

Numerics and Statistics of the Local Density of States

Diogo Pinheiro

Mestrado em Física

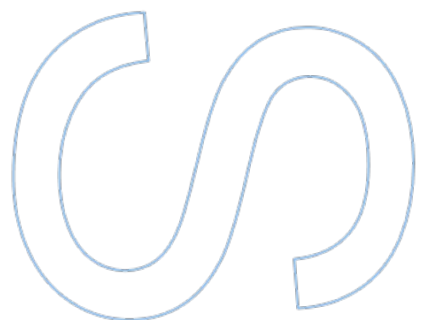
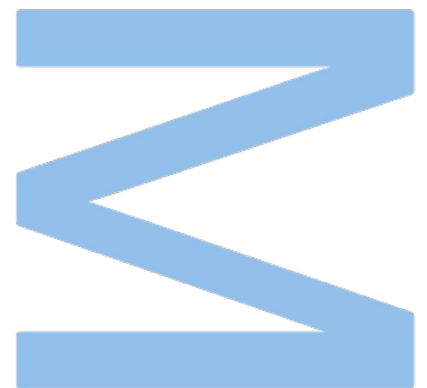
Departamento de Física e Astronomia
2023

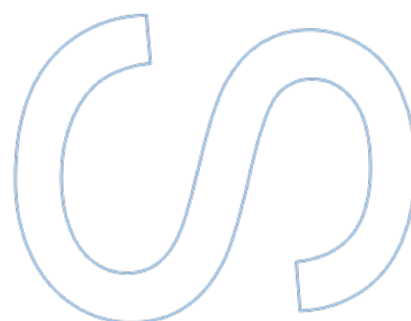
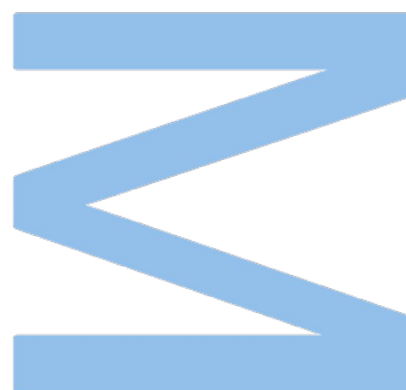
Orientador

João Manuel Parente Viana Lopes, Professor Auxiliar, Faculdade
de Ciências da Universidade do Porto

Coorientador

João Pedro Santos Pires, Consultant at Deloitte Risk Advisory





Life before death, strength before weakness, journey before destination.

-Brandon Sanderson

[This page was intentionally left blank]

Acknowledgments

Before we begin there is some due acknowledgments. This has been a hard five year journey that would be impossible without the support and knowledge of the people I met along, and, for that, I would like to express my utmost gratitude.

To my supervisor, Prof. João Viana Lopes, for his unwavering availability to answer my questions and providing essential knowledge to solve the scientific challenges that arose along the way and pushed me to ever greater heights, but who also likes to daily remind me that, due to the city i was born, disliking ham is bizarre.

To my co-supervisor, Dr. João Pedro Pires, for also being always available to discuss and provide insightful critiques concerning our research, which it was born from his prior contributions, but who also is a great friend.

To my parents, Maria Tavares and José Pinheiro, despite not understanding a single word about the work I do, for their undying support.

To João “O Calvo” Teixeira and Helena Silva, for being the greatest friends I ever made on this journey and whose ears are sore for hearing me complain. For their unending support, moments of uncontrollable laughter, and, most importantly, for never letting me collapse on the hardest moments.

To Júlio Oliveira and Nicolau Sobrosa, whom I met in the beginning of the course and never departed. For their friendship and fruitful conversations that helped me on my studies and research.

And a special thanks to Patrícia, Luís, Raul, Manecas, Titi, David, Ana Cacho, João Silva, Ricardo, Henrique, Cristina, Adão, Rocha, Gonçalo, Inês, Afonso and Sofia for the precious moments we had.

And, of course, my cat, Shirley.

[This page was intentionally left blank]

Abstract

Ever since Anderson’s seminal work [1], the study of localization properties in quantum systems has been recognized to be best performed through a statistical analysis of local physical quantities. The Local Density of States (LDoS) is a perfect candidate for it is one of the simplest observable one can build and, furthermore, it can be measured experimentally with Scanning Tunneling Microscope (STM) techniques [2]. In 1D and quasi-1D/2D systems, earlier analytical studies have been conducted by Altshuler et. al. [3] and Mirlin et. al. [4], respectively. In general, it was widely shown that the distribution of the LDoS follows a Log-Normal when disorder is present. At the Anderson transition from Metal-to-Anderson Insulator, the distribution undergoes a distinct change from Normal to a Log-Normal. On the numerical side, Schubert et. al. [5] performed calculations of this distribution with the Kernel Polynomial Method (KPM) technique [6] and also sees the emergence of this change of the distributions on the Anderson transition and that it can be used to identify the presence of extended/localized regimes.

This thesis is devoted to the investigation of the role of disorder and Anderson localization in the statistical properties of the local density of states of lattice systems. In order to perform this study, a new set of numerical methods are introduced allowing the evaluation of the LDoS with an efficiency hitherto unseen in the literature. The Kernel Polynomial Method is a numerical method with linear complexity in the size of the system, i.e. $\mathcal{O}(L)$. The study of these distributions requires the accumulation of many LDoS points throughout the lattice, and thus, the effective computational complexity of the numerical method is elevated to $\mathcal{O}(L^2)$. Given this, we present a new method to circumvent this problem inspired by the ideas behind the stochastic computation of traces of operators [6]. We effectively create a novel method to compute the LDoS across the entire lattice in linear time, $\mathcal{O}(L)$.

Moreover, recent advance in the use of Graphics Processing Units (GPUs) with the Kernel Polynomial Method [7] have drawn our attention. We revisit this topic and develop a brand new and original software programmed in CUDA that can take best advantage of the most recent Nvidia GPUs. With this we are able to personalize it to maximize the performance, resulting in a software capable of running the KPM hundreds of times faster than a CPU-based implementation.

By combining our new linear-scaling method for computing LDoS across the lattice with GPU-accelerated software, we are able to obtain the statistics of the LDoS in unprecedented scale. With this, we examined the one-dimensional systems and employing the findings of Altshuler et al. [3] as a benchmark for a comprehensive study. Additionally, we explored the interplay among the parameters governing the simulation of the one-dimensional system and established the presence of an effective length and a crucial parameter for ensuring the robust convergence of the LDoS statistics.

This breakthroughs not only advances the field but also has promising implications for future research. With the ease of obtaining LDoS statistics, we could study other special systems, such as the Weyl Semimetals where the transition from Semimetal-to-Metal is an open problem and begs the use of LDoS distributions to properly comprehend it.

[This page was intentionally left blank]

Resumo

Desde o trabalho seminal de Anderson [1], o estudo das propriedades de localização em sistemas quânticos têm sido reconhecidos como melhor compreendidos por meio de uma análise estatística de quantidades físicas locais. A Densidade Local de Estados (LDoS) é uma candidata perfeita para isso, uma vez que pode ser medida experimentalmente com técnicas de Microscopia de Varrimento por Efeito Túnel (STM) [2]. Em sistemas 1D e quasi-1D/2D, estudos analíticos foram feitos por Altshuler et al. [3] e Mirlin et al. [4], respectivamente. De maneira geral, foi amplamente demonstrado que a distribuição da LDoS segue uma distribuição Log-Normal na presença de desordem. Na transição de Metal para Isolador de Anderson, a distribuição sofre uma mudança distinta de Normal para Log-Normal. Do ponto de vista numérico, Schubert et al. [5] realizaram cálculos dessa distribuição com a técnica do Kernel Polynomial Method (KPM) [6] e observaram também o surgimento dessa mudança nas distribuições na transição de Anderson, que pode ser usada para identificar a presença de regimes estendidos/localizados.

Esta tese é dedicada à investigação do papel da desordem e da localização de Anderson nas propriedades estatísticas da LDoS de sistemas em redes discretas. Para realizar este estudo, um novo conjunto de métodos numéricos é introduzido, permitindo a avaliação da LDoS com uma eficiência até então nunca vista na literatura. O Kernel Polynomial Method é um método numérico com complexidade linear, $\mathcal{O}(L)$. O estudo dessas distribuições requer a acumulação de muitos pontos de LDoS ao longo da rede, e, portanto, a complexidade efetiva do método numérico é elevada para $\mathcal{O}(L^2)$. Dado isso, apresentamos um novo método para contornar esse problema, inspirado nas ideias por trás do cálculo estocástico de traços de operadores [6]. Efetivamente, criamos um método inovador para calcular a LDoS em toda a rede em tempo linear, $\mathcal{O}(L)$.

Além disso, avanços recentes no uso de Unidades de Processamento Gráfico (GPUs) com o Kernel Polynomial Method [7] chamaram a nossa atenção. Revisitamos esse tópico e desenvolvemos um novo e original software programado em CUDA que pode tirar o melhor proveito das GPUs Nvidia mais recentes. Com isso, somos capazes de personalizá-lo para maximizar o desempenho, resultando em um software capaz de executar o KPM centenas de vezes mais rápido do que uma implementação baseada em CPU.

Ao combinar o nosso novo método de escalonamento linear para calcular a LDoS em toda a rede com um software acelerado por GPU, somos capazes de obter estatísticas da LDoS em uma escala sem precedentes. Com isso, examinamos os sistemas unidimensionais e utilizamos os resultados de Altshuler et al. [3] como referência para um estudo abrangente. Além disso, exploramos a interação entre os parâmetros que governam a simulação do sistema unidimensional e estabelecemos a presença de um comprimento efetivo e um parâmetro crucial para garantir a convergência robusta das estatísticas da LDoS.

Essas descobertas não apenas avançam o campo, mas também têm implicações promissoras para pesquisas futuras. Com a facilidade de obtenção de estatísticas da LDoS, poderíamos estudar outros sistemas especiais, como os Semimetais de Weyl, onde a transição de Semimetal para Metal é um problema em aberto e exige o uso de distribuições de LDoS para compreendê-lo adequadamente.

[This page was intentionally left blank]

Contents

Acknowledgements	iii
Abstract	v
Resumo	vii
List of Figures	xiv
1 Localization in Disordered Systems	1
1.1 Localization	1
1.2 Identifying Localization	2
1.2.1 Decay of the Wavefunction	2
1.2.2 Inverse Participation Ratio	3
1.2.3 Statistics of Local Density of States	4
1.3 Thesis Overview	10
2 Spectral Methods	11
2.1 Chebyshev Polynomials	11
2.2 Expansions of Common Functions	13
2.3 Kernels	19
2.4 Expansion for Operators	24
2.4.1 Basic Notions	24
2.4.2 A Trick to Compute Moments	26
2.4.3 Error Analysis for the Trace of an Operator	27
2.5 Self-Averaging of the Density of States	29
2.5.1 Error Analysis of the Density of States	29
2.5.2 Example of KPM in the DoS	31
3 Implementation in Graphical Processing Units	33
3.1 Basic of GPU Architecture	33
3.2 The Basic Algorithms for the KPM in GPU	34
3.2.1 Summation of Vectors	34
3.2.2 Reduction Algorithm	35
3.2.3 Memory Optimization	38

3.3	Performance Study	40
4	Computational Methods for the Local Density of States	42
4.1	Computational Complexity of the LDoS with KPM	42
4.2	Stochastic LDoS	43
4.3	Measuring the LDoS through the Error Bar	45
4.4	Importance of the Twisted Boundary Conditions in Statistics	47
4.5	Parameters and Details of the LDoS Statistics	48
5	Results	50
5.1	Local DoS Fluctuations in One-Dimensional Systems	50
5.2	Finite Size Effects	54
5.3	Additional Insights on the Dirac- δ Regularization	59
5.4	Basic Scales	61
6	Conclusions	62
A	Calculation of the Variance of the DoS and LDoS	64
A.1	Density of States	64
A.2	Local Density of States	66
A.3	Correlators	67
	Bibliography	70

