On The Random Euclidean Assignment Problem in one dimension

Università degli Studi di Milano

Physics Department

Candidate

Giuseppe Del Vecchio Del Vecchio

Supervisor
Prof. Sergio Caracciolo

Assistant Supervisor
Phd. Enrico Maria Malatesta

16 October 2017

Abstract

This thesis work deals with the Random Euclidean Assignment Problem. Many contemporary and past authors were involved in its study both for its challenging nature and for its infinite applications. Basically, given a set of N workers and a set of N jobs one aims to find the best way to assign bijectively workers to jobs. Our target will be the optimal minimum cost. Differently, with respect the classical linear version, we have artificially introduced a disorder, using an a priori chosen probability distribution. Mapping workers and jobs to vertices of a complete bipartite graph we would like to study the average properties of the system, searching for interesting behavior as we do in physical ststems with a huge numebr of particles. We have computed the analytical expression of the ground state of the system, averaging over all possible configurations, up to the first order corrections. Explicit expressions for the correlation function is given when the cost is a function of the eucliadean distance between points. The Hungarian algorithm typical complexity is studied with numerical simulations.

To my family
and Giulia

Contents

Chapter 1

Introduction

This thesis work is organized in three main sections, each of one is divided in three subsections. In the section 2.1, **General Structure**,the problem is presented in its generality. Here,in 2.1.1, introductory concepts of Graph Theory are given, without asking for the maximum generality, but to explain, as simply as possible, what is the mathematical background to understand the problem. The main reason to study this kind of problem is that there are great similarities between disordered physical systems and the ideas developed in the general theory of random optimisation problems. Indeed, it is possible to think to the energy as a "cost" which must be minimised to reach an equilibrium position. By the way, statistical physics methods were applied succesfully to optimisation problems and so in 2.1.2 it is given a brief sketch about the celebrated *replica trick* and it is introduced the thermodynamic limit, which we will use to compute average properties of our system. Following, in 2.1.3 the computational aspects are analysed. Today, a lot of algorithms to solve linear problems are available on the market: we choose to use and to describe the Hungarian Algorithm. We studied it, reinterpreting the calssical complexity theory, which is interested in study the worst case, looking for average properties. The programming language we have used to check theoretical results is **Julia**, a high-level, high-performance dynamic language ,relatively new, in the world of computing [2] [3]. The section 2.2, **Optimal Solutions**, deal with fundamental results that allow us to explicitly compute the average properties we are interested in. Mainly, we follow the same approach of [6]. We will see that for particular values of the parameter by which the cost depends, optimal solutions are *ordered*. In the last section 2.3, **Tackling the Prob-**

lem, we effectively try to derive analytical expressions and each theoretical expression is checked by numerical simulations. In 2.3.1, it is presented a fundamental theorem ,discovered by Donsker, which allows us to use the theory of a particular stochastic process, the Brownian Bridge, to average over the disorder of the system. In 2.3.2 and in 2.3.3 the average mean cost and the correlation function, respectively, are computed, up to the first order corrections $o\left(\frac{1}{N}\right)$. It will be a challenging integral "divertissement".

Chapter 2

The Random Euclidean Assignment Problem in one dimension

2.1 General Structure

2.1.1 A Random Graph

Graph Theory has its roots at the beginning of the 19^{TH} century with the Four-Color Conjecture. In the 1976 its resolution by K. Appel and W. Hanken using a computer-assisted proof marked a turning point in the theory development [20]. Its modern foundation is based on the seminar work of Claude Berge, Paul Erdös, Bille Tutte and others. Most of real world problems have a natural interpretation in term of a graphs. In fact a graph is the mathematical abstraction of the associating things with other things. Examples could be the genealogical tree of a given person or the railway system thought with its railways and stations. An easy introduction full off examples may be found in [4]. Let us start with the basic definition of what precisely a graph is:

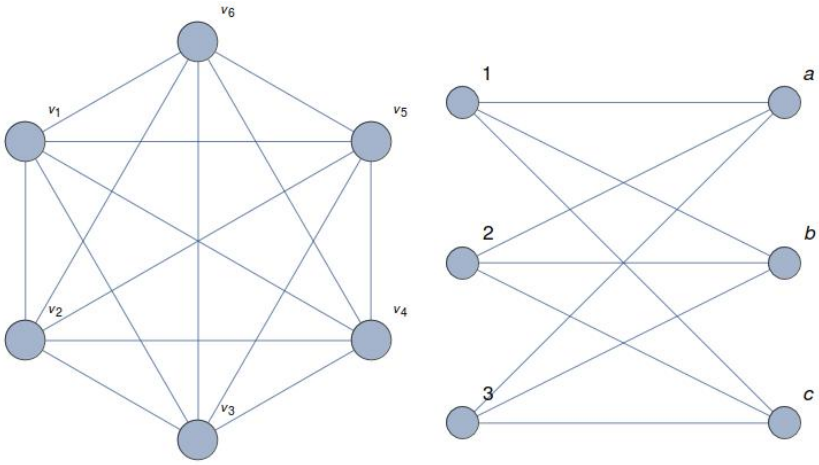
Definition 1. *Given sets $\mathcal{V}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ we define a graph $\mathcal{G}$ the pair $(\mathcal{V}, \mathcal{E})$ and an **incidence** map $\phi_{\mathcal{G}} \colon \mathcal{E} \to \mathcal{V} \times \mathcal{V}$ such that if e is an edge and u, v are vertices $\phi_{\mathcal{G}}(e) = (u, v)$.*

We always write simply $\mathcal{G} = Graph(\mathcal{V}, \mathcal{E})$ understanding the incidence map $\phi_{\mathcal{G}}$. The set $\mathcal{V}$ is the *vertices* (or nodes) set while $\mathcal{E}$ is the *edges* (or lines) set. Then given two vertices $u, v \in \mathcal{V}$ an edge between

u and v is the pair (u, v). The number of vertices is $V = |\mathcal{V}|$ and is called **order** of $\mathcal{G}$. Given an edge e and an vertex $v \in e$ we say that e is **incident** with v. We call $\mathcal{E}(v)$ the set of incidence edges at v and we $|\mathcal{E}(v)|$ the **degree** of v. If $(u, v) \in e$ we say that u and v are **adjacent**. We call **complete** graph $\mathcal{K}_{\mathcal{V}}$ the graph with V vertices in which all of them are adjacent to all the others. Basically speaking a complete graph is one where all vertices are linked with all the others. A complete **bipartite** graph $\mathcal{K}_{\mathcal{V}_1,\mathcal{V}_2}$ is a complete graph $\mathcal{G} = Graph(\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E})$ where $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$. Fig. 2.1 shows examples of such graphs. The incidence relation can be ordered leading to the definition af a digraph:

Definition 2. *A digraph $\mathcal{D} = DiGraph(\mathcal{V}, \mathcal{E})$ is a graph in which edges are ordered i.e. $(u, v) \neq (v, u)$.*

In case of a digraph we depict an edge as $\vec{e} = (\vec{u, v})$ and we say that u is the **head** and v is the **tail**. The most important tool we have to understand the properties of a graph is its graphical interpretation so that in most cases we consider the sketch of a graph the graph iteself. The diagram of a graph merely depicts the incidence relation between its vertices and edges.



(a) $\mathcal{K}_{\mathcal{V}}$ (b) $\mathcal{K}_{\mathcal{V}_1,\mathcal{V}_2}$

Figure 2.1: A complete graph a) and a complete bipartite graph b)

In order to say why graphs are the natural mathematical background of our problem we need some other basic definitions. It is possible to define different notions of subgraph. Actually, there is a general simple

way to contruct a graph from an other, that is removing or adding vertices or edges. Given a graph, by using these simple operations we can contruct every subgraphs or supergraphs.

Definition 3. *We say that a graph $\mathcal{G}' = Graph(\mathcal{V}', \mathcal{E}')$ is a subgraph of $\mathcal{G} = Graph(\mathcal{V}, \mathcal{E})$ if $\mathcal{V}' \subseteq \mathcal{V}$ and $\mathcal{E}' \subseteq \mathcal{E}$. In such case we write $\mathcal{G}' \subseteq \mathcal{G}$. It is understood that the incidence map it given by $\phi_{\mathcal{G}'} : \mathcal{E}|_{\mathcal{E}'} \to \mathcal{V}' \times \mathcal{V}'$.*

We say that $\mathcal{G}'$ is a *spanning* subgraph of $\mathcal{G}$ if $\mathcal{V}' = \mathcal{V}$. A **vertex cover** of $\mathcal{G} = Graph(\mathcal{V}, \mathcal{E})$ is a subset of $\mathcal{V}$ such that any edge of $\mathcal{G}$ has at least one endpoint in it. The smallest possible size of a vertex cover is called *vertex covering number* $c_V(\mathcal{G})$. In the same way we define an **edge cover** aa a subset of $\mathcal{E}$ such that any vertex of $\mathcal{G}$ is the end point of at least one edge in it. The smallest possible size of an edge cover is the *edge covering number* $c_E(\mathcal{G})$.

Now a fundamental definition for our aims:

Definition 4. *We say that $\mathcal{M} = Graph(\mathcal{V}_\mathcal{M}, \mathcal{E}_\mathcal{M}) \subseteq \mathcal{G} = Graph(\mathcal{V}, \mathcal{E})$ is a* matching *of size $|\mathcal{E}_\mathcal{M}|$ in $\mathcal{G}$ if given two edges in $\mathcal{M}$ they have no vertex in common.*

The size of the largest matching (*maximum matching*) is called *matching number* $m(\mathcal{G})$. Whenever $\mathcal{V}_\mathcal{M} = \mathcal{V}$ we say that the matching is perfect.

Here a fundamental result proved in 1931 by Konig:

Theorem 1 (Konig's Theorem)**.** *Let $\mathcal{G}$ be a bipartite graph. Then it has a matching if and only if*

$$c_V(\mathcal{G}) = m(\mathcal{G}) \tag{2.1}$$

and an immediate corollary:

Corollary 1 (Marriage Theorem)**.** *Let $\mathcal{G} = Graph(\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E})$ be a bipartite graph. Then it admits a perfect matching if and only a matching exists and*

$$|\mathcal{V}_1| = |\mathcal{V}_2| \tag{2.2}$$

Each times in a bipartite graph we have $|\mathcal{V}_1| = |\mathcal{V}_2| = V$, labelling or numbering its vertices in different ways, a matching can be expressed as a permutation of V elements [17]. The Assignment Problem is a classical linear problem defined on a *complete bipartite graph* $\mathcal{K}_{V_1,V_2} = Graph(\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E})$, see 2.1. Given a set of N black points $\mathbf{X} = \{X_i\}_{i=1}^N$

and a set of N white points $\mathbf{Y} = \{Y_i\}_{i=1}^N$, the embedding of the vertices on the open unitary interval $\Lambda = (0, 1)$ will be perfomed by the following (invertible) map

$$\psi \colon \mathcal{E} \to \Lambda \times \Lambda \tag{2.3a}$$
$$e \mapsto \psi(e) = (x, y) \tag{2.3b}$$

which relates each same-color points set respectively to $\mathcal{V}_1$ and $\mathcal{V}_2$. We will label points, withouth loss of generality, in such a way that:

$$x_1 \leq \ldots \leq x_i \leq \ldots \leq x_N \tag{2.4a}$$
$$y_1 \leq \ldots \leq y_i \leq \ldots \leq y_N \tag{2.4b}$$

To introduce a disorder in the system we assume that each point is a random variable with PDF given by:

$$\rho(x) = \chi_{(0,1)}(x) \tag{2.5}$$

where

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & otherwise \end{cases}$$

is the characteristic function of the set A. Random Assignment Problems are concerned with the study of the average properties of the system when $N \to \infty$ and when points are distibuted according to a given PDF. So thinking to the space of all possible REAMs we are considering just an instace, given by the choice (2.5). An optimisation problem is a couple $(\mathfrak{F}, \mathfrak{C})$ where $\mathfrak{F} \neq \emptyset$ is the **space of feasible solutions** and a **cost function** $\mathfrak{C} \colon \ \to \mathbb{R}$. The target of the problem is usually the minimisation (or maximisation) of the cost function, defined on the space of all possible initial configurations which represents the problem itself. A **globally optimal solution** $\mathrm{X}_0 \in \mathfrak{F}$ is such that

$$\mathfrak{C}[\mathrm{X}_0] = \min_{\mathrm{X} \in \mathfrak{F}} \mathfrak{C}[\mathrm{X}] \tag{2.6}$$

It is important to mention that the cardinality of $\mathfrak{F}$ and the *existence* of an optimal solution is not guaranteed.

