Università degli studi di Trieste

Corso di laurea magistrale in Fisica della materia

Laboratorio di Fisica Computazionale

Unit IV
Random Walks

Davide Bernocco
Matricola s285296

Programming language used:
Python

Contents

1 1D RW: properties; comparison of numerical and analytical results; convergence

We are going simulate 1D Random Walks by a Monte Carlo approach, comparing numerical key quantities as $\langle x_i \rangle$, $\langle x_i^2 \rangle$ and $\langle (\Delta x)_i^2 \rangle$ averaged over several independent walkers with the theoretical expectations. For sake of clarity, except where differently specified, we assume the following simplification:

- Initial position corresponding to the origin for all the walkers
- Fixed step length of $l = 1$
- $p_{\leftarrow} = p_{\rightarrow} = 1/2$

1.1 Instantaneous position and square-position

The first thing we can do is plotting together different walkers just to get an idea of how they progressively spread in the space with each iteration. Numerically, the time flow is encoded in a loop index which goes from 0 to a certain N fixed.
Here are the two plots of the instantaneous covered distance x_i first and the instantaneous square-distance x_i^2 with $N = 64$:

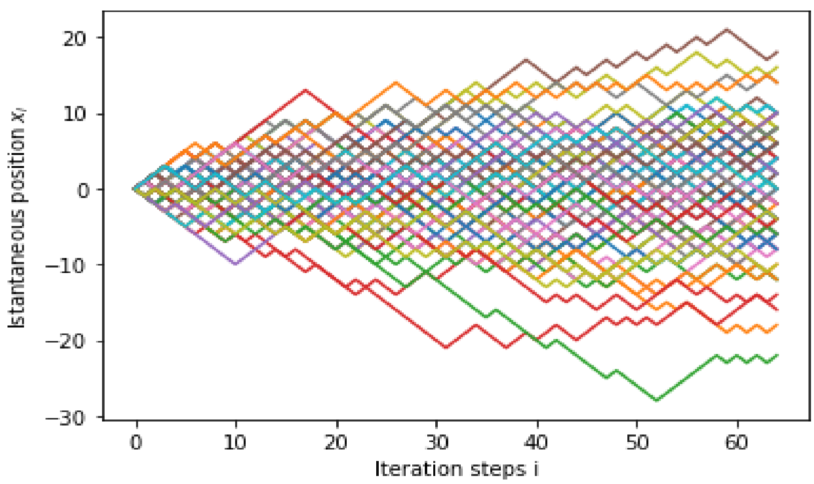


Fig 1.1 x_i for 100 independent walkers

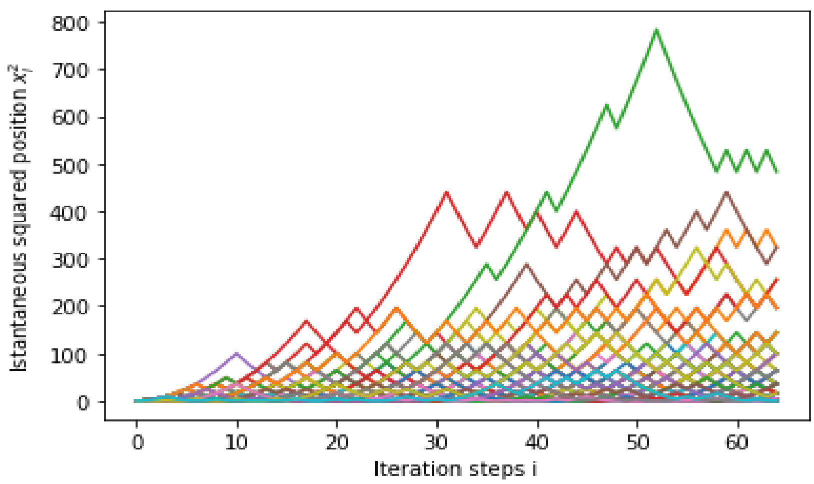


Fig 1.2 x_i^2 for 100 independent walkers

As it can be seen, even though they start at the same point, the walkers progressively tend to cover larger and larger distances with time. This peculiar behaviour makes RW the first simplified model one often uses to simulate the phenomenon of `diffusion`.

1.2 Average position and square-position

Now, if we want to make quantitative comparisons with the theoretical results that can be found in literature, we have to look at averaged quantities. In fact, taking the averages $\langle x_i \rangle$ and $\langle x_i^2 \rangle$ over all the considered walkers, instant by instant, should make us able to numerically reproduce the analytical expected values (remember l was set to 1):

$$\langle x_i \rangle^{th} = 0 \qquad\qquad \langle x_i^2 \rangle^{th} = i$$

The following plots shows the same set of walkers with the averaged quantities of interest for each $i \leqslant N$:

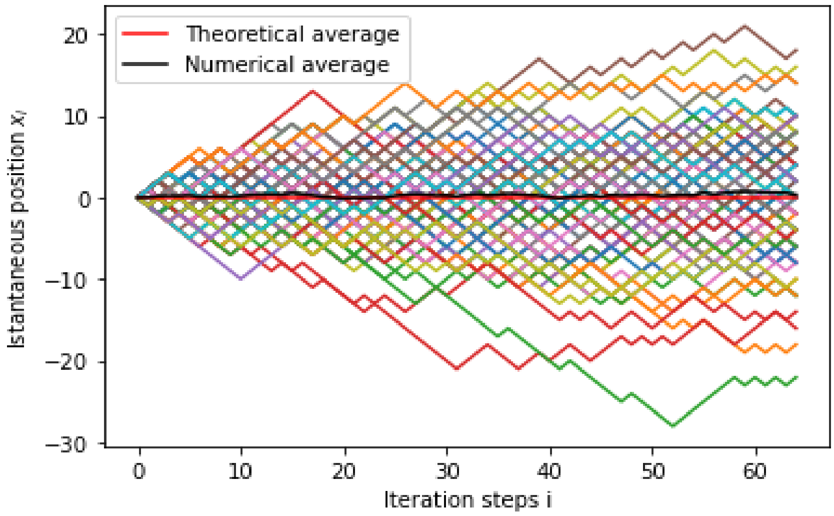


Fig 1.3 x_i and $\langle x_i \rangle$ for 100 independent walkers

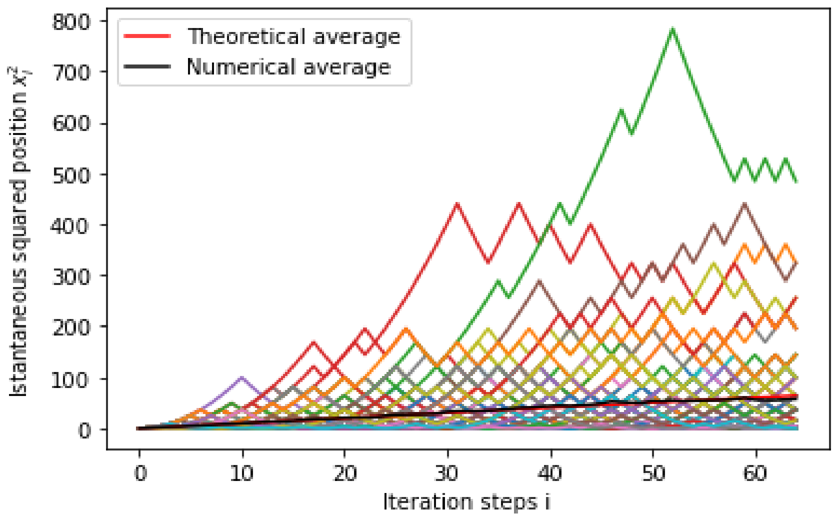


Fig 1.4 x_i^2 and $\langle x_i^2 \rangle$ for 100 independent walkers

Performing a linear regression $y = a + bx$ on the two quantities, we finally get:

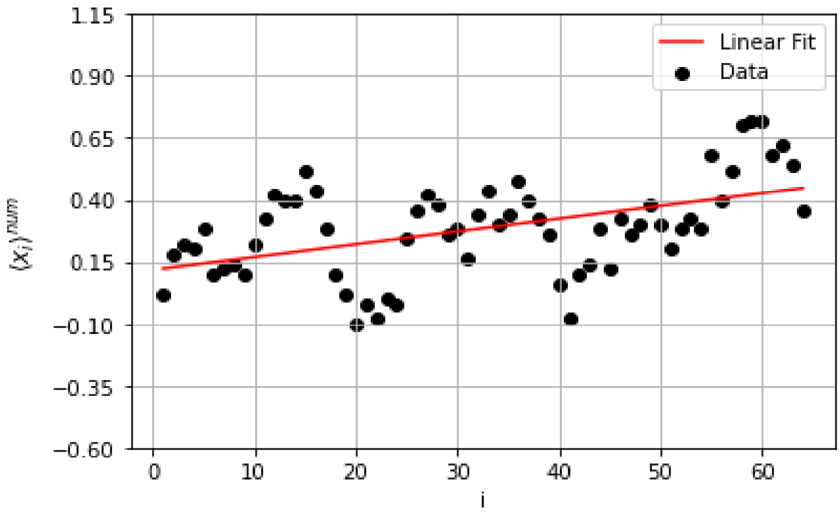


Fig 1.5 Linear regression of $\langle x_i \rangle$ for 100 independent walkers

In the case shown in *Fig. 1.5*, the parameters returned by the fit are:

$$a = 0,12 \pm 0,04$$
$$b = 0,005 \pm 0,001$$

which give a raw accordance with the theoretical trend within the 3σ[1].

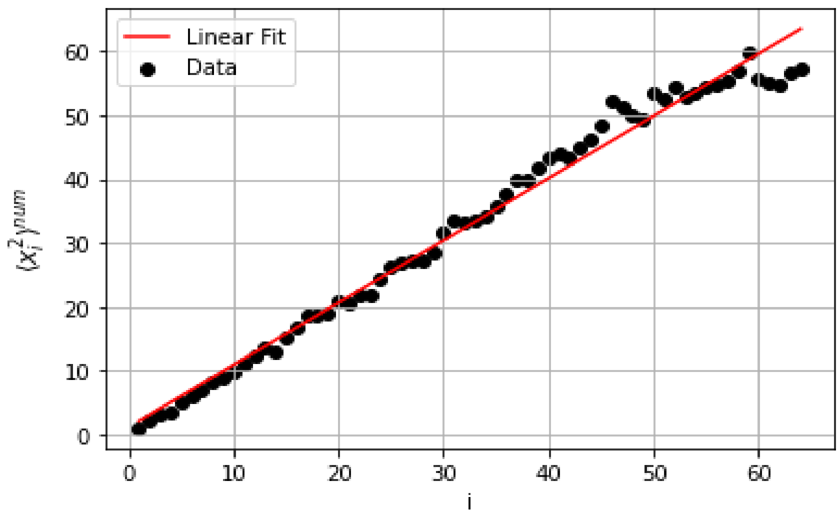


Fig 1.6 Linear regression for the averaged square-position of 100 independent walkers

While in the case corresponding to *Fig. i.6* the coefficients returned by the fit are:

$$a = 1.2 \pm 0.6$$
$$b = 0,97 \pm 0,02$$

which give the same raw accordance with the theoretical trend within the 3σ.

