

## Transport cost estimates for random measures in dimension one\*

Martin Huesmann<sup>†</sup>

### Abstract

We show that there is a sharp threshold in dimension one for the transport cost between the Lebesgue measure  $\lambda$  and an invariant random measure  $\mu$  of unit intensity to be finite. We show that for *any* such random measure the  $L^1$  cost is infinite provided that the first central moments  $\mathbb{E}[|n - \mu([0, n])|]$  diverge. Furthermore, we establish simple and sharp criteria, based on the variance of  $\mu([0, n])$ , for the  $L^p$  cost to be finite for  $0 < p < 1$ .

**Keywords:** optimal transport; random measures; shift-coupling; allocation; extra head scheme.

**AMS MSC 2010:** Primary 60G57; 60G55, Secondary 49Q20.

Submitted to ECP on September 28, 2015, final version accepted on April 19, 2016.

Supersedes arXiv:1510.03601.

## 1 Introduction

In [7, 6] it was shown that there is a unique optimal coupling between the Lebesgue measure  $\lambda^d$  on  $\mathbb{R}^d$  and an invariant random measure  $\mu$  on  $\mathbb{R}^d$  of unit intensity *provided* that the *asymptotic mean transportation cost*

$$c_\infty = \liminf_{n \rightarrow \infty} \inf_{q \in \text{Cpl}(\lambda^d, \mu)} \frac{1}{n^d} \mathbb{E} \left[ \int_{\mathbb{R}^d \times [0, n]^d} \vartheta(|x - y|) q(dx, dy) \right] \quad (1.1)$$

is finite, where  $\text{Cpl}(\lambda^d, \mu)$  denotes the set of all couplings between  $\lambda^d$  and  $\mu$  and  $\vartheta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a strictly increasing and diverging function. Moreover, as the optimal coupling  $\hat{q}$  is concentrated on the graph of a random map  $T$ , i.e.  $\hat{q} = (\text{Id}, T)_* \lambda^d$ , a posteriori it can be shown that

$$c_\infty = \inf_{S, S_* \lambda^d = \mu} \mathbb{E}[\vartheta(|0 - S(0)|)]. \quad (1.2)$$

In principle, these results give a black box construction of allocations and invariant couplings suitable for applications, e.g. modelling of cellular structure via Laguerre tessellation [12] (and references therein) or the recent construction of unbiased shifts [10]. However, both conditions (1.1) and (1.2) are difficult to verify, mainly, because optimal couplings are highly non-local objects. For instance, consider the optimal semicoupling (cf. Section 2) between  $\lambda^d$  and a Poisson point process on  $B_n = [0, n]^d$ . It is

---

\*The author thanks Theo Sturm for valuable discussions and gratefully acknowledges funding by the CRC 1060 and support by the Hausdorff center for mathematics.

<sup>†</sup>Universität Bonn, Germany. E-mail: huesmann@iam.uni-bonn.de

an open problem to estimate the amount of mass that is transported from outside of  $B_n$  into  $B_n$ , for fixed  $n$  as well as asymptotically as  $n$  tends to  $\infty$ .

The aim of this note is to give in dimension one sharp and easily checkable conditions for the asymptotic mean transportation cost to be finite. For ease of exposition, in this note we focus on  $L^p$  cost, i.e. we consider  $\vartheta_p(r) := r^p$  for  $p > 0$ , and put

$$c_\infty(p) = \inf_{S, S_* \lambda^d = \mu} \mathbb{E}[\vartheta_p(|0 - S(0)|)] = \inf_{S, S_* \lambda^d = \mu} \mathbb{E}[|S(0)|^p].$$

We remark that similar results can be obtained for Orlicz-type cost functions as introduced in [17].

We denote by  $\text{Var}(Z)$  the variance of a random variable  $Z$ . We say that a random measure  $\mu$  satisfies a CLT if the sequence  $\left( (\mu([0, n]) - \mathbb{E}[\mu([0, n])]) / \sqrt{\text{Var}(\mu([0, n]))} \right)_n$  weakly converges to a standard normal distribution. We say a random measure  $\mu$  has a regular variance if  $f(n) := \text{Var}(\mu([0, n]))$  satisfies

$$\lim_{n \rightarrow \infty} \frac{a_n}{n} = 0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \frac{f(a_n)}{f(n)} = 0.$$

Our first result states

**Theorem 1.1.** Fix  $0 < p < 1$  and let  $\mu$  be an invariant random measure of unit intensity.

- i) If  $\limsup_{n \rightarrow \infty} \sqrt{\text{Var}(\mu([0, n]))} \cdot n^{p-1} = 0$ , then  $c_\infty(q) < \infty$  for all  $0 < q < p$ .
- ii) Assume that  $\mu$  has a regular variance and satisfies a CLT. If  $\limsup_{n \rightarrow \infty} \sqrt{\text{Var}(\mu([0, n]))} \cdot n^{p-1} > 0$ , then  $c_\infty(p) = \infty$ .

For the question of finiteness of  $c_\infty(p)$  or otherwise only the tail of  $\vartheta_p$  is relevant. Therefore,  $c_\infty(p) = \infty$  implies  $c_\infty(p') = \infty$  for all  $p' > p$  (see also [7, Lemma 5.1]).

**Remark 1.2.**  $\mu$  has a regular variance, if for example  $f$  is convex (recall  $f(n) = \text{Var}(\mu([0, n]))$ ) or if  $f$  is concave and there is  $p > 0$  such that  $f^p$  is convex. Indeed, assume that  $f$  is convex and assume for contradiction that  $1 \geq \liminf_{n \rightarrow \infty} \frac{f(a_n)}{f(n)} \geq c > 0$ . Then, we have (denoting by  $g$  the concave inverse function of  $f$ , i.e.  $g \circ f = f \circ g = \text{Id}$ , with  $f(0) = g(0) = 0$ ) for large  $n$  and some  $c' < c \leq 1$

$$a_n \geq g(c'f(n)) = g(c'f(n) + (1 - c')0) \geq c'g(f(n)) = c'n,$$

which is a contradiction to  $a_n \in o(n)$ . In the second case we can use the same argument by considering  $f = (f^p)^{1/p}$  and using the monotonicity of  $x \mapsto x^{1/p}$ .

