European Journal of Applied Mathematics (2026), **37**, pp. 925–967
doi:10.1017/S0956792525100144

CAMBRIDGE UNIVERSITY PRESS

PAPER

Semi-discrete unbalanced optimal transport and quantization

David P. Bourne[1], Bernhard Schmitzer[2] and Benedikt Wirth[3]

[1]Maxwell Institute for Mathematical Sciences and Department of Mathematics, Heriot-Watt University, Edinburgh, UK
[2]University of Göttingen, Göttingen, Germany
[3]University of Münster, Münster, Germany
Corresponding author: David P. Bourne; Email: d.bourne@hw.ac.uk

Received: 09 July 2024; **Revised:** 17 August 2025; **Accepted:** 18 August 2025; **First published online:** 17 October 2025*

Keywords: quantization; semi-discrete optimal transport; unbalanced optimal transport; crystallization; Laguerre diagrams

2020 Mathematics Subject Classification: 49Q22, 49K99, 49J99

Abstract

In this paper we study the class of optimal entropy-transport problems introduced by Liero, Mielke and Savaré in Inventiones Mathematicae 211 in 2018. This class of unbalanced transport metrics allows for transport between measures of different total mass, unlike classical optimal transport where both measures must have the same total mass. In particular, we develop the theory for the important subclass of semi-discrete unbalanced transport problems, where one of the measures is diffuse (absolutely continuous with respect to the Lebesgue measure) and the other is discrete (a sum of Dirac masses). We characterize the optimal solutions and show they can be written in terms of generalized Laguerre diagrams. We use this to develop an efficient method for solving the semi-discrete unbalanced transport problem numerically. As an application, we study the unbalanced quantization problem, where one looks for the best approximation of a diffuse measure by a discrete measure with respect to an unbalanced transport metric. We prove a type of crystallization result in two dimensions – optimality of a locally triangular lattice with spatially varying density – and compute the asymptotic quantization error as the number of Dirac masses tends to infinity.

1. Introduction

In this paper we study *semi-discrete unbalanced* optimal transport problems: What is the optimal way of transporting a diffuse measure to a discrete measure (hence the name *semi-discrete*), where the two measures may have different total mass (hence the name *unbalanced*)? As an application, we study the *unbalanced quantization problem*: What is the best approximation of a diffuse measure by a discrete measure with respect to an unbalanced transport metric?

1.1. Unbalanced optimal transport

Classical optimal transport theory asks for the most efficient way to rearrange mass between two given probability distributions. Its origin goes back to 1781 and the French engineer Gaspard Monge, who was interested in the question of how to transport and reshape a pile of earth to form an embankment at minimal effort. It took over 200 years to develop a complete mathematical understanding of this problem, even to answer the question of whether there exists an optimal way of redistributing mass. Since the mathematical breakthroughs of the 1980s and 1990s, the field of optimal transport theory has thrived and found applications in crowd and traffic dynamics, economics, geometry, image and signal

*Article updated since first published online. HTML & PDF versions have been updated.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

processing, machine learning and data science, PDEs and statistics. Depending on the context, mass may represent the distribution of particles (people or cars), supply and demand, population densities, etc. For thorough introductions see, e.g., [27, 58, 61, 69].

In classical optimal transport theory the initial and target measures must have the same total mass. In applications this is not always natural. Changes in mass may occur due to creation or annihilation of particles or a mismatch between supply and demand. Therefore so-called *unbalanced* transport problems, accounting for such differences, have recently received increased attention [18, 25, 39, 45]. Brief overviews and discussions of various formulations can be found, for instance, in [19, 64]. Further theoretical properties are examined in [41, 46], examples for applications in data analysis can be found in [17, 42, 66]. In this article, we study the class of unbalanced transport problems called *optimal entropy-transport problems* from [45]; see Definition 2.4. In particular, we develop this theory for the special case of semi-discrete transport.

1.2. Semi-discrete transport

Semi-discrete optimal transport theory is about the best way to transport a diffuse measure, $\mu \in L^1(\Omega)$, $\Omega \subset \mathbb{R}^d$, to a discrete measure, $\nu = \sum_{i=1}^M m_i \delta_{x_i}$. These type of problems arise naturally, for instance, in economics in computing the distance between a population with density μ and a resource with distribution $\nu = \sum_{i=1}^M m_i \delta_{x_i}$, where $x_i \in \Omega$ represent the locations of the resource and $m_i > 0$ represent the size or capacity of the resource. The classical semi-discrete optimal transport problem, where μ and ν are probability measures, has a nice geometric characterization. For example, for $p \in [1, \infty)$, the Wasserstein-p metric W_p is defined by

$$W_p(\mu, \nu) = \min \left\{ \sum_{i=1}^M \int_{T^{-1}(x_i)} |x - x_i|^p \mu(x)\, \mathrm{d}x \;\middle|\; T : \Omega \to \{x_i\}_{i=1}^M,\ \int_{T^{-1}(x_i)} \mu(x)\, \mathrm{d}x = m_i \right\}^{1/p}$$

where $\sum_{i=1}^M m_i = \int_\Omega \mu(x)\, \mathrm{d}x = 1$. This is an optimal partitioning (or assignment) problem, where the domain Ω is partitioned into the regions $T^{-1}(x_i)$ of mass m_i, $i \in \{1, \dots, M\}$, and each point $x \in T^{-1}(x_i)$ is assigned to point x_i. For example, in two dimensions, Ω could represent a city, μ the population density of children, x_i and m_i the location and size of schools, $T^{-1}(x_i)$ the catchment areas of the schools, and $W_p(\mu, \nu)$ the cost of transporting the children to their assigned schools. If $p = 2$, it turns out that the optimal partition $\{T^{-1}(x_i)\}_{i=1}^M$ is a *Laguerre diagram* or *power diagram*, which is a type of weighted Voronoi diagram: There exist weights $w_1, \dots, w_M \in \mathbb{R}$ such that

$$\overline{T^{-1}(x_i)} = \{x \in \Omega \mid |x - x_i|^2 - w_i \le |x - x_j|^2 - w_j \ \forall\, j \in \{1, \dots, M\}\}.$$

The transport cells $T^{-1}(x_i)$ are the intersection of convex polytopes (polygons if $d = 2$, polyhedra if $d = 3$) with Ω. The weights $w_1, \dots, w_M \in \mathbb{R}$ can be found by solving an unconstrained concave maximization problem. If $p = 1$, the optimal partition $\{T^{-1}(x_i)\}_{i=1}^M$ in an Apollonius diagram. See, e.g., [3, Sec. 6.4], [27, Chap. 5], [37, 52], [58, Chap. 5], and Section 2.3 below, where we summarize the main results from classical semi-discrete optimal transport theory. Applications of semi-discrete transport include fluid mechanics [28, 29], microstructure modelling [10, 15], optics [53] and the Lagrangian discretization of Wasserstein gradient flows [43] and mean field games [63].

In Section 3, we extend these results to unbalanced transport, where μ and ν no longer need to have the same total mass, and the Wasserstein-p metric is replaced by the unbalanced transport metric W from Definition 2.4. We prove that, also in the unbalanced case, the optimal partition is a type of generalized Laguerre diagram, and it can be found by solving a concave maximization problem for a set of weights $w_1, \dots, w_M$; see Theorems 3.1 and 3.3. This problem is natural from a modelling perspective, for example to describe a mismatch between the demand of a population μ and the supply of a resource ν and to model the prioritization of high-density regions at the expense of areas with a low population density.

For unbalanced transport, there is no one, definitive transport cost, but many models are conceivable. As a first application of our theory of semi-discrete unbalanced transport, in Examples 3.14 and 3.15,

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

we use it to compare different unbalanced transport models. As a second application, in Section 4, we apply it to the quantization problem.

1.3. Quantization

Quantization of measures refers to the problem of finding the best approximation of a diffuse measure by a discrete measure [32], [35, Sec. 33]. For example, the classical quantization problem with respect to the Wasserstein-p metric, $p \in [1, \infty)$, is the following: Given $\mu \in L^1(\Omega)$, $\Omega \subset \mathbb{R}^d$, $\int_\Omega \mu(x)\,\mathrm{d}x = 1$, find a discrete probability measure $\nu = \sum_{i=1}^M m_i \delta_{x_i}$ that gives the best approximation of μ in the Wasserstein-p metric,

$$Q_p^M(\mu) = \min \left\{ W_p^p(\mu, \nu) \;\middle|\; \nu = \sum_{i=1}^M m_i \delta_{x_i},\ x_1, \ldots, x_M \in \Omega,\ m_i > 0,\ \sum_{i=1}^M m_i = 1 \right\}. \tag{1.1}$$

We call Q_p^M the *quantization error*. Problems of this form arise in a wide range of applications including economic planning and optimal location problems [7, 8, 14], finance [57], numerical integration [21, Sec. 2.2], [57, Sec. 2.3], energy-driven pattern formation [11, 40] and approximation of initial data for particle (meshfree) methods for PDEs. An approach to quantization using gradient flows is given in [16, 36]. We mention a few important variations on the classical quantization problem. The case where the masses $m_1, \ldots, m_M$ are fixed and the minimization in (1.1) is only taken over $x_1, \ldots, x_M$ is considered for example in [10, 51, 70]. The case where μ is a discrete measure, with support of cardinality $N \gg M$, has applications in image and signal compression [22, 31] and data clustering (k-means clustering) [49, 65]. If ν is a one-dimensional measure (supported on a set of Hausdorff dimension 1), then the quantization problem is known as the *irrigation problem* [48, 55]. In this paper, we consider the variation where the Wasserstein-p metric in (1.1) is replaced by an unbalanced transport metric.

It can be shown that the quantization problem (1.1) can be rewritten as an optimization problem in terms of the particle locations $\{x_i\}_{i=1}^M$ and their Voronoi tessellation:

$$Q_p^M(\mu) = \min \{ J(x_1, \ldots, x_M) \mid x_1, \ldots, x_M \in \Omega \} \tag{1.2}$$

where

$$J(x_1, \ldots, x_M) = \sum_{i=1}^M \int_{V_i(x_1,\ldots,x_M)} |x - x_i|^p \mu(x)\,\mathrm{d}x$$

and where $\{V_i\}_{i=1}^M$ is the Voronoi diagram generated by $\{x_i\}_{i=1}^M$,

$$V_i = V_i(x_1, \ldots, x_M) = \left\{ x \in \Omega \;\middle|\; |x - x_i| \le |x - x_j| \text{ for all } j \in \{1, \ldots, M\} \right\}.$$

If $(x_1, \ldots, x_M)$ is a global minimizer of J, then $\sum_{i=1}^M \left(\int_{V_i} \mu\,\mathrm{d}x \right) \delta_{x_i}$ is an optimal quantizer of μ with respect to the Wasserstein-p metric. See for instance [11, Sec. 4.1], [38, Sec. 7] and Theorem 4.2. In the vector quantization (electrical engineering), literature J is known as the *distortion* of the quantizer [31].

The quantization problem with respect to the Wasserstein-2 metric is particularly well studied. In this case, it can be shown that critical points of J are generators of *centroidal Voronoi tessellations* (CVTs) of M points [21]; this means that $\nabla J(x_1, \ldots, x_M) = 0$ if and only if x_i is the centre of mass of its own Voronoi cell V_i for all i,

$$x_i = \frac{\displaystyle\int_{V_i(x_1,\ldots,x_M)} x\mu(x)\,\mathrm{d}x}{\displaystyle\int_{V_i(x_1,\ldots,x_M)} \mu(x)\,\mathrm{d}x}, \quad i \in \{1, \ldots, M\}. \tag{1.3}$$

In general, there does not exist a unique CVT of M points, as illustrated in Figure 1, and J is non-convex with many local minimizers for large M. Equation (1.3) is a nonlinear system of equations for $x_1, \ldots, x_M$. A simple and popular method for computing CVTs is Lloyd's algorithm [21, 24, 47, 60], which is a fixed point method for solving the Euler–Lagrange equations (1.3).

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

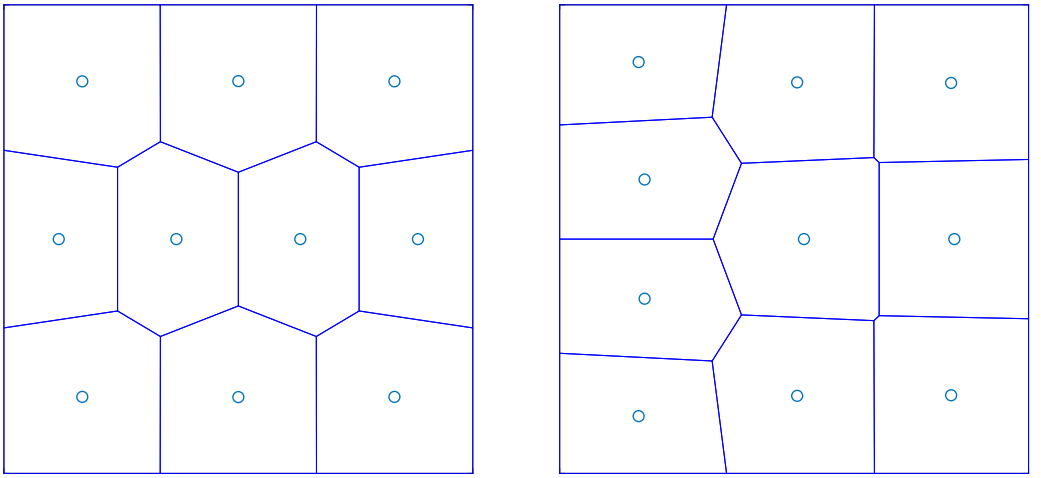


Figure 1. *Two (approximate) centroidal Voronoi tessellations (CVTs) of 10 points for the uniform density $\mu = 1$ on a unit square. The polygons are the centroidal Voronoi cells V_i and the circles are the generators x_i. The CVTs were computed using Lloyd's algorithm. The CVT on the left has a lower energy J than the CVT on the right. The corresponding quantizer $\nu = \sum_{i=1}^{10} m_i \delta_{x_i}$ of μ is reconstructed from the CVT by taking m_i as the areas of the centroidal Voronoi cells and x_i as their generators.*

In Sections 4.1 and 4.2 we extend these results to unbalanced quantization, where the Wasserstein-p metric in (1.1) is replaced by the unbalanced transport metric W (defined in equation (2.6)), and where μ and ν need not have the same total mass. In Theorem 4.2 we prove an expression of the form (1.2), which states that the unbalanced quantization problem can be reduced to an optimization problem for the locations $x_1, \ldots, x_M$ of the Dirac masses. This optimization problem is again formulated in terms of the Voronoi diagram generated by $x_1, \ldots, x_M$. In Section 4.2 we solve the unbalanced quantization problem numerically, which includes extending Lloyd's algorithm to the unbalanced case.

We conclude the paper in subsection 4.3 by studying the asymptotic unbalanced quantization problem: What is the optimal configuration of the particles $x_1, \ldots, x_M$ as $M \to \infty$ and how does the quantization error scale in M? Consider for example the classical quantization problem (1.1) with $p = 2$, $|\Omega| = 1$, $\mu = 1$ (i.e., μ is the Lebesgue measure on Ω), and M fixed. From above, we know that an optimal quantizer ν corresponds to an optimal CVT of M points, where *optimal* means that the CVT has lowest energy J amongst all CVTs of M points. Gersho [30] conjectured that, as $M \to \infty$, the Voronoi cells of the optimal CVT asymptotically have the same shape, i.e., asymptotically they are translations and rescalings of a single polytope. In two dimensions ($d = 2$), various versions of Gersho's Conjecture have been proved independently by several authors [7, 33, 54, 56, 67, 68]. Roughly speaking, it has been shown that the hexagonal tiling is optimal as $M \to \infty$. In other words, arranging the seeds $x_1, \ldots, x_M$ in a regular triangular lattice is asymptotically optimal. This crystallization result can be stated more precisely as follows: If Ω is a convex polygon with at most 6 sides, then

$$J(x_1, \ldots, x_M) \geq \frac{5\sqrt{3}}{54} \frac{1}{M} \tag{1.4}$$

where the right-hand side is the energy of a regular triangular lattice of M points such that the Voronoi cells V_i are regular hexagons of area $1/M$. In general, this lower bound is not attained for finite M (unless Ω is a regular hexagon and $M = 1$), but it is attained in limit $M \to \infty$:

$$\lim_{M \to \infty} M \cdot Q_2^M(1) = \lim_{M \to \infty} M \cdot \min_{x_i \in \Omega} J(x_1, \ldots, x_M) = \frac{5\sqrt{3}}{54}. \tag{1.5}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

See the references above or [12, Thm. 5]. We generalize (1.4) and (1.5) to the unbalanced quantization problem in Theorem 4.7 and Theorem 4.12, respectively. Roughly speaking, we show that again for $\mu = 1$ the triangular lattice is optimal in the limit $M \to \infty$. For general $\mu \in L^1(\Omega)$, it is asymptotically optimal for the particles to locally form a triangular lattice with density determined by a nonlocal function of μ.

While our quantization results are limited to two dimensions, this is also largely true for the classical quantization problem. In three dimensions, it is not known whether Gersho's conjecture holds, although there is some numerical evidence for the case $p = 2$ that optimal CVTs of M points tend as $M \to \infty$ to the Voronoi diagram of the body-centered cubic (BCC) lattice, where each Voronoi cell is congruent to a truncated octahedron [23]. See also [20]. For $p = 2$, it has been proved that, *amongst lattices*, the BCC lattice is optimal [5].

For general p, d and μ, the scaling of the quantization error is known even if the optimal quantizer is not; Zador's theorem [71], [35, Cor. 33.3] states that

$$\lim_{M\to\infty} M^{\frac{p}{d}} \cdot Q_p^M(\mu) = c(p,d)\, \|\mu\|_{L^{\frac{d}{d+p}}(\Omega)} \tag{1.6}$$

where the constant $c(p, d)$ is characterized by

$$c(p,d) = \lim_{M\to\infty} M^{\frac{p}{d}} \cdot Q_p^M(\mathcal{L}\llcorner[0,1]^d),$$

and where $\mathcal{L}$ is the d-dimensional Lebesgue measure. For a modern proof using Γ-convergence see [9] and [62, Proposition 7.21]. For generalizations to quantization on Riemannian manifolds see [34], [38, Thm. 1.2] and [4]. It is an open problem to compute the optimal constant $c(p, d)$ except for $d = 1$ and $d = 2$, where

$$c(p,1) = \int_{-1/2}^{1/2} |x|^p \,\mathrm{d}x, \qquad c(p,2) = \int_{H(1)} |x|^p \,\mathrm{d}x, \tag{1.7}$$

where $H(1)$ is a regular hexagon of area 1 centred at the origin 0. We recover Zador's theorem for the case $d = 2$, along with the optimal constant $c(p, 2)$, as a special case of Theorem 4.12; see Example 4.18.

1.4. Outline and contribution

Section 2 collects relevant results from classical, unbalanced and semi-discrete transport, which will be generalized in Section 3 to the case of semi-discrete unbalanced transport. Finally, Section 4 considers the unbalanced quantization problem.

In more detail, the contributions of this article are the following.

- Section 3.1: We extend semi-discrete transport theory to the unbalanced case, most importantly a simple, geometric tessellation formulation (Theorem 3.1), optimality conditions that fully characterize primal and dual solutions (Theorem 3.3) and additional different primal and dual convex formulations. Unlike in the balanced case, the dual potentials associated with the discrete mass locations do not only determine the tessellation of the continuous measure but also the density of the optimal transport plan. Particular attention needs to be paid to areas where the ground transport cost function becomes infinite. Special cases of these results were derived in [43, 63] to study a Lagrangian discretization of Wasserstein gradient flows and variational mean field games.
- Section 3.2: We develop numerical algorithms for solving the semi-discrete unbalanced transport problem and numerically illustrate novel phenomena of unbalanced transport (Example 3.14). In particular, we show qualitative differences between different unbalanced transport models and examine the effect of changing the length scale, which typically is intrinsic to unbalanced transport models.
- Sections 4.1 and 4.2: We extend the theory of optimal transport-based quantization of measures to unbalanced transport, deriving in particular an equivalent Voronoi tessellation problem (Theorem 4.2), which turns out to be a natural generalization of the known corresponding formulation in classical transport. The interesting fact here is that the simple geometric Voronoi

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

tessellation structure survives when passing from balanced to unbalanced transport, but the mass of the generating points now depends in a more complex way on the mass within their cells. We also illustrate unbalanced quantization numerically, extending the standard algorithms (including Lloyd's algorithm) to the unbalanced case.

- Section 4.3: In two spatial dimensions, where crystallization results from discrete geometry are available, we derive the optimal asymptotic quantization cost and the optimal asymptotic point density for quantizing a given measure μ using unbalanced transport (Theorem 4.12). Our result includes Zador's theorem for classical, balanced quantization as a special case; see Example 4.18. As is common in asymptotic quantization, we consider a spatial rescaling of the domain as the number of points increases and the most interesting regime is where the rescaled point density converges to a non-zero, finite limit. While in the balanced case the rescaled asymptotic cost only depends on the growth behaviour of the transport ground cost function, in the unbalanced setting we now observe an interplay between the rescaled point density and the intrinsic length scale of unbalanced transport. An interesting, novel effect in this unbalanced setting is that the optimal point density depends nonlocally on the global mass distribution in such a way that whole regions with positive measure may be completely neglected in favour of regions with higher mass.

1.5. Setting and notation

Throughout this article we work in a domain $\Omega = \overline{U}$ for $U \subset \mathbb{R}^d$ open. (In principle, the results could be extended to more general metric spaces such as Riemannian manifolds.) The Euclidean distance on $\mathbb{R}^d$ is denoted $d(\cdot, \cdot)$, and we will write $\pi_i : \Omega \times \Omega \to \Omega$ for the projections $\pi_i(x_1, x_2) = x_i$, $i = 1, 2$. The (d-dimensional) Lebesgue measure of a measurable set $A \subset \mathbb{R}^d$ will be indicated by $\mathcal{L}(A)$ or $|A|$ for short, its diameter by $\operatorname{diam}(A)$.

By $\mathcal{M}_+(\Omega)$ we denote the set of nonnegative Radon measures on Ω, and $\mathcal{P}(\Omega) \subset \mathcal{M}_+(\Omega)$ is the subset of probability measures. The notation $\mu \ll \nu$ for two measures $\mu, \nu \in \mathcal{M}_+(\Omega)$ indicates absolute continuity of μ with respect to ν, and the corresponding Radon–Nikodym derivative is written as $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$. The restriction of $\mu \in \mathcal{M}_+(\Omega)$ to a measurable set $A \subset \mathbb{R}^d$ is denoted $\mu \llcorner A$, and its support is denoted $\operatorname{spt} \mu$. For a Dirac measure at a point $x \in \mathbb{R}^d$ we write δ_x. The pushforward of a measure μ under a measurable map T is denoted $T_\# \mu$.

The spaces of Lebesgue integrable functions on U or of μ-integrable functions with $\mu \in \mathcal{M}_+(\Omega)$ are denoted $L^1(U)$ and $L^1(\mu)$, respectively. Continuous functions on Ω are denoted by $\mathcal{C}(\Omega)$.

2. Background

The purpose of this section is a short introduction to classical, unbalanced and semi-discrete transport.

2.1. Optimal transport

Here we briefly recall the basic setting of optimal transport. For a thorough introduction we refer, for instance, to [61, 69]. For $\mu, \nu \in \mathcal{P}(\Omega)$ the set

$$\Gamma(\mu, \nu) = \{\gamma \in \mathcal{P}(\Omega \times \Omega) \mid \pi_{1\#}\gamma = \mu,\ \pi_{2\#}\gamma = \nu\} \tag{2.1}$$

is called the *couplings* or *transport plans* between μ and ν. A measure $\gamma \in \Gamma(\mu, \nu)$ can be interpreted as a rearrangement of the mass of μ into ν where $\gamma(x, y)$ intuitively describes how much mass is taken from x to y. The total cost associated to a coupling γ is given by

$$\int_{\Omega \times \Omega} c(x, y)\, \mathrm{d}\gamma(x, y) \tag{2.2}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

where $c : \Omega \times \Omega \to [0, \infty]$ and $c(x, y)$ specifies the cost of moving one unit of mass from x to y. The *optimal transport problem* asks for finding a γ that minimizes (2.2) among all couplings $\Gamma(\mu, \nu)$,

$$W_{\mathrm{OT}}(\mu, \nu) = \inf \left\{ \int_{\Omega \times \Omega} c \, \mathrm{d}\gamma \;\middle|\; \gamma \in \Gamma(\mu, \nu) \right\} . \tag{2.3}$$

Under suitable regularity assumptions on c, existence of minimizers follows from standard compactness and lower semi-continuity arguments.