List of Figures

1.1	Amplitude of the state with energy E_α closest to 0 for a one-dimensional nearest neighbor tight binding system with size $L = 4096$; Left) $W = 0.01$. Right) $W = 1$.	2
1.2	Upper) IPR (P^{-1}) for a 3D tight-binding nearest neighbor model with periodic boundary conditions averaged over 1000 realizations of disorders. The lines are system sizes, going from black to dark green, $L = 8, 9, 10, 12, 14, 16, 20$. Lower) Probability distribution of the IPR at the band center $E = 0$ (dark line) and near the band edges (blue dashed line). The blue line is a system Taken from [8].	3
1.3	Results taken by [8] by where we have the normalized standard deviation of the IPR, $\Delta P/P$, in the band center $E = 0$ of the 3D tight-binding model as a function of disorder with periodic boundary conditions. The obtained results are averaged of 1000 realizations of disorder.	4
1.4	Evolution of the Density of States at the band center, $\rho(E = 0)$, for several disorders. This shows smooth behavior when we cross the critical disorder $W_c \sim 16.5t$.	5
1.5	Schematic representation of the LDoS. If one assumes a regularization for the $\delta(E - E_\alpha)$ we see that the LDoS is a weighted average of the wavefunctions.	5
1.6	Numerical result showcasing the visual change of the LDoS distribution for a slice of a 3D tight-binding system with $L = 512^3$ and periodic boundary conditions. The Anderson disorder is $W = 10.0$. Right: $E = 0$; Left: $E = 7.69$. The inset is a 50^2 zoom. This change of character of the distribution indicates a phase transition. Taken from [9].	6
1.7	Examples of the probability distribution for the LDoS of a 3D cubic tight-binding lattice for several systems sizes, obtained and taken from [5].	7
1.8	Density of states and Typical density of states for a nearest neighbor tight binding cubic lattice model with lateral size $L = 50$. Showing the vanishing of TDoS for $W \geq 16$. Taken from [8].	8
1.9	Plot of the order parameter $R(E) = \rho_{\text{typ}}(E) / \rho_{\text{av}}(E)$ for a cubic tight binding model, $L = 50^3$, in green versus other methods of studying localization. Triangles-Decay of the wave function. Diamonds-Inverse Participation Ratio. This shows that R works with the same precision for identification of localization. Taken from [8].	9
1.10	Phase diagram of the Anderson model on a 3D cubic lattice tight-binding model. Shown are the mobility edge (solid line), obtained the calibrating the effective order parameter R to the known critical disorder, $W_c \sim 16.5t$ and the Lifshitz boundaries (dashed lines). This is obtained and taken from [8].	9
2.1	Example of Chebyshev Polynomials.	13

2.2	Expansion of a Delta-Dirac $\delta(x)$ with different number of Moments.	14
2.3	Example of complex exponential Chebyshev expansion for $x = 0.5$ and $a = 1$ with 512 moments. Left) Real part Right) Imaginary part.	15
2.4	Example of Green's function Chebyshev expansion for $z = -0.02i$ with 512 moments. Left) Real part Right) Imaginary part.	16
2.5	Difference of representations, Gaussian and Lorentzian, for the $\delta_{\text{reg}}(x, \eta)$ with width $\eta = 0.1$	17
2.6	Left) Example of a Chebyshev expansion of the integrand in Eq. (2.2.20) with width $\eta = 0.01$, $\mu = 0$ and cutting the domain to $[-10\eta, 10\eta]$ for $M = 310$. Right) The same integrand in the full domain $[-1, 1]$. As you can see Right) would require many more points of integration to capture the important features, while on left we can reduce the number of points and have a very good approximation for the value of the integrals.	18
2.7	Gaussian Chebyshev expansion of width $\eta = 0.001$, $\mu = 0$ and $M = 10000$. Left) Magnitude of the moments Right) Reconstruction of the Gaussian. This expansion agrees remarkably well up to 10^{-13}	18
2.8	Lorentzian Chebyshev expansion of width $\eta = 0.001$, $\mu = 0$ and $M = 10000$. Left) Magnitude of the moments Right) Reconstruction of the Gaussian. This expansion agrees remarkably well up to 10^{-13}	19
2.9	Left) Example of a Chebyshev expansion of $\delta(x)$ with and without the Jackson kernel, with the same number of moments, 100 Right) Suppressed positive Gibbs oscillations in a high order Chebyshev expansion of $\delta(x)$	21
2.10	Expansion of a Delta-Dirac $\delta(x - \mu)$ for different μ and $M = 100$. This exemplifies the change in width σ of the overlapping Gaussian, as we approach the boundaries it starts to narrow greatly.	22
2.11	Expansion of a Delta-Dirac $\delta(x)$ with 512 moments. This figure shows the remarkable difference in the tails when we expand and compare directly to its Gaussian as discussed.	24
2.12	Example of a reconstruction of DoS for the clean tight-binding nearest neighbor model for a cubic lattice. There is only a single sample on the random vectors and number of disorders.	32
3.1	Very simple schematic of a GPU architecture and kernel execution. Kernel is just a piece of code that is executed in the GPU [10]. Image taken from [11].	33
3.2	Difference between pieces of code that perform the summation of vectors; Left) CPU; Right) GPU, each thread executes it. Notice the jump in the looping variable, for the CPU it is just a skip of 1 while in the GPU it is $\text{blockDim} * \text{gridDim}$, which is the number of threads. Taken from [10].	34
3.3	Schematics for the parallelization of vector summation in CPUs.	35
3.4	Schematics for the parallelization of vector summation in GPUs.	35
3.5	Schematics for the operation of the dot product between vectors before the reduction. Each thread will accumulate its partial sum P_n while iterating in the vector in chunks of size $\text{blockDim} * \text{gridDim}$. Each color represents a block, which has shared memory, of 4 threads, while in real life GPUs its on the order of $128 - 512$ threads.	36
3.6	Visual representation of thread activation in a block with shared memory while performing the reduction. Taken from [10]	36

3.7	Left: Sane Python code for the dot product. Right: Real CUDA C code showing how to apply the reduction algorithm. Black box: Accumulates the partial sums of the vector to vector elements products; Blue box: Loading to corresponding index of shared memory of the block; Red box: Algorithm collapsing the partial sums into a single number for each block in $\text{blockDim}/2$ loops. Taken from [10].	37
3.8	Visual representation of memory speed between components in a computer. CPU to GPU (RAM to Global Memory) is the slowest, while the access of the GPU to its global memory is of moderate speed. The fastest memory is not represented, since it is the Shared Memory, that is on chip.	39
3.9	Times between different KPM programs. The computation example was of the DoS of 3D tight-binding lattice with 1024 moments.	40
4.1	Top) Absolute error of the LDoS measured through the sample mean. Bottom) Relative error of the LDoS measured through the sample mean. It is 64×64 Square Lattice Tight-Binding nearest neighbor model. The disorder strength is $W = 6.0$ and 30 disorder realizations with 512 random vectors. The light dark filled region is a 3σ error bar, corresponding to a 99% certainty.	45
4.2	Top) Absolute error of the LDoS measured through the sample variance. Bottom) Relative error of the LDoS measured through the sample variance. It is 64×64 Square Lattice Tight-Binding nearest neighbor model. The disorder strenght is $W = 6.0$ and 30 disorder realizations with 512 random vectors.	46
4.3	Distribution of the relative error of the $R(N)$ estimator for the LDoS with several different number of averages. The inset shows the progression of the width Δ of such distributions. The simulated system is a 64×64 Square Lattice Tight-Binding nearest neighbor model with disorder strength $W = 6.0$ and 50 realizations of disorder.	47
4.4	Diagonalization of a 2D tight-binding system with Left) PBC Right) TBC; The plotted gaussian represents the width σ of the kernel chosen, the states that fall in between are the selected.	48
5.1	Example of the behavior shift of the Inverse-Gaussian distribution with $\tau = 1$ and varying σ	51
5.2	Local density of states statistics for a periodic 1D tight-binding chain nearest neighbor of size $L = 512000$ and Anderson disorder $W = 0.3$ with increasing broadening of the Lorentzian. Top) Intermediate and small broadening. Bottom) Large broadening.	52
5.3	Top) Localization length obtained by a fit of the Inverse-Gaussian to the data. Bottom) Collapse of the curves. Localization length rescaled obtained by a fit of the Inverse-Gaussian to the data.	53
5.4	Statistics of the LDoS $\ln(\tilde{\rho}_r)$ with increasing system size, L , for a fixed Lorentzian width, $\sigma = 0.05$ and localization length, $\xi_{\text{Thouless}} = 16384$, corresponding to a disorder of $W = 0.076$	54
5.5	Statistics of the LDoS $\ln(\tilde{\rho}_r)$ with increasing system size, L , for a fixed product $\sigma \cdot \xi_{\text{Thouless}} = 1$. Here we do $W = 0.3$, which corresponds to $\xi_{\text{Thouless}} \sim 1024$. Top: Statistics divided by a normalization factor $\langle \rho_r \rangle$. Bottom: Statistics of each system divided by its LDoS mean. This results uses TBC.	55

5.6	Top: Samples of the LDoS in Fig. 5.5 for $L=16$, evenly separated. Bottom: Difference between the low and high LDoS samples and its corresponding DoS, the systems chosen was the first one for the high LDoS and 7th and 11th for the low LDoS.	56
5.7	Top: Statistics for $\sigma \cdot \xi = 1.0$. Bottom: Statistics for $\sigma \cdot \xi = 4.0$. The sizes are sequentially encoded in the colors, being $L = 2^j$, starting at $L = 16$ and ending in $L = 32768$. While the shape, triangles, crosses and squares, has the localization length information, being 256, 512 an 1024 respectfully.	57
5.8	Distribution of the local density of states statistics for a one-dimensional tight-binding chain nearest neighbor for several sizes and Anderson disorder $W = 0.3$ and different product $\sigma \cdot \xi$ values, being from $\sigma \cdot \xi = 0.1 - 51.2$, always doubling, for the KPM (Colored curves). The grey curves are obtained from exact diagonalization from sizes ranging from $8 - 1024$, localization lengths from $64 - 1024$ and different L_{keff} . This shows that the L_{keff} is the important parameter. Top) Small L_{keff} Down) Intermediate-Big L_{keff} . The triangles were simulated with $L = 1024$ while the square with $L = 2048$	58
5.9	Position of the maximum of the distributions in Fig. 5.8. The diagonalization samples corresponds to different disorders realizations.	59
5.10	LDoS distribution for a system size $L = 16384$ and $W = 0.3$ for small N_k	59
5.11	Example of the difference between the statistics of the LDoS for a system size $L = 512000$ and $W = 1.0$ using; squares: Lorentzian; triangles: Gaussian; regularization of the Dirac- δ	60

[This page was intentionally left blank]

Chapter 1

Localization in Disordered Systems

1.1 Localization

The localization of quantum particles in disordered systems is one of the most studied topics in condensed matter physics. The disorder in solid state systems can come in several ways, such as small defects in the crystal lattice of the system, vacancies (i.e., the absence of atoms in an otherwise perfect crystal), onsite impurities, mechanical tensions among other effects that can alter the perfect crystalline lattice.

One of the consequences of disorder is the localization of states. In the absence of disorder the single-electron states will be Bloch states,

$$\psi_{\mathbf{k}} = u_{\mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}), \quad (1.1.1)$$

where $\Psi_{\mathbf{k}}$ is the eigenstate with momentum $\mathbf{k}$, $\mathbf{r}$ is the position vector and $u_{\mathbf{k}}(\mathbf{r})$ a function with the same periodicity as the crystal lattice. These are propagating states that extend throughout the entire system, even when it becomes infinite. The introduction of disorder breaks the translation invariance and the Bloch theorem is no longer valid. Sufficiently strong disorder causes the states to localize in a region, having tails that decay exponentially in space with a characteristic length scale, ξ , called localization length, $|\psi(r)| \sim \exp(-|r-r_0|/\xi)$. The simplest way to understand the effects of disorder in metallic systems is by studying the single electron problem which is described by the following Hamiltonian,

$$\mathcal{H} = \sum_{\mathbf{r}} \xi_{\mathbf{r}} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} + \sum_{\mathbf{r} \neq \mathbf{r}'} t_{\mathbf{r}\mathbf{r}'} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}'}, \quad (1.1.2)$$

where $\xi_{\mathbf{r}}$ are the local on-site energies, $t_{\mathbf{r}\mathbf{r}'}$ the hoppings between position $\mathbf{r}$ and $\mathbf{r}'$, and $c_{\mathbf{r}}^{\dagger}/c_{\mathbf{r}}$ the fermionic creation/annihilation operators. This will describe the motion of the electrons in the crystal lattice and we can draw conclusions about the localization depending on the behavior of the solution.

Disorder in theoretical computations can be modeled by means of random local energies, $\xi_{\mathbf{r}}$, to simulate lattice defects. One of the most common being the Anderson disorder model, in which the on-site energies are drawn from an uniform distribution of $[-W/2, W/2]$, where W measures the disorder strength.

The equation of motion becomes, from the Heisenberg equation of motion $i\dot{\mathcal{A}} = [\mathcal{A}, \mathcal{H}]$, assuming $\hbar = 1$,

$$i \frac{d}{dt} c_{\mathbf{r}} = \xi_{\mathbf{r}} c_{\mathbf{r}} + \sum_{\mathbf{r} \neq \mathbf{r}'} t_{\mathbf{r}\mathbf{r}'} c_{\mathbf{r}'}, \quad (1.1.3)$$

Localization can be understood from the time asymptotics behavior of the solution $c_r(t)$ of Eq. (1.1.3) [1]. If the solution $c_{r_0}(t)$ has the property $\lim_{t \rightarrow \infty} c_{r_0}(t) \neq 0$ with the initial conditions $c_r(0) = \delta(\mathbf{r} - \mathbf{r}_0)$, then, localization means that there is always a non zero probability that the electron returns to $\mathbf{r}_0$, starting at the same position $\mathbf{r}_0$. This localization has impact on physical observables because it will spatially confine the wavefunctions, such as the conductivity of the material [12–14], further justifying the study of disorder.