Formally taking  $p = 1$  in Theorem 1.1 ii) indicates that  $c_\infty(1)$  might be infinite if  $\limsup_{n \rightarrow \infty} \sqrt{\text{Var}(\mu([0, n]))} = \infty$ . Unfortunately, the proof of Theorem 1.1 breaks down at  $p = 1$ . However, following Liggett [13, Section 3] and combining this with [11, Proposition 4.5] we get

**Theorem 1.3.** Let  $\mu$  be an invariant random measure of unit intensity.

If  $\limsup_{n \rightarrow \infty} \mathbb{E}[|n - \mu([0, n])|] = \infty$ , then  $c_\infty(1) = \infty$ .

Note that if  $\mu$  satisfies a CLT in  $L^1$  then  $\mathbb{E}[|n - \mu([0, n])|]$  behaves asymptotically like  $\sqrt{\text{Var}(\mu([0, n]))}$ .

Here are a few examples to which our results apply:

- i) The Poisson point process has finite transport cost iff  $p < 1/2$ . In particular, we recover the second part of Theorem 3.1 of [13].
- ii) Invariant determinantal random point fields [15] yield a wide and well studied class of random measures to which our results apply. Many of them satisfy a central limit theorem [16]. The behaviour of  $\text{Var}(\mu([0, n]))$  can be expressed nicely via the

integral kernel [15, Lemma 6]. For instance the determinantal random point field associated to the sine kernel

$$K(x, y) = \frac{\sin(\pi(x - y))}{\pi(x - y)}$$

satisfies  $\text{Var}(\mu([0, n])) \sim \log(n)$ . Hence, the transport cost are finite iff  $p < 1$  (see next point for the only if statement). This behaviour of the variance is not prototypical for determinantal point processes; for each  $0 < \beta < 1$  there is a determinantal point process with  $\text{Var}(\mu([0, n])) \sim n^\beta$ , see the last paragraph of Section 3 in [15].

- iii) The  $\text{Sine}_\beta$  point processes introduced in [20] appear as the limit of the bulk of eigenvalues of  $\beta$ -ensembles.  $\text{Sine}_\beta$  are translation invariant, satisfy a central limit theorem [8] and  $\text{Var}(\mu([0, n])) \sim 1/\beta \log(n)$ . From the large deviation result [3] it is possible to deduce that the assumption of Theorem 1.3 is satisfied. Hence, the transport cost are finite iff  $p < 1$ . Note that  $\text{Sine}_2$  is the determinantal process associated to the sine kernel.

Holroyd and Peres [5] call  $X = T(0)$  an *extra head scheme*. In view of (1.2) our results give moment properties for extra head schemes. In the case of a Poisson process precise moment properties of extra head schemes have been established in [4, 5, 7] in arbitrary dimensions. For other random measures (in the continuous setup of this article) this is—to the best of our knowledge—still open.

A natural interpretation of our results is to think of  $p^* := \sup\{p, c_\infty(p) < \infty\}$  as a measure of regularity of the random measure. For example in the case of the sine kernel process the repulsion of the particles causes a rigid behaviour reflected in the logarithmic growth of the variance, and hence in the transport cost estimates. Similar estimates in higher dimensions could be very useful to detect possible phase transitions, e.g. a phase transition in the parameter  $\beta$  for the equilibrium measures of the infinite dimensional system of interacting SDEs studied by Osada [14] (in dimension one these measures are conjectured to be—and proven to be for  $\beta = 1, 2, 4$ —the  $\text{Sine}_\beta$  processes). Therefore we end the introduction with the following challenging open problem:

**Open problem.** Is it possible to establish similar results in higher dimension; e.g. reducing the question of finiteness or infiniteness of transportation cost to the question of asymptotics of moments?

## 2 Preliminaries

We write  $\lambda^1 = \lambda$  and denote by  $(\Omega, \mathcal{F}, \mathbb{P})$  a generic probability space on which our random elements are defined. Given a map  $S$  and a measure  $\rho$  we denote the push-forward of  $\rho$  by  $S$  by  $S_*\rho = \rho \circ S^{-1}$ . The set of all  $\sigma$ -finite measures on a space  $X$  will be denoted by  $\mathcal{M}(X)$ . For a Polish space  $X$  we denote by  $\mathcal{B}(X)$  its Borel  $\sigma$ -algebra. For  $X = X_1 \times X_2$  we denote the projection on  $X_i$  by  $\text{proj}_i$ .

### 2.1 Random measures

Let  $\mu$  be a random  $\sigma$ -finite measure on  $\mathbb{R}$ , i.e. a measurable map  $\mu : \Omega \rightarrow \mathcal{M}(\mathbb{R})$ . We assume that  $\mathbb{R}$  acts on  $(\Omega, \mathcal{F})$  via a *measurable flow*  $\theta_t : \Omega \rightarrow \Omega, t \in \mathbb{R}$ , i.e. the mapping  $(\omega, t) \mapsto \theta_t\omega$  is  $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}) - \mathcal{F}$  measurable with  $\theta_0 = \text{Id}$  and  $\theta_t \circ \theta_s = \theta_{t+s}$  for  $s, t \in \mathbb{R}$ . A random measure  $\mu$  on  $\mathbb{R}$  is then called *invariant* (sometimes also *equivariant*) if for  $A \in \mathcal{B}(\mathbb{R}), t \in \mathbb{R}$  and  $\omega \in \Omega$  it holds that

$$\mu(\theta_t\omega, A - t) = \mu(\omega, A).$$

A random measure  $q$  on  $\mathbb{R} \times \mathbb{R}$  will be called invariant if for all  $A, B \in \mathcal{B}(\mathbb{R}), t \in \mathbb{R}$  and  $\omega \in \Omega$  it holds that

$$q(\theta_t \omega, A - t, B - t) = q(\omega, A, B).$$

For an invariant measure  $\mu$  we sometimes write  $\theta_t \mu(\omega) = \mu(\theta_t \omega)$ .

