# Quantum harmonic oscillator

**Computational physics laboratory**

**Fabio Coldani**
**814127**

**October 2, 2023**

Academic year 2021/2022

**Abstract**

We solved the 1D quantum harmonic oscillator with the lattice path integral approach through numerical techniques. By generating Feynman paths using a Markov chain Monte Carlo method and exploring various lattice configurations, we computed the energy difference and position operator matrix element between the ground state and the first excited state. The best results were $\Delta\tilde{E} = 0.9998(5)$ and $\left|\left\langle \tilde{E}_0 \middle| \hat{x} \middle| \tilde{E}_1 \right\rangle\right|^2 = 0.4998(3)$ for theoretical values of 1 and 0.5. The calculations were conducted on five lattices of progressively smaller spacing, whose results were extrapolated to the continuum with linear fits. In addition, our analysis focused on general features of Monte Carlo simulations of quantum theories.

## Contents

# 1 Introduction

The harmonic oscillator is likely the simplest physical system that is also worth studying. Despite said simplicity, a deep understanding of the harmonic oscillator is the key to understand complex topics such as classical mechanics and quantum field theory.

The path integral approach is particularly relevant in computational physics, as it provides the blueprint for the study of quantum field theories on the lattice. The steps taken to compute the energy gap between the ground and the first excited state of the harmonic oscillator are the same ones that are required to compute the mass of the lightest particle with a particular set of quantum numbers of a given theory – *e.g.* the proton in QCD – and the results present the great advantage of being able to be confronted with known analytical values, both on the lattice and in the continuum. This allows us to verify that the lattice numerical resolution, which we implemented with Monte Carlo techniques, is correct in both cases.

## 1.1 One dimensional quantum harmonic oscillator

The hamiltonian of a one dimensional quantum harmonic oscillator of mass $m$ and frequency $\omega/2\pi$ is

$$\hat{\mathrm{H}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{1}$$

with the position and momentum operators obeying the canonical commutation relations $\left[\hat{x}, \hat{p}\right] = i$[1] and its eigenvalues $E_n$ and eigenstates $|E_n\rangle$ defined by the equation

$$\hat{\mathrm{H}}\,|E_n\rangle = E_n\,|E_n\rangle\,. \tag{2}$$

It is convenient to introduce the ladder operators $\hat{a}^\dagger$ and $\hat{a}$

$$\hat{a}^\dagger \equiv \sqrt{\frac{m\omega}{2}}\left(\hat{x} - i\frac{\hat{p}}{m\omega}\right), \tag{3}$$

$$\hat{a} \equiv \sqrt{\frac{m\omega}{2}}\left(\hat{x} + i\frac{\hat{p}}{m\omega}\right), \tag{4}$$

$$\left[\hat{a}, \hat{a}^\dagger\right] = 1, \tag{5}$$

$$\hat{a}\,|E_n\rangle = \sqrt{n}\,|E_{n-1}\rangle\,, \tag{6}$$

$$\hat{a}^\dagger\,|E_n\rangle = \sqrt{n+1}\,|E_{n+1}\rangle \tag{7}$$

which can be used to rewrite $\hat{\mathrm{H}}$ and easily find its eigenvalues

$$\hat{\mathrm{H}} = \omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right),\ E_n = \omega\left(n + \frac{1}{2}\right). \tag{8}$$

[1] In *natural units* where $\hbar = c = K_B = 1$.

Rewriting $\hat{x}$ in terms of $\hat{a}$ and $\hat{a}^\dagger$, the computation of matrix elements is also straightforward, in particular we will make use of

$$\langle E_0|\hat{x}|E_0\rangle = 0, \quad \langle E_0|\hat{x}|E_1\rangle = \frac{1}{\sqrt{2m\omega}},$$
$$\langle E_0|\hat{x}^2|E_0\rangle = \frac{1}{2m\omega}. \tag{9}$$

## 1.2 Lattice path integral for the harmonic oscillator

Feynman's formulation of quantum mechanics reveals a connection between the quantum theory and statistical mechanics through the equivalence of the Euclidean path integral and the partition function of a statistical system with Boltzmann weight $e^{-S}$, where $S$ is the classical action we intend to quantize. As a result, we are able to apply the well known Monte Carlo techniques of statistical physics to quantum mechanics[2], provided we can rigorously define the path integral.

The only way to do so is to perform a discretization of Euclidean spacetime with cells of size $a$, put it in a box of finite size $Na$ and work at finite values of $a$ and $N$, then – at a later stage – take the limits $a \to 0$, $N \to \infty$ of the results[3].

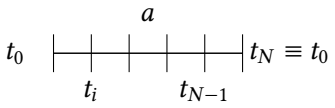


Figure 1: Discretization of the time interval – we have defined a one dimensional lattice of $N$ sites and spacing $a$ with periodic boundary conditions.

For one dimensional systems, we are left with a lattice with $N$ discrete time coordinates $t_i$, with $i = 0, \ldots, N-1$. Therefore we define the position coordinate at time $t_i$ to be $x_i \equiv x(t_i)$, from which we define the discretized versions of the harmonic oscillator Langrangian and the corresponding action

$$\mathcal{L}(x_{i+1}, x_i) = \frac{m}{2}\left(\frac{x_{i+1} - x_i}{a}\right)^2 + \frac{1}{4}m\omega^2\left(x_{i+1}^2 + x_i^2\right) \tag{10}$$

$$S = a\sum_{i=0}^{N-1}\mathcal{L}(x_{i+1}, x_i). \tag{11}$$

The lattice partition function of such a system is

$$Z(T, a) \equiv \left(\frac{m}{2\pi a}\right)^{\frac{N}{2}}\int\prod_{i=0}^{N-1}\mathrm{d}x_i\, e^{-S}$$
$$= \mathrm{Tr}\left[\hat{\mathrm{T}}_a^N\right], \tag{12}$$

[2] And potentially to quantum field theories.

[3] Eventually we can *rotate* back to the Minkowski spacetime.

where $\hat{\mathrm{T}}_a$ is the transfer operator

$$\hat{\mathrm{T}}_a = \exp\left[-\frac{a}{2}V(\hat{x})\right]\exp\left(-a\frac{\hat{p}^2}{2m}\right)\exp\left[-\frac{a}{2}V(\hat{x})\right],$$
$$V(\hat{x}) = \frac{1}{2}m\omega^2\hat{x}^2 \tag{13}$$

which evolves the system on the lattice of one step of size $a$. This is the analogous to the time evolution operator of the continuum theory and, as it is explained in [CF81], its diagonalization is equivalent to finding the time evolution of the discretized system.

