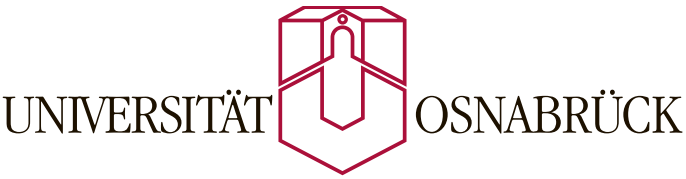

Eigenstate thermalization hypothesis beyond exact diagonalization: Study of random matrix behaviour using natural operator spin bases

A thesis submitted in partial satisfaction of the
requirements for the degree of
Master of Science
in
Physics
in the
School of Mathematics/Computer Science/Physics
of the
University of Osnabrück

Submitted to: Professor Jochen Gemmer, Ph. D. and Professor Robin Steinigeweg, Ph. D.

Submitted by: Fynn Zöllner

Date: 03.07.2023

Contents

1 Introduction

Thermalization in many body systems, is a macroscopic phenomenon which can be described using statistical mechanics. In the framework of classical mechanics this phenomenon and in particular how macroscopic behaviour follows from microscopic behaviour, is well understood. Even though classical mechanics has had a huge success in explaining macroscopic behaviour, the microscopic description has to be done using quantum mechanics. However the microscopic description of quantum many body systems which recovers the established macroscopic behaviour, is still a topic under discussion. The reason why the transition from classical mechanics to quantum mechanics is so difficult in this case is, that the phase space description that is being used in classical mechanics does not translate to quantum mechanics because momentum and position are non commutative. Therefore the central concepts, ergodicity and chaos, which are defined in the phase space, do not directly translate from classical to quantum mechanics.
While the discussion of this topic originates back to Schrödinger [1] and Von Neumann [2], this problem has gotten a significant upsurge of interest over the last decade in theoretical and experimental physics [3–5]. While over the years many approaches were suggested, there was a significant problem in testing hypotheses, since numerical calculations of many body quantum systems become very intensive for increasing system sizes and directly simulating such systems require very high fidelity. The technological improvement has led to such calculations and experiments being practical for relevant system sizes.
Even though thermalization, from a microscopic quantum mechanical standpoint, is not fully understood, there are hypotheses which makes the description possible, at least under a few reasonable assumptions. When it comes to the Hamiltonian H a concept that is widely used is that of quantum chaos. For the observable O however, one important hypothesis is the so called Eigenstate Themalization Hypothesis (ETH). This hypothesis assumes a specific structure of observable matrix elements $O_{ij} = \langle i|O|j\rangle$ in the energy eigenbasis $|i\rangle$ of a generic Hamiltonian H. Specifically, the ETH ansatz reads

$$O_{ij} = \mathcal{O}(\overline{E})\delta_{ij} + \frac{1}{\sqrt{\Omega(\overline{E})}} f(\overline{E}, \omega) r_{ij}, \tag{1}$$

where $\omega = E_i - E_j$ is the transition frequency between eigenstates $|E_i\rangle$ and $|E_j\rangle$ with mean energy $\overline{E} = (E_i + E_j)/2$. $f(\overline{E}, \omega)$ and $\mathcal{O}(\overline{E})$ are smooth functions of there arguments, $\Omega(\overline{E})$ is the density of states and $r_{ij} = r_{ij}^*$ are generally assumed to be Gaussian distributed independent random numbers with zero mean and unit variance.
Assuming this hypothesis to be valid, one can derive the agreement between

long time expectation values of an observable with its predictions from ensemble descriptions, as we will see in more detail in later sections. Also a parallel can be drawn between classical mechanics with dynamical chaos and ergodicity and quantum mechanics with quantum chaos and the ETH.
Even though a direct proof of this ansatz does not exist, its validity was confirmed in a variety of models [6–14]. Specifically, numerical studies showed convincing agreements with the predictions of the ETH regarding the Gaussianity of the r_{ij} [14], the distribution of the transition strengths $|O_{ij}|^2$[15] and the ratio of variances of diagonal to off-diagonal matrix elements[5]. In general the ETH is assumed to be fulfilled for chaotic systems and few-body observables. Important cases in which the hypothesis is known to break down, are integrable models because of the extensive number of constants of motion [16] and strongly disordered models which undergo a transition to a many body localized phase in finite one-dimensional systems [17]. It should be noted that there are also systems which only show weak ETH violation as for example, models that feature so called many body 'quantum scars' [18].
Even though the ETH was confirmed to be valid in a variety of systems, there is still debate regarding the ansatz, specifically about the randomness of off-diagonal elements. In fact it is clear, that for the full observable matrix the off-diagonal elements can not be given by truly independent random numbers. Consider for example a spin chain system and an observable, which is a Pauli operator, acting on a lattice site. The eigenvalues of this observable are clearly defined as ± 1 and therefore the matrix elements of the observable in the energy eigenbasis need to have correlations. However when considering a narrow energy window and therefore a submatrix of the observable, the matrix elements are expected to become independent random numbers for a sufficiently small window.
Also, there are still technical difficulties revolving around the numerics. Since these hypotheses are assumed to be valid in the thermodynamic limit, one would want to study as large systems as possible. However, fully diagonalizing many body quantum systems Hamiltonians becomes exponentially more complex with growing system size, the dimension of a Hamiltonian of a is given by $((2^L)^2)$. Hence, one wants to calculate large systems but is limited through exact diagonalization. Therefore different methods were investigated and put forward, which allow one to calculate these quantities for system sizes that lay outside of the possibilities using exact diagonalization [19].
In this work we will investigate a new approach to the so called relation of moments indicator which probes the independence of off-diagonal matrix elements. This method revolves around calculating moments of observables in a particular operator basis, that has a set of properties, that we can then make use of numerically. In Sec. 2, we will begin to discuss the theoretical foundation of this work, starting with the main concepts of the ETH, quantum chaos

and integrability and then going over to method specific concepts like series of moments and Pauli matrices. Having set the theoretical framework, we will then go over to present our considered system and study it with respect to integrability and ETH using established indicators and methods in Sec. 3. In this section only exact diagonalization will be used. Thereafter, in Sec. 4, our new method will be discussed, including its theoretical concept as well as its application that we used here to calculate the relation of moments indicator. The results we got with this method, can be seen in Sec. 5. We will then continue by discussing the findings in Sec. 6, and presenting possible improvements as well as a number of different potential applications we see using this method, in Sec. 7.
Note that our chosen procedure can be understood as a proof of concept. Central goal is to show, that the method works and that it can be applied to cases of interest, as well as pointing out the methods potentials.

2 Theoretical basics

We will begin by discussing the major theoretical concepts important for the purpose at hand. This comprises central ideas for thermalization in isolated quantum mechanical many body systems, as well as important foundational knowledge relevant for the presented method. Namely the Eigenstate Thermalization Hypothesis as main concept of this work, but also quantum chaos, integrability, series of moments and Pauli matrices. Thereby, note that these concepts are not meant to be discussed in their full complexity, but rather presented in a way to introduce the concept and only go into detail in aspects that are of relevance for this work.

2.1 Eigenstate Thermalization Hypothesis

The Eigenstate Thermalization Hypothesis (ETH) describes a set of ideas which try to explain how an isolated quantum mechanical many-body system in an out of equilibrium initial state, can reach equilibrium after sufficient time. The same behaviour is well understood in the classical domain, where it can be explained through the properties of dynamical chaos and ergodicity. If the ergodic hypothesis holds, the system gets arbitrarily close to every point on the constant energy surface in phase space, after sufficient time. This leads to the fact that such a dynamical system in fact takes on all possible states. More precisely it follows, that time averages become equal to averages over the phase space and therefore ensemble averages.
This behaviour in the domain of quantum mechanics is still not fully understood. The classical explanation is based on the phase space description of the system, which is a concept which does not translate to quantum mechanics, because position and momentum are non commutative. Therefore it is far from trivial how thermalization emerges from quantum mechanics.
To get a more in depth understanding of the problem, we need to oppose the ensemble average to the time average of an isolated system. Consider a quantum many-body system, the microcanonical ensemble is based on the assumption, that every microstate in the energy window has equal probability. Making use of the expectation value of an observable O in the energy eigenbasis, the so called ensemble average $\langle . \rangle_{micro}$, can be written as

$$\langle O \rangle_{micro} = \frac{1}{N} \sum_{E_i \in [E_0 - \frac{\Delta}{2}, E_0 + \frac{\Delta}{2}]} \langle E_i | O | E_i \rangle \,, \tag{2}$$

where N is the number of states in the energy window, E_0 is the centre of the energy window, Δ is the width of the energy window and $|E_i\rangle$ are the energy eigenstates with eigenvalues E_i.
Now consider the same isolated system in an initial state $|\Psi(t_0)\rangle$, where the

expectation value of the observable O in this state, is far from its corresponding ensemble average. Without loss of generality we can assume this initial state to be a superposition of energy eigenstates with a set of coefficients c_i

$$|\Psi(t_0)\rangle = \sum_i c_i |E_i\rangle . \tag{3}$$

From the Schrödinger equation now follows the time evolution as $|\Psi(t)\rangle = U(t,t_0)|\Psi(t_0)\rangle$. The time evolution operator $U(t,t_0)$ is given by $U(t,t_0) = e^{-iH(t-t_0)}$, for a time independent Hamiltonian. The time average $\langle . \rangle_T$ of the observable O in that system, is therefore given by

$$\langle O \rangle_T = \frac{1}{T}\int_0^T d\tau \, \langle \Psi(\tau)|O|\Psi(\tau)\rangle = \frac{1}{T}\int_0^T d\tau \sum_{ij} c_i^* c_j e^{\frac{i}{\hbar}(E_i - E_j)\tau} O_{ij}. \tag{4}$$

The problem now really becomes apparent when comparing Eq. (2) and (4). The ensemble average is only dependent on the energy of the system.
Comparing this with the time average of the observable O starting from an out of equilibrium state, raises a fundamental question. Even if we assume the eigenvalues to be incommensurate, the long time averages is still directly dependent of the coefficients c_i and therefore the details of the initial state

$$\langle O \rangle_T \overset{T\to\infty}{\approx} \sum_i |c_i|^2 O_{ii}. \tag{5}$$

From Eq. (5) follows, that if there is no additional information about the O_{ii} the long-time average is strongly dependent of the choice of coefficients c_i. The long-time average can therefore very strongly depend on the initial conditions, even in the same energy window.
How does it now come, that the system does reach equilibrium after sufficient long time and is accurately described by the microcanonical ensemble without any detailed information of the initial state?
The eigenstate thermalization hypothesis now tries to do this by assuming a set of properties for the matrix elements of observables O in the energy eigenbasis.
Consider a system with the Hamiltonian H. The eigenstates $|E_i\rangle$ and their corresponding eigenvalues E_i satisfy the eigenvalue equation $H|E_i\rangle = E_i|E_i\rangle$. The eigenstates of the Hamiltonian can be ordered with respect to their corresponding eigenvalues. Firstly, the ETH assumes that the expectation values of an observable O in each energy eigenstate only vary an neglectable amount between neighbouring eigenstates. The diagonal of the observable matrix O

can then be written as a smooth function of the energy $\mathcal{O}(\overline{E})$, together with a fluctuation term that is exponentially small in system size. For the off diagonal elements, it is assumed that they are much smaller compared to the diagonal elements and more specifically exponentially small in system size. Furthermore, they are assumed to be Gaussian distributed uncorrelated random numbers.
These conditions can be mathematically implemented by the following equation for the matrix elements

$$O_{ij} = \langle E_i|O|E_j\rangle = \mathcal{O}(\overline{E})\delta_{ij} + \frac{1}{\sqrt{\Omega(\overline{E})}} f(\overline{E},\omega) r_{ij}, \tag{6}$$

where δ_{ij} is the Kronecker-delta, $\overline{E} = \frac{E_i+E_j}{2}$ is the average energy, $\omega = E_i - E_j$ is the transition frequency, $\Omega(\overline{E})$ the density of states and $f(\overline{E},\omega)$ is also a smooth function of its arguments. The $r_{ij} = r_{ji}^*$ are in general assumed to be independent normal distributed random numbers with zero mean and unit variance. This ansatz assumes the observable matrix to take the form of a diagonal ensemble perturbed by a random matrix . This formulation can therefore be seen as an extension of Random Matrix Theory (RMT) applied to observables [20, 21].
This approach leads to the ensemble average being equal to the long time average. Inserting the ETH ansatz into the long time average, Eq. (5), yields

$$\langle O\rangle_{T\to\infty} = \sum_i |c_i|^2 \mathcal{O}(\overline{E}). \tag{7}$$

One assumes the coefficients c_i to only take non-neglectable values in a narrow energy window (E_0, Δ). Since $\mathcal{O}(\overline{E})$ is a smooth function and changes only a neglectable amount between neighbouring energy eigenstates, it stays approximately constant throughout the energy window (E_0, Δ). This constant value $\mathcal{O}$ can then be pulled out of the sum leading to

$$\langle O\rangle_{T\to\infty} \approx \mathcal{O} \sum_i |c_i|^2. \tag{8}$$

Because of the normalization of the initial state $\sum_i |c_i|^2 = 1$, Eq. (8) yields the same value as the ensemble average, where the same approximation can be applied

$$\langle O\rangle_{T\to\infty} \approx \mathcal{O} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{O} \approx \langle O\rangle_{micro}\,. \tag{9}$$

Since the ETH leads to the agreement of time averages with ensemble averages it can be interpreted as the quantum mechanical manifestation of ergodicity, since in classical mechanics, ergodicity leads to the same property [22].

In order to understand the choice for the off diagonal elements in Eq. (6) in more detail, we consider the autocorrelation function of an observable O in the Heisenberg picture

$$Tr\{O(t)O\} = \sum_i \langle E_i|O(t)O|E_i\rangle \tag{10}$$

where $|E_i\rangle$ correspond to the energy eigenstates of the respected system. Inserting a basis projection $1 = \sum_j |E_j\rangle \langle E_j|$ and the time evolution operator $e^{-\frac{i}{\hbar}Ht}$ for a time independent Hamiltonian H, yields

$$\begin{aligned} Tr\{O(t)O\} &= \sum_{ij} \langle E_i|e^{\frac{i}{\hbar}Ht}Oe^{-\frac{i}{\hbar}Ht}|E_j\rangle \langle E_j|O|E_i\rangle \\ &= \sum_{ij} e^{\frac{i}{\hbar}(E_i-E_j)t}|O_{ij}|^2. \end{aligned} \tag{11}$$

Now separating the sum in diagonal and off-diagonal elements ans substituting the ETH ansatz Eq. (6) yields

$$Tr\{O(t)O\} = \sum_i |O(i,i)|^2 + \sum_{ij;i\neq j} \frac{1}{\Omega(i,j)} f^2(i,j) r_{ij}^2 e^{\frac{i}{\hbar}(E_i-E_j)t}. \tag{12}$$

The diagonal part can now be approximated as a constant C. Substitute the arguments i and j for the average energy $\overline{E} = \frac{E_i+E_j}{2}$ and the transition frequency $\omega = E_i - E_j$. The square of the random real numbers r_{ij} will average to one, because of the law of large numbers. Assume now the system to be sufficiently large, so that the energy level spacing can be approximated to be continuous. The sums can then be replaced with integrations over $\bar{E}$ and ω,

$$Tr\{O(t)O\} = \int_{\overline{E},\omega} d\overline{E}d\omega f^2(\overline{E},\omega)e^{\frac{i}{\hbar}\omega t} + C(\mathcal{O}(\overline{E})), \tag{13}$$

where the volume element is the density of states $\Omega(\overline{E})$. For simplicity reasons the case is considered in which the density of states is constant. In this case the volume element cancels out with $1/\Omega(\overline{E})$ from Eq. (12). The integration corresponds to a Fourier transformation from the frequency domain into the time domain. Therefore, the function $f(\bar{E},\omega)$ describes the Fourier transform

of a function $\tilde{f}(t)$, which determines the time evolution of the operator O, and therefore its equilibration dynamics.
From this follows, that $f(\bar{E}, \omega)$ inherits the conditions bound to a time evolution and equilibration dynamics, like for example time reversal invariance.

When looking at the actual nature of the physical systems this approach is applied on, it becomes apparent that the observable can not be truly given by uncorrelated random numbers. Consider a generic system of spins. The Hamiltonian of the system, as well as an observable in general, is given by Pauli matrices which have to have correlated matrix elements, in order to yield the eigenvalues ± 1. It also has been argued, that correlation of matrix elements is necessary to explain properties of the so called out of time ordered correlation function (OTOC), which is connected to the property of scrambling in quantum many body systems [23]. The question therefore remains, to what extent the matrix elements can be taken as uncorrelated random numbers.
As described above, the matrix elements of the full matrix of an observable in the energy eigenbasis, can not be truly uncorrelated. One could now ask what would happen if one considers a submatrix of the Hamiltonian. This Hamiltonian submatrix can described by a energy window ΔE which is related to an inverse time scaling $1/T$. In this picture, a small energy window then corresponds to a long time scale and the full matrix corresponds to a very short time scale.
Now, at short times thermodynamic behaviour is not expected anyway. The thermodynamic behaviour at long times and therefore ETH, or rather RMT behaviour, is expected to emerge in small energy windows. The energy scale, at which the matrix elements can be taken as independent random numbers, is denoted as ΔE_{RMT}. Determining this limit under which the observable can be described by a GOE is still under discussion and different indicators were suggested.