Theorem 2.1 ([69, Thm. 4.1]). *If $c : \Omega \times \Omega \to [0, \infty]$ is lower semi-continuous, then minimizers of (2.3) exist. The minimal value may be $+\infty$.*

2.2. Unbalanced transport

The optimal transport problem (2.3) only allows the comparison of measures μ, ν with equal mass. Otherwise, the feasible set $\Gamma(\mu, \nu)$ is empty. Therefore, so-called unbalanced transport problems have been studied, where mass may be created or annihilated during transport and thus measures of different total mass can be compared in a meaningful way. See Section 1 for context and references.

Throughout this article we consider unbalanced optimal entropy-transport problems as studied in [45]. The basic idea is to replace the hard marginal constraints $\pi_{1\#}\gamma = \mu$, $\pi_{2\#}\gamma = \nu$ in (2.1) with soft constraints where the deviation between the marginals of γ and the measures μ and ν is penalized by a marginal discrepancy function. This allows more flexibility for feasible γ. We focus on a subset of the family of marginal discrepancies considered in [45].

Definition 2.2 (Marginal discrepancy). *Let $F : [0, \infty) \to [0, \infty]$ be proper, convex and lower semi-continuous with $\lim_{s\to\infty} \frac{F(s)}{s} = \infty$. For a given measure $\mu \in \mathcal{M}_+(\Omega)$, the function F induces a marginal discrepancy $\mathcal{F}(\,\cdot\,|\mu) : \mathcal{M}_+(\Omega) \to [0, \infty]$ via*

$$\mathcal{F}(\rho|\mu) = \begin{cases} \displaystyle\int_\Omega F\left(\frac{\mathrm{d}\rho}{\mathrm{d}\mu}\right) \mathrm{d}\mu & \text{if } \rho \ll \mu, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.4}$$

Note that the integrand is only defined μ-almost everywhere. $\mathcal{F}$ is (sequentially) weakly-$*$ lower semi-continuous [1, Thm. 2.34].

We extend the domain of definition of F to $\mathbb{R}$ by setting $F(s) = \infty$ for $s < 0$. The Fenchel–Legendre conjugate of F is then the convex function $F^* : \mathbb{R} \to (-\infty, +\infty]$ defined by

$$F^*(z) = \sup_{s\in\mathbb{R}} (z \cdot s - F(s)) = \sup_{s\geq 0} (z \cdot s - F(s)) .$$

Example 2.3 (Kullback–Leibler divergence). *The Kullback–Leibler divergence is an example of Definition 2.2 for the choice $F_{\mathrm{KL}} : [0, \infty) \to [0, \infty)$,*

$$F_{\mathrm{KL}}(s) = \begin{cases} s \log s - s + 1 & \text{if } s > 0, \\ 1 & \text{if } s = 0. \end{cases}$$

The Fenchel–Legendre conjugate is given by $F^*_{\mathrm{KL}}(z) = e^z - 1$.

Definition 2.4 (Unbalanced optimal transport problem). *Let F be as in Definition 2.2 and let $\mathcal{F}$ be the induced marginal discrepancy. Let μ, $\nu \in \mathcal{M}_+(\Omega)$ and $c : \Omega \times \Omega \to [0, \infty]$ be lower semi-continuous. The corresponding unbalanced transport cost $\mathcal{E} : \mathcal{M}_+(\Omega \times \Omega) \to [0, \infty]$ is given by*

$$\mathcal{E}(\gamma) = \int_{\Omega\times\Omega} c \, \mathrm{d}\gamma + \mathcal{F}(\pi_{1\#}\gamma|\mu) + \mathcal{F}(\pi_{2\#}\gamma|\nu) \tag{2.5}$$

and induces the optimization problem

$$W(\mu, \nu) = \inf \left\{ \mathcal{E}(\gamma) \;\middle|\; \gamma \in \mathcal{M}_+(\Omega \times \Omega) \right\} . \tag{2.6}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

The interaction between the terms in (2.5) that penalize transport and mass change introduces an intrinsic length scale for transport that you do not see in classical balanced transport. This is discussed in Example 3.15.

Theorem 2.5 ([45, Thm. 3.3]). *Minimizers of (2.6) exist. The minimal value may be $+\infty$.*

Remark 2.6. *Observe that $\mathcal{F}(\rho|\mu) = \infty$ whenever $\rho \not\ll \mu$ and $\mathcal{F}(\rho|\nu) = \infty$ whenever $\rho \not\ll \nu$. This guarantees that $\pi_{1\#}\gamma \ll \mu$ and $\pi_{2\#}\gamma \ll \nu$ for all feasible γ, where feasible means that $\mathcal{E}(\gamma) < \infty$. Thus, when $\mu \ll \mathcal{L}$ and ν is discrete, as in the semi-discrete setting (which will be discussed in the following section), then the first and second marginal of any feasible γ will share these properties.*

Remark 2.7. *For simplicity we assume that the same marginal discrepancy is applied to both marginals in (2.5), but of course in some cases it may be more appropriate to consider two different discrepancies. All results in this article generalize to this case in a canonical way.*

In this article we focus on cost functions c that can be written as increasing functions of the distance between x and y.

Definition 2.8 (Radial cost). *A cost function $c : \Omega \times \Omega \to [0, \infty]$ is called radial if it can be written as $c(x, y) = \ell(d(x, y))$ for a strictly increasing function $\ell : [0, \infty) \to [0, \infty]$, continuous on its domain with $\ell(0) = 0$.*

Note that the cost c need not be *twisted* (twistedness means that $y \mapsto \nabla_x c(x, y)$ is injective for all x, see [61, Definition 1.16]), which leads to some technical complications. The following examples shall be used throughout for illustration. They all feature a *radial* transport cost c in the sense of Definition 2.8.

Example 2.9 (Unbalanced transport models).

(a) ***Standard Wasserstein-2 distance (W2).*** *Classical balanced optimal transport can be recovered as a special case of Definition 2.4 by choosing $\mathcal{F}(\rho|\mu) = 0$ if $\rho = \mu$ and ∞ otherwise. This corresponds to*

$$F(s) = \iota_{\{1\}}(s) = \begin{cases} 0 & \text{if } s = 1, \\ \infty & \text{otherwise,} \end{cases} \qquad F^*(z) = z\,.$$

Then $\mathcal{E}(\gamma) < \infty$ only if $\gamma \in \Gamma(\mu, \nu)$, and therefore (2.6) reduces to (2.3). In particular, the Wasserstein-2 setting is obtained for $c(x, y) = d(x, y)^2$, and the Wasserstein-2 distance is defined by $W_2(\mu, \nu) = \sqrt{W(\mu, \nu)}$.

(b) ***Gaussian Hellinger–Kantorovich distance (GHK).*** *This distance is introduced in [45, Thm. 7.25] using*

$$F(s) = F_{\mathrm{KL}}(s) = \begin{cases} s \log s - s + 1 & \text{if } s > 0, \\ 1 & \text{if } s = 0, \end{cases} \qquad F^*(z) = e^z - 1\,, \qquad c(x, y) = d(x, y)^2\,.$$

(c) ***Hellinger–Kantorovich distance (HK).*** *This important instance of unbalanced transport was introduced in different formulations in [18, 39, 45] whose mutual relations are described in [19]. In Definition 2.4 one chooses*

$$F(s) = F_{\mathrm{KL}}(s)\,, \quad F^*(z) = e^z - 1\,,$$

$$c(x, y) = c_{\mathrm{HK}}(x, y) = \begin{cases} -2 \log\left[\cos\left(d(x, y)\right)\right] & \text{if } d(x, y) < \frac{\pi}{2}, \\ \infty & \text{otherwise,} \end{cases}$$

and the HK distance is defined by $\mathrm{HK}(\mu, \nu) = \sqrt{W(\mu, \nu)}$. The distance HK is actually a geodesic distance on the space of non-negative measures over a metric base space. From $c_{\mathrm{HK}}(x, y) = \infty$ for $d(x, y) \geq \frac{\pi}{2}$, we learn that mass is never transported further than $\frac{\pi}{2}$ in this setting.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

(d) ***Quadratic regularization (QR).*** *The Kullback–Leibler discrepancy F_{KL} used in both previous examples has an infinite slope at 0, which in Definition 2.4 leads to a strong incentive to achieve $\pi_{1\#}\gamma \gg \mu$ and $\pi_{2\#}\gamma \gg \nu$. The following mere quadratic discrepancy does not have this property,*

$$F(s) = (s-1)^2\,, \qquad F^*(z) = \begin{cases} \frac{z^2}{4} + z & \text{if } z \geq -2, \\ -1 & \text{otherwise,} \end{cases} \qquad c(x,y) = d(x,y)^2\,.$$

Unsurprisingly, the structure of the function F has a great influence on the behaviour of the unbalanced optimization problem (2.6). Often it is helpful to analyse corresponding dual problems where the conjugate function F^* appears. We gather some properties of F^*, implied by the assumptions on F in Definition 2.2 and on some additional assumptions that we will occasionally make in this article.

Lemma 2.10 (Properties of F^*). *Let F satisfy the assumptions given in Definition 2.2. Then*

(i) $F^*(z) > -\infty$ *for* $z \in \mathbb{R}$;
(ii) F^* *is increasing;*
(iii) $F^*(z) \leq 0$ *for* $z \leq 0$;
(iv) $F^*(z) < \infty$ *for* $z \in (0, \infty)$;
(v) F^* *is real-valued and continuous on* $\mathbb{R}$;
(vi) if F is strictly convex on its domain, then F^* is continuously differentiable on $\mathbb{R}$;
(vii) $F^*(z) \geq -F(0)$ *for all* $z \in \mathbb{R}$;
(viii) if $\inf\{x \geq 0 | F(x) < \infty\} = 0$ *(which holds in particular when $F(0) < \infty$), then*

$$\lim_{z\to-\infty} \min \partial F^*(z) = \lim_{z\to-\infty} \max \partial F^*(z) = 0.$$

Proof. **(i)** Since F is proper, we can find $s \in (0, \infty)$ with $F(s) < \infty$. Then for all $z \in \mathbb{R}$, $F^*(z) = \sup_{x\geq 0}(z \cdot x - F(x)) \geq z \cdot s - F(s) > -\infty$.

(ii) Let $z_1 \leq z_2$. Then $F^*(z_2) = \sup_{x\geq 0}(z_2 \cdot x - F(x)) \geq \sup_{x\geq 0}(z_1 \cdot x - F(x)) = F^*(z_1)$.

(iii) Let $z \leq 0$. Since $F \geq 0$, then $F^*(z) = \sup_{x\geq 0}(z \cdot x - F(x)) \leq \sup_{x\geq 0} z \cdot x = 0$.

(iv) Let $z \in (0, \infty)$. Since $F \geq 0$, $F^*(z) = \infty$ is only possible if any maximizing sequence $x_1, x_2, \ldots$ for $F^*(z) = \sup_{x\geq 0}(z \cdot x - F(x))$ diverges to ∞. However, $\lim_{n\to\infty}(z \cdot x_n - F(x_n)) = \lim_{x\to\infty} x\big(z - \frac{F(x)}{x}\big) = -\infty$ since $\lim_{s\to\infty} \frac{F(s)}{s} = \infty$. So $F^*(z) < \infty$.

(v) (i), (iv) and (iii) imply $\mathrm{dom}(F^*) = \mathbb{R}$. By convexity, F^* is therefore continuous.

(vi) This is a special case of a classical result in convex analysis, which can be found, for instance, in [59, Thm. 26.3].

(vii) Let $z \in \mathbb{R}$. Then $F^*(z) = \sup_{x\geq 0}(z \cdot x - F(x)) \geq -F(0)$.

(viii) Let $z_1, z_2, \ldots$ and $u_1, u_2, \ldots$ be sequences with $z_n \to -\infty$ as $n \to \infty$ and $u_n \in \partial F^*(z_n)$. By monotonicity of F^*, (ii), we have $u_n \geq 0$. By (iii) and convexity one finds for any $a \geq 0$ with $F(a) < \infty$ that $0 \geq F^*(0) \geq F^*(z_n) + u_n \cdot (0 - z_n) \geq a \cdot z_n - F(a) + u_n \cdot |z_n|$, which implies that $u_n \leq F(a)/|z_n| + a$. This implies that $\limsup_n u_n \leq a$. Sending now $a \to 0$ yields the claim. □

Remark 2.11 (Feasibility for finite $F(0)$). *Note that for $F(0) < \infty$ the trivial transport plan $\gamma = 0$ leads to a finite cost in (2.5) so that $W(\mu, \nu) < \infty$ for all $\mu, \nu \in \mathcal{M}_+(\Omega)$.*

2.3. Semi-discrete transport

An important special case of the classical balanced optimal transport problem (2.3) is the case where μ is absolutely continuous with respect to the Lebesgue measure,

$$\mu \ll \mathcal{L}\,, \tag{2.7a}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

and ν is a discrete measure,

$$\nu = \sum_{i=1}^{M} m_i \delta_{x_i}\,, \tag{2.7b}$$

with $m_i > 0$, $x_i \in \Omega$ and $x_i \neq x_j$ for $i \neq j$. See Section 1 for context and references. In this section we review the special structure of problem (2.3) that follows from (2.7). For instance, optimal couplings for (2.3) turn out to have a very particular form: the domain Ω is partitioned into cells, one cell for each discrete point x_i, and mass will only be transported from each cell to its corresponding discrete point. The shape of the cells is determined by μ, ν and the cost function c and can be expressed with the aid of Definition 2.12. Problem (2.3) can be rewritten explicitly as an optimization problem in terms of the cells. This tessellation formulation is given in Theorem 2.14, and its optimality conditions are described in Theorem 2.16.

Definition 2.12 (Generalized Laguerre cells). *Given a transportation cost c and points $x_1, \ldots, x_M \in \Omega$, we define the generalized Laguerre cells corresponding to the weight vector $w \in \mathbb{R}^M$ by*

$$C_i(w) = \left\{ x \in \Omega \;\middle|\; c(x, x_i) < \infty,\ c(x, x_i) - w_i \leq c(x, x_j) - w_j \text{ for all } j \in \{1, \ldots, M\} \right\} \tag{2.8}$$

for $i \in \{1, \ldots, M\}$. The residual of Ω, the set not covered by any of the cells C_i, is defined by

$$R = \left\{ x \in \Omega \;\middle|\; c(x, x_i) = \infty \text{ for all } i \in \{1, \ldots, M\} \right\}. \tag{2.9}$$

Note that R can also be written as $R = \Omega \setminus \left(\bigcup_{i=1}^{M} C_i(w) \right)$, which does not depend on $w \in \mathbb{R}^M$. Note also that, if $a = \lambda(1, 1, \ldots, 1) \in \mathbb{R}^M$ is a vector with all components equal, then $C_i(w + a) = C_i(w)$ for all $i \in \{1, \ldots, M\}$.

Example 2.13 (Generalized Laguerre cells [3]).

(a) ***Voronoi diagrams.*** *If c is radial (see Definition 2.8) and finite, then the collection of generalized Laguerre cells with weight vector $0 \in \mathbb{R}^M$, $\{C_i(0)\}_{i=1}^M$, is just the Voronoi diagram generated by the points $x_1, \ldots, x_M$. The residual set $R = \emptyset$.*

(b) ***Laguerre diagrams or power diagrams.*** *If $c(x, y) = |x - y|^2$, then the collection of generalized Laguerre cells $\{C_i(w)\}_{i=1}^M$ is known as the Laguerre diagram or power diagram generated by the weighted points $(x_1, w_1), \ldots, (x_M, w_M)$. The cells C_i are the intersection of convex polytopes with Ω. The residual set $R = \emptyset$.*

(c) ***Apollonius diagrams.*** *If $c(x, y) = |x - y|$, then the collection of generalized Laguerre cells $\{C_i(w)\}_{i=1}^M$ is known as the Apollonius diagram generated by the weighted points $(x_1, w_1), \ldots, (x_M, w_M)$. The cells C_i are the intersection of star-shaped sets with Ω, and in two dimensions, the boundaries between cells are arcs of hyperbolas. Again, $R = \emptyset$.*

Theorem 2.14 (Dual tessellation formulation for semi-discrete transport). *Assume that μ and ν satisfy (2.7) and $\mu(\Omega) = \nu(\Omega)$. Let the cost function c be radial (see Definition 2.8) and $W_{OT}(\mu, \nu) < \infty$. Then*

$$W_{\mathrm{OT}}(\mu, \nu) = \sup \left\{ \sum_{i=1}^{M} \int_{C_i(w)} c(x, x_i)\, \mathrm{d}\mu(x) + \sum_{i=1}^{M} \left(m_i - \mu(C_i(w)) \right) \cdot w_i \;\middle|\; w \in \mathbb{R}^M \right\}. \tag{2.10}$$

Remark 2.15 (Existence of optimal weights). *Maximizers for (2.10) do not always exist, even when $W_{\mathrm{OT}}(\mu, \nu) < \infty$. A simple sufficient condition for existence is that c is bounded from above on $\Omega \times \Omega$. More details can be found, for instance, in [69, Thm. 5.10].*

Theorem 2.16 (Optimality conditions). *Under the conditions of Theorem 2.14, a coupling $\gamma \in \Gamma(\mu, \nu)$ and a vector $w \in \mathbb{R}^M$ are optimal for $W_{\mathrm{OT}}(\mu, \nu)$ in (2.3) and (2.10) respectively, if and only if*

$$\gamma = \sum_{i=1}^{M} \left(\mu \llcorner C_i(w) \otimes \delta_{x_i} \right), \qquad \mu(C_i(w)) = m_i \text{ for } i \in \{1, \ldots, M\}. \tag{2.11}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Proofs of Theorem 2.14 and Theorem 2.16 can be found below and for example in [37] and [52, Section 4] for twisted costs c. We provide proofs of Theorems 2.14 and 2.16 for two reasons. They serve as preparation for the proof of Theorems 3.1 and 3.3 in the case of semi-discrete *unbalanced* transport, which generalize Theorems 2.14 and 2.16. In addition, they deal with the technical aspect that our cost function c is not necessarily twisted and may take the value $+\infty$ at finite distances. In particular, c does not satisfy the assumptions in [37, 52]. We rely on the following lemma, which essentially provides the existence of a Monge map in the semi-discrete setting (Corollary 2.18). For twisted costs, this result can be found in [52, Proposition 37].

Lemma 2.17 (Laguerre cell boundaries). *Let the cost function c be radial in the sense of Definition 2.8 and let $\{x_i\}_{i=1}^M$ be M distinct points in Ω. The induced generalized Laguerre cells satisfy $|C_i(w) \cap C_j(w)| = 0$ for $i \neq j$ and any $w \in \mathbb{R}^M$.*

Proof. Fix $i \neq j$ and $w \in \mathbb{R}^M$ and recall that $c(x, y) = \ell(d(x, y))$. We have

$$C_i(w) \cap C_j(w) = \bigcup_{n\in\mathbb{N}} A_n \quad \text{for } A_n = \{x \in \Omega \mid c(x, x_i) - w_i = c(x, x_j) - w_j,\ c(x, x_i) \leq n\}\,,$$

and we will show that the d-dimensional Hausdorff measure of each A_n is zero, $\mathcal{H}^d(A_n) = 0$, which implies $|A_n| = 0$ and thus also $|C_i(w) \cap C_j(w)| = 0$. Indeed, abbreviating $f = d(\cdot, x_i)$ and noting that the Jacobian of f equals 1 almost everywhere, the coarea formula [1, Thm. 2.93] yields

$$\mathcal{H}^d(A_n) = \int_{A_n} 1 \,\mathrm{d}\mathcal{H}^d = \int_{\mathbb{R}} \mathcal{H}^{d-1}(A_n \cap f^{-1}(t))\,\mathrm{d}t = \int_0^{\ell^{-1}(n)} \mathcal{H}^{d-1}(A_n \cap f^{-1}(t))\,\mathrm{d}t\,.$$

Now, for $t \in [0, \ell^{-1}(n)]$,

$$A_n \cap f^{-1}(t) = \{x \in \Omega \mid d(x, x_i) = t \text{ and } d(x, x_j) \in \ell^{-1}(\ell(d(x, x_i)) + w_j - w_i)\},$$

where $\ell^{-1}(\ell(d(x, x_i)) + w_j - w_i)$ is either empty or single-valued due to the strict monotonicity of ℓ. Hence, $A_n \cap f^{-1}(t)$ is contained in the intersection of two non-concentric $(d-1)$-dimensional spheres and thus is $\mathcal{H}^{d-1}$-negligible. □

Proof of Theorem 2.14. By Kantorovich duality [69, Thm. 5.10] we can write

$$W_{\mathrm{OT}}(\mu, \nu) = \sup\left\{ \int_\Omega \phi \,\mathrm{d}\mu + \int_\Omega \psi \,\mathrm{d}\nu \;\middle|\; \phi \in L^1(\mu),\ \psi \in L^1(\nu),\right.$$
$$\left. \phi(x) + \psi(y) \leq c(x, y)\ \forall\,(x, y) \in \Omega \times \Omega \right\}\,. \tag{2.12}$$

Since ν is discrete, $L^1(\nu)$ is isomorphic to $\mathbb{R}^M$ under the isomorphism $I{:}L^1(\nu) \to \mathbb{R}^M$, $\psi \mapsto (\psi(x_1), \ldots, \psi(x_M))$. The above dual problem thus becomes

$$W_{\mathrm{OT}}(\mu, \nu) = \sup\left\{ \int_\Omega \phi \,\mathrm{d}\mu + \sum_{i=1}^M w_i\, m_i \;\middle|\; \phi \in L^1(\mu),\ w \in \mathbb{R}^M,\right.$$
$$\left. \phi(x) + w_i \leq c(x, x_i)\ \forall\, x \in \Omega, i \in \{1, \ldots, M\} \right\}\,.$$

Next, for fixed w, one can explicitly maximize over ϕ, which corresponds to pointwise maximization subject to the constraint. We denote the maximizer by ϕ_w to emphasize the dependency on w,

$$\phi_w(x) = \min\left\{c(x, x_i) - w_i \mid i = 1, \ldots, M\right\}\,. \tag{2.13}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Since $\infty > W_{\mathrm{OT}}(\mu,\nu) - \sum_{i=1}^M w_i\, m_i \geq \int_\Omega \phi_w \,\mathrm{d}\mu$ and ϕ_w is bounded from below (as c is so in our setting) one must have $\phi_w \in L^1(\mu)$ for all $w \in \mathbb{R}^M$, and we find

$$W_{\mathrm{OT}}(\mu,\nu) = \sup\{\mathcal{E}_{\mathrm{SD}}(w) \mid w \in \mathbb{R}^M\} \quad \text{with} \quad \mathcal{E}_{\mathrm{SD}}(w) = \int_\Omega \phi_w(x)\,\mathrm{d}\mu(x) + \sum_{i=1}^M w_i\, m_i\,. \tag{2.14}$$

Since $\phi_w \in L^1(\mu)$ for any $w \in \mathbb{R}^M$, the residual set R must be μ-negligible; likewise, the intersection of generalized Laguerre cells is μ-negligible by Lemma 2.17. Consequently,

$$\mathcal{E}_{\mathrm{SD}}(w) = \sum_{i=1}^M \int_{C_i(w)} \phi_w(x)\,\mathrm{d}\mu(x) + \sum_{i=1}^M w_i\, m_i = \sum_{i=1}^M \int_{C_i(w)} [c(x,x_i) - w_i]\,\mathrm{d}\mu(x) + \sum_{i=1}^M w_i\, m_i\,,$$

which leads to the desired result. □

Proof of Theorem 2.16. The condition $\gamma \in \Gamma(\mu,\nu)$ implies that γ can be written as $\gamma = \sum_{i=1}^M \gamma_i \otimes \delta_{x_i}$ where $\gamma_i \in \mathcal{M}_+(\Omega)$, $\gamma_i(A) := \gamma(A \times \{x_i\})$. Observe that $\sum_{i=1}^M \gamma_i = \mu$ and $\gamma_i(\Omega) = m_i$. We obtain

$$W_{\mathrm{OT}}(\mu,\nu) \leq \int_{\Omega\times\Omega} c\,\mathrm{d}\gamma = \sum_{i=1}^M \int_\Omega c(x,x_i)\,\mathrm{d}\gamma_i(x)\,, \tag{2.15}$$

where the inequality is an equality if and only if γ is optimal. Let $w \in \mathbb{R}^M$. From (2.14) with ϕ_w given by (2.13), we find

$$W_{\mathrm{OT}}(\mu,\nu) \geq \int_\Omega \phi_w(x)\,\mathrm{d}\mu(x) + \sum_{i=1}^M w_i\, m_i = \sum_{i=1}^M \int_\Omega [\phi_w(x) + w_i]\;\mathrm{d}\gamma_i(x)\,, \tag{2.16}$$

where the inequality is an equality if and only if w is optimal. Subtracting (2.16) from (2.15) yields

$$0 \leq \sum_{i=1}^M \int_\Omega [c(x,x_i) - w_i - \phi_w(x)]\;\mathrm{d}\gamma_i(x) \tag{2.17}$$

with equality if and only if γ and w are optimal. By definition of ϕ_w, the integrand in each term of the sum is nonnegative and strictly positive for $x \notin C_i(w)$. Therefore (2.17) is an equality if and only if γ_i is concentrated on $C_i(w)$ for all $i \in \{1,\dots,M\}$. Combining absolute continuity with respect to the Lebesgue measure of μ and γ_i and Lemma 2.17 implies that the unique choice is $\gamma_i = \mu \llcorner C_i(w)$. Due to the second marginal constraint, this implies $\mu(C_i(w)) = \gamma_i(\Omega) = m_i$. □

The above results imply the existence of an optimal Monge map for the semi-discrete problem.

