UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE MATEMATICHE,
FISICHE E NATURALI

Corso di Laurea in Fisica

Critical Behaviour of a Marriage Problem

Relatore: Prof. Sergio Caracciolo
Correlatore: Dott. Andrea Sportiello

Tesi di Laurea di:
Elena Boniolo
Matricola: 506125

Codice PACS: 02.50.-r

Anno Accademico: 2011-2012

02.50.-r Probability theory, stochastic processes, and statistics

Contents

Introduction

Over the last decades, it has been recognised that combinatorial optimisation is connected with statistical mechanics in a natural way: if we identify the instances of the optimisation problem with the configurations of the model of statistical mechanics and the cost function with the energy, finding a solution to a problem of minimisation corresponds to finding the ground state of the thermodynamic system in the limit of zero temperature.

This analogy allows us to adapt the ideas and tools of statistical mechanics, such as the universality of the critical exponents or the techniques of the Renormalisation Group, to discuss optimisation problems and develop new algorithms. Conversely, it is possible to study optimisation problem as testing ground to devise and experiment techniques useful in the study of complex physical models, such as disordered systems (spin glasses or polymer networks, for instance, but also, in different areas, collective social behaviour or financial markets).

In this thesis, we will consider a problem known as the "Grid-Poisson Marriage". This is defined as the optimal matching between N lattice points and M random points in the continuum, taken with Poisson distribution. As we will see, it is a typical assignment problem, where the weight function is given by the sum of the Euclidean distances between the couples of points. In this case, the problem is trivial at very high or very low densities $\rho = M/N$, while the system becomes critical near $\rho = 1$, where there is the symmetry in the exchange of N with M. Owing also to the underlying geometrical structure, this is a critical phenomenon in all respects, as at $\rho = 1$, the searching process of an "ideal partner" extends to all length scales and the energy density diverges.

The importance of this model lies in its showing a non-trivial critical behaviour, and still being simpler to study than analogous problems. Indeed, in this case we can make use of a very powerful weapon, the Hungarian algorithm,

which provides us with a solution for the assignment problem in polynomial time ($O(N^3)$), so allowing us to perform numerical simulations at appreciable sizes in relatively short times.

As a final remark, we want to observe that the system is not invariant under translation, and in this sense the GPM can be used as a simple model in the study of physical disordered systems, which are in general very difficult to deal with.

This problem, and related models such as the Poisson-Poisson matching, have been studied in the past and there are already some important results (see, for instance, [20], [19], or [21]). In particular, Elia Zarinelli has discussed this subject in his Master's thesis [1], and the present work can be considered as a development of his own.

Structure of the thesis

In this thesis, we study some properties of the GPM from a theoretical point of view, and we compare them with data from numerical simulations on lattices of different sizes and with different densities of Poisson points.

Chapter 1

We recall some basic concepts of statistical mechanics and combinatorial optimisation and formalise the link between these two areas of study.

Chapter 2

We give some definitions and properties of random processes, which will be useful in the following, and we define the model under investigation.

Chapter 3

We use the properties of stochastic processes, in particular Wiener processes and Brownian bridges, to find the exact solution for the correlation function in one dimension at density $\rho = 1$.

We show that, if we consider as a weight function to minimise the sum of the squares of the distances, the analytic solution found is in very good agreement with the results from numerical simulations.

We also show the numerical curves at $\rho \neq 1$ and give a brief qualitative description of them.

Chapter 4

We state the fundamental ideas of Finite-Size Scaling of thermodynamic systems, and apply this technique to the 2-dimensonal GPM. In particular, by means of simulations at the critical point $\rho = 1$, we give a numerical estimate of the scaling of the wall-to-wall correlation function.

We show numerically the shape of the correlation function near the critical point, give two possible definitions of correlation length and examine their behaviour.

Chapter 5

We study the probability distribution of the edge lengths from simulations at different sizes and densities.

In addition, we examine the mean of these lengths as a function of the size and we verify the results presented in literature on the subject.

1

Statistical mechanics and combinatorial optimisation

1.1 Statistical mechanics and critical phenomena

In this section, we will introduce some basic concepts of statistical mechanics and give a brief account of the phenomenon of phase transitions, with particular stress on the subjects of universality classes and critical exponents.

1.1.1 Classical equilibrium statistical mechanics

The main aim of statistical mechanics is to predict the relations between the observable macroscopic properties of a mechanical system, given only a knowledge of the microscopic forces between the large number of particles composing it (as an example, the molecules of a gas or the spins of a magnet).

The normal approach of Hamiltonian mechanics in this context is unreasonable, due to the large number of degrees of freedom, and it becomes necessary to make use of the tools of probability theory.

Let us denote by $\mathcal{C}$ a generic state of the system (for example, in a ferromagnetic substance, a state is specified once the orientation of each magnetic dipole is known). Assume that the total number N of the spins is sufficiently large and that the system is at thermal equilibrium. Within these assumptions, the probability that a given configuration $\mathcal{C}$ of the system is realised, is

given by the Boltzmann law

$$P[\mathcal{C}] = \frac{e^{-\beta E(\mathcal{C})}}{Z}, \tag{1.1}$$

where $E(\mathcal{C})$ is the energy of the configuration $\mathcal{C}$, and $\beta = 1/k_B T$ (T being the absolute temperature).

The expectation value of any physical observable $\mathcal{O}$ is then expressed by the statistical average on all configurations, with weights given by the Boltzmann law

$$\langle \mathcal{O} \rangle = \frac{\sum_{\mathcal{C}} \mathcal{O}(\mathcal{C}) e^{-\beta E(\mathcal{C})}}{Z}, \tag{1.2}$$

where the quantity Z in the denominator is a normalisation factor defined by

$$Z(N, \beta) = \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})}. \tag{1.3}$$

Z is called the *partition function* of the system, while the factor $e^{-\beta E(\mathcal{C})}$ that gives the weight of each configuration is the *Gibbs factor*.

The thermodynamic quantities typical of the system at equilibrium can then be written in terms of Z. For example, for the free energy F and the internal energy U we have

$$F = -\frac{1}{\beta} \log Z \tag{1.4}$$

$$U = -T^2 \frac{\partial}{\partial T} \frac{F}{T} \tag{1.5}$$

1.1.2 Phase transitions

When we consider systems with a macroscopic number of degrees of freedom ($N \to \infty$), a phenomenon can occur which has no counterpart in both classical and quantum mechanics of finite degrees of freedom: systems ruled by the same Hamiltonian, can coexist in different phases, or undergo sudden transitions between phases as the temperature changes. This phenomenon is known as *phase transition*.

Two typical examples are the transition water-vapour at 100℃ and the magnetisation of a metallic (for instance iron) bar in the presence of an external magnetic field $\vec{B}$. If we examine the latter, we see that the magnetisation is

discontinuous at zero external field, as we are in the presence of a residual finite magnetisation M_0, and it is possible to have at the same time different regions of the sample in different magnetisation states. Moreover, if we study the trend of the magnetisation along this discontinuity as a function of the temperature, we see that it decreases as the temperature increases, until it reaches the so called *critical temperature* T_c of the material.

Both the non-analyticity of the thermodynamic quantities and the coexistence of different phases are particularly surprising. They can be understood if we suppose the system has more than one possible local equilibrium state (minimum of the free energy) and can go from a minimum to another only by sudden global changes. The points in the space of parameters where the transition occur are called *critical points* and it is customary to use, instead of the parameter T (which for historical reasons is treated as a temperature, but can be any parameter of the system), the adimensional parameter called *reduced temperature*

$$t = \frac{T - T_c}{T_c}. \tag{1.6}$$

Very often, the discontinuities of the physical observables near the critical point show a power law in their dominant part, with possibly non-integer exponent. These exponents are called *critical exponents* and have an important role in the classification of the different systems that show phase transitions. For example, the anomalous behaviour of the magnetisation is parametrised by the critical exponent β:

$$M(B = 0, t) \sim (-t)^{\beta} \quad \text{if } t \to 0^{-}.$$

One of the main peculiarities of the systems that show phase transitions is that local fluctuations present a radius of influence on the system which diverges in the proximity of the critical point. This radius is called *correlation length* and is a a typical size of the system which is a measure of how far the cooperative effects of the interaction go. Once chosen an *order parameter* σ that represents the quantity we are interested in (for example, the spin in the magnetic sample), it is possible to define a function that estimates these cooperative effects between the variables of the system. This function is known as the *(connected) correlation function*:

$$G(x, x') = \langle \sigma_x \sigma_{x'} \rangle - \langle \sigma_x \rangle \langle \sigma_{x'} \rangle. \tag{1.7}$$

Its behaviour at long distances can be written as

$$G(|x-x'|=r) \sim \begin{cases} e^{-\frac{r}{\xi}}, & T \neq T_c \\ \frac{1}{r^{d-2+\eta}}, & T = T_c \end{cases} \tag{1.8}$$

or, in one expression,

$$G(r) \sim \frac{1}{r^{d-2+\eta}} f\left(\frac{r}{\xi}\right) \tag{1.9}$$

where ξ is the correlation length, d is the dimensionality of the system, and η is a critical exponent known as the *anomalous dimension* of the order parameter. This formula involves the scaling function $f(x)$ that depends only on the dimensionless ratio $x = r/\xi$. For large x, this function has the asymptotic behaviour $f(x) \sim e^{-x}$, while its value at $x = 0$ simply fixes the normalisation of this quantity, which can always be chosen as $f(0) = 1$. It is worth stressing that the temperature enters the correlation functions only through the correlation length $\xi(T)$. The critical exponent associated with ξ is usually denoted by ν:

$$\xi(T) \sim \begin{cases} \xi_+ t^{-\nu}, & T > T_c \\ \xi_- (-t)^{-\nu}, & T < T_c \end{cases} \tag{1.10}$$

The list of critical exponents is summarised in Table 1.1.

The exponents α, β, δ, γ, η and ν defined above, are not all independent. It has been observed that they satisfy the algebraic conditions

$$\alpha + 2\beta + \gamma = 2$$

$$\alpha + \beta\delta + \beta = 2$$

$$\nu(2-\eta) = \gamma$$

$$\alpha + \nu d = 2,$$

so that it is sufficient to determine only two critical exponents in order to fix all the others. Moreover, the existence of these algebraic equations suggests a *scaling hypothesis*, namely that the thermodynamic quantities of the system are functions of B and T in which these variables enter only in homogeneous combinations (in other words, they satisfy scaling laws).

Exponent	Definition	Condition
α	$C \sim \lvert T - T_c \rvert^{-\alpha}$	$B = 0$
β	$M \sim (T - T_c)^{-\beta}$	$t < T_c,\ B = 0$
γ	$\chi \sim \lvert T - T_c \rvert^{-\gamma}$	$B = 0$
δ	$B \sim \lvert M \rvert^{\delta}$	$T = T_c$
ν	$\xi \sim \lvert T - T_c \rvert^{-\nu}$	$B = 0$
η	$G \sim r^{-(d-2+\eta)}$	$T = T_c$

Table 1.1: Definition of the critical exponents. C is the specific heat of the system and χ its magnetic susceptibility.

Another important idea in the study of critical phenomena is the *universality hypothesis*. This can be put as an hypothesis on physical grounds, which nowadays takes a stronger justification within the context of Renormalisation Group. The hypothesis states that the kind of singularity at the critical points is determined only by general properties of the configuration space and of the Hamiltonian (dimensionality of the underlying space, range of interaction, symmetry properties of the variables involved, and so on) and do not depend on the details of the interaction.

This hypothesis justifies an abstract mathematical approach to critical phenomena: the study of idealised models reveals the critical properties also of the potentially complicated concrete physical systems which share the same universality characteristics of the model, that is, which belong to the same *universality class*.

1.2 Combinatorial optimisation

Combinatorial optimisation problems are very commonly encountered in everyday life and have been studied for centuries. Think about perhaps the most famous one, the *Travelling Salesman Problem* (TSP): given a certain set of cities and the distances between them, a travelling salesman must find a tour, as short as possible, in which he visits all the cities and goes back to his starting point. What we would like to find is an algorithm which is able to find

the shortest tour for any instance of the problem in a reasonable amount of (computer) time.

By considering this example, we can define the basic ingredients of a combinatorial optimisation problem:

(1) The *domain* of the problem, which is the family of possible instances (e.g. for the TSP, any ensemble of points and matrix of distances between them).

(2) The rules which define a configuration (for the TSP, a configuration is a tour).

(3) A *cost function* which allows to compute the cost of any configuration (for the TSP, the length of the tour).

An algorithm is a sequence of elementary instructions understandable by an appropriate automated machine, such that, given some input data, in a finite number of steps it generates some output. An algorithm solves the combinatorial problem if it is able to find the configuration of lowest cost for any instance.

It is clear that any problem can be solved. It would be sufficient to naïvely list all configurations and choose the best one. However, suppose this procedure takes one second on a given machine to solve a problem with $N!$ possible configurations when $N = 20$. Then to solve a problem with $N = 40$ it will need $40!/20! = 3.35 \cdot 10^{29}$ seconds $\approx 10^{22}$ years! Obviously, this is not what we would mean by "solvable" and such an algorithm would be useless.

For this reason, a natural classification among the algorithms is according to the time they take, and the variation of this time with the size of the instance one is solving. A very coarse-grained distinction is between polynomial algorithms and exponential ones, depending on whether the computer time grows as a power of the size or exponentially.

1.2.1 Classification of optimisation problems

The classification of optimisation problems is as follows: first there are the "simple" ones which are solved by a polynomial algorithm. They form the class "P" of *polynomial* problems. A much wider class is the "NP" class of *non-deterministic polynomial* problems, which can be solved in a polynomial time by a non deterministic algorithm. A non deterministic-algorithm is, roughly speaking, an algorithm that can run in parallel on an arbitrarily large number

of processors. We say that a problem is in NP class if, given a feasible solution, the check that it is a solution can be done in polynomial time. Obviously NP$\supseteq$P, but the question whether NP$\supset$P or NP=P is still open.

Among the NP problems one can introduce an order relation. One says that problem P_1 is at least as hard as P_2 when the following statement is true: "If P_1 can be solved in a polynomial time, so can P_2". *NP complete* problems are NP problems which are at least as hard as any other NP problem. It has been shown that there exist such problems. To prove that a new problem is NP complete, it is then enough to show that it is at least as hard as one of the already known NP complete problems. Finally there are problems, even harder than NP, which are called, accordingly, *NP-hard*.

Practically, NP complete problems require a prohibitive computer time, growing exponentially with the size of the problem. Therefore one must settle for second best and look for algorithms (named heuristics) which provide an approximate solution of the problem: they find configurations which have a cost nearly equal to the optimal one. In many applications this can be enough.

1.2.2 Some examples of classical problems

Some of the most famous optimisation problems can be described in in mathematical terms with the use of graph theory[1]. Here we give a list of optimisation problems defined on a connected graph G of vertices V and weighted edges E.

The following are problems for which an algorithm has been found.

- *Minimum Cut Problem*: we want to find the cut of minimum cost for the graph G.
- *Minimum Spanning Tree Problem*: we want to find the minimum cost spanning tree subset of G.
- *Chinese Postman Problem*: we want to find the tour (closed path) of minimum length that passes through every edge at least once.
- *Eulerian Circuit*: this problem consists, given a graph, in finding if there is a circuit that visits all the edges exactly once and returns to the starting point. Such a circuit is called "Eulerian" because the problem was first discussed by Euler in 1736, while solving the famous Seven Bridges of Königsberg problem. Euler proved that in a connected graph there exists an Eulerian circuit if and only if every vertex has even degree.

[1]For some basic notions on graph theory see Appendix B.

- *The assignment problem*, which will be treated in the following 1.2.3.

On the contrary, the TSP is an example of a problem which cannot (yet) be solved in polynomial time. Other problems in this category are, for instance:

- *Maximum Cut Problem*: we want to find the cut with maximal cost for the graph G.
- *Hamiltonian Cycle*: given a graph, it consists in finding if there exists an Hamiltonian cycle (a tour that visits every edge exactly once).
- *K-Satifiability Problem*: given a set of N boolean variables and M clauses, each of them involving exactly K literals, the problem consists in finding a configuration of the variables such that every clause is satisfied. Equally important is to determine whether no such assignments exist.

1.2.3 The assignment problem

The *assignment problem* is a special case of the transportation problem, which was first formalised by the French mathematician Gaspard Monge in 1781 [5], while dealing with minimising the cost of refilling n places with surplus brash from m other places. Transportation theory deals with the study of optimal transportation and allocation of resources.

There are many ways to describe the assignment problem. A common one is to consider the problem of assigning N jobs to N workers, one each, given a set of costs $\{\epsilon_{ik}\}$ for the kth worker to perform the ith job, where the goal is to minimise the sum of all the N costs.

So, a valid assignment consists in a one-to-one mapping of jobs onto workers, that is, a permutation π of the indices k, and the cost of π is thus encoded in the cost function

$$\mathcal{H}_\epsilon(\pi) = \sum_i \epsilon_{i,\pi(i)} \tag{1.11}$$

One can give a representation of this problem in terms of graphs. Given $K_{N,N}$, the complete bipartite graph of order N, one can identify the two sets of N vertices, V_j and V_w, as the "jobs" and the "workers", and naturally assign weights $\{\epsilon_{ik}\}$ to the edges (ik) with $i \in V_j$ and $k \in V_w$. Then, a valid assignment consists of a matching M on the graph, i.e. a subset of the edge set $E = V_j \times V_w$ such that each vertex has degree one. The weight of the matching is the sum of the weights on the occupied edges.

The Hungarian algorithm

A classical algorithm for the assignment problem which finds an optimal matching in worst-case polynomial time ($O(N^3)$) is due to Harold Kuhn [6], who called it *Hungarian algorithm* as a tribute to the mathematicians, Kőnig and Egerváry, authors of previous results on which it is based.

We prefer not to go into detail about this procedure, which is a subject amply covered in literature (starting with Kuhn's original article [6]). In the following, we will use it as a black box that, for any given instance of Poisson points, will provide us with the optimal matching with the grid points (see 2.3).

1.2.4 Combinatorial optimisation and statistical mechanics at zero T

We have already explained the parallel between combinatorial optimisation and statistical physics. Here, following [7], we only want to expressly give the correspondence between the two terminologies.

Firstly, with the language of statistical mechanics we can write the partition function of the combinatorial problem at the "temperature" $T = 1/\beta$ as

$$Z = \sum_{\text{configurations}} e^{-\beta \cdot \text{cost(configuration)}} \tag{1.12}$$

Then can schematise the other relevant concepts in the following table:

OPTIMISATION	STATISTICAL MECHANICS
instance	sample
cost function	energy
optimal configuration	ground state
minimal cost	ground state energy

and analogously for the thermodynamic quantities such as the internal energy U, which can be seen as the averaged cost of the configurations weighted with their Boltzmann Gibbs probabilities, or the entropy, which corresponds to the logarithm of the number of configurations which contribute at a fixed energy.

2

The Grid-Poisson Marriage

2.1 Marriage problems

Suppose there are, for simplicity, N bachelors and N girls of "marriageable age", and we want to find the best solution in order to secure N successful marriages.

Each boy will have his own ordered list of preferences, and the same applies to the girls. If we want the marriages to last, it is necessary that, in the end, there does not exist a boy and a girl who prefer each other to their respective partners, otherwise two divorces and a new marriage are in view.

This is a representation of the so-called *stable marriage* problem, which was first introduced by Gale and Shapley in 1962 [2], and has been amply studied in literature (see, for instance, [4]). The algorithm that solves this problem is simple, and can be easily extended to similar problems, like college admissions (the difference being that colleges can accept more than a single student).

The stable marriage solution involves the presence of very "happy" people and people that, on the contrary, had to make do with a partner they may not like at all. The "global happiness" of the marriages is not necessarily the best possible, in such a way as to benefit the group on the whole.

Alternatively, suppose we have a measure of how unhappy every possible marriage would be, how can we choose the N couples, in such a way that the sum of these values is minimum? Now, that is a typical assignment problem that can be expressed in terms of weights (or costs) to minimise. As we will see, we will deal with the particular case of "marrying" points on a grid with random points in the continuum, with costs given by their Euclidean distances.

However, before we can proceed on this line, we have to give some basic concepts about random processes that will be necessary to define rigorously and study our model. For a complete coverage, we refer, for example, to [8] or [10].

2.2 Stochastic processes

The theory of stochastic processes is concerned with the study of experiments whose outcomes are random; that is, they cannot be predicted with certainty. Not only is it important to analyse a wide range of physical phenomena, but it has become essential also in many models of economics, engineering, biology and social sciences.

Examples of stochastic processes are

- The number of nuclear decays in a radioactive sample, as registered by a Geiger–Muller counter during a period of time.
- The size of an animal or human population at a certain time t, randomly fluctuating due to environmental stochasticity in birth, death, or migration.
- The number of customers who arrive during a certain time Δt at a service counter, or the number of those who are served during that period.

2.2.1 Random variables

Consider a random experiment. The collection Ω of all possible outcomes is called a *sample space*. An element ω of Ω is called an *elementary event*, or a *sample point*.

A *random variable* (r.v.) $X = X(\omega)$ is a single-valued real function that assigns a real number called the *value* of $X(\omega)$ to each sample point ω of Ω. The terminology used here is traditional. Note that a random variable is not a variable in the usual sense, and it is a function. It also customary to shorten the notation $\{\omega : X(\omega) = x\}$, say, by $\{X = x\}$.

A mapping $\vec{X} : \Omega \to \mathbb{R}^d$, $\vec{X} = (X_1, X_2, \ldots, X_d)$, is called *random vector* if for every k, $1 \leq k \leq d$, X_k is a random variable.

Let X be a random variable with cdf[1] $F_X(x)$. If $F_X(x)$ changes values only in jumps (at most a countable number of them) and is constant between jumps (that is, $F_X(x)$ is a staircase function), then X is called a *discrete* random variable. Alternatively, X is a discrete random variable only if its range contains a finite or countably infinite number of points.

Suppose the jumps in $F_X(x)$ occur at the points $x_1, x_2, \ldots$ and we assume $x_i < x_j$ if $i < j$. Then

$$F_X(x_i) - F_X(x_j) = P(X \leq x_i) - P(X \leq x_j) = P(X = x_i). \tag{2.1}$$

Let X be a random variable with cdf $F_X(x)$. If $F_X(x)$ is continuous and also has a derivative $dF_X(x)/dx$ which exists everywhere except at possibly a finite number of points and is piecewise continuous, then X is called a *continuous* random variable. Alternatively, X is a continuous random variable only if its range contains an interval (either finite or infinite) of real numbers. Thus, if X is a continuous random variable, then

$$P(X = x) = 0. \tag{2.2}$$

Note that this does not mean $X = x$ is the impossible event $\varnothing$.

2.2.2 Random processes

A *random process*, or *stochastic process*, is a family of random variables $\{X(t), t \in T\}$ indexed by the parameter t, where t varies over an index set T.

As a random variable is a function defined on the sample space Ω , a random process is really a function of two arguments, $\{X(t, \omega),\ t \in T,\ \omega \in \Omega\}$. For a fixed $t = t_k$, $X(t_k, \omega) = X_k(\omega)$ is a random variable denoted by $X(t_k)$, as ω varies over the sample space Ω. On the other hand, for a fixed sample point $\omega_i \in \Omega$, $X(t_k, \omega) = X_i(t)$ is a single function of t, called a *sample function* or a *realisation* of the process. The totality of all sample functions is called an *ensemble*.

In a random process $\{X(t), t \in T\}$, the index set is called *parameter set* of the random process. The values assumed by $X(t)$ are called *states*, and the set of all possible values forms the *state space* E of the random process. If the index set T of a random process is discrete, then the process is called a *discrete-parameter* (or *discrete-time*) process. A discrete-parameter process is

[1]For the definitions of distribution functions and their basic properties, see Appendix C.

also called a *random sequence* and is denoted by $\{X_n, n = 1, 2, \ldots\}$. If T is continuous, then we have a *continuous-parameter* (or *continuous-time*) process. If the state space E of a random process is discrete, then the process is called a *discrete-state* process, often referred to as a *chain*. In this case, the state space E is often assumed to be $\{0, 1, 2, \ldots\}$. If the state space E is continuous, then we have a *continuous-state* process.

Classification of random processes

- A random process $\{X(t), t \in T\}$ is said to be *stationary* or *strict-sense stationary* if, for all n and for every set of time instants $\{t_i \in T, i = 1, 2, ..., n\}$,

$$F_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = F_X(x_1, \ldots, x_n; t_1 + \tau, \ldots, t_n + \tau) \quad (2.3)$$

  for any τ. Hence, the distribution of a stationary process will be unaffected by a shift in the time origin, and $X(t)$ and $X(t + \tau)$ will have the same distributions for any τ. Thus,

$$F_X(x; t) = F_X(x; t + \tau) = F_X(x)$$

$$f_X(x; t) = f_X(x)$$

$$\mu_X(t) = E[X(t)] = \text{constant}$$

$$\text{var}[X(t)] = \text{constant}$$

  And similarly,

$$F_X(x_1, x_2, ; t_1, t_2) = F_X(x_1, x_2; t_2 - t_1)$$

$$f_X(x_1, x_2, ; t_1, t_2) = f_X(x_1, x_2; t_2 - t_1)$$

- If condition (2.3) only holds for $n \leq k$, then $X(t)$ is *stationary to order k*. If $k = 2$, then $X(t)$ is said to be *wide-sense stationary* or *weak stationary*, and we have $E[X(t)] = \text{constant}$.
- In a random process $X(t)$, if $X(t_i)$ for $i = 1, \ldots, n$ are independent r.v.'s, so that for $n = 2, 3, \ldots$,

$$F_X(x_1, \ldots, x_n; t_1, \ldots, t_n) = \prod_{i=1}^{n} F_X(x_i; t_i) \quad (2.4)$$

  then we call $X(t)$ an *independent random process*.

- A random process $\{X(t), t \geq 0\}$ is said to have *independent increments* if whenever $0 < t_1 < t_2 < \ldots < t_n$, $X(0)$, $X(t_1) - X(0)$, $X(t_2) - X(t_1)$, $\ldots$, $X(t_n) - X(t_{n-1})$ are independent.

- If $\{X(t), t \geq 0\}$ has independent increments and $X(t) - X(s)$ has the same distribution as $X(t+h) - X(s+h)$ for all $s, t, h \geq 0, s < t$, then the process $X(t)$ is said to have *stationary independent increments*.

  Let $\{X(t), t \geq 0\}$ be a random process with stationary independent increments and assume that $X(0) = 0$. Then

  $$E[X(t)] = E[X(1)]\ t$$

  $$\text{var}[X(t)] = \text{var}[X(1)]\ t$$

- A random process $\{X(t), t \in T\}$ is called a *normal* or *Gaussian process* if for any integer $n > 1$ and any finite sequence $t_1 < t_2 < \ldots t_n$ from T the r.v.'s $X(t_1)$, $\ldots$, $X(t_n)$ are jointly normally distributed.

  Equivalently, a stochastic process $\{X(t), t \in T\}$ is called a *Gaussian process* if every finite linear combination of the r.v.'s $X(t)$, $t \in T$, is normally distributed.

- A random process $\{X(t), t \geq 0\}$ is said to be a *Markov process* if the conditional probability distribution of future states of the process depends only upon the present state, not on the past history. That is, whenever $t_1 < t_2 < \ldots < t_n < t_{n+1}$,

  $$\begin{aligned} P\{X(t_{n+1}) \leq x_{n+1} | X(t_1) = x_1, X(t_2) = x_2, \ldots, X(t_n) = x_n\} \\ = P\{X(t_{n+1}) \leq x_{n+1} | X(t_n) = x_n\} \end{aligned} \tag{2.5}$$

  A discrete-state Markov process is called a *Markov chain*. For a discrete-parameter Markov chain $\{X_n, n \geq 0\}$, we have for every n

  $$P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_n = i) = P(X_{n+1} = j | X_n = i) \tag{2.6}$$

  Equation (2.5) or Eq. (2.6) is referred to as the *Markov property* (which is also known as the *memoryless property*).

  It is possible to show (see e.g. [8]) that every stochastic process $X(t)$, $t \geq 0$, with independent increments has the Markov property.

Random walks

An important example of Markov chain is is the random walk.

Let $\{J_n, n \geq 1\}$ be a sequence of independent identically distributed (i.i.d.) r.v.'s taking values in the d-dimensional Euclidean space $\mathbb{R}^d$, and X_0 a fixed vector in $\mathbb{R}^d$. The stochastic process $X = \{X_n, n \geq 0\}$ defined by

$$X_n = X_0 + J_1 + \cdots + J_n, \qquad n \geq 1, \tag{2.7}$$

is called a *d-dimensional random walk*. If the vector X_0 and the r.v.'s J_n take values in $\mathbb{Z}^d$, then $\{X_n\}$ is called a *d-dimensional lattice random walk*. In the lattice walk case, if we allow only the jumps J_n from $\vec{x} = (x_1, \ldots, x_d)$ to $\vec{y} = (x_1 + \epsilon_1, \ldots, x_d + \epsilon_d)$ where $x \in \mathbb{Z}^d$ and e_k = -1 or 1, $1 \leq k \leq d$, then the corresponding walk is called a *simple random walk*. If each of the $2d$ moves at any given jump in a simple random walk occurs with equal probability $p = (1/2d)$, then X is called a *symmetric random walk*. In all these cases, if the jumps J_n are only independent but not necessarily identically distributed, then X is called a *nonhomogeneous random walk*.

A picturesque way of thinking of a 2-dimensional random walk is by imagining a drunkard walking randomly in an idealised infinite city. The city is arranged in a square grid, and at every step, the drunkard chooses one of the four possible directions with equal probability.

The scaling limit of a random walk in dimension 1 is a Wiener process. This means that a random walk with a large number of very small steps is an approximation to a Wiener process. Wiener processes will be important in the next chapter (Chap. 3) and will be further discussed there.

Poisson processes

Let t represent a time variable. Suppose an experiment begins at $t = 0$. Events of a particular kind occur randomly, the first at T_1, the second at T_2, and so on. The r.v. T_i denotes the time at which the ith event occurs, and the values t_i of $T_i (i = 1, 2, \ldots)$ are called *points of occurrence*.

Let

$$Z_n = T_n - T_{n-1} \tag{2.8}$$

and $T_0 = 0$. Then Z_n denotes the time between the $(n-1)$st and the nth events. The sequence of ordered r.v.'s $\{Z_N, n \geq 0\}$ is sometimes called an

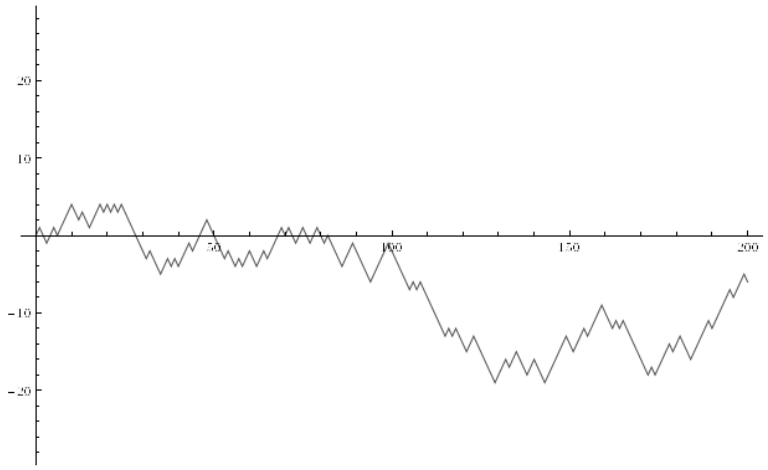

Figure 2.1: An example of a simple random walk of 200 steps in dimension 1.

interarrival process. If all r.v.'s Z_n are independent and identically distributed, then $\{Z_N, n \geq 0\}$ is called a *renewal process* or a *recurrent process*. From Eq. (2.8), we see that

$$T_n = Z_1 + Z_2 + \cdots + Z_n \tag{2.9}$$

where T_n denotes the time from the beginning until the occurrence of the nth event. Thus, $\{T_n, n \geq 0\}$ is sometimes called an *arrival process*.

A continuous-time stochastic process $X(t), t \geq 0$ with values in the state space $\Omega = (0, 1, 2, \dots)$ is called a *counting process* if $X(t)$, for any t, represents the total number of "events" that have occurred during the time period $[0, t]$.

A counting process $\{X(t), t \geq 0\}$ is said to be a *Poisson process* with rate $\lambda > 0$ if:

1. $X(0) = 0$
2. $X(t)$ is a process with independent increments
3. the number of events in any interval of length t is Poisson distributed with rate λt, that is, for all $s, t \geq 0$,

$$P\{X(t+s) - X(s) = x\} = e^{-\lambda t}\frac{(\lambda t)^x}{x!}, \qquad x = 0, 1, 2, \dots \tag{2.10}$$

Equivalently[2], a counting process $\{X(t), t \geq 0\}$ is said to be a *Poisson process* with rate $\lambda > 0$ if:

[2]For a proof, see e.g. [8]

1. $X(0) = 0$
2. $X(t)$ is a process with independent and stationary increments
3. the following relations hold:

$$P\{X(t+h) - X(t) = 1\} = Xh + o(h) \tag{2.11}$$

$$P\{X(t+h) - X(t) \geq 2\} = o(h), \tag{2.12}$$

(where a function $f(x)$ is said to be of order $o(h)$ if $\lim_{h\to 0} f(h)/h = 0$.)

It is possible to show that if $\{X(t), t \geq 0\}$ is a Poisson process with rate λ, the corresponding sequence of successive interarrival times $\{t_n, n \geq 1\}$ are independent identically distributed r.v.'s obeying an exponential density with mean λ^{-1}. The proof is given, e.g., in [8], but it is straightforward to see that the probability law for first interarrival time is

$$P\{t_1 > \tau\} = P\{X(\tau) = 0\} = e^{-\lambda} \tag{2.13}$$

as the event $\{t_1 > \tau\}$ occurs if and only if no Poisson event has occurred in the interval $[0, t]$.

2.2.3 Point processes

The definition of our model relies on the concept of point process on a compact subset of $\mathbb{R}^n$.

Let E be a subset of $\mathbb{R}^n$. We assume[3] that $\{X_n, n \geq 0\}$ are random elements of E, which represent points in the state space E. Next, we define an indicator random variable 1_{X_n} by

$$1_{X_n}(A) = \begin{cases} 1, & \text{if } X_n \in A \\ 0, & \text{if } X_n \notin A \end{cases} \tag{2.14}$$

Note, therefore, that 1_{X_n} is a function whose domain is the subsets of E, and whose range is $\{0, 1\}$, and that it takes the value one whenever X_n is in the subset of interest. Other common notations for an indicator random variable are I and χ.

