On the one dimensional Euclidean matching problem: exact solutions, correlation
functions and universality

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

Gabriele Sicuro
Dipartimento di Fisica “E. Fermi”, University of Pisa and INFN,
Largo Bruno Pontecorvo, 3, I-56127 Pisa, Italy
(Dated: August 16, 2014)

We discuss the equivalence relation between the Euclidean bipartite matching problem on the line
and on the circumference and the Brownian bridge process on the same domains. The equivalence
allows us to compute the correlation function and the optimal cost of the original combinatoric
problem in the thermodynamic limit; moreover, we solve also the minimax problem on the line and
on the circumference. The properties of the average cost and correlation functions are discussed.

I. INTRODUCTION

Let us consider a bipartite complete graph with vertex
set \( V \) and edge set \( E \), \( \mathbb{R}_{N,N} = \text{Graph}(V,E) \), whose vertex
set can be partitioned in two disjoint subsets of the same
cardinality, \( V = \mathcal{R} \cup \mathcal{B} \), \( |\mathcal{R}| = |\mathcal{B}| = N \in \mathbb{N} \), \( \mathcal{R} := \{r_1, \ldots, r_N\}, \mathcal{B} := \{b_1, \ldots, b_N\} \). Let us also introduce a
weight function \( w: \mathcal{E} \rightarrow \mathbb{R} \), \((r_ib_j) \in \mathcal{E} \rightarrow w_{ij}\). In the
weighted bipartite matching problem, we are interested in the permutation \( \pi \) of \( N \) elements such that a certain
cost function

\[
E[\pi] \equiv \sum_{i=1}^{N} w_{ij} \quad \text{f: } \mathbb{R}^N \rightarrow \mathbb{R}^+, \tag{1}
\]

is minimized. The most common form adopted in the
literature for the function \( f \) is simply \( f(x_1, \ldots, x_N) : = \sum_{i=1}^{N} |x_i| \): in this case, the assignment problem can be solved in polynomial time using the Khun–Munkres algorithm [1, 2] that, in the Edmonds and Karp’s version [3], has \( O(N^3) \) running time.

As a variation of the problem, sometimes random
weights \( \{w_{ij}\} \) are considered: in this case the average
properties of the solution are of great interest. In the
hypothesis of independently and identically distributed weights, the problem was studied using arguments
borrowed both from probability theory [4] and from the
theory of disordered systems [5]. Finally, if the vertices \( V \)
of the graph are associated to points randomly generated
in the hypercube \( \Omega = [0,1]^d \) in \( d \) dimensions and \( w_{ij} \) is a function of the Euclidean distance between the \( r_i \) vertex and the \( b_j \) vertex, the problem is called Euclidean bipartite matching problem [6-8].

In the present paper we will consider the so-called
grid-Poisson matching problem in one dimension both
with open boundary conditions (OBC) and with periodic
boundary conditions (PBC). The set of vertices \( \mathcal{R} \) is asso-
ciated to a set of fixed points on the interval \( \Omega : = [0,1] \)
and in particular \( r_i \rightarrow x_i \equiv \frac{2i-1}{2N}, i = 1, \ldots, N \), whilst the set of vertices \( \mathcal{B} \) is associated to a set of \( N \) points, \( \{y_i\}_{i=1,\ldots,N} \), randomly generated in the interval, such that \( b_i \rightarrow y_i \). We will suppose the \( \mathcal{B} \)-vertices indexed in such a way that \( i \leq j \Rightarrow y_i \leq y_j \). Finally, we will consider the following cost functional

\[
E_p[\pi] = \sum_{i=1}^{N} w \left( |x_i - y_{\pi(i)}| \right)^p, \quad p \in (1, +\infty), \tag{2}
\]

in which the function \( w: [0,1] \rightarrow [0,1] \) is defined as below:

\[
w(x) = \begin{cases} x & \text{for OBC;} \\
\frac{x}{\theta \left( \frac{1}{2} - x \right) + (1 - x) \theta \left( x - \frac{1}{2} \right)} & \text{for PBC,} \tag{3}
\end{cases}
\]

where \( \theta(x) \) is the Heaviside function. In the following,

\[
e_p \equiv \sqrt{N} \min_{\pi} E_p[\pi], \tag{4}
\]

and

\[
e_p \equiv \sqrt{N} \min_{\pi} \left( E_p[\pi] \right)^p. \tag{5}
\]

