Scaling hypothesis for the Euclidean bipartite matching problem. II. Correlation functions

Sergio Caracciolo

Dipartimento di Fisica, University of Milan and INFN, via Celoria 16, I-20133 Milan, Italy

Gabriele Sicuro

Centro Brasileiro de Pesquisas Físicas, Rua Xavier Sigaud 150, 22290-180 Rio de Janeiro, Brazil

(Received 2 April 2015; published 18 June 2015)

We analyze the random Euclidean bipartite matching problem on the hypertorus in $d$ dimensions with quadratic cost and we derive the two-point correlation function for the optimal matching, using a proper ansatz introduced by Caracciolo et al. [Phys. Rev. E 90, 012118 (2014)] to evaluate the average optimal matching cost. We consider both the grid-Poisson matching problem and the Poisson-Poisson matching problem. We also show that the correlation function is strictly related to the Green’s function of the Laplace operator on the hypertorus.

DOI: 10.1103/PhysRevE.91.062125

PACS numbers: 02.50.Cw, 02.10.Ox

I. THE MATCHING PROBLEM: AN INTRODUCTION

The assignment problem, or bipartite matching problem, is a classical combinatorial optimization problem in which two sets of $N$ elements, $B = \{b_i\}_{i=1}^N$ and $R = \{r_i\}_{i=1}^N$, are considered. An assignment is an element $\pi \in \mathcal{S}_N$ of the set $\mathcal{S}_N$ of permutations of $N$ elements, such that $b_i \mapsto r_{\pi(i)}$. Moreover, a cost function, $w: \mathcal{R} \times B \rightarrow \mathbb{R}^+$, $(r_i, b_j) \mapsto w_{ij}$, is given, from which we can define the total cost of a certain assignment $\pi$ as

$$E_N[\pi; w] := \frac{1}{N} \sum_{i=1}^N w_{i\pi(i)}. \quad (1)$$

In the bipartite matching problem we want to find the permutation $\pi^*$ that minimizes the previous quantity for a given function $w$. From a computational point of view, the problem belongs to the $P$ computational complexity class and it can be efficiently solved using fast algorithms [1–3].

If random instances are considered, i.e., $w_{ij}$ are random quantities, we are usually interested on the average optimal cost,

$$E_N := E_N[\pi^*; w], \quad (2)$$

where we denoted by $\mathbb{E}$ the expectation over all the possible instances $w$. If the values $\{w_{ij}\}$ are independent and identically distributed random variables, the problem is usually called random assignment problem: in this case, the average optimal cost and its properties in the large $N$ limit were investigated both with statistical physics techniques [4] and probability arguments [5].

In a more complicated variation of the random assignment problem, the so-called Euclidean bipartite matching problem (EBMP), the two sets $\mathcal{R}$ and $\mathcal{B}$ are in one-to-one correspondence with uniformly generated random points on the unit hypercube $\Omega_d := [0,1]^d \subset \mathbb{R}^d$, while the weight $w_{ij}$ is a function of the Euclidean distance $\|r_i - b_j\|$ (for simplicity, we identify the elements $b_i$ and $r_j$ with the corresponding geometric points in $\Omega_d$): in this case correlations between different values $\{w_{ij}\}$ appear and a proper mathematical treatment is more complicated. In the following we will consider weight functions in the form

$$w_{ij} := \|r_i - b_j\|^p, \quad p \in \mathbb{R}^+. \quad (3)$$

Mézard and Parisi [6] considered the previous problem on the hypercube $\Omega_d$ for any value of $d$ in the large $N$ limit, assuming that correlations can be treated as perturbations to the purely random case and evaluating approximately the average optimal cost through replica arguments. An exact solution to the problem for the $d = 1$ and $p > 1$ case is provided in [7,8], where the average optimal cost and correlation functions are computed; moreover, the correspondence between the matching problem and the Brownian bridge process on the line and the circle is proved.

Denoting by $\pi^*$ the optimal permutation for a given instance, we introduce the optimal matching ray:

$$\mathbf{m}(r_i) := b_{\pi^*(i)} - r_i, \quad i = 1, \ldots, N. \quad (4)$$

The optimal cost is

$$E_N^{(p)}[\mathbf{m}; \{\mathcal{R}, \mathcal{B}\}] := \frac{1}{N} \sum_{i=1}^N \|\mathbf{m}(r_i)\|^p, \quad (5)$$

$$E_N^{(p)}(d) := \frac{1}{\mathbb{E}_N^{(p)}[\mathbf{m}; \{\mathcal{R}, \mathcal{B}\}]. \quad (6)$$

where $d$ is the dimensionality of the Euclidean space and the average $\mathbb{E}$ is performed over the positions of the points.

