Non-existence of the 2D quadratic Poisson optimal matching

Carlos A.C.C. Perello

Reuben College

University of Oxford

Supervised by Prof. James Martin

A thesis submitted for the degree of

Master of Science

Trinity 2024

Contents

List of Figures

Chapter 1

Introduction

This dissertation aims to introduce the reader to the so-called optimal matching problem and give an expository and transparently presented proof of [1, Theorem 1.1].

First, the (Poisson) optimal matching problem will be introduced in bounded and unbounded domains, alongside its connection to optimal transportation theory. Subsequently, we will discuss the linearisation of the Monge-Ampère equation near the uniform measure, which gives rise to the so-called *harmonic estimate* of the optimal matching between points, which states that under certain conditions, the displacement of the optimal matching can be well-approximated by the gradient of a harmonic function. Following that, the derivation of the upper bound utilised in [1, Theorem 1.1] will be presented, split into smaller components to enhance the clarity of the arguments employed. Afterwards, we will mention [2], where the optimal matching cost is no longer quadratic. We note that arguments presented in this dissertation may be employed for the non-quadratic matching case, at the expense of losing much of the proof's geometric interpretation. Finally, some avenues of further work are stated, alongside a small discussion about possible extensions of this result using multi-marginal optimal transport.

1.1 The optimal matching problem

We first introduce the (Poisson) *optimal matching* problem (sometimes called the *bipartite matching problem* or the *optimal assignment problem*). In a bounded domain, the problem goes as follows: consider a Borel-measurable and bounded subset[1] $B \subset \mathbb{R}^d$, and let $X = (X_i)_{i=1}^n$ and $Y = (Y_i)_{i=1}^n$ be i.i.d on B. Furthermore, let $\mu = \sum_{i=1}^N \delta_{X_i}$ and $\nu = \sum_{i=1}^N \delta_{Y_i}$ be the counting measures of X and Y, respectively. We define $\mathcal{T}$ to be the set of all possible matchings, i.e. $\mathcal{T} := \{T : \mathbb{R}^d \to \mathbb{R}^d \text{ s.t. } T\#\mu = \nu\}^2$. We now introduce the γ-cost of a matching:

[1] In fact, this problem can also be formulated for B a compact manifold, but we take $B \subset \mathbb{R}^d$ for simplicity.

[2] Recall that T *pushes forward* μ *into* ν if $T\#\mu = \nu$ iff $X \sim \mu \implies T^\circ(X) \sim \nu$.

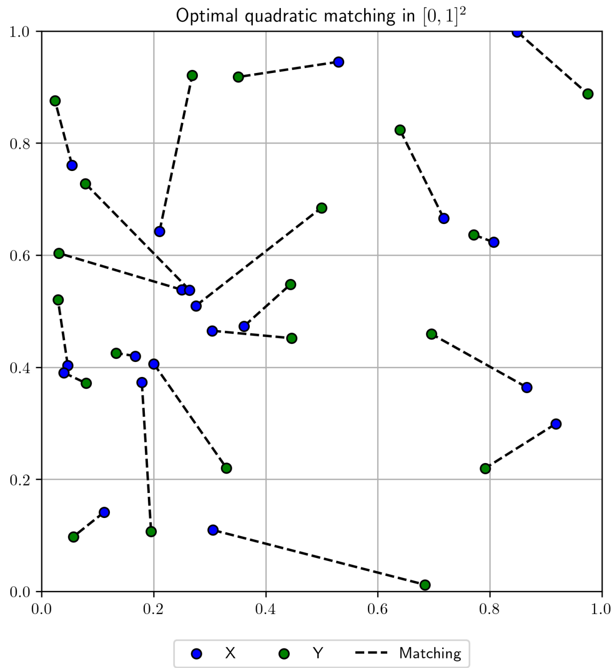

Figure 1.1: Quadratic optimal matching on the bi-dimensional unit square with 20 points.

Definition 1.1 (γ-cost of a matching)**.** For a matching $T \in \mathcal{T}$, the γ-*cost* $C_\gamma(T)$ of the matching T is defined as:

$$C_\gamma(T) := \sum_{i=1}^{N} |T(X_i) - X_i|^\gamma$$

Furthermore, a matching T^* *solves the (bounded)* γ*-optimal matching problem* if it satisfies:

$$T^*_\gamma = \inf_{T \in \mathcal{T}} C_\gamma(T) \tag{1.1}$$

which is finite as we are considering a finite number of samples in a bounded domain. When γ is not specified, we take $\gamma = 2$ so that $C(T) = C_2(T)$.

It is known that different values of γ change the geometric structure of T^*_γ; we will not be focusing on the geometric properties of T^*_γ, but we refer the reader to [3] for a discussion on how changing γ alters the behaviour of T^*_γ. Figure 1.1 displays the quadratic optimal matching $T^*_2 = T^*$ on $[0,\, 1]^2$. We are now interested in the case where the points in X and Y are i.i.d. distributed on $\mathbb{R}^d$. To arrive at a *canonical* matching, we restrict our set of possible matchings $\mathcal{T}$ to matchings that "look the same everywhere", which we now define precisely. Let X and Y be independent and homogeneous Poisson Point Processes (PPPs)[3] on $\mathbb{R}^d$. Unless otherwise specified, all the PPPs we consider throughout this dissertation will have unit intensity. Fix $a \in \mathbb{Z}^d$ and let us define the triple $\mathcal{S} := [(X_i), (Y_i), T(\cdot)]$. Consider the action $\oplus_a$ which translates the aforementioned triple by a:

$$\oplus_a(\mathcal{S}) := [(X_i + a), (Y_i + a), T(\cdot - a) + a]$$

[3]For a primer on spatial Poisson point processes, we refer the reader to [4].

Here and throughout, let the probability measure $\mathbb{P}$ be the probability measure induced by the PPPs' counting measures μ and ν, as well as the distribution on the set of matchings $\mathcal{T}$, and $\mathbb{E}$ be the probability measure's associated expectation. To arrive at a "canonical" matching, we would like this action to be *stationary* and *ergodic*:

Definition 1.2 (Stationarity & ergodicity)**.** Given the triple $\mathcal{S}$ and $a \in \mathbb{Z}^d$, the translation action $\oplus_a$ defined above is

1. $\oplus_a$ is *stationary* if for all events E generated by the σ-algebra $\sigma(\mathcal{S})$ and the corresponding shifted event $\oplus_a(E)$, then $\mathbb{P}(E) = \mathbb{P}(\oplus_a \cdot E)$.[4]

2. $\oplus_a$ is *ergodic* if for all $A \in \sigma(\mathcal{S})$ such that $\oplus_a \cdot A = A$, $\mathbb{P}(A) = 0$ or 1.

Additionally, if $\oplus_a$ is stationary or ergodic for all $a \in \mathbb{Z}^d$, then we call T, the matching in the triple $\mathcal{S}$, stationary or ergodic, respectively.

Observe that the PPPs X and Y we are considering are homogeneous, imposing that the action $\oplus_a$ is stationary and ergodic only affects the set of matchings $\mathcal{T}$. In simpler terms, the matching T is stationary if it looks the same regardless of what portion of $\mathbb{R}^d$ we consider it in, whilst T is ergodic if all of the translation-invariant events concerning T have trivial probabilities. We denote the set of all stationary and ergodic matchings in $\mathcal{T}$ by $\mathcal{T}^\circ$.

With this in mind, we would like to discern amongst all possible matchings in $\mathcal{T}^\circ$ a quadratic optimal matching that, in a sense, minimises $C(T)$. However, one cannot simply define T^* as in (1.1), as $C(T^*)$ is a.s. infinite in this case. This can be readily observed by considering $\mathbb{R}^d$ as a union of translated hypercubes and considering the cost of the matching T^* restricted to such a hypercube. As the matching T^* is both stationary and ergodic, the sum of the cost over N hypercubes grows at least as fast as N, and therefore, by letting $N \to \infty$, it follows that there are a.s. no stationary and ergodic maps with finite cost. The next best thing one can hope for is *local optimality*:

Definition 1.3 (Local optimality)**.** A matching T between two PPPs X and Y on $\mathbb{R}^d$ is *locally optimal* if, for any other matching $\tilde{T}$ such that $T(X) \neq \tilde{T}(X)$ on finitely many points, we have:

$$C(T^\circ) - C(\tilde{T}) \leq 0$$

Notice that due to the infinite number of cancellations, the quantity on the left-hand side above is finite.

We now aim to answer the question:

Given two homogeneous PPPs X and Y in $\mathbb{R}^2$, does there exist a stationary, ergodic and locally optimal matching T°?

[4]This is equivalent to saying that the push-forward of the map induced by the group action is trivial.

This question, despite seeming highly combinatorial is, in fact, deeply related to optimal transport. To see this, we introduce the so-called Monge problem:

Definition 1.4 (Monge problem)**.** Given two measures μ and ν supported on $\mathcal{X}$ and $\mathcal{Y}$ respectively, a map T^* is said to solve the *Monge problem with cost function* $c(x, y)$ if it satisfies:

$$T^* = \inf_{T \in \mathcal{T}} \int_{\mathcal{X}} c(x, T(x)) d\mu \tag{1.2}$$

where we recall that $\mathcal{T} = \{T : \mathcal{X} \to \mathcal{Y} \mid T\#\mu = \nu\}$.

Indeed, if we consider $B = \mathbb{R}^d$ and μ and ν to be the counting measures of X and Y and if we do not require T to be stationary and ergodic, then this problem can be thought of as finding a solution to the Monge problem between μ and ν with quadratic cost. This relationship is also hinted at by the fact that local optimality is, in fact, equivalent to *cyclical monotonicity* [5, 6]:

Definition 1.5 (Cyclical monotonicity)**.** A matching $T \in \mathcal{T}$ is *cyclically monotone* if for any finite collection of points $\{X_i\}_{i=1}^N \subset X$ in its support, we have:

$$\sum_{i=1}^{N} T(X_i)^\top (X_i - X_{i-1}) \geq 0, \qquad \text{where } X_0 := X_N$$

Cyclical monotonicity is a fundamental notion in optimal transportation theory, especially when considering optimal transport problems with quadratic cost, as all optimal transport plans have cyclically monotone support [5, Section 2.3]. However, cyclical monotonicity is much less intuitive than local optimality and therefore, in practice, it is easier to work with local optimality. Finally, the total cost of a matching $C_\gamma(T)$ is related to the γ-Wasserstein distance:

Definition 1.6 (γ-Wasserstein distance)**.** Let p and q be two probability measures on $\mathcal{X}$ and $\mathcal{Y}$ and $\gamma > 1$. Furthermore, let X and Y have finite γ moments. Then, their γ-Wasserstein distance is defined by:

$$W_\gamma(p, q) = \left(\inf_{\pi \in \Pi(p, q)} \int_{\mathcal{X} \times \mathcal{Y}} |x - y|^\gamma d\pi(x, y) \right)^{1/\gamma}$$

Where $\Pi(p, q)$ is the set of all couplings between p and q, i.e. $\pi \in \Pi(p, q)$ satisfies $\pi(\mathcal{X} \times B) = q(B)$ and $\pi(A \times \mathcal{Y}) = p(A)$.

Indeed, notice that if we take p and q to be counting measures of the PPPs X and Y restricted to some finite Borel set $B \subset \mathbb{R}^d$ and if $Y = T^*(X)$, then $W_\gamma^\gamma(p, q) = C_\gamma|_B (T^*)$, the γ-cost of T restricted to the set B. The γ-Wasserstein distance is a metric on the space of measures with finite γ-moments, which conveniently defines a notion of "closeness" of measures. To see the relationship between the Monge problem and the γ-Wasserstein distance, if T^* solves the Monge problem for $c(x, y) = |x - y|^\gamma$, then by definition of the Monge problem, coupling $\pi^* := (\text{Id}, T^*)\#(\mu, \nu)$. With this, we are ready to briefly introduce the Monge-Ampère equation and its connections to optimal transportation.

1.2 The Monge-Ampère equation

The Monge-Ampère equation is perhaps one of the most studied fully nonlinear PDEs of the last century. This may be partly due to the vast links this equation has to geometric analysis, as the Monge-Ampère equation can be used to derive meaningful results about the classification of affine spheres [7], the existence of surfaces with prescribed Gaussian curvature[5] [8], the existence of local isometric embeddings of a 2-manifold into $\mathbb{R}^3$ [9] and, of course, optimal transport [5, 6].

In essence, the Monge-Ampère prescribes the determinant of the Hessian of its solution:

$$\det(D^2u(x)) = f(x,\, u,\, \nabla u) \tag{1.3}$$

The equation above is equivalent to prescribing the product of the eigenvalues of the Hessian of u, hinting its connection to the Poisson equation, which prescribes the *sum* of the eigenvalues of the Hessian. In fact, both equations are intrinsically linked, and it is precisely this connection that yields the *harmonic estimate* derived in [10] via the linearisation of (1.3).

We now give a heuristic demonstration of how the Monge-Ampère arises from the optimal transport problem and discuss the linearisation mentioned above, mimicking the arguments in [5, Section 4.1] and [11]. Suppose we are given two probability measures μ and ν on $A \subset \mathbb{R}^d$ (for simplicity, let $A = [0,\, 1]^d$) with densities $\mu(x)$ and $\nu(x)$, respectively, and consider a measure-preserving map T such that

$$\mu(x) = \nu(T(x)) \det J_T(x) \tag{1.4}$$

where $J_T(X)$ is the Jacobian of the map T. Now, suppose $T(X) = T^*(X)$ is the solution of the quadratic Monge problem (1.2) for μ and ν. It is widely known that (subject to some regularity conditions on μ and ν) such a solution of the Monge problem is of the form $T^*(X) = \nabla\varphi(x)$ where $\varphi(x)$ is a convex function[6] [5, Theorem 2.12]. Then, we can plug this into (1.4) to obtain:

$$\mu(x) = \nu(\nabla\varphi(x)) \det\left(\nabla^2\varphi(x)\right) \tag{1.5}$$

This is the Monge-Ampère equation (1.3) with $f(x,\, u,\, Du) = \frac{\mu(x)}{\nu(\nabla\varphi(x))}$.