[3] Here, we are following [12]

Next, we note that by taking the sum over n, we find the total number of the points $\{X_n\}$ contained in the set A. Therefore, we define the counting measure N by

$$N := \sum_n 1_{X_n} \tag{2.15}$$

so that for $A \subset E$,

$$N(A) := \sum_n 1_{X_n}(A) \tag{2.16}$$

gives the total number of points in $A \subset E$.

The function N is called a *point process*, and $\{X_n\}$ are called the *points*. If the X_n's are almost surely distinct, then the point process is known as *simple*.

We note that as N depends explicitly on the values of the points, X_n, it is natural to call such an object a random measure.

We will make the running assumption that bounded regions of A must always contain a finite number of points with a probability of one. That is, for any bounded set A, $P\{N(A) < \infty\} = 1$.

The simplest example of a point process (and the one that will be used in our work) is the Poisson point process, which is a spatial generalisation of the Poisson process described above. Namely, we say that a point process N is a (homogeneous) *Poisson point process* or a *Poisson random measure* if the joint distributions of the counts $N(A_i)$ on bounded disjoint sets A_i satisfy

$$P[N(A_i) = k_i, i = 1, \dots n] = \prod_i e^{-\lambda\mu(A_i)} \frac{(\lambda\mu(A_i))^{k_i}}{k_i!}. \tag{2.17}$$

where $k_1, \dots, k_n$ are non-negative integers and μ denotes the Lebesgue measure. The constant λ is called the *intensity* or *rate* of the Poisson point process. An inhomogeneous Poisson point process is defined as above but by replacing $\lambda\mu(A_i)$ with $\int_{A_i} \lambda(x)dx$ where λ is a non-negative function on $\mathbb{R}^d$.

2.3 The model

Consider the hypercube[4] $[0, L]^d \subset \mathbb{R}^d$, with $L \in \mathbb{Z}^+$, with the Euclidean distance $\text{dist}(a, b) = [(x_1^b - x_1^a)^2 + (x_2^b - x_2^a)^2 + \ldots + (x_d^b - x_d^a)^2]^{1/2}$.

[4]Note that, in this work, d will either be 1 or 2, therefore we will consider only intervals in $\mathbb{R}$ or squares on the plane.

We will call *grid points* the discrete subset of points of the hypercube, defined by

$$\mathcal{G} = \{(i_1 - 0.5,\ i_2 - 0.5,\ \ldots,\ i_d - 0.5) \in [0, L]^d, \text{with } i_k = 1, 2, \ldots, L\} \quad (2.18)$$

The number of grid points is $N = |\mathcal{G}| = L^d$.

Let $\mathfrak{P}$ be a simple point process of finite intensity ρ in $[0, L]^d$. The support of $\mathfrak{P}$ is the discrete random set

$$\mathcal{P} := \{\vec{x} \in \mathbb{R}^d : \mathfrak{P}(\{\vec{x}\}) = 1\}. \quad (2.19)$$

We define $\mathcal{P}$ as an instance of *Poisson points*. The coordinates of the Poisson points are independent and identically distributed random variables with uniform distribution in $[0, L]$.

We define M as the number of Poisson points, $M = |\mathcal{P}|$.

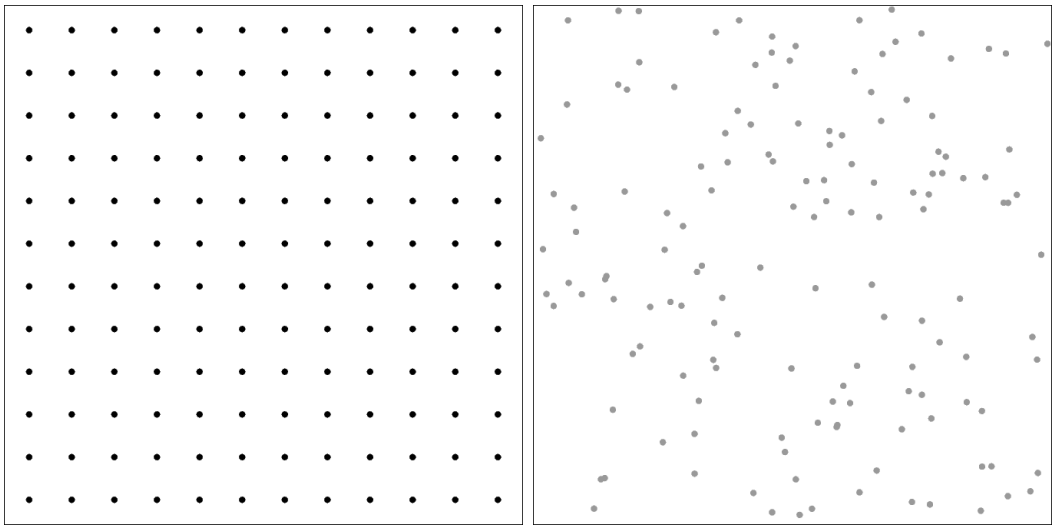

Figure 2.2: Left: grid points on the square $[0, L] \times [0, L]$, with $L = 12$. Right: an example of an instance of 144 Poisson points on the same square.

Given an instance of Poisson points, we define a *marriage* between grid and Poisson points as a function $\pi : \mathcal{G}' \subset \mathcal{G} \longrightarrow \mathcal{P}$, with $|\mathcal{G}'| = \min(N, M)$, that matches a grid point to a Poisson point, in such a way that every Poisson point is "married" to no more than one grid point. In the terminology of graph theory, a marriage is a maximum bipartite matching between $\mathcal{G}$ and $\mathcal{P}$ (i.e. a maximum matching of $\mathcal{G} \cup \mathcal{P}$ where all the edges are from $\mathcal{G}$ to $\mathcal{P}$).

We define a weight function on the edges of the matching as the length of the edge. The *energy* associated with a marriage π is then defined as the sum of the distances between matched pairs:

$$\mathcal{H}_{\mathcal{P}}(\pi) := \sum_{i=1}^{N} \mathrm{dist}(i, \pi(i)) \tag{2.20}$$

We will call π_{opt} the marriage with minimum energy and we define the *energy* (or *cost*) of an instance of Poisson points as the energy of π_{opt}

$$\mathcal{H}(\mathcal{P}) := \mathcal{H}_{\mathcal{P}}(\pi_{\mathrm{opt}}) = \min_{\pi} \mathcal{H}_{\mathcal{P}}(\pi) \tag{2.21}$$

If the instance has a single marriage π with minimum energy, we say that it is *non-degenerate*, otherwise it is *degenerate*.

Again, if we consider the complete bipartite, this time weighed, graph $K_{N,M}$, with $V(K_{N,M}) = \mathcal{G} \cup \mathcal{P}$ and weight function $w(i,j) = \mathrm{dist}(i,j)$, what we want to find is the *optimal maximum matching* of $K_{N,M}$.

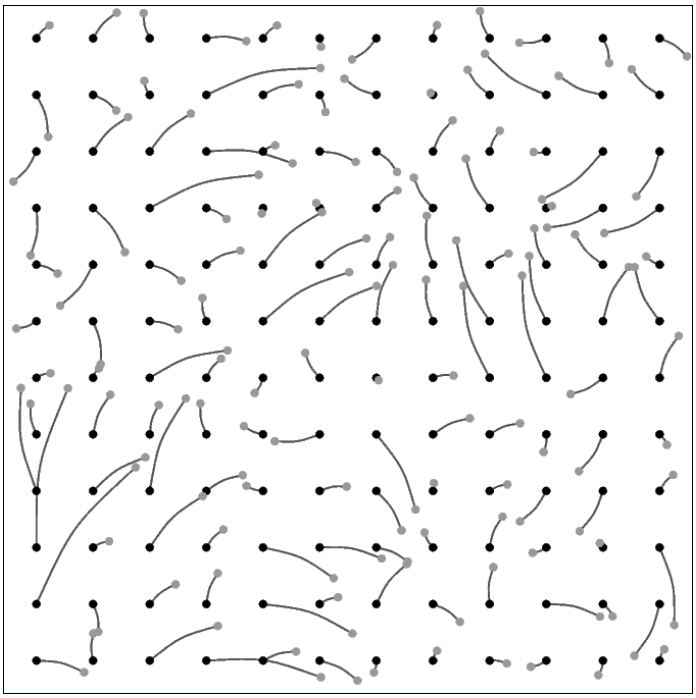

Figure 2.3: The optimal marriage for the points in Fig. 2.2.

On the other hand, in the lexicon of combinatorial optimisation, the GPM is a particular case of assignment problem. The space of possible configurations

is the set of all possible marriages (if $|\mathcal{G}| = |\mathcal{P}| = N$, the number of possible marriages is $N!$), and the cost function to minimise is the energy $\mathcal{H}_{\mathcal{P}}(\pi)$.

However, in this case, the costs ϵ_{ij} are not independent from each other, as they were in the problem job-workers, and they cannot, for example, be chosen randomly, but are subject to geometrical constraints. It is clear that, if A, B an C are three points on the plane, then $\text{dist}(A, B)$ cannot be independent from $\text{dist}(A, C)$ and $\text{dist}(C, B)$. These constraints make it difficult to perform statistical averages and prevent us from using powerful tools like mean field theory or the cavity method. And not only does this make the analysis more difficult, but also changes the characteristics of the problem. In the random assignment problem, it has been shown [7] that as $N \to \infty$ the energy becomes a constant, while, as we will see in Chap. 5, in the GPM (like in the Poisson-Poisson marriage) the energy grows faster than the number of couples.

3

The marriage problem in one dimension

3.1 Exact solution for the correlation function at density one

We want to describe the characteristics of the optimal marriage on a linear lattice at density $\rho = 1$.

In our numerical simulations, we can only have a discrete system with a finite number of points. However, for simplicity of calculation, we deal with the problem in the continuum limit, that is we imagine to take the lattice spacing to zero while holding the density fixed. The real system will hopefully tend to this model as the number of points increases.

With this premise in mind, we start by introducing some mathematical tools we will need in our analysis.

3.1.1 Definitions

Wiener process

A standard one-dimensional *Wiener process* (also called *Brownian motion process*, as it was first introduced to describe the natural phenomenon of the same name) is a stochastic process $W(t)$: $t \in \mathbb{R}, t \geq 0$, with the following properties:

(1) $W(0) = 0$

(2) The function $t \to W(t)$ is almost surely continuous

(3) The process $W(t)$ has stationary, independent increments

(4) The increment $W(t) - W(s)$ is normally distributed with expected value 0 and variance $t - s$

Basic properties of the Wiener process

- $W(t)$ is a *Gaussian process*, that is for all n and times $t_1, \ldots, t_n$, the linear combination of $W(t_1), \ldots, W(t_n)$ is normally distributed
- The unconditional probability density function at a fixed time t is given by

$$p_{W(t)}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \tag{3.1}$$

- $\forall t$, the expectation is zero:

$$E[W(t)] = 0 \tag{3.2}$$

- The variance:

$$\text{var}[W(t)] = E[W^2(t)] - E^2[W(t)] = E[W^2(t)] = t \tag{3.3}$$

- The covariance[1]:

$$\text{cov}[W(s), W(t)] = \min(s, t) \tag{3.4}$$

The area of a Wiener process, defined by

$$W^{(-1)}(t) := \int_0^t \mathrm{d}s\, W(s), \tag{3.5}$$

[1] To see this, let us suppose $s \leq t$. Then

$$\begin{aligned}
\text{cov}[W(s), W(t)] &= E[(W(s) - E[(W(s)]) \cdot (W(t) - E[(W(t)])] \\
&= E[W(s) \cdot W(t)] = E[W(s) \cdot ((W(t) - W(s)) + W(s))] \\
&= E[W(s) \cdot (W(t) - W(s))] + E[W^2(s)] = s
\end{aligned}$$

is itself a Wiener process (as a linear combination of Wiener processes) characterised by its expected value and variance:

$$E[W^{(-1)}(t)] = \int_0^t \mathrm{d}s\, E[W(s)] = 0 \tag{3.6}$$

$$\begin{aligned}\mathrm{var}[W^{(-1)}(t)] &= E\left[\int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' W(s)W(s')\right] = \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \,\mathrm{cov}(W_s, W_s') \\ &= \int_0^t \mathrm{d}s \left(\int_0^s \mathrm{d}s' \,\min(s,s') + \int_s^t \mathrm{d}s' \,\min(s,s')\right) = \frac{t^3}{3}.\end{aligned} \tag{3.7}$$

Brownian bridge

A standard *Brownian bridge* $B(t)$ over the interval $[0,1]$ is a standard Wiener process conditioned to have $B(1) = B(0) = 0$.
Now, if we have a Wiener process $W(t)$, the linear combination

$$B(t) := W(t) - tW(1) \tag{3.8}$$

is a Brownian bridge with expectation, variance and covariance:

$$E[B(t)] = 0 \tag{3.9}$$

$$\begin{aligned}\mathrm{var}[B(t)] &= E[(W(t) - tW(1))^2] \\ &= E[W^2(t)] - 2tE[W(1)\cdot W(t)] + t^2 E[W^2(1)] \\ &= t(1-t)\end{aligned} \tag{3.10}$$

$$\begin{aligned}\mathrm{cov}[B(t), B(s)] &= E[(W(t) - tW(1))\cdot(W(s) - sW(1))] \\ &= \min(t,s) - ts\end{aligned} \tag{3.11}$$

The area of a Brownian bridge, defined by

$$B^{(-1)}(t) := \int_0^t \mathrm{d}s\, B(s), \tag{3.12}$$

is, again, a Gaussian variable (as a linear combination of Wiener processes) characterised by its expected value and variance:

$$E[B^{(-1)}(t)] = \int_0^t \mathrm{d}s\, E[B(s)] = 0 \tag{3.13}$$

$$\begin{aligned}\mathrm{var}[B^{(-1)}(t)] &= \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \ \mathrm{cov}[B(s), B(s')] \\ &= \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \ (\min(s, s') - ss') \\ &= \frac{t^3}{3} - \frac{t^4}{4}. \end{aligned} \tag{3.14}$$

In particular, if $t = 1$,

$$\mathrm{var}[B^{(-1)}(1)] = \frac{1}{12} \tag{3.15}$$

and the covariance between $B^{(-1)}(1)$ and $B(t)$ is

$$\begin{aligned}\mathrm{cov}[B^{(-1)}(1), B(t)] &= \int_0^1 \mathrm{d}s \ \mathrm{cov}[B(s), B(t)] = \int_0^1 \mathrm{d}s \ (\min(s, t) - st) \\ &= \frac{1}{2} t(1 - t). \end{aligned} \tag{3.16}$$

Let us consider two Brownian bridges, $B(s)$ and $B(t)$, and let us assume that $s \leq t$. The covariance matrix is then

$$C = \begin{pmatrix} s(1-s) & s(1-t) \\ \\ s(1-t) & t(1-t) \end{pmatrix}. \tag{3.17}$$

The distribution is Gaussian, which means the density function is given by

$$p_A(x_1, x_2) = \sqrt{\det A} \ \frac{e^{-\frac{1}{2}\sum_{i=1}^{2} x_i A_{ij} x_j}}{2\pi}, \tag{3.18}$$

with (see, for example, [9])

$$A = C^{-1} = \begin{pmatrix} \frac{t}{s(t-s)} & -\frac{1}{t-s} \\ \\ -\frac{1}{t-s} & \frac{1-s}{(1-t)(t-s)} \end{pmatrix}. \tag{3.19}$$

3.1.2 Open boundary conditions

In our problem, we consider a linear lattice of size L and parameter 1, or, equivalently, a linear lattice of size 1 and parameter $1/L$, and we generate L random points uniformly distributed on the interval $[0, L]$ (equivalently, $[0, 1]$).

In dimension one, for open boundary conditions, the optimal marriage is a forced choice, as the first point of the grid must clearly be matched with the first random point, the second with the second and so on. That is to say, for all n, the nth point of the grid will be matched with the nth random point and, as the random points are uniformly distributed, the (1-dimensional) vector connecting the couple will be a random variable with zero mean and variance $= n$.

In the limit $L \to \infty$, this situation can be represented by a Wiener process $W(t)$. More precisely, as there are no points for $t > 1$, we can describe it as a Brownian bridge $B(t)$ over the interval $[0, 1]$.

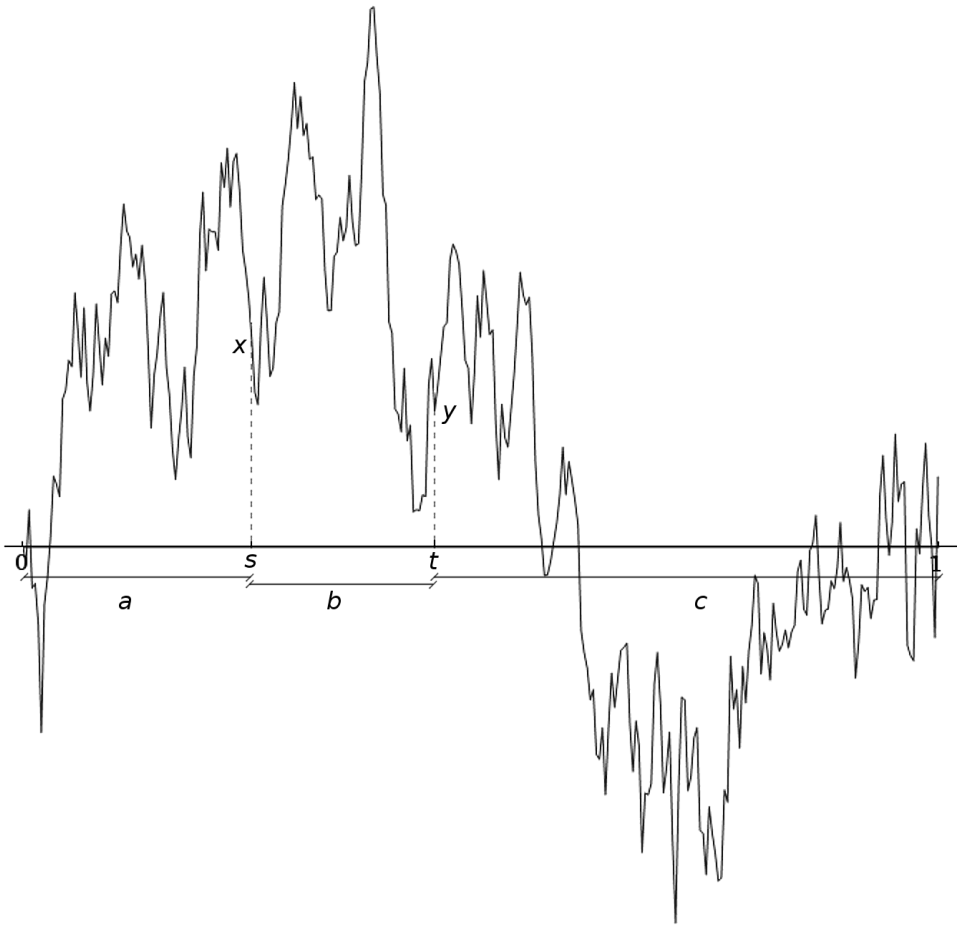

Figure 3.1: An example of Brownian bridge $B(t)$ over the interval $[0, 1]$ with intermediate times s and t

Now, let us consider two intermediate times, s and t, with $0 < s < t < 1$. The probability that the process started at the origin arrives at x after a time

s is Gaussian with zero mean and variance $= s$:

$$p_{W(s)}(x) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}}. \tag{3.20}$$

Similarly, to move from x to y in the interval $(t-s)$:

$$p_{W(t-s)}(y-x) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(y-x)^2}{2(t-s)}} \tag{3.21}$$

and, finally, to move from y to 0 in the interval $(1-t)$:

$$p_{W(1-t)}(y) = \frac{1}{\sqrt{2\pi(1-t)}} e^{-\frac{y^2}{2(1-t)}}. \tag{3.22}$$

By a change of parameters, if we consider the three segments of length

$$\begin{aligned} a &= s \\ b &= t-s \\ c &= 1-t, \end{aligned} \tag{3.23}$$

the matrix A in (3.19) becomes

$$A = \begin{pmatrix} \frac{1}{a}+\frac{1}{b} & -\frac{1}{b} \\ -\frac{1}{b} & \frac{1}{c}+\frac{1}{b} \end{pmatrix}. \tag{3.24}$$

In general, since the distribution is Gaussian, we know from (3.18) that

$$\frac{\sqrt{\det A}}{2\pi} \int\int \mathrm{d}x\, \mathrm{d}y\, e^{-\frac{x^2}{2a}-\frac{(x-y)^2}{2b}-\frac{y^2}{2c}} = 1, \tag{3.25}$$

with

$$\det A = \frac{a+b+c}{abc}. \tag{3.26}$$

The joint probability distribution for the random variables x and y is then

$$p_{a,b,c}(x,y) = \sqrt{2\pi}\,\sqrt{a+b+c}\,\frac{e^{-\frac{x^2}{2a}-\frac{(x-y)^2}{2b}-\frac{y^2}{2c}}}{\sqrt{2\pi a}\,\sqrt{2\pi b}\,\sqrt{2\pi c}}. \tag{3.27}$$

Correlation function

Given an instance of Poisson points and found the optimal marriage as seen, we can define the quantity $\varphi(t)$ as the distance between the grid point in t and the Poisson point associated to it by the marriage. If we define a correlation function between two point s and t as

$$G_1(s,t) = \frac{\varphi(s)\cdot\varphi(t)}{|\varphi(s)|\cdot|\varphi(t)|}, \tag{3.28}$$

its value is obviously

$$G_1(s,t) = \operatorname{sgn}(\varphi(s))\cdot\operatorname{sgn}(\varphi(t)) = \operatorname{sgn}(\varphi(s)\cdot\varphi(t)). \tag{3.29}$$

Therefore, the quantity we want to calculate is the expected value of the sign function under the measure (3.27). That is, with the substitution (3.23),

$$\begin{aligned} G_2(a,b,c) &= \int\int \mathrm{d}x\,\mathrm{d}y\, p_{a,b,c}(x,y)\operatorname{sgn}(x\cdot y) \\ &= \int\int \mathrm{d}x\,\mathrm{d}y\,\sqrt{2\pi}\,\sqrt{a+b+c}\,\frac{e^{-\frac{x^2}{2a}-\frac{(x-y)^2}{2b}-\frac{y^2}{2c}}}{\sqrt{2\pi a}\,\sqrt{2\pi b}\,\sqrt{2\pi c}}\,\operatorname{sgn}(x\cdot y) \end{aligned} \tag{3.30}$$

If we define

$$\alpha(a,b,c) := \int_{x\geq 0}\int_{y\geq 0} \mathrm{d}x\,\mathrm{d}y\; p_{a,b,c}(x,y) \tag{3.31}$$

and

$$\beta(a,b,c) := \int_{x\geq 0}\int_{y\leq 0} \mathrm{d}x\,\mathrm{d}y\; p_{a,b,c}(x,y), \tag{3.32}$$

then

$$G_2(a,b,c) = 2\alpha(a,b,c) - 2\beta(a,b,c). \tag{3.33}$$

In addition, since $p_{a,b,c}(x,y)$ is Gaussian, we know that

$$2\alpha(a,b,c) + 2\beta(a,b,c) = 1, \tag{3.34}$$

then

$$G_2(a,b,c) = 4\alpha(a,b,c) - 1. \tag{3.35}$$

By performing the integral (3.31), we find

$$\alpha(a,b,c) = \frac{1}{4} + \frac{1}{2\pi}\arctan\sqrt{\frac{ac}{b(a+b+c)}}, \tag{3.36}$$

and then

$$G_2(a,b,c) = \frac{2}{\pi}\arctan\sqrt{\frac{ac}{b(a+b+c)}}, \tag{3.37}$$

with, in this case,

$$\begin{aligned} &a,\ b,\ c \geq 0 \\ &a+b+c = 1. \end{aligned} \tag{3.38}$$

If we keep the distance b between the two points constant, and we calculate the mean over the interval $[0,1]$, we finally obtain

$$G_{obc}(b) = \frac{2}{\pi}\ \frac{1}{1-b}\ \int_0^{1-b} \mathrm{d}a\ \arctan\sqrt{\frac{a(1-a-b)}{b}} = \frac{1-\sqrt{b}}{1+\sqrt{b}}. \tag{3.39}$$

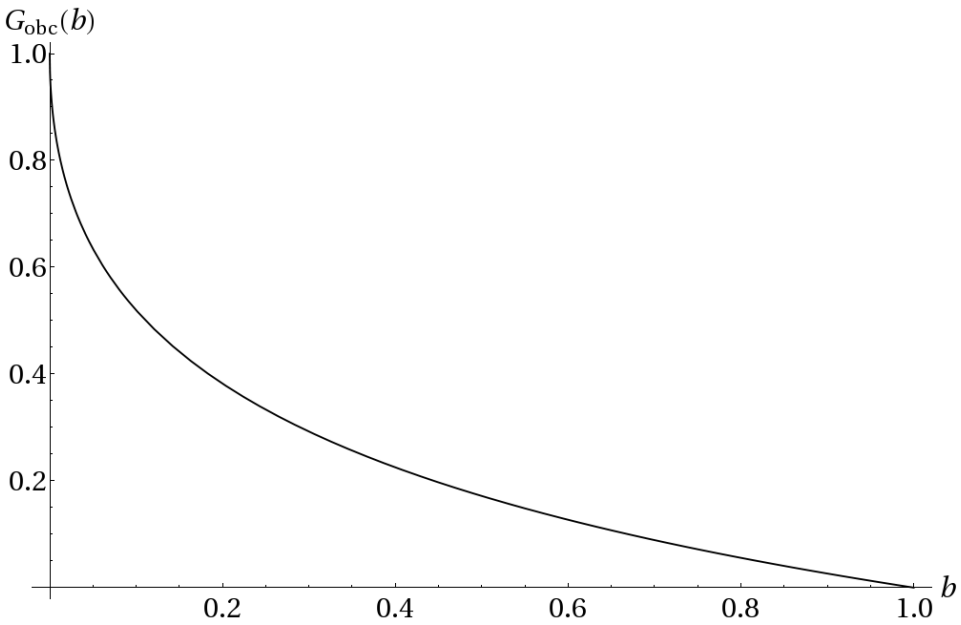

Figure 3.2: The theoretical correlation function in one dimension for open boundary conditions.

3.1.3 Periodic boundary conditions

In the periodic problem, we can still represent our process with a Brownian bridge, but this time we do not know to which random point the first grid point is matched. However, we do know that once the first couple is formed, all others follow necessarily in the same way as the previous case.

This corresponds to a situation where the first point is not mapped to zero, but to some constant k. That is to say, a Brownian bridge with initial and final value k: $B(0) = B(1) = k$.

We choose to parametrise the family of the possible processes of this kind by the area under their path. We therefore take as our new variable:

$$B^{(-1)}(1) - B(t). \tag{3.40}$$

This is, again, a Gaussian random variable with zero expected value. Its covariance can be easily calculated as follows:

$$\mathrm{cov}[B^{(-1)}(1)-B(s), B^{(-1)}(1)-B(t)] = \frac{1}{12}-\frac{1}{2}t(1-t)-\frac{1}{2}s(1-s)+\min(s,t)-st. \tag{3.41}$$

If we assume $s \leq t$, this expression becomes

$$\mathrm{cov}[B^{(-1)}(1) - B(s), B^{(-1)}(1) - B(t)] = \frac{1}{12} - \frac{1}{2}(t-s)(1-(t-s)), \tag{3.42}$$

which, as we could expect, is no longer a function of s and t separately, but rather of their difference, and depends symmetrically on $t-s$ and $1-(t-s)$.

If we define

$$\tau = t - s \tag{3.43}$$

and

$$\lambda = \tau(1-\tau) = (t-s)(1-(t-s)), \tag{3.44}$$

the covariance matrix can be written as

$$C_{pbc} = \begin{pmatrix} \frac{1}{12} & \frac{1}{12} - \frac{1}{2}\lambda \\ \frac{1}{12} - \frac{1}{2}\lambda & \frac{1}{12} \end{pmatrix} \tag{3.45}$$

and therefore

$$A_{pbc} = C_{pbc}^{-1} = \frac{1}{\lambda(1-3\lambda)} \begin{pmatrix} 1 & -1+6\lambda \\ -1+6\lambda & 1 \end{pmatrix}. \tag{3.46}$$

By comparing (3.19) with (3.46), we find

$$\begin{aligned} b &= \tfrac{\lambda(1-3\lambda)}{1-6\lambda} \\ a = c &= \tfrac{1-3\lambda}{6}. \end{aligned} \tag{3.47}$$

The important difference with respect to the non-periodic case is that

$$a + b + c = \frac{(1-3\lambda)^2}{3(1-6\lambda)}, \tag{3.48}$$

which is in general $\neq 1$. Moreover, b and $a+b+c$ can now have a negative sign:

$$\begin{aligned} &b < 0 \\ &a + b + c < 0 \end{aligned} \quad \text{if} \quad \lambda > \frac{1}{6}. \tag{3.49}$$

If $\lambda < \frac{1}{6}$, the result in (3.36) is still valid, while if $\lambda > \frac{1}{6}$, we obtain[2]

$$\alpha(a,b,c) = \frac{1}{2\pi} \arctan \sqrt{\frac{b(a+b+c)}{ac}} = \frac{1}{2\pi} \left[\arctan \left(-\sqrt{\frac{ac}{b(a+b+c)}} \right) + \frac{\pi}{2} \right], \tag{3.50}$$

which leads to

$$\begin{aligned} G_3(\lambda) &= \begin{cases} \frac{2}{\pi} \arctan \left(\frac{|1-6\lambda|}{\sqrt{12\lambda(1-3\lambda)}} \right) & \text{if} \quad \lambda < \frac{1}{6} \\ \frac{2}{\pi} \arctan \left(-\frac{|1-6\lambda|}{\sqrt{12\lambda(1-3\lambda)}} \right) & \text{if} \quad \lambda > \frac{1}{6} \end{cases} \\ &= \frac{2}{\pi} \arctan \left(\frac{1-6\lambda}{\sqrt{12\lambda(1-3\lambda)}} \right). \end{aligned} \tag{3.51}$$

Or, as a function of τ,

$$G_{pbc}(\tau) = \frac{2}{\pi} \arctan \left(\frac{1-6\tau(1-\tau)}{\sqrt{12\tau(1-\tau)(1-3\tau(1-\tau))}} \right). \tag{3.52}$$

[2]Here we have used the trigonometric identities: $\arctan x + \arctan(\frac{1}{x}) = \frac{\pi}{2}$ and $\arctan(-x) = -\arctan x$.

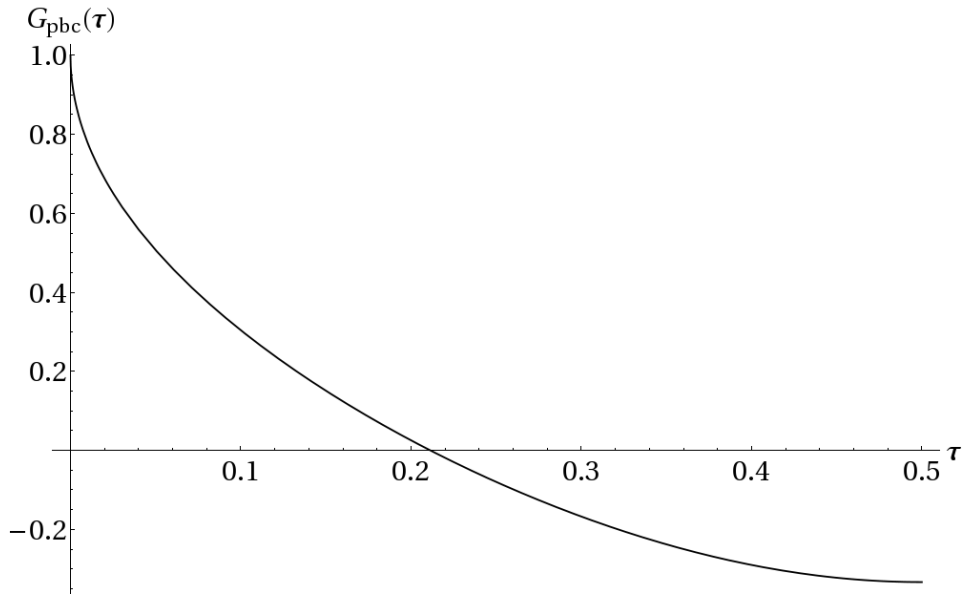

Figure 3.3: The theoretical correlation function in one dimension for periodic boundary conditions.

3.2 Numerical simulations

3.2.1 Choice of the weight function

As we explained in Sect. 2.3, in two dimensions the optimal marriage we have considered is the one which minimises the sum of the distances between matched pairs.

However, in one dimension it was necessary to make a different choice and minimise the sum of the squares of those distances. We had two reasons for doing so.

The first is non-uniqueness in the definition of the optimal matching itself. In one dimension, by using the distances as weights, it is far from exceptional to meet with situations in which two or more matchings have the same energy, that is, to have degenerate instances. To clarify this concept we have shown a couple of examples in Figures 3.4 and 3.5.

Although this ambiguity itself does not affect the quantitative results for the correlation function, it does affect other properties of the problem (such as the distribution of the edge lengths, see Chap. 5).

The second and more important reason for minimising the sum of the squares arises in the case of periodic boundary conditions and lies in the assumptions we have made.

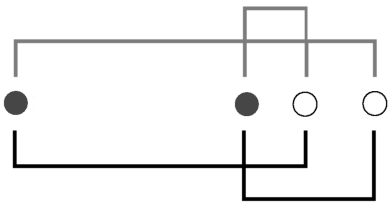

Figure 3.4: The two equivalent matchings for a particular size-2 instance in one dimension.

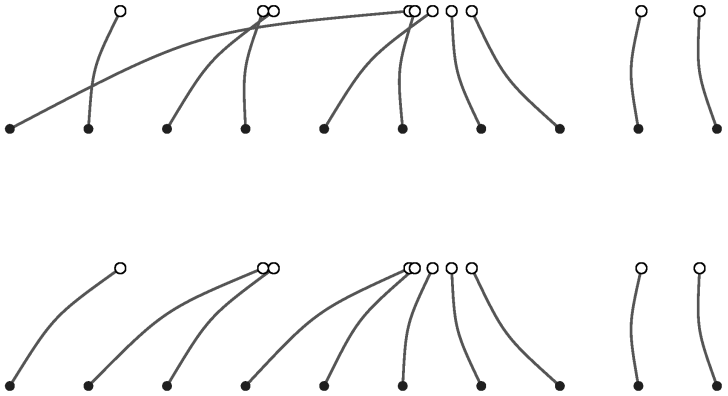

Figure 3.5: Two possible matchings of the same energy for a sample instance at size 10 (only for clarity, we have drawn the Poisson points on a different line).

