



## **The Sinkhorn algorithm, parabolic optimal transport and geometric Monge–Ampère equations**

Downloaded from: <https://research.chalmers.se>, 2024-09-16 03:38 UTC

Citation for the original published paper (version of record):

Berman, R. (2020). The Sinkhorn algorithm, parabolic optimal transport and geometric Monge–Ampère equations. *Numerische Mathematik*, 145(4): 771-836.  
<http://dx.doi.org/10.1007/s00211-020-01127-x>

N.B. When citing this work, cite the original published paper.



# The Sinkhorn algorithm, parabolic optimal transport and geometric Monge–Ampère equations

Robert J. Berman<sup>1</sup>

Received: 2 March 2018 / Revised: 10 March 2020 / Published online: 27 June 2020  
© The Author(s) 2020

## Abstract

We show that the discrete Sinkhorn algorithm—as applied in the setting of Optimal Transport on a compact manifold—converges to the solution of a fully non-linear parabolic PDE of Monge–Ampère type, in a large-scale limit. The latter evolution equation has previously appeared in different contexts (e.g. on the torus it can be identified with the Ricci flow). This leads to algorithmic approximations of the potential of the Optimal Transport map, as well as the Optimal Transport distance, with explicit bounds on the arithmetic complexity of the construction and the approximation errors. As applications we obtain explicit schemes of nearly linear complexity, at each iteration, for optimal transport on the torus and the two-sphere, as well as the far-field antenna problem. Connections to Quasi-Monte Carlo methods are exploited.

**Mathematics Subject Classification** 35J60 · 35K55 · 90C08 · 65N99

## 1 Introduction

The theory of Optimal Transport [67,68] is used in a multitude of applications ranging from economics, statistics, cosmology, geometric optics and meteorology to more recent applications in data sciences (including machine learning, vision, graphics and imaging; see [27,63,64]). In the last years there has been a flurry of numerical work applying the Sinkhorn algorithm [61] (aka the Iterative Proportional Fitting Procedure [55]) as a fast and efficient way of computing approximations to optimal transport maps, or equivalently, solutions to certain geometric Monge–Ampère type equations. This is motivated by applications to machine learning [26] (concerning optimal transport in Euclidean  $\mathbb{R}^n$ ) and computer graphics and image processing [63] (where the general setting of optimal transport on Riemannian manifolds is considered). The

---

✉ Robert J. Berman  
robertb@chalmers.se

<sup>1</sup> Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden

key advantage of the Sinkhorn algorithm in this context is its favorable large-scale computational properties (parallelization, linear time convergence, etc [3]).

The main aim of the present paper is to show that, in a large-scale limit, the Sinkhorn algorithm converges towards the solution of a parabolic PDE of Monge–Ampère type, which, incidentally, previously has appeared in [44,45,57] and is called the *parabolic optimal transport equation* in [44]. The convergence is shown with explicit error estimates. This leads, in particular, to the first constructive approximation of the potential of the optimal transport map with explicit bounds on the time-complexity of the construction *and* the approximation errors introduced by the discretization.

## 1.1 Background and setup

### 1.1.1 The Sinkhorn iteration

Let  $p$  and  $q$  be two vectors in  $\mathbb{R}_+^n$  whose entries sum to one. Given any matrix  $K \in \mathbb{R}_+^N \times \mathbb{R}_+^N$  there exists, by Sinkhorn's theorem [61], two diagonal positive matrices  $D_a$  and  $D_b$  with diagonal vectors  $a$  and  $b$  in  $\mathbb{R}_+^N$  such that the matrix

$$B := D_b K D_a \tag{1.1}$$

has the property that the rows sum to  $p$  and the columns sum to  $q$ . The diagonal matrices  $D_a$  and  $D_b$  are uniquely determined up to scaling  $D_a$  and  $D_b^{-1}$  by the same positive number. Moreover,  $B$  can be obtained as the limit of the algorithm defined by alternately normalizing the rows and columns of the matrix. In other words,

$$B = \lim_{m \rightarrow \infty} B(m), \quad B(m) = D_{b(m)} K D_{a(m)},$$

where the pair of positive vectors  $(a(m), b(m))$  are defined by the following recursion, formulated in terms of matrix vector multiplications and component-wise division of vectors:

$$\begin{aligned} b(m) &= p/K \cdot a(m) \\ a(m+1) &= q/K^T \cdot b(m) \end{aligned}$$

with initial data  $a(0)$  taken as the vector with entries 1. In fact, any initial positive vector  $a(0)$  will do and the vectors  $a(m)$  and  $b(m)$  are convergent as  $m \rightarrow \infty$  (without any need of scaling) as follows from Theorem 2.9. The corresponding iteration

$$a(m+1) := q/K^T \cdot (p/K \cdot a(m))$$

will be called the Sinkhorn *iteration*. Its fixed point (uniquely determined up to scaling) is thus the vector  $a$  appearing in formula 1.1.

The same algorithm has appeared in various fields (economics, traffic planning, statistics,...; see [27]). In its most general (infinite dimensional) form, known as the *Iterative Proportional Fitting Procedure* in the statistics literature, the roles of  $p$  and  $q$

are played by probability measures on two (possible non-finite) topological spaces. In this setting the corresponding convergence of  $B(m)$  towards a limit  $B$  was established in [55] using a maximum entropy characterization of  $B$ , which in the discrete setting above says that  $B$  is the unique element realizing the infimum

$$\inf_{\gamma \in \Pi(p,q)} \mathcal{I}(\gamma|K)$$

where  $\mathcal{I}$  denotes the Kullback–Leibler divergence of  $\gamma$  relative to  $K$ , when  $\gamma$  and  $K$  are identified with measures on the discrete product  $\{1, \dots, N\}^2$  and  $\Pi(p, q)$  denotes the set of all matrices  $\gamma$  in  $\mathbb{R}_+^N \times \mathbb{R}_+^N$  with row sum  $p$  and column sum  $q$  (i.e. the corresponding measures on  $\{1, \dots, N\}^2$  have marginals  $p$  and  $q$ , respectively). Since  $-\mathcal{I}(\gamma|K)$  is the “physical” entropy of  $\gamma$  relative to  $K$  this can indeed be viewed as a maximum entropy characterization of  $B$ . An alternative proof of the convergence follows from Theorem 2.9 below, which also shows that  $a(m)$  and  $b(m)$  have unique limits (determined by the initial value  $a(0)$ ).

### 1.1.2 Discrete optimal transport

Now replace  $K$  with a family of matrices  $K_\epsilon$  of the form

$$(K_\epsilon)_{ij} = e^{-\epsilon^{-1}C_{ij}},$$

for a given matrix  $C_{ij}$  and parametrized by a positive number  $\epsilon$ . Then the corresponding matrix  $B_\epsilon \in \Pi(p, q)$  furnished by Sinkhorn’s theorem converges, as  $\epsilon \rightarrow 0$ , to a matrix  $B_0$  realizing the infimum

$$C := \inf_{\gamma \in \Pi(p,q)} \langle C, \gamma \rangle. \tag{1.2}$$

In the terminology of discrete optimal transport theory [27,67] this means that  $B_0$  is an *optimal transport plan (coupling) between  $p$  and  $q$ , with respect to the cost matrix  $C$* . The convergence follows from the maximum entropy characterization of  $B_\epsilon$  (recalled in the previous section), which reveals that  $B_\epsilon$  realizes the perturbed minimum

$$C_\epsilon := \inf_{\gamma \in \Pi(p,q)} \langle c, \gamma \rangle + \epsilon \mathcal{I}(\gamma|I).$$

While the matrix  $B_0$  is sparse (and is typically supported on the graph of a transport map) the approximation  $B_\epsilon$  always has full support (by Sinkhorn’s theorem) and is thus more “regular” than  $B_0$ . Accordingly, the small parameter  $\epsilon$  is sometimes referred to as the *entropic regularization parameter*. This is illustrated by the simulations in [3] for the case when  $p$  and  $q$  represent the discretization of two probability measures on the unit-interval in  $\mathbb{R}$ , using a large number  $N$  of points and with  $C_{ij}$  the cost matrix defined by to the squared distance function on  $\mathbb{R}$ . When  $\epsilon$  is taken to be of the order  $1/N$  [3, Fig. 1] shows how the discrete probability measures on  $\mathbb{R} \times \mathbb{R}$  appear as smoothed out versions of the graph of the corresponding optimal transport map.

It should also be pointed out that the entropy minimization problem above can be traced back to the work by Schrödinger on Quantum Mechanics in the 30s [60] (see the survey [48], where the connection to optimal transport is emphasized).

### 1.1.3 Discretization of optimal transport on the torus

Let now  $X$  be a compact manifold endowed with a cost function  $c(x, y)$ . To keep things as simple as possible we will start by taking the manifold  $X$  to be the  $n$ -dimensional torus

$$T^n := \left( \frac{\mathbb{R}}{\mathbb{Z}} \right)^n$$

endowed with the standard distance function  $d_{T^n}(x, y)$ , induced from the Euclidean distance function on  $\mathbb{R}^n$ . Let  $\mu$  and  $\nu$  be two probability measure on  $T^n$  (which thus correspond to two periodic measures on  $\mathbb{R}^n$ ) with Hölder continuous and strictly positive densities  $e^{-f}$  and  $e^{-g}$ , respectively

$$\mu = e^{-f} dV, \quad \nu = e^{-g} dV \quad (1.3)$$

where  $dV$  is the Riemannian normalized volume form on  $T^n$ . Define the *cost function*  $c(x, y)$  by

$$c(x, y) := d_{T^n}(x, y)^2/2.$$

As is well-known, a continuous self-map  $F$  of  $T^n$  transporting (pushing forward)  $\mu$  to  $\nu$  is *optimal* with respect to this cost function, i.e. the corresponding transport plan  $\gamma_F := (I \times F)_* \mu$  realizes the infimum

$$d^2(\mu, \nu)/2 := \inf_{\gamma \in \Pi(\mu, \nu)} \langle c, \gamma \rangle, \quad (1.4)$$

if and only if  $F$  can be expressed in terms of a *potential*  $u \in C^2(T^n)$  :

$$F(x) := x + \nabla u(x), \quad T^n \rightarrow T^n,$$

which is strictly *quasi-convex* in the sense that the symmetric matrix  $\nabla^2 u + I$  is positive definite:

$$\nabla^2 u + I > 0$$

(we identify  $u$  with a  $\mathbb{Z}^n$ -periodic function on  $\mathbb{R}^n$  so that  $F(x)$  descends to define a self-map of  $T^n$ ). The function  $u$  is uniquely determined, up to an additive constant, by the following Monge–Ampère equation

$$\exp(-g(x + \nabla u(x))) \det(I + \nabla^2 u(x)) = \exp(-f(x)). \quad (1.5)$$

We recall that the distance  $d(\mu, \nu)$  between  $\mu$  and  $\nu$ , defined by formula 1.4, is usually called the *Wasserstein  $L^2$ -distance* or the *Optimal Transport distance*.

**Main results in the torus setting**

For notational reasons it will be convenient to express the entropic regularization parameter as

$$\epsilon := k^{-1}$$

for  $k$  a positive integer (but the results apply to any real parameter  $k \geq k_0 > 0$ ). For each  $k$  we fix a positive integer  $N_k$  and denote by  $\Lambda_k$  the corresponding discrete torus in  $T^n$  with  $N_k$  points. In other words,  $\Lambda_k$  is the grid on  $T^n$  with edge-length  $N_k^{-1/n}$  :

$$\Lambda_k := \left( \frac{(N_k^{-1/n} \mathbb{Z})}{\mathbb{Z}} \right)^n \subset T^n.$$

We will assume that

$$\lim_{k \rightarrow \infty} N_k = \infty.$$

Denote by  $p^{(k)}$  and  $q^{(k)}$  the corresponding discrete approximations in  $\mathbb{R}^{N_k}$  of  $\mu$  and  $\nu$ , defined by the normalized values of the densities of  $\mu$  and  $\nu$  at the points in  $\Lambda_k$ . Defining a sequence of  $N_k \times N_k$  matrices  $K^{(k)}$  by

$$K_{ij}^{(k)} := \mathcal{K}^{(k)}(x_i^{(k)}, x_j^{(k)}), \quad \mathcal{K}^{(k)}(x, y) := e^{-kd(x,y)^2/2} \tag{1.6}$$

and applying Sinkhorn’s theorem to the triple  $(K^{(k)}, p^{(k)}, q^{(k)})$  furnishes two positive vectors  $a^{(k)}$  and  $b^{(k)}$  in  $\mathbb{R}^{N_k}$ , uniquely determined by the normalization condition that  $a_{i_k}^{(k)} = 0$  for the index  $i_k$  corresponding to the point  $x_{i_k} = 0$  in  $\Lambda_k$ .

Our first result shows that the potential  $u$  for the optimal transport problem between  $\mu$  and  $\nu$  can be recovered from the positive vectors  $a^{(k)}$  and  $b^{(k)}$ , furnished by the Sinkhorn theorem, i.e. the fixed points of the corresponding iteration:

**Theorem 1.1** (Static case) *If  $x_{i_k}^{(k)}$  is a sequence of points in the discrete torus  $\Lambda_k$  converging to the point  $x$  in the torus  $T^n$ , as  $k \rightarrow \infty$ , then*

$$- \lim_{k \rightarrow \infty} k^{-1} \log a_{i_k}^{(k)} = u(x)$$

where  $u$  is the unique optimal transport potential solving the Monge–Ampère equation 1.5 and normalized by  $u(0) = 0$ .

This convergence result should come at no surprise and it holds in a very general setting (see Theorems 3.3, 3.4 and 3.6). But the main point of the present paper is

that the Sinkhorn algorithm itself, when viewed as a discrete dynamical system for the positive vectors  $a_{i_k}^{(k)}$ , also admits a continuous large-scale limit  $u_t(x)$ , evolving according to the following fully non-linear parabolic PDE:

$$\frac{\partial u_t(x)}{\partial t} = \log \det(I + \nabla^2 u_t(x)) - g(x + \nabla u_t(x)) + f(x), \quad u_0 = 0 \tag{1.7}$$

The existence of a  $C^4$ -smooth solution  $u_t$  to this PDE, given  $f, g \in C^{2,\alpha}(T^n)$ , essentially follows from the results in [44,45,57] (for completeness a proof is provided in ‘‘Appendix B’’). In order to formulate the convergence in the next theorem we first observe that for  $m \geq 1$  the function

$$u_m^{(k)}(x_i) = -k^{-1} \log \frac{a_i^{(k)}(m)}{p_i} \tag{1.8}$$

on the discrete torus  $\Lambda_k$  admits a canonical extension, defining a quasi-convex function on  $X$  :

$$u_m^{(k)}(x) := k^{-1} \log \sum_{i=1}^{N_k} e^{-kd(x,y_i^{(k)})^2/2} b_i^{(k)}(m-1) \tag{1.9}$$

expressed in terms of a Fourier/Gauss type sum, with  $k$  playing the role of the bandwidth.

**Theorem 1.2** (Dynamic case) *Assume that  $f$  and  $g$  are in  $C^{2,\alpha}(T^n)$  for some  $\alpha > 0$  and that  $N_k = k^n$ , i.e. the edge-length of the grid on  $T^n$  is equal to  $k^{-1}$ . For any sequence of discrete times  $m_k$  (iterations) such that  $m_k/k \rightarrow t$  we have*

$$\lim_{k \rightarrow \infty} u_{m_k}^{(k)} = u_t$$

*uniformly on  $T^n$ , where  $u_t$  is the smooth and strictly quasi-convex solution of the parabolic PDE 1.7 with initial data  $u_0 = 0$ . More precisely, for any positive integers  $k$  and  $m$*

$$\sup_{T^n} \left| u_m^{(k)} - u_{m/k} \right| \leq C \frac{m}{k} k^{-1},$$

*for a constant  $C$  independent of  $t$ . More generally, if at the initial discrete time  $m = 0$ ,*

$$u_{0|\Lambda_k}^{(k)} = u_{0|\Lambda_k},$$

*for a given strictly quasi-convex function  $u_0$  in  $C^{4,\alpha}(T^n)$ , then the corresponding result still holds (with a constant  $C$  depending on  $u_0$ ).*

An immediate consequence is that if the number  $m_k$  of iterations is too small, of the order  $o(k)$ , then  $u_{m_k}^{(k)} \rightarrow 0$  under the Sinkhorn iterations, i.e. ‘‘nothing happens’’. As discussed in Sect. 4.2 the estimate in the previous theorem can be expected to be sharp

under the regularity assumption in the theorem. Moreover, the proof of the theorem yields an essentially explicit control on the constant  $C$  appearing in the error bound.

Using that  $u_t$  converges exponentially to a potential  $u$  for the optimal transport problem, as  $t \rightarrow \infty$  (which follows from results in [44]) we deduce the following

**Corollary 1.3** (Constructive approximation of the potential  $u$ ) *Assume that  $f$  and  $g$  are in  $C^{2,\alpha}(T^n)$  for some  $\alpha > 0$  and that  $N_k = k^n$ , i.e. the “edge length” of the grid on  $T^n$  is equal to  $k^{-1}$ . There exists a positive constant  $A_0$  such that for any  $A > A_0$  the following holds: after  $m_k = \lfloor Ak \log k \rfloor$  iterations the corresponding quasi-convex functions  $u_k(x) := u_{m_k}^{(k)}(x)$  satisfy the estimate*

$$\sup_{T^n} |u_k - u| \leq Ck^{-1} \log k, \tag{1.10}$$

for some constant  $C$  (depending on  $A$ ), where  $u$  is a potential for the corresponding optimal transport map. Moreover, the discrete probability measures  $\gamma_k$  on  $T^n \times T^n$ , determined by the Sinkhorn algorithm, converge weakly towards the corresponding optimal transport plan  $(I \times (\nabla u + I))_*\mu$ , concentrating exponentially on the graph  $\Gamma$  of the transport map  $\nabla u + I$  :

$$\gamma_k \leq pk^p e^{-kd_{\Gamma_u}^2/p} \delta_{\Lambda_k}, \tag{1.11}$$

for some positive constant  $p$ , where  $d_\Gamma$  denotes the vertical distance to the graph  $\Gamma$  in  $T^n \times T^n$ , i.e.  $d_\Gamma(x, y) := d_{T^n}(y, \nabla u(x) + x)$  and  $\delta_{\Lambda_k}$  denotes the discrete uniform probability measure on the discrete torus  $\Lambda_k$ .

More generally, we will show in Sect. 5, that if  $f$  and  $g$  are in  $C^\infty(T^n)$ , then the previous theorem and its corollary still hold as long as the number of discretization points in the grid satisfies

$$N_k \geq C_\delta k^{n/2(1+\delta)}, \quad C_\delta > 0$$

for some  $\delta \in ]0, 1/2]$ . In particular, this means that one can then take a grid with larger edge-length

$$h = O(1/k^{1/2+\delta})$$

without affecting the quality of the approximation  $u_k$ , that is, without affecting the order  $O(k^{-1} \log k)$  of the corresponding error terms. In other words, for smooth data the error terms are nearly of order  $O(h^2)$  if the entropic regularization parameter is taken to be close to the square of the edge-length  $h$  of the grid.

Since each iteration in the Sinkhorn algorithm may be formulated in terms of matrix-vector multiplication, which requires  $O(N^2)$  arithmetic operators, and since the direct construction of the function  $u_k$  uses, in general,  $O(kN_k^2 \log k)$  elementary arithmetic operations. Moreover, in the present case of the torus the matrix-vector operations in question are discrete convolutions and can thus be performed using



merely  $O(N \log N)$  arithmetic operations, by using the Fast Fourier Transform (or  $O(N^{1+1/n})$  operations, using separability; see Sect. 6.3.1). Thus the construction of  $u_k$  in the previous corollary requires merely  $O(kN_k(\log N_k) \log k)$  arithmetic operations to obtain on error term of order  $O(k^{-1} \log k)$ . The same number of arithmetic operations (modulo a negligible term  $O(N_k)$  needed to for the summing) yields the following constructive approximation of the squared Optimal Transport distance  $d(\mu, \nu)$  (formula 1.4) with an additive error of the order  $O(k^{-1} \log k)$ :

**Corollary 1.4** (Constructive approximation of  $d(\mu, \nu)^2$ ) *Assume that  $f$  and  $g$  are in  $C^{2,\alpha}(T^n)$  for some  $\alpha > 0$ . There exists a positive constant  $A$  such that after  $m_k = \lfloor Ak \log k \rfloor$  iterations*

$$\frac{1}{2}d(\mu, \nu)^2 = \frac{1}{k} \sum_{i=1}^{N_k} p_i^{(k)} \log a_i^{(k)}(m_k) + \frac{1}{k} \sum_{i=1}^{N_k} q_i^{(k)} \log b_i^{(k)}(m_k) + O\left(\frac{1}{k} \log k\right)$$

and, as a consequence, such an equality also holds for the squared discrete Optimal Transport distance  $d(\mu^{(k)}, \nu^{(k)})^2$  (defined by formula 1.2 with  $C_{ij} = d(x_i, x_j)^2/2$ ).

Finally we point out that, by symmetry, Theorem 1.2 also shows that, on the one hand, the functions

$$v_{m_k}^{(k)}(x_i) := -k^{-1} \log \frac{b_i^{(k)}(m_k)}{q_i} \tag{1.12}$$

converge, as  $k \rightarrow \infty$ , towards the solution  $v_t$  of the parabolic equation obtained by interchanging the roles of  $\mu$  and  $\nu$ . On the other hand, by Lemma 3.1, the function  $v_{m_k}^{(k)}$  is equal to the Legendre transform (in the space variable) of  $u_{m_k}^{(k)}$ , up to a negligible  $O(k^{-1} \log k)$  error term. Thus Theorem 1.2 is consistent, as it must, with the fact that the Legendre transform of the solution  $u_t$  of Eq. 1.7 solves the parabolic equation obtained by interchanging the roles of  $\mu$  and  $\nu$  (as can be checked by a direct calculation).

## 1.2 Generalizations

### 1.2.1 The static case

The result in the static setting is shown to hold in a very general setting of optimal transport between two probability measures  $\mu$  and  $\nu$  defined on compact topological spaces  $X$  and  $Y$ , respectively: see Theorems 3.3, 3.4 and also Theorem 3.6 which, in particular, applies to the classical Euclidean setting where  $X$  and  $Y$  are convex domains in  $\mathbb{R}^n$  and  $c(x, y) := -x \cdot y$ . In the case when  $Y$  is convex the corresponding limit  $\phi$  obtained from the Sinkhorn iteration is the unique convex normalized solution to the second boundary value problem for Monge–Ampère equation in the interior of  $X$  :

$$e^{-g(\nabla\phi)} \det(\nabla^2\phi)dx = \mu, \quad (\partial\phi)(X) \subset Y. \tag{1.13}$$

The result holds without any regularity assumptions on  $\mu$  and  $g$  if Alexandrov’s classical notion of a Monge–Ampère measure is employed (see Sect. 3.2). Moreover, the

convergence holds even when the target set  $Y$  is not convex, but then  $(\partial\phi)(X)$  is contained in the convex hull of  $Y$  (see the discussion in Sect. 3.2).

In the general setting the roles of the positive vectors  $p^{(k)}$  and  $q^{(k)}$ , discretizing  $\mu$  and  $\nu$ , are played by two sequences  $\mu^{(k)}$  and  $\nu^{(k)}$ , satisfying certain density properties with respect to  $\mu$  and  $\nu$  (which are almost always satisfied in practice). Moreover, the cost function  $c$  is merely assumed to be continuous and can even be replaced by any sequence  $c_k$  converging uniformly to  $c$ , as the inverse  $k$  of the entropic regularization parameter tends to infinity (which applies, in particular, to the convolutional Wasserstein distances introduced in [63], where the matrix in formula 1.6 is replaced by a heat kernel; see Sect. 3.3).

### 1.2.2 The dynamic case

The corresponding result in the dynamic setting (Theorem 5.7) requires that  $X$  and  $Y$  be compact Riemannian manifolds and further regularity assumptions on  $c$ ,  $\mu$  and  $\nu$  (the case when  $X$  and  $Y$  have boundaries is left for the future). A “local density property” on the approximations  $\mu^{(k)}$  and  $\nu^{(k)}$  is required, roughly saying that the approximations of  $\mu$  and  $\nu$  hold up to length scales of the order  $k^{-1/2}$ , with an  $O(1/k)$ -error term. Interestingly, the local density properties turn out to be satisfied when the approximations  $\mu^{(k)}$  and  $\nu^{(k)}$  are defined by (weighted) point clouds, generated using *Quasi-Monte Carlo* methods for numerical integration [12,13,17]. The general results are applied to the case of the round *two-sphere* endowed with the two different cost functions: (1)  $d(x, y)^2$  and (2)  $-\log|x - y|$ . These two cases appear, for example, in applications to (1) computer graphics (texture mapping), medical imaging [30,72], mesh adaptation for global weather and climate prediction [70] and (2) the reflector antenna problem in geometric optics [37,69]. Nearly linear complexity of the corresponding Sinkhorn iteration is achieved in both cases, using fast transforms and  $O(N^{3/2})$ -complexity using separability.

### 1.3 Relation to previous results

To the best of the authors knowledge these are the first convergence results concerning the Sinkhorn algorithm (and its fixed points) in the limit when the number  $N$  of points and the inverse of the regularization parameter  $\epsilon$  jointly tend to infinity (see [21] and references therein for the static case when only  $\epsilon^{-1}(= k)$  tends to infinity in the Euclidean  $\mathbb{R}^n$ -setting and [47,48] for a very general setting). This kind of joint limit is, in practice, what is studied in numerical simulations in the context of geometric optimal transport (see for example [63] and the in-depth study in [35,59] on CPU and GPU hardware, respectively). Thus the convergence analysis in the present paper provides a theoretical bases for these simulations and yields concrete rates, under appropriate regularity assumptions (see Sect. 4.2 for a comparison with previous rates for the Sinkhorn algorithm). In particular, the present results rigorously establish and quantify the experimental observations in [59, Fig. 2], that the Sinkhorn algorithm converges after essentially  $O(\epsilon^{-1})$  iterations. Moreover, the experimental findings in [59, Fig. 2] that  $\epsilon$  can be taken to be close to the order  $O(h^2)$  on a grid with “edge

length”  $h$  are confirmed, if the data is assumed to be  $C^\infty$ -smooth (note that the data in [59, Fig. 2] is even real-analytic).