Note the analogy with physics: a physical system tends to reach equilibrium making minimum its energy. In our case possible configurations

are given by pairs of random variables (X, Y) uniformly and indipendently distributed on the unitary interval. We do the following assumption about the cost function:

$$c(x, y) = c(|x - y|) \tag{2.7}$$

In the subsequent development we will call 2.7 *cost* instead of the more abstract one in 2.6 understanding that actually $c = \mathfrak{C} \circ \psi$, where ψ is defined in 2.3. Specifically, in our case, we choose:

$$c(z) = (|z| - \alpha)^2 \quad \Lambda \ni z = x - y \quad \alpha \in \mathbb{R} \tag{2.8}$$

It follows immediately that due to the dependence by the relative distance between points, the cost is a homogeneus and roto-traslationally invariant function. This means that its simmetry group is $SO(3, \mathbb{R})$ union the traslations. A hypotetical physical system described by this model has a degenerate spectrum as happens for the Coulomb potential. In addition, as points on the real line are completely equivalent there is also an exchange symmetry as happens for identical particles. In typical situations a particle couples or interacts with particles in its immediatly neighborhood as for example in the **Ising Model**[18]. This does not happen in our case: it is no more true that energy is locally minimised by such couplings. A given particle may quietly choose to couple with a farer one. Such systems are called **frustrated** due to this not "free" behavior. A question one may try to answer is: what is the probability that, given a black point, it couples with the k-th white point to minimise the total energy? This probability can be computed in the large N limit adn but the work is a little bit tedious. By (2.8) in our case we have the cost matrix:

$$c_{ij} = c(|x_i - y_j|) \tag{2.9}$$

It is important to say that due to the embedding of points in a geometric space correlations between entries of the cost matrix arises. For example points must satisfy the triangle intequality. It is clear, from the symmetries, that 2.9 is symmetric. Recalling Definition 4 and denoting as $\mathcal{S}$ the symmetric group (the set of all the permutations of N objects), an optimal solution for the REAM is given by a *perfect matching* $\mathrm{M}_0 \subseteq \mathcal{G}$ which minimise the total cost *per edge* i.e. such that:

$$\epsilon_N^{\alpha}(\sigma) = \frac{1}{N} \min_{\pi \in \mathcal{S}_N} \sum_{i=1}^{N} (|\varphi_i| - \alpha)^2 \tag{2.10}$$

where

$$\varphi_i \doteq x_i - y_{\pi(i)} \tag{2.11}$$

is the displacement field at the point x_i. Note that as $|\mathcal{V}_1| = |\mathcal{V}_2|$ the *exitence* of such a matching is guaranteed by the Marriage Theorem. What about randomness? An instance of the problem is an immutable configuaration of the black and white points randomly distributed on the interval. For each fixed the configuration the minimum optimal cost can be found. Clearly, for each configuration the optimal cost 2.7 will be different and this behavior does not allow us to extract satisfactory informations. By the way, we will consider a large number of instances of the same random problem, then performing an average operation over all possible initial configuration, we will get informations about average properties of the system. Indicating as $\langle\cdot\rangle$ the average over the disorder, our targets will be

$$E^*_{N,\alpha} = \langle E_{N,\alpha}\rangle = \langle \min_{\pi\in\mathcal{P}_N} \sum_{i=1}^{N} (|\varphi_i| - \alpha)^2 \rangle \tag{2.12}$$

and the correlation function given by

$$\chi_N(r) = \frac{\sum_{(x,y):|(x,y)|=r} G(x,y)}{\sum_{(x,y):|(x,y)|=r} 1} \tag{2.13}$$

where

$$G(x,y) = \sum_{\sigma} \sigma_x \sigma_y = \langle \sigma_x \sigma_y \rangle \tag{2.14}$$

and

$$\sigma_i = \frac{\varphi_i}{|\varphi_i|} \tag{2.15}$$

2.1.2 The Thermodynamic Limit

In statistical mechanics one is interested in finding *average* properties of the system. At the beginning of the 19^{TH} L.E. Boltzmann understood that macroscopic properties of objects around us are just average over a miscroscopic complex behavior with generally different laws. In a system with N possible states, in thermal equilibrium with an external

bath, within the canonical formalism it is defined the partition function as:

$$Z(\beta; N) = \sum_n e^{-\beta E_n} \tag{2.16}$$

where β is the inverse temperature of the system (or the bath) and n is a label that runs over all possible states. By the knowledge of this function it is possible to derive a set of relations which can be used to compute all macroscopic quantities of interest. In our case, when points on the interval Λ are generated, they have no kinetic energy i.e. are considered frozen in their positions. To recover this situation the average energy will be computed as:

$$\langle E \rangle = \lim_{\beta \to \infty} -\frac{\partial}{\partial \beta} \langle \log Z(\beta, N) \rangle \tag{2.17}$$

where it is understood that E already is intended to be the *thermal mean* of the energy. The average in 2.17 is still hard to be computed analytically and it has been approached using the celebrated **replica trick**[15][8]. Basically one may consider n identical *replicas* of the system and then average over it. The mean of a logarithm is not an easy object to manipulate and one should try to use some smart method. We can write:

$$\begin{aligned} & x^n = e^{n \log x} \\ & \Rightarrow e^{n \log x} \approx 1 + n \log x \quad as \quad n \to 0 \\ & \Rightarrow \log x = \lim_{n \to 0} \frac{x^n - 1}{n} \end{aligned} \tag{2.18}$$

Then,

$$\langle \log(Z(\beta; N)) \rangle = \lim_{n \to 0} \frac{\langle Z(\beta; N)^n - 1 \rangle}{n} = \lim_{n \to 0} \frac{\langle Z(\beta; N)^n \rangle - 1}{n} \tag{2.19}$$

Moreover, using

$$nx \approx \log(1 + nx) \quad as \quad nx \ll 1$$

we get,

$$n\langle\log(Z(\beta;N))\rangle \approx \log(1+n\langle\log(Z(\beta;N)\rangle) \qquad (2.20)$$
$$as \quad n\to 0$$

and by 2.18

$$\begin{aligned}\langle\log(Z(\beta;N))\rangle &= \lim_{n\to 0}\frac{1}{n}\log(1+n\langle\log(Z(\beta;N))\rangle)\\ &= \lim_{n\to 0}\frac{1}{n}\log[1+n(\frac{\langle Z(\beta;N)^n\rangle-1}{n})]\\ &= \lim_{n\to 0}\frac{1}{n}\log(\langle Z(\beta;N)^n\rangle) \qquad (2.21)\end{aligned}$$

The mindful reader should be aware as $n\in\mathbb{N}$ then in performing the limit 2.18 one must specify how to take this limit due to the fact that 0 is not an accumulation point for the natural numbers set $\mathbb{N}$. To avoid this mathematical problem one considers the analytical continuation of $Z(\beta;N)^n$, where $n\mapsto x\in\mathbb{R}$, and then performs the limit.

2.1.3 The Hungarian Method

In this section we will face with the computational aspects of the problem. Given an instance of the problem, there are a lot of algorithms able to find the optimal cost[10][14][1]. Every problem has to do with linearity is easier to be solved and so the huge interest in linear programming. Algorithms are routiney improved and an extensive study of their complexity is done. An instance of the REAM is just its classical linear version. But how to formalise precisely a linear problem? Hiterto, We have talked about such a problem but until this point we have not stated its precise definition. For an optimisation problem two possibilities exist:

- Maximisation
- Minimisation

In our study we consider the minimisation of a function, the *cost*, then let us give the definition in this this case [9]:

Definition 5 (Standard Minimum Linear Problem)**.** *A **linear problem** is the research of a vector of numbers* $\boldsymbol{x}$ *such that*

$$\boldsymbol{x}^T\boldsymbol{b} = \min_{\boldsymbol{y}} \sum_{n \in I \subseteq \mathbb{N}} y_n b_n \tag{2.22a}$$

$$\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{c} \tag{2.22b}$$

$$\boldsymbol{x} \geq 0 \tag{2.22c}$$

where $\boldsymbol{A}$*,* $\boldsymbol{c}$ *and* $\boldsymbol{b}$ *are a matrix and vectors respectively (we suppose the dimensions coherent with standard matrix multiplication).*

By Def. 5 we can say that a linear problem is defined by a set of inequalities (which may be equalities of course) as 2.22b and 2.22c, called **constraints**, and a *target* function as 2.22a to minimise (maximise). A *solution* **x** is called **feasible** if it satisfies the constraints. In the same way a problem is said to be **feasible** if there exist at least one feasible solution. Moreover, we say that a problem is **unbounded** if the solution can assume arbitrarily large values otherwise is **bounded**. Among the wide literature on the subject we choose the **Hungarian** method to solve computationally the Assignment Problem. As anticipated, for each instance we compute the minimum cost and then we average over all realisations. The Hungarian algorithm was developed by Harold Kuhn in 1955, who based its study on the previous work of hungarians mathematicians Dénes König and Jenó Egerváry. The complexity class of the implemented algorithm is $O(N^3)$ as we will verify later. This is not the unique possible existing implementation but it is surely the best one. It is an example of procedure which uses an *augmenting path* subroutine to speedup the research of the optimal solution. Mathematically speaking is not hard to illustrate how and why the algorithm works, we need just a couple of very easy definitions. With reference to definitions in Sec. 2.1.1 we define:

Definition 6. *Let* $\mathcal{G} = Graph(\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E})$ *be a weighted bipartite graph with weight function* w*.*

i) A feasible labeling *is a function* $l \colon \mathcal{V} \to \mathbb{R}$ *such that*

$$l(v_1) + l(v_2) \leq w(v_1, v_2) \tag{2.23}$$

ii) The equality graph $\mathcal{G} = Graph(\mathcal{V}, \mathcal{E}_l)$ *(with respect to 2.23) is defined as*

$$\mathcal{E}_l = \{(v_1, v_2) \colon l(v_1) + l(v_2) = w(v_1, v_2)\} \tag{2.24}$$

Here the promised easy-to-prove result which shows **why** the Hungarian algorithm works:

Theorem 2 (Kuhn-Munkres)**.** *Given a bipartite graph $\mathcal{G} = Graph(\mathcal{V}, \mathcal{E})$, with $|\mathcal{V}| = |\mathcal{V}_1| + |\mathcal{V}_2|$, if l is a* feasible labeling *and $\mathcal{M}$ is a perfect matching in $\mathcal{G}_l$, then $\mathcal{M}$ is a minimum-weight matching in $\mathcal{G}$.*

Proof. Denote an edge $e \in \mathcal{E}$ as $e = (v_1, v_2)$. Let $\mathcal{M}'$ be any perfect matching in $\mathcal{G}$. By definition every $v \in \mathcal{V}$ is covered exactly once by an edge in $\mathcal{M}$. Then:

$$\sum_{e \in \mathcal{M}'} w(e) \geq \sum_{(v_1, v_2) \in \mathcal{M}'} (l(v_1) + l(v_2)) = \sum_{v \in \mathcal{V}} l(v)$$

Now let $\mathcal{M}$ be a perfect matching in $\mathcal{G}_l$. By the definition in 2.24 the thesis is proved. □

To give an idea of **how** the algorithm works here there is an intuitive explanation:

Algorithm 1 Hungarian Algorithm (augmenting-path)

```
  function F(I)ND() Find an augmenting path for M ⊆ G_l
  and increases the size of M
  end function
Require: A feasible labeling l and some matching M ⊆ G_l
  while M ⊆ G_l is not perfect do
  FIND()
      if No augmenting path exists then Improve l to l'
      end if
  end while
```

Note that the process will continue to find augmenting paths but as the size of the the graph is finite the execution must terminate. The result will be of course the optimal matching. To test the mean time distribution of the Hungarian algorithm,for fixed input N, we generated $5 \cdot 10^4$ different configurations, recovering at each time the duration of a single computation of the optimal solution. After, an average over all recovered times gives the mean time. Repeating that procedure a lot of times, 10^4 is good again, we are able to trace the distribution. It is important to say that we work with a physical situation in mind. A measure of the time distribution may give useful information about the energy of the simulated system. Programs ran on a machine with an Intel

Core i7-5500U CPU @ 2.40GHz x 4 (Turbo Boost up to 3.0 GHz) and a total of 8 GB of RAM.