We will show that, for the cost functional (2), a solution
of the problem can be obtained in the continuum limit,
\( N \rightarrow \infty \), not only for \( p = 2 \) (as already shown in Ref. [7])
but also in the \( p \rightarrow \infty \) limit using well known properties of
the Brownian bridge process. In fact, in the limit
\( p \rightarrow \infty \),

\[
e_p \rightarrow \sqrt{N} \max_{\pi} \min_{i} w \left( |x_i - y_{\pi(i)}| \right), \tag{6}
\]

i.e., the problem reduces to the minimax grid-Poisson
matching problem in one dimension. The problem was
studied by Leighton and Shor [9] for \( d = 2 \) and Shor and
Yukich [10] for \( d \geq 3 \). The minimax problem is related
to a lot of different computational problems and the evaluation
of the scaling of its cost gives directly useful information
on the computational cost of other algorithms
(for a discussion of the related problems in \( d = 2 \) see for example Ref. [9]).

---

*Electronic address: sergio.caracciolo@mi.infn.it

Electronic address: gabriele.sicuro@for.unipi.it
II. OPTIMAL COST AND CORRELATION FUNCTION ON THE INTERVAL

The solution of the grid–Poisson matching problem in one dimension for OBC is easily found by simple arguments [7] for $p > 1$ and cost functional (2): in this case, in fact, the optimal matching is always ordered, i.e. $\pi(i) = i$, in the hypothesis above that $i < j \Rightarrow y_i \leq y_j$. Remarkably, the solution is very different if $p \in (0, 1)$, as discussed in [7, 11]: indeed, in that case the solution is expressed in terms of non intersecting arcs joining the matched points on the segment. Note that, due to the monotony of the function $x \mapsto x^p$, the optimal solution for the cost $E_p$ is also optimal for the cost $(E_p)^p$. The probability density distribution for the position of the $i$-th $\mathcal{B}$-point is:

$$\Pr (y_i, \in d y) = \left( \frac{N}{i} \right) y^i (1 - y)^{N - i} d y, \quad (7)$$

where we used the short-hand notation $x \in d z \Leftrightarrow x \in [z, z + d z]$. In the $N \to \infty$ limit, a non trivial result is obtained introducing the variable $m(y)$

$$m(y) := \sqrt{N} M(y), \quad M(y) := \frac{i}{N} - y \quad (8)$$

expressing the rescaled (signed) distance between a $\mathcal{B}$-point in $[y, y + d y]$ and its corresponding $\mathcal{R}$-point in the optimal matching. We finally obtain a distribution for the variable $m(y)$ depending on the position on the interval $y \in [0, 1]$:

$$\Pr (m(y) \in d m) = \frac{e^{-\frac{\pi m^2}{2}}}{\sqrt{2\pi y(1 - y)}} d m. \quad (9)$$

The distribution (9) is the one of a Brownian bridge on the interval $[0, 1]$, a continuous time stochastic process defined as

$$\mathcal{B}(t) := \mathcal{W}(t) - t \mathcal{W}(1), \quad t \in [0, 1], \quad (10)$$

where $\mathcal{W}(t)$ is a Wiener process. The joint distribution of the process can be derived similarly (see Appendix). In particular, the covariance matrix for the 2-points joint distribution has the form, for $s, t \in [0, 1]$ (see eq. (A6)),

$$\Sigma_2 = \begin{pmatrix} 2\phi(s) & \phi(s) + \phi(t) - \phi(t - s) \\ \phi(s) + \phi(t) - \phi(t - s) & 2\phi(t) \end{pmatrix}, \quad (11)$$

where we introduced the function

$$\phi(x) := \frac{1 - |x|}{2}. \quad (12)$$

Averaging over the positions $s, t$ and fixing the distance $\tau := |t - s|$, we have

$$\Sigma_2(\tau) = \begin{pmatrix} c & c - \phi(\tau) \\ c - \phi(\tau) & c \end{pmatrix}, \quad c = \frac{1}{6}. \quad (13)$$

![Figure 1: Plot of $m(y_i) = \sqrt{N} (\frac{1}{2} - y_i)$ for a certain realisation of the problem with $N = 4096$.](image)