2.2 Quantum chaos and random matrices

Even though the ETH is an hypothesis, it has been confirmed in a variety of models. It is also believed, that the ETH is valid in chaotic systems, similar to ergodicity being implied from chaos in classical mechanics.
In order to get a better understanding of the concept of chaos, its classical notion is briefly discussed. As described above, in classical mechanics, the dynamics are often described through the phase space description. In that, chaos is a very closely related concept to ergodicity. Chaos describes the property of a system, that an arbitrarily small perturbation in initial conditions will lead to an exponential deviation of orbit in phase space.
Even though most systems are chaotic and ergodic at the same time, ergodicity and chaos are not equivalent. For example, the one dimensional harmonic

oscillator is ergodic but not chaotic, since close trajectories do not diverge from each other. It should also be noted, that chaos is the stronger property which means, that chaos implies ergodicity but not the other way around, as we have just seen in the example of the harmonic oscillator.
As seen above, in classical mechanics chaos is a statement about how trajectories behave in the phase space description of the system. When transitioning to quantum mechanics, this leads to a problem. The notion of phase space does not make sense in quantum mechanics, as position and momentum do not commute. The correspondence principal says, that if the ratio of the Planck constant to the action of the system tends to zero, the predictions of quantum dynamics need to agree with classical predictions. The concept of chaos in classical dynamics therefore has to emerge from quantum dynamics in some way.
Therefore, a convenient way to investigate how chaos manifests in the quantum domain, is by studying systems, that are expected to be very chaotic. Such a system would be a strongly interacting, isolated many body system. A very good example of such a system, that is also historically one of the first systems that was studied in this context, is a heavy nucleus.
When the energy levels of heavy nucleus were first examined, QCD where yet unknown. The energy levels of the nucleus where found to not follow a simple formula, unlike the electrons in the shell of the atom. In 1955, Wigner found, that the distribution of energy level spacing of large nuclei, are sufficiently well described by the formula

$$P(s) = \frac{\pi}{2} s e^{-\pi \frac{s^2}{4}}. \tag{14}$$

This distribution corresponds to the level spacing distribution of the eigenvalues of random matrices drawn from the so called Gaussian Orthogonal Ensemble. This implies energy level repulsion and therefore the absence of energy degeneracies. Even though QCD, and therefore the explicit interactions within the system, were unknown at the time, the description of heavy nuclei with random matrices fit the data very well.
Over the years other systems of this kind where investigated and numerically calculated. These experiments and calculations showed a strong agreement with the preceded findings of similarities between systems that are expected to be chaotic and random matrices. These findings led to the Bohigas-Giannoni-Schmit conjecture: Spectra of time reversal-invariant systems whose classical analogues are chaotic show the same fluctuation properties as predicted by GOE.

In mathematics, a random matrix is a matrix-valued analogue of a random variable. The most commonly investigated distributions in this context are

the Gaussian ensembles. The probability measure

$$\mu = C_{\mathrm{GUE}}(n) e^{-\frac{n}{2} tr\{H^2\}} \tag{15}$$

over the space of $n \times n$ hermitian matrices $H_{ij} = H^*_{ji}$ describes the so called Gaussian unitary ensemble (GUE). Where both the real and imaginary part of the matrix elements H_{ij} are uncorrelated random numbers of zero mean and unit variance and $C_{\mathrm{GUE}}(n)$ is a normalization constant, so that the integral over the full probability space equals 1. In physics, this ensemble is generally associated with Hamiltonians that lack time reversal invariance.
The probability measure

$$\mu = C_{\mathrm{GOE}}(n) e^{-\frac{n}{4} tr\{H^2\}} \tag{16}$$

over the same space of hermitian matrices on the other hand, describes the so called Gaussian orthogonal ensemble (GOE). The matrix elements H_{ij} here, are uncorrelated random numbers of zero mean and a variance of 2 and again $C_{\mathrm{GOE}}(n)$ insures its normalization property needed for it to be a probability measure. In physics, this ensemble is associated with Hamiltonians of systems implementing time reversal invariance.

2.3 Symmetry and integrability

Integrability is a property of dynamical systems. Informally speaking, a integrable system is a system that has a sufficient amount of symmetries and therefore conserved quantities, so that the degrees of freedom are much lower than the dimension of the system. In classical mechanics this leads to the fact, that the dynamics of a system is restricted to a submanifold of its full phase space. Since in quantum mechanics, the concept of a phase space like in classical mechanics, does not make sense, due to the fact that space and momentum do not commute, the notion of integrability differs in quantum mechanics. In this section the concept of integrability will be discussed in the context of spin chains. Since integrability and symmetry are closely related concepts, we will start from a discussion about symmetry in the context at hand.
Consider a 1-dimensional Heisenberg chain with length L. The Hamiltonian H is given by

$$H = J \sum_r \vec{s}_r \cdot \vec{s}_{r+1}, \tag{17}$$

with J being the interaction strength and

$$\vec{s_r} = \begin{bmatrix} s_r^x \\ s_r^y \\ s_r^z \end{bmatrix}, \tag{18}$$

being the spin operator acting on lattice site r. With the ladder operators $s^+ = s^x + is^y$ and $s^- = s^x - is^y$ the Hamiltonian can be written as

$$H = J \sum \frac{1}{2}(s_r^- s_{r+1}^+ + H.c.) + s_r^z s_{r+1}^z. \tag{19}$$

The spin operators s^x and s^y can therefore be written as a spin flip of two neighbouring spins with respect to the $z-$direction. Consider the full $d = 2^L$ corresponding spin states

$$|\uparrow\downarrow\downarrow \ldots\rangle, |\downarrow\uparrow\downarrow \ldots\rangle, \ldots \tag{20}$$

As it can be easily seen the spin states are not eigenstates of the Hamiltonian. Therefore in order to solve the eigenvalue problem of the Hamiltonian, the naive approach would be to calculate the full matrix of H in the spin basis and then diagonalize it, in order to obtain energy eigenvalues E_i and eigenstates $|E_i\rangle$. This approach would take into account every transition from two arbitrary spin states. Solving this problem for a system of non negligible size would be of such a high complexity, that it could only be accessible through numerical methods.
In fact this Hamiltonian has a variety of symmetries that can be used, to massively reduce the amount of transitions that need to be considered, simplifying the problem significantly. Consider the subset of spin states that are fully polarised downwards, except for one spin. As stated above these states on their own are not eigenstates of the Hamiltonian. However given the translation operator T which shifts the up spin by one position

$$T\,|\uparrow\downarrow\downarrow \ldots \downarrow\rangle = |\downarrow\uparrow\downarrow \ldots \downarrow\rangle, \tag{21}$$

it can be seen, that applying the translation operator to one of those states, does not change the action of the Hamiltonian on that state

$$[H, T] = 0. \tag{22}$$

The system is invariant under translation. This symmetry induces a conserved quantity. From this follows, that a state with a specific value for this

conserved quantity can only transition to states with the same value. The full problem therefore can be reduced to independent subspaces, where every subspace contains all states with the value for the respected conserved quantity. The problems in each subspace is independent from all other subspaces, since states are bound to that subspace. The full Hamiltonian can be simplified to 'blocks', where each 'block' can be solved independently.

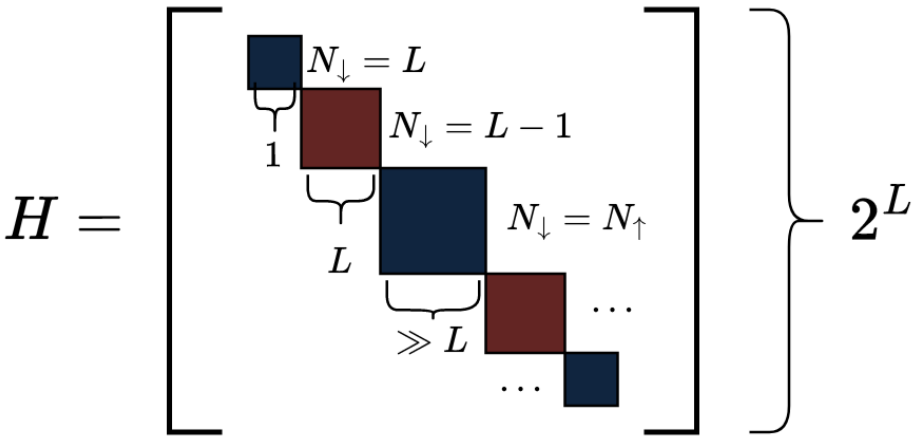

Figure 1: Visual illustration of the independent blocks of matrix elements resulting from additional symmetries in a system. This example is for a spin chain with translational symmetry. The symmetry induces a constant of motion, which for this case is the magnetization which can be written as the total number of down spins $N_\downarrow$. Every value of this constant of motion has a correspondent subspace which is independent of all other subspaces. Note that this is just a sketch and is not meant to depict the matrix structure exactly.[24]

In every subspace common eigenstates of the Hamiltonian and the corresponding symmetry operator, can be found. Now since all the subsystems are uncorrelated, it is probable for the energy eigenvalues to be degenerate and in fact this property can be used to quantify a systems integrability off of its spectrum. In contrast to chaotic systems, as it is discussed in Sec. 2.2, integrable systems are expected to have energy degeneracies. Considering the energy level spacing $s = |E_i - E_j|$ this behaviour would show in the form of a Poisson distribution

$$P(s) = e^{-s} \tag{23}$$

even though there are exceptions[25].
Every block of equal total spin can be solved through the so called Bethe ansatz. Even though it was initially invented to find the exact eigenvalues and

eigenstate of the one-dimensional anti-ferromagnetic Heisenberg model Hamiltonian [26], the method has been extended to a variety of other one dimensional models like the XXZ-model or the Hubbard-model.
Even though the above considerations have parallels to the classical formulation of integrability, the reduction of the complexity of the problem through making use of symmetries, the definition around constants of motions does not translate to quantum mechanics. The reason for that, is that in quantum mechanics every quantity that can be expressed through energy eigenstates is a constant of motion, which are in general infinitely many.
Also, integrable systems have other symmetries which can often not even be explicitly written out, in contrast to the global case here. Furthermore, there a systems which have global symmetries of for example translation, and the symmetry subspace of the global symmetry is actually non-integrable. Therefore in the quantum mechanical case, the notion of conservation laws has to be specialized to local conservation laws. In fact global symmetries do not have a direct connection to integrability and chaos, but considering their implications clarify how symmetries in general manifest in such systems.
Since the definition of integrability in quantum mechanics is much less clear then in classical mechanics because of the aforementioned reasons, it is often just defined by its outcome. Namely, the possibility to exactly solve the system through Bethe ansatz.

2.4 Sequences of moments and relation of moments indicator

In order to describe the probability distributions of random variables, a concept that is often used, is that of moments. The moments of a probability distribution describe different properties of the distribution. The k-th moment M_k of a random variable X is defined as the expectation value $\langle . \rangle$ of the random variables k-th power

$$M_k := \langle X^k \rangle \,. \tag{24}$$

The first moment of a random variable corresponds to its expectation value. The centralized moment is now defined by

$$M_k^c := \langle (X - M_0)^k \rangle \tag{25}$$

so the calculation is centred around its expectation value. The second centralized moment is then described as the variance, the third moment as skewness and the fourth moment as kurtosis.
One can now ask if, given a sequence of moments M_k, the reversed mapping to a probability measure exists and if so, if it is unique. There are three dif-

ferent so called moment problems, which differ by the considered support of the probability measure μ. Whereas the Hamburg moment problem and the Stieltjes moment problem consider open unbound intervals , $(-\infty, \infty)$ and $[0, \infty)$, the Hausdorf moment problem considers a closed interval as support of the probability measure. In contrast to the Stieltjes and Hamburg moment problems, the Hausdorf moment problem has a unique solution if it exists.
Considering the probability measure μ to be only different from zero in a closed interval. Without loss of generality this interval can be taken to be $[0, 1]$. Given a series of moments M_k, Hausdorff showed, that there exists a unique solution to the moment problem if and only if the series is completely monotonic [27]. Therefore, M_k needs to fulfil the equation

$$(-1)^n(\Delta^n M)_k \geq 0, \tag{26}$$

with the difference operator

$$(\Delta M)_k = M_{k+1} - M_k. \tag{27}$$

Consider now a random matrix H. The sequence of moments M_k of a random matrix is given by

$$M_k = \begin{cases} C_{\frac{k}{2}}(M_2)^{\frac{k}{2}}, & \text{for even } k. \\ 0, & \text{for odd } k \end{cases} \tag{28}$$

where C_n are the Catalan numbers which are defined by

$$C_n = \frac{1}{n+1}\binom{2n}{n} \quad , n \geq 0. \tag{29}$$

As it can be seen in Sec. 2.2, the probability measure of random matrices are given by Gaussian distributions. Since the Gaussian distribution is heavily focused around its expectation value, the interval that it is not mapped to zero by this measure, can be assumed to be bounded and closed.
The sequence of moments, Eq. (28), also implements the condition of complete monotony in the Hausdorff moment problem, as it can be easily seen. Therefore, given a sequence of moments N_k of a matrix O, if it can be shown, that this sequence of moments is equal to Eq. (28), then the matrix O is a random matrix with the same probability measure, following from the uniqueness.

Consider the ETH ansatz Eq. (6), the off-diagonal elements are mainly given

through independent random numbers with zero mean and unit variance r_{ij}. As described above, the sequence of moments of this ensemble is given by Eq. (28) and its reversed mapping from the series of moments to the probability measure exists and is unique.
Given an arbitrary observable in the EEB, the off-diagonal part of the ETH ansatz can therefore be checked, by the agreement with the sequence of moments of this observable with Eq. (28). In first approximation this "level of agreement" can be calculated through the relation of the first two non zero moments. As it can be easily seen this corresponds to the quotient of the second and fourth moment

$$\Lambda = \frac{M_2^2}{M_4}. \tag{30}$$

The second moment in Eq. (30) is squared since we need to consider the same power in the observable. Inserting Eq. (28) into Eq. (30) yields

$$\Lambda = \frac{1}{2}. \tag{31}$$

as prediction for the indicator for GOE random matrices.
Furthermore, the value of 1/2 for the relation of moments indicator, can be directly concluded form the ETH ansatz [28]. Consider the 4th moment of an observable O in the energy eigenbasis

$$M_4 = \frac{1}{d} \sum_{m,n,k,l} O_{mn} O_{nk} O_{kl} O_{lm} \quad \text{with } m,n,k,l : E_m \in [E_0 - \Delta/2, E_0 + \Delta/2] \tag{32}$$

Assuming the system to fulfill the ETH, the off diagonal elements of the observable is given through independent random numbers with zero mean. In general summing over a large number of such random numbers would yield zero, but there are special index combinations that lead to non vanishing terms. These are the index combination where the term can be written as squares. Namely, these are the terms where $k = m$ and $l = n$

$$M_4 = \frac{1}{d} \sum_{m,n,l} |O_{mn}|^2 |O_{lm}|^2 + \frac{1}{d} \sum_{m,n,k} |O_{mn}|^2 |O_{kn}|^2 \tag{33}$$

This equation can be written as

$$M_4 = \frac{2}{d}\sum_m (\sum_n |O_{mn}|^2)^2 = \frac{2}{d}\sum_m (\langle m|O^2|m\rangle)^2 \tag{34}$$

Making use of the diagonal part of the ETH ansatz with respect to O^2, namely that the diagonal is approximately constant over a sufficiently small energy window, yields

$$\begin{aligned} M_4 \approx & \frac{2}{d}\sum_m (\frac{1}{d}\sum_{m'} \langle m'|O^2|m'\rangle)^2 \\ = & \frac{2}{d}\sum_m (M_2)^2 \\ = & 2M_2^2 \end{aligned} \tag{35}$$

Inserting this into Eq. (30), leads to the same prediction as derived for random matrices

$$\Lambda \approx \frac{1}{2}. \tag{36}$$

Here it also should be noted, that from the calculation above it can be seen easily, that the odd part of the series of moments of the ETH ansatz agrees with that of random matrices. The calculation is based on the fact, that all terms vanish because of the random numbers cancelling each other out, except for those cases where all matrix elements can be grouped to squares. This is only possible because there are an even amount of matrix elements in the sum. If there is an odd amount of matrix elements all terms cancel each other out, which is the case for all odd moments. Therefore, the ETH ansatz agrees with the odd case of the sequence of moments of random matrices, Eq. (28).

2.5 Pauli matrices

The Pauli matrices are a set of matrices which are used to represent the spin operator with its commutation rules. The spin component operators can then be written as a real number times the corresponding pauli matrix

$$S^i = \frac{\hbar}{2}\sigma^i \tag{37}$$

for a spin 1/2 particle. With that the spin operator $\vec{S}$ can be written as

$$\vec{S} = \frac{\hbar}{2} \begin{pmatrix} \sigma^x \\ \sigma^y \\ \sigma^z \end{pmatrix} \tag{38}$$

The Pauli operators, including the identity matrix, are defined by

$$1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma^y = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \quad \sigma^z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{39}$$

With this definition and the matrix multiplication the Pauli matrices follow the following algebraic relation

$$\sigma^i \cdot \sigma^j = \delta_{ij} + i \sum_{k=1}^{3} \epsilon_{ijk} \sigma_k \tag{40}$$

with ϵ_{ijk} being the Levi-Cevita-Symbol. From this immediately follows

$$\sigma^x \sigma^y \sigma^z = i. \tag{41}$$

Furthermore, the Pauli matrices are traceless, except for the identity matrix, and fulfill the spin commutation rules.
Consider a term consisting of multiple products of Pauli matrices. If all Pauli matrices are only of two kinds, x and z for example, the needed properties reduce to

$$(\sigma^x)^2 = (\sigma^z)^2 = (1)^2 = 1, \quad \sigma^i \sigma^j = -\sigma^j \sigma^i \tag{42}$$

since the only way, that Eq. (41) can be used to reduce Pauli matrices to the identity, is by using the property $\sigma^z \sigma^x = i\sigma_y$ and another pair of σ^x and σ^z. If we have two pairs of σ^z and σ^x the same result can be accomplished under the use of Eq. (42). Since every transformation including the use of Eq. (41) can be also done using only Eq. (42), this subset of equations is sufficient in describing the algebraic relations. This property will become helpful later in this work.

3 System and classical methods

In this section we will study an isolated quantum many body system with respect to thermalization using established methods. Here the considered system as well as the calculated quantities are presented, and the obtained results are discussed. In this section we will make use of established indicators and use exact diagonalization.

3.1 Ising model

The system of interest in this work is a 1-dimensional Ising chain in a magnetic field with $x-$ and $z-$components with periodic boundary conditions. The Hamiltonian of the system is given by the sum over the local Hamiltonians

$$H = \sum_{l=1}^{L} H_l \tag{43}$$

where L is the chain length. The local Hamiltonians are given by the interaction of each spin with the magnetic fields and the nearest-neighbour interaction

$$H_l = J\sigma_l^z \sigma_{l+1}^z + h_x \sigma_x + h_z \sigma_z \tag{44}$$

where J is the interaction strength and h_x and h_z are magnetic field strengths. σ_l^z is the $z-$Pauli operator at the lattice site l.
Even though most calculations will be performed on the clean model, two defect terms $h_2\sigma_z^2$ and $h_5\sigma_z^5$ can be added in order to lift translational and parity symmetries. This model is nonintegrable [28], and exhibits diffusive energy transport [28].
Even though the Ising model is a relatively simple model, it has broad use cases [29], where in particular it serves as a mathematical model for ferro magnetism in solid state crystals.