The transfer operator is unitary by construction, therefore it is possible to express it as the exponentiation of a self adjoint operator with eigenvalues $\varepsilon_n$ and eigenstates $|\varepsilon_n\rangle$

$$\hat{\mathrm{T}}_a|\varepsilon_n\rangle = e^{-a\varepsilon_n}|\varepsilon_n\rangle. \tag{14}$$

To diagonalize it we introduce a hamiltonian $\hat{\bar{\mathrm{H}}}$

$$\hat{\bar{\mathrm{H}}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\bar{\omega}^2\hat{x}^2, \tag{15}$$
$$\bar{\omega}^2 \equiv \omega^2\left(1 + \frac{a^2\omega^2}{4}\right) \tag{16}$$

where $\omega$ is the same frequency that appears in the Lagrangian 10. This is an auxiliary harmonic oscillator on the continuum, whose frequency $\bar{\omega}$ depends on $a$, with energy levels

$$\tilde{E}_n = \tilde{\omega}\left(n + \frac{1}{2}\right), \tag{17}$$
$$\tilde{\omega} = \frac{1}{a}\log\left(1 + a\bar{\omega} + \frac{a^2\omega^2}{2}\right) \tag{18}$$

which, in the limit $a \to 0$, rebuild the spectrum of the *original* harmonic oscillator.

Since $\hat{\bar{\mathrm{H}}}$ and $\hat{\mathrm{T}}_a$ commute, they share a basis of eigenstates $\left\{\left|\tilde{E}_n\right\rangle\right\}$ with eigenvalues that are respectively $\tilde{E}_n$ and $e^{-a\tilde{E}_n}$. Therefore combining equation 14 and the definition 12 we explicitely find the partition function both on the lattice and in the continuum

$$Z(T,a) = \sum_{n=0}^{\infty} e^{-a\tilde{E}_n N} \underbrace{\longrightarrow}_{\substack{N\to\infty \\ a\to 0}} Z(T) = \sum_{n=0}^{\infty} e^{-E_n T}. \tag{19}$$

Finally, simply by substituting $\bar{\omega}$ to $\omega$ in equations 9 we find the useful matrix elements

$$\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_0\right\rangle = 0, \quad \left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle = \frac{1}{\sqrt{2m\bar{\omega}}},$$
$$\left\langle\tilde{E}_0\middle|\hat{x}^2\middle|\tilde{E}_0\right\rangle = \frac{1}{2m\bar{\omega}}. \tag{20}$$

## 1.3 Purpose of the simulation

The core of this work is the numerical simulation of a quantum harmonic oscillator on the lattice, with the purpose of recovering the analytical results both for the discretized theory and in the continuum. Namely, we will compute the observables

$$\Delta\tilde{E} \equiv \tilde{E}_1 - \tilde{E}_0 = \tilde{\omega} \tag{21}$$
$$\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2 = \frac{1}{2m\bar{\omega}} \tag{22}$$

and use them to extrapolate the equivalent quantities of the continuum theory sending $a \to 0$.

In this limit, the lattice results differ from the continuum quantities by linear corrections in $a^2$

$$\Delta\tilde{E} \approx \omega - \frac{\omega^3}{24}a^2 \tag{23}$$
$$\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2 \approx \frac{1}{2m\omega} - \frac{\omega}{16m}a^2 \tag{24}$$

so we need to choose a small enough lattice spacing to approximate the continuum limit, while keeping the product $Na$ sufficiently large that the ground state properties of the system can still be observed. If we define $T_E \equiv {}^{2\pi}/_{E_0}$, a sensible choice according to [CF81] is to consider $a \sim 0.1, 0.2 \times T_E$ and $Na \sim 3, 10 \times T_E$. In this work we kept $Na = 64$ and we performed the same simulation on 5 different lattices with progressively smaller spacing $a$, eventually obtaining the continuum ($a = 0$) results after a linear fit on $a^2$.

These quantum properties can be extracted from the two points correlation function of the position operator: if we select two sites $l$ and $k$ on the lattice, such that $t_{lk} \equiv a|l-k|$[4], we can define

$$c(t_{lk}) \equiv \langle x_l x_k\rangle = \frac{1}{Z(T,a)}\int\prod_{i=0}^{N-1}\mathrm{d}x_i\, x_l x_k e^{-S} \tag{25}$$

as our primary observables, which we intend to evaluate with numerical integration methods. The integral is invariant under translation, therefore the correlator is independent of the particular sites $l$, $k$ and actually depends on their distance $|l-k| = {}^{t_{lk}}/_a$.

We can rewrite the correlator with the operator formalism

$$c(t_{lk}) = \frac{1}{\mathrm{Tr}\left[\hat{\mathrm{T}}_a^N\right]}\mathrm{Tr}\left[\hat{\mathrm{T}}_a^{N-|l-k|}\hat{x}\hat{\mathrm{T}}_a^{|l-k|}\hat{x}\right]; \tag{26}$$

and, after inserting two complete sets of $\hat{\mathrm{T}}_a$ eigenstates, the dominant contribution in the large $|l-k|$

[4] Physical time $t$ will always appear in unit of $a$.

and $N$ limit is found to be

$$c(t_{lk}) \approx 2\left|\left\langle \tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2 \exp\left(-\frac{N}{2}a\Delta\tilde{E}\right)\times \times \cosh\left[\left(\frac{N}{2}-|l-k|\right)a\Delta\tilde{E}\right]. \quad (27)$$

This expression is even with respect to $|l-k| = N/2$, so we can limit our analysis to the correlators whose distance $|l-k|$ is at most $N/2$.
To find the energy gap without having to fit the data using equation 27 as the model, we consider a combination of correlators at different physical times and invert it to find the analytic expression

$$\Delta\tilde{E} = \frac{1}{a}\cosh^{-1}\left[\frac{c(t_{lk}+1)+c(t_{lk}-1)}{2c(t_{lk})}\right] \quad (28)$$

which can be plugged back into equation 27 to find the matrix element

$$\left|\left\langle \tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2 = \frac{c(t_{lk})\exp\left(\frac{N}{2}a\Delta\tilde{E}\right)}{2\cosh\left[\left(\frac{N}{2}-|l-k|\right)a\Delta\tilde{E}\right]}. \quad (29)$$

Moreover, we can also prove that the noise to signal ratio of the correlator grows exponentially with $t_{lk}$ and therefore only small values of $t_{lk}$ are going to be useful for our analysis. This was highlighted by Giorgio Parisi in 1983 and was given the name of *exponential problem*.