Now, suppose that both the source and target densities are "close" to each other in the sense that

$$\nu(x) = \nu_\varepsilon(x) = (1 + \varepsilon h(x) + O(\varepsilon^2))\mu(x) \tag{1.6}$$

Where, for simplicity, $h(x) \in C^\infty(A)$ and $\mu(x) \in C^1$. In this situation, we expect the transport map $T^*(X)$ to be close to the identity map, and as $T^*(X) = \nabla\varphi(x)$, then we

[5] This is also known as the *Minkowski problem.*

[6] This result is also known as *Brenier's theorem.*

can justify making the ansatz

$$\varphi(x) = \varphi_\varepsilon(x) = \frac{|x|^2}{2} + \varepsilon\psi(x) + O(\varepsilon^2) \tag{1.7}$$

where we assume $\psi(x) \in C^\infty(A)$. Plugging this ansatz into (1.5) and discarding second-order terms gives:

$$\mu(x) = (1 + \varepsilon h(x + \varepsilon\nabla\psi(x)))\mu(x + \varepsilon\nabla\psi(x)) \det(I + \varepsilon\nabla^2\psi(x))$$

Taking a first-order Taylor expansion in x for h and μ and once again discarding any second-order terms yields:

$$\mu(x) = (1 + \varepsilon h(x)(\mu(x) + \varepsilon\nabla\mu(x)\nabla\psi(x)) \det(I + \varepsilon\nabla^2\psi(x))$$

Expanding the first two terms and keeping first-order terms results in:

$$\mu(x) = (\mu(x) + \varepsilon\nabla\mu(x)\nabla\psi(x) + \varepsilon\mu(x)h(x)) \det(I + \varepsilon\nabla^2\psi(x))$$

Finally, we add and subtract $\det(I)$ to $\det(I + \varepsilon\nabla^2\psi(x))$

$$\begin{aligned}\mu(x) = (\mu(x) + \varepsilon\nabla\mu(x)\nabla\psi(x) + \varepsilon\mu(x)h(x))(\det(I + \varepsilon\nabla^2\psi(x)) - \det(I))\\ + \mu(x) + \varepsilon\nabla\mu(x)\nabla\psi(x) + \varepsilon\mu(x)h(x)\end{aligned}$$

Observe that we may cancel a $\mu(x)$ term. Recalling Jacobi's formula, specifically that the derivative of the determinant function at I is the trace, i.e. $(\det)'_I(M) = \mathrm{tr}(M)$ [12, Section 8.3], we now divide both sides by ε and take the limit as $\varepsilon \downarrow 0$ to obtain:

$$0 = \mu(x)\det'_I(\nabla^2(\psi(x))) + \nabla\mu(x)\nabla\psi(x) + \mu(x)h(x)$$

Notice that $\mathrm{tr}(\nabla^2(\psi(x))) = \Delta\psi$; dividing the equation above by $\mu(x)$, re-expressing the gradients and rearranging yields:

$$-\Delta\psi(x) - \nabla(\log(\mu(x)))\nabla\psi(x) = h(x)$$

We impose one further assumption: suppose that not only are our densities close to each other, but they are also close to the Lebesgue measure; that is

$$\mu(x) = 1 + \delta\mu'(x) \tag{1.8}$$

where $\delta \ll \sup_A |\mu(x)|$. Then, taking a first-order Taylor expansion, $\nabla(\log(\mu(x))) \approx \delta\mu'(x) \ll 1$, and therefore the Monge-Ampère equation (1.5), when linearised near the Lebesgue measure, becomes the Poisson equation:

$$-\Delta\psi(x) = h(x)$$

This heuristic derivation illustrates how the Monge-Ampère equation "becomes" the Poisson equation when both the source and target densities are close to the Lebesgue measure. However, there are multiple problems with the arguments presented above, the most prominent being that this only holds for probability measures that have a density[7]. Crucially, in the Poisson optimal matching problem, we consider counting measures, which do not have a density. Furthermore, we would like to make the notion of "closeness" more rigorous and obtain *quantitative* estimates on the quality of this linearisation, e.g. a result that states that given that the two measures μ and ν are quantitatively close in some sense, then the displacement $y - x$ between points distributed according to μ and ν is quantifiably close (again, in a notion that will be made precise) to a bounded harmonic gradient field $\nabla\Phi(x)$, analogous to how $T^*(x) - x = \nabla\left[\varphi(x) - x^2/2\right]$ for a quadratically optimal transport map T^*. The quantitative control we desire is exactly [10, Theorem 1.4], which will be essential in proving the non-existence of T°.

1.3 Proof outline

So far, we have touched upon several introductory notions in the theory of Poisson random matchings, optimal transport and its connection to the Monge-Ampère and Poisson equations. We recall that the main goal of this dissertation is to answer the following question in the negative:

Given two homogeneous PPPs X and Y in $\mathbb{R}^2$, does there exist a stationary, ergodic and locally optimal matching T°?

We will employ the same arguments as in [1], possibly presenting them differently so they are easier to understand. The proof of the non-existence of T° is based on a classical local 1-energy lower bound for bipartite matchings; if X and Y are two PPPs in $\mathbb{R}^d$ and $T \in \mathcal{T}$ is a matching between then, then we have[8]:

$$\frac{1}{R^d} \sum_{X \text{ or } T(X) \in B_R} |T(X) - X| \geq \Omega(\sqrt{\log R}) \tag{1.9}$$

where B_R is the ball of radius R. Of particular interest to us is the case when $d = 2$. We will not address this lower bound in great detail, but it will be briefly discussed after the proof, as this lower bound is essentially the same one derived in [13] around 50 years ago.

On the other hand, by using arguments that fundamentally rely on stationarity and ergodicity, as well as the harmonic approximation developed in [10], it is shown in [1] that:

$$\frac{1}{R^2} \sum_{X \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X| \leq o(\sqrt{\log R}) \tag{1.10}$$

[7] Unless specified otherwise, all densities considered are with respect to the Lebesgue measure.
[8] Note that we do not require the stationarity and ergodicity assumptions for the lower bound to hold.

which is incompatible with the lower bound (1.9). We briefly illustrate the quantities we must control to arrive at said upper bound. First, fix $L \gg 1$, which should be treated as a constant we can make arbitrarily large. Then, we divide the sum in the left-hand side of (1.10) into "long" (larger than L) and "short" (smaller than L) edges in the matching:

$$\frac{1}{R^2} \sum_{X \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X| = \frac{1}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| \le L}} |T^\circ(X) - X| + \frac{1}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| > L}} |T^\circ(X) - X|$$

Now apply the obvious bound to the first sum and Cauchy-Schwarz to the second sum to recover a quadratic cost term:

$$\begin{aligned} \frac{1}{R^2} \sum_{X \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X| \le \\ \frac{L}{R^2} |\{X : X \text{ or } T^\circ(X) \in B_R \text{ and } |T^\circ(X) - X| \le L\}| \\ + \left[\frac{1}{R^2} |\{X : X \text{ or } T^\circ(X) \in B_R \text{ and } |T^\circ(X) - X| > L\}|\right]^{1/2} \\ \times \left[\frac{1}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| \ge L}} |T^\circ(X) - X|^2\right]^{1/2} \end{aligned} \tag{1.11}$$

By inspecting the above equation, we can obtain an upper bound of the form (1.10) if we successfully control the following three quantities in terms of L and, crucially, R:

1. The *number of short edges*, corresponding to the term

$$A_1 := |\{X : X \text{ or } T^\circ(X) \in B_R \text{ and } |T^\circ(X) - X| \le L\}|$$

2. The *number of long edges*, corresponding to the term

$$A_2 := |\{X : X \text{ or } T^\circ(X) \in B_R \text{ and } |T^\circ(X) - X| > L\}|$$

3. The *local energy of the long edges*, corresponding to the term

$$A_3 := \frac{1}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| > L}} |T^\circ(X) - X|^2$$

Using these definitions, (1.11) becomes

$$\frac{1}{R^2} \sum_{X \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X| \le \frac{L}{R^2} A_1 + \sqrt{\frac{A_2}{R^2}} \times \sqrt{A_3} \tag{1.12}$$

which implies that, by successfully bounding A_1, A_2 and A_3 in terms of R, we arrive at (1.10) by a particular choice of L. In the sequel, we will observe that both A_1 and A_2 can be controlled using geometric arguments, but A_3 requires the harmonic estimate to approximate the displacement $|T^\circ(X) - X|$ by the gradient of a harmonic function, $\nabla\Phi(X)$, which is bounded. Additionally, we note that A_2 heavily relies on the stationarity and ergodicity assumptions made on T°, and that our control of A_3 requires both the bounds on A_1 and A_2 to hold.

This dissertation is structured as follows: Chapter 2 is devoted to bounding the term A_1 by controlling the *local distance of scale* R between the PPP counting measures μ and ν. More specifically, we show that, in $\mathbb{R}^2$, this quantity grows slower than $\log R$, a phenomenon that does not occur in higher dimensions. The criticality of dimension $d = 2$ which is why this result has only been proven, to date, for PPPs in $\mathbb{R}^2$. To show this, tools from harmonic analysis and martingale inequalities will be used, as well as multiple geometric arguments.

Chapter 3 addresses the bound on A_2, achieved using Birkhoff's ergodic theorem [14, Theorem 10.6]. This theorem, together with the stationarity and ergodicity assumptions on T°, yields the desired control on A_2.

Chapter 4 introduces the harmonic approximation result developed in [10] and shows how to apply it to this problem to find a bound for A_3. By analysing this problem at the mesoscopic scale, we also show that the harmonic approximation theorem holds in our context. To achieve this, we attain control of the *local energy at scale* R of the matching T°.

Chapter 5 combines all the bounds derived alongside a particular choice of L to show that the bound (1.10) indeed holds. The $\Omega(\sqrt{\log R})$ lower bound (1.9) is also touched upon, and the generalisation of the non-existence of a stationary, ergodic and locally optimal matching for costs $\gamma \in [1, \infty]$ introduced in [2] is briefly touched upon. Finally, some open problems in the Poisson matching area are introduced and discussed, as well as multi-marginal optimal transport, which may be the subsequent key development needed to extend this result to matchings of more than 2 PPPs.

Chapter 2

Local distances

2.1 Growth rate

We first bound the number of close edges, A_1, whose definition we recall below:

$$A_1 := |\{X :\ X \text{ or } T^\circ(X) \in B_R \text{ and } |T^\circ(X) - X| \leq L\}|$$

In fact, we will bound A_1 by proving the following lemma:

Lemma 2.1. *For any matching $T \in \mathcal{T}$, $\exists C > 0$ and an a.s. finite random radius r_* such that $\forall R \geq r_*$:*

$$|\{X :\ X \in B_R\} \cup \{T(X) :\ T(X) \in B_R\}| \leq CR^2 \tag{2.1}$$

Observe that the above Lemma applies for all $T \in \mathcal{T}$, so we do not require the stationarity and ergodicity assumption on T. Also, note that the quantity on the left-hand side of Lemma 2.1 is larger than A_1, which implies:

Corollary 2.2. $\exists C > 0$ *and an a.s. finite random radius r_* such that $\forall R \geq r_*$:*

$$A_1 \leq CR^2$$

Lemma 2.1 can also be used to control A_2, but doing so would not yield the desired upper bound (1.10). We can afford to use such a crude estimate for A_1 as in (1.12), A_1 is multiplied by L, a large constant we can choose; the fact that we can pick L allows us to circumvent the crudeness of the bound by carefully choosing L.

In the sequel, it will be helpful to introduce the *μ-density number of scale R*, $n_{R,\mu}$; this quantity is useful as if $n_{R,\mu} \approx 1$, then the point process corresponding to μ close to the expected number of points of a unit intensity PPP in B_R:

Definition 2.3 (Density number)**.** For a PPP X on $\mathbb{R}^d$ with corresponding counting measure μ, the *μ-density number of scale R*, $n_{R,\mu}$, is given by:

$$n_{R,\mu} := \frac{\mu(B_R)}{|B_R|} = \frac{|\{X :\ X \in B_R\}|}{|B_R|}$$

In case another set E_R of scale R is used instead of B_R, we use the notation

$$n_{E_R,\,\mu} := \frac{\mu(E_R)}{|E_R|}$$

We also introduce *local distance of scale* R, $D(R)$, which quantifies how the close counting measures of two PPPs are to each other and to the uniform measure at scale R. This is achieved by comparing the PPPs' counting measures μ and ν using Wasserstein distances and their density numbers $n_{R,\,\mu}$ and $n_{R,\,\nu}$:

Definition 2.4 (Local distance)**.** For $R > 0$ and two point process X and Y on $\mathbb{R}^d$ with counting measures μ and ν respectively, the *local distance of scale* R, $D(R)$ is defined as

$$\begin{aligned} D(R) := {} & \frac{1}{R^d}\, W_2^2\big|_{(-R,\,R)^d}(\mu,\, n_{R,\,\mu}) + \frac{R^2}{n_{R,\,\mu}}(n_{R,\,\mu}-1)^2 \\ & + \frac{1}{R^d}\, W_2^2\big|_{(-R,\,R)^d}(\nu,\, n_{R,\,\nu}) + \frac{R^2}{n_{R,\,\nu}}(n_{R,\,\nu}-1)^2 \end{aligned}$$

where $W_2|_{(-R,\,R)^d}(\mu,\,\nu)$ is the 2-Wasserstein distance where the measures considered are restricted to $(-R,\,R)^d$:

$$W_2|_{(-R,\,R)^d}(\mu,\,\nu) = W_2\left(\mu|_{(-R,\,R)^d},\, \nu|_{(-R,\,R)^d}\right)$$

and where, if $n \in \mathbb{R}_+$, then $W_2|_{(-R,\,R)^d}(\mu,\,n) := W_2|_{(-R,\,R)^d}(\mu,\,n\,dx)$, where dx is the Lebesgue measure.