3.2 Observable

The local observable investigated here is the measurement of one spin in z-direction

$$O = \sigma_q^z \tag{45}$$

where q is a natural number from the interval $[0, L]$ which describes the lattice site on which the spin is measured.

Here we consider the operator projected into a narrow energy window with respect to the system Hamiltonian. Therefore we define O^T as the projected operator. Let $|n\rangle$ be the energy eigenstates with the eigenenergies E_n, so that they fulfill the following equation $H|n\rangle = E_n|n\rangle$. The matrix elements of the projected operator, in this energy eigenbasis, are defined by

$$O^T_{mn} = \langle m|P_T O P_T|n\rangle \tag{46}$$

where P_T is the projection operator onto a submatrix of the Hamiltonian

$$P_T = \sum_{|E_n - E_0| \leq \Delta/2} |n\rangle \langle n| \tag{47}$$

the dimensions of the submatrix is given by the two parameters Δ and E_0. E_0 is the centre of the respected energy window and ΔE is the width of the window. For the whole work we consider the energy center to be $E_0 = 0$. The projection operator takes over all matrix elements, that lay inside the defined window and neglects all matrix elements that lay outside of it.

3.3 Calculated quantities

In order to gain information about the system, specific quantities need to be calculated, from that information about different aspects of the system can be derived. In this section the most important quantities for the purpose at hand will be discussed.

First of all the system should be investigated with respect to integrability and chaos. These properties express themselves through the distribution of energy level spacings. The straight forward approach would now be, to calculate this distribution and assign it to a Wigner-Dyson or Poisson distribution.

However, calculating the energy level spacing distribution requires some preliminary work especially with Hamiltonians that have additional symmetries like translation or parity, which would be the case for the Hamiltonian considered here without defect terms.

First of all, if the system exhibits additional global symmetries, a shared symmetry subspace needs to be considered. For translational symmetry, this is done through first defining the translation operator T, which shifts the position of every spin in a spin state by one, or more precise

$$T|\uparrow\downarrow\downarrow\ldots\downarrow\rangle = |\downarrow\uparrow\downarrow\ldots\downarrow\rangle\,. \tag{48}$$

Evaluating the matrix of T and diagonalizing it will result in a set of eigenstates

$|\Psi_j^k\rangle$ with eigenvalues $e^{-i\frac{2\pi}{L}k}$. These eigenstates will have degeneracy, namely there are a number of eigenstates with the same eigenvalue, these eigenstates span the symmetry subspace of the respected eigenvalue. The Hamiltonian, and therefore the system, can now be evaluated in one of these subspaces

$$(H^{k_0})_{ij} = \langle \Psi_i^{k_0} | H | \Psi_j^{k_0} \rangle \,, \tag{49}$$

where k_0 is the chosen parameter, that implies the eigenvalue of the considered subspace. Important in the choice of symmetry subspaces is, that the considered symmetry subspace does not have other global symmetries. Also the different subspaces have different dimensions. Therefore mostly the subspace with the highest dimension is chosen.
Furthermore, for studying the spectrum directly a procedure is necessary, that is called unfolding of the spectrum. For the simplest case, of no symmetries except for the energy, the unfolding process rescales the local average level space to 1. This can be done by rescaling the energy spectrum by

$$\tilde{E_m} = E_m \rho(\overline{E_m}), \tag{50}$$

where E_m are the energy eigenvalues and $\rho(\overline{E_m})$ is the density of states in a narrow energy window surrounding E_m [30].
To avoid this unfolding process a different method is used here. This method uses the ratio of adjacent level spacing r_m to make conclusions about the spectrum. The ratio of adjacent level spacing is defined as

$$r_m = \frac{min(s_m, s_{m-1})}{max(s_m, s_{m-1})}, \tag{51}$$

where $s_m = E_{m+1} - E_m$ are the adjacent level spacings. The mean value of this quantity $\langle r_m \rangle$ now differs for a Poisson distribution and a Wigner-Dyson distribution of level spacings

$$\langle r_m \rangle \approx \begin{cases} 0.39 & \text{Poisson} \\ 0.53 & \text{Wigner-Dyson.} \end{cases} \tag{52}$$

Through calculating $\langle r_m \rangle$ for the considered Hamiltonian, it can be decided if the system is integrable or non-integrable [31].

In order to check if the ETH is fulfilled, a variety of indicators were suggested . Generally there are two kinds of ETH indicators, the ones that check the

diagonal part of the ETH ansatz and those which check the off diagonal part. For the diagonal part of the ETH, the indicator chosen here revolves around the variance of the diagonal matrix elements. The ETH states, that the diagonal matrix elements in the energy eigenbasis O_{nn} become a smooth function in the thermodynamic limit $L \to \infty$. The variance of the diagonal elements should therefore, decrease rapidly with system size. The variance $\sigma_d^2(E_0)$ of the diagonal elements of an observable O in an energy window $[E_0, \Delta]$ is defined as

$$\sigma_d^2(E_0) = \frac{1}{N_{E_0}} \sum_m [O_{mm}]^2 - \left(\frac{1}{N_{E_0}} \sum_m O_{mm} \right)^2, \tag{53}$$

where the sums run over all eigenstates $|E_n\rangle$ with corresponding energy eigenvalues E_n for which $E_n \in [E_0 - \Delta/2, E_0 + \Delta/2]$ holds. N_{E_0} is the number of states in this window. For non-integrable systems $\sigma_d^2(E_0)$ is expected to fall off exponentially with L, while for integrable systems a polynomial dependence is expected [5].
For the off diagonal part of the ETH ansatz, multiple aspects can be investigated, like the behaviour and shape of the function $f(\bar{E}, \omega)$ or the distribution of the r_{ij}. Although different indicators will be briefly mentioned and discussed, main emphasis here lies on the relation of moments indicator, which probes the off diagonal elements with regard to their correlations. For this indicator, the relation between the second moment squared to the fourth moment

$$\Lambda^T = \frac{M_2^2}{M_4}, \quad M_k = \frac{Tr[(O^T)^k]}{d} \tag{54}$$

is calculated, where M_k is the k-th central moment and d is the dimension of the system.
The ETH assumes the off-diagonal matrix to be a random matrix. The sequence of moments of a random matrix is clearly defined. The relation of the second moment squared to the fourth moment has the constant value

$$\Lambda^T = \frac{1}{2}$$

If the relation of moments for our system is approximately $\frac{1}{2}$, the off-diagonal matrix of the projected operator shows random matrix behaviour and fulfills the ETH in this regard.
In order to probe the function $f(\bar{E}, \omega)$, the matrix elements O_{mn} can be considered in an sufficiently small energy window, which is centred around $\bar{E}$, in

that the density of states $\Omega(\bar{E})$ is approximately constant. This allows, on the one hand, to analyse the ω dependence of $f(\bar{E}, \omega)$ and on the other hand the distribution of the O_{mn}. Consider the average over $|O_{mn}|^2$ in a small $\tilde{\omega}$ interval of width $\Delta\omega \ll \tilde{\omega}$

$$\overline{|O_{mn}|^2}(\tilde{\omega}) = \frac{1}{N_{\tilde{\omega}}} \sum_{nm} |O_{mn}|^2 \tag{55}$$

where the sum runs over all matrix elements where the transition frequency $\omega = E_m - E_n$ lays inside the $\tilde{\omega}$ interval, $\omega \in [\tilde{\omega} - \Delta\omega/2, \tilde{\omega} + \Delta\omega/2]$. $N_{\tilde{\omega}}$ is the number of matrix elements for which the above condition holds. Inserting Eq. (1) into Eq. (55) shows, that $\overline{|O_{mn}|^2}(\tilde{\omega})$ yields $f^2(\bar{E}, \omega)$ except for a prefactor [14].
Furthermore, if $\overline{O_{mn}} = 0$, the quantity

$$\Gamma(\tilde{\omega}) = \frac{\overline{|O_{mn}|^2}}{\overline{|O_{mn}|}^2}, \tag{56}$$

with the averages being defined analogous to Eq. (55), allows to investigate the distribution of the r_{mn}. If the distribution of the O_{mn} matches a Gaussian distribution, $\Gamma(\tilde{\omega}) = \pi/2$ [14].

The autocorrelation function $C(t)$ describes the correlation of an observable O with itself, developed in time. It is defined as

$$C(t) = \langle O(t)O \rangle = \frac{1}{d} \sum_{n: E_n \in [E_0 - \frac{\Delta}{2}, E_0 + \frac{\Delta}{2}]} \langle n|O(t)O|n\rangle \,. \tag{57}$$

For a time independent Hamiltonian, the time evolution operator $U(t)$ is given by

$$U(t) = e^{-\frac{i}{\hbar}Ht}. \tag{58}$$

Inserting the time evolution operator in the definition of the autocorrelation function and inserting $\sum_m |m\rangle \langle m| = 1$ yields

$$C(t) = \frac{1}{d} \sum_{n,m: E_n, E_m \in [E_0 - \frac{\Delta}{2}, E_0 + \frac{\Delta}{2}]} e^{\frac{i}{\hbar}\omega t} |O_{nm}|^2 \tag{59}$$

where $\omega = (E_n - E_m)$ is the transition frequency and $O_{nm} = \langle n|O|m\rangle$ are the matrix elements of the observable in the energy eigenbasis.

The autocorrelation function $C(t)$ now not only shows the time evolution of an observable O for a centred observable $O_c = O - \langle O \rangle$ but also shows specifically, how much the value of the observable differs from its ensemble average, for different times t.
We define the thermalization time τ_{th} as the time where the quantity $\tilde{C(t)} = [C(t) - C(t \to \infty)]/[C(0) - C(t \to \infty)]$ has decayed to a value of $\tilde{C(t)} < 0.01$ and stays below this threshold[28].

3.4 Implementation

In this section the classical numeric approach that was implemented here is presented. For this implementation numerical methods are being used to calculate the full matrix of the Hamiltonian H and observables O. This is done by first calculating the Hamiltonian and observable matrix elements using a spin basis

$$|\uparrow\downarrow\downarrow ...\rangle, |\downarrow\uparrow\downarrow ...\rangle, ... \tag{60}$$

Since Hamiltonian and observable are only given through spin operators and coefficients, the action of the spin operators on the spin states is apparent. Through solving the eigenvalue problem of the Hamiltonian with common numerical methods, energy eigenvalues and vectors can be generated. A basis transformation of the Hamiltonian and observable matrix into the energy eigenbasis can now be performed via

$$O_{eeb} = T^{-1} O_{sb} T, \tag{61}$$

with

$$T = \begin{bmatrix} e_1 & e_2 & e_3 & \dots & e_d, \end{bmatrix} \tag{62}$$

where e_i are the energy eigenvectors.

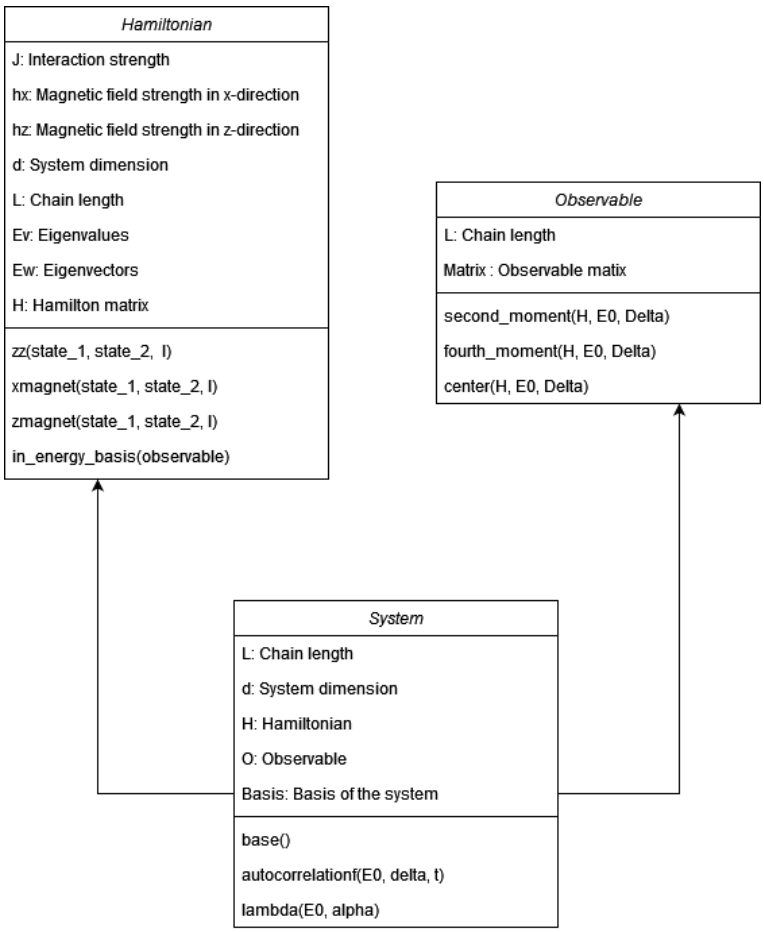

Figure 2: General structure and dependences of the classes which forms the program, that models the system considered and is able to calculate important quantises through the use of classic numerical methods. It consists of three classes. One for the Hamiltonian, which calculates the Hamiltonian matrix, diagonalize it and has some additional functionalities that are connected to the Hamiltonian. One for the Observable, which has analogue functionalities like the Hamiltonian, and a system class which manages the Hamiltonian and observable classes and calculates desired quantities.

For the implementation itself, a class based framework is being used. A visual illustration of the classes implemented can be seen in Fig. 2. In this diagram every box represents one class, where the attributes are listed in the first section and the methods in the second. The arrows between the classes describe a generic connection between the two classes.
The first class is the representative of the Hamiltonian. Firstly, it holds the corresponding parameters, including the interaction strength J and the magnetic field strength h_x. Also, it holds the basic parameters of the system like the chain length L and the dimension of the system which is just $d = L^2$. After calculating the matrix elements of the Hamiltonian, the full Hamiltonian matrix H is also held as a parameter with its eigenvalues E_v and eigenvectors E_w.
The methods of the Hamiltonian are then implementing the action, of the dif-

ferent terms in the Hamiltonian, on the spin basis states. Where the parameter l describes the lattice position. With the full matrix of the Hamiltonian and its energy eigenvalues and vectors, the Hamiltonian class is also able to transform a given observable matrix into the energy eigenbasis through Eq. (61).
Necessary for the system is now also the representation of an observable. The observable class holds, just like the Hamiltonian, the chain length L and its matrix as parameters.
Aside from its calculation of its matrix elements, the observable class calculates the second and fourth moment of itself, where it takes the Hamiltonian of the system H and the mean energy E_0 and the energy width Δ.
Now the last class is the system class which manages the Hamiltonian and observable. Therefore, again it holds the chain length and system dimension. Now also the spin basis needs to be implemented. Every possible spin state can be implemented easily through generating all binary numbers in the interval $[0, d]$. All basis states are then also saved. With that it creates instances of the Hamiltonian and observable classes and holds them as parameters.
The system class also implements all desired quantities like the autocorrelation function C and the relation of moments indicator Λ_T.

3.5 Results

In the following, results are presented from the study of the described system. The system is studied with respect to integrability and chaos, fulfilment of ETH and thermalization behaviour through the quantities described in Sec. 3.3. For the basic parameters of the Hamiltonian $J = h_x = 1$, $h_z = 0.5$, $h_2 = 0.165$ and $h_5 = -0.24$ was chosen.
First the ratio of adjacent level spacing for quantum chaos is calculated. For a chain length of $L = 11$ the average yielded a value of 0.518 which is in good agreement with the expectation for a Wigner-Dyson distribution for the energy level spacing, see Eq. (52). The Hamiltonian therefore shows chaotic properties for chain lengths larger then $L = 11$.
With respect to ETH, the system is studied with respect to the diagonal and off-diagonal aspects for the observable defined in Sec. 3.2.

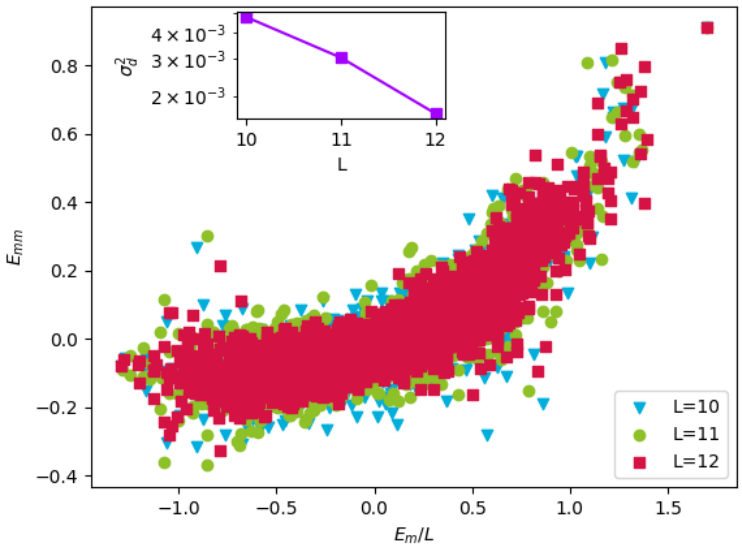

Figure 3: Diagonal matrix elements of the observable $O = \sigma_0^z$ in the energy eigenbasis, for system sizes $L = 10, 11, 12$. The inset shows the variance of the diagonal matrix elements as function of the system size for an energy window of width $\Delta = 4$ centred around $E_0 = 0$.

In Fig. 3 the diagonal matrix elements O_{mm} are plotted as function of the corresponding energy divided by the chain length E_m/L. The rescaling of the energy eigenvalues by the chain length allows to directly compare values obtained for different system sizes. In the diagram data is showed for chain lengths from $L = 10$ up to $L = 12$.

The ETH predicts, that the diagonal elements should converge to a smooth function of the average energy $\overline{E}$, in the thermodynamic limit $L \to \infty$. Even though the diagram spreads at the edges, the diagonal elements converge to a smooth curve for increasing system size. Especially for a narrow energy window centred around the average energy $\overline{E} = 0$ the ETH is well fulfilled with respect to the smoothness of the diagonal elements.