Assuming the observables $c(t_{lk})$ are averaged over $N_{config}$ configurations, the variance is

$$\sigma^2[c(t_{lk})] = \frac{1}{N_{config}}\left[\left\langle x_l^2 x_k^2\right\rangle - \langle x_l x_k\rangle^2\right], \quad (30)$$

where $\langle x_l x_k\rangle$ is precisely $c(t_{lk})$, while $\left\langle x_l^2 x_k^2\right\rangle$ is the two points correlation function of the $\hat{x}^2$ operator.
Like we did for the correlator of $\hat{x}$, we can write $\left\langle x_l^2 x_k^2\right\rangle$ using the operator formalism and, following the same steps, recover its dominant contribution for large $N$

$$\left\langle x_l^2 x_k^2\right\rangle = \frac{1}{\mathrm{Tr}\left[\hat{\mathrm{T}}_a^N\right]}\mathrm{Tr}\left[\hat{\mathrm{T}}_a^{N-|l-k|}\hat{x}^2\hat{\mathrm{T}}_a^{|l-k|}\hat{x}^2\right] \approx \left|\left\langle \tilde{E}_0\middle|\hat{x}^2\middle|\tilde{E}_0\right\rangle\right|^2. \quad (31)$$

Before we continue discussing the exponential problem, it is worth noting that for large values of $N$, as the distance $|l-k|$ grows, the variance (equation 30) will be dominated by the constant value of $\left\langle x_l^2 x_k^2\right\rangle$[5]

$$\sigma^2[c(t_{lk})] \approx \frac{\left|\left\langle \tilde{E}_0\middle|\hat{x}^2\middle|\tilde{E}_0\right\rangle\right|^2}{N_{config}}, \quad (32)$$

meaning that it would be possible to use the variance of our numerical computation to extract the ground state expectation value of $\hat{x}^2$.
On the other hand, if we focus on the region where $N$ and $|l-k|$ are large but $|l-k| \ll N/2$, equation 27 implies that the correlators decrease very rapidly with $|l-k|$ [6]

$$c(t_{lk}) \approx \left|\left\langle \tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2 e^{-a|l-k|\Delta\tilde{E}}. \quad (33)$$

The noise to signal ratio, *i.e.* the relative error, therefore grows exponentially with the distance of the sources

$$\frac{\sigma^2[c(t_{lk})]}{c(t_{lk})^2} \approx \frac{1}{N_{config}}\frac{\left|\left\langle \tilde{E}_0\middle|\hat{x}^2\middle|\tilde{E}_0\right\rangle\right|^2}{\left|\left\langle \tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^4}e^{2a|l-k|\Delta\tilde{E}}. \quad (34)$$

This behaviour greatly limits the number of correlators we can actually use to calculate the energy gaps and matrix elements (equations 28 and 29) for a given number of configurations, as the value of the observables is eventually going to be dominated by their error. A possible solution would be to exponentially increase the number of configurations, which would heavily worsen the computational cost.

More generally, the origin of this problem in QFTs is to be found in the quantization of the energy spectrum which leads to the presence of a mass gap in the theory.

## 2 Numerical simulation

In order to extract the quantum properties of the discretized system (equations 28 and 29) and recover the continuum results (equations 21 and 22) we have to compute the correlators $c(t_{lk})$, defined as the $N-$dimensional integral in equation 25. We proceed numerically, via the importance sampling Monte Carlo method, using the Metropolis algorithm to design an ergodic Markov chain with asymptotic probability distribution $e^{-S}/Z$. The chain will be used to extract the random variables needed for the integration, which are physically equivalent to considering different Feynman paths.

In this section we discuss the Metropolis algorithm and the subsequent Markov chain in greater detail. The data shown, unless explicitely specified, have been computed using the following lattice parameters.

### 2.1 Monte Carlo method

As highlighted in the previous section, the initial goal of the simulation is the numerical evaluation of the integral in equation 25. The random variables needed

[5] Equation 27 shows that $c(t_{lk})$ is suppressed exponentially with $N$ when $|l-k|$ is large enough.

[6] $\cosh(x) \sim e^x/2$ as $x \to \infty$.

| Parameters | | | |
|---|---|---|---|
| $N$ | 256 | $m$ | 1 |
| $a$ | 0.25 | $\omega$ | 1 |
| $N_{sweep}$ | 1e6 | $\Delta$ | 1 |

Table 1: Parameters used for the analysis shown in the next sections.

for the integration are the Feynman paths $E_J \equiv \{x_0, \dots, x_{N-1}\}_J$.

To ensure that each new configuration is extracted following the right probability distribution $e^{-S}/Z$, we choose to work with an ergodic Markov chain. Indeed, an ergodic Markov chain with probability transition matrix $P_{JK}$ is guaranteed to have an asymptotic distribution $\pi_J$ whose form does not depend on the starting configuration; thus, we just need to design an ergodic Markov chain such that $\pi_J = e^{-s}/Z$.

A sufficient condition for a Markov chain to be ergodic is for it to obey the detailed balance $\pi_J P_{JK} = \pi_K P_{KJ}$ and a sufficient condition for a Markov chain to satisfy the detailed balance is to be generated by the Metropolis algorithm. Therefore, it is enough for us to implement the Metropolis algorithm to eventually generate suitable configurations – *i.e.* suitable Feynman paths.

The Metropolis Monte Carlo algorithm is composed of two steps

1. **Proposal:** given an initial state $E_J$, a state $E_K$ with transition probability $Q_{JK}$ satisfying the microreversibility condition[7] is proposed;

2. **Accept/Reject:** to generate events with asymptotic probability distribution $\pi_J = {}^{R_J}/_{\sum_J R_J}$[8], the new configuration is accepted with probability 1 if $R_K \geq R_J$ and with probability ${}^{R_K}/_{R_J}$ otherwise. Thus, we construct a transition probability matrix

$$P_{JK} = \begin{cases} Q_{JK} \text{ if } R_K \geq R_J \\ Q_{JK} \dfrac{R_K}{R_J} \text{ if } R_K < R_J \end{cases} . \tag{35}$$

If a new configuration is rejected, the algorithm restarts from $E_J$ and a new proposal is made.

In our implementation the asymptotic distribution is $e^{-S}/Z$, so a new proposal $E_K$ is certainly accepted if the action decreases[9] and accepted with probability $e^{-(S[E_K]-S[E_J])}$ if the action increases. Configurations that correspond to larger actions are still allowed, but they are at a disadvantage compared to those that lead to smaller actions. The algorithm makes a new proposal for each coordinate $x_i$, one at a time, of the starting configuration. That is

$$\begin{aligned} E_K &= \{x_0, \dots, x_i, \dots, x_{N-1}\} \mapsto \\ E_{K+1} &= \{x'_0, \dots, x_i, \dots x_{N-1}\} \mapsto \\ &\vdots \\ E_{K+N} &= \{x'_0, \dots, x'_i, \dots, x'_{N-1}\}. \end{aligned} \tag{36}$$

Once every coordinate has undergone the accept/reject phase we say that a *sweep* has happened and *Markovian time* increases by one unit (figure 2).