The introduction of this quantity is motivated by the linearisation of the Monge-Ampère equation; in fact, $D(R)$ being small is a vital part of making the heuristic assumptions made in Equations (1.6) and (1.8) rigorous, where we imposed that both the source and target measures be close to each other and to the uniform measure. It is clear that the Wasserstein distance terms in $D(R)$ quantify the closeness of the counting measures to a multiple of the Lebesgue measure (namely, $n_{R,\,\mu}dx$ and $n_{R,\,\nu}dx$), whilst the square-distance terms impose that said multiples be close to 1; these conditions together yield that μ and ν are close each other and are near-uniform in B_R. Finally, note that although we introduce this quantity for PPPs in $\mathbb{R}^d$, the results we develop concerning $D(R)$ are in dimension $d = 2$; we introduce this quantity for general d as it has multiple applications throughout the analysis of the Monge-Ampère equation and its linearisation.

We begin by showing the following bound on the expected local 2-Wasserstein distance:

Lemma 2.5 (Expected W_2 bound, [1, Lemma 2.8.])**.** *If μ is the counting density associated to a PPP X in $\mathbb{R}^2$, then for R large enough the following bound holds:*

$$\mathbb{E}[W_2^2|_{(0,\,R)^2}(\mu,\, n_{(0,\,R)^2,\,\mu})] \lesssim R^2 \log R \tag{2.2}$$

where $A \lesssim B \iff \exists C$ depending solely on d such that $A \leq C \cdot B$ and where we recall that $n_{(0,\,R)^2,\,\mu} = \frac{\mu((0,\,R)^2)}{|(0,\,R)^2|} = \frac{\mu((0,\,R)^2)}{R^2}$

Proof. Without loss of generality, we may assume $R \in \mathcal{D}$ throughout the rest of the proof, where $\mathcal{D}$ is the set of all dyadic natural numbers, as for every R we can find $\bar{R} \in \mathcal{D}$ such that $\bar{R} \geq R$. Now consider all dyadic squares $Q \subset (0,\, R)^2$ with dyadic integer side length, and define the square's number density n_Q as:

$$n_Q := \frac{\mu(Q)}{|Q|}$$

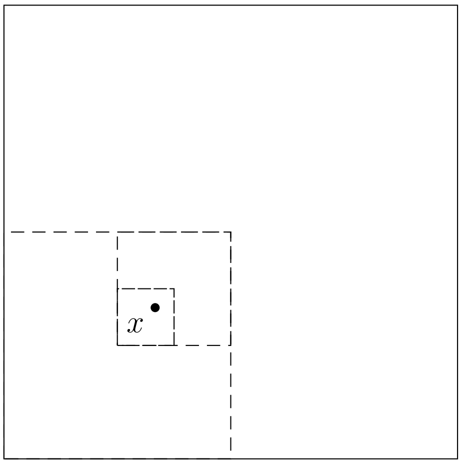

Figure 2.1: Sequence of three dyadic squares homing onto point x.

For a fixed point $x \in \mathbb{R}^2$, consider the sequence of nested dyadic squares $(Q_i)_{i=1}^{\infty}$ that contain x, i.e. $x \in Q_1 \subset Q_2 \subset \ldots$ Figure 2.1 displays the first three dyadic squares for a fixed $x \in \mathbb{R}^2$. Then, the corresponding random sequence of density numbers $(n_{Q_i})_{i=1}^{\infty}$ is a martingale, as it can be deduced from properties of PPPs that $n_Q|Q|$ follows a Poisson distribution with mean $|Q|$. We now will stop this subdivision when n_Q leaves the range of "moderate" values $\left[\frac{1}{2},\, 2\right]$, as if R is large enough, then $n_{R,\mu}$ concentrates around its expected value of 1 due to the Chernhoff bound applied to Poisson random variables [15, Proposition 11.15]. Then, for every $x \in \mathbb{R}^2$, we can define the *stopping scale*, $r_*(x)$:

$$r_*(x) := 2 \sup_{r \in \mathcal{D}} \left\{ r \text{ is the side length of square } Q_i \ni x \text{ s.t. } n_Q \notin \left[\frac{1}{2},\, 2\right] \right\}$$

Observe that for a square Q^* of side length $1/2$ we have that:

$$n_{Q^*} = 4\mu(Q^*) \in 4\mathbb{N}_0 \notin \left[\frac{1}{2},\, 2\right]$$

which implies $r_*(x) \geq 1$. Now, we show that $r_*^6(x)$ is $O(1)$ in expectation:

$$\mathbb{E}[r_*^6(x)] \lesssim 1$$

Indeed, by the definition of r_*, we have:

$$\mathbb{P}(r_*(x) > \rho) \leq \sum_{Q \ni x,\, r_Q \geq \rho} \mathbb{P}\left(n_{Q_\rho} \notin \left[\frac{1}{2},\, 2\right]\right)$$

where r_Q is the side length of the square Q, and Q_r is a square with side length r. Applying the Chernhoff bounds once more yields:

$$\mathbb{P}(r_*(x) > \rho) \lesssim \sum_{\text{dyadic } r \geq \rho} \exp(-Cr^2) \lesssim \exp(-C\rho^2)$$

We use the bound above to estimate $\mathbb{E}[r_*^6(x)]$ using the layer-cake representation and a change of variables:

$$\mathbb{E}[r_*^6(x)] = 6\int_0^\infty \rho^5 \mathbb{P}(r_*(x) > \rho)d\rho \lesssim \int_0^\infty \rho^5 \exp(-C\rho^2)d\rho \lesssim 1 \tag{2.3}$$

We now consider the following two disjoint cases; when $r_* \leq R$ on all of $(0,\, R)^2$ (which we call event E_R) and when $\exists\, y \in (0,\, R)^2$ s.t. $r_* > R$ (which will be denoted E_R^c).

In the latter, there must exist a cube $Q' \ni y$ of length $r_{Q'} \geq R$ such that $n_{Q'} \notin \left[\frac{1}{2},\, 2\right]$, and therefore $r_* > R$ on all of $(0,\, R)^2$. Fix $y \in (0,\, R)^2$. Notice that $\pi = \mu \otimes n_{(0,\,R)^2}dx$ is a coupling of μ and $n_{(0,\,R)^2}dx$, and that if $x,\, y \in (0,\, R)^2$, $\|x - y\|^2 \leq 2R^2$. Then:

$$\begin{aligned}
W_2^2\big|_{(0,\,R)^2}(\mu,\, n_{(0,\,R)^2,\,\mu}) &\leq \int_{(0,R)^2\times(0,R)^2} \|x-y\|^2 [d\mu(x) \otimes n_{(0,\,R)^2,\,\mu}dy] \\
&\leq 2R^2 \int_{(0,\,R)^2\times(0,\,R)^2} [d\mu(x) \otimes n_{(0,\,R)^2,\,\mu}dy] \\
&= 2R^2R^2\mu((0,\, R))n_{(0,\,R)^2,\,\mu} \\
&= 2R^6 n^2_{(0,\,R)^2,\,\mu}
\end{aligned}$$

which, combined with the fact that $(0,\, R)^2 \subset (0,\, r_*(y))^2$, gives

$$W_2^2\big|_{(0,\,R)^2}(\mu,\, n_{(0,\,R)^2,\,\mu}) \leq n^2_{(0,\,r_*^2(y)),\,\mu} r_*^2(y) 2R^4$$

Now, by definition of $r_*(y)$, $\frac{1}{2} \leq n_{(0,\,r_*(y))^2,\,\mu} \leq 2$ and thus

$$W_2^2\big|_{(0,\,R)^2}(\mu,\, n_{(0,\,R)^2,\,\mu}) \leq 4r_*^2(y)2R^4 \lesssim r_*^6(y)$$

holds whenever $r_*(y) > R$. Using indicator functions, taking expectations and applying (2.3) yields the estimate:

$$\mathbb{E}[W_2^2\big|_{(0,\,R)^2}(\mu,\, n_{(0,\,R)^2,\,\mu})\chi(E_R^c)] \lesssim 1 \tag{2.4}$$

where $\chi(E)$ denotes the indicator function of the event E.

We now turn to the case when $r_* \leq R$ on all of $(0,\, R)^2$. We now define a partition $\mathcal{Q}$ of $(0,\, R)^2$; a dyadic square $Q_* \in \mathcal{Q}$ if its number density n_{Q_*} is within the "moderate" value range $\left[\frac{1}{2},\, 2\right]$, but the number density of at least one of its four children leaves this range. Using the definition of $r_*(x)$, this is equivalent to imposing the following condition:

$$r_{Q_*} = \max\{r_*(x) \,:\, x \in Q_*\}$$

Once more, by definition, one of the children of Q_* (which we denote $Q_{*,c}$) will be such that $n_{Q_{*,c}} \notin \left[\frac{1}{2}, 2\right]$. Therefore, for $x' \in Q_{*,c}$, $r_*(x') = r_{Q_*}$, and thus

$$\fint_{Q_*} r_*^2(x)dx \geq \frac{1}{r_{Q_*}^2} \int_{Q_{*,c}} r_*^2(x)dx = \frac{1}{4} r_{Q_*}^2 \tag{2.5}$$

We now consider the transportation cost between $n_{Q_*}dx$ and μ. By construction, n_{Q_*} takes values near its mean; this implies that there exists a coupling $\tilde{\pi}$ between μ and $n_{Q_*}dx$ such that no points are distributed between different squares in $\mathcal{Q}$; i.e. all of the mass is redistributed within the squares Q_*. In each square, the redistribution can have at most length $r_{Q_*}\sqrt{2}$, which gives

$$\begin{aligned} W_2^2\big|_{(0,R)^2\times(0,R)^2}(\mu, n_{Q^*}) &\leq \int_{(0,R)^2} \|x-y\|^2 d\tilde{\pi} \\ &= \sum_{Q_*\in\mathcal{Q}} \int_{Q^*\times Q^*} \|x-y\|^2 d(\tilde{\pi}|_{Q_*}) \\ &\leq \sum_{Q_*\in\mathcal{Q}} 2r_{Q_*}^2 n_{Q_*} |Q_*| \\ &= \sum_{Q_*\in\mathcal{Q}} 2r_{Q_*}^4 n_{Q_*} \end{aligned}$$

where $\tilde{\pi}|_{Q^*}$ is the coupling π restricted to the square Q^*. We apply (2.5) to the inequality above and the fact that $n_{Q_*} \leq 2$ to obtain

$$W_2^2\big|_{(0,R)^2}(\mu, n_{Q^*}) \leq \sum_{Q_*\in\mathcal{Q}} 16 \int_{Q_*} r_*^2(x)dx = 16 \int_{(0,R)^2} r_*^2(x)dx$$

where we have used that the squares Q_* form a partition of $(0, R)^2$. Now, as $r_*(x) \geq 1$, $r_*^2(x) \leq r_*^6(x)$; using this and taking expectations yields:

$$\mathbb{E}[W_2^2\big|_{(0,R)^2}(\mu, n_{Q^*})] \leq 16 \int_{(0,R)^2} \mathbb{E}[r_*^2(x)]dx \leq 16 \int_{(0,R)^2} \mathbb{E}[r_*^6(x)]dx$$

Finally, using (2.3) and using indicator functions gives

$$\mathbb{E}[W_2^2\big|_{(0,R)^2}(\mu, n_{Q^*})\chi(E_R)] \lesssim 1 \tag{2.6}$$

Now, it only remains to show that, for any square $Q_* \in \mathcal{Q}$,

$$\mathbb{E}[W_2^2\big|_{(0,R)^2}(n_{Q_*}, n_{(0,R)^2,\mu})\chi(E_R)] \lesssim R^2 \log R \tag{2.7}$$

This inequality, together with (2.4) and the triangle inequality for transportation distances, implies (2.2). We now introduce the *Eulerian* or *fluid mechanics* formulation of the Wasserstein 2-distance as done in [16] and discussed further in [17]:

Theorem 2.6. *For two measures p and q on $\mathbb{R}^n$ with finite second moments, the following equality holds:*

$$\begin{aligned} W_2(p,\, q) &= \inf_{(\rho,\, j)} \int_0^1 \int_{\mathbb{R}^n} |j(x,\, t)|^2 \rho(x,\, t) dx dt \\ s.t.\ \frac{\partial \rho}{\partial t} &+ \nabla \cdot (\rho j) = 0,\ \rho(x,\, 0) = p(x),\ \rho(x,\, 1) = q(x), \quad \textit{weakly} \end{aligned} \tag{2.8}$$

It is clear that the theorem above also holds if we restrict p and q to some measurable support $X \subset \mathbb{R}^d$, which in our case is $(0,\, R)^2 \subset \mathbb{R}^2$. Observe that an admissible ρ would be a linear interpolation between the measures, i.e.

$$\rho(x,\, t) = t n_{(0,\, R)^2,\, \mu}(x) + (1 - t) n_{Q_*}(x) \tag{2.9}$$

By construction we have the bounds $\frac{1}{2} \leq n_{Q^*},\, n_{(0,\, R)^2,\, \mu} \leq 2$, so $\frac{1}{2} \leq \rho(x,\, t) \leq 2$. Thus, if $j(x,\, t) = j(x)$, the Eulerian formulation of the W_2 implies the bound

$$W_2^2\big|_{(0,\, R)^2} (n_{Q_*},\, n_{(0,\, R)^2,\, \mu}) \leq 2 \int_{(0,\, R)^2} |j(x)|^2 dx \tag{2.10}$$

For all j which satisfy the continuity equation conditions in (2.8) which, due to our choice of ρ, become all j such that

$$\nabla \cdot j = n_{(0,\, R)^2,\, \mu} - n_{Q_*} \text{ on } (0,\, R)^2, \quad \nu \cdot j = 0 \text{ on } \partial(0,\, R)^2 \tag{2.11}$$

where ν is the outward normal of $(0,\, R)^2$. We will now construct a $j(x)$ satisfying (2.11), from which the bound (2.7) will follow.