Furthermore, in the inset of the aforementioned figure the variance of the diagonal elements σ_d^2, defined in Eq. (53), is shown as function of the chain length L, for an energy window of width $\Delta = 4$ centred around $E_0 = 0$. As expected, the variance also decreases exponentially with system size.

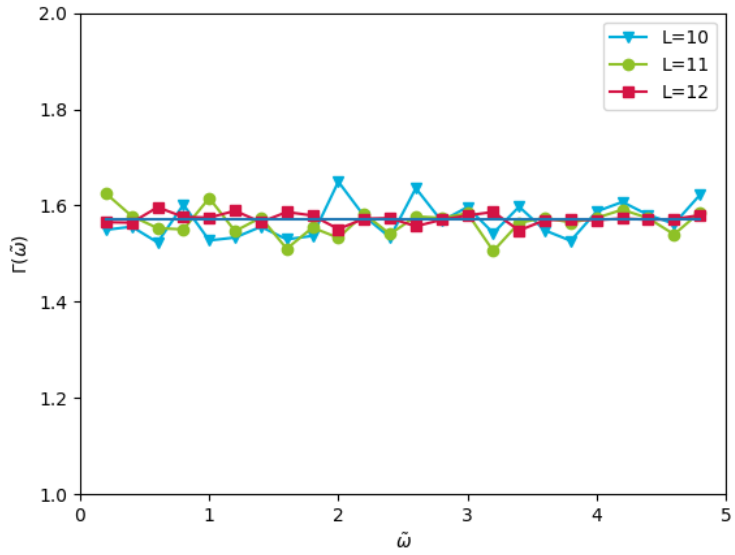

Figure 4: Frequency-dependent ration $\Gamma(\tilde{\omega})$ for a mean energy $\bar{E} \in [-0.5, 0.5]$, with a frequency width of $\Delta\omega = 0.05$ for system sizes $L = 10, 11, 12$.

In order to study the distribution of the off-diagonal elements of the ETH we study the frequency-dependent ration $\Gamma(\tilde{\omega})$, described in Sec. 3.3. Fig. 4 shows $\Gamma(\tilde{\omega})$ for an average energy of $\bar{E} \in [-0.5, 0.5]$ and a frequency width of $\Delta\omega = 0.05$ for different system sizes $L = 10, 11, 12$. Even though $\Gamma(\tilde{\omega})$ is fluctuating around the Gaussian value of $\pi/2$, the fluctuations decrease with increasing system size, indicating a Gaussian distribution of the O_{nm} for a wide range of frequencies in sufficiently large systems.

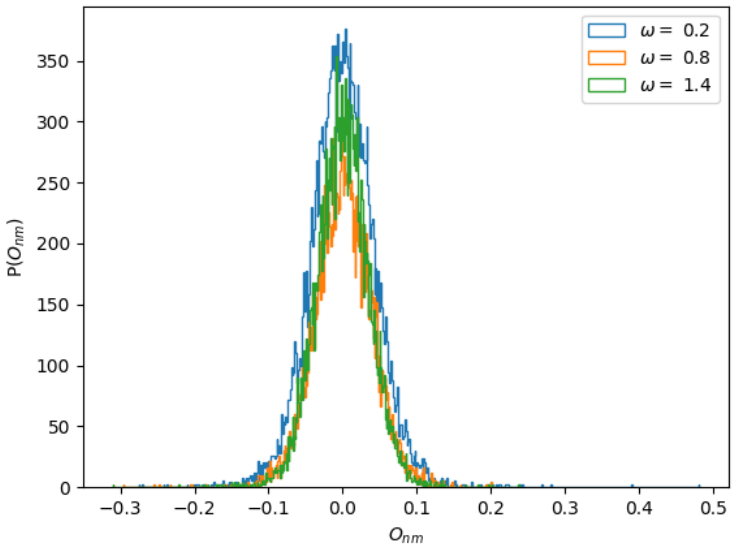

Figure 5: Full distribution of the O_{nm} in for a variety of frequencies and a frequency width of $\Delta\omega = 0.05$.

Fig. 5 also shows the full distribution $P(O_{nm})$ for different ω while the same frequency width was used as above. The distribution shows a Gaussian shape for a variety of frequencies.

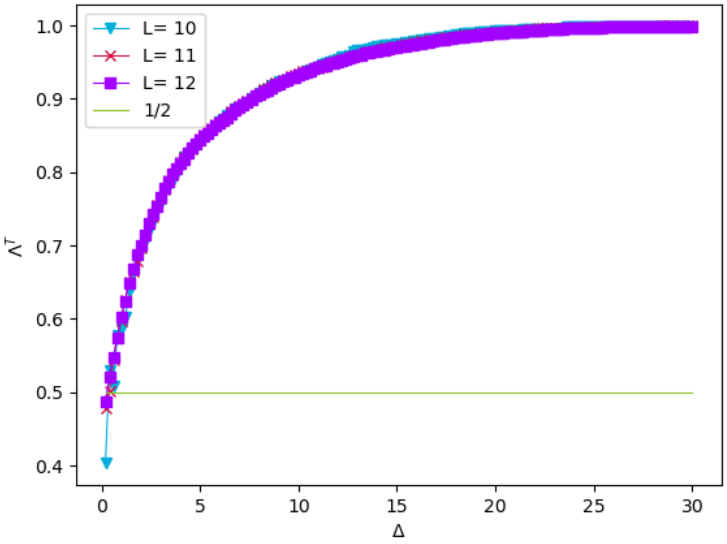

Figure 6: Relation of moments Λ^T as function of the energy window width Δ. Theoretical RMT value of 1/2 is indicated as the green line.

The independence of the off-diagonal elements of the ETH is studied through the relation of moments indicator. Fig. 6 shows the relation of moments indicator for an energy window centred around $E_0 = 0$ as function of Δ, for chain

lengths of $L = 10$ up to $L = 12$. Additionally, the theoretical value predicted by RMT of $1/2$ is plotted as reference.
For increasing energy window width Δ the relation of moments indicator converges to a value of 1, in which it meets the exact value, when Δ comprises the full energy spectrum. For narrower energy window widths the indicator approaches the theoretical value of $1/2$. Comparing the two indicator for diagonal and off-diagonal part of the ETH, the off-diagonal part demands a more narrow energy window, than the diagonal part. While the diagonal part is in good agreement with the ETH at a width of $\Delta \approx 5$, the off-diagonal part demands a width of at least $\Delta \approx 2$ to be in rough agreement and around $\Delta \approx 1$ for good agreement, with the value predicted by RMT. For increasing system size the shape of the results do not change, aside from becoming more pronounced.
At this point it should be noted, that for too narrow energy widths, these calculations become non representative, since the considered quantities measure stochastic properties of the matrix elements of the submatrix. In order for statistical phenomena to occur, the energy window needs to comprise a certain amount of energy eigenstates. This is one reason why the graph in Fig. 6 does not meet the exact value of $1/2$ for $\Delta \to 0$.

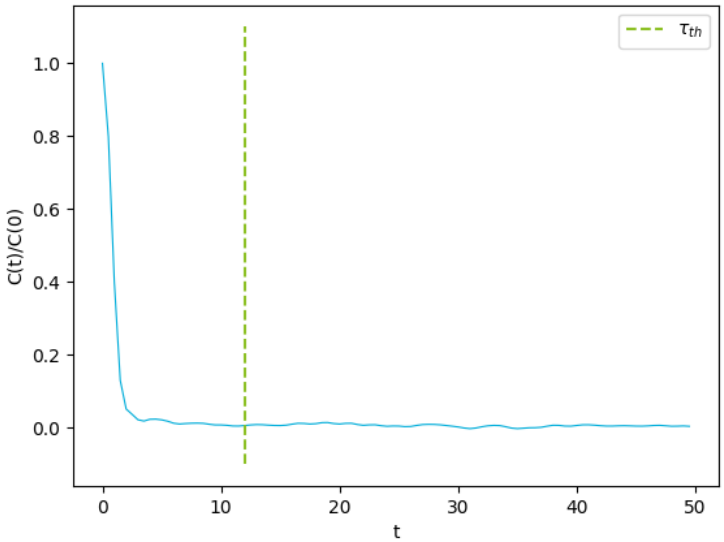

Figure 7: Normalized autocorrelation function for the centred observable as function of the time. The dashed line indicates the thermalization time, where the autocorrelation function stays under 10 percent of its initial value.

The thermalization behaviour can directly be seen through the autocorrelation function defined in Eq. (57). In Fig. 7, the autocorrelation function $C(t)$ divided by its initial value $C(0)$ for the local observable $O = \sigma_0^z$ centred around

its ensemble average $\langle O \rangle_{mc}$, is plotted as function of time t. The data is obtained for a chain length of $L = 11$ and an energy window of $(E_0 = 0, \Delta = 2)$. The autocorrelation function shows a typical relaxation dynamics. Starting at the initial value 1 for $t = 0$, the autocorrelation function collapses to zero after slight oscillations. Since the observable is centred, the relaxation to zero indicates the observable meeting its value predicted by the microcanonical ensemble. The ensemble description, describes many-body systems in an equilibrium state. The autocorrelation function therefore shows the equilibration dynamics of the observable O.

4 Operator Basis Norm Method

The operator basis norm method for calculating the relation of moments indicator Λ^T for random matrix behaviour, Sec. 2.4, now revolves around making use of the commutation rules and properties of the Pauli operators. In this method, the calculation of the moments of the projected operator is done through a counting of terms, this method promises to be much lower in computation time which leads to the possibility to calculate the relation of moments indicator for much larger system sizes.
In order to make use of the properties of the Pauli operators the projection operator $P_T = \sum_{|E_m - E_0| \leq \Delta E} |m\rangle \langle m|$ is approximated by a polynomial filter $\mathcal{P}$. For simplicity reasons we will for now consider the most simple case of a parabolic filter

$$P_T = \mathcal{P} = 1 - \alpha \tilde{H}^2, \quad \alpha \in [0, 1] \tag{63}$$

where α is a parameter and

$$\tilde{H} = \frac{H - b}{a}$$

with $a = \frac{E_{max} - E_{min}}{2}$ and $b = \frac{E_{max} + E_{min}}{2}$, is the rescaled Hamiltonian. These definitions become clear, when viewing P_T in the energy eigenbasis. In the EEB the Hamiltonian can just be replaced with energy eigenvalues E_n. Inserting the largest energy eigenvalue E_{max} into the definition of $\tilde{H}$ then yields 1. On the other hand, inserting the smallest energy eigenvalue E_{min} into the definition of $\tilde{H}$ yields -1. Therefore the spectrum of the Hamiltonian is rescaled to the interval $[-1, 1]$.
Now again viewing the projection operator Eq. (63) in the EEB, the second term can be seen as a weighting with respect to the energy gap to the average energy b. Energy eigenstates which are close in energy to the mean energy are highly weighted while the bigger the gap becomes, the states are weighted lower with respect to the quadratic dependence. The parameter α now controls the slope of the quadratic weighting, where $\alpha = 0$ corresponds to no weighting
.
In order to visualize this filter, we consider $\mathcal{P}$ in the energy eigenbasis. The filter is now just a function of the energy $\mathcal{P}(E_n)$. In Fig. 8 the parabolic filter from Eq. (63) can be seen for the system described in the previous section with a system size of $L = 11$

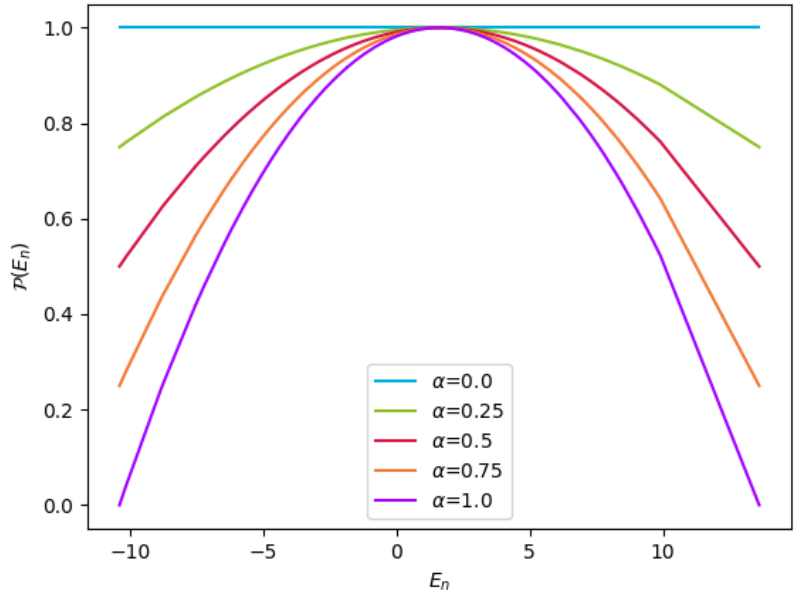

Figure 8: Parabolic filter plotted over the energy for a variety of parameter values α.

Inserting the approximated projection operator into Eq. (46) yields

$$O_T = (1 - \alpha\tilde{H}^2)O(1 - \alpha\tilde{H}^2). \tag{64}$$

since the rescaled Hamiltonian $\tilde{H}$ and the observable O are just given by Pauli operators or real numbers, the properties of the Pauli matrices can be taken advantage of.

Consider the second moment of the projected operator

$$M_2 = \frac{1}{d}Tr\{O_T^2\}. \tag{65}$$

Inserting the projected operator O_T with the approximated projection operator Eq. (63) and executing all products and sums, yields a trace over a sum of terms, which only consist of products of Pauli operators with coefficients. As mentioned in Sec. 2.5 all Pauli matrices, and therefore also all products of Pauli matrices, are traceless. Therefore the only terms that contribute to the second moment, Eq. (65), are those who can be reduced to the identity matrix, through the properties of the Pauli operators described in Sec. 2.5. The value of the contribution is then given by the coefficient of the respected term. This coefficient describes how many times this term occurs in the initial sum. It should be noted, that a special property of the system used here simplifies this

problem. As described in chapter 2.5 the only way to reduce a term of Pauli matrices only consisting of two kinds, is by $(\sigma^i)^2 = 1$.

The first moment of the operator, which is the expectation value, is always zero for the considered system. Therefore when generating all terms in Eq. (64), there is no term that can be reduced to the identity. Now consider the projected operator fully written out

$$O^T = c_1 \cdot \sigma_q^z + c_2 \cdot \sigma_q^z \sigma_0^z \sigma_1^z ..., \tag{66}$$

This equation can be interpreted as the form the operator O^T takes, in an operator basis. Each term then consists of a basis operator, which are the Pauli matrices and a respected coefficient. The operator O^T can therefore be written like

$$O^T = C \cdot [e_1 + e_2 + ...] \tag{67}$$

with

$$C = \begin{bmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_d \end{bmatrix} \tag{68}$$

and

$$e_1 = \sigma_q^z \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \quad e_2 = \sigma_q^z \sigma_0^z \sigma_1^z \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \dots \tag{69}$$

It should be noted, that this representation is based on the fact, that the trace is taken over the projected operator O^T. The products of Pauli matrices fulfil the conditions of basis states in this context as well as normalisation

$$Tr\{e_i \cdot e_j\} = \delta_{ij}. \tag{70}$$

If we now want to calculate the second moment we can just take the O^T operator in the operator basis. The trace of the squared O^T operator will now only have contributions of the products of the basis states with itself. Therefore

the second moment has the form of a squared 2-norm in the operator basis

$$M_2(O^T) = \frac{1}{d} Tr\{(POP)^2\} = \sum_{i=1}^{d} c_i^2 = ||O^T||_2^2. \tag{71}$$

As we can see, in order to calculate the second moment of O^T we just need to calculate the coefficient matrix of O^T in the operator basis described above. This procedure can be generalized for arbitrary moments of even power. Every even k-Moment of the operator O^T is fully determined by the coefficient matrix of the operator $(O^T)^{(k/2)}$ in its natural operator basis, so generalizing Eq. (71) for an arbitrary operator O in the natural operator basis, leads to

$$M_k(O) = ||O^{\frac{k}{2}}||_2^2 \quad \text{for even } k \tag{72}$$

4.1 Counting method

The counting method does take a straight forward approach. It generates all terms of the respected operator and transforms the resulting sum into the minimal form. From that, the coefficient matrix can be extracted and with that the norm and the moment at hand, can be calculated.

4.1.1 Concept

In order to be able to decide if two terms are equal, we need to define a minimal form, which is unique for every product of Pauli matrices. The minimal form used here sets an order for the Pauli matrices from left to right. Firstly the matrices are ordered ascending in kind with $(1, \sigma^x, \sigma^y, \sigma^z) \rightarrow (\sigma^0, \sigma^1, \sigma^2, \sigma^3)$ and secondly ascending in lattice position. The term needs to be ordered in this way through making use of the rule $\sigma_l^i \sigma_m^j = (-1)^{\delta_{lm}} \sigma_m^j \sigma_l^i$. If the term is ordered, all the same Pauli matrices are next to each other and with $(\sigma^i)^2 = 1$, the term can be reduced to the minimal form. This minimal form is unique for every product of Pauli matrices. For example the term

$$\sigma_1^z \sigma_0^z \sigma_0^x \sigma_1^z, \tag{73}$$

transforms first to

$$-\sigma_0^x \sigma_0^z \sigma_1^z \sigma_1^z \tag{74}$$

and then to

$$-\sigma_0^x \sigma_0^z, \tag{75}$$

in the minimal form.

Now we need to be able to define a Hamiltonian and an operator in the form of Pauli matrices. The given terms should implement the respected algebra. On the other hand we need to be able to make use of the commutation rules and properties of the Pauli operators, in order to change the order of the operators and reduce the terms to the minimal form.
After generating the minimal form of the operator, the last challenge lays in the counting process to generate the coefficient matrix. With that the rest of the calculation is straight forward by just calculating the sum over the squared coefficients for the POP and $(POP)^2$ operator for the second and fourth moment.