The other important variable we are interested into is the `mean square displacement` $\langle (\Delta x_N)^2 \rangle$. Indeed, in one of the most immediate applications of RW, According to Einstein's interpretation of

[1]Not for b, even though it is reasonably close to the expected value of 0. Anyway, We will later illustrate a possible explanation of this discrepancy.

Brownian motion, this measurable quantity is directly related to the diffusion coefficient. Its expected value is ($l = 1$):

$$\langle (\Delta x_i)^2 \rangle^{th} = i$$

From the numerical point of view, it can be determined instant by instant as $\langle (\Delta x_i)^2 \rangle = \langle x_i^2 \rangle - \langle x_i \rangle^2$. To compare the analytical result to the one obtained in the simulation we need again a linear fit:

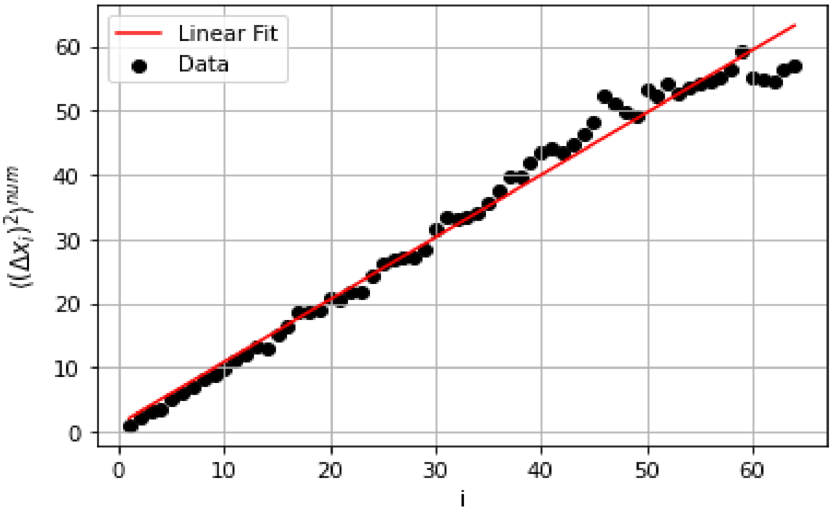


Fig 1.7 Linear regression for the variance of 100 independent walkers

Here are the coefficient are:

$$a = 1.16 \pm 0.6$$
$$b = 0,97 \pm 0,02$$

for whom the same previous considerations apply.

1.3 Accuracy of the mean squared distance

Clearly, we expect the grater the number of walkers over which averages are made the more accurate the numerical estimate of the mean square displacement is. Precisely, what we mean by *accuracy* is:

$$\Delta = \left| \frac{\langle (\Delta x_N)^2 \rangle^{num}}{\langle (\Delta x_N)^2 \rangle^{th}} - 1 \right| \tag{1}$$

As a consequence, an interesting question that arises in terms of computational costs is: "Which is the `minimum` number of walkers needed to obtain a desired accuracy, for instance of $\Delta \leqslant 5\%$?
We evaluated the Δ for an increasing number of walkers, then performing the averages over sets of increasing size. As soon an acceptable level of accuracy was reached, the corresponding value of *nwalkers* was saved; then all the procedure was repeated for other 10^5 times so that it was possible to estimate a reasonable average:

$$N_{walkers}^{min} \sim 160$$

From now on, the optimal value of 160 walkers will be held fixed.
In hindsight, this also tells us that the number of 100 walkers used before was underestimated. This is probably the reason why the numerical interpolations over the averaged quantities were not so satisfying: an accuracy issue.

A question which comes almost spontaneous now is: "What about N? Does the amount of walkers necessary to get the desired accuracy on Δ depend on the number of RW steps?
We tried to give a graphical qualitative answer, exploiting the same code used for the result right above and just varying N:

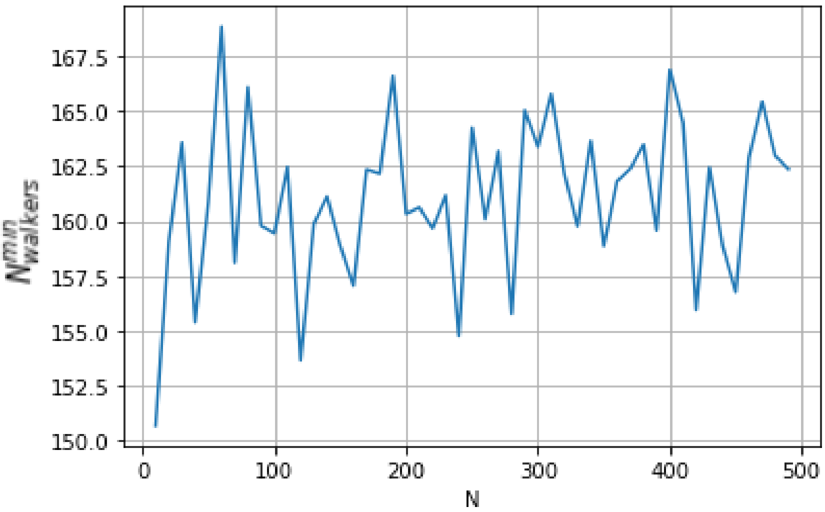


Fig 1.8 Average minimum number of walkers vs N

Apart from local fluctuations, it looks like the global trend settles around 160, at least up to $N = 500$. We can thus conclude the accuracy does not depend on N.

1.4 Dependence of $\langle(\Delta x_N)^2\rangle$ on N

Once we have set the optimal number of walkers $N_{walkers} = 160$ so that the accuracy is at least of the 5%, we can now study how the `mean square distance` varies with N.
On the basis of what reported at §1.2, we expect a linear dependence:

$$\langle(\Delta x_N)^2\rangle \sim N$$

Using the same code applied in the second subsection, we Plotted the results for N = 8, 16, 32, 64. 128 and subsequently operated a linear regression as shown in *Fig. 1.9.*

The fit parameter associated to the power of N is compatible with the expected value of 1, confirming our guess:

$$b = 1,00 \pm 0,05$$
$$a = 0.0 \pm 0.2$$

1.5 Distribution $P_N(x)$: numerical and theoretical comparison

If we store all the final positions reached after N steps by each one of the simulated walkers, we can statistically study how likely a certain place is to be occupied after a fixed number of steps. In other words we should watch at the *probability distribution* $P_N(x)$.

For the 1D case there is an exact analytical formula which can be well approximated for sufficiently large N by a gaussian function. The aim of this subsection is verifying from a qualitative point of view that this approximation holds; the function we are going to compare with the numerical data is:

$$P_N(x) \sim \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp -\frac{(x - \langle x\rangle)^2}{2\sigma^2} \tag{2}$$

where $\langle x\rangle$ and σ are estimated with the same code used in §1.2 for $N \in \{8, 16, 32, 64\}$. The origin of the approximation can be demonstrated considering the problem of RW in terms of a discretized differential

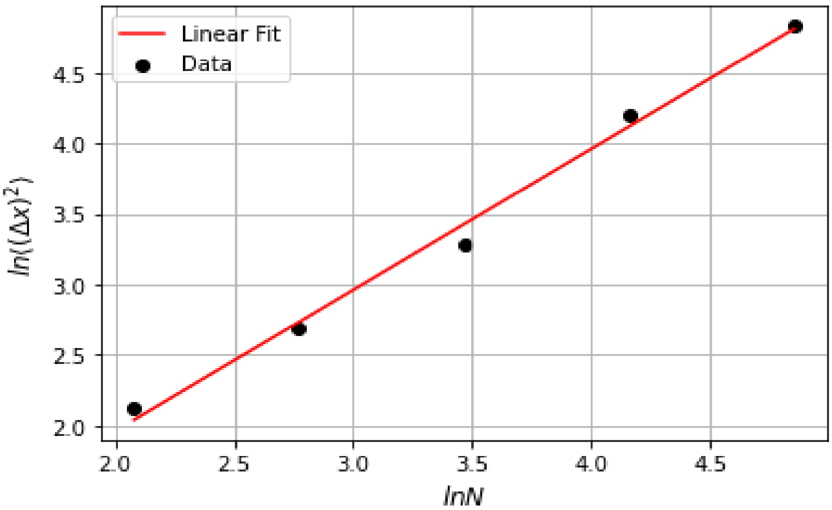


Fig 1.9 Mean square displacement vs N, log-log scale

equation which trace back to the typical diffusion equation in the continuum limit. In so doing, the standard deviation σ turns out to be proportional to the diffusion coefficient which is linked to the mean square distance as we already mentioned:

$$\sigma^2 \propto Diffusion\ \ coefficient \propto \langle (\Delta x_N)^2 \rangle$$

Adjusting the normalization so that the shapes of the superposed quantities dimensionally coincide[2], here are the outcomes. Since the additional computational time required was almost negligible, it was used 10000 as number of averaged walkers instead of 160. The outcomes are reported in *Fig. 1.10*. The correspondence is rather evident.

If we now take a look at the shapes, we notice that as N increases the height of the maximum decreases and the FWHM increases (and so the σ, i.e. $\langle (\Delta x_N)^2 \rangle$), coherently with what observed in the previous subsection.
From a physical point of view, this behaviour of this simple model well suits to describe the diffusive motion, where we expect the trajectories of suspended particles can cover wider and wider portion of space as time goes by and where numerically the *diffusion coefficient* can be indeed estimated through the $\langle (\Delta x_N)^2 \rangle$.