The *intensity* of an invariant random measure  $\mu$  on  $\mathbb{R}$  is defined as  $\mathbb{E}[\mu([0, 1])]$ ;  $\mu$  has unit intensity if  $\mathbb{E}[\mu([0, 1])] = 1$ .

A measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$  is called *stationary* if it is invariant under the action of  $\theta$ , i.e.  $\mathbb{P} \circ \theta_t = \mathbb{P}$  for all  $t \in \mathbb{R}$ .

**Remark 2.1.** We can think of  $\theta_t \omega$  as  $\omega$  shifted by  $-t$ , see Example 2.1 in [11].

From now on we will always assume to be in the setting described above.

So, let  $\mathbb{P}$  be a stationary probability measure and  $\mu$  be an invariant random measure. Let  $B \in \mathcal{B}(\mathbb{R})$  with  $0 < \lambda(B) < \infty$ . The *Palm measure*  $\mathbb{P}_\mu$  of  $\mu$  (with respect to  $\mathbb{P}$ ) is the measure on  $(\Omega, \mathcal{F})$  defined by

$$\mathbb{P}_\mu(A) := \frac{1}{\lambda(B)} \mathbb{E} \int_B 1_A(\theta_t \omega) \mu(\omega, dt).$$

As this is independent of  $B$ , we can deduce by a monotone class argument the *refined Campbell theorem*

$$\mathbb{E} \int f(\theta_t \omega, t) \mu(\omega, dt) = \int_\Omega \int_{\mathbb{R}} f(\eta, s) ds \mathbb{P}_\mu(d\eta) \tag{2.1}$$

for bounded and measurable  $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ . We refer to [19, Chapter 8] and [9] for more details on Palm theory.

Last and Thorisson [11, Proposition 4.5] show the following remarkable result which is crucial for the proof of Theorem 1.3.

**Theorem 2.2.** Consider two invariant random measures  $\xi$  and  $\eta$  and let  $T : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$  be measurable and satisfy

$$T(\theta_t \omega, s - t) = T(\omega, s) - t \quad s, t \in \mathbb{R}, \omega \in \Omega.$$

Then  $\mathbb{P}$ -a.s.  $T_* \xi = \eta$  iff for all  $A \in \mathcal{F}$

$$\mathbb{P}_\xi(\theta_{T(0)} \omega \in A) = \mathbb{P}_\eta(A).$$

Any map  $T$  as in the theorem will be called *allocation rule* or invariant transport map.

**Example 2.3.** If  $\mathbb{P}$  is stationary, the constant invariant random measure  $\lambda$  has Palm measure  $\mathbb{P}_\lambda = \mathbb{P}$ . In particular, given an invariant random measure  $\mu$  with unit intensity and an invariant transport map from  $\lambda$  to  $\mu$  which is measurably dependent only on the  $\sigma$ -algebra generated by  $\mu$  Theorem 2.2 yields a *shift-coupling*, see [1] and [18], between  $\mathbb{P}$  and  $\mathbb{P}_\mu$ , i.e. for all  $A \in \mathcal{F}$  it holds that  $\mathbb{P}[\theta_{T(0)} \omega \in A] = \mathbb{P}_\mu[A]$ . By considering the image measure  $\mathbb{P} \circ \mu^{-1}$  we can assume w.l.o.g. that  $(\Omega, \mathcal{F})$  is the canonical probability space  $(\mathcal{M}(\mathbb{R}), \mathcal{B}(\mathcal{M}(\mathbb{R})))$  and  $\mu$  the identity map. Then, Theorem 2.2 can be read as a shift-coupling between  $\mu$  and  $\mathbb{P}_\mu$ :

$$\mathbb{P}[\theta_{T(0)} \mu \in \cdot] = \mathbb{P}_\mu[\cdot].$$

## 2.2 Optimal transport between random measures

A *semicoupling* between two measures  $\nu$  and  $\eta$  on  $\mathbb{R}$  is a measure  $q$  on  $\mathbb{R} \times \mathbb{R}$  such that  $(\text{proj}_1)_* q \leq \nu$  and  $(\text{proj}_2)_* q = \eta$ . It is called *coupling* if additionally  $(\text{proj}_1)_* q = \nu$ . A semicoupling between  $\lambda$  and a random measure  $\mu$  is a random measure  $q : \Omega \rightarrow \mathcal{M}(\mathbb{R} \times \mathbb{R})$

such that for all  $\omega \in \Omega$  the measure  $q_\omega$  is a semicoupling between  $\lambda$  and  $\mu_\omega$ . It is called coupling if additionally  $q_\omega$  is a coupling between  $\lambda$  and  $\mu_\omega$  for all  $\omega \in \Omega$ . We denote the set of all couplings (resp. semicouplings) between  $\lambda$  and  $\mu$  by  $\text{Cpl}(\lambda, \mu)$  (resp.  $\text{SCpl}(\lambda, \mu)$ ).

Considering the cost-function  $c_p(x, y) = |x - y|^p$  for  $0 < p \leq 1$  we will be interested in the cost functional

$$\mathbb{W}_p(\nu, \eta) := \inf_{q \in \text{SCpl}(\nu, \eta)} \mathbb{E} \int |x - y|^p q(dx, dy).$$

By standard results in optimal transport, e.g. [21, Section 7.1],  $\mathbb{W}_p$  constitutes a metric on the space of probability measures and therefore also on the space of measures with fixed and finite mass equal to  $m \in (0, \infty)$ .