If $X^\gamma = |\varphi(x)|^\gamma$, $\gamma > 0$, is our weight function, in order to determine the optimal matching we have to minimise $\sum_{\text{all pairs}} X^\gamma$, which means to find a stationary point:

$$\sum_{\text{all pairs}} X^{\gamma-1} = 0. \tag{3.53}$$

In Section 3.1.3, we chose to parametrise our family of processes by the area under their path (3.40), which became our new Gaussian variable.

This assumption corresponds to the case $\gamma = 2$:

$$\sum X^{\gamma-1} = \sum X = 0 \tag{3.54}$$

For any $\gamma \neq 2$ the result of the derivative is not a linear function of the Brownian motion and in this case we do not know the exact solution to the problem, even though we do not expect it to differ to a great extent.

3.2.2 Numerical results at the critical point

We show the results of our simulations for open (Fig. 3.6) and periodic (Fig. 3.7) boundary conditions. The curves in colour represent the experimental data for the correlation function, while the black ones are the plot of the theoretical functions for different values of the system size.

We found that the experimental curves, both for open and periodic boundary conditions, are in good agreement with the theoretical predictions, even at sizes as small as $L = 100$.

Comparison for different weight functions

In Figure 3.8 we show the difference between the correlation function for the optimal marriages obtained by minimising $\sum X^\gamma$, with $\gamma = 1, 2, 3, 4$. We see that the theoretical solution found is in agreement with experimental data for $\gamma = 2$.

3.2.3 Near the critical point

We could not find an exact solution for the correlation function at $\rho \neq 1$ and we do not have an ansatz on how it may be like. However, in Fig. 3.9 and Fig. 3.10 we wish to show the qualitative behaviour of the curve as derived from our simulations.

We see that both in the case of open and periodic boundary conditions, the curves move up as the density increases, reach a peak at $\rho = 1$ and then move down again.

Moreover, the shape and position of the curves are similar for the same value of $|t| = |\rho - 1|$, regardless of the sign of t.

Finally, we observe that this function presents two ranges of behaviour. If $|t| < \bar{t}$, with $\bar{t} \approx 0.01$, the function is strictly decreasing with x, and has a minimum at $x = L$, for open boundary conditions, or $x = \frac{L}{2}$ fo periodic bundary conditions. On the other hand, if $|t| > \bar{t}$, the shape is different. It reaches a minimum at an intermediate value x', then goes up again approaching zero as $x \to$ Ł, or $x \to \frac{L}{2}$ (depending on the boundary conditions).

If we define $\bar{x}$ as the point where the curve has a zero, we find that $\bar{x}(\rho)$ has a maximum at density $\rho = 1$ (see Fig. 3.11).

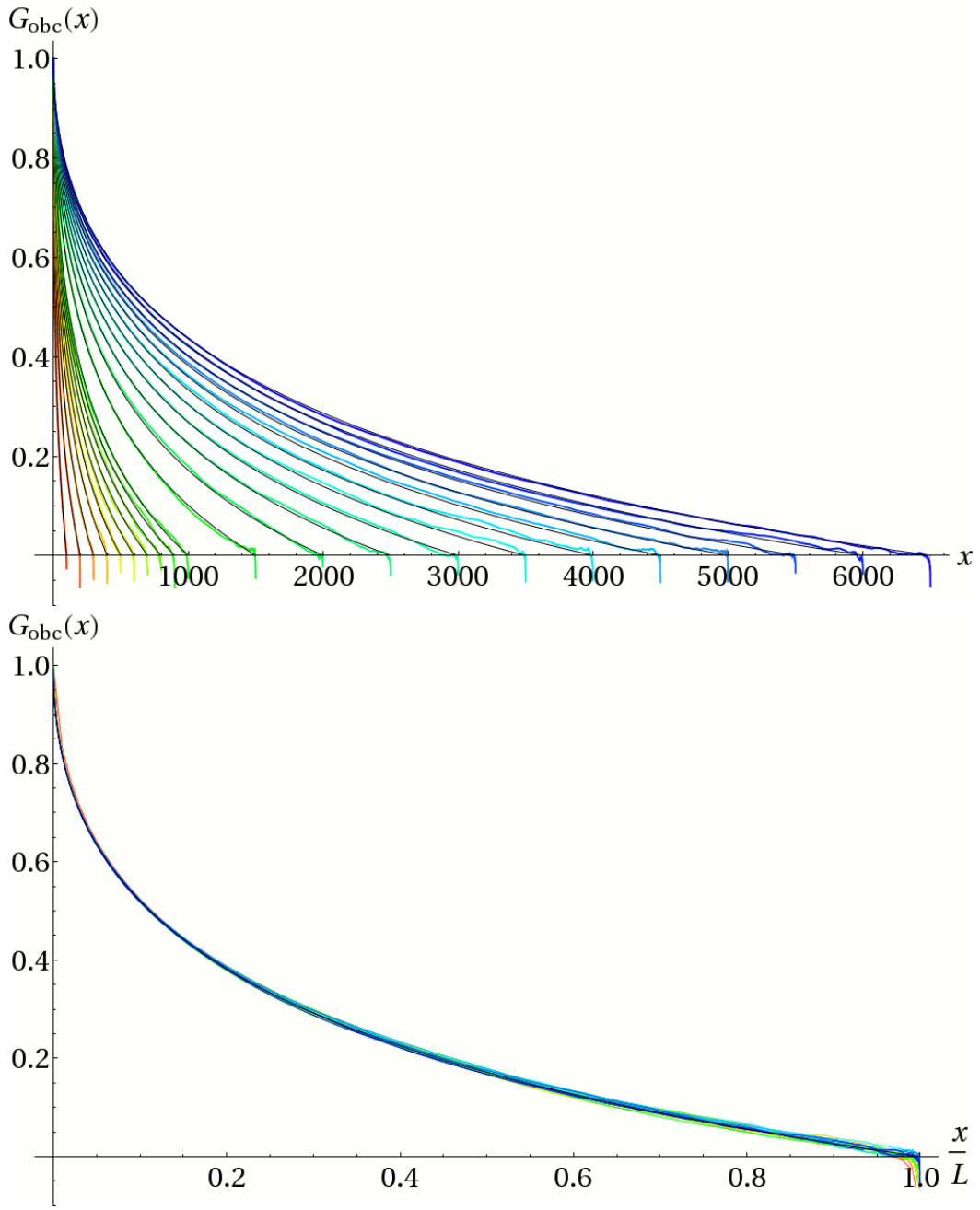

Figure 3.6: The mean of the experimental correlation function over 10^4 instances of the optimal marriage in dimension one with open boundary conditions. Top: increasing sizes are represented from red ($L = 100$) to blue ($L = 6500$); Bottom: the same experimental curves, rescaled to show the agreement with the theoretical function, regardless of the size.

The value of $\bar{x}_L/L$ at $(t = 0)$ is known for all L, as in this case we have an

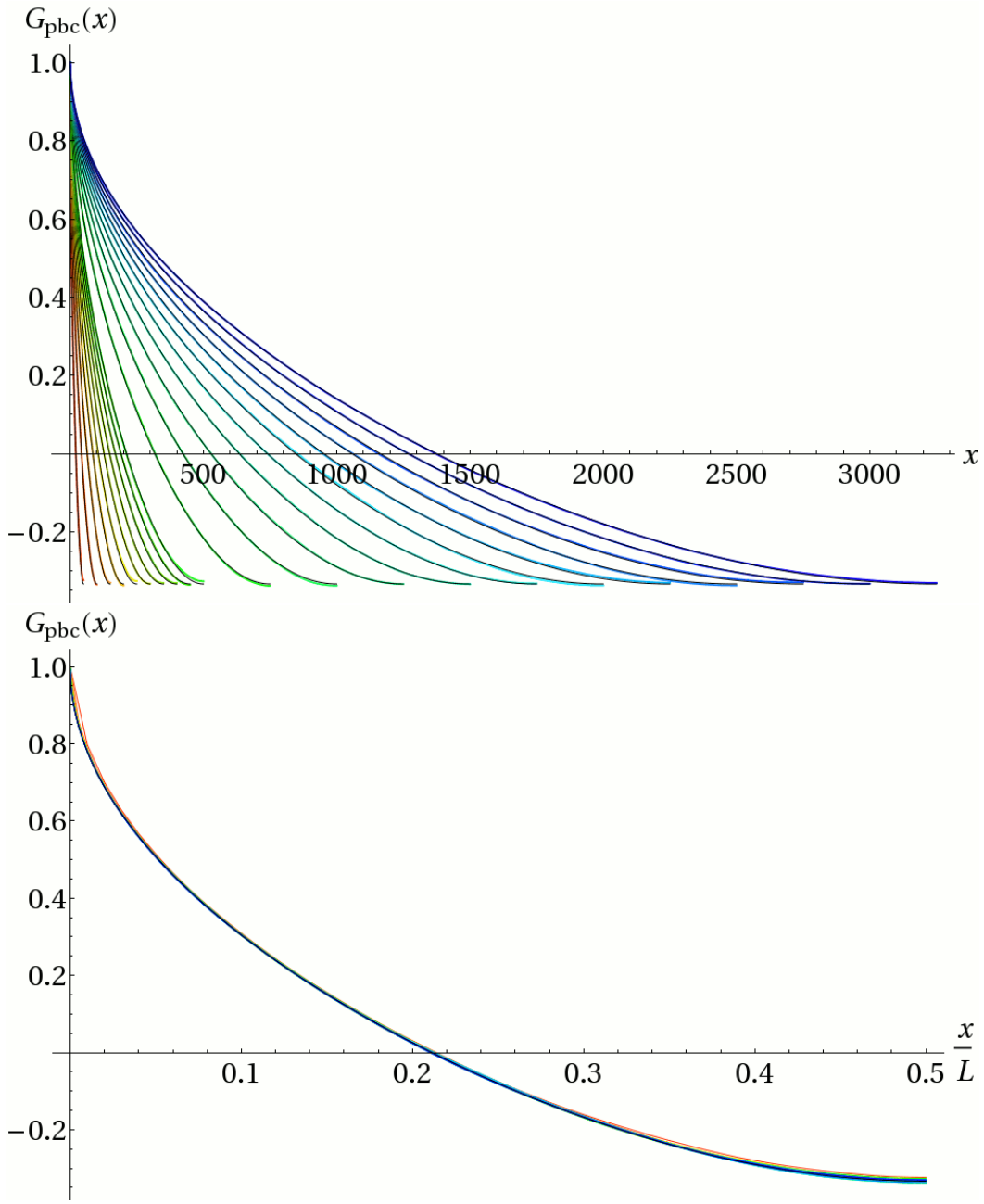

Figure 3.7: The mean of the experimental correlation function over 10^4 instances of the optimal marriage in dimension one with periodic boundary conditions. Top: increasing sizes are represented from red ($L = 100$) to blue ($L = 6500$); Bottom: the same experimental curves, rescaled to show the agreement with the theoretical function, regardless of the size.

exact solution.

$$\frac{\bar{x}_L(t=0)}{L} = 0.211325. \tag{3.55}$$

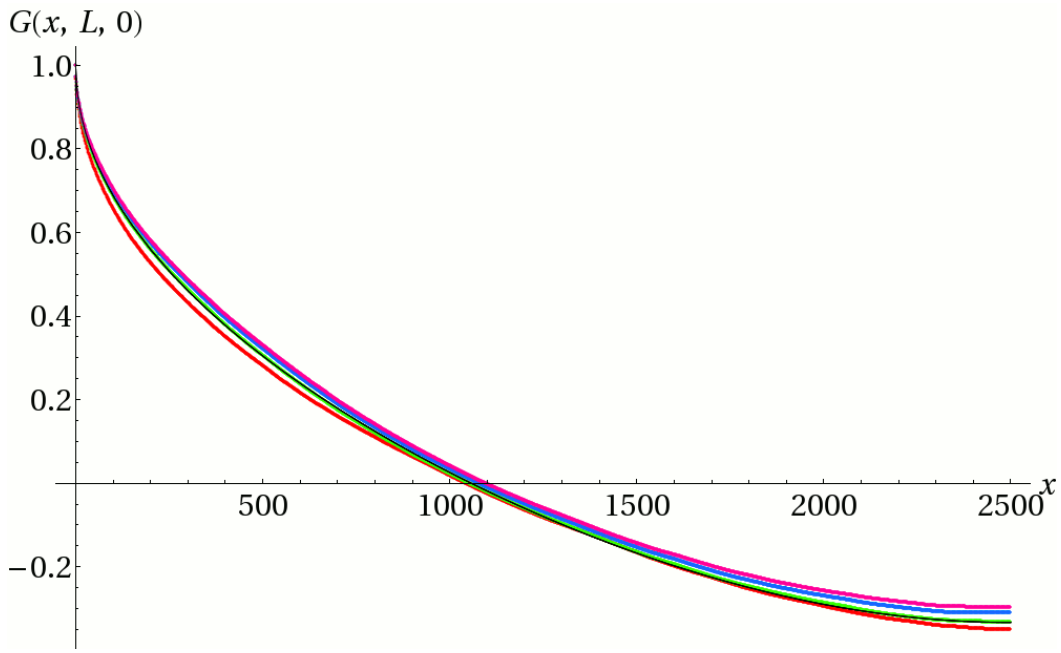

Figure 3.8: Comparison of the correlation function at $\rho = 1$ for the optimal marriages obtained by minimising $\sum X^\gamma$, with different values of γ ($L = 5000$). From red to purple, $\gamma = 1, 2, 3, 4$ (green is $\gamma = 2$). Each curve is the mean over 10^3 instances of optimal marriage. The black thinner curve is the plot of the theoretical function.

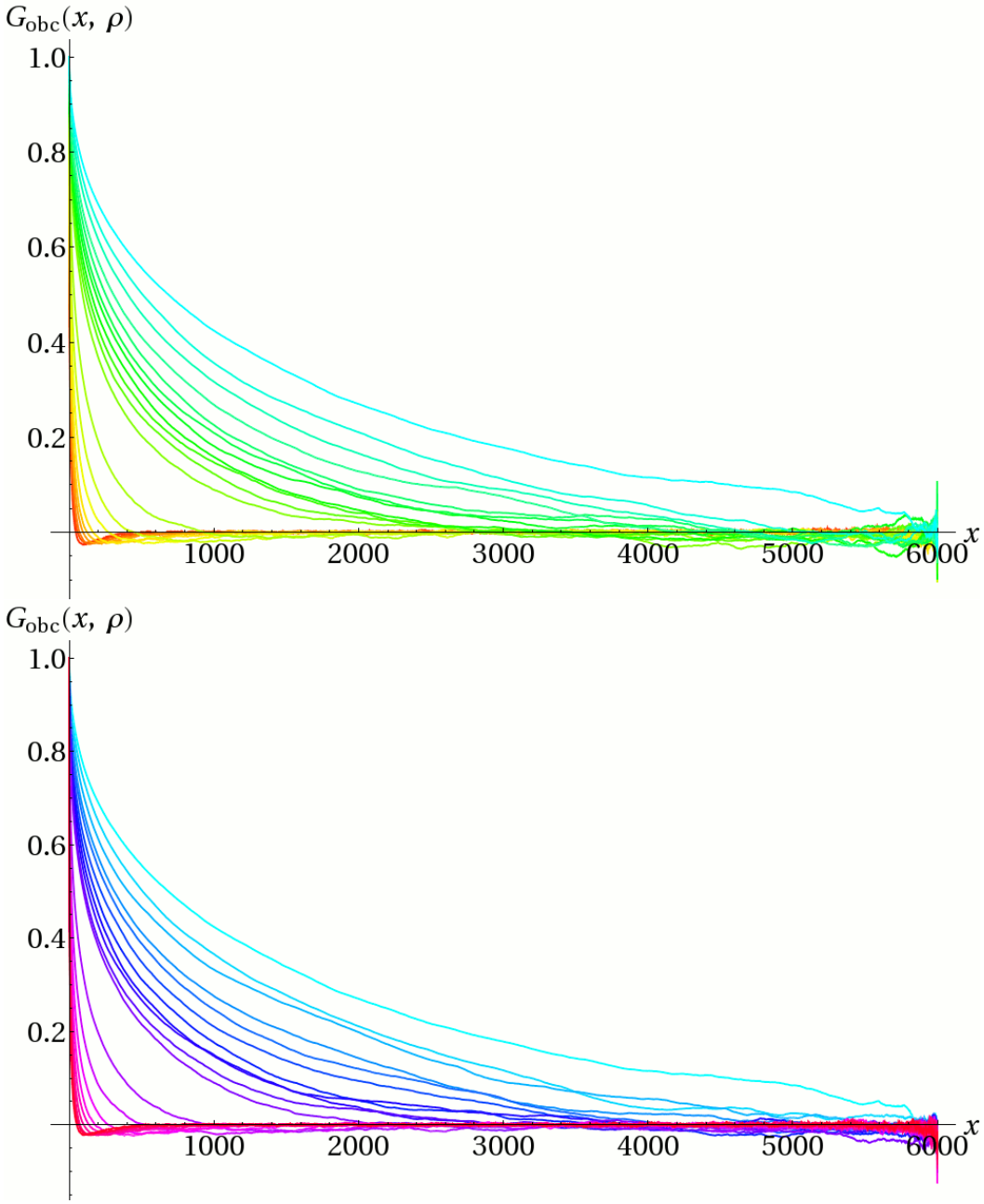

Figure 3.9: The correlation function near the critical point at size 6000 with open boundary conditions.
Top: from red to cyan, ρ = 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.991, 0.992, 0.993, 0.994, 0.995, 0.996, 0.997, 0.998, 0.999, 1.
Bottom: from cyan to purple, ρ =1, 1.001, 1.002, 1.003, 1.004, 1.005, 1.006, 1.007, 1.008, 1.009, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08, 1.09, 1.1. Each curve is the mean over 10^3 instances of optimal marriage.

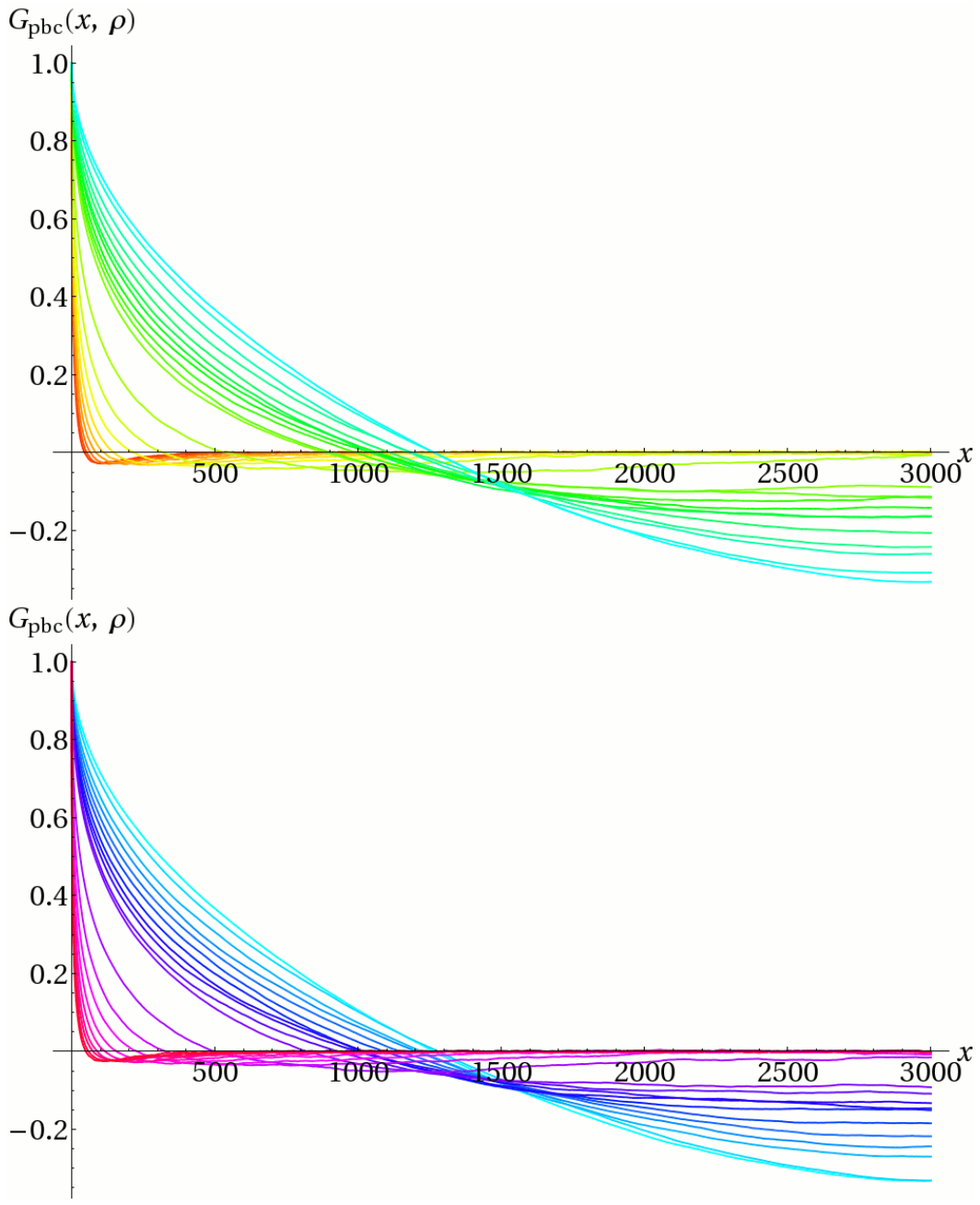

Figure 3.10: The correlation function near the critical point at size 6000 with periodic boundary conditions.
Top: from red to cyan, ρ = 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.991, 0.992, 0.993, 0.994, 0.995, 0.996, 0.997, 0.998, 0.999, 1.
Bottom: from cyan to purple, ρ =1, 1.001, 1.002, 1.003, 1.004, 1.005, 1.006, 1.007, 1.008, 1.009, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08, 1.09, 1.1. Each curve is the mean over 10^3 instances of optimal marriage.

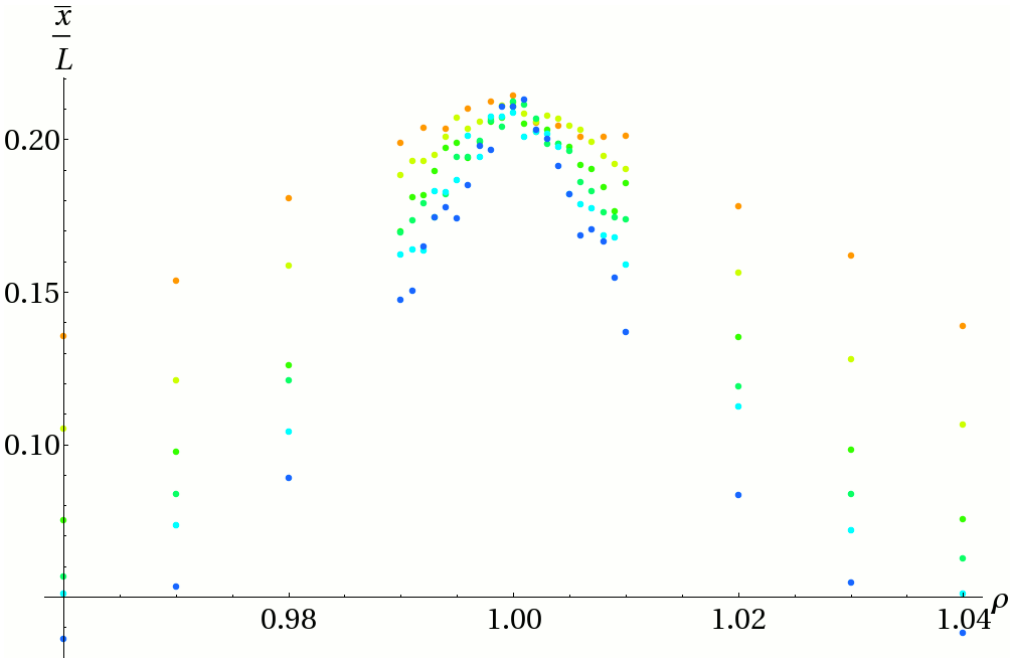

Figure 3.11: The experimental rescaled intersection with the x-axis ($\bar{x}_L/L$) as a function of the density (ρ). From orange to blue, $L = 500, 1000, 2000, 3000, 4000, 6000$. For all L, $\bar{x}_L(t=0)/L \approx 0.211325$.

4

Finite-Size Scaling in two dimensions

Over the last decades, the widening availability of more and more powerful computers and specially-designed processors has led to an ever-increasing role of numerical simulations in studies of equilibrium and non-equilibrium problems of statistical mechanics.

In addition, it has become clear that for complex models such as those of critical phenomena, analytic determinations are particularly difficult to achieve and numerical results can bring about a great amount of useful qualitative and quantitative information.

However, no matter how powerful the technology gets, computers will always be able to deal only with a finite number of elements. Consequently, it is possible to work only on lattices of finite sizes, while we know from statistical mechanics that a phase transition can only occur in the thermodynamic limit, that is, for systems where the number of particles tends to infinity.

Here comes into play the theory of finite-size scaling (FSS), which provides a powerful technique to extrapolate to the thermodynamic limit from finite-size results, which can be obtained from computer simulations. In this way the behaviour of large systems can be inferred from the known behaviour of relatively small systems.

The ideas of FSS were first introduced by Fisher [17], and Fisher and Barber (see [14]), in the early seventies, and thoroughly studied in ensuing papers.

In the next section, we will give a brief introduction to the theory of FSS, partly following [16]. Our description will be limited to the concepts necessary to our purposes, while for a complete coverage and theoretical justification, we refer to [13] and [14].

4.1 Theory of finite-size scaling

We shall consider the framework of second-order phase transitions, and a system controlled by a single scalar parameter T (which is usually assimilated to a temperature, while in our problem will correspond to a density).

4.1.1 Thermodynamic limit

We shall assume it is possible to give a thermodynamic description of our model on a discrete lattice in a finite box Λ. That is to say, given any observable $\mathcal{O}$, we can calculate its value on the thermodynamic state determined by T and Λ as

$$\mathcal{O}_\Lambda(T) := \langle \mathcal{O} \rangle_\Lambda(T)$$

where $\langle \cdot \rangle_\Lambda(T)$ is the appropriate averaging. In addition, we shall assume the existence of a *thermodynamic limit*. Usually, a model is taken to the thermodynamic limit by increasing the volume together with the particle number while keeping the particle number density constant. A common way of doing so is to consider an increasing sequence of boxes, $\{\Lambda_n\}_n$. The value of the observable in the infinite-volume is then given by

$$\mathcal{O}_\infty = \lim_{n \to \infty} \mathcal{O}_{\Lambda_n}(T)$$

This limit exists in a wide range of cases and usually does not depend on the shape of the boxes. Only in the thermodynamic limit can the singularities in the thermodynamic functions associated with a critical point occur. In this case, in systems characterized by short-range interactions, if we define a (connected) correlation function of the order parameter $\phi(x)$ as

$$G_{\phi,\infty}(x) := \langle \phi(x); \phi(0) \rangle_\infty = \langle \phi(0)\phi(x) \rangle_\infty - \langle \phi(0) \rangle_\infty \langle \phi(x) \rangle_\infty,$$

we expect $G_{\phi,\infty}(x)$ to show an exponential decay at long distances

$$G_{\phi,\infty}(x) \underset{x \to \infty}{\sim} e^{-\frac{|x|}{\xi_\infty}}$$

As we saw in Chap. 1, the length scale governing this exponential decay is the correlation length ξ_∞:

$$\xi_\infty := - \lim_{|x| \to \infty} \frac{|x|}{\log |G_{\phi,\infty}(x)|}$$

In principle, the correlation length may depend on the direction along wich this limit is taken. However, for simplicity, in the following we shall only take into consideration systems in which ξ does not depend on this direction and are thus said to undergo an *isotropic* phase transition. In this case, we shall consider boxes of size L in all directions. We shall also only consider the case of zero external ordering field: $H = 0$.

4.1.2 Scaling hypothesis

In the proximity of the critical point at $t = 0$, where we have defined the reduced temperature

$$t := \frac{T - T_c}{T_c} \tag{4.1}$$

(T_c being the critical temperature, at which the second-order phase transition occurs), several thermodynamic quantities diverge. Their bulk ($L = \infty$) critical behaviour is given by

$$\mathcal{O}_\infty(t) \sim |t|^{-x_\mathcal{O}} \qquad \text{for } t \to 0, \tag{4.2}$$

where $\sim$ means that $\lim_{t\to 0} \mathcal{O}_\infty(t)/|t|^{-x_\mathcal{O}}$ is finite, and $x_\mathcal{O}$ is the critical exponent of $\mathcal{O}$. As seen, the critical exponent of the correlation length is traditionally denoted by ν:

$$\xi_\infty \sim |t|^{-\nu} \qquad \text{for } t \to 0, \tag{4.3}$$

As we have already pointed out, a phase transition is only possible in the thermodynamic limit, that is, for infinite volumes. This means that

$$\lim_{t\to 0} \lim_{L\to\infty} \mathcal{O}_L(t)/|t|^{-x_\mathcal{O}} \neq \lim_{L\to\infty} \lim_{t\to 0} \mathcal{O}_L(t)/|t|^{-x_\mathcal{O}} = 0, \tag{4.4}$$

if, for example, $x_\mathcal{O} > 0$.

The finite-size scaling theory predicts the behaviour of the function $\mathcal{O}_L(t)$ near T_c and for L large enough compared to the lattice spacing. In this region, called the *critical region*, it is possible to write $\mathcal{O}_L(t)$ as a function of ξ_∞ and L:

$$\mathcal{O}_L(t) \approx f_\mathcal{O}(\xi_\infty(t), L),$$

where $\approx$ means that we are omitting lower-order terms as $L \to \infty$.

As $\xi_\infty(t)$ and L are the only dimensionful quantities, the function $f_\mathcal{O}(x, y)$ must be a homogeneous function of its arguments, then

$$\mathcal{O}_L(t) \approx f_\mathcal{O}(\xi_\infty(t), L) = \xi_\infty(t)^{y_\mathcal{O}} F_\mathcal{O}^{(1)}\left(\frac{\xi_\infty(t)}{L}\right) \tag{4.5}$$

or, equivalently,

$$\mathcal{O}_L(t) \approx L^{y_\mathcal{O}} F_\mathcal{O}^{(2)}\left(\frac{\xi_\infty(t)}{L}\right) \tag{4.6}$$

The exponent $y_\mathcal{O}$, that is the degree of the homogeneous function, can be determined as $x_\mathcal{O}/\nu$, since[1]

$$\mathcal{O}_\infty(t) = \lim_{L\to\infty} \mathcal{O}_L(t) = \xi_\infty(t)^{y_\mathcal{O}} F_\mathcal{O}^{(1)}(0) \sim |t|^{-\nu y_\mathcal{O}}$$

In this way, for a good definition of the finite-size correlation length, we must have

$$\xi_L(t) \approx L \cdot g\left(\frac{\xi_\infty(t)}{L}\right), \tag{4.7}$$

since, in this case, $x_\xi = \nu$.

4.1.3 Asymptotic form of FSS

At this point, we can replace the unknown quantity ξ_∞ with known parameters wich can be tuned in experiments: t and L.

$$\mathcal{O}_L(t) \approx L^{x_\mathcal{O}/\nu} F_\mathcal{O}^{(2)}\left(\frac{\xi_\infty(t)}{L}\right) \approx L^{x_\mathcal{O}/\nu} F_\mathcal{O}^{(2)}\left(\frac{t^{-\nu}}{L}\right) = L^{x_\mathcal{O}/\nu} G_\mathcal{O}\left(tL^{1/\nu}\right) \tag{4.8}$$

The function $G_\mathcal{O}(z)$ is finite and non-vanishing in zero, and should satisfy

$$G_\mathcal{O}(z) \sim |z|^{-x_\mathcal{O}} \qquad \text{for } z \to \infty. \tag{4.9}$$

This simple form of the FSS hypothesis heavily relies on the assumption that the bulk critical temperature is known or can be evaluated. In addition, the unknown critical exponent of ξ_∞ is still present.

[1] We wish to point out that the existence of a finite limit for $F_\mathcal{O}^{(1)}(z)$ for $z \to 0$ is based on the hypotheses of the existence of a thermodynamic limit for $\mathcal{O}_L$ and the possibility to interchange this limit with the FSS limit. We will assume these are verified. For a deeper analysis on this subject, we refer to [16].

An important consequence of (4.8) is that, once we get to know T_c, it tells us that there is a well defined functional dependence between $\mathcal{O}_L(t)/L^{x_\mathcal{O}/\nu}$ and $tL^{1/\nu}$ over the critical region. This suggests the possibility of estimating $x_\mathcal{O}$ and ν by carrying out experiments at different sizes and temperatures and plotting the results with different values of the exponents until the set of points collapse into a single curve.

4.1.4 Extrapolating to infinite volume

As explained in [15], there is another strategy to take advantage of FSS relations in such a way as to obtain a relation which can be profitably used in numerical experiments.

In the critical region, for any fixed t, (4.7) can be written as

$$\frac{\xi_L}{L} \approx g\left(\frac{\xi_\infty}{L}\right)$$

or, by inverting the functional relation,

$$\frac{\xi_\infty}{L} \approx h\left(\frac{\xi_L}{L}\right). \tag{4.10}$$

This means that

$$\mathcal{O}_L(t) \approx L^{x_\mathcal{O}/\nu} F_\mathcal{O}^{(3)}\left(\frac{\xi_L(t)}{L}\right) \tag{4.11}$$

and therefore, if we take the ratio of $\mathcal{O}_L(t)$ at two different sizes L and αL, we have

$$\frac{\mathcal{O}_{\alpha L}(t)}{\mathcal{O}_L(t)} \approx F_\mathcal{O}\left(\frac{\xi_L(t)}{L}\right), \tag{4.12}$$

where the critical exponents disappear and only measurable quantities of the finite system are present.