It should also be pointed out that the “change of variables” in formula 1.8 which plays an important role in the present paper, is also crucial for numerical simulations as it ensures numerical stability, as emphasized in [27,35,59]. As stressed in [35] the corresponding log-sum-exp KeOps routines [22], used in [35] to implement the iteration  $u_m \rightarrow u_{m+1}$  on GPU hardware, are just as efficient as matrix-vector products with the kernel matrix  $K_{ij}$ . Moreover, the present results provide a theoretical justification for the kernel truncations employed in [35,59]. To briefly explain this we recall that the starting point of the stabilization scheme advocated in [59] (see also [27, Remark 4.22]) is to write the iteration in terms of the variables  $(u_{m+1} - u_m, v_{m+1} - v_m)$ . In the setting of a general cost function  $c(x, y)$  this corresponds to replacing the matrix kernel  $e^{-kc(x,y)}$  with the “stabilized” kernel

$$e^{-kc(x,y)} e^{-ku_{m-1}(x)} e^{-kv_{m-1}(y)}$$

By Theorems 1.2, 5.7 (and the argument in the proof 1.11) the latter kernel is, for  $k$  large, exponentially concentrated on the graph of  $\Gamma_t$  in  $X \times Y$  of the diffeomorphism  $F_{u_t}$  corresponding to the parabolic solution  $u_t$  for  $t = (m - 1)/k$ . This means that the corresponding  $N_k \times N_k$  matrices  $(e^{-kc(x_i,y_j)} e^{-ku_{m-1}(x_i)} e^{-kv_{m-1}(y_j)})$  are effectively sparse. Building on the results in the present paper this is exploited in [6] to modify the Sinkhorn algorithm in order to obtain a numerically stable algorithm on the torus, which is shown to have nearly  $O(N)$ -complexity at each iteration.

### 1.3.1 Comparison with other numerical schemes for optimal transport

The quantitative convergence in Corollary 1.3 should be compared with previous results in the rapidly growing literature on numerical approximations schemes for solutions to Optimal Transport problems and the corresponding Monge–Ampère type equations. However, the author is not aware of any previous results providing both complexity bounds (in terms of  $N$ ) and a quantified rate of convergence of the error of the approximate solution, as  $N \rightarrow \infty$ . Recall that a time-honored approach is to approximate the optimal transport potential  $u$  by solving the linear program which is dual to the discretized optimal transport problem. This can be done using combinatorial algorithms. However, they do not scale well for large  $N$  (see, for example, the exposition in [14], where applications to cosmology are given in the periodic setting). Moreover, it is not clear how to establish quantitative convergence rates for the convergence towards  $u$ . Another influential approach is the Benamou–Brenier Augmented Lagrangian approach, using computational fluid mechanics, introduced in [2] in the periodic setting (numerical experiments suggest that it has  $O(N^3)$  time-complexity, as pointed out in [4]). There is also a rapidly expanding literature concerning other discretization schemes in the field of numerical analysis of PDEs, mainly concerning the case of domains in  $\mathbb{R}^n$  (as in Sect. 3.2) and the periodic case of the torus (as in Sect. 1.1.3). Here we will focus on the schemes where convergence results—analogueous to Theorems 3.6, 1.1 - have been established. The mostly studied approaches are based on either *finite differences* [4,5] or *semi-discrete Optimal Transport* [46] (which goes

back to the classical work of Alexandrov and Pogorelov). Then the fully non-linear Monge–Ampère equation for the potential  $u$  (and the corresponding boundary conditions 1.13 or periodicity conditions) is replaced with a finite dimensional non-linear algebraic equation for a “discrete” function  $u_h$ , where  $h$  denotes the corresponding spatial resolution. In our torus setting we thus have  $h = 1/k$  (defined as the entropic regularization parameter) when the data is  $C^{2,\alpha}$  smooth and for  $C^\infty$ -data  $h$  can be taken arbitrarily close to  $1/k^{1/2}$ . The equation for  $u_h$  may be expressed as the fixed point condition for a non-linear map  $S_h$  :

$$S_h(u_h) = u_h,$$

(whose role in our setting is played by the scaled logarithm of the Sinkhorn operator defined by formula 2.20). In practice, the map  $S_h$  is often taken to be a Newton type iteration. In experiments it has been computed in almost linear time-complexity  $O(N_h)$  [4,5]. Moreover, merely a few dozen iterations  $S_h^{(m_h)}(u_0)$  appear to be needed in order to obtain a good approximation of  $u_h$ . In the semi-discrete approach the convergence of  $u_h$  towards  $u$ , as  $h \rightarrow 0$ , follows directly from basic stability results for optimal transport plans. Moreover, in the case of finite difference approach the convergence is established in [4] in the setting of domains. Another discretization of the Optimal Transport problem in domains, called the *logarithmic discrete Monge–Ampère optimization problem*, is introduced in [49] and the corresponding solution  $u_h$  is shown to converge towards  $u$ . Very recently, a general convergence framework is introduced in [38], showing how to slightly modify a range of existing numerical schemes (including [5]) to establish the convergence of the corresponding discrete solutions  $u_h$  towards  $u$  in the setting of domains.

However, the problem of establishing convergence rates as  $h \rightarrow 0$  is still open in these schemes (see the discussion in the survey [34]). From this point of view one of the main points of the present paper is to rigorously quantify how many iterations  $m_h$  are needed in the setting of the Sinkhorn iteration in order to approximate the solution  $u$  to nearly order  $O(h)$ , for  $C^{2,\alpha}$ -data (and nearly order  $O(h^2)$  for  $C^\infty$ -data). The answer is that nearly  $O(h^{-1})$  iterations are needed, according to Corollary 1.3 (see the discussion in Sect. 4.2). This is in line with the experimental findings in [35,59] (see, in particular, [35, Figure 3.19] where  $\epsilon := k^{-1} = 10^{-4}$  and [59, Fig. 2]).

**Remark 1.5** Interestingly, as discussed in detail in [35,59], heuristics inspired by simulated annealing in numerics (aka  $\epsilon$ -scaling) and multi-scale techniques can be used to effectively reduce the number of Sinkhorn iterations needed to approximate the optimal transport potential  $u$  to a few dozen. In a nutshell, the idea is to fine-tune and gradually decrease both the parameter  $\epsilon$  and the spatial resolution  $h$  during the Sinkhorn iterations (see [35, Figure 3.26]). One can anticipate that variants of Theorems 1.2, 5.7 will play a role in making these heuristics and experimental findings mathematically rigorous with quantitative error estimates. The local density property in Definition 5.6 should be useful in this regard (note that the length scale  $k^{-1/2}$  ( $= \epsilon^{1/2}$ ) appearing in Definition 5.6 corresponds to the notion of blurring scale in [35]). On the other hand, as discussed in Sect. 7, it is of independent interest to be able to approximate the solution  $u_t$  of the parabolic optimal transport equations at any finite time  $t$

and then  $\epsilon$  must be kept fixed during the iterations (so that an  $O(\epsilon)$ -approximation of  $u_t$  is obtained after  $O(\epsilon^{-1}t)$  iterations, if logarithmic factors are ignored).

In the setting of semi-discrete optimal transport the convergence of a damped Newton iteration towards  $u_h$  is established in [46] (under similar assumptions as in Theorem 5.10) at a linear rate (in the exponential sense). Furthermore, rates of convergence of  $u_h$  towards  $u$  are established in [11]. However, the damping parameter and the rate established in [46] depends on the discrete solution  $u_h$  (and hence on  $h$ ) in a rather complicated way and degenerates as  $N_h$  is increased, i.e. when  $h$  is decreased; see [46, Remark 1.3] and compare also with the discussion in Sect. 4.2. This means that, for the moment, the analog of Corollary 1.3 in the semi-discrete setting seems to be out of reach.

A damped Newton approach has also been applied directly to the Monge–Ampère equation and shown to converge to  $u$  at a linear rate in the periodic setting in [51] when the target density is constant and then in [58] in the general periodic setting. In these approaches the iteration  $u^{(m+1)}$  is obtained by inverting a second order linear elliptic operator (in non-divergence form) depending on  $u^{(m)}$ . Various discretizations of these schemes are studied experimentally in [42, 51, 58, 70].

One advantage of the Sinkhorn framework over many other approaches, when applied to general manifolds, is that it is meshfree. In other words, it does not require generating a grid or a polyhedral tessellation of the manifolds, but only a suitable point cloud, which can be efficiently generated using Quasi-Monte Carlo methods. In the case of the round sphere various different numerical algorithms have previously been explored in the literature: see [30, 70, 72] for experimental work on the case of the cost function  $d(x, y)^2$  and [18, 28] for the case of the cost function  $-\log|x - y|$ , as applied to the reflector antenna problem in geometric optics.

### 1.3.2 Kähler geometry

The present results are very much inspired by an analogous setup which appears in complex (Kähler) geometry. Briefly, the role of the Sinkhorn algorithm is then played by Donaldson’s iteration, introduced in [32], whose fixed points are called balanced metrics and  $k$  appears as the power of a given ample line bundle  $L$  over  $X$  with  $e^{-ku}$  playing the role of a Hermitian metric on  $L$ . Moreover, the role of the function  $\rho_{ku}$  (formula 2.23) is played by the (normalized) point-wise norm of the Bergman kernel on the diagonal, induced by the pair  $(u, \mu)$ . From this point of view the “static case” of Theorem 1.1 (and its generalization Theorem 3.3) is the analog of [10, Thm B] and the “dynamic case” of Theorem 1.2 is the analog of the result in [8] showing that Donaldson’s iteration converges to the Kähler-Ricci flow [20], as conjectured in [32]. In fact, identifying the real torus  $T^n$  with a reduction of the complex torus  $X := \mathbb{C}^n / (\mathbb{Z}^n + i\mathbb{Z}^n)$  the parabolic flow 1.7 can, in the case when  $g$  is constant be identified with a twisted Kähler-Ricci flow [20] whose stationary solutions are Kähler potentials solving the corresponding complex Monge–Ampère equation (known as the Calabi–Yau equation in this context).<sup>1</sup> But it should be stressed that a new feature of

<sup>1</sup> When  $f$  and  $g$  are constant the corresponding twisted Kähler-Ricci flow coincides with Hamilton’s Ricci flow restricted to the space of Kähler metrics.

the analysis in the present paper, compared to the usual situation in Kähler geometry (apart from allowing a non-uniform target measure  $\nu$ , i.e. a non-constant  $g$ ) is that the source measure  $\mu$  is taken to be *discrete* and depend on  $k$ , i.e. it is given by a discrete sequence  $\mu^{(k)}$ . In practice, such discretizations are used in implementations of Donaldson's iteration, such as the experimental work [33], motivated by String Theory. The discreteness of  $\mu^{(k)}$  leads to various technical complications, that do not seem to have been studied rigorously in the Kähler geometry setting. Interestingly, the density condition on the sequence  $\mu^{(k)}$  appearing in Lemma 3.1 can be viewed as a real analog of the Bernstein–Markov property for a sequence  $\mu^{(k)}$ , as studied in the complex geometric and pluripotential theoretic setting (see the discussion on page 8 in [9]). The relations between the real and complex settings will be expanded on elsewhere.

## 1.4 Organization

In Sect. 2 a general setting for iterations on  $C(X)$ , generalizing the Sinkhorn algorithm, is introduced. The iteration in question, which is determined by a triple  $(\mu, \nu, c)$ , is essentially equivalent to the Iterative Proportional Fitting Procedure and the results in Sect. 2 are probably more or less well-known (except perhaps Theorem 2.9). But one point of the presentation is to exploit the variational structure. It can be viewed as a real analogue of the formalism introduced in [8], in the setting of Donaldson's iteration [32] and it lends itself to various generalizations of the optimal transport problem (such as Monge–Ampère equations with exponential non-linearities). In Sect. 3 the variational structure is used to give a general convergence result for the Sinkhorn fixed points, which when specialized to the torus setting yields Theorem 1.1. Applications to the second boundary value problem for the Monge–Ampère operator in  $\mathbb{R}^n$  and to convolutional Wasserstein distances are also given. Then in Sect. 4 the convergence of the Sinkhorn iteration towards the parabolic Optimal Transport equations on the torus is shown (Theorem 1.2). The proof leverages the regularity theory and a priori estimates of the corresponding parabolic PDE on the torus (shown in “Appendix B”). In the following Sect. 5 the result is generalized to a rather general setting of optimal transport on compact manifolds. In Sect. 6 it is shown that nearly linear complexity can be achieved in the case of optimal transport on the torus and the sphere (which applies, in particular, to the reflector antenna problem). Section 7 gives an outlook on relations to singularity formation in the parabolic optimal transport equations. In particular, the variational approach to the Sinkhorn iteration introduced in Sect. 2 is exploited in order to propose a generalized notion of solution to the corresponding parabolic PDE. In “Appendix A” a proof of a discrete version of the classical stationary phase approximation is provided.

## 2 General setup and preliminaries

If  $Z$  is a compact topological space then we will denote by  $C(Z)$  the space of continuous functions on  $Z$  endowed with the sup-norm and by  $\mathcal{P}(Z)$  the space of all (Borel)

probability measures on  $Z$ , endowed with the weak topology. Given a subset  $S$  of  $Z$  we will denote by  $\chi_S$  the function which is equal to 0 on  $S$  and infinity otherwise.

Throughout the paper we will assume given a triple  $(\mu, \nu, c)$  where  $\mu$  and  $\nu$  are probability measures on compact topological spaces  $X$  and  $Y$ , respectively and a function  $c$  on  $X \times Y$ . The function  $c$  will be assumed to be continuous in all sections except in Sect. 5 (where we assume that  $c$  is lower semi-continuous). The supports of  $\mu$  and  $\nu$  will be denote by  $X_\mu$  and  $Y_\nu$ , respectively. Given  $u \in C(X)$  and  $v \in C(Y)$  we will, abusing notation slightly, identify  $u$  and  $v$  with their pull-backs to  $X \times Y$ .

## 2.1 Recap of optimal transport and the $c$ -Legendre transform

Let us start by recalling the standard setup for optimal transport (see the book [67] for further background). A probability measure  $\gamma \in \mathcal{P}(X \times Y)$  is said to be a *transport plan (or coupling) between  $\mu$  and  $\nu$*  if the the push forwards of  $\gamma$  to  $X$  and  $Y$  are equal to  $\mu$  and  $\nu$ , respectively. The subspace of all such probability measures in  $\mathcal{P}(X \times Y)$  will be denote by  $\Pi(\mu, \nu)$ . A transport plan in  $\Pi(\mu, \nu)$  is said to be *optimal wrt the cost function  $c$* , if it realizes the following infimum:

$$\inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c \gamma$$

By weak compactness such an optimal transport plan always exists. The  *$c$ -Legendre transform  $u^c$*  of a function  $u \in C(X)$  is defined as the following function in  $C(Y)$

$$u^c(y) := \sup_{x \in X} (-c(x, y) - u(x)).$$

Similarly, if  $v \in C(Y)$  then  $v^c$  is the function on  $C(X)$  defined by replacing  $u$  in the previous formula with  $v$  and taking the sup over  $Y$ . A function  $u \in C(X)$  is said to be  *$c$ -convex* if

$$(u^c)^c = u$$

Equivalently,  $u$  is  $c$ -convex iff there exists some  $v \in C(Y)$  such that  $u = v^c$ . Indeed, this follows from the observation that  $u^{ccc} = u^c$  for any  $u \in C(X)$ , which in turn follows from  $u^{cc} \leq u$ . The following functional on  $C(X)$  will be called the *Kantorovich functional*:

$$J(u) := \int u \mu + \int u^c \nu. \quad (2.1)$$

**Proposition 2.1** (“Optimality criterion”) *A transport plan  $\gamma \in \Pi(\mu, \nu)$  is optimal iff there exists  $u \in C(X)$  which is  $c$ -convex and such that  $\gamma$  is supported in*

$$\Gamma_u := \{(x, y) \in X \times Y : u(x) + u^c(y) + c(x, y) = 0\} \quad (2.2)$$

Moreover, if this is the case then

$$\int_{X \times Y} c\gamma = -J(u)$$

**Proof** This is standard and known as the Knott–Smith optimality criterion (in the Euclidean setting) [67]. For completeness we provide the simple proof of the direction that we shall use later on. Since  $u + u^c + c \geq 0$  on  $X \times Y$  the following lower bound holds for any given  $\gamma \in \Pi(\mu, \nu)$

$$\int_{X \times Y} c\gamma \geq - \inf_{u \in C(X)} J(u) \tag{2.3}$$

Now, if  $\gamma$  is supported in  $\Gamma_u$  it follows directly that  $\int_{X \times Y} c\gamma = -J(u)$  and hence  $\gamma$  attains the lower bound above, i.e.  $\gamma$  is optimal.  $\square$

**Remark 2.2** A byproduct of Theorem 3.3 below (applied to the case when  $\mu_k = \mu$  and  $\nu_k = \nu$  for all  $k$ ) is a proof that there always exists a transport plan  $\gamma_*$  with support in  $\Gamma_u$  for some  $c$ -convex function. Since  $\gamma_*$  saturates the lower bound 2.3 it follows that taking the infimum over all  $\gamma$  in  $\Pi(\mu, \nu)$  yields equality in 2.3. As a consequence,

$$\begin{aligned} \inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c\gamma &= \sup_{(u, v) \in \Phi_c} \int u\mu \\ &+ \int vv, \quad \Phi_c := \{(u, v) \in C(X) \times C(Y) : u + v \leq c\} \end{aligned} \tag{2.4}$$

This is the content of Kantorovich duality, which is usually shown using Rockafeller-Fenchel duality in topological vector spaces [67].

### 2.1.1 The torus setting

In section we consider the case when  $X = Y = T^n := \mathbb{R}^n / \mathbb{Z}^n$  and the cost function  $c := d_{T^n}^2 / 2$  is half the squared standard distance function on  $T^n$ . We will identify a function  $u$  on  $T^n$  with a  $\mathbb{Z}^n$ -periodic function on  $\mathbb{R}^n$  in the usual way. Similarly, we identify the cost function  $c$  on  $T^n$  with a function on  $\mathbb{R}^n \times \mathbb{R}^n$ , which is  $\mathbb{Z}^n$ -periodic in each argument

$$c(x, y) := \frac{1}{2} d_{T^n}(x, y)^2 := \frac{1}{2} \inf_{m \in \mathbb{Z}^n} |x + m - y|^2 \tag{2.5}$$

Note that  $c(x, \cdot)$  is Lipschitz with Lipschitz constant  $\sqrt{n}$  on  $T^n$  (endowed with its standard metric). As a consequence any  $c$ -convex function  $u$  on  $T^n$  is also Lipschitz with Lipschitz constant  $\sqrt{n}$ . In this particular case a  $c$ -convex function will be called *quasi-convex* and we will say that  $u \in C^2(T^n)$  is *strictly quasi-convex* if  $\nabla^2 u + I > 0$ .



**Lemma 2.3** *Let  $u \in C^2(T^n)$  be strictly quasi-convex. Then*

- the map

$$x \mapsto y_x := x + (\nabla u)(x) \tag{2.6}$$

*defines a  $C^1$ -diffeomorphism of  $T^n$ .*

- $u^c$  is also a strictly quasi-convex  $C^2$ -function on  $T^n$  and the corresponding map

$$y \mapsto x_y := y + (\nabla u^c)(y) \tag{2.7}$$

*is the inverse of the map 2.6 and the following matrix relation holds*

$$(\nabla^2 u + I)(x_y)^{-1} = (\nabla^2 u^c + I)(y) \tag{2.8}$$

*Conversely, if  $u \in C^2(T^n)$  is quasi-convex and the map 2.6 is a  $C^1$ -diffeomorphism of  $T^n$ , then  $u$  is strictly quasi-convex. Moreover, if  $u \in C^k(T^n)$  is strictly quasi-convex, for  $k$  a positive integer and  $k \geq 2$ , then  $u^c \in C^k(T^n)$ .*

**Proof** Given a function  $\phi$  on  $\mathbb{R}^n$  we denote by  $\phi^*$  its classical Legendre transform:

$$\phi^*(y) := \sup_{x \in \mathbb{R}^n} x \cdot y - \phi(x) \tag{2.9}$$

(in other words, this is the  $c$ -Legendre transform wrt  $c(x, y) := -x \cdot y$ ). Given a  $\mathbb{Z}^n$ -invariant quasi-convex function  $u$  on  $\mathbb{R}^n$  we set  $\phi(x) := u(x) + |x|^2/2$ . Then it follows directly from the definitions that  $\phi$  is convex and

$$\phi^*(y) = u^c(y) + |y|^2/2, \quad x + (\nabla u)(x) = \nabla \phi(x), \quad \nabla^2 u + I = \nabla^2 \phi \tag{2.10}$$

Next note that since  $u^c$  is continuous and periodic it is bounded and hence  $\phi^*$  is finite on all of  $\mathbb{R}^n$  with quadratic growth. As a consequence, given  $y \in \mathbb{R}^n$ , the function  $x \mapsto x \cdot y - \phi(x)$  attains its sup at some point  $x_y \in \mathbb{R}^n$

$$\phi^*(y) = x_y \cdot y - \phi(x_y) \tag{2.11}$$

and since the point is a local maximum we have  $y = \nabla \phi(x_y)$ . This shows that  $\nabla \phi$  maps  $\mathbb{R}^n$  surjectively onto  $\mathbb{R}^n$ . Moreover, since  $\nabla^2 \phi > 0$  the  $\phi$  function is strictly convex and  $x_y$  is uniquely determined. Thus the  $C^1$ -map  $\nabla \phi$  is a bijection and moreover its inverse  $y \mapsto x_y$  is also a  $C^1$ -map (by the inverse function theorem), which proves the first claim in the lemma. Moreover, since  $x_y$  is the unique minimizer of the sup defining  $\phi^*$  the function  $\phi^*$  is differentiable with gradient  $x_y$  at  $y$ . Hence, the  $C^1$ -inverse of  $\nabla \phi$  is given by  $\nabla \phi^*$ , showing that  $\phi^*$  is in  $C^2(\mathbb{R}^n)$ . Differentiating the identity  $\nabla \phi^* \circ \nabla \phi = I$  finally proves 2.8 and the last statement follows from the implicit function theorem:  $\nabla \phi$  is  $C^{k-1}$  implies (since  $\nabla^2 \phi > 0$ ) that its inverse  $\nabla \phi^*$  is also  $C^{k-1}$ . Finally, if  $u \in C^2(T^n)$  is quasi-convex and the map 2.6 is a diffeomorphism of  $T^n$ , then differentiating  $(\nabla \phi)^{-1} \circ \nabla \phi = I$  reveals that the non-negative matrix  $\nabla^2 \phi$  is non-degenerate, hence strictly positive, i.e  $u$  is strictly quasi-convex. □

We will also have use for the following

**Lemma 2.4** *Assume that  $u$  is  $C^1$ -smooth and strictly quasi-convex. Then, for any fixed  $y \in T^n$ , the unique infimum of the function  $x \mapsto c(x, y) + u(x)$  on  $T^n$  is attained at  $x = x_y$  (defined by formula 2.7). Moreover, the function  $x \mapsto c(x, y)$  is smooth on some neighborhood of  $x_y$  in  $T^n$  and its Hessian is equal to the identity there.*

**Proof** First observe that, given  $y \in T^n$ , the infimum in question is attained at  $x_y$  (defined by formula 2.7), as follows directly from combining formulas 2.11 and 2.10. Representing  $x_y$  and  $y$  with points in  $\mathbb{R}^n$  we thus have

$$d_{T^n}(x_y, y)^2 = \inf_{m \in \mathbb{Z}^n} |x_y + m - y|^2 = |x_y + m_0 - y|^2$$

for some  $m_0 \in \mathbb{Z}^n$ . We claim that, under the assumptions of the lemma, the inf above is uniquely attained at  $m_0$ , i.e.

$$m \neq m_0 \implies |x_y + m - y|^2 > |x_y + m_0 - y|^2.$$

To see this we note that, since  $u$  is periodic, when viewed as a function on  $\mathbb{R}^n$ , we have

$$\inf_{x \in \mathbb{R}^n} d_{T^n}(x, y)^2/2 + u(x) = \inf_{x \in \mathbb{R}^n} |x - y|^2/2 + u(x) \tag{2.12}$$

Since the inf in the left hand side above is attained at  $x_y$  so is the inf in the right hand side. Now assume, to get a contradiction, that the claim above does not hold, i.e. there exists a non-zero  $m \in \mathbb{Z}^n$  such that  $|x_y + m - y| = |x_y + m_0 - y|$ . This implies that the inf in the right hand side in formula 2.12 is attained both at  $x_y$  and at  $x_y + m$  (since  $u$  is periodic). But this contradicts the fact that the function  $x \mapsto |x - y|^2/2 + u(x)$  is strictly convex on  $\mathbb{R}^n$  (by the assumed strict quasi-convexity of  $u$  on  $T^n$ ). Finally, the claim shows, since the inequality in the claim is preserved when  $\bar{x}$  is perturbed slightly, that  $d_{T^n}(x, y)^2 = |x - y|^2$  for all  $x$  sufficiently close to  $x_y$ . Hence,  $x \mapsto d_{T^n}(x, y)^2/2$  is smooth there and its Hessian is constant, as desired.  $\square$

### 2.2 The log Sinkhorn iteration on $C(X)$

In this section we will consider an iteration on  $C(X)$ , which can be viewed as a reformulation of the Sinkhorn algorithm and the Iterative Proportional Fitting Procedure, recalled in Sect. 1.1.1 (see Sect. 2.3.1). Given data  $(\mu, \nu, c)$ , as in Sect. 2.1, we first introduce the following maps

$$T_\mu : C(X) \rightarrow C(Y), \quad u \mapsto v[u] := \log \int e^{-c(x, \cdot) - u(x)} \mu(x)$$

and

$$T_\nu : C(Y) \rightarrow C(X), \quad v \mapsto u[v] := \log \int e^{-c(\cdot, y) - v(y)} \nu(y)$$

(abusing notation slightly we will write  $T_\mu(u) = v[u]$  etc). This yields an iteration on  $C(X)$  defined by

$$u_{m+1} := S[u_m], \quad (2.13)$$

where  $S$  is defined as the the composed operator  $T_v \circ T_\mu$  on  $C(X)$  :

$$S : C(X) \rightarrow C(X), \quad u \mapsto u[v[u]]$$

In lack for a better name the iteration 2.13 will be called the *log Sinkhorn iteration* and the operator  $S$  will be called the *log Sinkhorn operator*. It will be convenient to rewrite it as the following difference equation:

$$u_{m+1} - u_m = \log(\rho_{u_m}), \quad (2.14)$$

where  $\rho_u$  is defined by

$$\rho_u := e^{S[u]-u} \quad (2.15)$$

and has the property that  $\rho_u \mu$  is a probability measure on  $X$  (as follows directly from the definitions).

In this section we will use a variational approach to study the log Sinkhorn iteration. An alternative approach will also be used in Sect. 4, which relies on the observation that the log Sinkhorn iteration contracts the  $L^\infty$ -distance on  $C(X)$  (see Step 2 in the proof of Lemma 4.4).

### 2.2.1 Existence and uniqueness of fixed points

Consider the following functional  $\mathcal{F}$  on  $C(X)$  :

$$\mathcal{F} := I_\mu - \mathcal{L}, \quad I_\mu(u) = \int_X u \mu, \quad \mathcal{L}(u) := - \int_Y v[u] \nu. \quad (2.16)$$

Note that  $I_\mu$  and  $\mathcal{L}$  are equivariant under the additive action of  $\mathbb{R}$  and hence  $\mathcal{F}$  is invariant.

