There is no stationary cyclically monotone Poisson matching in 2d

Martin Huesmann · Francesco Mattesini · Felix Otto

Abstract
We show that there is no cyclically monotone stationary matching of two independent Poisson processes in dimension $d = 2$. The proof combines the harmonic approximation result from Goldman et al. (Commun. Pure Appl. Math. 74:2483–2560, 2021) with local asymptotics for the two-dimensional matching problem for which we give a new self-contained proof using martingale arguments.

Mathematics Subject Classification
60D05 · 60G55 · 52A22 · 49Q22

1 Introduction

Consider two locally finite point sets $\{X\}, \{Y\} \subset \mathbb{R}^d$ in $d$-dimensional space. We are interested in their matching, which we think of as a bijection $T$ from $\{X\}$ onto $\{Y\}$. More specifically, we are interested in their matching by cyclically monotone maps

---

1 And thus countable.

All authors are supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the SPP 2265 Random Geometric Systems. MH and FM have been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2044 -390685587, Mathematics Münster: Dynamics–Geometry–Structure.
$T$, which means that for any finite subset $^2 \{X_n\}_{n=1}^N$ we have
\[ \sum_{n=1}^{N} T(X_n) \cdot (X_n - X_{n-1}) \geq 0 \quad \text{with} \quad X_0 := X_N. \] (1.1)

It is elementary to see that (1.1) is equivalent to local optimality, meaning
\[ \sum_{X} (|T(X) - X|^2 - |	ilde{T}(X) - X|^2) \leq 0 \] (1.2)
for any other bijection $\tilde{T}$ that differs from $T$ only on a finite number of points. $^3$ This makes a connection to the optimal transportation between the measures
\[ \mu = \sum_{X} \delta_X \quad \text{and} \quad \nu = \sum_{Y} \delta_Y \] (1.3)
related via $T\#\mu = \nu$, which we shall explore in this paper. Note however that because of the (typically) infinite number of points, we cannot view $T$ as a minimizer.

Before proceeding to the random setting, we make two simple observations that show that the set of $(\{X\}, \{Y\}, T)$ is rich: For $d = 1$, (1.1) is easily seen to be equivalent to plain monotonicity. Every single matching $T(X_0) = Y_0$ can obviously be extended in a unique way to a monotone bijection $T$ of $\{X\}$ and $\{Y\}$, so that for $d = 1$, the set of monotone bijections $T$ has the same magnitude as $\{X\}$ itself. Returning to general $d$, we note that for any monotone bijection $T$ of $\{X\}$ and $\{Y\}$, and for any two shift vectors $\tilde{x}, \tilde{y} \in \mathbb{R}^d$, the map $x \mapsto T(x - \tilde{x}) + \tilde{y}$ is a monotone bijection of the shifted point sets $\{\tilde{x} + X\}$ and $\{\tilde{y} + Y\}$.

We are interested in the situation when the sets $\{X\}, \{Y\}$ and their cyclically monotone bijection $T$ are random. More precisely, we consider the case when $\{X\}$ and $\{Y\}$ are independent Poisson point processes of unit intensity. We assume that the $\sigma$-algebra for $(\{X\}, \{Y\}, T)$ is rich enough so that the following elementary observables are measurable, namely the number $N_{U, V}$ of matched pairs $(X, Y) \in U \times V$ for any two Lebesgue-measurable sets $U, V \subset \mathbb{R}^d$ (with $U$ or $V$ having finite Lebesgue measure$^4$):
\[ N_{U, V} := \#\{(X, Y) \in U \times V | Y = T(X)\} \in \{0, 1, \ldots\}. \]

We now come to the crucial assumption on the ensemble: In view of the above remark, the additive group $\mathbb{Z}^d \ni \tilde{x}$ acts on $(\{X\}, \{Y\}, T)$ via
\[ (\{X\}, \{Y\}, T) \mapsto ([\tilde{x} + X], \{\tilde{x} + Y\}, T(\cdot - \tilde{x}) + \tilde{x}). \]

$^2$ Which we momentarily enumerate.
$^3$ So that the sum is effectively finite.
$^4$ So that the following number is finite.
We assume that this action is stationary and ergodic. On the one hand, stationarity is a structural assumption; we shall only use it in following form: For any shift vector \( \bar{x} \), the random natural numbers \( N_{\bar{x} + U, \bar{x} + V} \) and \( N_{U, V} \) have the same distribution. On the other hand, ergodicity is a qualitative assumption; we will only use it in form of the following application of Birkhoff’s ergodic theorem:

\[
\lim_{R \uparrow \infty} \frac{1}{R^d} \sum_{\bar{x} \in \mathbb{Z}^d \cap [0, R]^d} N_{\bar{x} + U, \bar{x} + V} = \mathbb{E} N_{U, V} \quad \text{almost surely.}
\]

**Theorem 1.1** For \( d \leq 2 \), there exists no stationary and ergodic ensemble of \( (\{X\}, \{Y\}, T) \), where \( \{X\}, \{Y\} \) are independent Poisson point processes and \( T \) is a cyclically monotone bijection of \( \{X\} \) and \( \{Y\} \).

Our interest in this problem is motivated on the one hand by work on geometric properties of matchings by Holroyd [13] and Holroyd et al. [14, 15], and on the other hand by work on optimally coupling random measures by the first author and Sturm [17] and the first author [16]. In [14], Holroyd, Janson, and Wästlund analyze (stationary) matchings satisfying the local optimality condition (1.2) with the exponent 2 replaced by \( \gamma \in [-\infty, \infty] \). They call matchings satisfying this condition \( \gamma \)-minimal and derive a precise description of the geometry of these matchings in dimension \( d = 1 \). In dimension \( d > 1 \) much less is known. In particular, in the critical dimension \( d = 2 \) they could only show existence of stationary \( \gamma \)-minimal matchings for \( \gamma < 1 \). The cases \( \gamma \geq 1 \) were left open, but see [15] and [13] for several open questions for \( d = 2 \) and in particular \( \gamma = 1 \). On the other hand, the first author and Sturm [16, 17] develop an optimal transport approach to these (and related) problems. They identify the point sets \( \{X\}, \{Y\} \) with the counting measures (1.3) and seek a stationary coupling \( Q \) minimizing the cost

\[
\mathbb{E} \int_{B_1 \times \mathbb{R}^d} |x - y|^{\gamma} dQ,
\]

where we denote by \( B_R \) the unit ball of radius \( R \). (Note that any (stationary) bijection \( T : \{X\} \to \{Y\} \) induces a (stationary) coupling between \( \mu \) and \( \nu \) by setting \( Q = (\text{id}, T)_\# \mu \).) If this cost functional is finite there exists a stationary coupling which is necessarily locally optimal in the sense of (1.2) with the exponent 2 replaced by \( \gamma \). In dimension 2 for \( \mu \) and \( \nu \) being two independent Poisson processes, this functional is finite if and only if \( \gamma < 1 \), which is in line with the results of [14].

In view of these results, it is natural to conjecture that Theorem 1.1 also holds for \( \gamma \)-minimal matchings with \( \gamma \geq 1 \). However, our proof crucially relies on the harmonic approximation result of [10] which so far is only available for \( \gamma = 2 \).

---

5 Recall that this is automatic for the two first components \((\{X\}, \{Y\})\) by the properties of the Poisson point process so that this amounts to a property of \( T \).

6 Paraphrased by saying that spatial averages are ensemble averages.

7 Recall that a coupling \( Q \) is said to be stationary if the law of the triple \((\{X\}, \{Y\}, Q)\) is invariant under the action of the additive group \( \mathbb{Z}^d \), i.e., for any \( \bar{x} \in \mathbb{Z}^d \) the vectors \((\{X\}, \{Y\}, Q)\) and \((\{\bar{x} + X\}, \{\bar{x} + Y\}, Q(\cdot - \bar{x}, \cdot - \bar{x}))\) have the same law.
Before we explain the main steps of the proof of Theorem 1.1 in Sect. 1.1 we would like to give a few remarks on extensions and variants of Theorem 1.1.

**Remark 1.2** Theorem 1.1 remains true if we replace the bijection $T$ by the a priori more general object of a stationary coupling $Q$. This can be seen by either using that matchings are extremal elements in the set of all couplings of point sets or by directly writing the proof in terms of couplings which essentially only requires notational changes.

**Remark 1.3** Very well studied siblings of stationary matchings are stationary allocations of a point process $\{X\}$, i.e. a stationary map $T : \mathbb{R}^d \to \{X\}$ such that $\text{Leb}(T^{-1}(X))$ equals $\mathbb{E}[\#\{X \in (0, 1)^d\}]^{-1}$. There are several constructions of such an allocation, for instance by using the stable marriage algorithm in [12], the flow lines of the gravitational force field exerted by $\{X\}$ in [8], an adaptation of the AKT scheme in [22], or by optimal transport methods in [17].

By essentially the same proof as for Theorem 1.1, one can show that in $d = 2$ there is no cyclical monotone stationary allocation to a Poisson process. The only place where we need to change something in the proof is the $L^\infty$ estimate Lemma 2.2.