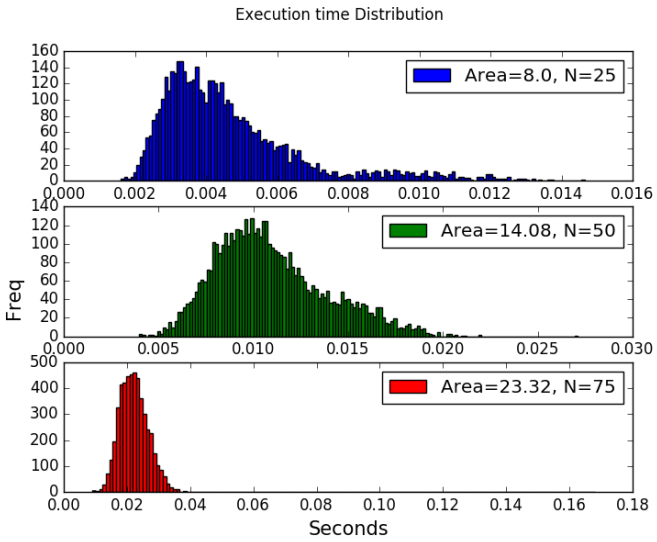


Figure 2.2: Some particular time distributions

In Fig. 2.2 there are the distributions for three different inputs. The first thing we note is that, for a small number of points to match, the shape of the distribution shows a second relative maximum near the absolue one. For $N = 50$ and $N = 75$ there is not such a maximum. How do we interpret this result? When points are less there must be certain configurations which raise up the average time. This means that not all configurations are equivalent but there are instead two calsses of them. When the number of points is greater then a certain threshold this splitting is no more present. Anyway, we are not interested in small-number confirurations, because of, as explained in section Sec. 2.1.2 we are studying the thermodynamic limit and so the system will be always far from this unconventional, thought interesting, situation. As we go up with the input size as the shape becomes smoother and regular. Contrary to the case of the previous discussed case when dealing with a big number of points we can be sure we are treating a calss of equivalent configurations. This happens because of the raising N, the number of all possible matchings becomes larger forcing the relative frequency of "strange" situations to be smaller. The result is that the algorithm 1 takes more time to augmenting path and then to output the optimal solution. In

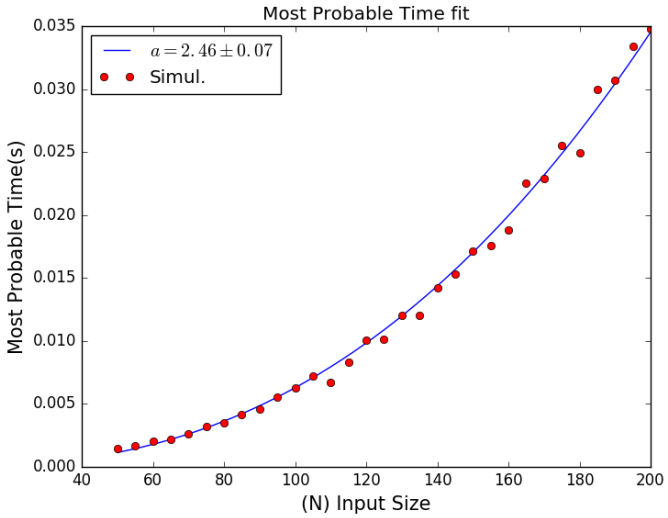


Figure 2.3: Here we fitted the position of the most probable time of the time distribution as the input size N grows.

any case, we can say more about the time distribution. Let be $f_H(N,t)$ the time distribution of the Hungarian algorithm when the size of the input is N. Clearly, as every respectable ditribution, it must satisfy a normalisation condition:

$$\int_{\mathbb{R}} f_H(N,t)dt = 1 \tag{2.25}$$

Suppose now that there exist a, b such that

$$\int_{\mathbb{R}} N^b f_H(1, t/N^a)dt = 1 \tag{2.26}$$

With a simple change of variable we find,

$$N^{a+b} \int_{\mathbb{R}} f_H(1,u)du = 1 \tag{2.27}$$

which implies that $a = -b$. Then we look for an universal exponent such that:

$$N^{-a} \int_{\mathbb{R}} f_H(1, t/N^a)dt = 1 \tag{2.28}$$

Fig. 2.3 shows the results from a simulation with $15 \cdot 10^3$ experiments on the distribution. The strategy of the fitting was to record the positions

of the most probable time or the average mean time as a function of the size N and then fit them with a non-linear method. Note that being the complexity of the Hungarian algorithm $O(N^3)$ we expect a coefficient smaller than three. Fig. 2.4 shows a good agreement with our ansatz. In Fig. 2.5 the experimental complexity $\mathcal{C}_{exp}$ of the algorithm is plotted. It is evident that $O(N^2) \leq \mathcal{C}_{exp} \leq O(N^3)$.

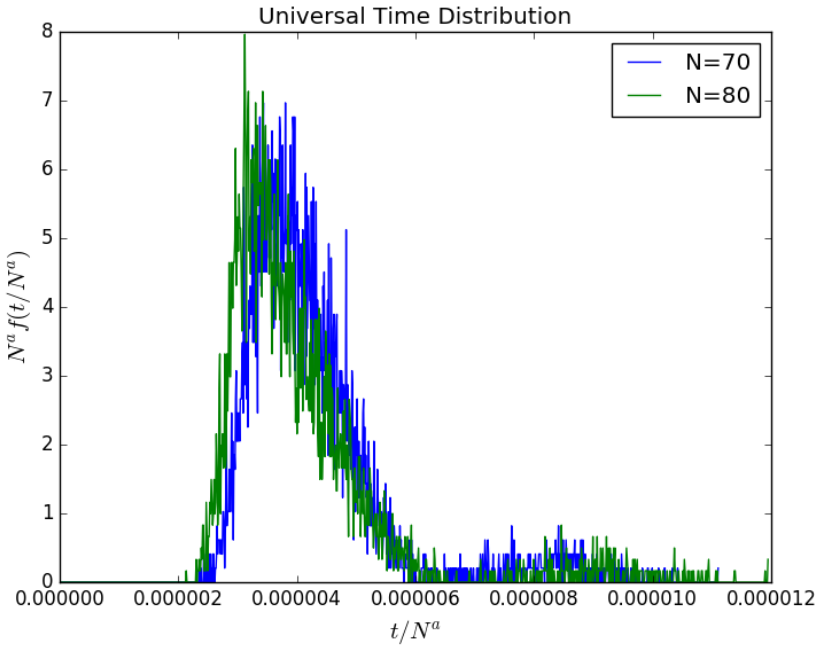


Figure 2.4: In the plot the hypotetical time distributions for $N = 70$ and $N = 80$. The Two curves tend to overlap with each other.

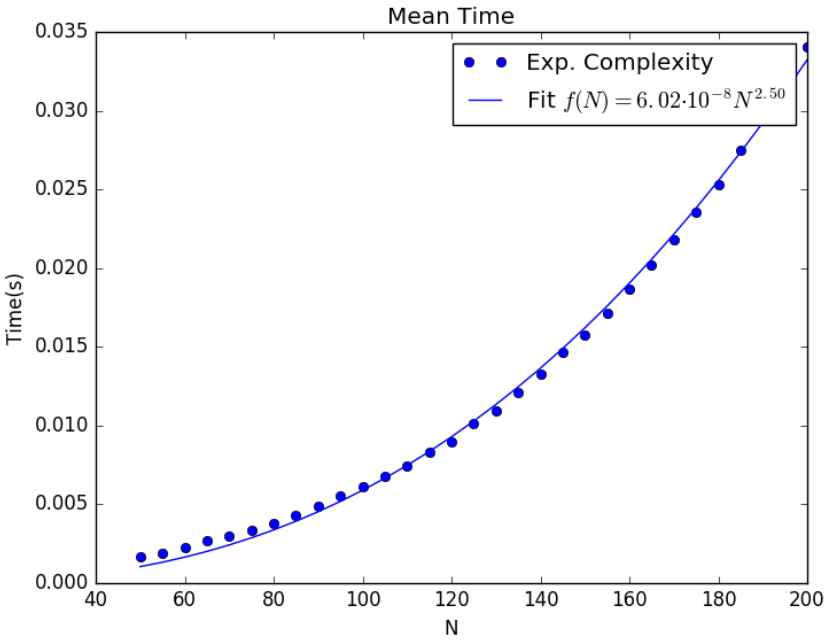


Figure 2.5: The experimental complexity is between $O(N^3)$ and $O(N^2)$

2.2 Optimal Solutions

In Sec. 2.1 we discussed the general structure of our problem. This section is devoted to the presentation of some results which will reveal fundamental to make explicit computations of the quantities of interest. Recall that our aim is the evaluation of the average minimum cost eq. 2.12 and the correlation function eq. 2.13. The explicit evaluation of the average minimum cost eq. 2.12 is not possible without knowing the permutation which makes the cost minimum. We have seen that the optimal matching is determined by a particular permutation. Of course we call that *optimal permutation*.

Definition 7. *Given a set of N elements, we say that a permutation $\pi \in \mathcal{S}_N$ belongs to $\mathsf{C}_N \subseteq \mathbb{N}$ if there exists $k \in \mathbb{N}$ such that $0 \leq k < N$ and*

$$\pi(i) = i + k \mod N, \quad i = 1, ..., N \tag{2.29}$$

This simply apparent definition will be fundamental in the future. In fact remembering that the cost function 2.8 depends parametrically by

α, we will show that the optimal permutation has a different structure depending on the values assumed by α. That is because such a parameter influences the shaping of the cost function whose properties influence the structure of the optimal permutation. The set of *cyclic* permutations C_N form an abelian group (the proof is straighforward and will be omitted) and has *order* N. Geometrically it corresponds to the N discrete rotations which leave a plane polygon invariant. Note also that $\mathsf{C}_2 = \mathcal{S}_2$ and $\mathsf{C}_3 = \mathcal{A}_3$, the alternating group of even permutations of three elements.

Definition 8. *Given a triple of three integers* $(i, j, k) \in [N]$, $[N] \doteq \{1, 2, ..., N\}$*, we say that the ordered triple of integers* (a, b, c) *is* cyclically co-oriented *with it it there esists an even permutation* $\pi \in \mathsf{C}_3$ *such that* $(\pi(a), \pi(b), \pi(c))$ *is in the same order of* (i, j, k) *respect to the order relation of integers.*

The following proposition is rather important:

Proposition 1 (Necessary and sufficient condition for cyclicity)**.** *A permutation* $\pi \in \mathcal{P}_N$ *belongs to* C_N *if and only if for any triple* (i, j, k) *the corresponding triple* $(\pi(i), \pi(j), \pi(k))$ *is cycliccally co-oriented with* (i, j, k).

Proof. Assume that $\pi \in \mathsf{C}_N$. Then π is order preserving ando so the *only if* holds. Nssume that each triple (i, j, k) is cyclically co-oriented with its image through the permutation π. Observe that for any couple $(i, j) \in [N]^2$ we have $\pi(i) - \pi(j) = i - j \mod N$, the $\pi \in \mathsf{C}_N$ To prove this we proceed by contradiction. Suppose that there exist at least one couple $(i, j) \in [N]^2$ such that $\pi(i) - \pi(j) \neq i - j \mod N$. If so, the sequences

$$I = (i, i + 1 \mod N, ..., j)$$
$$J = (\pi(i), \pi(i) + 1, ..., \pi(j))$$

do not have the same cardinality. This implies that there exists k such that either $k \in I$ and $\pi(k) \notin J$ or $k \notin I$ and $\pi(k) \in J$. By consequence, the triples (i, j, k) are not cyclically co-oriented and the theorem is proved. □

The proof of the statement can be found in [5]. The results means that to test wether a permutation is cyclic we need just to check if it leaves invariant (respect to the standard order relation) triples of integers.

Definition 9 (Crossing and Planar Matchings)**.** *Consider two sets of points* $\boldsymbol{X} = \{X_i\}_{i=1}^N$ *and* $\boldsymbol{Y} = \{Y_i\}_{i=1}^N$ *on the interval* $\Lambda = (0, 1)$ *labeled as in 2.4. A matching between* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *is said to be* planar or non-crossing *if, given the corresponding permutation* π *adn two pair of points* $(x_i, y_{\pi(i)})$ *and* $(x_j, y_{\pi(j)})$*, the corresponding intervals are either disjoint* $x_i < y_{\pi(i)} < x_j < y_{\pi(j)}$*, or nested* $x_i < x_j < y_{\pi(j)} < x_{\pi(i)}$*. Otherwise the matching is said to be* crossing.

From a pictorial point of view a matching is crossing if drawing a semi-circle in the upper half-plane which connects a pair of matched points, it intersects one that connects another pair of points.