Corollary 2.18 (Existence of Monge map). *If a maximizer $w \in \mathbb{R}^M$ of (2.10) exists (cf. Remark 2.15), then the optimal coupling γ in (2.3) is induced by a transport map $T : \Omega \to \{x_i\}_{i=1}^M \subset \Omega$, $\gamma = (\mathrm{Id} \times T)_\# \mu$, defined by $T(x) = x_i$ when $x \in C_i(w)$. By virtue of Lemma 2.17 and since $\mu \ll \mathcal{L}$, T is well defined μ-almost everywhere.*

Example 2.19 (Optimal tessellations for Wasserstein distances). *Let μ and ν satisfy (2.7).*

(a) ***Wasserstein-2 distance.*** *Let $c(x,y) = |x-y|^2$. If T is an optimal Monge map, then the optimal transport cells $T^{-1}(\{x_i\})$ are the Laguerre cells (or power cells) $C_i(w)$ with weight vector $w = (\psi(x_1),\dots,\psi(x_M))$, where $\psi : \Omega \to \mathbb{R}$ is an optimal Kantorovich potential for the dual transport problem (2.12).*

(b) ***Wasserstein-1 distance.*** *Let $c(x,y) = |x-y|$. If T is an optimal Monge map, then the optimal transport cells $T^{-1}(\{x_i\})$ are the Apollonius cells $C_i(w)$ with weight vector $w = (\psi(x_1),\dots,\psi(x_M))$, where ψ is an optimal Kantorovich potential.*

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

3. Semi-discrete unbalanced transport

In this section we consider *semi-discrete unbalanced transport*. That is, we study (2.6) for the cases where μ is absolutely continuous with respect to the Lebesgue measure and ν is discrete, as stated in (2.7), and we do not require that $\mu(\Omega) = \nu(\Omega)$. Semi-discrete unbalanced transport models the situation where there is a mismatch between the capacity of a discrete resource ν and the demand of a population μ.

3.1. Tessellation formulation

The main results of this Section are Theorems 3.1 and 3.3, which generalize Theorems 2.14 and 2.16 to the unbalanced setting. Furthermore, in Corollary 3.6, we state a 'primal' counterpart of Theorem 3.1, which is somewhat pathological in the classical, balanced optimal transport setting, but quite natural in the unbalanced case.

The following result generalizes Theorem 2.14 to unbalanced transport.

Theorem 3.1 (Tessellation formulation for semi-discrete unbalanced transport). *Let the cost function c be radial (see Definition 2.8). Given $\mu, \nu \in \mathcal{M}_+(\Omega)$ satisfying (2.7), define $\mathcal{G} : \mathbb{R}^M \to (-\infty, \infty]$ by*

$$\mathcal{G}(w) = -\sum_{i=1}^{M} \left(\int_{C_i(w)} F^*\big(-c(x, x_i) + w_i\big)\, \mathrm{d}\mu(x) + F^*(-w_i) \cdot m_i \right) + F(0) \cdot \mu(R) \tag{3.1a}$$

with the convention $\infty \cdot 0 = 0$. Then the unbalanced optimal transport distance can be obtained via

$$W(\mu, \nu) = \sup \left\{ \mathcal{G}(w) \,\middle|\, w \in \mathbb{R}^M \right\}. \tag{3.1b}$$

This is a concave maximization problem.

Proof. In analogy to the Kantorovich duality (2.12) for the classical optimal transport problem (2.3), we make use of a corresponding duality result for the unbalanced transport problem (2.6),

$$W(\mu, \nu) = \sup \left\{ -\int_\Omega F^*(-\phi(x))\, \mathrm{d}\mu(x) - \int_\Omega F^*(-\psi(x))\, \mathrm{d}\nu(x) \,\middle|\, \phi, \psi \in \mathcal{C}(\Omega), \right.$$
$$\left. \phi(x) + \psi(y) \le c(x, y)\ \forall\, (x, y) \in \Omega \times \Omega \right\}.$$

This follows from [45, Thm. 4.11 and Cor. 4.12], where the former establishes the duality formula with ϕ and ψ ranging over all lower semi-continuous simple functions and the latter allows us to use continuous functions instead, exploiting the fact that F^* is continuous on $\mathbb{R}$ by Lemma 2.10 (v). Analogously to the proof of Theorem 2.14 (dual tessellation formulation), we now parameterize the function ψ on the set $\{x_i\}_{i=1}^M$ by a vector $w \in \mathbb{R}^M$, $w_i = \psi(x_i)$, and obtain

$$W(\mu, \nu) = \sup \left\{ -\int_\Omega F^*(-\phi(x))\, \mathrm{d}\mu(x) - \sum_{i=1}^{M} m_i F^*(-w_i) \,\middle|\, \phi \in \mathcal{C}(\Omega), w \in \mathbb{R}^M, \right.$$
$$\left. \phi(x) + w_i \le c(x, x_i)\ \forall\, x \in \Omega, i \in \{1, \dots, M\} \right\}. \tag{3.2}$$

Next, given $w \in \mathbb{R}^M$ we would like to optimize for ϕ as we did in (2.13). Note though that $\phi_w = \infty$ on the residual set R, which in unbalanced transport may be nonnegligible despite finite $W(\mu, \nu)$. For this reason we argue by truncation: For given $w \in \mathbb{R}^M$ and $n \in \mathbb{N}$, the function $\phi = \phi_{w,n}$ with

$$\phi_{w,n} : \Omega \to \mathbb{R}, \quad \phi_{w,n}(x) = \min\{n, \min\{c(x, x_i) - w_i \mid i \in \{1, \dots, M\}\}\}$$

lies in $\mathcal{C}(\Omega)$ and is feasible in (3.2). Moreover, for fixed w, the sequence $(\phi_{w,n})_{n\in\mathbb{N}}$ is a maximizing sequence for the maximization over ϕ, and it converges pointwise monotonically to the function ϕ_w

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

defined in (2.13). By Lemma 2.10 (ii) and (v), $z \mapsto -F^*(-z)$ is real-valued, continuous, and increasing. Since $\phi_{w,n}(x) \geq -\max_i w_i$, $-F^*(-\phi_{w,n}(x))$ is uniformly bounded from below with respect to n and x. Consequently, also ϕ_w is bounded from below. Therefore, the monotone convergence theorem implies that

$$\lim_{n\to\infty} -\int_\Omega F^*(-\phi_{w,n}(x))\,\mathrm{d}\mu(x) = -\int_\Omega F^*(-\phi_w(x))\,\mathrm{d}\mu(x),$$

where by convention $F^*(-\infty) = \lim_{z\to-\infty} F^*(z) = -F(0)$ (see Lemma 2.10). With this, (3.2) finally becomes

$$W(\mu,\nu) = \sup\left\{ -\int_\Omega F^*(-\phi_w(x))\,\mathrm{d}\mu(x) - \sum_{i=1}^M m_i F^*(-w_i) \,\middle|\, w \in \mathbb{R}^M \right\}. \tag{3.3}$$

Note that this objective is $> -\infty$ for all $w \in \mathbb{R}^M$ by the lower bound on ϕ_w and the monotonicity and real-valuedness of F^*. Now we decompose the integration domain Ω into $\{C_i(w)\}_{i=1}^M$ and R (using once more $\mu \ll \mathcal{L}$ and Lemma 2.17; it is only here that we use the radiality of the cost function c so that the cells $C_i(w)$ are well defined with negligible overlap). For $x \in C_i(w)$ one finds $\phi_w(x) = c(x,x_i) - w_i$, while for $x \in R$ one obtains $\phi_w(x) = \infty$ and therefore $F^*(-\phi_w(x)) = -F(0)$. This leads to expression (3.1a) and also implies that $\mathcal{G}(w) > -\infty$.

For fixed $x \in \Omega$ the map $w \mapsto \phi_w(x)$ is concave (since it is a minimum over affine functions). Moreover, the map $z \mapsto -F^*(-z)$ is concave and increasing (cf. Lemma 2.10 (ii)). Therefore, the objective function in (3.3) and consequently $\mathcal{G}$ are concave functions of w. □

Remark 3.2 (Finiteness of $\mathcal{G}$). *$\mathcal{G} \not\equiv \infty$ if and only if $\mathcal{G}$ is finite everywhere. Indeed, the last summand in (3.1a) is independent of w, and $-\sum_{i=1}^M F^*(-w_i)\,m_i$ is finite by* Lemma 2.10 *(v). Finally, if $\tilde{\mathcal{G}}(w) < \infty$ for some $w \in \mathbb{R}^M$, where $\tilde{\mathcal{G}}(w) = -\sum_{i=1}^M \int_{C_i(w)} F^*\big(-c(x,x_i)+w_i\big)\,\mathrm{d}\mu(x)$, then for $g_i(x) \in \partial F^*(-c(x,x_i)+w_i)$ and $\tilde{g}_i \in \partial F^*(w_i)$ we have*

$$\begin{aligned}\tilde{\mathcal{G}}(\tilde{w}) &\leq -\sum_{i=1}^M \int_{C_i(\tilde{w})} F^*\big(-c(x,x_i)+w_i\big) + g_i(x)(\tilde{w}_i - w_i)\,\mathrm{d}\mu(x)\\ &\leq \tilde{\mathcal{G}}(w) + \sum_{i=1}^M |\tilde{w}_i - w_i| \int_{C_i(w)} \tilde{g}_i\,\mathrm{d}\mu(x) < \infty\end{aligned}$$

for all $\tilde{w} \in \mathbb{R}^M$ by the convexity and monotonicity of F^* from Lemma 2.10 *(i).*

$\mathcal{G} \not\equiv \infty$ is for instance ensured by $F(0) < \infty$ or by boundedness of ℓ (which implies $R = \emptyset$). Indeed, the former implies boundedness of F^* from below by Lemma 2.10 *(vii), while the latter implies uniform boundedness of $-c(x,x_i)+w_i$ from below so that in either case the integrand of $\tilde{\mathcal{G}}$ is uniformly bounded from below.*

The following result generalizes the optimality conditions of Theorem 2.16 to unbalanced transport.

Theorem 3.3 (Optimality conditions). *Let $\gamma \in \mathcal{M}_+(\Omega\times\Omega)$, $w \in \mathbb{R}^M$, and set $\rho = \pi_{1\#}\gamma$. If $W(\mu,\nu) < \infty$ and γ and w are optimal for $W(\mu,\nu)$ in (2.6) and (3.1), respectively, then*

$$\gamma = \sum_{i=1}^M \rho \llcorner C_i(w) \otimes \delta_{x_i}, \tag{3.4a}$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) \in \partial F^*(-c(x,x_i)+w_i) \text{ for } \mu\text{-a.e. } x \in C_i(w), \quad \frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) = 0 \text{ for } x \in R, \tag{3.4b}$$

$$\frac{\rho(C_i(w))}{m_i} \in \partial F^*(-w_i) \text{ for } i \in \{1,\dots,M\}. \tag{3.4c}$$

Conversely, if γ and w satisfy (3.4), then they are optimal in (2.6) and (3.1), respectively.

Proof. Let $\gamma \in \mathcal{M}_+(\Omega\times\Omega)$ be such that $\mathcal{E}(\gamma)$ in (2.5) is finite. This implies that γ can be written as $\gamma = \sum_{i=1}^M \gamma_i \otimes \delta_{x_i}$ for $\gamma_i \in \mathcal{M}_+(\Omega)$, $\sum_{i=1}^M \gamma_i = \pi_{1\#}\gamma = \rho \ll \mu$ and $\rho(R) = 0$. (Note that the same holds

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

true if (3.4) is assumed instead of $\mathcal{E}(\gamma) < \infty$.) We obtain

$$\begin{aligned}\mathcal{E}(\gamma) &= \int_{\Omega\times\Omega} c \,\mathrm{d}\gamma + \mathcal{F}(\rho|\mu) + \mathcal{F}(\pi_{2\#}\gamma|\nu)\\ &= \sum_{i=1}^{M} \int_{\Omega\setminus R} c(x,x_i)\,\mathrm{d}\gamma_i(x) + \int_{\Omega\setminus R} F\left(\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x)\right)\mathrm{d}\mu(x) + F(0)\cdot\mu(R) + \sum_{i=1}^{M} F\left(\frac{\gamma_i(\Omega)}{m_i}\right)\cdot m_i\end{aligned}$$

so that the duality gap between the primal and dual formulations (2.6) and (3.1) reads

$$\begin{aligned}\mathcal{E}(\gamma) - \mathcal{G}(w) = \sum_{i=1}^{M} \int_{\Omega\setminus R} c(x,x_i)\,\mathrm{d}\gamma_i(x) + \int_{\Omega\setminus R}\left[F\left(\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x)\right) + F^*(-\phi_w(x))\right]\mathrm{d}\mu(x)\\ + \sum_{i=1}^{M}\left(F\left(\frac{\gamma_i(\Omega)}{m_i}\right) + F^*(-w_i)\right)\cdot m_i\,.\end{aligned}$$

Using the Fenchel–Young inequality, which states that $F(s) + F^*(z) \geq s\cdot z$ with equality if and only if $z \in \partial F(s)$ or equivalently $s \in \partial F^*(z)$ [6, Prop. 13.13 and Thm. 16.23], we obtain the lower bound

$$\begin{aligned}\mathcal{E}(\gamma) - \mathcal{G}(w) &\geq \sum_{i=1}^{M} \int_{\Omega\setminus R} c(x,x_i)\,\mathrm{d}\gamma_i(x) - \int_{\Omega\setminus R} \phi_w(x)\,\mathrm{d}\rho(x) - \sum_{i=1}^{M} w_i\cdot\gamma_i(\Omega)\\ &= \sum_{i=1}^{M} \int_{\Omega\setminus R} [c(x,x_i) - w_i - \phi_w(x)]\,\mathrm{d}\gamma_i(x) \geq 0\,,\end{aligned}$$

where the first inequality is an equality if and only if $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) \in \partial F^*(-\phi_w(x))$ for μ-almost every $x \in \Omega\setminus R$ and $\frac{\gamma_i(\Omega)}{m_i} \in \partial F^*(-w_i)$ for $i = 1,\dots,M$, and where the second inequality is an equality if and only if $\operatorname{spt}\gamma_i \subset C_i(w)$ and thus $\gamma_i = \rho\llcorner C_i(w)$ for $i = 1,\dots,M$. As a consequence, we have $\mathcal{E}(\gamma) - \mathcal{G}(w) = 0$ if and only if (3.4) holds.

Now let $W(\mu,\nu) < \infty$ and γ and w be optimal in (2.6) and (3.1) so that $W(\mu,\nu) = \mathcal{E}(\gamma) = \mathcal{G}(w) < \infty$. Then necessarily $\mathcal{E}(\gamma) - \mathcal{G}(w) = 0$ and so (3.4) holds. Conversely, if (3.4) holds, then if $\mathcal{E}(\gamma) < \infty$ or $\mathcal{G}(w) < \infty$ (so that the difference $\mathcal{E}(\gamma) - \mathcal{G}(w)$ is well defined), the above argument shows that $\mathcal{E}(\gamma) - \mathcal{G}(w) = 0$, which due to $\mathcal{E}(\gamma) \geq W(\mu,\nu) \geq \mathcal{G}(w)$ implies $W(\mu,\nu) = \mathcal{E}(\gamma) = \mathcal{G}(w)$ and thus the optimality of γ and w. If on the other hand $\mathcal{E}(\gamma) = \mathcal{G}(w) = \infty$, then $W(\mu,\nu) = \infty$ so that γ and w are trivially optimal. □

Corollary 3.4 (Uniqueness of coupling). *Let $W(\mu,\nu) < \infty$ and w be optimal for (3.1). Then the unique minimizer γ for (2.6) is given by (3.4a), where ρ is uniquely determined by (3.4b) and automatically satisfies (3.4c).*

Proof. We first show that (3.4b) fully specifies ρ. Let S be the set where ∂F^* is not a singleton. By convexity, S is countable. In analogy to Lemma 2.17, for any $s \in S$, the set $\{x \in \mathbb{R} \mid -c(x,x_i) + w_i = s\}$ is Lebesgue-negligible. Since S is countable, the set $\{x \in \mathbb{R} \mid -c(x,x_i) + w_i \in S\}$ is Lebesgue-negligible and thus also μ-negligible. Consequently, $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}$ is uniquely defined by (3.4b) on Ω up to a μ-negligible set.

For $W(\mu,\nu) < \infty$, conditions (3.4) are necessary and must therefore be satisfied by any minimizer γ (which exists by Theorem 2.5). Therefore, as ρ is uniquely determined by (3.4b), so is γ by (3.4a). Optimality of γ and w ensures that (3.4c) also holds. □

To gain some intuition, we will illustrate the previous results with numerical examples in the next section. Here we just spell out consistency with the balanced transport setting.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Remark 3.5 (Balanced transport). *For classical optimal transport with $F = \iota_{\{1\}}$ (such as the Wasserstein-2 distance from* Example 2.9 *(a)) one obtains $-F^*(-z) = z$. Then (3.1) becomes* (2.10) *(and finiteness of $W_{\mathrm{OT}}(\mu, \nu)$ implies that $\mu(R) = 0$). Furthermore, with $\partial F^*(z) = 1$ for all z, equation* (3.4b) *implies $\rho = \mu \llcorner (\Omega \setminus R) = \mu$. Then* (3.4a) *and* (3.4c) *become* (2.11).

From the derivation of (3.1) we learned that it can be interpreted as a variant of the dual problem to (2.6), where one of the dual variables is parametrized by w. Given the form of primal optimizers γ according to Theorem 3.3, we can formulate a corresponding variant of the primal problem.

Corollary 3.6 (Primal tessellation formulation of semi-discrete unbalanced transport). *Assume $W(\mu, \nu) < \infty$ and that optimizers of the unbalanced primal and dual problems (2.6) and (3.1) exist. Then*

$$W(\mu, \nu) = \min \left\{ \sum_{i=1}^{M} \int_{C_i(w)} c(x, x_i) \, \mathrm{d}\rho(x) + \mathcal{F}(\rho|\mu) + \sum_{i=1}^{M} F\left(\frac{\rho(C_i(w))}{m_i} \right) \cdot m_i \;\middle|\; w \in \mathbb{R}^M,\ \rho \in \mathcal{M}_+(\Omega),\ \rho \llcorner R = 0 \right\}. \tag{3.5}$$

If γ and w are optimal in (2.6) and (3.1), respectively, then w and $\rho = \pi_{1\#}\gamma$ are optimal in (3.5). Conversely, if w and ρ are optimal in (3.5), then (3.4a) defines an optimal γ for (2.6).

Proof. For any $w \in \mathbb{R}^M$ and $\rho \in \mathcal{M}_+(\Omega)$ with $\rho \llcorner R = 0$, the objective function in (3.5) is equal to $\mathcal{E}(\gamma)$ for $\gamma = \sum_{i=1}^{M} \rho \llcorner C_i(w) \otimes \delta_{x_i}$. Therefore, minimizing (3.5) corresponds to minimizing $\mathcal{E}$ over a particular subset of $\mathcal{M}_+(\Omega \times \Omega)$, which implies that the right-hand side of (3.5) is no smaller than $W(\mu, \nu)$. Now, if γ and w are a pair of optimizers for (2.6) and (3.1), then by (3.4), the objective function in (3.5) for w and $\rho = \pi_{1\#}\gamma$ becomes $\mathcal{E}(\gamma) = W(\mu, \nu)$ so that the right-hand side of (3.5) actually equals $W(\mu, \nu)$ and w and ρ are minimizers of (3.5).

Conversely, if w and ρ minimize (3.5), the induced γ must minimize $\mathcal{E}$. □

Remark 3.7 (Optimality of dual variable). *The converse conclusion that optimal w in* (3.5) *are optimal in (3.1) is in general not true. Indeed,* (3.5) *only depends on w via the cells $\{C_i(w)\}_{i=1}^M$ and therefore is invariant with respect to adding the same constant to all components of w, which does not change the cells. In the balanced case, when $F^*(z) = z$, then (3.1) is also invariant under such transformations (as long as $\mu(\Omega) = \sum_{i=1}^{M} m_i$). However, for general F, F^* is nonlinear, and the objective function of (3.1) is no longer invariant.*

Similarly, when the support of the optimal ρ in (3.5) *is bounded away from the boundary of some $C_i(w)$ (see Fig. 2, right), then slightly changing the corresponding w_i will not affect the value of* (3.5), *whereas (3.1) will usually not exhibit this invariance.*

Finally, if $c(x, x_i)$ becomes infinite for sufficiently small $d(x, x_i)$, then there exists an isolated cell $C_i(w)$ that is strictly bounded away from any other cell (see Fig. 4, right). In that case, none of the cells $\{C_j(w)\}_{j=1}^M$ depend on w_i, and so neither does (3.5). *However, the objective function of (3.1) in general still depends on w_i via F^*.*

Remark 3.8 (Primal tessellation formulation for classical optimal transport). *For classical optimal transport with $F = \iota_{\{1\}}$, the term $\mathcal{F}(\rho|\mu)$ in* (3.5) *is finite (and zero) if and only if $\rho = \mu$. Likewise, $\sum_{i=1}^{M} F\left(\frac{\rho(C_i(w))}{m_i}\right) \cdot m_i$ is finite (and zero) if and only if $\rho(C_i(w)) = m_i$. These are the optimality conditions given in* Theorem 2.16. *Thus, the objective function in* (3.5) *is finite only where it is optimal, making it somewhat pathological.*

Even though (3.5) is less pathological for more general unbalanced transport problems, we focus on (3.1) for numerical optimization.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

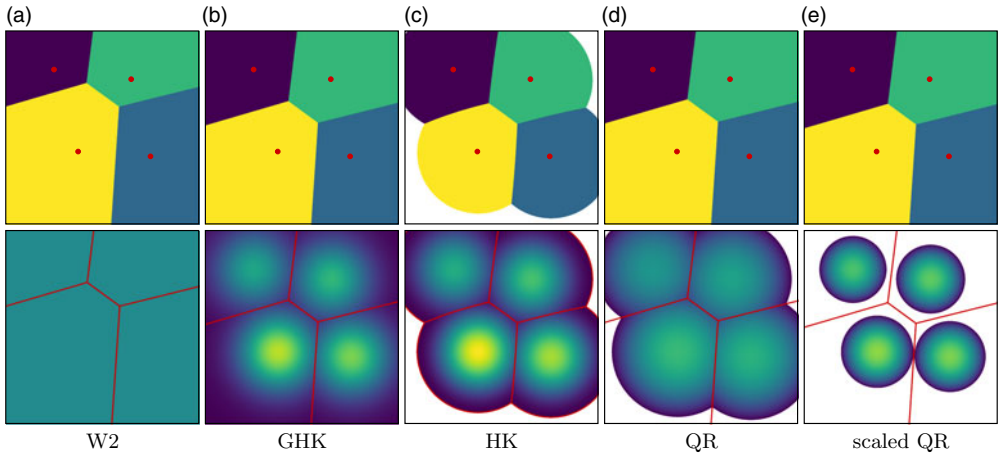


Figure 2. *Semi-discrete transport between the Lebesgue measure on* $\Omega = [0, L]^2$, $L = 5$ *and a discrete measure with* $M = 4$ *Dirac masses of locations* $(x_1, x_2, x_3, x_4) = L \cdot ((0.375, 0.375), (0.75, 0.35), (0.65, 0.75), (0.25, 0.8))$ *and weights* $(m_1, m_2, m_3, m_4) = |\Omega| \cdot (0.38, 0.29, 0.19, 0.14)$. *Top row: optimal cells* $\{C_i(w)\}_{i=1}^M$*; the residual set* R *is represented by white; location of the discrete points* $(x_1, \dots, x_M)$ *is indicated with red dots. Bottom row: optimal marginal* ρ *(identical colour scale in all figures; regions with* $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) = 0$ *are white for emphasis) and boundaries of cells* $\{C_i(w)\}_{i=1}^M$ *(red) are shown for models (a)–(b) from Examples 2.9 and 3.14. Figure (e) shows the same model as (d), only with* $c(x, y) = [d(x, y)/2]^2$ *instead of* $c(x, y) = d(x, y)^2$*; on some cells* spt ρ *is now strictly bounded away from the boundaries of* $C_i(w)$.