1.2 Identifying Localization

Initially conceptualized by P. W. Anderson [1], the existence of disorder can drastically change how the electrons behave in the crystal lattice. As disorder intensifies, all single-particle states exponentially localize in space. While this behavior is general for disordered lattices, the Anderson localization as a function of disorder strength depends on system's dimensionality. For one-dimensional systems, it is known that an infinitesimally small amount of disorder will cause all states to localize. [15,16], while for three-dimensions a system is expected to undergo a phase transition from metallic to localized at a critical disorder W_c [5,17–19].

The Anderson transition, from Metal-to-Anderson Insulator, is expected to exist in the dimensions higher than 2, $d > 2$, by the Scaling Theory of Localization [20,21]. In this introductory chapter, we explore the several ways one can do to identify the localization of states and the transition within a numerical treatment.

1.2.1 Decay of the Wavefunction

The obvious thought one has when we want to measure the localization length ξ is to look directly at the wavefunction. The localization length ξ is infinite in the extended phase and finite in the localized regime. For localized states, the modulus of the envelope of the wave function goes as

$$|\Psi_\alpha(\mathbf{r})| \sim \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_{\max}|}{\xi}\right), \quad (1.2.1)$$

where $\Psi_\alpha(\mathbf{r})$ is the α -th state of the system with energy E_α and $\mathbf{r}_{\max}$ is the position of the maximum of $|\Psi_\alpha(\mathbf{r})|$. This result is asymptotic in the tails and does not indicate the initial behavior from $\mathbf{r}_{\max}$. Therefore, one can diagonalize the Hamiltonian matrix, observe the decay of the squared-eigenstates and directly measure the localization length.

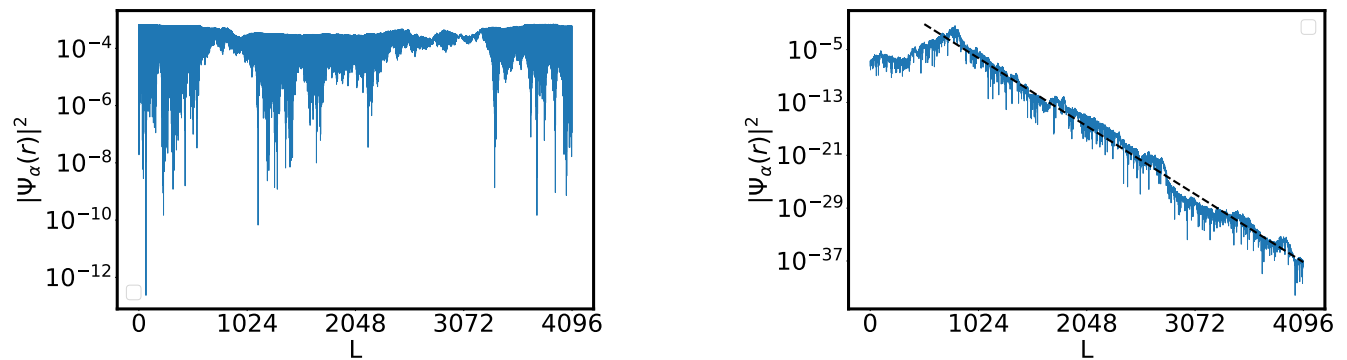


Figure 1.1: Amplitude of the state with energy E_α closest to 0 for a one-dimensional nearest neighbor tight binding system with size $L = 4096$; Left) $W = 0.01$. Right) $W = 1$.

In Fig. 1.1 Right), we clearly identify an exponentially decaying tail in the wavefunction from which one may obtain ξ by means of a fit to Eq. (1.2.1). If the disorder strength is high enough we see clearly the tails appearing, while if the localization length is too big (very low disorder) we might not have enough system size L to see the decay. This way of observing the localization is not computationally efficient since it requires the exact diagonalization and further analysis of the entire wavefunction.

1.2.2 Inverse Participation Ratio

Another quantity that measures the Localization of a state is the Inverse Participation Ratio, defined by

$$\text{IPR}_\alpha = \frac{\sum_{\mathbf{r}}^L |\Psi_\alpha(\mathbf{r})|^4}{\left(\sum_{\mathbf{r}}^L |\Psi_\alpha(\mathbf{r})|^2\right)^2},$$

For delocalized states, the wavefunction goes as $|\Psi_\alpha(\mathbf{r})|^2 \sim 1/L$, where L is the lateral system size, and, therefore, the IPR will also go as $\text{IPR}_\alpha \sim 1/L$, and vanishes in the thermodynamic limit. Alternatively, for spatially exponentially localized states, the wavefunction extends within a finite region, resulting in a finite IPR_α . This is an interesting behavior that allows us to identify the localization of a wave function.

This behavior is shown, for example, in the Fig. 1.2 for the simple cubic tight binding model with nearest-neighbor hoppings, where there is a known Anderson transition [21] at a disorder strength of about $W \sim 16.5t$. For $W = 10t$ we see the aforementioned behavior in IPR (P^{-1}), that decreases following $1/L$. The distribution in this region, $W = 10t$, is centered at a lower value and with a lower variance. Meanwhile, at $W = 18t$, where we are in the localized regime, the IPR (P^{-1}) scaling with lateral size L goes at first with $1/L$ and then it stagnates at a value $1/L_0$, for $L > L_0$ we see no change in the behavior. The distribution is centered at a higher value and the variance is greater, having longer tails in the region of high IPR (P^{-1}). The distribution changes at the transition, being sharply peaked in the extended region and broad in the localized regime.

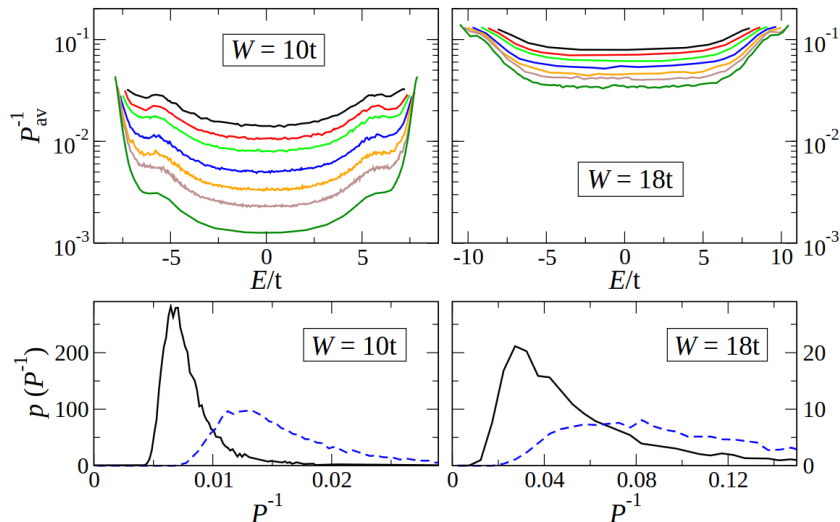


Figure 1.2: Upper) IPR (P^{-1}) for a 3D tight-binding nearest neighbor model with periodic boundary conditions averaged over 1000 realizations of disorders. The lines are system sizes, going from black to dark green, $L = 8, 9, 10, 12, 14, 16, 20$. Lower) Probability distribution of the IPR at the band center $E = 0$ (dark line) and near the band edges (blue dashed line). The blue line is a system Taken from [8].

Based on the distribution of the IPR an alternative numerical approach was proposed [22] for identifying the Anderson transition. Malyshev et al. [22] suggests that studying the ratio of the standard deviation, ΔP , to the mean, P_{av} , of the IPR distribution provides a system size independent parameter at the Anderson transition. Although the authors tested this hypothesis for one-dimensional models with Anderson disorder and power-law long-range hopping, $1/r^a$, because it exhibits localization-delocalization transitions [23–25], Schubert et al. [8] tested this for the 3D tight-binding model, having observed the data in Fig. 1.3. It indeed appears that there is a point near W_c that the relative standard deviation of the IPR is independent of system size, although it exists some disparity as shown in the inset. After some finite size analysis the authors were able to arrive at a critical disorder $W_c = 16.1 \pm 0.8$, which agrees with the W_c obtained through independent measurements. This type of analysis suggests that the study of distributions of certain local quantities can reveal deep information about localization and the Anderson transition.

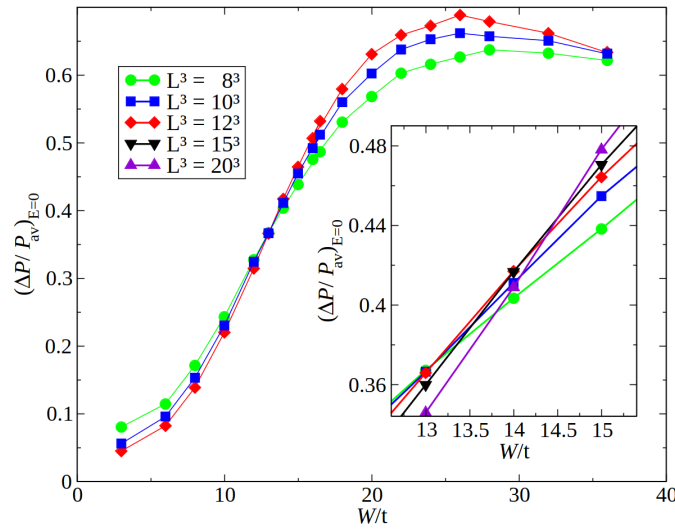


Figure 1.3: Results taken by [8] by where we have the normalized standard deviation of the IPR, $\Delta P/P$, in the band center $E = 0$ of the 3D tight-binding model as a function of disorder with periodic boundary conditions. The obtained results are averaged of 1000 realizations of disorder.

1.2.3 Statistics of Local Density of States

The density of states (DoS) is a global quantity, defined by

$$\rho(E) = \frac{1}{D} \sum_{\mathbf{r}} \langle \mathbf{r} | \delta(E - H) | \mathbf{r} \rangle = \frac{1}{D} \sum_{\alpha} \delta(E - E_{\alpha}), \quad (1.2.2)$$

where D is the dimension of the Hilbert space and H the Hamiltonian with corresponding E_{α} eigenenergies. This quantity, as per the name indicates, measures the number of eigenstates, per unit volume and unit energy. This global quantity is of utmost importance in the context of condensed matter physics, but it tell us very little about localization. This could be used to identify gaps and band behaviors, like for example the linear dispersion $\rho(E) \sim E$ in the 2D Dirac Cone or quadratic asymptotics $\rho(E) \sim E^2$ in 3D, but it tell us nothing about the nature of the states itself. As you can see, another example, is shown in Fig. 1.4, the DoS (red line) is always finite and it does not undergoes a critical behavior when we cross the critical disorder, $W_c \sim 16.5t$ for the 3D cubic lattice.

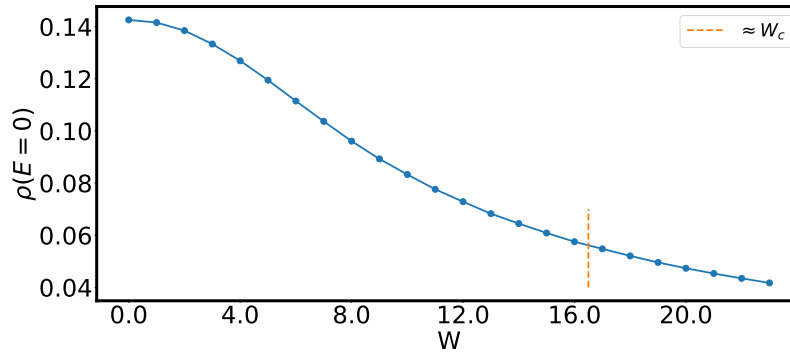


Figure 1.4: Evolution of the Density of States at the band center, $\rho(E=0)$, for several disorders. This shows smooth behavior when we cross the critical disorder $W_c \sim 16.5t$.

Anderson [1] already pointed out that to in order to study and interpret the transition from metal-to-insulator, from a extended to localized region, one needs to discuss the distribution of local quantities. One of the local quantities that has a critical behavior in the Anderson transition [26] is the Local Density of States (LDoS), defined by,

$$\rho_r(E) = \frac{1}{D} \langle \mathbf{r} | \delta(E - H) | \mathbf{r} \rangle = \frac{1}{D} \sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^2 \delta(E - E_{\alpha}), \quad (1.2.3)$$

where D is the Hilbert space dimension, $\Psi_{\alpha}(\mathbf{r})$ the wavefunctions and E_{α} the corresponding eigenvalue. The advantage in using the LDoS as a parameter is the measurability by Scanning Tunneling Microscopy (STM) [2, 27, 28].