Consider a dyadic square and its four children Q_c. On each of the four children, we find a weak solution to the following Poisson equation with piecewise-constant right-hand side and Neumann boundary conditions

$$-\Delta \varphi_Q = n_Q - n_{Q_c} \text{ on } Q_c,\ \nu \cdot \nabla \varphi_Q = 0 \text{ on } \partial Q$$

which, by the Fredholm alternative and the self-adjointness of the Laplacian [18], has a solution, as

$$\int_Q n_Q dx - \int_Q n_{Q_c} dx = \mu(Q) - \sum_{Q_c \text{ children of } Q} \mu(Q_c) = 0$$

We define the flux $j^*(x)$ as:

$$j^*(x) := - \sum_{Q \cap Q_* \ni x,\, r_Q \in [2 r_{Q_*},\, R]} \nabla \varphi_Q(x)$$

This choice of flux $j^*(x)$ is motivated by the Neumann boundary conditions imposed on ∂Q, which avoid singular contributions to the above sum across square boundaries. Furthermore, we must restrict the side lengths of the squares we sum to be strictly larger than r_{Q^*}, as we require the bound $\rho(x,\, t) \leq 2$ to hold for the inequality (2.10) to be true, and therefore all the squares we may consider must be coarser than r_{Q_*}. Indeed, it is also straightforward to see that the boundary conditions of the original problem (2.10) hold, showing that a $j^*(x)$ is an admissible choice of flux.

We now show that (2.7) holds for $j^*(x)$. Indeed, by the Poincaré inequality, we have:

$$\frac{1}{r_Q^2}\int_Q |\nabla\varphi_Q|^2 dx \lesssim \int_Q |\Delta\varphi_Q|^2 dx = |Q| \sum_{Q_c \text{ children of } Q} (n_{Q_c} - n_Q)^2 \tag{2.12}$$

As n_Q is in the "moderate" range of values, we employ the crude estimate

$$\sum_{Q_c \text{ children of } Q} (n_{Q_c} - n_Q)^2 \lesssim \sum_{Q_c \text{ children of } Q} (n_{Q_c} - 1)^2$$

which is a useful bound, as we note that

$$\mathbb{E}[(n_Q - 1)^2] = \frac{1}{|Q|^2}\mathbb{E}[\mu(Q)^2] - \frac{2}{|Q|}\mathbb{E}[\mu(Q)] + 1 \lesssim \frac{1}{|Q|}$$

meaning that applying (2.12) yields

$$\mathbb{E}\left[\int_Q |\nabla\varphi_Q|^2 dx\right] \lesssim |Q|^2 \sum_{Q_c \text{ children of } Q} \mathbb{E}[(n_{Q_c} - 1)^2] \lesssim |Q| \tag{2.13}$$

Now fix x and consider all the squares $Q \ni x$. Obviously, $\nabla\varphi_Q(x)$ only depends on the PPP through $\{n_{Q_c}\}$. Furthermore, by the Poisson equation, $\mathbb{E}[\nabla\varphi_Q(x)] = 0$. We now claim that

$$M_r := \sum_{Q \ni x,\, r_Q \in [r,\, R]} \nabla\varphi_Q(x),$$

where r takes values in the set $\{R \cdot 2^{-n} \mid n \in \mathbb{N}\}$, is a discrete martingale wrt the filtration generated by $\{n_Q\}$. Denoting by $\mathcal{F}_r$ the filtration generated by $\{n_Q\}$ where $r_Q = r$, one has

$$\begin{aligned}
\mathbb{E}[M_{r/2}|\mathcal{F}_r] &= \sum_{Q \ni x,\, r_Q \in [r,\, R]} \nabla\varphi_Q(x) + \mathbb{E}\left[\sum_{Q \ni x,\, r_Q \in [r/2,\, r]} \nabla\varphi_Q(x) \middle| \mathcal{F}_r\right] \\
&= M_r + \sum_{Q \ni x,\, r_Q \in [r/2,\, r]} \mathbb{E}[\nabla\varphi_Q(x)|\mathcal{F}_r] \\
&= M_r
\end{aligned}$$

where we have used that M_r is $\mathcal{F}_r$-measurable, that the second term is independent of $\mathcal{F}_r$ and that the term $\nabla\varphi_Q(x)$ has zero mean. Therefore, we can exploit the martingale structure of M_r to find a bound for the flux $j^*(x)$. Observe that as $r_* \geq 1$, we have that

$$\mathbb{E}[|j^*(x)|^2] \leq \mathbb{E}\left[\sup_{r \geq 1} |M_r|^2\right]$$

We may now apply the Burkholder-Davis-Gundy inequality [19, Section 11.5] to M_r to obtain

$$\mathbb{E}[|j^*(x)|^2] \lesssim \mathbb{E}\left[\sum_{Q \ni x,\, r_Q \in [1,\, R]} |\nabla \varphi_Q(x)|^2\right] \leq \mathbb{E}\left[\sum_{Q:\, r_Q \in [1,\, R]} |\nabla \varphi_Q(x)|^2\right]$$

which, after integrating over $(0,\, R)^2$, becomes

$$\mathbb{E}\left[\int_{(0,\, R)^2} |j^*(x)|^2 dx\right] \lesssim \sum_{Q:\, r_Q \in [1,\, R]} \mathbb{E}\left[\int_Q |\nabla \varphi_Q(x)|^2 dx\right],$$

using indicator functions and putting this together with (2.9) and (2.13) gives:

$$\mathbb{E}[W_2^2\big|_{(0,\, R)^2}(n_{Q_*},\, n_{(0,\, R)^2,\, \mu})\chi(E_R^c)] \lesssim \sum_{Q:\, r_Q \in [1,\, R]} |Q| \lesssim R^2 \log R \tag{2.14}$$

Then, we have

$$\begin{aligned}
&\mathbb{E}[W_2^2\big|_{(0,\, R)^2}(\mu,\, n_{(0,\, R)^2,\, \mu})] \\
&= \mathbb{E}[W_2^2\big|_{(0,\, R)^2}(\mu,\, n_{(0,\, R)^2,\, \mu})\chi(E_R)] + \mathbb{E}[W_2^2\big|_{(0,\, R)^2}(\mu,\, n_{(0,\, R)^2,\, \mu})\chi(E_R^C)] \\
&\overset{(2.4)+\triangle-\text{ineq}}{\lesssim} 1 + \mathbb{E}[W_2^2\big|_{(0,\, R)^2}(\mu,\, n_{Q_*})\chi(E_R^C)] + \mathbb{E}[W_2^2\big|_{(0,\, R)^2}(n_{Q_*},\, n_{(0,\, R)^2,\, \mu})\chi(E_R^C)] \\
&\overset{(2.6)+(2.14)}{\lesssim} 1 + 1 + R^2 \log R \lesssim R^2 \log R
\end{aligned}$$

yielding (2.2), as required. □

Remark 2.7. In the proof above, the only term that gives rise to the $R^2 \log R$ upper bound is (2.14), which quantifies the 2-Wasserstein distance at small scales, as the other terms used in the triangle inequality bound contribute $O(1)$.

We are now in a good position to derive a key estimate of $D(R)$ in $\mathbb{R}^2$, which corresponds to [1, Lemma 2.6]:

Lemma 2.8 (Growth of $D(R)$ when $d = 2$)**.** *If X and Y are two PPPs on $\mathbb{R}^2$, then $\exists C > 0$ and an a.s. finite random radius r_* such $\forall R \geq r_*$, $D(R)$ satisfies:*

$$D(R) \leq C \log R \tag{2.15}$$

Proof. We will show that

$$\frac{1}{R^2}\, W_2^2\big|_{(-R,\, R)^2}(\mu,\, n_{(-R,\, R),\, \mu}) + \frac{R^2}{n_{(-R,\, R),\, \mu}}(n_{(-R,\, R),\, \mu} - 1)^2 \lesssim \log R \tag{2.16}$$

First, we prove that $\exists C > 0$ and an a.s. finite r_* such that $\forall$ dyadic $R \geq r_*$

$$\frac{1}{R^2}\, W_2^2\big|_{(-R,\, R)^2}(\mu,\, n_{(-R,\, R)^2,\, \mu}) \leq C \log R \tag{2.17}$$

Pick R large enough such that Lemma 2.5 holds in the square $(0,\, 2R)^2$; by stationarity of the PPP X, we may translate this to the symmetric square $(-R,\, R)^2$ to obtain

$$\frac{1}{R^2}\mathbb{E}[W_2^2\big|_{(-R,\,R)^2}(\mu,\, n_{(-R,\,R)^2,\,\mu})] \lesssim \log R$$

for R large enough. We now appeal to [20, Proposition 2.7], which by clever choice of M can be interpreted as a concentration inequality for the Wasserstein distance:

Proposition 2.9 (Goldman, Huesmann, and Otto [20])**.** *For a Poisson process μ on $\mathbb{R}^2$, and $\varepsilon > 0$, $\exists C > 0$ independent of R and ε such that:*

$$\mathbb{P}\left(\frac{1}{R^2 \log R}\left|W_2^2\big|_{(-R,\,R)^2}(\mu,\, n_{(-R,\,R)^2,\,\mu}) - \mathbb{E}[W_2^2\big|_{(-R,\,R)^2}(\mu,\, n_{(-R,\,R)^2,\,\mu})]\right| > \varepsilon\right) \lesssim e^{-C\varepsilon \log R}$$

The proposition above, together with the Borel-Cantelli lemma, gives (2.17), as it is clear that the right-hand side of the concentration inequality is summable for dyadic R.

Finally, we now show that $\exists C > 0$ and an a.s. finite $r_* < \infty$ such that $\forall$ dyadic $R \geq r_*$

$$\frac{R^2}{n_{(-R,\,R)^2,\,\mu}}(n_{(-R,\,R)^2,\,\mu} - 1)^2 \leq C \log R \tag{2.18}$$

As once more we may assume that R is large enough such that $n_{(-R,\,R)^2,\,\mu} \in \left[\frac{1}{2},\, 2\right]$, the above is equivalent to showing:

$$R^2(n_{(-R,\,R)^2,\,\mu} - 1)^2 \lesssim \log R$$

Observe that since $n4R^2$ is a Poisson random variable with mean $4R^2$, by the Chernhoff tail bounds:

$$\mathbb{P}\left(R^2(n_{(-R,\,R)^2,\,\mu} - 1)^2 > \log R\right) = \mathbb{P}\left(\left|n4R^2 - 4R^2\right| > CR\sqrt{\log R}\right) \lesssim e^{-C \log R}$$

Applying the Borel-Cantelli lemma once more gives that (2.18). Therefore, we have shown that (2.16) holds for μ. Now, applying this inequality to both μ and ν with a.s. finite radii $r_{\mu,*}$ and $r_{\nu,*}$ respectively; taking $r_* = \max(r_{\mu,*},\, r_{\nu,*})$ gives (2.15), as claimed. □

Now, returning to A_1, we may use (2.15) to obtain a bound on (2.1); indeed, notice that by (2.15), $\forall \varepsilon > 0, \exists R > 0$ such that

$$D(R) \leq \varepsilon R^2$$

In addition, let R be large enough such that $n_{R,\,\mu},\, n_{R,\,\nu} \in \left[\frac{1}{2},\, 2\right]$. Then, unpacking the definition of $D(R)$ and keeping only the Euclidean-square distance terms gives:

$$\frac{R^2}{n_{R,\mu}}(n_{R,\mu} - 1)^2 + \frac{R^2}{n_{R,\nu}}(n_{R,\nu} - 1)^2 \leq D(R) \tag{2.19}$$

Recalling the definition of $n_{R,\mu}$ and applying it to the first term in the right-hand side of (2.19) gives the yields bound

$$\frac{\pi R^4}{\mu(B_R)}\left(\frac{\mu(B_R)^2}{\pi^2 R^4}-2\frac{\mu(B_R)}{\pi R^2}+1\right)=\frac{\mu(B_R)}{\pi}-2R^2+R^2\frac{1}{n_{R,\,\mu}}\geq\frac{\mu(B_R)}{\pi}-\frac{3}{2}R^2$$

Applying the same procedure to $n_{R,\,\nu}$ implies that $\exists C>0$ such that

$$\mu(B_R)+\nu(B_R)\leq CR^2$$

Finally, notice that if we take $T=T^\circ\in\mathcal{T}^\circ$, then:

$$A_1\leq|\{X:\ X\in B_R\}\cup\{T^\circ(X):\ T^\circ(X)\in B_R\}|=\mu(B_R)+\nu(B_R)\leq CR^2 \tag{2.20}$$

Thus bounding A_1.