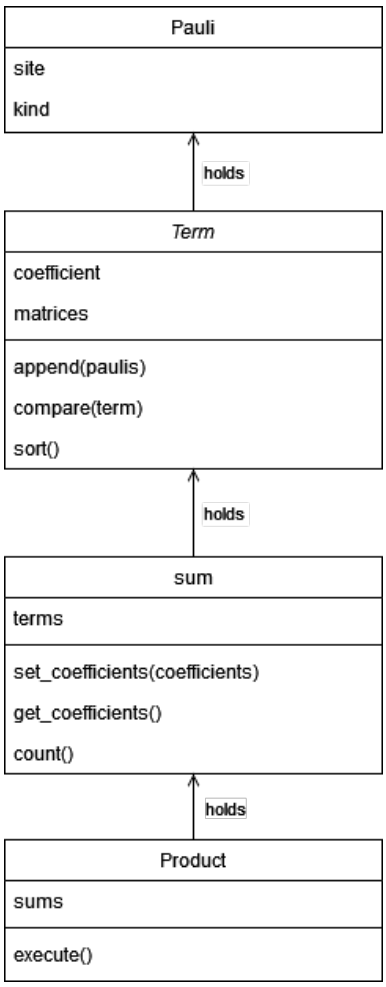

Figure 9: Visual illustration of the hierarchy of classes which form the program that implements the OBM-Method. Every block represents one class where the first section lists the attributes and the second section lists the methods of the respected class. Every class is a direct representation of the corresponding mathematical concept where every functionality, that is important for the task at hand, is implemented. Every object essentially consists of a list of objects from the class one step lower in the hierarchy.

The program consists of a hierarchy of objects where each object is just a list with extended functionality. At the bottom we need a representative for the Pauli operator. The only information about each Pauli operator that is needed for this implementation, is its kind and on which lattice site it is acting.
The next bigger object is the term. A term holds a list of Pauli operators and a coefficient. This object also needs to be able to transform itself into the minimal form, in accord with the commutation rules and properties of the Pauli operator.
Next comes the representative of a sum, which holds a list of terms. The sum object needs to be able to perform the counting process, when all terms are in the standard form.
The last object is now the product. it holds two sums and is able to perform

the product of a sum and therefore returns a sum in the process. These objects now build the basic framework of the program.

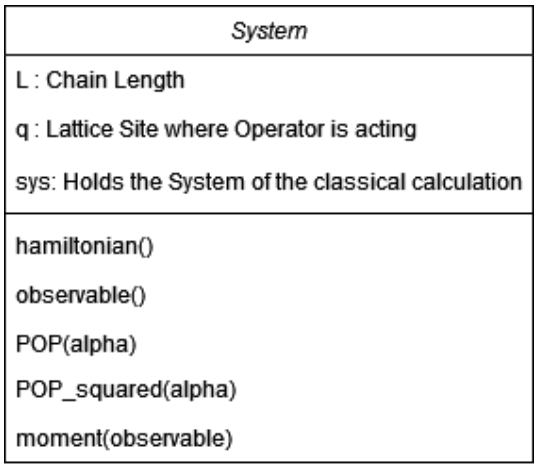

Figure 10: Description of the system class, which analogous to the classic model, manages the implemented framework described above. It creates the Hamiltonian and observable as well as the projected operator and its square under the use of the above framework and with that calculates the second and fourth moment.

The next step is now to actually create the system, the Hamiltonian H and the operators O and P. Since we have implemented direct representatives of terms, sums and product of sums this process is relatively straight forward since its definitions, Eq. (43), (44) and (45), can directly be created as instances of sums or products. With the maximum and minimum energy eigenvalues from the classical calculations the Hamiltonian can be scaled. Now also the projection operator P and the two operators POP and $(POP)^2$ can be created.
In order to calculate the second and fourth moment we need to calculate the coefficients in the operator basis as described in Sec. 4. Therefore all terms in the sums POP and $(POP)^2$ need to be transformed into the standard form and all equal terms need to be counted. With the resulting coefficients the second and fourth moments, and therefore the relation of moments indicator Λ^T, can be calculated with Eq. (71).

4.1.2 Implementation

The counting of terms and therefore the calculation of the coefficients is the most computation intensive part of the program. Therefore efforts should be made to make this process as efficient and fast as possible.
Here an algorithm was implemented, that revolves around a hash map. Given a sum object with all terms being in the minimal form, a unique string is generated for every term. This string can for example list the pauli matrices kind and lattice site in the right order. This unique string can be used as a

hash function. While looping over the terms in the sum, for every term its unique hash is generated. With a dictionary object, it can be checked if this hash already exists as key in the dictionary. If it does not exist, a new entry is appended with the hash as key and the coefficient as value. If the hash already exists as key, the coefficient of the considered term is added to the value in the dictionary. Since the hash is a unique mapping to a series of pauli matrices, a counted sum object can be recovered from the dictionary, through reverse mapping the keys and values in the dictionary, to term objects and join these into a sum.
This algorithm has multiple benefits. Not only is it relatively simple to implement, but it is also relatively fast. Since only one loop over the sum is necessary and dictionaries have constant lookup times, the runtime is $\mathcal{O}(n)$ with n being the number of terms in the sum.
As described in the previous section, the transformation of every term into the minimal form can be implemented through a sorting of the Pauli matrices. The sorting was implemented here through an algorithm akin to the bubble sort algorithm. Similar to this known algorithm, a pointer is going through the considered term from left to right. In the traversing of the term, the pointed at matrix is compared to its right neighbour, and its decided if the two matrices need to exchange their position, depending on the definition on the minimal form, or are the same matrix. If they are same matrix, the matrices are removed from the term. If they are not equal and they are in the right order, the pointer shifts by one position to the right. Lastly, if the matrices need to exchange positions, their position is swapped and if necessary the coefficient is multiplied by -1, in accord with Eq. (42). This algorithm is repeated until the end of the term is reached, then if the term was modified since it was last put in the starting position the algorithm is started again from the starting position. If the term is fully traversed without a needed modification the term is in the minimal form and the algorithm is terminated. This implements the commutation rule in Eq. (42) and fully transform the term into the minimal form.
The benefit of this sorting algorithm for the purpose at hand is, that it terminates very fast, if the term is already sorted. Since in the program, the sorting algorithm is often called on a sum, that was only mildly changed, the case of the term already being sorted, or at least only needs slight adjustment, is a very common case. Therefore, even though there are better performing algorithms on the average case, this algorithm behaves particularly well in the best case scenario having a runtime of up to $\mathcal{O}(n)$, which is why we chose this algorithm here.

In order to generate the sums POP and $(POP)^2$ the operators O, H^2 and therefore also H are needed. These operators represent the considered system

and therefore have to be implemented by hand with the classes described in Sec. 4.1.1.

Now when generating the projected operators POP and $(POP)^2$ efforts should be made to already minimize the size of the operators in order to minimize computational effort when counting the resulting sum. First of all the projection operator P which is defined by Eq. (64) can also directly be implemented with its definition and H^2. The full projected operator can be generated by creating a sum object which is equal to P. After this initialisation the sum is multiplied with the next operator in the first case O. This process is continued until the operator POP or $POPPOP$ is created. After every iteration of multiplying the initial sum with the respected next operator the result is already put into the minimal form for every term in the sum and counted to minimize the size of the sum.

With the operators POP and $(POP)^2$ in the operator basis the second and fourth moments and with that relation of moments indicator Λ^T, can be calculated with Eq. (71) and (54). This calculation can now be performed with different values for α.

5 Results

In this section the results of the OBN counting method, described in Sec. 4, for the relation of moments indicator will be discussed. First, the approach, which makes use of the pauli operator computation rules, will be compared to the results obtained by exact diagonalization, for a very simple case. Thereby, we will show, that both approaches in fact yield equal results. After that, the OBN method will be compared to the hard cut energy window. We will go through a number of different approaches of how to construct a polynomial filter, that on the one hand shows agreeing results with the hard cut energy window, while on the other hand keep the computational costs as low as possbile.

Firstly the in Sec. 4.1 described OBN counting method will be compared to exact diagonalization by calculating the relation of moments indicator, using Eq. (72). In Fig. 11 the results for the relation of moments indicator is shown for a variety of system sizes, starting at $L = 2$ up to $L = 6$. The upper figure resulted from the counting method, while for the lower figure the full matrices were being used. For simplicity reasons a polynomial filter of second power was used for this computation. As it can be easily seen both methods show the exact same results over a variety of system sizes.

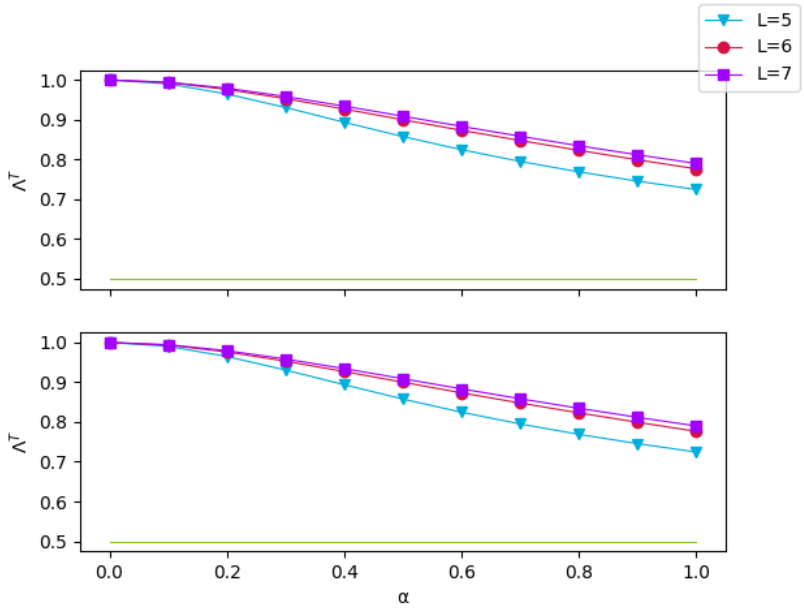

Figure 11: Comparison of the relation of moments indicator with a parabolic filter, as function of the parameter α, using different methods for a variety of system sizes L. The upper graph shows the result from the OBN method, while the lower graph shows the results from exact diagonalization.

Even though relatively small system sizes were considered, Fig. 11 already

suggests, the results from the second power polynomial filter does not show the desired behaviour for the relation of moments indicator. For $\alpha = 0$, meaning filtering turned completely off, the method yields the expected value of $\Lambda^T = 1$. Increasing the filter parameter α leads to a drop of the graph, however for reaching $\alpha = 1$, the relation of moments does not converge to $1/2$. For increasing system size the value for $\alpha = 1$ in fact becomes increasingly closer to 1.

For that reason we will now consider a different approach for the shape of the polynomial filter. The most straight forward approach, after the second power parabolic filter, would now be to consider higher even powers. However, doing this leads to a significant problem. Changing the power of the parabolic filter will change the slope of the filter and therefore the shape of the filter will get closer to the shape of a hard cut filter. However, the filter with the rescaled Hamiltonian $\tilde{H}$ is constructed in a way so that it always only takes values from the Interval $[0, 1]$ in the spectrum of the Hamiltonian. For stronger slopes the rescaling leads to a stretching of the filter shape, resulting in an effective energy window, which is more 'hard-cut-shaped' but wider, than more weaker slopes. This can be seen in Fig. 12 where filter shapes are plotted over the energy, for parabolic filters and different powers.

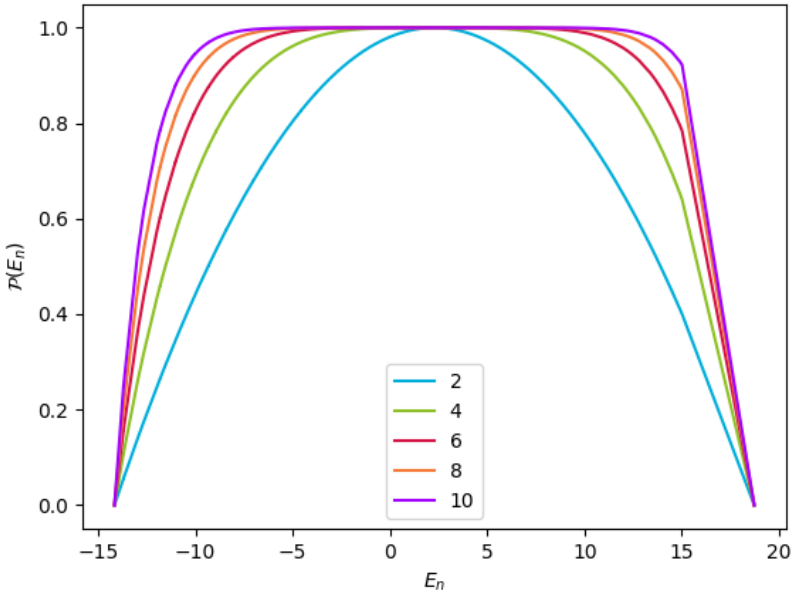

Figure 12: Polynomial filter as function of the energy E_n for a parabolic filter shape. The graph shows the shape of the filter for $\alpha = 1$ for different powers H^2 up to H^{10}.

In order to resolve this problem we chose a different approach for the filter shape, that has a series of benefits over the intuitive approach. We used a Gaussian shaped filter, where the projection operator is then given by

$$\mathcal{P} = e^{-\frac{1}{2}(\frac{H}{\sigma})^2}. \tag{76}$$

In Fig. 13 the shape of the filter can be seen in the energy eigenbasis as function of the energy E_n.

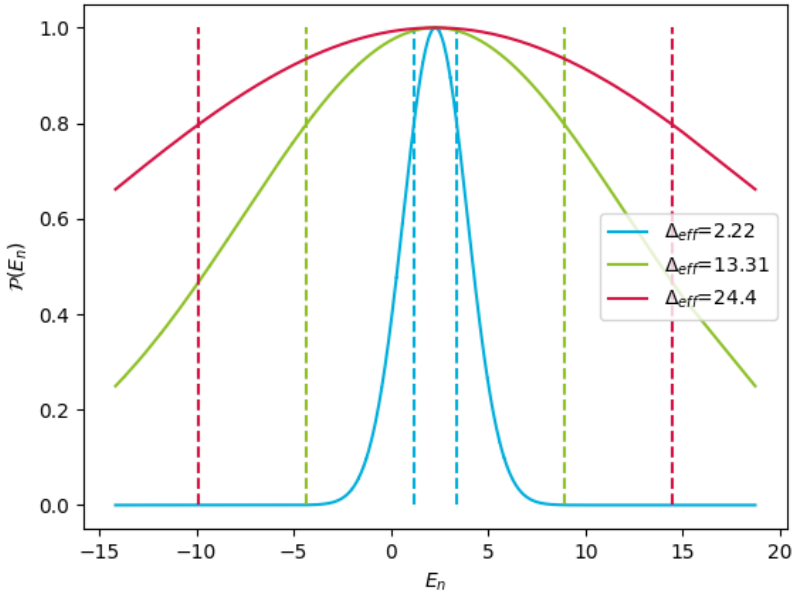

Figure 13: Polynomial filter as function of the energy E_n for a gaussian filter shape. The graph shows the shape of the filter for a number of different effective energy window widths Δ_{eff}. The dashed lines indicate the effective energy window of the respected filter.

With the standard deviation of the Gaussian σ^2 we can now control the width and the slope of the filter. We define an effective energy window width Δ_{eff} from the standard deviation of the Gaussian filter as

$$\Delta_{\text{eff}} = 2 \cdot 0.674490 \cdot \sigma. \tag{77}$$

With this definition, the integral over the Gaussian filter for the interval $[E_0 - \Delta_{\text{eff}/2}, E_0 - \Delta_{\text{eff}}/2]$ corresponds to half of the full integral over the real line. In other words, the effective energy window is constructed in a way so that it defines a centred interval around the mean, so that an integral over the interval yields half the value of a full integral over the interval $[-\infty, \infty]$.
Also we define an effective dimension d_{eff}. When defining the effective dimension one needs to be careful to find an accurate representation of the dimension of the subspace the projection operator projects into. Also, in general the definition should become equal to the exact definition in the hard cut energy

window, when transitioning from a Gaussian to a hard cut window. We therefore made use of the fact, that the trace over a projection operator is equal to the dimension of the subspace the operator projects into. We therefore defined the effective dimension as

$$d_{\text{eff}} = Tr\{\mathcal{P}\}. \tag{78}$$

Another benefit of this definition is, with our purpose at hand in mind, that since the trace is taken over the filter, the effective dimension can be easily calculated without the use of exact diagonalization. Furthermore the definition is independent of the particular kind of filter one considers.
Through making use of the series definition of the exponential function, Eq. (76) can be written as a polynomial

$$\mathcal{P} = \sum_{n=1}^{\infty} \frac{-1^n}{2^n \sigma^{2n} n!} H^{2n}. \tag{79}$$

The sum in Eq. (79) can be cut off when a sufficient approximation is accomplished over the spectrum of the Hamiltonian. So Eq. (79), can be calculated through the counting method.
Fig. 14 shows the relation of moments indicator Λ^T for the exact method in comparison with the OBN method with the Gaussian filter. The results are generated for a system size of $L = 11$. The results obtained from the OBN method show a slightly larger value from the limit value of 1 for very large windows. This is probably related to the fact, that the Gaussian falls off to a value of around 0.8 towards the edges of the effective window. If one considers energy windows areound the size of the full spectrum, the exact window covers the full spectrum while the Gaussian falls of towards the edges, which leads to a deviation in the result. Towards lower energy window widths, the slope for the Gaussian filter stays close, to the exact results while it starts to fall off slightly harder until it crosses the exact results at around $\Delta_{(\text{eff})} \approx 11$. After that point, the results for the Gaussian filter stay closely under the exact results slowly falling of lesser until it crosses over the exact results again at around $\Delta_{(\text{eff})} \approx 3.5$. From there the results obtained again are slightly larger, compared to the exact results while falling of slightly faster. Finally both results meet each other for $\Delta_{(\text{eff})} \to 0$ at the theoretical value of $\Lambda^T \approx 0.5$. All in all the results for the Gaussian filter stays in close proximity to the exact results over the full range of energy windows, while it 'slowly oscillates' around it. This 'slow oscillation' in particular shows as slightly larger values for large energy windows, slightly lower values in medium sized windows and again slightly larger values for small energy windows.