**Remark 2.5** This functional can be viewed as an analog of the Kantorovich functional  $J(u)$  (formula 2.1), where the  $c$ -Legendre transform  $u^c$  is replaced by  $v[u]$ . From a numerical perspective this amounts to replacing the supremum defining  $u^c$  by a “soft max” [27]. It is well-known that  $\mathcal{F}$  decreases under the log Sinkhorn iteration, i.e. that  $\mathcal{F}(Su) \leq \mathcal{F}(u)$ . This is usually shown using block coordinate descent on the “dual functional” to the “primal” minimization problem in Prop 2.12; see [27, Prop 4.21] and [66, Section 3.1]. But here we observe that, in fact, *both* functionals  $I_\mu$  and  $-\mathcal{L}$  decrease along the iteration (see Step 1 in the proof of Theorem 2.9). This will be important in the proof of the last step of Theorem 2.9 and in the proof of Prop 7.1.

**Lemma 2.6** *The following is equivalent:*

- $u$  is a critical point for the functional  $\mathcal{F}$  on  $C(X)$
- $\rho_u = 1$  a.e. with respect to  $\mu$

Moreover, if  $u$  is a critical point, then  $u_* := S(u)$  is a fixed point for the operator  $S$  on  $C(X)$

**Proof** First observe that the differential of the functional  $\mathcal{L}$  defined in formula 2.16, at an element  $u \in C(X)$ , is represented by the probability measure  $\rho_u \mu$ , where  $\rho_u$  is defined by formula 2.15. This means that for any  $\dot{u} \in C(X)$

$$\frac{d(\mathcal{L}(u + t\dot{u}))}{dt} \Big|_{t=0} = \int \dot{u} \rho_u \mu.$$

This follows readily from the definitions by differentiating  $t \mapsto v[(u + t\dot{u})]$  to get an integral over  $(X, \mu)$  and then switching the order of integration. As a consequence,  $u$  is a critical point of the functional  $\mathcal{F}$  on  $C^0(X)$  iff  $\rho_u \mu = \mu$ , i.e. iff  $\rho_u = 1$  a.e. with respect to  $\mu$ . Finally, if this is the case then  $S(u) = u$  a.e wrt  $\mu$  and hence  $S(S(u)) = S(u)$  (since  $S(f)$  only depends on  $f$  viewed as an element in  $L^1(X, \mu)$ ). □

The following basic compactness property holds:

**Lemma 2.7** *Given a point  $x_0 \in X$  the subset  $\mathcal{K}_{x_0}$  of  $C(X)$  defined as all elements  $u$  in the image of  $S$  satisfying  $u(x_0) = 0$  is compact in  $C(X)$ .*

**Proof** First observe that, since  $X \times Y$  is assumed compact, the continuous function  $c$  is, in fact, uniformly continuous on  $X$ . Hence, it follows from the very definition of  $S$  that  $S(C(X))$  is an equicontinuous family of continuous functions on  $X$ . By Arzela-Ascoli theorem it follows that the set  $\mathcal{K}_{x_0}$  is compact in  $C(X)$ . □

Using the previous two lemmas gives the following

**Proposition 2.8** *The operator  $S$  has a fixed point  $u_*$  in  $C(X)$ . Moreover,  $u_*$  is uniquely determined a.e. wrt  $\mu$  up to an additive constant and  $u_*$  minimizes the functional  $\mathcal{F}$ . More precisely, there exists a unique fixed point in  $S(C(X))/\mathbb{R}$ .*

**Proof** We start by noting that

$$\mathcal{F}(Su) \leq \mathcal{F}(u)$$

(this is shown in the first step of Theorem 2.9 below). Since  $\mathcal{F}$  is invariant under the natural  $\mathbb{R}$ -action we conclude that

$$\inf_{C(X)} \mathcal{F} = \inf_{\mathcal{K}_0} \mathcal{F},$$

where  $\mathcal{K}_0$  denotes the compact subset of  $C(X)$  appearing in Lemma 2.7. Since  $\mathcal{F}$  is clearly continuous on  $C(X)$  this implies the existence of a minimizer of  $\mathcal{F}$  which is moreover in  $\mathcal{K}_0$ .

Next observe that  $\mathcal{F}$  is convex on  $C(X)$ . Indeed, for any fixed  $y \in Y$ ,  $u \mapsto v[u](y)$  is convex on  $C(X)$ , as follows directly from Jensen’s inequality. Hence,  $-\mathcal{L}$  is convex and since  $I_\mu$  is affine we conclude that  $\mathcal{F}$  is convex. More precisely, Jensen’s (or

Hölder’s inequality implies that  $\mathcal{F}$  is strictly convex on  $C(X)/\mathbb{R}$  viewed as a subset of  $L^1(\mu)/\mathbb{R}$ . Hence, if  $u_0$  and  $u_1$  are two minimizers, then there exists a constant  $C$  such that  $u_0 = u_1 + C$  a.e. wrt  $\mu$ . In particular, if  $C = 0$  then  $u_* := S(u_0) = S(u_1)$  gives the same fixed point of  $S$ .  $\square$

### 2.2.2 Monotonicity and convergence properties of the iteration

We next establish the following result, which can be seen as a refinement, in the present setting, of the convergence of the general Iterative Proportional Fitting Procedure established in [55]. The result will be used in the proof of Proposition 7.1.

**Theorem 2.9** *Given  $u_0 \in C(X)$  the corresponding iteration  $u_m := S^m u_0$  converges uniformly to a fixed point  $u_\infty$  of  $S$ .*

**Proof** *Step 1*  $I_\mu$  and  $-\mathcal{L}$  are decreasing along the iteration and hence  $\mathcal{F}$  is also decreasing. The functionals are strictly decreasing at  $u_m$  unless  $S(u_*) = u_*$  for  $u_* := S(u_m)$ .

Using the difference equation 2.14 for  $u_m$  and Jensen’s inequality, we have

$$I_\mu(u_{m+1}) - I_\mu(u_m) = \int \log \rho_{u_m} \mu \leq \log \int \rho_{u_m} \mu = \log 1 = 0$$

Moreover, equality holds unless  $\rho_{u_m} = 1$  a.e wrt  $\mu$  i.e.  $S(u_m) = u_m$  and  $S(u_*) = u_*$  everywhere on  $X$ . Similarly, by symmetry,

$$\mathcal{L}(u_m) - \mathcal{L}(u_{m+1}) = \int \log \rho_{v_m} v \leq \log \int \rho_{v_m} v = \log 1 = 0,$$

where now  $\rho_v$ , for  $v \in C(Y)$ , denotes the probability measure on  $Y$  defined as in formula 2.15, but with the roles of  $\mu$  and  $v$  interchanged.

*Step 2 Convergence in  $C(X)/\mathbb{R}$ .*

Given the initial data  $u_0$  we denote by  $\mathcal{K}_{u_0}$  the closure of the orbit of  $u_0$  in  $C(X)$  under repeated application of  $S$ . By Lemma 2.7  $\mathcal{K}_{u_0}/\mathbb{R}$  is compact in  $C(X)/\mathbb{R}$ . Hence, after perhaps passing to a subsequence,  $u_m \rightarrow u_\infty$  in  $C(X)/\mathbb{R}$ . Now, since  $\mathcal{F}$  is decreasing along the orbit we have

$$\mathcal{F}(u_\infty) = \inf_{\mathcal{K}_0} \mathcal{F}.$$

Hence, by the condition for strict monotonicity it must be that  $Su_\infty = u_\infty$  a.e. wrt  $\mu$  and hence, since  $u_\infty$  is the image of  $S$ , it follows that  $Su_\infty = u_\infty$  on all of  $X$ . It then follows from Proposition 2.8 that  $u_\infty$  is uniquely determined in  $C(X)/\mathbb{R}$  (by the initial data  $u_0$ ), i.e. the whole sequence converges in  $C(X)/\mathbb{R}$ .

*Step 3 Convergence in  $C(X)$*

Let us first show that there exists a number  $\lambda \in \mathbb{R}$  such that

$$\lim_{m \rightarrow \infty} I_\mu(u_m) = \lambda. \tag{2.17}$$

By Step 1  $I_\mu$  is decreasing and hence it is enough to show that  $I_\mu(u_m)$  is bounded from below. But  $I_\mu = \mathcal{F} + \mathcal{L}$ , where, by Prop 2.8 (or the previous step)  $\mathcal{F}$  is bounded from below (by  $\mathcal{F}(u_\infty)$ ). Moreover, by the first step  $\mathcal{L}(u_m) \geq \mathcal{L}(u_0)$ , which concludes the proof of 2.17. Next, decompose

$$u_m = \tilde{u}_m + u_m(x_0),$$

By Lemma 2.7 the sequence  $(\tilde{u}_m)$  is relatively compact in  $C(X)$  and we claim that  $|u_m(x_0)| \leq C$  for some constant  $C$ . Indeed, if this is not the case then there is a subsequence  $u_{m_j}$  such that  $|u_{m_j}| \rightarrow \infty$  uniformly on  $X$ . But this contradicts that  $I_\mu(u_m)$  is uniformly bounded (by 2.17). It follows that the sequence  $(u_m)$  is also relatively compact. Hence, by the previous step the whole sequence  $u_m$  converges to the unique minimizer  $u_*$  of  $\mathcal{F}$  in  $S(C(X))$  satisfying  $I_\mu(u_*) = \lambda$ .  $\square$

**Remark 2.10** The convergence result in [55] is, in the present setting, equivalent to convergence of the induced iteration on  $C(X)/\mathbb{R}$ . In fact, the latter convergence holds at linear rate, i.e. there exists a norm  $\|\cdot\|_{C(X)/\mathbb{R}}$  on  $C(X)/\mathbb{R}$  and a positive number  $\delta$  such that  $\|u - u_\infty\|_{C(X)/\mathbb{R}} \leq e^{-\delta m}$ . Indeed, setting  $\|u - u'\|_{C(X)/\mathbb{R}} := \|\sup(u - u') - \inf(u - u')\|_{C(X)}$  (which corresponds, under  $u \mapsto e^{-u}$ , to the Hilbert metric on the cone of positive functions in  $C(X)$ ) this follows from Birkhoff’s theorem about positive operators on cones, precisely as in the finite dimensional situation of the Sinkhorn iteration considered in [36]; see also [27, Thm 4.2] and [66].

### 2.2.3 The induced discrete evolution on $C(X) \times C(Y)$

Fixing an initial function  $u_0 \in C(X)$  the corresponding evolution  $m \mapsto u_m$  induces a sequence of pairs  $(u_m, v_m) \in C(X) \times C(Y)$  defined by the following recursion:

$$(u_{m+1}, v_{m+1}) := (u[v_{m+1}], v[u_m])$$

### 2.2.4 The induced discrete evolution on $\mathcal{P}(X \times Y)$ and entropy

Let us briefly explain the dual point of view involving the space  $\mathcal{M}(X \times Y)$  of measures on  $X \times Y$  (which, however, is not needed for the proofs of the main results). The data  $(\mu, \nu, c)$  induces the following element  $\gamma_c \in \mathcal{M}(X \times Y)$  :

$$\gamma_c := e^{-c} \mu \otimes \nu$$

Given a function  $u \in C(X)$  we will write

$$\gamma_u := e^{-(u+v[u])} \gamma_c \tag{2.18}$$

**Lemma 2.11**  $u$  satisfies  $S(u) = u$  a.e. wrt  $\mu$  iff  $\gamma_u \in \Pi(\mu, \nu)$ .

**Proof** A direct computation reveals that the push-forwards of  $e^{-(u+v[u])}\gamma_c$  to  $X$  and  $Y$ , respectively, are given by

$$\int_X e^{-(u+v[u])}\gamma_c = \nu, \quad \int_Y e^{-(u+v[u])}\gamma_c = \rho_u\mu$$

Hence,  $\gamma_u \in \Pi(\mu, \nu)$  iff  $\rho_u\mu = \mu$ , which, by the definition 2.15 of  $\rho_u$ , concludes the proof. □

The discrete dynamical system  $u_m$  induces a sequence

$$\gamma_m := \gamma_{u_m} (= e^{-u_m(x)} e^{-v_m(x)} \gamma_c) \in \mathcal{P}(X \times Y)$$

**Proposition 2.12** *The unique minimizer  $\gamma_*$  of the functional  $\mathcal{I}(\cdot|\gamma_c)$  on  $\Pi(\mu, \nu)$  is characterized by the property that it has the form*

$$\gamma_* = e^{-\Phi} \gamma_c$$

for some  $\Phi \in C(X) + C(Y)$ . Moreover,  $\gamma_* = \gamma_{u_*}$ , where  $u_*$  is a fixed point for  $S$  on  $C(X)$  (or more generally, on  $L^1(X, \mu)$ ) and

$$\inf_{\Pi(\mu, \nu)} \mathcal{I}(\cdot|\gamma_c) = \inf_{C(X) \times C(Y)} \mathcal{F} \tag{2.19}$$

and given any function  $u_0 \in C(X)$ , the corresponding sequence  $\gamma_m$  converges in  $L^1$  (i.e. in variation norm) towards  $\gamma_*$  (and moreover  $\mathcal{I}(\gamma_m|\gamma_*) \rightarrow 0$ ).

**Proof** By construction  $\gamma_* := \gamma_{u_*}$  has the property that

$$\gamma_* = e^{-\Phi} \gamma_c, \quad \gamma_* \in \Pi(\mu, \nu)$$

for some  $\Phi \in L^\infty(X) + L^\infty(Y)$ . But a standard calculus argument reveals that any such  $\gamma_*$  is the unique minimizer of the restriction of  $\mathcal{I}$  to the affine subspace  $\Pi(\mu, \nu)$  of  $\mathcal{P}(X \times Y)$  (using that  $\mathcal{I}$  is strictly convex). The last convergence statement then follows directly from Theorem 2.9 (only the easier convergence in  $C(X)/\mathbb{R}$  is needed). □

Rewriting

$$k^{-1}\mathcal{I}(\gamma|\gamma_{kc}) = \int c\gamma + k^{-1}\mathcal{I}(\gamma|\mu \otimes \nu),$$

the equality 2.19 can be viewed as an entropic variant of Kantorovich duality 2.4 in the limit when  $c$  is replaced by  $kc$  for a large positive number  $k$ . In fact, it follows from Theorem 3.3 applied to  $\mu_k = \mu$  and  $\nu_k = \nu$  that

$$\lim_{k \rightarrow \infty} \inf_{\gamma \in \Pi(\mu, \nu)} k^{-1}\mathcal{I}(\gamma|\gamma_{kc}) = \inf_{\gamma \in \Pi(\mu, \nu)} \int c\gamma = \sup_{\Phi_c} \int u\mu + \int \nu v,$$

as in the Kantorovich duality 2.4. In the next section we will consider the setting where  $\mu$  and  $\nu$  also change with  $k$ .

### 2.3 The scaled setting and discretization

Let us next consider the following variant of the previous setting, parametrized by a parameter  $k$  (which is the parameter that will later on tend to infinity and which corresponds to the entropic regularization parameter  $\epsilon := k^{-1}$ ). This means that we replace the triple  $(\mu, \nu, c)$  with a sequence  $(\mu^{(k)}, \nu^{(k)}, kc)$ . As explained in Sect. 1.1.2 replacing  $c$  with  $kc$  corresponds to introducing the entropic regularization parameter  $\epsilon = k^{-1}$ . We then rescale the functions in  $C(X)$  and  $C(Y)$  by  $k$  and consider the corresponding rescaled operators:

$$\begin{aligned} v^{(k)}[u] &:= k^{-1} \log \int e^{-kc(x,\cdot)-ku(x)} \mu^{(k)}(x) \\ u^{(k)}[v] &:= k^{-1} \log \int e^{-kc(\cdot,y)-kv(y)} \nu^{(k)}(y) \\ S^{(k)}(u) &:= k^{-1} S(ku) \end{aligned} \tag{2.20}$$

etc. The corresponding rescaled iteration is thus defined by the iteration

$$u_{m+1}^{(k)} := S^{(k)} u_m^{(k)} \in C(X), \tag{2.21}$$

given the initial value  $u_0^{(k)} \in C(X)$ . It will be called the *scaled log Sinkhorn iteration (at level  $k$ )*. Equivalently,

$$u_{m+1}^{(k)} - u_m^{(k)} = k^{-1} \log(\rho_{ku_m^{(k)}}), \tag{2.22}$$

where

$$\rho_{ku}(x) := \frac{e^{ku^{(k)}[v^{(k)}[u]](x)}}{e^{ku}}(x), \tag{2.23}$$

which can be explicitly expressed as

$$\rho_{ku}(x) = \int_Y \frac{e^{-kc(x,y)-ku(x)}}{\int_X e^{-kc(x',y)-ku(x')} \mu^{(k)}(x')} \nu^{(k)}(y)$$

We also set

$$\mathcal{F}^{(k)}(u) := k^{-1} \mathcal{F}(ku) = \int u \mu + \int v^{(k)}[u] \nu. \tag{2.24}$$

By Theorem 2.9 (applied to a fixed  $k$ ) as  $m \rightarrow \infty$  the iteration  $u_m^{(k)}$  converges in  $C(X)$  to a fixed point  $u^{(k)}$  of the operator  $S^{(k)}$  (uniquely determined by the initial value  $u_0^{(k)}$ ).

We observe that the following compactness property holds (and is proved exactly as in Lemma 2.7):



**Lemma 2.13** *The union  $\bigcup_{k \geq 0} S^{(k)}$  is relatively compact in  $C(X)/\mathbb{R}$  (identifying  $C(X)/\mathbb{R}$  with the set of all continuous functions vanishing at a given point  $x_0$ )*

### 2.3.1 Discretization and the Sinkhorn algorithm

Now assume that  $\mu^{(k)}$  and  $\nu^{(k)}$  are discrete probability measures whose supports are finite sets

$$X^{(k)} := \{x_i^{(k)}\}_{i=1}^{N_k}, \quad Y^{(k)} := \{y_i^{(k)}\}_{i=1}^{N_k}$$

of the same number  $N_k$  of points in  $X$  and  $Y$ , respectively. This means that there exist vectors  $p^{(k)}$  and  $q^{(k)}$  in  $\mathbb{R}^{N_k}$  such that

$$\mu^{(k)} = \sum_{i=1}^{N_k} \delta_{x_i^{(k)}} p_i^{(k)}, \quad \nu^{(k)} = \sum_{i=1}^{N_k} \delta_{y_i^{(k)}} q_i^{(k)}.$$

Moreover, since  $\mu^{(k)}$  and  $\nu^{(k)}$  are probability measures the vectors  $p^{(k)}$  and  $q^{(k)}$  are elements in the simplex  $\Sigma_{N_k}$  in  $\mathbb{R}^{N_k}$  defined by

$$\Sigma_N := \left\{ v \in \mathbb{R}^N : v_i \geq 0, \sum_{i=1}^N v_i = 1 \right\}, \tag{2.25}$$

which we identify with  $\mathcal{P}(\{1, \dots, N\})$ . Similarly, we identify the discrete measure

$$\gamma_c^{(k)} := e^{-kc} \mu^{(k)} \otimes \nu^{(k)}$$

on  $X \times Y$  with the matrix  $\tilde{K} \in \mathbb{R}^{N_k} \times \mathbb{R}^{N_k}$  defined by

$$\tilde{K}_{ij} := K_{ij}^{(k)} p_i^{(k)} q_j^{(k)}, \quad K_{ij}^{(k)} := \exp(-kC_{ij}), \quad C_{ij} := c(x_i^{(k)}, y_j^{(k)}),$$

where  $C_{ij}$  is viewed as a cost function on  $\{1, \dots, N\}^2$ . Under the identifications

$$C(X^{(k)}) \leftrightarrow \mathbb{R}_+^{N_k}, \quad u \mapsto a, \quad a_i := e^{-ku(x_i^{(k)})} p_i^{(k)}$$

and

$$C(Y^{(k)}) \leftrightarrow \mathbb{R}_+^{N_k}, \quad v \mapsto b, \quad b_i := e^{-kv(y_j^{(k)})} q_i^{(k)}$$

the scaled iteration 2.21 gets identified with the recursion  $a^{(k)}(m)$  defined by the Sinkhorn algorithm determined by the matrix  $K^{(k)}$  and the positive vectors  $p^{(k)}$  and  $q^{(k)}$  (see Sect. 1.1.1). Given an initial positive vector  $a^{(k)}(0)$  Theorem 2.9 thus shows that  $(a^{(k)}(m), b^{(k)}(m))$  converges, as  $m \rightarrow \infty$ , to a pair of positive vectors  $(a^{(k)}, b^{(k)})$  such that the scaled matrix  $D_b K^{(k)} D_a$  has the property that the rows sum to  $p^{(k)}$  and the columns sum to  $q^{(k)}$ .

**Remark 2.14** By construction, the functions  $u_m^{(k)}(x)$  on  $X$  can be expressed in terms of a Fourier type sum:

$$u_m^{(k)}(x) = k^{-1} \log \sum_{i=1}^{N_k} e^{-kc(x, y_i^{(k)})} b_i^{(k)}(m - 1)$$

where the ‘‘Fourier coefficients’’  $b_i^{(k)}(m - 1)$  are given by the Sinkhorn algorithm. In the case when  $X$  and  $Y$  are domains in  $\mathbb{R}^n$ , with  $c(x, y) = -x \cdot y$ , this is the analytic continuation to  $i\mathbb{R}^n$  of a bona fide Fourier sum with Fourier coefficients in  $k$  times the support of  $v^{(k)}$ . Hence,  $k$  plays the role of the ‘‘band-width’’.

### 3 Convergence of the fixed points

In this section we will prove various generalizations of Theorem 1.1, stated in the introduction. Throughout the section we will consider the parametrized setting in Sect. 2.3 and assume that the sequences  $\mu^{(k)}$  and  $v^{(k)}$  converge to  $\mu$  and  $v$  in  $\mathcal{P}(X)$  and  $\mathcal{P}(Y)$ , respectively (in the standard weak topology). We will denote by  $u^{(k)}$  the fixed point of the corresponding operator  $S^{(k)}$  on  $C(X)$ , uniquely determined by the normalization condition  $u^{(k)}(x_0) = 0$ , at a given point  $x_0$  in  $X$  and set  $v^{(k)} := T_\mu u^{(k)} =: v^{(k)}[u^{(k)}]$ , which is a fixed point of the corresponding operator  $S^{(k)}$  on  $C(Y)$ .

#### 3.1 A general convergence result for the fixed points

We start by giving a density condition on  $\mu^{(k)}$  ensuring that  $v^{(k)}[u]$  converges uniformly to the  $c$ -Legendre transform  $u^c$  of  $u$ , when  $\mu$  has full support:

**Lemma 3.1** *Assume that the sequence  $\mu^{(k)}$  converging to  $\mu$  in  $\mathcal{P}(X)$  has the following ‘‘density property’’: for any given open subset  $U$  intersecting the support  $X_\mu$  of  $\mu$*

$$\liminf_{k \rightarrow \infty} k^{-1} \log \mu^{(k)}(U) \geq 0 \tag{3.1}$$

*Then, for any given  $u \in C(X)$ , the sequence  $v^{(k)}[u]$  converges uniformly to  $(\chi_{X_\mu} + u)^c$  in  $C(Y)$ .*

**Proof** Replacing the integral over  $\mu^{(k)}$  with a sup directly gives

$$v^{(k)}[u](y) \leq (\chi_{X_\mu} + u)^c(y) \tag{3.2}$$

for any  $y \in Y$ . To prove a reversed inequality let  $x_y$  be a point in  $X_\mu$  where the sup defining  $(\chi_{X_\mu} + u)^c(y)$  is attained and  $U_\delta$  a neighborhood of  $x_y$  where the oscillation of  $c(\cdot, y) + u$  is bounded from above by  $\delta$  (the existence of  $U_\delta$  is ensured by the continuity of  $c$  and the compactness of  $X$  and  $Y$ ). Then

$$v^{(k)}[u](y) \geq k^{-1} \log \int_{U_\delta} e^{-k(c(x, y) + u(x))} \mu^{(k)}(x) \geq k^{-1} \log \mu^{(k)}(U_\delta) + u^c(y) - \delta$$

Hence, as  $k \rightarrow \infty$ ,  $v^{(k)}[u](y) \rightarrow u^c(y)$  and since  $v^{(k)}[u]$  is equicontinuous (by the assumed compactness of  $X \times Y$  and the continuity of  $c$ ) this implies the desired uniform convergence. □

**Example 3.2** (Weighted point clouds) If  $\mu_k = \mu$  for any  $k$  then the density property is trivially satisfied. More generally, the density property 3.1 is satisfied by any reasonable approximation  $\mu^{(k)}$ . For example, in the discrete case where  $\mu^{(k)} = \sum_{i=1}^{N_k} w_i^{(k)} \delta_{x_i^{(k)}}$  the property in question holds if  $\sup_i 1/w_i^{(k)}$  and the inverse of the number of points  $x_i^{(k)}$  in any given open set  $U$  intersecting  $X_\mu$  have sub-exponential growth in  $k$ .

**Theorem 3.3** *Suppose that  $\mu^{(k)} \rightarrow \mu$  and  $v^{(k)} \rightarrow \nu$  in  $\mathcal{P}(X)$  and  $\mathcal{P}(Y)$ , respectively and assume that  $\mu^{(k)}$  and  $v^{(k)}$  satisfy the density property 3.1. Let  $u^{(k)}$  be the normalized fixed point for the scaled log Sinkhorn operator  $S^{(k)}$  on  $C(X)$ . Then, after perhaps passing to a subsequence, the following holds:*

$$u^{(k)} \rightarrow u$$

uniformly on  $X$ , where  $u$  is a  $c$ -convex minimizer of the Kantorovich functional  $J$  (formula 2.1) satisfying

$$u = (\chi_{Y_\nu} + (\chi_{X_\mu} + u)^c)^c \tag{3.3}$$

As a consequence, the corresponding probability measures

$$\gamma^{(k)} := e^{-k(u^{(k)}+v^{(k)})} e^{-kc} \mu^{(k)} \otimes v^{(k)} \in \mathcal{P}(X \times Y)$$

converge weakly to a transport plan  $\gamma$  between  $\mu$  and  $\nu$ , which is optimal wrt the cost function  $c$ .

**Proof Step 1: Convergence of a subsequence of  $u^{(k)}$**

In the following all functions will be normalized by demanding that the values vanish at a given point. By Lemma 2.13 we may, after perhaps passing to a subsequence, assume that  $u^{(k)} \rightarrow u^{(\infty)}$  uniformly on  $X$ , for some element  $u^{(\infty)}$  in  $C(X)$ . By the previous lemma, for any given  $u \in C(X)$  we have

$$\mathcal{F}^{(k)}(u) = J(\chi_{X_\mu} + u) + o(1), \tag{3.4}$$

where  $\mathcal{F}^{(k)}$  is defined by formula 2.24 and  $J$  denotes the Kantorovich functional, defined by formula 2.1. Now take a sequence  $\epsilon_k$  of positive numbers tending to zero such that

$$u^{(\infty)} - \epsilon_k \leq u^{(k)} \leq u^{(\infty)} + \epsilon_k \tag{3.5}$$

Since  $u \mapsto v^{(k)}[u]$  is decreasing it follows that

$$\mathcal{F}^{(k)}(u^{(\infty)}) - 2\epsilon_k \leq \mathcal{F}^{(k)}(u^{(k)}) + 2\epsilon_k \tag{3.6}$$

Next note that since  $u^{(k)}$  minimizes the functional  $\mathcal{F}^{(k)}$  (by Prop 2.8) we have  $\mathcal{F}^{(k)}(u^{(k)}) \leq \mathcal{F}^{(k)}(u)$  for any given  $u$  in  $C(X)$ . Hence, combining 3.6 and 3.4 and letting  $k \rightarrow \infty$  gives

$$J(u^{(\infty)}) \leq \inf_{u \in C(X)} J(u),$$

showing that  $u^{(\infty)}$  minimizes  $J$  on  $C(X)$ . To see that  $u^{(\infty)}$  is  $c$ -convex first recall that, by definition,  $u^{(k)}$  satisfies

$$u^{(k)} = u^{(k)}[v^{(k)}[u^{(k)}]].$$

Hence, combing 3.5 with the previous lemma, applied twice, gives

$$u^{(k)} = u^{(k)}[(\chi_{X_\mu} + u^{(\infty)})^c] + o(1) = ((\chi_{Y_\mu} + (\chi_{X_\mu} + u^{(\infty)})^c)^c + o(1)$$

This shows that  $u^{(\infty)} = ((\chi_{Y_\mu} + (\chi_{X_\mu} + u^{(\infty)})^c)^c$ , proving that  $u^{(\infty)} = f^c$  for some  $f \in C(Y)$ . Hence  $u^{(\infty)}$  is  $c$ -convex.