Chapter 3

Ergodicity & Stationarity

3.1 Birkhoff's theorem

We now seek to control the quantity A_2, which we achieve by using methods from ergodic theory, namely Birkhoff's ergodic theorem. First, we define the *point count* of the set $U \times V$:

Definition 3.1 (Point count of $U \times V$)**.** Let X, Y be PPPs on $\mathbb{R}^d$, which are matched by $T \in \mathcal{T}$. For U, V Lebesgue-measurable sets with at least one of U or V having finite Lebesgue measure, we define the *point count of* $U \times V$, $N_{U,V}$, as:

$$N_{U,V} := |\{(X, T(X)) : X \in U, T(X) \in V\}|$$

We now assume that T is stationary and ergodic, i.e. $T = T^\circ \in \mathcal{T}^\circ$. Furthermore, to give meaningful results about the matching problem, we must assume that our sigma algebra $\sigma(\mathcal{S})$ is nontrivial in the sense that it allows basic quantities such as $N_{U,V}$ to be measurable. Then, for any $a \in \mathbb{Z}^d$, due to stationarity of $\oplus_a$ we have

$$N_{U,V} \overset{\text{Law}}{=} N_{a+U,\, a+V}$$

which, alongside the ergodicity of the action $\oplus_a$, will allow us to employ Birkhoff's ergodic theorem:

Theorem 3.2 (Birkhoff's ergodic theorem [14, Theorem 10.6])**.** *If $f \in L^1(X, B, \mu)$ and if the measure μ is g-invariant where g is an action acting on X, and if g is ergodic, then:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(g^n \cdot x) = \mathbb{E}[f(x)] \qquad a.e.$$

We will apply Birkhoff's ergodic theorem to obtain:

$$\lim_{R\uparrow\infty} \frac{1}{R^d} \sum_{a\in\mathbb{Z}^d\cap[0,R^d)} N_{U+a,\, V+a} = \mathbb{E}[N_{U,V}] \qquad a.s.$$

The above limit will allow us to estimate A_2 as $A_2 = N_{A,B}$ for a particular choice of A and B.

3.2 Long-distance bounds

With this in mind, we first recall the definition of A_2 below:

$$A_2 := |\{X : \ X \text{ or } T^\circ(X) \in B_R \text{ and } |T^\circ(X) - X| > L\}|$$

We will show the following lemma:

Lemma 3.3 (Ergodic estimate, [1, Lemma 2.1]). *For any $\varepsilon > 0$, $\exists L > 0$ deterministic and an a.s. finite random radius r_* such that $\forall R \geq r_*$, we have:*

$$|\{X \in (-R,\, R)^d : |T^\circ(X) - X| > L\}| \leq (\varepsilon R)^d \tag{3.1}$$

This immediately shows that $A_2 \leq o(R^d)$, as the hypercube $(-R,\, R)^d$ contains the hypersphere of radius R. We present this result for general $\mathbb{R}^d$ to highlight its applicability to higher dimensions, but we will only apply it in the setting where $d = 2$.

Proof. First, we rewrite the set in the left-hand side of (3.1) as $N_{A,\,B}$; writing $Q_R = (-R,\, R)^d$ gives and $B_L(Q_R)$ for the set of points a distance L from Q_R gives:

$$N_{Q_R,\,\mathbb{R}^d \setminus B_L(Q_R)} = \{X \in (-R,\, R)^d : |T^\circ(X) - X| > L\}$$

We begin by showing the following ergodic result:

$$\lim_{R\uparrow\infty} \frac{1}{R^d} N_{Q_R,\,\mathbb{R}^d \setminus B_L(Q_R)} = \mathbb{E}[N_{Q_1,\,\mathbb{R}^d \setminus B_L(Q_1)}] \qquad a.s. \tag{3.2}$$

Then, we show that taking $L \to \infty$ gives

$$\mathbb{E}[N_{Q_1,\,\mathbb{R}^d \setminus B_L(Q_1)}] \xrightarrow{L\to\infty} 0 \tag{3.3}$$

which and together imply the existence of an a.s. finite random radius r_* and a deterministic L such that Lemma 3.3 holds. This can be achieved as follows: fix $\varepsilon > 0$ and first pick L large enough such that

$$\mathbb{E}[N_{Q_1,\,\mathbb{R}^d \setminus B_L(Q_1)}] \leq \frac{\varepsilon^d}{2} \tag{3.4}$$

holds. Then, choose r_* large enough such that

$$\left| \frac{1}{R^d} N_{Q_R,\,\mathbb{R}^d \setminus B_L(Q_R)} - \mathbb{E}[N_{Q_1,\,\mathbb{R}^d \setminus B_L(Q_1)}] \right| \leq \frac{\varepsilon^d}{2} \tag{3.5}$$

is satisfied. Combining (3.4) and (3.5) gives that

$$N_{Q_R,\,\mathbb{R}^d \setminus B_L(Q_R)} \leq (\varepsilon R)^d$$

which is precisely (3.1).

We first show (3.2); observe that

$$N_{Q_R,\,\mathbb{R}^d\setminus B_L(Q_R)} = \sum_{|i|<R} N_{Q_1+i,\,\mathbb{R}^d\setminus B_L(Q_1+i)}$$
$$= \sum_{|i|<R} \{X \in (i_1-1,\, i_1+1)\times\cdots\times(i_d-1,\, i_d+1) :\ |T^\circ(X)-X| > L\}$$

As μ, the counting measure of the PPP X, and T° are stationary and ergodic, then by Birkhoff's ergodic theorem we have that, for a sequence $(R_n)_{n=1}^\infty \subset \mathbb{N}$ such that $R_n \uparrow \infty$ as $n \to \infty$ the following holds:

$$\lim_{n\to\infty} \frac{1}{R_n^d} \sum_{|i|<R_n} N_{Q_1+i,\,\mathbb{R}^d\setminus B_L(Q_1+i)} = \mathbb{E}[N_{Q_1,\,\mathbb{R}^d\setminus B_L(Q_1)}] \qquad a.s. \tag{3.6}$$

For integer radii, note that the following equality holds:

$$\frac{1}{R_n^d} \sum_{|i|<R_n} N_{Q_1+i,\,\mathbb{R}^d\setminus B_L(Q_1+i)} = \frac{1}{R_n^d} N_{Q_{R_n},\,\mathbb{R}^d\setminus B_L(Q_{R_n})}$$

Therefore, (3.6) becomes (3.2) when $R \to \infty$ along integer values. Furthermore, by the squeeze theorem, the above also holds when R diverges along real values, so we may assume without loss of generality that $(R_n) \subset \mathbb{R}$, proving (3.2).

We now turn to show (3.3). We show that

$$N_{Q_1,\,\mathbb{R}^d\setminus B_L(Q_1)} \to 0 \qquad a.s. \tag{3.7}$$

Observe that $\forall L > 0$, $N_{Q_1,\,\mathbb{R}^d\setminus B_L(Q_1)} \le N_{Q_1,\,\mathbb{R}^d}$. As Q_1 has finite Lebesgue measure, $N_{Q_1,\,\mathbb{R}^d}$ has finite expectation; therefore, if (3.7) holds, then we may apply the Lebesgue dominated convergence theorem to arrive at (3.3). (3.7) holds as $N_{Q_1,\,\mathbb{R}^d\setminus B_L(Q_1)}$ is a.s. finite, and therefore there exists L large enough such that the equality $N_{Q_1,\,\mathbb{R}^d\setminus B_L(Q_1)} = 0$ holds almost surely. This, together with the fact that $\mathbb{E}[N_{Q_1,\,\mathbb{R}^d\setminus B_L(Q_1)}]$ is finite by standard theory of PPPs then implies that (3.3) holds. □

Chapter 4

Local energies

4.1 Energy bounds

We now turn to bounding the final term in (1.11), A_3. The main goal of this chapter is to prove the following lemma:

Lemma 4.1. *Let X, Y be PPPs on $\mathbb{R}^2$ that are matched according to a locally optimal, stationary and ergodic matching T°. Then, $\exists \beta > 0$ and an a.s. finite random radius r_* such that $\forall R \geq r_*$, the following holds:*

$$A_3 := \frac{1}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| > L}} |T^\circ(X) - X|^2 \leq \beta \log R$$

It is useful to introduce the *local energy at scale R*, $E(R)$, as is done in [10], as this quantity plays a central role in proving Lemma 4.1:

Definition 4.2 (Local energy)**.** For a matching T between two point processes X and Y on $\mathbb{R}^d$ and a given $R > 0$, the *local energy of T at scale R*, $E(R)$, is defined as

$$E(R) := \frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^2$$

Similarly to $D(R)$, the introduction of $E(R)$ can be motivated in terms of the heuristics discussed in 1.2. Recall that the optimal map between $T^*(X)$ is always given by the gradient of a convex function, so $T^*(X) = \nabla\varphi(x)$ for some convex φ. Indeed, $\frac{1}{R^2}E(R)$ being small is the quantitative condition needed to justify the ansatz (1.7) so that $T^*(X) \approx x + \varepsilon\nabla\psi(x)$, as outlined in [10]. Again, as with $D(R)$, even though we introduce $E(R)$ for PPPs in $\mathbb{R}^d$, all of the results concerning $E(R)$ will be derived in $\mathbb{R}^2$.

Remark 4.3. It is useful to think of the harmonic function $\Phi(x)$ as being "equal" to $\varepsilon\psi(x)$ for $\varepsilon \ll 1$; this is not technically correct as $\varepsilon\psi(x)$ comes from heuristic arguments, but both functions play the same role in the linearisation of the Monge-Ampère equation.

We now introduce the harmonic approximation theorem [10, Theorem 1.4], which quantitatively describes the quality of the approximation $\nabla\Phi(x) \approx T^*(x) - x$ given that $\frac{1}{R^2}E(R)$ and $\frac{1}{R^2}D(R)$ are both small. Furthermore, we expect to be able to control $|\Phi(x)|$ at a scale of order $O(R)$; such control is also explicitly achieved by this theorem.

Theorem 4.4 (Harmonic approximation). *Let X and Y be PPPs on $\mathbb{R}^d$ and T^* be a locally optimal matching between X and Y. Suppose that $\forall \tau > 0$ s.t. $\tau \ll 1$, $\exists \varepsilon_\tau > 0$ such that if the following holds*

$$\frac{1}{R^2}E(4R) + \frac{1}{R^2}D(4R) \leq \varepsilon_\tau \tag{4.1}$$

then $\exists C_\tau > 0$ and a harmonic function $\Phi(x)$ such that

$$\begin{aligned} \frac{1}{R^d} \sum_{X \in B_R \text{ or } T^*(X) \in B_R} |T^*(X) - X - \nabla\Phi(X)|^2 \leq \tau E(4R) + C_\tau D(4R) \\ \sup_{B_{2R}} |\nabla\Phi|^2 \leq C_\tau(E(4R) + D(4R)) \end{aligned} \tag{4.2}$$

i.e. $\nabla\Phi(X)$ is small in norm in B_{2R} and locally approximates the displacement $T^*(X) - X$

Note that the factor of 4 in $E(4R)$ and $D(4R)$ is not significant and is a mere technicality; $E(4R)$ and $D(4R)$ still capture the behaviour of the matching and the PPPs at a scale $O(R)$. In fact, the original result [10, Theorem 1.4] uses a factor of 6 as opposed to 4, which was recently improved in [21, Theorem 1.1]. Although this version of the theorem is slightly different to its original formulations in [10] and [21], it readily follows from setting the measures μ and ν as the counting measures of the PPPs X and Y, respectively. W

e now introduce a crucial proposition that, together with Theorem 4.4, will allow us to prove Lemma 4.1:

Proposition 4.5 (Edges are of length $< R$). *There is R large enough such that if $T^\circ(X) \in B_R$, then $X \in B_{2R}$.*

We will now first prove Lemma 4.1 and then show a result that both justifies that we can indeed apply Theorem 4.4 (by showing (4.1) holds) and verifies Proposition 4.5.