As anticipated in this section when the parameter α varies, the shape of the cost function changes. An approach to find explicit solutions is to characterize the structure of the optimal permutation in dependence of the functional properties of the cost eq. 2.8. For the further development, it is useful a little analysis of our cost function, given by:

$$c(z) = (|z| - \alpha)^2$$

In ref. [6] it was studied the cost function $c(z) = |z|^p$ and pretty results were obtained. In the article, two different regions are taken into account:

- $p \in (-\infty, 0)$, where the optimal solution is the identity permutation
- $p \in (1, +\infty)$, where the optimal solution is a cyclic permutation

The same approach has been used here and it has revealed to show excellent results in our case too.

2.2.1 Strictly Increasing Convex Functions

It is trivial to prove that the considered cost function is convex and strictly increasign when $\alpha < 0$. The shape of c is shown in Fig. 2.6:

Proposition 2. *The function* $c(z) = (|z| - \alpha)^2$ *is a strictly increasing and convex for* $\alpha < 0$ *when* $z \in \Lambda$.

i) Strictly increasing

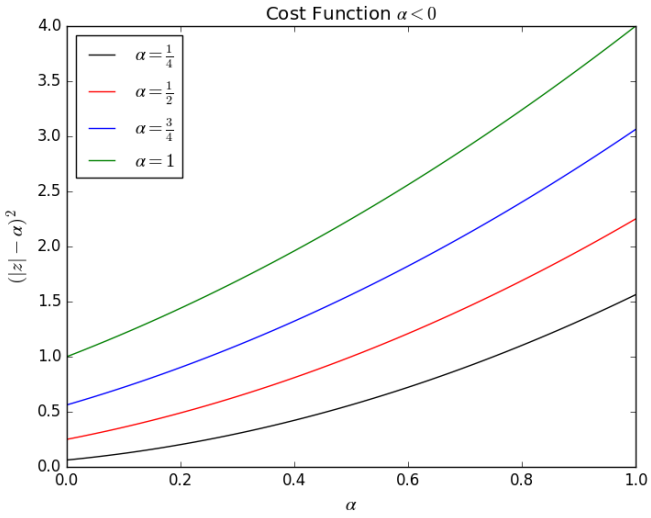


Figure 2.6: The cost for some values of α

Proof. Using the definition, taking $z_1, z_2 \in \Lambda \quad z_1 < z_2$:

$$
\begin{aligned}
(z_1 - \alpha)^2 &< (z_2 - \alpha)^2 \\
z_1^2 + \alpha^2 - 2z_1\alpha &< z_2^2 + \alpha^2 - 2z_2\alpha \\
z_1^2 - z_2^2 &< 2\alpha(z_1 - z_2) \\
\alpha &< 0
\end{aligned}
$$

because $z_1^2 - z_2^2 > 0$ and $z_1 - z_2) < 0$. □

ii) Convexity

Proof. The second derivative is always positive. □

At the beginning of this section we anticipated that analytical properties of the cost force the structure of the optimal permutation. In fact the following theorem holds:

Theorem 3. *Given the assignment problem on the unitary interval, if the cost function is strictly increasing and convex, then the optimal permutation is the identity permutation.*

Having proved that our cost function is strictly increasing and convex in the region $\alpha < 0$, this theorem says precisely what is the permutation which minimises the total cost. We will use explicitly this result in the last section.

2.2.2 C-Functions

What is a C-function and where this name come from?

Definition 10 (C-function)**.** *We say that a function* $f\colon (0,1) \to \mathbb{R}$ *is a* C-function *if for any* $\eta \in [0, 1 - z_2], \hat{\eta} \in [z_2, 1]$ *and* $0 < z_1 < z_2 < 1$ *:*

$$f(z_2) - f(z_1) \leq f(\eta + z_2) - f(\eta + z_1) \tag{2.30a}$$

$$f(z_2) - f(z_1) \leq f(\hat{\eta} - z_2) - f(\hat{\eta} - z_1) \tag{2.30b}$$

Eq. 2.30a implies that $\Psi_\eta(z) := f(\eta + z) - f(z)$ is an increasing function on the interal $(0, 1-\eta)$, for any value of $\eta \in (0, 1)$. In ref. [6] it is shown that if f is continuous eq. 2.30a is equivalent to convexity. Eq. 2.30b implies that the function $\Phi_\eta(z) := f(\eta - z) - f(z)$ is increasing on the interval $(0, \eta)$ for any $\eta \in (0, 1)$. If f is differentiable, this can be written as:

$$f'(\eta - z) + f'(z) \leq 0 \quad z \in (0, \eta) \quad \eta \in (0, 1)$$

and for $\eta \to 0$ we have:

$$f'(1 - z) + f'(z) \leq 0 \tag{2.31}$$

Using these facts it is immediate to show the following:

Proposition 3. *The function 2.8 is a* C*-function when* $\alpha > \frac{1}{2}$.

Proof. Being a polynomial continuity is guaranteed and convexity is independent of α as shown in Proposition 2 and so 2.30a is satisfied. For the second inequality using 2.31 we have, $f'(1 - z) + f'(z) = 2(1 - z - \alpha) + 2(z - \alpha) = 2 - 4\alpha \leq 0$. Then $\alpha > \frac{1}{2}$. □

The name C-function is justified by the following results:

Theorem 4. *Given the assignment problem on the unitary interval if the cost function is a* C*-function , then the optimal permutation belongs to* C_N*, i.e. is a cyclic permutation.*

So, in our case, for $\alpha > \frac{1}{2}$ the optimal solution is a cyclic permutation. It is important to note that Theorem 4 does not say what is the precise optimal solution. But instead of searching the optimal solution in the set of all permutations $\mathcal{S}_N$, which has cardinality N!, we can restrict our attention to C_N, which has cardinality N. In computational terms this is a great advantage and in fact it turned out to be fundamental in avoiding simulation times of the order of 24 hours. In addition because the rapid explosion of N!, it would be hard to check our theoretical results if we are interested in the thermodynamic limit. Theorem 4 allows us to decrease dramatically the number of operations needed to compute the quantities of interest.

2.3 Tackling the Problem

Finally we arrived at the conclusive point of this thesis. Here we will show a comparison between theorical predictions and numerical results. For the most part of what follows we will consider the *mean cost per edge* given by:

$$\epsilon_{N,\alpha} := \frac{1}{N} E_{N,\alpha} \tag{2.32}$$

We compute the leading and the subleading term up to the order $o(\frac{1}{N})$ of eq. 2.32 while for the eq. 2.13 the results is correct in accord to Donsker's Theorem of the following section.

2.3.1 Donsker's Theorem

Recall that in the REAM we have a cost functions whose variables are actually randomly generated according to the distribution eq. 2.5 of Sec. 2.1.1. The dispacement field is connected, in the thermodynamic limit, with a stochastic process called Brownian Bridge, which is a linear combination of Wiener Processes. **Donsker's Theorem** makes this connection quantitative and allows us to do an analytical forecasting of the quantities in eq. 2.12 and in eq. 2.13. A brief introduction to these processes is given in Appendix A.

Theorem 5 (Donsker)**.** *For any $N \in \mathbb{N}$ there exists a probability space Ω_N such that we can define on it the random variable $\boldsymbol{X}_N := (X_i)_i, \boldsymbol{X}_N :$*

$\Omega_N \to \Lambda^N$, *each component being a random variable uniformly distributed on the unitary interval* Λ. *Moreover, let us consider the corresponding N-th empirical process*

$$F_N(t, \boldsymbol{X}_N) := \frac{1}{N}\sum_{i=1}^{N}\theta(t - x_i) - t \tag{2.33}$$

Then we can find a ***path-continuous*** *Brownian Bridge process on* Λ, *defined on the same probability space* Ω_N, $B_N : \Lambda \times \Omega_N \to B_N(t, \omega)$ *such that for all* $\epsilon > 0$

$$\lim_{N\to\infty} \mathbb{P}[\sup_{t\in[0,1]} |\sqrt{N}F_N(t, \boldsymbol{X}_N(\omega)) - B_N(t, \omega)| > \epsilon] = 0 \tag{2.34}$$

In terms of Appendix A we can say that

$$\sup_{t\in[0,1]} |\sqrt{N}F_N(t, \mathbf{X}_N(\omega)) - B_N(t, \omega)| \to 0$$

in probability. In [7][13] the convergence rate has been studied:

$$\sup_{t\in[0,1]} |\sqrt{N}F_N(t, \mathbf{X}_N(\omega)) - B_N(t, \omega)| = O(\frac{\log N}{\sqrt{N}}) \tag{2.35}$$

Let us now see why Donsker's Theorem is so fruitful for the study of our problem. Given an instance of the problem $\omega \in \Omega_N$, take a realisation of points distrubution $\Xi_N = \mathbf{X}_N(\omega) = \{x_i\}_{i=1}^N$. The empirical process in Donsker's Theorem is :

$$F_N(t, \mathbf{X}_N(\omega)) = \frac{1}{N}\sum_{i=1}^{N}\theta(t - x_i) - t \tag{2.36}$$

Now considering the labeling order given in 2.4, we have:

$$\lim_{t\to x_i} F_N(t, \mathbf{X}_N(\omega)) = \frac{i}{N} - x_i \tag{2.37}$$

Therefore, for two different realisations $\Xi_N = \mathbf{X}_N(\omega) = \{x_i\}_{i=1}^N$ and $\Gamma_N = \mathbf{X}_N(\hat{\omega}) = \{y_i\}_{i=1}^N$, we can write:

$$y_j - x_i = \frac{j - i}{N} + \lim_{t\to x_i^+} F_N(t, \mathbf{X}_N(\omega)) - \lim_{t\to y_j^+} F_N(t, \mathbf{X}_N(\hat{\omega})) \tag{2.38}$$

At this point we do practically the thermodynamic limit: take $i = Nu + \frac{1}{2}$ and $j = Nv + \frac{1}{2}$ with $u, v \in (0, 1)$. In this way we are using the fact

that rational numbers are dense on the real line allowing us to work with continuous indices u, v instead of discrete ones i, j. The shift $\frac{1}{2}$ is just convenient to do calculations and obviously does not affect final results. By using this transormation it is possible to choose every generated point in the interval Λ. For example, the first point is taken if we choose $u = \frac{1}{2^N}$ and the last one is taken when $u = 1 - \frac{1}{2^N}$. Now, the probability to find a given point x_i in the interval $[x, x + dx]$ is given by:

$$\begin{aligned}\mathbb{P}(x_i \in dx) &= \frac{x^{i-1}}{(i-1)!}\frac{x^{(N-i)}}{(N-i)!}\frac{1}{\frac{x^N}{N!}} = \binom{N}{i}x^i(1-x)^{N-i}\frac{i}{x}dx \\ &= B_i(N,x)\frac{i}{x}dx \end{aligned} \tag{2.39}$$

being $B_i(N, x)$ the binomial distribution. This implies that, in the large N limit,

$$\mathbb{P}(|x_i - u| \geq \epsilon) = \exp\left(-\frac{N\epsilon^2}{2u(1-u)}\right) \tag{2.40}$$

Then by Donsker's Theorem we can write:

$$\sqrt{N}(y_{Nv+\frac{1}{2}} - x_{Nu+\frac{1}{2}} + u - v) \to B(u,\omega) - B(v,\hat{\omega}) \quad , \quad N \to \infty \tag{2.41}$$

The fundamental property of Brownian bridge processes we will use is:

$$\langle B(s,\omega)B(t,\omega)\rangle = \min(s,t) - st \tag{2.42}$$

For the proof of this property go to Appendix A.