In particular, this relation is valid for the correlation length. At fixed t

$$\frac{\xi_{\alpha L}(t)}{\xi_L(t)} \approx F_\xi\left(\frac{\xi_L(t)}{L}\right), \tag{4.13}$$

As a result, we find that in a regime where FSS is proved to hold, there must be a universal function which, if we are able to estimate it, allows us to extrapolate

values of $\xi_{\alpha^n L}$ and $\mathcal{O}_{\alpha^n L}$ for arbitrary n, until we get to the limiting values ξ_∞ and $\mathcal{O}_\infty$ (save, of course, for systematic and statistical errors).

The idea is that we make numerical experiments at finite L and we want to extrapolate to $L \to \infty$. However, we do not want to do so in some naïve way (that is, simply extrapolating from a series of points), but we want to avail ourselves of this structure. By doing so, we are really capable of extrapolating to much larger values of L than we can simulate, and we are using a finer approach.

Clearly, these expressions are not exact. The corrections can be evaluated by means of the techniques of the Renormalization Group, but this is beyond the scope of this work.

4.2 Numerical simulations

In the following, we will apply the FSS theory to the marriage problem in two dimensions and verify whether the scaling relations (especially (4.8)) hold in this context.

Firstly, we need to define the order parameter and the corresponding correlation function we want to analyse.

4.2.1 Definitions

Let us consider the GPM problem on a square lattice of side L. Given an instance of Poisson points and found the optimal marriage as seen (see 2.3), we can define a map $\vec{\varphi}$ that relates each matched grid point of coordinates $\vec{x}$ to a vector $\vec{\varphi}(\vec{x})$ with the tail in $\vec{x}$ and the head in the matched Poisson point.

We then define a finite-size correlation function between two grid points $\vec{x}$ and $\vec{x}'$ as[2]

$$G_{pp}(\vec{x}, \vec{x}') = \frac{\vec{\varphi}(\vec{x}) \cdot \vec{\varphi}(\vec{x'})}{|\vec{\varphi}(\vec{x})| \cdot \left|\vec{\varphi}(\vec{x'})\right|} = \cos\theta(\vec{\varphi}(\vec{x}), \vec{\varphi}(\vec{x'})), \tag{4.14}$$

[2]In the following, we will imply that we are taking the average over the probability distribution of the Poisson points $\langle\cdot\rangle_{\mathcal{P}}$. In concrete terms, we will be averaging over from 10^3 to $5 \cdot 10^4$ instances of Poisson points.

(where $\theta\left(\vec{\varphi}, \vec{\varphi'}\right)$ is the angle between the two vectors), and a finite-size correlation function at distance $\vec{r}$:

$$G(\vec{r}) = \left\langle \frac{\vec{\varphi}(\vec{x}) \cdot \vec{\varphi}(\vec{x}+\vec{r})}{|\vec{\varphi}(\vec{x})| \cdot |\vec{\varphi}(\vec{x}+\vec{r})|} \right\rangle_{\vec{x}}, \tag{4.15}$$

where $\langle\cdot\rangle_{\vec{x}}$ means we are taking the average over the coordinates $\vec{x}$ of all matched grid points.

The wall-to-wall correlation function

In this work, we have focused our attention on the *wall-to-wall* correlation function, which describes the correlation of all points that belong to the columns (rows) at a fixed horizontal (vertical) distance z ($z \in [0, L-1], z \in \mathbb{N}$).

More precisely, we define the horizontal wall-to-wall correlation function as

$$G^h_L(x, z) = \left\langle \frac{\vec{\varphi}(x, y) \cdot \vec{\varphi}(x+z, y')}{|\vec{\varphi}(x, y)| \cdot |\vec{\varphi}(x+z, y')|} \right\rangle_{yy'}, \tag{4.16}$$

where $\vec{x} = (x, y)$, and $\langle\cdot\rangle_{yy'}$ means we are taking the average over the vertical coordinates y and y' of the matched pairs in the columns.

In a more operative form

$$G^h_L(x, z) = \frac{1}{n} \sum_{\substack{y,y'=1 \\ y,y' \text{ of matched pairs}}}^{L} \frac{\vec{\varphi}(x, y) \cdot \vec{\varphi}(x+z, y')}{|\vec{\varphi}(x, y)| \cdot |\vec{\varphi}(x+z, y')|}, \tag{4.17}$$

where $n = L^2$ if all grid points in the columns are matched.

The vertical wall-to-wall correlation function $G^v_L(y, z)$ is defined equivalently.

The quantity we studied was the symmetrisation at distance z of the average of the horizontal and vertical wall-to-wall function over all coordinates x and y:

$$G(z, L) = \frac{1}{2}\left(\langle G^h_L(x, z)\rangle_x + \langle G^v_L(y, z)\rangle_y\right) \tag{4.18}$$

The reason why we chose to study this function, instead of (4.15) is that it is easier to handle, as it is a function of only one integer variable.

In addition, we have decided to focus only on the case of periodic boundary conditions (that is to say, on a torus), in order to avoid border effects.

Our function will then depend on two parameters, the side of lattice L and the reduced temperature $t = \frac{\rho-\rho_c}{\rho_c} = \rho - 1$, and one variable x.

Moreover, by construction, it will be symmetric in the exchange of x with $L - x$:

$$G(x, L, t) = G(L - x, L, t) \tag{4.19}$$

and consequently, only values of x in $[0, \frac{L}{2}]$ will be considered.

4.2.2 Qualitative behaviour of the correlation function

In Fig. 4.1, we show the behaviour of the curves for a sample side $L = 60$ at different values of the parameter t.

Similarly to what we found in one dimension, we can see that there are essentially two ranges of behaviour:

- if $|t| < \bar{t}$, with $\bar{t} \approx 0.01$, the function is strictly decreasing with x, and has a minimum at $x = \frac{L}{2}$.
- if $|t| > \bar{t}$, on the other hand, the shape is different. It reaches a minimum at an intermediate value x' ($0 < x' < \frac{L}{2}$), then goes up again and presents a maximum at $x = \frac{L}{2}$.

We are interested in the critical behaviour, that is, near $t = 0$. Therefore in the following we will only consider the smallest values of $|t|$, where the curves are of the first type.

Moreover, we see that the experimental value of G as we have defined it is to some extent altered at $x = 0$ by the presence in the sum of the terms of self-correlation, $G_{pp}(\vec{x}, \vec{x}) = 1$, which raise abnormally the figure. For this reason, in our following analyses, we will not consider this point.

By looking at the curves in Fig. 4.1 and Fig. 4.2, we also observe that, on first approximation, the shape of the curve depends on the absolute value of t, but is very similar for positive and negative signs.

As a last preliminary note, we see that the point of intersection of the curves and the x-axis increases with t for $t < 0$, has a maximum at $t = 0$ and then decreases with t for $t > 0$.

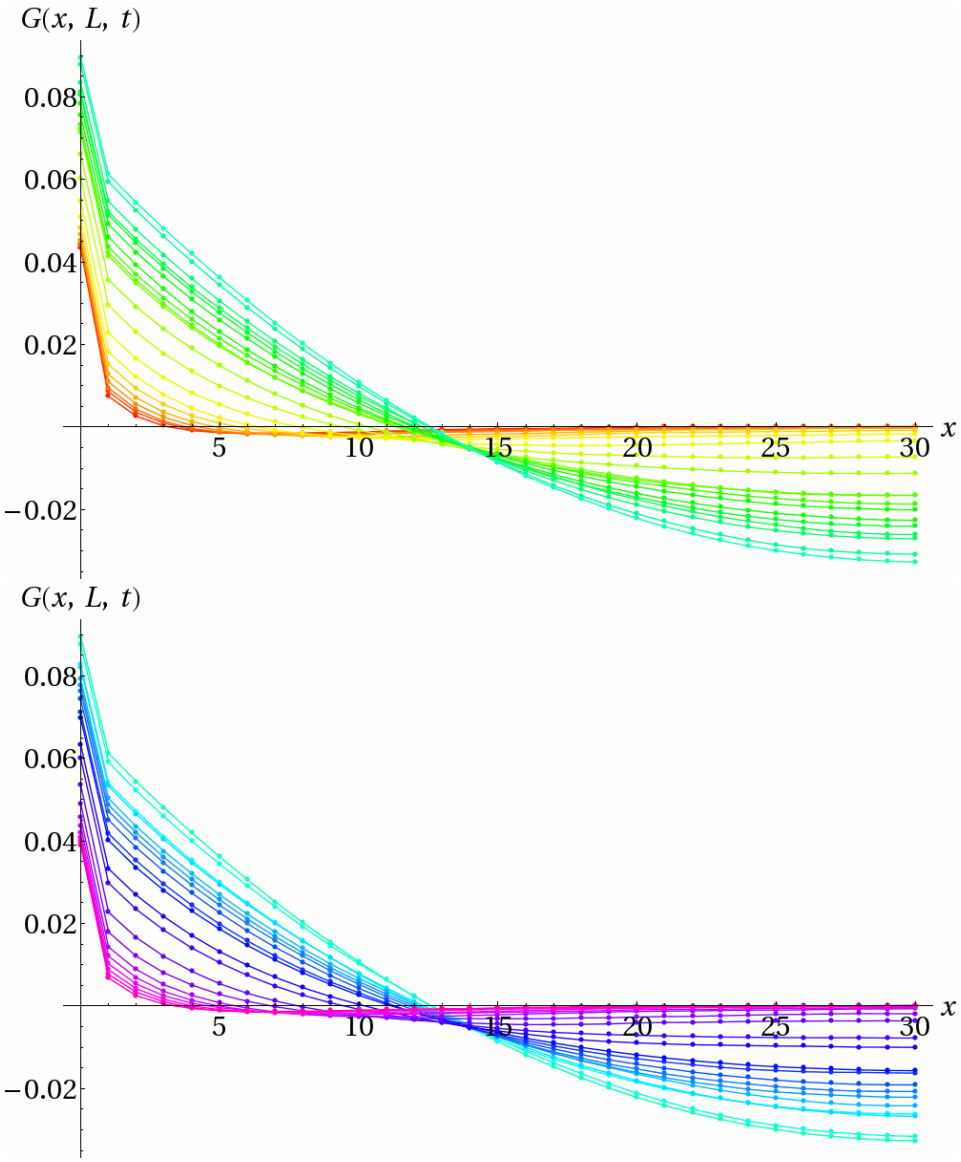

Figure 4.1: The wall-to-wall correlation function at side $L = 60$ for different values of the density.
Top: from red to cyan, $\rho = 0.9$, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.986, 0.99, 0.991, 0.992, 0.993, 0.994, 0.996, 0.997, 0.998, 0.999, 1.
Bottom: from cyan to purple, $\rho = 1$, 1.001, 1.002, 1.003, 1.004, 1.006, 1.007, 1.008, 1.009, 1.01, 1.015, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08, 1.09, 1.1.

4.2.3 Curve-fitting ansatz

We want to make an educated guess about what mathematical function can describe the behaviour we expect from the system and be best in accordance

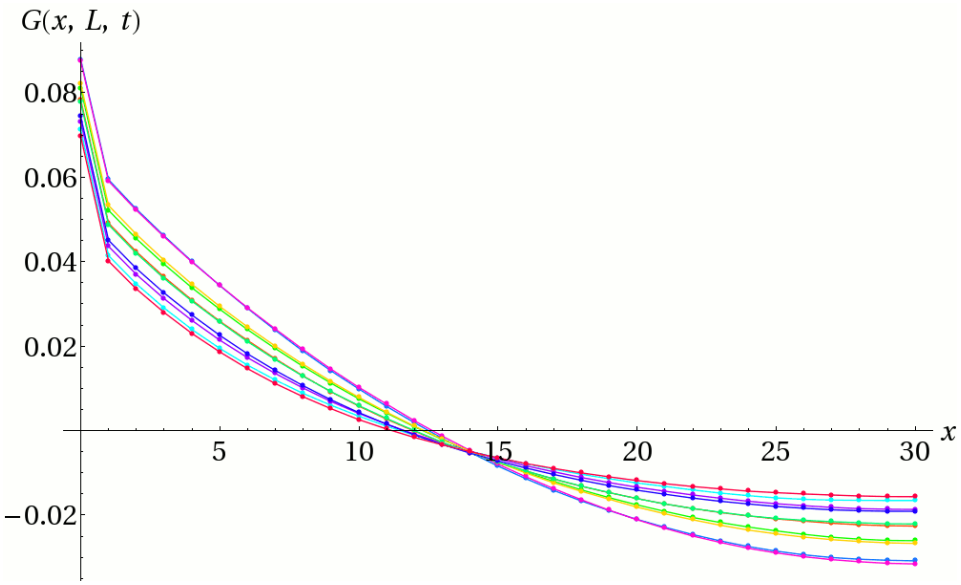

Figure 4.2: Comparison of the wall-to-wall correlation function at $t = \pm 0.001, \pm 0.003, \pm 0.006, \pm 0.008$ and $t = \pm 0.01$ ($L = 60$).

with the experimental data. In particular, we are looking for a function which shows an exponential decay with the distance, x, but is symmetric in the exchange $x \leftrightarrow L - x$.

Our first choice is then the hyperbolic cosine:

$$G(x, L, t) = a \cdot \cosh\left[c \cdot \left(x - \frac{L}{2}\right)\right] + b, \tag{4.20}$$

where $a = a(L, t), b = b(L, t)$ and $c = c(L, t)$ are the fitting parameters we would like to estimate.

For small values of $|t|$, $t \neq 0$, the experimental points are fitted well by this function, as we show in Fig. 4.3.

On the contrary, at precisely $t = 0$ it has been impossible to find a clear fit with the function (4.20), as the parameter c tends to vanish.

Since for small values of its argument, the hyperbolic cosine can be approximated by a parabola, on the critical point we chose to fit our data to the function:

$$G(x, L, t = 0) = a \cdot \left(x - \frac{L}{2}\right)^2 + b, \tag{4.21}$$

with $a = a(L, t)$ and $b = b(L, t)$ (see Fig. 4.4).

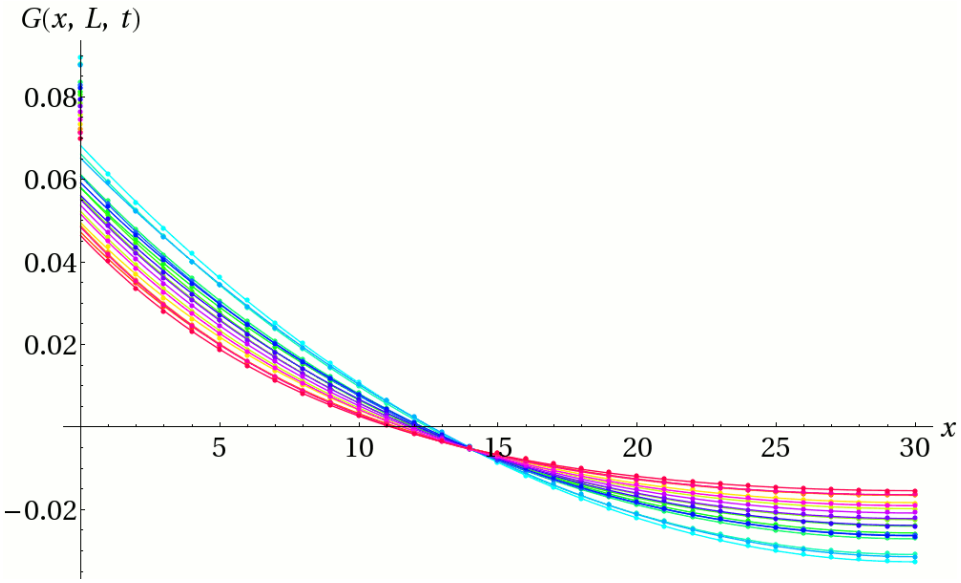

Figure 4.3: The experimental points for the correlation function for $L = 60$ near $t = 0$ shown together with their respective fitted curves (hyperbolic cosines).

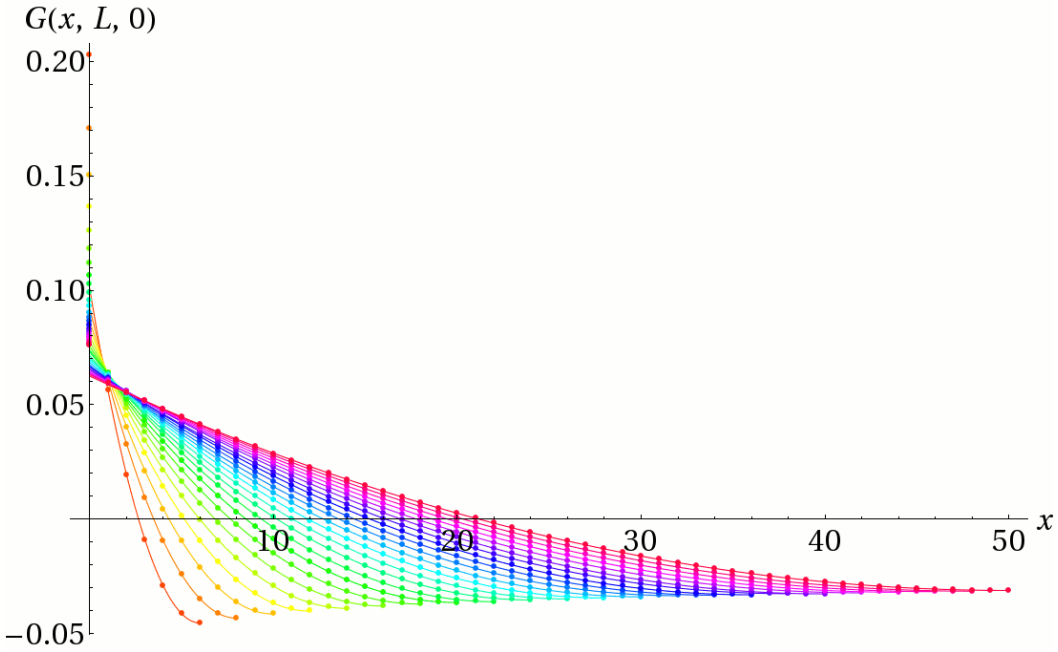

Figure 4.4: The experimental points for the correlation function at $t = 0$ shown together with the respective fitted curves (parabolas). From red to purple, $L = 12, 16, 20, \ldots, 96, 100$.

We would like to observe that, similarly to what happened with the function we found as an exact solution in one dimension with periodic boundary conditions (see Sect. 3.1.3), in this context the dependence on the rescaled distance $\frac{x}{L}$ is through an expression which depends symmetrically on $\frac{x}{L}$ and $(1-\frac{x}{L})$, as it is easily seen:

$$\left(\frac{L}{2}-x\right)^2 = L^2\left(\frac{1}{4}-\frac{x}{L}\left(1-\frac{x}{L}\right)\right), \tag{4.22}$$

4.2.4 Scaling ansatz at the critical point

If t is small enough, theory tells us there must be scaling. This means that $G(x,L,t)$ must be a homogeneous function, for instance[3]:

$$G(x,L,t) \approx L^{\alpha} F\left(\frac{x}{L}, tL^{1/\nu}\right) \tag{4.23}$$

where, in the notation of the previous sections, $\alpha = x_G/\nu$ (cf (4.8)).

We are interested in knowing what the function F is like and the values of the parameters α and ν. However, in this form, G can be difficult to study, as it depends on two arguments. For this reason, in the first place, we have studied the behaviour of this function at $t=0$, as we do not meet with any difficulties in performing simulations exactly on the critical point.

It is clear that, on first approximation, if $t=0$,

$$G(x,L,t=0) \approx L^{\alpha} f\left(\frac{x}{L}\right). \tag{4.24}$$

where $f(z)$ is a universal function.

4.2.5 Determination of the parameter α

We see that in (4.24) the exponent ν is no longer present. We can then focus on finding an estimate for α.

[3] In general, F depends on x/L^{β}, but we will see in the following that our simplified assumption in this case holds.

First estimate of α

A very simple way of determining the parameter α is the following. Let us consider a certain fixed value of x such that $G(x, L, t = 0)$ is not too close to zero (any value, except $x = 0$ for the reason stated above). For instance, let us take the point $x = \frac{L}{2}$. We can therefore write

$$G(x = \frac{L}{2}, L, t = 0) \approx L^{\alpha} f\left(\frac{1}{2}\right). \tag{4.25}$$

valid for all even L. Then

$$\begin{aligned} \log\left[-G\left(\frac{L}{2}, L, 0\right)\right] &\approx \log\left[L^{\alpha}\left(-f\left(\frac{1}{2}\right)\right)\right] \\ &\approx \alpha \log L + \log\left[-f\left(\frac{1}{2}\right)\right] \end{aligned} \tag{4.26}$$

In this way, if we plot $\log\left[-G\left(\frac{L}{2}, L, 0\right)\right]$ against $\log L$ and calculate the parameter of the linear fit (Fig. 4.5), this gives us a first rough estimate for α as

$$\alpha \approx -0.176. \tag{4.27}$$

In order to have an idea of the range of sides L in which this result is (approximately) correct, we plot $G\left(\frac{L}{2}, L, 0\right)/L^{\alpha}$ against L (see Fig. 4.6).

There are some fluctuations in these values, but for the time being we do not see any well-defined systematic trends.

A preliminary test of our ansatz comes from the simple procedure of plotting $G\left(x, L, 0\right)/L^{\alpha}$ as a function of $\frac{x}{L}$ (see Fig. 4.7). As we hoped, all points collapse into a single curve, regardless of the value of L. This curve is, as already seen (Sect. 4.2.3), a parabola.

$$G(x, L, 0) = a(L) \cdot \left(x - \frac{L}{2}\right)^2 + b(L) \approx L^{\alpha} f\left(\frac{x}{L}\right), \tag{4.28}$$

which means that

$$f\left(\frac{x}{L}\right) \approx A\left(\frac{x}{L} - \frac{1}{2}\right)^2 + B, \tag{4.29}$$

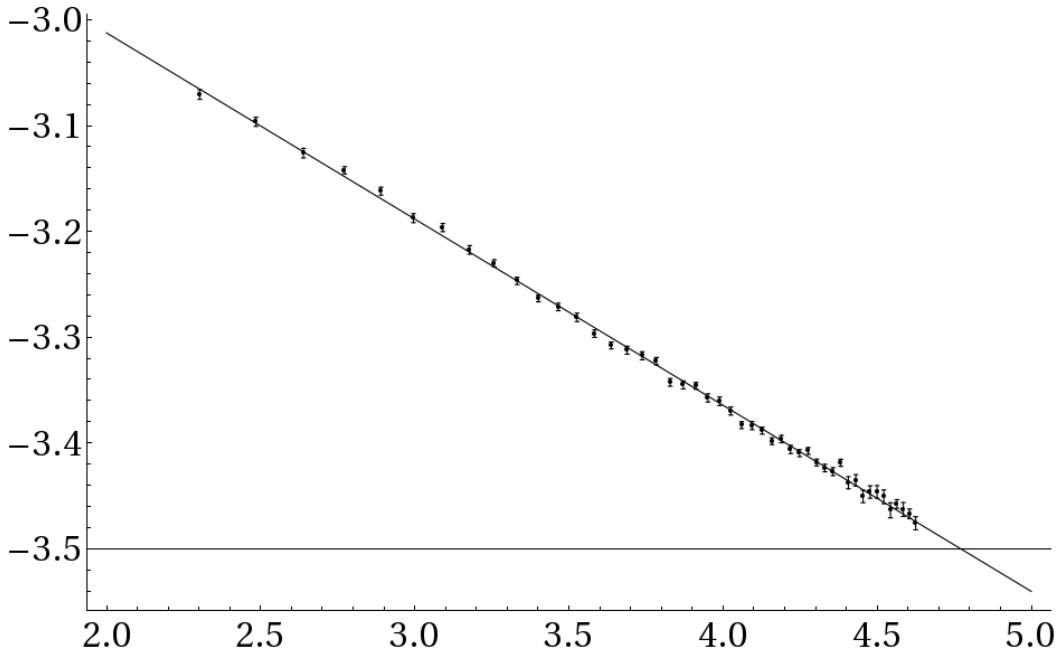

Figure 4.5: Plot of $\log\left[-G\left(\frac{L}{2}, L, 0\right)\right]$ as a function of $\log L$ together with the fitted line.

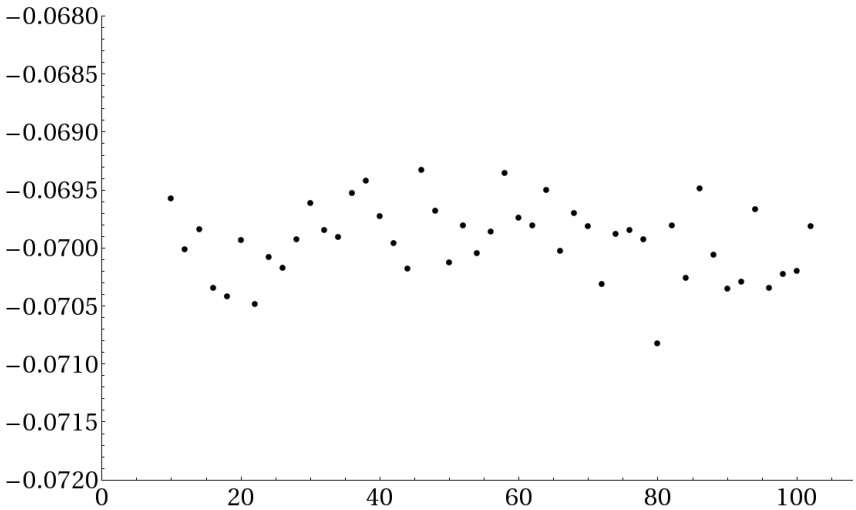

Figure 4.6: Plot of $G\left(\frac{L}{2}, L, 0\right)/L^{\alpha}$ as a function of L.

with $a(L) = AL^{\alpha-2}$ and $b(L) = BL^{\alpha}$.

There is a gap for small x and small L, which was expected and suggests it can be appropriate to add some corrections to the (4.24).

It is important to observe that any estimate of α is affected by errors of two different natures: statistical and systematic. The first component comes from

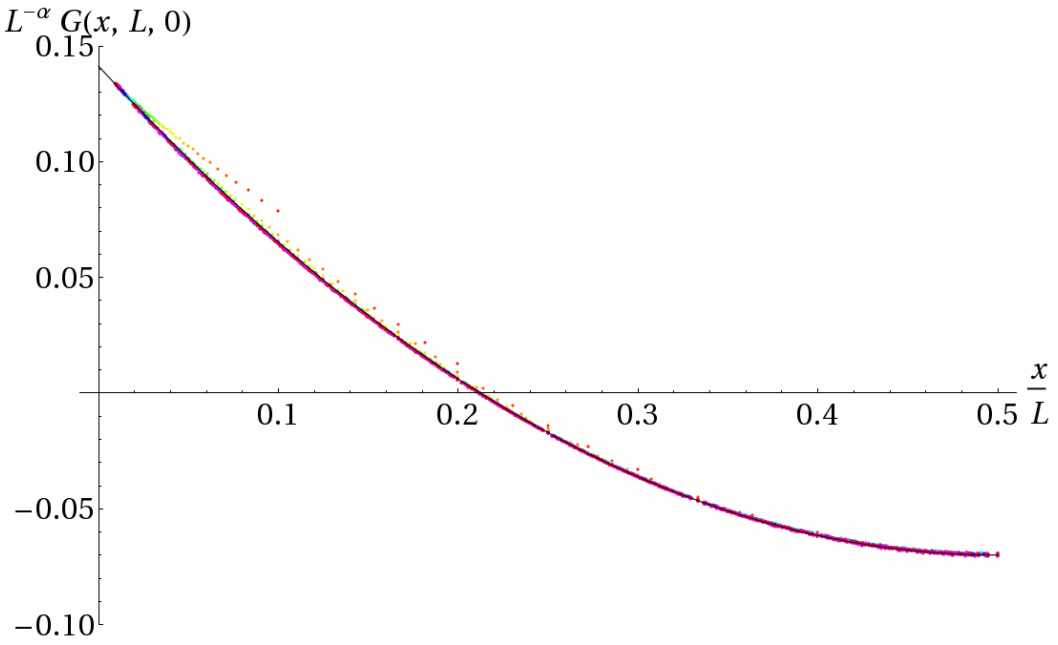

Figure 4.7: Plot of $G\left(x, L, 0\right)/L^{\alpha}$ as a function of x/L. In black, the plot of $f(x) = a \cdot \left(x - \frac{1}{2}\right)^2 + b$, with $a = -0.06997$ and $b = 0.8452$.

the statistical fluctuations that our points exhibit and it can be reduced by making a larger number of simulations. The second component, on the other hand, is due to the ansatz we are making about the formula. By using different curves, we can try to better fit our data, and for this reason in the following we will make some assumptions about the shape of $f\left(\frac{x}{L}\right)$.

Ansatz on the corrections

The hypothesis we make on the form of the correction is the following:

$$G(x, L, 0) \approx L^{\alpha} f\left(\frac{x}{L}\right) + L^{\alpha-2} h\left(\frac{x}{L}\right), \tag{4.30}$$

First of all, we make a test at the point $x = L/2$.

$$G(L/2, L, 0) \approx L^{\alpha} f\left(\frac{1}{2}\right) + L^{\alpha-2} h\left(\frac{1}{2}\right) = L^{\alpha}(B + QL^{-2}), \tag{4.31}$$

We found $\alpha = -0.1766$, $B = -0.07011$ and $Q = 0.01494$. If we plot $G(L/2, L, 0)$ as a function of L with the fitted curve, we can verify whether our assumption was correct (Fig. 4.8).

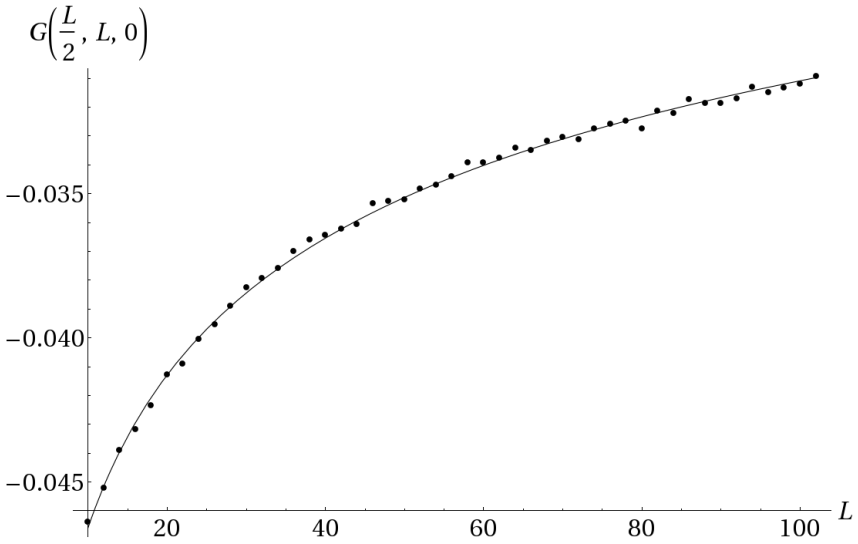

Figure 4.8: Plot of $G\left(\frac{L}{2}, L, 0\right)$ against L for all L, together with the graph of the fitted curve $y(L) = L^{\alpha}(B + QL^{-2})$, with $\alpha = -0.1766$, $B = -0.07011$ and $Q = 0.01494$.

Investigation on the shape of the corrections

We want to make a basic estimate of the shape of $h\left(\frac{x}{L}\right)$. In order to do so, we fix the parameter α at the approximate value we have found in (4.27) and consider $G(x, L, 0)$ at different values of $x = mL$, $m \in \left(0, \frac{1}{2}\right]$. For each m, we take $G(mL, L, 0)$ at different L (those L such that mL is integer) and make a fit

$$G(mL, L, 0) = aL^{\alpha} + bL^{\alpha-2} \tag{4.32}$$

where the fit parameters we are computing are the values of f and h at the point m.

Finally, we plot the results so obtained as m changes and examine the graphs (Fig. 4.9 and Fig. 4.10).

For $f(m)$, we find again the parabola we expected (Fig. 4.9), while the graph obtained for $h(m)$ is as shown in Fig. 4.10. We see that these points also lie approximately on a curve of the shape of a parabola having vertex at $m = \frac{1}{2}$. We then write

$$f\left(\frac{x}{L}\right) = A\left(\frac{x}{L} - \frac{1}{2}\right)^2 + B \tag{4.33}$$

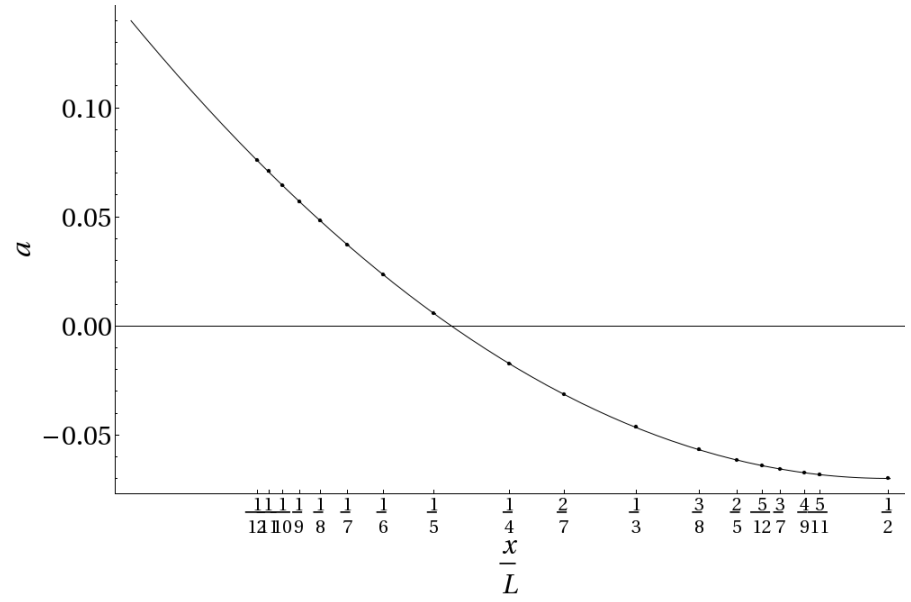

Figure 4.9: Plot of the fit parameter a in (4.32) at different values of m. The equation of the fitted curve shown is $y = -0.06987 + 0.8391\left(x - \frac{1}{2}\right)^2$.