Let  $\mu$  be an invariant random measure with unit intensity. For  $q \in \text{SCpl}(\lambda, \mu)$  we set

$$\mathfrak{C}(q) = \sup_{n \geq 1} \frac{1}{n} \int_{\mathbb{R} \times [0, n]} |x - y|^p q(dx, dy).$$

By [6, Corollary 6.5], we have

$$\begin{aligned} c_\infty &:= \inf_{q \in \text{SCpl}(\lambda, \mu)} \mathfrak{C}(q) \\ &= \liminf_{n \rightarrow \infty} \frac{1}{n} \inf_{q \in \text{SCpl}(\lambda, \mu)} \int_{\mathbb{R} \times [0, n]} |x - y|^p q(dx, dy) \\ &= \inf_{q \in \text{SCpl}(\lambda, \mu)} \liminf_{n \rightarrow \infty} \frac{1}{n} \int_{\mathbb{R} \times [0, n]} |x - y|^p q(dx, dy). \end{aligned}$$

We sometimes write  $c_\infty(p)$  to stress the dependence on  $p$ .

**Definition 2.4.** Let  $\mu$  be an invariant random measure with unit intensity. A (semi) coupling  $q$  between  $\lambda$  and  $\mu$  is called

- asymptotically optimal if  $\mathfrak{C}(q) = c_\infty$ .
- optimal if it is asymptotically optimal and invariant.

The main results of [7, 6] show that there is a unique optimal coupling between  $\lambda$  and  $\mu$  provided that  $c_\infty < \infty$ . In particular, even though there are arbitrarily many asymptotically optimal couplings there is a unique invariant one. Moreover, the optimal coupling  $\hat{q}$  is concentrated on an invariant transport map  $T$ , i.e.  $\hat{q} = (\text{Id}, T)_* \lambda$ , which is measurably only dependent on the  $\sigma$ -algebra generated by the random measure  $\mu$ .

### 3 Proof of Theorem 1.1

#### 3.1 Proof of Theorem 1.1 i)

The strategy is to construct a coupling between  $\lambda$  and  $\mu$  which is not optimal but whose cost can be controlled nicely. To this end, we set  $Z_n := \mu([0, 2^n])$  and put

$$\bar{c}_n := 2^{-n} \mathbb{W}_p(\mathbb{1}_{[0, Z_n]} \lambda, \mathbb{1}_{[0, 2^n]} \mu).$$

By invariance of  $\lambda$  and  $\mu$  this equals  $\frac{1}{2}(\bar{c}_n + \bar{c}'_n)$  with  $Z'_n := \mu([2^n, 2^{n+1}]) = Z_{n+1} - Z_n$  and

$$\bar{c}'_n := 2^{-n} \mathbb{W}_p(\mathbb{1}_{[2^n, 2^n + Z'_n]} \lambda, \mathbb{1}_{[2^n, 2^{n+1}]} \mu).$$

By the triangle inequality for  $\mathbb{W}_p$  we have

$$\begin{aligned} &\mathbb{W}_p(\mathbb{1}_{[0, Z_{n+1}]} \lambda, \mathbb{1}_{[0, 2^{n+1}]} \mu) \\ &\leq \mathbb{W}_p(\mathbb{1}_{[0, Z_{n+1}]} \lambda, \mathbb{1}_{[0, Z_n]} \lambda + \mathbb{1}_{[2^n, 2^n + Z'_n]} \lambda) \\ &\quad + \mathbb{W}_p(\mathbb{1}_{[0, Z_n]} \lambda + \mathbb{1}_{[2^n, 2^n + Z'_n]} \lambda, \mathbb{1}_{[0, 2^n]} \mu + \mathbb{1}_{[2^n, 2^{n+1}]} \mu) \\ &\leq \mathbb{W}_p(\mathbb{1}_{[0, Z_{n+1}]} \lambda, \mathbb{1}_{[0, Z_n]} \lambda + \mathbb{1}_{[2^n, 2^n + Z'_n]} \lambda) \\ &\quad + \mathbb{W}_p(\mathbb{1}_{[0, Z_n]} \lambda, \mathbb{1}_{[0, 2^n]} \mu) + \mathbb{W}_p(\mathbb{1}_{[2^n, 2^n + Z'_n]} \lambda, \mathbb{1}_{[2^n, 2^{n+1}]} \mu), \end{aligned}$$

where the last inequality comes from forcing a certain allocation of the mass of the coupling between  $\mathbb{1}_{[0, Z_n)}\lambda + \mathbb{1}_{[2^n, 2^n + Z'_n)}\lambda$  and  $\mathbb{1}_{[0, 2^{n+1})}\mu$ . Together with the definition of  $\bar{c}_n$  and  $\bar{c}'_n$  this implies

$$\begin{aligned} 2^{n+1}(\bar{c}_{n+1} - \bar{c}_n) &= 2^{n+1}(\bar{c}_{n+1} - \frac{1}{2}(\bar{c}_n + \bar{c}'_n)) \\ &= W_p(\mathbb{1}_{[0, Z_{n+1})}\lambda, \mathbb{1}_{[0, 2^{n+1})}\mu) - W_p(\mathbb{1}_{[0, Z_n)}\lambda, \mathbb{1}_{[0, 2^n)}\mu) - W_p(\mathbb{1}_{[2^n, 2^n + Z'_n)}\lambda, \mathbb{1}_{[2^n, 2^{n+1})}\mu) \\ &\leq W_p(\mathbb{1}_{[0, Z_{n+1})}\lambda, \mathbb{1}_{[0, Z_n)}\lambda + \mathbb{1}_{[2^n, 2^n + Z'_n)}\lambda). \end{aligned}$$