**Remark 1.4** We do not use many particular features of the Poisson measures $\mu$ and $\nu$ in the proof of Theorem 1.1 since ergodicity and stationarity allow us to argue on a pathwise level via the harmonic approximation result (cf. Sect. 1.1).

We use two properties of the Poisson measure. The first property is concentration around the mean. The second property is more involved. Denote by $W_p$ the $L^p$ Wasserstein distance. We use that $\frac{1}{R^d} W_2(\mu \otimes B_R, \frac{\mu(B_R)}{|B_R|} \text{Leb})$ diverges at the same rate for $R \to \infty$ as $\frac{1}{R^d} W_2(\mu \otimes B_R, \frac{\mu(B_R)}{|B_R|} \text{Leb})$ for some $\varepsilon > 0$ (here we use $\varepsilon = 1$).

As Remark 1.3 indicates stationary matchings are closely related to the bipartite matching problem, which is the natural variant of the problem studied in this paper with only a finite number of points $\{X_1, \ldots, X_n\}, \{Y_1, \ldots, Y_n\}$. Note that then the local optimality condition (1.2) turns into a global optimality condition. This (finite) bipartite matching problem has been the subject of intense research in the last 30 years, see e.g. [1] for the first proof of the rate of convergence in the case of iid points in dimension $d = 2$, [24] for sharp integrability properties for matchings in $d \geq 3$, [7] for a new approach based on the linearization of the Monge-Ampère equation, and [5] for a thorough analysis of the bipartite matching problem in $d = 1$.

We rely on this connection in two ways. On the one hand we use the asymptotics of the cost of the bipartite matching which are known since [1]. However, we need a local version for which we give a self-contained proof using martingale techniques (see Sect. 2.5) which is new and interesting on its own. On the other hand we exploit a large scale regularity result for optimal couplings, the harmonic approximation result, developed in [10]. This regularity result was inspired by the PDE approach proposed by [7], see also [2, 3, 11, 21] for remarkable results for the bipartite matching problem using this approach.

---

8 Given a measure $\mu$ on a $\mathbb{R}^d$ and a subset $A \subseteq \mathbb{R}^d$ we denote its restriction to $A$ by $\mu \otimes A(\cdot) := \mu(\cdot \cap A)$.

9 Which in measure theoretic formulation used in optimal transport is nothing but the cyclical monotonicity of the support of the corresponding coupling.
1.1 Main steps in the proof of Theorem 1.1

In the following we will describe the main steps in the proof of Theorem 1.1. For the detailed proofs we refer to Sect. 2.

We argue by contradiction. We assume that there is a locally optimal stationary matching \( T \) between \( \{X\} \) and \( \{Y\} \). On the one hand, we will show that

\[
\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \leq o(\ln \frac{1}{2} R). \tag{1.4}
\]

On the other hand, it is known (we will prove the local version needed for our purpose) that any bipartite matching satisfies

\[
\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \geq \Omega(\ln \frac{1}{2} R) \tag{1.5}
\]

leading to the desired contradiction.

Let us now describe the different steps leading to (1.4) and (1.5) in more detail. Our starting point is the observation that by stationarity and ergodicity the following \( L^0 \)-estimate on the displacement \( T(X) - X \) holds

\[
\# \left\{ X \in (-R, R)^d : |T(X) - X| \gg 1 \right\} \leq o(R^d), \tag{1.6}
\]

see Lemma 2.1 for a precise statement. This is the only place where stationarity and ergodicity enter. Since \( T \) is locally optimal, its support is in particular monotone, which means that for any \( X, X' \in \{X\} \) we have

\[
(T(X') - T(X)) \cdot (X' - X) \geq 0.
\]

By (1.6), we know that most of the points in \((-R, R)^d\) are not transported by a large distance. Combining this with monotonicity allows us to also shield the remaining points from being transported by a distance of order \( R \), so that

\[
|T(X) - X| \leq o(R) \text{ provided that } X \in (-R, R)^d, \tag{1.7}
\]

see Lemma 2.2 for a precise statement. By concentration properties of the Poisson process we may assume that \( \frac{\# \{ X \in B_R \} |B_R|}{|B_R|} \leq \left[ \frac{1}{2}, 2 \right] \) for \( R \gg 1 \). Summing (1.7) over \( B_R \) we obtain

\[
\frac{1}{R^d} \sum_{X \in B_R} |T(X) - X|^2 \leq o(R^2). \tag{1.8}
\]

Now comes the key step of the proof. We want to exploit regularity of \( T \) to upgrade (1.8) to an \( O(\ln R) \) bound. The tool that allows us to do this is the harmonic approximation result [10, Theorem 1.4] (see also [19] for a simplified exposition of the theory) which
quantifies the closeness of the displacement $T(X) - X$ to a harmonic gradient field $\nabla \varphi$, taking as input only the local energy

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

(1.9)

and the distance of $\mu \perp (-R, R)^d$ and $\nu \perp (-R, R)^d$ to the Lebesgue measure on $B_R$

$$D(R) := \frac{1}{R^d} W^2_{(-R, R)^d} (\mu, n_\mu) + \frac{R^2}{n_\mu} (n_\mu - 1)^2 + \frac{1}{R^d} W^2_{(-R, R)^d} (\nu, n_\nu) + \frac{R^2}{n_\nu} (n_\nu - 1)^2,$$

(1.10)

where $n_\mu = \frac{|\{X \in B_R\}|}{|B_R|}$ and $n_\nu = \frac{|\{Y \in B_R\}|}{|B_R|}$, and $W_B (\mu, n) = W_2 (\mu \perp B, n \text{Leb} \perp B)$. By (1.8) (together with its counter part arising from exchanging the roles of $\{X\}$ and $\{Y\}$) we have $E(R) \leq o(R^2)$. By the well-known bound for the matching problem in $d = 2$ and by concentration properties of the Poisson process we have $D(R) \leq O(\ln R)$, see Lemma 2.6. Iteratively exploiting the harmonic approximation result on an increasing sequence of scales we obtain that the local energy $E$ inherits the asymptotic of $D$:

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^2 \leq O(\ln R),$$

(1.11)

see Lemma 2.3. Combining this with the $L^0$-estimate yields

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \leq o(\ln^{\frac{1}{2}} R),$$

see Lemma 2.4.

It remains to establish the lower bound (1.5), which is essentially known. However, for our purpose we need the local version (1.5). Our proof is very similar to the proof of the lower bound in the seminal paper [1]. Both approaches construct candidates for the dual problem based on dyadic partitions of the cube. However, instead of using a quantitative embedding result into a Gaussian process as in [1] we use a natural martingale structure together with a concentration argument. More precisely, we show that there exists $\zeta$ with $\text{supp} \zeta \in (0, R)^2$ and $|\nabla \zeta| \leq 1$ such that

$$\frac{1}{R^d} \sum_{X \in (0, R)^2 \text{ or } T(X) \in (0, R)^2} \zeta (T(X)) - \zeta (X) \geq \Omega(\ln^{\frac{1}{2}} R),$$
see Lemma 2.5. Note that this is sufficient for the contradiction stated at the beginning of the subsection, indeed

\[
O(\ln \frac{1}{2} R) \leq \frac{1}{R^d} \sum_{X \in (0, R)^2 \text{ or } T(X) \in (0, R)^2} \zeta (T(X)) - \zeta (X)
\]

\[
\leq \frac{1}{R^d} \sum_{X \in (0, R)^2 \text{ or } T(X) \in (0, R)^2} |T(X) - X|
\]

\[
\leq \frac{1}{R^d} \sum_{X \in B_{\sqrt{2} R} \text{ or } T(X) \in B_{\sqrt{2} R}} |T(X) - X|
\]

\[
\leq o(\ln \frac{1}{2} R).
\]

\section{Proofs}

For the rest of the paper we let \{X\} and \{Y\} be two Poisson point processes and we let \(T\) be their matching, i.e. a bijection from \{X\} to \{Y\}.

\subsection{The ergodic estimate}

\textbf{Lemma 2.1} For any \(\varepsilon > 0\) there exist a deterministic \(L\) and a random radius \(r_* < \infty \) a.s. such that for all \(R \geq r_*\)

\[
\# \left\{ X \in (-R, R)^d \mid |T(X) - X| > L \right\} \leq (\varepsilon R)^d. \tag{2.1}
\]

\textbf{Proof} Let \(Q_R = (-R, R)^d\) and consider the number of points in \(Q_R\) which are transported by a distance greater than \(L\), namely

\[
N^{>L}_{Q_R} := \# \{ X \in Q_R \mid |T(X) - X| > L \}.
\]

We show that stationarity together with the ergodic theorem implies as \(R \to \infty\) that

\[
\frac{1}{R^d} N^{>L}_{Q_R} \to \mathbb{E} N^{>L}_{Q_1} \quad \text{a.s.}. \tag{2.2}
\]

Then, taking \(L \to \infty\) we have

\[
\mathbb{E} N^{>L}_{Q_1} \to 0. \tag{2.3}
\]

Note that (2.2) together with (2.3) imply the existence of a random radius \(r_*\) and a deterministic \(L\) as in the Lemma. Indeed, for any fixed \(\varepsilon > 0\) we can choose \(L\) large enough so that \(\mathbb{E} N^{>L}_{Q_1} \leq \varepsilon^d 2\), and then choose \(r_*\) large enough so that for \(R \geq r_*\)

\[
\left| \frac{1}{R^d} N^{>L}_{Q_R} - \mathbb{E} N^{>L}_{Q_1} \right| \leq \frac{\varepsilon^d}{2},
\]
which turns into the equivalent assertion $\frac{1}{R^d}N_{Q_R}^{>L} \leq \varepsilon^d$.