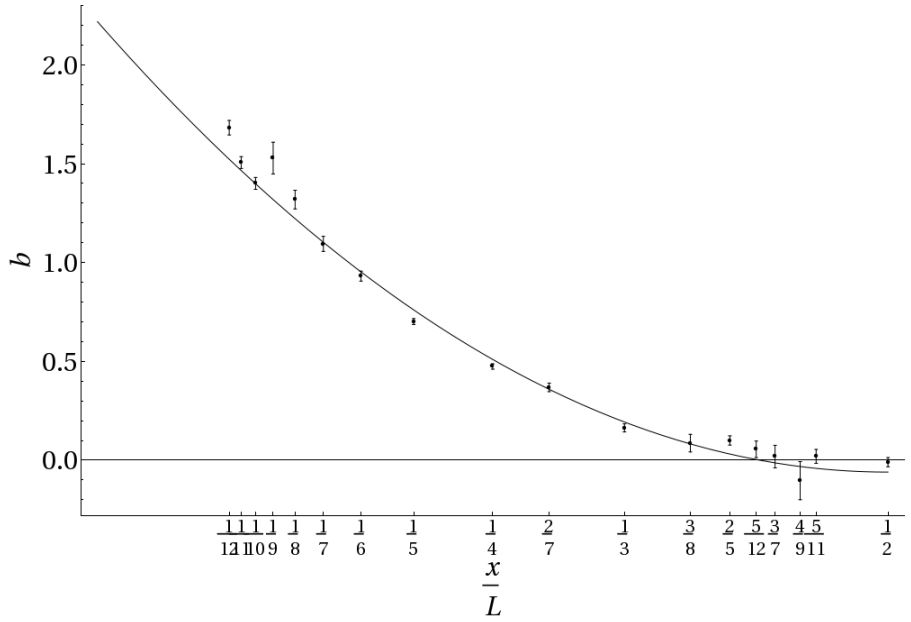

Figure 4.10: Plot of the fit parameter b in (4.32) at different values of m. The equation of the fitted curve shown is $y = -0.06010 + 9.114\left(x - \frac{1}{2}\right)^2$.

and

$$h\left(\frac{x}{L}\right) = C\left(\frac{x}{L} - \frac{1}{2}\right)^2 + D \tag{4.34}$$

Final ansatz on the form of the correlation function and calculation of the fit parameters

At this stage, we can write a second ansatz that takes into account the previous correction.

$$\begin{aligned} G(x, L, 0) &\approx L^{\alpha} f\left(\frac{x}{L}\right) + L^{\alpha-2} h\left(\frac{x}{L}\right) \\ &\approx L^{\alpha}\left[A\left(\frac{x}{L} - \frac{1}{2}\right)^2 + B + L^{-2}\left(C\left(\frac{x}{L} - \frac{1}{2}\right)^2 + D\right)\right]. \end{aligned} \tag{4.35}$$

In order to get the best estimate for all parameters at play in (4.35), the last step is to put together all the information gathered so far and make a non-linear fit for our correlation function with all available data, that is for all x and L.

In Tab. 4.1, we present the values obtained from the fit with the two functions:

$$f_1\left(\frac{x}{L}\right) = L^{\alpha}\left[A\left(\frac{x}{L} - \frac{1}{2}\right)^2 + B\right] \tag{4.36}$$

$$f_2\left(\frac{x}{L}\right) = L^{\alpha}\left[A\left(\frac{x}{L} - \frac{1}{2}\right)^2 + B + L^{-2}\left(C\left(\frac{x}{L} - \frac{1}{2}\right)^2 + D\right)\right] \tag{4.37}$$

with their respective $\tilde{\chi}^2$ tests.

The data we have analysed come from simulations on $5 \cdot 10^4$ instances of optimal marriage for all L, $10 \leq L \leq 80$, and $2 \cdot 10^4$ instances for all L, $81 \leq L \leq 102$.

Our final estimate of the parameter α is then

$$\alpha = -0.1752 \pm 0.0003. \tag{4.38}$$

In Fig. 4.11 we show the three-dimensional plot of the points at all x and L (blue), together with the graph of the fitted curves for the two functions considered (red).

	f_1	f_2
α	-0.1862 ± 0.00016	-0.1752 ± 0.0003
A	0.8805 ± 0.0006	0.837 ± 0.001
B	-0.07294 ± 0.00005	-0.0697 ± 0.0001
C	–	10.2 ± 0.1
D	–	-0.12 ± 0.01
$\tilde{\chi}^2$	18.73	2.63

Table 4.1: Table of the fit parameters for the functions in (4.36) and (4.37).

4.2.6 Behaviour near the critical point: definition and study of the correlation length

We want to give a good definition of the correlation length ξ_L at finite size L and analyse its behaviour in the proximity of the critical point in order to verify whether the FSS relations hold.

The first candidate of a correlation length we have examined is the inverse of the fit parameter c in the equation of the correlation function as a hyperbolic cosine:

$$G(x, L, t) = a \cdot \cosh\left[c \cdot \left(x - \frac{L}{2}\right)\right] + b \tag{4.39}$$

where $a = a(L, t), b = b(L, t)$ and $c = c(L, t)$.

$$\xi_1(L, t) = (c(L, t))^{-1} \tag{4.40}$$

As we observed, this definition is not applicable at $t = 0$ since the parameter c tend to vanish as the curve becomes a parabola, and then a different approach becomes necessary.

No matter what definition we choose, the bulk correlation length must be such as to diverge for $t \to \infty$. Moreover, as explained in Sect. 4.1.4, we can replace the dependence of a generic observable $\mathcal{O}_L(t)$ on $\xi_\infty(t)$ with the

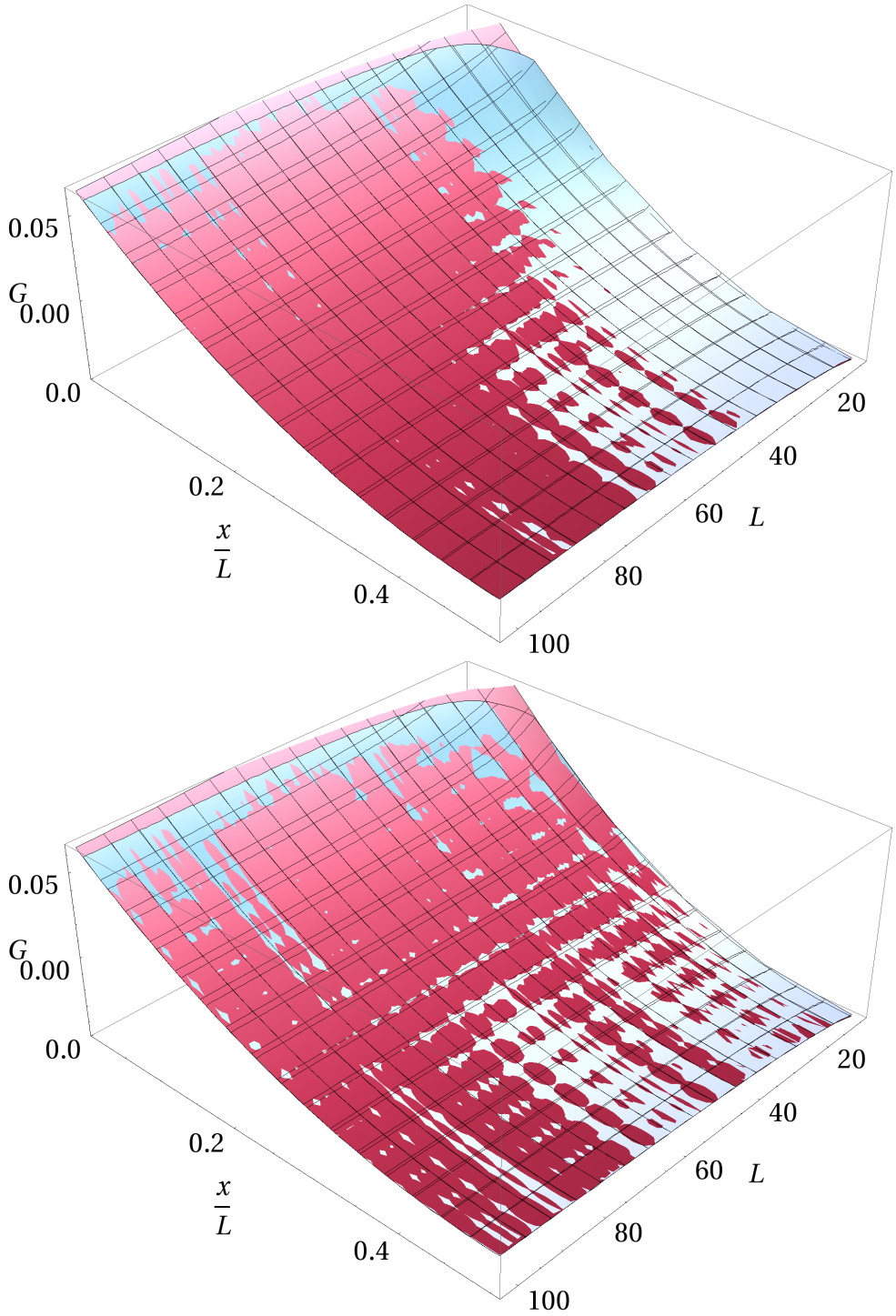

Figure 4.11: 3-D plot of the points at all x and L (blue), together with the graph of the fitted curves (red) for $f_1\left(\frac{x}{L}\right)$ (top) and $f_2\left(\frac{x}{L}\right)$ (bottom).

dependence on $\xi_L(t)$, where $\xi_L(t)$ is any appropriate correlation length at finite size L.

We then observe that in all the experimental curves, there is one (and only one) point $\bar{x}$ such that

$$G(\bar{x}, L, t) = 0 \tag{4.41}$$

If the scaling ansatz holds, (4.41) is equivalent to

$$\mathcal{F}\left(\frac{\bar{x}}{L}, \frac{\xi_\infty}{L}\right) = 0 \tag{4.42}$$

or

$$\tilde{\mathcal{F}}\left(\frac{\bar{x}}{L}, \frac{\xi_L}{L}\right) = 0 \tag{4.43}$$

That is, $\frac{\bar{x}}{L}$ is a function of $\frac{\xi_\infty}{L}$ or $\frac{\xi_L}{L}$

$$\frac{\bar{x}}{L} = H\left(\frac{\xi_\infty}{L}\right) = K\left(\frac{\xi_L}{L}\right) \tag{4.44}$$

and therefore we can take $\bar{x}$ as a definition of correlation length at finite size.

The value of $\bar{x}$ is determined in the following way. We select the experimental point nearest to zero, one point to its immediate left and one point to its immediate right. We then make a fit of these three points with a polynomial of degree two and take the intersection of the resulting curve with the x-axis.

By making use of the value of $\bar{x}$ and some properties of the curves, we can give another estimate of the correlation length 4.40, independent from the fitted curve. We proceed as follows. We consider a slightly, but equivalent, parametrisation of the function 4.39, that is:

$$G(x, L, t) = a\left\{\cosh\left[\frac{1}{\xi}\left(x - \frac{L}{2}\right)\right] - b\right\}. \tag{4.45}$$

We then observe that the curves have zero integral (we see this by analysing the experimental data), then:

$$\int_0^L G(x, L, t) = a\left[2\left(\sinh\frac{L}{2\xi}\right)\xi - bL\right] = 0 \tag{4.46}$$

and use this property to write

$$b = \frac{2\xi}{L}\sinh\frac{L}{2\xi}. \tag{4.47}$$

Moreover, we have defined $\bar{x}$ as the point where $G(\bar{x}, L, t) = 0$, then

$$\cosh\left(\frac{\bar{x}}{\xi} - \frac{L}{2\xi}\right) = b, \tag{4.48}$$

which leads to the new definition of correlation length ξ_2:

$$\xi_2(L,t) = \left\{\xi \mid \frac{L}{2} - \xi \operatorname{arccosh}\left[\frac{2\xi}{L}\sinh\left(\frac{L}{2\xi}\right)\right] = \bar{x}\right\} \tag{4.49}$$

In Fig. 4.12 we show the ratio between the two definitions of correlation length. We see that they are approximately equivalent.

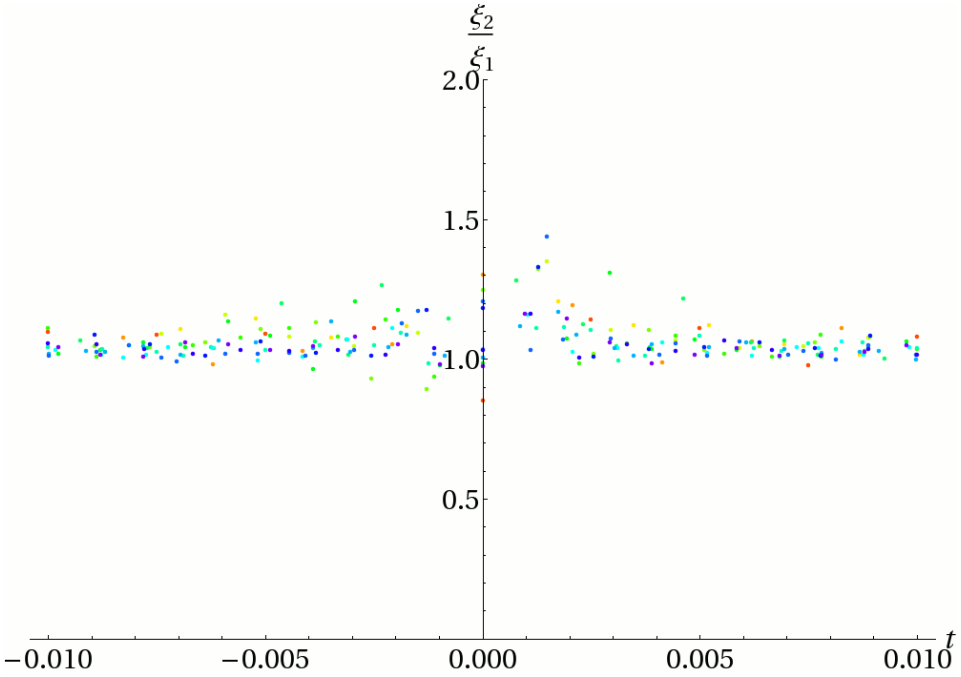

Figure 4.12: The ratio between ξ_2 and ξ_1 as a function of t. From red to purple, $L = 20, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64$.

In Fig. 4.13 and Fig. 4.14 we show the dependence of $\bar{x}/L$ and L/ξ_2 on t, while in Fig. 4.15 we show the ratio between $\bar{x}(L,t)$ and $\xi_2(L,t)$.

At the critical point $t = 0$

In Fig. 4.16 we show $\bar{x}(t=0)/L$ as a function of L. The correlation length is a typical size of the problem and at $t = 0$ must diverge for $L \to \infty$. We found that

$$\bar{x} \sim L \qquad \text{for} \quad L \to \infty \tag{4.50}$$

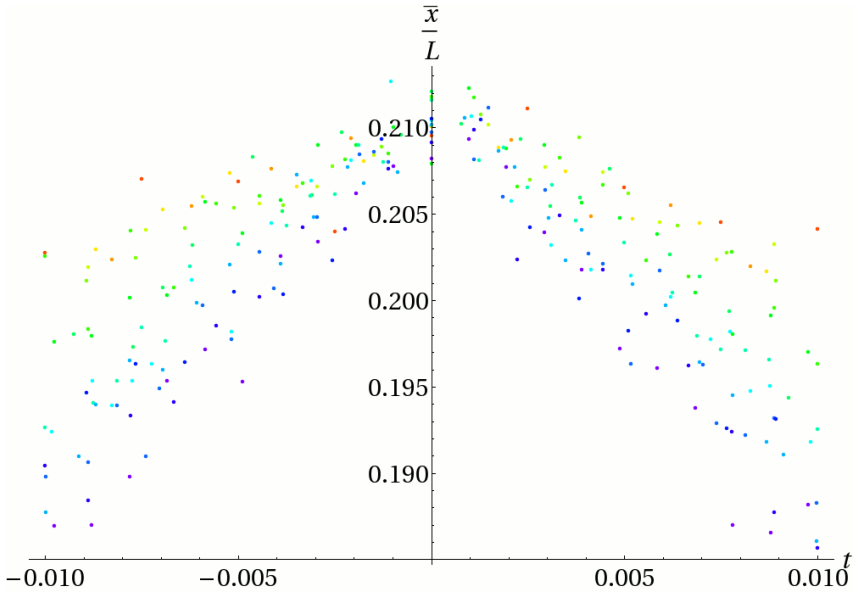

Figure 4.13: $\bar{x}(L,t)/L$ as a function of t. From red to purple, $L = 20$, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64.

Moreover, we observe that as L tends to ∞, the ratio $\bar{x}/L$ tends to a quantity very close to the value of the intersection with the x-axis of the exact solution at $t = 0$ in one dimension, as calculated in 3.1.3 (cf (3.55)).

$$\frac{\bar{x}}{L} \to k = \frac{\bar{x}_{\text{dim1}}}{L} = 0.211325 \qquad \text{for} \quad L \to \infty \tag{4.51}$$

Near the critical point, $t \sim 0, t \neq 0$

Firstly, we use the (4.10) to replace ξ_∞/L with $\bar{x}/L$ in (4.6) and write

$$G(x,L,t) = L^\alpha F_G\left(\frac{x}{L}, \frac{\bar{x}(L,t)}{L}\right). \tag{4.52}$$

Now we can look at what happens at $x = L/2$ by plotting $G(L/2,L,t)/L^\alpha$ against $\bar{x}(L,t)/L$, with α as calculated in Sect. 4.2.5. In Fig. 4.17 we see that the points fall approximately on the same curve, regardless of L. There are some fluctuations within the data for the same L at different t.

Next, following Eq. (4.8), we plot $G(L/2,L,t)/L^\alpha$ against $tL^{1/\nu}$ for different values of ν. When all points collapse into a single curve regardless of L, we

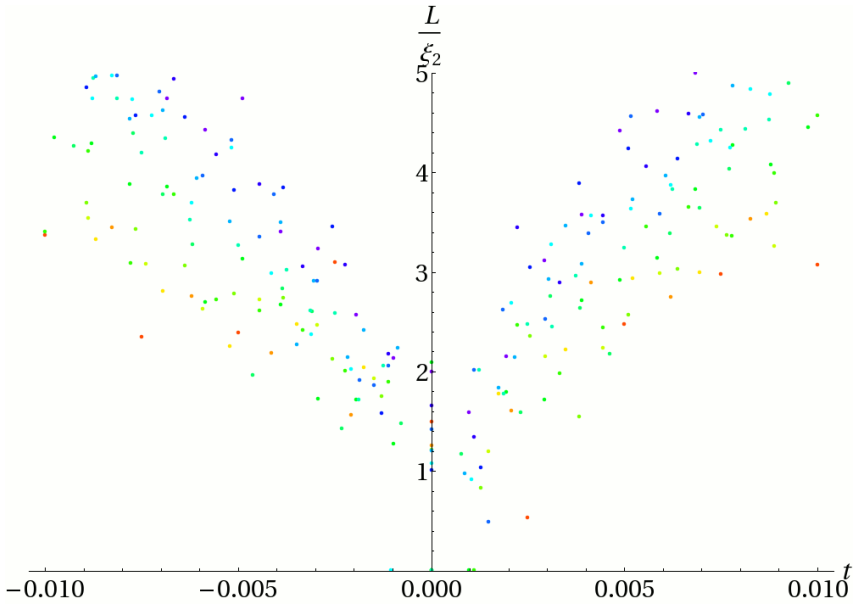

Figure 4.14: $L/\xi_2(L,t)$ as a function of t. From red to purple, $L = 20$, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64.

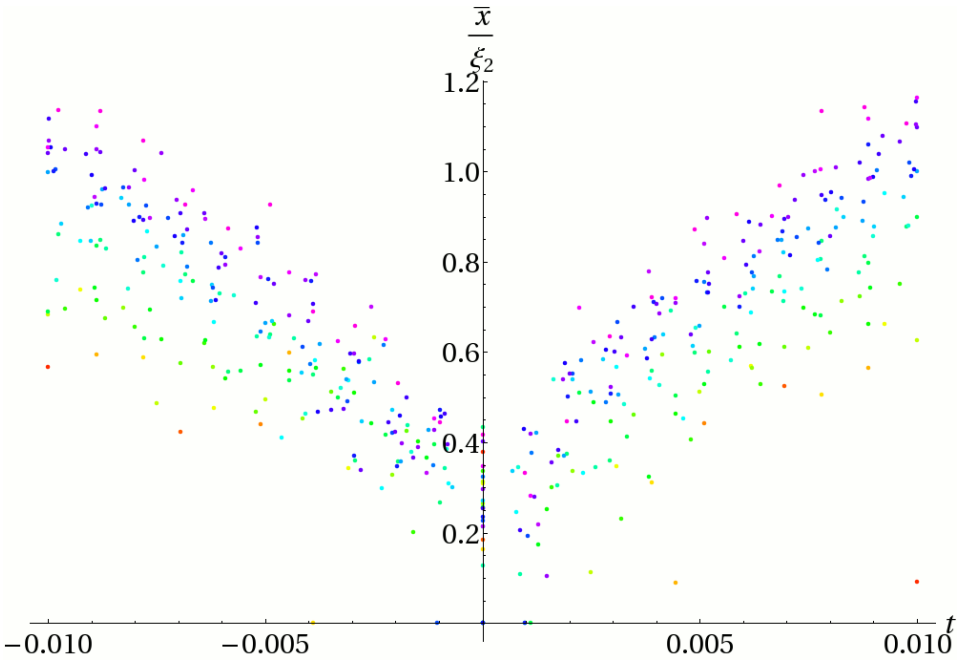

Figure 4.15: $\bar{x}(L,t)/\xi_2(L,t)$ as a function of t. From red to blue, $L = 20$, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64.

have found the correct value of ν. The results are shown in Fig. 4.18 and seem

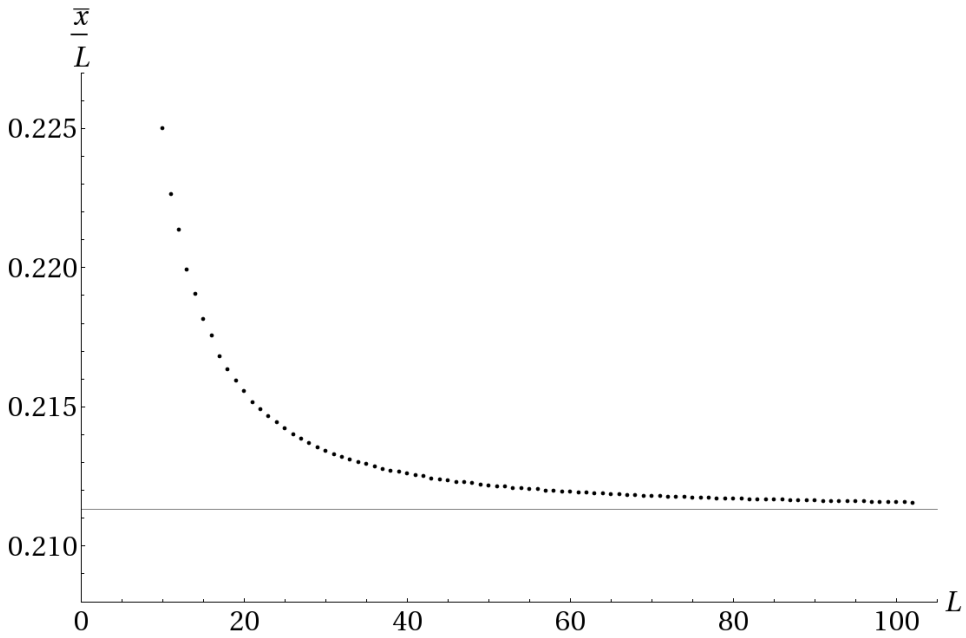

Figure 4.16: $\bar{x}(t=0)/L$ as a function of L. The line shown is at $k = 0.211325$ and represents the value of the intersection with the x-axis of the theoretical curve at $t = 0$ in one dimension.

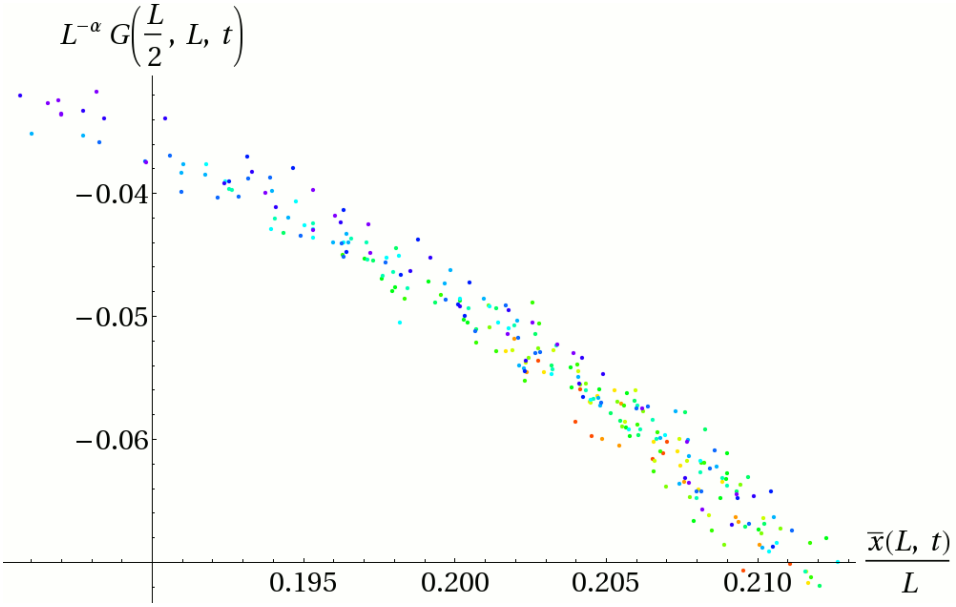

Figure 4.17: $G(L/2, L, t)/L^{\alpha}$ as a function of $\bar{x}(L,t)/L$. From red to blue, $L =$ 20, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64.

to suggest a value of the critical exponent ν of

$$\nu \approx 1. \tag{4.53}$$

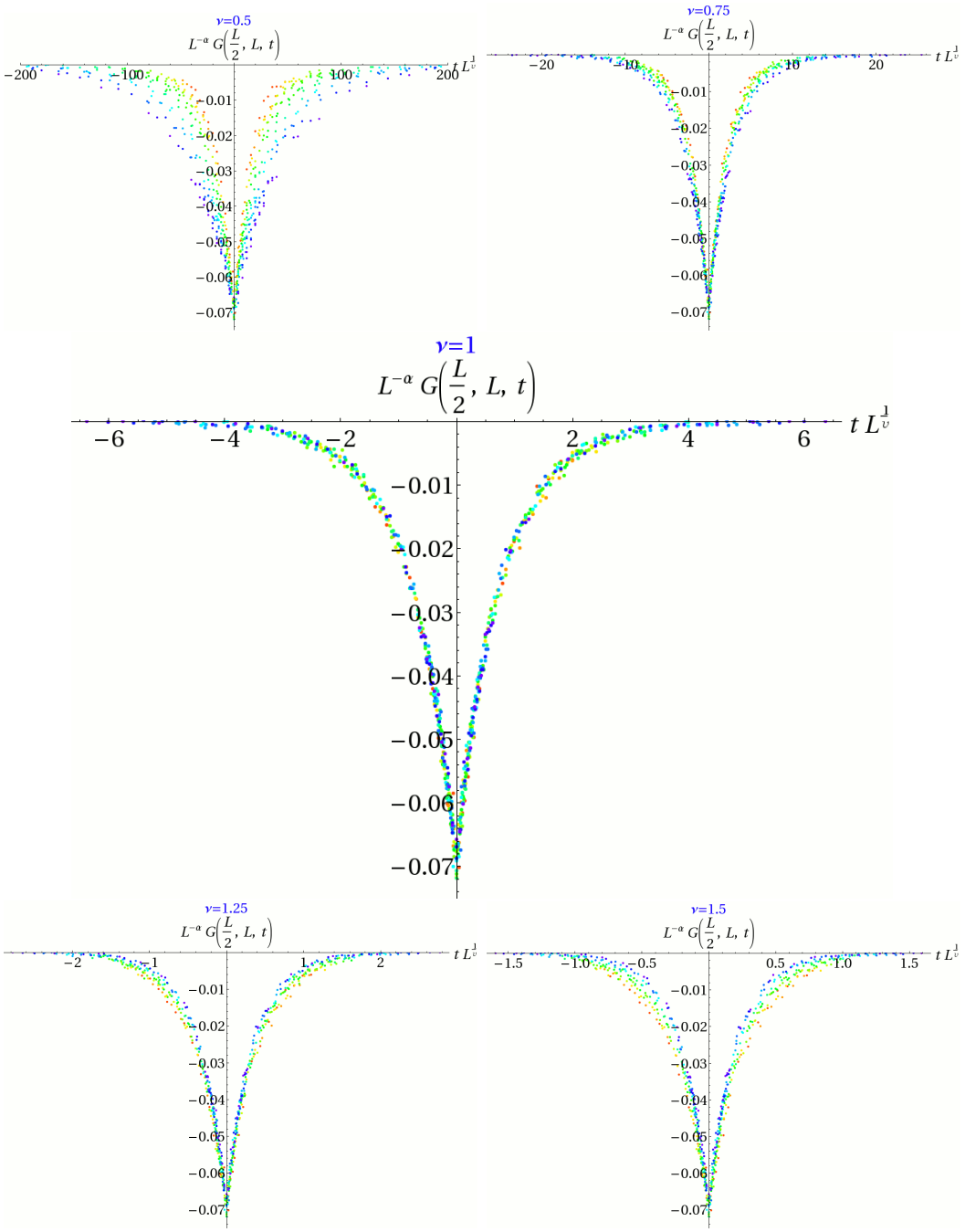

Figure 4.18: $G(L/2, L, t)/L^{\alpha}$ as a function of $tL^{1/\nu}$, for different values of ν. From red to purple, $L = 20$, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64.

Another test to find ν is by plotting $\bar{x}/L$ as a function of $tL^{1/\nu}$, as (4.42) is equivalent to

$$\overline{\mathcal{F}}\left(\frac{\bar{x}}{L}, tL^{1/\nu}\right) = 0. \tag{4.54}$$

We show the results in Fig. 4.19.

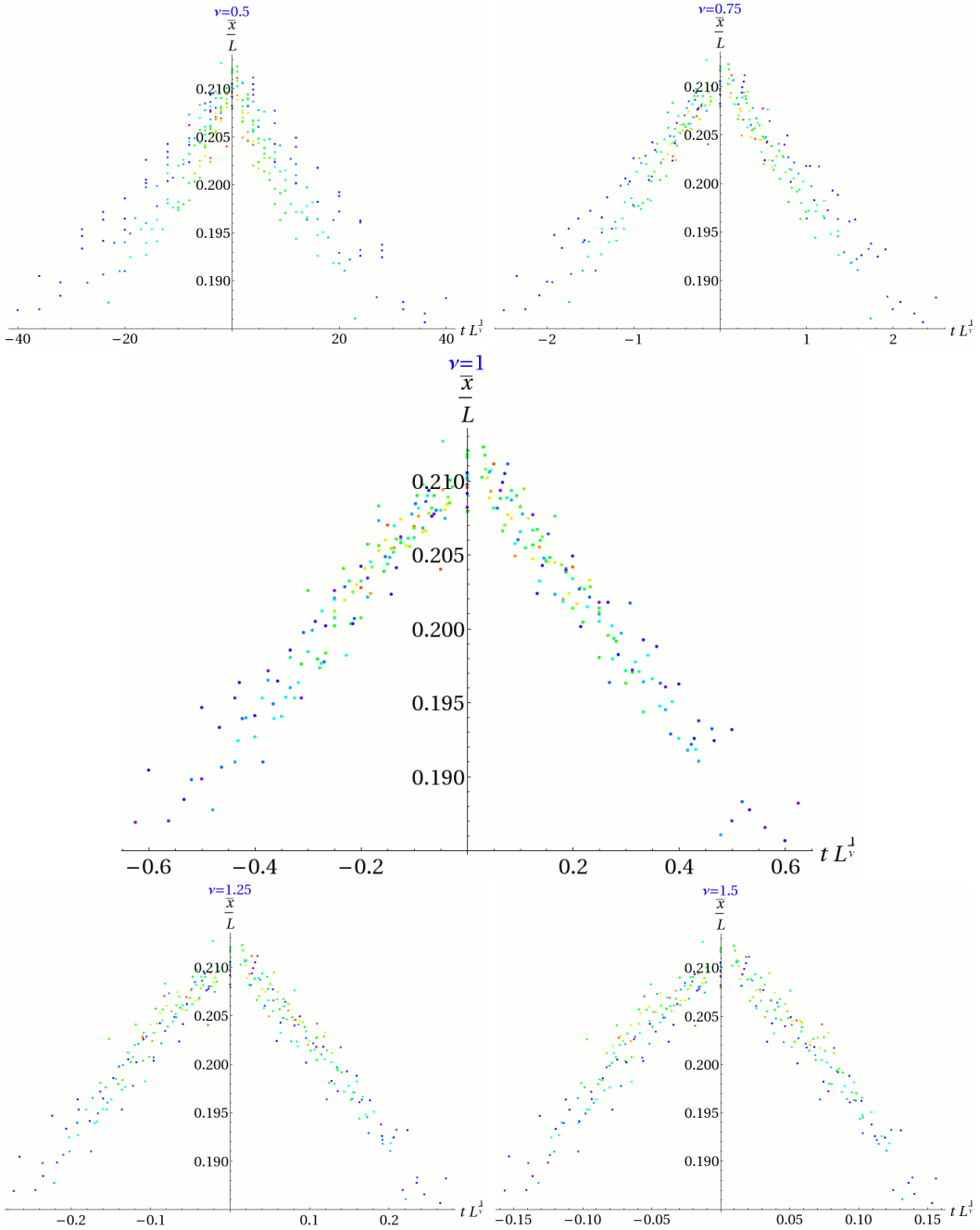

Figure 4.19: $\bar{x}/L$ as a function of $tL^{1/\nu}$, for different values of ν. From red to blue, $L = 20, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64$.

Similarly, in Fig. 4.20, we show the plots of $L/\xi_2 = L/\xi_2(L,t)$ as a function of $tL^{1/\nu}$.