[2]There are a factor $N_{walkers}$ over the histogram to normalize it to 1 and a factor 2 in the normalization coefficient of the gaussian function as solution to the differential equation

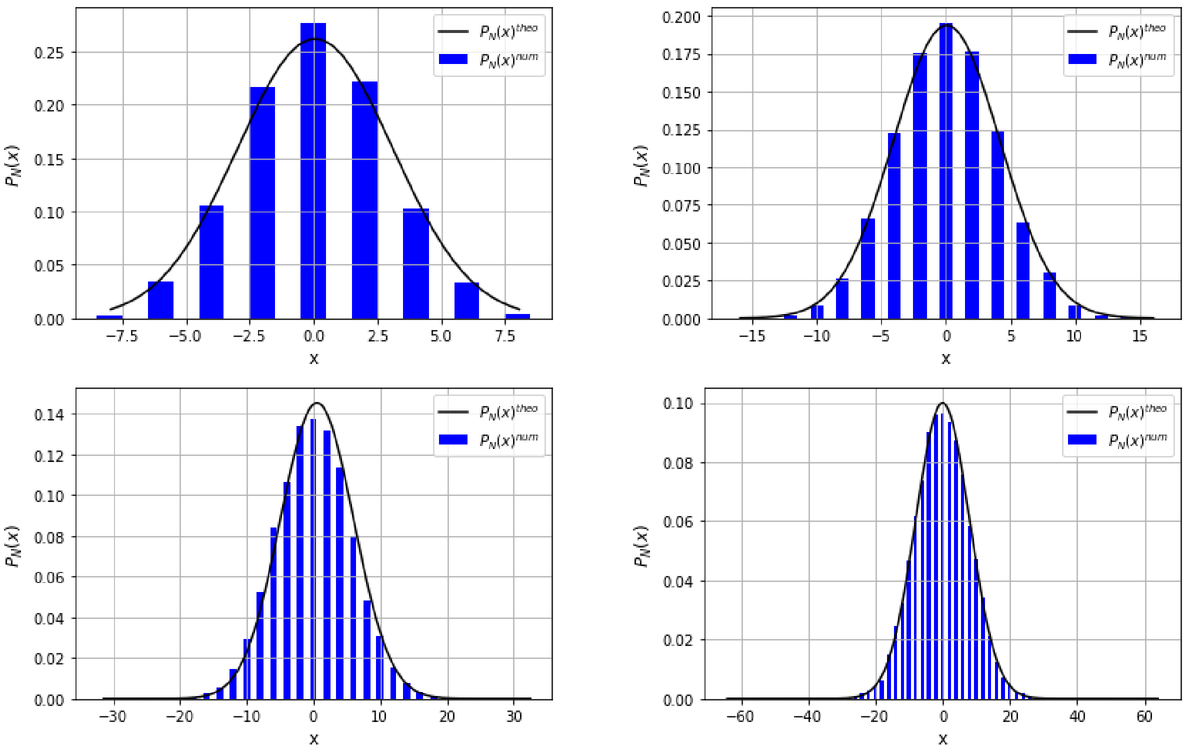


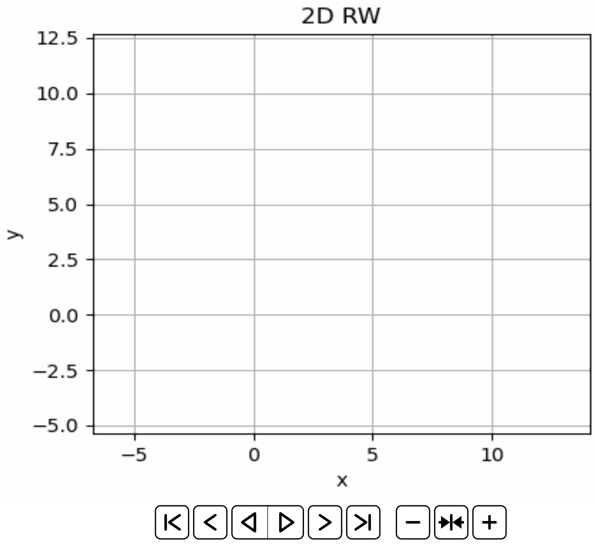
Fig. 1.10 Simulated probability distribution and asymptotic theoretical trend for a) $N = 8$, b) $N = 16$, c) $N = 32$, d) $N = 64$.

2 2D Random Walks (optional)

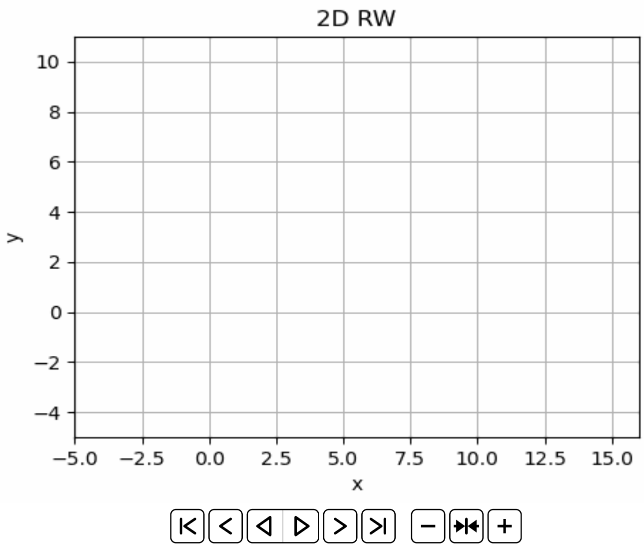
We are going to give here a rough idea of what happens when we extend the RW model to 2D. The algorithm used here to implement the two-dimensional walks essentially follows what had been done for the 1D case, with some proper changes.
For the sake of simplicity we focus now our attention on two basic examples:

1. 2D RW with unit displacement and arbitrary direction chosen uniformly in the interval $[0, 2\pi[$
2. 2D RW with unit displacement and arbitrary direction chosen with equal probability among *up, down, right and left* on a square lattice

You can find below the corresponding animations associated to a simulation with 4 walkers and N=64 steps.