3.2. Numerical examples and different models

Depending on the choice of the cost function c and the marginal discrepancy $\mathcal{F}$, the semi-discrete unbalanced transport problem exhibits several qualitatively different regimes, which we will illustrate in this section. The discussion will be complemented with numerical examples.

Problem (3.1) is an unconstrained, finite-dimensional maximization problem over a concave objective. For simplicity, throughout this section we shall assume that the cost c is radial and F^* is differentiable or equivalently F is strictly convex (those assumptions are satisfied for the models from Example 2.9). This allows us to derive the objective function gradient in Theorem 3.10 and to treat the optimization problem with methods of smooth (as opposed to nonsmooth) optimization. A simple discretization scheme is given in Remark 3.12. The resulting discrete problem is solved with an L-BFGS quasi-Newton method [72]. As stated in Remark 3.13, the quality of the obtained solution can easily be verified via the primal-dual gap between (3.1) and (3.5). The special case of balanced optimal transport is discussed in Remark 3.11. Afterwards, we provide numerical illustrations for several examples of different unbalanced models.

To calculate the gradient of $\mathcal{G}$ we make use of the following lemma.

Lemma 3.9 (Derivative of integral functionals). *Let* $f : \Omega \times \mathbb{R}^M \to \mathbb{R}$ *be uniformly Lipschitz in its second argument, and let* $\mu \in \mathcal{M}_+(\Omega)$ *and* $u \in \mathbb{R}^M$ *be such that* $\mathbb{R}^M \ni \tilde{u} \mapsto f(x, \tilde{u})$ *is differentiable at* $\tilde{u} = u$ *for* μ*-almost all* $x \in \Omega$. *Define* $\mathcal{H}{:}\mathbb{R}^M \to \mathbb{R}$ *by* $\mathcal{H}(\tilde{u}) = \int_\Omega f(x, \tilde{u})\, \mathrm{d}\mu(x)$. *If* $|\mathcal{H}(u)| < \infty$, *then* $\mathcal{H}$ *is differentiable at* $\tilde{u} = u$ *with*

$$\frac{\partial \mathcal{H}}{\partial \tilde{u}}(u) = \int_\Omega \frac{\partial f}{\partial \tilde{u}}(x, u)\, \mathrm{d}\mu(x)\,.$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Proof. This is similar to the proof of Theorem 2.27 in [26], which however only covers the case where $M = 1$ and $f(x, \cdot)$ is differentiable for all $x \in \Omega$. We show that the directional derivative of $\mathcal{H}$ in an arbitrary direction $\hat{u} \in \mathbb{R}^M$ exists and is of the desired form. Indeed, let $L > 0$ be the Lipschitz constant of f in its second argument. By assumption, there exists $S \subset \Omega$ Lebesgue-negligible such that $f(x, \cdot)$ is differentiable at u for all $x \in \Omega \setminus S$. Now for $t \neq 0$,

$$\frac{\mathcal{H}(u + t\hat{u}) - \mathcal{H}(u)}{t} = \int_{\Omega \setminus S} \frac{f(x, u + t\hat{u}) - f(x, u)}{t} \, \mathrm{d}\mu(x) \, .$$

Since the integrand is bounded in absolute value by $L\|\hat{u}\|$ and since it converges pointwise to $\frac{\partial f}{\partial u}(x, u) \cdot \hat{u}$ as $t \to 0$, by the Dominated Convergence Theorem, we have

$$\lim_{t \to 0} \frac{\mathcal{H}(u + t\hat{u}) - \mathcal{H}(u)}{t} = \int_{\Omega \setminus S} \frac{\partial f}{\partial \tilde{u}}(x, u) \cdot \hat{u} \, \mathrm{d}\mu(x) = \int_{\Omega} \frac{\partial f}{\partial \tilde{u}}(x, u) \, \mathrm{d}\mu(x) \cdot \hat{u} \, .$$

The arbitrariness of $\hat{u}$, the linearity of the directional derivative, and the Lipschitz continuity of $\mathcal{H}$ imply that $\mathcal{H}$ is differentiable and has the desired form. □

Note that one may also replace the Lipschitz condition by a convexity or concavity condition, in which case dominated convergence would be replaced with monotone convergence; such a procedure would allow one to treat different examples in the later application Lemma 4.5.

Theorem 3.10 (Gradient of dual tessellation formulation). *If F is strictly convex and $\mathcal{G}$ from Theorem 3.1 is finite (e.g., if $F(0)$ is finite or ℓ is bounded, see Remark 3.2) then $\mathcal{G}$ is differentiable with*

$$\frac{\partial \mathcal{G}}{\partial w_i}(w) = (F^*)'(-w_i) \cdot m_i - \int_{C_i(w)} (F^*)'(-c(x, x_i) + w_i) \, \mathrm{d}\mu(x) \, . \tag{3.6}$$

Proof. Define

$$f(x, w) = \min\{-F^*(-c(x, x_j) + w_j) \mid j = 1, \ldots, M\}.$$

Since F^* is increasing by Lemma 2.10 (ii), then $f(x, w) = -F^*(-c(x, x_i) + w_i)$ for $x \in C_i(w)$. If ℓ is bounded, the residual set R is empty; otherwise, we have $f(x, w) = -F^*(-\infty) = F(0)$ for $x \in R$. Therefore

$$\mathcal{G}(w) = \int_{\Omega} f(x, w) \, \mathrm{d}\mu(x) - \sum_{i=1}^{M} F^*(-w_i) \cdot m_i.$$

By Remark 3.2, $\mathcal{G}(w)$ is finite for all $w \in \mathbb{R}^M$. Thus, there exists $\tilde{\Omega} \subset \Omega$ with $\mu(\Omega \setminus \tilde{\Omega}) = 0$ such that $f(x, w)$ is finite for all $x \in \tilde{\Omega}$ and $w \in \mathbb{R}^M$. Now consider the function $\Omega \times \mathbb{R} \ni (x, v) \mapsto f_i(x, v) = -F^*(-c(x, x_i) + v)$, where $i \in \{1, \ldots, M\}$. Furthermore, by Lemma 2.10 (vi) the strict convexity of F implies continuous differentiability of its conjugate F^* so that $f_i(x, \cdot)$ is differentiable for any $x \in \Omega$ with $f_i(x, \cdot)$ finite. Moreover, since F^* is convex and increasing, $\partial f_i / \partial v$ is nonpositive and decreasing so that $f_i(x, \cdot)$ (if finite) is Lipschitz on $(-\infty, \omega]$ for any $\omega \in \mathbb{R}$ with Lipschitz constant $L \leq -\frac{\partial f_i}{\partial v}(x, \omega) \leq (F^*)'(\omega)$. Consequently, $(-\infty, \omega]^M \ni \hat{w} \mapsto f(x, \hat{w})$ is Lipschitz with constant $\sqrt{M}(F^*)'(\omega)$ for all $x \in \tilde{\Omega}$ and differentiable in $\hat{w}$ for all $x \in \tilde{\Omega} \setminus S$, where $S = \bigcup_{i=1}^{M} \partial C_i(w)$ is Lebesgue-negligible and thus also μ-negligible. Thus, by the previous Lemma, $\mathcal{G}$ is differentiable with

$$\frac{\partial \mathcal{G}(w)}{\partial w_i} = (F^*)'(-w_i) \cdot m_i + \int_{\Omega} \frac{\partial f}{\partial w_i}(x, w) \, \mathrm{d}\mu(x) \, ,$$

where $\frac{\partial f}{\partial w_i}(x, w) = -(F^*)'(-c(x, x_i) + w_i)$ for μ-almost all $x \in C_i(w)$ and $\frac{\partial f}{\partial w_i}(x, w) = 0$ for μ-almost all $x \notin C_i(w)$. □

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Remark 3.11 (Balanced transport). *For classical optimal transport with $F = \iota_{\{1\}}$ as in* Remark 3.5, Theorem 3.10 *reduces to well-known results. In particular,* (3.6) *becomes*

$$\frac{\partial \mathcal{G}(w)}{\partial w_i} = m_i - \mu(C_i(w)) \,. \tag{3.7}$$

For more details we refer, for example, to [37, Thm. 1.1] or [52, Thm. 40]. For marginals $\mu = \tilde{\mu}\mathcal{L}$ with $\tilde{\mu} \in \mathcal{C}(\Omega)$ the Hessian

$$\frac{\partial^2 \mathcal{G}(w)}{\partial w_i \partial w_j} = -\frac{\partial \mu(C_i(w))}{\partial w_j} \tag{3.8}$$

can also be computed explicitly in terms of face integrals (see, for instance, [37, Thm. 1.3] and [52, Thm. 45]). Therefore (2.10) *lends itself to efficient numerical optimization [2, 37, 44, 50]. For special cost functions, most prominently for the squared Euclidean distance, the gradient* (3.7) *and Hessian* (3.8) *can be evaluated numerically efficiently and with high precision, allowing the application of Newton's method [37].*

The semi-discrete unbalanced problem (3.1) is more complicated due to the influence of the marginal fidelity $\mathcal{F}$ and since we are often interested in non-standard cost functions such as c_{HK}. Generalizing the above methods for balanced transport to the unbalanced case is therefore beyond the scope of this article.

Remark 3.12 (Discretization). *Problem (3.1) is already finite-dimensional. We must however evaluate the integrals over $C_i(w)$. For classical optimal transport and special cost functions c, these integrals can be evaluated essentially in closed form (see* Remark 3.11*). For simplicity, in this section, we approximate (Ω, μ) with Dirac masses on a fine Cartesian grid. The cells $\{C_i(w)\}_{i=1}^M$ are approximated using brute force by computing $c(x, x_i) - w_i$ for each point x in the Cartesian grid for each $i \in \{1, \dots, M\}$. Points x on the common boundaries of several cells $\{C_i(w)\}_{i=1}^M$ are arbitrarily assigned to one of those cells. (An efficient GPU-implementation of this brute force method can be found in [15].) Note that for the special cost $c(x, y) = |x - y|^2$, the Laguerre diagram $\{C_i(w)\}_{i=1}^M$ can be computed exactly, up to machine precision, and much more efficiently using, e.g., the lifting method [3, Sec. 6.2.2], which has complexity $\mathcal{O}(M \log M)$ in $\mathbb{R}^2$ and $\mathcal{O}(M^2)$ in $\mathbb{R}^3$. Our discretization yields an approximation of $\mathcal{G}(w)$ from* (3.1a) *and of $\nabla \mathcal{G}(w)$ from* (3.6)*, as required for the quasi-Newton method. In the numerical examples below, we use $\Omega = [0, L]^2$ for some $L > 0$ and approximate it by a regular Cartesian grid with* 1000 *points along each dimension.*

Remark 3.13 (Primal-dual gap). *The sub-optimality of any vector $w \in \mathbb{R}^M$ for (3.1) can be bounded by the primal-dual gap between* (3.1a) *and the objective of* (3.5)*. We avoid the remaining optimization over ρ in* (3.5) *by generating a feasible candidate via* (3.4b)*. Corollaries 3.4 and 3.6 guarantee that the primal-dual gap vanishes for optimal w.*

In the remainder of the section, we illustrate semi-discrete unbalanced transport by numerical examples. In particular, we showcase qualitative differences between different models as well as phenomena due to model-inherent length scales, which do not occur in classical, balanced transport.

Example 3.14 (Comparison of unbalanced transport models). *The structure of the optimal unbalanced coupling γ in (2.6) and its first marginal $\rho = \pi_{1\#}\gamma$ can vary substantially, depending on the choices for c and F. Below we discuss the models from Example 2.9 with a corresponding numerical illustration in Fig. 2.*

(a) ***Standard Wasserstein-2 distance (W2, Fig. 2(a)).*** *Since this is an instance of balanced transport, necessarily we have $\rho = \mu$. Furthermore, the cells $\{C_i(w)\}_{i=1}^M$ are standard, polygonal Laguerre cells and $R = \emptyset$.*

(b) ***Gaussian Hellinger–Kantorovich distance (GHK, Fig. 2(b)).*** *The cells are still standard polygonal Laguerre cells with $R = \emptyset$. This time, however, we usually have $\rho \neq \mu$. Nevertheless, we find $\operatorname{spt}\rho = \operatorname{spt}\mu$ since (3.4b) with $(F_{\mathrm{KL}}^*)'(z) = e^z > 0$ implies $\frac{\mathrm{d}\rho}{\mathrm{d}\mu} > 0$. This behaviour essentially originates from the infinite slope of F_{KL} in 0. Since $c(x, y) = d(x, y)^2$, the density $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}$ is piecewise Gaussian.*

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

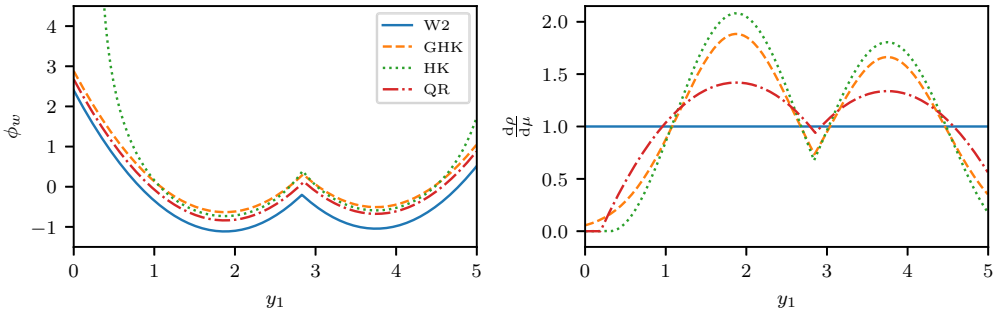


Figure 3. *One-dimensional slices of computational results from Figure 2 along* $[0, L] \times \{0.375\, L\}$ *with* $L = 5$. *Left:* ϕ_w *for optimal* $w \in \mathbb{R}^M$. *For models (a), (b), and (d),* ϕ_w *is piecewise quadratic; for (c) the profile is determined by* c_{HK} *and* $\phi_w = \infty$ *on* $R \neq \emptyset$. *Right: Optimal density* $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}$, *where* $\frac{\mathrm{d}\rho}{\mathrm{d}\mu} = (F^*)'(-\phi_w)$ *on* $\Omega \setminus R$ *and* 0 *elsewhere by (3.4b). For (a) the density is constant, for (b) it is piecewise Gaussian, for (c) it is piecewise given by* $\cos(d(y, x_i))^2$ *on* $\Omega \setminus R$ *and* 0 *on R, and for (d) it is given by truncated paraboloids.*

(c) ***Hellinger–Kantorovich distance (HK, Fig. 2(c)).*** *This time, the generalized Laguerre cells have curved boundaries, and also R is in general no longer empty, as* $c_{\mathrm{HK}}(x, y) = +\infty$ *for* $d(x, y) \geq \frac{\pi}{2}$. *Thus,* $\rho = 0$ *on R by (3.4b), independent of* μ. *However, similarly to (b) we have* $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) > 0$ *on the complement of R, the union of all generalized Laguerre cells.*

(d) ***Quadratic regularization (QR, Fig. 2(d)-(e)).*** *Since* $c(x, y) = d(x, y)^2$, *once more the cells are polygonal Laguerre cells and* $R = \emptyset$. *However, (3.4b) together with* $(F^*)'(z) = 0$ *for* $z \leq -2$ *implies* $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) = 0$ *whenever* $\phi_w(x) = \min\left\{c(x, x_i) - w_i \mid i = 1, \ldots, M\right\} \geq 2$, *even on* $\Omega \setminus R$. *Intuitively, this is possible since F and its right derivative are finite at* $z = 0$ *so that, for large transport costs* $c(x, x_i)$, *mass removal may be more profitable than transport.*

We emphasize that the reasons for $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) = 0$ *between models (c) and (d) are different: In the Hellinger–Kantorovich case,* $c(x, x_i) = \infty$ *for* $x \in R$ *prohibits any transport. In the quadratic case, despite finite transport cost and* $R = \emptyset$, *it may still be cheaper to remove and create mass via the fidelity F, due to its behaviour at* $z = 0$. *Also the slope at which* $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}$ *approaches zero is different for both models, as can be seen in the one-dimensional slice visualized in Fig. 3.*

Example 3.15 (Varying transport length scales). *As illustrated in the previous comparison of different models, unbalanced transport models typically have an intrinsic length scale that determines how far mass is optimally transported. Varying this length scale for fixed* μ *and* ν *changes the behaviour of the semi-discrete transport. For illustration, we concentrate on the Hellinger–Kantorovich distance and vary its intrinsic length scale by replacing* $c(x, y) = c_{\mathrm{HK}}(x, y)$ *with*

$$c(x, y) = c^{\varepsilon}_{\mathrm{HK}}(x, y) = c_{\mathrm{HK}}\left(\frac{x}{\varepsilon}, \frac{y}{\varepsilon}\right) = \begin{cases} -2\log\left[\cos\left(d(x, y)/\varepsilon\right)\right] & \text{if } d(x, y) < \frac{\pi}{2}\varepsilon, \\ \infty & \text{otherwise,} \end{cases}$$

that is, we set

$$\mathrm{HK}^{\varepsilon}(\mu, \nu)^2 = \inf\left\{\int_{\Omega\times\Omega} c^{\varepsilon}_{\mathrm{HK}}\mathrm{d}\gamma + \int_{\Omega} F_{\mathrm{KL}}\left(\frac{\mathrm{d}\pi_{1\#}\gamma}{\mathrm{d}\mu}\right)\mathrm{d}\mu + \int_{\Omega} F_{\mathrm{KL}}\left(\frac{\mathrm{d}\pi_{2\#}\gamma}{\mathrm{d}\nu}\right)\mathrm{d}\nu \,\middle|\, \gamma \in \mathcal{M}_+(\Omega \times \Omega)\right\}.$$

Note that this is equivalent to rescaling the domain Ω *by the factor* $\frac{1}{\varepsilon}$ *and simultaneously replacing the measures* μ *and* ν *by their pushforwards under* $x \mapsto \frac{x}{\varepsilon}$.

For large ε, *transport becomes very cheap relative to mass changes and thus asymptotically, as* $\varepsilon \to \infty$, *one recovers the Wasserstein-2 distance:* $\lim_{\varepsilon\to\infty} \varepsilon \mathrm{HK}^{\varepsilon}(\mu, \nu) = W_2(\mu, \nu)$ *by [45, Thm. 7.24]. In*

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

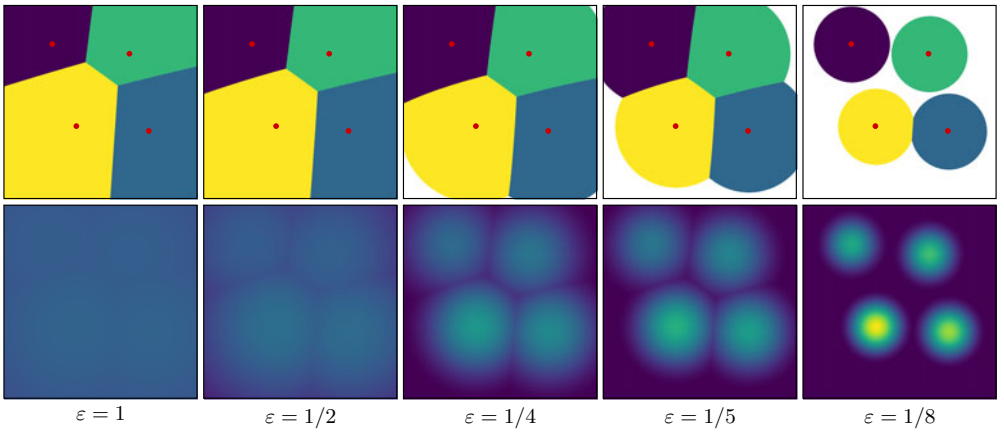


Figure 4. *Semi-discrete Hellinger–Kantorovich transport on* $\Omega = [0,1]^2$ *(using the same values for* x_i/L *and* $m_i/|\Omega|$ *as in Figure 2) for different length scales* ε*. Top row: optimal cells* $\{C_i(w)\}_{i=1}^M$*; the residual set* R *is represented by white; location of the discrete points* $(x_1, \dots, x_M)$ *is indicated with red dots. Bottom row: optimal marginal* ρ *(using the same colour scale for all images). For large* ε*, the behaviour is similar to that of the standard semi-discrete Wasserstein-2 distance. As* ε *decreases, the effects of unbalanced transport become increasingly prominent.*

particular, the distance diverges when $\mu(\Omega) \neq \nu(\Omega)$*. Conversely, as* $\varepsilon \searrow 0$*, transport becomes increasingly expensive and mass change is preferred. Asymptotically one obtains* $\lim_{\varepsilon \searrow 0} \mathrm{HK}^\varepsilon(\mu, \nu) = \mathrm{Hell}(\mu, \nu)$ *[45, Thm. 7.22], where* Hell *denotes the Hellinger distance*

$$\mathrm{Hell}(\mu, \nu)^2 = \int_\Omega \left(\sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\sigma}} - \sqrt{\frac{\mathrm{d}\nu}{\mathrm{d}\sigma}} \right)^2 \mathrm{d}\sigma$$

for $\sigma \in \mathcal{M}_+(\Omega)$ *an arbitrary reference measure with* $\mu, \nu \ll \sigma$ *(for instance* $|\mu| + |\nu|$ *with* $|\cdot|$ *indicating the total variation measure). By positive one-homogeneity of the function* $(m_1, m_2) \mapsto (\sqrt{m_1} - \sqrt{m_2})^2$*, the value of* $\mathrm{Hell}(\mu, \nu)$ *does not depend on the choice of* σ*. In our semi-discrete setting,* μ *and* ν *are always mutually singular so that* $\mathrm{Hell}(\mu, \nu)^2 = \mu(\Omega) + \nu(\Omega)$*.*

Figure 4 illustrates the optimal cells $\{C_i(w)\}_{i=1}^M$ *and marginal densities* $\rho = \pi_{1\#}\gamma$ *between the uniform volume measure* $\mu = \mathcal{L}$ *on* $\Omega = [0,1]^2$ *and a discrete measure* $\nu = \sum_{i=1}^M m_i \, \delta_{x_i}$ *for* $M = 4$*, using different values of the intrinsic length scale* ε *(the same experiment with* $M = 128$ *discrete points is shown in Fig. 5). As expected, for large* ε *the cells* $\{C_i(w)\}_{i=1}^M$ *look very similar to standard, polygonal Laguerre cells for the squared Euclidean distance* $c(x, y) = d(x, y)^2$*, and the residual set* R *is empty. The optimal* ρ *is essentially equal to* μ*, as dictated by balanced transport. As* ε *decreases, the boundaries between the cells become curved. Eventually,* R *becomes non-empty, and finally, the cells start to decompose into disjoint discs. In accordance, the density of the optimal marginal* ρ *is given on each cell* $C_i(w)$ *by* $\cos(d(x, x_i)/\varepsilon)^2 \cdot e^{w_i}$*. The interpolatory behaviour of* HK^ε *between the Wasserstein-2 distance* W_2 *and the Hellinger distance* Hell *for* $\varepsilon \to \infty$ *and* $\varepsilon \searrow 0$ *is numerically verified in Fig. 6.*

4. Unbalanced quantization

In this section we study the unbalanced quantization problem: we aim to approximate a given Lebesgue-continuous measure $\mu \in \mathcal{M}_+(\Omega)$ by a discrete, quantized measure $\nu = \sum_{i=1}^M m_i \cdot \delta_{x_i}$ with at most $M \in \mathbb{N}$ Dirac masses, where the unbalanced transport cost serves as a measure of approximation quality. To be

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

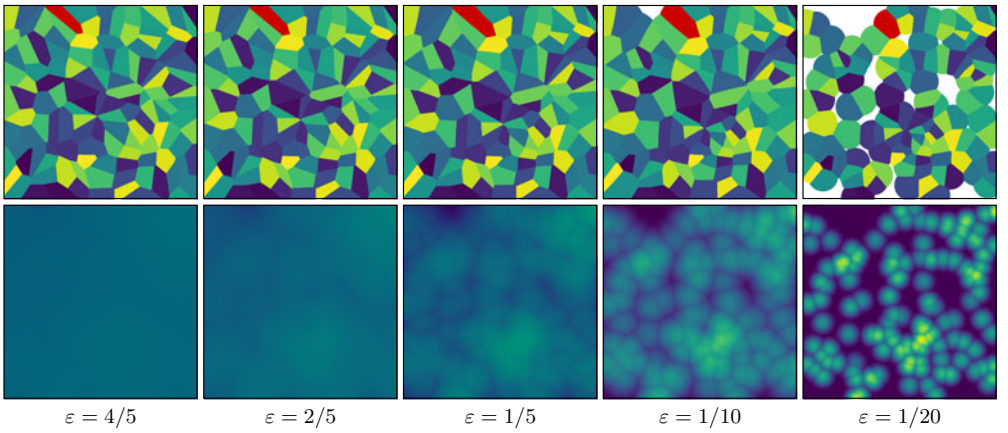


Figure 5. *Semi-discrete Hellinger–Kantorovich distance on* $\Omega = [0, 1]^2$ *for different length scales* ε*, as in Figure 4, but for* $M = 128$*. The evolution of one cell* $C_i(w)$ *for fixed* i *is highlighted in red (top row). For large* ε*,* $C_i(w)$ *is essentially the standard Wasserstein-2 Laguerre cell, not necessarily containing* x_i*. For small* ε*,* $C_i(w)$ *becomes (a fraction of) the open ball* $B_{\varepsilon\pi/2}(x_i)$*.*

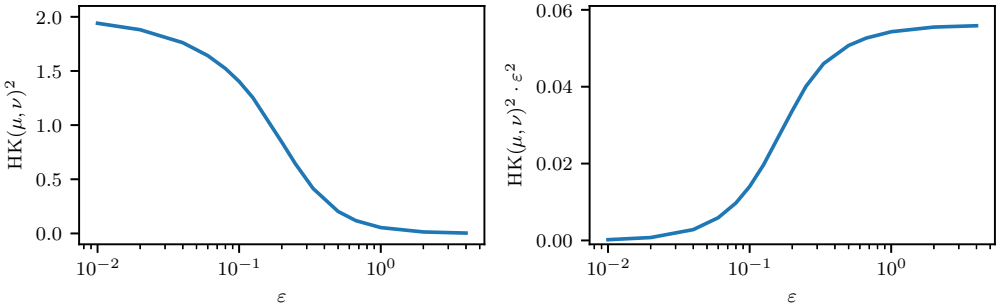


Figure 6. $\mathrm{HK}^\varepsilon(\mu,\nu)^2$ *for different length scales* ε *for the setup from Figure 4. Left: as* $\varepsilon \searrow 0$*,* $\mathrm{HK}^\varepsilon(\mu,\nu)^2$ *tends to* $\mathrm{Hell}(\mu,\nu)^2 = 2$*. Right: as* $\varepsilon \to \infty$*,* $\varepsilon^2\mathrm{HK}^\varepsilon(\mu,\nu)^2$ *tends to* $W_2(\mu,\nu)^2$*.*

precise, we consider the optimization problem

$$\min\left\{ W(\mu,\nu) \;\middle|\; \nu = \sum_{i=1}^{M} m_i\delta_{x_i},\ x_1,\dots,x_M \in \Omega,\ m_1,\dots,m_M \geq 0 \right\}. \tag{4.1}$$

Existence of minimizers will be established in Theorem 4.2. Applications include optimal location problems (economic planning), information theory (vector quantization) and particle methods for PDEs (approximation of continuous initial data by particles). We first characterize optimal particle configurations in terms of Voronoi diagrams, then consider a corresponding numerical scheme, and finally prove the optimal energy scaling of the quantization problem in terms of M for the case $d = 2$. The procedure essentially follows the one known for classical optimal transport; the important fact is that the Voronoi tessellation structure survives if mass changes are allowed.