2.3.2 The Ground State

In Sec. 2.2 we saw that analytical properties of the cost function 2.8 are different in dependence of the values of α. By the way, structure of the optimal permutation is different depending on α. Then, it is smart to consider two different intervals of solution:

i) $\alpha < 0$

ii) $\alpha > \frac{1}{2}$

Let us start with the case i). Consider eq. 2.32. By Theorem 2 we have:

$$\epsilon_{N,\alpha} = \frac{1}{N}\sum_{k=1}^{N}(|\varphi_k| - \alpha)^2 = \alpha^2 - \frac{2\alpha}{N}\sum_{k=1}^{N}|\varphi_k| + \frac{1}{N}\sum_{k=1}^{N}|\varphi_k|^2 \tag{2.43}$$

This expression, written as above, is rather difficult to tackle, because of the presence of the disorder. Using the fact that black and white points are random variables independently and identically distributed, we can average over all possible configurations using the following expression:

$$\mathbb{P}(\varphi_k \in d\varphi) = d\varphi \binom{N}{k}^2 k^2 \times \\ \times \iint_0^1 \delta(\varphi - y + x)(xy)^{k-1}[(1-x)(1-y)]^{N-k} dxdy \tag{2.44}$$

Now defining $\phi(s) = \sqrt{N}\varphi_{Ns+\frac{1}{2}}$ and taking the limit $N \to \infty$ we get:

$$\mathbb{P}(\phi(s) \in d\phi) = d\phi \frac{e^{-\frac{\phi^2}{4s(1-s)}}}{2\sqrt{\pi s(1-s)}} \{1 + \frac{s(1-s)+1}{8Ns(1-s)} + \frac{7s(1-s)-2}{8Ns^2(1-s)^2}\phi^2 \\ + \frac{1-5s(1-s)}{32Ns^3(1-s)^3}\phi^4 + o(\frac{1}{N})\} \tag{2.45}$$

This result may appear rather complicated but actually it is not too difficult to work with it. It is convenient to introduce the following notation:

$$\frac{1}{N}\sum_{k=1}^{N} |\varphi_k|^p := \epsilon_N^{(p)} \tag{2.46}$$

The average mean cost per edge eq. 2.32 becomes:

$$\epsilon_{N,\alpha} = \alpha^2 - 2\alpha\epsilon_N^{(1)} + \epsilon_N^{(2)} \tag{2.47}$$

that means that in the present case, $\alpha < 0$, the cost is a quadratic function in α, at least in the leading term. Following [6] and using the well-known properties of the Γ function, it is easy to derive the expression below, as N in large:

$$N^{\frac{p}{2}}\epsilon_N^{(p)} ds = \int_0^1 \langle |\phi(s)|^p \rangle ds = \\ = \frac{\Gamma(\frac{p}{2}+1)}{p+1}\Big(1 - \frac{p(p+2)}{8N}\Big) + o\Big(\frac{1}{N}\Big) \tag{2.48}$$

Now using eqs. 2.47 and 2.48, we get:

$$\epsilon_{N,\alpha} = \alpha^2 + \frac{1}{3N}\Big(1 - \frac{1}{N}\Big) - \frac{\alpha\sqrt{\pi}}{2\sqrt{N}}\Big(1 - \frac{3}{8N}\Big) + o\Big(\frac{1}{N}\Big) = \\ = \alpha^2 - \frac{\alpha\sqrt{\pi}}{2\sqrt{N}} + \frac{1}{3N} + o\Big(\frac{1}{N}\Big) \tag{2.49}$$

and the total cost:

$$E_{N,\alpha} = N\epsilon_{N,\alpha} = \\ = N\Big(\alpha^2 - \frac{\alpha\sqrt{\pi}}{2\sqrt{N}} + \frac{1}{3N} + o\Big(\frac{1}{N}\Big)\Big) \tag{2.50}$$

Note the expected parabolic dependence by α of the leading term. The subleading term is linear in α. Here an interesting thing happens: the first correction to the leading term is of order $o\Big(\frac{1}{\sqrt{N}}\Big)$. This happens because by eq. 2.48 there is a mismatch in the scaling of $\epsilon_N^{(1)}$ and $\epsilon_N^{(2)}$. The following plots shows the agreement between theory and simulations. In particular Fig. 2.7 shows the finite size corrections. We resctricted our attention to the first order term i.e. we have considered

$$E_{N,\alpha} = N\Big(\alpha^2 - \frac{\alpha\sqrt{\pi}}{2\sqrt{N}} + o\Big(\frac{1}{\sqrt{N}}\Big)\Big) \tag{2.51}$$

Further corrections can be verified in the same way as shown in the figure 2.8.

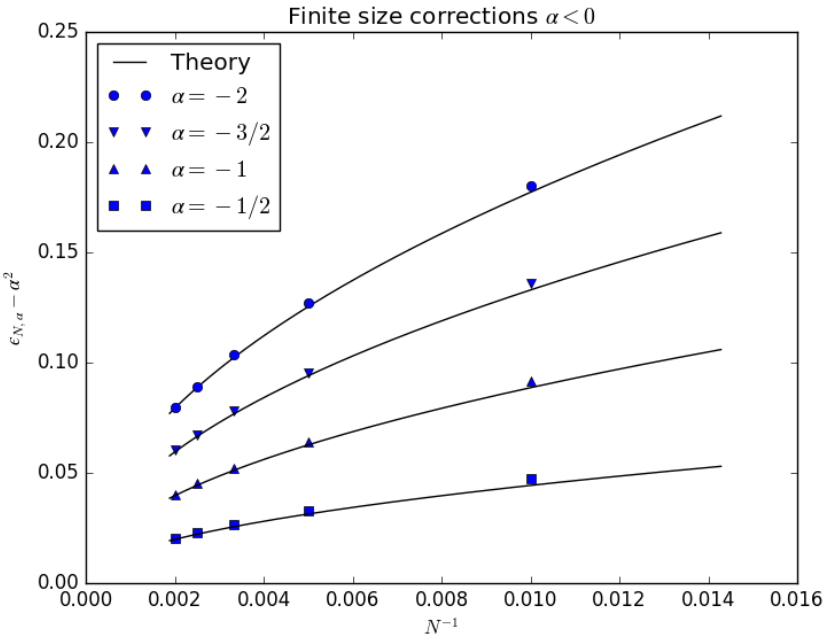


Figure 2.7: In this plot we see the agreement between the theoretical prediction the finite size corrections in eq. 2.51. Simulation is done over 10^5 instances. The agreement with the thoeretical prediction is good and error bars are typically smaller then the markers size.

Let us now try to do a similar computation when $\alpha > \frac{1}{2}$, i.e. the case ii). This is the more challenging case because Theorem 4 does not say what is the precise optimal permutation, we know just that $\pi_{opt} \in \mathsf{C}_N$. Define the quantity:

$$\frac{1}{\sqrt{N}}\phi_t^{(N)}(s) := y_{N((s+t) \mod 1)+1/2} - x_{Ns+1/2} - \sigma(s,t) \tag{2.52}$$

where $Ns + 1/2 = k \in [N]$, $Nt \in [N]$ and

$$\sigma(s,t) := ((s+t) \mod 1) - s \tag{2.53}$$

In the above expression we have used the continuum limit approximation. The shift variable t is the continuum counterpart of k in Definition 7. We can write

$$\epsilon_N^{(p)} = \frac{1}{N}\sum_{k=1}^{N}\left|\sigma\Big(\frac{k-1/2}{N}, t\Big) + \frac{1}{\sqrt{N}}\phi_t^{(N)}\Big(\frac{k-1/2}{N}\Big)\right|^p \tag{2.54}$$

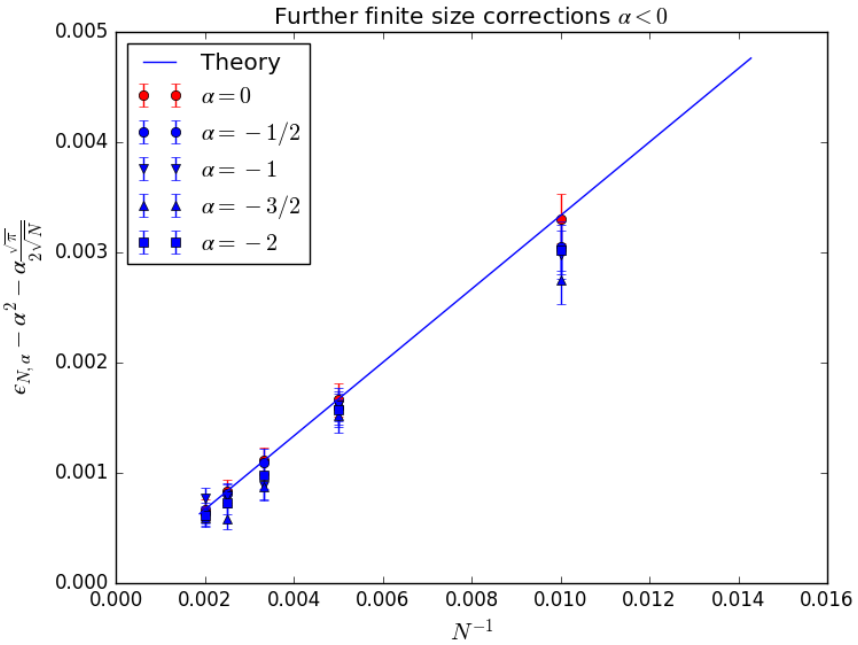


Figure 2.8: Checking the corrections of order $o\left(\frac{1}{N}\right)$. Simulation is done over 10^5 instances.

and using again Donsker's Theorem we have

$$\phi_t^{(N)}(s) \to \phi_t^{(N)} := B(s,\omega) - B((s+t) \mod 1), \hat{\omega}), \quad N \to \infty \tag{2.55}$$

Now eq. 2.54 as $N \to \infty$ becomes:

$$\epsilon_N^{(p)} = \int_0^1 \left|\sigma(s,t) + \frac{1}{\sqrt{N}}\phi_t(s)\right|^p ds \tag{2.56}$$

Now to estimate the mean shift we proceed perturbatively:

$$\lim_{N\to\infty} \int_0^1 \left|\sigma(s,t) + \frac{1}{\sqrt{N}}\phi_t(s)\right|^p ds = t^p(1-t) + t(1-t)^p \tag{2.57}$$

Putting the above expression, specialised for $p = 1, 2$, in eq. 2.47, we

get:

$$\epsilon_{N,\alpha} = \alpha^2 - \frac{4\alpha}{N}(t(1-t)) + \frac{1}{N}(t^2(1-t) + t(1-t)^2) = \\ = \frac{t - 4t\alpha + N\alpha^2 + t^2(4\alpha - 1)}{N} \tag{2.58}$$

and minimising with respect to t we get,

$$t = 1/2 \tag{2.59}$$

which means that the shift is independent by α. Following again [6] we suppose that

$$t = 1/2 + \frac{\tau}{\sqrt{N}} \tag{2.60}$$

where τ depends both on α and on the particular configuration of the problem. The figure above shows the results of a numerical simulation.

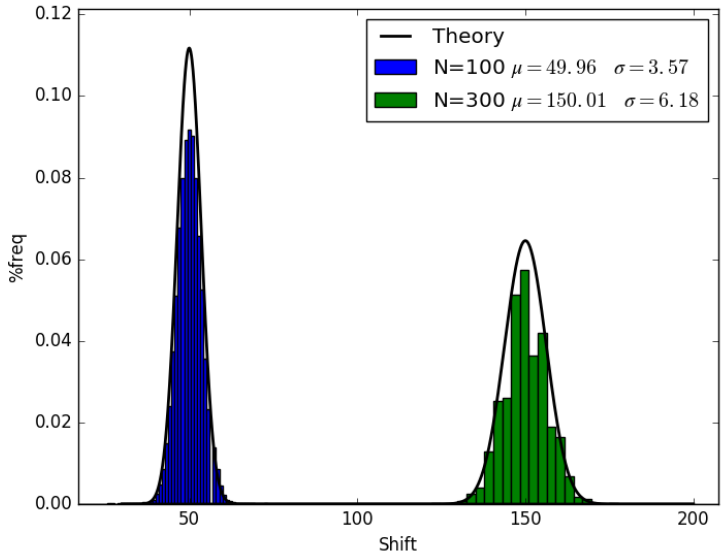


Figure 2.9: As expected the mean shift which results from the ansatz 2.60 is $N/2$.