*Step 2: Convergence of  $\gamma^{(k)}$  (for the subsequence in Step 1) towards an optimizer*

By Lemma 2.11  $\gamma^{(k)}$  is in  $\Pi(\mu, \nu)$ . Hence, by weak compactness, we may assume that  $\gamma^{(k)}$  converges towards an element  $\gamma^{(\infty)}$  in  $\mathcal{P}(X \times Y)$ . By Prop 2.1 it will thus be enough to show that  $\gamma^{(\infty)}$  is supported in  $\Gamma_{u^{(\infty)}}$ . To this end let  $\Gamma_\delta$  be the closed subset of  $X \times Y$  where  $u + u^c + c \geq \delta > 0$  for  $u := u^{(\infty)}$ . By the previous lemma  $\gamma^{(k)} \leq e^{-k\delta/2} \mu^{(k)} \otimes \nu^{(k)}$  on  $\Gamma_\delta$ , when  $k$  is sufficiently large and hence the limit  $\gamma^{(\infty)}$  is indeed supported on  $\Gamma_{u^{(\infty)}}$ . □

In order to ensure that the whole sequence  $u^{(k)}$  is convergent some conditions on the cost function  $c$  and the measures  $\mu$  and  $\nu$  need to be imposed. Exploiting well-known uniqueness result for optimal transport plans/maps this can, in particular, be achieved in the following Riemannian setting.

**Theorem 3.4** *Let  $M$  be a Riemannian manifold and denote by  $d$  the Riemannian distance function. Let  $X$  and  $Y$  be compact subsets of  $M$  such that  $Y$  is a topological domain, i.e.  $Y$  is equal to the closure of the interior of  $Y$  and take  $c(x, y)$  to be the restriction of  $d(x, y)^2/2$  to  $X \times Y$ . Assume that  $\nu$  is absolutely continuous wrt the Riemannian volume form and has support  $Y$  and that  $\mu^{(k)} \rightarrow \mu$  and  $\nu^{(k)} \rightarrow \nu$  in  $\mathcal{P}(X)$ . Denote by  $u^{(k)}$  the normalized fixed point of the scaled log Sinkhorn operator  $S^{(k)}$  on  $C(X)$ . Then*

- $v^{(k)}$  converges uniformly in  $Y$  to a  $c$ -convex function  $v$ , which is a potential for the unique optimal Borel map transporting  $\nu$  to  $\mu$ , i.e. the map that can be expressed as

$$y \mapsto x_y := \exp_y(\nabla v), \tag{3.7}$$

(which means that  $x_x$  is obtained by transporting  $y$  along a unit-length geodesic in the direction of  $(\nabla v)(y)$ ).

- $u^{(k)}$  converges uniformly on  $X$  towards the  $c$ -convex function  $u$  given by the  $c$ -Legendre transform  $v^c$  of  $v$ . Moreover,  $u$  and  $v$  satisfy

$$u = (u + \chi_{X_v})^{cc}, \quad v = (u + \chi_{X_v})^c = u^c = v$$

- If  $\mu$  is absolutely continuous wrt the Riemannian volume form, then  $x \mapsto \exp_y(\nabla u)$  defines the optimal transport of  $\mu$  to  $v$ .

**Proof** This will be shown to follow from the previous theorem combined with results in [68], generalizing Brenier’s theorem in  $\mathbb{R}^n$  [16] and its Riemannian version in [52] (and, in particular, [23], concerning the torus case). After passing to a subsequence, as in the previous theorem, we may assume that  $u^{(k)} \rightarrow u := u^{(\infty)}$  and that  $v^{(k)} \rightarrow v := (\chi_{X_\mu} + u)^c$ . In particular,  $v$  is  $c$ -convex. Denote by  $\gamma$  the corresponding optimal transport plan, furnished by the previous theorem, which is supported in the subset of  $X \times Y$  where  $u + v + c = 0$ . Hence, it follows from [68, Thm 10.41] (and its proof) that the Borel map  $y \mapsto \exp_y(\nabla v)$  is the unique optimal transport (Borel) map, pushing forward  $v$  to  $\mu$ . Since  $Y$  is assumed to be a topological domain it follows that  $v$  is uniquely determined on  $Y$ , modulo additive constants (see [68, Remark 10.30]). Now, by formula 3.3 we have  $u = v^c$  and since  $u(x_0) = 0$  it follows that  $u$  is uniquely determined. But, as shown in the proof of the previous theorem, we have  $v = (u + \chi_{X_v})^c$  and hence  $v$  is also uniquely determined (i.e. not only determined modulo an additive constant). Next, since, by assumption,  $Y = Y_v$  formula 3.3 says that  $(u + \chi_{X_v})^{cc} = u$ . In general,  $w^{cc} = w^c$  for any function  $w$  and hence it follows that  $(u + \chi_{X_v})^c = u^c$  which shows that  $v = u^c$ . □

The previous theorem applies more generally as soon as a unique Borel optimal map exists (see for example [68, Thm 10.38] for conditions on  $c$  ensuring that this is the case).

### 3.1.1 The torus case: proof of Theorem 1.1

First assume only that the probability measure  $\nu$  is absolutely continuous wrt Lebesgue measure. By the previous theorem  $u^{(k)}$  then converges uniformly towards a  $c$ -convex function  $u$  such that

$$(\nabla u^c + I)_* \nu = \mu$$

If  $\mu$  moreover and  $\nu$  have densities  $e^{-g}$  and  $e^{-g}$  which are Hölder continuous, then it is well-known that there exists a unique optimal transport map and its potential (which is uniquely determined up to a constant) is in  $C^{2,\alpha}(T)$  for some  $\alpha > 0$  (see [23] where this is deduced from the regularity results of Caffarelli  $\mathbb{R}^n$ ). It follows that  $u^c$  and hence also  $u$  is  $C^2$ -smooth and strictly quasi-convex and solves the Monge–Ampère equation 1.5.

### 3.2 Application to the second boundary value problem for the Monge–Ampère operator in $\mathbb{R}^n$

Now consider the Euclidean case of Theorem 3.4, i.e. the case where  $M = \mathbb{R}^n$  and  $d(x, y) = |x - y|$  and assume that  $X$  and  $Y$  are compact convex domains (and, in particular, topological domains) and take  $x_0 = 0$ . As before we also assume that the support of  $\nu$  is equal to  $Y$  and that  $\nu$  is absolutely continuous wrt  $dx$  and fix discretizations  $\mu^{(k)}$  and  $\nu^{(k)}$  satisfying the density property 3.1.

In order to conform to classical notation in  $\mathbb{R}^n$  we set  $\phi^{(k)}(x) := u^{(k)}(x) + |x|^2/2$  and  $\psi^{(k)} = v^{(k)}(y) + |y|^2$  etc. This corresponds to replacing the cost function  $d^2/2$  with

$$c(x, y) := -x \cdot y.$$

We will use the classical notation  $\phi^*$  for the corresponding Legendre transform (formula 2.9). Then, by Theorem 3.4,  $\phi^{(k)}$  converges uniformly to a convex function  $\phi$  on  $X$  and

$$(\nabla\psi)_*\nu = \mu, \quad \psi := (\chi_X + \phi)^* \tag{3.8}$$

We next observe that this means that  $\phi$  satisfies the following Monge–Ampère equation on  $\Omega$

$$MA_\nu(\phi) = \mu, \tag{3.9}$$

where  $MA_\nu(\phi)$  denote the Monge–Ampère measure of  $\phi$  relative the target measure  $\nu$ , in the sense of Alexandrov. We recall that if  $\nu$  is a given probability measure on  $\mathbb{R}^n$  which is absolutely continuous wrt  $dx$  and  $\phi$  is a finite convex function on a convex open set  $\Omega$  in  $\mathbb{R}^n$ , then  $MA_\nu(\phi)$  is the Borel measure on  $\Omega$  defined by

$$\int_E MA_\nu(\phi) = \int_{(\partial\phi)(E)} \nu,$$

where  $E$  is a given Borel set in  $\Omega$  and  $(\partial\phi)(E)$  denotes the image of  $E$  under the multivalued sub-gradient map  $\partial\phi$  [67]. As is well-known, the assumption that  $\nu$  is absolutely continuous wrt  $dx$  ensures that  $MA_\nu(\phi)$ , as defined above, is indeed a measure on  $\Omega$ , i.e. countably additive (see [53, Section 3] for a more general setting involving a cost function  $c$ ). Moreover, if  $\phi$  is  $C^2$ -smooth and strictly convex, then making the change of variables  $y = \nabla\phi(x)$  reveals that

$$MA_\nu(\phi) = \rho_\nu(\nabla\phi) \det(\nabla^2\phi), \tag{3.10}$$

where  $\rho_\nu$  denotes the density of  $\nu$  wrt  $dx$ . We will use the following general representation of the the Monge–Ampère measure (see [11, Section 2.2] and references therein).

**Lemma 3.5** *Let  $\phi$  be a bounded finite convex function in  $\Omega$ . Then*

$$MA_\nu(\phi) = (\nabla\psi)_*\nu, \quad \psi := (\chi_X + \phi)^*$$

Hence, by formula 3.8, the limit  $\phi$  of  $\phi^{(k)}$  indeed solves the Monge–Ampère equation 3.9. It should be stressed that, in general, there can be many different convex solutions to the Monge–Ampère equation (which not only differ by an additive constant). But the point is that the Sinkhorn iteration singles out a particular solution  $\phi$ . On the other hand, when  $Y$  is convex we have the following well-known uniqueness result:

Let  $\Omega$  be a convex open set in  $\mathbb{R}^n$  and  $\mu$  and  $\nu$  probability measures on  $\mathbb{R}^n$  such that  $\mu$  is supported in  $\Omega$ ,  $\nu$  is absolutely continuous wrt  $dx$  and the support  $Y$  of  $\nu$  is a convex body, i.e. a compact convex set with non-empty interior. Then a solution  $\phi$  to the second boundary value problem

$$MA_\nu(\phi) = \mu, \quad (\partial\phi)(\Omega) \subset Y \tag{3.11}$$

is uniquely determined up to an additive constant.

**Proof** A simple proof is given in [11], which we recall here since it fits well with the spirit of the present paper (see also [53, Thm 3.1] for a different proof in the setting of more general cost function). The starting point is the observation that any convex function  $\phi$  on an open convex set  $\Omega$  with the property that  $(\partial\phi)(\Omega) \subset Y$ , for a convex body  $Y$ , may be expressed as  $\phi = (\chi_Y + (\chi_X + \phi)^*)$ . But, by the previous lemma, the equation  $MA_\nu(\phi) = \mu$  implies that the function  $\psi := (\chi_X + \phi)^*$  in  $Y$  is uniquely determined up to an additive constant. Hence, so is  $\phi$ .  $\square$

We thus arrive at the following result, which also applies to the second boundary value problem in all of  $\mathbb{R}^n$  :

**Theorem 3.6** *Let  $X$  and  $Y$  be compact convex domains in  $\mathbb{R}^n$  endowed with probability measure  $\mu$  and  $\nu$  respectively. Assume that the support of  $\nu$  is equal to  $Y$  and that  $\nu$  is absolutely continuous wrt  $dx$ . Let  $\mu^{(k)}$  and  $\nu^{(k)}$  be sequences of probability measures on  $X$  and  $Y$  converging weakly towards  $\mu$  and  $\nu$  respectively and satisfying the density property 3.1. Denote by  $\phi^{(k)}$  the unique normalized fixed point of the scaled log Sinkhorn operator on  $C(X)$  corresponding to the cost function  $c(x, y) = -x \cdot y$ . Then*

$$\phi^{(k)} \rightarrow \phi, \quad k \rightarrow \infty \tag{3.12}$$

*uniformly on  $X$ , where  $\phi$  is the unique normalized convex solution to the second boundary value problem 3.11 in the interior of  $X$ . More generally, the corresponding result holds when  $X$  is replaced by  $\mathbb{R}^n$  under the assumption that there exists a compact subset containing the support of  $\mu^{(k)}$  for all  $k$ . Then the corresponding convergence 3.12 is uniform on compact subsets of  $\mathbb{R}^n$ .*

**Proof** First assume that  $X$  is compact. As explained above it then follows from Theorem 3.4 that  $\phi^{(k)}$  converges uniformly to a normalized convex function  $\phi$  on  $X$ , satisfying  $MA_\nu(\phi) = \mu$ . Next, note that, it follows directly from the definition of the fixed point  $\phi^{(k)}$  that  $\phi^{(k)}$  may be expressed as

$$\phi^{(k)}(x) = k^{-1} \log \int_Y e^{kx \cdot y} \nu^{(k)}(y)$$

for a measure  $\nu'$  supported on  $Y$ . As a consequence,  $\phi^{(k)}$  is smooth and  $\nabla\phi^{(k)}$  is contained in the convex hull of  $Y$ , i.e. in  $Y$ , since  $Y$  is assumed to be convex. But then it follows that  $(\partial\phi)(\Omega) \subset Y$ , which concludes the proof in the case when  $X$  is compact. Note that, alternatively, we could have used directly that, by Theorem 3.4, the limit  $\phi$  satisfies  $\phi = (\chi_Y + (\chi_X + \phi)^*)$ , which is the unique normalized solution to the second boundary problem in question (by the proof of the previous lemma). To prove the non-compact case when  $X$  is replaced by  $\mathbb{R}^n$  denote by  $X_R$  the ball or radius  $R$  centered at 0 and assume that  $R$  is sufficiently large to ensure that  $\mu$  and all  $\mu^{(k)}$  are supported in the interior of  $X_R$ . Denote by  $\phi_R^{(k)}(x)$  the normalized fixed point of the corresponding iteration on  $C(X_R)$ . Since  $\mu^{(k)}$  is supported in  $X_R$  it follows, from the very definition of the iteration, that  $\phi_R^{(k)}(x)$  is, in fact, independent of  $R$ . Accordingly, we may define a normalized convex function  $\phi^{(k)}$  on  $\mathbb{R}^n$  by setting  $\phi^{(k)} := \phi_R^{(k)}$  on  $X_R$  for any  $R$  sufficiently large. Then  $\phi^{(k)}$  is the unique normalized fixed point of the corresponding iteration on  $C(\mathbb{R}^n)$ . Moreover, since  $X_R$  is compact  $\phi^{(k)}$  converges uniformly on  $X_R$  to a convex function  $\phi$  solving the second boundary value problem 3.11 on  $X_R$ . Since  $R$  is arbitrary this shows that  $\phi$ , in fact, solves the problem on all of  $\mathbb{R}^n$ .  $\square$

### 3.3 Application to convolutional Wasserstein distances

Theorem 3.3 holds more generally (with essentially the same proof) when the function  $c$  is replaced by a sequence  $c_k$  such that

$$\|c_k - c\|_{L^\infty(X \times Y)} \rightarrow 0 \tag{3.13}$$

For example, in the Riemannian setting of Theorem 3.4. denoting by  $\mathcal{K}_t(x, y)$  the corresponding heat kernel and setting  $t := 2k^{-1}$ , the sequence

$$c_k := -t^{-1} \log \mathcal{K}_t(x, y) \tag{3.14}$$

satisfies 3.13, by Varadhan’s formula (which holds more generally on Lipschitz Riemannian manifolds [54]). Replacing  $c$  by  $c_k$  in this setting thus has the effect of replacing the matrix  $A_{ij} := e^{-kd^2(x_i, x_j)/2}$  appearing in the corresponding Sinkhorn algorithm with the heat kernel matrix  $\mathcal{K}_{2k^{-1}}(x_i, x_j)$  which, as emphasized in [63], has computational advantages. Following [63] we consider the squared *convolutional Wasserstein distance* between  $\mu$  and  $\nu$  :

$$\mathcal{W}_{(k)}^2(\mu, \nu) := k^{-1} \inf_{\gamma \in \Pi(\mu^{(k)}, \nu^{(k)})} \mathcal{I}(\gamma, \mathcal{K}_{2k^{-1}}\mu^{(k)} \otimes \nu^{(k)}),$$

defined wrt approximations  $\mu^{(k)}$  and  $\nu^{(k)}$ , for example given by weighted point clouds, as in Example 3.2. In [63, Page 3], the problem of developing conditions for the convergence of  $\mathcal{W}_{(k)}^2(\mu, \nu)$  was posed. The following result provides an answer:



**Theorem 3.7** *Let  $X$  be a compact Riemannian manifold (possibly with boundary) and set  $c(x, y) := d(x, y)^2/2$ , where  $d$  is the Riemannian distance function. Suppose that  $\mu^{(k)} \rightarrow \mu$  and  $\nu^{(k)} \rightarrow \nu$  in  $\mathcal{P}(X)$  and that  $\mu^{(k)}$  and  $\nu^{(k)}$  satisfy the density property 3.1. Then*

$$\lim_{k \rightarrow \infty} \mathcal{W}_{(k)}^2(\mu, \nu) = \mathcal{W}^2(\mu, \nu),$$

where  $\mathcal{W}^2(\mu, \nu)$  denotes the squared  $L^2$ -Wasserstein distance between  $\mu$  and  $\nu$ .

**Proof** Repeating the argument in the proof of Theorem 3.3, with  $c$  replaced by  $c_k$  as above, gives

$$\lim_{k \rightarrow \infty} \inf_{u \in C^0(X)} \mathcal{F}^{(k)} = \inf_{u \in C(X)} J(u)$$

According to formula 2.12 the infimum appearing in the left hand side above is precisely  $\mathcal{W}_{(k)}^2(\mu, \nu)$ . Since the infimum in the right hand side above is equal to  $\mathcal{W}^2(\mu, \nu)$ , by Kantorovich duality (formula 2.4), the result follows. □

## 4 Convergence of the iteration towards parabolic optimal transport equations on the torus

### 4.1 Proof of Theorem 1.2

We will denote by  $\delta_{\Lambda_k}$  the uniform discrete probability measure supported on the discrete torus  $\Lambda_k$  with edge-length  $1/k$  :

$$\delta_{\Lambda_k} := \frac{1}{N_k} \sum_{x_i \in \Lambda_k} \delta_{x_i}$$

Given two probability measures  $\mu = e^{-f} dx$  and  $\nu = e^{-g} dy$  we can then define their discretizations as the probability measures

$$\mu^{(k)} = \frac{1}{\int_{T^n} e^{-f} \delta_{\Lambda_k}} e^{-f} \delta_{\Lambda_k}, \quad \nu^{(k)} = \frac{1}{\int_{T^n} e^{-g} \delta_{\Lambda_k}} e^{-g} \delta_{\Lambda_k}$$

Note that the normalization constants have the asymptotics

$$\left| \int_{T^n} e^{-f} \delta_{\Lambda_k} - 1 \right| \leq C_f k^{-1}, \quad \left| \int_{T^n} e^{-g} \delta_{\Lambda_k} - 1 \right| \leq C_g k^{-1} \tag{4.1}$$

where  $C$  only depends on an upper bound on  $|\nabla f|$  on  $T^n$  (and similarly for  $g$ ); see [13, Page 2].

**Remark 4.1** As will be clear from the proof, Theorem 1.2 also holds with the simpler discretizations  $e^{-f} \delta_{\Lambda_k}$  and  $e^{-g} \delta_{\Lambda_k}$  (which, in general, are not probability measures).

We start with the following discrete version of the classical Laplace method of integration, proved in “Appendix A”:

**Lemma 4.2** *Let  $\alpha$  be a lower-semicontinuous (lsc) function on  $T^n$  with a unique minimum at  $x_0$  and assume that  $\alpha$  is  $C^4$ -smooth on a neighborhood of  $x_0$  and  $\nabla^2\alpha(x_0) > 0$ . Then, if  $h$  is  $C^2$ -smooth*

$$k^{n/2} \int e^{-k\alpha} h \delta_{\Lambda_k} = (2\pi)^{n/2} e^{-k\alpha(x_0)} \frac{h(x_0)}{\sqrt{\det(\nabla^2\alpha(x_0))}} (1 + \epsilon_k), \quad |\epsilon_k| \leq Ck^{-1}$$

where the constant  $C$  only depends on an upper bounds on the  $C^4$ -norm of  $\alpha$ , the  $C^2$ -norm of  $h$  and a strict lower bound on the smallest eigenvalue of  $\nabla^2\alpha$  on a neighborhood of  $x_0$ .

We next prove the key asymptotic result that will be used in the proof of Theorem 1.2 giving the asymptotics of the function  $\rho_{ku}(x)$ , defined by formula 2.23 (the result can be viewed as a refinement of Lemma 3.1).

**Proposition 4.3** *Let  $u$  be a strictly quasi-convex function in  $C^4(T^n)$ . Then the following asymptotics hold*

$$\rho_{ku}(x) = \det(I + \nabla^2u(x)) e^{f(x)-g(x+\nabla u(x))} (1 + \epsilon_k), \quad |\epsilon_k| \leq Ck^{-1}$$

where the constant  $C$  only depends on upper bounds on the  $C^4$ -norm of  $u$  and the  $C^2$ -norms of  $f$  and  $g$  and a strict positive lower bound on the eigenvalues of the matrix  $(I + \nabla^2u(x))$ .

**Proof** First recall that, by Lemma 2.3,  $u^c$  is also strictly quasi-convex and in  $C^4(T^n)$ . Set  $c(x, y) = d_{T^n}^2(x, y)/2$ . In the proof we will denote by  $O(k^{-1})$  any sequence of functions satisfying  $|O(k^{-1})| \leq Ck^{-1}$ , for a constant  $C$  depending on data as in the statement of the proposition. Fix  $y \in Y$ . The function  $x \mapsto c(x, y) + u(x)$  is lsc on  $T^n$  and has a unique minimum  $x_y$  and it is  $C^4$ -smooth close to  $x_y$  (by Lemma 2.4). Applying the asymptotics 4.1 and Lemma 4.2 thus gives

$$k^{n/2} e^{kv^{(k)}[u](y)} := k^{n/2} \int e^{-k(c(x,y)+u(x))} \mu^{(k)}(x) = e^{ku^c(y)} \left( h(y) + O(1)k^{-1} \right),$$

where

$$h(y) := (2\pi)^{n/2} \frac{\exp(-f(x_y))}{\sqrt{\det(I + \nabla^2u(x_y))}}.$$

As a consequence, the inverse of  $k^{n/2} e^{kv^{(k)}[u](y)}$  is given by  $e^{-ku^c(y)} (h(y)^{-1} + O(1)k^{-1})$  and hence

$$\begin{aligned} e^{ku^{(k)}[v^{(k)}[u]](x)} &:= k^{n/2} \int e^{-kc(x,y)} \left( \frac{1}{k^{n/2} e^{kv^{(k)}[u](y)}} \right) v^{(k)}(y) \\ &= k^{n/2} \int e^{-k(c(x,y)+u(x)+u^c(y))} h(y)^{-1} v^{(k)}(y) + R_k(x), \end{aligned} \tag{4.2}$$

where

$$R_k(x) \leq O(k^{-1})k^{n/2} \int e^{-k(c(x,y)+u(x)+u^c(y))} v^{(k)}(y).$$

By Lemma 4.2 (and the asymptotics 4.1) we have

$$k^{n/2} \int e^{-k(c(x,y)+u(x)+u^c(y))} v^{(k)}(y) \leq C'$$

and hence it follows that

$$R_k(y) = O(k^{-1}).$$

Now, the same localization argument as above shows that the integral over  $v^{(k)}(y)$  in formula 4.2 localizes around a small neighborhood  $V$  of  $y = y_x$ . Hence, applying Lemma 4.2 (and the asymptotics 4.1) again gives

$$\begin{aligned} k^{n/2} \int e^{-k(d(x,y)^2/2+u(x)+u^c(y))} h(y)^{-1} v^{(k)}(y) \\ = e^{k((u^c)^c)(x)} h^{-1}(y_x) \frac{(2\pi)^{n/2} \exp(-g(y_x))}{\sqrt{\det(I + \nabla^2 u^c(y_x))}} (1 + O(k^{-1})). \end{aligned}$$

All in all, since  $((u^c)^c) = u$ , this shows that

$$e^{ku^{(k)}[v^{(k)}[u]](x)} = e^{ku(x)} h^{-1}(y_x) \frac{(2\pi)^{n/2} \exp(-g(y_x))}{\sqrt{\det(I + \nabla^2 u^c(y_x))}} (1 + O(k^{-1})).$$

The proof is thus concluded by invoking the inverse properties of the Hessians in Lemma 2.3. □

**Lemma 4.4** *Let  $X$  be a compact topological space and consider the following family of difference equations on  $C(X)$ , parametrized by a positive number  $k$  and a discrete time  $m$  :*

$$u_{m+1}^{(k)} - u_m^{(k)} = k^{-1} D^{(k)}(u_m^{(k)}), \tag{4.3}$$

where  $D^{(k)}$  is an operator on  $C(X)$ , which descends to  $C(X)/\mathbb{R}$  and with the property that  $I + k^{-1} D^{(k)}$  is an increasing operator (wrt the usual order relation on  $C(X)$ ).