The quantization cost will essentially depend on the function $-F^* \circ (-\ell)$ for ℓ from Definition 2.8 (see Theorem 4.2), which is why we briefly list a few of its relevant properties. We will mention below when these properties are assumed or used.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Lemma 4.1 (Properties of $-F^* \circ (-\ell)$). *Let ℓ define a radial cost and F a marginal discrepancy and consider the following conditions on them,*

$$F(0) < \infty \qquad or \qquad \ell(s) < \infty \; \forall \, s \in [0, \infty), \tag{4.2}$$

$$F(0) > 0\,, \tag{4.3}$$

$$F(1) = 0\,, \tag{4.4}$$

$$\lim_{s \to \infty} \ell(s) = \infty\,, \tag{4.5}$$

$$\ell \text{ is Lipschitz on } [0, \operatorname{diam}(\Omega)] \text{ or on } [0, \infty) \text{ when } \Omega \text{ is unbounded.} \tag{4.6}$$

Under these conditions, $-F^* \circ (-\ell)$ satisfies the following properties:

(P1) $-F^* \circ (-\ell)$ *is nondecreasing (increasing before it potentially reaches its maximum),*

(P2) $-F^* \circ (-\ell)$ *is continuous,*

(P3) $-F^* \circ (-\ell) < \infty$ *on* $(0, \infty)$,

(P4) $-F^* \circ (-\ell) > 0$ *on* $(0, \infty)$,

(P5) $-F^* \circ (-\ell)(0) = 0$,

(P6) $\lim_{s \to \infty} -F^* \circ (-\ell)(s) = F(0)$,

(P7) $-F^* \circ (-\ell)$ *is Lipschitz on* $[0, \operatorname{diam}(\Omega)]$ *or on* $[0, \infty)$ *when* Ω *is unbounded.*

More specifically, (P1) and (P2) hold by properties of F and ℓ according to Definitions 2.2 and 2.8, (4.2)⇒ (P3), (4.3)⇒ (P4), (4.4)⇒ (P5), (4.5)⇒ (P6), and (4.6)⇒ (P7).

We leave the proof as a straightforward exercise, it essentially being a direct consequence of Lemma 2.10 and Definitions 2.2 and 2.8. Note that the conditions are not necessary for the properties (P1)-(P7) to hold, which is why in the remainder of the section we will solely refer to these properties rather than to conditions on F and ℓ. However, the above conditions on F and ℓ are natural: (4.2) expresses that it is always possible to either completely remove or transport mass at any given location with finite cost ($W(\mu, \nu)$ may still be infinite, if $F(0) = \infty$, $\sup_z \ell(z) = \infty$, and μ has strong tails), (4.3) expresses that complete mass removal has a positive cost, (4.4) expresses that there is zero cost for not changing the mass, (4.5) expresses that the transport cost becomes infinite for infinite distances, and (4.6) ensures that the transport cost does not increase superlinearly. Of course (4.6) is not necessary for (P7), since the growth of ℓ can be compensated by a sufficiently slow increase of $-F^*(-\cdot)$, which corresponds to F growing only slowly (or being bounded) near zero. Examples for this are the GHK and HK distances; see Example 4.3. Throughout this section, we will assume that zero mass change induces zero cost, i.e. (4.4). This is the natural choice for approximating μ, as it implies a preference for $\pi_{1\#}\gamma = \mu$ in the first marginal fidelity term $\mathcal{F}$ of (2.5). Since $F(z) \geq 0$ by Definition 2.2, a consequence is

$$0 \in \partial F(1) \qquad \Leftrightarrow \qquad 1 \in \partial F^*(0) \qquad \text{and} \qquad F^*(0) = 0. \tag{4.7}$$

Also note that for the quantization problem only the behaviour of F on $[0, 1]$ is relevant since by a simple comparison argument one can see that any minimizer ν in (4.1) and any associated coupling γ in Definition 2.4 satisfy $\pi_{2\#}\gamma = \nu$ and $\pi_{1\#}\gamma \leq \mu$.

4.1. Unbalanced quantization as a Voronoi tessellation problem

The following theorem shows that the quantization problem can equivalently be formulated as an optimization of the points $x_1, \ldots, x_M$ with a functional depending on the Voronoi tessellation induced by $(x_1, \ldots, x_M)$.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Theorem 4.2 (Tessellation formulation of quantization problem). *For F satisfying (4.4), the unbalanced quantization problem (4.1) is equivalent to the minimization problem*

$$\min\left\{J(x_1,\dots,x_M)\,\middle|\,x_1,\dots,x_M\in\Omega\right\} \tag{4.8}$$

where

$$J(x_1,\dots,x_M)=\sum_{i=1}^{M}\int_{V_i(x_1,\dots,x_M)} -F^*(-c(x,x_i))\,\mathrm{d}\mu(x)$$

and where $V_i(x_1,\dots,x_M)=\{x\in\Omega\mid d(x,x_i)\le d(x,x_j)\ \text{for } j=1,\dots,M\}$ is the Voronoi cell associated with x_i, and we adopt the convention $-F^*(-\infty)=F(0)$ and $\partial F^*(-\infty)=\{0\}$ (cf. Lemma 2.10). Indeed, the minimum values coincide and, if $(x_1,\dots,x_M)$ minimizes (4.8) and the minimal value is finite, then $(x_1,\dots,x_M,m_1,\dots,m_M)$ minimizes (4.1) for

$$m_i=\int_{V_i(x_1,\dots,x_M)}\partial F^*(-c(x,x_i))\,\mathrm{d}\mu(x)\,,\quad i=1,\dots,M. \tag{4.9}$$

(By the proof of Corollary 3.4 the subgradient $\partial F^*(-c(x,x_i))$ contains a unique element for μ-almost every x and so the m_i are well defined.) Furthermore, the optimal transport plan γ associated with $W(\mu,\nu)$ only transports mass from each Voronoi cell $V_i(x_1,\dots,x_M)$ to the corresponding point x_i.

Example 4.3 (Tessellation formulation for unbalanced transport models). *The cost functional in (4.8) and the masses in (4.9) for the models from Example 2.9 are*

W2:
$$\begin{cases} J=\displaystyle\sum_{i=1}^{M}\int_{V_i} d(x,x_i)^2\,\mathrm{d}\mu(x)\,,\\ m_i=\mu(V_i)\,,\end{cases}$$

GHK:
$$\begin{cases} J=\displaystyle\sum_{i=1}^{M}\int_{V_i}\left[1-e^{-d(x,x_i)^2}\right]\mathrm{d}\mu(x)\,,\\ m_i=\displaystyle\int_{V_i} e^{-d(x,x_i)^2}\,\mathrm{d}\mu(x)\,,\end{cases}$$

HK:
$$\begin{cases} J=\displaystyle\sum_{i=1}^{M}\int_{V_i}\sin^2\left(\min\left\{d(x,x_i),\frac{\pi}{2}\right\}\right)\mathrm{d}\mu(x)\,,\\ m_i=\displaystyle\int_{V_i}\cos^2\left(\min\left\{d(x,x_i),\frac{\pi}{2}\right\}\right)\mathrm{d}\mu(x)\,,\end{cases}$$

QR:
$$\begin{cases} J=\displaystyle\sum_{i=1}^{M}\int_{V_i\cap B_{\sqrt{2}}(x_i)}\left[d(x,x_i)^2-\frac{d(x,x_i)^4}{4}\right]\mathrm{d}\mu(x)+\mu(V_i\setminus B_{\sqrt{2}}(x_i))\,,\\ m_i=\displaystyle\int_{V_i}\max\left\{1-\frac{d(x,x_i)^2}{2},0\right\}\mathrm{d}\mu(x).\end{cases}$$

Remark 4.4. *An intuitive strategy for proving* Theorem 4.2 *could be as follows. One starts from the primal tessellation formulation in* Corollary 3.6 *and in addition minimizes over masses $(m_1,\dots,m_M)$ and positions $(x_1,\dots,x_M)$. By (4.4) we find that minimizing masses are given by $m_i=\rho(C_i(w))$. Next, only the transport term depends on the weights w, and since the cost c is a strictly increasing function of distance, the term is minimized for $w=0$, thus essentially reducing the generalized Laguerre cells $C_i(w)$ into (truncated) Voronoi cells. Finally, the remaining minimization over ρ can be handled with arguments from convex analysis, similar to those of* Theorem 3.3, *thus arriving at* (4.8). *We give a shorter proof, using results from the dual tessellation formulation and its optimality conditions.*

Proof of Theorem 4.2. First we consider minimization over the masses $(m_i)_i$ for fixed positions $(x_i)_i$. Let $\nu=\sum_{i=1}^{M} m_i\cdot\delta_{x_i}$ be any admissible measure for (4.1). From (3.1b) we find $W(\mu,\nu)\ge\mathcal{G}(0)$ for any

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

positions $x_1, \dots, x_M$ and masses $m_1, \dots, m_M$. Note that $\mathcal{G}(0)$ does not depend on $m_1, \dots, m_M$ since we assume $F^*(0) = 0$, (4.7). Consider now the case where $\mathcal{G}(0) < \infty$ (and hence $\mathcal{G}$ is finite for all values of w, see Remark 3.2). We now show $W(\mu, \nu) = \mathcal{G}(0)$ for a particular choice of $m_1, \dots, m_M$, which therefore must be optimal (for given locations $x_1, \dots, x_M$). We first define ρ via (3.4b) and then γ via (3.4a) for $w = 0$ (ρ and γ are fully determined, see Corollary 3.4). Furthermore, since $1 \in \partial F^*(0)$ by (4.7), equation (3.4c) is satisfied by the choice $m_i = \rho(C_i(0))$. (Note that had we chosen $F(\hat{z}) = 0$ instead of $F(1) = 0$, one would simply use $m_i = \rho(C_i(0))/\hat{z}$ so that (4.9) would change by the factor $\hat{z}$.) By Theorem 3.3 (optimality conditions), γ and w are optimizers of $\mathcal{E}$ and $\mathcal{G}$ for these mass coefficients, which implies that $W(\mu, \nu) = \mathcal{G}(0)$. Using $F^*(0) = 0$ from (4.7), we have

$$\min_{(m_1,\dots,m_M)} W(\mu, \nu) = \mathcal{G}(0) = -\sum_{i=1}^{M} \int_{C_i(0)} F^*\big(-c(x, x_i)\big)\, \mathrm{d}\mu(x) + F(0) \cdot \mu(R)\,.$$

Since $c(x, y)$ is a strictly increasing function of the distance $d(x, y)$, for $w = 0$ we find $C_i(0) \subset V_i(x_1, \dots, x_M)$. With the convention $-F^*(-\infty) = F(0)$ (cf. Lemma 2.10), the term $F(0) \cdot \mu(R)$ becomes $\int_R -F^*(-\phi_0(x))\, \mathrm{d}\mu(x)$, where ϕ_0 was defined in equation (2.13). Since $\mu \ll \mathcal{L}$, integrating over R and $\Omega \setminus R$ is equivalent to integrating over all Voronoi cells $\{V_i(x_1, \dots, x_M)\}_{i=1}^M$, and we arrive at

$$\min_{(m_1,\dots,m_M)} W(\mu, \nu) = -\sum_{i=1}^{M} \int_{V_i(x_1,\dots,x_M)} F^*\big(-c(x, x_i)\big)\, \mathrm{d}\mu(x)\,, \tag{4.10}$$

which establishes equivalence between (4.1) and (4.8).

Finally, with $m_i = \rho(C_i(0))$ and ρ given by (3.4b) one obtains (4.9), where the integral runs over $C_i(0)$ instead of $V_i(x_1, \dots, x_M)$. If the minimum is finite, then either $\mu(R) = 0$ or $F(0)$ is finite, which implies the convention $(F^*)'(-\infty) = 0$ (cf. Lemma 2.10 (viii)). In both cases we can extend the area of integration to $V_i(x_1, \dots, x_M)$ without changing its value. Equation (3.4a) implies that mass is only transported from each Voronoi cell $V_i(x_1, \dots, x_M) \supset C_i(0)$ to the corresponding point x_i.

If on the other hand $\mathcal{G}(0) = \infty$, then by (3.1b) for all ν concentrated on the positions $(x_i)_i$ one has $W(\mu, \nu) = \infty$. This establishes the equality of the minimal values in both the finite and infinite cases.

Now we consider minimization over the positions $(x_i)_i$. We merely need to consider the case where the minimal value is finite, as otherwise any configuration is optimal. Let $((x_i^k)_{i=1}^M)_{k\in\mathbb{N}}$ be a minimizing sequence of points for the right-hand side of (4.10). After selection of a subsequence and relabelling the order of the points, there will be an integer $N \in \{0, \dots, M\}$ and points $(x_i)_{i=1}^N$ such that $\lim_k x_i^k = x_i$ for $i \in \{1, \dots, N\}$ and $\lim_k |x_i^k| = \infty$ for $i > N$. Let $(\rho^k)_{k\in\mathbb{N}}$ be the sequence induced via (3.4b) from $((x_i^k)_i)_k$ (for $w = 0$). Using non-negativity of c, convexity of F^* and the fact that $1 \in \partial F^*(0)$, we get that $\rho^k \le \mu$ for all k and hence the sequence is tight, allowing extraction of a weak (i.e. in duality with bounded continuous functions) cluster point ρ. Let $m_i^k = \rho^k(V_i((x_j^k)_j))$ be the corresponding optimal masses, as above, and let $m_i = \rho(V_i((x_j)_{j=1}^N))$ be the masses induced by ρ for $i \in \{1, \dots, N\}$. Using weak convergence of $\rho^k \to \rho$ and absolute continuity of ρ one finds that $m_i^k \to m_i$ for $i \in \{1, \dots, N\}$ and $m_i^k \to 0$ for $i > N$. One then has that $\nu^k := \sum_{i=1}^M m_i^k \cdot \delta_{x_i^k}$ converges weakly to $\nu := \sum_{i=1}^N m_i \cdot \delta_{x_i}$. By [45, Lemma 3.9], W is weakly lower-semicontinuous and therefore the points $(x_i)_{i=1}^N$ are candidates for a minimizer in (4.10). However, if $N < M$, they are too few. But adding arbitrary points on the right-hand side of (4.10) will not increase the objective, and thus any such extension must yield a minimizer. □

4.2. A numerical method: Lloyd's algorithm and Quasi-Newton variant

Formulation (4.8) has the advantage over (4.1) that it does not contain an inner minimization to find the optimal transport coupling. Thus, we aim to solve (4.8) numerically. To this end, we compute the gradient $\partial_{x_j} J$ (see also analogous derivatives for similar functionals as for instance in [13]).

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Lemma 4.5 (Derivative of the cost functional J). *Let $\mu \in \mathcal{M}_+(\Omega)$ be absolutely continuous, and let (P1) and (P7) hold. Then for $j = 1, \dots, M$,*

$$\partial_{x_j} J(x_1, \dots, x_M) = \int_{V_j(x_1,\dots,x_M)} r(d(x, x_j))(x_j - x)\, \mathrm{d}\mu(x)$$

where

$$r(s) = \frac{[-F^* \circ (-\ell)]'(s)}{s}$$

(note that $-F^* \circ (-\ell)$ is differentiable for almost every $s \in [0, \operatorname{diam}(\Omega)]$ so that r and $r(d(\cdot, x_j))$ are well defined almost everywhere).

Example 4.6 (Cost derivative for unbalanced transport models). *For the models from Example 2.9 one can readily check*

$$\begin{aligned}
&\text{W2:} && r(s) = 2\,,\\
&\text{GHK:} && r(s) = 2e^{-s^2}\,,\\
&\text{HK:} && r(s) = \sin(2s)/s \text{ if } s \le \frac{\pi}{2} \text{ and } 0 \text{ otherwise,}\\
&\text{QR:} && r(s) = \max\{2 - s^2, 0\}\,.
\end{aligned}$$

Note that W2 *only satisfies assumption (P7) on bounded domains.*

Proof of Lemma 4.5. Note that $J(x_1, \dots, x_M) = \int_\Omega f(x, (x_1, \dots, x_M))\, \mathrm{d}\mu(x)$ with

$$f(x, (x_1, \dots, x_M)) = \min\{-F^*(-\ell(d(x, x_i))) \mid i = 1, \dots, M\}$$

since $-F^* \circ (-\ell)$ is nondecreasing by (P1). By the Lipschitz assumption (P7) on $F^* \circ (-\ell)$, f is Lipschitz in its second argument. Furthermore, $F^* \circ (-\ell)$ is differentiable almost everywhere, and $d(x, x_i)$ is differentiable in its second argument for all $x \neq x_i$. Therefore, f is differentiable in its second argument at $(x_1, \dots, x_M)$ for almost all $x \in \Omega$ (thus for μ-almost all $x \in \Omega$) with

$$\partial_{x_j} f(x, (x_1, \dots, x_M)) = \begin{cases} r(d(x, x_j))(x_j - x) & \text{if } x \in V_j(x_1, \dots, x_M),\\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3.9 now implies the desired result. □

To find a minimizer of J and thus a solution to the optimality condition $\partial_{x_j} J = 0$ for $j = 1, \dots, M$, one can perform the following fixed point iteration associated with the optimality conditions,

$$x_i^{(k+1)} = \frac{\displaystyle\int_{V_i(x_1^{(k)},\dots,x_M^{(k)})} x r(d(x_i^{(k)}, x))\, \mathrm{d}\mu(x)}{\displaystyle\int_{V_i(x_1^{(k)},\dots,x_M^{(k)})} r(d(x_i^{(k)}, x))\, \mathrm{d}\mu(x)}\,, \quad i = 1, \dots, M,$$

starting from some initialization $x_1^{(0)}, \dots, x_M^{(0)} \in \Omega$. This iteration is well defined as long as the denominator is nonzero, for instance if μ is strictly positive on Ω (recall that $-F^*(-\ell(s))$ is strictly increasing for small s by (P1)). This is a generalization of Lloyd's algorithm for computing centroidal Voronoi Tessellations [21], which are critical points of the function

$$\tilde{J}(x_1, \dots, x_M) = \sum_{i=1}^{M} \int_{V_i(x_1,\dots,x_M)} |x - x_i|^2\, \mathrm{d}\mu(x)\,.$$

Its convergence has been proven in a number of settings [13, 24, 60], which also cover many possible choices for our μ, c, and F. Since the algorithm is based solely on the first variation one can expect linear

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

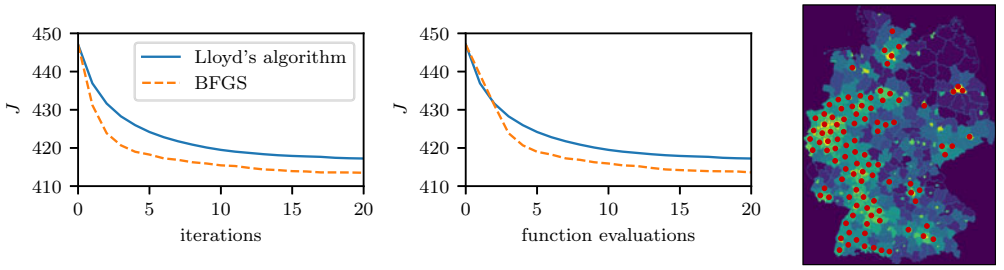


Figure 7. *Quantization energy decrease of Lloyd's algorithm and a BFGS method versus number of iterations (left) and function evaluations (centre; for the BFGS method function evaluations differ from iterations due to additional evaluations in the stepsize control) for the example shown on the right. Right: Input density μ and optimal locations $(x_1, \dots, x_M)$ for $M = 100$, where μ is population density in Germany 2015 (published by the Federal Statistical Office of Germany in the "Regional Atlas" http://www.destatis.de/regionalatlas). The computations use the Hellinger–Kantorovich model.*

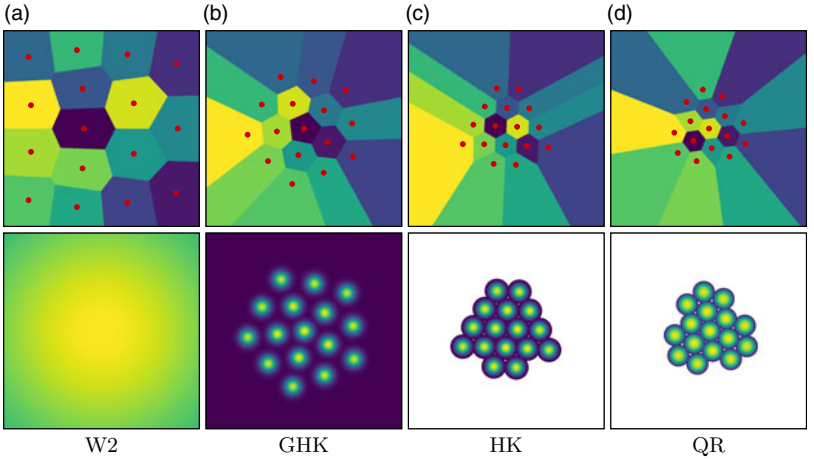


Figure 8. *Quantization results for $\mu = (1 + \exp(-\frac{|x|^2}{2(4\pi)^2})) \cdot \mathcal{L} \llcorner \Omega$ and $\Omega = [-4\pi, 4\pi]^2$, $M = 16$ on the same models as in Figure 2. Top row: optimal locations $x_1, \dots, x_M$ and Voronoi cells $\{V_i(x_1, \dots, x_M)\}_{i=1}^M$. Bottom row: optimal marginal $\rho = \pi_{1\#}\gamma$ (identical colour scale in all figures; regions with $\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x) = 0$ are white for emphasis). For (a) we have $\rho = \mu$.*

convergence. To achieve faster convergence one may use a quasi-Newton method for the minimization of J instead, which seems particularly well suited since the optimization is performed over a finite-dimensional space.