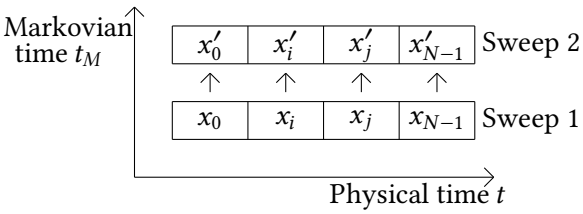


Figure 2: The elementary step of the Markov chain consists of the (potential) update of one coordinate, after the elementary step has happened for every coordinate a sweep is completed. The sweep index is called *Markovian time*. After a sufficiently large number of sweeps, the chain is said to be thermalized and each new random $x_i$ will be extracted following $e^{-S}/Z$. Physically, a new sweep is a new Feynman path.

The proposal for an updated coordinate $x'_i \leftarrow\!\shortmid\, x_i$ is chosen in the closed interval $[x_i - \Delta, x_i + \Delta]$ ($\Delta$ is a positive, real algorithm parameter that does not impact the physics) and it depends on a random number $r_1$, generated with uniform distribution between 0 and 1[10]

$$x'_i = x_i + 2\Delta\left(r_1 - \frac{1}{2}\right). \tag{37}$$

The accept/reject phase of the algorithm is implemented with a second random number $r_2$, also generated with uniform distribution between 0 and 1

$$\begin{cases} \text{if } e^{-\Delta S} \geq r_2 \text{ accept } x'_i & \implies x_i \mapsto x'_i \neq x_i \\ \text{if } e^{-\Delta S} < r_2 \text{ reject } x'_i & \implies x_i \mapsto x'_i \equiv x_i \end{cases} . \tag{38}$$

Since the chain only tries to update one coordinate at a time, if we call the coordinate shift $\delta x_i$, the corresponding variation of the action is

$$\Delta S = \frac{m}{2a}\delta x_i\left[\left(2 + a^2\omega^2\right)(2x_i + \delta x_i) - 2(x_{i+1} + x_{i-1})\right]. \tag{39}$$

[7] $Q_{JK} = Q_{KJ}$

[8] Note that this has the same form of $e^{-S}/Z$.

[9] $\Delta S = S[E_K] - S[E_J] < 0$.

[10] As $r_1$ is chosen with a flat distribution microreversibility is guaranted – *i.e.* the transition $x_i \mapsto x'_i$ is equiprobable to the transition $x'_i \mapsto x_i$.

Now the discussion of the experiment can begin, starting from the thermalization of the chain.

## 2.2 Thermalization of the chain

Before the chain reaches its asymptotic probability distribution we need to let it evolve for a number of sweeps. Figure 3 shows that the thermalization value is reached independently of the initialization values of $E_J$ for a given lattice, which we initalized in four ways:

1. with random numbers between $-1$ and 1;
2. with random numbers between $-5$ and 5;
3. setting $x_i = 0\ \forall i$ (cold initialization);
4. setting $x_i = 10\ \forall i$ (hot initialization).

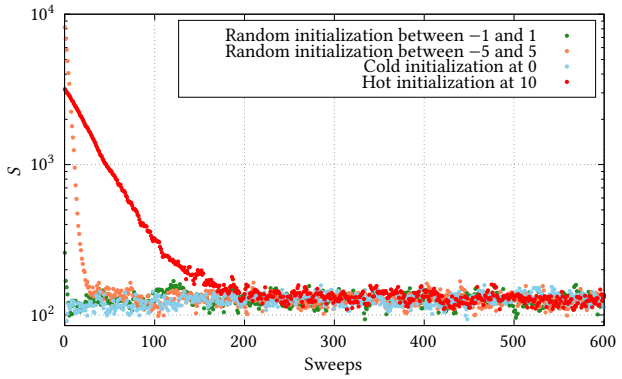


Figure 3: Thermalization of different Markov chains. Regardless of their initialization, they eventually start oscillating around an equilibrium value.

It is worth noting that while the thermalization value is the same, the speed is different. The *colder* initialized chains take $\sim 30, 50$ sweeps to thermalize, while the *hot* chain needs around $\sim 200$. As a safety measure, while a colder initialization was generally preferred, due to the low computational costs of sweeping we decided to let each chain thermalize for $N_{therm} = 1000$ sweeps[11] throughout the whole simulation.

As we will see in subsection 2.5, the width of the interval used to extract the coordinate update proposals $x_i'$ has an effect on the thermalization speed.

Once the chain is thermalized, it is sensible to compute the correlators.

[11] This proved to be more than enough for every lattice we considered.

## 2.3 Improved observables

The integrals we are interested in are of the form

$$\langle O(x_l)O(x_k)\rangle = \frac{1}{Z(T,a)} \int \prod_{i=0}^{N-1} \mathrm{d}x_i\, e^{-S} O(x_l)O(x_k), \tag{40}$$

whose best estimate is

$$\overline{O(x_l)O(x_k)} = \frac{1}{N_{config}} \sum_{n=1}^{N_{config}} O(x_l)_n O(x_k)_n \tag{41}$$

in the limit of large $N_{config}$. Clearly, the value of the integral is

$$\langle O(x_l)O(x_k)\rangle = \lim_{N_{config}\to\infty} \overline{O(x_l)O(x_k)}. \tag{42}$$

Furthermore, we can exploit the path integral's translational invariance to define improved observables with smaller error. For each configuration we define

$$O(t_{lk})_n = \frac{1}{N} \sum_{i=0}^{N-1} O(x_{l+i})_n O(x_{k+i})_n \tag{43}$$

so that we identify $\overline{O(t_{lk})} = \overline{O(x_l)O(x_k)}$ and find an improved version of the observable 41 as

$$\overline{O(t_{lk})} = \frac{1}{N_{config}} \sum_{n=1}^{N_{config}} O(t_{lk})_n. \tag{44}$$

Thus, the best estimate of the integral 25 is

$$\overline{c(t_{lk})} = \langle x_l x_k\rangle = \frac{1}{N_{sweep}} \sum_{n=1}^{Nsweep} c(t_{lk})_n, \tag{45}$$

where the improved observables are, at fixed physical time $t_{lk}$ and for each Feynman path,

$$c(t_{lk})_n = \frac{1}{N} \sum_{i=0}^{N-1} \left[x_i x_{i+t_{lk}}\right]_n \tag{46}$$

and $n$ labels a thermalized sweep of the Markov chain.