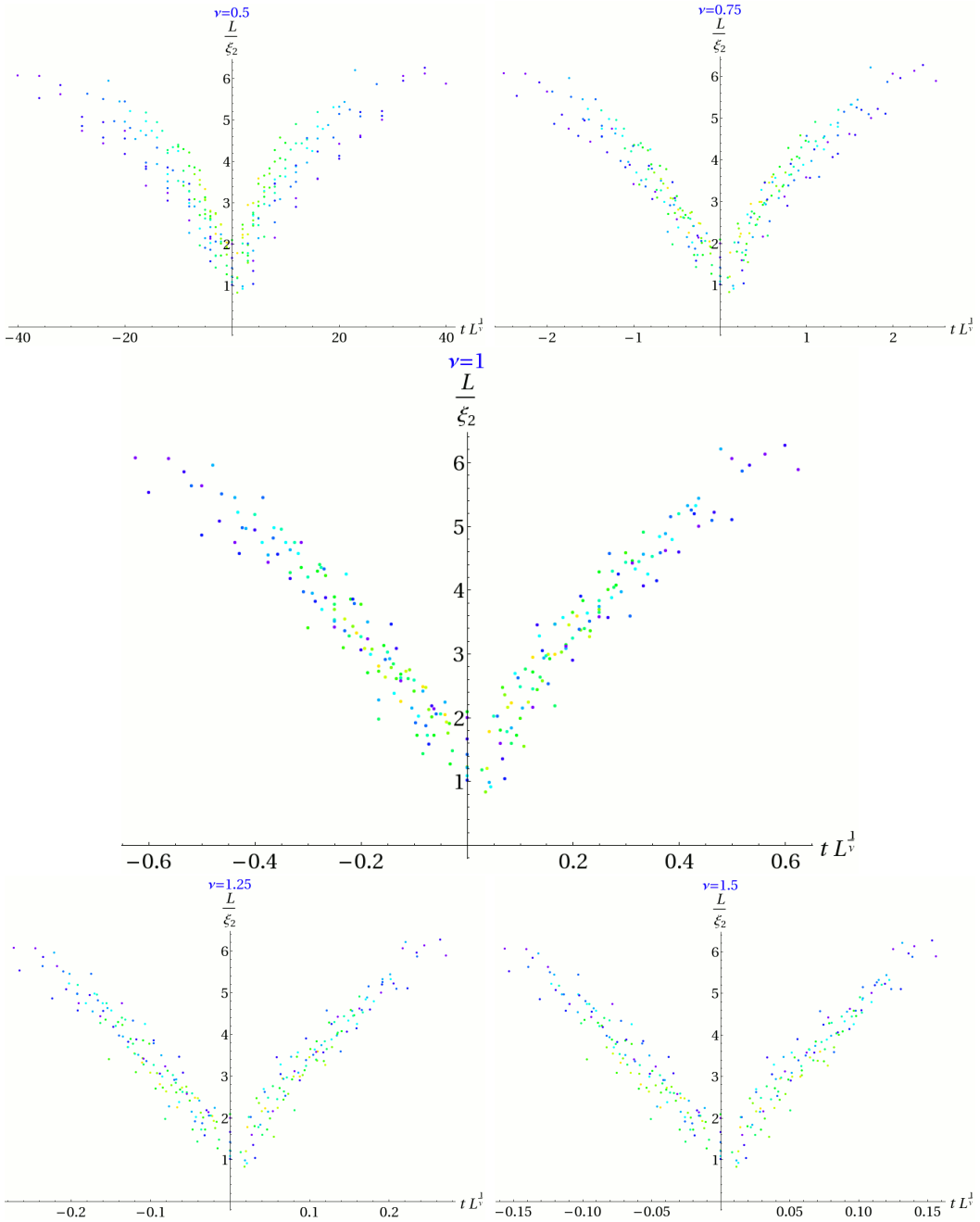

Figure 4.20: L/ξ_2 as a function of $tL^{1/\nu}$, for different values of ν. From red to blue, $L = 20, 22, 24, 26, 28, 30, 32, 36, 40, 44, 48, 52, 56, 60, 64$.

We see that the data for $\bar{x}$ and ξ_2 are still very noisy and this does not allow us to extract from them useful information on the correlation length. Further investigation is necessary in this area.

5

Probability density of the distance between matched pairs

In this chapter we shall give an overall phenomenological description of the behaviour of the length of the vectors that link each grid point to its respective matched Poisson point. That is to say, the quantity that in the previous chapters was referred to as $|\vec{\varphi}(\vec{x})|$. For simplicity, in the following we will merely call it "X":

$$X = |\vec{\varphi}(\vec{x})| \,.$$

This problem has already been addressed theoretically (for example, by Holroyd et al. [19]) for a generic perfect matching (bipartite or not), and also for the stable marriage and other matching schemes, while many papers have been written on the mean distance between matched pairs (see below, Sect. 5.1.1). Here, we will focus on a numerical analysis in the specific case of the optimal marriage as described in Chap. 2.

All data presented come from samples of 10^3 instances of optimal marriage, and, as a consequence of the large amount of information produced in our numerical experiments, it would be difficult (and not very significant) to consider every single value of the quantity X. For this reason we decided, instead, to focus on the probability density function, and to do so by making a histogram of our data.

In the following graphics, the point at $x = X$ represents the number of vectors whose module falls between X and $X + \Delta$, where Δ is the width of the bins. In two dimensions, the width we considered was a hundredth of the

lattice constant: $\Delta_2 = 1/100$, while in one dimension, where, as we shall see, the edge lengths tend to fall in a wider interval, we had to take a bin width of one tenth of the lattice constant: $\Delta_1 = 1/10$.

We also want to point out that, as already explained in Chap. 3, the optimal marriage considered in one dimension is the one which minimises the sum of the squares of the edge lengths, instead of the usual edge lengths.

5.1 Numerical results in two dimensions

5.1.1 At the critical point

In Fig. 5.1 we show the histogram of our measurements at density $\rho = 1$ for different sizes L of the system with a bin width of $\Delta_2 = 1/100$. We also show the cumulative distribution function, that is $p(X > r)$ (Fig. 5.2).

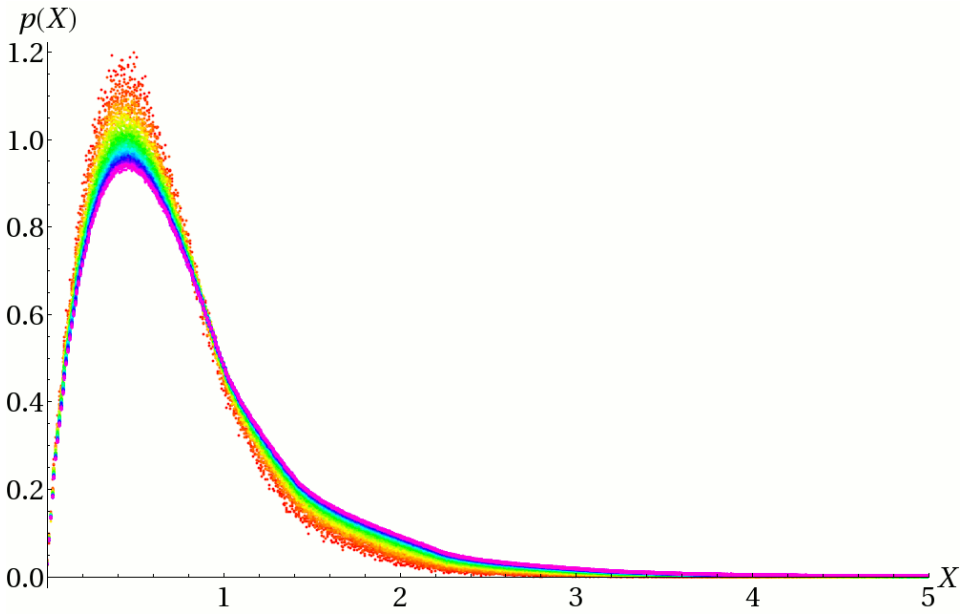

Figure 5.1: Normalised histogram of the edge length at density $\rho = 1$. From red to purple, all L between 10 and 102.

We see that the distribution has a well-defined peak that typically falls between 0.4 and 0.5 and does not seem to depend on the value of L (Fig. 5.3).

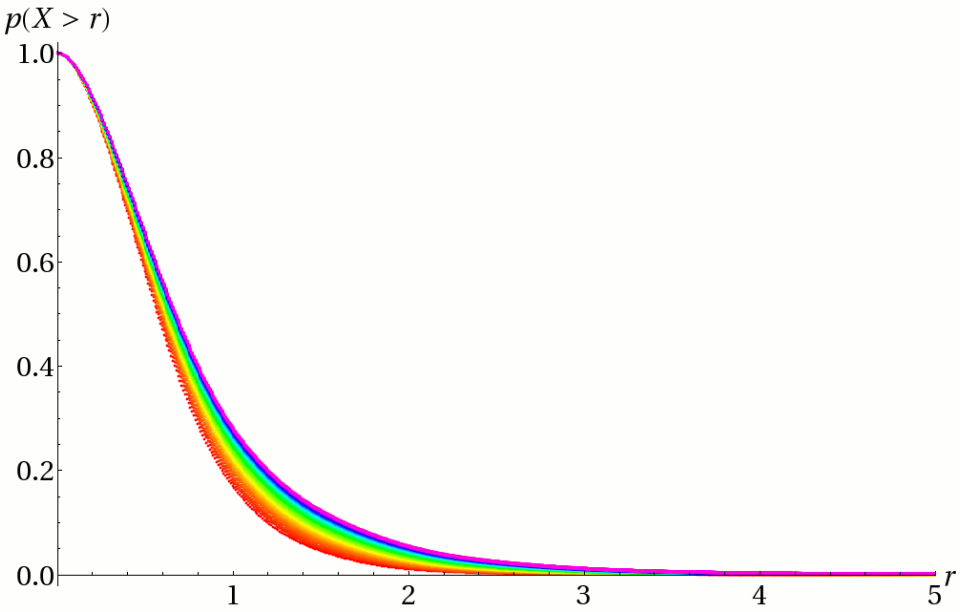

Figure 5.2: Cumulative distribution function. From red to purple, all L between 10 and 102.

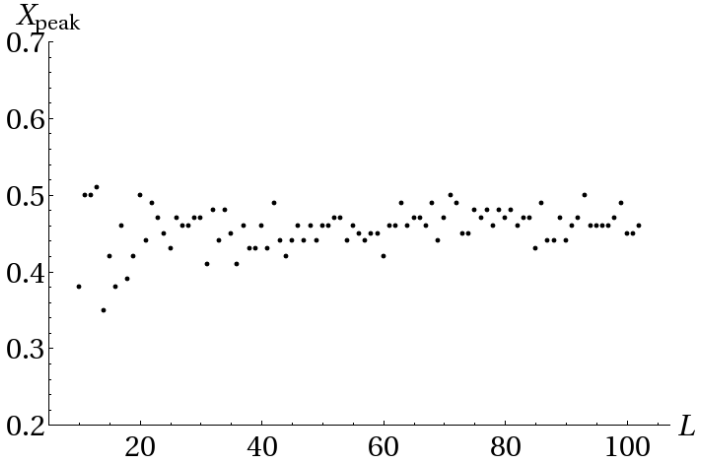

Figure 5.3: Position of the peak of the previous curves as a function of the size of the system L.

The distribution tends to widen as L increases, as can be estimated by the growth of the mean value of X and the variance[1] $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$. (Fig. 5.4).

Finally, we observe the presence of small bends in the curves at precise points.

[1]In order to get a more precise number, the mean values of X and X^2 were calculated as the average over every single measure of X, rather than from the histogram.

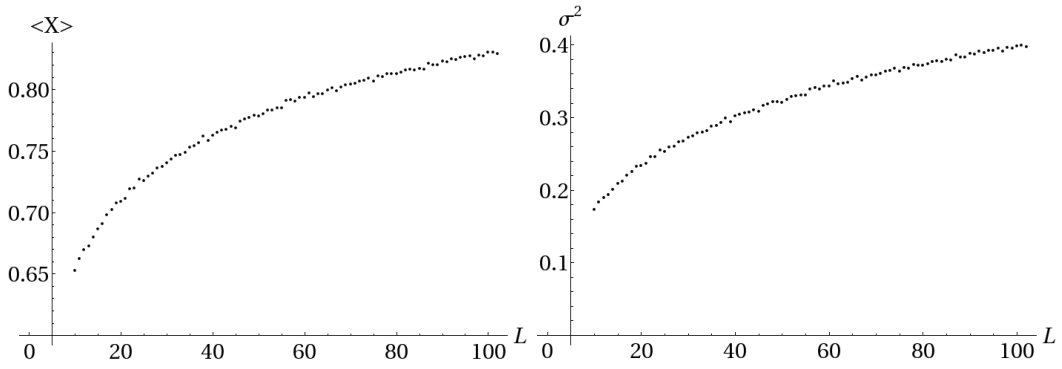

Figure 5.4: Mean value of X, $\langle X \rangle$, and variance $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$ as a function of the size of the system L.

They can be explained by the characteristics of the grid: the distance $X = \sqrt{i}$, $i \in \mathbb{Z}^+$, corresponds to the position of another grid point, which is a strong "attractor" for Poisson points in the immediate neighbourhood.

Mean edge length

Studying the mean edge length is clearly the same as studying the energy or, in other words, the *transportation cost* of the matching, defined as

$$T_N = \sum_{i=1}^{i=N} X_i, \tag{5.1}$$

where $N = L^2$, is the number of pairs, X_i is the edge length of the ith matched pair, and, obviously, $\langle X \rangle = T_N/N$.

This problem has been profusely treated in the past (see, for example, [21], [22], [23], [24], and, for the related problem of the minimax matching length, [25]) and it has been proved ([21]) that, for N couples in the unit square, there exist two constant c, C such that[2],

$$c(N \log N)^{1/2} < T_N < C(N \log N)^{1/2} \tag{5.2}$$

or, equivalently,

$$\langle X \rangle = O(\sqrt{\log N/N}). \tag{5.3}$$

[2]The proof is given for two collections of random points, instead of a set of random points and a set of grid points, but this changes only constant multiplicative factors, as shown in [22].

The above result is related to a 1×1 square. By rescaling to $L \times L$, we can finally write:

$$\langle X \rangle = O(\sqrt{\log L^2}) \tag{5.4}$$

In order to verify this assumptions, we made a fit of our experimental data with $f(L) = a\sqrt{\log L^2} + b$. The result is

$$\langle X \rangle = 0.202\sqrt{\log L^2} + 0.216, \tag{5.5}$$

and the fitted curve is shown in Fig. 5.5.

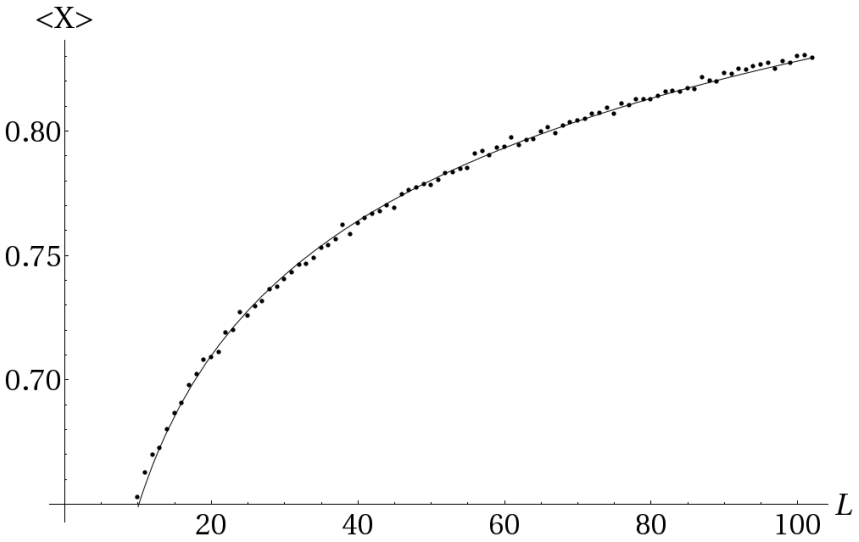

Figure 5.5: Mean value of X, $\langle X \rangle$, as a function of the size of the system L, together with the fitted curve $f(L) = a\sqrt{\log L^2} + b$, with $a = 0.202$ and $b = 0.216$.

Tail behaviour

We want to analyse how the probability density function $p(X)$ decreases as X increases. In order to do so, we take the logarithm of $p(X)$ and examine the resulting curves.

We see that, at some distance from the peak, the logarithm of the curve becomes, on first approximation, a linear function of X:

$$\log(p_L(X)) = A_L + B_L X \tag{5.6}$$

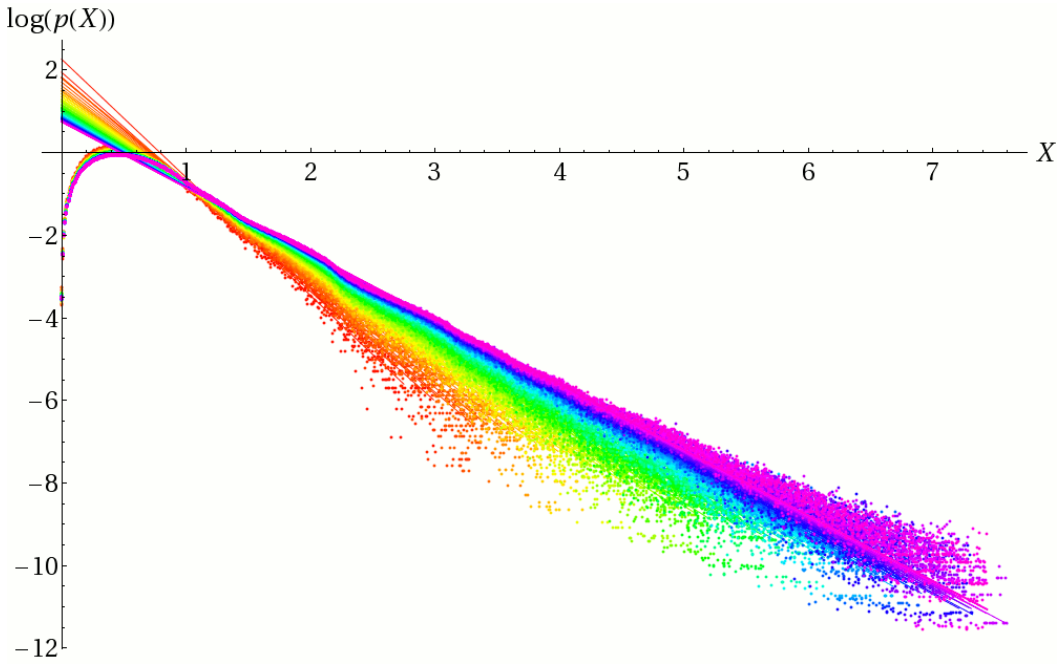

Figure 5.6: Logarithm of the probability density, shown with the fitted lines. From red to purple, all L between 10 and 102.

In Fig. 5.6 we show the points together with the fitted lines.

At different sizes L, the slope of the line changes, and this can be quantitatively evaluated by considering the value of the parameter B_L of the fitting function $y = A_L + B_L x$, with $y = \log(p_L(X))$, $x = X$, and $A_L = A(L, \rho = 1)$, $B_L = B(L, \rho = 1)$.

We found that, on first approximation, the behaviour of B_L as a function of L can be described by (see Fig. 5.7)

$$B(L, \rho = 1) = -C - D\frac{\log L}{L}. \tag{5.7}$$

with $C = 1.290$ and $D = 6.58$.

Similarly, for A_L (see Fig. 5.8):

$$A(L, \rho = 1) = F + G\frac{\log L}{L}. \tag{5.8}$$

with $F = 0.388$ and $G = 7.83$.

By taking into account both results, we found that, if X is far enough from the position of the peak ($X \gtrsim 1.5$), the probability density function behaves approximately as

$$p(X, L, \rho = 1) = F' L^{\frac{1}{L}(G - DX)} e^{-CX}. \tag{5.9}$$

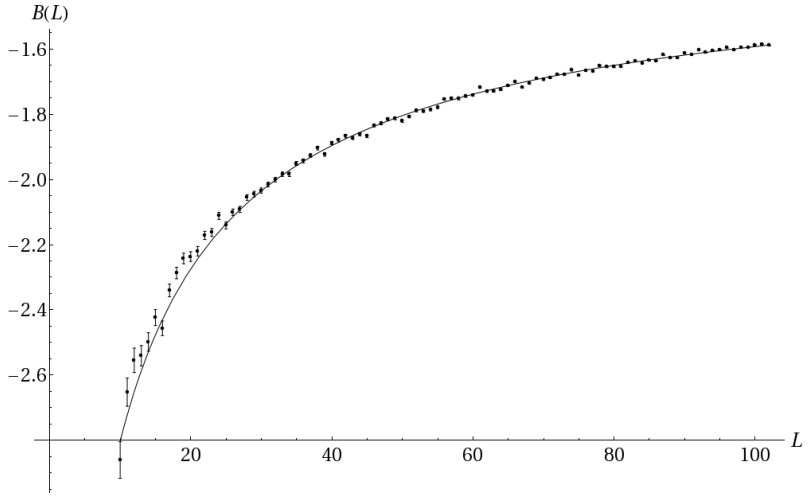

Figure 5.7: The fit parameter B_L as a function of L with the fitted curve $C + D\log(L)/L$, with $C = 1.290$ and $D = 6.58$.

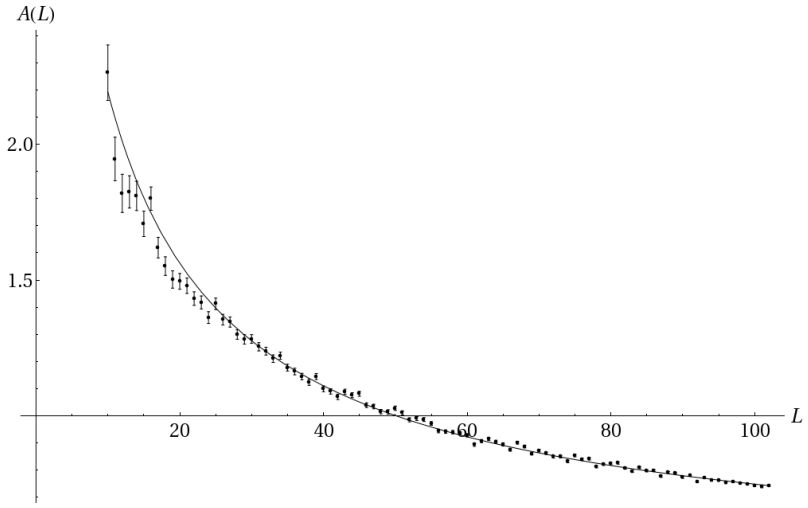

Figure 5.8: The fit parameter A_L as a function of L with the fitted curve $F + G\log(L)/L$, with $F = 0.388$ and $G = 7.83$.

A fit of all available data with the function in (5.9) gives the following results

for the fit parameters:

$$F' = 1.467 \pm 0.003$$

$$C = 1.287 \pm 0.001$$

$$D = 6.84 \pm 0.02$$

$$G = 8.27 \pm 0.04$$

$$\tilde{\chi}^2 = 3.43$$

A 3-D plot of the data and the fitted curve is shown in Fig. 5.9.

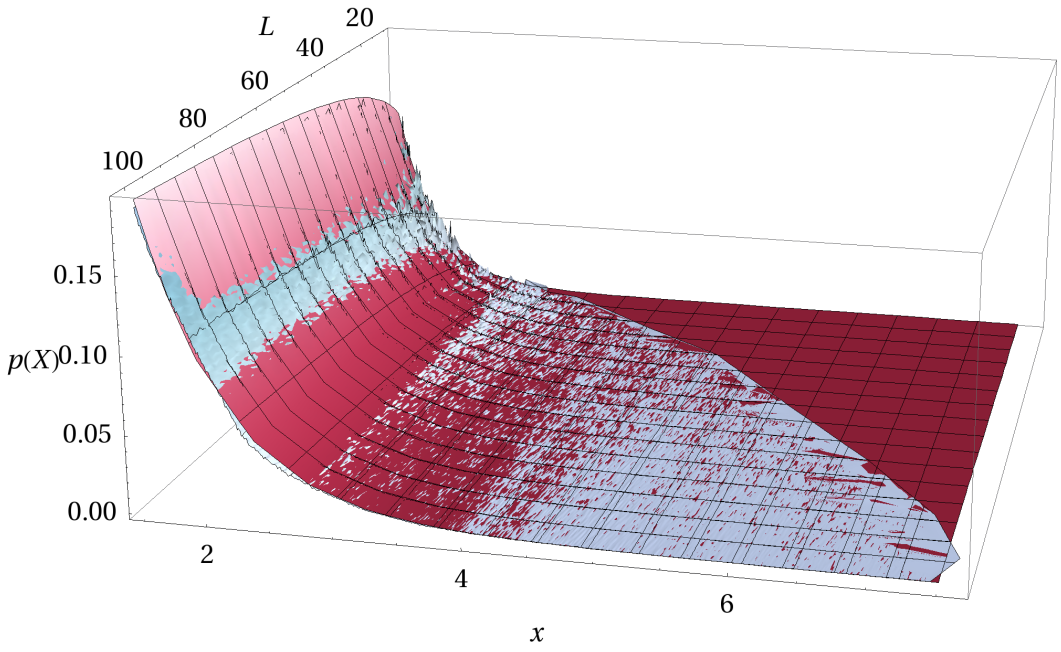

Figure 5.9: Experimental data for $p(X, L, \rho = 1)$, $X > 1.5$ (blue), shown together with the fitted curve (5.9) (red).

We also tried to perform a more precise fit, by considering the logarithm of the cumulative density function and a different fitting function:

$$\log p(r > X, L, \rho = 1) = A_L - \sqrt{A_L^2 + (B_L x)^{1.5}}. \tag{5.10}$$

In Fig. 5.10 we show the experimental data with the new fitted curves, while Fig. 5.10 shows a plot of A_L and B_L as a function of L.

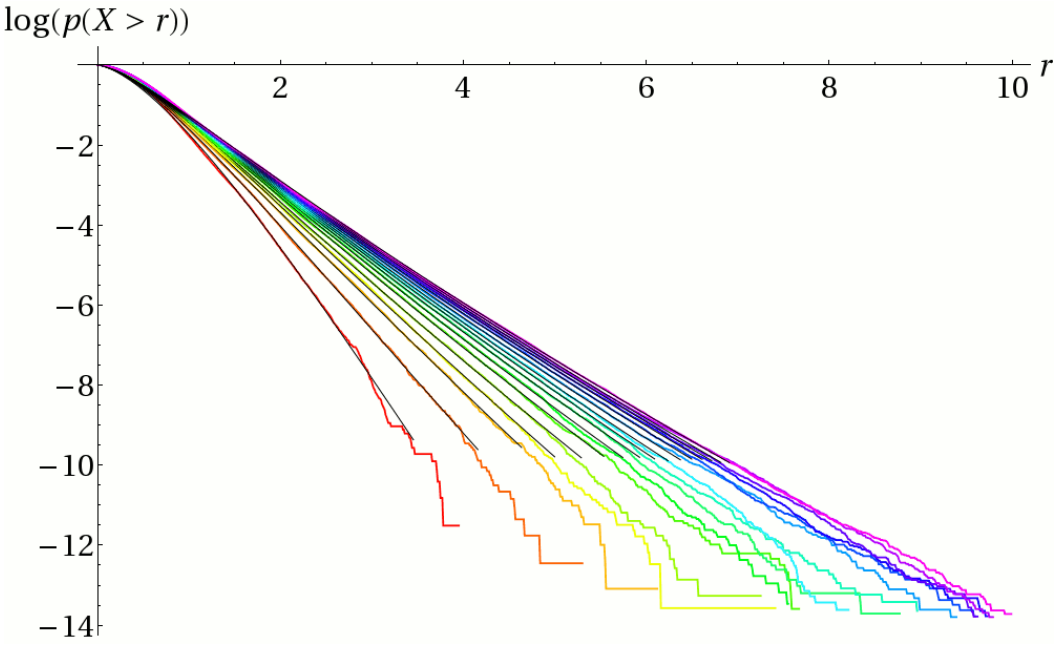

Figure 5.10: Experimental data for $\log p(r > X, L, \rho = 1)$ (in colour), together with the fitted curves (5.10) (black).

By plotting $1/A_L$ and $1/B_L$ as a function of $\log L$ (Fig. 5.10), we have estimated:

$$\frac{1}{A_L} = \alpha_0 + \alpha_1 \log L \tag{5.11}$$

$$\frac{1}{B_L} = \beta_0 + \beta_1 \log L \tag{5.12}$$

with $\alpha_0 = -0.3233$, $\alpha_1 = 0.1537$, $\beta_0 = -0.1247$ and $\beta_1 = 0.0808$ (Figures 5.11, 5.12 and 5.13).

Shape for small X

We want to have a closer look at the first points of the histogram, and try to give a first estimate of the shape of the curves.

If we consider small distances from a certain grid point, we expect that, if there is a Poisson point, probably it will be matched to the grid point considered. As random points are distributed according to a Poisson point process, if we consider the radial distribution we have

$$\mathrm{d}\, p(r_{\min} = r) = \mathrm{d}\, r \; 2\pi\, r\, e^{-\pi r^2} \tag{5.13}$$

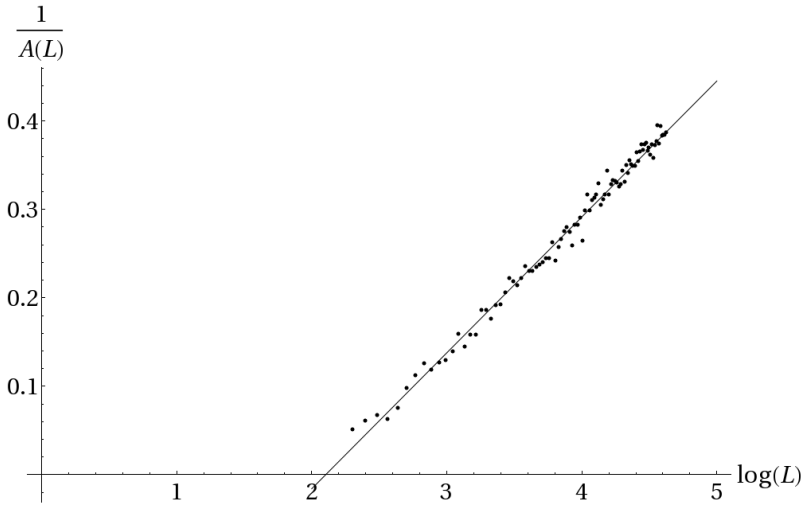

Figure 5.11: $1/A_L$ as a function of $\log L$, with the fitted curve (5.11).

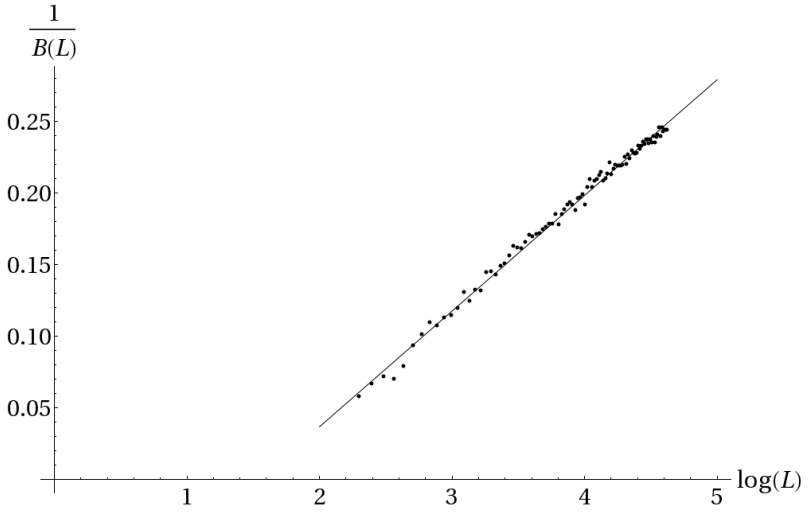

Figure 5.12: $1/B_L$ as a function of $\log L$, with the fitted curve (5.12).

and then we expect the cdf to be, on first approximation:

$$p(r_{\min} > r) = \int_r^\infty \mathrm{d}\, r'\; 2\pi\, r'\, e^{-\pi\, r'^2} = e^{-\pi\, r^2} \tag{5.14}$$

However, the situation described does not happen in all cases. If there are many long edges, and one of them passes near the potential couple, it may be energetically favourable to have two medium-length edges than a very small and a very long one. If we look at Fig. 5.14, we see that the dashed-line matching becomes preferable over the solid-line one when $A+B+2\epsilon > C+D$,

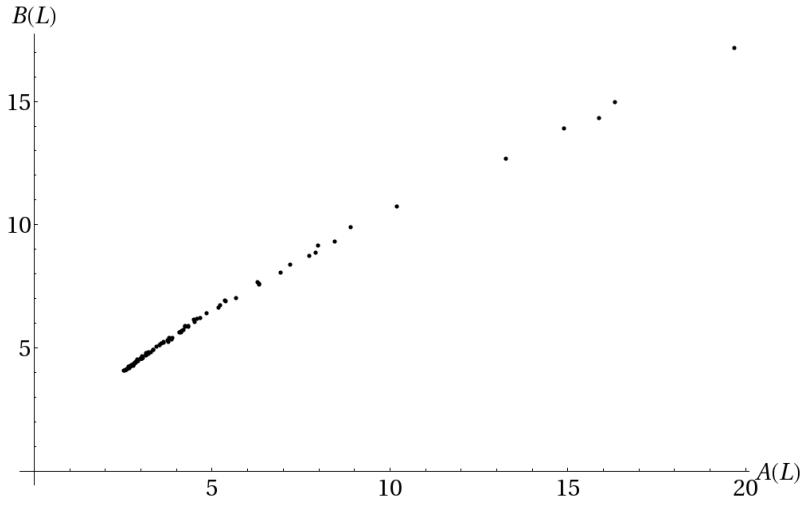

Figure 5.13: B as a function of A.

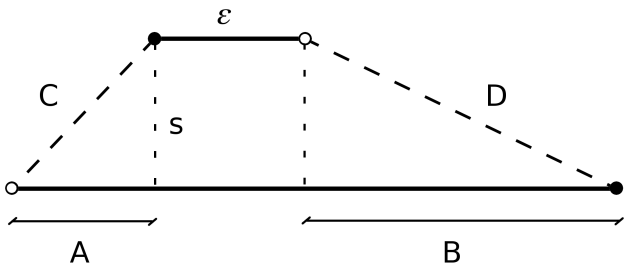

Figure 5.14: The dashed-line matching is preferable over the solid-line one when $A + B + 2\epsilon > C + D$

that is, when[3] $s \lesssim \sqrt{\epsilon}$.