Assume that there exists a subset  $\mathcal{H}$  of  $C(X)$  and an operator  $D$  on  $\mathcal{H}$  such that for any  $u$  in the class

$$\left| D^{(k)}(u) - D(u) \right| \leq C_u \epsilon_k, \tag{4.4}$$

where  $C_u$  is a positive constant depending on  $u$  and  $\epsilon_k$  is a sequence of positive numbers converging towards zero. Assume that  $u_0^{(k)} = u_0$  where  $u_0$  is a fixed function in  $\mathcal{H}$  and that there exists a family  $u_t \in \mathcal{H}$ , which is two times differentiable wrt  $t$  and solving

$$\frac{\partial u_t}{\partial t} = D(u_t), \quad (u_t)|_{t=0} = u_0, \tag{4.5}$$

for  $t \in [0, T]$ . Then, for any  $(k, m)$  such that  $m/k \in [0, T]$ ,

$$\sup_{T^n} \left| u_m^{(k)} - u_{m/k} \right| \leq C_T \frac{m}{k} \cdot 2 \max\{\epsilon_k, k^{-1}\}, \quad C_T := \max \left\{ \sup_{t \in [0, T]} C_{u_t}, \sup_{X \times [0, T]} \frac{\partial^2 u_t}{\partial^2 t} \right\}.$$

**Proof** We will write  $\psi_{k,m} = u_{m/k}$ .

*Step 1: The following holds for all  $(k, m)$*

$$\sup \left| \psi_{k,m+1} - \psi_{k,m} - k^{-1} D^{(k)}(\psi_{k,m}) \right| \leq k^{-1} C_T \epsilon'_k, \quad \epsilon'_k := 2 \max\{\epsilon_k, k^{-1}\}.$$

Indeed, using the mean value theorem we can write

$$\psi_{k,m+1} - \psi_{k,m} = \frac{1}{k} \left( \frac{u_{m/k+1/k} - u_{m/k}}{1/k} \right) = \frac{1}{k} \left( \frac{\partial u_t}{\partial t} \Big|_{t=m/k} + O(k^{-1}) \right),$$

where the term  $O(k^{-1})$  may be estimated as  $|O(k^{-1})| \leq A_T k^{-1}$ , where  $A_T = \sup_{X \times [0, T]} \left| \frac{\partial^2 u_t}{\partial^2 t} \right|$ . Using the evolution equation for  $u_t$  and applying formula 4.4 thus proves Step 1.

*Step 2: The discrete evolution on  $C(X)$  defined by the difference equation 4.3 decreases the sup-norm, i.e.*

$$\sup_X |\phi_{m+1} - \psi_{m+1}| \leq \sup_X |\phi_m - \psi_m|$$

if  $\phi_m$  and  $\psi_m$  satisfy the difference equation 4.3 for a fixed  $k$ .

To see this set  $C := \sup |\phi_m - \psi_m|$ . Then,  $\phi_m \leq \psi_m + C$  and hence, since  $I + k^{-1} D^{(k)}$  is assumed to be increasing,

$$\phi_{m+1} = \phi_m + k^{-1} D^{(k)}(\phi_m) \leq \psi_m + C + k^{-1} D^{(k)}(\psi_m + C) = \psi_{m+1} + C$$

In particular,  $\sup(\phi_{m+1} - \psi_{m+1}) \leq C$ . Applying the same argument with the roles of  $\phi$  and  $\psi$  interchanged concludes the proof.

*Step 3: Conclusion:*

$$\sup_X \left| u_m^{(k)} - \psi_{k,m} \right| \leq C_T \frac{m}{k} \epsilon'_k. \tag{4.6}$$

We will prove this by induction over  $m$  (for  $k$  fixed) the statement being trivially true for  $m = 0$ . We fix the integer  $k$  and assume as an induction hypothesis that 4.6 holds for  $m$  with  $C$  the constant in the previous inequality. Applying first Step 2 and then the induction hypothesis thus gives

$$\begin{aligned} & \sup_X \left| \left( \psi_{k,m} + k^{-1} D^{(k)}(\psi_{k,m}) \right) - \left( u_m^{(k)} + k^{-1} D^{(k)}(u_m^{(k)}) \right) \right| \\ & \leq \sup_X |\psi_{k,m} - u_m^{(k)}| \leq C_T \frac{m}{k} \epsilon'_k. \end{aligned}$$

Now, by Step 1,

$$\sup_X |\psi_{k,m+1} - (\psi_{k,m} + k^{-1} D^{(k)}(\psi_{k,m}))| \leq \frac{C_T}{k} \epsilon'_k$$

for all  $(m, k)$ . Hence,

$$\sup_X |\psi_{k,m+1} - u_{m+1}^{(k)}| \leq C_T \frac{m}{k} \epsilon'_k + C_T \frac{1}{k} \epsilon'_k = C_T \frac{(m+1)}{k} \epsilon'_k,$$

proving the induction step and hence the final Step 3. □

In the present setting  $\mathcal{H}$  will be taken as subspace of  $C^4(T^n)$  consisting of all strictly quasi-convex functions  $u$  and

$$D(u)(x) := \log(\det(\nabla^2 u(x) + I)) - g(\nabla u(x) + x) + f(x).$$

The following proposition essentially follows from the results in [44,45,57] (for completeness a proof is provided in ‘‘Appendix B’’). The space  $C^{k,\alpha}(T^n)$  denotes, as usual, space of all functions such that the  $k$ th order derivatives are Hölder continuous with Hölder exponent  $\alpha \in ]0, 1[$ .

**Proposition 4.5** *Let  $f$  and  $g$  be two given functions in  $C^{2,\alpha}(T^n)$ . Then, for initial data  $u_0 \in C^{4,\alpha}(T^n)$  which is strictly quasi-convex there exists a solution  $u(x, t)$  to the corresponding parabolic PDE 1.7 in  $C^2([0, \infty[ \times T^n)$  such that for any  $t > 0$   $u_t \in C^4(T^n)$  and  $u_t$  is strictly quasi-convex. Moreover,*

- *There exists a positive constant  $C_1$ —only depending on upper bounds on the  $C^2$ -norms of  $u_0, f$  and  $g$  and a strict positive lower bound on the eigenvalues of the matrix  $(I + \nabla^2 u_0(x))$ —such that*

$$C_1^{-1} I \leq \left( \nabla^2 u_t(x) + I \right) \leq C_1 I$$

*and a constant  $C_2$  which, moreover, depends on the  $C^{4,\alpha}$ -norm of  $u_0$  and the  $C^{2,\alpha}$ -norms of  $f$  and  $g$  such that*

$$\|u_t\|_{C^4(T^n)} \leq C_2$$

- There exist positive constants  $A_1$  and  $A_0$  such that

$$\sup_{T^n} |u_t - u_\infty| \leq A_1 e^{-t/A_0} \tag{4.7}$$

where  $u_\infty$  is a potential solving the corresponding optimal transport problem (i.e. a solution to the corresponding stationary equation).

- if  $u_0, f$  and  $g$  are in  $C^\infty(T^n)$ , then so is  $u_t$  and  $\|u_t\|_{C^s(T^n)} \leq C_s$  for any positive integer  $s$ .

Note that differentiating the parabolic PDE 1.7 gives

$$\frac{\partial^2 u_t}{\partial t^2} = L_{u_t}[D(u_t)],$$

where  $L_{u_t}$  is the linearization of  $D$  at  $t_t$ . By the first point in the previous theorem  $L_{u_t}$  is a uniformly elliptic second order operator (see formula 9.2 in ‘‘Appendix B’’) and hence

$$\sup_{X \times [0, \infty[} \left| \frac{\partial^2 u_t}{\partial t^2} \right| \leq C_3$$

where  $C_3$  only depends on the constants  $C_1$  and  $C_2$  in the previous proposition and the  $C^2$ -norms of  $f$  and  $g$ .

### 4.1.1 Conclusion of proof of Theorem 1.2 and Corollary 1.3

Consider the log Sinkhorn iteration and define  $D^{(k)}(u)$  to be  $\log(\rho_{ku})$ , as in Eq. 2.22. The proof of Theorem 1.2 follows directly from combining Lemma 4.4 with Propositions 4.3, 4.5 and also using that the corresponding operator  $S^{(k)} = I + k^{-1}D^{(k)}$ , defined by formula 2.20, is clearly increasing and invariant under the additive  $\mathbb{R}$ -action. Finally, Corollary 1.3 follows by combining Theorem 1.2 with the exponential convergence in formula 4.7. Indeed, setting  $m_k = kt_k$  with  $t_k := A \log k$ , where  $A$  is the constant appearing formula 4.7, gives

$$\begin{aligned} \|u_{m_k}^{(k)} - u_\infty\|_{C(X)} &\leq \|u_{m_k}^{(k)} - u_{t_k}\|_{C(X)} + \|u_{t_k} - u_\infty\|_{C(X)} \\ &\leq ACk^{-1} \log k + A_1(k^{-1})^{A/A_0} \end{aligned}$$

where  $C$  is bounded uniformly from above, independently of  $t$  as follows from 1.2 (using the uniform bounds on  $u_t$  in Prop 4.5, which ensure that  $C$  is bounded from above, independently of  $t$ ). Setting  $u_k := u_{m_k}^{(k)}$  this proves the estimate 1.10 when  $A > A_0$ .

Next note that the estimate 1.10 implies the estimate 1.11 for  $\gamma_k$ . Indeed, by definition,  $\gamma_k := e^{-kd_{T^n}^2(x,y)^2/2} e^{-ku_k(x)} e^{-kv_k(y)} \mu^{(k)} \otimes \nu^{(k)}$ , where  $v_k := \nu[u_k]$ . By the estimate 1.10 and Lemma 8.1 there exists a positive number  $C$  such that

$$u_k(x) + v_k(y) \geq u(x) + u^c(y) + Ck^{-1} \log k.$$

The proof is thus concluded by invoking the following elementary inequality, which we claim holds for any strictly quasi-convex function  $u$  in  $C^2(T^n)$  :

$$d_{T^n}^2(x, y)^2/2 + u(x) + u^c(y) \geq \frac{\delta}{2} d_{T^n}(y, F_u(x))^2, \quad F_u(x) = x + (\nabla u)(x), \quad (4.8)$$

under the assumption that

$$0 < \delta I \leq \left( \nabla^2 u(x) + I \right) \leq \delta^{-1} I.$$

To see this we identify  $u$  and  $u^c$  with  $\mathbb{Z}^n$ -periodic functions on  $\mathbb{R}^n$  and set  $\phi(x) := u(x) + |x|^2/2$ . Then  $\phi^*(y) = u^c(y) + |y|^2/2$ , where  $\phi^*$  is the classical Legendre-transform on  $\mathbb{R}^n$  (compare with the proof of Lemma 2.3). Hence, it will be enough to show that

$$-x \cdot y + \phi(x) + \phi^*(y) \geq \frac{\delta}{2} |y - \nabla \phi(x)|^2. \quad (4.9)$$

Indeed, the claimed inequality 4.8 follows from the latter one after replacing  $x$  with  $x + m$  and taking the infimum over all  $m \in \mathbb{Z}^n$ . Now, by assumption  $\nabla^2 \phi \leq \delta^{-1} I$  and hence  $\nabla^2 \phi^* \geq \delta I$  (by 2.8). As a consequence,  $\phi^*(y) \geq \phi^*(y - t) + t \cdot \nabla \phi^*(y - t) + \delta |t|^2/2$  for any  $t \in \mathbb{R}^n$ . Setting  $t := y - \nabla \phi(x)$  and using that  $\phi^*(\nabla \phi(x)) = \nabla \phi(x) \cdot x - \phi(x)$  and  $(\nabla \phi^*)(\nabla \phi(x)) = x$  this implies the desired inequality 4.9.

### 4.1.2 Proof of Corollary 1.4

By Prop 2.1  $C(\mu, \nu) = -J(u)$ , where  $u$  is an optimal transport potential and  $J$  is the Kantorovich functional. Now, since  $u$  and  $u^c$  are Lipschitz continuous (with Lipschitz constant  $\sqrt{n}$ ) we can approximate the integrals defining  $J(u)$  to get

$$-C(\mu, \nu) = J(u) := \int u \mu + \int u^c \nu = \int u \mu^{(k)} + \int u^c \nu^{(k)} + O(k^{-1})$$

(see [13, Page 2]). Next, by Corollary 1.3  $u = u_{m_k}^{(k)} + O(k^{-1} \log k)$ . Similarly, applying Corollary 1.3 to the situation where the roles of  $\mu$  and  $\nu$  have been reversed shows that  $v = v_{m_k}^{(k)} + O(k^{-1} \log k)$  and hence

$$C(\mu, \nu) = - \sum_{x_i \in \Lambda_k} p_i u_{m_k}^{(k)}(x_i) - \sum_{y_i \in \Lambda_k} q_i v_{m_k}^{(k)}(y_i) + O(k^{-1} \log k)$$

The proof is thus concluded by expressing  $u_{m_k}^{(k)}$  and  $v_{m_k}^{(k)}(y_i)$  in terms of  $a^{(k)}(m)$  and  $b^{(k)}(m)$ , respectively (formulas 1.8, 1.12) and using that  $k^{-1} \log p_i = k^{-1} \log(e^{-f}(x_i) (1 + O(k^{-1}))) = O(k^{-1})$  and similarly for  $q_i$ . This proves the first statement in Corollary 1.4. The last statement then follows from the observation that

$$\frac{1}{2}d(\mu^{(k)}, \nu^{(k)})^2 = \frac{1}{2}d(\mu, \nu)^2 + O(k^{-1})$$

This observation is without doubt standard but, for completeness, we provide a proof. Denote by  $X_k$  and  $Y_k$  the support of  $\mu^{(k)}$  and  $\nu^{(k)}$ , respectively. By definition,  $\frac{1}{2}d(\mu^{(k)}, \nu^{(k)})^2$  is the sup of  $\langle c, \gamma \rangle$  over all probability measures on  $X_k \times Y_k$  with marginals  $\mu^{(k)}$  and  $\nu^{(k)}$ , respectively (with the usual identifications). Since any probability measure  $\gamma$  on  $T^n \times T^n$  with marginals  $\mu^{(k)}$  and  $\nu^{(k)}$  is automatically supported on  $X_k \times Y_k$  we may as well take the sup over all such  $\gamma$ . Hence,  $\frac{1}{2}d(\mu^{(k)}, \nu^{(k)})^2 = C(\mu^{(k)}, \nu^{(k)})$  on  $T^n \times T^n$ . Kantorovich duality then gives

$$\begin{aligned} -\frac{1}{2}d(\mu^{(k)}, \nu^{(k)})^2 &= \inf_{u: u^{cc}=u} \left( \int u \mu^{(k)} + \int u^c \nu^{(k)} \right) \\ &= \inf_{u: u^{cc}=u} \left( \int u \mu + \int u^c \nu \right) + O(k^{-1}), \end{aligned}$$

where the last equality follows from the argument in the beginning of the proof of the corollary. Using Kantorovich duality again thus concludes the proof.

### 4.2 Comparison with previous convergence rates for the Sinkhorn algorithm and sharpness

As pointed out in Remark 2.10 it is well-known that, for  $k$  fixed,  $u_m^{(k)}$  converges exponentially in  $C(X)/\mathbb{R}$  towards a fixed point  $u_\infty^{(k)}$  :

$$\left\| u_m^{(k)} - u_\infty^{(k)} \right\|_{C(X)/\mathbb{R}} \leq C_0 e^{-\delta_k m}$$

(i.e. it converges at *linear rate* in the terminology of numerical analysis). The general estimate on  $e^{-\delta_k}$  provided in [36] gives that  $1 - e^{-\delta_k}$  is comparable to  $1 - 2e^{-k}$  if we assume, for simplicity, that  $c$  is Lipschitz continuous with Lipschitz constant 1 and that the diameter of  $X$  is equal one (see [66]). Hence,  $\delta_k$  is comparable to  $e^{-k}$ , which means that, for  $k$  large, one would need to run  $O(e^k)$  iterations to get close to the fixed point  $u_\infty^{(k)}$ . On the other hand experimental findings reported in [27, Remark 4.15] and [35] suggest that in the “geometric settings”, such as in present torus setting,  $\delta_k$  is of the order  $O(k^{-1})$ . Hence, only  $O(k)$  iterations should be needed to get close to the limiting fixed point (see also [59, Prop 18], where this is shown for a modified asymmetric version of the Sinkhorn algorithm, inspired by the auction algorithm). This is confirmed by Theorem 1.2 (and the proof of Corollary 1.3). Indeed, setting  $m_k = t_k k$  and applying Theorem 1.2 gives

$$\left\| u_{m_k}^{(k)} - u_\infty^{(k)} \right\|_{C(X)} \leq C t_k k^{-1} + \|u_{t_k} - u_\infty\|_{C(X)} + \|u_{t_k} - u_\infty\|_{C(X)}$$



Hence, by the exponential convergence in Prop. 4.5 and Theorem 1.1

$$\|u_{m_k}^{(k)} - u_\infty^{(k)}\|_{C(X)} \leq Ctkk^{-1} + Ae^{-tk/A} + \epsilon_k.$$

In particular, by taking  $m_k \sim Tk$  for  $T$  sufficiently large fixed number one can make the right hand side above arbitrarily small, for  $k$  large. But an important point of Corollary 1.3 is that it provides a *quantative* estimate on  $\|u_{m_k}^{(k)} - u_\infty\|_{C(X)}$  when  $m_k \sim Ak \log k$ , where  $u_\infty$  is the optimal transport potential, which arises as the limit of the corresponding parabolic equation. It should also be pointed out that it can be shown that the asymptotics in Prop 4.3 are sharp in the sense that error term can not be improved to  $o(k^{-1})$ , in general. Moreover, inspection of the proof of the Laplace method suggests that  $f$  and  $g$  should have at least two derivatives to get the order  $O(k^{-1})$ . Hence, the rate of convergence in Theorem 1.2 can be expected to be sharp.

### 5 Convergence towards parabolic optimal transport equations on compact manifolds

Let  $X$  and  $Y$  be compact smooth manifolds (without boundary) and  $c$  a lsc function on  $X \times Y$ , taking values in  $] - \infty, \infty]$  which is smooth on the complement of a closed proper subset, denoted by  $\text{sing}(c)$ .

**Remark 5.1** Note that, in contrast to previous sections, we do not assume that  $c$  is continuous on all of  $X \times Y$ . See for example Sect. 6.3.3 for a relevant example where  $c(x, y) = -\log|x - y|$  for  $X$  and  $Y$  given by the unit-sphere in  $\mathbb{R}^{n+1}$  and (hence  $\text{sing}(c)$  is the diagonal in  $X \times Y$ ). In the Riemannian case when  $X = Y$  and  $c = d^2/s$  (where  $d$  is the Riemannian distance function)  $c$  is, of course, continuous on all of  $X \times Y$  and  $\text{sing}(c)$  is non-empty, due to the presence of cut-locus.

We will denote by  $\partial_x$  the vector of partial derivatives defined wrt a choice of local coordinates around a fixed point  $x \in X$ . Given two normalized volume forms  $\mu$  and  $\nu$  in  $\mathcal{P}(X)$  and  $\mathcal{P}(Y)$  we locally express

$$\mu = e^{-f} dx, \quad \nu = e^{-g} dy$$

in terms of the local volume forms  $dx$  and  $dy$  determined by a choice of local coordinates. We assume that  $f$  and  $g$  are  $C^\infty$ -smooth.

Following standard practice we will assume that the cost function satisfies the following assumptions

- (A1) (“Twist condition”) The map  $y \mapsto \partial_x c(x, y)$  is injective for any  $(x, y) \in X \times Y - \text{sing}(c)$
- (A2) (“Non-degeneracy”)  $\det(\partial_{x_i} \partial_{y_j} c)(x, y) \neq 0$  for any  $(x, y) \in X \times Y - \text{sing}(c)$

**Remark 5.2** See [68] for an in depth discussion of various assumption on cost functions. In [68] A1+A2 is called the strong twist condition and as pointed out in [68, Remark

12.23] it holds for the cost function derived from any well-behaved Lagrangian, including the Riemannian setting where  $c = d^2/2$ ).

**Definition 5.3** The space  $\mathcal{H}(X)$  (or  $\mathcal{H}$  for short) of all *smooth potentials* on  $X$  is defined as the subspace of  $C^\infty(X)$  consisting of all  $c$ -convex (i.e. such that  $(u^c)^c = u$ ) smooth functions  $u$  on  $X$  with the property that the subset  $\Gamma_u$  of  $X \times Y$  defined by formula 2.2 is the graph of a diffeomorphism, denoted by  $F_u$  and  $c$  is smooth in a neighborhood of  $\Gamma_u$ .

The definition has been made so that, if  $u \in \mathcal{H}$  and  $(F_u)_*\mu = \nu$ , then  $F_u$  is an optimal map (diffeomorphism) wrt the cost function  $c$  (by Prop 2.1). Accordingly, we will call  $u$  the *potential* of the map  $F_u$ .

**Lemma 5.4** Assume that  $c$  satisfies condition A1. If  $u \in \mathcal{H}$ , then

$$y = F_u(x) \iff (\partial_x c)(x, y) + (\partial_x u)(x) = 0 \tag{5.1}$$

**Proof** By the very definition of  $F_u$  we have that  $y = F_u(x)$  iff  $u^c(y) = -c(x, y) - u(x)$ . In turn, by the definition of  $u^c$  this equivalently means that  $x$  maximizes the usc function  $x' \mapsto \mathcal{C}(x') := -c(x', y) - u(x')$  on  $X$ . Now assume first that  $y = F_u(x)$ . Then, by assumption, the function  $\mathcal{C}$  is smooth close that  $x$  and hence the differential of  $\mathcal{C}$  vanishes at  $x$ , i.e.  $(\partial_x c)(x, y) + (\partial_x u)(x) = 0$  at  $x$ . Conversely, if the latter equation holds then, by the twist assumption A1,  $y$  is uniquely determined by  $x$  and since (as explained above)  $F_u(x)$  satisfies the equation in question it follows that  $y = F_u(x)$ .  $\square$

**Example 5.5** Assume that  $X = Y$  and that  $c(x, y) \geq 0$  with equality iff  $x = y$  and that  $c$  is smooth in a neighborhood of the diagonal. Then  $u = 0$  is in  $\mathcal{H}$  (with  $F_0$  given by the identity) and more examples of potentials are obtained by using that  $\mathcal{H}$  is open in the  $C^\infty$ -topology, in general. In particular, this applies in the ‘‘Riemannian setting’’, where  $c = d^2/2$  on a compact Riemannian manifold  $X$ .

### 5.1 Parabolic optimal transport equations

Consider now the following parabolic PDE, introduced in [44]:

$$\frac{\partial u_t(x)}{\partial t} = \log \det (\partial_x F_{u_t}) - g(F_{u_t}(x)) + f(x) \tag{5.2}$$

expressed in terms of a choice of local coordinates, where  $\det (\partial_x F_u)$  denotes the local Jacobian of the map  $x \mapsto F_u(x)$ . We note that

- The right hand side in the equation 5.2 is globally well-defined (i.e. independent of the choice of local coordinates around  $(x, F_u(x))$  in  $X \times Y$ ). Indeed, it is equal to the logarithm of the quotient of  $(F_u)_*\mu$  and  $\nu$ . Accordingly,  $u$  is a stationary solution iff it is the potential of an optimal transport map.
- Differentiating the Eq. 5.1 reveals that (see [68, Section 12])

$$\det(\partial_x F_u) = \frac{\det ((\partial_x^2 c)(x, F_u(x)) + (\partial_x^2 u)(x))}{\det ((\partial_x \partial_y c)(x, F_u(x)))}. \tag{5.3}$$

### 5.2 Convergence of the Sinkhorn algorithm towards parabolic optimal transport

We next introduce the following stronger local form of the “global” density property appearing in Lemma 3.1.

**Definition 5.6** A sequence (or family) of probability measures  $\mu^{(k)}$  on  $X$ , converging weakly towards a measure  $\mu$ , is said to have the *local density property at order  $s$  (and length scale  $k^{-1/2}$ )* if there exists a positive integer  $s \in [2, \infty[$  and constants  $C_1$  and  $C_2$  such for any fixed  $x_0 \in X$  there exist local coordinates  $\xi := (\xi_1, \dots, \xi_n)$  centered at  $x_0$  with the following property: for any sequence  $h_k$  defined on the polydisc  $2D_k$  of radius  $2 \log k$ , centered at 0 in  $\mathbb{R}^n$  and satisfying  $|\partial^{|\alpha|} h_k|(x) \leq C_1 e^{-|x|^2/C_1}$  on  $2D_k$  for all multiindices  $\alpha$  satisfying  $|\alpha| \leq s$  the bound

$$k^{n/2} \int_{D_k} h_k(F_{x_0}^{(k)})_*(\mu^{(k)} - \mu) \leq C_2 k^{-1},$$

holds, where  $F_{x_0}^{(k)}$  is the scaled coordinate map from a neighborhood of  $x_0$  in  $X$  into  $\mathbb{R}^n$  defined by  $F_{x_0}^{(k)}(x) := (k^{1/2}\xi(x))$ . If this property holds for some  $s \in [2, \infty[$  we will simply say that  $\mu^{(k)}$  has the local density property.

We are now ready for the following generalization of Theorem 1.2 stated in the introduction (as in the torus setting in Sect. 1.1.2 the entropic regularization parameter is expressed as  $\epsilon = k^{-1}$ , where  $k$  is a positive number).

**Theorem 5.7** *Let  $c$  be a function satisfying the assumptions A1 and A2 and  $\mu^{(k)}$  and  $\nu^{(k)}$  be two sequences converging towards  $\mu$  and  $\nu$  in  $\mathcal{P}(X)$  and  $\mathcal{P}(Y)$ , respectively, satisfying the local density property. Given  $u_0 \in \mathcal{H}$  assume that there exists a solution  $u_t$  in  $C^2(X \times [0, T])$  of the parabolic PDE 5.2 with initial condition  $u_0$  and such that  $u_t \in \mathcal{H}$  for any  $t \in [0, T]$ . Denote by  $u_m^{(k)}$  the iteration 2.21 defined by the data  $(\mu^{(k)}, \nu^{(k)}, c)$  and such that  $u_0^{(k)} = u_0$  for any given  $k$ . Then there exists a constant  $C$  such that for any  $(m, k)$  satisfying  $m/k \in [0, T]$*

$$\sup_{T^n} \left| u_m^{(k)} - u_{m/k} \right| \leq C \frac{m}{k} k^{-1},$$

**Proof** The assumptions have been made precisely to ensure that the proof of Theorem 1.2 can be generalized, almost verbatim. Hence, we will be rather brief.

*Step 1:* Let  $\alpha$  be a lower-semicontinuous (lsc) function on  $X$  with a unique minimum at  $x_0$  and assume that  $\alpha$  is  $C^4$ -smooth on a neighborhood of  $x_0$  and  $\nabla^2 \alpha(x_0) > 0$  (in local coordinates centered at  $x_0$ ). Then, in local coordinates centered at  $x_0$ ,

$$k^{n/2} \int_X e^{-k\alpha} h \mu^{(k)} = (2\pi)^{n/2} e^{-k\alpha(x_0)} \frac{h(x_0) e^{-f(x_0)}}{\sqrt{\det(\partial_x^2 \alpha(x_0))}} (1 + O(k^{-1})). \tag{5.4}$$

Using the local density assumption as a replacement of Lemma 8.1, this is shown essentially as before. The point is that the derivatives  $\partial^\alpha$  of any fixed order of  $h_k(x) :=$

$((F_{x_0}^{(k)})^{-1})^*$   $(fe^{-k\alpha})$  are bounded by  $Ce^{-|x|^2/C}$  when  $|x_i| \leq k^{1/2}$  (and in particular, for  $|x_i| \leq 2 \log k$ ). This follows readily from the chain rule, as before. In fact, the proof of formula 5.4 is even a bit simpler than in the torus case, since the last step 3 is not needed when the local density assumption is made at any point  $x_0$ .