The Euclidean matching problem on the interval $[0, 1]$ with open boundary conditions and cost functional (2) is therefore related to the Brownian bridge in the $N \to \infty$ limit. By using this correspondence, the correlation function $\forall p > 1$ is computed as

$$\langle m(t) m(t + \tau) \rangle = \frac{1}{6} - \phi(\tau). \quad (14)$$

where the average $\langle \cdot \rangle$ is intended on the position $t$, whilst we denoted by $\langle \cdot \rangle$ the average over different realisations of the problem. This theoretical prediction was confirmed numerically. Introducing the normalised variable

$$\sigma(t) = \frac{m(t)}{|m(t)|} = \text{sign}(m(t)), \quad (15)$$

Bonito et al. [7] computed also the correlation function for this quantity, finding

$$\langle \sigma(s) \sigma(t) \rangle = \frac{2}{\pi} \arctan \sqrt{\frac{\min(s, t)(1 - \max(s, t))}{|t - s|}} \Rightarrow \int_0^{1-t} \langle \sigma(s) \sigma(s + t) \rangle d s = \frac{1 - \sqrt{t}}{1 + \sqrt{t}}. \quad (16)$$

Both formulas were confirmed numerically. Note that all the results above holds $\forall p > 1$ in the case of open boundary conditions.

Let us now compute the average cost of the matching. From eq. (9) we obtained that

$$\langle e_p \rangle \xrightarrow{N \to \infty} \int_0^1 \langle |B(t)|^p \rangle d t = \frac{1}{2\pi (p + 1)} \Gamma \left( \frac{p}{2} + 1 \right). \quad (17)$$
The distribution of the supremum of the absolute value of the Brownian bridge is the well known Kolmogorov distribution [13]

$$\Pr \left( \sup_{t \in [0,1]} |B(t)| < u \right) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2u^2}$$  \hspace{1cm} (20)

and therefore

$$\langle \epsilon_p \rangle \overset{N \to \infty}{\longrightarrow} \frac{\sqrt{\pi}}{2} \ln 2.$$  \hspace{1cm} (21)

In figure 3 we plotted $\rho_p(u) := \frac{d}{du} [\Pr (\epsilon_p \leq u)]$ for different values of $p$: observe that $\rho_p$ approaches the Kolmogorov distribution in the large $p$ limit.

Finally, observe also that the variance of $\epsilon_p$

$$\langle (\epsilon_p - \langle \epsilon_p \rangle)^2 \rangle = \frac{(1 + p)\Gamma(p + 2) - (2p + 1)\Gamma \left( \frac{p}{2} + 1 \right)}{2p(p + 1)^2(2p + 1)}$$  \hspace{1cm} (22)

increases with $p$ and that

$$\frac{\langle (\epsilon_p - \langle \epsilon_p \rangle)^2 \rangle}{\langle \epsilon_p \rangle^2} = \frac{(p + 1)\Gamma(p + 2)}{(2p + 1)\Gamma \left( \frac{p}{2} + 1 \right)} - 1 \to +\infty.$$  \hspace{1cm} (23)

From a numerical point of view, this means that a computation of $\langle \epsilon_p \rangle$ requires a larger amount of iterations as $p$ increases and fluctuations around the mean value become extremely large in the $p \to \infty$ limit. On the other hand, fluctuations of the optimal cost $\epsilon_p$ around its mean value $\langle \epsilon_p \rangle$ for $p \to \infty$ remain finite

$$\langle (\epsilon_p - \langle \epsilon_p \rangle)^2 \rangle \overset{p \to \infty}{\longrightarrow} \frac{\pi^2}{12} - \frac{\pi}{2} \ln^2 2.$$  \hspace{1cm} (24)

This fact allows to perform a precise computation of $\langle \epsilon_p \rangle$ for large $p$.

III. OPTIMAL COST AND CORRELATION FUNCTION ON THE CIRCUMFERENCE

Let us now consider the case of periodic boundary conditions. As discussed in Ref. [7], the solution for both the cost $E_p$ and the cost $(E_p)^p$, with $p \in (1, +\infty)$, is again ordered; however, in this case the mapping is $i \mapsto \pi(i) = i + \lambda \mod N$ for a certain $\lambda \in \{0, 1, \ldots, N-1 \}$. In the continuum limit, the solution is a generalised Brownian bridge, $m_p(t) = B(t) + \lambda_p$, $t \in [0,1]$, for a certain constant $\lambda_p \in \mathbb{R}$ depending on $p$. The constant $\lambda_p$ can be found by optimality condition on the cost functional (2):

$$\frac{\partial}{\partial \lambda_p} \left( \int_0^1 |B(t) + \lambda_p|^p \, dt \right)^{\frac{1}{p}} = 0.$$  \hspace{1cm} (25)