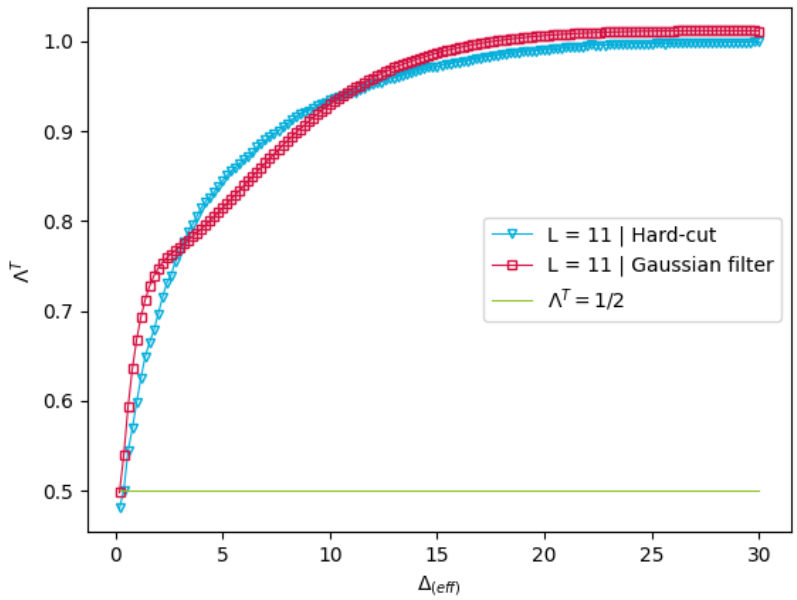

Figure 14: Results for the relation of moments indicator calculated for a Gaussian filter which was approximated using a Taylor series, in comparison to the hard cut energy window, for system size $L = 11$.

Since the Gaussian filter gives good results however small deviations in specific parts of the range, we will now look at the differences in filter shapes in more detail. Since the case of large energy windows was already covered above, we will now look at the differences between the hard cut energy window and the Gaussian for medium sized and small sized windows.

In Fig. 15 the results for both filter shapes can be seen again for a system size of $L = 11$. In the graph two values, $\Delta_{\text{eff}} = 0.5$ and $\Delta_{\text{eff}} = 5$, are highlighted through dashed lines. At the top left and top right the Gaussian filter is plotted over the energy in comparison to the hard cut energy window, for the two highlighted values respectively.

For the medium sized window $\Delta_{\text{eff}} = 5$, the Gaussian and the hard cut window show relatively large differences. Not only does the Gaussian not match the window function inside the window well, but in particular the slow decay outside of the window still shows relatively large values for a much larger interval.For the small window $\Delta_{\text{eff}} = 0.5$, the filter shapes become increasingly alike, inside of the effective energy window the gaussian serves as as a good approximation, the decay outside of the effective window is concentrated relatively close to the edges of the window, however it is relevant. These differences between the exact hard cut window and the Gaussian for small and medium sized windows seem to result in the small deviations from the exact results.

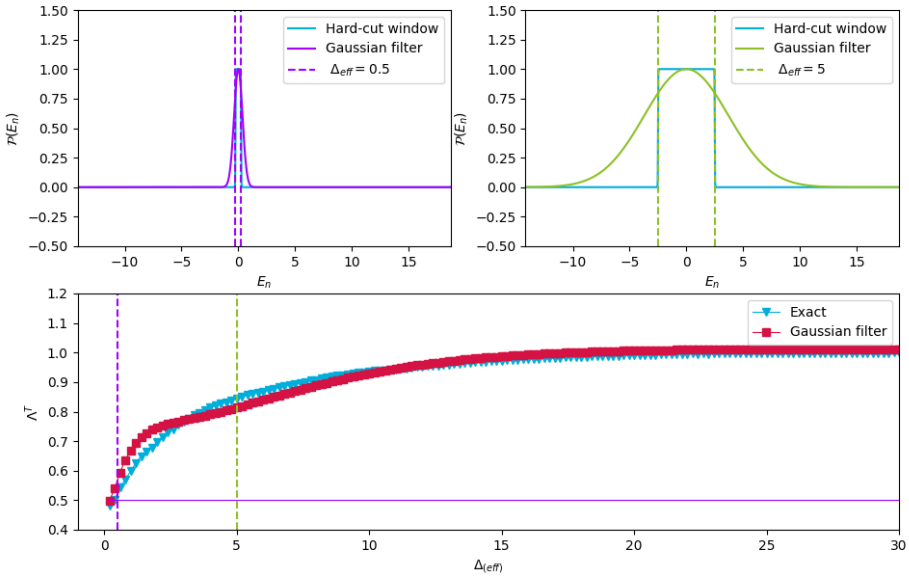

Figure 15: Bottom graphic: Results for the relation of moments indicator calculated for a Gaussian filter, which was approximated using a Taylor series, in comparison to the hard cut energy window, for system size $L = 11$. Top graphic: Shape of the used polynomial filter in comparison to the exact hard cut energy window for two points $\Delta_{\text{eff}} = 0.5$ and $\Delta_{\text{eff}} = 5$, which are indicated in the results as dashed lines.

Even though this approach shows good results for small energy windows, it has a significant problem. Given a fixed number of terms in the Taylor series of the exponential function Eq. (79), the interval over which the approximation is satisfactory is proportional to the variance of the Gaussian. That means, that a very small number of terms shows good results for large windows, however the smaller windows one considers, the more terms are needed to cover the full spectrum. In order to get accurate values for windows as small as $\Delta_{\text{eff}} \approx 0.5$ we needed to consider around a 1000 terms. Using the OBN method with a polynomial of such a high order goes against the central goal of increasing the performance. If the approximation is not accurate over the full energy spectrum, the filter starts to diverge to ∞ or $-\infty$ at the edges of the spectrum.
For these reason we explored a further approach of numerically fitting a polynomial to a desired filter shape, so that, for the example of the Gaussian filter, the resulting filter is not exactly a Gaussian, however it looks very similar but behaves better for lower orders. We used a function from the numpy library which allows one to fit a polynomial to a function. Put simply, this function is handed a series of points over an interval x, the function to be approximated evaluated at these points $f(x)$, and the desired degree of the resulting polynomial. The parameters a of a polynomial p of the desired degree are then optimized, so that

$$|f(x) - p(x)| \tag{80}$$

is minimal for all points x and the found parameters a are then returned.
We used this method for two filter shapes, a Gaussian filter shape as well as directly approximating the hard cut energy window. In Fig. 16 the results are shown for the Gaussian filter fit in comparison with the exact energy window, for a polynomial order of 50. The results where obtained for a system size of $L = 11$. The results are in good agreement with those for the Taylor series up to a effective window width of ≈ 1.5 where the results start to diverge. In the top left and right we see the polynomial fit in comparison to the exact Gaussian, for a window width where the fit yields bad, values $\Delta_{\text{eff}} = 0.4$ and good values, $\Delta_{\text{eff}} = 2$.
For $\Delta_{\text{eff}} = 2$, the polynomial fit shows a good agreement with the exact Gaussian inside of the effective energy window. The polynomial fit almost perfectly matches the exact Gaussian filter, only very small oscillations can be seen at the upper edge of the spectrum. For the $\Delta_{\text{eff}} = 0.4$ the polynomial fit oscillates around 0 with small amplitude fading out towards the edges, outside of the effective energy window. For the window itself however, the fit shows a much wider shape while reaching a peak of only ≈ 0.4. The fit does not approximate the Gaussian to a satisfactory amount for this value.

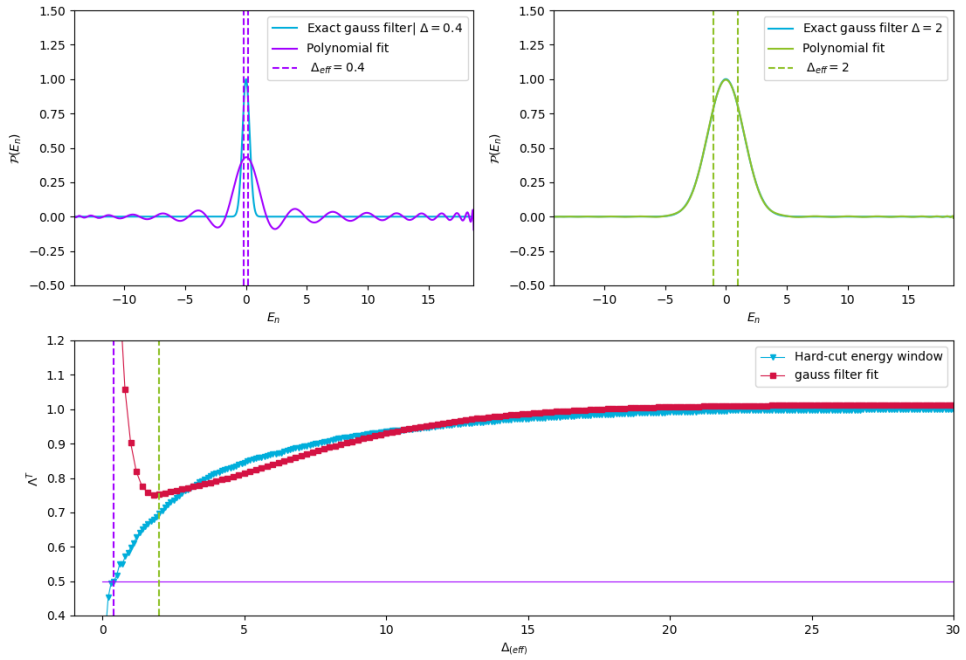

Figure 16: Bottom graphic: Results for the relation of moments indicator calculated for a polynomial fit of a Gaussian filter, in comparison to the hard cut energy window, for system size $L = 11$. Top graphic: Shape of the used polynomial filter fit in comparison to the exact Gaussian for two points $\Delta_{\text{eff}} = 0.4$ and $\Delta_{\text{eff}} = 2$ which are indicated in the results as dashed lines.

Next we will look at the polynomial fit of the exact hard cut energy window. In Fig. 17 the results are shown for the hard cut window filter fit in comparison with the exact energy window, for a polynomial order of 50. The polynomial shows good agreement with the exact window for large effective energy windows. However for narrower effective energy windows the slope falls off slightly harder as for the exact filter, also the graph starts to increasingly oscillate. At a value of $\Delta_{\text{eff}} \approx 3$ the results start to diverge, similar as for the Gaussian filter fit. In the top left and right we see the polynomial fit in comparison to the exact window filter, for window widths $\Delta_{\text{eff}} = 1$ and $\Delta_{\text{eff}} = 3.3$.
For $\Delta_{\text{eff}} = 3.3$ the results shows low amplitude oscillations around 0 fading out towards the edges outside of the effective energy window. For the window itself, the fit shows a 'Gaussian-type' shape which fits the hard cut window relatively well, however it peaks over the limit value of 1 up to around 1.25. For $\Delta_{\text{eff}} = 1$ the fit shows very similar problems as the Gaussian fit; the energy window is to wide and the peak is to small while the outside of the window shows the known oscillation behaviour.

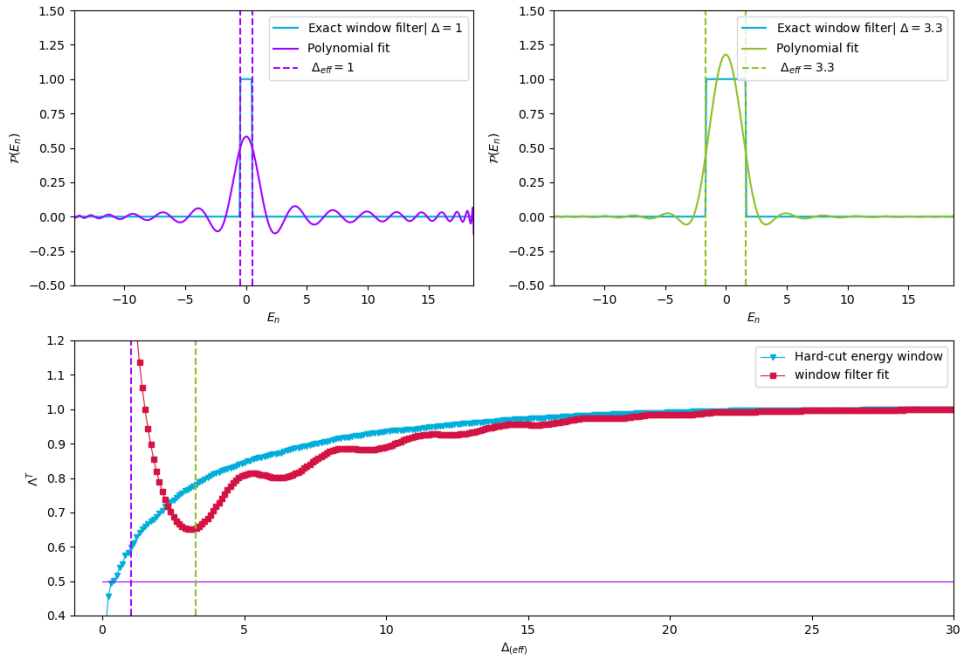

Figure 17: Bottom graphic: Results for the relation of moments indicator, calculated for a polynomial fit of a window filter, in comparison to the hard cut energy window, for system size $L = 11$. Top graphic: Shape of the used polynomial filter fit in comparison to the exact hard cut energy window for two points $\Delta_{\text{eff}} = 1$ and $\Delta_{\text{eff}} = 3.3$, which are indicated in the results as dashed lines.

Since we looked at the issue of small windows before, we will now look more into detail about the occurring oscillation of the results for the hard cut window fit. In Fig. 18 the results are shown again, with this time filter graphs for $\Delta_{\text{eff}} = 3.3$ and $\Delta_{\text{eff}} = 5$, which correspond to two turning points in the oscillating graph for the relation of moments indicator.
For $\Delta_{\text{eff}} = 5$ the shape inside the window does not show a Gaussian type shape, in contrast to $\Delta_{\text{eff}} = 3.3$. Instead one can see a fast rise up to a value of 1 at the edges of the energy window, then the polynomial fit oscillates around this value until it drops of again at the other edge of the window. Outside of the window we again see small oscillations around 0 which fade out towards the edges, however compared to $\Delta_{\text{eff}} = 3.3$ the oscillations are relatively large. Even though for $\Delta_{\text{eff}} = 5$ the shape matches the exact filter comparably poorly outside of the effective energy window, since it approximates the window itself better then the filter for $\Delta_{\text{eff}} = 3.3$ it yields closer results to the exact solution.

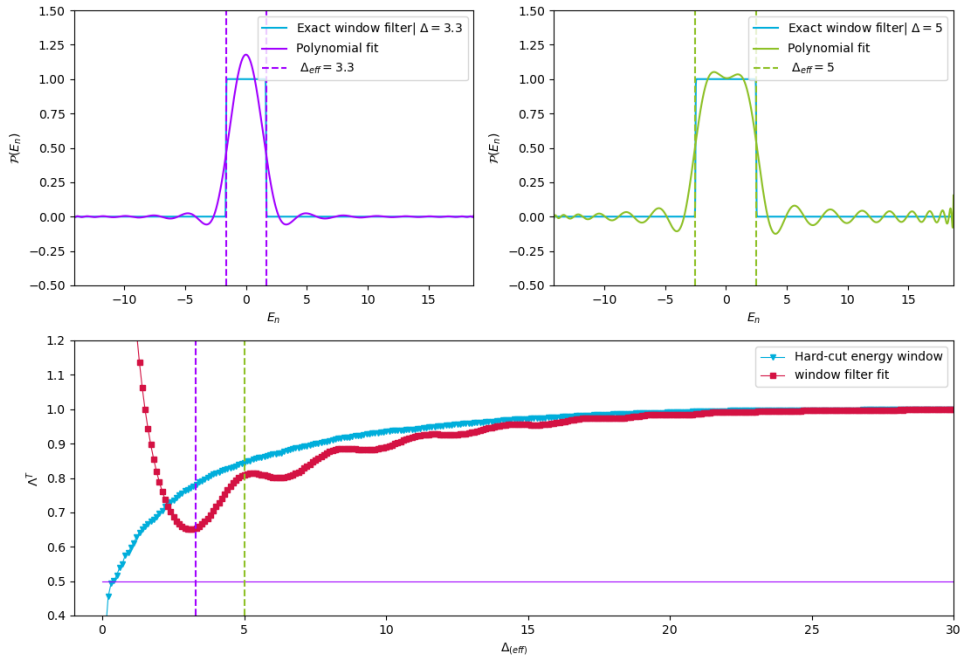

Figure 18: Bottom graphic: Results for the relation of moments indicator calculated for a polynomial fit of a window filter, in comparison to the hard cut energy window, for system size $L = 11$. Top graphic: Shape of the used polynomial filter fit in comparison to the exact hard cut energy window for two points $\Delta_{\text{eff}} = 3.3$ and $\Delta_{\text{eff}} = 5$ which are indicated in the results as dashed lines.

It should be noted here, that increasing the order of the polynomial fit, did not lead to significantly better results. The same results that are presented here, where also evaluated up to polynomial orders of 200 while no significant changes in the results where apparent.

The main problems for the divergence for low window widths is, that the effective window width Δ_{eff} for the exact filters do not match the filter shape of the resulting fits for small windows.

In order to tackle the problems that occurred with the polynomial fit onto the exact window, we rescaled the polynomial filter to a peak value of 1, as well as redefined the effective window width. Instead of using the exact window width, we use the definition from equation (77). The obtained results are shown in Fig. 19.

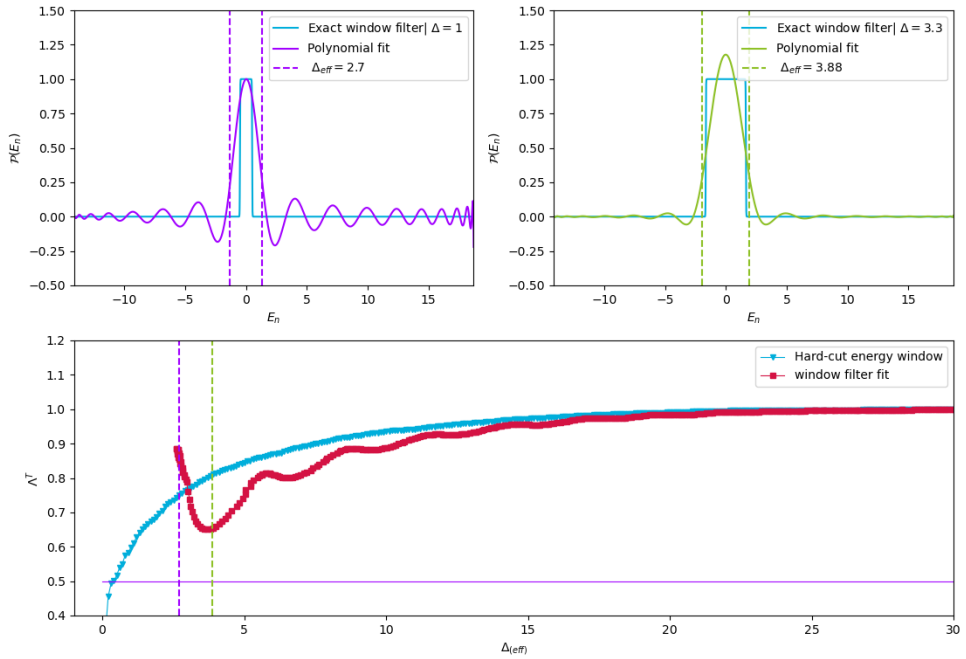

Figure 19: Bottom graphic: Results for the relation of moments indicator calculated for a polynomial fit of a window filter, in comparison to the hard cut energy window, for system size $L = 11$. Top graphic: Shape of the used polynomial filter fit in comparison to the exact hard cut energy window for two points $\Delta_{\text{eff}} = 3.3$ and $\Delta_{\text{eff}} = 5$ which are indicated in the results as dashed lines.