Proof of Lemma 4.1. The goal behind introducing the harmonic approximation theorem is as follows: suppose we may apply (4.2); we then have

$$\begin{aligned} A_3 &= \frac{1}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| > L}} |T^\circ(X) - X|^2 \\ &\leq \frac{2}{R^2} \sum_{X \in B_R \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X - \nabla\Phi(X)|^2 + \frac{2}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| > L}} |\nabla\Phi(X)|^2 \\ &\overset{(4.2)}{\leq} 2\tau E(6R) + 2C_\tau D(6R) + \frac{2}{R^2} \sum_{\substack{X \in B_R \text{ or } T^\circ(X) \in B_R \\ \text{and } |T^\circ(X) - X| > L}} |\nabla\Phi(X)|^2 \end{aligned} \tag{4.3}$$

Then, as we have assumed that both (4.2) and Proposition 4.5 hold, we have

$$\sum_{\substack{X\in B_R \text{ or } T^\circ(X)\in B_R\\ \text{and } |T^\circ(X)-X|>L}} |\nabla\Phi(X)|^2 \leq A_2 \sup_{X\in B_{2R}} |\nabla\Phi(X)|^2$$

which we can control thanks to the second line of (4.2); using this estimate gives

$$A_3 \leq 2\tau E(4R) + 2C_\tau D(4R) + \frac{2}{R^2} A_2 (C_\tau E(4R) + D(4R))$$

By Lemma 3.3, we may assume that R is large enough such that that the bound on A_2, (3.1), holds with τ instead of ε; note that by doing this, we have fixed a choice of $L = L_\tau$. Note that in this step the assumption that T° is stationary and ergodic is crucial, as we need these assumptions for Lemma 3.3 to hold. Therefore, we have

$$A_3 \leq 2\tau E(4R) + 2C_\tau D(4R) + 2\tau C_\tau (E(4R) + D(4R))$$

Notice that as in the proof of Lemma 3.3 we first choose L so that (3.4) holds and afterwards we pick R such that (3.5) holds, we may pick r_* large enough such that (3.5) still holds and such that, for all $R \geq r_*$,

$$L_\tau^2 \leq \log R \tag{4.4}$$

holds. This, together with the fact that

$$\begin{aligned} E(R) &= \frac{1}{R^2} \sum_{\substack{X\in B_R \text{ or } T^\circ(X)\in B_R\\ \text{and } |T^\circ(X)-X|>L_\tau}} |T^\circ(X)-X|^2 + \frac{1}{R^2} \sum_{\substack{X\in B_R \text{ or } T^\circ(X)\in B_R\\ \text{and } |T^\circ(X)-X|\leq L_\tau}} |T^\circ(X)-X|^2 \\ &= A_3 + \frac{1}{R^2} \sum_{\substack{X\in B_R \text{ or } T^\circ(X)\in B_R\\ \text{and } |T^\circ(X)-X|\leq L_\tau}} |T^\circ(X)-X|^2 \\ &\leq A_3 + \frac{L_\tau^2}{R^2} A_1 \end{aligned}$$

As we are in the case where $d = 2$, the estimate (2.20) holds. This, in conjunction with (4.3) and (4.4) gives:

$$E(R) \leq C\log R + 2\tau E(4R) + 2C_\tau D(4R) + 2\tau C_\tau (E(4R) + D(4R)) \tag{4.5}$$

Once more, as $d = 2$, we may enlarge R the inequality (2.15) holds with constant M instead of C, and such that:

$$D(4R) \leq \frac{\varepsilon}{2} R^2 \tag{4.6}$$

which holds as $\frac{\log R}{R^2} \to 0$ as $R \to \infty$; this will be useful when showing (4.1) holds. Therefore, rearranging (4.5) and applying (2.15) yields:

$$\begin{aligned} E(R) &\leq C\log R + 2\tau(1 + C_\tau)E(4R) + 2C_\tau(1+\tau)D(4R) \\ &\leq 2\tau(1+C_\tau)E(4R) + (2C_\tau(1+\tau) + K)\log R \end{aligned} \tag{4.7}$$

Where C, $M \leq K$ and where we have implicitly used that $R \geq \log 6$, which we can assume without loss of generality by possible enlargement of R. In the sequel, we now relabel τ and C_τ such that (4.7) holds in the form:

$$E(R) \leq \tau E(4R) + C_\tau \log R \tag{4.8}$$

Observe that (4.8) is a recursive inequality for $E(R)$, which we may exploit. Indeed, using (4.8) multiple times with $E(R)$, $E(4R)$, $E(4^2R)$, ... yields:

$$\begin{aligned} E(R) &\leq \tau E(4R) + C_\tau \log R \leq \tau^2 E(4^2R) + \tau C_\tau \log R + C_\tau \log R \\ &\leq \tau^k E(4^kR) + C_\tau \log R \sum_{i=0}^{k-1} \tau^i \end{aligned} \tag{4.9}$$

We now apply (4.1) to $\frac{1}{R^2}E(4R)$ with $R = 4^{k-1}R'$ to obtain

$$\frac{1}{4^{2k-2}R'^2}E(4^2kR') \leq \varepsilon \implies E(6^kR') \leq \frac{16^k}{16}\varepsilon R'^2 \leq 16^k \varepsilon R'^2$$

which, after relabelling R' and combined with (4.9) gives:

$$E(R) \leq \varepsilon(16\tau)^k R^2 + C_\tau \log R \sum_{i=0}^{k-1} \tau^i$$

Now, pick $\tau < \frac{1}{16}$ in Theorem (4.4) and send $k \to \infty$ to obtain

$$E(R) \leq C_\tau \log R \sum_{i=0}^{\infty} \tau^i \leq \beta \log R$$

where $\beta \geq C_\tau \sum_{i=0}^{\infty} \tau^i$, thus giving a bound on $E(R)$. As $E(R)$ dominates A_3, also shows that $A_3 \leq \beta \log R$. □

4.2 Edge control

Now, we must check the validity of Proposition 4.5 and that we may indeed apply Theorem 4.4 by showing (4.1) holds. These two claims follow from the lemma below, which shows that if $X \in (-R, R)^2$, then the matchings cannot be too long:

Lemma 4.6 (Edge control, [1, Lemma 2.2])**.** *$\forall \varepsilon > 0$, $\exists$ an a.s. finite random radius r_* such that $\forall R \geq r_*$, if $X \in (-R, R)^2$, then*

$$|T^\circ(X) - X| \leq \varepsilon R$$

Indeed, if we show Lemma 4.6, then we may apply it with $4R$ instead of R to obtain

$$\frac{1}{R^2} \sum_{X \in B_{4R}} |T^\circ(X) - X|^2 \leq \frac{|X \in B_{4R}|}{R^2} \varepsilon^2 R^2 \lesssim n_{R,\mu} \varepsilon^2 R^2$$

where we have used the definition of $n_{R,\,\mu}$. As we may assume without loss of generality that R is large enough such that $n_{R,\,\mu} \in \left[\frac{1}{2},\, 2\right]$, relabelling ε gives:

$$\frac{1}{R^2} \sum_{X \in B_{4R}} |T^\circ(X) - X|^2 \le \frac{\varepsilon}{4} R^2 \tag{4.10}$$

Now, by symmetry, we may also assume (possibly enlarging R once more) that Lemma 4.6 holds with $T^\circ(X)$ replaced with X and with $\varepsilon < 1$. This yields:

$$\frac{1}{R^2} \sum_{T^\circ(X) \in B_{6R}} |T^\circ(X) - X|^2 \le \frac{\varepsilon}{4} R^2$$

which, combined with (4.10), gives:

$$E(4R) = \frac{1}{R^2} \sum_{X \in B_{4R}} |T^\circ(X) - X|^2 \le \frac{\varepsilon}{4} R^2 + \frac{1}{R^2} \sum_{T^\circ(X) \in B_{4R}} |T^\circ(X) - X|^2 \le \frac{\varepsilon}{2} R^2$$

This inequality, together with (4.6), shows that (4.1) holds for $d = 2$, justifying our application of Theorem 4.2 in the proof of Lemma 4.1.

Furthermore, by choice of R and because $T^\circ(X) \in B_R \subset (-R,\, R)^2$, then:

$$|X| \le |T^\circ(X) - X| + |T^\circ(X)| \le (\varepsilon + 1)R < 2R \implies X \in B_{2R}$$

proving Proposition 4.5.

We have seen that verifying Lemma 4.6 is sufficient to validate our proof of Lemma 4.1; hence, we devote the remainder of this chapter to showing Lemma 4.6.

Proof of Lemma 4.6. Fix a small $\varepsilon > 0$, i.e. $0 < \varepsilon \ll 1$. We begin by assuming r_* is large enough such that $\forall R \ge r_*$:

1. Lemma 3.3 holds for ε with $(-2R,\, 2R)^2$ instead of B_R,
2. (2.16) holds with $2R$ instead of R,
3. $L \le \varepsilon r_*$

where we will use the first two conditions in the following form

$$\begin{aligned} &|\{X \in (-2R,\, 2R)^2 \text{ and } |T^\circ(X) - X| > L\}| \le (\varepsilon 4R)^2 \\ &W_2^2\big|_{(-2R,\, 2R)^2} (\mu,\, n_{(-2R,\, 2R),\, \mu}) + \frac{(4R)^4}{n_{(-2R,\, 2R),\, \mu}} (n_{(-2R,\, 2R),\, \mu} - 1)^2 \le (\varepsilon 4R)^4 \end{aligned} \tag{4.11}$$

Observe that the second line of (4.11) immediately implies that $n_{(-2R,\, 2R),\, \mu} \in \left[\frac{1}{2},\, 2\right]$. Furthermore, note we are once more using stationarity and ergodicity assumptions on T° when bounding A_3, as Lemma 3.3 requires these assumptions to hold.

Hereon, we fix a realisation and an $R \geq r_*$. We now show that for a square Q of *mesoscopic* side length r, that is, r satisfies

$$\varepsilon R \ll r < 2R, \tag{4.12}$$

one has:

$$r^2 \lesssim |\{X \in Q\}| \tag{4.13}$$

Let η be a Lipschitz function with Lipschitz constant $M_\eta \lesssim \frac{1}{r}$ that is also compactly supported in Q and such that $0 \leq \eta \leq 1$ and $\eta \geq \frac{1}{2}$ on $Q_{r/2}$, the square of side length $r/2$; in particular, these conditions translate to the following inequalities:

$$\begin{aligned} \int_Q \eta dx &\geq \int_{Q_{r/2}} \eta dx \geq \frac{1}{4} r^2 \\ \int_Q \eta d\mu &\leq \int d\mu = |\{X \in Q\}| \end{aligned} \tag{4.14}$$

Now let π be a coupling between μ and $n_{(-2R,\,2R),\,\mu} dx$; then

$$\begin{aligned} \left| \int_Q \eta d\mu - \int_Q \eta n_{(-2R,\,2R),\,\mu} dx \right| &= \left| \int_Q \eta(x) - \eta(y) d\pi(x,\,y) \right| \\ &\leq \int_Q |\eta(x) - \eta(y)| d\pi(x,\,y) \\ &\leq M_\eta \int_Q |x-y| d\pi(x,\,y) \\ &\leq M_\eta \left(\int_Q |x-y|^2 d\pi(x,\,y) \right)^{1/2} \left(\int_Q 1 d\pi(x,\,y) \right)^{1/2} \\ &\leq M_\eta \, W_2|_{(-2R,\,2R)^2} (\mu,\, n_{(-2R,\,2R),\,\mu}) \left(\int_Q d\mu + n_{(-2R,\,2R),\,\mu} |Q| \right)^{1/2} \end{aligned} \tag{4.15}$$

where we have used the triangle inequality, the Lipschitz property, Minkowski's inequality and the definitions of the 2-Wasserstein distance and the coupling π, in that order. We may develop (4.15) further to recover a W_2^2 term, as

$$\begin{aligned} &M_\eta \, W_2|_{(-2R,\,2R)^2} (\mu,\, n_{(-2R,\,2R),\,\mu}) \left(\int_Q d\mu + n_{(-2R,\,2R),\,\mu} |Q| \right)^{1/2} \\ &\leq M_\eta^2 \, W_2^2\big|_{(-2R,\,2R)^2} (\mu,\, n_{(-2R,\,2R),\,\mu}) + |\{X \in Q\}| + n_{(-2R,\,2R),\,\mu} |Q| \end{aligned}$$

which when combined with (4.14) and (4.15) gives

$$\frac{r^2}{4} \leq M_\eta^2 \, W_2^2\big|_{(-2R,\,2R)^2} (\mu,\, n_{(-2R,\,2R),\,\mu}) + 2|\{X \in Q\}| + n_{(-2R,\,2R),\,\mu} r^2$$

an inequality that, together with the fact that R is large enough such that (4.11) holds and due to our choice of η, gives

$$\frac{r^2}{4} \lesssim \frac{(4\varepsilon R)^4}{r^2} + 2|\{X \in Q\}| + 2$$

which, when combined with (4.12) gives (4.13).

It should be noted that (4.13) implies that for all squares Q of side length r satisfying (4.12), one has

$$\exists X \in Q \text{ s.t. } |T^\circ(X) - X| \leq L \tag{4.16}$$

This result follows by contradiction: suppose (4.16) does not hold for some square Q of side length satisfying (4.12). Then, all of the Poisson points in Q are transported a distance greater than L; i.e.

$$|\{X \in Q\}| = |\{X \in Q \text{ s.t. } |T^\circ(X) - X| > L\}|$$

As (4.13) holds for such a square Q and r is such that (4.12) then, for such a Q

$$\begin{aligned} (4\varepsilon R)^2 &< |\{X \in Q \text{ s.t. } |T^\circ(X) - X| > L\}| \\ &\overset{(4.12)}{\leq} |\{X \in (-2R,\, 2R)^2 \text{ s.t. } |T^\circ(X) - X| > L\}| \\ &\overset{(4.11)}{\leq} (4\varepsilon R)^2 \end{aligned}$$

giving a contradiction, so (4.16) holds.

Now, we will bolster (4.16) by showing that, given $X \in (-R,\, R)^d$, there are at least three Poisson points transported less than distance L that are evenly spread out around X and whose distance from X is $O(r)$, where r is as in (4.12). Indeed, we will show the existence of three Poisson points X_1, X_2 and X_3 and a small $\rho \ll 1$ such that $\forall n = 1,\, 2,\, 3$

$$|T^\circ(X_n) - X_n| \leq L \tag{4.17}$$

$$|X_n - X| \sim r \tag{4.18}$$

$$B_\rho \subset \text{Hull}\left(\left\{\frac{X_n - X}{|X_n - X|}\right\}_{n=1}^{3}\right) \tag{4.19}$$

The proof of this statement is highly geometric; we begin by considering the symmetric trisection[1] of the plane at X. Now rotate the trisection by a small angle α, namely $\alpha \ll \frac{\pi}{3}$; we will call the small cones formed by the area between these two planar trisections σ_i. Furthermore, let $\zeta_i := \sigma_i \cap (B_{2r}(X) \setminus B_r(X))$, so that ζ_i is the intersection of σ_i with the annulus of radii r and $2r$ centred at X; this construction is shown in Figure 4.1. As α is small enough, i.e. if $\alpha \ll 1$, then if we select three unit vectors e_1, e_2 and e_3 such that $e_i \in \sigma_i$, then $\exists \rho \ll 1$ such that $B_\rho(X) \subset \text{Hull}(\{e_1,\, e_2,\, e_3\})$; such a $B_\rho(X)$ is displayed in Figure 4.2. Notice that ρ is bounded away from zero and independent from r. Now observe ζ_i contains a square of side length $O(r)$, and clearly $\zeta_i \subset (-2R,\, 2R)^2$. Therefore, we may apply (4.16) to each of the ζ_i to conclude that $\exists X_i \in \zeta_i$, where X_i is a Poisson point, showing (4.17). Furthermore, as $\zeta_i \subset (B_{2r}(X) \setminus B_r(X))$, (4.18) holds. Finally, picking the unit vectors $e_i = \frac{X_i - X}{|X_i - X|}$ verifies (4.19).