The scaling properties of the optimal matching ray, and therefore of the optimal cost, are known to the literature for $p > 1$ [8–10], being

$$\|\mathbf{m}(x)\| \sim \begin{cases} \frac{1}{\sqrt{N}} & \text{for } d = 1, \\ \frac{\ln N}{N} & \text{for } d = 2, \\ \frac{1}{\sqrt{N}} & \text{for } d \geq 3. \end{cases} \quad (7)$$

In the present paper we are interested in the correlation function of the optimal matching ray $\mathbf{m}$ in the large $N$ limit; we will assume periodic boundary conditions on the hypercube.
forms of the correlation functions obtained for 
Finally, we will give numerical evidences that the functional 
normalized optimal matching ray 

Ω_d; i.e., we will consider the problem on the flat hypertorus $\mathbb{T}^d$ in $d$ dimensions. We will analyze both the case in which two sets of random points are considered and the case in which one set of points is supposedly fixed on a regular hypercubic lattice, while the second set is obtained from a Poisson process (see, e.g., Fig. 1 for a pictorial representation of a realization of the two-dimensional problem). To introduce our results, in Sec. II we will review the Monge-Kantorovič formulation of the optimal transport problem, from which a suitable general ansatz, already used by Caracciolo et al. [11], is derived for the expression of the optimal matching ray in the continuum limit for the $p = 2$ case. Using this working ansatz, in Sec. II B we will consider the EBMP with quadratic cost on $\mathbb{T}^d$, and we will give evidences that, in the large $N$ limit, the correlation function

$$C_d(x) = \frac{m(r_1) \cdot m(r_2)}{\|m(r_1)\|}$$ \hspace{1cm} (8)$$

is related to the Green’s function of the Laplacian operator on $\mathbb{T}^d$. We will consider also the correlation function for the normalized optimal matching ray

$$\sigma(r_1) := \frac{m(r_1)}{\|m(r_1)\|}.$$ \hspace{1cm} (9)$$

Finally, we will give numerical evidences that the functional forms of the correlation functions obtained for $p = 2$ in the two-dimensional case are in good agreement with the numerical results for $p = 1$ and $p = 3$ in the same dimension.

To our knowledge, these results are new to the literature, where only the $d = 1$ case is evaluated explicitly [7,8]. The present work can be seen as a natural expansion and completion of a previous work of Caracciolo et al. [11].

II. THE MONGE-KANTOROVICH MASS TRANSFER PROBLEM AND CORRELATION FUNCTIONS

A. The Monge-Kantorovič problem

The Monge-Kantorovič transfer problem is a well-studied problem in measure theory [12], in the context of transportation theory. Let us suppose that we are given two nonnegative smooth density measures $\rho_1$ and $\rho_2$ on a compact subset $\Omega \subset \mathbb{R}^d$, $\int_{\Omega} \rho_1(x) \, dx = \int_{\Omega} \rho_2(x) \, dx$. Suppose also that a work function $w : \Omega \times \Omega \to \mathbb{R}^+$ is given. We want to find a one-to-one mapping $M : \Omega \to \Omega$, called optimal transport map, such that the following conditions are satisfied:

1. $M \in T$, $T$ set of suitable transport maps, defined as follows:

$$T := \left\{ M : \Omega \to \Omega \left| \int_{\Omega} \rho_1(x) \, dx = \int_{\Omega} \rho_2(x) \, dx \right. \right\}.$$ \hspace{1cm} (10)$$

2. $M$ minimizes a certain cost functional,

$$E[M; w] := \int_{\Omega} w(x, M(x)) \rho_1(x) \, dx,$$ \hspace{1cm} (11)$$

where $w : \Omega \times \Omega \to \mathbb{R}^+$ is a transportation cost function, i.e., $E[M; w] = \min_{M \in T} E[M; w]$.