These considerations lead to a correction term of order $\sqrt{r}$:

$$\mathrm{d}\, p(r_{\min} = r) = \mathrm{d}\, r \; 2\pi\, r\, e^{-\pi r^2}(1 - \alpha\sqrt{r}) \tag{5.15}$$

and then

$$p(r_{\min} > r) = \int_r^\infty \mathrm{d}\, r' \; 2\pi\, r'\, e^{-\pi r'^2}(1 - \alpha\sqrt{r}) \;\sim\; e^{-\pi r^2 + a x^{5/2}} \tag{5.16}$$

or, equivalently,

$$\log p(r_{\min} > r) = -\pi r^2 + a x^{5/2} \tag{5.17}$$

By plotting $\log p(X > r)$ together with the fitted curves obtained from (5.17) at small X ($X < 0.3$), we can have an estimate of the goodness of our assumptions (Fig. 5.15).

[3] $2\epsilon > C + D - A - B = \sqrt{s^2 + A^2} + \sqrt{s^2 + B^2} - A - B \approx \frac{1}{2} s^2 \left(\frac{1}{A} + \frac{1}{B}\right)$

In Fig. 5.16 we show a plot of the parameter a as a function of L. The fitted curve shown is $a(L) = b - c\sqrt{\frac{\log L}{L}}$, with $b = 2.872$ and $c = 3.045$.

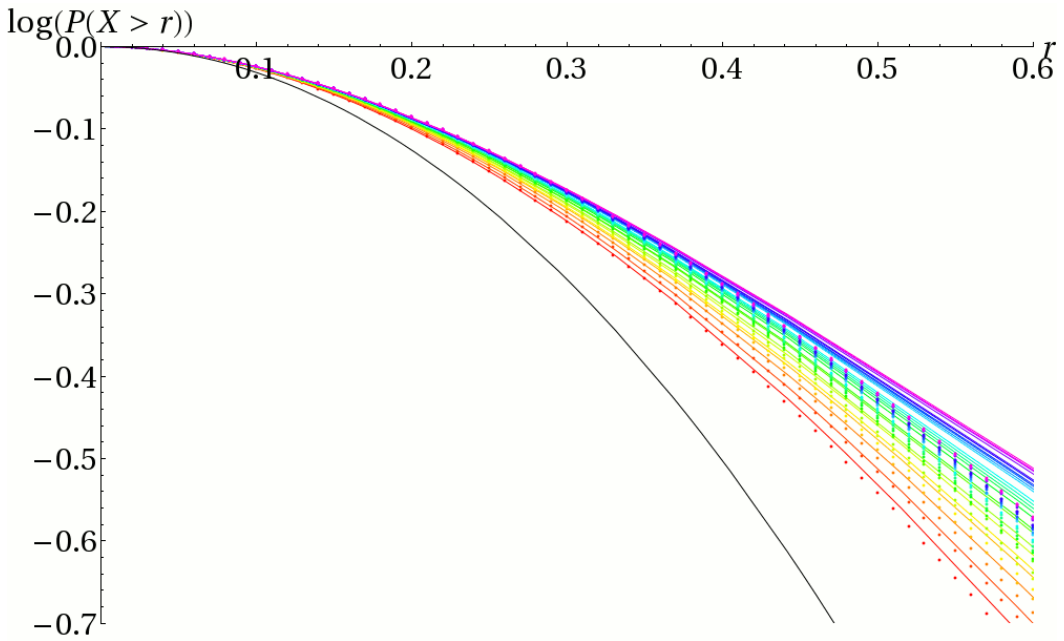

Figure 5.15: $\log p(X > r)$ with the fitted curves (5.17). In black, the plot of $y = -\pi r^2$.

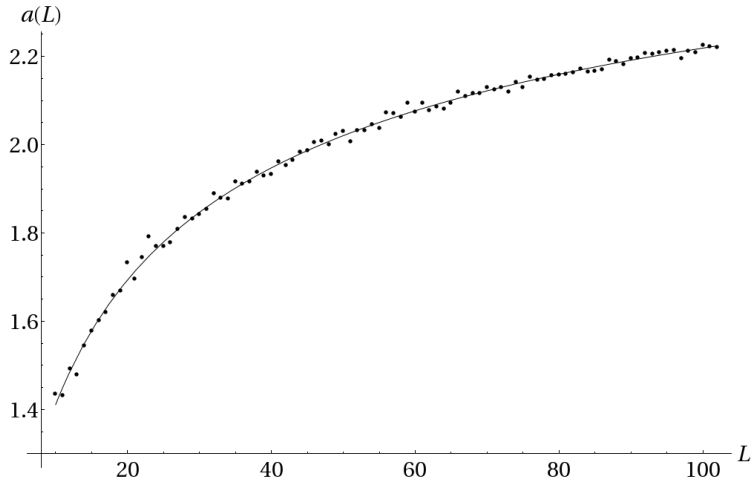

Figure 5.16: The parameter a in (5.17) as a function of L with the fitted curve $a(L) = 2.872 - 3.045\sqrt{\frac{\log L}{L}}$.

5.1.2 Near the critical point

As the density ρ grows away from $\rho = 1$, we expect the distance between matched points, on average, to decrease. For $\rho \gg 1$, every grid point will simply tend to be matched to the closest Poisson point, and the other way round for $\rho \ll 1$.

In Fig. 5.17 we show the probability density for the sample case $L = 60$ at different values of the density ρ . We see that, as expected, the curves are more

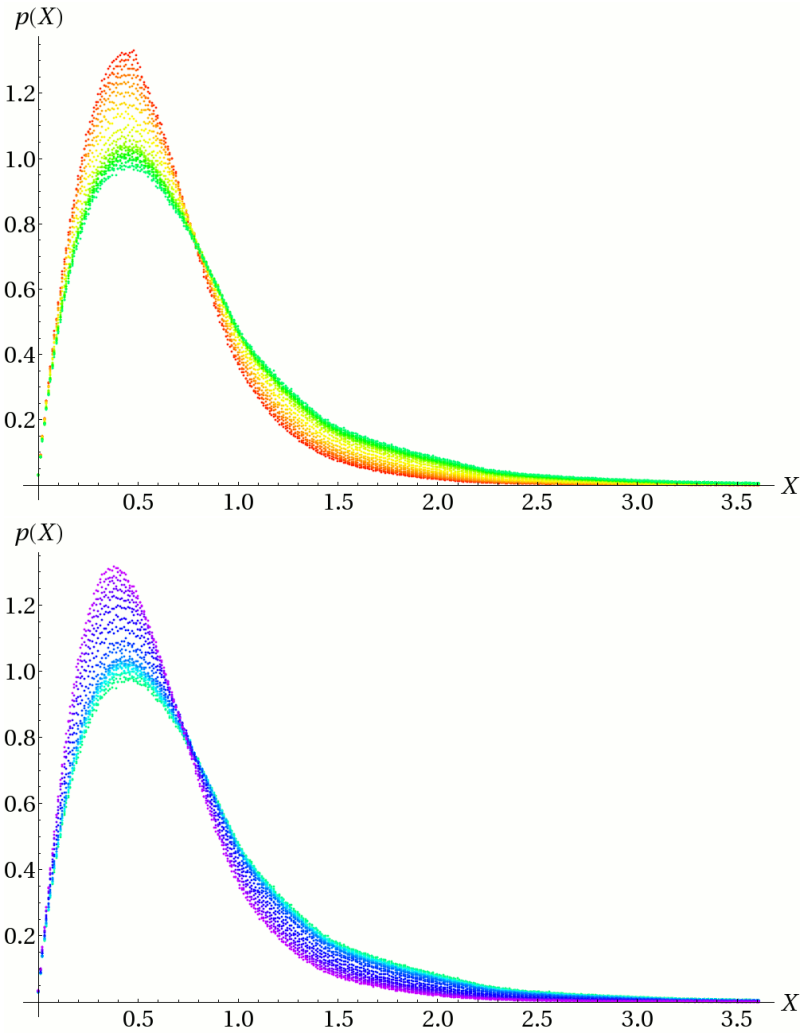

Figure 5.17: $p_{L=60,\rho}(X)$, for different values of ρ. Top: from red to green, $\rho =$ 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.986, 0.99, 0.991, 0.992, 0.993, 0.994, 0.996, 0.997, 0.998, 0.999, 1.
Bottom: from green to purple, $\rho =$ 1, 1.001, 1.002, 1.003, 1.004, 1.006, 1.007, 1.008, 1.009, 1.01, 1.016, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08, 1.09, 1.1.

and more peaked as ρ grows away from 1. Moreover, if $\rho > 1$, the position of the peak tends to get closer to zero as ρ increases (Fig. 5.18).

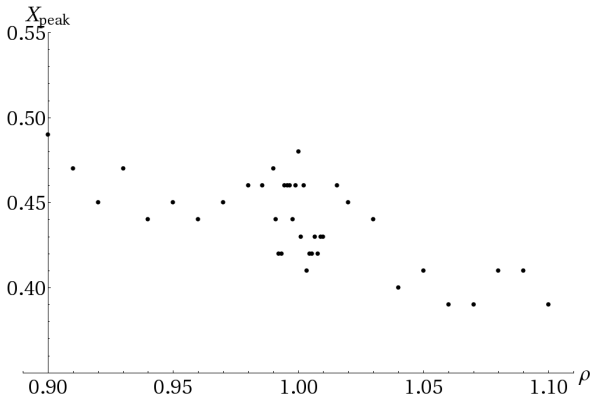

Figure 5.18: Position of the peak as a function of ρ (L=60).

Fig. 5.19 shows the mean and variance for X at different values of ρ, for $L = 60$ fixed.

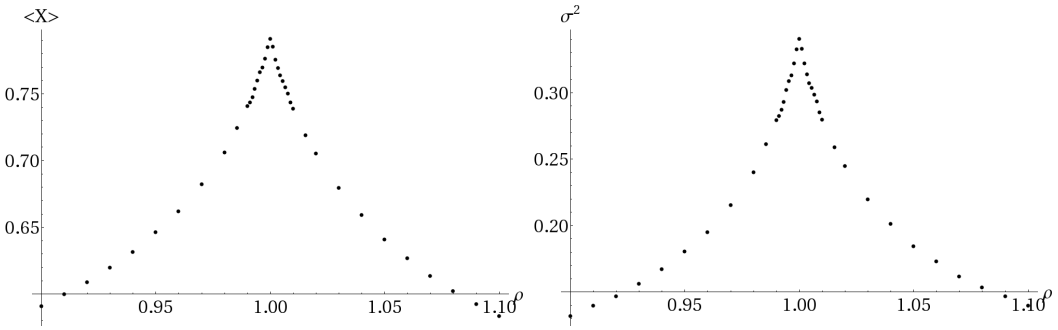

Figure 5.19: $\langle X \rangle$ (left) and $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$ (right) as a function of ρ (L=60).

The limit values we expect for $\langle X \rangle$ are[4]:

$$\langle X \rangle \to 0 \qquad \text{as} \quad \rho \to \infty$$

$$\langle X \rangle \to 0.3825 \qquad \text{as} \quad \rho \to 0.$$

To test this property we computed the mean value of X at $L = 10$, for different densities from $\rho = 0.01$ to $\rho = 110$. The results are shown in Fig. 5.20.

[4]In the limit of one Poisson point in the lattice, the nearest grid point will be, on average, at distance $\int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \mathrm{d}x\, \mathrm{d}y\, \sqrt{x^2 + y^2} = 0.3825$.

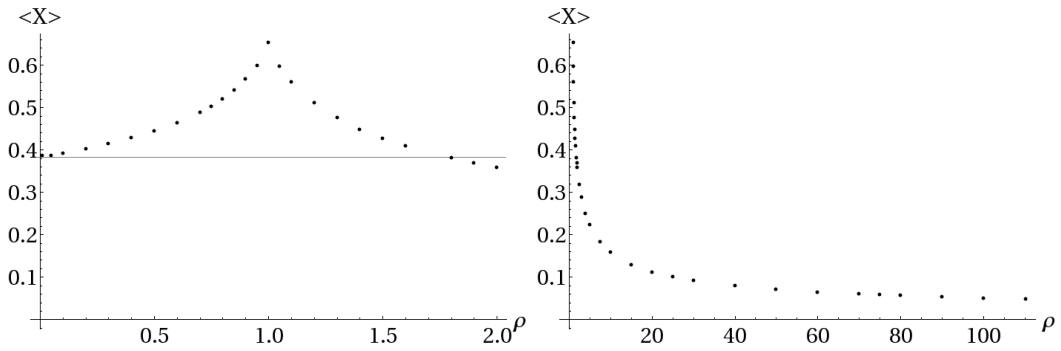

Figure 5.20: $\langle X \rangle$ as a function of ρ. Left: ρ from 0.01 to 2. Right: ρ from 1 to 110.

Tail behaviour

Also in this case, we examine the logarithm of the curves for $X \gtrsim 1.5$ and we make a fit with a line $y = A_\rho + B_\rho x$, with $y = \log(p_\rho(X))$, $x = X$. This time $L = \bar{L}$ is fixed and $A_\rho = A(\bar{L}, \rho)$, $B_\rho = B(\bar{L}, \rho)$.

If we plot the values of the fit parameters as a function of ρ, we see that, on first approximation, both $A_{L=\bar{L}}(\rho)$ and $B_{L=\bar{L}}(\rho)$ can be fitted by a polynomial of degree 2 in the module of the reduced temperature $|t| = |\rho - 1|$. In Fig. 5.22 we show the fitted curves for, as an example, $L = 60$.

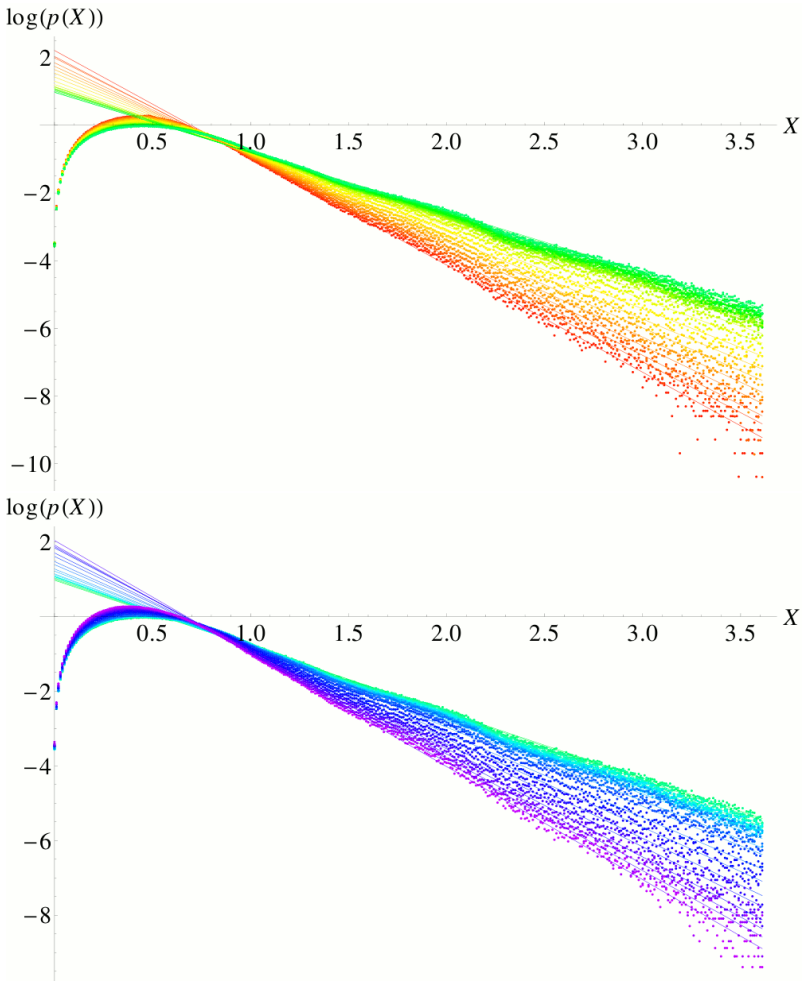

Figure 5.21: The logarithm of the histograms in Fig. 5.17, shown together with the fitted lines.

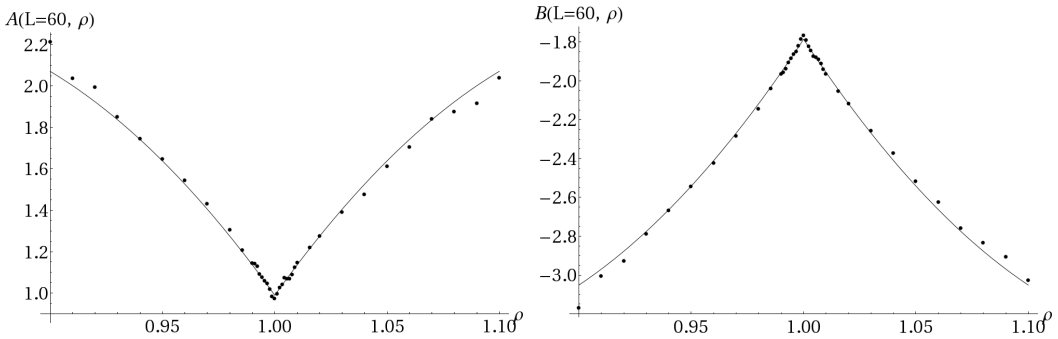

Figure 5.22: $A(L = 60, \rho)$ and $B(L = 60, \rho)$ as a function of ρ with the fitted curves (polynomial of degree 2 in $|\rho - 1|$). The fitted curves are: $A(L = 60, \rho) = (0.989115 + 15.1458|\rho - 1| - 43.2854|\rho - 1|^2)$ and $B(L = 60, \rho) = (-1.78577 - 17.5973|\rho - 1| + 49.3519|\rho - 1|^2)$.

5.2 Numerical results in one dimension

As previously mentioned, we found that in one dimension the distribution of the lengths between matched points is considerably larger than in two dimensions. For this reason, in this case we shall consider a broader histogram bin, specifically, one tenth of the lattice parameter: $\Delta_1 = 1/10$. In addition, we observe that these results have been obtained by minimising the sum of the squares of the distance between matched pairs.

5.2.1 At the critical point

In Figures 5.23 and 5.24 we show the experimental curves at density $\rho = 1$ for different sizes of the system. Again, the distribution tends to widen as the

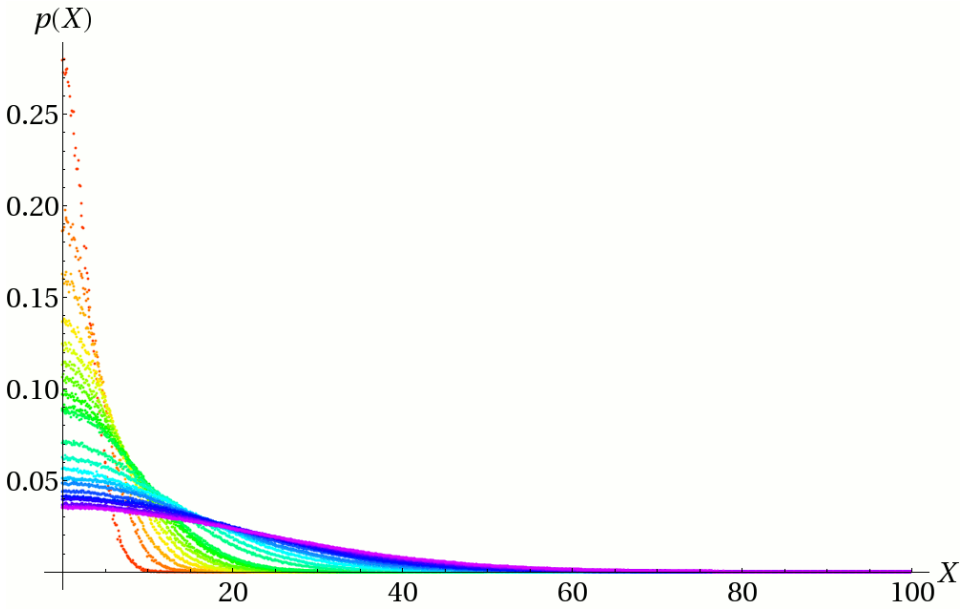

Figure 5.23: Normalised histograms of the edge length at density $\rho = 1$. From red to purple, L =100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500.

size, L, increases (in Fig. 5.25, $\langle X \rangle$ and σ^2), while we cannot determine with precision the position of the peak, as it appears to fall in the first bin of the histogram. It may be possible to establish its value by considering narrower bins.

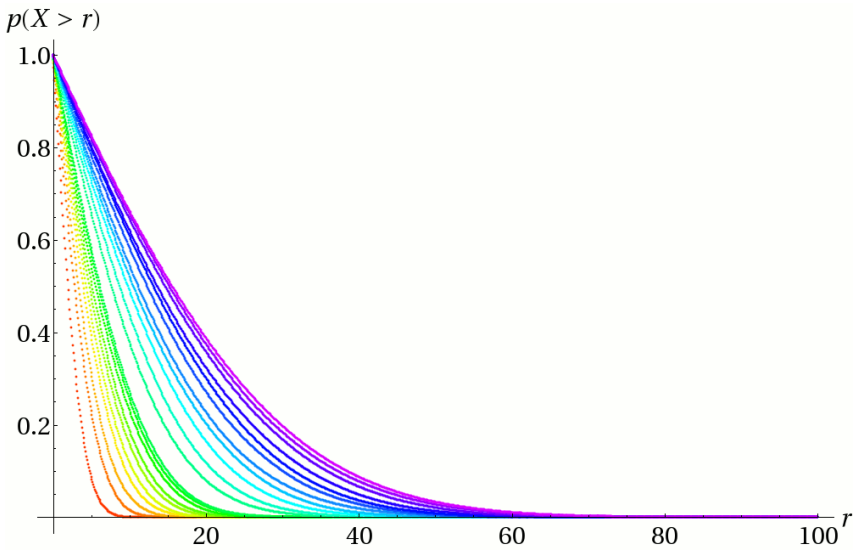

Figure 5.24: Cumulative distribution function. From red to purple, L between 100 and 6500.

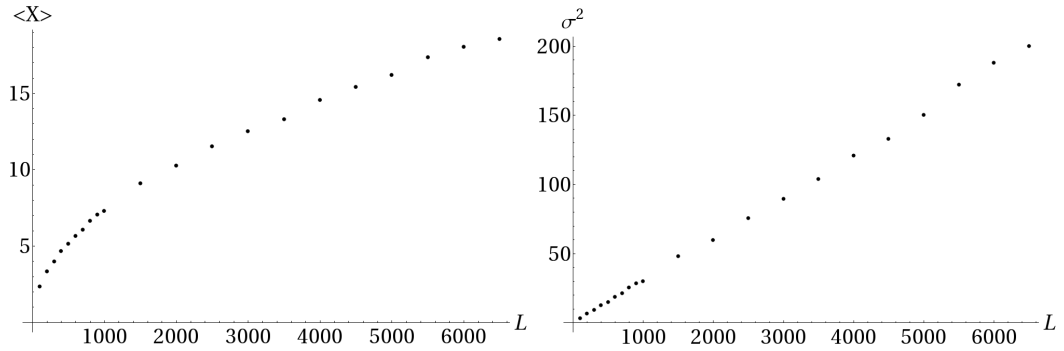

Figure 5.25: Mean value of X and variance $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$ as a function of the size of the system L.

Tail behaviour

Similarly to what we did in 2 dimensions, we take the logarithm of the probability density, but this time we find a different behaviour (Fig. 5.26).

Therefore, we try taking the squares of the abscissas, and in this way the resulting curves become linear functions of X^2 (Fig. 5.27):

$$\log(p_L(X)) = A_L + B_L X^2 \tag{5.18}$$

In Fig. 5.28. we show a plot of $A(L)$ and $B(L)$ as a function of L. This time we could not find a suitable function to fit these points.

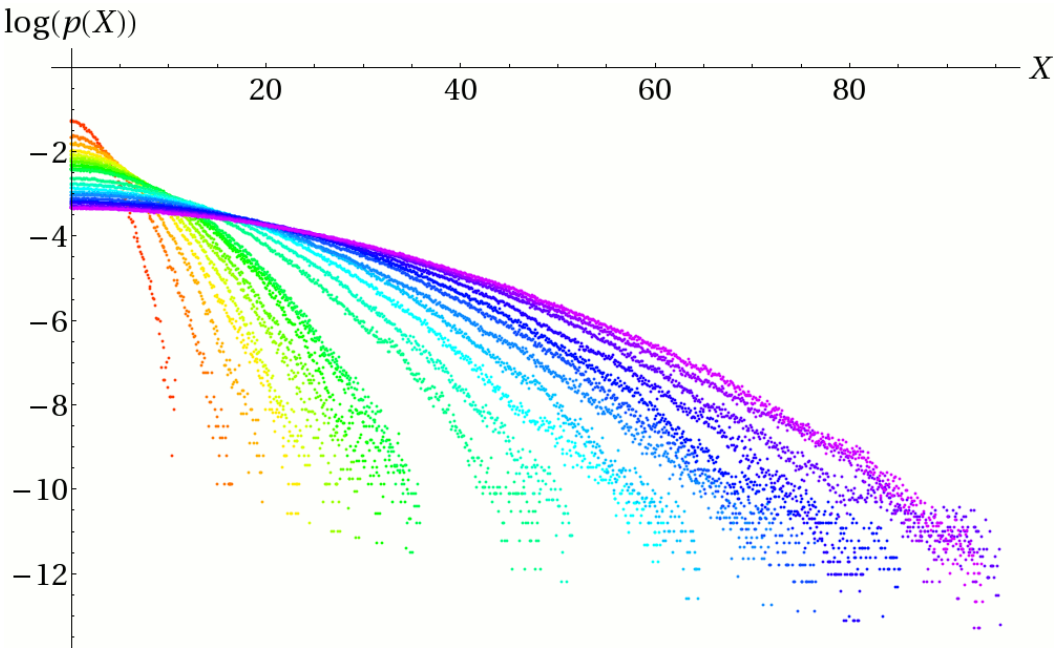

Figure 5.26: Logarithm of the probability distribution. From red to purple, L =100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500.

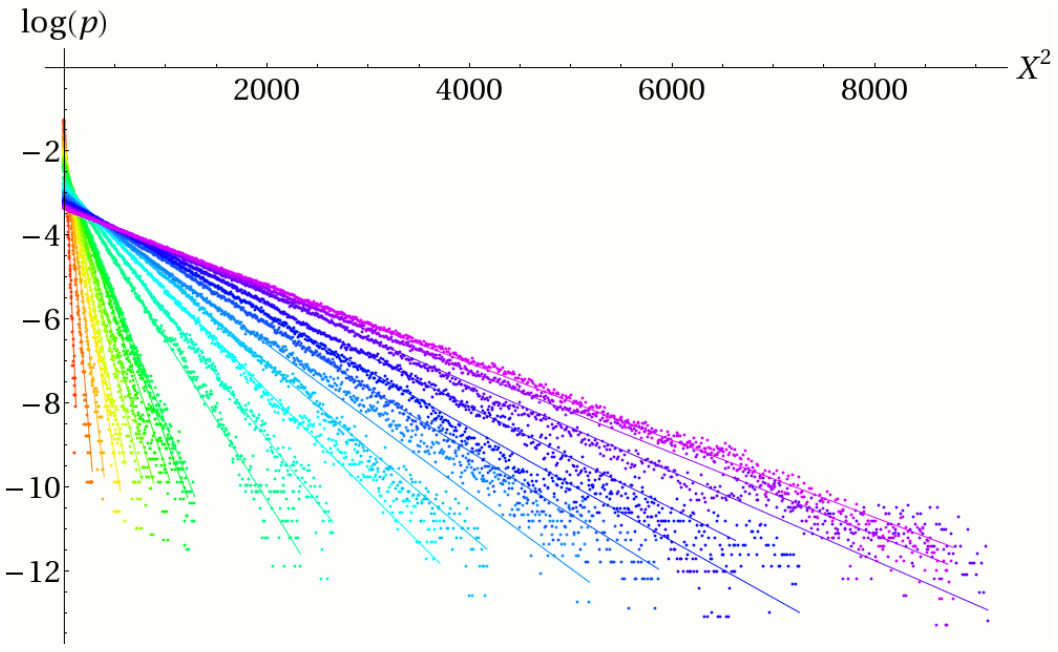

Figure 5.27: Logarithm of the probability distribution as a function of X^2.

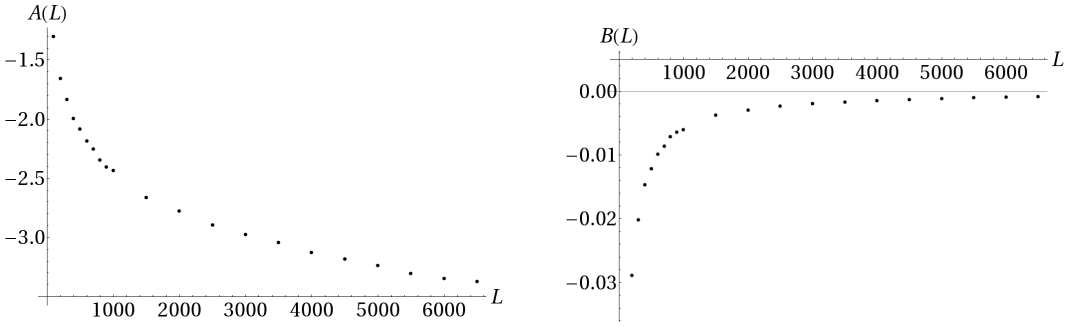

Figure 5.28: The fit parameters A_L and B_L in $\log p_L(X) = A_L + B_L X^2$ as a function of L.

5.2.2 Near the critical point

The comments we made for two dimensions hold also in one, and we expect the mean edge length to decrease and the curves to narrow as the density grows farther from $\rho = 1$ (see Fig. 5.29). In this case, the limit values we expect for

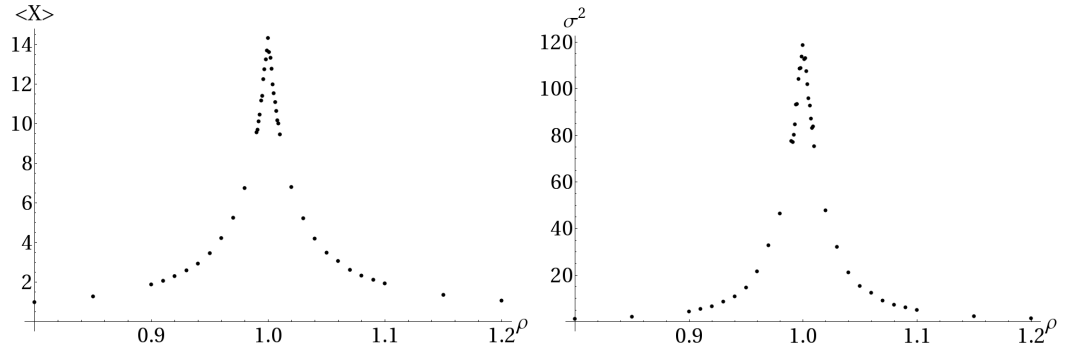

Figure 5.29: $\langle X \rangle$ (left) and $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$ (right) as a function of ρ (L=4000).

the mean edge length are:

$$\langle X \rangle \to 0 \qquad \text{as} \quad \rho \to \infty$$

$$\langle X \rangle \to 0.25 \qquad \text{as} \quad \rho \to 0,$$

as we have verified in Fig. 5.30.

In Fig. 5.31 we show the probability density for the sample case $L = 4000$ at different values of the density ρ, with $0.8 < \rho < 1.2$. We then focus only on the values closest to the critical point: $0.99 < \rho < 1.01$ (Fig. 5.32). We see that, as expected, the curves are more and more peaked as ρ grows away from 1.

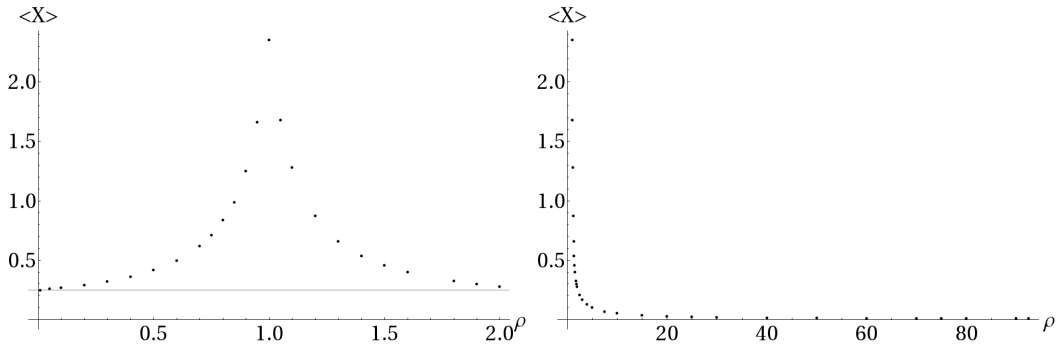

Figure 5.30: $\langle X \rangle$ as a function of ρ for $L = 100$. Left: ρ from 0.01 to 2. Right: ρ from 1 to 92.5.

Fig. 5.33 shows the logarithm of the previous curves and, similarly to what already done at $\rho = 1$, by taking the squares of the abscissas, the resulting curves become linear functions of X^2 (Fig. 5.34):

Finally, in Fig. 5.35 we plot the values of the fit parameters and we find that, again, they seem to be approximately a function of $|t| = |\rho - 1|$, even though the dependence is significantly less precise than in the previous case.