Simulation of 2D RW with unit random direction in $[0, 2\pi[$



Simulation of 2D RW with unit random direction on a square lattice

2.1 Mean square distance

As before, the quantity we are mainly interested in is the *mean square distance* $\langle(\Delta R_N)^2\rangle$. Clearly we expect the defining statistical properties of 2D RWs emerge when considering averages on more than four walkers! Hence the "reasonable" number of walkers found in §1.3 gets back into the game.
We thus numerically evaluate

$$\langle(\Delta R_N)^2\rangle = \langle x_N^2\rangle + \langle x_N^2\rangle - \langle x_N\rangle^2 - \langle y_N\rangle^2$$

for both of the considered cases, applying a linear fit to the logarithm of the data:

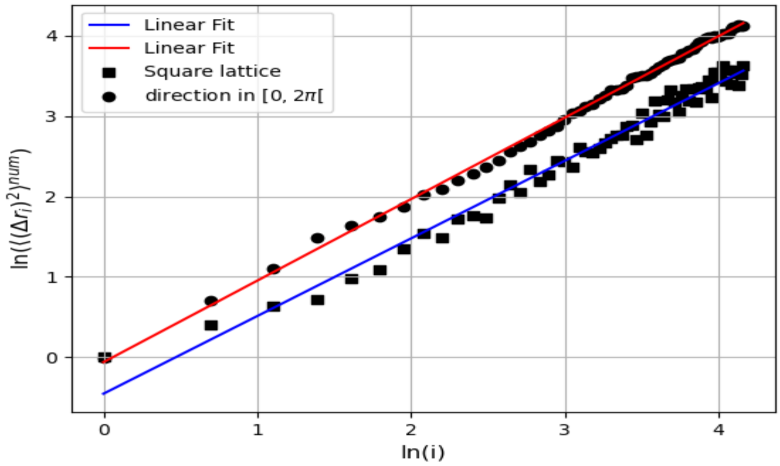


Fig 2.1 Mean square displacement vs N, log-log scale

$$\text{slope for the square lattice case} = 0.97 \pm 0,02$$
$$\text{slope for the random unit step case} = 1.02 \pm 0.01$$

Within the 3σ, this coefficients shows in a satisfactory way how the MSDs seem to follow the linear behaviour in N predicted by the theory for RWs in a generic dimension.

2.2 Probability distribution on the square lattice

Let's start from the example with the RWs on a square lattice. Which is the distribution followed by the radial displacement R_N?
The key idea is enumerating all the possible paths of fixed length that lead to a specific point. As a consequence, if the four different direction can be taken with equal probabilities at each step:

$$p((x_P, y_P), N) = \frac{n\ \ paths\ \ of\ \ length\ \ N\ \ to\ \ point\ \ P}{4^4}$$

Fig. 2.1 shows the relevant points on the lattice[3] for $N = 8$ alongside a numerical label. It represents all the possible numbers of horizontal steps that can be applied, starting from the origin, in order to reach that specific position.

Without getting into the details, a little bit of combinatorics[4] is enough to help us get through this task, finally obtaining the formula for the number of walks long N bringing to the point P. As a consequence, the discrete probability distribution (just for the points in the quadrant slice here) is:

[3]Thanks to symmetry arguments it is possible to restrict ourselves to the lower slice of the first quadrant and then recover the overall outcome as function of radial distance with just a bunch of trivial multiplications.

[4]Think: how many paths link the origin to the point P with s horizontal steps? You can think each path as a set of N instruction encoded in two coupled arrays, one for the horizontal moves and the other for the vertical ones.
For example P(1,1), N=4 and s=1: x->(0,1,0,0), y->(1,0,-1,1).
In general you will have two complementary N-ple; they will be made of 0, 1 and -1 each repeated a certain number of times!

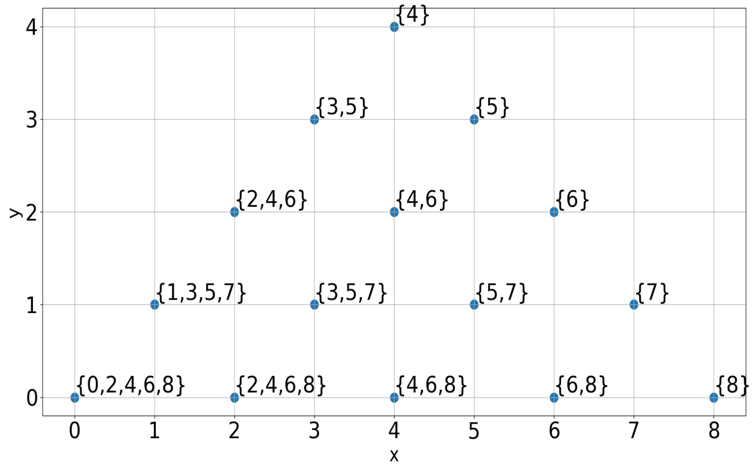


Fig 2.1 Square lattice, sufficient points and their number of allowed horizontal steps for $N = 8$.

$$p((x_P, y_P), N) = \frac{1}{4^4} \sum_s \frac{N!}{(\frac{s+x_P}{2})!(\frac{s-x_P}{2})!(N-s)!} * \frac{(N-s)!}{(\frac{N-s+y_P}{2})!(\frac{N-s-y_P}{2})!} \tag{3}$$

In *Fig. 2.2* follows a comparison between numerical results from simulations and analytical solutions of the above formula[5].