The previous equation can be solved only for $p = 2$ and $p \to +\infty$. For $p = 2$, $\lambda_2 = -\int_0^1 B(\tau) \, d\tau$; therefore,
\[ m_2(t) = B(t) - \int_0^1 B(t) \, dt \] and
\[ \langle m_2(t) m_2(t + \tau) \rangle = \frac{1}{12} - \phi(\tau). \] (26)

For \( p \to \infty \), \( \left( \int_0^1 |B(t) + \lambda \rho(t) \, dt \right)^\frac{1}{2} \xrightarrow{p \to \infty} \sup_{t \in [0,1]} |B(t) + \lambda \sup_{t \in [0,1]} B(t) + \inf_{t \in [0,1]} B(t) + \lambda | \) and therefore the optimality condition becomes
\[ \frac{\partial}{\partial \lambda} \left( \sup_{t \in [0,1]} |B(t) + \lambda| \right) = 0 \]
\[ \Rightarrow \lambda_{\infty} = -\frac{\sup_{t \in [0,1]} B(t) + \inf_{t \in [0,1]} B(t)}{2}. \] (27)

Indeed, we have that
\[ \sup_{t \in [0,1]} |B(t) + \lambda| = \sup_{t \in [0,1]} B(t) + \lambda \theta(\sup_{t \in [0,1]} B(t) + \lambda) + \inf_{t \in [0,1]} B(t) + \lambda \theta(\inf_{t \in [0,1]} B(t) + \lambda) \] (28)

from which the eq. (27) is derived. We have therefore, for \( t \in [0,1] \),
\[ m_{\infty}(t) = B(t) - \frac{\sup_{s \in [0,1]} B(s) + \inf_{s \in [0,1]} B(s)}{2}. \] (29)

The correlation function can be directly found using the known joint distributions for the Brownian bridge and its sup, eq. (B2), and for the sup and inf of a Brownian bridge, eq. (B1). After some calculations we obtain
\[ \langle m_{\infty}(t) m_{\infty}(t + \tau) \rangle = \frac{2 - \zeta(2)}{4} - \phi(\tau), \] (30)

where \( \zeta(2) := \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \). The value \( \langle m_{\infty}^2(t) \rangle = \frac{2 - \zeta(2)}{4} = 0.0887665 \ldots \) is very close to the value obtained for \( p = 2 \), \( \langle m_{\infty}^2(t) \rangle = \frac{1}{12} = 0.083 \). In figure 5 we plot the values of \( \langle m_{\infty}^2(t) \rangle \) as function of \( p \). Note, finally, that due to the fact that we have imposed PBC, \( \otimes \equiv \heartsuit \) holds in all the previous formulas.

Let us now introduce the normalised transport field
\[ \sigma_p(t) := \frac{m_p(t)}{|m_p(t)|} = \text{sign}(m_p(t)). \] (31)

The correlation function \( \langle \sigma_p(s) \sigma_p(t) \rangle \) can be computed from the covariance matrix observing that the process is still Gaussian. The correlation function is found in the
form
\[
\langle \sigma_p(s) \sigma_p(t) \rangle = \frac{2}{\pi} \arctan \frac{c_p - \phi(t - s)}{\sqrt{\phi(t - s) (2c_p - \phi(t - s))}}
\]  
(32)
where \(c_2 = \frac{1}{12}\) for \(p = 2\) and \(c_\infty = 2 - \zeta(2)\).

The optimal cost in the \(p \to \infty\) limit can be evaluated as the average spread of the Brownian bridge. Denoting
\[
\xi := \sup_{s \in [0,1]} B(s) - \inf_{s \in [0,1]} B(s),
\]  
(33)
the distribution of the spread \(\xi\) is given by \([13]\)
\[
\Pr(\xi < u) = \theta_3(e^{-2u^2}) + u \frac{d}{du} \theta_3(e^{-2u^2}),
\]  
(34)
where
\[
\theta_3(x) \equiv \theta_3(0, x), \quad \theta_3(z, q) := 1 + 2 \sum_{n=1}^{\infty} q^{n^2} \cos(2nz)
\]  
(35)
is the third Jacobi theta function. From eq. (34) the distribution of the optimal cost in the \(p \to \infty\) limit is easily derived. Moreover,
\[
\langle \epsilon_p \rangle \xrightarrow{N \to \infty} \langle \epsilon_\infty \rangle = \frac{\langle \xi \rangle}{2} = \frac{1}{2} \left( \sup_{s \in [0,1]} B(s) - \inf_{s \in [0,1]} B(s) \right)
\]  
(36)
with corresponding variance
\[
\langle (\epsilon_p - \langle \epsilon_p \rangle)^2 \rangle \xrightarrow{p \to \infty} \frac{\pi^2}{24} - \frac{3\pi}{24}.
\]  
(37)