Step 2: If  $u \in \mathcal{H}$ , then the following asymptotics holds

$$\rho_{ku}(x) = \det(\partial_x F_u(x))e^{f(x)-g(F_u(x))}(1 + O(k^{-1}))$$

To prove this first observe that if  $u \in \mathcal{H}(X)$ , then  $u^c \in \mathcal{H}(Y)$ . Indeed, by assumption there is a unique  $x(= x_y)$  such that  $y = F_u(x)$ . Moreover, by the very definition of  $F_u$  (Definition 5.3) we can express

$$u^c(y) = -c(x_y, F_u(x_y)) - u(x_y).$$

Since  $c$  is assumed to be smooth in a neighborhood of  $\Gamma_u$  the right hand side above defines a smooth function in  $x$  and since  $F$  is a diffeomorphism it follows that  $u^c(y)$  is smooth. Moreover, by symmetry  $\Gamma_{u^c} = \Gamma_u$ , which can be identified with the graph of the diffeomorphism  $F_u^{-1}$ . This shows that  $u^c \in \mathcal{H}(Y)$  and

$$F_{u^c} = (F_u)^{-1} \tag{5.5}$$

Setting  $y_x := F_u(x)$  and  $x_y = F_{u^c}(y)$  we can now apply the previous step, essentially as before, to get

$$\rho_{ku}(x) = \sqrt{\frac{\det((\partial_x^2 c)(x_y, x) + \partial_x^2 u(x_y))}{\det((\partial_y^2 c)(x, y_x) + \partial_y^2 u^c(y_x))}} e^{f(x)-g(F_u(x))}(1 + O(k^{-1}))$$

Finally, differentiating the relation 5.5 reveals that

$$\det((\partial_y F_{u^c})(y_x)) = \det((F_u(x))^{-1})$$

and hence using Eq. 5.3 and symmetry (which ensures that the denominator appearing in Eq. 5.3 coincides with the one appearing obtained when  $u$  is replaced by  $u^c$ ) concludes the proof of Step 2.

Step 3: Conclusion of proof

The proof is concluded, as before, by invoking Lemma 4.4. □

As pointed out in [44] it follows from standard short-time existence results for parabolic PDEs that the existence of a solution  $u_t$  as in the previous theorem holds for some  $T > 0$  (see, for example, [1, Main Thm 1]). Moreover, by [44] long-time existence, i.e.  $T = \infty$ , holds under the following further assumptions on  $c$  :

- (A3) (“Stay-away property”) For any  $0 < \lambda_1, \lambda_2$  there exists  $\epsilon > 0$  only depending on  $\lambda_1, \lambda_2$  such that  $\lambda_1 \leq |\det \partial_x F_u| \leq \lambda_2 \implies \text{dist}(\Gamma_u, \text{sing}(c)) \geq \epsilon$  for any  $u \in \mathcal{H}$

- (A4) (“Semi-concavity”)  $c$  is locally semi-concave, i.e. the sum of a concave and a smooth function on the domain where it is finite.
- (A5) (“Strong MAW-condition”) The Ma-Wang-Trudinger tensor of  $c$  is bounded from below on  $X \times Y - \text{sing}(c)$  by a uniform positive constant  $\delta$ .

**Theorem 5.8** (Kim-Streets-Warren [44]) *Assume that  $c$  satisfies the assumptions A1–A5. Then, for any given  $u_0 \in \mathcal{H}$  there exists a solution  $u(x, t)$  in  $C^\infty(X \times [0, \infty[)$  of the parabolic PDE 5.2 with initial condition  $u_0$  and such that  $u_t \in \mathcal{H}$  for any  $t > 0$ . Moreover,  $u_t$  converges exponentially in  $C^0(X)$ , as  $t \rightarrow \infty$ , to a potential  $u \in \mathcal{H}$  of a diffeomorphism  $F_u$  transporting  $\mu$  to  $\nu$ , which is optimal wrt the cost function  $c$ .*

**Proof** Let us explain how to translate the result in [44] to the present setting. Following [44] a function  $u \in C^2(X)$  is said to be *locally strictly  $c$ -convex*, if, in local coordinates, the matrix  $(\partial_x^2 c)(x, F_u(x)) + (\partial_x^2 u)(x)$  is positive definite. This condition is independent of the choice of local coordinates. Indeed, it equivalently means that any given  $x_0 \in X$  is a non-degenerate local minimum for the function

$$x \mapsto c(x, F(x_0)) + u(x) \text{ on } X. \tag{5.6}$$

It follows that for any such  $u$  the corresponding map  $F_u$  is a local diffeomorphism. The main result in [44] says that, under the assumptions on  $c$  in the statement above, for any initial datum  $u_0 \in C^\infty(X)$  which is locally strictly  $c$ -convex, there exists a solution  $u(x, t)$  in  $C^\infty(X \times ]0, T])$  which is also locally strictly  $c$ -convex. To make the connection to the present setting first note that if  $u \in \mathcal{H}$  then  $u_0$  is even an absolute minimum for the function 5.6, which is non-degenerate (since  $F_u$  is a diffeomorphism) and hence  $u$  is locally strictly  $c$ -convex. Conversely, if  $u$  is locally strictly  $c$ -convex then [44, Cor 7.1] says that  $u \in C^2(X)$  is  $c$ -convex (i.e.  $(u^c)^c = u$ ) and the proof given in [44, Cor 7.1] moreover shows that  $F_u$  is a global  $C^1$ -diffeomorphism. It then follows from the Assumption A3 that  $c$  is smooth in a neighborhood of  $\Gamma_u$ . Hence,  $u \in C^\infty(X)$  is locally strictly  $c$ -convex iff  $u \in \mathcal{H}$ , which concludes the proof of the theorem. □

**Remark 5.9** Under the assumptions in the previous theorem it follows, in particular, that the optimal transport map is smooth. Conversely, the assumptions are “almost necessary” for regularity of the optimal transport map (see [68, Chapter 12] and reference therein). Also note that the semi-concavity assumption is always satisfied in the case when  $X = Y$  is a compact Riemannian manifold and  $c = d^2/2$  [68, (b), Page 278].

Combining the exponential large-time convergence of  $u_t$ , in the previous theorem, with Theorem 5.7 gives, just as in the torus setting, the following

**Corollary 5.10** *Assume that the cost function  $c$  satisfies the assumptions in Theorem 5.7 (or that  $c = d_{T^n}^2/2$  in the case of the torus  $T^n$ ). Assume, moreover, that  $\mu^{(k)}$  and  $\nu^{(k)}$  satisfy the local density property. Then, for any given  $u_0 \in \mathcal{H}$ , there exists a positive constant  $A_0$  such that for any  $A > A_0$  the following holds:  $u_k(x) := u_{m_k}^{(k)}(x_{i_k})$ ,*

with  $m_k := \lfloor Ak \log k \rfloor$ , converges uniformly to the optimal transport potential  $u(x)$ . More precisely, there exists a constant  $C$  (depending on  $A$ ) such that

$$\sup_{T^n} |u_k - u| \leq Ck^{-1} \log k$$

In turn, this corollary implies, just as before, that the analog of Corollary 1.4 holds (under the same assumptions as in the previous corollary).

**Example 5.11** The assumptions on  $c$  in the previous corollary are satisfied when  $X = Y$  is the  $n$ -sphere and  $c(x, y) = d^2(x, y)/2$  for the standard round metric or  $c(x, y) = -\log|x - y|$ , where  $|x - y|$  denotes the chordal distance (see [44] and references therein). The latter case appears in the reflector antenna problem, as explained in Sect. 6.3.3.

### 5.3 Constructing discretizations using Quasi-Monte Carlo systems

In order to construct discretizations satisfying the local density property (Definition 5.6) on general compact manifolds we will employ Quasi-Monte Carlo systems, familiar from the theory of numerical integration. These can be viewed as generalizations of the standard grids with  $N$  points on the torus.

Let  $(X, g)$  be an  $n$ -dimensional compact Riemannian manifold and denote by  $dV$  the corresponding normalized volume form on  $X$ . Following [12] (and [17] in the case of a sphere) the *worst case error* of integration of points  $\{x_i\}_{i=1}^N \in X$  and weights  $\{w_i\}_{i=1}^N \in \mathbb{R}_+$  (assumed to sum to one) with respect to some Banach space  $W$  of continuous functions on  $X$ , is defined as

$$\text{wce}(\{(x_i, w_i)\}_{i=1}^N) := \sup \left\{ \left| \int f dV - \sum_{i=1}^N f(x_i)w_i \right| : f \in W, \|f\| \leq 1 \right\}$$

We will use the shorthand  $X_N$  for the *weighted point cloud*  $\{(x_i, w_i)\}_{i=1}^N$  and call the corresponding discrete probability measure

$$\lambda_{X_N} := \sum_{i=1}^N w_i \delta_{x_i} \tag{5.7}$$

for the *sampling measure* associated to  $X_N$ . Let now  $W := W_p^s(X)$  be the Sobolev space of all functions  $f$  on  $X$  such that all (fractional) distributional derivatives of order  $s$  are in  $L^p(X)$  and assume that  $s > n/p$  and  $p \in [1, \infty]$  (which ensures that  $W_p^s(X) \subset C^0(X)$ ). Then a sequence of weighted point clouds  $X_N := \{(x_i, w_i)\}_{i=1}^N$  is said to be a *quasi-Monte Carlo system* for  $W_p^s(X)$  if

$$\text{wce}(X_N) \leq \frac{C}{(N^{1/n})^s}$$

for some uniform constant  $C$  (the corresponding lower bound holds for any sequence  $X_N$ ). In view of the applications to the present setting we introduce the following definition (modeled on the corresponding definition for  $p = 2$  in [17]):

**Definition 5.12** Given  $p \in [1, \infty]$ , a sequence of weighted point clouds  $X_N := \{(x_i, w_i)\}_{i=1}^N$  is said to be a *generic Quasi-Monte Carlo  $p$ -system* if  $X_N$  is a quasi-Monte Carlo system for  $W_p^s(X)$  for any positive integer  $s > n/p$ . Moreover,  $X_N$  is a *completely generic QMC system* if it is a generic Quasi-Monte Carlo  $p$ -system for any  $p \in [1, \infty]$ .

**Example 5.13** When  $M = T^n$  the standard grid with  $N$  points on  $T^n$  (i.e. with “edge length”  $1/N^{1/n}$ ) and with all weights taken to be equal to  $1/N$  defines a completely generic Quasi-Monte Carlo system  $X_N$  [13]. Moreover, in the case of the standard  $n$ -dimensional sphere  $S^n$  it follows from [15,17] that there exists a sequence of completely generic quasi-Monte Carlo system with all weights equal to  $1/N$ . The corresponding point sets are taken as *spherical  $t$ -designs* with  $t \sim N^{1/n}$  (these are defined as quadrature points, discussed in Sect. 5.3.1 below, with all weights equal). Such points have been generated for large values of  $N$  [71]. But allowing different weights has the advantage that *explicit* completely generic QMC systems can be constructed, as discussed below.

**Remark 5.14** Recall that the covering radius  $\rho_N$  of a point cloud  $\{x_1, \dots, x_N\}$  on  $(X, g)$  is defined by  $\sup_{x \in X} \min_{i \leq N} d(x, x_i)$ . By [12, Prop 4.1, 4.3], if  $X_N$  is a weighted generic QMC 1-system then the corresponding covering radius  $\rho_N$  is comparable to  $N^{-1/n}$ .

In order to apply this setup to the present setting of the Sinkhorn iteration we will need to adapt the number  $N$  of points to the value of the parameter  $k$  (defined as the invers of the entropic regularization parameter). This is made precise by the following lemma, quantifying how large  $N (= N_k)$  needs to be:

**Lemma 5.15** *Let  $s$  be a positive integer. Assume that  $X_{N_k} := \{(x_i^{(k)}, w_i^{(k)})\}_{i=1}^{N_k}$  is a sequence of quasi-Monte Carlo systems for  $W_p^s(X)$  (where  $s > n/p$ ) indexed by a parameter  $k$ . Then the corresponding sampling measures  $\lambda_k (:= \lambda_{N_k})$  have the local density property at order  $s$  at length scales  $k^{-1/2}$  (Definition 5.6) if the “number condition”*

$$N_k^{-1/n} \leq \frac{C}{k^{1/2}} \frac{1}{k^{(1+\frac{n}{2}(1-1/p))/s}} \tag{5.8}$$

holds for some positive constant  $C$ .

**Proof** Given a sequence  $h_k$  as in Definition 5.6 set  $f_k := (F_{x_0}^{(k)})^*(\chi_k h_k)$ , where  $\chi_k(x) := \chi(x/\log k)$  for a given smooth function  $\chi$  in  $\mathbb{R}^n$  equal to one on the polydisc  $D_1$  centered at 0 and with support in  $2D_1$ . By the assumed quasi-Monte Carlo property there exists a constant  $C$  such that

$$\left| \int_X f_k(\lambda_k - dV) \right| \leq C \frac{1}{N_k^{s/n}} \sum_{|\alpha| \leq s} \left( \int |\partial^\alpha f_k|^p dV \right)^{1/p}.$$

Multiplying both sides with  $k^{n/2}$  and using the chain rule thus gives

$$\begin{aligned} k^{n/2} \left| \int_{D_k} h_k(F_{x_0}^{(k)})_*(\lambda_k - dV) \right| &\leq k^{n/2} \left| \int_{2D_k} \chi_k h_k(F_{x_0}^{(k)})_*(\lambda_k - dV) \right| \\ &= \frac{Ck^{n/2}}{N_k^{s/n}} \left( \sum_{|\alpha| \leq s} k^{|\alpha|/2} \int_{D_{2k}} |\partial^\alpha (\chi_k h_k)|^p (F_{x_0}^{(k)})_* dV \right)^{1/p} \\ &\leq C' \frac{k^{n/2} k^{-n/2p}}{N_k^{s/n}} k^{s/2} \sum_{|\alpha| \leq s} \left( \int_{D_{2k}} |\partial^\alpha (\chi_k h_k)|^p dx \right)^{1/p}. \end{aligned}$$

Now, by assumption, all  $|\partial^\alpha h_k|(x) \leq C e^{-|x|^2/C}$  on  $D_{2k}$ . Moreover,  $|\partial^\alpha \chi_k|$  is uniformly bounded (since  $\partial^\alpha \chi_k = \partial \chi / (\log k)^{|\alpha|}$ ) and hence the sum in the right hand side above is uniformly bounded. Since the number condition ensures that the factor in front of the sum is bounded by a constant times  $k^{-1}$  this concludes the proof.  $\square$

**Remark 5.16** When  $X = T^1$  and  $s = 2$  and  $p = 1$  (which satisfies  $s > n/p$ ) the previous lemma is closely related to the case  $n = 1$  of Lemma 8.1 (recall that the case  $X = T^n$  can be reduced to the case when  $n = 1$ ).

### 5.3.1 Constructing completely generic QMC systems

Finally, we recall how to construct completely generic QMC systems on any  $n$ -dimensional compact Riemannian manifold  $(X, g)$ , generalizing the standard grid on the torus, following [13]. Given a positive number  $W$  (the “bandwidth”) we denote by  $H_{\leq W}(X, g)$  the finite dimensional subspace of  $C^\infty(X)$  consisting of all eigenfunctions of the Laplacian with eigenvalue bounded from above by  $W^2$ . A weighted point cloud  $X_N$  is said to consist of *weighted quadrature points* for  $H_{\leq W}(X, g)$  if the corresponding numerical integration error vanishes for any  $f \in H_{\leq W}(X)$ . For any sufficiently large constant  $C_X$  there exists a sequence of weighted quadrature points  $X_N$  for  $H_{\leq C_X N^{1/n}}(X)$  (as follows from [13, Cor 2.11] and the Weyl asymptotics, saying that the dimension of  $H_{\leq W}(X)$  grows like a constant times  $W^n$ ). Moreover, by [13, Cor 2.13] such a sequence  $X_N$  defines a completely generic QMC system.

**Example 5.17** When  $(X, g)$  is the round two-sphere  $S^2$  it is customary to rewrite  $W^2 = l(l + 1)$ . Letting  $l$  range over all non-negative integers then enumerates all eigenvalues of the Laplacian and the corresponding space  $H_{\leq W}(X)$  then coincides with the space of all spherical polynomials of degree  $l$  (see Sect. 6.3.2). In this case the most commonly used explicit weighted quadrature points are obtained using various longitude-latitude rules; see [40, Section 4.1] and also [31, Thm 3] (applied to  $(l, m) = (0, 0)$ ) for the “equi-angular” case.



### 5.4 Reducing the numbers $N_k$ of discretization points by exploiting higher regularity of the data

Coming back to the setup in the beginning of Sect. 5 we fix two sequences  $X_N$  and  $Y_N$  of  $N$  weighted points on  $X$  and  $Y$ , respectively. We will assume that  $X_N$  and  $Y_N$  are completely generic QMC systems defined with respect to some Riemannian volume forms  $dV_X$  and  $dV_Y$  on  $X$  and  $Y$ , respectively. Fix  $\delta \in ]0, 1/2]$  and index the number of points  $N$  by the parameter  $k$  so that

$$N_k^{1/n} \geq C_\delta k^{1/2+\delta}, \quad \delta > 0, \quad C_\delta > 0 \tag{5.9}$$

(by Remark 5.14 this equivalently means that the covering radius of the corresponding point clouds is of the order  $O(1/k^{1/2+\delta})$ ). We then define a family of “discretizations”  $\mu^{(k)}$  and  $\nu^{(k)}$  of  $\mu$  and  $\nu$ , by proceeding as in the torus case, but replacing the sampling measure  $\delta_{\Lambda_k}$  with the sampling measures corresponding to  $X_N$  and  $Y_N$ . In other words, setting  $\rho_X := \mu/dV_X$  the discrete probability measure  $\mu^{(k)}$  is defined as

$$\mu^{(k)} := \rho_X \lambda_{X_{N_k}} / Z_{N_k},$$

where  $Z_{N_k}$  is the normalizing constant (and  $\nu$  is defined in a similar fashion).

**Lemma 5.18** *The sequences  $\mu^{(k)}$  and  $\nu^{(k)}$  have the local density property.*

**Proof** First observe that it follows directly from the assumptions on  $X_N$  and  $Y_N$  that the corresponding norming constants are equal to  $1 + O(k^{-1})$  (in fact, the error terms are of the order  $O(k^{-\infty})$  since the densities are assumed to be smooth). Hence, it will be enough to show that local density property holds for the sequences  $\rho_X \lambda_{X_{N_k}}$  and  $\rho_Y \lambda_{Y_{N_k}}$ . But then the result is reduced to Lemma 3.1 by simply replacing  $h_k$  with  $\tilde{h}_k := h_k(F_{x_0}^{(k)})^* \rho$ , where  $\rho$  is the density of  $\mu$  or  $\nu$ . Indeed, by Leibniz rule and the chain rule  $\tilde{h}_k$  also satisfies the estimates in the lemma. Thus taking  $s = 1/\delta$  and  $p = 1$  and invoking Lemma 3.1 concludes the proof.  $\square$

This means that Theorem 5.7 and Corollary 5.10 apply to the discretization scheme above, as long as the number  $N_k$  of points satisfies the bound 5.9 for some  $\delta > 0$ .

**Remark 5.19** The previous argument shows that if the densities of  $\mu$  and  $\nu$  are in  $C^s(X)$  for some integer  $s \in [2, \infty[$  and  $s > n/p$  for  $p \in [1, \infty[$  then the local density condition (at order  $s$ ) holds if  $N_k$  satisfies the condition 5.8. In particular, if  $s > n$  one can take  $p = 1$  and then the condition is that  $N_k^{1/n} \geq k^{1/2+1/s}$ . In the particular case of a uniform grid on the torus the condition is thus that the edge-length of the grid is bounded from below by  $Ck^{-(1/2+1/s)}$  (the assumption  $s > n$  is not needed in this case, since one can reduce to the case when  $n = 1$ ).

## 6 Nearly linear complexity on the torus and the sphere

In this section we start by showing that the convergence results in Sect. 5 hold in a more general setting where the kernel  $\mathcal{K}^{(k)}(x, y) := e^{-kc(x,y)}$  is replaced with an

appropriate approximate kernel. This extra flexibility is then applied in the setting of optimal transport on the two-sphere, using “band-limited” heat-kernels, where it leads to a nearly linear algorithmic cost for the corresponding Sinkhorn iterations.

### 6.1 Sequences $c_k$ and approximate kernels $K_k$

Just as in the generalization of the (static) Theorem 3.3, considered in Section 3.7, the (dynamic) Theorem 5.7 can be generalized by replacing the cost function  $c$  with a suitable sequence  $c_k$ . But then the uniform convergence of  $c_k$  towards  $c$  (formula 3.13) has to be supplemented with further asymptotic properties on the complement of the singularity locus of  $c$ . For example, the proof of Theorem 5.7 goes through, almost word for word, if the upper bound corresponding to 3.13 holds globally, i.e.:

$$e^{-kc_k(x,y)} \leq O(e^{\epsilon k})e^{-kc(x,y)} \tag{6.1}$$

(where  $O(e^{\epsilon k})$  denotes a sequence of sub-exponential growth) and  $c_k$  has the following further property: on any given compact subset in the complement of  $\text{sing}(c)$  there exists a strictly positive smooth function  $h_0(x, y)$  and a uniformly bounded sequence  $r_k(x, y)$  of functions such that

$$\mathcal{K}^{(k)}(x, y) := e^{-kc_k(x,y)} = e^{-kc(x,y)}(h_0(x, y) + k^{-1}r_k(x, y)) \tag{6.2}$$

This implies, in particular, that if Theorem 3.3 holds for a given kernel  $\mathcal{K}^{(k)}$ , then it also holds for any other kernel  $\tilde{\mathcal{K}}^{(k)}$  which has error  $O(k^{-1})$  as an approximation relative to  $\mathcal{K}^{(k)}$ , i.e. such that

$$|\mathcal{K}^{(k)} - \tilde{\mathcal{K}}^{(k)}| \leq Ck^{-1}\mathcal{K}^{(k)} \tag{6.3}$$

or such that  $\tilde{\mathcal{K}}^{(k)}$  has absolute error  $e^{-Ck}$ , for  $C$  sufficiently large, i.e.

$$|\mathcal{K}^{(k)} - \tilde{\mathcal{K}}^{(k)}| \leq e^{-Ck}, \quad C > \inf_{X \times Y} c \tag{6.4}$$

### 6.2 Heat kernel approximations in the Riemannian setting

Consider now the Riemannian setting where  $X$  is a compact Riemannian manifold (without boundary) and  $c = d^2/2$  and  $c_k$  is defined in terms of heat kernel (formula 3.14):

**Theorem 6.1** *Let  $X$  be a compact Riemannian manifold (without boundary) and set  $c(x, y) := d(x, y)^2/2$ , where  $d$  is the Riemannian distance function. Then the results in Theorem 5.7 and Corollary 5.10 still hold when the matrix kernel  $e^{-kd^2(x,y)/2}$  is replaced with the heat kernel  $\mathcal{K}_{2k-1}(x, x)$  (at time  $t = 2k^{-1}$ )*

**Proof** As discussed above this follows from the following heat kernel asymptotics (which are a special case of [7, Thm 3.1] and more generally hold for the heat kernel

associated to a suitable hypoelliptic operator). Assume that  $x$  and  $y$  are contained in a compact subset of the complement of the cut-locus. Then

$$\mathcal{K}_t(x, y) = t^{-n/2} e^{-t^{-1}d^2(x,y)/4} (h_0(x, y) + tr_1(t, x, y)),$$

where  $h_0$  is smooth and  $h_0 > 0$  and  $r_1$  is smooth and uniformly bounded on  $]0, t_0] \times X$ . This is not exactly of the form 6.2 due to the presence of the factor  $t^{-n/2} := A_k$ . But it is, in fact, enough to know that 6.2 holds when the right hand side is multiplied with a sequence  $A_k$ , only depending on  $k$ . Indeed, the iteration  $u_m^{(k)}$  is unaltered when the cost function  $c_k(x, y)$  is replaced by  $c_k(x, y) + C_k$  for some constant  $C_k$  (which is consistent, as it must, with the fact that the parabolic equation 5.2 is unaltered when a constant is added to  $c$ ).  $\square$

The use of the heat kernel in the Sinkhorn algorithm for optimal transport on Riemannian manifolds was advocated in [63], where it was found numerically that discretized heat kernels provide substantial speedups, when compared to other methods. The previous theorem offers a theoretical basis for the experimental findings in [63], as long as the discretized heat kernels  $\tilde{\mathcal{K}}^{(k)}$  satisfy one of the approximation properties 6.3 and 6.4 (when compared with the corresponding bona fide heat kernel). However, the author is not aware of any general such approximation results in the discretized setting (but see [24] and references therein for various numerical approaches to discretizations of heat kernels). We will instead follow a different route, based on “band-limited” heat kernels and fast Fourier type transforms, applied to the case when  $X$  is the two-sphere.

### 6.3 Near linear complexity using fast transforms

Each iteration in the Sinkhorn algorithm amounts to computing two vector-matrix products of the form

$$a_i = \sum_{j=1}^N \mathcal{K}(x_i, y_j) b_j, \quad i = 1, \dots, N, \quad (6.5)$$

for a given function  $\mathcal{K}$  on  $X \times Y$  (followed by  $N$  inversions), where  $b$  and  $a$  denote generic “input vector” and “output vectors”, respectively. In general, this requires  $O(N^2)$  arithmetic operations. But, as we will next exploit, in the presence of suitably symmetry fast summations techniques can be used to lower the complexity to nearly linear, i.e. to at most  $CN(\log N)^p$  operations (for some positive constants  $C$  and  $p$ ). Alternatively, separability properties of the kernels in question can often be exploited to directly decrease the complexity (as in [27, Remark 4.17]). For example, consider the case when  $X = X_1 \times X_2$  and  $Y = Y_1 \times Y_2$  and denote by  $\pi_1$  and  $\pi_2$  the projections on the first and second factors, respectively. If  $\mathcal{K}(x, y)$  is *separable*, i.e. factors as

$$\mathcal{K}(x, y) = \mathcal{K}_1(\pi_1(x), \pi_1(y)) \mathcal{K}_2(\pi_2(x), \pi_2(y))$$

one can take the point clouds  $(x_i)_{i \leq 1}$  and  $(y_i)_{i \leq N}$  as the “grid”s induced by given point clouds on each factor with (consisting of  $N_1$  and  $N_2$  points, respectively). Then by first summing over the first factor the complexity of the computation is reduced to  $O(N(N^1 + N^2))$  operations. More generally, if the separability holds with  $r$  factors, then, by induction, the complexity becomes  $O(N(\max_{i \leq r} N_i))$ .