We now turn to (2.2). For $i \in \mathbb{Z}^d$ note that

$$N_{Q_R}^{>L} = \sum_{|i|<R} N_{Q_{1+i}}^{>L} = \sum_{|i|<R} \# \{ X \in (i_1 - 1, i_1 + 1) \times \cdots \times (i_d - 1, i_d + 1) \mid |T(X) - X| > L \} .$$

Since $\mu$ is stationary and ergodic, by the Birkhoff-von Neumann ergodic theorem [18, Theorem 9.6] for any divergent sequence of integer radii

$$\frac{1}{R^d} \sum_{|i|<R} N_{Q_{1+i}}^{>L} \to \mathbb{E} N_{Q_1}^{>L} . \tag{2.4}$$

Note that for integer $R$ the term on the left hand side amounts to $\frac{1}{R^d} N_{Q_R}^{>L}$. Since for every real $R$ there exists an integer $\tilde{R}$ such that $\tilde{R} \leq R \leq \tilde{R} + 1$ and the ratio between $\tilde{R}$ and $\tilde{R} + 1$ goes to 1 as $R \to \infty$, (2.4) holds also for any divergent sequences of real radii.

Finally we turn to (2.3). By dominated convergence it suffices to show that

$$N_{Q_1}^{>L} \to 0 \text{ a. s..} \tag{2.5}$$

Indeed, since $N_{Q_1}^{>L} \leq \# \{ X \in Q_1 \}$ by dominated convergence (2.3) holds. Note that $N_{Q_1}^{>L}$ is finite for every realization of $\{X\}$ and there exists $L$ large enough such that $N_{Q_1}^{>L} = 0$, which implies the almost sure convergence (2.5) since $\mathbb{E} N_{Q_1}^{>L}$ is finite by the properties of the Poisson process.

$\square$

### 2.2 The $L^\infty$-estimate

The proof is very similar to the $L^\infty$ estimate in [10, Lemma 2.9]. However, note that in [10, Theorem 1.4] a local $L^2$ estimate was turned into a $L^\infty$ estimate whereas in the current setup we want to turn a $L^0$ estimate into a $L^\infty$ estimate. The key property that allows us to do this is the monotonicity of the support of $T$. This translates the partial control given by (2.1) into the claimed $L^\infty$ estimate.

**Lemma 2.2** For every $\varepsilon > 0$ there exists a random radius $r_* < \infty$ a. s. such that for every $R \geq r_*$

$$|T(X) - X| \leq \varepsilon R \text{ provided that } X \in (-R, R)^d . \tag{2.6}$$

**Proof** Step 1. Definition of $r_* = r_*(\varepsilon)$ as the maximum of three random scales. Fix $0 < \varepsilon \ll 1$. First, by Lemma 2.1, there exists a (deterministic) length $L < \infty$ and a (random) length $r_* < \infty$ such that for $4R \geq r_*$, the number of the Poisson points in $(-2R, 2R)^d$ transported further than the “moderate distance” $L$ is small in the sense
of

\[ \#\{ X \in (-2R, 2R)^d | |T(X) - X| > L \} \leq (\varepsilon 4R)^d. \tag{2.7} \]

Second, by Lemma 2.6 we may also assume that \( r_* \) is so large that for \( R \geq r_* \), the non-dimensionalized transportation distance of \( \mu \) to its number density

\[ n_{Q_{2R}} := \frac{\mu((-2R, 2R)^d)}{(4R)^d} \]

is small, and that \( n_{Q_{2R}} \approx 1 \), in the sense of

\[ W^2_{(-2R,2R)^d}(\mu, n_{Q_{2R}}) + \frac{(4R)^{d+2}}{n_{Q_{2R}}}(n_{Q_{2R}} - 1)^2 \leq (4R)^{d+2}. \tag{2.8} \]

Third, w. l. o. g. we may assume that \( r_* \) is so large that

\[ L \leq \varepsilon r_. \tag{2.9} \]

We now fix a realization and \( R \geq r_* \).

**Step 2.** There are enough Poisson points on mesoscopic scales. We claim that for any cube \( Q \subset (-2R, 2R)^d \) of “mesoscopic” side length

\[ r \gg \varepsilon R \tag{2.10} \]

we have

\[ \#\{ X \in Q \} \gtrsim r^d. \tag{2.11} \]

Indeed, it follows from the definition of \( W_{(-2R,2R)}(\mu, n_{Q_{2R}}) \) that for any Lipschitz function \( \eta \) with support in \( Q \) we have

\[ \left| \int \eta d\mu - \int \eta n_{Q_{2R}} dy \right| \leq (\text{Lip}\eta) \left( \int_Q d\mu + n_{Q_{2R}}|Q| \right)^{1/2} W_{(-2R,2R)^d}(\mu, n_{Q_{2R}}). \]

We now specify to an \( \eta \leq 1 \) supported in \( Q \), to the effect of \( \int \eta d\mu \leq \int d\mu = \#\{ X \in Q \} \), so that by Young’s inequality

\[ \int \eta n_{Q_{2R}} dy \lesssim \#\{ X \in Q \} + (\text{Lip}\eta)^2 W^2_{(-2R,2R)^d}(\mu, n_{Q_{2R}}) \]

\[ + (\text{Lip}\eta)(n_{Q_{2R}}|Q|)^{1/2} W_{(-2R,2R)^d}(\mu, n_{Q_{2R}}). \tag{2.12} \]

---

10 The notation \( A \gtrsim B \) means that there exists a constant \( C > 0 \), which may only depend on \( d \), such that \( A \geq CB \). Similarly \( A \lesssim B \) means that there exists a constant \( C > 0 \), which may only depend on \( d \), such that \( A \leq CB \).
At the same time, we may ensure $\int_{(-2R,2R)^d} \eta \gtrsim r^d$ and $\text{Lip}(\eta) \lesssim r^{-1}$, so that by (2.8), which in particular ensures $n_{Q_{2R}} \approx 1$, (2.12) turns into

$$r^d \lesssim \#\{X \in Q\} + r^{-2}(\varepsilon R)^{d+2} + r^{-\frac{d}{2} - 1}(\varepsilon R)^{\frac{d}{2} + 1}.$$  

Thanks to assumption (2.10) we obtain (2.11).

**Step 3.** Iteration. There are enough Poisson points of moderate transport distance on mesoscopic scales. We claim that for any cube $Q \subset (-2R, 2R)^d$ of side-length satisfying (2.10) we have

there exists $X \in Q$ with $|T(X) - X| \leq L$. \hspace{1cm} (2.13)

We suppose that (2.13) were violated for some cube $Q$. By (2.11), there are $r^d$ of such points. By assumption (2.10), there are thus $\gg (\varepsilon R)^d$ Poisson points in $(-2R, 2R)^d$ that get transported by a distance $> L$, which contradicts (2.7).

**Step 4.** At mesoscopic distance around a given point $X \in (-R, R)^d$, there are sufficiently many Poisson points that are transported only over a moderate distance. More precisely, we claim that provided (2.10) holds, there exist $d+1$ Poisson points $\{X_n\}_{n=1}^{d+1}$ that are transported over a moderate distance, i.e.

$$|T(X_n) - X_n| \leq L, \hspace{1cm} (2.14)$$

but on the other hand lie in “sufficiently general” directions around $X$, meaning that

the convex hull of $\left\{\frac{X_n - X}{|X_n - X|}\right\}_{n=1}^{d+1}$ contains $B_\rho$ \hspace{1cm} (2.15)

for $\rho \ll 1$, while the distances to $X$ are of order $r$

$$|X_n - X| \sim r. \hspace{1cm} (2.16)$$

Indeed, this can be seen as follows: Consider the symmetric tetrahedron\footnote{Using the $d = 3$-language, for $d = 2$ it is the equilateral triangle.} with barycenter at $X$. For each of its $d+1$ vertices, consider a rotationally symmetric cone with apex at $X$ and with axis passing through the vertex. Provided the opening angles $\alpha$ are $\ll 1$, by continuity, any selection $e_n$ of unit vectors in this cones still has the property that their convex hull contains $B_\rho$ for $\rho \ll 1$. Consider the intersection of these cones with the (dyadic) annulus centered at $X$ of radii $r$ and $2r$. These $d+1$ intersections are contained in $(-2R, 2R)^d$, and each contains a cube of side-length $\sim r$. Hence (2.13) applies and we may pick a Poisson point $X_n$ with (2.14) in each of these intersections, see Fig. 1. Condition (2.16) is satisfied because the points $X_n$ lie within the chosen annulus, and condition (2.15) is satisfied because the points $X_n$ lie within the chosen cones.
Step 5. All Poisson points are transported over distances $\ll R$. We claim that for all Poisson points $X$

$$|T(X) - X| \lesssim \varepsilon R \quad \text{provided } X \in (-R, R)^d. \quad (2.17)$$