**IV. GENERAL SOLUTION: A CONTINUUM APPROACH**

In the present section we will justify and generalise the previous results looking at a continuum version of the problem, the so called *Monge–Kantorovič problem*.

Let us consider the interval \(\Omega := [0, L] \subset \mathbb{R}\), \(L \in \mathbb{R}^+\) and let us suppose also that two different measures are given on \(\Omega\), i.e., the uniform (Lebesgue) measure
\[
dm(x) := \frac{1}{L} \, dx,
\]  
(38)
and a non uniform measure \(dn(x)\) with measure density \(\nu(x)\),
\[
dn(x) := \nu(x) \, dx = \frac{dx}{L} + \frac{1}{2} \sum_{k=1}^{\infty} \left( \nu_1(k) \cos \frac{2\pi k x}{L} + \nu_2(k) \sin \frac{2\pi k x}{L} \right).
\]  
(39)
We ask for the optimal map \(\mu(x): \Omega \rightarrow \Omega\) such that the following transport condition is satisfied
\[
\frac{1}{L} \int_{\Omega} dx = \int_{A} dn(x) \quad \forall A \subset \Omega \text{ measurable}. \quad (40)
\]  
It can be proved \([14]\) that for \(p > 1\) eq. (40) can be rewritten as a change-of-variable formula:
\[
L \, dn(x) = d \mu(x).
\]  
(42)
We will restrict therefore to the case \(p > 1\). Imposing PBC, that is \(\mu_p(0) = \mu_p(L) = L\), the solution of (42) determines the optimal map up to a constant \(\mu_p(0)\) as
\[
\mu_p(x) = x + \mu_p(0) + \Phi(x).
\]  
(43)
In the previous equation we have introduced \(\Phi(x)\)
\[
\Phi(x) := \sum_{k=1}^{\infty} \nu_1(k) \frac{L^2 \sin \left( \frac{\pi k x}{L} \right)}{\pi^2 k^2} \cos \left( \frac{k \pi x}{L} \right) + \sum_{k=1}^{\infty} \nu_2(k) \frac{L^2 \sin^2 \left( \frac{k \pi x}{L} \right)}{\pi^2 k^2}. \quad (44)
\]
Note that \(\Phi(0) = \Phi(L) = 0\). The value of \(\mu_p(0)\) must be determined requiring that the functional (41) is minimum: we have that
\[
\frac{\partial}{\partial \mu_p(0)} \int_0^L |\mu_p(0) + \Phi(x)|^p \, dn(x) = \int_0^L \nu_1(0) + \Phi(x) + 2 \mu_p(0) + \Phi(x)|^{p-1} \, dn(x) = 0.
\]  
(45)
If instead OBC are considered, then \(\mu_p(0) = 0\) and the solution is obtained explicitly \(\forall p > 1\).
Let us now suppose that \( L \equiv N \in \mathbb{N} \) and that the measure \( n(dx) \) is obtained as a limit measure of a random atomic measure of the form

\[
d n_N(x) := \frac{d x}{N} \sum_{i=1}^{N} \delta(x - \eta_i),
\]

where \( \{\eta_i\}_{i=1, \ldots, N} \) is a set of \( N \) points uniformly randomly distributed in \( \Omega \). The previous measure can be written as

\[
n_N(x) = \frac{1}{\sqrt{2\pi N}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left( \frac{x - \eta_i}{\sqrt{N}} \right)^n.
\]

Observe now that \( Z_k(x) \) is a sum of independent identically distributed random variables. From the central limit theorem, we have that \( Z_k(x) \) is normally distributed as

\[
Z_k \sim \mathcal{N}(0, 1) \quad \forall k \in \mathbb{N}.
\]