The oscillations of the graph are not significantly reduced. However, the polynomial fit did not produce filter shapes with an effective window width smaller then $\Delta_{\text{eff}} = 2.7$. Therefore the changes did not particularly increased the quality of the results.

6 Discussion

In this work, we began by studying the observable $O = \sigma_x^0$ in the eigenbasis of a mixed field Ising chain in the presence of two symmetry braking defect terms.
The non-integrability of the Hamiltonian H was verified through identifying the distribution of energy level-spacings with a Wigner-Dyson distribution. We then proceeded by studying the observable matrix in the energy eigenbasis, with respect to its compliance with the ETH using established methods. We found, that for a energy window almost at the centre of the spectrum, the observable is in good agreement with the predictions of the ETH, which confirms previous works [19, 28]. Also conventional expectations are met, that for local observables and a non-integrable system, the ETH is satisfied. The found agreement with the ETH comprises, that the variance of the diagonal elements decreases exponentially with system size, the off-diagonal elements follow a Gaussian distribution and correlations between off-diagonal matrix elements vanish for increasingly narrower energy window widths.
Main emphasis was laid on the correlations between off-diagonal matrix elements, which we studied through the relation of moments indicator. In agreement with previous works [19, 28], we found that the energy window width, at which correlations between off-diagonal matrix elements vanish, is significantly smaller then for other indicators to be in good agreement with the ETH.
The central concept in this work was then introduced in the OBN method and its numerical implementation in the term counting algorithm. This method tries to tackle a main problem of numerics in quantum thermodynamics. It is that, the mathematical predictions on systems rely on the fact that the considered system is large enough, so that statistical phenomenon can manifest. In particular most of these predictions are only expected to be completely fulfilled in the thermodynamic limit. However fully diagonalizing Hamiltonians becomes exponentially computational intensive with increasing system sizes. Spin lattice models are therefore only accessible below a number of 20 spins, using exact diagonalization. So one needs to calculate systems with large number of spins, but is bound to relatively small system sizes by exact diagonalization.
The OBN method tries to approach this problem, here in particular with respect to the relation of moments indicator. The indicator is given through the relation between the second and fourth moment of an observable. If this relation agrees with the predictions of random matrix theory, the series of moments of the given observable agrees up to at least $4-$th order with the series of moments of a random matrix. As we described in Sec. 2.4, the two matrices with the same series of moments have the same probability distribution. The given observable therefore then shows random matrix behaviour, in particular

the off-diagonal elements are uncorrelated.
The central quantity thereby are moments which in essence are just traces over operators which yield a number. Since the result is just a number, the basis in which the calculation is performed does not matter, even though the ETH is a statement over matrix elements in the energy eigenbasis. We therefore chose a basis, which on the one hand does not need exact diagonalization and on the other hand has beneficial properties in the application of calculating traces. In the natural spin operator basis, described in Sec. 4, the entire trace operation contains only products of Pauli matrices, which allows one to use the properties of the Pauli matrices in order to simplify the trace calculation. The trace calculation can then be implemented, using the OBN method, through a term counting algorithm.
However, one needs to approximate the hard-cut energy window used in the exact method, by a polynomial function in order to be able to apply the OBN method. In order to demonstrate the potentials of the OBN method, we first started with showing, that the term counting method in fact yield the same results, as obtained using exact diagonalization. We used a simple parabolic filter of second order and compared the results obtained from the OBN method and the ones from exact diagonalization. We found that both methods yield the exact same results.
We then continued by finding a good approximation for the hard-cut energy window, while keeping the approximation polynomial as simple as possible in order to increase performance of the term counting method. The first intuitive approach of higher order parabolas did not lead to useful results, since the effective energy window the polynomial filter described, can not be controlled in a way useful for the purpose at hand. Therefore, more sophisticated polynomial approximation where considered.
First a Gaussian filter was considered, in that the exponential function was approximated using its series definition up to a sufficient degree. The case of particular interest is the case of small windows, since there the observable is expected to show random matrix behaviour. We found, that the Gaussian filter leads to results very close to the exact results over the full range of Δ_{eff}, while slowly oscillating around them. All in all, the deviations from the exact results are small over the full range and the graph portrays the behaviour of the exact results, therefore the Gaussian filter serves as a good approximation for the hard cut energy window. Note that a crucial part of comparing the Gaussian filter with the exact hard-cut shaped window, is the definition of the effective window width Δ_{eff}, Eq. (77), as well as the effective dimension d_{eff}, which we described in Sec. 5.
Even though we found very good results in particular the case of interest, the calculation of the polynomial approximation of the exponential function demands up to 1000 terms to yield useful results for very narrow windows. In

order to increase the performance of the method, we used a polynomial fit of the Gaussian. In that, a polynomial of given order is optimized, so that it fits a given function shape. Benefit of this approach is, that one can calculate useful values for small window widths with a reasonable polynomial order, however this comes at the cost, that the obtained filter is not necessarily an exact Gaussian. Using this polynomial fit, we found good results down to a window width of $\Delta_{\text{eff}} \approx 1.5$ using a polynomial fit of order 50. Under a window width of $\Delta_{\text{eff}} \approx 1.5$, the polynomial fit failed to replicate the width of the exact Gaussian on the one hand and on the other hand disturbances became an increasingly large problem. These problems sustained even for much higher polynomial orders, with the method we used here. The values for the relation of moments indicator for effective energy window widths under $\Delta_{\text{eff}} \approx 1.5$ started to increase, in opposite to the exact Gaussian.
We then continued with using the polynomial fit method, to directly approximate the hard cut energy window instead of a Gaussian. For the effective window width Δ_{eff}, the width of the window the polynomial was fit to, was used. For the effective dimension d_{eff} we used the number of micro-states in that window. For large window, we found a good agreement of the results for the exact window and the polynomial fit. At an effective window width of $\Delta_{\text{eff}} \approx 14$ however the results obtained with the polynomial fit, started to show oscillations while also falling off slightly faster then for the exact window. Below $\Delta_{\text{eff}} \approx 3$ the polynomial fit runs into similar problems like the Gaussian polynomial fit. Looking at the obtained filter shapes for narrow windows, the central problems of the polynomial fit is that the polynomial fit did not accurately replicated the window width of the exact window, large oscillations outside of the window and too low or too high peaks compared to the goal peak value of 1.
These problems where then tackled, by rescaling the obtained polynomial fit to a peak value of 1, as well as redefining the effective window width. The problems of the oscillating graph and the diverging results below a certain threshold could not be significantly reduced this way. The polynomial fit did not produce filters with an effective window width below $\Delta_{\text{eff}} = 2.7$ and the graph still falls off faster then for the exact window.

7 Outlook

In the following we will discuss problems we encountered and how one could try to solve those problems. Also a few suggestions will be given for further applications of the OBN method.

7.1 System

First of all different aspects of the system considered in this work can be changed. Different operators can be considered as well as bigger system sizes and different Hamiltonians. Thereby one should pay attention that the Ising model considered here, significantly simplifies the method, since it only consists of two types of Pauli operators. When considering more complex Hamiltonians which consists of all kinds of Pauli matrices, all commutation rules need to be considered.

7.2 Computation

In this work the counting method was actually outperformed by the classical method. For that reason the results shown in Sec. 5 for larger system sizes were generated using exact diagonalization. Reason for the low performance of the counting method compared to the classical method was, that the numerics used here are written in python. For the computations of the classical methods, libraries can be used, that are heavily optimized. Python libraries like python are particularly optimized to handle and diagonalize matrices, while also operating in C. A object oriented framework entirely written in python was used for the counting method. We tried to increase the algorithms used to a significant degree, however the use of external libraries was not efficiently possible since the algorithm handled self defined objects which are foreign to libraries. Since python is a interpreted high-level programming language the performance suffered from this approach. Choosing a different approach for the implementation, for example using more fundamental data types, so that external libraries can be used in order to up-scale the program, should lead to a significant performance increase, making system sizes accessible beyond exact diagonaliation.
It should be noted, that calculating system sizes larger then those which are accessible by diagonalization, one still needs the boundaries of the energy spectrum for the rescaling of the Hamiltonian. However, one could extrapolate the maximum and minimum energy of the considered Hamiltonian from those system sizes that are accessible with exact diagonalization.

7.3 Counting method

When calculating the relation of moments indicator with the counting method, a Gaussian shape was used for the polynomial filter $\mathcal{P}$. In order to calculate the moments with the OBN method, the exponential function was approximated through its Taylor series, so that it takes the form of a polynomial. Thereby the approximation depth that is necessary, so that the approximation is accurate over the full spectrum of the Hamiltonian depends on the boundaries of the spectrum E_{max} and E_{min} and the variance of the Gaussian σ^2. The smaller the variance σ^2, and therefore the narrower the effective energy window Δ_{eff}, the more terms are needed, so that the approximation is accurate over the full spectrum. This leads to a problem, because small effective energy windows are in particular, the case of interest. For the results presented in Sec. 5, we calculated up to a sufficient approximation depth however this led to a decrease in performance.
We took the exponential function as basis in order to find a polynomial which is very similar to a Gaussian but behaves better for a lower amount of terms. This is in theory always possible, one can fix a limit for the amount of terms in the polynomial, for example up to H^{10} and then optimize the set of parameters, so that the function shows the desired shape. This process also becomes increasingly intensive for higher powers, however this can be seen as preliminary work and does not directly impair the performance of the moment calculation itself. The results that where obtained here did not show a significantly better behaviour for small energy windows, since the algorithm struggled to accurately recreate Gaussian with very small variances. One of the reasons for that could have been, the chosen algorithm which was used in order to construct such a polynomial. Using an algorithm that is better suited for the purpose at hand, could however lead to better results.

Next to the Gaussian filter, there are a lot of different functions that can be used to get even closer to the hard cut energy window. In particular the Stone-Weierstrass theorem states, that all continuous functions defined on a closed interval can be approximated by a polynomial as closed as desired. This can be done through Taylor expanding the desired function. Even though the Taylor series maybe includes an infinite sum, the series can always be cut off when the approximation extends over the full interval defined by the boundaries of the energy spectrum of the considered Hamiltonian $[E_{min}, E_{max}]$.
In particular, one could also directly approximate the exact energy window. Overlapping two Heaviside step functions $H(x)$ which are defined by

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 1, \end{cases} \tag{81}$$

one can write a hard cut energy window $\Delta_E(E)$ as a function

$$\Delta_E(E) = H[E - (E_0 - \frac{\Delta}{2})] \cdot H[-E + (E_0 + \frac{\Delta}{2})], \tag{82}$$

where E_0 is the centre and Δ is the width of the energy window. Inserting the smooth approximation of the step function

$$H(x) = \lim_{k\to\infty} \frac{1}{2} + \frac{1}{2} tanh(kx) = \lim_{k\to\infty} \frac{1}{1 + e^{-2kx}} \tag{83}$$

allows to write Eq. (82) as a Taylor series and therefore as a polynomial. With the parameter k the slope of the edges can be controlled while $\lim_{k\to\infty}$ corresponds to a step function. With Eq. (83) and (82), the hard cut energy window can then be written as

$$\Delta_E(E) = \lim_{k\to\infty} \frac{1}{1 + e^{-2k[-E+(E_0+\frac{\Delta}{2})]} + e^{-2k[E-(E_0-\frac{\Delta}{2})]} + e^{-2k\Delta}}. \tag{84}$$

7.4 Possible further applications of the OBN method

Next to calculating moments of observables through counting terms, we see further possible applications for the OBN method. In this section we will present some of these ideas in order to show potentials and further paths for the OBN method.

7.4.1 Algorithmization of analytical derivation

The derivation of an analytical expression for a central moment in the OBN method can be formalized. For simplicity reasons, consider the approximated projection operator with the parabolic filter, Eq. (63) and an Ising model Hamiltonian, with a single field in z-direction

$$H = \sum_l J\sigma_l^z \sigma_{l+1}^z + h_z \sigma_z \tag{85}$$

where J is the interaction strength and h_z is the magnetic field strength. In order to derive an analytical expression for the second central moment M_2 of an observable $O = \sigma_z^0$, we have to find all distinct terms of O_T

$$O_T = POP = (1 - \alpha H^2)O(1 - \alpha H^2). \tag{86}$$

First consider the squared Hamiltonian

$$\begin{aligned} H^2 =&(\sum_{l=0}^{L} \sigma_z^l \sigma_z^{l+1} + h_z \sigma_z^l)(\sum_{m=0}^{L} \sigma_z^m \sigma_z^{m+1} + h_z \sigma_z^m) \\ =& \sum_{m,l=0}^{L} [\sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1} + h_x \sigma_z^l \sigma_z^{l+1} \sigma_z^m + h_x \sigma_z^l \sigma_z^m \sigma_z^{m+1} + h_z^2 \sigma_z^l \sigma_z^m]. \end{aligned} \tag{87}$$

We will number the terms in H^2 with small Roman letters, so that the first term is i the second ii and so on. Now execute the products in Eq. (86)

$$POP = O - \alpha H^2 O - \alpha O H^2 + \alpha^2 H^2 O H^2, \tag{88}$$

where we again number the terms with Roman letters. Here we will start at the second term, since the first term is trivial, with upper Roman letters, so that $-\alpha H^2 O$ is the term I, $-\alpha O H^2$ is term II and so on.
If one now insert Eq. (87) into Eq. (88) and executes all the products, one gets a number of different sums, where we can refer to each one with a combination of Roman letters. We will now continue, by deriving the analytical expression for M_2 for only the first non-trivial sum which is Ii

$$'Ii' = \alpha[\sum_{l,m=0}^{L} \sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1}] O. \tag{89}$$

First only consider the sum within the brackets. Even though the sum runs over all values of l and m from 0 up to L, one can identify special index combinations, where we can apply the algebraic relations of the Pauli matrices. Since we have two Pauli matrix pairs, which always apply to the sites next to each other, there are three special index combinations. The case in which the pairs both act on the same two sites $l = m$, and where only one of each pair act on the same site $l + 1 = m$ and $l = m + 1$. We can now split the sum off into four different sums, where in three of those sums a special index combination applies, while the last one contains the remaining terms

$$\begin{aligned} 'Ii' =& \alpha \Bigg[\sum_{l,m=0;l=m}^{L} 1 + \sum_{l,m=0;l+1=m}^{L} \sigma_z^l \sigma_z^{l+2} + \sum_{l,m=0;l=m+1}^{L} \sigma_z^{l-1} \sigma_z^{l+1} \\ &+ \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1}^{L} \sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1} \Bigg] O. \end{aligned} \tag{90}$$

Note, that the sums in which the special index combinations hold, one sum collapses and it is only a summation over one index, in this case l. Through

index shifting the third term, one can merge the second and third term, and the fourth term is symmetrical, therefore one can apply the condition $m > l+1$ and multiply the term by 2

$$\begin{aligned} 'Ii' =&\alpha\Bigg[\sum_{l=0}^{L} 1 + 2\sum_{l=0}^{L} \sigma_z^l \sigma_z^{l+2} \\ &+ 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1}^{L} \sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1}\Bigg] O. \end{aligned} \tag{91}$$

The product with the observable O can now be performed and $O = \sigma_z^0$ can be inserted, leading to

$$'Ii' = \alpha\Bigg[L\sigma_z^0 + 2\sum_{l=0}^{L} \sigma_z^l \sigma_z^{l+2} \sigma_z^0 + 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1}^{L} \sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1} \sigma_z^0\Bigg] \tag{92}$$

where we executed the first sum which just yields L. Again we are now at a similar situation as with Eq. (89), however this time the observable has a fixed site it is acting on. One can now go over both sums and find every special index combination, split the sum into new sums in which these condition applies and one in which they not apply. Continuing with this procedure leads to

$$\begin{aligned} 'Ii' =&\alpha\Bigg[L\sigma_z^0 + 2 \sum_{l=0;l,l+2\neq 0}^{L} \sigma_z^l \sigma_z^{l+2} \sigma_z^0 + 2\sigma_z^2 + 2\sigma_z^{L-1} \\ &+ 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1}^{L} \sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1} \sigma_z^0 \\ &+ 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1,l=0}^{L} \sigma_z^1 \sigma_z^m \sigma_z^{m+1} \\ &+ 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1,l+1=0}^{L} \sigma_z^L \sigma_z^m \sigma_z^{m+1} \\ &+ 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1,m+1=0}^{L} \sigma_z^l \sigma_z^{l+1} \sigma_z^L\Bigg]. \end{aligned} \tag{93}$$

Note that the case in which $m = 0$ is not possible since $l + 1 < m$, however $m + 1 = 0$ is possible in the case of $m = L$ because of periodic boundary conditions. The last two sums produce the same terms. Consider the last sum, in which $m = L$ for all terms and also the condition $l \neq m$ and $l+1 \neq m$ applies for all terms. Therefore σ_z^L can be interchange positions with all Pauli

operators in the sum without the need of commutation rules. An analogous reasoning can be applied to the second last sum. Therefore the final form of the sum is given by

$$\begin{aligned} {}'Ii' =& \alpha \Bigg[L\sigma_z^0 + 2 \sum_{l=0;l,l+2\neq 0}^{L} \sigma_z^l \sigma_z^{l+2} \sigma_z^0 + 2\sigma_z^2 + 2\sigma_z^{L-1} \\ &+ 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1}^{L} \sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1} \sigma_z^0 \\ &+ 2 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1,l=0}^{L} \sigma_z^1 \sigma_z^m \sigma_z^{m+1} \\ &+ 4 \sum_{l,m=0;l\neq m,l+1\neq m,l\neq m+1,m>l+1,l+1=0}^{L} \sigma_z^L \sigma_z^m \sigma_z^{m+1} \Bigg]. \end{aligned} \tag{94}$$

All terms in Eq. (94) are unique and cant be reduced further, while the coefficient of every term is known. The last step in order to calculate the second moment M_2 for the term Ii, is to derive the complexity of the sums in Eq. (94), that is how many individual terms the sum produces when carried out.