## 2.4 Autocorrelation and binning

At this point we are confronted with a problem. The best estimate of the correlator is an average over the number of thermalized configurations, but each *next* configuration is generated, site by site, from the *previous*. In other words, the configurations corresponding to subsequent sweeps are not independent: the observables $c(t_{lk})_n$ are correlated in Markovian time

$n$. As a consequence, in the limit $N_{sweep} \to \infty$, the variance of our numerical computation is

$$\sigma^2\left[\overline{c(t_{lk})}\right] = \frac{\sigma^2[c(t_{lk})]}{N_{sweep}}\left\{1 + 2\sum_{n=1}^{\infty}\frac{\Gamma_{t_{lk}}(n)}{\Gamma_{t_{lk}}(0)}\right\} \tag{47}$$

where $\Gamma_{t_{lk}}(n) \equiv \langle c(t_{lk})_m c(t_{lk})_{m+n}\rangle_m - \overline{c(t_{lk})}^2$ is the autocorrelation function at (Markovian) time $n$[12]. The ratio $\Gamma_{t_{lk}}(n)/\Gamma_{t_{lk}}(0)$ is a measure of the correlation between two observables $c(t_{lk})_m$ that are $n$ sweeps apart[13]. Indeed, if the ratio was 0 we would recover the variance on the mean of uncorrelated measures that we would expect.

For a Markov chain,

$$\frac{\Gamma_{t_{lk}}(n)}{\Gamma_{t_{lk}}(0)} \sim \exp\left(-\frac{n}{\tau}\right); \tag{48}$$

we call $\tau$ the *autocorrelation time* and note that after $\sim 4,5\tau$ sweeps the correlation between two observables becomes negligible. Therefore, it is possible

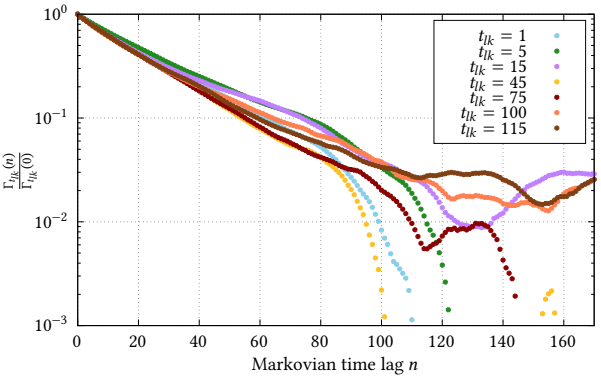


Figure 4: Autocorrelation as a function of Markovian lag for observables at different fixed physical times. Its value decreases exponentially with $n$ – $y$ axis is in log scale – until it starts oscillating. Observables corresponding to larger physical times (equivalent to the distance of the sites on the lattice) need more sweeps to decorrelate.

to decorrelate the measures through a binning procedure, in order to minimize the impact of consequent Feynman paths. The value of $\tau$ can be determined from an exponential fit of ratio 48 and used to group $D_{bin} \sim 10, 100\tau$ observables together in a single bin, of which we take the average value. As a consequence, we will now work with $N_{bin} = N_{sweep}/D_{bin}$ decorrelated observables

$$c(t_{lk})_n = \frac{1}{D_{bin}}\sum_{j=1}^{D_{bin}} c(t_{lk})_j;\ n = 1, \dots, N_{bin} \tag{49}$$

on which we can apply the *usual* statistical analysis over $N_{bin}$ Feynman paths, thus obtaining the best estimation of the correlator 25 as

$$c(t_{lk}) = \frac{1}{N_{bin}}\sum_{n=1}^{N_{bin}} c(t_{lk})_n, \tag{50}$$

with variance[14]

$$\sigma^2[c(t_{lk})] = \frac{1}{N_{bin}}\left[\left\langle x^2_{l,binned} x^2_{k,binned}\right\rangle - c(t_{lk})^2\right]. \tag{51}$$

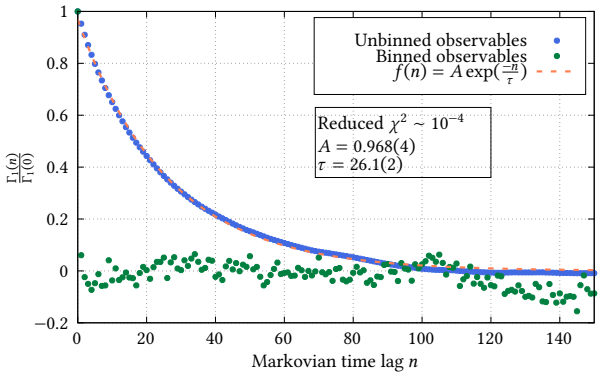


Figure 5: Autocorrelation for binned and unbinned correlators; the exponential fit has been done with Gnuplot. The value of $\tau = 26$ was then used to determine the bin width $D_{bin} = 500$. The binned observables, shown here in green, are decorrelated from the start.

In figure 5 we show the autocorrelation for both the unbinned and the binned correlator at $t_{lk} = 1$. It takes $\tau \sim 26$ sweeps for the unbinned observables' autocorrelation to decrease to $1/e$, therefore we set $D_{bin} = 500$. As expected, the binned observables among different sweeps are not correlated from the start.

## 2.5 The $\Delta$ parameter

Now we are ready to investigate the effects of the different choices of the algorithm parameter $\Delta$ on the thermalization speed and autocorrelation time of the chain. The first concept we need to introduce is the acceptance.

The proposal for an updated coordinate (equation 37) selects an $x_i'$ inside a closed interval of center $x_i$ and width $\Delta$: it is reasonable that such parameter influences the efficiency of the Metropolis algorithm. We define *acceptance rate* or simply *acceptance* the frequency of a new proposal getting accepted.

The higher the acceptance, the more likely it is that a new proposal is accepted, which is desirable

[12] It is worth noting that $\Gamma_{t_{lk}}(0) = \sigma^2[c(t_{lk})]$

[13] $n$ is also known as *lag*.

[14] The binning procedure on $x^2$ happened *after* the coordinates were squared.

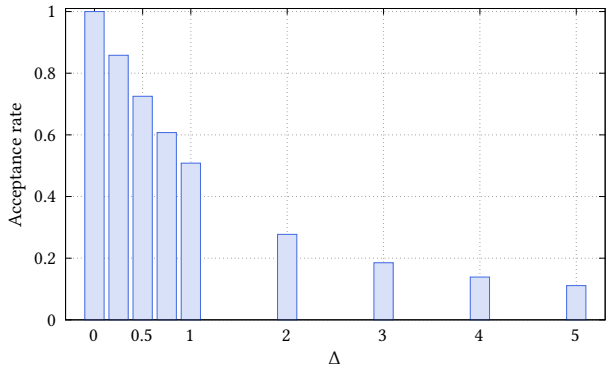


Figure 6: Acceptance rate plotted against the value of Δ. Acceptance decreases for larger Δ values.