Remarkably the previous distribution does not depend on \( x \). Moreover, the \( Z_k \) and \( Z_i \) are independent random variables for \( k \neq l \) being Gaussian distributed and \( \langle Z_i Z_k \rangle = 0 \), where the average \( \langle \cdot \rangle \) is intended over the possible values \( \{\eta_i\} \). In eq. (47) the Karhunen–Loève expansion for the Brownian bridge [15] on the interval \([0, N]\) appears:

\[
B_N(x) := \sum_{k=1}^{\infty} \sqrt{\frac{2}{\pi N}} \frac{Z_k}{\pi k} \sin \left( \frac{\pi k x}{N} \right), \quad Z_k \sim \mathcal{N}(0, 1) \quad \forall k \in \mathbb{N}.
\]

It follows that \( n_N(x) \) can be written for large \( N \), up to irrelevant additive constants, as

\[
n_N(x) \approx \frac{1}{N} B_N(x) + \frac{x}{N}
\]

and therefore we cannot associate a density measure to it, due to the fact that \( n_N(x) \) is not differentiable.

However the solution of the matching problem in the continuum can be still obtained directly from eq. (39); considering PBC, then

\[
\mu_p(x) = \mu_p(0) + x + B_N(x), \quad x \in [0, N].
\]

Denoting by

\[
M_p(x) := \mu_p(x) - x,
\]

it follows that \( \forall p > 1 \)

\[
\langle M_p(x)M_p(y) \rangle = c_p(N) - N \phi \left( \frac{x - y}{N} \right).
\]

where \( c_p(N) \) is a constant depending on \( N \) and \( p \). Adopting the notation

\[
F(N) \sim G(N) \iff 0 < \lim_{N \to \infty} \frac{F(N)}{G(N)} < +\infty,
\]

for two positive real functions \( F \) and \( G \) depending on \( N \), note that

\[
c_p(N) \sim c_p N
\]

for some positive constant \( c_p \) depending on \( p \). Indeed, in the discrete case, from the fact that the solution must be ordered, in the large \( N \) limit it can be easily seen that

\[
\min_{\mu} E_p[\mu] \sim N^{\frac{3}{2}},
\]

where \( E_p[\mu] \) is the cost functional (41) in which the measure (46) is adopted. Therefore, \( \mu_2(0) = M_p(0) \sim M_p(x) \sim \sqrt{N} \). Moreover, note also that \( \langle M^2(t) \rangle \leq N^2 \Rightarrow c_p(0) = 0 \). For \( p = 2 \), eq. (45) becomes

\[
\mu_2(0) = -\frac{1}{N} \int_0^N B_N(x) \, d B_N(x) - \frac{1}{N} \int_0^N B_N(x) \, d x = -\frac{1}{N} \int_0^N B_N(x) \, d x
\]

where the first integral is intended in the Stratonovitch sense. The result is in agreement with the one presented in Section III. If we consider the transport cost functional

\[
E_p[\tilde{\mu}] := \frac{1}{p} \int_0^N \frac{1}{p} (x - \tilde{\mu}(x))^p \, d n(x)
\]

the matching problem has the same solution obtained for the cost (41) for all finite values of \( p \). Let \( \mu_p(x) = \tilde{\mu}_p(x) \Rightarrow \mu_p(0) \equiv \tilde{\mu}_p(0) \), due to the fact that the function \( f(x) = \frac{x}{\sqrt{N}} \) is monotone. However, for the functional cost (50), in the \( p \to \infty \) limit, we can reproduce the computations of Section III, obtaining

\[
\lim_{p \to \infty} \mu_p(0) = \frac{\sup_{s \in [0, N]} B_N(s) + \inf_{s \in [0, N]} B_N(s)}{2}.
\]

If OBC are considered, then \( \mu_p(0) = 0 \) and we have simply, \( \forall p > 1 \),

\[
M_p(x) \equiv M(x) = B_N(x), \quad x \in [0, N];
\]

It can be easily seen that

\[
\frac{1}{N} \int_0^N \langle M(x)M(x + r) \rangle \, d x = c(N) + N \phi \left( \frac{r}{N} \right)
\]

where \( c(N) = \frac{N}{6} \).
V. CONCLUSIONS

In the present work we solved the Euclidean bipartite matching problem on the interval [0, 1], using the cost functional (2) in the continuum limit, \( N \to \infty \), for \( p \to \infty \), both with open boundary conditions and with periodic boundary conditions. The solution is based on the exact correspondence between the optimal map and a Brownian bridge process on the same interval. Moreover, we computed the correlation function for the optimal map and we observed that in all considered cases it has the form