### 6.3.1 Optimal transport on the flat torus

Let us first come back to the case of the flat torus  $T^n$  discretized by the discrete torus  $\Lambda_k$ , considered in Sect. 1.1.3. Since  $\mathcal{K}(x, y) := e^{-kd^2(x,y)}$  is invariant under the diagonal action of the torus  $T^n$  it follows from standard arguments that the sums 6.5 can be computed in  $O(N)(\log N)$  arithmetic operations. Indeed, using the group structure on  $T^n$  we can write  $\mathcal{K}(x, y) = h(x - y)$ , for some function  $h$  on  $\Lambda_k$ . Then the classical convolution theorem of Fourier Analysis, on the discrete torus  $\Lambda_k$  (viewed as an abelian finite group), gives (with  $m_1, \dots, m_N$  the points of the dual discrete torus, identified with  $\Lambda_k$ ) :

$$a_i = \sum_{j=1}^N h(x_i - y_j)b_j = \sum_{j=1}^N \hat{h}(m_j)\hat{b}(m_j)e^{2\pi i m_j \cdot x_i}, \quad \hat{f}(m_j) := \sum_{i=1}^N f_i e^{-2\pi i x_i \cdot m_j}$$

This requires evaluating two Discrete Fourier Transforms (DFT) at the  $N = k^n$  points  $m_1, \dots, m_N$ . Using the Fast Fourier Transform (FFT) this can be done in  $O(N)(\log N)$  arithmetic operations. Note that, since the heat kernel is also torus invariant, the same argument can also be used for the kernel appearing in Theorem 6.1, in the torus case. Alternatively, using that  $\mathcal{K}(x, y)$  is separable on  $T^n$  (since, in general, the squared Riemannian distance function on a Riemannian product is the sum of the squared distances of the factors), the summing can directly be achieved with complexity  $O(N^{1+1/n})$  for any fixed point cloud on  $S^1$ .

### 6.3.2 Optimal transport on the round two-sphere

Consider the round two-sphere  $S^2$  embedded as the unit-sphere in  $\mathbb{R}^3$ . Removing the north and south pole on  $S^2$  we have the standard spherical (longitude-colatitude) coordinates  $(\varphi, \theta) \in [0, 2\pi[ \times ]-\pi, \pi[$ . A complete set of (non-normalized) eigenfunctions for the Laplacian on  $L^2(S^2)$  is given by the spherical harmonics

$$Y_l^m(\varphi, \theta) := e^{im\varphi} P_l^m(\cos \theta), \quad |m| \leq l,$$

which has eigenvalue  $\lambda_{l,m}^2 := l(l + 1)$ . Here  $P_l^m(z)$  denotes, for  $z \in [-1, 1]$ , as usual, the Legendre function of degree  $l$  and order  $m$  (aka the associated Legendre polynomial); see, for example, [31,40].

Given a positive number  $W$  (the “band-width”) we consider the *band-limited heat kernel* on the two-sphere:

$$\mathcal{K}_t(x, y)_W := \sum_{|m| \leq l \leq W} c_{m,l} Y_l^m(x) \overline{Y_l^m(y)}, \quad c_{m,l} := e^{-tl(l+1)} \|Y_l^m\|_{L^2}^{-2} \tag{6.6}$$

(By the spectral theorem this means that  $\mathcal{K}_t(x, y)_W$  is the integral kernel of  $e^{-t\Delta} \Pi_W$  where  $\Pi_W$  is the orthogonal projection onto the space of all band-limited functions).

**Theorem 6.2** *Consider the two-sphere, discretized by a given good Quasi-Monte Carlo (QMC) system and take  $R$  such that  $R > 1$ . Then the analog of all the results in Sect. 1.1.3 are valid when the matrix kernel  $e^{-kd^2(x,y)/2}$  is replaced by the band-limited heat kernel  $\mathcal{K}_{2k^{-1}}(x, y)_{Rk}$ . Moreover, the arithmetic complexity of each Sinkhorn iteration is  $O(N^{3/2})$ .*

**Proof** As recalled in Example 5.11 the cost function  $d(x, y)^2$  on the sphere satisfies the assumptions in Theorem 5.7 (with  $t = \infty$ ) and Corollary 5.10.

*Step 1: The asymptotics 6.1 and 6.2 are satisfied.*

By Theorem 6.1 it is enough to observe that the following basic estimate holds if  $t = 2k^{-1}$  and  $W = Rk$  :

$$|\mathcal{K}_t(x, y) - \mathcal{K}_t(x, y)_W| \leq C_\delta e^{-2R^2k(1-\delta)}$$

for any given  $\delta \in ]0, 1[$ . To prove the estimate note that

$$\begin{aligned} |\mathcal{K}_t(x, y) - \mathcal{K}_t(x, y)_W| &\leq \sum_{l > W} e^{-2k^{-1}l(l+1)} \frac{\|Y_l^m\|_{L^\infty}^2}{\|Y_l^m\|_{L^2}^2} \\ &\leq 2Ck^3 \sum_{l/k > R} e^{-2k(\frac{l}{k})^2} \frac{(l+1)^2}{k^2} \frac{1}{k}, \end{aligned}$$

using that the quotient involving  $Y_l^m$  is dominated by  $Cl$  (and that a given  $l$  corresponds to  $2l + 1$   $ms$ ). Indeed, this is a special case of the the universal  $L^2$ -estimates for an eigenfunction  $\Psi_\lambda$  of the Laplacian (with eigenvalue  $\lambda^2$ ) on a general  $n$ -dimensional Riemannian manifold [62], which gives the growth factor  $C\lambda^{n-1}$ . Finally, dominating the Riemann Gaussian sum above with the integral of the function  $e^{-ks^2} s^2$  over  $[R, \infty[$  concludes the proof.

*Step 3: Complexity analysis*

Using formula 6.6 gives

$$a_i = \sum_{|m| \leq l \leq W} c_{m,l} \hat{b}_{l,m} Y_l^m(x_i), \quad c_{m,l} = e^{-tl(l+1)} \|Y_l^m\|_{L^2}^{-1} \hat{b}_{l,m}, \quad \hat{b}_{l,m} := \sum_{j=1}^N b_j \overline{Y_l^m(x_j)}$$

$\hat{b}_{l,m}$  is the “forward discrete spherical Fourier transform” evaluated at  $(l, m)$ . Once it has been computed for all  $(l, m)$   $a_i$  becomes an “inverse discrete spherical Fourier

transform” (with coefficients  $c_{m,l}\hat{b}_{l,m}$ ). By separation of variables, both these transforms can be computed using a total of  $O(k^3)(= O(N^{3/2}))$  arithmetic operations (see the discussion after formula 1.9 in [56]).  $\square$

In the special case when the good Quasi-Monte Carlo system is defined by an “equi-angular” colatitude-longitude rule of  $N$  weighted points on  $S^2$  the arithmetic complexity of each iteration can be reduced to  $O(N)(\log N)^2$  operations, using a fast discrete spherical Fourier transform [39].

### 6.3.3 Application to the reflector antenna problem

The extensively studied *far field reflector antenna problem* appears when  $X = Y = S^n$  is the  $n$ -dimensional sphere  $S^n$ , embedded as the unit-sphere in  $\mathbb{R}^{n+1}$  and the cost function is taken as  $c(x, y) := -\log|x - y|$  [37,69]. Briefly, the problem is to design a perfectly reflecting surface  $\Sigma$  in  $\mathbb{R}^{n+1}$  with the following property: when  $\Sigma$  is illuminated with light emitted from the origin with intensity  $\mu \in \mathcal{P}(S^n)$  the output reflected intensity becomes  $\nu \in \mathcal{P}(S^n)$  (of course,  $n = 2$  in the usual applications). Representing  $\Sigma$  as a radial graph over  $S^n$  :

$$\Sigma := \{h(x)x\}, \quad x \in S^n,$$

for a positive function  $h$  on  $S^n$  it follows from the reflection law and conservation of energy that  $h$  satisfies the following Monge–Ampère type equation, expressed in terms of the covariant derivatives  $\nabla_i$  in local orthonormal coordinates:

$$\frac{\det(-\nabla_i \nabla_j h + 2h^{-1} \nabla_i h \nabla_j h + (h - \eta)\delta_{ij})}{((|\nabla h|^2 + h^2)/2h)^n} = e^{g(F_h(x)) - f(x)}, \tag{6.7}$$

where  $F_h(x)$  denotes the reflected direction of the ray emitted in the direction  $x$  (and  $\mu$  and  $\nu$  are represented as in 1.3). The equation is also supplemented with the “second boundary value condition” that  $F_h$  maps the support of  $\mu$  onto the support of  $\nu$ . Assuming that  $f$  and  $g$  are smooth there exists a smooth solution  $h$ , which is unique up to scaling (see [19] and references therein).

**Theorem 6.3** *Consider the two-sphere, discretized by a given good Quasi-Monte Carlo (QMC) system. Let  $K^{(k)}$  be the  $N_k \times N_k$  matrix defined by*

$$K_{ij}^{(k)} = |x_i^{(k)} - x_j^{(k)}|^k$$

*Consider the Sinkhorn algorithm associated to  $(p^{(k)}, q^{(k)}, K^{(k)})$ . There exists a positive constant  $A_0$  such that for any  $A > A_0$  the following holds: after  $m_k = \lfloor Ak \log k \rfloor$  Sinkhorn iterations the function  $h_k$  on  $S^n$  defined by the  $k$  th root of  $a_k := a^{(k)}(m_k)$  converges uniformly, as  $k \rightarrow \infty$ , towards a solution  $h$  of the antenna equation 6.7 satisfying the corresponding second boundary value condition. More precisely, there exists a constant  $C$  (depending on  $A$ ) such that*

$$\sup_{S_N} |h_k - h| \leq Ck^{-1} \log k,$$

Moreover, the arithmetic complexity of each iteration is  $O(N^{3/2})$  in general and  $O(N)(\log N)^2$  in the case of an equi-angular grid.

**Proof** As recalled in Example 5.11 the cost function  $-\log|x - y|$  on the sphere satisfies the assumptions in Corollary 5.10. For the complexity analysis we first recall the general fact that any kernel  $K^{(k)}(x, y)$  which is radial, i.e. only depends on  $|x - y|$ , may be expressed as

$$K^{(k)}(x, y) = \sum_{l=1}^{\infty} C_{m,l} Y_l^m(x) \overline{Y_l^m(y)} \tag{6.8}$$

for some positive constants  $C_{m,l}$  (a proof will be given below). By the argument in the proof of Step 2 in Theorem 6.2, it will be enough to show that  $C_{m,l} = 0$  when  $l > k$ , i.e. that  $K^{(k)}(x, y)$  is already band-limited with  $W = k$ . To this end we follow the general approach in [43]. First observe that when  $x$  and  $y$  are in  $S^2$  we can write  $|x - y|^2 = 2(1 - x \cdot y)$ . Hence,  $K_k(x, y) = 2^k f^{(k)}(x \cdot y)$ , where  $f^{(k)}(s) = (1 - s)^k$  for  $s \in [-1, 1]$ . The Legendre polynomials  $p_l (= p_l^0)$  form a base in the space of all polynomials of degree at most  $k$  (which is orthogonal wrt Lebesgue measure on  $[0, 1]$ ) and hence we can decompose

$$2^k f^{(k)} = \sum_{l=1}^k c_l^{(k)} p_l.$$

Formula 6.8 now follows from the classical Spherical Harmonic (Legendre) addition theorem:

$$p_l(x \cdot y) = \frac{4\pi}{2l + 1} \sum_{|m| \leq l} Y_l^m(x) \overline{Y_l^m(y)}.$$

□

## 7 Outlook

### 7.1 Generalized parabolic optimal transport and singularity formation

Consider the setting in Sect. 5 with a cost function  $c$  satisfying the assumptions A1 and A2, but assume for simplicity that  $c$  is globally continuous (for example,  $c = d^2/2$  in the Riemannian setting). Recall that, given initial data  $u_0 \in \mathcal{H}$  and volume forms  $\mu$  and  $\nu$ , the parabolic equation 5.2 admits a smooth solution  $u_t$  on some maximal time-interval  $[0, T[$  and the corresponding maps  $F_{u_t}$  give an evolution of diffeomorphisms of  $X$ . It does not seem to be known whether  $T = \infty$ , in general, i.e. it could be that there are no solutions in  $C^\infty(X \times ]0, \infty[)$ , in general. Still, using the corresponding iteration  $u_k^{(m)}$  (say, defined with respect to  $\mu_k = \mu$  and  $\nu_k = \nu$ ) a generalized notion of solution can be defined:

**Proposition 7.1** *Given a  $c$ -convex function  $u_0$ , define the following curve  $u_t$  of functions on  $X$ , emanating from  $u_0$  :*

$$u_t := \sup \left\{ u_k^{(m)} : (m, k) : m/k \rightarrow t, k \rightarrow \infty \right\}$$

*Then  $u_t$  is  $c$ -convex for any fixed  $t$  (and, in particular, continuous) and there exists a constant  $C$  such that  $\sup_{X \times ]0, \infty[} |u_t(x)| \leq C$ .*

**Proof** *Step 1 there exists a constant such that  $|u_k^{(m)}| \leq C$ .*

By the argument in Step 3 in the proof of Theorem 2.9 we have

$$\mathcal{F}^{(k)}(u^{(k)}) + \mathcal{L}^{(k)}(u_0) \leq I_\mu(u_m^{(k)}) \leq I_\mu(u_0)$$

By Lemma 3.1  $\mathcal{L}^{(k)}(u_0) \rightarrow -\int u_0^c v$  and by Theorem 3.3  $\mathcal{F}^{(k)}(u_m^{(k)}) \rightarrow \inf_{C^0(X)} J$  and hence the lhs above is uniformly bounded in  $k$ . Thus, there exists a constant  $C$  such that  $-C \leq I_\mu(u_m^{(k)}) \leq C$ . The proof of Step 1 is now concluded by observing that there exist constants  $A_1$  and  $A_2$  such that, for any  $c$ -convex function  $u$ ,

$$\sup_X u \leq I_\mu(u) + A_1, \quad \inf_X u \geq I_\mu(u) - A_2.$$

Indeed, both functionals  $f_1(u) := \sup_X u - I_\mu(u)$  and  $f_2(u) := \inf_X u - I_\mu(u)$  are continuous on  $C(X)$  and descend to  $C(X)/\mathbb{R}$ . But the space of  $c$ -convex functions is compact in  $C(X)/\mathbb{R}$  (as is shown precisely as in Lemma 2.7) and hence any continuous functional on the space is uniformly bounded, which implies the two inequalities above.

*Step 2: If  $\{u_\alpha\}_{\alpha \in A}$  is a finite family of  $c$ -convex functions, then  $u := \max\{u_\alpha\}_{\alpha \in A}$  is  $c$ -convex.*

It is enough to find a function  $v \in C(X)$  such that  $u = v^c$ . We will show that  $v := \min\{u_\alpha^c\}_{\alpha \in A}$  does the job. To this end first observe that  $u \mapsto u^c$  is order preserving. Hence,  $u_\alpha \leq u$  implies that  $u_\alpha^c \geq u^c$ , giving  $v \geq u^c$ . Applying the  $c$ -Legendre transform again thus gives  $v^c \leq u^{cc} = u$ . To prove the reversed inequality first observe that, by definition,  $u_\alpha^c \geq v$  and hence  $u_\alpha = (u_\alpha^c)^c \leq v^c$ . Finally, taking the sup over all  $\alpha$  proves the desired reversed inequality.

*Step 3: Conclusion*

Denote by  $\mathcal{K}_t$  the closure in  $C(X)$  of the set  $S_t$  of all  $u_m^{(k)}$  such that  $m/k \rightarrow t$  and  $k \rightarrow \infty$ . By Step 1 and Lemma 2.13  $\mathcal{K}_t$  is compact. Let  $u_1, \dots, u_m$  be the limit points of  $S_t$ . By the argument towards the end of Step 1 in the proof of Theorem 3.3,  $u_i$  is  $c$ -convex. Hence, by Step 2, so is  $u := \max\{u_i\}$ . □

The curve  $u_t$  5.2 is well-defined for any probability measure  $\mu$  and  $\nu$  on compact topological spaces  $X$  and  $Y$  and for any continuous cost function  $c$ . Moreover, if  $\mu$  and  $\nu$  are normalized volume forms on compact manifolds, assumptions A1 and A2 hold and  $u_0 \in \mathcal{H}$ , then, by Theorem 2.9,  $u_t$  coincides with the classical solution of the parabolic equation 5.2, as long as such a solution exists in  $\mathcal{H}$ , i.e. as long as  $F_{u_t}$  is a well-defined diffeomorphism. This makes the curve  $u_t$  a candidate for a solution



to the problem posed in [29, Problem 9] of defining some kind of weak solution to the parabolic equation 5.2, without making assumptions on the MTW-tensor etc (as in Theorem 5.8). The connection to the Sinkhorn algorithm also opens the possibility of numerically exploring singularity formation of classical solutions  $u_t$  to the parabolic equation 5.2 as  $t \rightarrow T$  (the maximal existence time). As indicated in [29, Problem 9] one could expect that the first derivatives of a classical solution  $u_t$  blow up along a subset  $S$  of  $X$  of measures zero as  $t \rightarrow T$  (moreover, in the light of the discussion in [29, Problem 8], the subset  $S$  might be expected to be rectifiable and of Hausdorff codimension at least one).

Finally, it may be illuminating to point out that, even if the construction of the generalized solution  $u_t$  may appear to be rather non-standard from a PDE point of view it bears some similarities to the method of “vanishing viscosity” for constructing solutions to PDEs by adding small regularizing terms. This is reinforced by the interpretation of the inverse of  $k$  as an “entropic regularization parameter” discussed in the introduction of the paper (also note that the approximations  $u_{m_k}^{(k)}$  are smooth when the heat kernel is used, as in Theorem 6.1). One is thus lead to ask whether, under suitable regularity assumptions on  $(\mu, \nu, c)$  the curve  $u_t$  is a viscosity solution of the parabolic PDE [25]?

### 8 “Appendix A”: proof of Lemma 4.2 (discrete Laplace method)

We start with the following elementary lemma:

**Lemma 8.1** *Let  $h_k$  be a sequence of continuous convex functions on the polydisc  $D_k$  in  $\mathbb{R}^n$  of radius  $\log k$  centered at 0 such that there exists a constant  $C$  such such that  $|h_k(x)| + |\nabla_{e_1} \nabla_{e_2} h_k|(x) \leq C e^{-|x|^2/C}$ , where  $e_1$  and  $e_2$  are any two unit-vectors. Then there exists a constant  $C'$  (only depending on  $C$ ) such that*

$$\left| k^{-n/2} \sum_{x_i^{(k)} \in D_k \cap (k^{-1/2}\mathbb{Z})^n} h_k(x_i^{(k)}) - \int_{D_k} h_k dx \right| \leq C'/k$$

**Proof** By restricting the integration to one variable at a time it is enough to consider the case when  $n = 1$ . Fix  $x_i^{(k)}$ , which, by symmetry, may be assumed non-negative. For any fixed  $x$  in the interval  $I_k(x_i^{(k)})$  centered at  $x_i^{(k)}$ , of length  $k^{-1/2}$ , Taylor expanding  $h_k$  gives

$$|h_k(x_i^{(k)}) - h_k(x) - (x_i^{(k)} - x)h'_k(x_i^{(k)})| \leq k^{-1} C e^{-(x_i^{(k)})^2/C} \leq C k^{-1} e^{-(x-1/2k^{1/2})^2/C},$$

using that  $e^{-x^2/C}$  is decreasing in the last step. By symmetry, the integral over  $I_k(x_i^{(k)})$  of the linear term  $(x_i^{(k)} - x)h'_k(x_i^{(k)})$  vanishes, giving

$$k^{-1/2} h_k(x_i^{(k)}) = \int_{I_k(x_i^{(k)})} h_k(x_i^{(k)}) dx = \int_{I_k(x_i^{(k)})} h_k(x) dx + \epsilon_i^{(k)},$$

where

$$|\epsilon_i^{(k)}| \leq Ck^{-1} \int_{I_k(x_i^{(k)})} e^{-(x-1/2k^{1/2})^2/C} dx$$

Hence, summing over all points  $x_i^{(k)} \in D_k \cap k^{-1/2}\mathbb{Z}$  except the end points and using that  $|h_k(x_i^{(k)})| \leq Ce^{-(\log k)^2/C} \leq C'/k^{-1}$  at the end points gives

$$\begin{aligned} & \left| k^{-1/2} \sum_{x_i^{(k)} \in D_k \cap k^{-1/2}\mathbb{Z}} h_k(x_i^{(k)}) - \int_{D_k} h_k dx \right| \\ & \leq C'k^{-1} + k^{-1}C'' \int_{0 \leq s \leq \log k} e^{-(s-1/2k^{1/2})^2/C} ds, \end{aligned}$$

which concludes the proof. □

In the sequel we will denote by  $+$  the ordinary group structure on  $T^n$  and by  $0$  the zero with respect to the group structure. Without loss of generality we may as well assume that  $\alpha(x_0) = 0$  and  $h(x_0) = 1$ . We will denote by  $O(k^{-1})$  any sequence of functions satisfying  $|O(k^{-1})| \leq Ck^{-1}$ , for a constant  $C$  depending on data as in the statement of the proposition to be proved.

**Step 1: Localization to a polydisc  $U_k$  of radius  $k^{-1/2} \log k$  centered at  $x_0$**

First fix a neighborhood  $U$  of  $x_0$ . Since  $\alpha$  is assumed lower-semicontinuous we have  $\alpha(x) \geq \delta > 0$  on  $U$  and hence

$$k^{n/2} \int_{T^n - U} e^{-k\alpha} h \delta_{\Lambda_k} \leq k^{n/2} \sup_{T^n} |h| e^{-k\delta} = k^{n/2} O(1)k^{-1},$$

using that  $\delta_{\Lambda_k}$  is a probability measure. Next, assume that  $U$  is a small coordinate neighborhood of  $x_0$  and denote by  $U_k$  the polydisc of radius  $k^{-1} \log k$  centered at  $x_0$ . By assumption we may assume that  $\alpha(x) \geq |x - x_0|^2/C$  on  $U$ . Hence,

$$k^{n/2} \int_{U - U_k} e^{-k\alpha} h \delta_{\Lambda_k} \leq k^{n/2} \sup_{T^n} |h| e^{-(\log k)^2} = O(1)k^{-1}$$

**Step 2: The case when  $x_0 = 0$  in  $T^n$**

Introducing the notation  $\alpha^{(k)}(x) := k\alpha(k^{-1/2}x)$  and  $h_k(x) := h(k^{-1/2}x) \exp(-\alpha^{(k)}(x))$  we can write

$$I_k := k^{n/2} \int_{U_k} e^{-k\alpha} h \delta_{\Lambda_k} = k^{-n/2} \sum_{x_i^{(k)} \in D_k \cap (k^{-1/2}\mathbb{Z})^n} h_k(x_i^{(k)})$$

Now, Taylor expanding  $\alpha$  and denoting by  $p^{(3)}$  the third order term (i.e. defining a polynomial of homogeneous degree three) gives, when  $|x| \leq k^{1/2}/C$  (and, in particular, when  $|x| \leq \log k$ )

$$\alpha^{(k)}(x) = Ax \cdot x/2 + k^{-1/2} p^{(3)} + k^{-1} O(|x|^4)$$

Thus  $h_k(x)$  may be Taylor expanded as follows

$$h_k(x) = e^{-Ax \cdot x/2} \left( 1 + k^{-1/2}(q^{(1)} + p^{(3)}) + k^{-1} O(|x|^4) \right), \quad q^{(1)}(x) := (\nabla h)(0) \cdot x$$

Next note that  $h_k(x)$  satisfies the assumptions of the previous lemma. Indeed, by the chain rule  $(\nabla_{e_1} \nabla_{e_2} \alpha^{(k)})(x)$  is, when  $|x_i| \leq k^{1/2}/C$ , equal to the corresponding second derivatives of  $\alpha$  on  $U$ , which are uniformly bounded on  $U$ , by assumption. Hence, applying the lemma in question gives

$$I_k = \int_{D_k} h_k dx + O(k^{-1})$$

This shows that in the present discrete setting we get the same result, up to the negligible error term  $O(k^{-1})$ , as the ordinary Laplace method of integration, which can hence be invoked to conclude. For completeness we provide an alternative direct argument. By the expansion above we have

$$I_k = \int_{D_k} e^{-Ax \cdot x/2} \left( 1 + k^{-1/2}(q^{(1)} + p^{(3)}) + k^{-1} O(|x|^4) \right)$$

Using the exponential decay of  $e^{-Ax \cdot x/2}$  the integral may be taken over all of  $\mathbb{R}^n$ , up to introducing an error term  $O(k^{-\infty})$ . Hence computing the Gaussian integral concludes the proof, once one has verified that the integrals over  $q^{(1)}$  and  $p^{(3)}$  vanish. In the case when  $A$  is the identity matrix the vanishing follows directly from the fact that  $q^{(1)}$  and  $p^{(3)}$  are odd. In the general case one first observes that the space of polynomials of homogeneous degree  $d$  on  $\mathbb{R}^n$  is invariant under the action of the space of invertible linear maps. Hence the problem reduces, by a linear change of variables, to the previous case of an identity matrix.

**Step 3: The case of a general  $x_0$**

Set  $\tilde{\alpha}(x) := \alpha(x + x_0)$  and  $\tilde{f}(x) := f(x + x_0)$  and decompose  $x_0 = m_k + r_k$  where  $m_k \in \Lambda_k$  and  $|r_k| \leq 1/k$  (where we have identified a small neighborhood of 0 in  $T^n$ , containing  $r_k$  with  $\mathbb{R}^n$ ). Then we can write

$$\int e^{-k\alpha} f \delta_{\Lambda_k} = \int e^{-k\tilde{\alpha}} \tilde{f} \delta_{(\Lambda_k - r_k)}$$

Indeed, for any function  $h$  on  $T^n$  we have, since  $m_k \in \Lambda_k$ ,

$$\sum_{x_i \in \Lambda_k} h(x_i + m_k + r_k) = \sum_{x_i} h(x_i + r_k)$$

Now, we note that the conclusion in the previous lemma remains true when  $\Lambda_k$  is replaced by the shifted set  $\Lambda_k - r_k$  (with essentially the same proof) and hence we can conclude as before.

### 9 “Appendix B”: Proof of Prop 4.5 (the parabolic PDE on the torus)

The main difference between Proposition 4.5, concerning the  $n$ -dimensional torus  $T^n$  and the corresponding result for cost functions  $c$  on general compact manifolds in [44] (see Theorem 5.8) is that in [44] it is assumed that the strong M-W-T condition holds, i.e. the Ma-Wang-Trudinger tensor of  $c$  is bounded from below by a *strictly* positive constant, while the Ma-Wang-Trudinger tensor vanishes identically for the standard cost function  $c$  on  $T^n$ . The only place where the strong M-W-T condition enters in [44] is in the proof of the  $C^2$ -estimate of  $u_t$ . Here we will provide a proof of the  $C^2$ -estimate in question on  $T^n$ , building on estimates in [45]. To make the connection to the setting in [45] we identify a solution  $u_t$  of the parabolic flow appearing in Prop 4.5 with a periodic function in  $\mathbb{R}^n$ , i.e.  $u_t$  is a  $\mathbb{Z}^n$ -invariant function in  $\mathbb{R}^n$ . Using the notation in [45] the parabolic equation in question may be expressed as

$$\frac{\partial u_t(x)}{\partial t} = D[u_t], \quad D[u_t] := \log \det \left( \nabla^2 u_t(x) - A \right) - \log B(x, \nabla u_t(x)) \quad (9.1)$$

with

$$A = -I, \quad \log B(x, \nabla u(x)) = g(x + \nabla u(x)) - f(x).$$

Recall that the initial data  $u_0$  is assumed to be a periodic function in  $C^{4,\alpha}(\mathbb{R}^n)$ , which is strictly quasi-convex, i.e. the matrix  $(\nabla^2 u_0)(x) + I$  is positive definite for all  $x$ . This is precisely the parabolic equation appearing in [45] in the case of the smooth cost function  $c(x, y) = |x - y|^2/2$  in  $\mathbb{R}^n \times \mathbb{R}^n$ . However, in [45] the equation is considered on a compact domain  $\Omega$  in  $\mathbb{R}^n$  with additional boundary conditions, while here we assume that  $u_t$  is a periodic function on  $\mathbb{R}^n$ . Denote by  $L_u$  the linearization of the operator  $D$  at  $u$ . A direct calculation reveals that

$$L_u = \Delta_w - (\nabla g) \cdot \nabla, \quad (9.2)$$

where  $\Delta_w$  denotes the Laplacian operator defined wrt the Hessian metric

$$(w_{ij}) = (\nabla^2 u_t)(x) + I, \quad (9.3)$$

i.e., denoting by  $(w^{ij})$  the inverse matrix,

$$\Delta_w = \sum_{i,j} w^{ij} \nabla_i \nabla_j$$

and the gradient  $\nabla g$  is evaluated at the point  $x + \nabla u(x)$  and  $(\nabla g) \cdot \nabla$  denotes the corresponding first order differential operator. Note that  $u$  is strictly quasi-convex iff  $L_u$  is an elliptic operator.