Interestingly, it can be proven that, if $w(x, y) = \|x - y\|^p$ with $p \in (1, +\infty)$ the condition $M \in T$ can be reexpressed as a change-of-variable formula [13],

$$\rho_1(x) = \rho_2[M(x)] \det J_M(x), \quad \forall x \in \Omega.$$ \hspace{1cm} (12)$$

where $J_M(x)$ is the Jacobian matrix of $M$, $[J_M(x)]_{ij} = \frac{\partial M_i}{\partial x_j}(x)$. Moreover, if $p = 2$, it can be proved [13] that the optimal transport map can be expressed as a gradient of a scalar potential $\varphi$, i.e.,

$$M(x) = \nabla \varphi(x).$$ \hspace{1cm} (13)$$

It follows that the solution $M$ of the transport problem has to be identified among the vector fields in the form of Eq. (13), with $\varphi$ satisfying the following relation:

$$\rho_1(x) = \rho_2[\nabla \varphi(x)] \det \text{Hess} \varphi(x), \quad \forall x \in \Omega.$$ \hspace{1cm} (14)$$

In the previous expression Hess $\varphi(x)$ is the Hessian matrix, $[\text{Hess} \varphi(x)]_{ij} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x)$. The nonlinear Eq. (14) is known to the literature as Monge-Ampère equation.

Consider now the Monge-Kantorovič problem in $\Omega \subset \mathbb{R}^d$, $\int_{\Omega} \rho_1(x) \, dx = \int_{\Omega} \rho_2(x) \, dx = 1$, with work function $w(x, y) = \|x - y\|^2$; suppose moreover that

$$\rho_1(x) = 1 + \delta \rho_1(x) \quad \text{and} \quad \rho_2(x) = 1 + \delta \rho_2(x),$$ \hspace{1cm} (15)$$

where

$$|\delta \rho_1(x)| \ll 1 \quad \text{and} \quad |\delta \rho_2(x)| \ll 1 \quad \forall x \in \Omega.$$ \hspace{1cm} (16)$$

We expect that $M(x) = x + m(x)$, $\|m(x)\| \ll 1 \quad \forall x \in \Omega$: in the first-order approximation, det $J_M(x) \approx 1 + \nabla \cdot m(x)$, so we have that

$$\nabla \cdot m(x) \approx \rho_1(x) - \rho_2(x) =: \rho(x);$$ \hspace{1cm} (17)$$

in particular, using the fact that $m = \nabla \varphi$, we have that in the limit of our approximation the Poisson equation holds:

$$\Delta \varphi(x) = \rho(x).$$ \hspace{1cm} (18)$$

Note that in this case the total cost of the transport is given by

$$E[M; \|\gamma\|^2] = \int_{\Omega} \|\nabla \varphi(x)\|^2 \, dx.$$ \hspace{1cm} (19)$$

In the following we will consider the unit hypercube, $\Omega \equiv [0, 1]^d$, with periodic boundary conditions, i.e., we will work on the flat hypertorus $\mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d$. Being $\int_{\Omega} \rho(x) \, dx = 0$, 062125-2
Eq. (18) has a unique solution on the compact manifold $T^d$, given by
\begin{equation}
\phi(x) = \int_{T^d} \rho(y) G_d(y, x) \, dy,
\end{equation}
where $G_d$ is the Green’s function for the Laplace operator $\Delta$ on $T^d$ defined by the relation
\begin{equation}
\Delta G_d(x, y) = \delta^{(d)}(x - y) - 1,
\end{equation}
the solution of which can be written as
\begin{equation}
G_d(x, y) \equiv G_d(x - y) = - \sum_{n \in \mathbb{Z}^d \setminus \{0\}} e^{2\pi in \cdot (x - y)} / 4\pi^2 |n|^2.
\end{equation}
In this linear approximation, the transport cost is given by
\begin{equation}
E[M, \|\bullet\|^2] = - \int_{T^d} \rho(x) G_d(x, y) \rho(y) \, dy \, dx
= \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \frac{\tilde{\beta}(n)^2}{4\pi^2 \|n\|^2},
\end{equation}
where
\begin{equation}
\tilde{\beta}(n) := \int_{T^d} \rho(x) e^{-2\pi i n \cdot x} \, dx.
\end{equation}

B. Correlation functions for the EBMP on the hypertorus

In the previous section we introduced the Monge-Kantorovič transport problem and we obtained also a set of simple results for the transport problem with quadratic cost between two almost uniform measures, through a proper linearization. We want to extract useful information about the discrete combinatorial problem from the continuum problem using the fact that, in the large $N$ limit, the EBMP between two sets of points on the hypertorus $T^d$ appears as a transport problem between two atomic measures that can be assumed as almost uniform measures on the domain of interest. This naive approach is justified a posteriori by the excellent agreement between theoretical predictions and numerical results.