Equation (1.2.3) is the exact definition of the LDoS, however, when dealing with simulations one often needs to regularize the Dirac- δ , $\delta(E - E_{\alpha})$, because of the discrete spectrum, as seen in Fig. 1.5.

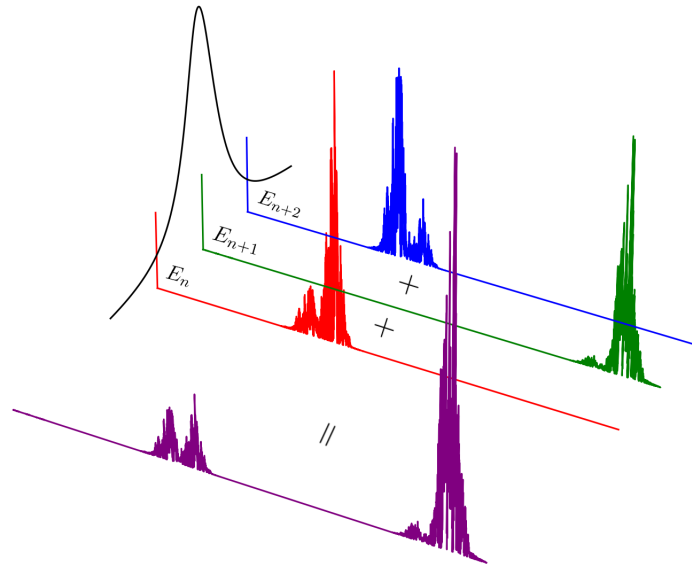


Figure 1.5: Schematic representation of the LDoS. If one assumes a regularization for the $\delta(E - E_{\alpha})$ we see that the LDoS is a weighted average of the wavefunctions.

This introduction of a regularization function for the Dirac- δ will introduce spectral resolution effects because

several eigenstates around energy E contributes to the LDoS with a spectral weight. Then, when dealing with simulations, the LDoS will be a weighted average of the eigenstates. As we increase the Hilbert space and reduce the broadening of the regularization function we approach the continuum. This is one of the main points in the upcoming sections of this work.

It is clear that $\rho_r(E)$ will be strongly fluctuating depending on the nature of the eigenstates. If the states with eigenvalue E_α are the localized, only a few lattice sites r will have very high LDoS, $\rho_r(E) > 0$, while most sites have close to null values. While if the state of E_α is extended all sites will have similar contributions and the LDoS will be fairly uniform throughout the system, as for example seen in Fig. 1.6. Therefore, to distinguish between regions of extended/localized one can compute $\rho_r(E)$ for many points and many disorder realizations and study its distributions. The behavior of the distribution with the system size is what tell us about the localization, not just the distribution itself.

In the extended region the distribution of the LDoS will be peaked around its mean and rather symmetric, due to the contributions of the LDoS being uniform in the lattice. This can be seen, for example, in Fig. 1.7 for a 3D lattice. Moreover, these distributions in the extended region does not change as we increase the system size. The absence of scaling is expected, since the extended states contribute with a probability $|\Psi_\alpha(r)|^2$ equally throughout the whole crystal lattice, therefore the statistics of the LDoS remains the same.

For localized states this is not the case. The probability distribution will be heavily asymmetric and peaked near zero, also shown in Fig. 1.7, causing the mean value to be much larger than the most probable one. In this case, the distribution becomes even more asymmetric as system size increases, which is due to the increase of the regions with low LDoS, since most of the LDoS will be highly concentrated around localized centers of the disorder landscape. We expect the LDoS statistics in this case to be nearer to 0 as we increase the system size.

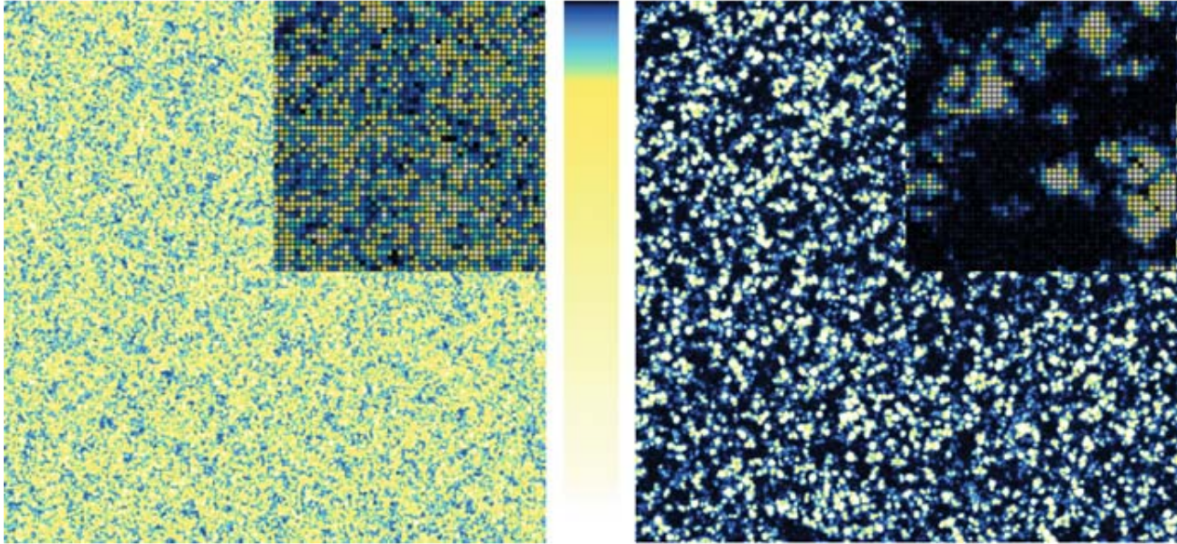


Figure 1.6: Numerical result showcasing the visual change of the LDoS distribution for a slice of a 3D tight-binding system with $L = 512^3$ and periodic boundary conditions. The Anderson disorder is $W = 10.0$. Right: $E = 0$; Left: $E = 7.69$. The inset is a 50^2 zoom. This change of character of the distribution indicates a phase transition. Taken from [9].

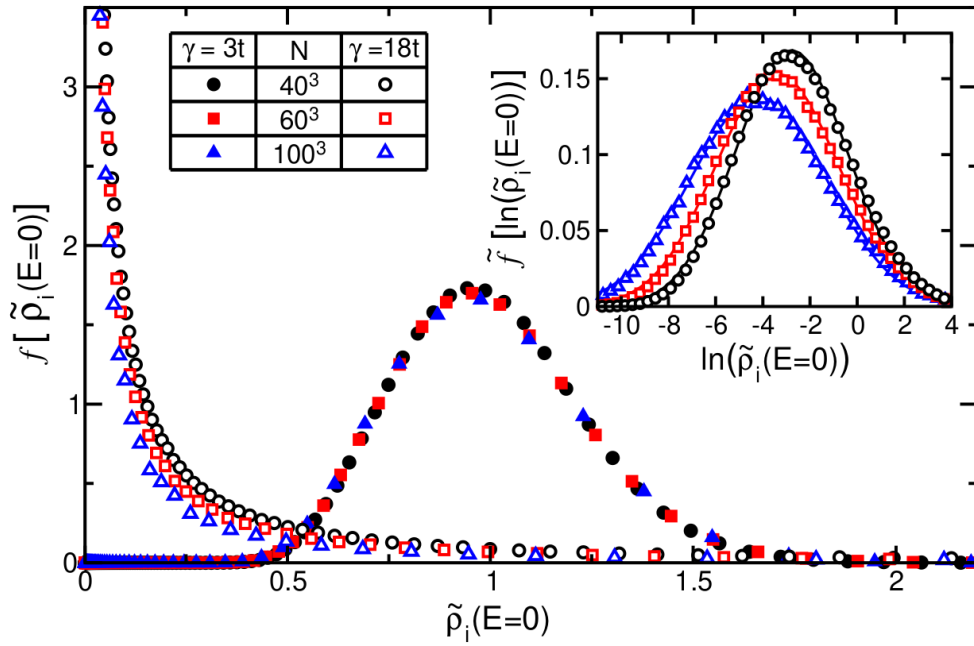


Figure 1.7: Examples of the probability distribution for the LDoS of a 3D cubic tight-binding lattice for several systems sizes, obtained and taken from [5].

Works by Mirlin et al. [4, 29–32] studied the LDoS statistics by using the supersymmetric σ -model and that $\tilde{\rho}_r = -\frac{1}{\pi N(E)} \Im(G_R(r, r, E))$, where G_R is the retarded Green function and $N(E)$ the mean density of states averaged over disorder realizations, and analytically obtained a log-normal distribution for the LDoS in the quasi-1D and 2D case,

$$f(\tilde{\rho}_r) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{\tilde{\rho}_r} \exp\left(-\frac{\ln^2(\tilde{\rho}_r/\tilde{\rho}_0)}{2\sigma^2}\right). \quad (1.2.4)$$

Alt'shuler et al. [3, 33, 34] also studied this problem analytically for the 1D case, with a different analytical technique than Mirlin et al. where it comprises in computing the moments $\langle \rho_r^n \rangle$ with the Berenzinskii diagrammatic. This concluded that in the infinite 1D system the LDoS distribution assumes an Inverse-Gaussian one,

$$f(\tilde{\rho}_r) = \sqrt{\frac{4\eta\tau}{\pi\tilde{\rho}_r^3}} \exp\left(-4\frac{(\tilde{\rho}_r - 1)^2}{\tilde{\rho}_r}\eta\tau\right), \quad (1.2.5)$$

where η is the broadening of the Lorentzian regularization of the Dirac- δ and τ the mean free path time of the electrons. This is of importance because it shows the LDoS distributions are dependent on the resolution η .

Having said that, the study of entire distributions is impracticable as a result of the computational complexity, we would have to compute the LDoS in many points to have a good distribution due to the nature of the tails in the probability distribution for the localized regime. Therefore, the work in [8] suggests that the relevant statistical information of these distributions are the arithmetic and geometric average, called mean DoS $\rho_{av}(E)$ and typical DoS $\rho_{typ}(E)$ (TDoS),

$$\rho_{av}(E) = \frac{1}{K \cdot N} \sum_{k=1}^K \sum_{r=1}^N \rho_r(E), \quad (1.2.6)$$

$$\rho_{\text{typ}}(E) = \exp \left\{ \frac{1}{K \cdot N} \sum_{k=1}^K \sum_{r=1}^N \ln \rho_r(E) \right\}, \quad (1.2.7)$$

where K is the number of realizations of disorder. These quantities stay finite in the extended phase while in the localized one only $\rho_{\text{av}}(E)$ is finite and $\rho_{\text{typ}}(E) \rightarrow 0$, as seen for example in Fig. 1.8. This fact suggested the creation of the a new effective order parameter $R(E) = \rho_{\text{typ}}(E)/\rho_{\text{av}}(E)$ which remains finite in the extended phase but scales down to zero in the thermodynamic limit of the localized phase. This allows us to compute more efficiently and identify with the same certainty, as seen in Fig. 1.9, the localized regions in the spectrum versus standard methods, such as the computation of IPR. With this knowledge we have another route to compute phase diagrams of localization. A phase diagram for extended-localized in 3D can be obtained with this method as shown in Fig. 1.10.

This study culminated in an another way of identifying the Metal-Anderson Insulator transition via the statistical study of the probability distribution of the LDoS [26, 35–40].

