* H_s \mu_i) + \frac{1}{N} \right) \\ &\leq C_2 \left(\mathbb{W}_2(\mu_i, H_s \mu_i) + \frac{1}{N} \right) \\ &\leq C_2 \left(\sqrt{s} + \frac{1}{N} \right) . \end{aligned}$$

Thus, the triangle inequality yields

$$\mathcal{W}_N(m_0^N, m_1^N) \leq \mathcal{W}_N(\tilde{m}_0^N, \tilde{m}_1^N) + C_2 \left(\sqrt{s} + \frac{1}{N} \right) ,$$

and by another application of Proposition 2.6,

$$\mathcal{W}_N^2(m_0^N, m_1^N) \leq \mathcal{W}_N^2(\tilde{m}_0^N, \tilde{m}_1^N) + C_2 \left(\sqrt{s} + \frac{1}{N} \right) . \quad (6.9)$$

Combining (6.6), (6.7), (6.8), and (6.9), we obtain

$$\mathcal{W}_N^2(m_0^N, m_1^N) \leq c^*(\theta, r) \mathbb{W}_2^2(\mu_0, \mu_1) + 3\varepsilon + \frac{C_1}{N} + C_2 \left(\sqrt{s} + \frac{1}{N} \right) .$$

Choosing s depending on ε we obtain the result. □

7. PROOF OF THE GROMOV–HAUSDORFF CONVERGENCE

We conclude this work with the proof of the Gromov–Hausdorff convergence in Theorem 1.1. First we recall one of the equivalent definitions; cf. [BBI01] for more details.

Definition 7.1 (Gromov–Hausdorff convergence). *A sequence of compact metric spaces $\{\mathcal{X}_N, d_N\}_N$ is said to converge in the sense of Gromov–Hausdorff to a compact metric space (\mathcal{X}, d) , if there exist maps $f_N : \mathcal{X} \rightarrow \mathcal{X}_N$ with the following properties:*

- ε -isometry: for any $\varepsilon > 0$ there exists $\bar{N} \in \mathbb{N}$ such that for any $N \geq \bar{N}$ and any $x, y \in \mathcal{X}$, we have:

$$|d_N(f_N(x), f_N(y)) - d(x, y)| \leq \varepsilon ;$$

- ε -surjectivity: for any $\varepsilon > 0$ there exists $\bar{N} \in \mathbb{N}$ such that for any $N \geq \bar{N}$ and any $z \in \mathcal{X}_N$ there exists $x \in \mathcal{X}$ satisfying

$$d_N(f_N(x), z) \leq \varepsilon .$$

Proof of Theorem 1.1. As the desired lower and upper bounds for the distance have been proved in Theorems 5.4 and 6.6, it remains to prove the Gromov–Hausdorff convergence.

Let $\varepsilon > 0$. It follows from Theorems 5.4 and 6.6 that there exists $\bar{N} \in \mathbb{N}$ such that, for any $N \geq \bar{N}$ and $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{S}^1)$,

$$\left| \mathcal{W}_N(P_N \mu_0, P_N \mu_1) - \sqrt{c^*(\theta, \mathcal{T})} \mathbb{W}_2(\mu_0, \mu_1) \right| \leq \varepsilon .$$

This shows that the map P_N is ε -isometric.

The ε -surjectivity of P_N holds trivially, since it is even surjective. \square

Acknowledgements. J.M. gratefully acknowledges support by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 716117). J.M and L.P. also acknowledge support from the Austrian Science Fund (FWF), grants No F65 and W1245. E.K. gratefully acknowledges support by the German Research Foundation through the Hausdorff Center for Mathematics and the Collaborative Research Center 1060. P.G. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 350398276.

REFERENCES

- [AGS08] 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 ed., 2008.
- [BB00] J.-D. BENAMOU and Y. BRENIER. A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84**, no. 3, (2000), 375–393.
- [BBI01] D. BURAGO, Y. BURAGO, and S. IVANOV. *A course in metric geometry*, vol. 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [CHLZ12] 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**, no. 3, (2012), 969–1008.
- [DFM18] P. DONDL, T. FRENZEL, and A. MIELKE, A gradient system with a wiggly energy and relaxed EDP-convergence, (2018), arXiv:1801.07144.
- [DL15] K. DISSER and M. LIERO. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *Netw. Heterog. Media* **10**, no. 2, (2015), 233–253.