The resemblance is quite evident. Zooming in on the tails, we could notice some of the predicted values are not observed in the simulations: this simply happens because the probabilities for large distances tend to be so tiny that it would take runs with more than 10^7 walkers to see some counts over there.

2.3 The continuum limit

What happens if we now apply the continuum limit to the basic model of RW on the square lattice? Does the radial distribution just found above approaches any known *probability density function*? To understand what happens, we will follow a method based upon the construction of a partial differential equation from some known properties that hold for the discrete lattice. In particular:

$$p((i,j),N) = \frac{1}{4}p((i+1,j),N-1) + \frac{1}{4}p((i-1,j),N-1) + \frac{1}{4}p((i,j+1),N-1) + \frac{1}{4}p((i,j-1),N-1) \tag{4}$$

Defining $x = il$, $y = jl$ and $t = N\tau$ where l is the generic the step length and τ the proper time between two consecutive steps:

$$4p((x,y),t) = p((x+l,y),t-\tau) + p((x-l,y),t-\tau) + p((x,y+l),t-\tau) + p((x,y-l),t-\tau) \tag{5}$$

Subtracting $4p((x,y),t-\tau)$ from both sides and dividing everything by τ:

$$\begin{aligned}\frac{4p((x,y),t) - 4p((x,y),t-\tau)}{\tau} &= \frac{1}{\tau}\Big(p((x+l,y),t-\tau) + p((x,y+l),t-\tau) \\ &\quad + p((x-l,y),t-\tau) + p((x,y-l),t-\tau) - 4p((x,y),t-\tau)\Big)\end{aligned} \tag{6}$$

We can interpret this equation as the discretized version of a partial differential equation where partial derivatives made with respect to time (left side) and to space doubly (right side, laplacian) show off:

$$\frac{\partial p((x,y),t)}{\partial t} \approx \frac{l^2}{4\tau}\nabla^2 p((x,y),t) \tag{7}$$

[5]Numerical averages are carried out over 10000 walkers, just to make it easier to catch the main characteristics at glance

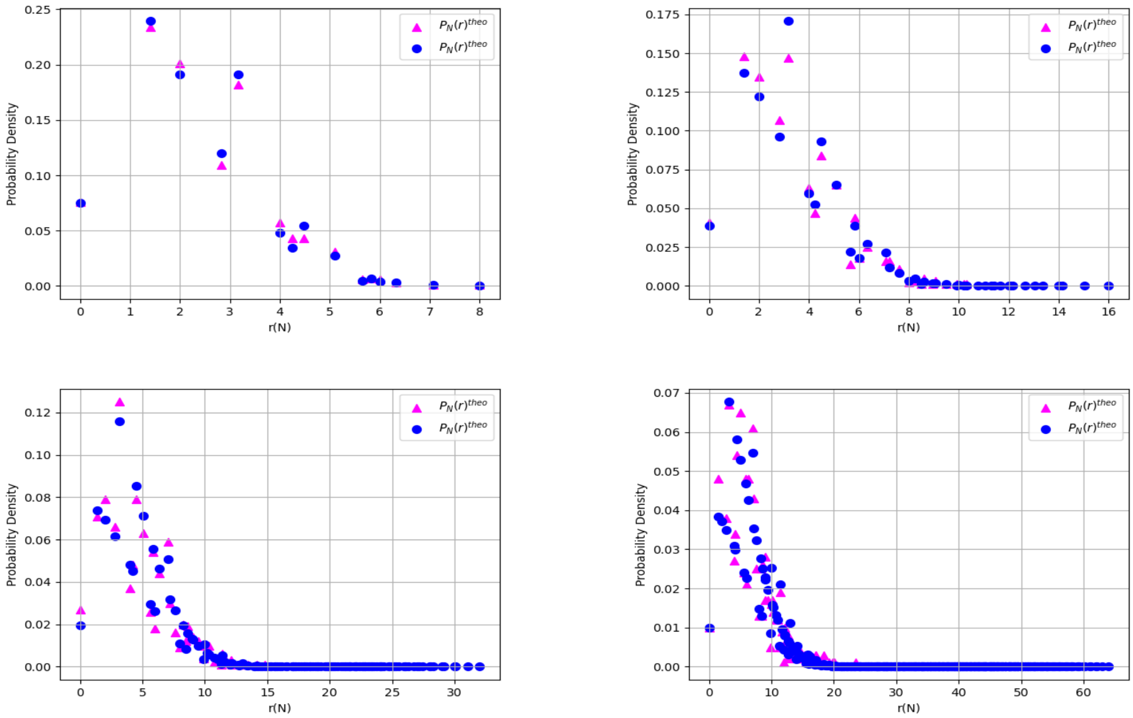


Fig. 2.2 Simulated 2D radial probability distribution and expected values for RWs on square lattice for a) $N = 8$, b) $N = 16$, c) $N = 32$, d) $N = 64$.