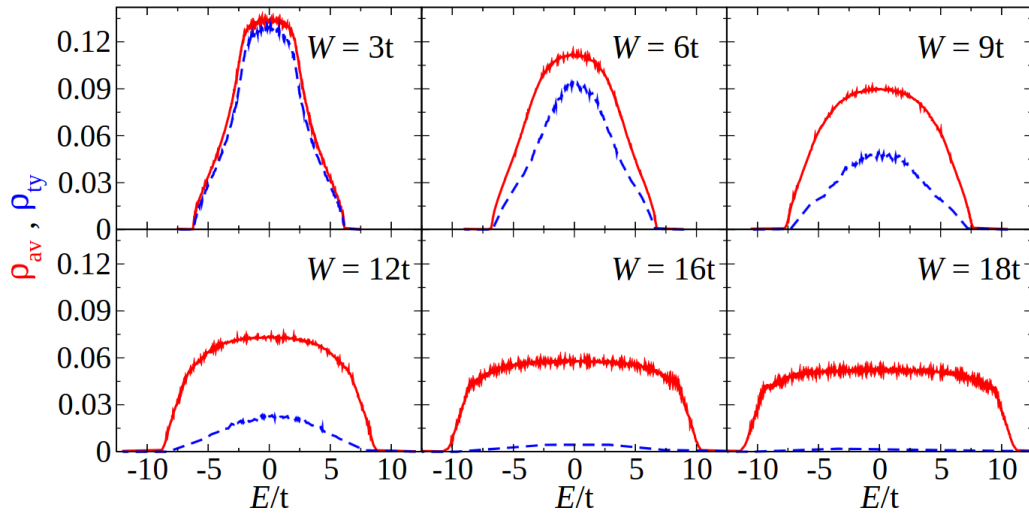


Figure 1.8: Density of states and Typical density of states for a nearest neighbor tight binding cubic lattice model with lateral size $L = 50$. Showing the vanishing of TDoS for $W \geq 16$. Taken from [8]

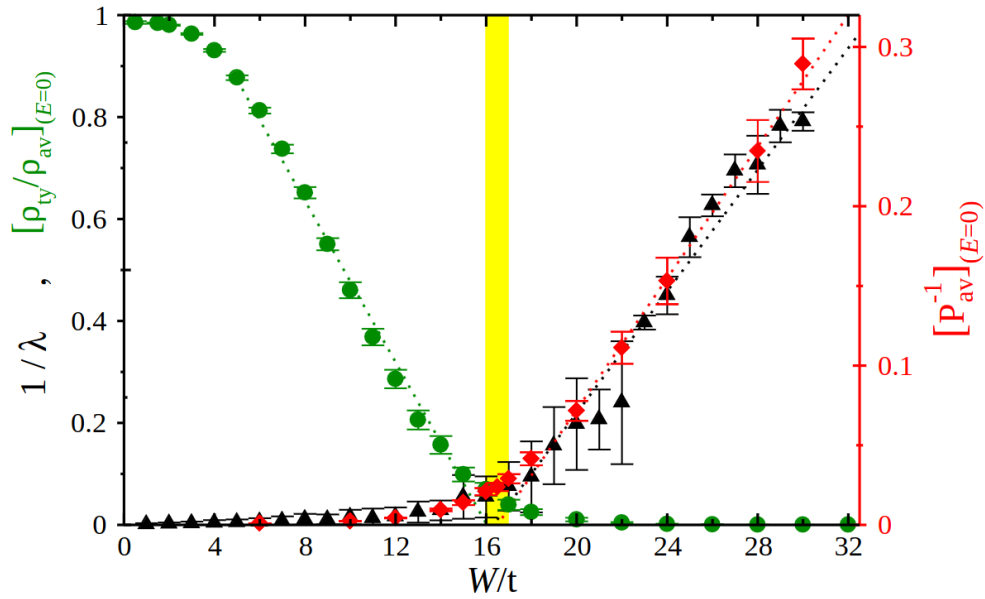


Figure 1.9: Plot of the order parameter $R(E) = \rho_{ty}(E)/\rho_{av}(E)$ for a cubic tight binding model, $L = 50^3$, in green versus other methods of studying localization. Triangles-Decay of the wave function. Diamonds-Inverse Participation Ratio. This shows that R works with the same precision for identification of localization. Taken from [8].

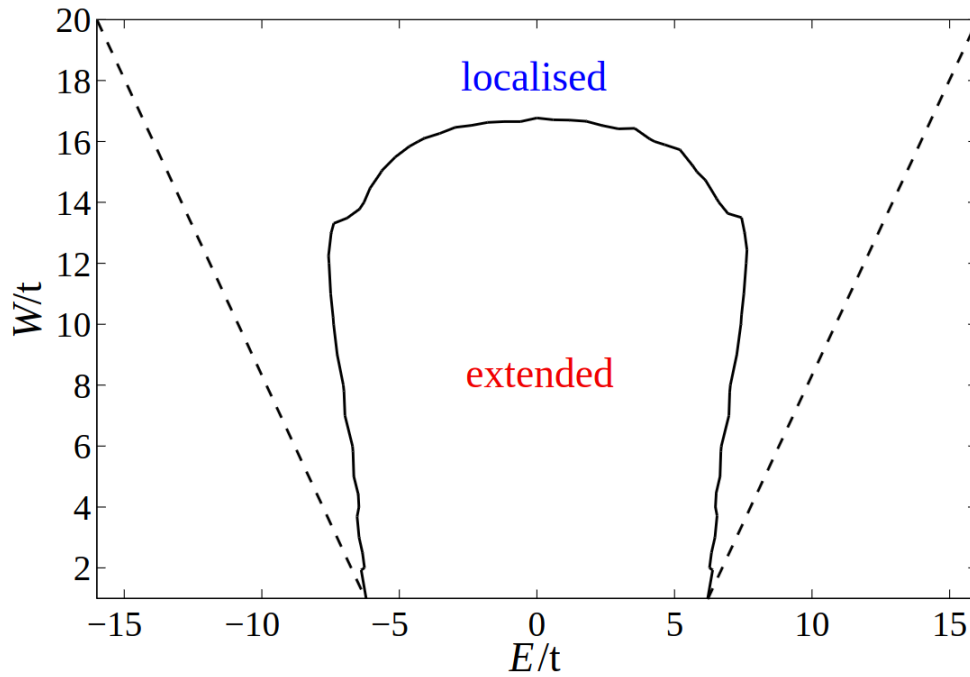


Figure 1.10: Phase diagram of the Anderson model on a 3D cubic lattice tight-binding model. Shown are the mobility edge (solid line), obtained the calibrating the effective order parameter R to the known critical disorder, $W_c \sim 16.5t$ and the Lifshitz boundaries (dashed lines). This is obtained and taken from [8].

1.3 Thesis Overview

This thesis was first envisioned on the study of the Weyl Semimetal-Metal transition. Weyl Semimetals [41] are a type of material that possess a linear dispersion relation at the Fermi level, $E = \hbar v_F |\mathbf{k}|$, where v_F is the Fermi velocity and $\mathbf{k}$ the momentum. It is known that at the node, $E = 0$, is robust to weak disorder [42, 43] and at the mean field level this type of material suffers from three phase transitions, Weyl Semimetal-Metal-Anderson Insulator (WS-M-AI). At the WS-M transition it was thought there existed a quantum critical point at finite disorder strength characterized by the sudden appearance of a non null Density of States [43–45]. Later it was numerically shown the appearance of rare states near $E = 0$ for weak disorder, thus dismissing a true quantum phase transition [46–49].

To study this transition we need a statistical approach of local quantities, such as the LDoS, because the rare states are heavily dependent on the disorder background, effectively being a infinite number of order parameters [50]. The study in the behavior of the LDoS statistics proved hard to understand. Consequently, we chose to thoroughly study the one-dimensional nearest neighbor tight-binding model to enhance our comprehension. Thus, this thesis is dedicated to understanding the LDoS numerics and its statistics.

Initially, we will review the spectral methods, namely the Kernel Polynomial Method (KPM), and focus on some technicalities. This will serve to point out some important details one must have in mind to pursue the following studies of the LDoS statistics and numerics.

Subsequently, we will provide a succinct overview of Graphical Processing Units (GPUs) and their integration into the implementation of the Kernel Polynomial Method (KPM). This has better promises in performance for matrix-vector operations and its a good candidate for the KPM hardware. This will encompass an examination of fundamental algorithms necessary for seamless functionality, as well as insights into memory optimization. Finalizing in the performance study when compared to CPU-based programs.

Thirdly, we will shift our focus to the computation of Local Density of States (LDoS) using the Kernel Polynomial Method (KPM). Here we introduce a novel algorithm designed for the calculation of LDoS across the entire lattice. Our examination will conclude by addressing the implications of twisted boundary conditions on LDoS statistics and providing nuanced insights into ensuring robust statistical comparisons across varying system sizes. This progression builds upon our earlier discussions on spectral methods and the GPU implementation.

In the final chapter, we will use the Kernel Polynomial Method (KPM) implementation on GPU and our newly developed algorithm for computing the Local Density of States (LDoS) to study the statistics of the LDoS in the one-dimensional tight binding nearest neighbor system. Through this analysis, we aim to draw meaningful conclusions about the inherent nature of LDoS distributions.

Chapter 2

Spectral Methods

In quantum systems simulations the main part of a numerical method is often the diagonalization of operators. The full diagonalization of large matrices is a numerically hard problem, which typically has a computational time complexity of $\mathcal{O}(D^3)$, where D is the matrix dimension. To overcome this issue, a significant research effort [51] have been placed in search of alternative methods that allows one to circumvent the explicit calculation of all the eigenpairs when accessing spectral properties, but offering a far more favorable scaling with the Hilbert Space dimension; One particularly successful algorithm is the Chebyshev spectral method [6, 52, 53]. In the following, we will put on the table the basics of this method and illustrate its usefulness by some common examples which often appears in quantum physics simulations. Then, we study the details of the method and touch on the concept of kernels which allows us to effectively use the method when the function has discontinuities. Lastly, we demonstrate how the method generalizes for operator-valued functions and see why this method is such a good one for this type of study.

2.1 Chebyshev Polynomials

The main object of the Kernel Polynomial Method (KPM) is the Chebyshev polynomials. These are defined as

$$T_n(x) = \cos(n \arccos(x)), \quad (2.1.1)$$

by expanding some polynomials we see that the first ones are

$$T_0(x) = 1, \quad (2.1.2)$$

$$T_1(x) = x,$$

$$T_2(x) = 2x^2 - 1,$$

$$T_3(x) = 4x^3 - 3x,$$

$$T_4(x) = 8x^4 - 8x^2 + 1. \quad (2.1.3)$$

These Chebyshev Polynomials also possess a recursion relation, which will be one of the pillars for the numerical method, given by

$$T_{m+2}(x) = 2xT_{m+1}(x) - T_m(x), \quad (2.1.4)$$

meaning we can easily construct the $T_m(x)$ polynomials for $m > 1$ using this recursion and with the initial conditions as

$$T_0 = 1 \quad T_1 = x. \quad (2.1.5)$$

They also satisfy an orthogonality relation in the range $x \in]-1, 1[$. More precisely, we have that

$$\int_{-1}^1 dx \cdot \frac{1}{\pi\sqrt{1-x^2}} T_n(x) T_m(x) = \frac{1+\delta_{n0}}{2} \delta_{nm}. \quad (2.1.6)$$

This can be simply proven by a trigonometric change of variables $x = \cos(z)$, then, the integral becomes

$$\frac{1}{\pi} \int_0^\pi dz T_n(\cos(z)) T_m(\cos(z)) = \frac{1}{\pi} \int_0^\pi dz \cos(nz) \cos(mz) = \frac{1+\delta_{n0}}{2} \delta_{nm} \quad (2.1.7)$$

The orthogonality means we can easily write functions as a series of Chebyshev polynomials, because if a function $f(x)$ can be expanded as

$$f(x) = \sum_{n=0}^{\infty} c_n T_n(x) \quad (2.1.8)$$

where c_m are unknown coefficients, then by doing the integral

$$\int_{-1}^1 dx \frac{f(x)}{\pi\sqrt{1-x^2}} T_m(x) = \int_{-1}^1 dx \frac{1}{\pi\sqrt{1-x^2}} \left\{ \sum_{n=0}^{\infty} c_n T_n(x) \right\} T_m(x) = c_m \frac{1+\delta_{m0}}{2}. \quad (2.1.9)$$

we can extract the coefficient because only one integral on the right-hand side will be nonzero, due to the orthogonality. These coefficients are correct for an expansion of the form

$$f(x) = c_0 + 2 \sum_{m=1}^{\infty} c_m T_m(x), \quad (2.1.10)$$

but for practical reasons, due to the divergence in $1/\sqrt{1-x^2}$, we rearrange the expansion in the form

$$f(x) = \frac{1}{\pi\sqrt{1-x^2}} \left\{ \alpha_0 + 2 \sum_{m=1}^{\infty} \alpha_m T_m(x) \right\}, \quad (2.1.11)$$

and the coefficients are now given by

$$\alpha_m = \int_{-1}^1 dx f(x) T_m(x). \quad (2.1.12)$$

This essentially comes from a different definition for the orthogonality relation, instead of Eq. (2.1.6), we have

$$\int_{-1}^1 dx \pi \sqrt{1-x^2} \phi_n(x) \phi_m(x) = \frac{1+\delta_{n0}}{2} \delta_{nm}, \quad (2.1.13)$$

where $\phi_n(x) = \frac{T_n(x)}{\pi\sqrt{1-x^2}}$ and, formally, the expansion becomes in the set of functions $\phi_n(x)$ with coefficients given by Eq. (2.1.12). It is worth defining a scalar product for these definitions, which we can alternatively