Given the Poisson point $X \in (-R, R)^d$, let $\{X_n\}_{n=1}^{d+1}$ be as in Step 4. By cyclical monotonicity (1.1), which implies monotonicity of the map $T$, we have $(T(X_n) - T(X)) \cdot (X_n - X) \geq 0$, which we use in form of

$$(T(X) - X) \cdot (X_n - X) \leq (T(X_n) - X_n) \cdot (X_n - X) + |X_n - X|^2 \lesssim |T(X_n) - X_n|^2 + |X_n - X|^2. \quad (2.18)$$

We now appeal to (2.14), which by (2.9) and and (2.10) implies

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

Inserting this and (2.16) into (2.18), we obtain

$$(T(X) - X) \cdot \frac{X_n - X}{|X_n - X|} \lesssim r$$

for all $n = 1, \ldots, d + 1$. Since by (2.15), any unit vector $e$ can be written as a linear combination of $\left\{ \frac{X_n - X}{|X_n - X|} \right\}_{n=1}^{d+1}$ with non-negative weights $\leq \frac{1}{\rho} \sim 1$, this implies $|T(X) - X| \lesssim r$. Since (2.10) was the only constraint on $r$, we obtain (2.17). \qed
2.3 Key step: Harmonic approximation

Lemma 2.3 There exist a constant $C^{12}$ and a random radius $r_* < \infty$ a. s. such that for every $R \geq r_*$ we have

$$\frac{1}{R} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^2 \leq C \ln R. \quad (2.19)$$

Proof of Lemma 2.3 The proof relies on the harmonic approximation result from [10, Theorem 1.4]. This result establishes that for any $0 < \tau \ll 1$, there exists an $\varepsilon > 0$ and a $C_\tau < \infty$ such that provided for some $R$

$$\frac{1}{R^2} E(6R) + \frac{1}{R^2} D(6R) \leq \varepsilon \quad (2.20)$$

(recall (1.9) and (1.10) for the definition) there exists a harmonic gradient field $\Phi$ such that

$$\frac{1}{R} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X - \nabla \Phi(X)|^2 \leq \tau E(6R) + C_\tau D(6R), \quad (2.21)$$

$$\sup_{B_{2R}} |\nabla \Phi|^2 \leq C_\tau (E(6R) + D(6R)).$$

The fraction $\tau$ will be chosen at the end of the proof. Note that in (1.10) $D(R)$ is defined on boxes $(-R, R)^d$ while [10, Theorem 1.4] (2.20) requires balls. Since $B_{6R} \subseteq (-6R, 6R)^d$ we may assume that (2.20) holds also for $B_{6R'}$ for $R'$ close to $R$, see [10, Lemma 2.10] for this type of restriction property.

Step 1. Definition of $r_*$ depending on $\tau$. For $0 < \tau \ll 1$ let $\varepsilon = \varepsilon(\tau)$ be as above. By Lemma 2.6 we may assume that $r_*$ is large enough so for any dyadic $R \geq r_*$

$$D(R) \leq C \ln R \quad (2.22)$$

and

$$\frac{D(6R)}{R^2} \leq \frac{\varepsilon}{2}. \quad (2.23)$$

Note that the estimate (2.23) is not sharp, but it is enough for our purpose. Moreover, only the bound (2.22) is specific to $d = 2$. From now on, we restrict ourselves to dyadic $R$, which we may do w. l. o. g. for (2.19). Note that by the bound on $D(6R)$ in (2.23) and the second and fourth term in the definition of $D(R)$ in (1.10)

$$\# \left( \{X \in B_R \} \cup \{T(X) \in B_R \} \right) \leq CR^d. \quad (2.24)$$

12 Here and throughout the paper $C$ denotes a positive constant, which may only depend on the dimension, the value of which may change from line to line.
Moreover, we may assume that $r_*$ is large enough so that (2.6) holds. Since $B_{6R} \subset (-6R, 6R)^d$ we may sum (2.6) over $B_R$ to obtain for $R \geq r_*$

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

By symmetry, potentially enlarging $r_*$, we may also assume that (2.6) holds with $X$ replaced by $T(X)$ so that both

$$
T(X) \in B_R \Rightarrow X \in B_{2R}
$$

and

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

thus

$$
E(6R) = \frac{1}{R^d} \sum_{X \in B_{6R} \text{ or } T(X) \in B_{6R}} |T(X) - X|^2 \leq \frac{\varepsilon}{2} R^2,
$$

and in particular (2.20) holds. Finally by Lemma 2.1 we may assume, possibly enlarging $r_*$, that there exists a deterministic constant $L_\tau$ and for $R \geq r_*$ we both have

$$
\#(\{X \in Q_R \mid |T(X) - X| > L_\tau\} \cup \{T(X) \in Q_R \mid |T(X) - X| > L_\tau\}) \leq \tau R^d.
$$

and

$$
L_\tau^2 \leq \ln R.
$$

Step 2. Application of harmonic approximation. For all $R \geq r_*$

$$
E(R) \leq \tau E(6R) + C_\tau \ln R.
$$

We split the sum according to whether the transportation distance is moderate or large. On the latter we use the harmonic approximation:

$$
\frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X) - X| > L_\tau} |T(X) - X|^2 \\
\leq \frac{2}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X - \nabla \Phi(X)|^2 \\
+ \frac{2}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X) - X| > L_\tau} |\nabla \Phi(X)|^2
$$

(2.21), (2.25), (2.26)

$$
\leq 2\tau E(6R) + 2C_\tau D(6R) + 2\tau C_\tau (E(6R) + D(6R))
$$
\[ = 2\tau (1 + C_\tau) E(6R) + 2C_\tau (1 + \tau) D(6R). \]

The last estimate combines to

\[
\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \]

\[= \frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X) - X| \leq L_\tau} |T(X) - X|^2 \]

\[+ \frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X) - X| > L_\tau} |T(X) - X|^2 \]

\[\leq C L_\tau^2 + 2\tau (1 + C_\tau) E(6R) + 2C_\tau (1 + \tau) D(6R) \]

\[\leq 2\tau (1 + C_\tau) E(6R) + (2C_\tau (1 + \tau) + C) \ln R. \]

Relabeling \(\tau\) and \(C_\tau\), this implies (2.28).

**Step 3. Iteration.** Iterating (2.28), we obtain for any \(k \geq 1\)

\[E(R) \leq \tau E(6R) + C_\tau \ln R \]

\[\leq \tau^2 E(6^2 R) + \tau C_\tau \ln R + C_\tau \ln R \]

\[\leq \tau^k E(6^k R) + C_\tau \sum_{l=0}^{k-1} \tau^l \ln R \]

\[\leq \epsilon (36\tau)^k R^2 + C_\tau \sum_{l=0}^{k-1} \tau^l \ln R. \]

We now fix \(\tau\) such that \(36\tau < 1\) to the effect of

\[E(R) \leq C \sum_{l=0}^{\infty} \tau^l \ln R \leq C \ln R. \]

\[\Box\]

### 2.4 Trading integrability against asymptotics

**Lemma 2.4** For every \(\epsilon > 0\) there exists a random radius \(r_* < \infty\) a. s. such that for every \(R \geq r_*\) we have

\[\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \leq \epsilon \ln^{1/2} R. \]
Proof By Lemma 2.3, we know that there exists a random radius \( r_\ast \) such that for \( R \geq r_\ast \) we have

\[
E (R) = \frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^2 \leq C \ln R. \tag{2.29}
\]

Let \( 0 < \varepsilon \ll 1 \). Possibly enlarging \( r_\ast \), we may also assume by Lemma 2.1 that there exists a deterministic constant \( L \) such that for \( R \geq r_\ast \)

\[
\# \{(X \in B_R \mid |T (X) - X| > L) \cup \{T(X) \in B_R \mid |T (X) - X| > L\}) \leq \varepsilon R^d. \tag{2.30}
\]

Furthermore, note that by Lemma 2.6 and the second and fourth term in the definition of \( D (R) \) in (1.10) we may also assume possibly enlarging \( r_\ast \) again that for \( R \geq r_\ast \)

\[
L \leq \varepsilon^{\frac{1}{2}} \ln \frac{1}{2} R. \tag{2.31}
\]

We split again the sum into moderate and large transportation distance and apply Cauchy-Schwarz:

\[
\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \leq \frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X) - X| \leq L} |T(X) - X| + \frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X) - X| > L} |T(X) - X| \leq C L + \varepsilon^{\frac{1}{2}} E(R) \frac{1}{2} \leq C \varepsilon^{\frac{1}{2}} \ln \frac{1}{2} R. \tag{2.24),(2.29),(2.30)
\]

Relabeling \( \varepsilon \) proves the claim. \( \square \)

2.5 Asymptotics for the bipartite matching problem in dimension \( d = 2 \)

In this section we give a self-contained proof of the upper and lower bounds on the asymptotics in the matching problem in the critical dimension \( d = 2 \).