\[
\langle m_p(t) m_p(t + \tau) \rangle = c_p - \phi(\tau), \quad \tau \in [0, 1],
\]

for some constant \( c_p \) depending on the adopted boundary conditions and on the value \( p \) for the optimal cost. Note that if we consider the problem of a matching of \( N \) random \( N \)-points to \( N \) lattice \( \mathcal{R} \)-points on the interval \([0, N]\), for \( s, t \in [0, N] \) the correlation function assumes the form

\[
G(\tau) := \langle M_p(t)M_p(t + \tau) \rangle = c_pN - \frac{\left\lfloor \frac{\tau}{2} \right\rfloor}{N} \left( 1 - \left\lfloor \frac{\tau}{2} \right\rfloor \right),
\]

where \( M_p(t) \) is the signed distance between \( t \in [0, N] \) and its corresponding inverse image under the action of the optimal map. It follows that for \( N \to \infty \) the correlation function \( G(\tau) \) has a divergent part, \( G(0) = c_pN \), depending through \( c_p \) on the specific details of the problem (e.g., the boundary conditions adopted or the value of \( p \)), a universal finite part \(-\frac{\left\lfloor \frac{\tau}{2} \right\rfloor}{N}\) and a (universal) finite size correction \( \frac{\tau^2}{N}\). We obtained also numerical evidences of the validity of eq. (64) for different values of \( p \) both with OBC and with PBC: an exact derivation of \( c_p \) for OBC \( \forall p > 1 \) and for \( c_2 \) and \( c_\infty \) for PBC was presented. This fact suggests that all Euclidean matching problems in one dimension with strictly convex cost functionals belong to the same universality class and that the specific details of the model determine only the value of the constant \( c_p \) in the divergent contribution \( G(0) \). Similarly, on the interval \([0, N]\) eq. (32) becomes

\[
\langle \sigma_p(t) \sigma_p(t + \tau) \rangle = \frac{2}{\pi} \arctan \frac{2c_pN - \frac{\tau}{2} \left( 1 - \frac{\tau}{2N} \right)}{\sqrt{\frac{\tau}{2} \left( 1 - \frac{\tau}{2N} \right) \left( 4c_pN - \tau \left( 1 - \frac{\tau}{2N} \right) \right)}}
= 1 - \frac{\tau}{\pi c_pN} + o \left( \frac{1}{\sqrt{N}} \right),
\]

in which the universal part is given by the constant 1 and finite size corrections scale as \( \frac{1}{\sqrt{N}} \) up to a scaling factor depending on \( c_p \).

VI. ACKNOWLEDGEMENTS

The authors are grateful to Luigi Ambrosio, from Scuola Normale Superiore in Pisa, and Andrea Sportiello, from Université Paris 13, for useful and stimulating discussions.

Appendix A: Joint distribution for the solution of the problem with open boundary conditions

In the present section we derive the joint distribution of the matching map \( m \) for the solution of the Euclidean bipartite matching problem on the line. In the hypothesis that \( 0 < t_1 < t_2 < \cdots < t_K < 1 \), \( K < N \), let us consider \( K \) \( B \)-points \( \{ y_k \}_{k=1,...,K} \), \( i_1 < i_2 < \cdots < i_K \), and evaluate the following quantity:

\[
\Pr \left( \{ y_k \in \mathcal{D} t_k \}_{k=1,...,K} \right) = \frac{N!(1 - t_K)^N - t_K}{(N - i_K)!} \prod_{k=1}^{K} \frac{(\Delta t_k)\Delta i_k - 1!}{(\Delta i_k - 1)!}
= \left( \frac{N}{\Delta i_1 \cdots \Delta i_{K+1}} \right) \prod_{k=1}^{K+1} \frac{\Delta i_{k+1} \Delta i_k}{\Delta y_k} \prod_{k=1}^{K} \Delta t_k
\]

(A1)

where \( i_0 := 0, i_{K+1} := N, t_0 := 0, t_{K+1} := 1 \) and

\[
\Delta t_k := t_k - t_{k-1}, \quad \Delta i_k := i_k - i_{k-1}.
\]

(A2)