[1]By the symmetric trisection at a point $\omega \in \mathbb{R}^2$ we mean three straight lines emerging from ω, each at an angle $\pi/3$ from each other.

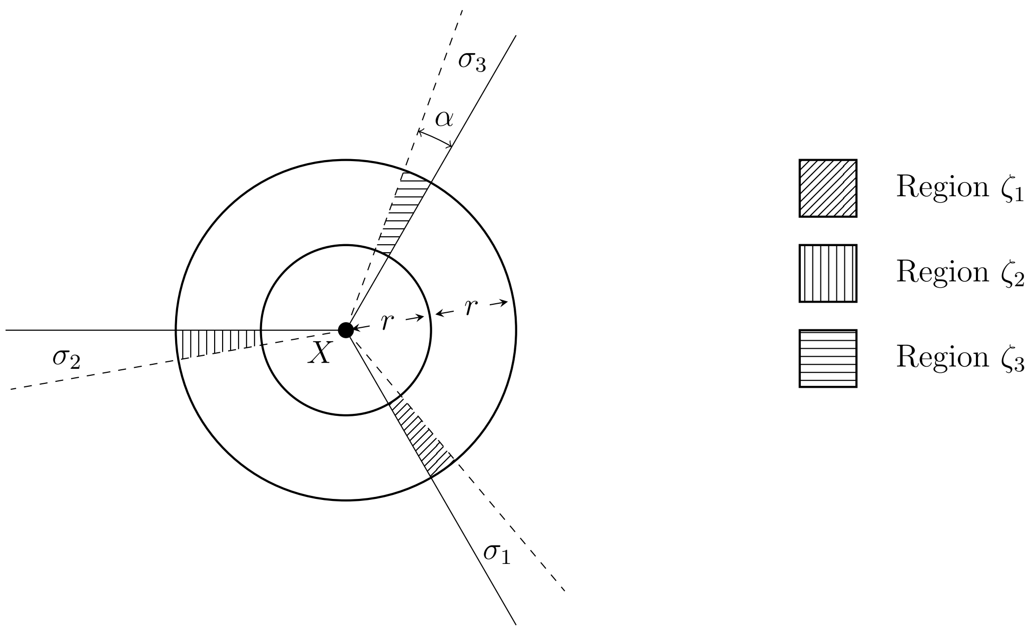

Figure 4.1: Diagram showing the symmetric trisection of the plane at X, rotated by angle α, illustrating the regions σ_i with their respective annular intersections.

With this in hand, we can now finish the proof and show Lemma 4.6. Let $X \in (-R, R)^2$ be a Poisson point, and find the three Poisson points X_1, X_2 and X_3 around it such that (4.17), (4.18) and (4.19) hold. As T is cyclically monotone, it is monotone; to see this, simply pick $N = 2$ in Definition 1.5 to obtain, for general $T \in \mathcal{T}$:

$$T(X_1)^\top (X_1 - X_2) + T(X_2)^\top (X_2 - X_1) \geq 0 \implies (T(X_1) - T(X_2))^\top (X_2 - X_1) \geq 0$$

We will use monotonicity as follows:

$$\begin{aligned}(T^\circ(X) - X)^\top (X_n - X) &= ([T^\circ(X_n) - X_n] + [T^\circ(X) - T^\circ(X_n)] + [X_n - X])^\top (X_n - X) \\ &\leq (T^\circ(X_n) - X_n)^\top (X_n - X) + |X_n - X|^2 \\ &\lesssim |T^\circ(X_n) - X_n|^2 + |X_n - X|^2\end{aligned} \tag{4.20}$$

where we have used Young's inequality in the last inequality. Now, as (4.17) holds and $L \leq \varepsilon R$, we use (4.12) to obtain

$$|T(X_n) - X_n| \leq r$$

which combined with (4.18) and (4.20) gives:

$$(T^\circ(X) - X)^\top \frac{X_n - X}{|X_n - X|} \lesssim |X_n - X| + \frac{|T^\circ(X_n) - X_n|^2}{|X_n - X|} \lesssim r$$

Finally, observe that as (4.19) holds, and as for any $v \in B_\rho(0)$, $v = \frac{1}{\rho} e$ for $e \in B_1(0)$, then:

$$e = \sum_{i=1}^{3} \lambda_n \frac{X_n - X}{|X_n - X|}, \quad \sum_{i}^{3} \lambda_i = \frac{1}{\rho}$$

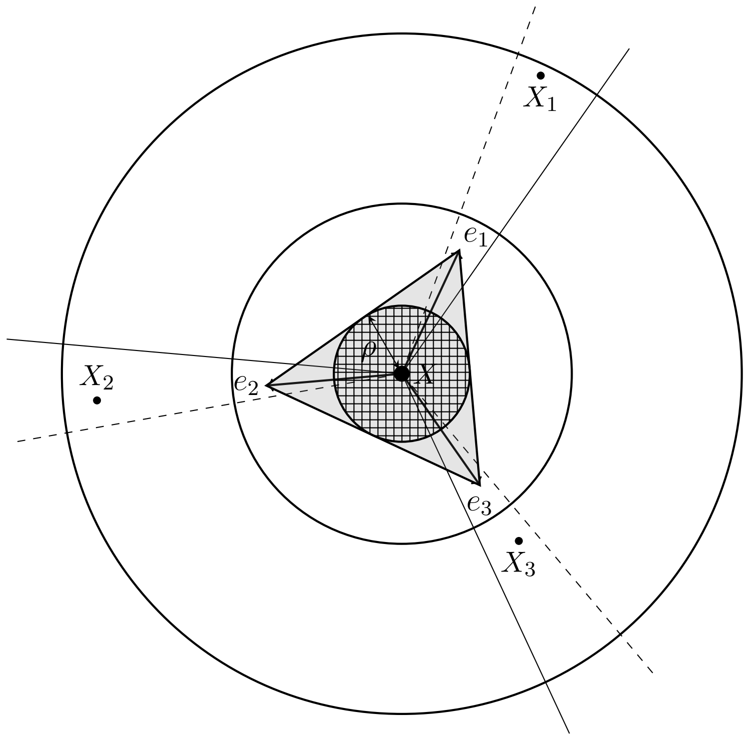

Figure 4.2: Illustration of convex hull and inscribed ball within the regions σ_i. The shaded area depicts the convex hull formed by the unit vectors e_1, e_2, and e_3. The grid-patterned circle highlights the ball with radius ρ and centred at X that can be inscribed within the hull of the unit vectors. The three Poisson points $X_i \in \zeta_i$ near X are also displayed.

and therefore, for all unit vectors e, we have:

$$(T^\circ(X) - X)^\top e = \sum_{i=1}^{3} \lambda_n (T^\circ(X) - X)^\top \frac{X_n - X}{|X_n - X|} \lesssim \frac{r}{\rho}$$

Now, as ρ is both bounded away from zero and independent from r, we may conclude that for all unit vectors e:

$$(T^\circ(X) - X)^\top e \lesssim r$$

which, together with the choice $e = \frac{T^\circ(X)-X}{|T^\circ(X)-X|}$, yields

$$|T^\circ(X) - X| \lesssim r$$

As this holds for all r such that (4.12) applies, we finally deduce that for all Poisson points $X \in (-R, R)$,

$$|T^\circ(X) - X| \leq \varepsilon R$$

proving Lemma 4.6. □

Therefore, our proof of Lemma 4.1 is now complete, successfully bounding A_3.

Chapter 5

Closing remarks & Further work

5.1 Remarks

5.1.1 Final estimate

We are now prepared to show the upper bound (1.10). First, fix $\varepsilon > 0$ and let $r_{*,1}$ be the random radius such that Corollary 2.2 holds. Furthermore, let $r_{*,2}$ and $r_{*,2}$ be the random radii such that Lemma 3.3 and Lemma 4.1 hold. Finally, there also exists an a.s. finite random radius $r_{*,4}$ such that $\forall R \geq r_{*,4}$, $L_\varepsilon \leq \sqrt{\varepsilon \log R}$. Then, for all R such that $R \geq \max(r_{*,1}, r_{*,2}, r_{*,3}, r_{*,4})$, applying all of the aforementioned bounds to (1.12) yields

$$\frac{1}{R^2} \sum_{X \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X| \leq CL + \sqrt{\varepsilon\beta \log R} \leq 2C\sqrt{\beta}\sqrt{\varepsilon}\sqrt{\log R}$$

Taking ε' such that $2C\sqrt{\beta}\sqrt{\varepsilon} = \varepsilon'$ yields:

$$\frac{1}{R^2} \sum_{X \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X| \leq CL + \sqrt{\varepsilon\beta \log R} \leq \varepsilon'\sqrt{\log R}$$

proving that indeed

$$\frac{1}{R^2} \sum_{X \text{ or } T^\circ(X) \in B_R} |T^\circ(X) - X| \leq o(\sqrt{\log R})$$

which, together with the lower bound (1.9) shows the nonexistence of T° in $\mathbb{R}^2$.

5.1.2 Lower bound

Notice that to show the non-existence of T°, it is crucial for the lower bound (1.9) to hold. Indeed, the lower bound is, essentially, the lower bound presented [13], adapted to the square $(0, R)^2$. The adapted bound is presented in [1] as

Lemma 5.1 (Lower bound [1, Lemma 2.5]). *Let X and Y be two PPPs on $\mathbb{R}^2$ with unit intensity. Then, $\exists C > 0$ and an a.s. finite random radius $r_* < \infty$ s.t. $\forall R \geq r_*$*

$$S_R(X, Y) := \sup_{\substack{\zeta \in C^\infty: \\ |\nabla\zeta| \leq 1, \\ \int \zeta\, dx = 0, \\ \operatorname{supp}(\zeta) \subset (0,R)^2}} \left[\frac{1}{R^2} \sum_{X,\, Y \in (0,\, R)^2} |\zeta(X) - \zeta(Y)| \right] \geq C\sqrt{\log R}$$

This lemma immediately implies the lower bound (1.9). We will not directly address the proof of this lemma here, as the novelty of [1] is the application of optimal transport methods and Monge-Ampère linearisation techniques to the Poisson matching problem. Still, we note that it is very similar to the proof of Lemma 2.5. Indeed, the lower bound is firstly shown in expectation by using dyadic square partitions of $(0,\, R)^2$, and then lifted to an a.s. bound via the following proposition [1]:

Proposition 5.2. *Let $S_R(X, Y)$ be as above. Then*

$$\lim_{R\uparrow\infty,\, R\in\mathcal{D}} \frac{1}{R^2\sqrt{\log R}} |S_R(X, Y) - \mathbb{E}[S_R(X, Y)]| = 0 \quad a.s.$$

The proposition above can easily be shown via a Borel-Cantelli argument, as S_R satisfies the following concentration inequalities [1]:

Proposition 5.3. *Let $S_R(X, Y)$ be as above, and $\mathbb{E}_X$ and $\mathbb{E}_Y$ be the expectations with respect to to the PPPs X and Y, accordingly. Then the following concentration inequalities hold:*

$$\mathbb{P}\left(\frac{1}{R^2\sqrt{\log R}} |S_R(X, Y) - \mathbb{E}[S_R(X, Y)]| > \varepsilon \right) \lesssim e^{-\frac{\varepsilon^2}{4}\log R}$$

$$\mathbb{P}\left(\frac{1}{R^2\sqrt{\log R}} |\mathbb{E}_X[S_R(X, Y)] - \mathbb{E}[S_R(X, Y)]| > \varepsilon \right) \lesssim e^{-\frac{\varepsilon^2}{4}\log R}$$

Where $\mathbb{E}_X$ is the expectation with respect to the PPP X only.

For more details on the lower bound, we encourage the reader to consult [13] and [1, Lemma 2.5].