The last expression can be estimated as follows. As  $r \mapsto r^p$  is concave (recall  $0 < p < 1$ ) the optimal coupling of two measures does not transport their common mass. Hence, in case that  $Z_n \leq 2^n$  we have to transport mass of amount  $2^n - Z_n$  at most distance  $2^n - Z_n + Z'_n$ . In case that  $Z_n > 2^n$  we have to transport mass of amount  $Z_n - 2^n$  at most distance  $Z_n - 2^n + Z'_n$ . Using the inequality  $(x + y)^p \leq x^p + y^p$  this implies

$$\begin{aligned} &W_p(\mathbb{1}_{[0, Z_{n+1})}\lambda, \mathbb{1}_{[0, Z_n)}\lambda + \mathbb{1}_{[2^n, 2^n + Z'_n)}\lambda) \\ &\leq \mathbb{E}[|Z_n - 2^n|(|Z_n - 2^n| + Z'_n)^p] \\ &\leq \mathbb{E}[|Z_n - 2^n|^{1+p} + |Z_n - 2^n|(Z'_n)^p] \end{aligned}$$

Using Jensen and Cauchy-Schwarz this can be further estimated by

$$\begin{aligned} &\mathbb{E}[|Z_n - 2^n|^{1+p} + |Z_n - 2^n|(Z'_n)^p] \\ &\leq \text{Var}(Z_n)^{(1+p)/2} + \text{Var}(Z_n)^{1/2}\mathbb{E}[(Z'_n)^{2p}]^{1/2} \\ &\leq \text{Var}(Z_n)^{(1+p)/2} + \text{Var}(Z_n)^{1/2}(\text{Var}(Z_n) + \mathbb{E}[Z_n^2])^{p/2} \\ &\leq \text{Var}(Z_n)^{(1+p)/2} + \text{Var}(Z_n)^{(1+p)/2} + \text{Var}(Z_n)^{1/2}2^{np}, \end{aligned}$$

where we used the identity  $\text{Var}(Z) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2$  in the second to last inequality and  $(x + y)^p \leq x^p + y^p$  in the last inequality. Therefore, we get

$$\bar{c}_{n+1} - \bar{c}_n \leq 2^{-n} \text{Var}(Z_n)^{(1+p)/2} + \frac{1}{2} \text{Var}(Z_n)^{1/2}2^{n(p-1)},$$

which readily implies the following lemma.

**Lemma 3.1.** *If  $\sum_{n \geq 1} (2^{-n} \text{Var}(Z_n)^{(1+p)/2} + \text{Var}(Z_n)^{1/2}2^{n(p-1)}) < \infty$  then  $c_\infty(p) < \infty$ .*

*Proof.* Put

$$c_n := \inf_{q \in \text{SCpl}(\lambda, \mu)} 2^{-n} \cdot \mathbb{E} \left[ \int_{\mathbb{R} \times [0, 2^n)} |x - y|^p q(dx, dy) \right].$$

Then, we have  $c_n \leq \bar{c}_n$  and hence  $c_\infty = \liminf_{n \rightarrow \infty} c_n \leq \liminf_{n \rightarrow \infty} \bar{c}_n =: \bar{c}_\infty$ . Therefore, it is sufficient to show that  $\bar{c}_\infty < \infty$ . However, this follows from

$$\liminf \bar{c}_n \leq \bar{c}_1 + \sum_{k \geq 1} \left( 2^{-k} \text{Var}(Z_k)^{(1+p)/2} + \text{Var}(Z_k)^{1/2}2^{k(p-1)} \right)$$

which is finite by assumption. □

*Proof of Theorem 1.1 i).* Assume that  $\limsup_{n \rightarrow \infty} \sqrt{\text{Var}(\mu([0, n]))} \cdot n^{p-1} = 0$ . We have to verify the condition of Lemma 3.1. By our assumption there is  $N \in \mathbb{N}$  such that for all  $n \geq N$  we have  $\text{Var}(Z_n) = \text{Var}(\mu([0, 2^n])) \leq 2^{2n(1-p)}$ . Hence, for  $0 < q < p$  we have

$$\begin{aligned} &\sum_{k \geq N} \left( 2^{-k} \text{Var}(Z_k)^{(1+q)/2} + \text{Var}(Z_k)^{1/2}2^{k(q-1)} \right) \\ &\leq \sum_{k \geq N} \left( 2^{-k} 2^{k(1-p)(1+q)} + 2^{k(q-1)-k(p-1)} \right) \\ &= \sum_{k \geq N} \left( 2^{k((1-p)(1+q)-1)} + 2^{k(q-p)} \right) < \infty, \end{aligned}$$

because  $(1 - p)(1 + q) < (1 - p)(1 + p) = 1 - p^2 < 1$  and  $q < p$ . Hence,  $\bar{c}_\infty(q) < \infty$ .  $\square$

**Remark 3.2.** Note that we have just shown that an equivalent condition in Lemma 3.1 would be that  $\sum_{k \geq 1} \text{Var}(Z_k)^{1/2} 2^{k(p-1)} < \infty$ . It is also not hard to see that the convergence of  $\sum_{k \geq 1} \text{Var}(Z_k)^{1/2} 2^{k(p-1)}$  is strictly stronger than the convergence of  $\sum_{k \geq N} 2^{-k} \text{Var}(Z_k)^{(1+q)/2}$  in the sense that the convergence of the second sum does not imply the convergence of the first.