The intuition why \( d = 2 \) is critical for optimal transportation is clear: the fluctuations of the number density of the Poisson point process of unit intensity, i.e. its deviation from 1 on some mesoscopic scale \( r \gg 1 \) are of size \( O \left( \frac{1}{\sqrt{r^d}} \right) \). To compensate these fluctuations, one has to displace particles by a distance \( O (r \times \frac{1}{\sqrt{r^d}}) \), which is \( O (1) \) iff \( d = 2 \). Hence taking care of the fluctuations on every dyadic scale \( r \) requires a displacement of \( O (1) \). Naively, this suggests a transportation cost per particle that is logarithmic in the ratio between the macroscopic scale \( R \) and the microscopic scale \( 1 \). However, by the independence properties of the Poisson point process, the displacements on every dyadic scale are essentially independent, so that there are the usual
cancellations when adding up the dyadic scales. Hence the displacement behaves like the square root of the logarithm.

The proof of the lower bound is similar to the original proof in [1]. However, in the final step we use martingale arguments instead of Gaussian embeddings. The proof of the upper bound has some similarities to the proof in [3]. Again we use martingale arguments and crucially the Burkholder-Davis-Gundy inequality. It would be interesting to see whether our technique for the upper bound allows to get the precise constant as in [3].

**Lemma 2.5** Let \( \mu, \nu \) denote two independent Poisson point processes in \( \mathbb{R}^2 \) of unit intensity. There exists a constant \( C \), and a random radius \( r_* < \infty \) a. s. such that for any dyadic radii \( R \geq r_* \)

\[
\sup \left\{ \int \zeta \, (d\mu - d\nu) \mid \text{supp} \zeta \subset (0, R)^2, \ |\nabla \zeta| \leq 1, \int \zeta \, dx = 0 \right\} \geq C R^2 \ln^{\frac{1}{2}} R.
\]

**Lemma 2.6** Let \( \mu, \nu \) be as in Lemma 2.5. Then there exists a constant \( C \), and there exists a random radius \( r_* < \infty \) a. s. such that for any dyadic radii \( R \geq r_* \)

\[
D(R) \leq C \ln^{\frac{1}{2}} R.
\]

For the lower as well as the upper bound, our proof proceeds in two stages. First we show the desired estimate in expectation. Then we use concentration arguments to lift it to a pathwise estimate.

### 2.5.1 Lower bound on \( W_1 \)

**Lemma 2.7** Let \( \mu \) denote the Poisson point process in \( \mathbb{R}^2 \) of unit intensity. Then it holds for \( R \gg 1 \)

\[
\mathbb{E}\sup \left\{ \int \zeta \, d\mu \mid \text{supp} \zeta \subset (0, R)^2, \ |\nabla \zeta| \leq 1, \int \zeta \, dx = 0 \right\} \geq R^2 \ln^{\frac{1}{2}} R. \quad (2.32)
\]

**Proof** W. l. o. g. we may assume that \( R \in 2^n \), and will consider the (finite) family of all dyadic cubes \( Q \subset (0, R)^2 \) of side-length \( \geq 1 \). For such \( Q \), we call \( Q_r \) and \( Q_l \) the right and left half of \( Q \) and consider the integer-valued random variable

\[
N_Q = \mu(Q_r) - \mu(Q_l). \quad (2.33)
\]

Fixing a smooth mask (or reference function) \( \hat{\zeta} \) with

\[
\text{supp} \hat{\zeta} \in (0, 1)^2, \quad \int \hat{\zeta}, d\hat{x} = 0, \quad \int_{\hat{x}_1 > \frac{1}{2}} \hat{\zeta} \, d\hat{x} - \int_{\hat{x}_1 < \frac{1}{2}} \hat{\zeta} \, d\hat{x} = 1, \quad (2.34)
\]
for every $Q$ we define $\zeta_Q$ via the simple transformation

$$\zeta_Q(A_Q \hat{x}) = \hat{\zeta}(\hat{x})$$

where $A_Q$ is the affine map with $Q = A_Q(0, 1)^2$. \hfill (2.35)

For later purpose we note that by standard properties of the Poisson process (see [20, equation (4.23)]) we have (recall that we consider dyadic cubes only)

$$\mathbb{E}N_Q^2 = |Q|, \quad \mathbb{E}N_QN_{Q'} = 0 \text{ for } Q \neq Q', \quad \mathbb{E}N_Q \int \zeta_Q d\mu = |Q|. \hfill (2.36)$$

For every dyadic $r = R, R^2, R^4, \ldots, 1$ we consider the function

$$\zeta_r := \sum_{Q \text{ of level } \geq r} N_Q \zeta_Q \hfill (2.37)$$

and note that for a fixed point $x \in (0, R)^d$, we have

$$\nabla \zeta_r(x) = \sum_{R \geq \rho \geq r} \sum_{Q \text{ of level } \rho} N_Q \nabla \zeta_Q(x), \hfill (2.38)$$

observing that the sum over $Q$ restricts to the one cube of dyadic level/side-length $\rho$ that contains $x$. We now argue that (2.38) is a martingale, where $r = R, R^2, R^4, \ldots$, plays the role of a discrete time. More precisely, it is a martingale with respect to the filtration generated by $\{\{N_Q\}_Q \text{ of level } r\}_r$. It suffices to show that when conditioned on $\{N_Q\}_Q$ of level $\rho$, for $\rho \geq 2r$, the expectation of $N_Q$ for every $Q$ of level $r$ vanishes. To this end, apply a Lebesgue-measure preserving transformation of $\mathbb{R}^2$ that swaps the left and right half of $Q$; when applied to point configurations, it preserves the law of the Poisson point process, it leaves $N_{\hat{Q}}$ for $\hat{Q}$ of level $\rho \geq 2r$, the expectation of $N_Q$ for every $Q$ of level $r$ invariant, in particular also the $\sigma-$algebra generated by these random variables, and it converts $N_{\hat{Q}}$ into $-N_Q$. Hence, $\mathbb{E}[N_Q|\{N_{\hat{Q}}\}_Q \text{ of level } \rho, \rho \geq 2r] = 0$.

As the sum over $Q$ in (2.38) reduces to a single summand, this martingale has (total) quadratic variation

$$\sum_{R \geq r \geq 1} \sum_{Q \text{ of level } r} N_Q^2 |\nabla \zeta_Q(x)|^2, \hfill (2.39)$$

and we claim that it satisfies

$$\mathbb{E} \int \sum_{R \geq r \geq 1} \sum_{Q \text{ of level } r} N_Q^2 |\nabla \zeta_Q(x)|^2 dx \lesssim R^2 \ln R. \hfill (2.39)$$

\hfill (2.39)
Indeed, by \((2.36)\) we obtain that the l. h. s. of \((2.39)\) is given by \(\sum_Q |Q| \int |\nabla \zeta_Q(x)|^2 \, dx\). By \((2.35)\), this in turn is estimated by \(\sum_Q |Q| = \sum_{R \geq r \geq 1} R^2 = R^2 \log_2 R\). We also claim that the last item in \((2.36)\) implies

\[
\mathbb{E} \int \zeta_1 \, d\mu \gtrsim R^2 \ln R. \tag{2.40}
\]

Indeed, the l. h. s. of \((2.40)\) is again given by \(\sum_Q |Q|\). We now appeal to the (quadratic) Burkholder inequality [23, Theorem 6.3.6] (exchanging \(\mathbb{E}\) and \(\int \, dx\)) to obtain from \((2.39)\)

\[
\mathbb{E} \int \sup_{R \geq r \geq 1} |\nabla \zeta_r(x)|^2 \, dx \lesssim R^2 \ln R. \tag{2.41}
\]

We now are in the position to define \(\zeta\) via a stopping “time”. Given an \(M < \infty\), which we think of as being large and that will be chosen later, we keep subdividing the dyadic cubes (recall that we restrict to those of side length \(\geq 1\)) as long as

\[
\int_Q |\nabla \sum_{Q \subset Q} N_Q \zeta_Q(x)|^2 \, dx \leq M |Q| \ln R \quad \text{and} \quad |N_Q|^2 \leq M |Q| \ln R. \tag{2.42}
\]

This defines a nested sub-family of dyadic cubes, and thereby a random (spatially piecewise constant) stopping scale \(r_\ast = r_\ast(x)\) (note that \(r_\frac{2}{\sqrt{r}}\) is a stopping time but we will not need that in this proof). We then set

\[
\zeta(x) := \zeta_{r_\ast(x)}(x). \tag{2.43}
\]

We first argue that we have the Lipschitz condition

\[
|\nabla \zeta|^2 \lesssim M \ln R. \tag{2.44}
\]

Indeed, consider one of the finest cubes \(Q\) in the family constructed in \((2.42)\); by definition \((2.43)\), the first item in \((2.42)\) amounts to

\[
\int_Q |\nabla \zeta(x)|^2 \, dx \leq M |Q| \ln R,
\]

so that \((2.44)\) follows once we show for arbitrary point \(x\)

\[
r_\ast(x)|\nabla^2 \zeta(x)| \lesssim \sqrt{M \ln R}. \tag{2.45}
\]

Indeed, by definitions \((2.35)\), \((2.37)\) and \((2.43)\), the l. h. s. is estimated by

\[
|\nabla^2 \zeta(x)| \leq \sum_{r_\ast(x) \leq r \leq R \ni x} \sum_{\text{level } r} |N_Q||\nabla^2 \zeta_Q(x)| \lesssim \sum_{r_\ast(x) \leq r \leq R} r^{-2} \sum_{Q \ni x \text{ level } r} |N_Q|. \tag{2.35}
\]
By construction of \( r_*(x) \) in form of the second item in (2.42) the latter is estimated by

\[
\sum_{r_*(x) \leq r \leq R} r^{-2} \sum_{Q \ni x \text{ level } r} |N_Q| \lesssim \sum_{r_*(x) \leq r \leq R} r^{-1} \sqrt{M \ln R}.
\]

This is a geometric series that is estimated by \( \lesssim r_*^{-1}(x) \sqrt{M \ln R} \), as desired.