The trivial case can be seen in the first term $L\sigma_z^0$. This term is singular, therefore its complexity is 1 and the term contributes a value of $L \cdot 1$. The first non-trivial case, is the term

$$2 \sum_{l=0;l,l+2\neq 0}^{L} \sigma_z^l \sigma_z^{l+2} \sigma_z^0 \tag{95}$$

This is a single index sum l which runs over the interval $[0, L]$. However the conditions lead to two indices-values being skipped, namely $l = 0$ and $l+2 = 0$. Therefore the complexity of the sum is given by $L-2$ and the sum contributes $2 \cdot (L-2)$. This procedure can be carried on for all sums resulting in a polynomial of L which describes the coefficient of the parameter α.

However the second moment of the term Ii is not really a quantity of use, one would need to apply this procedure for all terms in Eq. (88) in order to get the analytical expression for $M_2(POP)$. For the terms I and II this calculations can be carried out by hand, however for the last term III the number of sums and conditions become very difficult to deal with by hand, at latest when considering the fourth moment $M_4 = M_2[(POP^2)]$ this method becomes impossible to be carried out by hand.

The calculation however does not necessarily become more complicated, it just becomes more complex. The procedure stays almost the same as in the case for the term Ii. Step 1 is, carrying out all products until one obtains a series

of sums. Step 2 is to then identify all special index combinations in that one can apply the Pauli matrix commutation rules. The sum can then be split into a set of new sums in which each of the special index combinations applies, and a sum with the remaining terms. This step is repeated for all sums until this no longer possible for any sums. In step 3 the complexity of every sum is determined leading to a polynomial function of α and L. This process can be written as a algorithm and be carried out numerically. We will therefore suggest a rough outline of a basic version of such a algorithm.

For this application a object oriented approach is advantageous. The two main building blocks in the program are sums and Paulis. A sum object consists of four parts, the indices it sums over, a set of conditions on the indices, a coefficient and the Paulis it holds. Pauli objects are not directly pauli operators, but they are modelling them for this specific case. There are three different kinds of Paulis which we will in the following refer to as

1. bound pairs : $\sigma_z^l \sigma_z^{l+1}$
2. free : σ_z^l
3. locked : σ_z^q

where l is a variable index and q is a constant number. Each Pauli object holds its kind, the site it acts on and the kind of its Pauli operator. One also need list objects which hold the sums and are able to execute products of sums, analogous to the product object described in the counting method in Sec. 4.1
One would now start with modelling Eq. (86) in terms of the objects described above. Step 1 now only consists of executing all products until one ends up with a series of sums.
Step 2 is the core of the algorithm, and it revolves around how different kinds of Pauli objects interact with each other when identifying special index combinations. Every kind of Pauli object has a fixed set of 'interactions' with other Pauli objects in a sum. These 'interactions' consist of two Pauli objects interacting resulting into a special index combination and a resulting set of Pauli objects. In Fig. 20 all possible interactions of Pauli objects are shown schematically.

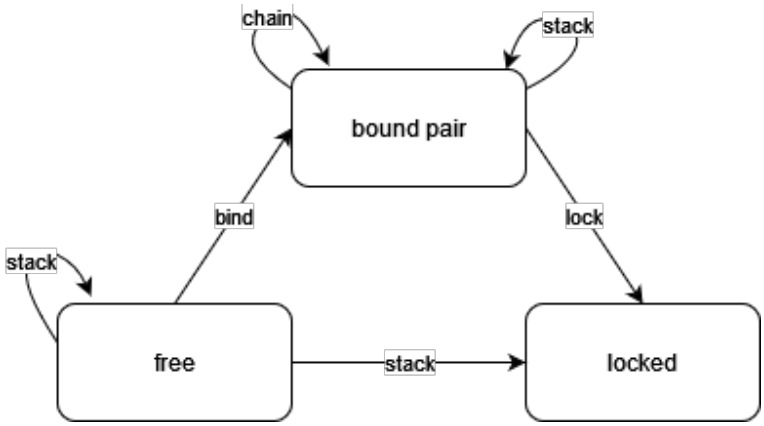

Figure 20: Scheme of all interactions between the three Pauli objects, in the derivation of analytical expression algorithm.

The 4 interactions stack, chain, bind and lock are given by

- stack: (bound pair) + (bound pair) $\rightarrow$ 1 — $\sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1} \xrightarrow{l=m} 1$
- chain:(bound pair) + (bound pair) $\rightarrow$ (bound pair) — $\sigma_z^l \sigma_z^{l+1} \sigma_z^m \sigma_z^{m+1} \xrightarrow{l+1=m} \sigma_z^l \sigma_z^{l+1}$
- bind: (bound pair) + (free) $\rightarrow$ (free) — $\sigma_z^m \sigma_z^{m+1} \sigma_z^l \xrightarrow{l=m} \sigma_z^{l+1}$
- lock: (bound pair) + (locked) $\rightarrow$ (locked) —$\sigma_z^l \sigma_z^{l+1} \sigma_z^q \xrightarrow{l=q} \sigma_z^{q+1}$.

Every arrow in the Fig. 20 defines exactly one interaction of two Pauli objects with their index combination. Because this is a fixed set, step 2 can now be carried out by choosing a sum and trying every possible interactions with the Paulis inside of the sum until no interactions are possible. One interaction is executed by taking two Paulis in the sum, chosing an interaction and then compare the implied index combination of that interaction with the conditions of the sum object. If it is possible the sum is split in two new sums. One in which the original Paulis are in, with the negated index combination added to the conditions to the sum and one in which the index combination is added to the conditions to the sum with the new Paulis, resulting from the interaction. This process is repeated until no interactions are possible in any sums.
In step 3 the complexity of all sum objects needs to be calculated. The complexity of a sum, without any conditions applied to it, is given by L to the power of the number of distinct indices the sum runs over. The conditions now lead to some terms being skipped and therefore the complexity reduces. We will discuss how one can calculate this value in general by an example. Consider the sum

$$\sum_{l,m=0; l\neq m, l\neq q, m\neq q}^{L} \sigma_z^l \sigma_z^m \sigma_z^q, \tag{96}$$

with $q \in [0, L]$. First of all, note that if a condition applies to two indices, this can be treated out of the perspective of just one. For example for the condition of $l \neq m$, this can be seen as a condition only on m, so that l is always fixed first and then m has the condition of not taking that value.
Beginning with the complexity C, for the term above withou considering the conditions, being given by

$$C = L \cdot L. \tag{97}$$

Applying the first condition $l \neq m$ will take one possible value for m, for every value of l, therefore the complexity yields

$$C = L \cdot (L - 1). \tag{98}$$

The next two conditions demand, that both indices don't take the constant value q, namely $l \neq q$ and $m \neq q$. This takes one possible value from both l and m, which leads to

$$\begin{aligned} C =& (L-1) \cdot (L-2) \\ =& L^2 - 3L + 2. \end{aligned} \tag{99}$$

Therefore one can calculate the complexity of an arbitrary sum from its indices and the conditions on these indices following the above scheme. This is based on the fact, that the terms in the sum do not vanish on their one, which is true for all sums generated from step 2.
In the end their is one very important step, step 4 and that is to identify sums that produce the same terms which was done on the fly before. In general two sums are equal if they hold the same set of Paulis and if their conditions are not mutually exclusive. Here it is important, that a index shift may be considered to identify two set of Paulis as equal, as in Eq. (91). Also one needs to consider sums that are symmetrical when summing over multiple indices, which can be resolved like in Eq. (91), since sums result in equal terms in themselves.
Aside from these special cases, if one finds two equal sums, one can drop one of the sums and add its coefficient to the other sum. The second moment is now given through the sum over the coefficients times the complexity of every sum. therefore one would obtain a polynomial M_2^{POP} in α where every coefficient is itself a polynomial in L

$$M_2^{POP}(\alpha, L) = A(L) + \alpha A^*(L) + \alpha^2 A^{**}(L) + \ldots, \tag{100}$$

where $A(L)$ refer to a polynomial function of L. Note that the asterisks in the superscript of A aims to signal that these are different functions.
One would be able to calculate the relation of moments indicator or other applications of the OBN method, for almost arbitrary system sizes since the result is just given by an analytical function. Also, this calculation just needs to be done once for a particular set of filter shape, observable and Hamiltonian. Also there are parallels to the counting method described in Sec. 4.1, however this method is a lot more sophisticated and powerful, if it can be implemented. Note that the algorithm described here is just a very basic version which would work for the specific example described above. For the considered system, Hamiltonian and observable only consist of a single kind of Pauli matrix, therefore one does not need to consider the case of commutation and the accompanying change in sign of the coefficient. For the system considered in the main part of this work, a mixed field Ising model Hamiltonian, one would need to also cover these cases. However, the goal here is to show the basic idea of how such a algorithm could look like, as well as giving specific suggestions of what framework and methods could be used.

7.4.2 Calculating generic traces

As described in Sec. 4 the OBN method was introduced in order to calculate even moments of operators. This is because, in the method we make use of the fact, that Pauli operators are traceless and from generating the full basis of an operator, we get all terms that don't vanish after squaring the operator and taking the trace. Squaring the operator is therefore a central step in this method.
Consider a operator A which is given by Pauli operator. With the OBN method one could now calculate even powers of the operator A. However, often the most interesting case is the first moment, which is the expectation value of that observable. This is especially interesting since it is the value that one would measure in a experimental context, when measuring the observable in the considered system. Through defining a root operator $A^{\frac{1}{2}}$, the calculation of the expectation value of A through OBN method becomes possible. The root operator is defined by

$$(A^{\frac{1}{2}})^2 = A, \tag{101}$$

therefore the second moment of the root operator is the expectation value of the operator A

$$M_2(A^{\frac{1}{2}}) = tr\{(A^{\frac{1}{2}})^2\} = tr\{A\} = \langle A \rangle \,. \tag{102}$$

Note that we, without loss of generality, considered the raw operator A, however the same considerations apply for the projected operator.
One would now first need to construct such a root operator for the desired observable A, and also this root operator needs to be written only in terms of a sum of pauli matrices. The power series definition of a root of an operator

$$A^{\frac{1}{2}} = \sum_{n=0}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} (\mathcal{I} - A)^n, \tag{103}$$

where $\mathcal{I}$ is the identity, would accomplish both these things, under the condition, that $\limsup_n ||(I - A)^n||^{\frac{1}{n}} < 1$.
Given an operator A which can be written as a sum of Pauli matrices, one could use the OBN method in order to calculate its expectation value through the second moment of its root operator. Note that one could also calculate higher odd moments of the operator A with this procedure. For a lot of operators, for example the one considered in this work, the expectation value as well as odd moments are trivial since they are zero. However, this method could be used in order to calculate generic traces of operators including more complex operators which can be written in terms of Pauli operators, where the expectation value in general is not zero. Also one could calculate autocorrelation functions $C(t)$, or even time evolution of operators in general $\langle A \rangle_t$ using this method, since the time evolution operator $U(t)$ can always be written in terms of Pauli operators for a spin chain system, by approximating the exponential function using its Taylor series.

References

[1] E. Schrödinger, *Annalen der physik 388, 956* (1927) (cited on page 1).

[2] J. von Neumann, *Proof of the ergodic theorem and the h-theorem in quantum mechanics* (The European Physical Journal H 35, 201, 2010) (cited on page 1).

[3] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, *Colloquium: nonequilibrium dynamics of closed interacting quantum systems* (Rev. Mod. Phys. 83, 863, 2011) (cited on page 1).

[4] C. Gogolin and J. Eisert, *Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems* (Rep. Prog. Phys. 79 056001, 2016) (cited on page 1).

[5] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Review; from quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics* (Advances in PhysicsVol. 65, No. 3, 239-362, 2016) (cited on pages 1–2, 21).

[6] L. F. Santos and M. Rigol, *Localization and the effects of symmetries in the thermalization properties of one-dimensional quantum systems* (Phys. Rev. E 82, 031130, 2010) (cited on page 2).

[7] W. Beugeling, R. Moessner, and M. Haquel, *Finite-size scaling of eigenstate thermalization* (Phys. Rev. E 89, 042112, 2014) (cited on page 2).

[8] W. Beugeling, R. Moessner, and M. Haquel, *Off-diagonal matrix elements of local operators in many-body quantum systems* (Phys. Rev. E 91, 012144, 2015) (cited on page 2).

[9] H. Kim, T. N. Ikeda, and D. A. Husel, *Testing whether all eigenstates obey the eigenstate thermalization hypothesis* (Phys. Rev. E 90, 052105, 2014) (cited on page 2).

[10] R. Mondaini, K. R. Fratus, M. Srednicki, and M. Rigol, *Eigenstate thermalization in the two-dimensional transverse field ising model* (Phys. Rev. E 93, 032104, 2016) (cited on page 2).

[11] R. Mondaini and M. Rigol, *Eigenstate thermalization in the two-dimensional transverse field ising model. ii. off-diagonal matrix elements of observables* (Phys. Rev. E 96, 012157, 2017) (cited on page 2).

[12] D. Jansen, J. Stolpp, L. Vidmar, and F. Heidrich-Meisner, *Eigenstate thermalization and quantum chaos in the holstein polaron model* (Phys. Rev. B 99, 155130, 2019) (cited on page 2).

[13] R. Steinigeweg, J. Herbrych, and P. Prelovšek, *Eigenstate thermalization within isolated spin-chain systems* (Phys. Rev. E 87, 012118, 2013) (cited on page 2).

[14] T. LeBlond, K. Mallayya, L. Vidmar, and M. Rigol, *Entanglement and matrix elements of observables in interacting integrable systems* (Phys. Rev. E 100, 062134, 2019) (cited on pages 2, 22).

[15] Y. Alhassid and R. D. Levine, *Transition-strength fluctuations and the onset of chaotic motion* (Phys. Rev. Lett. 57, 2879, 1986) (cited on page 2).

[16] F. H. L. Essler and M. Fagotti, *Quench dynamics and relaxation in isolated integrable quantum spin chains* (J. Stat. Mech. 064002, 2016) (cited on page 2).

[17] R. Nandkishore and D. A. Huse, *Many-body localization and thermalization in quantum statistical mechanics* (Annu. Rev. Condens. Matter Phsy. 6, 15, 2015) (cited on page 2).

[18] N. Shiraishi and T. Mori, *Systematic construction of counterexamples to the eigenstate thermalization hypothesis* (Phys. Rev. Lett. 119, 030601, 2017) (cited on page 2).

[19] J. Richter, A. Dymarsky, R. Steinigeweg, and J. Gemmer, *Eigenstate thermalization hypothesis beyond standard indicators: emergence of random-matrix behavior at small frequencies* (Phys. Rev. E 102, 042127, 2020) (cited on pages 2, 51).

[20] J. M. Deutsch, *Quantum statistical mechanics in a closed system* (Phys. Rev. A 43, 2046, 1991) (cited on page 6).

[21] M. Srednicki, *Chaos and quantum thermalization* (Phys. Rev. E 50, 888, 1994) (cited on page 6).

[22] J. M. Deutsch, *Eigenstate thermalization hypothesis* (arXiv:1805.01616, 2018) (cited on page 7).

[23] L. Foini and J. Kurchan, *Eigenstate thermalization hypothesis and out of time order correlators* (Phys. Rev. E 99, 042139, 2019) (cited on page 8).

[24] R. Steinigewg, *Lecture: numerical physics of condensed matter* (University of Osnabrück, 2021) (cited on page 12).

[25] J. A. Scaramazza, B. S. Shastry, and E. A. Yuzbashyan, *Integrable matrix theory: level statistics* (PHYSICAL REVIEW E 94, 032106, 2016) (cited on page 12).

[26] H Bethe, *Zur theorie der metalle. i. eigenwerte und eigenfunktionen der linearen atomkette* (Zeitschrift für Physik. 71 (3–4): 205–226, 1931) (cited on page 13).

[27] F Hausdorff, *Summationsmethoden und momentfolgen. i. und ii.* (Mathematische Zeitschrift 9, 74–109 and 280–299, 1921) (cited on page 14).

[28] J. Wang, M. H. Lamann, J. Richter, R. Steinigeweg, A. Dymarsky, and J. Gemmer, *Eigenstate thermalization hypothesis and its deviations from random-matrix theory beyond the thermalization time* (Phy. Rev. Letters, 2020) (cited on pages 15, 18, 23, 51).

[29] B. A. Cipra, *An introduction to the ising model* (The American Mathematical Monthly, Vol. 94, No. 10, pp. 937-959, 1987) (cited on page 18).

[30] A. A. Abul-Magd and A. Y. Abul-Magd, *Unfolding of the spectrum for chaotic and mixed systems* (Physica A 396 185-194, 2014) (cited on page 20).

[31] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, *Distribution of the ratio of consecutive level spacings in random matrix ensembles* (Phys. Rev. Lett. 110, 084101, 2013) (cited on page 20).

[32] M. Rigol, V. Dunkjo, and M. Olshanii, *Thermalization and its mechanism for generic isolated quantum systems* (Nature 452, 854-858, 2008).

[33] G. Livan, M. Novaes, and P. Vivo, *Introduction to random matrices, theory and practice* (https://arxiv.org/abs/1712.07903, 2017).

[34] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes* (Fortschritte der Physik 61, 781, 2013).

[35] S. Khlebnikov and M. Kruczenski, *Locality, entanglement, and thermalization of isolated quantum systems* (Physical Review E 90, 050101, 2014).

[36] M. Brenes, T. LeBlond, J. Goold, and M. Rigol, *Eigenstate thermalization in a locally perturbed integrable system* (Phys. Rev. Lett. 125, 070605, 2020).

[37] M. Brenes, J. Goold, and M. Rigol, *Low-frequency behavior of off-diagonal matrix elements in the integrable xxz chain and in a locally perturbed quantum-chaotic xxz chain* (Phys. Rev. B 102, 075127, 2020).