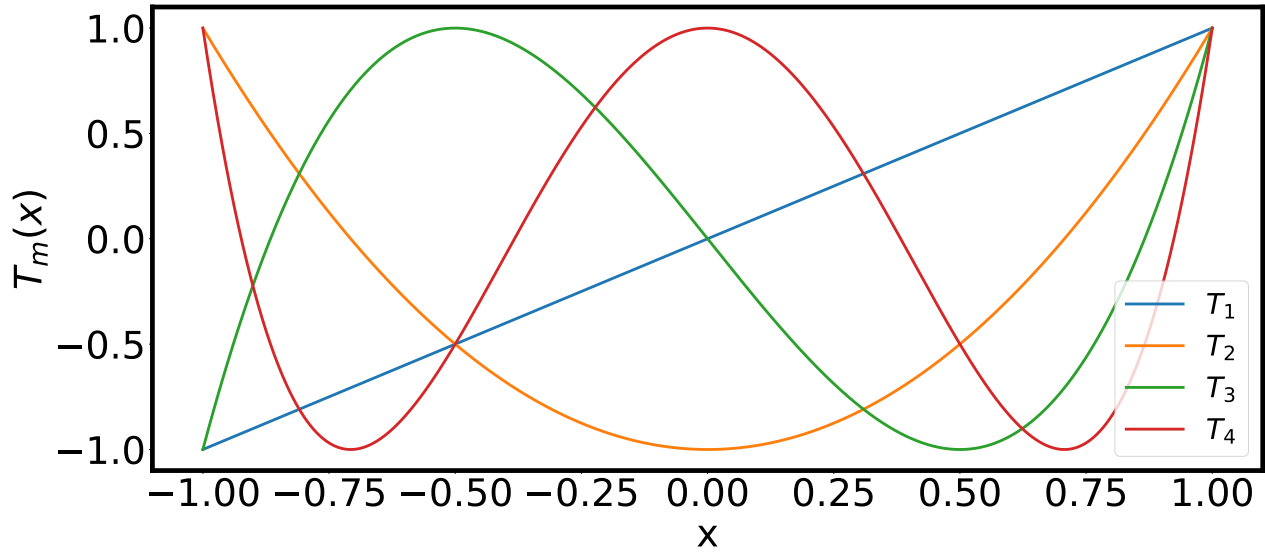


Figure 2.1: Example of Chebyshev Polynomials.

use each on if the integrals prove easier in these forms.

$$\xi = \int_{-1}^1 dx \cdot \frac{1}{\pi\sqrt{1-x^2}} f(x) T_m(x) \quad \zeta = \int_{-1}^1 dx \cdot \pi\sqrt{1-x^2} f(x) \phi_n(x). \quad (2.1.14)$$

2.2 Expansions of Common Functions

Expansion of a Dirac- δ

The use of a Dirac- δ , $\delta(x-y)$, appears in a wide range of physical quantities, essential for us to study its expansion in this polynomials. The coefficients α_δ^m , using the second definition of the expansion given by Eq. (2.1.11), are

$$\alpha_\delta^m = \int_{-1}^1 dx f(x) T_m(x) = \int_{-1}^1 dx \delta(x-y) T_m(x) = T_m(y). \quad (2.2.1)$$

Assuming we expand it in x and that $x, y \in]-1, 1[$. Then the expansion becomes, $T_0(y) = 1$,

$$\delta(x-y) = \frac{1}{\pi\sqrt{1-x^2}} \left\{ 1 + 2 \sum_{m=1}^{\infty} T_m(y) T_m(x) \right\}. \quad (2.2.2)$$

As you can see by Fig. 2.2 the truncated expansion in Chebyshev polynomials of the Dirac- δ gets closer to the $\delta(x)$, although the oscillations that arise at the discontinuities are exuberant and growing with increasing number of polynomials M . To mitigate these oscillations we have to resort to the concept of Kernels, which we will study later.

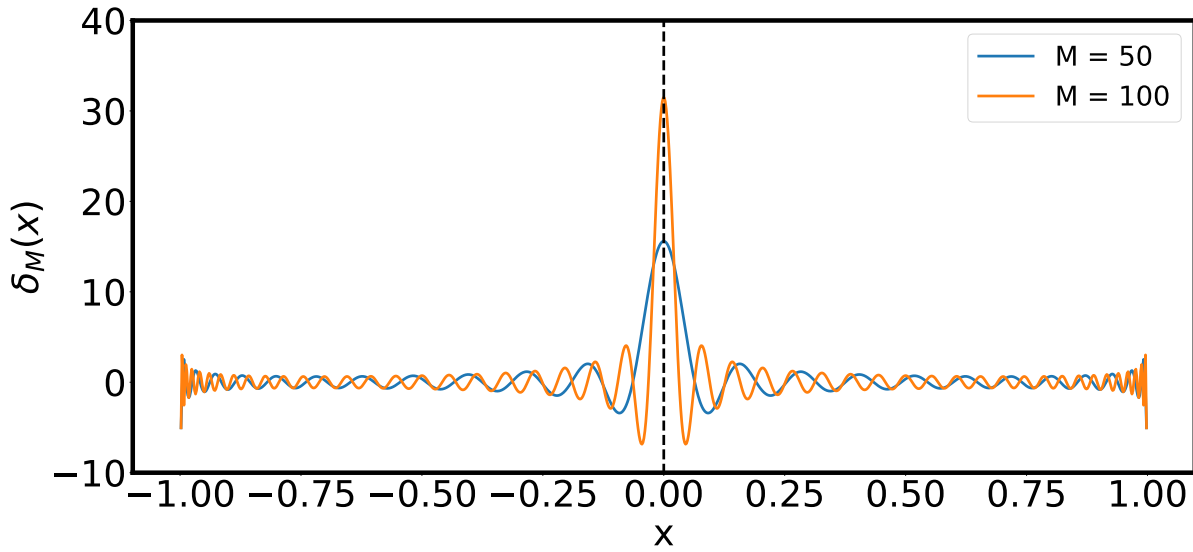


Figure 2.2: Expansion of a Delta-Dirac $\delta(x)$ with different number of Moments.

Expansion of a Complex Exponential

The complex exponential has a lot of interest because it can describe the time evolution of quantum states with a given Hamiltonian $\mathcal{H}$. We are interested in expanding in Chebyshev polynomials the complex exponential

$$f(x) = e^{-iaxt}, \quad (2.2.3)$$

where a and t are scalar parameters, with t normally being the time. The coefficients are computed, using now the first definition for the scalar product,

$$\alpha_e^m = \int_{-1}^1 dx \cdot \frac{1}{\pi\sqrt{1-x^2}} \cdot e^{-iaxt} \cdot T_m(x), \quad (2.2.4)$$

where the subscript α_e^m indicates which functions we are expanding, subscript e for exponential. We chose this definition because if we use the substitution $x = \cos(\theta)$ we have for the integral

$$\alpha_e^m = \frac{1}{\pi} \int_0^\pi d\theta \cdot e^{-iat \cdot \cos(\theta)} \cdot \cos(m\theta), \quad (2.2.5)$$

which is one the integral definitions for the Bessel functions of the first kind, $J_m(x)$,

$$J_m(x) = \frac{1}{\pi i^m} \int_0^\pi d\theta \cdot e^{ix \cdot \cos(\theta)} \cdot \cos(m\theta), \quad (2.2.6)$$

Therefore, the coefficients are given by

$$\alpha_e^m = (-i)^m J_m(at) \quad (2.2.7)$$

where we used that $J_m(-x) = (-1)^m J_m(x)$, the expansion, which we call $\tilde{f}$, becomes

$$\tilde{f}(x) = J_0(at) + 2 \cdot \sum_{m=1}^M (-i)^m J_m(at) T_m(x). \quad (2.2.8)$$

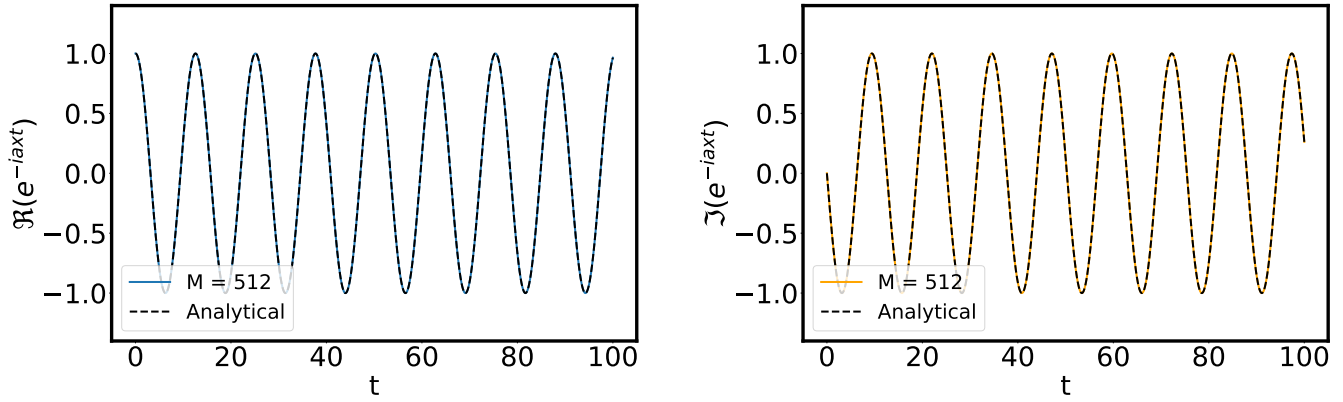


Figure 2.3: Example of complex exponential Chebyshev expansion for $x = 0.5$ and $a = 1$ with 512 moments. Left) Real part Right) Imaginary part.

As you can see by Fig. 2.3, even with $M = 512$ moments the expansion has a remarkable agreement with the analytical form.

Expansion of a Green's Function

One of the other important functions to study its expansion is the Green's function, we of course do the scalar form

$$g(z) = \frac{1}{z - x}, \quad (2.2.9)$$

where z is a complex variable and x a real one, belonging to $[-1, 1]$. The coefficients are

$$\alpha_g^m = \int_{-1}^1 dx \cdot \frac{1}{\pi \sqrt{1-x^2}} \cdot \frac{1}{z-x} \cdot T_m(x), \quad (2.2.10)$$

and once again doing the transformation $x = \cos(\theta)$ we have the integral

$$\alpha_g^m = \frac{1}{\pi} \int_0^\pi d\theta \cdot \frac{\cos(m\theta)}{z - \cos(\theta)}. \quad (2.2.11)$$

This integral can be done with the aid of complex analysis, by exploring the fact that the integrand is even and doing the transformation $w = e^{i\theta}$ and $\cos(\theta) = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) = \frac{1}{2}(w + w^{-1})$ we get

$$\alpha_g^m = \frac{1}{4\pi i} \oint_{|w|=1} dw \cdot \Re \left\{ \frac{1}{w} \cdot \frac{w^m}{z - \frac{1}{2}(w + w^{-1})} \right\}. \quad (2.2.12)$$

By the Residue Theorem,

$$\alpha_g^m = \frac{1}{2} \sum_{a_k} \text{Res}(f, a_k), \quad (2.2.13)$$

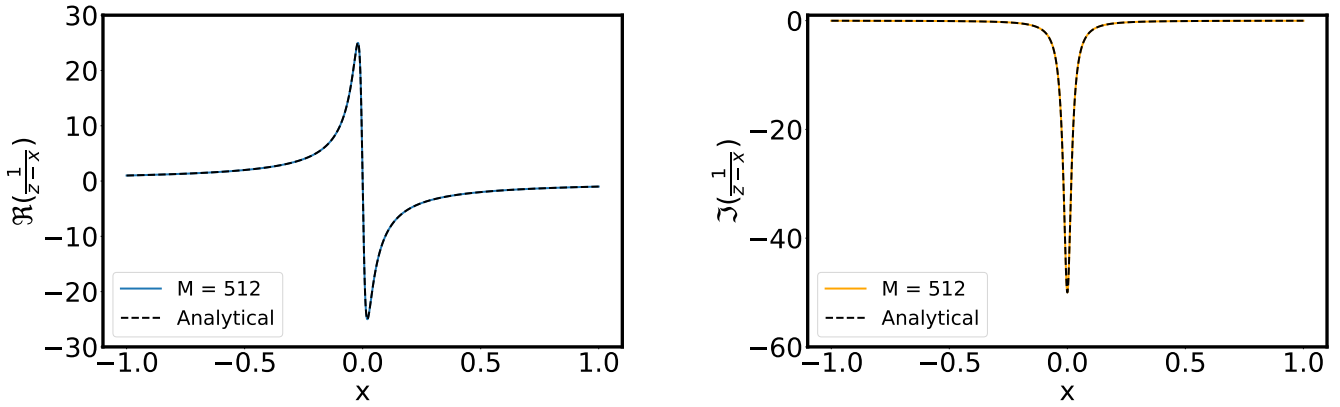


Figure 2.4: Example of Green's function Chebyshev expansion for $z = -0.02i$ with 512 moments. Left) Real part Right) Imaginary part.

where we sum over the residues of $f(w) = \frac{1}{w} \cdot \frac{w^m + w^{-m}}{z - \frac{1}{2}\{w + w^{-1}\}}$ inside of the unit circle. After some careful algebra and assuming $z = \epsilon + \eta i$, the integral is equal to

$$\alpha_g^m = \frac{e^{im \cdot \text{sign}(\eta) \cdot \arccos(z)}}{\sqrt{z^2 - 1}}. \quad (2.2.14)$$

The expansion then becomes

$$\tilde{g}(z) = \frac{-i \cdot \text{sign}(\eta)}{\sqrt{1 - z^2}} \left\{ 1 + 2 \cdot \sum_{m=1}^M e^{im \cdot \text{sign}(\eta) \cdot \arccos(z)} T_m(x) \right\}. \quad (2.2.15)$$

I want to remind the reader that in the limit of $\eta \rightarrow 0^\pm$, which is what we do for the Green function, we have the following understanding of the expression

$$\lim_{\eta \rightarrow 0^\pm} \frac{1}{a - x + \eta i} = \mathcal{P} \left\{ \frac{1}{a - x} \right\} - \text{sign}(\eta) \cdot \pi \delta(a - x), \quad (2.2.16)$$

where $\mathcal{P}$ denotes the principal value. Fig. 2.4 show that this expansion has a remarkable agreement.