In particular, let us denote by $B := \{b_i\}_{i=1, \ldots, N} \subset \Omega_d$ and $R := \{r_i\}_{i=1, \ldots, N} \subset \Omega_d$ two sets of points in $\Omega_d$, each set of cardinality $N$. The optimal cost of the matching on the flat hypertorus $T^d$ with quadratic cost is therefore given, in the notation above, by
\begin{equation}
E^2_N[\rho; \{R, B\}] = \frac{1}{N} \sum_{i=1}^N \|\mathbf{m}(r_i)\|^2.
\end{equation}
In the previous formula, $\mathbf{m}$ is the (geodesic) optimal matching ray on $T^d$. Let us now introduce two atomic measure densities:
\begin{equation}
\rho_R(x) = \frac{1}{N} \sum_{i=1}^N \delta^{(d)}(x - r_i),
\end{equation}
\begin{equation}
\rho_B(x) = \frac{1}{N} \sum_{i=1}^N \delta^{(d)}(x - b_i).
\end{equation}
In the following, we will assume, as a working ansatz, that, in the large $N$ limit, $\mathbf{m} \rightarrow \mathbf{m}(x) \equiv \nabla \phi(x)$ in such a way that $\mathbf{M}(x) = x + \mathbf{m}(x)$ is an optimal transport map between the two almost uniform measures above. Under the hypothesis that at least one set of points is randomly generated, we introduce the following correlation function:
\begin{equation}
C_d(x, y) := \nabla \phi(x) \cdot \nabla \phi(y),
\end{equation}
and, by Eq. (20), we obtain
\begin{equation}
C_d(x, y) \equiv C_d(x - y) = \int \nabla_x G_d(z - x) \cdot \nabla_w G_d(w - y) \rho(z) \rho(w) \, dz \, dw,
\end{equation}
where we denoted by
\begin{equation}
\rho(x) := \rho_R(x) - \rho_B(x) = \frac{1}{N} \sum_{i=1}^N \left[ \delta^{(d)}(x - r_i) - \delta^{(d)}(x - b_i) \right],
\end{equation}
and the average $\bar{\rho}$ is intended over all possible instances. Observe now that $C_d(0)$ is, in the large $N$ limit, the average optimal cost for the Euclidean bipartite matching problem; using this simple correspondence, Caracciolo et al. [11] derived the correct scaling of the optimal cost and, through a proper regularization procedure, the finite-size corrections to the average optimal cost for any dimension. In the following we will consider the complete correlation function $C_d(x)$ in any dimension and we will derive it using the same ansatz successfully adopted by Caracciolo et al. [11] to obtain the scaling of the average optimal cost.

We will distinguish two different cases. Poisson-Poisson Euclidean matching problem. In the Poisson-Poisson (PP) Euclidean matching problem both the points of $R$ and the points of $B$ are random points uniformly distributed within $\Omega_d$. In this case we obtain
\begin{equation}
\rho(x) \rho(y) = \frac{2}{N} \left[ \delta^{(d)}(x - y) - 1 \right],
\end{equation}
and therefore the correlation function is
\begin{equation}
C_d(x - y) = - \frac{2}{N} G_d(x - y).
\end{equation}
As anticipated, the average optimal cost is given by
\begin{equation}
E^2_N(d) := E^2_N[\rho; \{R, B\}] = C_d(0).
\end{equation}
Grid-Poisson Euclidean matching problem. In the grid-Poisson (GP) Euclidean matching problem we suppose that $N = L^d$ for some natural number $L \in \mathbb{N}$ and that one set of points, e.g., the set $R = \{r_i\}_{i=1, \ldots, N}$, is fixed on the vertices of an hypercubic lattice, in such a way that $R = \{ \frac{k}{L} \in \mathbb{Z} \cap [0, L^d] : k \in \mathbb{N}^d \}$, while the set $B = \{b_i\}_{i=1, \ldots, N} \subset \Omega_d$ is obtained as before considering randomly generated points in $\Omega_d$. We have
\begin{equation}
\rho(x) \rho(y) = \frac{1}{N} \delta^{(d)}(x - y) + \frac{N^2 - N}{N^2} + \frac{1}{N^2} \sum_{ij} \delta^{(d)}(x - r_i) \delta^{(d)}(y - r_j)
- \frac{1}{N} \sum_i \left[ \delta^{(d)}(x - r_i) + \delta^{(d)}(y - r_i) \right].
\end{equation}
In this case the correlation function is therefore
\[ c_{d}(x - y) = \frac{1}{N} G_{d}(x - y). \] (35)

Being the average optimal cost of the matching in the grid-
Poison case,
\[ E_{N}^{(2; \text{GP})}(d) := E_{N}^{(2)}[m; \{ R \text{ fixed}, B \}] = c_{d}(0), \] (36)
we expect that in this case it will be asymptotically one half of the
PP case.