**3.2 Proof of Theorem 1.1 ii)**

Denote by  $q_n$  the optimal semicoupling between  $\lambda$  and  $\mathbb{1}_{[0,n]}\mu$ . By Proposition 4.2 in [6], there is a transport map  $T_n$  and a random density  $\rho_n$  such that  $q_n = (\text{Id}, T_n)_*(\rho_n \lambda)$ . Put  $l_n := \inf\{x : \rho_n(x) > 0\}$  and  $r_n := \sup\{x : \rho_n(x) > 0\}$ . If  $l_n < 0$  (resp.  $r_n > n$ ) it follows by optimality that  $\rho_n = 1$  on  $[l_n, 0]$  (resp.  $[n, r_n]$ ). In that case, we put

$$a_n := T_n(l_n/2), \quad (\text{resp. } b_n := T_n(n + \frac{1}{2}(r_n - n))).$$

If  $l_n \geq 0$  (resp.  $r_n \leq n$ ) we put  $a_n = 0$  (resp.  $b_n = n$ ).

We claim that there exists a sequence of events  $(A_n)_n$  s.t.

- a)  $\liminf_{n \rightarrow \infty} \mathbb{P}[A_n] \geq c > 0$
- b) on  $A_n$  either  $|l_n| \geq 2\sqrt{\text{Var}(\mu([0, n]))}$  or  $|r_n - n| \geq 2\sqrt{\text{Var}(\mu([0, n]))}$
- c) on  $A_n$  there exists  $1 > \kappa > 0$  such that for large  $n$  either  $a_n \geq \kappa n$  or  $c_n := n - b_n \geq \kappa n$ , i.e.  $\liminf_{n \rightarrow \infty} (a_n + c_n)/n \geq \kappa$ .

As a consequence of concavity of  $r \mapsto r^p$  we have  $T_n(x) \geq T_n(y)$  for all  $l_n \leq x \leq y \leq 0$  (resp.  $T_n(x) \leq T_n(y)$  for all  $n \leq y \leq x \leq r_n$ ), see e.g. [2]. Hence, assuming a),b) and c), we can argue

$$\begin{aligned} c_\infty &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[ \mathbb{1}_{A_n} \int_{[l_n, \frac{l_n}{2}] \cup [n + \frac{r_n - n}{2}, r_n]} |x - T_n(x)|^p \lambda(dx) \right] \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \kappa^p n^p \sqrt{\text{Var}(\mu([0, n]))} \mathbb{P}[A_n] \\ &\geq \liminf_{n \rightarrow \infty} \kappa^p n^{p-1} \sqrt{\text{Var}(\mu([0, n]))} \cdot c = \infty, \end{aligned}$$

by assumption. Hence, it remains to establish the claim.

We put  $Y_n := \mu([0, n])$  and set

$$\tilde{A}_n = \{Y_n \geq n + 4\sqrt{\text{Var}(Y_n)}\}.$$

By the CLT, it follows that  $\liminf_{n \rightarrow \infty} \mathbb{P}[\tilde{A}_n] \geq \tilde{c} > 0$  so that a) holds. On  $\tilde{A}_n$  we have to transport mass of amount at least  $4\sqrt{\text{Var}(Y_n)}$  into the interval  $[0, n]$ . Hence, either  $|l_n| \geq 2\sqrt{\text{Var}(Y_n)}$  or  $|r_n - n| \geq 2\sqrt{\text{Var}(Y_n)}$  so that b) holds also. It remains to show c). We will show that on  $\tilde{A}_n$  it is not possible that both  $(a_k)_k \in o(k)$  and  $(c_k)_k \in o(k)$ . Put  $Y_{a_n} = \mu([0, a_n])$  and  $Y'_{c_n} = \mu([b_n, n])$ . Then, we have

$$\begin{aligned} &\mathbb{P}[\tilde{A}_n, (a_k)_k \in o(k), (c_k)_k \in o(k)] \\ &\leq \mathbb{P}[Y_{a_n} + Y'_{c_n} \geq a_n + c_n + 2\sqrt{\text{Var}(Y_n)}, (a_k)_k \in o(k), (c_k)_k \in o(k)], \end{aligned}$$

since on  $\tilde{A}_n$  there is no transport from outside of  $(T_n(l_n), T_n(r_n))$  into  $(T_n(l_n), T_n(r_n))$ , by concavity of the cost function, and at most half of the Lebesgue mass that is transported from outside of  $[0, n]$  (the total excess is at least  $4\sqrt{\text{Var}(\mu([0, n]))}$ ) is transported into

$(\tilde{a}_n, T(l_n)] \cup [T(r_n), \tilde{b}_n)$  (where  $\tilde{a}_n = a_n$  if  $a_n > 0$  and  $\tilde{a}_n = T(l_n)$  otherwise and similarly for  $\tilde{b}_n$ ). Hence,

$$\begin{aligned} & \mathbb{P}[\tilde{A}_n, (a_k)_k \in o(k), (c_k)_k \in o(k)] \\ & \leq \mathbb{P}[Y_{a_n} \geq a_n + \sqrt{\text{Var}(Y_n)}, (a_k)_k \in o(k)] + \mathbb{P}[Y'_{c_n} \geq c_n + \sqrt{\text{Var}(Y_n)}, (c_k)_k \in o(k)] \end{aligned}$$

We consider these two terms separately and start with the first one. We put  $\mathbb{P}_n^a := (a_n)_* (\mathbb{1}_{(a_k)_k \in o(k)} \mathbb{P})$  and set  $a_n^* := \sup\{x : x \in \text{supp}(\mathbb{P}_n^a)\} \in o(n)$ . Then, we have

$$\begin{aligned} & \mathbb{P}[Y_{a_n} \geq a_n + \sqrt{\text{Var}(Y_n)}, (a_k)_k \in o(k)] \\ & \leq \frac{1}{\text{Var}(Y_n)} \mathbb{E}[(Y_{a_n} - a_n)^2, (a_k)_k \in o(k)] \\ & \leq \frac{1}{\text{Var}(Y_n)} \int \text{Var}(Y_t) \mathbb{P}_n^a(dt) \\ & \leq \frac{\text{Var}(Y_{a_n^*})}{\text{Var}(Y_n)} = \frac{f(a_n^*)}{f(n)}, \end{aligned}$$

which goes to zero by the assumption that  $\mu$  has a regular variance.