Expansion of a Gaussian and Lorentzian

Sometimes we cannot directly expand the $\delta(x - \mu)$ due to the Gibbs oscillations, which can contaminate our results. To go around this problem we could regularize the Dirac- δ as a normalized distribution function and expand it in Chebyshev Polynomials, the most common ones being the Gaussian and a Lorentzian with a given width η and mean μ ,

$$\delta_{\text{reg}}^G(x - \mu, \eta) = \frac{1}{\sqrt{2\pi\eta}} \exp \left\{ -\frac{1}{2} \cdot \frac{(x - \mu)^2}{\eta^2} \right\}, \quad (2.2.17)$$

$$\delta_{\text{reg}}^L(x - \mu, \eta) = \frac{1}{\pi} \cdot \frac{\eta}{(x - \mu)^2 + \eta^2}, \quad (2.2.18)$$

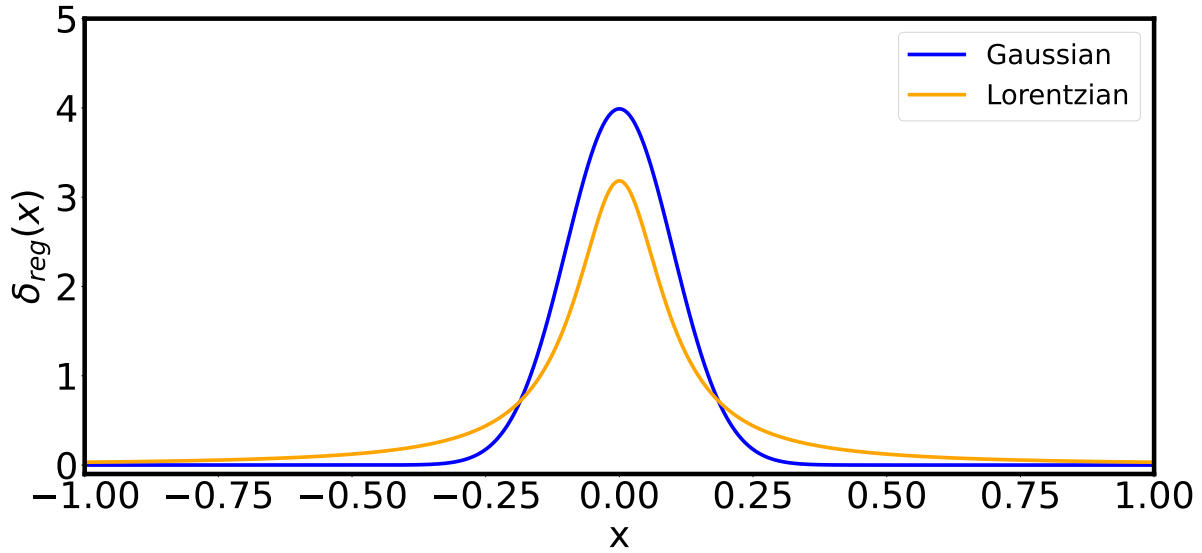


Figure 2.5: Difference of representations, Gaussian and Lorentzian, for the $\delta_{\text{reg}}(x, \eta)$ with width $\eta = 0.1$.

they both converge to the Dirac- δ in the limit, $\lim_{\eta \rightarrow 0} \delta_{\text{reg}}^{G,L}(x - \mu, \eta) = \delta(x - \mu)$. The main difference between them are the tails of the distribution, whereas the Gaussian decays exponentially the Lorentz have heavy tails, meaning it decays more slowly.

The task then becomes how to approach the Chebyshev expansion for such representations. Recalling that coefficients, using the first definition of the orthogonality, are given by

$$\alpha_m = \int_{-1}^1 dx \cdot \frac{1}{\pi \sqrt{1-x^2}} f(x) T_m(x), \quad (2.2.19)$$

we just have to perform the integral, first and foremost, this integral should be done analytical whenever possible, the following integrals could be done analytically with complex analysis, but if it proves too hard we have to resort to numerical integration, and this is what we want to explore in this section. The integral we want to compute is, by expressing the full terms, is

$$\alpha_m = \int_{-1}^1 dx \cdot \frac{1}{\pi \sqrt{1-x^2}} \frac{1}{\sqrt{2\pi}\eta} \exp\left\{-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\eta^2}\right\} \cos(m \cdot \arccos(x)), \quad (2.2.20)$$

and the main problem of this integrals in the numerical aspect lies in the oscillations of $\cos(m\theta)$. One of the best integrals techniques for this type is the Chebyshev-Gauss quadrature, which requires the integral to be written as

$$I = \int_{-1}^1 dx \cdot \frac{1}{\sqrt{1-x^2}} h(x), \quad (2.2.21)$$

where $h(x)$ is a general function. These integration methods heavily relies on the number points used to partition the domain. Due to heavy oscillatory nature of the $T_m(x)$ polynomials, which gets worse as we increase m , and the Gaussian envelope with width η , we easily see the problem, for higher order moments m and small η we can miss the important features if we misjudge the number of points necessary for the integration of the integrand function, rendering the numerical integration useless.

To avoid this problem we could crank up heavily on the number of partition points of the domain, but that

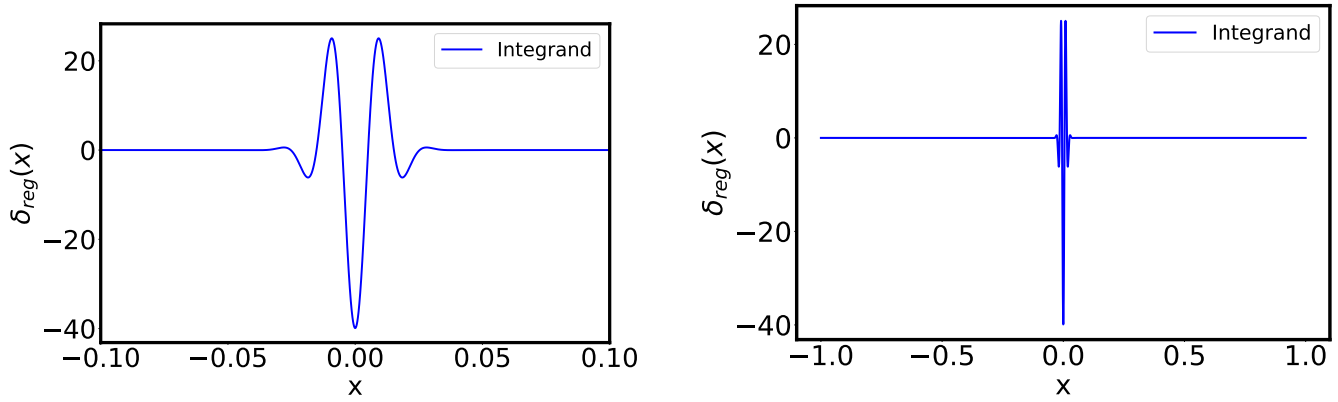


Figure 2.6: Left) Example of a Chebyshev expansion of the integrand in Eq. (2.2.20) with width $\eta = 0.01$, $\mu = 0$ and cutting the domain to $[-10\eta, 10\eta]$ for $M = 310$. Right) The same integrand in the full domain $[-1, 1]$. As you can see Right) would require many more points of integration to capture the important features, while on left we can reduce the number of points and have a very good approximation for the value of the integrals.

would require more computation power. A smarter way is to notice that the integrand decays exponentially fast, especially for small η . Therefore we could choose only the integration points in the range of $[-10\eta + \mu, 10\eta - \mu]$, which corresponds to 10η of the Gaussian, and integrate just that small portion.

We now devised a method which allows us only to pick the important integration points, now we need to see how many do we need nonetheless. The oscillations of Chebyshev polynomials, $T_m(\cos(\theta)) = \cos(m\theta)$, have a period of $T = 2\pi/n$, therefore, in addition of picking points in the range $[-10\eta + \mu, 10\eta - \mu]$ we need at least to partition the original domain with more than m points, so we can capture the oscillations.

After a careful examination of how to numerically perform the heavy oscillatory integrals we are capable of expanding Eq. (2.2.17).

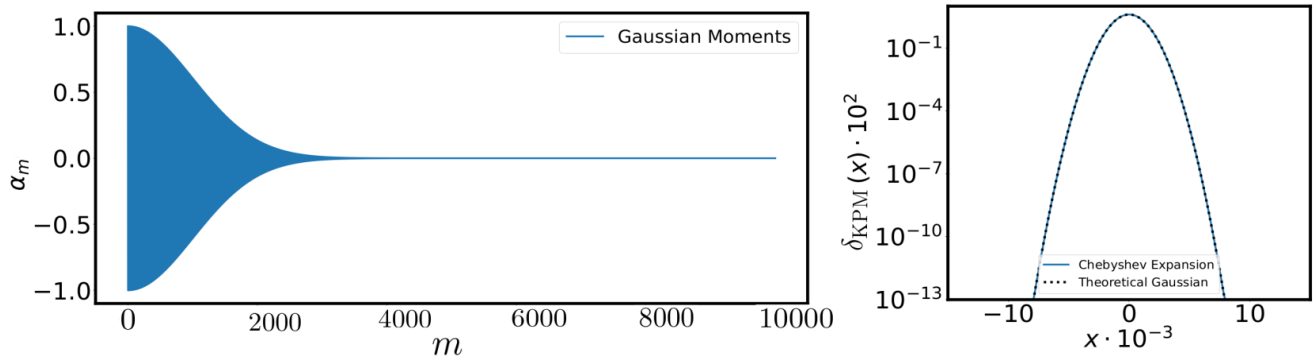


Figure 2.7: Gaussian Chebyshev expansion of width $\eta = 0.001$, $\mu = 0$ and $M = 10000$. Left) Magnitude of the moments Right) Reconstruction of the Gaussian. This expansion agrees remarkably well up to 10^{-13} .

As you can see by Fig. 2.7 Right) the expansion agrees very well and does not show any sign the oscillations we see when we directly expand the Dirac- δ . The reason comes from the magnitude of the expansion coefficients, while for a Gaussian $\alpha_m \xrightarrow{m \rightarrow \infty} 0$, as you can see in Fig. 2.7 Left), for a Dirac- δ they never converge to a value, but instead are $\alpha_m = T_m(\mu)$, which never decays with increasing m and therefore the oscillations are

not dampened.

One additional commentary for these expansion, for a good agreement we always want to compute as many moments as possible until we reach a good bound for them, normally very close to zero. For the above example we computed 10 times as many moments as its inverse of width, this choice was rather random but always guaranteed a sufficient number of moments as you can see in Fig. 2.7 Left). Of course, a better way can be simply to set a routine to compute until we reach a certain bound value $|\alpha_m| < 10^{-a}$, where a will control the required precision, this is another way we can determine the number of moments until we are satisfied.

After this careful analysis we are ready to tackle the Lorentzian with the same procedure. Remember that the moments of the Lorentzian can be done analytically, $-\Im \left\{ \frac{e^{im \cdot \text{sign}(\eta) \cdot \arccos(z)}}{\sqrt{z^2 - 1}} \right\}$, where $z = \epsilon + \eta i$. We chose to explain here some details about its numerical integration because it will be important for other similar functions, which we will use in this thesis, such as the square root of the Lorentzian.

Since this distribution has heavy tails, we set the integration to many times larger than with the Gaussian. For practical reasons we shall not limit the integration since the tails for reasonable enough width, $\eta \sim 0.001$, don't justify this. We have to, unfortunately, set a very high number of integration points. The moment integral is

$$\alpha_m = \frac{1}{\pi} \int_{-1}^1 dx \cdot \frac{1}{\pi \sqrt{1-x^2}} \frac{\eta}{x^2 + \eta^2} \cos(m \cdot \arccos(x)).$$

Fig. 2.8 Right) also shows the good agreement of the expansion with the numerical integration, and the moments for the Lorentzian suffers the same fate $\alpha_m \xrightarrow{m \rightarrow \infty} 0$, as shown in Fig. 2.8 Left), avoiding the Gibbs oscillations.