We now recognize in this equation the typical *2D diffusion equation*, which has the well known solution with coefficient $D = l^2/4\tau$:

$$p((x,y),t) = \frac{1}{4\pi Dt} e^{-\frac{(x^2+y^2)}{4Dt}} \tag{8}$$

We immediately observe it is separable into a product of two independent 1D gaussian distribution, one for each spacial direction.
In addition, for our purpose, since it comes more natural dealing with N on a computer simulation, we just make the solution depend explicitly on N rather than on time:

$$p((x,y),t) \to p((x,y),N) = \frac{1}{\pi N} e^{-\frac{(x^2+y^2)}{N} \equiv p(x,N)*p(y,N)} \tag{9}$$

We are almost done: we would like to find the probability density function associated to the variable $R \equiv \sqrt{X^2+Y^2}$ with X and Y that follow the normal distribution $p_X(x,N)$ and $p_Y(y,N)$. We must pass through the *cumulative distribution* of R, $F_R(r,N)$:

$$F_R(r,N) = \iint_{D_r} p_X(x,N) p_Y(y,N) d\mathbf{S} \qquad D_r = \{(x,y)\ \sqrt{x^2+y^2} \leqslant r\} \tag{10}$$

Writing everything in polar coordinates and remembering that, by definition, the probability density function $p_R(r,N)$ is the derivative of the cumulative distribution, by the fundamental calculus theorem we finally get:

$$p_R(r,N) = \frac{2r}{N} e^{-r^2/N} \tag{11}$$

known as *"Rayleigh probability density function*

If we build a properly normalized histogram with the data from the case with random unit displacement, and fit it with the Rayleigh function we see a satisfactory accordance:

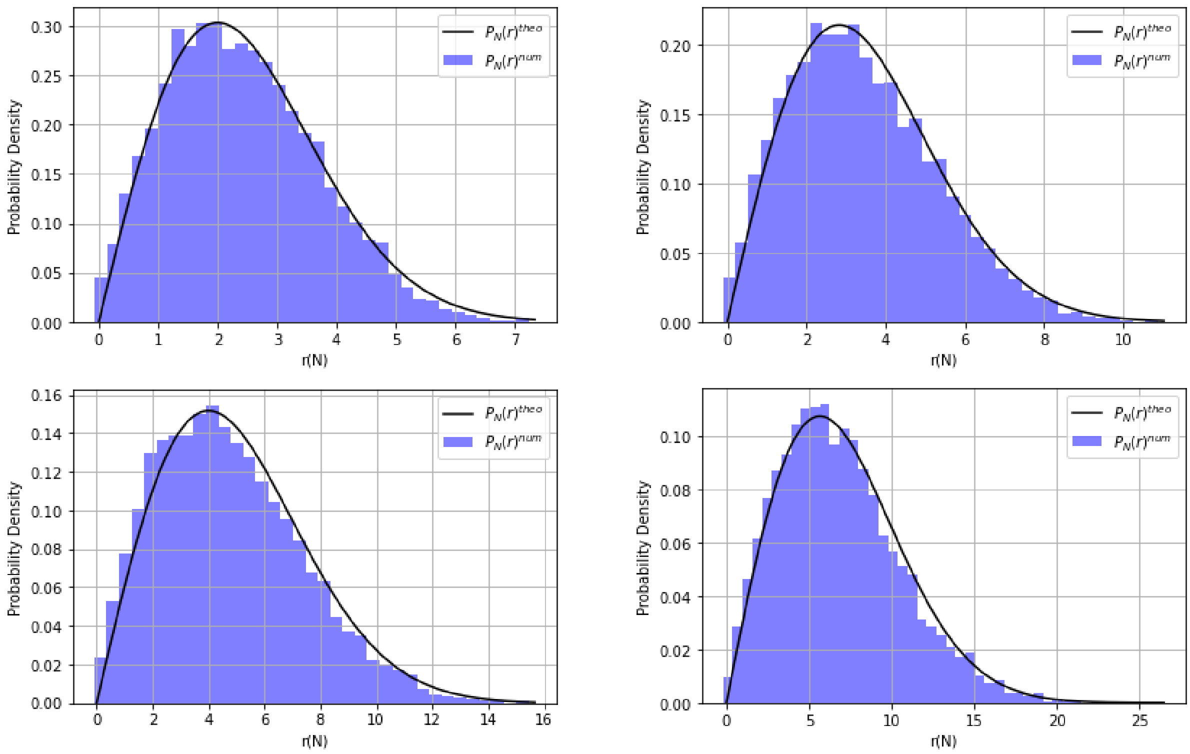


Fig. 2.3 Simulated 2D radial probability distribution and asymptotic theoretical trend for a) $N = 8$, *b)* $N = 16$, *c)* $N = 32$, *d)* $N = 64$.

The same goes for the case implemented on the square lattice: if we took the data showed in *Fig 2.2* and built the corresponding histograms, we would get shapes that progressively approaches the Rayleigh function as N grows (keeping fixed the number of walkers).

This qualitative check concludes our brief insight into the 2D Random Walks models.