We will consider also the correlation function for the
normalized transport field, i.e., the following quantity:
\[ c_{d}(x - y) = \sigma(x) \cdot \sigma(y), \] (37)
in which the correlation between the values normalized
transport field,
\[ \sigma(x) := \frac{m(x)}{\|m(x)\|} = \frac{\nabla_{x} \phi(x)}{\|\nabla_{x} \phi(x)\|}, \] (38)
in different positions is evaluated. Note that \( \sigma \) lives on the \( d \)-
dimensional unit sphere. To compute the correlation function Eq. (37) for the normalized field in the PP case, we assume a
Gaussian behavior for the joint probability distribution of two
values of the optimal transport field, and therefore we have

\[ c_{d}(x - y) = \int \int d^{d} m_{1} d^{d} m_{2} \frac{m_{1} \cdot m_{2}}{\|m_{1}\| \|m_{2}\|} e^{-\frac{1}{2} \langle m_{1} - m_{2} \rangle \Sigma^{-1}(x,y) \langle m_{1} \rangle} (2 \pi \sqrt{\det \Sigma})^{d/2}, \] (39)

where \( \Sigma(x,y) \) is the covariance matrix,
\[ \Sigma(x,y) := \begin{pmatrix} C_{d}(0) & C_{d}(x-y) \\ C_{d}(x-y) & C_{d}(0) \end{pmatrix}. \] (40)

For \( d \geq 2 \) (the case \( d = 1 \) was studied in Ref. [7]), introducing
\[ \begin{align*}
A & := \frac{C_{d}(0)}{\det \Sigma(x,y)}, \\
B & := \frac{C_{d}(x-y)}{\det \Sigma(x,y)},
\end{align*} \] (42a)
observe that \( \frac{B}{A} \to 0^{+} \) for \( N \to \infty \), being \( NC_{d}(x) \) finite for \( x \neq 0 \) and \( NC_{d}(0) \sim N^{1 - \frac{d}{2}} \) for \( d > 2 \), \( NC_{d}(0) \sim \ln N \) for \( d = 2 \).
We have therefore that, in the notation above,
\[ \det \Sigma = \frac{1}{A^{2} - B^{2}} \] (43)
and
\[ c_{d}(x,y) = \left( \sqrt{A^{2} - B^{2}} \right)^{d} 2 \pi^{\frac{d}{2}} \Gamma \left( \frac{d+1}{2} \right) \Gamma \left( \frac{d-1}{2} \right) \frac{1}{A^{d-1}} \int_{0}^{\pi} \sin^{d-2} \theta \cos \theta \int_{0}^{\infty} d m_{1} \int_{0}^{\infty} d m_{2} m_{1}^{d-1} m_{2}^{d-1} e^{-\frac{1}{2} (m_{1}^{2} + m_{2}^{2}) B m_{1} m_{2} \cos \theta} \] (44)

\[ = B \frac{2 \Gamma^{2} \left( \frac{d+1}{2} \right)}{A d \Gamma^{2} \left( \frac{d}{2} \right)} \left( 1 - \frac{B^{2}}{A^{2}} \right)^{\frac{d}{2}} \right) \frac{2}{d} \frac{2}{d+1} \frac{B^{2}}{A^{2}} \] (45)

\[ \left[ \sum_{n=0}^{\infty} \frac{(a)_{n} (b)_{n} z^{n}}{(c)_{n} n!} \right] = \frac{\Gamma(a+1)}{\Gamma(a-n+1)}. \] (46)

Observe that we can reproduce exactly the same calculation for the normalized field in the GP, obtaining
\[ c_{d}^{\text{GP}}(x - y) = \frac{2}{d} \left( \frac{\Gamma \left( \frac{d+1}{2} \right)}{\Gamma \left( \frac{d}{2} \right)} \right)^{2} \frac{c_{d}(x - y)}{E_{N}^{(2; \text{GP})}(d)}. \] (47)