- [EF18] M. ERBAR and M. FATHI. Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. *J. Funct. Anal.* **274**, no. 11, (2018), 3056–3089.
- [EGH00] R. EYMARD, T. GALLOUËT, and R. HERBIN. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, 713–1020. North-Holland, Amsterdam, 2000.
- [EHMT17] M. ERBAR, C. HENDERSON, G. MENZ, and P. TETALI. Ricci curvature bounds for weakly interacting Markov chains. *Electron. J. Probab.* **22**, (2017), Paper No. 40, 23.
- [EM12] M. ERBAR and J. MAAS. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.* **206**, no. 3, (2012), 997–1038.
- [EM14] M. ERBAR and J. MAAS. Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Syst. Syst.* **34**, no. 4, (2014), 1355–1374.
- [EMT15] M. ERBAR, J. MAAS, and P. TETALI. Discrete Ricci curvature bounds for Bernoulli-Laplace and random transposition models. *Ann. Fac. Sci. Toulouse Math. (6)* **24**, no. 4, (2015), 781–800.
- [FM16] M. FATHI and J. MAAS. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.* **26**, no. 3, (2016), 1774–1806.
- [GKM18] P. GLADBACH, E. KOPFER, and J. MAAS. Scaling limits of discrete optimal transport, (2018), arXiv:1809.01092.
- [GM13] N. GIGLI and J. MAAS. Gromov–Hausdorff convergence of discrete transportation metrics. *SIAM Journal on Mathematical Analysis* **45**, no. 2, (2013), 879–899.
- [GT17] N. GARCÍA TRILLOS, Gromov-Hausdorff limit of Wasserstein spaces on point clouds, (2017), arXiv:1702.03464.
- [JKO98] R. JORDAN, D. KINDERLEHRER, and F. OTTO. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29**, no. 1, (1998), 1–17.
- [Maa11] J. MAAS. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* **261**, no. 8, (2011), 2250–2292.
- [Mie11] A. MIELKE. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity* **24**, no. 4, (2011), 1329–1346.
- [Mie13] A. MIELKE. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations* **48**, no. 1-2, (2013), 1–31.
- [Mie16a] A. MIELKE. Deriving effective models for multiscale systems via evolutionary Γ -convergence. In *Control of self-organizing nonlinear systems*, Underst. Complex Syst., 235–251. Springer, [Cham], 2016.
- [Mie16b] A. MIELKE. On evolutionary Γ -convergence for gradient systems. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, vol. 3 of *Lect. Notes Appl. Math. Mech.*, 187–249. Springer, [Cham], 2016.
- [PC19] G. PEYRÉ and M. CUTURI. Computational optimal transport. *Foundations and Trends® in Machine Learning* **11**, no. 5-6, (2019), 355–607.
- [San15] F. SANTAMBROGIO. *Optimal transport for applied mathematicians*, vol. 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
- [Vil03] C. VILLANI. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [Vil08] C. VILLANI. *Optimal transport: old and new*, vol. 338. Springer Science & Business Media, 2008.

MATHEMATISCHES INSTITUT, UNIVERSITÄT LEIPZIG, AUGUSTUSPLATZ 5, 04103 LEIPZIG
E-mail address: gladbach@math.uni-leipzig.de

INSTITUT FÜR ANGEWANDTE MATHEMATIK, UNIVERSITÄT BONN, ENDENICHER ALLEE 60, 53115
BONN, GERMANY

E-mail address: eva.kopfer@iam.uni-bonn.de

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (IST AUSTRIA), AM CAMPUS 1, 3400
KLOSTERNEUBURG, AUSTRIA

E-mail address: jan.maas@ist.ac.at

E-mail address: lorenzo.portinale@ist.ac.at