We now argue that the “exceptional” set

\[
E := \{ x \in (0, R)^2 | r_*(x) > 1 \},
\]

where the dyadic decomposition stops before reaching the minimal scale \( r = 1 \), has small volume fraction in expectation:

\[
M E |E| \lesssim R^2.
\]  

(2.46)

Indeed, by definitions (2.37) and (2.42), \( E \) is the disjoint union of dyadic cubes \( Q \) that have at least one (out of four) children \( Q' \) of level \( r = \frac{r_*(x)}{2} \) with

\[
\int_{Q'} |\nabla \zeta_r(x)|^2 dx > M |Q'| \ln R \quad \text{or} \quad |N_{Q'}|^2 > M |Q'| \ln R,
\]

which combines to

\[
|N_{Q'}|^2 + \int_{Q'} \sup_{1 \leq r \leq R} |\nabla \zeta_r(x)|^2 dx > M \frac{|Q|}{4} \ln R.
\]

Summing over all \( Q \) covering \( E \), this yields

\[
\sum_{\text{all } Q} |N_{Q'}|^2 + \int \sup_{1 \leq r \leq R} |\nabla \zeta_r(x)|^2 dx > M \frac{|E|}{4} \ln R.
\]

Taking the expectation we obtain from (2.36) and (2.41)

\[
R^2 \ln R \gtrsim M E |E| \ln R,
\]

which yields (2.46).

We finally argue that

\[
M^{\frac{1}{4}} E |\int (\zeta_1 - \zeta) d\mu| \lesssim R^2 \ln R.
\]  

(2.47)

Together with (2.40) this implies \( E \int \zeta d\mu \gtrsim R^2 \ln R \) for \( M \gg 1 \), which we fix now. In combination with (2.44) this implies the claim of the lemma. Indeed, \( \frac{\zeta}{\sqrt{r}} \) is an admissible candidate for (2.32) and the statement follows once \( M \) in (2.47) is
chosen to be large enough. The dyadic decomposition defines a family of exceptional cubes \( Q \) (a cube \( Q \) is exceptional if and only if for any \( x \in Q \) we have that its level \( r \) satisfies \( r < r_s(x) \)). In view of (2.37) and (2.43), we thus have to estimate 
\[
\sum_Q I(Q \text{ exceptional}) N_Q \int \zeta_Q d\mu.
\]
Applying \( E|\cdot| \) and using Hölder’s inequality in probability we obtain
\[
E|\int (\zeta_1 - \zeta) d\mu| \leq \sum_Q (P Q \text{ exceptional})^{\frac{1}{4}} (E N_Q^4)^{\frac{1}{4}} (E(\int \zeta_Q d\mu)^2)^{\frac{1}{2}}.
\]
As for (2.36), we have for the last factor
\[
E(\int \zeta_Q d\mu)^2 = \int \zeta_Q^2 \lesssim |Q|. \tag{2.49}
\]
Indeed, write \( F = \{ \zeta : \text{supp} \zeta \subset (0, R)^2, |\nabla \zeta| \leq 1, \int \zeta d\mu = 0 \} \) and \( S = \sup_{\zeta \in F} \int \zeta d\mu \), where the supremum has to be understood as an essential supremum. By basic results on essential suprema, there exists a countable subset \( \{ \zeta_n, n \geq 1 \} \subset F \) such that \( S = \sup_{n \geq 1} \int \zeta_n d\mu \). Setting, \( S_n = \max_{k \leq n} \int \zeta_k d\mu \) we have \( S_n \not\to S \) and by monotone convergence \( E S_n \not\to E S \). Hence, there is \( n \) such that \( E S_n \gtrsim R^2 \ln^{\frac{1}{2}} R \) by Lemma 2.7.

By definition of \( S_n \), \( S_n = \int \zeta d\mu \) for some \( F \)-valued random variable \( \zeta \) (that takes finitely many values \( \zeta_1, \ldots, \zeta_n \) w. r. t. \( \mu \)), which is admissible in (2.49). Clearly, \( \zeta \) is measurable w. r. t. \( \mu \) alone. However, \( \zeta \) is measurable only dependent on \( \mu \) and not

\[\Box \text{ Springer}\]
on $\nu$. Since $\nu$ has intensity measure $dx$ we obtain by independence of $\mu$ and $\nu$
\begin{align*}
\mathbb{E} \left[ \int \zeta (d\mu - d\nu) \right] &= \mathbb{E}_\mu \left[ \mathbb{E}_\nu \left[ \int \zeta (d\mu - d\nu) \right] \right] = \mathbb{E}_\mu \left[ \int \zeta d\mu - \int \zeta dx \right] \\
&= \mathbb{E}_\mu \left[ \int \zeta d\mu \right] \gtrsim R^2 \ln \frac{1}{2} R \end{align*}
which proves (2.49).

**Step 2.** Concentration around expectation. We claim that

\[ S_R (\mu, \nu) := \sup \left\{ \int \zeta \ (d\mu - d\nu) \mid \text{supp}\zeta \subset (0, R)^2, \ |\nabla \zeta| \leq 1, \ \int \zeta \ dx = 0 \right\} \]

satisfies

\[ \lim_{R \to \infty} \frac{1}{R^2 \ln \frac{1}{2} R} |S_R (\mu, \nu) - \mathbb{E} S_R (\mu, \nu)| = 0 \ \text{P-a.s.} \quad (2.51) \]

Indeed, by the triangle inequality, we split the statement into two:

\[ |S_R (\mu, \nu) - \mathbb{E} S_R (\mu, \nu)| \leq |S_R (\mu, \nu) - \mathbb{E}_\mu S_R (\mu, \nu)| + |\mathbb{E}_\mu S_R (\mu, \nu) - \mathbb{E} S_R (\mu, \nu)|. \]

By a Borel-Cantelli argument it suffices to show that for any fixed $\varepsilon > 0$ and $R \gg 1$ the following statements hold

\begin{align*}
\mathbb{P} \left( \frac{1}{R^2 \ln \frac{1}{2} R} |S_R (\mu, \nu) - \mathbb{E}_\mu S_R (\mu, \nu)| > \varepsilon \right) &\lesssim \exp \left( -\frac{\varepsilon^2}{4} \ln R \right), \quad (2.52) \\
\mathbb{P} \left( \frac{1}{R^2 \ln \frac{1}{2} R} |\mathbb{E}_\mu S_R (\mu, \nu) - \mathbb{E} S_R (\mu, \nu)| > \varepsilon \right) &\lesssim \exp \left( -\frac{\varepsilon^2}{4} \ln R \right). \quad (2.53)
\end{align*}

We first turn to (2.52). For $z \in \mathbb{R}^2$ we consider the difference operator

\[ D_z S_R (\mu, \nu) := S_R (\mu + \delta_z, \nu) - S_R (\mu, \nu). \]

Since the $\zeta$’s in the definition (2.50) satisfy $\sup |\zeta| \leq \sqrt{2} R$, we have

\[ |D_z S_R (\mu, \nu)| \leq \sup \left\{ |\zeta (z)| \mid \text{supp}\zeta \subset (0, R)^2, \ |\nabla \zeta| \leq 1, \ \int \zeta \ dx = 0 \right\} \leq R. \quad (2.54) \]

Hence,

\[ \int_{\mathbb{R}^2} (D_z S_R (\mu, \nu))^2 \ dz \leq \int_{(0, R)^2} R^2 \ dz \leq R^4. \]
Applying [26, Proposition 3.1] with $\beta = R$ and $\alpha^2 = R^4$ we obtain for $R \gg 1$

$$
\mathbb{P} \left( \left| S_R (\mu, v) - \mathbb{E}_\mu S_R (\mu, v) \right| > \varepsilon R^2 \ln \frac{1}{\varepsilon} R \right) \lesssim \exp \left( -\frac{\varepsilon R^2 \ln \frac{1}{\varepsilon} R}{\beta} \ln \left( 1 + \frac{\beta \varepsilon R^2 \ln \frac{1}{\varepsilon} R}{\alpha^2} \right) \right) = \exp \left( -\frac{\varepsilon R \ln \frac{1}{\varepsilon} R}{2} \ln \left( 1 + \frac{\varepsilon R \ln \frac{1}{\varepsilon} R}{R} \right) \right) \lesssim \exp \left( -\frac{\varepsilon^2}{4} \ln R \right).
$$

The argument for (2.53) is almost identical: For arbitrary $w \in \mathbb{R}^2$ we need to consider

$$
D_w \mathbb{E}_\mu S_R (\mu, v) = \mathbb{E}_\mu S_R (\mu, v + \delta_w) - \mathbb{E}_\mu S_R (\mu, v).
$$

Because of $D_w \mathbb{E}_\mu S_R (\mu, v) = \mathbb{E}_\mu D_w S_R (\mu, v)$ we obtain as above

$$
\left| D_w \mathbb{E}_\mu S_R (\mu, v) \right| = \left| \mathbb{E}_\mu D_w S_R (\mu, v) \right| \leq R.
$$

Then applying once more [26, Proposition 3.1] with $\beta = R$ and $\alpha = R^2$ implies (2.53).

\[\square\]

### 2.5.2 Upper Bound

**Lemma 2.8** Let $\mu$ denote the Poisson point process in $\mathbb{R}^2$ of unit intensity. Then it holds for $R \gg 1$

$$
\mathbb{E} W^2_{(0,R)^2}(\mu, n) \lesssim R^2 \ln R,
$$

where $n = \frac{\mu((0,R)^2)}{R^2}$ is the (random) number density.