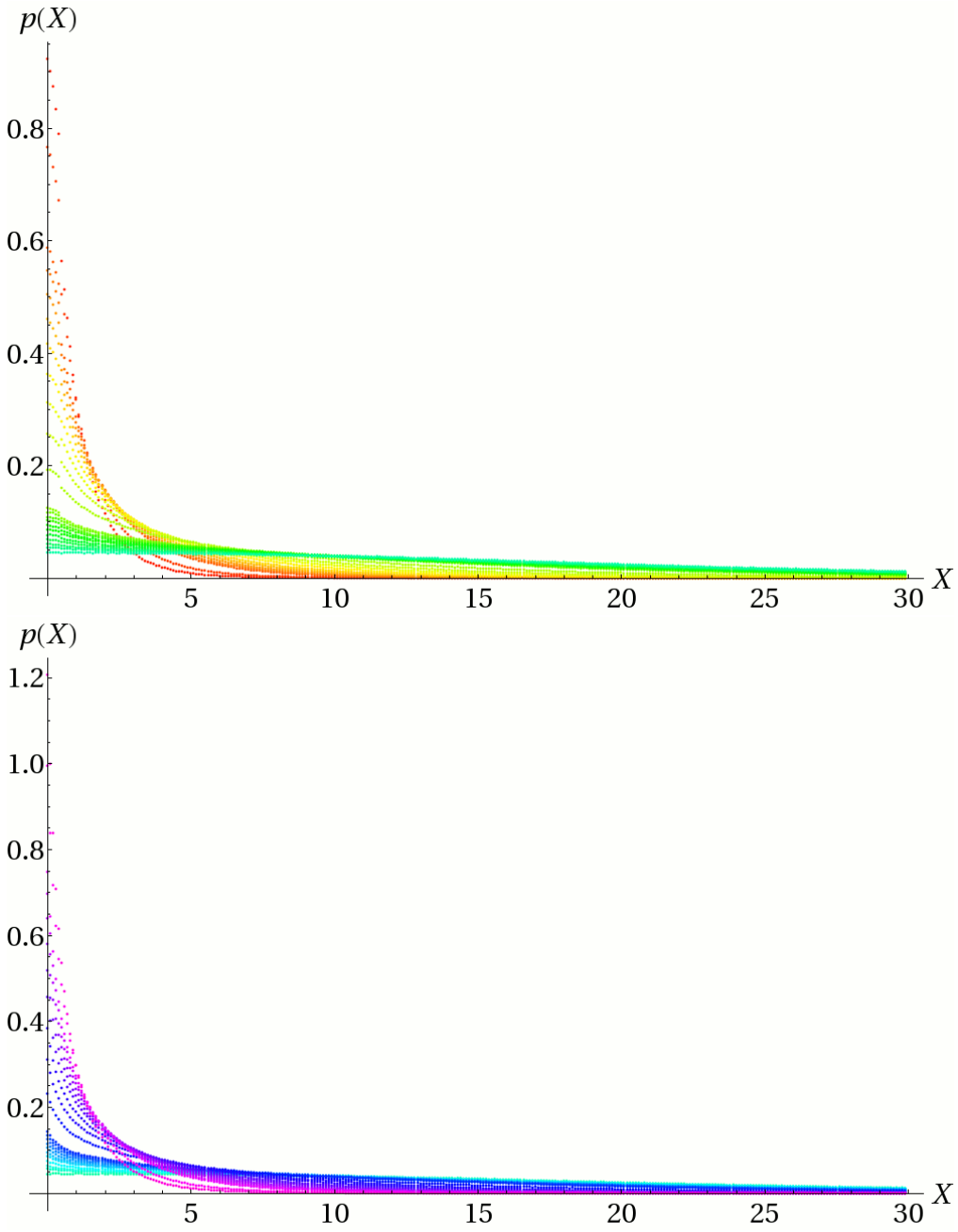

Figure 5.31: $p_{L=4000,\rho}(X)$, at different values of ρ. Top: from red to green, ρ = 0.8, 0.85, 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.991, 0.992, 0.993, 0.994, 0.995, 0.996, 0.997, 0.998, 0.999, 1. Bottom: from green to purple, ρ = 1.001, 1.002, 1.003, 1.004, 1.005, 1.006, 1.007, 1.008, 1.009, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08, 1.09, 1.1, 1.15, 1.2.

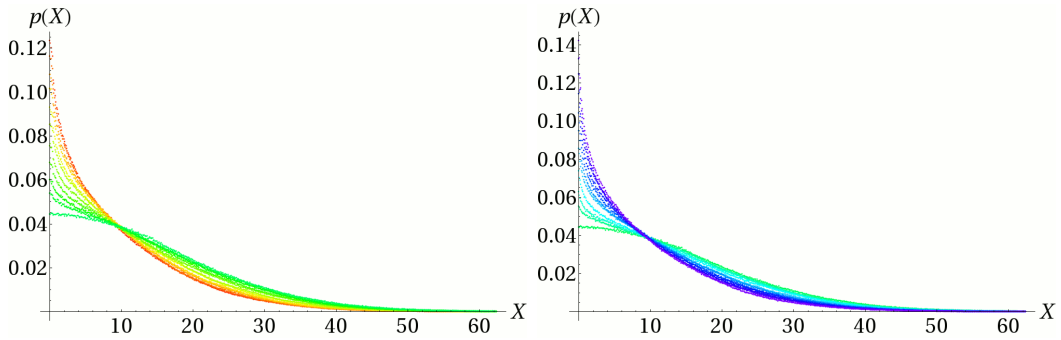

Figure 5.32: $p_{L=4000,\rho}(X)$, at different values of ρ. From red to green, $\rho = 0.99$, 0.991, 0.992, 0.993, 0.994, 0.995, 0.996, 0.997, 0.998, 0.999, 1. From green to purple, $\rho = 1$, 1.001, 1.002, 1.003, 1.004, 1.005, 1.006, 1.007, 1.008, 1.009, 1.01.

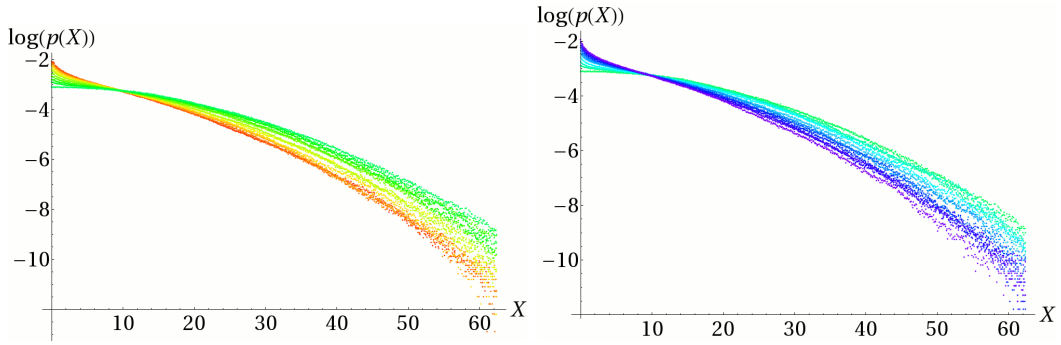

Figure 5.33: Logarithm of the curves above.

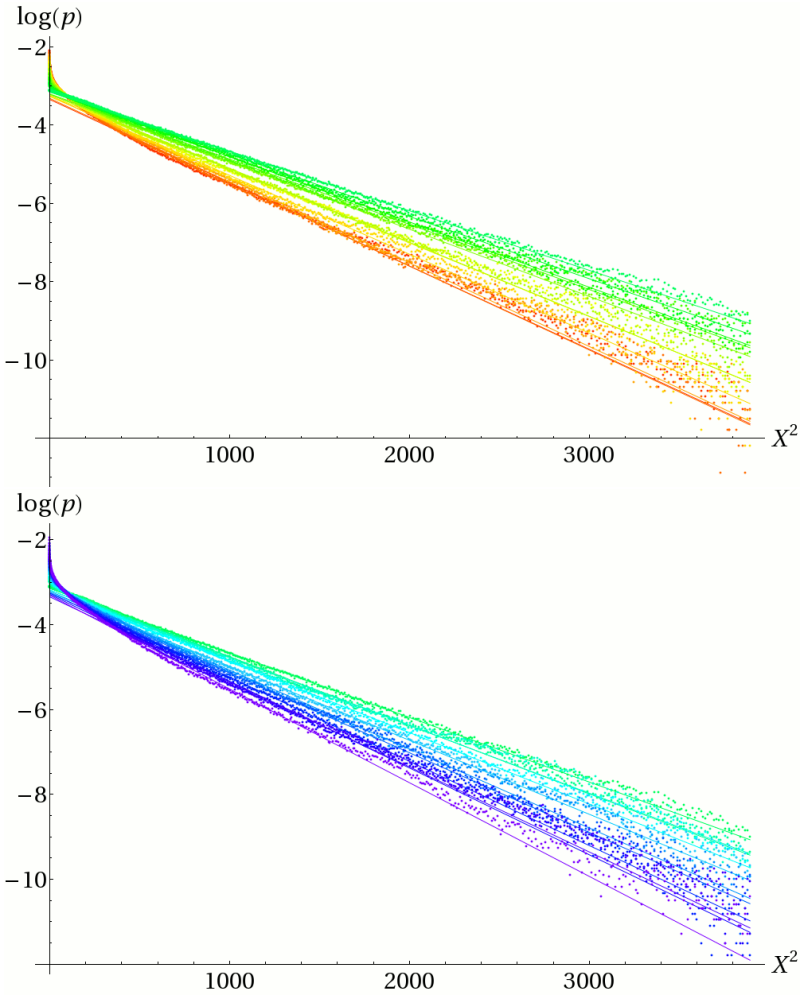

Figure 5.34: Logarithm of the probability distribution as a function of X^2, together with the fitted lines.

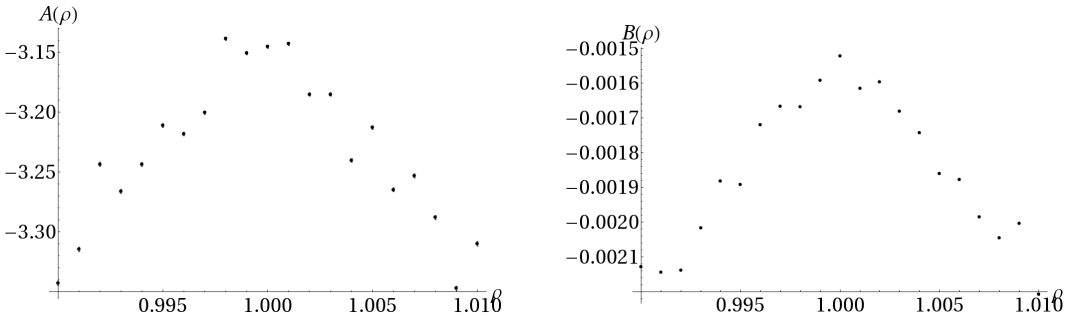

Figure 5.35: The fit parameters $A = A(L, \rho)$ and $B = B(L, \rho)$ for $\log(p_L(X)) = A + BX^2$ as a function of ρ ($L = 4000$ fixed).

Conclusions and future developments

In this thesis, we have focused in particular on the behaviour of the Grid-Poisson Marriage on the critical point, that is at $\rho = 1$.

Regarding the study of the correlation function, we have arrived at the exact solution in one dimension at $\rho = 1$, and, by means of the theory of Finite-Size Scaling, we have given an estimate of how this function scales in two dimensions.

The results at $\rho \neq 1$, are less definite, particularly in one dimension, where we do not have an ansatz for the form of the correlation function.

In two dimensions, despite having such an ansatz, the numerical simulations have not enabled us to have precise numeric details on the correlation length, which would have allowed us to use more sophisticated techniques for the determination of the critical exponents.

In this sense, the question of a definition of correlation function convenient from the point of view of a numerical determination, is still open to new ideas and research.

Concerning the study of the length of the edges of the optimal matching, the theoretical predictions about the mean value have been confirmed numerically, to the best of our knowledge, for the first time, and we have come to an estimate of the scaling of the energy with the number of couples for the two-dimensional GPM on the critical point.

We have also given a preliminary description of the density distribution, whose characteristics still need to be further investigated.

In conclusion, in this work the tools of statistical mechanics, and in particular the techniques of the Finite-Size Scaling, have been profitably put into use in the study of a problem of combinatorial optimisation, so confirming the strong link between these two areas of study.

A

Technical tools

As the backbone of our work was the computer simulations we devised to examine the statistical properties of our model, in this appendix we wish to give a brief description of the main technical tools we have used.

We also want to point out that all the simulations performed for this thesis, including the necessary tests and inevitable dead ends, add up to 233800 hours of CPU process time.

A.1 Hardware

We were only able to carry out this large number of complex simulations thanks to the facilities of the Laboratorio di Calcolo e Multimedia, that is to say, the computer laboratory of the Physics Department of the University of Milan.

We could make use of a cluster of 13 client machines and a main server, which performs two kinds of services: networking and computing[1].

Networking services

- NFS/Cluster diskless-root – The operating system is not present on the client nodes. Instead, it is mounted remotely through the Network File System protocol.

[1] The following information has been kindly provided by S. Mandelli, staff member of the laboratory.

- DHCP – The Dynamic Host Configuration Protocol is used to assign an IP address to any client that makes a request.
- PXE – The Preboot eXecution Environment enables the server to remotely provide the clients with the kernel they need in order to boot.
- BIND – The DNS server is needed to resolve the names of the nodes into their IP addresses.

Computing services

- PBS – The Portable Batch System is a service that performs job scheduling. This task is accomplished by means of two programs:
  - Torque 3.0.4, a distributed resource manager providing control over batch jobs and distributed nodes.
  - Maui 3.3, a cluster job scheduler.
- Recently, five nodes have been added to an *InfiniBand* (very high speed) subnetwork in order to enhance the simulations with MPI/OpenMPI (Message Passing Interface), which require the utilization of several nodes in parallel.

General notes on the cluster

- The operating system on both the server and the nodes is Debian Linux Stable (Squeeze 6.0).
- All nodes are Dell PowerEdge.
- All processors are Intel Xeon. In particular, there are four X5650, while the least powerful are some X5450.
- On the whole, by summing up the resources of all nodes, we get
  - number of CPUs: 208
  - total memory (RAM): 293 GB

A.2 Software

The programs we run were written in the C programming language. As a starting-point, we took the source files by Elia Zarinelli, whose work of thesis [1] was the basis for our own, modifying them as it was necessary for our purposes.

This code makes use in its core function of the implementation in C written by Donald E. Knuth [3] of the Hungarian algorithm.

All data analyses and calculations have been performed by means of the mathematical software program *Wolfram Mathematica* v. 8.0.

Finally, this thesis has been written using the LaTeX 2_ε typesetting system and all original images have been created or modified either by the above-mentioned Mathematica program or with the *GIMP* graphics editor.

B

Basic concepts of graph theory

B.1 Graphs and subgraphs

Let V be a finite set, and refer to the subsets of two (not necessarily distinct) elements of V as

$$E(V) = \{\{u, v\} | u, v \in V\}.$$

A pair $G = (V, E)$ with $E \subseteq E(V)$ is called a *graph* (on V).

The elements of V are called the *vertices* (or *nodes*, or *points*) of G, and those of E the *edges* (or *lines*, or *arcs*) of G.

We denote by $V(G)$ the vertex set of G, and by $E(G)$ the edge set. Their cardinalities $|V(G)|$ and $|E(G)|$ are called, respectively, the *order* and the *size* of the graph.

An edge with endvertices u and v is denoted by uv. The edge uv is said to be *incident* to its *ends* u and v. Vertices u and v are *adjacent* or *neighbours*, if $uv \in G$. If the endpoints are distinct, the edge is called a *link*, otherwise, it is called a *loop*.

If the graph G contains loops vv and *parallel* (or *multiple*) edges between vertices, it is called a *multigraph*, otherwise it is a *simple graph*.

In graphical representations, a vertex is usually drawn as a node or a dot, while an edge is drawn as a line connecting two vertices.

If the pairs $\{u, v\}$ are ordered (that is, $E \subseteq V \times V$), $G(V, E)$ is called a *directed graph*, or *digraph*. In this case, $uv \neq vu$. The *directed edge* is drawn

as an arrow between the endpoints, and for this reason they are called the *tail* and *head* of the edge. A digraph can contain edges uv and vu of opposite directions.

Let $v \in G$ be a vertex of a graph G. The *neighbourhood* of v is the set

$$N_G(v) = \{u \in G | vu \in G\}.$$

The *degree* (or *valency*) of v is the number of its neighbours:

$$d_G(v) = |N_G(v)|.$$

If $d_G(v) = 0$, then v is said to be *isolated* in G, and if $d_G(v) = 1$, then v is a *leaf* of the graph.

It is also possible to consider graphs with infinitely many vertices or edges or both. In this case the graph is said to be *infinite*. An infinite graph where every vertex has finite degree is called *locally finite*.

The graph $G = K_V$ is the *complete graph* on V, if every two vertices are adjacent. All complete graphs of order n are isomorphic with each other, and they are denoted by K_n.

A graph G is said to be *regular*, if every vertex of G has the same degree. If this degree is equal to r, then G is *r-regular* or *regular of degree* r. A complete graph K_n is $(n-1)$-regular.

We can also we define a *weight function* $w : E(G) \to \mathbb{R}$, which assign a "weight" $\in \mathbb{R}$ to each edge of the graph. In this case, we say that the graph is *(edge-)weighted*.

A graph G is called *bipartite*, if V_G has a partition to two subsets X and Y such that each edge $uv \in G$ connects a vertex of X and a vertex of Y. In this case, G is said to be (X, Y)-*bipartite*.

A bipartite graph G is *complete* (m, k)-*bipartite*, if $|X| = m$, $|Y| = k$, and $uv \in G$ for all $u \in X$ and $v \in Y$. All complete (m, k)-bipartite graphs are isomorphic and denoted as $K_{m,k}$.

A graph H is a *subgraph* of a graph G, denoted by $H \subseteq G$, if $V_H \subseteq V_G$ and $E_H \subseteq E_G$. A subgraph $H \subseteq G$ *spans* G (and H is a *spanning subgraph* of G), if every vertex of G is in H, that is, $V_H = V_G$.

For a graph G, a subset $M \subseteq E_G$ is a *matching* of G, if M contains no adjacent edges. The two ends of an edge $e \in M$ are matched under M. A matching M is a *maximum matching*, if for no matching $M', |M| < |M'|$.

A matching M *saturates* $v \in G$, if v is an end of an edge in M. Also, M saturates $A \subseteq V_G$, if it saturates every $v \in A$. If M saturates V_G , then M is a *perfect matching*.

B.2 Paths and cycles

Let $e_i = u_i u_{i+1} \in G$ be (adjacent) edges of G for $i \in [1, k]$. The sequence $W = \{e_1, e_2, \dots e_k\}$ is a *walk* of length k from u_1 to u_{k+1}.

W is said to be

- *closed* (or a *loop*), if $u_1 = u_{k+1}$
- a *path*, if $u_i \neq u_j$ for all $i \neq j$
- a *cycle*, if it is closed, $k \geq 2$, and $u_i \neq u_j$ for $i \neq j$ except that $u_1 = u_{k+1}$

A graph G is said to be *connected*, if for all $u, v \in V(G)$, u and v can be connected by a walk; otherwise, it is disconnected. The maximal connected subgraphs of G are its *connected components*.

If we denote with $c(G)$ the number of connected components of G, and with $L(G)$ the number of independent loops, then

$$|V(G)| - |E(G)| + L(G) = c(G),$$

which is known as *Euler's formula*.

A vertex $v \in G$ is a *cut vertex*, if $c(G - v) > c(G)$, that is, when it is removed from G the graph becomes disconnected. Analogously, a *cut edge* or *bridge* is an edge whose removal disconnects a graph.

A subset of arcs is called a *cut* if it is such that when these arcs are removed from G, the graph becomes disconnected.

If a graph contains no cycles, it is said to be *acyclic*, or a *forest*. A connected acyclic graph is a *tree*.

C

Distribution functions of random variables

C.1 Distribution functions

Let X be a random variable.

The *cumulative distribution function* (cdf) of X is the function defined by

$$F_X(x) = P(X \leq x) \qquad -\infty < x < \infty. \tag{C.1}$$

We have

- $0 \leq F_X(x) \leq 1$
- $F_X(x_1) \leq F_X(x_2) \qquad \text{if} \quad x_1 < x_2$
- $\lim_{x \to \infty} F_X(x) = 1$
- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to a^+} F_X(x) = F_X(a)$
- $P(a < X \leq b) = F_X(b) - F_X(a)$
- $P(X > a) = 1 - F_X(a)$
- $P(X < b) = F_X(b^-) \qquad b^- = \lim_{0 < \varepsilon \to 0} b - \varepsilon$

Let X be a discrete random variable. If we define

$$p_X(x) = P(X = x), \tag{C.2}$$

$p_X(x)$ is called the *probability mass function* (pmf) of the discrete random variable X, and

- $0 \leq p_X(x_k) \leq 1 \qquad k = 1, 2, \ldots$
- $p_X(x) = 0 \qquad \text{if} \quad x \neq x_k (k = 1, 2, \ldots)$
- $\sum\limits_k p_X(x_k) = 1$
- $F_X(x) = P(X \leq x) = \sum\limits_{x_k \leq x} p_X(x_k)$

Let X be a continuous r.v. and $f_X(x) = \dfrac{dF_X(x)}{dx}$. The function $f_X(x)$ is called the *probability density function* (pdf) of X, and it has the following properties

- $f_X(x) \geq 0$
- $\int\limits_{-\infty}^{\infty} f_X(x)dx = 1$
- $f_X(x)$ is piecewise continuous
- $F_X(x) = P(X \leq x) = \int\limits_{-\infty}^{x} f_X(x')dx'$
- $P(a < X \leq b) = \int\limits_{a}^{b} f_X(x)dx = F_X(b) - F_X(a)$

Let $\vec{X} : \Omega \to \mathbb{R}^d$, $\vec{X} = (X_1, X_2, \ldots, X_d)$. The (*d-variate*) *distribution function* of $\vec{X}$ or the *joint distribution function* of $X_1, X_2, \ldots, X_d$ is defined by

$$F(x_1, \ldots, x_d) = P\{X_1 \leq x_1, \ldots, X_d \leq x_d\} \tag{C.3}$$

for $d \geq 1$, $x_k \in \mathbb{R}$, $1 \leq k \leq d$.

If $f(x_1, \ldots, x_d) = \partial^d F/(\partial x_1 \ldots \partial x_d)$ exists for all $(x_1, \ldots, x_d) \in \mathbb{R}^d$, then the function $f(x_1, \ldots, x_d)$ is called the *joint density function* of $F(x_1, \ldots, x_d)$ or $(X_1, X_2, \ldots, X_d)$, and

$$F(x_1, \ldots, x_d) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_d} f(x'_1, \ldots, x'_d)\, dx'_d \ldots dx'_1 \tag{C.4}$$

C.2 Mean, variance and covariance

- The *mean* (or *expected value*) of a r.v. X, denoted by μ_X or $E[X]$, is defined by

$$\mu_X = E[X] = \begin{cases} \sum_k x_k p_k(x_k) & X : \text{discrete} \\ \int_{-\infty}^{\infty} x f_X(x)dx & X : \text{continuous} \end{cases} \tag{C.5}$$

- The *variance*, denote by σ_X^2 or $\text{var}[X]$, is defined by

$$\sigma_X^2 = \text{var}[X] = E[(X - E[X])^2] \tag{C.6}$$

and therefore

$$\sigma_X^2 = \begin{cases} \sum_k (x_k - \mu_k)^2 p_k(x_k) & X : \text{discrete} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx & X : \text{continuous} \end{cases} \tag{C.7}$$

- For any two random variables X and Y with finite variances σ_X^2 and σ_Y^2, the *correlation* of X and Y is defined by $E[XY]$, and the *covariance* $\text{cov}(X, Y)$ is defined by

$$\text{cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \tag{C.8}$$

C.3 Some special distributions

Uniform distribution A r.v. X is called a *uniform* r.v. over the interval (a, b) if its pdf is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases} \tag{C.9}$$

Normal (or Gaussian) distribution A r.v. X is called a *normal* or *Gaussian* r.v. if its pdf is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{C.10}$$

The mean and variance are

$$\mu_X = E[X] = \mu$$

$$\sigma_X^2 = \text{var}[X] = \sigma^2$$

It is customary to write $N(\mu; \sigma^2)$ to denote a r.v. which is normal with mean μ and variance σ^2.

Poisson distribution A discrete r.v. X is called a *Poisson* r.v. with parameter λ if its pmf is given by

$$p_X(k) = P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \qquad k = 0, 1, \ldots. \tag{C.11}$$

The mean and variance are

$$\mu_X = E[X] = \lambda$$

$$\sigma_X^2 = \text{var}[X] = \lambda.$$

Exponential distribution A r.v. is called an *exponential* r.v. with parameter $\lambda(> 0)$ if its pdf is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > o \\ 0, & x < 0 \end{cases} \tag{C.12}$$

The mean and variance are

$$\mu_X = E[X] = \frac{1}{\lambda}$$

$$\sigma_X^2 = \text{var}[X] = \frac{1}{\lambda^2}$$

Riassunto in italiano

Il lavoro svolto in questa tesi concerne le proprietà statistiche di un problema di ottimizzazione combinatoria conosciuto come "Grid-Poisson Marriage" (GPM), ovvero il matching ottimale fra N punti di un reticolo e M punti scelti a caso nel continuo con distribuzione Poissoniana.

Negli ultimi decenni si è compreso come i problemi di ottimizzazione siano connessi a quelli di meccanica statistica in modo naturale: identificando la funzione costo con l'energia, si tratta di trovare lo stato fondamentale a temperatura nulla di un sistema a molti gradi di libertà. Questa analogia ci permette di adattare le idee e gli strumenti della meccanica statistica, andando nel limite di temperatura zero, al fine di affrontare problemi di ottimizzazione.

Il modello da noi considerato è un tipico problema di assegnazione in cui la funzione peso è data dalla somma delle distanze euclidee fra le coppie di punti. In questo caso, il problema è banale per densità $\rho = M/N$ molto grandi o molto piccole, mentre il sistema diventa critico intorno a $\rho = 1$, dove abbiamo la simmetria nello scambio di N con M. Questo è un fenomeno critico a tutti gli effetti, in quanto a $\rho = 1$ il processo di "ricerca del partner ideale" propaga a tutte le scale di lunghezza e la densità di energia diverge.

In questa tesi abbiamo studiato alcune caratteristiche del GPM dal punto di vista teorico e le abbiamo confrontate con i dati ottenuti da simulazioni numeriche su reticoli di diverse taglie e con diverse densità di punti casuali.

Schema della tesi

Capitolo 1

Richiamiamo i concetti fondamentali di meccanica statistica e di ottimizzazione combinatoria e spieghiamo il legame fra questi due ambiti di studio.

Capitolo 2

Diamo le definizioni essenziali sui processi stocastici che ci serviranno nel seguito e definiamo il modello oggetto di studio di questa tesi.

Capitolo 3

L'osservabile su cui ci concentriamo inizialmente è il versore che unisce i punti accoppiati, mentre la funzione di correlazione che prendiamo in considerazione è il loro prodotto scalare.

Per prima cosa, arriviamo alla soluzione esatta a densità $\rho = 1$ per la funzione di correlazione in dimensione 1, che in questo caso si riduce al prodotto dei segni. Modellizzando il comportamento del parametro d'ordine come un ponte Browniano e mediando opportunamente sulla distribuzione di probabilità così ricavata, otteniamo per la la funzione di correlazione $G(x)$, nel caso di condizioni al contorno aperte, l'espressione:

$$G_{cca}(x) = \frac{1-\sqrt{x}}{1+\sqrt{x}}$$

mentre nel caso di condizioni al contorno periodiche:

$$G_{ccp}(x) = \frac{2}{\pi} \arctan\left(\frac{1-6x(1-x)}{\sqrt{12x(1-x)(1-3x(1-x))}} \right).$$

Consideriamo alcune variazioni della funzione peso e mostriamo come la soluzione analitica proposta sia in ottimo accordo con i risultati ottenuti dalle simulazioni numeriche, a patto di scegliere come funzione peso da minimizzare la somma dei quadrati delle distanze fra i punti. Descriviamo poi qualitativamente le curve numeriche nel caso $\rho \neq 1$.

Capitolo 4

Nel caso del GPM bidimensionale, applichiamo le idee fondamentali della teoria del Finite-Size Scaling dei sistemi termodinamici allo studio della funzione di correlazione wall-to-wall per scriverla come

$$G(x, L, t) \approx L^{\alpha} F\left(\frac{x}{L}, tL^{1/\nu}\right),$$

dove L è la taglia del sistema e $t = \rho - 1$.

Attraverso simulazioni sul punto critico $t = 0$, diamo una stima numerica del valore di α.

Fuori dal punto critico, mostriamo numericamente l'andamento della funzione di correlazione. Diamo due possibili definizioni di lunghezza di correlazione e ne esaminiamo il comportamento.

Capitolo 5

Infine, prendiamo in considerazione il modulo del vettore che unisce le coppie e ne studiamo la distribuzione di probabilità ottenuta dalle simulazioni, per diverse taglie del sistema e a diverse densità. Inoltre, esaminando la media $\langle X \rangle$ di queste lunghezze in funzione della taglia del sistema, verifichiamo i risultati presenti in letteratura: $\langle X \rangle = O(\sqrt{\log N/N})$, che corrisponde a un andamento dell'energia come: $E = O(\sqrt{N \log N})$.

Ringraziamenti

Primo fra tutti, vorrei ringraziare il Professor Caracciolo, non solo per tutto ciò che mi ha insegnato e su cui mi ha spinto a riflettere, ma anche per la sua infinita gentilezza e pazienza, e per le molte parole di incoraggiamento che mi ha rivolto nel corso di questi mesi.

Un ringraziamento particolare devo anche al Dottor Sportiello, per la sua disponibilità, il suo competente aiuto e il suo contagioso entusiasmo.

Vorrei inoltre ringraziare Marco Gherardi per i suoi suggerimenti, e gli amministratori del Laboratorio di Calcolo e Multimedia, in particolare Stefano Mandelli e Roberto Palazzi, per l'assistenza che mi hanno fornito nell'utilizzo della farm.

A livello personale, devo molto al supporto e all'incoraggiamento di alcune persone. Fra questi, vorrei citare la mia cara amica Santa, che mi ha appoggiato nei momenti difficili, e Monica e Roberto, con i quali ho condiviso molte stimolanti conversazioni. Un grazie gigante anche a Tiziano, per il suo aiuto e i suoi preziosi consigli.

Ultimi, ma non ultimi per importanza, ringrazio Simone per aver sopportato così a lungo la mia lontananza, e i miei genitori, senza il supporto dei quali non avrei potuto completare questo percorso.

Bibliography

[1] E. Zarinelli, *Statistical properties of the random 2-dimensional euclidean Grid-Poisson Marriage problem*, Master's Thesis, Università degli Studi di Milano, 2007-2008.

[2] D. Gale and L. Shapley, *College admissions and stability of marriage*, Amer. Math. Monthly, 69(1): 9–15, 1962.

[3] D. E. Knuth, *The Stanford GraphBase: A Platform for the Combinatorial computing*, Addison-Wesley, 1993.

[4] D. E. Knuth, *Mariages Stables et leurs relations avec d'autres problèmes combinatoires*, Montréal, Les Presses de l'Université de Montréal, 1976.

[5] Gaspard Monge, *Mémoire sur la théorie des déblais et des remblais*, in *Histoire de l'Académie Royale des Sciences, Année MDCCLXXXI. Avec les Mémoires de Mathématiques et de Physique pour la même année*, Paris, 1784.

[6] H. Kuhn, *The Hungarian Method for the assignment problem*, Naval Research Logistics Quarterly, 2:83–97, 1955.

[7] M. Mézard, G. Parisi, M. A. Virasoro, *Spin Glass Theory and Beyond*, Word Scientic, Singapore, 1987.

[8] D. Kannan, *An Introduction to Stochastic Processes*, North Holland, New York, 1979.

[9] R. Durrett, *Probability: Theory and Examples*, 4th Edition, Cambridge University Press, 2010.

[10] Hwei P. Hsu, *Schaum's Outline of Theory and Problems of Probability, Random Variables, and Random Processes*, McGraw-Hill, 1997.

[11] J. Chang, *Stochastic Processes*, book in preparation, retrieved from `http://www.stat.yale.edu/~jtc5/251/stochastic-processes.pdf` on 17$^{\text{th}}$ April, 2012.

[12] D. F. Anderson, *Introduction to Stochastic Processes with Applications in the Biosciences*, Lecture notes, University of Wisconsin at Madison, 2011. Available at http://www.math.wisc.edu/~anderson/605F11/Notes/StochBio.pdf.

[13] V. Privman (editor), *Finite Size Scaling and Numerical Simulation of Statistical Systems*, World Scientific, Singapore, 1990.

[14] J. L. Cardy (editor), *Finite-Size Scaling*, North-Holland, Amsterdam, 1988.

[15] S. Caracciolo, R. G. Edwards, S. J. Ferreira, A. Pelissetto, and A. D. Sokal, *Extrapolating Monte Carlo Simulations to Infinite Volume: Finite-Size Scaling at* $\frac{\xi}{L} \gg 1$, Phys. Rev. Lett. 74, 2969, 1995, e-print hep-lat/9409004.

[16] M. Gubinelli, *Finite-Size Scaling in Non-Equilibrium Critical Phenomena*, Ph.D. Thesis, Università degli Studi di Pisa, 2002.

[17] M. E. Fisher, in: *Critical Phenomena, Proc. 51st "Enrico Fermi" Summer School*, Varenna, Italy, edited by M. S. Green, Academic Press, New York, 1972.

[18] D. Fichera, *Cavity Methods in Optimization Problems: Exact and Approximated Algorithms*, Ph.D. Thesis, Università degli Studi di Milano, 2007-2008.

[19] A. Holroyd, R. Pemantle, Y. Peres, O. Schramm, *Poisson Matching*, arXiv:0712.1867v, 2007.

[20] C. Hoffman, A. Holroyd, Y. Peres, *A stable marriage of Poisson and Lebesgue*, Annals of Probability 2006, Vol. 34, No. 4, 1241–1272; arXiv:math/0505668v4 [math.PR].

[21] M. Ajtai, J. Komlós and G. Tusnády, *On optimal matchings*, Combinatorica, Vol. 4, No. 4, 259–264, 1984.

[22] P. W. Shor, *Random Planar Matching and Bin Packing*, Ph.D. thesis, MIT, September 1985

[23] P. W. Shor and J. E. Yukich, *Minimax grid matching and empirical measures*, Annals of Probability, 19, 1338–1348, 1991.

[24] J. E. Yukich, *Optimal Matching and Empirical Measures*, Proceedings of the American Mathematical Society, Vol. 107, No. 4, pp. 1051–1059, December 1989.

[25] T. Leighton and P. Shor, *Tight bounds for minimax grid matching with applications to the average case analysis of algorithms*, Combinatorica, Vol. 9, No. 2, 161–187, 1989.

[26] T. Harju, *Lecture Notes on Graph Theory*, Department of Mathematics, University of Turku, Finland, 1994-2011. Available at http://users.utu.fi/harju/graphtheory/graphtheory.pdf.