The term  $\mathbb{P}[Y'_{c_n} \geq c_n + \sqrt{\text{Var}(Y_n)}, (c_k)_k \in o(k)]$  can be treated analogously. Hence,

$$\mathbb{P}[\tilde{A}_n, (a_k)_k \in o(k), (c_k)_k \in o(k)] \rightarrow 0.$$

By making the sets  $\tilde{A}_n$  slightly smaller yielding sets  $A'_n$  we can therefore assume that for large  $n$ , say  $n > N$ , on  $A'_n$  either  $(a_k)_k \in \Theta(k)$  or  $(c_k)_k \in \Theta(k)$  (since  $a_n, c_n \leq n$ ), property b) holds and  $\liminf P[A'_n] \geq \tilde{c}/2 = c'$ . This means that for any  $\omega \in A'_n$  there is  $\kappa'(\omega) > 0$  such that for large  $n$  we have either  $a_n(\omega) \geq \kappa'(\omega)n$  or  $c_n(\omega) \geq \kappa'(\omega)n$ . In particular,  $\{\kappa' > 0\} \supset \tilde{A}_n$  for all  $n > N$ . Take  $\kappa > 0$  such that  $\mathbb{P}[\kappa' < \kappa] \leq c'/2$  and set  $A_n := A'_n \cap \{\kappa' \geq \kappa\}$ . Then  $(A_n)_{n \geq N}$  satisfy the required properties a), b) and c).

### 4 Proof of Theorem 1.3

As indicated in the introduction the proof follows from the reasoning as in Section 3 of [13] together with Proposition 4.5 of [11]. Let  $X$  be some real valued random variable and  $\mathbb{P}' := (\theta_{-X})_* \mathbb{P} = \mathbb{P} \circ \theta_X$ , i.e.  $\mathbb{P}$  and  $\mathbb{P}'$  are *shift-coupled* by  $X$  (cf. [1, 19]). Then we have for any measurable  $f : \Omega \rightarrow [-1, 1]$

$$\begin{aligned} & \left| \frac{1}{t} \int_0^t f(\omega) ((\theta_{-s})_* \mathbb{P}(d\omega) - (\theta_{-s})_* \mathbb{P}'(d\omega)) ds \right| \\ & = \left| \frac{1}{t} \int_0^t \mathbb{E}[f(\theta_{-s}\omega) - f(\theta_{-s-X}\omega)] ds \right| \\ & = \left| \frac{1}{t} \mathbb{E} \left[ \int_{\mathbb{R}} f(\theta_{-s}\omega) (\mathbb{1}_{[0,t]}(s) - \mathbb{1}_{[X, X+t]}(s)) ds \right] \right| \\ & \leq \frac{1}{t} \mathbb{E} \left[ \int_{\mathbb{R}} |\mathbb{1}_{[0,t]}(s) - \mathbb{1}_{[X, X+t]}(s)| ds \right] \\ & = \frac{2}{t} \mathbb{E}[|X| \wedge t]. \end{aligned}$$

Hence, we have derived the *shift-coupling inequality* (cf. [18] and [19, p. 166])

$$\left\| \frac{1}{t} \int_0^t ((\theta_{-s})_* \mathbb{P}) - (\theta_{-s})_* \mathbb{P}' ds \right\| \leq \frac{2}{t} \mathbb{E}[|X| \wedge t],$$

where  $\|\cdot\|$  denotes the total variation distance.



By Theorem 2.2, any invariant transport map  $T$  balancing  $\lambda$  and  $\mu$ , i.e. transporting  $\lambda$  to  $\mu$ , induces a shift-coupling of  $\mathbb{P}$  with its Palm-measure  $\mathbb{P}_\mu$ . By (1.2),  $c_\infty = \inf_{T, T_*\lambda=\mu} \mathbb{E}[|T(0)|]$  and, by the results of [6], the infimum is attained by a unique map  $\hat{T}$  which is measurably dependent only on the  $\sigma$ -algebra generated by  $\mu$ . Hence,  $X := \hat{T}(0)$  shift-couples  $\mathbb{P}$  and  $\mathbb{P}_\mu$  and, by (1.2), we need to show that  $\mathbb{E}[|X|] = \infty$ .

By stationarity of  $\mathbb{P}$  we have  $\frac{1}{t} \int_0^t (\theta_{-s})_* \mathbb{P} ds = \mathbb{P}$  and by the refined Campbell theorem (2.1) it follows that for any bounded and non-negative measurable function  $f : \Omega \rightarrow \mathbb{R}$  and  $g(\omega, s) := \mathbb{1}_{[0,t]}(s)f(\theta_{-s}\omega)$  we have

$$\begin{aligned} \int_\Omega \int_0^t f(\omega) (\theta_{-s})_* \mathbb{P}_\mu(d\omega) ds &= \int_\Omega \int_0^t f(\theta_{-s}\omega) \mathbb{P}_\mu(d\omega) ds \\ &= \int_\Omega \int_{\mathbb{R}} g(\omega, s) \mathbb{P}_\mu(d\omega) ds = \int f(\omega) \mu_\omega([0, t]) \mathbb{P}(d\omega) \end{aligned}$$

Hence, we have  $\int_0^t (\theta_{-s})_* \mathbb{P}_\mu(d\omega) ds = \mu_\omega([0, t])\mathbb{P}(d\omega)$ . Putting everything together, we get (recall  $\|f d\nu - g d\nu\| = \int |f - g| d\nu$ )

$$\left\| \frac{1}{t} \int_0^t ((\theta_{-s})_* \mathbb{P} - (\theta_{-s})_* \mathbb{P}_\mu) ds \right\| = \mathbb{E} \left[ \left| 1 - \frac{\mu([0, t])}{t} \right| \right] \leq \frac{2}{t} \mathbb{E}[|X| \wedge t].$$

By assumption, we have  $\limsup_{t \rightarrow \infty} \mathbb{E}[|t - \mu([0, t])|] = \infty$ . This implies

$$\mathbb{E}[|X|] \geq \limsup_{t \rightarrow \infty} \mathbb{E}[|X| \wedge t] \geq \limsup_{t \rightarrow \infty} \frac{1}{2} \mathbb{E}[|t - \mu_\omega([0, t])|] = \infty,$$

which proves the result.