Finally, for \( d \geq 2 \) we can compute also the so called wall-to-wall correlation function for the PP case:
\[ W_{d}(r) := \prod_{i=2}^{d} \left( \int_{0}^{r} d x_{i} \right) c_{d}(r, x_{2}, \ldots, x_{d}) = \frac{4}{dN} \left( \frac{\Gamma \left( \frac{d+1}{2} \right)}{\Gamma \left( \frac{d}{2} \right)} \right)^{2} \frac{G_{1}(r)}{E_{N}^{(2; \text{GP})}(d)}. \] (48)
Similarly, the computation for the GP case gives
\[
W_d^{GP}(r) := \prod_{i=2}^{d} \left( \int_{0}^{1} \frac{d x_{i}}{d} \right) \left( \sum_{n \neq 0} \frac{1}{4 \pi^{2} n^{2}} \right) e^{2 \pi i n r} = - \frac{2}{dN} \left( \frac{\Gamma \left( \frac{d+1}{2} \right)}{\Gamma \left( \frac{d}{2} \right)} \right)^{2} G_{1}(r) \frac{G_{N}(x_{d})}{E_{N}^{2}(x_{d})}.
\]

III. NUMERICAL RESULTS

In the following we consider explicitly the cases \( d = 1, d = 2, \) and \( d = 3 \) and we numerically verify the results presented above.

A. Case \( d = 1 \)

For \( d = 1 \) we have that
\[
G_{1}(r) = - \sum_{n \neq 0} \frac{1}{4 \pi^{2} n^{2}} e^{2 \pi i n r} = - \frac{1}{12} + \frac{|r|}{2} \left( 1 - |r| \right).
\]

It follows from Eq. (32) that
\[
C_{1}(x - y) = \frac{1}{N} \left[ \frac{1}{2} - |x - y| \left( 1 - |x - y| \right) \right],
\]

moreover, note that the average optimal cost is given by
\[
E_{N}^{2}(1) = C_{1}(0) = \frac{1}{6N}.
\]

In the GP case we obtain from Eq. (35)
\[
C_{1}^{GP}(x - y) = \frac{1}{N} \left[ \frac{1}{12} - |x - y| \left( 1 - |x - y| \right) \frac{2}{2} \right].
\]

The average total cost of the optimal matching is given by
\[
E_{N}^{2,GP}(1) = C_{1}^{GP}(0) = \frac{1}{12N}.
\]

The previous results for the \( d = 1 \) case are known to the literature [7,8,10,11], although the correlation function was derived using a different probabilistic approach. In Ref. [7] the correlation function for the normalized transport field is evaluated as
\[
c_{1}(x) = c_{1}^{GP}(x) = \frac{2}{\pi} \arctan \left( \frac{1 - 6x(1 - x)}{\sqrt{12x(1 - x) [1 - 3x(1 - x)]}} \right).
\]

B. Case \( d = 2 \)

For \( d = 2 \), denoting by \( x = (x_{1}, x_{2}) \) and by \( y = (y_{1}, y_{2}) \) two points on the unit flat torus \( T^{2} \), the Laplacian Green’s function can be written in terms of special functions as [14]
\[
G_{2}(x - y) = - \sum_{n \neq 0} \frac{1}{4 \pi^{2} |n|^{2}} e^{2 \pi i n (x-y)} = \frac{1}{2\pi} \ln \left| \frac{2 \pi^{3/4} \beta_{1}(\pi z)}{\Gamma \left( \frac{1}{4} \right)} \right|_{z = (x_{1} - y_{1}) + i(x_{2} - y_{2})} - \frac{(x_{2} - y_{2})}{2}.
\]

where we introduced the first Jacobi \( \beta \) function:
\[
\beta_{1}(z|x) := 2 e^{\pi z} \sum_{n=0}^{\infty} (-1)^{n} e^{i \pi n (x+1)} \sin[(2n+1)z].
\]

From Eq. (32) we have simply
\[
C_{2}(x) = - \frac{1}{N} \left[ \frac{1}{2\pi} \ln \left| \frac{2 \pi^{3/4} \beta_{1}(\pi z)}{\Gamma \left( \frac{1}{4} \right)} \right|_{z = x_{1} + i x_{2}} - \frac{x_{2}^{2}}{2} \right].
\]

In the GP case, we have as usual
\[
C_{2}^{GP}(x - y) = \frac{1}{4} C_{2}(x - y).
\]

Observe that the previous expressions contains no free parameters and therefore a direct comparison with numerical data is possible. We present our numerical results both for the GP case and the PP case in Fig. 2(a). The average optimal cost for the PP EBMP is given by \( E_{N}^{2}(2) = C_{2}(0) \); however, \( G_{2}(x) \) is divergent for \( x = 0 \). Analyzing the scaling of the total cost and performing a proper regularization of the previous quantity, Caracciolo et al. [11] obtained
\[
C_{2}(0) = \frac{1}{N} \left[ \ln N + \beta_{PP} + o \left( \frac{1}{N} \right) \right], \quad \beta_{PP} = 0.1332(5).
\]