Our numerical implementation is performed in Matlab. The integrals over a Voronoi cell $V_i(x_1, \dots, x_M)$ are evaluated using Gaussian quadrature on the triangulation that is obtained by connecting each vertex of $V_i(x_1, \dots, x_M)$ with x_i. The Voronoi cells themselves are computed using the built-in function `voronoin`. Figure 7 shows a slightly faster convergence of the BFGS method compared to Lloyd's algorithm, while Figure 8 shows quantization results for the same models as in Figure 2, resulting in different point distributions. Similarly, Figure 9 shows quantization results for the same input marginal μ and the Hellinger–Kantorovich model, but for varying length scales.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

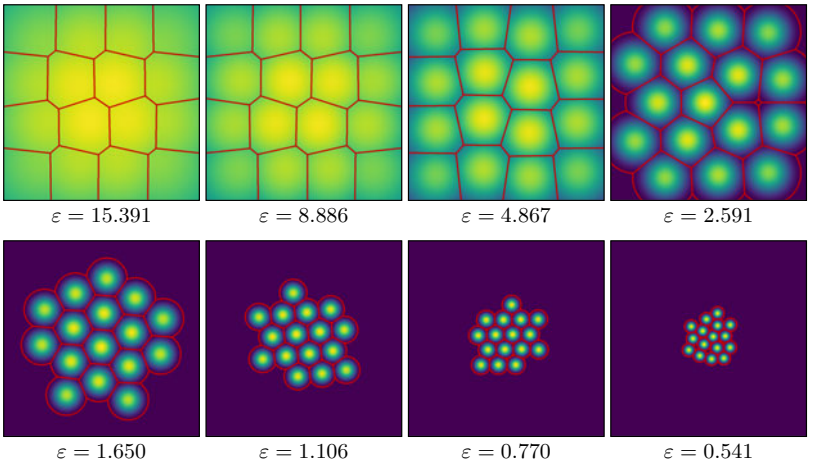


Figure 9. *Quantization results for the Hellinger–Kantorovich model and different length scales, showing the optimal Laguerre cells $C_i(0)$ (which coincide with the optimal Voronoi cells up to the set R from (2.9)) and the optimal marginals $\rho = \pi_{1\#}\gamma$ (same domain and μ as in Figure 8; identical colour scale in all figures).*

4.3. Crystallization in two dimensions

In this section we consider the asymptotic behaviour of the unbalanced quantization problem in the limit of infinitely many points, $M \to \infty$, in two dimensions, $\Omega \subset \mathbb{R}^2$, in which case crystallization results from discrete geometry are available.

To simplify the exposition in this section we assume (P3) so that the unbalanced transport cost is always finite. Additionally we assume (P4), which simply ensures that the quantization problem is not trivially degenerate. The situation without these conditions can in principle be treated similarly, but requires a number of technical case distinctions (such as whether the domain of $(-F^* \circ (-\ell))$ is open or closed).

As we increase M, the average distance between points of Ω and their nearest discrete point x_i decreases so that the (balanced) transport cost from μ onto ν vanishes in the limit, whereas the cost for changing mass remains unchanged. Therefore, in the limit $M \to \infty$, the interplay of transport and mass change in *unbalanced* transport would not be visible. To avoid this, we will rescale the metric on the domain Ω as M grows and study the resulting different regimes, depending on the scaling. Consequently, in this section, we consider the scaled cost

$$J_\varepsilon^M(x_1, \dots, x_M) = \sum_{i=1}^{M} \int_{V_i(x_1,\dots,x_M)} -F^*\left(-\ell\left(\frac{d(x, x_i)}{\varepsilon}\right)\right) \mathrm{d}\mu(x) \tag{4.11}$$

for $M \in \mathbb{N}$, $\varepsilon \in (0, \infty)$.

We first prove a lower bound on the quantization cost J_ε^M for the Lebesgue measure, which corresponds to a perfect triangular lattice. Then a corresponding upper bound is derived. Finally, for μ with Lipschitz continuous Lebesgue density, we show that these two bounds imply that asymptotically a locally regular triangular lattice becomes an optimal quantization configuration, where the local density of points depends on the density of μ.

Theorem 4.7 (Lower bound for quantization of the Lebesgue measure). *Let $\Omega \subset \mathbb{R}^2$ be a convex polygon with at most six sides, let μ be the Lebesgue measure on Ω, and let (P1) hold. A lower bound on (4.11)*

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

is given by

$$\min_{x_1,\dots,x_M\in\mathbb{R}^2} J_\varepsilon^M(x_1,\dots,x_M) \geq M \int_{H(|\Omega|/M)} -F^*\left(-\ell\left(\frac{d(x,0)}{\varepsilon}\right)\right)\,\mathrm{d}x\,, \tag{4.12}$$

where $H(|\Omega|/M)$ is a regular hexagon of area $|\Omega|/M = \mathcal{L}(\Omega)/M$ centred at the origin 0.

Remark 4.8 (Cost of the triangular lattice). *Comparing with* Theorem 4.2, *the lower bound is exactly the unbalanced transportation cost $W(\mu,\nu)$ from a regular triangular lattice ν of M Dirac measures of mass*

$$m = \int_{H(|\Omega|/M)} \partial F^*\left(-\ell\left(\frac{d(x,0)}{\varepsilon}\right)\right)\,\mathrm{d}x\,,$$

whose Voronoi cells are translations of $H(|\Omega|/M)$, onto μ the Lebesgue measure on the union of these Voronoi cells.

Proof of Theorem 4.7. Since $-F^*(-\ell(\cdot/\varepsilon))$ is increasing by (P1), for $x_1,\dots,x_M\in\Omega$, we have

$$\begin{aligned} J_\varepsilon^M(x_1,\dots,x_M) &= \sum_{i=1}^M \int_{V_i(x_1,\dots,x_M)} -F^*\left(-\ell\left(\frac{d(x,x_i)}{\varepsilon}\right)\right)\,\mathrm{d}x \\ &= \int_\Omega \min_{i=1,\dots,M} -F^*\left(-\ell\left(\frac{d(x,x_i)}{\varepsilon}\right)\right)\,\mathrm{d}x\,, \end{aligned}$$

and the result follows immediately from L. Fejes Tóth's Theorem on Sums of Moments [33] (see also [54, 68]). □

Remark 4.9 (Degeneracy of minimizers). *As opposed to the quantization problem for classical optimal transport, the set of minimizers in the unbalanced transport case can exhibit strong degeneracies. As an example, consider the case of Hellinger–Kantorovich transport with $M \ll 4\,|\Omega|/(\pi^3\varepsilon^2)$. Let $x_1,\dots,x_M$ be any arrangement of the point masses such that the balls $B_{\varepsilon\,\pi/2}(x_i)$ are pairwise disjoint and included in Ω (which necessarily implies $M \leq 4\,|\Omega|/(\pi^3\varepsilon^2)$). Then $(x_1,\dots,x_M)$ achieves the lower bound since*

$$\begin{aligned} J_\varepsilon^M(x_1,\dots,x_M) &= \sum_{i=1}^M \int_{V_i(x_1,\dots,x_M)} -F^*\left(-\ell\left(\frac{d(x,x_i)}{\varepsilon}\right)\right)\,\mathrm{d}x \\ &= \sum_{i=1}^M \int_{V_i(x_1,\dots,x_M)} \sin^2\left(\min\left\{d(x,x_i)/\varepsilon, \frac{\pi}{2}\right\}\right)\,\mathrm{d}x \\ &= |\Omega| - M\varepsilon^2\,\pi^3/4 + \sum_{i=1}^M \int_{B_{\varepsilon\,\pi/2}(x_i)} \sin^2\left(d(x,x_i)/\varepsilon\right)\,\mathrm{d}x \\ &= M\int_{H(|\Omega|/M)} -F^*(-\ell(d(x,0)/\varepsilon))\,\mathrm{d}x\,, \end{aligned}$$

where we used $H(|\Omega|/M) \supset B_{\varepsilon\,\pi/2}(0)$. Indeed, the energy does not discriminate between different solutions because $-F^*\circ(-\ell)$ is constant for distances larger than $\frac{\epsilon\pi}{2}$.

Theorem 4.10 (Upper bound for quantization of the Lebesgue measure). *Let $\Omega\subset\mathbb{R}^2$ be a square domain, let μ be the Lebesgue measure, and let (P1) hold. Let $x_1,\dots,x_M$ be a regular triangular arrangement of points in the following sense: Let $G\subset\mathbb{R}^2$ be a regular triangular lattice with lattice spacing $\sqrt{\frac{2\,|\Omega|}{\sqrt{3}\,M}}$, such that the corresponding Voronoi cells are regular hexagons with area $|\Omega|/M$ and side length $L = \sqrt{\frac{2\,|\Omega|}{3\sqrt{3}M}}$. Let $\{x_1,\dots,x_{\tilde{M}}\}\subset G$ be those points for which the corresponding hexagon H_i is*

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

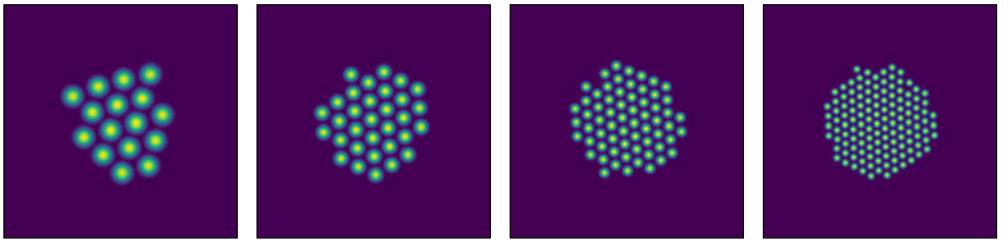


Figure 10. *Quantization results for the Hellinger–Kantorovich model using different length scales and numbers of discrete points, with constant total point density $\varepsilon_M^2 M$. The optimal marginals $\rho = \pi_{1\#}\gamma$ are shown (same domain and μ as in Figure 8; identical colour scale in all figures).*

fully contained in Ω. Assume that M is sufficiently large so that $\hat{M} \geq 1$. If $\hat{M} < M$, pick $\{x_{\hat{M}+1}, \dots, x_M\}$ arbitrarily from Ω. Then

$$J_\varepsilon^M(x_1, \dots, x_M) \leq M \int_{H(\frac{|\Omega|}{M})} -F^*\left(-\ell\left(\frac{d(x,0)}{\varepsilon}\right)\right) \mathrm{d}x - |\partial\Omega| \sqrt{\frac{8|\Omega|}{3\sqrt{3}M}}\, F^*\left(-\ell\left(\sqrt{\frac{C_Q\,|\Omega|}{\varepsilon^2 M}}\right)\right), \quad (4.13)$$

where $C_Q = 2(2\sqrt{2}+1)^2/(3\sqrt{3})$, $|\partial\Omega|$ denotes the one-dimensional Hausdorff measure of $\partial\Omega$ and $|\Omega| = \mathcal{L}(\Omega)$.

Proof. Let $S = \Omega \setminus \bigcup_{i=1}^{\hat{M}} H_i$ be those points in Ω that are not covered by any hexagon H_i. Note that all $x \in S$ lie no further away from $\partial\Omega$ than the diameter of a hexagon, $2L = \sqrt{\frac{8\,|\Omega|}{3\sqrt{3}M}}$. Since Ω is convex, we thus have $|S| \leq |\Omega \cap \bigcup_{x\in\partial\Omega} B_{2L}(x)| \leq 2\,L\,|\partial\Omega|$. Likewise, any point $x \in S$ lies no further away from the union of all hexagons H_i than $2\sqrt{2}L$ and thereby no further away from $\{x_1, \dots, x_M\}$ than $(2\sqrt{2}+1)L$, thus $\min_i d(x, x_i) \leq (2\sqrt{2}+1)L$.

Note that $V_i(x_1, \dots, x_M) \setminus S \subseteq H_i$ for $i = 1, \dots, \hat{M}$ and that $-F^* \circ (-\ell)$ is monotonously increasing so that we find

$$\begin{aligned} J_\varepsilon^M(x_1, \dots, x_M) &\leq \sum_{i=1}^{\hat{M}} \int_{H_i} -F^*\left(-\ell\left(\frac{d(x,x_i)}{\varepsilon}\right)\right) \mathrm{d}x + \int_S -F^*\left(-\ell\left(\frac{\min_i d(x,x_i)}{\varepsilon}\right)\right) \mathrm{d}x \\ &\leq M \cdot \int_{H(|\Omega|/M)} -F^*\left(-\ell\left(\frac{d(x,0)}{\varepsilon}\right)\right) \mathrm{d}x - |S| F^*\left(-\ell\left(\frac{(2\sqrt{2}+1)L}{\varepsilon}\right)\right). \end{aligned}$$

Substituting the value of L and the above bound for $|S|$ proves the claim. □

Remark 4.11 (A priori estimate). *From (P1) we also have the estimate*

$$\min J_\varepsilon^M \leq \int_\Omega -F^*(-\ell(\mathrm{diam}(\Omega)/\varepsilon))\,\mathrm{d}\mu \leq \mu(\Omega) \cdot (-F^*(-\ell(\mathrm{diam}(\Omega)/\varepsilon))),$$

whose right-hand side may be further bounded by the potentially infinite $\mu(\Omega)F(0) = W(\mu, 0)$ (the latter bound is directly obtained by choosing $v = 0$ as a quantization candidate in (4.1)).

Let now $(\varepsilon_M)_{M\in\mathbb{N}}$ be a positive, decreasing sequence of scaling factors. We use Theorems 4.7 and 4.10 to study the asymptotic quantization behaviour of the sequence of functionals $(J_{\varepsilon_M}^M)_M$ as $M \to \infty$ for a non-uniform mass distribution μ with Lipschitz Lebesgue density m. We identify three different regimes, depending on the behaviour of the sequence $\varepsilon_M^2 M$ (the quantity $\varepsilon_M^2 M$ indicates something like the average point density). A corresponding numerical illustration for the case of constant average point density is provided in Figure 10.

Theorem 4.12 (Asymptotic quantization). *Let $\Omega \subset \mathbb{R}^2$ be a closed Lipschitz domain (a domain whose boundary is locally the graph of a Lipschitz function with the domain lying on one side) and $\mu = m \cdot$*

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

$(\mathcal{L}\llcorner\Omega)$ for $\mathcal{L}$ the Lebesgue measure and $m:\Omega\to[0,\infty)$ a Lipschitz-continuous density. Let (P1)-(P4) hold. For any sequence $\varepsilon_1,\varepsilon_2,\ldots>0$ with $\varepsilon_M\searrow 0$ as $M\to\infty$ the following holds:

1. *If $\lim_{M\to\infty}\varepsilon_M^2 M=\infty$, then $\lim_{M\to\infty}\min J^M_{\varepsilon_M}=-F^*(-\ell(0))\cdot\mu(\Omega)$, which under (P5) equals 0.*
2. *If $\lim_{M\to\infty}\varepsilon_M^2 M=0$, then $\lim_{M\to\infty}\min J^M_{\varepsilon_M}=\mu(\Omega)\lim_{s\to\infty}-F^*(-\ell(s))$, which under (P6) equals $\mu(\Omega)F(0)=W(\mu,0)$.*
3. *If $\lim_{M\to\infty}\varepsilon_M^2 M=P\in(0,\infty)$, then*

$$\lim_{M\to\infty}\min J^M_{\varepsilon_M}=\left[\kappa\mapsto\int_\Omega B^*(\kappa/m(x))\,\mathrm{d}\mu(x)\right]^*(P),$$

where B:$(-\infty,\infty)\to(0,\infty]$,

$$B(z)=z\cdot\int_{H(1/z)}-F^*(-\ell(d(x,0)))\,\mathrm{d}x\quad\text{for } z>0,$$

$B(0)=\lim_{s\to\infty}-F^*(-\ell(s))$, and $B(z)=\infty$ for $z<0$. Furthermore, there exists a unique constant $\lambda<0$ and some measurable function $D:\Omega\to[0,\infty)$ such that

$$\lim_{M\to\infty}\min J^M_{\varepsilon_M}=\int_\Omega B(D(x))\,\mathrm{d}\mu(x),$$

and

$$D(x)\in\partial B^*(\lambda/m(x))\text{ for almost all }x\in\Omega,\qquad P=\int_\Omega D(x)\,\mathrm{d}x \tag{4.14}$$

(by convention, for $m(x)=0$ we set $D(x)=0$). That is, D can be interpreted as (being proportional to) the spatially varying point density of the asymptotically optimal local triangular grid.

Remark 4.13 (Limit cases). Theorem 4.12 *(1) and (2) can in fact be recovered as the special cases $P=\infty$ and $P=0$ of* Theorem 4.12 *(3) if we set $(\lambda,D)\equiv(0,\infty)$ or $(\lambda,D)\equiv(-\infty,0)$, respectively. However, it is simpler to treat them separately.*

Before stating a few more remarks and proving the asymptotic result, we analyse the function B, which represents the cell problem of quantizing a hexagon by a single Dirac mass.

Lemma 4.14 (Properties of the cell problem). *Assume (P1)-(P4). On $(0,\infty)$, the function B from Theorem 4.12 is finite, positive, decreasing and convex with continuous derivative*

$$B'(z)=\frac{1}{z}\left[B(z)-\frac{1}{|\partial H(1/z)|}\int_{\partial H(1/z)}-F^*(-\ell(d(x,0)))\,\mathrm{d}x\right]=:G(z).$$

Further, $B(z)\to B(0)$ as $z\searrow 0$, $G(z)\to 0$ as $z\to\infty$, and there exists some $Z\geq 0$ such that G is constant on $(0,Z]$ and strictly increasing on (Z,∞), where $Z>0$ if $-F\circ(-\ell)$ achieves a maximum. With $r=\lim_{z\searrow Z}B'(z)\in[-\infty,0)$ we can summarize $r<G(z)<0$ for $z>Z$ and

$$\partial B(z)=\begin{cases}\emptyset & \text{for } z<0,\\ (-\infty,r] & \text{for } z=0,\\ \{r\} & \text{for } z\in(0,Z],\\ \{G(z)\} & \text{for } z>Z,\end{cases}\qquad \partial(B^*)(s)=\begin{cases}\{0\} & \text{for } s<r,\\ [0,Z] & \text{for } s=r,\\ \{G^{-1}(s)\} & \text{for } s\in(r,0),\\ \emptyset & \text{for } s\geq 0.\end{cases}$$

Example 4.15 (Balanced quantization). *We consider the case of the standard Wasserstein-2 distance, where $\ell(t)=t^2$, $F^*(z)=z$ and $F(0)=\infty$. Then for $z>0$,*

$$B(z)=z\int_{H(1/z)}|x|^2\mathrm{d}x=\frac{5\sqrt{3}}{54}\frac{1}{z},\qquad B'(z)=G(z)=-\frac{5\sqrt{3}}{54}\frac{1}{z^2}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

and so $Z = 0$, $r = -\infty$. *For* $s < 0$,

$$B^*(s) = -2\sqrt{-\frac{5\sqrt{3}}{54}\,s}, \qquad (B^*)'(s) = \sqrt{-\frac{5\sqrt{3}}{54}\frac{1}{s}} = G^{-1}(s).$$

Remark 4.16 (Bounds in terms of cell problem). $B(z)$ *can be interpreted as the energy density associated with a regular triangular lattice with point density* z *(that is, each Voronoi cell occupies an area of* $1/z$*). The energy of such a lattice with* M *cells with total area* $|\Omega|$ *will be given by* $B(M/|\Omega|)\cdot|\Omega|$. *Taking into account the scaling factor* ε, *we can restate the bounds* (4.12) *and* (4.13) *as*

$$\min_{x_1,\dots,x_M\in\Omega} J_\varepsilon^M(x_1,\dots,x_M) \geq |\Omega|\cdot B\left(\frac{\varepsilon^2 M}{|\Omega|}\right) \qquad \text{and}$$

$$J_\varepsilon^M(x_1,\dots,x_M) \leq |\Omega|\cdot B\left(\frac{\varepsilon^2 M}{|\Omega|}\right) - |\partial\Omega|\sqrt{\frac{8|\Omega|}{3\sqrt{3}M}}\,F^*\left(-\ell\left(\sqrt{\frac{C_Q\,|\Omega|}{\varepsilon^2 M}}\right)\right).$$

Proof of Lemma 4.14. By (P4), $B(z) > 0$ for $z > 0$. Likewise, $B(z) < \infty$ by (P1) and (P3). Now observe that B yields the average value of $-F^*(-\ell(d(\cdot,0)))$ over $H(1/z)$. Hence, the monotonicity (P1) of $-F^*\circ(-\ell)$ implies that B is decreasing. Further,

$$\lim_{z\searrow 0} B(z) = \lim_{z\searrow 0}\int_{H(1)} -F^*(-\ell(d(y/\sqrt{z},0)))\,\mathrm{d}y = -F^*(-\ell(\infty)) = B(0).$$

Since $-F^*\circ(-\ell)$ is continuous by (P2), the integral in the definition of B is differentiable with respect to z by the Leibniz integral rule, and we have

$$\begin{aligned} B'(z) &= \frac{\mathrm{d}}{\mathrm{d}z}\left[z\int_{H(1/z)} -F^*(-\ell(d(x,0)))\,\mathrm{d}x\right] \\ &= \int_{H(1/z)} -F^*(-\ell(d(x,0)))\,\mathrm{d}x + z\int_{\partial H(1/z)} -F^*(-\ell(d(x,0)))v_n(z)\,\mathrm{d}x\cdot\left(-\frac{1}{z^2}\right), \end{aligned}$$

where $v_n(z) = 1/|\partial H(1/z)|$ is the normal velocity of the hexagonal boundary as the area of the hexagon is increased at rate 1. This coincides with the formula provided in the statement. To check convexity, we first assume that $-F^*\circ(-\ell)$ is differentiable. In the following we use the notation

$$⨍_{\partial H(1/z)} f\,\mathrm{d}x = \frac{1}{|\partial H(1/z)|}\int_{\partial H(1/z)} f\,\mathrm{d}x$$

and calculate

$$\begin{aligned} B''(z) &= \\ &= -\frac{1}{z^2}\left[B(z) - ⨍_{\partial H(1/z)} -F^*(-\ell(d(x,0)))\,\mathrm{d}x\right] \\ &\quad + \frac{1}{z}\left[\frac{1}{z}\left(B(z) - ⨍_{\partial H(1/z)} -F^*(-\ell(d(x,0)))\,\mathrm{d}x\right) - \frac{\mathrm{d}}{\mathrm{d}z}\left(⨍_{\partial H(1/z)} -F^*(-\ell(d(x,0)))\,\mathrm{d}x\right)\right] \\ &= -\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\left(⨍_{\partial H(1/z)} -F^*(-\ell(d(x,0)))\,\mathrm{d}x\right) \\ &\geq 0 \end{aligned}$$

since $-F^*\circ(-\ell)$ is nondecreasing. Therefore B is convex. By (P1) there exists some $R\in(0,\infty]$ such that $-F^*\circ(-\ell)$ is strictly increasing on $[0,R)$ and constant on $[R,\infty)$. Thus we see $B'' > 0$ on (Z,∞) for some $Z \geq 0$ and $B''(z) = 0$ for $z < Z$. The monotonicity properties of B' without assuming differentiability of $-F^*\circ(-\ell)$ now follow by a standard approximation argument. Note that positivity and

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

monotonicity of B imply $B'(z) \to 0$ as $z \to \infty$. We leave it as an easy exercise in convex analysis to check the expressions for the subdifferentials ∂B and $\partial(B^*)$. □

Remark 4.17 (Calculation of asymptotic density). *Given a density m, the asymptotically optimal point density D can be computed numerically based on the function B using*

$$D(x) \in \partial B^*(\lambda/m(x)) = \begin{cases} \{0\} & \text{if } \lambda/m(x) < r, \\ [0, Z] & \text{if } \lambda/m(x) = r, \\ \{(B')^{-1}(\lambda/m(x))\} & \text{if } \lambda/m(x) \in (r, 0), \\ \emptyset & \text{otherwise,} \end{cases}$$

where r was defined in Lemma 4.14.