**Proof** Clearly, the intuition mentioned at the beginning of Sect. 2.5 suggests to uncover a martingale structure. W.l.o.g. we may assume that $R \in 2\mathbb{N}$, and will consider the family of all dyadic squares $Q$. For such $Q$, we consider the number density

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

Note that $n_Q |Q| \in \mathbb{N}_0$ is Poisson distributed with expectation $|Q|$. Note that if, for a fixed point $x$, we consider the sequence of nested squares $Q$ that contain $x$, the corresponding random sequence $n_Q$ is a martingale. We want to stop the dyadic subdivision just before the number density leaves the interval $[\frac{1}{2}, 2]$ of moderate values. This defines a scale:

$$
r_*(x) := 2 \sup \{r | \text{is the sidelength of dyadic square } Q \ni x \text{ with } n_Q \notin \left[ \frac{1}{2}, 2 \right] \}.
$$
Since on a square $Q$ of side length $r = \frac{1}{2}$ we have $n_Q \in 4\mathbb{N}_0$ we trivially obtain
\[ r_*(x) \geq 1. \tag{2.58} \]

As we shall argue now, it follows from the properties of the Poisson distribution that (the stationary) $r_*(x)$ is $O(1)$ with overwhelming probability, in particular
\[ \mathbb{E}r_*^4(x) \lesssim 1. \tag{2.59} \]

Indeed, by the concentration properties of the Poisson distribution we have for any $\rho \in 2\mathbb{N}$
\[
\mathbb{P}(r_*(x) > \rho) \leq \sum_{Q \ni x, r_Q \geq \rho} \mathbb{P}\left(n_Q \notin \left[\frac{1}{2}, 2\right]\right)
\lesssim \sum_{r \geq \rho \text{ dyadic}} \exp\left(-C r^2\right) \lesssim \exp\left(-C \rho^2\right).
\]

Thus we may estimate
\[
\mathbb{E}r_*^4(x) = 4 \int_0^\infty \rho^3 \mathbb{P}(r_*(x) > \rho) \, d\rho \lesssim \int_0^\infty \rho^3 \exp\left(-C \rho^2\right) \, d\rho \lesssim 1.
\]

We now distinguish the case of $r_* \leq R$ on $(0, R)^2$ and its complement. In the last case, there exists a $y \in (0, R)^2$ such that $r_*(y) > R$, which by (2.57) means that there exists a dyadic cube $Q \ni y$ of sidelenent $r_Q \geq R$ and $n_Q \notin \left[\frac{1}{2}, 2\right]$, which in turn implies that $r_* > R$ on the entire $(0, R)^2$. Now fix a deterministic $y \in (0, R)^2$. Since $n_{(0,R)^2} R^2$ is the number of particles we use the brutal estimate of the transportation distance
\[ W^2_{(0,R)^2}(\mu, n_{(0,R)^2}) \leq n_{(0,R)^2} R^2 2R^2. \tag{2.60} \]

Since by assumption, $(0, R)^2 \subset (0, r_*(y))^2$, this yields $W^2_{(0,R)^2}(\mu, n_{(0,R)^2}) \leq n_{(0,r_*(y))^2} r_*^2(y) 2R^2$. Hence by definition of $r_*(y)$, see (2.57) we obtain
\[ W^2_{(0,R)^2}(\mu, n_{(0,R)^2}) \leq 2r_*^2(y) 2R^2 \lesssim r_*^4(y). \]

By (2.59), taking the expectation yields as desired
\[
\mathbb{E}W^2_{(0,R)^2}(\mu, n_{(0,R)^2}) I (r_*(y) > R) \lesssim 1,
\]

where $I (r_*(y) > R)$ denotes the indicator function of the event $\{r_*(y) > R\}$.

In the more interesting case of $r_* \leq R$, we define a partition of $(0, R)^2$. A cube $Q_*$ is an element of the partition if and only if its sidelenent $r_{Q_*}$ satisfies
\[ r_{Q_*} = \max\{r_*(x) : x \in Q_\} \tag{2.61}. \]
In words, this means that the number density of \( Q^* \) is still within \( \left[ \frac{1}{2}, 2 \right] \) but the number density of at least one of its four children leaves the interval \( \left[ \frac{1}{2}, 2 \right] \). This implies the following reverse of (2.61)

\[
\frac{1}{|Q^*|} \int_{Q^*} r^2_* \geq \frac{1}{4} r^2_{\bar{Q}_*}.
\]

(2.62)

Equipped with this partition, we define the density \( \lambda \) by

\[
\lambda = n_{Q^*} \text{ on } Q^*.
\]

We consider the transportation distance between \( \mu \) and the measure \( \lambda \, dx \). By definition of \( \lambda \), there exists a coupling where the mass is only redistributed within the partition \( \{Q^*\} \). This implies the inequality

\[
W^2_{(0,R)^2}(\mu, \lambda) \leq \sum_{Q^*} 2r^2_{\bar{Q}_*}n_{Q^*}|Q^*| \leq 16 \sum_{Q^*} \int_{Q^*} r^2_* = 16 \int_{(0,R)^2} r^2_*.
\]

(2.57), (2.62)

Taking the expectation, by (2.59) and Jensen’s inequality,

\[
\mathbb{E} W^2_{(0,R)^2}(\mu, \lambda) I (r_* \leq R) \lesssim 1.
\]

Hence in view of (2.59) and the triangle inequality for the transportation distance it remains to show

\[
\mathbb{E} W^2_{(0,R)^2}(\lambda, n_{(0,R)^2}) I (r_* \leq R) \lesssim R^2 \ln R.
\]

(2.63)

The purpose of the stopping (2.57) is that we have the lower bound \( \lambda, n_{(0,R)^2} \geq \frac{1}{2} \) on the two densities, which implies the inequality

\[
W^2_{(0,R)^2}(\lambda, n_{(0,R)^2}) \leq 2 \int_{(0,R)^2} |j|^2
\]

(2.64)

for all distributional solutions \( j \) of

\[
\nabla \cdot j = n_{(0,R)^2} - \lambda \text{ in } (0,R)^2, \quad \nu \cdot j = 0 \text{ on } \partial (0,R)^2.
\]

(2.65)

Inequality (2.64) can be easily derived from the Eulerian description of optimal transportation (see for instance [4, Proposition 1.1] or [25, Theorem 8.1]). Indeed, the couple \((\rho, j)\), with \( \rho = tn_{(0,R)^2} + (1-t) \lambda \geq \frac{1}{2} \) and \( j \) satisfying (2.65) is an admissible candidate for the Eulerian formulation of \( W^2_{(0,R)^2}(\lambda, n_{(0,R)^2}) \).

We now construct a \( j \): For a dyadic square \( Q \) we define its four children \( Q' \) as the four squares that we obtain by dyadically decomposing \( Q \) and we (distributionally) solve the Poisson equation with piecewise constant r. h. s. and no-flux boundary data

\[
-\Delta \varphi_Q = n_Q - n_{Q'} \text{ in } Q', \quad \nu \cdot \nabla \varphi_Q = 0 \text{ on } \partial Q.
\]

(2.66)
Note that (2.66) admits a solution since the integral of the r. h. s. over $Q$ vanishes by definition of $\{n_Q\}$. Since the no-flux boundary condition allows for concatenation of $-\nabla \varphi_Q$ without creating singular contributions to the divergence,

$$j(x) := -\sum_{Q \ni x, R \geq r_Q \geq 2r_{Q_*}} \nabla \varphi_Q(x)$$  \hspace{1cm} (2.67)

which means that the sum\textsuperscript{14} extends over all dyadic cubes $Q \subset (0, R)^2$ containing $x$ and being strictly coarser than $\{Q_*\}$, defines indeed a distributional solution of (2.65).