Expanding eq. 2.57 in $t = 1/2 + \frac{\tau}{\sqrt{N}}$ in the large N limit we have:

$$\int_0^1 \left|\sigma(s,t) + \frac{1}{\sqrt{N}}\phi_t(s)\right|^p ds =$$
$$= \int_0^{1/2-\frac{\tau}{\sqrt{N}}} \left|1/2 + \frac{\tau}{\sqrt{N}} + \frac{1}{\sqrt{N}}\phi_{t+\frac{\tau}{\sqrt{N}}}(s)\right|^p + \int_{1/2-\frac{\tau}{\sqrt{N}}}^1 \left|1/2 - \frac{\tau}{\sqrt{N}} + \frac{1}{\sqrt{N}}\phi_{1/2+\frac{\tau}{\sqrt{N}}}(s)\right|^p =$$
$$= \frac{1}{2^p} + p\int_0^1 \frac{\text{sign}(1/2-s)}{2^{p-1}\sqrt{N}}\phi_{1/2}(s)ds - \frac{p\tau(\tau+\phi_{1/2}(1/2))}{2^{p-2}N}$$
$$+ \frac{p(p-1)}{2^{p-1}N}\int_0^1 (\tau+\phi_{1/2}(s))^2 ds + \tau p\int_0^1 \frac{\text{sign}(1/2-s)}{2^{p-1}N}\partial_t\phi_t(s)|_{t=1/2}ds + o(\frac{1}{N}) \tag{2.61}$$

where we have used

$$\sigma(s, 1/2+\frac{\tau}{\sqrt{N}}) = \begin{cases} 1/2+\frac{\tau}{\sqrt{N}} & s < 1/2-\frac{\tau}{\sqrt{N}} \\ 1/2-\frac{\tau}{\sqrt{N}} & s > 1/2-\frac{\tau}{\sqrt{N}} \end{cases} \tag{2.62a}$$

$$|1+x|^p = 1 + px + \frac{p(p-1)}{2}x^2 + o(x^2) \quad x \to 0 \tag{2.62b}$$

$$\int_0^\epsilon f(x)dx = \epsilon f(0) + o(\epsilon) \quad \epsilon \to 0 \tag{2.62c}$$

$$\frac{1}{\sqrt{N}}\phi_{1/2+\frac{\tau}{\sqrt{N}}}(s) = \frac{1}{\sqrt{N}}\phi_{1/2}(s) + \frac{\tau}{N}\partial_t\phi_t(s)|_{t=1/2} + o(\frac{1}{N}) \quad N \to \infty \tag{2.62d}$$

Now minimising this expression with respect to τ we find:

$$\tau = \frac{(1-2\alpha)(\phi_{1/2}(1/2)+\phi_{1/2}(0)) - \int_0^1 \phi_{1/2}(s)ds}{4\alpha - 1} \tag{2.63}$$

Remember that until this moment no average over the disorder has been done. To do this note that:

$$\begin{aligned}\langle\phi_{1/2}(s)\phi_{1/2}(t)\rangle &= \min(s,t) - st \\ &+ \min\left(s + \frac{\text{sign}(1/2-s)}{2}\right), t + \frac{\text{sign}(1/2-t)}{2}\Big) \\ &- \Big(s + \frac{\text{sign}(1/2-s)}{2}\Big)\Big(t + \frac{\text{sign}(1/2-t)}{2}\Big)\end{aligned} \tag{2.64}$$

This relation is trivial to be proved using the fundamental property of brownian Bridge processes eq. 2.42. The first term in eq. 2.61 requires

to proceed carefully in performing the average operation because of it is of order $\frac{1}{\sqrt{N}}$ and we must take into account the correction provided by the limiting Brownian Bridge distribution. Introducing the variable $\zeta(s) := s + \sigma(s, 1/2)$ we get,

$$\begin{aligned}\mathbb{P}(\phi_{1/2}(s) \in d\phi) = &\frac{\exp\Big(-\frac{\phi^2}{2s(1-s)+2\zeta(1-\zeta)}\Big)}{\sqrt{2\pi}\sqrt{s(1-s)+\zeta(1-\zeta)}} \\ &\times \Big[1 + \frac{s-\zeta}{\sqrt{N}}\Big(\frac{(s+\zeta-1)^2}{(s(1-s)+\zeta(1-\zeta))^2}\phi \\ &- \frac{1-(s-\zeta)^2-3s(1-s)-3\zeta(1-\zeta)}{3(s(1-s)+\zeta(1-\zeta))^3}\phi^3\Big) + o(\frac{1}{\sqrt{N}})\Big]\end{aligned} \tag{2.65}$$

Performing this integral using the *saddle-point method* (see Appendix B) it provides:

$$\langle \phi_{1/2}(s) \rangle = -\frac{1}{\sqrt{N}}\sigma(s, 1/2) \tag{2.66}$$

Using this result and remembering that the integral and the average operator $\langle \cdot \rangle$ commute, we can perform all averages we need. For example,

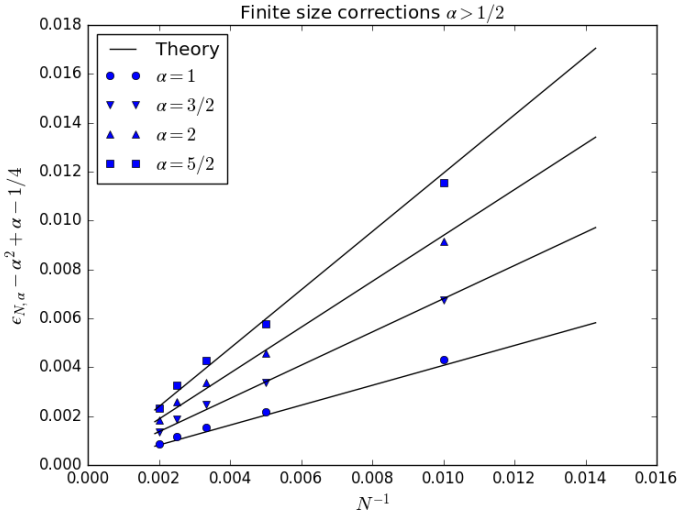


Figure 2.10: Simulation is done over 10^5 instances. The agreement with the thoeretical prediction is good and error bars are typically smaller then the markers size.

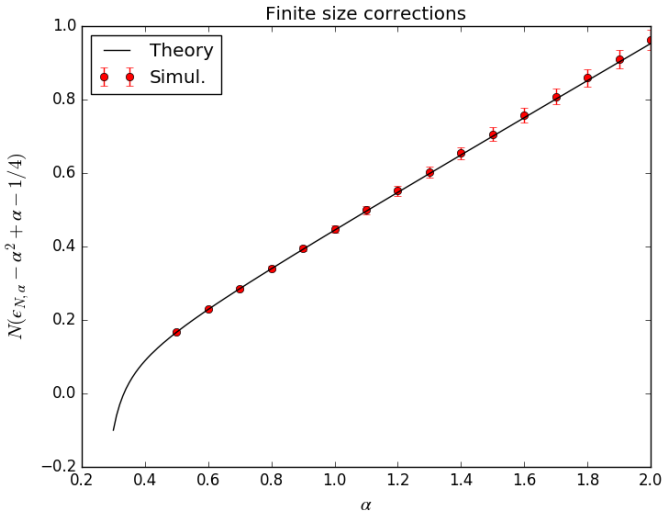


Figure 2.11: Corrections from a different point of view.

it is easy to find:

$$\int_0^1 \langle \phi_{1/2}(s) \rangle ds = 0 \tag{2.67}$$

After some integration exercise we finally get:

$$\langle \tau \rangle = 0 \tag{2.68a}$$

$$\langle \tau^2 \rangle = \frac{12\alpha^2 - 6\alpha + 1}{6(4\alpha - 1)^2} \tag{2.68b}$$

This yields,

$$\epsilon_{N,\alpha} = (\alpha - 1/2)^2 + \frac{2\alpha(3\alpha - 1)}{3(4\alpha - 1)N} + o\Big(\frac{1}{N}\Big) \tag{2.69}$$

and for the total cost,

$$E_{N,\alpha} = N\Big((\alpha - 1/2)^2 + \frac{2\alpha(3\alpha - 1)}{3(4\alpha - 1)N} + o\Big(\frac{1}{N}\Big)\Big) \tag{2.70}$$

We have verified our analytical computations and also in this case we find a good agreement between theory and simulations.

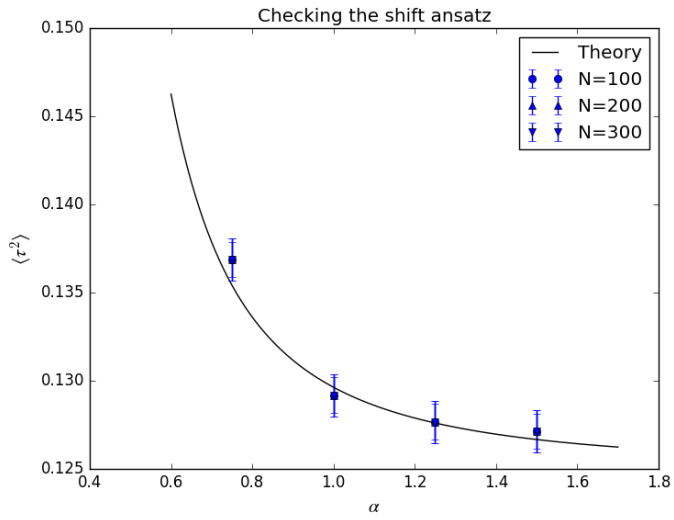


Figure 2.12: In the plot it is possible to see a good agreement for the ansatz in 2.60

2.3.3 The Correlation Function

As anticipated, when the cost function depends explicitly on the relative distance between points, euclidean correlations may arise. This quantity gives a measure of how two quantities depends by each other. In the continuum limit, the same as in the previous section, the correlation function when $\alpha < 0$ is easily computed, as in [6]

$$\begin{aligned}
\chi_N(r) &= \frac{1}{\mathcal{N}(r)} \iint_0^1 \langle \phi(s)\phi(t) \rangle \delta(|s-t|-r) ds dt = \\
&= \frac{1}{\mathcal{N}(r)} \iint_0^1 \langle (B(s,\omega) - B(s,\hat{\omega}))(B(t,\omega) - B(t,\hat{\omega})) \rangle \delta(|s-t|-r) ds dt = \\
&= \frac{1}{\mathcal{N}(r)} \iint_0^1 2(\min(s,t) - st) \delta(|s-t|-r) ds dt = \\
&= \frac{2}{\mathcal{N}(r)} \int_0^1 dt \Big[(1-t) \int_0^t s\delta(s-t+r) ds + t \int_t^1 (1-s)\delta(s-t-r) ds \Big] = \\
&= \frac{2}{\mathcal{N}(r)} \Big[\int_0^1 (t-r)\theta(t-r) dt + \int_0^1 (1-t-r)\theta(1-r-t) dt \Big] \\
&= \frac{2}{3\mathcal{N}(r)} (r-1)^3 \qquad (2.71)
\end{aligned}$$

and using the same trick of splitting the inner integral the normalisation constant is given by,

$$\mathcal{N}(r) = \iint_0^1 \delta(|s-t|-r)dsdt = 2(1-r) \tag{2.72}$$

Then the correlation function is given by:

$$\chi_N(r) = \frac{1}{3}(1-r)^2, \quad r \in \Lambda \tag{2.73}$$

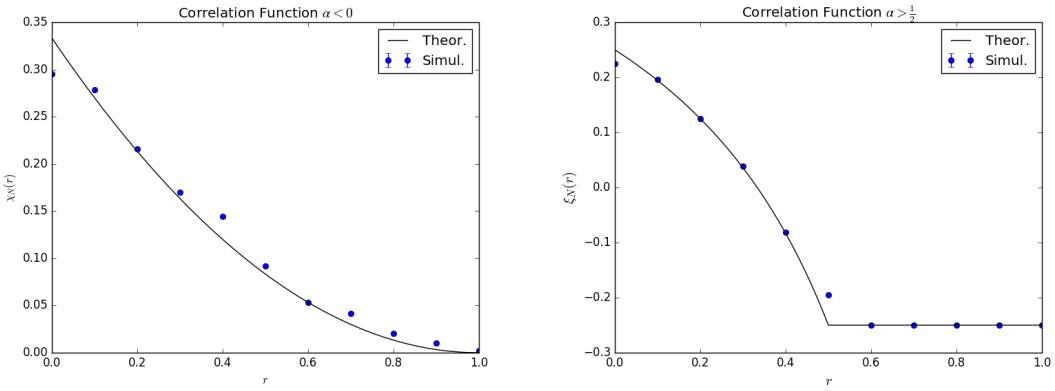
The case $\alpha > \frac{1}{2}$ is similar but not the same. Firs of all it is necessary to note that as $N \to \infty$ the matching field becomes,

$$\mu_k = y_{\pi(k)} - x_k \to \mu(s) = \begin{cases} 1/2 & 0 \leq s < 1/2 \\ -1/2 & 1/2 < s \leq 1 \end{cases} \tag{2.74}$$

Now in the same way as in eq. 2.71 we have,

$$\begin{aligned} \xi_N(r) = \frac{1}{\mathcal{N}(r)} \iint_0^1 \langle \mu(s)\mu(t) \rangle \delta(|s-t|-r)dsdt = \\ = \begin{cases} 3/4 - \frac{1}{2(1-r)} & 0 \leq r < 1/2 \\ -1/4 & 1/2 < r \leq 1 \end{cases} \end{aligned} \tag{2.75}$$

In the following numerical simulations ,again, show a good agreement



(a) (b)

Figure 2.13: Numerical check of eqs. 2.73 and 2.75.

with theoretical results. The numerical computation of the correlation function is not an easy matter. Despite of its easy and intuitive definition, when considering a single bin of a finite size, the number of instances to get a reasonable agreement with the theoretical result may be high. In this section we verified what was done in [6] about the correlation function. Correlation functions is a measure of how points in the system re-organize theirselves when a perturbation is applied. As the correlation is positive as the system respond accordingly to the perturbation. As the correlation is bigger as the system responds strongly.