A numerical fit of the optimal costs for \( d = 2 \) for the GP EBMP gives
\[
C_{2}^{GP}(0) = \frac{1}{2N} \left[ \ln N + \beta_{GP} + o \left( \frac{1}{N} \right) \right], \quad \beta_{GP} = 0.3758(5).
\]

The correlation function, Eq. (37), for the normalized matching field in the PP case has the expression, Eq. (44),
\[
c_{2}(x - y) = \frac{\pi}{4} C_{2}(x - y).
\]

Observe that the only free parameter in this quantity is \( C_{2}(0) \); inserting the value obtained by Caracciolo et al. [11], Eq. (59), we obtain the theoretical prediction in Fig. 2(b), where we also present some numerical results for \( c_{2}(x,y) \) that show the agreement with the theoretical curve.

Using the value Eq. (60) in Eq. (46) for \( d = 2 \) we also obtain the theoretical curve for the grid-Poisson problem depicted in Fig. 2(b), where, once again, an excellent agreement is found with numerical data.

Finally, let us compute the wall-to-wall correlation function for the PP Euclidean matching problem on the flat torus. The theoretical prediction is given by Eq. (47),
\[
W_{2}(r) = - \frac{\pi}{2NC_{2}(0)} G_{1}(r).
\]

In the GP case, instead, we have
\[
W_{2}^{GP}(r) = - \frac{\pi}{4NC_{2}^{GP}(0)} G_{1}(r).
\]

Numerical results both for the PP case and for the GP case are presented in Figs. 2(c) and 2(d). Once again observe that the values of the average optimal cost in the corresponding
cases, Eqs. (59) and (60), fix completely the expression of the wall-to-wall correlation function.

C. The case \( p \neq 2 \)

Up to now, we analyzed the correlation functions for \( p = 2 \) and \( d = 2 \) obtained from the linearized Eq. (18). We present some numerical results for \( p = 1 \) and \( p = 3 \) and \( N = 3600 \) both in the GP case and in the PP in the two-dimensional case. In particular, we analyzed the wall-to-wall correlation function for different values of \( p \) and we obtained numerical evidences of a functional form for it of the type

\[
W_2(r; p, N) = w(p, N) \left[ \frac{1}{6} - r(1 - r) \right].
\]  

(64)

both in the PP and in the GP case. Inspired by the obtained expression for the \( p = 2 \) case, Eq. (62), we assumed for the global factor \( w(p) \) the following dependence on the size \( N \) of the considered system,

\[
w(p, N) = \frac{\alpha(p) \pi^2}{2 \ln N + 4\pi\beta}.
\]

(65)

where \( \alpha(p) \) depends only on the weight exponent \( p \) and \( \beta \equiv \beta_\text{GP} \) if we are considering a GP matching and \( \beta \equiv \beta_\text{PP} \) if we are dealing with a PP matching. We expected that \( \alpha(2) = 1 \). We performed a numerical fit using the previous expression also for \( p = 2 \), obtaining the results presented in Table I. Numerical data are in excellent agreement with Eq. (64), suggesting
TABLE I. Fit results for the wall-to-wall correlation function of the Euclidean matching problem for \( d = 2 \) and \( p = 1, 2, 3 \) with reference to the notation of the fitting curve, Eq. (64). Observe that we expected \( \alpha(2) = 1 \).

| \( p \) | \( \alpha(p) \) | \( \beta \) fixed |
| --- | --- | --- |
| \( p = 1 \) | 0.860(5) | --- |
| \( p = 2 \) | 0.996(5) | \( \beta \equiv \beta_{pp} = 0.1332(5) \) |
| \( p = 3 \) | 0.96(1) | --- |

The general expressions, Eqs. (32) and (35), for the correlation functions presented above can be applied in any dimensionality in the \( N \to \infty \) limit. We show here some numerical results for \( d = 3 \) for the grid-Poisson case, taking into account, however, that the precision of the data is lower, since the computation has complexity \( O(N^3) \) (i.e., the computational complexity of the Hungarian algorithm in the Edmonds and Karp [3] version), where the number of points \( N \) has to scale exponentially with the system dimensionality in order to maintain the same accuracy.