The Poincaré inequality gives the universal bound

$$\int_Q dx \frac{1}{r_Q} |\nabla \varphi_Q|^2 \lesssim |Q| \sum_{Q' \text{ child of } Q} (n_{Q'} - n_Q)^2.$$  

Appealing to the trivial estimate $\sum_{Q' \text{ child of } Q} (n_{Q'} - n_Q)^2 \lesssim \sum_{Q' \text{ child of } Q} (n_{Q'} - 1)^2$, and to the standard estimate on the variance of the Poisson process, namely $\mathbb{E}(n_Q - 1)^2 \lesssim \frac{1}{|Q|}$, we obtain

$$\mathbb{E} \int_Q |\nabla \varphi_Q|^2 \lesssim |Q|. \hspace{1cm} (2.68)$$

Now comes the crucial observation: We momentarily fix a point $x$ and consider all squares $Q \ni x$. We note that $\nabla \varphi_Q(x)$ depends on the Poisson point process only through $\{n_{Q'}\}$ where $Q'$ runs through the children of $Q$. Moreover, since the expectation of the r. h. s. of the Poisson equation (2.66) vanishes, also the expectation of $\nabla \varphi_Q(x)$ vanishes. Hence the sum

$$\sum_{Q \ni x, R \geq r_Q \geq r} \nabla \varphi_Q(x)$$

is a martingale in the scale parameter $r = R, \frac{R}{2}, \frac{R}{4}, \ldots$ w. r. t. to the filtration generated by the $\{n_Q\}$. Since $r_* \geq 1$, we thus obtain by the Burkholder inequality [23, Theorem 6.3.6]

$$\mathbb{E} \left| \sum_{Q \ni x, R \geq r_Q \geq 2r_{Q_*}} \nabla \varphi_Q(x) \right|^2 \leq \mathbb{E} \sup_{r \geq 1} \left| \sum_{Q \ni x, R \geq r_Q \geq r} \nabla \varphi_Q(x) \right|^2 \lesssim \mathbb{E} \sum_{Q \ni x, R \geq r_Q \geq 1} |\nabla \varphi_Q(x)|^2.$$

\textsuperscript{14} With the understanding that $j(x) = 0$ when the sum is empty.
Inserting definition (2.67) into the l. h. s., using triangle inequality, integrating over \(x \in (0, R)^2\), and using (2.58) on the r. h. s. we obtain

\[
\mathbb{E} \int_{(0,R)^2} |j|^2 \lesssim \sum_{Q, R \ge r_Q \ge 1} \mathbb{E} \int_Q |\nabla \varphi_Q|^2.
\]

Using (2.64) on the l. h. s. and (2.68) on the r. h. s. yields (2.63). \(\square\)

**Remark 2.9** The same argument in \(d > 2\) yields the bound \(\mathbb{E} W_{(0,R)^d} (\mu, n) \lesssim 1\). However, the interesting (well known) information from the proof is that in \(d > 2\) the main contribution comes from the term \(W_{(0,R)^d} (\mu, \lambda)\) which collects the contributions on the small scales.

**Proof of Lemma 2.6** W.l.o.g. it suffices to prove that for the Poisson point process \(\mu\) there exists a random radius \(r_*\) such that for any dyadic radii \(R \ge r_*\)

\[
\frac{1}{R^d} W_{(-R,R)^d} (\mu, n) + \frac{R^2}{n} (n - 1)^2 \lesssim \ln R. \tag{2.69}
\]

The statement will follow by choosing the maximum of this random radius and the one pertaining to \(\nu\).

**Step 1.** We claim that there exists a constant \(C\) and a random radius \(r_* < \infty\) a. s. such that for any dyadic radii \(R \ge r_*\)

\[
\frac{1}{R^2} W_{(-R,R)^2} (\mu, n) \le C \ln R. \tag{2.70}
\]

By Lemma 2.8 we may assume that (2.55) holds with \((0, R)^2\) replaced by \((0, 2R)^2\). Furthermore by stationarity of the Poisson point process we may assume that (2.55) holds in the form

\[
\frac{1}{R^2} \mathbb{E} W_{(-R,R)^2} (\mu, n) \lesssim \ln R, \tag{2.71}
\]

provided that \(R \gg 1\). By [9, Proposition 2.7] it follows that for any \(\varepsilon > 0\),

\[
\mathbb{P} \left( \left| \frac{1}{R^2 \ln R} \left[ W_{(-R,R)^2} (\mu, n) - \mathbb{E} W_{(-R,R)^2} (\mu, n) \right] > \varepsilon \right| \right) \lesssim \exp (-C \varepsilon \ln R).
\]

Then (2.70) follows by arbitrariness of \(\varepsilon\) together with a Borel-Cantelli argument.

**Step 2.** We show that there exists a constant \(C\) and random radius \(r_* < \infty\) a. s. such that for any dyadic radii \(R \ge r_*\)

\[
\frac{R^2}{n} (n - 1)^2 \le C \ln R. \tag{2.72}
\]

Since for large \(R \gg 1\) we have \(\frac{\ln R}{R^2} \ll 1\) (2.72) is equivalent to

\[
R^2 (n - 1)^2 \lesssim \ln R.
\]

\(\Diamond\) Springer
Since $n4R^2$ is Poisson distributed with parameter $4R^2$ by Cramér-Chernoff’s bounds [6, Theorem 1]

$$
P \left( R^2 \left(n - 1\right)^2 > \ln R \right) = P \left( \left| n4R^2 - 4R^2 \right| > CR \ln \frac{1}{2} R \right) \lesssim \exp \left( -C \ln R \right).$$

Finally, by a Borel-Cantelli argument (2.72) holds for any dyadic radii $R \geq r_\ast$. 

\section*{Funding} Open Access funding enabled and organized by Projekt DEAL.

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

\section*{References}

1. Ajtai, M., Komlós, J., Tusnády, G.: On optimal matchings. Combinatorica 4, 259–264 (1984)
2. Ambrosio, L., Glaudo, F., Trevisan, D.: On the optimal map in the 2-dimensional random matching problem. Discrete Contin. Dyn. Syst. 39(12), 7291–7308 (2019)
3. Ambrosio, L., Stra, F., Trevisan, D.: A PDE approach to a 2-dimensional matching problem. Probab. Theory Relat. Fields 173(1–2), 433–477 (2019)
4. Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. Numer. Math. 84(3), 375–393 (2000)
5. Bobkov, S.G., Ledoux, M.: A simple Fourier analytic proof of the AKT optimal matching theorem. Ann. Appl. Probab. (2019). https://api.semanticscholar.org/CorpusID:202573077
6. Boucheron, S., Lugosi, G., Massart, P.: Concentration Inequalities. Oxford University Press, Oxford (2013)
7. Caracciolo, S., Lucibello, C., Parisi, G., Sicuro, G.: Scaling hypothesis for the Euclidean bipartite matching problem. Phys. Rev. E 90(1), 012118 (2014)
8. Chatterjee, S., Peled, R., Peres, Y., Romik, D.: Gravitational allocation to Poisson points. Ann. Math. 172, 617–671 (2010)
9. Goldman, M., Huesmann, M., Otto, F.: A large-scale regularity theory for the Monge–Ampère equation with rough data and application to the optimal matching problem. arXiv:1808.09250 (2018)
10. Goldman, M., Huesmann, M., Otto, F.: Quantitative linearization results for the Monge–Ampère equation. Commun. Pure Appl. Math. 74, 2483–2560 (2021)
11. Goldman, M., Trevisan, D.: Convergence of asymptotic costs for random Euclidean matching problems. Probab. Math. Phys. (2020). https://api.semanticscholar.org/CorpusID:221555337
12. Hoffman, C., Holroyd, A.E., Peres, Y.: A stable marriage of Poisson and Lebesgue. Ann. Probab. 34(4), 1241–1272 (2006)
13. Holroyd, A.E.: Geometric properties of Poisson matchings. Probab. Theory Relat. Fields 150(3), 511–527 (2011)
14. Holroyd, A.E., Janson, S., Wästlund, J.: Minimal matchings of point processes. arXiv preprint arXiv:2012.07129 (2020)
15. Holroyd, A.E., Pemantle, R., Peres, Y., Schramm, O.: Poisson matching. Ann. l’I.H.P. Probab. Stat. 45(1), 266–287 (2009)
16. Huesmann, M.: Optimal transport between random measures. Ann. l’Institut Henri Poincaré Probab. Stat. 52(1), 196–232 (2016)
17. Huesmann, M., Sturm, K.-T.: Optimal transport from Lebesgue to Poisson. Ann. Probab. 41(4), 2426–2478 (2013)
18. Kallenberg, O.: Foundations of Modern Probability. Probability and its Applications (New York), 2nd edn. Springer-Verlag, New York (2002)
19. Koch, L., Otto, F.: Lecture notes on the harmonic approximation to quadratic optimal transport. arXiv preprint arXiv:2303.14462 (2023)
20. Last, G., Penrose, M.: Lectures on the Poisson Process. Institute of Mathematical Statistics Textbooks. Cambridge University Press (2017)
21. Ledoux, M.: On optimal matching of Gaussian samples. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI) 457 (Veroyatnost’ i Statistika. 25), 226–264 (2017)
22. Markó, R., Timár, A.: A Poisson allocation of optimal tail. Ann. Probab. 44(2), 1285–1307 (2016)
23. Stroock, D.W.: Probability Theory: An Analytic View, 2nd edn. Cambridge University Press (2010)
24. Talagrand, M.: The transportation cost from the uniform measure to the empirical measure in dimension $\geq 3$. Ann. Probab. 22(2), 919–959 (1994)
25. Villani, C.: Topics in Optimal Transportation, volume 58 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (2003)
26. Wu, L.: A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. Probab. Theory Relat. Fields 118(3), 427–438 (2000)

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