**Remark 4.1.** Following the argumentation in [13, Section 3] we can recover the assertion of Theorem 1.1 ii) in the setting of Theorem 1.3 assuming additionally that  $\mathbb{E}[|t - \mu([0, t])|] \sim \sqrt{\text{Var}(\mu([0, t]))}$  for large  $t$ . Indeed, we have with  $Y_t = \mu([0, t])$

$$\frac{2}{t} \mathbb{E}[|X| \wedge t] \geq \frac{1}{t} \mathbb{E}[|t - Y_t|] \sim \frac{1}{t} \sqrt{\text{Var}(Y_t)}.$$

By assumption, we have  $\limsup_{t \rightarrow \infty} \sqrt{\text{Var}(Y_t)} t^{p-1} \geq C > 0$ . Therefore, we have for  $t$  large enough

$$\sqrt{\text{Var}(Y_t)} \geq C' t^{1-p}.$$

This implies

$$0 < C' \leq \mathbb{E} \left[ \frac{|X| \wedge t}{t^{1-p}} \right] \leq \mathbb{E}[|X|^p],$$

since  $\frac{|X| \wedge t}{t^{1-p}} \leq |X|^p$ . Assuming  $\mathbb{E}[|X|^p] < \infty$  implies by the dominated convergence theorem that

$$0 < C' \leq \mathbb{E} \left[ \frac{|X| \wedge t}{t^{1-p}} \right] \rightarrow 0 \quad \text{as } t \rightarrow \infty,$$

which is a contradiction.

## References

- [1] D. J. Aldous and H. Thorisson, *Shift-coupling*, Stochastic Process. Appl. **44** (1993), no. 1, 1–14. MR-1198659
- [2] W. Gangbo and R. J. McCann, *The geometry of optimal transportation.*, Acta Mathematica **177** (1996), no. 2, 113–161. MR-1440931
- [3] D. Holcomb and B. Valkó, *Large deviations for the sine $_\beta$  and sch $_\tau$  processes*, Probab. Theory Related Fields **163** (2015), no. 1–2, 339–378. MR-3405620

- [4] A.E. Holroyd and T.M. Liggett, *How to find an extra head: optimal random shifts of Bernoulli and Poisson random fields*, Ann. Probab. **29** (2001), no. 4, 1405–1425. MR-1880225
- [5] A.E. Holroyd and Y. Peres, *Extra heads and invariant allocations*, Ann. Probab. **33** (2005), no. 1, 31–52. MR-2118858
- [6] M. Huesmann, *Optimal transport between random measures*, Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 1, 196–232. MR-3449301
- [7] M. Huesmann and K.T. Sturm, *Optimal transport from Lebesgue to Poisson*, Ann. Probab. **41** (2013), no. 4, 2426–2478. MR-3112922
- [8] E. Kritchevski, B. Valkó, and B. Virág, *The scaling limit of the critical one-dimensional random Schrödinger operator*, Comm. Math. Phys. **314** (2012), no. 3, 775–806. MR-2964774
- [9] G. Last, *Modern random measures: Palm theory and related models*, New perspectives in stochastic geometry, Oxford Univ. Press, Oxford, 2010, pp. 77–110. MR-2654676
- [10] G. Last, P. Mörters, and H. Thorisson, *Unbiased shifts of Brownian motion*, Ann. Probab. **42** (2014), no. 2, 431–463. MR-3178463
- [11] G. Last and H. Thorisson, *Invariant transports of stationary random measures and mass-stationarity*, Ann. Probab. **37** (2009), no. 2, 790–813. MR-2510024
- [12] C. Lautensack and S. Zuyev, *Random Laguerre tessellations*, Adv. in Appl. Probab. **40** (2008), no. 3, 630–650. MR-2454026
- [13] T. M. Liggett, *Tagged particle distributions or how to choose a head at random.*, In and out of equilibrium. Probability with a physics flavor. Papers from the 4th Brazilian school of probability, Mambucaba, Brazil, August 14–19, 2000, Boston: Birkhäuser, 2002, pp. 133–162. MR-1901951
- [14] H. Osada, *Infinite-dimensional stochastic differential equations related to random matrices*, Probab. Theory Related Fields **153** (2012), no. 3–4, 471–509. MR-2948684
- [15] A. Soshnikov, *Determinantal random point fields*, Uspekhi Mat. Nauk **55** (2000), no. 5(335), 107–160. MR-1799012
- [16] A. Soshnikov, *Gaussian limit for determinantal random point fields*, Ann. Probab. **30** (2002), no. 1, 171–187. MR-1894104
- [17] K.-T. Sturm, *Generalized Orlicz spaces and Wasserstein distances for convex–concave scale functions*, Bull. Sci. Math. **135** (2011), no. 6, 795–802. MR-2838101
- [18] H. Thorisson, *Shift-coupling in continuous time*, Probab. Theory Related Fields **99** (1994), no. 4, 477–483. MR-1288066
- [19] H. Thorisson, *Coupling, stationarity, and regeneration*, Springer New York, 2000. MR-1741181
- [20] B. Valkó and B. Virág, *Continuum limits of random matrices and the Brownian carousel*, Invent. Math. **177** (2009), no. 3, 463–508. MR-2534097
- [21] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR-1964483