For \( d = 3 \), Eqs. (22) and (32) give

\[
C_3(x - y) = \frac{1}{2\pi^2 N} \sum_{n \in \mathbb{Z}^1(0)} \frac{1}{||n||^2} e^{2\pi i n \cdot (x - y)}. \tag{66}
\]

Clearly the previous function cannot be represented in a plot. From the correlation function \( C_3(x) \), the wall-to-wall correlation function can be obtained as before in the form

\[
W_3(r) = -\frac{16}{3\pi N C_3(0)} G_1(r). \tag{67}
\]

As in the previous cases, \( C_3(0) \) can be evaluated from the cost fit [11] and it is equal to \( C_3(0) = 0.66251(2) N^{-\frac{3}{2}} - 0.45157\ldots \)

(note that an exact formula for the coefficient of the \( \frac{1}{N} \) correction to \( C_3(0) \) is provided in Ref. [11] in terms of an Epstein function).

Following the same procedure of the PP case, we can compute the wall-to-wall correlation function on the unit hypercube in \( d = 3 \) for the GP matching problem. Reproducing the computations of the \( d = 2 \) case we have

\[
W_3^{GP}(r) = -\frac{8}{3\pi N C_3^{GP}(0)} G_1(r). \tag{68}
\]

We evaluated \( C_3^{GP}(0) \) from the cost scaling, obtaining

\[
C_3^{GP}(0) = 0.4893(4) N^{-\frac{3}{2}} - 0.23(5) \frac{1}{N}. \tag{69}
\]

D. Case \( d = 3 \)

The prediction obtained and the numerical data are presented in Fig. 3.

IV. CONCLUSIONS

In the present work we adopted the scaling ansatz proposed by Caracciolo et al. [11] to compute the correlation function for the optimal matching ray and for the normalized optimal matching ray in the Euclidean bipartite matching problem on the \( d \)-dimensional flat hypertorus with quadratic cost. We showed also that the correlation function is strictly related to the Green’s function of the Laplacian operator on the flat hypertorus itself in the large \( N \) limit. Given the value the average optimal cost at fixed size \( N \), the obtained expressions have no free parameters and were directly compared with the results of numerical simulations, showing an excellent agreement. For \( d = 2 \) and \( d = 3 \) we computed also the wall-to-wall correlation function: for \( d = 2 \) in particular we give numerical evidences that, for \( p \neq 2 \), the wall-to-wall correlation function has the same form obtained for \( p = 2 \) up to a global multiplicative constant.

All previous results suggest that the EBMP with quadratic cost, in the large \( N \) limit, appears as a Gaussian free theory on the \( d \)-dimensional flat hypertorus, in such a way that the correlation function of the matching ray is related directly to the free propagator of the theory itself. In subsequent publications we will investigate this crucial aspect of the problem and its implications on the universal behavior for different values of the exponent \( p \) in the cost functional.

ACKNOWLEDGMENTS

We thank Luigi Ambrosio, from Scuola Normale Superiore di Pisa, Carlo Lucibello, from Politecnico di Torino, and Giorgio Parisi, from University of Rome “La Sapienza,” for useful discussions. G.S. acknowledges partial financial support from the John Templeton Foundation.
[1] H. W. Kuhn, Naval Res. Logist. Quart. 2, 83 (1955).
[2] J. Munkres, J. Soc. Indust. Appl. Math. 5, 32 (1957).
[3] J. Edmonds and R. M. Karp, J. ACM 19, 248 (1972).
[4] M. Mézard and G. Parisi, J. Phys. Lett. 46, 771 (1985).
[5] D. J. Aldous, Random Struct. Alg. 18, 381 (2001).
[6] M. Mézard and G. Parisi, J. Phys. 49, 2019 (1988).
[7] E. Boniolo, S. Caracciolo, and A. Sportiello, J. Stat. Mech. (2014) P11023.
[8] S. Caracciolo and G. Sicuro, Phys. Rev. E 90, 042112 (2014).
[9] M. Ajtai, J. Komlós, and G. Tusnády, Combinatorica 4, 259 (1984).
[10] M. Talagrand, Ann. Probabil. 2, 846 (1992).
[11] S. Caracciolo, C. Lucibello, G. Parisi, and G. Sicuro, Phys. Rev. E 90, 012118 (2014).
[12] L. C. Evans, in Current Developments in Mathematics, 1997 (International Press, Boston, MA, 1999), pp. 65–126.
[13] C. Villani, Optimal Transport: Old and New, Grundlehren der mathematischen Wissenschaften (Springer, Berlin, 2008).
[14] C.-S. Lin and C.-L. Wang, Ann. Math. 172, 911 (2010).