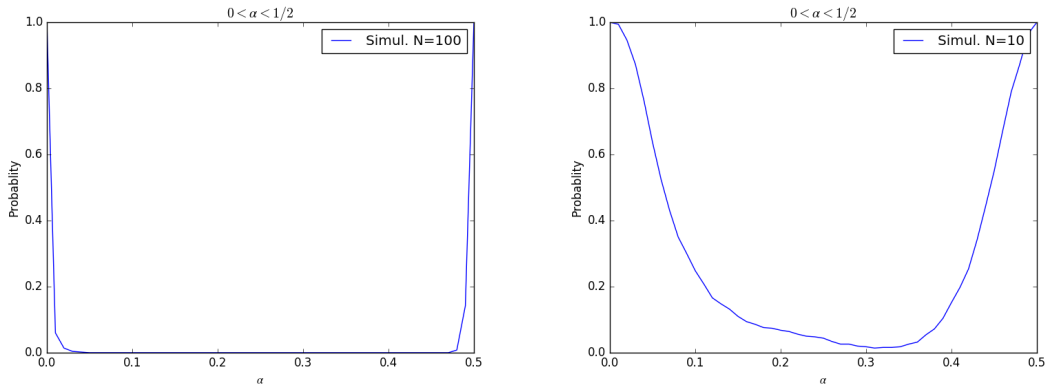
Chapter 3

Conclusions

In this thesis work we have faced the Random Euclideam Assignment Problem. Quantities of interest were average properties of the system and the correlation function. We were able to give analytical expression for these functions in dependence of the parameter of the system up to the first order term. This was possible in all regions where we knew how the optimal permutation is made, thanks to theorems of Sec. 2.2. We found that in the case $\alpha \leq 0$ there are corrections of order $\frac{1}{\sqrt{N}}$ as well as of order $\frac{1}{N}$. In the case $\alpha \geq 1/2$ corrections to the total cost do depends by $\frac{1}{N}$. Then, regarding the cost function we can conclude that we have a good control on whats going on, at least in regions where the optimal permutation is known. In the remaining interval $(0, 1/2)$ we have found an interesting information about the probability of finding as a solution a cyclic permutation: almost everywhere this probability is zero in the large N limit. We verified the appearing of a non trivial correlation for the dispacement field, due to the underlying geometry of the embedding space. The correlation functions, in each case, are positive and decreasing as points are far from the origin. Physically, this means that the system, always responds to an external perturbation accordingly to the perturbation and with less intensity as we go against the endpoint of the interval. An interesting result is obtained regarding the Hungarian algorithm. We studied its complexity and we tried to fit its universal time distibution. The ansatz we did was oversimplified and it will be interesting in further studies to analyse better the typical behavior of the algorithm.

3.1 Futher Studies

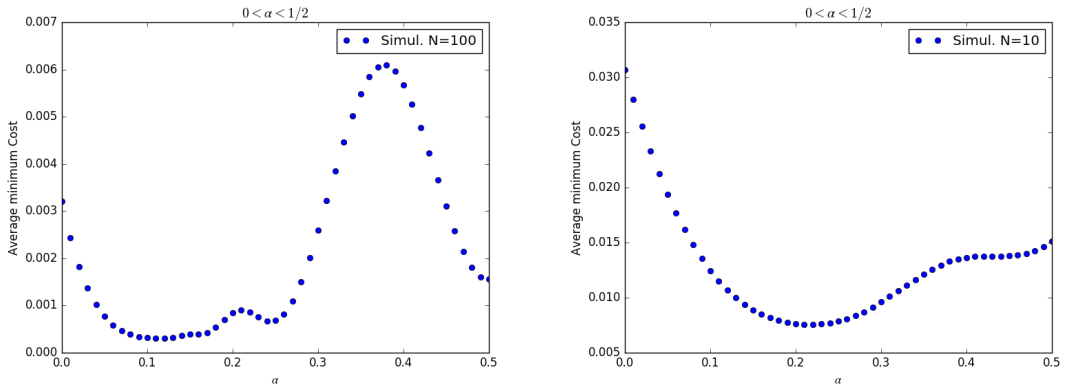
Maybe the most interesting case of this particular problem in when $0 < \alpha < 1/2$. That is because we **do not know** how the general structure of the optimal solution is, a fact that makes any attempt to do an analytical computation unsuccesful. If on one hand (the anlytical one) it is difficult to face this situation, on the other hand (the computational one) something may be exctracted by simulations. In fact, if for $\alpha < 0$ and $\alpha > 1/2$ the optimal permutation is *always* cyclic (the identity permutation is a cyclic permutation with zero shift), the first question one may ask is: what about the structure of the optimal solutions when $0 < \alpha < 1/2$? To answer this question we plotted the probability distribution of the event *"the optimal solution is a cyclic permutation"*. In the figure above we can see the form of the distribution.



(a) (b)

Figure 3.1: $\mathbb{P}$("the optimal solution is a cyclic permutation")

As we can see the the probabilty to find a cyclic permutation in this interval is almost surely zero. Going up with the size of the system the probability distribution becomes more and more smashed on zero. In addition if we try to solve for the optimal cost in this region we get:



(a) (b)

Figure 3.2: Optimal Cost ,for $N = 10$ (a) and $N = 100$ (b), in the "unkown" region $0 < \alpha < 1/2$. Note how complicated is the shape of the average minimum cost.

3.2 Acknowledgments

The resolution of the problem I choosed to discuss has not been so simply as one may think. When I faced my supervisor Caracciolo for the first time he asked me to read some articles before to getting started. After, when we met to talk about If I was sure about this study, the only thing I was able to say was: «I think this is too hard to me"». The Prof. Caracciolo answered: «Yes, it is hard, but you surely will find a way». It has been a steep stairway but ,at the end, I have effectively found a way. First of all I would like to say thank to my family, who supported me during all the period of my studies. Without their help I may not be arrived at this point of my life. The second person who deserves a great thank is my girlfriend Giulia. Even if she has not been physically present during the writing of this work, she always supported me as she could, despite miles away. Living alone, far away the family is a hard matter. A big thank goes to my friends which during these years became my second family. When I started to attend the physics department I was alone. My classmates have been of great help for my studies, mostly when I faced difficult subjects, where comparison between ideas are necessary and formative. Obviously, I would like to thank to my supervisor: I learnt more during these few months then in all my years of study. Furthermore, I would like to apologize with him for all the stress I caused. Finally, I thank Matteo d'Achille: without his help I would not have

completed my work. Thank for all the advices necessary to compute the very hard integrals and to find ,using simulations, sensitive quantities as the correlation functions.

Appendix A

Continuous Time Stochastic Processes

All phenomena have a stochastic component. Exact macroscopic laws of nature are the result of an average of a completely different microscopic behavior. A pollen grain on the surface of the water follows a random paths due to the collision with liquid molecules. The societies are regulated by complex mechanisms which we do not know exactly. Stock prices goes up and down following non-smooth curves which make them unpredictable and can cause cathastrofic economic breakdowns. Mathematics is surely an abstract subject but it gives us powerful tools to produce quantitative answers. Indeed mathematical thinking is the actual state of art of the human mind, although there are scientist who are working on different kinds of investigation tools [19]. In such a big framework Probability theory gives us the chance to understand globally what we do not understand totally. Let us recall some familiar concept on the basics. The interpretation of probability is matter of long, sometimes philosophycal, discussions that, although interesting, are not the main subject of this chapter and so we recommend [12].

Definition 11. *Given a set Ω , a family of its subsets $\mathcal{F}$ is said to be a σ-algebra on Ω if*

i) $\emptyset, \Omega \in \mathcal{F}$

ii) $A^C \in \mathcal{F} \quad \forall A \in \mathcal{F}$

iii) $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$

Definition 12. *Given a set Ω and σalgebra $\mathcal{F}$ on it, a map $\mathbb{P}\colon \mathcal{F} \to [0,1]$ such that it is normalized to unity and countably additive is called probability measure on Ω. In formulae:*

i) $\mathbb{P}(\Omega) = 1$

ii) $\mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} A_n \quad \forall \{A_n\}_{n\in\mathbb{N}}\colon \quad A_i \cap A_j = \emptyset \quad \forall i,j$

Definition 13. *Given a non empty set Ω, a probability space is a triple $(\Omega, \mathcal{F}(\Omega), \mathbb{P})$ where $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is a probability measure on Ω.*

Traditionally, in probability theory, the set Ω is called *samples space* and $\mathcal{F}(\Omega)$ is the *event space*. The structure of σ-algebra is needed to combine sample events. Note that logically speaking a set is equivalent to a proposition to which we assign a non-negative number, i.e. a probability. This fact follows immediately from *De Morgan's* laws. Combining events is then the same as combining proposition and using the probability function we ask for a measure of how much reasonably an event occurs.

Definition 14. *A map $X\colon (\Omega, \mathcal{F}(\Omega)) \to (\Omega', \mathcal{F}'(\Omega'))$ such that $X^{-1}(A') \in \mathcal{F} \quad \forall A' \in \mathcal{F}'(\Omega')$ is said to be measurable. Given a set Ω, a real-valued random varaible is a measurable map $X\colon (\Omega, \mathcal{F}(\Omega)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ where $\mathcal{B}(\mathbb{R})$ is the **Borel** σ-algebra of all open subsets of $\mathbb{R}$.*

We indicate as $< X >= \int_\Omega X d\mathbb{P}(x)$ the *mean value* of X, $\sigma_X^2 = \int_\Omega (X - \langle X \rangle)^2 d\mathbb{P}(x) = \langle X^2 \rangle - \langle X \rangle^2$ the *variance* and $\sigma_X = \sqrt{\sigma_X^2}$ the *fluctuation*. Note that the for traditional random variables as for example a gaussian one, we can formally write $d\mathbb{P}(x) = \rho_X(x)dx$ where $\rho_X(x)$ is the probability density function (PDF) of X. In fact, more precisely, the PDF is defined as a push-forward operation: by $\mathbb{P}(X^{-1}(B)) = \int_B \rho_X(x)dx \quad \forall B \in \mathcal{B}(\mathbb{R})$. The cumulative density function (CDF) is defined as $F_X(s) = \mathbb{P}(X^{-1}(-\infty, s])$. More precisely the measure $\mu(B) = \mathbb{P}(X^{-1}(B))$, which is called **law** of X, defines a measure on $\mathcal{B}(\mathbb{R})$ i.e. on the Borel σ-algebra. The relation between ρ and F is trivial but rather important: $\rho = F'$. For further reasons we anticipate that given a real-valued random variable X, its probability of being in $[x, x+dx]$ is indicated by the shorthand $\mathbb{P}(X \in dx)$. A set of n random variables is called **identically distributed** if each random variable hase the same PDF. Moreover, it is called **independent** if the **joint** density

factorizes in the product of the **cumulatives**: $\rho_X(x) = \prod_n \rho_{X_n}(x_n)$. The **covariance** between two random variables is given by $Cov(X,Y) = \langle XY \rangle - \langle X \rangle \langle Y \rangle$ and the **correlation coefficient** by $c_{(X,Y)} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$.

Proposition 4. *Given independent random variables X, Y then $Cov(X,Y) = 0$. In this case we say that variables are uncorrelated.*

An interesting quantity for our discussion is the following:

Definition 15. *Given two real-valued random variables X, Y on the same probability space $(\Omega, \mathcal{F}(\Omega), \mathbb{P})$ with PDFs ρ_X and ρ_Y, the space-shifted correlation function is defined as*

$$c_{X,Y}(r) = \int_\Omega \rho_X(x)\rho_Y(x+r)dx \quad r \in \mathbb{R} \tag{A.1}$$

An useful quantity in probability theory is the so called *characteristic function*. In practice it is the Fourier transform of the PDF of a given random variable except for the normalisation factor.

Definition 16. *Given a random variable $X \colon \Omega \to \mathbb{R}$ the characteristic function is defined as*

$$\phi_X(s) = \langle e^{isX} \rangle = \int_\Omega e^{isx}\rho_X(x)dx \tag{A.2}$$

This definition will be useful in the sebsequently discussion. The n-th moment of a random variable is defined as $\mu_n = \int_\Omega X^n d\mathbb{P}(x)$ and using Fourier transform properties it followes imemdiately that:

Proposition 5.

$$i^n \mu_n = \partial_s \phi_X(s)|_{t=0} \tag{A.3}$$

In statistical mechanics ,where disorder is present, both classical and quantum, this is atypical quantity of interest [16]. The set $\mathcal{L}$ of all real-valued random variables is an algebra i.e .a *vector space* closed with respect a product. Many results can be proved and many techniques from functional analysis provides very powerful theorems [11]. Here we develop just notions to reach our aims. There are a lot of convergence definitions in $\mathcal{L}$. The one we are interested in is:

Definition 17 (Convergence in probability)**.** *We say that a sequence X_n of random variables converges in probability to X if*

$$\mathbb{P}(|X_n - X| \geq \epsilon) \to 0 \quad \forall \epsilon > 0 \tag{A.4}$$

Let us now introduce the most important concept of this chapter. Basically speaking a stochastic process is a random variable with a parametric time dependence. Here we discuss continuous time stochastic processes, then $t \in [0, +\infty] \subset \mathbb{R}$.