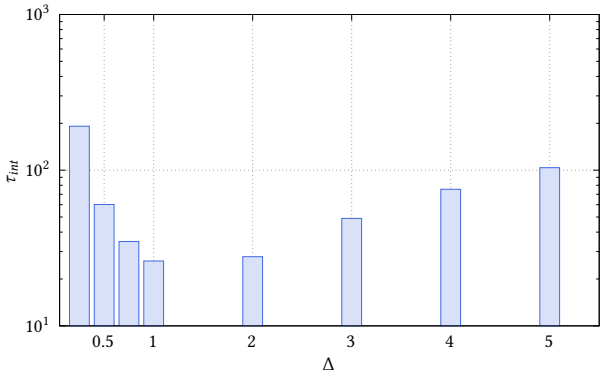


Figure 8: Interaction time in log scale for different values of Δ. Δ = 0 is not considered as every sweep simply copies the previous.

from a computational cost standpoint since it implies less *void* iterations of the algorithm. On the other hand, the purpose of the Metropolis algorithm is to iteratively refine the transition probability of a Markov chain, therefore we would like a somewhat strict selection. Figure 6 gives some insight on the acceptance rate for different values of Δ.

The limit case with Δ = 0, acceptance = 1, is a good example of the concept above: if Δ = 0, every sweep would be an exact copy of the previous so the chain will never reach the desired thermalization value.

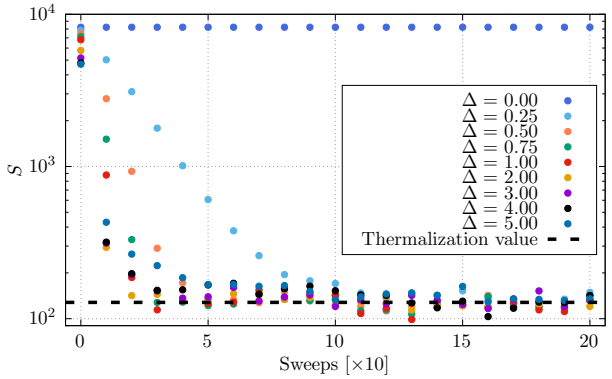


Figure 7: Thermalization speed for different values of Δ. The action is plotted once every 10 sweeps. If Δ ≠ 0, the thermalization value is reached with variable speed: the sweet spot seems to be around $1 \leq \Delta \leq 3$.

Figure 7 shows the values of the action as a function of Markovian time, plotted once every 10 sweeps, for different values of Δ. The thermalization time decreases up to Δ = 3 and then it starts slowly growing again.

Since the composition of each configuration heavily depends on Δ, the parameter also influences the autocorrelation time $\tau_{int}$. This dependence is analyzed in figure 8.

The interaction time decreases from Δ = 0.25 to Δ = 1, where it reaches its minimum, then starts growing again: this is the parameter we decided to prioritize. Autocorrelation time is used to determine the dimension of each bin and, as a consequence, the number of of bins for a fixed number of sweeps. In order to maximize the number of bins, we kept Δ = 1 for the entire simulation.

# 3 Physical results

In this section we report the results of the simulation. We start discussing the lattice results for the correlation functions, along with their error and noise to signal ratio, then we present the energy gap $\Delta\tilde{E}$ and the matrix element; finally, we show the continuum limit extrapolations. The physical parameters $m$ and $\omega$ were both fixed at 1 for every lattice and time $t$ is in units of the spacing $a$.

## 3.1 Lattice results

Keeping a lattice with $N = 256$ and $a = 0.25$ as a case study, we computed the correlation functions $c(t_{lk})$ (equation 50) using $N_{bin} = 10000$ bins of dimension $D_{bin} = 500$.

Figures 9 and 10 show the correlators and their error.

We note that the plots demonstrate the expected features discussed in subsection 1.3: the correlators are symmetric with respect to the central site of the lattice (equation 27) and the error is constant when the distance between the lattice sites is large enough (equation 32). Specifically, the errors cluster around $\sigma[c(t_{lk})] = 0.004962$, which is the value we expect from the square root of equation 32 for a lattice of spacing $a = 0.25$ on which we considered $N_{bin} = 10000$ configurations. Even though this is not enough to draw any definitive conclusions on $|\langle \tilde{E}_0 | \hat{x}^2 | \tilde{E}_0 \rangle|$, as that would take

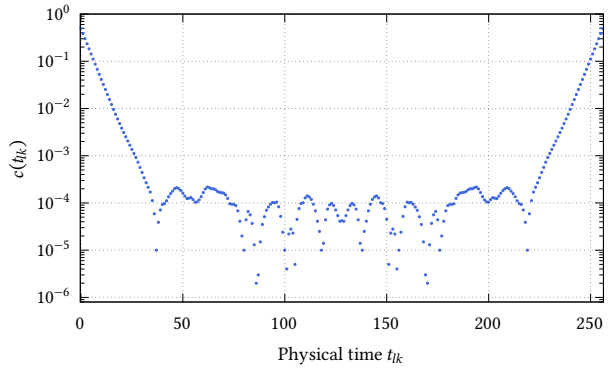


Figure 9: Logarithmic plot of the correlators as a function of physical time. In this plot we highlight their symmetry with respect to $t_{lk} = 128$, more generally they are expected to be symmetric with respect to $N/2$.

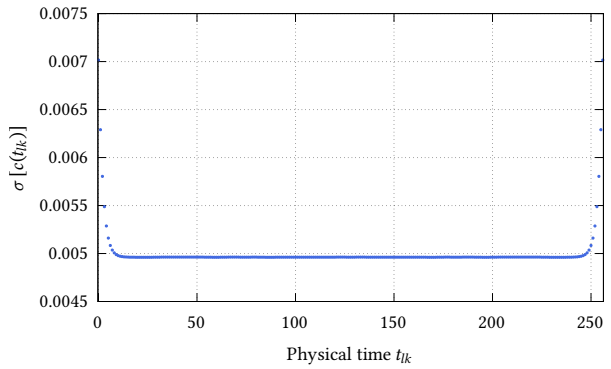


Figure 10: Standard error on the correlators. When $|l-k|$ is large, since $c(t_{lk})$ is exponentially decreasing (equation 27), the error (equation 51) is dominated by $\left\langle x_l^2 x_k^2 \right\rangle$ which is constant (equation 31).

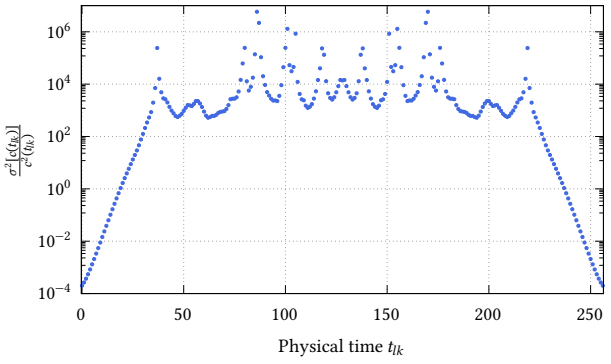


Figure 11: Logarithmic plot of the relative error squared. The exponential increase when $|l-k|$ grows but $N/2 \gg |l-k|$ is a manifestation of the so called exponential problem.

a deeper analysis which is outside the scope of this experiment, we take it as a strong check of our results. The symmetry allows us to discard correlators that depend on $t_{lk} > N/2$ when computing $\Delta\tilde{E}$ and $\left|\left\langle \tilde{E}_0 \middle| \hat{x} \middle| \tilde{E}_1 \right\rangle\right|^2$, as $c(N/2+i)$ is identically equal to $c(N/2-i)$ for any give site $i$.