Example 4.18 (Zador's Theorem is a special case of Theorem 4.12). *We show that Zador's Theorem [34, 71] in two dimensions (see equation (1.6) with $d = 2$) can be recovered from Theorem 4.12 by taking $\ell(t) = t^p$ and $F(s) = \iota_{\{1\}}(s)$. In this case, $\min J_{\varepsilon_M}$ is just the standard (balanced) optimal quantization error with respect to the Wasserstein-p distance. Note that $F(0) = +\infty$ but the transport cost ℓ is finite and so assumption (4.2) is satisfied. We have $F^*(z) = z$ and*

$$B(z) = c_6(p) z^{-\frac{p}{2}} \quad \text{for } z > 0 \text{ and } +\infty \text{ otherwise,}$$

where

$$c_6(p) = \int_{H(1)} |y|^p \, \mathrm{d}y.$$

Therefore, for $z > 0$, $s < 0$,

$$B'(z) = -\frac{p}{2} c_6(p) z^{-\frac{p}{2}-1}, \qquad (B')^{-1}(s) = \left(-\frac{p c_6(p)}{2s}\right)^{\frac{2}{p+2}}.$$

Assume that we are in Regime 3: $\lim_{M\to\infty} \varepsilon_M^2 M = P$. Note that B' is nowhere constant and so $Z = 0$ and $r = \lim_{z\to 0} B'(z) = -\infty$. By Remark 4.17, if $m(x) > 0$,

$$D(x) = (B')^{-1}(\lambda/m(x)) = \left(-\frac{p c_6(p) m(x)}{2\lambda}\right)^{\frac{2}{p+2}} \tag{4.15}$$

where $\lambda < 0$ is the constant given by Theorem 4.12. Then

$$P = \int_\Omega D(x) \, \mathrm{d}x = \left(-\frac{p c_6(p)}{2\lambda}\right)^{\frac{2}{p+2}} \int_\Omega m(x)^{\frac{2}{p+2}} \, \mathrm{d}x. \tag{4.16}$$

We can eliminate λ from (4.15) and (4.16) to write D in terms of P:

$$D(x) = P m^{\frac{2}{p+2}} \left(\int_\Omega m(x)^{\frac{2}{p+2}} \, \mathrm{d}x\right)^{-1}.$$

Therefore, by Theorem 4.12,

$$\begin{aligned} \lim_{M\to\infty} \varepsilon_M^{-p} \min_{\{x_i\}_{i=1}^M} \sum_{i=1}^M \int_{V_i} |x - x_i|^p \, m(x) \mathrm{d}x &= \lim_{M\to\infty} \min J^M_{\varepsilon_M} \\ &= \int_\Omega B(D(x)) \, m(x) \mathrm{d}x \\ &= P^{-\frac{p}{2}} c_6(p) \left(\int_\Omega m(x)^{\frac{2}{p+2}} \, \mathrm{d}x\right)^{\frac{p+2}{2}}. \end{aligned}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

Since $\lim_{M\to\infty} \varepsilon_M^2 M = P$, *we can rewrite this as*

$$\lim_{M\to\infty} M^{\frac{p}{2}} \min_{\{x_i\}_{i=1}^M} \sum_{i=1}^M \int_{V_i} |x - x_i|^p \, m(x)\mathrm{d}x = c_6(p) \left(\int_\Omega m(x)^{\frac{2}{p+2}} \, \mathrm{d}x \right)^{\frac{p+2}{2}}$$

which is exactly Zador's Theorem in two dimensions; see equations (1.6) *and* (1.7).

Proof of Theorem 4.12. **Regime 1:** $\lim_{M\to\infty} \varepsilon_M^2 M = \infty$**.** We can pick a sequence of radii r_M such that $r_M/\varepsilon_M \to 0$, but still $r_M^2 M \to \infty$ (for instance pick $r_M = \varepsilon_M^{1/2}/M^{1/4}$). Since Ω is a Lipschitz domain, for M large and thus r_M small enough, Ω can be covered with $K_M \le 2|\Omega|/r_M^2$ squares of side length r_M (for instance cover Ω with a regular square grid of lattice spacing r_M). Now position the points $x_1, \dots, x_M$ arbitrarily with the only condition that each square contains at least one point (this is possible for M large enough, since $K_M/M \to 0$ as $M \to \infty$). Obviously, Ω is covered by the balls $B_{2r_M}(x_i)$ of radius $2r_M$ centred at the x_i, as $B_{2r_M}(x_i)$ includes the whole square containing x_i. Thus, we find $V_i(x_1, \dots, x_M) \subset B_{2r_M}(x_i)$ for $i = 1, \dots, M$. Then

$$\min J_{\varepsilon_M}^M \le J_{\varepsilon_M}^M(x_1, \dots, x_M) = \sum_{i=1}^M \int_{V_i(x_1,\dots,x_M)} -F^*\left(-\ell\left(\frac{d(x,x_i)}{\varepsilon_M}\right)\right) \mathrm{d}\mu(x)$$

$$\le -F^*\left(-\ell\left(\frac{2r_M}{\varepsilon_M}\right)\right) \cdot \mu(\Omega) \to -F^*(-\ell(0)) \cdot \mu(\Omega) \quad \text{as } M \to \infty$$

since $[0,\infty) \ni z \mapsto -F^*(-\ell(z))$ is continuous in $z = 0$ by (P2). $J_{\varepsilon_M}^M \ge -F^*(-\ell(0)) \cdot \mu(\Omega)$ follows from (P1), the monotonicity of $-F^* \circ (-\ell)$.

Regime 2: $\lim_{M\to\infty} \varepsilon_M^2 M = 0$**.** Remark 4.11 yields $\min J_{\varepsilon_M}^M \le \mu(\Omega) \cdot \lim_{s\to\infty} -F^*(-\ell(s))$ (which may be infinite). Let now $r_1, r_2, \dots$ be a positive sequence such that $r_M^2 \cdot M \to 0$ and $r_M/\varepsilon_M \to \infty$ as $M \to \infty$. Let $x_1, \dots, x_M$ be arbitrary distinct points in Ω and set $S = \Omega \cap \bigcup_{i=1}^M B_{r_M}(x_i)$. Note that, since $r_M^2 \cdot M \to 0$ and $\mu \ll \mathcal{L}$, $\mu(S) \to 0$ as $M \to \infty$. Clearly,

$$\min_{i\in\{1,\dots,M\}} -F^*\left(-\ell\left(\frac{d(x,x_i)}{\varepsilon_M}\right)\right) \ge \begin{cases} -F^*\left(-\ell\left(\frac{r_M}{\varepsilon_M}\right)\right) & \text{for } x \in \Omega \setminus S, \\ 0 & \text{for } x \in S. \end{cases}$$

Therefore,

$$J_{\varepsilon_M}^M(x_1, \dots, x_M) \ge -F^*\left(-\ell\left(\frac{r_M}{\varepsilon_M}\right)\right) \cdot \mu(\Omega \setminus S) \to \lim_{s\to\infty} -F^*(-\ell(s)) \cdot \mu(\Omega) \quad \text{as } M \to \infty.$$

Regime 3: $\lim_{M\to\infty} \varepsilon_M^2 M = P \in (0,\infty)$**. Part 3.1: lower bound.** For $\delta \in (0,1)$ cover Ω by a regular grid of squares with edge length $\delta + \delta^2$. Denote by $\{S_i\}_{i=1}^N$ the N squares that are fully contained in Ω. For each S_i denote by $\hat{S}_i$ the square of edge length δ that lies centered within S_i such that $\mathrm{dist}(\hat{S}_i, S_j) \ge \delta^2/2$ for $i \ne j$ (the boundary layers of width δ^2 will be necessary to control the interaction between neighbouring squares). Denote the union of all $\hat{S}_i$ by $\hat{S} = \bigcup_{i=1}^N \hat{S}_i$. Since Ω is a Lipschitz domain and $\mu \ll \mathcal{L}$, $\mu(\Omega \setminus \hat{S}) \to 0$ as $\delta \to 0$.

Let $x_1, \dots, x_M$ be M points from Ω and denote by M_i the number of points in a square S_i (points on the square boundaries are assigned to precisely one square). Obviously, $\sum_{i=1}^N M_i \le M$. Now pick an arbitrary $\eta > 0$ (we will later send $\eta \to \infty$; it has the role to regularize the integrand and thus the cell problem function B). Note that for $x \in \hat{S}_i$ and M large enough (depending on δ), we have

$$\min_j d(x, x_j) \ge \min\left\{ \min_{j: x_j \in S_i} d(x, \pi_{\hat{S}_i}(x_j)), \frac{\delta^2}{2} \right\} \ge \min\left\{ \min_{j: x_j \in S_i} d(x, \pi_{\hat{S}_i}(x_j)), \varepsilon_M \eta \right\}$$

where $\pi_{\hat{S}_i}$ denotes the orthogonal projection onto $\hat{S}_i$. By the nonnegativity and monotonicity of $-F^*(-\ell(\,\cdot\,))$, one thus obtains

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

$$J^M_{\varepsilon_M}(x_1,\dots,x_M) = \int_\Omega \min_j -F^*\left(-\ell\left(\frac{d(x,x_j)}{\varepsilon_M}\right)\right) m(x)\,\mathrm{d}x$$
$$\geq \sum_{i=1}^N m_i \int_{\hat S_i} \min_{j:x_j\in S_i} -F^*\left(-\ell\left(\min\left\{\frac{d(x,\pi_{\hat S_i}(x_j))}{\varepsilon_M},\eta\right\}\right)\right)\,\mathrm{d}x$$

with $m_i := \inf_{x\in S_i} m(x)$. Next introduce

$$\phi_\eta(t) = -F^*\left(-\ell(\min\{t,\eta\})\right).$$

One can readily verify that actually $\phi_\eta = -F^*_\eta \circ (-\ell)$, where F_η is the lower semi-continuous convex envelope of a modification of F,

$$F_\eta = \tilde F^{**} \qquad \text{with } \tilde F(s) = \begin{cases} F(s) & \text{if } s\neq 0,\\ -F^*\left(-\ell\left(\eta\right)\right) & \text{else.}\end{cases}$$

Let

$$\hat M_i = \#\{\pi_{\hat S_i}(x_j) : x_j \in S_i\} \leq M_i$$

be the number of points in S_i that are distinct after projecting onto $\hat S_i$. We can now apply Theorem 4.7 (lower bound for quantization of the Lebesgue measure) or Remark 4.16 on each square $\hat S_i$ separately (the boundary layers of width δ^2 are necessary to control the interaction between neighbouring squares),

$$J^M_{\varepsilon_M}(x_1,\dots,x_M) \geq \sum_{i=1}^N m_i \int_{\hat S_i} \min_{j:x_j\in S_i} \phi_\eta\left(d(x,\pi_{\hat S_i}(x_j))/\varepsilon_M\right)\,\mathrm{d}x$$
$$\geq \sum_{i=1}^N m_i\,\hat M_i \int_{H(|\hat S_i|/\hat M_i)} \phi_\eta\left(d(x,0)/\varepsilon_M\right)\,\mathrm{d}x$$
$$= \sum_{i=1}^N m_i\,|\hat S_i| B_\eta\left(\frac{\varepsilon_M^2\,\hat M_i}{|\hat S_i|}\right)$$

where B_η is defined as in Lemma 4.14, only with $-F^*\circ(-\ell)$ replaced by ϕ_η or equivalently F by F_η. Now set

$$E(x) := \begin{cases} \frac{\varepsilon_M^2\,\hat M_i}{|\hat S_i|} & \text{for } x\in\hat S_i,\\ 0 & \text{otherwise,}\end{cases}$$

and let L_m denote the Lipschitz constant of the density m. Then

$$\int_\Omega B_\eta(E(x))\,\mathrm{d}\mu(x) = \sum_{i=1}^N \int_{\hat S_i} B_\eta\left(\frac{\varepsilon_M^2\,\hat M_i}{\delta^2}\right) m(x)\,\mathrm{d}x + \int_{\Omega\setminus\hat S} B_\eta(0) m(x)\,\mathrm{d}x$$
$$= \sum_{i=1}^N \int_{\hat S_i} B_\eta\left(\frac{\varepsilon_M^2\,\hat M_i}{\delta^2}\right) m_i\,\mathrm{d}x$$
$$+ \sum_{i=1}^N \int_{\hat S_i} B_\eta\left(\frac{\varepsilon_M^2\,\hat M_i}{\delta^2}\right)\left(m(x)-m_i\right)\,\mathrm{d}x + \mu(\Omega\setminus\hat S)\cdot B_\eta(0)$$
$$\leq \sum_{i=1}^N m_i|\hat S_i| B_\eta\left(\frac{\varepsilon_M^2\,\hat M_i}{\delta^2}\right) + L_m\cdot\sqrt{2}\,\delta\cdot B_\eta(0)\cdot|\Omega| + \mu(\Omega\setminus\hat S)\cdot B_\eta(0)$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

where in the last step we used $B_\eta(z) \le B_\eta(0) < \infty$ for $z \ge 0$. Abbreviate the last two summands (that do not depend on M) as $C_\delta B_\eta(0)$ and note that $C_\delta \to 0$ as $\delta \to 0$, then in summary we have arrived at

$$J^M_{\varepsilon_M}(x_1, \dots, x_M) \ge \int_\Omega B_\eta(E(x))\, \mathrm{d}\mu(x) - C_\delta B_\eta(0). \tag{4.17}$$

The function E satisfies $\int_\Omega E(x)\,\mathrm{d}x = \varepsilon_M^2 \sum_{i=1}^N \hat{M}_i \le \varepsilon_M^2 M$. By minimizing over all such functions, we thus obtain a lower bound for the minimum,

$$\min J^M_{\varepsilon_M} \ge \inf \left\{ \int_\Omega B_\eta(E(x))\, m(x)\,\mathrm{d}x \;\middle|\; E \in L^1(\Omega;[0,\infty)), \int_\Omega E(x)\,\mathrm{d}x \le \varepsilon_M^2 \cdot M \right\} - C_\delta B_\eta(0).$$

Since B_η is nonincreasing, the estimate can be rewritten as

$$\min J^M_{\varepsilon_M} \ge \inf \left\{ \int_\Omega B_\eta(E(x))\, m(x)\,\mathrm{d}x \;\middle|\; E \in L^1(\Omega;[0,\infty)), \int_\Omega E(x)\,\mathrm{d}x = \varepsilon_M^2 \cdot M \right\} - C_\delta B_\eta(0).$$

Now denote by L_B the Lipschitz constant of B_η on $[0,\infty)$ (which exists by Lemma 4.14). If $E \in L^1(\Omega;[0,\infty))$ satisfies $\int_\Omega E(x)\,\mathrm{d}x = \varepsilon_M^2 \cdot M$, then $\tilde{E} := \frac{P}{(\varepsilon_M^2 M)} E$ satisfies $\int_\Omega \tilde{E}(x)\,\mathrm{d}x = P$ and

$$\begin{aligned}\left| \int_\Omega B_\eta(\tilde{E}(x))\,\mathrm{d}\mu(x) - \int_\Omega B_\eta(E(x))\,\mathrm{d}\mu(x) \right| &\le L_B \int_\Omega |E(x) - \tilde{E}(x)|\,\mathrm{d}\mu(x) \\ &\le L_B |\varepsilon_M^2 M - P| \max_{x\in\Omega} m(x) =: c_M,\end{aligned}$$

which converges to zero as $M \to \infty$. Summarizing, we obtain

$$\min J^M_{\varepsilon_M} \ge \inf \left\{ \int_\Omega B_\eta(E(x))\,\mathrm{d}\mu(x) \;\middle|\; E \in L^1(\Omega;[0,\infty)), \int_\Omega E(x)\,\mathrm{d}x = P \right\} - C_\delta B_\eta(0) - c_M.$$

Now let first $M \to \infty$ and then $\delta \to 0$. Introducing a Lagrange multiplier λ for the constraint on E, we thus find

$$\begin{aligned}\liminf_{M\to\infty} \min J^M_{\varepsilon_M} &\ge \inf_{E\in L^1(\Omega;[0,\infty))} \sup_{\lambda\in\mathbb{R}} \int_\Omega \left[B_\eta(E(x))\, m(x) - \lambda\, E(x) \right] \mathrm{d}x + \lambda \cdot P \\ &\ge \sup_{\lambda\in\mathbb{R}} \inf_{E\in L^1(\Omega;[0,\infty))} \int_\Omega \left[B_\eta(E(x))\, m(x) - \lambda\, E(x) \right] \mathrm{d}x + \lambda \cdot P \\ &\ge \sup_{\lambda\in\mathbb{R}} \int_\Omega \inf_{E\ge 0} \left[B_\eta(E)\, m(x) - \lambda\, E \right] \mathrm{d}x + \lambda \cdot P \\ &= \sup_{\lambda\in\mathbb{R}} \int_\Omega \left[-m(x) \cdot B_\eta^*(\lambda/m(x)) \right] \mathrm{d}x + \lambda \cdot P.\end{aligned}$$

Now as $\eta \to \infty$ one has the pointwise convergence $\phi_\eta \nearrow -F^* \circ (-\ell)$ and thus $B_\eta \nearrow B$ pointwise by the monotone convergence theorem. In fact, $\phi_\eta(r) = -F^*(-\ell(r))$ for all $r < \eta$ and therefore $B_\eta(z) = B(z)$ for all $z > 2/(3\sqrt{3}\eta^2)$. Consequently, also $B_\eta^* \searrow B^*$ monotonously so that one obtains sharper bounds with increasing η. Thus, using again the monotone convergence theorem,

$$\begin{aligned}\liminf_{M\to\infty} \min J^M_{\varepsilon_M} &\ge \sup_{\eta>0} \sup_{\lambda\in\mathbb{R}} \int_\Omega \left[-m(x) \cdot B_\eta^*(\lambda/m(x)) \right] \mathrm{d}x + \lambda \cdot P \\ &= \sup_{\lambda\in\mathbb{R}} \sup_{\eta>0} \int_\Omega \left[-m(x) \cdot B_\eta^*(\lambda/m(x)) \right] \mathrm{d}x + \lambda \cdot P \\ &= \sup_{\lambda\in\mathbb{R}} \int_\Omega \left[-m(x) \cdot B^*(\lambda/m(x)) \right] \mathrm{d}x + \lambda \cdot P.\end{aligned}$$

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

By evaluating the supremum over λ, we finally obtain

$$\liminf_{M\to\infty} \min J^M_{\varepsilon_M} \geq \left[\kappa \mapsto \int_\Omega B^*(\kappa/m(x))\,\mathrm{d}\mu(x)\right]^*(P)\,.$$

Part 3.2: optimal limit density. Observe that $B^*(z)=\infty$ for $z>0$, B^* is convex, lower semi-continuous, nondecreasing, satisfies $\lim_{z\to-\infty} B^*(z)/|z|=0$ (which follows from $s\cdot z - B^*(z)\leq B(s)<\infty$ for all $s>0$, $z<0$) and has infinite left derivative at 0 (by Lemma 4.14). Therefore the map

$$\mathbb{R}\ni\lambda\mapsto \int_\Omega \left[-m(x)\cdot B^*(\lambda/m(x))\right]\,\mathrm{d}x + \lambda\cdot P$$

is concave and there exists a maximizing $\lambda<0$ satisfying the necessary and sufficient optimality condition

$$0\in\partial\left[\kappa\mapsto\int_\Omega m(x)\cdot B^*\left(\frac{\kappa}{m(x)}\right)\,\mathrm{d}x-\kappa\cdot P\right](\lambda) \quad\Longleftrightarrow$$
$$P\in\partial\left[\kappa\mapsto\int_\Omega m(x)B^*\left(\frac{\kappa}{m(x)}\right)\,\mathrm{d}x\right](\lambda).$$

We next aim to find a function $D_\xi:\Omega\to[0,\infty)$ satisfying $D_\xi(x)\in\partial B^*(\lambda/m(x))$ as well as $\int_\Omega D_\xi(x)\,\mathrm{d}x=P$. To this end, recall Z and r from Lemma 4.14 and define

$$D_\xi(x)=\begin{cases}(B^*)'(\lambda/m(x)) & \text{if } m(x)>\lambda/r,\\ \xi\in[0,Z] & \text{if } m(x)=\lambda/r,\\ 0 & \text{otherwise.}\end{cases}$$

Since m is continuous, its superlevel sets are Lebesgue measurable. D_ξ is constructed by assigning new values to the level sets of m in a monotone way (by the monotonicity of $(B^*)'$, see Lemma 4.14); hence it is also Lebesgue measurable. We now pick $\xi(P)\in[0,Z]$ such that $\int_\Omega D_{\xi(P)}(x)\,\mathrm{d}x=P$. (Note that necessarily $\xi(P)=0$ if $Z=0$, and the choice of ξ is irrelevant if $\{x\in\Omega\mid m(x)=\lambda/r\}$ is a nullset. The following argument still applies in these cases.) Such a $\xi(P)$ exists due to $\int_\Omega D_0(x)\,\mathrm{d}x\leq P$ and $\int_\Omega D_Z(x)\,\mathrm{d}x\geq P$, as we now show. Indeed, note by Lemma 4.14 that for all $x\in\Omega$ the function $D_0(x)$ equals the left derivative of B^* at $\lambda/m(x)$ (which by convention shall be 0 for $m(x)=0$), while $D_Z(x)$ equals the right derivative. Beppo Levi's monotone convergence theorem thus yields

$$\begin{aligned}\int_\Omega D_0(x)\,\mathrm{d}x &= \int_\Omega \lim_{\tilde\lambda\nearrow\lambda}\frac{m(x)B^*\left(\frac{\tilde\lambda}{m(x)}\right)-m(x)B^*\left(\frac{\lambda}{m(x)}\right)}{\tilde\lambda-\lambda}\,\mathrm{d}x\\ &= \lim_{\tilde\lambda\nearrow\lambda}\frac{\displaystyle\int_\Omega m(x)B^*\left(\frac{\tilde\lambda}{m(x)}\right)\,\mathrm{d}x-\int_\Omega m(x)B^*\left(\frac{\lambda}{m(x)}\right)\,\mathrm{d}x}{\tilde\lambda-\lambda}\\ &\leq P \qquad (4.18)\end{aligned}$$

since $P\in\partial\left[\kappa\mapsto\int_\Omega m(x)B^*\left(\frac{\kappa}{m(x)}\right)\,\mathrm{d}x\right](\lambda)$ and since (4.18) is the left derivative of $\lambda\mapsto\int_\Omega m(x)B^*\left(\frac{\lambda}{m(x)}\right)\,\mathrm{d}x$. The inequality $\int_\Omega D_Z(x)\,\mathrm{d}x\geq P$ follows analogously. Writing $D=D_{\xi(P)}$, we finally obtain (4.14) and

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

$$\lim_{M\to\infty} \min J^M_{\varepsilon_M} \geq \int_\Omega -m(x)B^*(\lambda/m(x))\,\mathrm{d}x + \lambda P$$
$$= \int_\Omega \frac{\lambda}{m(x)} D(x) - B^*(\lambda/m(x))\,\mathrm{d}\mu(x)$$
$$= \int_\Omega B(D(x))\,\mathrm{d}\mu(x)\,,$$

where the last equality follows from the Moreau–Fenchel identity [6, Prop. 16.9], which states that $B(s) + B^*(t) = st \iff s \in \partial B^*(t) \iff t \in \partial B(s)$.

Part 3.3: upper bound. Finally, we derive the corresponding upper bound, essentially by constructing a piecewise triangular point configuration with point density approximating the expected point density D from the above lower bound proof. Fix $M \in \mathbb{N}$, $\delta > 0$, and $\eta \in (0, 1)$. Denote the 1-neighbourhood of Ω by $\underline{\Omega} = \Omega + B_1(0)$ and extend the mass density m and the expected point density D to $\underline{\Omega} \setminus \Omega$ by zero. Cover Ω with a tessellation of squares of side length δ (we will later send $\delta \to 0$). We keep the squares $\{S_i\}_{i=1}^{N_\delta}$ that intersect Ω. We may assume δ to be small enough such that all squares lie within $\underline{\Omega}$.