Definition 18. *A real-valued stochastic process is a collection of real-valued random variables* $X_t \quad t \in \mathbb{R}$.

We consider continuous time measurable stochastic processes only, i.e. mappings $X\colon \mathcal{B}([0,+\infty]) \times \mathcal{F}(\Omega) \to \mathbb{R}$. Keeping $\omega \in \Omega$ fixed we can consider $f_X(t) = X(t,\omega)$, a family of functions called **sample functions** of X. If the sample functions are continuous $\forall \omega \in \Omega$ then the process is said to be **path-continuous**. If $X_{t+h} \to X_t$ as $h \to 0$ in probability the process is called **continuous in probability**. If a process is continuous *both* in probability and path-continuous it is said to be **continuous**. Properties of sample functions are of fundamental interest in stochastic processes study. In the figure below there is an example of Random Walks, one of the first studied random processes.

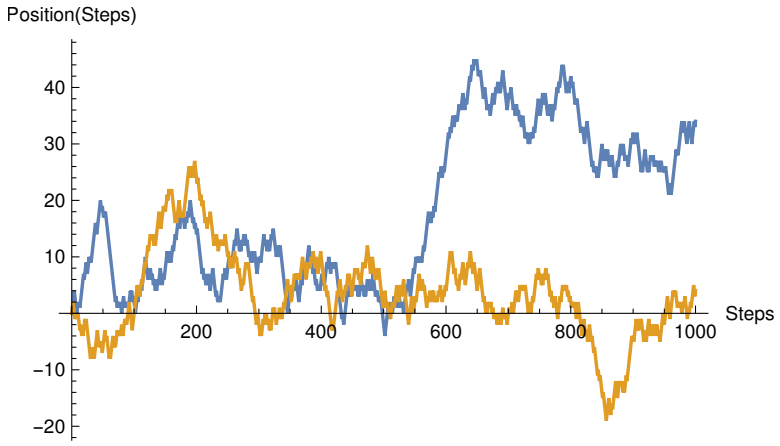


Figure A.1: A pair of Random Walks Processes with parameter $p = \frac{1}{2}$

To define precisely what a Brownian motion is we should mention **random vectors**. Random vectors are collections of random variables so one could think that a gaussian vector is a set of gaussian ones. The right definition is:

Definition 19. *An* $\mathbb{R}^n$*-valued random vector* X *is called* ***Gaussian*** *if it has multidimensional characteristic function given by:*

$$\phi_X(s) = \langle e^{is\cdot X} \rangle = e^{-\frac{1}{2}s\cdot Vs + im\cdot s} \tag{A.5}$$

where $m = \langle X \rangle$ is the mean vector and $V_{ij} = Cov(X_i, X_j)$ is a non singular $n \times n$ symmetric matrix called covariance matrix.

Definition 20. *A random vector is said to be jointly Gaussian if any linear combination $\sum_{i=0}^{n} a_i X_i$ is a Gaussian random variable for any $a_i \in \mathbb{R}$.*

Proposition 6. *If a random vector is jointly Gaussian then it is Gaussian.*

And immediately follows:

Proposition 7. *If V is a non singular symetric matrix then, the random vector*

$$X(x) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(V)}} e^{-\frac{(x-m)\cdot V^{-1}(x-m)}{2}} \tag{A.6}$$

on Ω is a Gaussian random variable.

Note that because V is symmetric one can diagonalize it and do computations in the diagonal basis so that the characteristic function reduces for normal random variables. For processes generalisations are straightforward:

Definition 21. *A Gaussian process is an $\mathbb{R}^n$-valued stochastic process with continuous time such that $(X_{t_1}, ..., X_{t_n})$ is jointly for any $t_1 \leq ... \leq t_n$. It is centered if $m_t = \langle X_t \rangle$*

Finally we have:

Definition 22. *An $\mathbb{R}^n$-valued continuous Gaussian process with mean $m_t =< X_t >$ and covariance matrix $V(s,t) = Cov(X_s, X_t) = \langle (X_s - m_s)(X_t - m_t) \rangle$ is a **Brownian Motion** or **Wiener process** if for any $t_1 \leq ... \leq t_n$ the random vectors $X_{t_0}, X_{t_{i+1}} - X_{t_i}$ are independent and the covariance matrix satisfy $V(s,t) = V(r,r)$ with $r = min(s,t)$.*

The Wiener Process is called **standard** if $m_t = 0 \ \forall t$ and $V(s,t) = \min(s,t)$. The first rigorous construction of the Brownian Motion was given by Norbert Wiener in 1923 by defining the wiener measure on the space of continuous bounded functions on an interval. We will not talk about this interesting contruction for space reasons. We refer to the bibliography for further deepening.

Here some basic properties fo the Wiener Process:

Theorem 6. *Given a Wiener Process W_t, then:*

i) ***Time Homogeneity***. $\forall t > 0 W_{t+s} - W(t+u) = W_s - W_u$

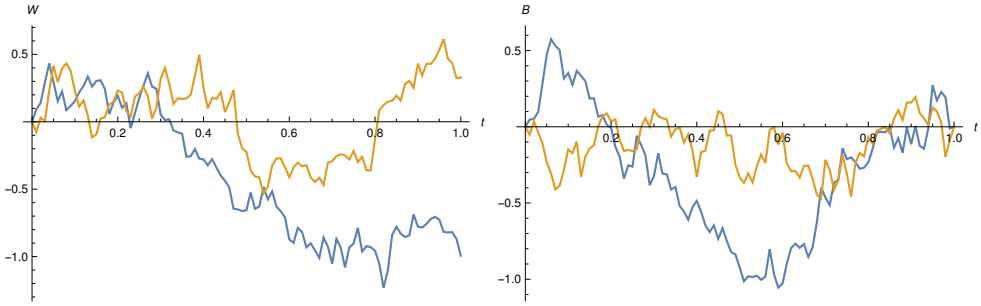
ii) ***Reflection simmetry***. *The process $-W_t$ is a Wiener process*

iii) ***Brownian scaling***. $\forall c > 0 W'_t = cW_{\frac{t}{c^2}}$ *is a Wiener process*

iv) ***Time inversion***. *The process $W_0 = 0$, $W'_t = W_{t\frac{1}{t}}$ for $t > 0$ is a Wiener process*

All of this seems to be not concerned with the existence of such a processes. In the past many authors as for example Albert Einstein showed an heuristic construction of the the Wiener Process in terms of Random Walks. We should be ensured by the following:

Theorem 7. *The Wiener Process exists.*



(a) W_t (b) B_t

Figure A.2: a) Wiener Processes and b) a Brownian Bridge Processes

Now with the existence result we can be quiete and talk about all of Wiener processes we want. There is one that is special for our purposes. This is a linear combination of Wiener processes:

Definition 23. *The process defined by $B_t = W_t - tW_1$ is called* ***Brownian Bridge***.

From the definition it is clear that a Brownian Bridge is a Wiener proces that starts at 0 and and ends at 1. In this sense it is a bridge between the two points.

Proposition 8. *Given a Brownian Bridge, if* $0 \leq s \leq 1$ *and* $0 \leq t \leq 1$ *the following holds:*

$$\langle B_s B_t \rangle = \min(s,t) - st \tag{A.7}$$

Proof. Using $\langle W_s W_t \rangle = \min(s,t)$ we have

$$\begin{aligned}
\langle B_s B_t \rangle =& \langle (W_s - sW_1)(W_t - tW_1) \rangle \\
&= \langle W_s W_t \rangle - t\langle W_s W_1 \rangle \\
&- s\langle W_t W_1 \rangle + st\langle W_1 W_1 \rangle \\
&= \min(s,t) - st - st + st \\
&= \min(s,t) - st \qquad \text{(A.8)} \\
& \qquad \text{(A.9)}
\end{aligned}$$

□

Appendix B

Saddle-point approximation

Here we give an heuristic derivation of the *saddle-point methos* in the real case. Consider the following integral:

$$I(\lambda) = \int_a^b f(x) e^{\lambda g(x)} dx \tag{B.1}$$

where f, g are two real valued functions and $a < b$. It is clear that, if g has a peaked maximum in $x_0 \in (a, b)$, the integral is dominated by the value around the point where the maximum is reached. Furthermore, suppose f and g are "regular" functions in such a way that it is possible to do a series expansion:

$$g(x) = g(x_0) - \frac{|g''(x_0)|}{2}(x - x_0)^2 + O((x - x_0)^3) \tag{B.2}$$

$$\tag{B.3}$$

Up to a change of variable, we can write,

$$I(\lambda) = e^{\lambda g(x_0)} \int_{u_a}^{u_b} \sqrt{\frac{2}{\lambda |g''(x_0)|}} f\Big(x_0 + \sqrt{\frac{2}{\lambda |g''(x_0)|}} u\Big) e^{-u^2 + R} dx \tag{B.4}$$

where $R \to 0$ as $x \to x_0$, $u_a = \sqrt{\frac{\lambda |g''(x_0)|}{2}}(a - x_0) < 0$ and $u_b = \sqrt{\frac{\lambda |g''(x_0)|}{2}}(b - x_0) > 0$. Now taking the limit $\lambda \to \infty$ and using $f(x_0 + \epsilon) = f(x_0) + \epsilon f'(x_0) + \frac{f''(x_0)}{2}\epsilon^2 + O(\epsilon^3)$ performing the gaussian integral

with standard methods we get,

$$I(\lambda) = e^{\lambda g(x_0)}\Big(f(x_0)\sqrt{\frac{2\pi}{\lambda|g''(x_0)|}}+\frac{1}{2}f''(x_0)\sqrt{\frac{2}{(\lambda|g''(x_0)|)^3}}+O(\lambda^{-5/2})\Big) \tag{B.5}$$

Bibliography

[1] Dimitri P. Bertsekas. A new algorith for the assignment problem. 1981.

[2] J. Bezanson, S. Karpinski, V. B. Shah, and A. Edelman. *julia: a fast dynamic language for technical computing*. 2012.

[3] J. Bezanson, S. Karpinski, V. B. Shah, and A. Edelman. *julia:a fresh approach to numerical computing*. 2014.

[4] J.A. Bondy and U.S.R. Murty. *Graph Theory*. Elsevier Science Ltd/North-Holland, 1993.

[5] Sergio Caracciolo, Elena Boniolo, and Andrea Sportiello. *correlation function for the grid-poisson euclidean matching on a line and on a circle*. 2014.

[6] Sergio Caracciolo, Matteo D'Achille, and Gabriele Sicuro. *euclidean matching problem in one dimension*. 2017.

[7] M. Cs˙ · *o*rgó and P. Révész. Strong approximations in probability and statistics. 1981.

[8] Victor. Dotsenko. One more discussion of the replica trick: the example of the exact solution. *Phylosophical Megazine*, 2011.

[9] T.S. Ferguson. Linear Programming. https://www.math.ucla.edu/ tom/LP.pdf.

[10] R. jonker and A. Volgenant. **??**A shortest augmenting path algorithm for dense and sparse linear assignment problems. 1987.

[11] Oliver Knill. *Probability and Stochastic Processes with Applications*. Overseas Press, 2009.

[12] A.N. Kolmogorov. *Foundations of the Theory of Probability*. Ams Chelsea Publishing, 1950.

[13] J. Komlós, P. Major, and G. Tusnády. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete. 1981.

[14] H.W. Kuhn. *the hungarian method for the assignment problem*. 1955.

[15] M. Mezárd, G. Parisi, N. Sourlas, G. Toulouse, and M. Virasoro. Replica symmetry breaking and the nature of the spin glass phase. *Journal de physique Paris*, 45(5), 1984.

[16] J. Sethna. *Entropy, Order, Parameter and Complexity*. Oxford, 1993.

[17] Gabriele Sicuro. *The Euclidean Matching Problem*. PhD thesis, Università di Pisa, 2016.

[18] M. van den Nest, W. D$\acute{}$ur, and H.J. Briegel. *completeness of the classical 2d ising model and universal quantum computation*. *Phys. Rev.*, 2008.

[19] Stephen Wolfram. *A new kind of Science*. Wolfram Research, 2002.

[20] Limin Xiang. *a formal proof of the four color theorem*. 2009.