Another expected feature, showcased by figure 11, is the behaviour of the relative error. This is a manifestation of the exponential problem (equation 34) and will limit the number of useful correlators in the following analysis.

After computing the correlators $c(t_{lk})$, we used them to calculate $\Delta\tilde{E}$ (equation 28) and $\left|\left\langle \tilde{E}_0 \middle| \hat{x} \middle| \tilde{E}_1 \right\rangle\right|^2$ (equation 29). This is when we had to consider two additional conditions: firstly, the observables are decorrelated on Markovian time but correlated on physical time; secondly, because of the exponential problem, as $t_{lk}$ grows so does the error, rendering later time calculations useless.

The physical time correlation is relevant because the energy gap, which is used itself to compute the matrix element, is a function of $c(t_{lk})$, $c(t_{lk}-1)$ and $c(t_{lk}+1)$: we dealt with it applying the Jackknife method to resample the observables and automate the statistical analysis for correlated variables. Regarding the exponential problem, we set a cutoff of 20% to the relative error of each energy gap and matrix element value; only the surviving values were kept and used for the analysis.

In figure 12 we show the 20 surviving values of the energy and their error. Figure 13 does the same for the 15 matrix elements.

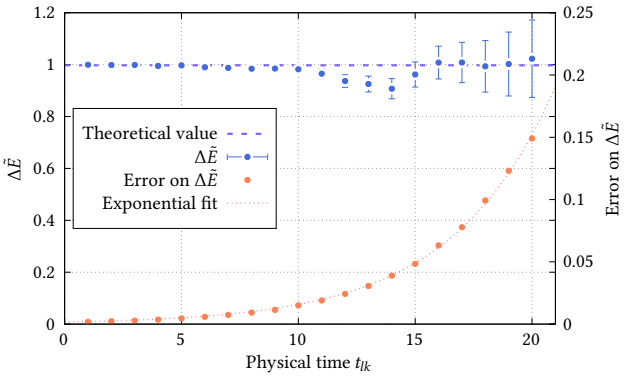


Figure 12: In blue, energy gap from $c(t_{lk})$ for different physical times; in orange, its error. The exponential growth of the error is highlighted.

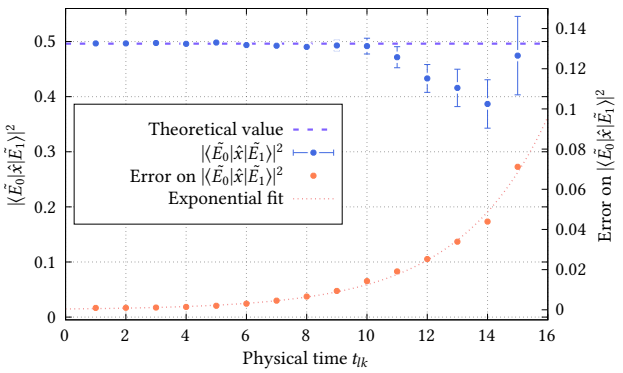


Figure 13: In blue, matrix element from $c(t_{lk})$ for different physical times, the error bars are doubled to improve visibility; in orange, its error. The exponential growth of the error is highlighted.

The best estimates for $\Delta\tilde{E}$ and $\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2$ are the weighted mean of the values that survive the cutoff and are collected in tables 3 and 4. The data shown in the figures correspond to the $A3$ lattice of table 2, which collects the parameters of the five lattices that were used in the simulation.

| Lattice | $N$ | $a$ | $D_{bin}$ | $N_{bin}$ |
|---|---|---|---|---|
| $A1$ | 64 | 1 | 50 | 20000 |
| $A2$ | 128 | 0.5 | 500 | 10000 |
| $A3$ | 256 | 0.25 | 500 | 10000 |
| $A4$ | 512 | 0.125 | 1000 | 2000 |
| $A5$ | 1024 | 0.0625 | 3000 | 2000 |

Table 2: Parameters used for the different lattices.

We kept the physical parameters $m = 1$, $\omega = 1$, $Na = 64$ fixed and repeated this procedure for each lattice of table 2. The more sites a lattice has, the higher the amount of values that survive the 20% relative error cutoff: from lattice $A1$ to $A5$ we were able to work with a number of data in the order of 5, 10, 20, 40, 80. The results are collected in table 3 for the energy gaps and table 4 for the matrix elements; the theoretical values correspond to equation 21 for $\Delta\tilde{E}$ and equation 22 for $\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2$.

### 3.2 Continuum limit

We used the data gathered in tables 3 and 4 to extract the physical properties of the continuum harmonic oscillator. For both the energy gap and the matrix elements, according to equations 23 and 24, we performed a linear fit $y = ma^2 + q$ to find $\Delta\tilde{E} = 0.9998(5)$ and $\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2 = 0.4998(3)$ as the interception with the $y$ axis. The results are in good agreement with the theoretical values of 1 and 0.5.

| Lattice | $\Delta\tilde{E}$ | |
|---|---|---|
| | Theoretical | Obtained |
| $A1$ | 0.9624 | 0.9634(17) |
| $A2$ | 0.9899 | 0.9887(13) |
| $A3$ | 0.9974 | 0.9970(11) |
| $A4$ | 0.99935 | 1.0002(6) |
| $A5$ | 0.9998 | 0.9993(5) |

Table 3: Energy gaps for the 5 lattices.

| Lattice | $\left\|\left\langle\tilde{E}_0\middle\|\hat{x}\middle\|\tilde{E}_1\right\rangle\right\|^2$ | |
|---|---|---|
| | Theoretical | Obtained |
| $A1$ | 0.4472 | 0.4477(7) |
| $A2$ | 0.4851 | 0.4850(5) |
| $A3$ | 0.4961 | 0.4965(5) |
| $A4$ | 0.4990 | 0.4991(3) |
| $A5$ | 0.4998 | 0.4999(3) |

Table 4: Matrix elements for the 5 lattices.

Figures 14 and 16 report the continuum limit for $\Delta\tilde{E}$ and $\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2$ respectively, while figures 15 and 17 show close-ups of the points near the origin.