### 9.1 The $C^2$ -estimate

First observe that

$$|\nabla u_t| \leq \sqrt{n}. \tag{9.4}$$

This follows directly from the fact that  $u_t$  is  $c$ -convex, when viewed as a function on the torus  $T^n$  (see the beginning of Sect. 2.1.1).

**Lemma 9.1** *Let  $u_t(x)$  be a solution to the parabolic equation 9.1 on  $T^n \times [0, t_{max}[$  which is  $C^2$ -smooth in  $x$  and jointly  $C^1$ -smooth in  $x$  and  $t$ . There exists a constant  $C_0$ , independent of  $t$ , such that*

$$|\nabla^2 u_t| \leq C_0.$$

*More precisely,  $C_0$  only depends on upper bounds on the  $C^2$ -norms of  $u_0, f$  and  $g$  and a strict positive lower bound on the eigenvalues of the matrix  $(I + \nabla^2 u_0(x))$ .*

**Proof** We will adapt the proof of the interior  $C^2$ -estimate in [45, Thm 10.1] to the present periodic setting (the estimate in [45, Thm 10.1] is a parabolic version of the corresponding estimate in the elliptic setting considered in [65]). Following the convention used in [45] we will omit the sum sign  $\sum$  from the formulas below. Moreover,  $C$  will denote a constant only depending on an upper bound on  $|\nabla u_t|$  and the first and second order derivatives of  $\log B$  (i.e. of  $g$  and  $f$ ), whose precise value may change from line to line. By the a priori estimate 9.4 these bounds are thus independent of  $t$ . Given a unit-vector  $\xi \in S^{n-1}$  and  $a > 0$  we consider the function

$$v(x, t) = w_{\xi\xi}(x) + a|\nabla u_t(x)|^2, \quad w_{\xi\xi}(x) := w_{ij}(x)\xi_i\xi_j,$$

which coincides with the function defined in the beginning of the proof of [45, Thm 10.1]. Fix a point  $x_0 \in \mathbb{R}^n$  and take  $\xi$  to be a unit-vector  $\xi_0$  maximizing  $w_{ij}(x_0)\xi_i\xi_j$ .

*Step 1:* The following inequality holds at the fixed point  $x_0$  if  $w_{ij}(x_0)\xi_i\xi_j \geq 1$

$$\left(-\frac{\partial}{\partial t} + L_{u_t}\right)[v] \geq (2a - C)w_{ii} - Cw^{ii} - Ca,$$

This inequality is shown in the course of the proof of [45, Thm 10.1], only using that the MTW-tensor is non-negative (in the present case where the matrix  $A$  is constant the MTW-tensor vanishes identically). Indeed, first it is shown, by differentiating the parabolic equation for  $u_t$ , that

$$\left(-\frac{\partial}{\partial t} + L_{u_t}\right)[v] \geq \frac{w^{il}w^{jk}\nabla_{\xi}w_{ij}\nabla_{\xi}w_{lk}}{(w_{\xi\xi})^2} - \frac{w^{ij}\nabla_iw_{\xi\xi}\nabla_jw_{\xi\xi}}{w_{\xi\xi}} + (2a - C)w_{ii} - Cw^{ii} - Ca.$$

Then it is shown that the sum of the first two terms is bounded from below by a constant times  $-w^{ii}$ . But, in fact, in the present case, where the matrix  $A$  is constant, the sum in question is actually non-negative (see the proof of [50, Thm 3.3] Theorem 3.3 for a different proof of this fact).

*Step 2:* The following inequality holds at any point  $x$

$$L_u[-u] \geq w^{ii} - C|\nabla u|.$$

To see this first observe that the following identity holds:

$$-\Delta_w u + n = w^{ii},$$

This is seen by multiplying the matrix identity 9.3 with the inverse  $(w^{ij})$  of  $(w_{ij})$  and taking the trace. Hence,

$$L[-u] \geq w^{ii} + (\nabla g) \cdot \nabla u$$

which concludes the proof of Step 2.

*Step 3:* Conclusion using the parabolic maximum principle

Now fix  $b > 0$  and consider the function

$$v_{a,\kappa}(x, t) = w_{ij}\xi_i\xi_j + a|\nabla u_t|^2 - b(u_t - u_t(0)).$$

Combining the previous steps we get, at the fixed point  $x_0$  :

$$\mathcal{L}_{a,\kappa}[v] \geq (2a - C)w_{ii} + (b - C)w^{ii} - C(a + b).$$

Take the parameters  $a$  and  $b$  sufficiently large to ensure that  $2a - C > 0$  and  $b - C > 0$  (this condition is independent of the choice of fixed point  $x_0$ ). Now fix  $T < t_{max}$  and consider the smooth function  $v_{a,\kappa}$  on  $\mathbb{R}^n \times [0, T] \times S^{n-1}$ . Since  $v_{a,\kappa}$  is periodic in the  $x$ -variable its sup is attained at some  $(x_0, t_0, \xi_0)$  in  $[0, 1]^n \times [0, T] \times S^{n-1}$ . Note that  $\xi_0$  maximizes  $w_{ij}(x_0)\xi_i\xi_j$ . We may assume that  $w_{ij}(x_0)\xi_i\xi_j \geq 1$  (otherwise we are already done). Now, if  $t_0 > 0$  then

$$0 + 0 \geq \left(-\frac{\partial}{\partial t} + L\right)[v_{a,\kappa}]$$

at  $(x_0, t_0)$ . But then the inequality in Step 1 implies an upper bound on  $w_{ii}$  at  $(x_0, t_0)$  only depending on  $C$ . Finally, since  $a|\nabla u_t|^2 - b(u_t - u_t(0))$  is uniformly bounded by a constant this shows that the sup of  $w_{ii}$  over  $\mathbb{R}^n \times [0, T]$  is also uniformly bounded by a constant. Finally, if  $t_0 = 0$  then we get an upper bound on  $w_{ii}$  in terms of the Hessian of  $u_0$ . □

### 9.2 Conclusion of the proof of Prop 4.5

The rest of the argument proceeds as in [44], but for the convenience of the reader we provide some details, highlighting the dependence on the regularity assumptions on the data  $u_0, f$  and  $g$ .

#### 9.2.1 Short-time existence

First assume that  $u_0 \in C^{2,\alpha}(T^n)$  and that  $u$  is strictly quasi-convex. Then the linearization  $L_{u_0}$  is uniformly elliptic and hence standard short-time existence results for parabolic equations imply that there exists a maximal  $t_{max} > 0$  and a unique solution  $u(x, t)$  to the equation 9.1 on  $T^n \times [0, t_{max}[$  with the property that  $u_t$  is in  $C^{2,\alpha}(X)$  and the corresponding Hölder derivatives of  $u_t$  are continuous as  $t \rightarrow 0$ . In particular, the norms  $\|u_t\|_{C^{2,\alpha}(T^n)}$  are uniformly bounded for  $t \in [0, T]$  for any given  $T < t_{max}$ . This follows, for example, from [1, Main Thm 1] or [41, Thm 1.2]. Next given two unit-vectors  $e_1$  and  $e_2$  denote by  $\nabla_1$  and  $\nabla_2$  the corresponding derivations. Applying  $\nabla_1$  to the equation 9.1 gives

$$\frac{\partial(\nabla_1 u_t)}{\partial t} = L_{u_t}[\nabla_1 u_t] - \nabla_1 f. \tag{9.5}$$

Next assume that  $u_0 \in C^{4,\alpha}(T^n)$ . Since  $L_{u_t}$  is uniformly elliptic when  $t \leq T < t_{max}$  and the coefficients of  $L_{u_t}$  are Hölder continuous (with exponent  $\alpha$ ) global linear parabolic Schauder estimates [41, Thm 2.3] yield, since  $e_1$  was arbitrary, that  $\nabla u_t \in C^{2,\alpha}(T^n)$  and that  $\|\nabla u_t\|_{C^{2,\alpha}(T^n)}$  is uniformly bounded for  $t \in [0, T]$ . Now, applying  $\nabla_2$  to the equation 9.5 gives

$$\begin{aligned} \frac{\partial(\nabla_2 \nabla_1 u_t)}{\partial t} &= L_{u_t}[\nabla_2 \nabla_1 u_t] - F_t, \quad F_t := \sum_{i,j} (\nabla_2 w^{ij}) \nabla_i \nabla_j (\nabla_1 u_t) \\ &\quad - \nabla_2 (\nabla g) \cdot \nabla (\nabla_1 u_t) - \nabla_2 \nabla_1 f. \end{aligned}$$

By assumption,  $\nabla^2 g$  and  $\nabla^2 f$  are Hölder continuous and hence (also using that  $\nabla u_t \in C^{2,\alpha}(T^n)$ ) the function  $F_t$  is Hölder continuous on  $T^n$  for some positive Hölder exponent  $\alpha' (\leq \alpha)$ . We can thus apply the global linear parabolic Schauder estimates again (now to the linear operator  $L_t[\cdot] - F_t$ ) and deduce that the norms  $\|\nabla^2 u_t\|_{C^{2,\alpha'}(T^n)}$  are uniformly bounded for  $t \in [0, T]$ .

#### 9.2.2 $t_0$ -independent a priori estimates

Next note that there exist constant  $C$  (independent of  $x$  and  $t$ ) such that

$$C^{-1}I \leq (w_{ij}(x, t)) \leq CI. \tag{9.6}$$

Indeed, the upper bound is the content of Lemma 9.1 and to prove the lower bound first note that the time derivative  $\dot{u}_t$  of  $u_t$  solves the linear parabolic equation

$$\frac{\partial U_t}{\partial t} = L_{u_t}[U_t], \tag{9.7}$$

Hence, by the parabolic maximum principle

$$\sup_{T^n \times [0, t_{max}[} |\dot{u}_t| \leq \sup_{T^n} |\dot{u}_0|.$$

Plugging this bound into the equation 9.1 this means that the determinant of  $w_{ij}(x, t)$  is uniformly bounded from below (by a constant only depending on the sup norm of  $g$  and the Hessian of  $u_0$ ). Combined with the upper bound in 9.6 this implies the lower bound in 9.6. Next note that the bound 9.6 says that the linearized operator  $L_{u_t}$  is uniformly elliptic (with constants independent of  $t$ ). Hence, the Krylov-Safanov theory for fully non-linear parabolic equations yield uniform interior  $C^{2,\alpha}(T^n)$ -estimates for  $u_t$  for some  $\alpha > 0$  (i.e. they hold for  $t \geq \epsilon > 0$  with constants only depending on  $\epsilon$ , the  $C^2$ -norms of  $u_t$  and  $f$  and  $g$ , i.e. on the previous uniform constant  $C$ ). It follows that  $t_{max} = \infty$  (otherwise we could restart the flow again). Then, by the argument in the previous section, using parabolic Schauder estimates, we get a bound

$$\left\| \nabla^2 u_t \right\|_{C^{2,\alpha'}(T^n)} \leq C'$$

for a constant independent of  $t \in [0, \infty[$ . In particular, such a bound holds for  $\|u_t\|_{C^4(T^n)}$ , which concludes the proof of the first point in Prop 4.5.

### 9.2.3 Exponential convergence

Finally, the exponential convergence follows from the generalization of the Li-Yau Harnack inequality in [40, Thm 5.2, Cor 5.3]. Briefly, by [40, Thm 5.2, Cor 5.3] the a priori bounds in the first point of Prop 4.5 imply that there exists a constant  $C'$  such that any positive solution  $U_t$  to the linear parabolic equation 9.7 satisfies

$$\sup_{T^n} U_{t+1/2} \leq C' \inf_{T^n} U_t$$

when  $n \geq 3$ . Using a standard induction argument this implies an exponential decay of the sup-norm of the oscillation of  $\dot{u}$ , which in turn implies the exponential convergence in Prop 4.5. The assumption that  $n \geq 3$  in [40, Cor 5.3] is then bypassed by taking a product of  $T^n$  with  $T^2$ , just as in [44, Section 7.1.2].

**Acknowledgements** Open access funding provided by Chalmers University of Technology. I am grateful to Klas Modin for many discussions and, in particular, for drawing my attention to the recent numerical works on the applications of the Sinkhorn algorithm to Optimal Transport, which was the starting point of the present paper. Also thanks to Gabriel Peyré and Jean-David Benamou for helpful comments and discussions, Jean Feydy for sharing his thesis draft and the two referees for their comments that greatly improved the exposition. Lacking background in Numerics and Optimal Transport I apologize for any omission in accrediting results properly. This work was supported by grants from the ERC, the KAW foundation and the Göran Gustafsson foundation.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Baker, C.: The mean curvature flow of submanifolds of high codimension. [arXiv:1104.4409](https://arxiv.org/abs/1104.4409)
2. Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84**(3), 375–393 (2000)
3. Benamou, J.-D., Carlier, G., Cuturi, M., Nenna, L., Peyré, G.: Iterative Bregman projections for regularized transportation problems. *SIAM J. Sci. Comput.* **37**(2), 1111–1138 (2015)
4. Benamou, J.D., Duval, V.: Minimal convex extensions and finite difference discretization of the quadratic Monge–Kantorovich problem. *Eur. J. Appl. Math.* **30**(6), 1041–1078 (2017)
5. Benamou, J.-D., Froese, B.-D., Oberman, A.M.: Numerical solution of the optimal transportation problem using the Monge–Ampère equation. *J. Comput. Phys.* **260**, 107–126 (2014)
6. Benamou, J.-D., Martinet, M.: Capacity Constrained Entropic Optimal Transport, Sinkhorn Saturated Domain Out-Summation and Vanishing Temperature. Preprint at <https://hal.archives-ouvertes.fr/hal-02563022>
7. Ben Arous, G.: Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Annales scientifiques de l'É.N.S. 4 e série, tome 21*(3), 307–331 (1988)
8. Berman, R.J.: Relative Kähler–Ricci flows and their quantization. *Anal. PDE* **6**(1), 131–180 (2013)
9. Berman, R., Boucksom, S., Witt Nyström, D.: Fekete points and convergence towards equilibrium measures on complex manifolds. *Acta Math.* **207**(1), 1–27 (2011)
10. Berman, R.J., Boucksom, S., Guedj, V.: Zeriahi: a variational approach to complex Monge–Ampère equations. *Publications math. de l'IHÉS* 1–67 (2012)
11. Berman, R.J.: Convergence rates for discretized Monge–Ampère equations and quantitative stability of optimal transport (2018). [arXiv:1803.00785](https://arxiv.org/abs/1803.00785)
12. Breger, A., Ehler, M., Graef, M.: Points on manifolds with asymptotically optimal covering radius. (2016) [arXiv:1607.06899](https://arxiv.org/abs/1607.06899)
13. Brandolini, L., Choirat, C., Colzani, L., Gigante, G., Seri, R., Travaglini, G.: Quadrature rules and distribution of points on manifolds. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **13**(4) (2014)
14. Brenier, Y., Frisch, U., Hénon, M., Loepe, G., Matarrese, S., Mohayaee, R., Sobolevskii, A.: Reconstruction of the early Universe as a convex optimization problem. *Mon. Not. R. Astron. Soc.* **346**(2), 501–524 (2003)
15. Bondarenko, A., Radchenko, D., Viazovska, M.: Optimal asymptotic bounds for spherical designs. *Ann. Math.* **178**(2), 443–452 (2013)
16. Brenier, Y.: Polar factorization and monotone rearrangement of vector-valued functions. *Commun. Pure Appl. Math.* **44**(4), 375–417 (1991)
17. Brauchart, J.S., Saff, E.B., Sloan, I., Womersley, R.S.: QMC designs: optimal order quasi Monte Carlo integration schemes on the sphere. *Math. Comput.* **83**(290), 2821–2851 (2014)
18. Caffarelli, L.A., Kochengin, S.A., Oliker, V.I.: On the numerical solution of the problem of reflector design with given far-field scattering data. In *Monge Ampère equation: applications to geometry and optimization* (Deerfield Beach, FL, 1997), vol. 226 of *Contemp. Math.*, pages 13–32. Amer. Math. Soc., Providence, RI, (1999)
19. Caffarelli, L.A., Gutiérrez, C.E., Huang, Q.: On the regularity of reflector antennas. *Ann. Math.* **167**(1), 299–323 (2008)
20. Cao, H.D.: Deformation of Kähler metrics to Kähler–Einstein metrics on compact Kähler manifolds. *Invent. Math.* **81**(2), 359–372 (1985)
21. Carlier, G., Duval, V., Peyré, G., Schmitzer, B.: Convergence of entropic schemes for optimal transport and gradient flows. *SIAM J. Math. Anal.* **49**(2), 1385–1418 (2017)

22. Charlier, B., Feydy, J., Glaunes, J.: Kernel operations on the GPU, with autodiff, without memory overflows. <http://www.kernel-operations.io/>
23. Cordero-Erausquin, D.: Sur le transport de mesures périodiques. *C.R. Acad. Sci. Paris Sér. I Math.* **329**, 199–202 (1999)
24. Crane, K., Weischedel, C., Wardetzky, M.: Geodesics in heat: a new approach to computing distance based on heat flow. *ACM Trans. Graph.* **32**(5), 1–11 (2013)
25. Crandall, M.G., Ishii, I., Lions, P.-L.: User's guide to viscosity solutions of second order partial differential equations. *Am. Math. Soc. Bull. New Ser.* **27**(1), 1–67 (1992)
26. Cuturi, M.: Sinkhorn distances: Lightspeed computation of optimal transport. In: *Advances in Neural Information Processing Systems (NIPS)*, vol. 26, pp. 2292–2300. MIT Press, Cambridge, MA (2013)
27. Cuturi, M., Peyré, G.: Computational Optimal Transport. *Foundations and Trends in Machine Learning*, vol. 11, no. 5–6, pp. 355–607 (2019). [arXiv:1803.00567](https://arxiv.org/abs/1803.00567)
28. de Castro, P.M.M., Mérigot, Q., Thiber, B.: Far-field reflector problem and intersection of paraboloids. *Numer. Math.* **134**(2), 389–411 (2016)
29. De Philippis, G., Figalli, A.: The Monge–Ampère equation and its link to optimal transportation. *Bull. Am. Math. Soc. (N.S.)* **51**(4), 527–580 (2014)
30. Dominitz, A., Tannenbaum, A.: Texture mapping via optimal mass transport. *IEEE Trans. Vis. Comput. Graph.* **16**(3), 419–433 (2010)
31. Driscoll, J., Healy Jr., D.M.: Computing Fourier transforms and convolutions on the 2-sphere. *Adv. Appl. Math.* **15**(2), 202–250 (1994)
32. Donaldson, S.K.: Some numerical results in complex differential geometry. *Pure Appl. Math. Q.* **5** (2009), no. 2, Special Issue: In honor of Friedrich Herzbruch. Part 1, 571–618
33. Douglas, M.R., Karp, R.L., Lukic, S., Reinbacher, R.: Numerical Calabi–Yau metrics. *J. Math. Phys.* **49**(3), 032302 (2008)
34. Feng, X., Glowinski, R., Neilan, : Recent developments in numerical methods for fully nonlinear second order partial differential equations. *SIAM Rev.* **55**(2), 205–267 (2013)
35. Feydy, J.: Geometric data analysis, beyond convolutions. Thesis draft at [https://www.math.ens.fr/~feidy/geometric\\_data\\_analysis\\_draft.pdf](https://www.math.ens.fr/~feidy/geometric_data_analysis_draft.pdf)
36. Franklin, J., Lorenz, J.: On the scaling of multidimensional matrices. *Linear Algebra Appl.* **114**(115), 717–735 (1989)
37. Glimm, T., Olikar, V.: Optical design of single reflector systems and the Monge–Kantorovich mass transfer problem. *Nonlinear problems and function theory. J. Math. Sci. (N. Y.)* **117**(3), 4096–4108 (2003)
38. Hamfeldt, B.F.: Convergence framework for the second boundary value problem for the Monge–Ampère equation. *SIAM J. Numer. Anal.* **57**, 945–971 (2019)
39. Healy Jr., D.M., Rockmore, D.N., Kostelec, P.J., Moore, S.: FFTs for the 2-sphere-improvements and variations. *J. Fourier Anal. Appl.* **9**(4), 341–385 (2003)
40. Hesse, K., Sloan, I.H., Womersley, R.S.: Numerical integration on the sphere. In: W. Freeden, M.Z. Nashed, T. Sonar (eds.) *Handbook of Geomathematics*, 2nd edn., pp. 2671–2710. Springer, Heidelberg (2010)
41. Huang, H.: The Cauchy problem for fully nonlinear parabolic systems on manifolds (2015) [arXiv:1506.05030](https://arxiv.org/abs/1506.05030)
42. Kawecki, E., Lakkis, O., Pryer, T.: A finite element method for the Monge–Ampère equation with transport boundary conditions. [arXiv:1807.03535](https://arxiv.org/abs/1807.03535)
43. Keiner, J., Kunis, S., Potts, D.: Fast summation of radial functions on the sphere. *Computing* **78**(1), 1–15 (2006)
44. Kim, Y.-H., Streets, J., Warren, M.: Parabolic optimal transport equations on manifolds. *Int. Math. Res. Not. IMRN* **19**, 4325–4340 (2012)
45. Kitagawa, J.: A parabolic flow toward solutions of the optimal transportation problem on domains with boundary. *J. Reine Angew. Math.* **672**, 127–160 (2012)
46. Kitagawa, J., Mérigot, Q., Thibert, B.: Convergence of a Newton algorithm for semi-discrete optimal transport. *J. Eur. Math. Soc.* (2019). <https://doi.org/10.4171/JEMS/889>
47. Léonard, C.: From the Schrödinger problem to the Monge–Kantorovich problem. *J. Funct. Anal.* **262**(4), 1879–1920 (2012)
48. Léonard, C.: A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.* **34**(4), 1533–1574 (2014)

49. Lindsey, M., Rubinstein, Y.A.: Optimal transport via a Monge–Ampère optimization problem. *SIAM J. Math. Anal.* **49**(4), 3073–3124 (2017)
50. Lions, P.-L., Trudinger, N.S., Urbas, J.I.E.: The Neumann problem for equations of Monge–Ampère type. *Commun. Pure Appl. Math.* **39**(4), 539–563 (1986)
51. Loeper, G., Rapetti, F.: Numerical solution of the Monge–Ampère equation by a Newton’s algorithm. *C. R. Math. Acad. Sci. Paris* **340**(4), 319–324 (2005)
52. McCann, R.J.: Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.* **11**(3), 589–608 (2001)
53. Ma, X.-N., Trudinger, N.S., Wang, X.-J.: Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.* **177**(2), 151–183 (2005)
54. Norris, J.R.: Heat kernel asymptotics and the distance function in Lipschitz Riemannian manifolds. *Acta Math.* **179**(1), 79–103 (1997)
55. Ruschendorf, L.: Convergence of the iterative proportional fitting procedure. *Ann. Stat.* **23**(4), 1160–1174 (1995)
56. Rokhlin, V., Tygert, M.: Fast algorithms for spherical harmonic expansions. *SIAM J. Sci. Comput.* **27**(6), 1903–1928 (2006)
57. Schnurer, Oliver C., Smoczyk, K.: Neumann and second boundary value problems for Hessian and Gauss curvature flows. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **20**(6), 1043–1073 (2003)
58. Saumier, L.-P., Agueh, M., Khouider, B.: An efficient numerical algorithm for the l2 optimal transport problem with periodic densities. *IMA J. Appl. Math.* **80**(1), 135–157 (2015)
59. Schmitzer, B.: Stabilized sparse scaling algorithms for entropy regularized transport problems. *SIAM J. Sci. Comput.* **41**(3), 1443–A1481 (2019)
60. Schrödinger, E.: Sur la théorie relativiste de l’électron et l’interprétation de la mécanique quantique. *Ann. Inst. H. Poincaré* **2**, 269–310 (1932)
61. Sinkhorn, R.: Diagonal equivalence to matrices with prescribed row and column sums. *Am. Math. Mon.* **74**, 402–405 (1967)
62. Sogge, C.D.: *Fourier Integrals in Classical Analysis*. Cambridge Tracts in Mathematics, vol. 105. Cambridge University Press, Cambridge (1993)
63. Solomon, J., De Goes, F., Peyré, G., Cuturi, M., Butscher, A., Nguyen, A., Du, T., Guibas, L.: Convolutional Wasserstein distances. *ACM Trans. Graph.* **34**(4), 66 (2015)
64. Solomon, J.: Optimal Transport on Discrete Domains: Proceedings of Symposia in Pure Mathematics (to appear); [http://people.csail.mit.edu/jsolomon/assets/optimal\\_transport.pdf](http://people.csail.mit.edu/jsolomon/assets/optimal_transport.pdf)
65. Trudinger, N.S., Wang, X.-J.: On the second boundary value problem for Monge–Ampère type equations and optimal transportation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **8**(1), 143–174 (2009)
66. Vialard, François-Xavier: An elementary introduction to entropic regularization and proximal methods for numerical optimal transport. Doctoral. France. (2019). hal-02303456
67. Villani, C.: *Topics in optimal transportation*. Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, xvi+370 pp (2003)
68. Villani, C.: *Optimal transport. Old and new*. Grundlehren der Mathematischen Wissenschaften 338. Springer, Berlin (2009)
69. Wang, X.-J.: On the design of a reflector antenna II. *Calc. Var. Part. Differ. Equ.* **20**(3), 329–341 (2004)
70. Wellera, H., Brownea, P., Buddb, C., Cullen, M.: Mesh adaptation on the sphere using optimal transport and the numerical solution of a Monge–Ampère type equation. *J. Comput. Phys.* **308**, 102–123 (2016)
71. Womersley, R.S.: Efficient spherical designs with good geometric properties. *Festschrift for the 80th Birthday of Ian H. Sloan*, Editors J. Dick et al (2018). See files at <http://web.maths.unsw.edu.au/~rsw/Sphere/EffSphDes/>
72. Zhao, X., Su, Z., Gu, X.D., Kaufman, A., Sun, J., Gao, J., Luo, F.: Area-preservation mapping using optimal mass transport. *IEEE Trans Vis Comput Gr.* **19**(12), 2838–2847 (2013)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.