5.1.3 Generalisations

Hereon, let T°_γ be a stationary, ergodic and locally optimal matching in $\mathbb{R}^d$ with respect to the γ-power of the norm, i.e. for any $\tilde{T}_\gamma$ differing from T°_γ on finitely many points, we have:

$$\sum_X |T^\circ_\gamma(X) - X|^\gamma - \sum_X |\tilde{T}_\gamma(X) - X|^\gamma \leq 0$$

Recently, it was shown in [2] that the same techniques applied to show the upper bound (1.10) can be applied to show the non-existence of T°_γ when $\gamma > 1$, as all of the arguments presented above generalise to arbitrary powers, including a bound on the local toric γ-Wasserstein distance of the form:

Lemma 5.4 (Locally toric γ-Wasserstein growth [2, Lemma 2.1])**.** *Let* $\widetilde{W}|_{[0,\,R)^d;\,\gamma}\,(\mu, n_{[0,R)^d),\,\mu})$ *be the γ-Wasserstein distance restricted to the d-dimensional torus* $[0,\, R)^d$*. Then*

$$\widetilde{W}|_{[0,\,R)^d;\,\gamma}\,(\mu, n_{[0,R)^d),\,\mu}) \leq CR^d \begin{cases} (\log R)^{\gamma/2} & \text{if } d = 2 \\ 1 & \text{otherwise} \end{cases}$$

This proposition, together with the following bound:

Proposition 5.5. *Let $\gamma > 1$ and $n_{R,\,\mu}$ be defined as before. Then:*

$$\frac{R^\gamma}{n_{R,\,\mu}}|n_{R,\,\mu} - 1|^\gamma \leq C \begin{cases} (\log R)^{\gamma/2} & \text{if } d = 2 \\ 1 & \text{otherwise} \end{cases}$$

Gives the full generalised version of Lemma 2.8:

Theorem 5.6 (Local γ-distance growth [2, Theorem 1.7])**.** *Let X and Y be two PPPs on $\mathbb{R}^d$ with counting measures μ and ν, and, for $\gamma > 1$, define the local γ-distance of scale R, $D_\gamma(R)$ as:*

$$D_\gamma(R) := \frac{1}{R^d}\, W_\gamma^\gamma\big|_{B_R}\,(\mu,\, n_{R,\,\mu}) + \frac{R^\gamma}{n_{R,\,\mu}}(n_{R,\,\mu} - 1)^2 + \frac{1}{R^d}\, W_\gamma^\gamma\big|_{B_R}\,(\nu,\, n_{R,\,\nu}) + \frac{R^\gamma}{n_{R,\,\nu}}(n_{R,\,\nu} - 1)^2$$

Then

$$D_\gamma(R) \leq CR^d \begin{cases} (\log R)^{\gamma/2} & \text{if } d = 2 \\ 1 & \text{otherwise} \end{cases}$$

A crucial step in generalising the methods outlined in our proof of the upper bound is an extension to the harmonic approximation theorem to γ-Wasserstein distances as presented in [21]; first, we introduce the *local γ-energy*, $E_\gamma(R)$, which plays a role akin to $E(R)$ in [2]:

Definition 5.7 (Local γ-energy)**.** For a matching T between two point processes X and Y on $\mathbb{R}^d$, a $\gamma > 1$ and a given $R \in \mathbb{R}$, the *local γenergy of scale R of T*, $E_\gamma(R)$, is defined as

$$E_\gamma(R) := \frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^\gamma$$

With this in hand, we introduce the γ-harmonic approximation theorem:

Theorem 5.8 (γ-Harmonic approximation)**.** *Let X and Y be PPPs on $\mathbb{R}^d$ and T^* be a locally optimal matching between X and Y. Suppose that $\forall \tau > 0$ s.t. $\tau \ll 1$, $\exists \varepsilon_\tau > 0$ such that if the following holds*

$$\frac{1}{R^\gamma}E_\gamma(4R) + \frac{1}{R^\gamma}D_\gamma(4R) \leq \varepsilon$$

then $\exists C_\tau > 0$ *and a* γ*-harmonic function*[1] $\Phi(x)$ *such that*

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T^\circ(X) \in B_R} \left| T^*(X) - X - |\nabla\Phi(X)|^{\gamma'-2}\nabla\Phi(X) \right|^\gamma \leq \tau E_\gamma(4R) + C_\tau D_\gamma(4R)$$

$$\sup_{B_{2R}} |\nabla\Phi|^{\gamma'} \leq C_\tau(E_\gamma(4R) + D_\gamma(4R))$$

where γ' *is the conjugate exponent of* γ*.*

This theorem, alongside the estimates on the number of long and short edges presented in Corollary 2.2 & Lemma 3.3 give an upper bound of $o((\log R)^{\gamma/2})$ when $d = 2$. In particular, it is worth noting that the techniques used to arrive at Corollary 2.2 are fundamentally different, as T°_γ no longer has cyclically monotone support for general $\gamma > 1$, so one cannot rely on the monotonicity properties of T°_2 employed in our proof.

Finally, the bound is arrived at by using Hölder's inequality for ℓ_p spaces in a similar fashion as in (1.12) and bounding A_1, A_2 and a quantity analogous to A_3, but with an exponent of γ instead of 2.

5.2 Further work

The non-existence of a quadratic and, more generally, γ-locally optimal (for $\gamma > 1$) stationary and ergodic matching between two PPPs in $\mathbb{R}^2$ presents a plethora of new avenues for exploration.

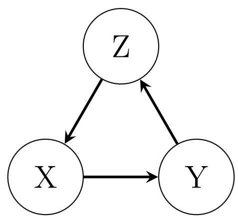

Figure 5.1: Directed graph showing the permissible matching directions between the PPPs X, Y and Z.

A natural extension of the non-existence of T° for two PPPs is to consider the same setting with three or more PPPs, which we call the *multiple PPP matching problem*; for example, let X, Y and Z be three homogeneous PPPs of unit intensity on $\mathbb{R}^2$ such that any point in X may be matched onto any point in Y or Z, and similarly for the remaining PPPs. One could then ask if a locally optimal stationary and ergodic matching exists between the points in X, Y and Z.

This question becomes trivial if we restrict the matching of the PPPs in a directed way; these restrictions can be represented as a directed graph with 3 nodes as done in Figure 5.1. In the particular setting of Figure 5.1, note that by the non-existence of T°,

[1]A function u is *γ-harmonic* if it solves $\Delta_\gamma u = \text{div}(|\nabla u|^{\gamma-2}\nabla u) = 0$

a stationary, ergodic and locally-optimal matching between X, Y and Z cannot exist, as by removing one of the PPPs from the diagram shows that the problem essentially decomposes into three instances of the Poisson matching problem with two PPPs. Nevertheless, if one makes the graph in Figure 5.1 undirected, the question becomes non-trivial because removing one of the PPPs does not necessarily yield two independent PPPs, as this removal may result in isolated points within the remaining point sets.

Another variant of the multiple PPP matching problem is to consider 3 PPPs X, Y and Z as before, but with different intensities. In this scenario, one can ask if there exists a stationary, ergodic and locally optimal matching between the three PPPs. Both of these topics, along with related questions of interest, are comprehensively examined in [22]. However, the local optimality of these matchings remains unaddressed, and optimal transport techniques have not been applied to these issues, highlighting the potential for further research in this area.

Furthermore, one can easily generalise the problem of matching multiple PPPs using a stationary, ergodic and locally optimal matching to higher-order types of matchings, not only to pairwise point matchings; in these higher-order matchings, one must define a notion of "cost of a matching", such as the sum of the distance between all of the points being matched, or the volume of the simplex of the region between the edges of the points being matched. Again, [22] explores higher-order matchings, but it refrains from assigning a cost to these matchings as their focus lies outside the realm of optimality. Just as the Poisson matching problem with two PPPs has multiple connections to the theory of bipartite graphs, extending the problem to more than PPPs and higher-order matchings gives clear connections to the n-colorability of hypergraphs.

The principal obstacle in applying the linearisation of the Monge-Ampère & optimal transport techniques presented throughout this dissertation to more than two PPPs and higher-order matchings is the relative youth of *multi-marginal optimal transport* [23], which generalises optimal transport theory to more than two marginal distributions. As in regular optimal transport, multi-marginal optimal transport typically refers to one of two problems, the *multi-marginal Kantorovich problem* or the *multi-marginal Monge problem*, the former being a relaxation of the latter. For our purposes, we are interested in solving the multi-marginal Monge problem, defined below:

Definition 5.9 (Multi-marginal Monge problem)**.** Let $\lambda_1, \ldots, \lambda_n$ probability measures on $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and $c(x_1, \ldots, x_n) : \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \to \mathbb{R}$ be a continuous cost function. Define $M(\lambda_1, \ldots, \lambda_n)$ to be the set of all $(n-1)$ tuple of maps $(F_2, \ldots, F_n)$ such that $F_i \# \lambda_1 = \lambda_i$. Then, an $(n-1)$-tuple of maps $(F_2^*, \ldots, F_n^*)$ solves the *multi-marginal Monge problem with cost c* if it satisfies:

$$(F_2^*, \ldots, F_n^*) = \inf_{(F_2, \ldots, F_n) \in M(\lambda_1, \ldots, \lambda_n)} \int_{\mathcal{X}_1 \times \cdots \times \mathcal{X}_n} c(x_1, F_2(x_2), \ldots, F_n(x_n)) d\lambda_1 d\lambda_2 \ldots d\lambda_n$$

A solution to the multi-marginal Monge optimal problem $(F_2^*, \ldots, F_n^*)$ is called a *Monge tuple*.

In our setting, given $X_1, \ldots, X_n$ PPPs of unit intensity, observe that each F_i^* in the Monge tuple is precisely the optimal map between the PPPs X_1 and X_i.

Nevertheless, to date, multi-marginal optimal transport lacks the same connections to the Monge-Ampère equation that regular optimal transport enjoys, which greatly obfuscates the intuition present in the original problem. In addition, no metric properties arising from multi-marginal optimal transport have proven useful in the Poisson optimal matching context. Moreover, a comprehensive analytic characterization of the constituents of the Monge tuple remains elusive, and while conditions for the existence of such a Monge tuple have been investigated [24], they are not yet exhaustive like in the regular optimal transport literature. Nevertheless, the conditions necessary for the existence of a Monge tuple have recently been established for cases where the involved measures are m-empirical (measures that are compactly supported on m points) [25], a result which is the first of its kind and could serve as a gateway to integrating multi-marginal OT techniques into the PPP matching literature.

Bibliography

[1] M. Huesmann, F. Mattesini, and F. Otto. "There is no stationary cyclically monotone Poisson matching in 2d". In: *Probability Theory and Related Fields* 187.3-4 (Dec. 2023), pp. 629–656.

[2] M. Huesmann, F. Mattesini, and F. Otto. "There is no stationary p-cyclically monotone Poisson matching in 2D". In: *arXiv: Probability* (Nov. 2023).

[3] A. E. Holroyd, S. Janson, and J. Wästlund. "Minimal matchings of point processes". In: *Probability Theory and Related Fields* 184.1-2 (Oct. 2022), pp. 571–611.

[4] D. Stoyan. *Stochastic geometry and its applications.* Ed. by S. N. Chiu, W. S. Kendall, and J. Mecke. Third edition. Chichester: Wiley, 2013.

[5] C. Villani. *Topics in optimal transportation.* Providence, RI: American Mathematical Society, 2003.

[6] L. A. Caffarelli. "The Monge-Ampère Equation and Optimal Transportation, an elementary review". In: *Optimal Transportation and Applications: Lectures given at the C.I.M.E. Summer School, held in Martina Franca, Italy, September 2-8, 2001.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 1–10.

[7] J. C. Loftin. "Survey on Affine Spheres". In: *arXiv: Differential Geometry* (2008).

[8] H. Minkowski. "Volumen und Oberfläche". In: *Mathematische Annalen* 57 (1903), pp. 447–495.

[9] E. V. Shikin. "The problem of regular isometric embeddings and the Monge-Ampère equation of hyperbolic type". In: *Journal of Mathematical Sciences* 94.2 (1999), pp. 1254–1260.

[10] M. Goldman, M. Huesmann, and F. Otto. "Quantitative Linearization Results for the Monge-Ampère Equation". In: *Communications on Pure and Applied Mathematics* 74 (2019).

[11] S. Caracciolo et al. "Scaling hypothesis for the Euclidean bipartite matching problem". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 90.1 (July 2014), p. 012118.

[12] J. R. Magnus. *Matrix differential calculus with applications in statistics and econometrics.* Ed. by H. Neudecker. Third edition. Hoboken, NJ: Wiley, 2019.

[13] M. Ajtai, J. Komlós, and G. Tusnády. "On optimal matchings". In: *Combinatorica* 4.4 (Dec. 1984), pp. 259–264.

[14] M. Pollicott. *Dynamical systems and ergodic theory.* Ed. by M. Yuri. Cambridge: Cambridge University Press, 1998.

[15] O. Goldreich. *Introduction to Property Testing.* Cambridge: Cambridge University Press, 2017.

[16] J. D. Benamou and Y. Brenier. “A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem”. In: *Numerische Mathematik* 84.3 (2000), pp. 375–393.

[17] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*. First edition. Basel: Birkhäuser Basel, 2006.

[18] L. C. Evans. *Partial Differential Equations*. Second edition. Providence. RI: American Mathematical Society, 2010.

[19] S. N. Cohen and R. J. Elliott. *Stochastic Calculus and Applications*. Second edition. New York City, NY: Springer New York, 2015.

[20] M. Goldman, M. Huesmann, and F. Otto. “A large-scale regularity theory for the Monge-Ampere equation with rough data and application to the optimal matching problem”. In: *arXiv: Analysis of PDEs* (2018).

[21] L. Koch. “Geometric linearisation for optimal transport with strongly p-convex cost”. In: *Calculus of Variations and Partial Differential Equations* 63.4 (Apr. 2024), pp. 1–33.

[22] G. Amir, O. Angel, and A. E. Holroyd. “Multicolour Poisson matching”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* 57.4 (Nov. 2016), pp. 1811–1833.

[23] B. Pass. “Multi-marginal optimal transport: theory and applications”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 49.6 (May 2014), pp. 1771–1790.

[24] B. Pass and A. Vargas-Jiménez. “Monge solutions and uniqueness in multi-marginal optimal transport: weaker conditions on the cost, stronger conditions on the marginals”. In: *arXiv: Optimisation and Control* (Feb. 2022).

[25] P. Emami and B. Pass. “On the existence of Monge solutions to multi-marginal optimal transport with quadratic cost and uniform discrete marginals”. In: *arXiv: Optimisation and Control* (Jan. 2024).