Define $D_\eta : \underline{\Omega} \to (0, \infty)$ to be a slight modification of the expected point density D,

$$D_\eta = (1-\eta)D + \eta \frac{P}{2|\underline{\Omega}|}.$$

The main role of the regularization parameter η is to ensure that we distribute particles throughout the whole domain Ω, even in regions where $m = 0$. We will send $\eta \to 0$ at the very end of the proof. For $i \in \{1, \ldots, N_\delta\}$, define the point number $M_i = M_i(M, \delta, \eta)$ associated with square S_i by

$$M_i = \left\lceil \frac{1}{\varepsilon_M^2} \int_{S_i} D_\eta(x)\,\mathrm{d}x \right\rceil.$$

Note that

$$\sum_{i=1}^{N_\delta} M_i \leq M \quad \text{if } M \text{ is sufficiently large.}$$

Indeed, we have

$$\varepsilon_M^2 \sum_{i=1}^{N_\delta} M_i \leq \varepsilon_M^2 \sum_{i=1}^{N_\delta} \left(\frac{1}{\varepsilon_M^2} \int_{S_i} D_\eta(x)\,\mathrm{d}x + 1 \right)$$
$$< \left((1-\eta) \int_\Omega D(x)\,\mathrm{d}x + \eta \frac{P}{2} \right) + \varepsilon_M^2 N_\delta = \left(1 - \frac{\eta}{2}\right) P + \varepsilon_M^2 N_\delta$$

so that

$$\varepsilon_M^2 \left(\sum_{i=1}^{N_\delta} M_i - M \right) < \left(1 - \frac{\eta}{2}\right) P - \varepsilon_M^2 M + \varepsilon_M^2 N_\delta \xrightarrow[M\to\infty]{} -\frac{\eta}{2} P < 0.$$

Next, on each S_i we choose M_i quantization points as in Theorem 4.10 (upper bound for quantization of the Lebesgue measure) and distribute the remaining $M - \sum_{i=1}^{N_\delta} M_i$ points arbitrarily in Ω (which does not increase the cost). By Remark 4.16 we thus obtain

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

$$J^M_{\varepsilon_M}(x_1,\dots,x_M) \le \sum_{i=1}^{N_\delta} \int_{S_i} \min_{k\in\{1,\dots,M\}:x_k\in S_i} -F^*\left(-\ell\left(\frac{d(x,x_k)}{\varepsilon_M}\right)\right) \mathrm{d}\mu(x)$$

$$\le \sum_{i=1}^{N_\delta} \max_{S_i} m \int_{S_i} \min_{k\in\{1,\dots,M\}:x_k\in S_i} -F^*\left(-\ell\left(\frac{d(x,x_k)}{\varepsilon_M}\right)\right) \mathrm{d}x$$

$$\le \sum_{i=1}^{N_\delta} \max_{S_i} m \left[|S_i| \cdot B\left(\frac{\varepsilon_M^2 M_i}{|S_i|}\right) - |\partial S_i| \sqrt{\frac{8|S_i|}{3\sqrt{3}M_i}}\, F^*\left(-\ell\left(\sqrt{\frac{C_Q\,|S_i|}{\varepsilon_M^2 M_i}}\right)\right)\right].$$

Exploiting the continuity of B and finiteness and monotonicity of $-F^*\circ(-\ell)$ on $(0,\infty)$ as well as

$$\lim_{M\to\infty} \varepsilon_M^2 M_i = \int_{S_i} D_\eta(x)\,\mathrm{d}x > 0$$

we thus obtain

$$\limsup_{M\to\infty} J^M_{\varepsilon_M}(x_1,\dots,x_M) \le \sum_{i=1}^{N_\delta} \max_{S_i} m\,|S_i| \cdot B\left(\frac{1}{|S_i|}\int_{S_i} D_\eta(x)\,\mathrm{d}x\right).$$

Define $E^\delta : \underline{\Omega} \to [0,\infty)$, $m^\delta : \underline{\Omega} \to [0,\infty)$ by

$$E^\delta(x) = \frac{1}{|S_i|}\int_{S_i} D_\eta(x)\,\mathrm{d}x \quad \text{if } x\in S_i \quad \text{and } E^\delta(x) = \frac{\eta P}{2|\underline{\Omega}|} \text{ else},$$

$$m^\delta(x) = \max_{S_i} m \quad \text{if } x\in S_i \quad \text{and } m^\delta(x) = 0 \text{ else}.$$

Then we can rewrite the previous inequality as follows:

$$\limsup_{M\to\infty} \min J^M_{\varepsilon_M} \le \limsup_{M\to\infty} J^M_{\varepsilon_M}(x_1,\dots,x_M) \le \int_{\underline{\Omega}} B\left(E^\delta(x)\right) m^\delta(x)\,\mathrm{d}x. \tag{4.19}$$

Next we pass to the limits $\delta\to 0$ and $\eta\to 0$ in that order. By the Lebesgue Differentiation Theorem, $\lim_{\delta\to 0} E^\delta = D_\eta$ pointwise almost everywhere. Since m is upper semi-continuous, then $\lim_{\delta\to 0} m^\delta = m$ pointwise. Moreover, $E^\delta \ge \eta P/2|\underline{\Omega}|$ and B is nonincreasing on $(0,\infty)$. Hence

$$B\left(E^\delta(x)\right) m^\delta(x) \le B\left(\frac{\eta P}{2|\underline{\Omega}|}\right) \max_{\Omega} m.$$

Therefore, by the Dominated Convergence Theorem,

$$\lim_{\delta\to 0^+} \int_{\underline{\Omega}} B\left(E^\delta(x)\right) m^\delta(x)\,\mathrm{d}x = \int_{\Omega} B\left(D_\eta(x)\right) m(x)\,\mathrm{d}x. \tag{4.20}$$

Finally, by the convexity of B,

$$\int_{\Omega} B\left(D_\eta(x)\right) m(x)\,\mathrm{d}x \le (1-\eta)\int_{\Omega} B\left(D(x)\right) m(x)\,\mathrm{d}x + \eta B\left(\frac{P}{2|\underline{\Omega}|}\right)\int_{\Omega} m(x)\,\mathrm{d}x. \tag{4.21}$$

By taking the limits $\delta\to 0$, then $\eta\to 0$ in (4.19) and using (4.20)-(4.21), we obtain the matching upper bound

$$\limsup_{M\to\infty} \min J^M_{\varepsilon_M} \le \int_{\Omega} B\left(D(x)\right) m(x)\,\mathrm{d}x$$

as required. □

Remark 4.19 (Lipschitz condition). *Inspecting the proof we see that the Lipschitz condition on m can actually be replaced by mere continuity; then all estimates based on the Lipschitz constant have to be replaced using the modulus of continuity of m.*

Remark 4.20 (Γ-convergence for unbalanced quantization). *We conjecture that* Theorem 4.12 *can be expanded into a* Γ*-convergence result in the spirit of [62, Proposition 7.18] for balanced optimal*

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

transport. For given $M \in \mathbb{N}$, define the functional

$$\mathcal{J}^M(\nu) := \begin{cases} J^M_{\varepsilon_M}(x_1, \dots, x_M) & \text{if } \nu = \varepsilon_M^2 \sum_{i=1}^M \delta_{x_i}, \\ +\infty & \text{otherwise,} \end{cases}$$

and let the limit functional be given by

$$\mathcal{J}(\nu) := \begin{cases} \int_\Omega B\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mathcal{L}}(x)\right) \mathrm{d}\mu(x) & \text{if } \nu \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

In the definition of $\mathcal{J}$, the part of ν that is singular with respect to $\mathcal{L}$ is simply discarded. It seems plausible that $\mathcal{J}^M$ Γ-converges to $\mathcal{J}$ with respect to the weak topology. Theorem 4.12 *(2) describes the special case of this result where the limit measure ν is* 0. Theorem 4.12 *(3) describes the special case of $\nu = D \cdot \mathcal{L} \neq 0$ that are minimizing for a prescribed mass P. In the regime of* Theorem 4.12 *(1), the sequence of measures is diverging and thus not described by such a Γ-convergence.*

Part 3.1 of the proof of Theorem 4.12 *establishes the lim-inf inequality for minimizing sequences. Part 3.2 identifies the optimal density D of the limit functional $\mathcal{J}$. Part 3.3 essentially derives the lim-sup condition for limit measures $\nu = D \cdot \mathcal{L}$, i.e. without a Lebesgue-singular part. The optimality of D is not leveraged in this part of the proof, and it applies to general densities. A complete Γ-convergence result would require an extension of the first part to configurations $(x_i)_{i=1}^M$ that are not (approximately) minimal and to include the Lebesgue-singular part of ν in both the lower and upper bounds. Given that the proof for the lower bound is already rather technical, we choose to not provide this extension here.*

Remark 4.21 (Quantization regimes). *The proof shows that the set of near- or asymptotically optimal point distributions for $\lim_{M\to\infty} \varepsilon_M^2 M \in \{0, \infty\}$ is quite degenerate. Indeed, if the limit is zero, then arbitrarily placed points $x_1, \dots, x_M \in \Omega$ were shown to asymptotically achieve the optimal energy. The interpretation is that in the limit $M \to \infty$ no transport takes place between μ and its discrete quantization approximation so that the quantization energy equals the cost for changing mass distribution μ to zero. If on the other hand the limit is infinite, then Dirac masses can be distributed over Ω in such a dense fashion that all transport distances and thus transport costs become negligibly small. Thus, to achieve the asymptotic cost* 0*, it suffices for instance to have a more or less uniform distribution of $x_1, \dots, x_M \in \Omega$, but otherwise the point arrangement does not matter. The case $\lim_{M\to\infty} \varepsilon_M^2 M \in (0, \infty)$ seems to be more rigid; here the optimal asymptotic cost is achieved by a construction which locally looks like a triangular lattice.*

Example 4.22 (Hellinger–Kantorovich). *The function B from Lemma 4.14 and its derivative B' can be computed numerically for different unbalanced transport models; we here consider the Hellinger-Kantorovich setting. In this case, computing the integral just on one triangular segment of $H(\frac{1}{z})$, we obtain*

$$\begin{aligned} B(z) &= z\left(6\int_{-\pi/6}^{\pi/6}\int_0^{L(\alpha,z)} \sin^2\left(\min\left\{r, \frac{\pi}{2}\right\}\right) r \,\mathrm{d}r \,\mathrm{d}\alpha\right) \\ &= 3\int_{-\pi/6}^{\pi/6} z \max\left\{\frac{1}{4} + \frac{L(\alpha,z)^2}{2} - \frac{\cos\left(2L(\alpha,z)\right)}{4} - \frac{L(\alpha,z)\sin\left(2L(\alpha,z)\right)}{2}, \frac{1}{2} - \frac{\pi^2}{8} + L(\alpha,z)^2\right\} \mathrm{d}\alpha \end{aligned}$$

for $L(\alpha, z) = 1/(\sqrt{2\sqrt{3}z}\cos\alpha)$ the length of the ray starting from the hexagon centre at angle α. The resulting B' (computed numerically) is shown in Figure 11. Thus, for a given mass distribution $\mu = m\mathcal{L}\llcorner\Omega$, we can compute the asymptotically optimal point density D of the quantization problem from Theorem 4.12 and Remark 4.17. Figure 11 shows computed examples for such asymptotic densities. One can see that the variations of μ are reduced for large values of P, but amplified for small values of P (in particular, large areas of Ω have zero point density).

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

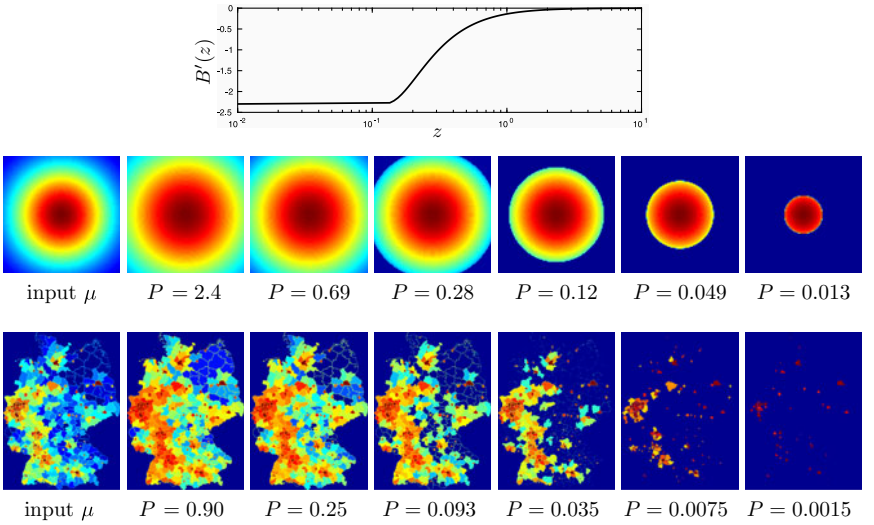


Figure 11. *Top row: B' from Lemma 4.14 for Hellinger–Kantorovich transport. Middle and bottom row: input distribution μ (a Gaussian and same data as in Figure 7) as well as asymptotically optimal point densities D for different values of $P = \lim_{M\to\infty} \varepsilon_M^2 M$ (colour-coding from blue for 0 to red for maximum value).*

Financial support. DPB acknowledges financial support from the Engineering and Physical Sciences Research Council (EPSRC) grants EP/R013527/1 Designer Microstructure via Optimal Transport Theory and EP/V00204X/1 Mathematical Theory of Polycrystalline Materials. BS was supported by the Emmy Noether programme of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 403056140. BW was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044-390685587, Mathematics Münster: Dynamics-Geometry-Structure.

Competing interests. None.

References

[1] Ambrosio, L., Fusco, N. & Pallara, D. (2000) *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press.
[2] Aurenhammer, F., Hoffmann, F. & Aronov, B. (1998) Minkowski-type theorems and least-squares clustering. *Algorithmica* **20** (1), 61–76.
[3] Aurenhammer, F., Klein, R. & Lee, D.-T. (2013) *Voronoi Diagrams and Delaunay Triangulations*, World Scientific Publishing Company.
[4] Aydın, A. D. & Iacobelli, M. (2025) Asymptotic quantization of measures on riemannian manifolds via covering growth estimates. *Adv. Math.* **474**, 110311.
[5] Barnes, E. S. & Sloane, N. J. A. (1983) The optimal lattice quantizer in three dimensions. *SIAM J. Algebraic Discrete Methods* **4** (1), 30–41.
[6] Bauschke, H. H. & Combettes, P. L. (2011) *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer.
[7] Bollobás, B. & Stern, N. (1972) The optimal structure of market areas. *J. Econ. Theory* **4** (2), 174–179.
[8] Bouchitté, G., Jimenez, C. & Mahadevan, R. (2011) Asymptotic analysis of a class of optimal location problems. *J. Math. Pures Appl.* **95** (4), 382–419.
[9] Bouchitté, G., Jimenez, C. & Rajesh, M. (2002) Asymptotique d'un problème de positionnement optimal. *C. R. Acad. Sci. Paris Ser. I Math.* **335**, 1–6.
[10] Bourne, D. P., Kok, P. J. J., Roper, S. M. & Spanjer, W. D. T. (2020) Laguerre tessellations and polycrystalline microstructures: A fast algorithm for generating grains of given volumes. *Phil. Mag.* **100** (21), 2677–2707.
[11] Bourne, D. P., Peletier, M. A. & Roper, S. M. (2014) Hexagonal patterns in a simplified model for block copolymers. *SIAM J. Appl. Math.* **74** (5), 1315–1337.
[12] Bourne, D. P., Peletier, M. A. & Theil, F. (2014) Optimality of the triangular lattice for a particle system with Wasserstein interaction. *Commun. Math. Phys.* **329** (1), 117–140.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

[13] Bourne, D. P. & Roper, S. M. (2015) Centroidal power diagrams, Lloyd's algorithm, and applications to optimal location problems. *SIAM J. Numer. Anal.* **53** (6), 2545–2569.

[14] Buttazzo, G. & Santambrogio, F. (2009) A mass transportation model for the optimal planning of an urban region. *SIAM Rev.* **51** (3), 593–610.

[15] Buze, M., Feydy, J., Roper, S. M., Sedighiani, K. & Bourne, D. P. (2024) Anisotropic power diagrams for polycrystal modelling: Efficient generation of curved grains via optimal transport. *Comput. Mater. Sci.* **245**, 113317.

[16] Caglioti, E., Golse, F. & Iacobelli, M. (2015) A gradient flow approach to quantization of measures. *Math. Models Methods Appl. Sci.* **25** (10), 1845–1885.

[17] Cai, T., Cheng, J., Schmitzer, B. & Thorpe, M. (2022) The linearized Hellinger–Kantorovich distance. *SIAM J. Imaging Sci.* **15** (1), 45–83.

[18] Chizat, L., Peyré, G., Schmitzer, B. & Vialard, F.-X. (2018) An interpolating distance between optimal transport and Fisher–Rao metrics. *Found. Comp. Math.* **18** (1), 1–44.

[19] Chizat, L., Peyré, G., Schmitzer, B. & Vialard, F.-X. (2018) Unbalanced optimal transport: Dynamic and Kantorovich formulations. *J. Funct. Anal.* **274** (11), 3090–3123.

[20] Choksi, R. & Lu, X. Y. (2020) Bounds on the geometric complexity of optimal Centroidal Voronoi Tesselations in 3D. *Comm. Math. Phys.* **377** (3), 2429–2450.

[21] Du, Q., Faber, V. & Gunzburger, M. (1999) Centroidal Voronoi tessellations: Applications and algorithms. *SIAM Rev.* **41** (4), 637–676.

[22] Du, Q., Gunzburger, M., Ju, L. & Wang, X. (2006) Centroidal Voronoi tessellation algorithms for image compression, segmentation, and multichannel restoration. *J. Math. Imaging Vis.* **24** (2), 177–194.

[23] Du, Q. & Wang, D. S. (2005) The optimal centroidal Voronoi tessellations and the Gersho's conjecture in the three-dimensional space. *Comput. Math. Appl.* **49** (9-10), 1355–1373.

[24] Emelianenko, M., Ju, L. & Rand, A. (2008) Nondegeneracy and weak global convergence of the Lloyd algorithm in $\mathbb{R}^d$. *SIAM J. Numer. Anal.* **46** (3), 1423–1441.

[25] Figalli, A. (2010) The optimal partial transport problem. *Arch. Ration. Mech. Anal.* **195** (2), 533–560.

[26] Folland, G. B. (1999) *Real Analysis: Modern Techniques and Their Applications*. 2nd ed., Wiley.

[27] Galichon, A. (2016) *Optimal Transport Methods in Economics*, Princeton University Press.

[28] Gallouët, T. O. & Mérigot, Q. (2018) A Lagrangian scheme à la Brenier for the incompressible Euler equations. *Found. Comput. Math.* **18** (4), 835–865.

[29] Gallouët, T. O., Mérigot, Q. & Natale, A. (2022) Convergence of a Lagrangian discretization for barotropic fluids and porous media flow. *SIAM J. Math. Anal.* **54** (3), 2990–3018.

[30] Gersho, A. (1979) Asymptotically optimal block quantization. *IEEE Trans. on Inform. Theory* **25** (4), 373–380.

[31] Gersho, A. & Gray, R. M. (1992) *Vector Quantization and Signal Compression*, Springer.

[32] Graf, S. & Luschgy, H. (2000) *Foundations of Quantization for Probability Distributions*, Springer.

[33] Gruber, P. M. (1999) A short analytic proof of Fejes Tóth's theorem on sums of moments. *Aequationes Math.* **58** (3), 291–295.

[34] Gruber, P. M. (2004) Optimum quantization and its applications. *Adv. Math.* **186** (2), 456–497.

[35] Gruber, P. M. (2007) *Convex and Discrete Geometry*, Springer.

[36] Iacobelli, M. (2018) A gradient flow perspective on the quantization problem. In Cardaliaguet, P., Porretta, A. & Salvarani, F. editors, *PDE Models for Multi-Agent Phenomena*, Vol. **28**, Springer, 145–165.

[37] Kitagawa, J., Mérigot, Q. & Thibert, B. (2019) Convergence of a Newton algorithm for semi-discrete optimal transport. *J. Eur. Math. Soc.* **21** (9), 2603–2651.

[38] Kloeckner, B. (2012) Approximation by finitely supported measures. *ESAIM Control Optim. Calc. Var.* **18**(2), 343–359.

[39] Kondratyev, S., Monsaingeon, L. & Vorotnikov, D. (2016) A new optimal transport distance on the space of finite Radon measures. *Adv. Differential Equ.* **21** (11-12), 1117–1164.

[40] Larsson, L., Choksi, R. & Nave, J. C. (2016) Geometric self-assembly of rigid shapes: A simple Voronoi approach. *SIAM J. Appl. Math.* **76**(3), 1101–1125.

[41] Laschos, V. & Mielke, A. (2019) Geometric properties of cones with applications on the Hellinger–Kantorovich space, and a new distance on the space of probability measures. *J. Funct. Anal.* **276** (11), 3529–3576.

[42] Lavenant, H., Zhang, S., Kim, Y.-H. & Schiebinger, G. (2024) Toward a mathematical theory of trajectory inference. *Ann. Appl. Probab.* **34** (1A), 428–500.

[43] Leclerc, H., Mérigot, Q., Santambrogio, F. & Stra, F. (2020) Lagrangian discretization of crowd motion and linear diffusion. *SIAM J. Numer. Anal.* **58** (4), 2093–2118.

[44] Lévy, B. (2015) A numerical algorithm for L2 semi-discrete optimal transport in 3D. *ESAIM Math. Model. Numer. Anal.* **49** (6), 1693–1715.

[45] Liero, M., Mielke, A. & Savaré, G. (2018) Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures. *Invent. Math.* **211** (3), 969–1117.

[46] Liero, M., Mielke, A. & Savaré, G. (2023) Fine properties of geodesics and geodesic lambda-convexity for the Hellinger–Kantorovich distance. *Arch. Ration. Mech. Anal.* **247**, 112.

[47] Lloyd, S. P. (1982) Least squares quantization in PCM. *IEEE Trans. Inform. Theory* **28** (2), 129–137.

[48] Lu, X. Y. & Slepčev, D. (2013) Properties of minimizers of average-distance problem via discrete approximation of measures. *SIAM J. Math. Anal.* **45** (5), 3114–3131.

[49] MacQueen, J. B. (1967) Some methods for classification and analysis of multivariate observations. In Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol. **1**, University of California Press, pp. 281–297.

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press

[50] Mérigot, Q. (2011) A multiscale approach to optimal transport. *Comput. Graph. Forum* **30** (5), 1583–1592.
[51] Mérigot, Q., Santambrogio, F. & Sarrazin, C. (2021) Non-asymptotic convergence bounds for Wasserstein approximation using point clouds. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang, and J. Wortman Vaughan, editors, Vol. 34 of *Advances in Neural Information Processing Systems*, Curran Associates, Inc., pp. 12810–12821.
[52] Mérigot, Q. & Thibert, B. (2021) Optimal transport: Discretization and algorithms. In Bonito, A. & Nochetto, R. H. editors, *Geometric Partial Differential Equations Part II*, Vol. **22** of Handbook of Numerical Analysis, pp. 133–212.
[53] Meyron, J., Mérigot, Q. & Thibert, B. (2018) Light in power: A general and parameter-free algorithm for caustic design. *ACM Trans. Graph.* **37** (6), 1–13.
[54] Morgan, F. & Bolton, R. (2002) Hexagonal economic regions solve the location problem. *Am. Math. Monthly* **109** (2), 165–172.
[55] Mosconi, S. & Tilli, P. (2005) Γ–convergence for the irrigation problem. *J. Conv. Anal.* **12** (1), 145–158.
[56] Newman, D. (1982) The hexagon theorem. *IEEE Trans. on Inform. Theory* **28** (2), 137–139.
[57] Pagès, G., Pham, H. & Printems, J. (2004) Optimal quantization methods and applications to numerical problems in finance. In Rachev, S. T. editor, *Handbook of Computational and Numerical Methods in Finance*, Birkhäuser.
[58] Peyré, G. & Cuturi, M. (2019) Computational optimal transport. *Found. Trends Mach. Learn.* **11** (5-6), 355–607.
[59] Rockafellar, R. T. (1972) *Convex Analysis*, 2nd ed., Princeton University Press.
[60] Sabin, M. & Gray, R. (1986) Global convergence and empirical consistency of the generalized Lloyd algorithm. *IEEE Trans. on Inform. Theory* **32** (2), 148–155.
[61] Santambrogio, F. (2015) *Optimal Transport for Applied Mathematicians*, Birkhäuser.
[62] Santambrogio, F. (2023) *A Course in the Calculus of Variations*, Springer.
[63] Sarrazin, C. (2022) Lagrangian discretization of variational mean field games. *SIAM J. Control Optim.* **60** (3), 1365–1392.
[64] Schmitzer, B. & Wirth, B. (2019) Dynamic models of Wasserstein-1-type unbalanced transport. *ESAIM Control Optim. Calc. Var.* **25**, 23.
[65] Thorpe, M., Theil, F., Johansen, A. M. & Cade, N. (2015) Convergence of the *k*-means minimization problem using Γ-convergence. *SIAM J. Appl. Math.* **75** (6), 2444–2474.
[66] Thual, A., Tran, Q. H., Zemskova, T., et al. (2022) Aligning individual brains with fused unbalanced Gromov–Wasserstein. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K. & Oh, A. editors, *Advances in Neural Information Processing Systems*, Vol. **35**, pp. 21792–21804.
[67] Tóth, G. F. (2001) A stability criterion to the moment theorem. *Studia Sci. Math. Hungar.* **38** (1), 209–224.
[68] Tóth, L. F. (1972) *Lagerungen in der Ebene auf der Kugel und im Raum*, Springer.
[69] Villani, C. (2009) *Optimal Transport: Old and New*, Springer.
[70] Xin, S.-Q., Lévy, B., Chen, Z., et al. (2016) Centroidal power diagrams with capacity constraints: Computation, applications, and extension. *ACM Trans. Graph.* **35** (6), 244:1–244:12.
[71] Zador, P. L. (1982) Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Trans. Inform. Theory* **28** (2), 139–149.
[72] Zhu, C., Byrd, R. H., Lu, P. & Nocedal, J. (1997) Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw.* **23** (4), 550–560.

Cite this article: Bourne D. P., Schmitzer B. and Wirth B. (2026). Semi-discrete unbalanced optimal transport and quantization. *European Journal of Applied Mathematics*, **37**, 925–967. https://doi.org/10.1017/S0956792525100144

https://doi.org/10.1017/S0956792525100144 Published online by Cambridge University Press