Note that the previous equation has the form of a multinomial distribution. Introducing the rescaled variable \( m_k(t_k) = \sqrt{N} \left( t_k - \frac{i_k}{N} \right) \), in the large \( N \) limit we obtain a multivariate Gaussian distribution in the variable \( \{ \Delta m_k \}_{k=1,...,K} \), \( \Delta m_k := m_k - m_{k-1}, m_0 := 0 \), whose covariance matrix is given by the degenerate matrix

\[
\Sigma_{ij}(\{ y \}) = \delta_{ij} \Delta t_i - \Delta t_i \Delta t_j.
\]

(A3)

Note that \( \det \Sigma = \left( 1 - \sum_{k=1}^{K+1} \Delta t_k \right) \prod_{k=1}^{K+1} \Delta t_k = 0 \): the constraint \( \sum_i \Delta t_i = 1 \) reduces the rank of the matrix from \( K + 1 \) to \( K \) (we have indeed \( K \) independent variables). In this case, no density distribution exists in the \((K+1)\)-dimensional space of variables \( \Delta m_1, \ldots, \Delta m_{K+1} \), due to the fact that an additional constraint, \( \sum_{k=1}^{K+1} \Delta m_k = 0 \), holds, and therefore the distribution is singular. We need to restrict therefore to the subspace of \( K \) variables \( \{ \Delta m_k \}_{k=1,...,K} \). The distribution is still a multivariate Gaussian but with covariance matrix \( \Sigma_{ij} = \Sigma_{ij} \) with \( i, j = 1, \ldots, K \). The covariance matrix is positive definite and non singular on this subspace, being \( \Sigma = \left( 1 - \sum_{k=1}^{K} \Delta t_k \right) \prod_{k=1}^{K} \Delta t_k = \prod_{k=1}^{K+1} \Delta t_k \) and

\[
[\Sigma(y)]_{ij} = \delta_{ij} \Delta t_i + \frac{1}{\Delta t_{K+1}}
\]

(A4)

The joint distribution is therefore

\[
\Pr \left( \{ m_k(t_k) \in \mathcal{D} m_k \}_{k=1,...,K} \right) = \frac{\prod_{k=1}^{K} d m_k}{\sqrt{(2\pi)^K \prod_{k=1}^{K+1} \Delta t_k}} \exp \left( -\sum_{k=1}^{K} \frac{(\Delta m_k)^2}{2\Delta t_k} - \frac{m_K^2}{2\Delta t_{K+1}} \right).
\]

(A5)
The previous distribution is exactly the joint distribution for the Brownian bridge process, proving the equivalence of the processes in the large $N$ limit. In particular, the two point joint distribution for the Brownian bridge is given by eq. (A5) for $K = 2$:

$$\Pr(m_1(t_1) \in d m_1, m_2(t_2) \in d m_2) = \frac{d m_1 d m_2}{2\pi \sqrt{(1-t_2)(t_2-t_1)t_1}} e^{-\frac{m_1^2}{2t_1} - \frac{(m_2-m_1)^2}{2(t_2-t_1)} - \frac{m_2^2}{2(1-t_2)}}$$

(A6)

where $0 < t_1 < t_2 < 1$ is assumed.

**Appendix B: Probability distributions for the Brownian bridge**

In the present section we briefly present, without proofs, some fundamental probability results on the Brownian bridge process and some noteworthy probability distributions related to it. As explained before, the distribution for the sup of the absolute value of a Brownian bridge is given by the well known Kolmogorov distribution, eq. (20). Using the reflection principle and Bayes’ theorems, it can be also proved that [13, 16]

$$\Pr\left(\sup_{t \in [0,1]} B(t) < M, \inf_{t \in [0,1]} B(t) > -m\right) = 1 + 2 \sum_{k=1}^{\infty} e^{-2k^2(M+m)^2} + \sum_{k=0}^{\infty} \left(e^{-2(km+(k+1)M)^2} + e^{-2(kM+(k+1)m)^2}\right).$$

(B1)

For $t \in (0,1)$ and $M > b$, it can be easily obtained that

$$\Pr\left(\sup_{s \in [0,1]} B(s) < M, B(t) \in d b\right) =$$

$$= e^{-\frac{b^2}{2(1-t)}} \left(1 - e^{-\frac{2M(M-b)}{b^2}}\right)\left(1 - e^{-\frac{2M(M-b)}{b^2}}\right) \frac{d b}{\sqrt{2\pi t(1-t)}}.$$

(B2)

The computation of (36) can be easily performed using the distribution of the sup of the Brownian bridge

$$\Pr\left(\sup_{t \in [0,1]} B(t) < M\right) = 1 - e^{-2M^2}.$$

(B3)