We also performed a *full fit* those same equations to recover the frequency $\omega$ and the mass $m$ of the oscillator. The results were satisfactory for the fit on $\Delta\tilde{E}$, which had $\omega$ and $A$ as parameters

$$\Delta\tilde{E} = \omega - \frac{\omega^3}{A}a^2, \tag{52}$$

while the results for $\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2$ were less good. The model we used was

$$\left|\left\langle\tilde{E}_0\middle|\hat{x}\middle|\tilde{E}_1\right\rangle\right|^2 = \frac{1}{2m\omega} - \frac{\omega}{16m}a^2 \tag{53}$$

with $\omega$ and $m$ as the fit parameters. Their values differ from the expected ones by 8 and 9 standard deviations, but the correlation matrix suggests a very large negative correlation between the fit parameters, which is not present in the theory

$$\mathrm{Corr}\,(\omega, m) = \begin{pmatrix} 1 & -0.999 \\ -0.999 & 1 \end{pmatrix}. \tag{54}$$

The results are collected in tables 5 and 6.

| Linear fit | | | |
|---|---|---|---|
| | Theoretical | Obtained | $\sigma_s$ |
| $\Delta\tilde{E}$ | 1 | 0.9998(5) | 0.4 |
| Reduced $\chi^2$ | | 1.58 | |
| **Full fit** | | | |
| | Theoretical | Obtained | $\sigma_s$ |
| $\omega$ | 1 | 0.9998(5) | 0.4 |
| $A$ | 24 | 26.9(1.5) | 1.9 |
| Reduced $\chi^2$ | | 1.58 | |

Table 5: Continuum energy gap, frequency and $A$ parameter. All the results are inside a $2\sigma$ neighbourhood of their theoretical value. All the fits were performed by Gnuplot.

| Linear fit | | | |
|---|---|---|---|
| | Theoretical | Obtained | $\sigma_s$ |
| $\left\|\left\langle \tilde{E}_0\middle\|\hat{x}\middle\|\tilde{E}_1\right\rangle\right\|^2$ | 0.5 | 0.4998(3) | 0.67 |
| Reduced $\chi^2$ | | 3.5 | |
| **Full fit** | | | |
| | Theoretical | Obtained | $\sigma_s$ |
| $\omega$ | 1 | 0.92(1) | 8 |
| $m$ | 1 | 1.09(1) | 9 |
| Reduced $\chi^2$ | | 3.5 | |

Table 6: Results for the matrix element. The continuum limit is satisfactory, while the full fit returns incompatible results. The parameters extracted from the full fit also present a very high negative correlation (covariance matrix 54), which could impact the result of the fit. All the fits were performed by Gnuplot.

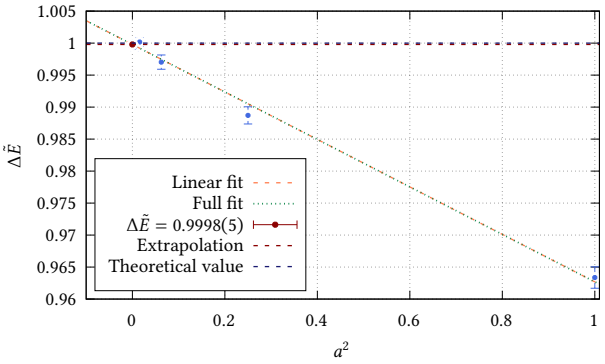


Figure 14: Energy gap continuum limit.

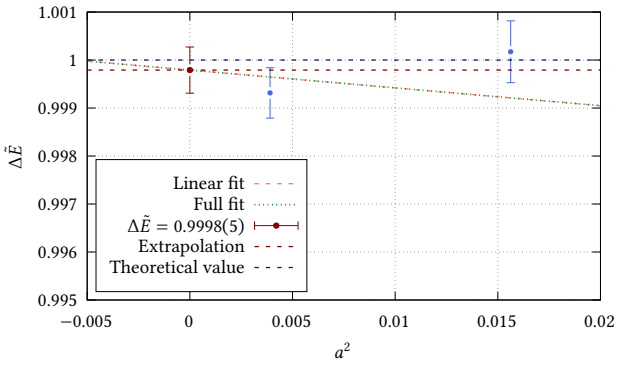


Figure 15: Close-up of the points next to the energy continuum limit.

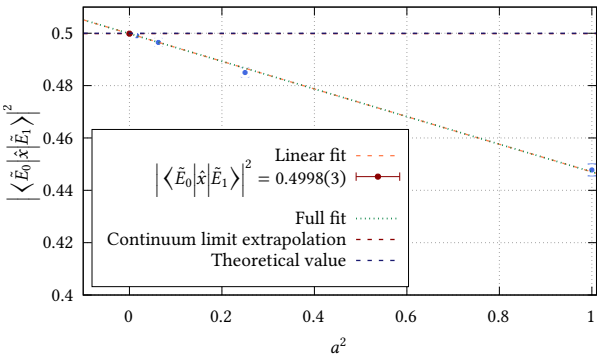


Figure 16: Matrix element continuum limit, error bars are doubled for visibility.

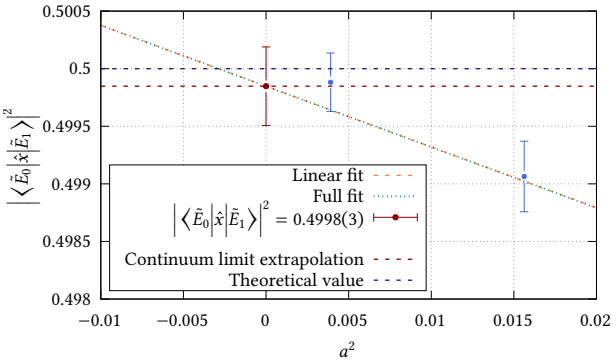


Figure 17: Close-up of the points next to the matrix element continuum limit.

## 4 Conclusions

In our study, we conducted an analysis of the thermalization and autocorrelation properties of the Metropolis Monte Carlo algorithm. We investigated how these properties are influenced by the algorithm parameter $\Delta$.

We found that the thermalization process is faster when starting from a cold Feynman path, which means

initializing the algorithm with a configuration that has lower energy. We kept $\Delta = 1$ as it provided a sweet spot for thermalization speed and autocorrelation time.

To reduce the correlation between consecutive Feynman paths, we averaged the correlation functions over $10,100\tau$ sweeps, where $\tau$ represents the autocorrelation parameter. This averaging suppresses the influence of previous paths on the current path and ensures correlation in Markovian time is negligible. The numerical evaluation of the correlators presented the expected features, up to their variance.

Finally, exploiting the simmetry of the correlators we computed $\Delta\tilde{E}$ and $\left|\left\langle \tilde{E}_0 \middle| \hat{x} \middle| \tilde{E}_1 \right\rangle\right|^2$ for five lattices, paying attention to the exponential problem. The statistical analysis, carried out with the jackknife method, gave us results in good agreement with the theory.

## Bibliography

[CF81] Michael Creutz and Barry A. Freedman. "A Statistical Approach to Quantum Mechanics". In: *Annals of Physics* 132 (1981), pp. 427–462.