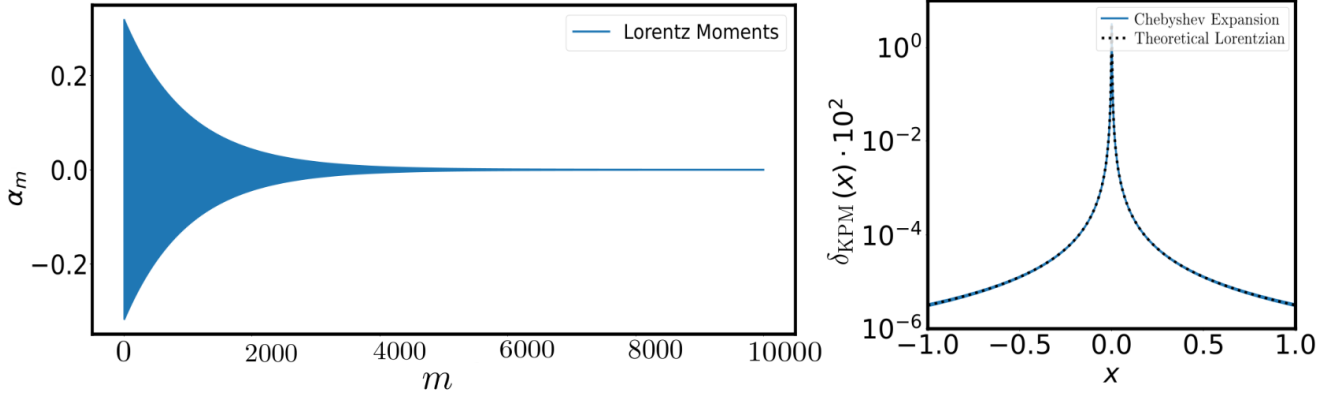


Figure 2.8: Lorentzian Chebyshev expansion of width $\eta = 0.001$, $\mu = 0$ and $M = 10000$. Left) Magnitude of the moments Right) Reconstruction of the Gaussian. This expansion agrees remarkably well up to 10^{-13} .

2.3 Kernels

As we saw beforehand, the expansion of discontinuous functions in Chebyshev polynomials at a finite number of moments M ,

$$f_{KPM}(x) = \frac{1}{\pi \sqrt{1-x^2}} \left\{ \alpha_0 + 2 \sum_{m=1}^{M-1} \alpha_m T_m(x) \right\}, \quad (2.3.1)$$

comes with Gibbs oscillations, which we cannot ignore. To work around this we modify the moments by a factor g_m ,

$$f_{KPM}(x) = \frac{1}{\pi\sqrt{1-x^2}} \left\{ g_0\alpha_0 + 2 \sum_{m=1}^{M-1} g_m\alpha_m T_m(x) \right\}, \quad (2.3.2)$$

and try to investigate the choice of g_m in the resulting expansion following Weisse et. al. [6]. We can define a new function $K_M(x, y)$, which we will denominate as Kernel, in the following way,

$$K_M(x, y) = g_0\phi_0(x)\phi_0(y) + 2 \sum_{m=1}^{M-1} g_m\phi_m(x)\phi_m(y), \quad (2.3.3)$$

where $\phi_m(x) = \frac{T_m(x)}{\pi\sqrt{1-x^2}}$. We can recover $f_{KPM}(x)$ by using the second orthogonality relation, given by Eq. (2.1.13),

$$f_{KPM}(x) = \int_{-1}^{-1} dy \cdot \pi\sqrt{1-y^2} K_M(x, y) f(y). \quad (2.3.4)$$

The question now reduces to finding the optimal kernel $K_M(x, y)$ which reduces the Gibbs oscillations.

The simplest Kernel, attributed to Dirichlet, is obtained by simply setting all the coefficients $g_m^D = 1$, which of course does nothing to control the Gibbs oscillations, it's Kernel $K_M^D(x, y)$ can be evaluated by the use of the Christoffel–Darboux identity

$$K_M^D(x, y) = \frac{\phi_M(x)\phi_{M-1}(y) - \phi_{M-1}(x)\phi_M(y)}{x - y}. \quad (2.3.5)$$

Through this simple choice we can already define and study the convergence. We say that

$$\int_{-1}^1 \pi\sqrt{1-x^2} \{f(x) - f_{KPM}(x)\}^2 \xrightarrow{M \rightarrow \infty} 0. \quad (2.3.6)$$

This alone does not tell us the convergence behavior for any given x , if it is uniform throughout x we would not have oscillations.

The next improvement is due to Féjer, details at ref. [6], which showed that for continuous functions the Kernel

$$K_M^F(x, y) = \frac{1}{M} \sum_{i=1}^M K_i^D(x, y), \quad (2.3.7)$$

converges uniformly in the interval $]-1 + \epsilon, 1 - \epsilon[$, where ϵ is just a rescaling factor to control the boundaries. It is showed that this kernel corresponds to the coefficients

$$g_m^F = 1 - \frac{m}{M}, \quad (2.3.8)$$

notice that the higher order terms, the ones that oscillate the most in the expansion, have a reduced weight, thus controlling the Gibbs oscillations.

Later, since we did not fully exhaust the constraints on the coefficients g_m , a better one was found by Jackson,

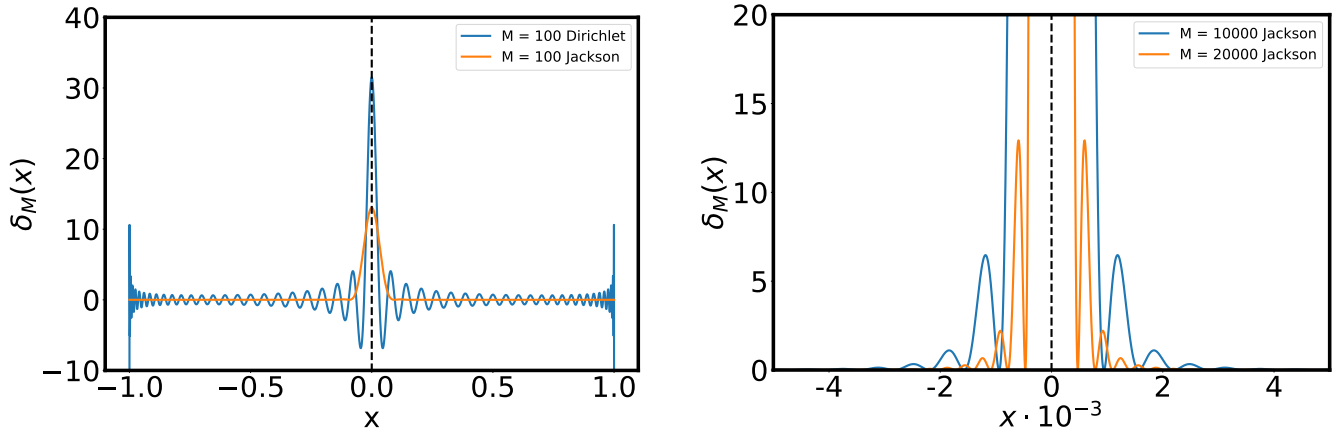


Figure 2.9: Left) Example of a Chebyshev expansion of $\delta(x)$ with and without the Jackson kernel, with the same number of moments, 100 Right) Suppressed positive Gibbs oscillations in a high order Chebyshev expansion of $\delta(x)$.

in which the coefficients are given by

$$g_m^J = \frac{(M - m + 1) \cos\left(\frac{\pi m}{M+1}\right) + \sin\left(\frac{\pi m}{M+1}\right) \cot\left(\frac{\pi}{M+1}\right)}{M + 1}, \quad (2.3.9)$$

and this is the one that converges best for most applications, although it does not truly kill the oscillations, as shown in Fig 2.9, instead they become strictly positive and very suppressed.

Since our initial motivation for introducing the regularizing kernels was to have a well behaved Chebyshev expansion of Dirac- δ functions, it is worth investigating the shape of the approximation as if it were a probability distribution, i.e., by measuring the variance and kurtosis.

We start by remembering that for the $\delta_{\text{KPM}}(x - y)$ expansion we have

$$\delta_{\text{KPM}}(x - y) = \frac{1}{\pi \sqrt{1 - x^2}} \left\{ g_0^J + 2 \sum_{m=1}^M g_m^J T_m(y) T_m(x) \right\}, \quad (2.3.10)$$

where we added the extra Jackson coefficients g_m^J . Thus, to compute the variance and kurtosis, we only need to evaluate the following integrals, which correspond to moments,

$$\langle x^n \rangle = \int_{-1}^1 dx \cdot x^n \delta_{\text{KPM}}(x - y), \quad (2.3.11)$$

for $n \in [1, 2, 3, 4]$ to compute the necessary central moments. For the variance we only need to compute $n = 1$ and $n = 2$. To these integrals we recall the orthogonality relation, Eq. (2.1.6), and that we can write x and x^2 in terms of Chebyshev polynomials by direct manipulation of Eq. (2.1.2). The relevant polynomials are

$$\begin{aligned} x &= T_1(x) \\ x^2 &= \frac{T_2(x) + T_0(x)}{2}, \end{aligned} \quad (2.3.12)$$

which allows the integrals $\langle x \rangle$ and $\langle x^2 \rangle$ to be trivially computed. They are

$$\langle x \rangle = \int_{-1}^1 dx \cdot x \cdot \delta_{\text{KPM}}(x - y) = g_1^J T_1(y), \quad (2.3.13)$$

$$\langle x^2 \rangle = \int_{-1}^1 dx \cdot x^2 \cdot \delta_{\text{KPM}}(x - y) = \frac{g_2^J T_2(y) + g_0^J T_0(y)}{2}, \quad (2.3.14)$$

then the variance becomes

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{1 - g_2^J}{2} + y^2 \left\{ g_2^J - [g_1^J]^2 \right\}, \quad (2.3.15)$$

and after using the definition of Jackson Kernel, we arrive at

$$\sigma^2 = \sin^2 \left(\frac{\pi}{M+1} \right) \frac{1}{M+1} \left\{ M - (M-1)y^2 \right\}, \quad (2.3.16)$$

since we will be dealing with a high number of moments M it makes perfect sense to use a Taylor expansion for the $\sin^2(x)$, we then have

$$\sigma^2 \sim \frac{\pi^2}{(M+1)^3} \left\{ M - (M-1)y^2 \right\}. \quad (2.3.17)$$

We can infer some results from this, at the center, $y = 0$, the width of the expansion at high order M goes like $\sigma \sim \pi/M$ while at the boundaries, $y = 1$, it goes like $\sigma \sim \pi/M^{3/2}$, as shown in Fig. 2.10. This is useful knowledge because if we want computations that guarantee the same width throughout the range $[-1, 1]$ we have to change the number of moments in the expansion.

It is reasonable to approximate the expansion of $\delta_{\text{KPM}}(x - y)$ as a Gaussian of width σ at the center.

$$\delta_{\text{KPM}}(x - y) \sim \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{1}{2} \left\{ \frac{x - y}{\sigma} \right\}^2 \right). \quad (2.3.18)$$

For the tails we need to investigate the fourth central moment.

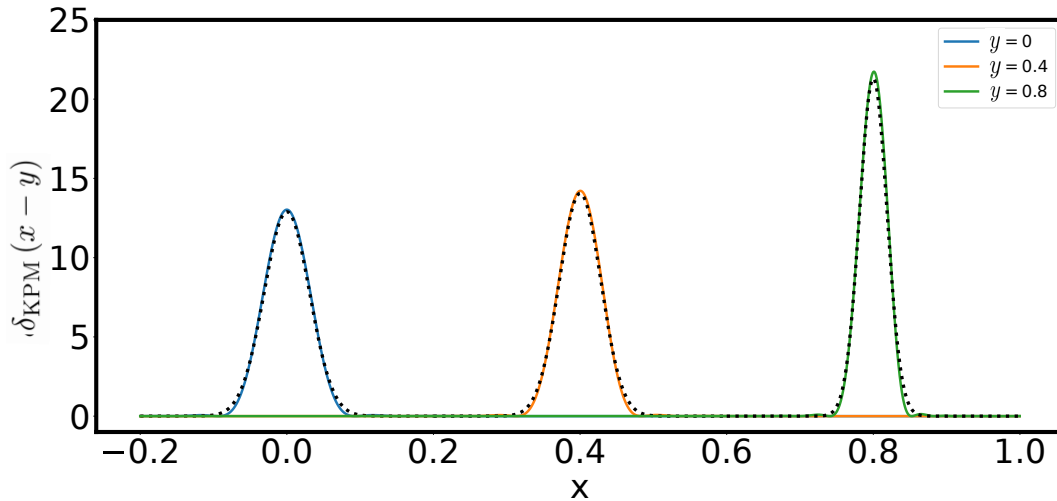


Figure 2.10: Expansion of a Delta-Dirac $\delta(x - \mu)$ for different μ and $M = 100$. This exemplifies the change in width σ of the overlapping Gaussian, as we approach the boundaries it starts to narrow greatly.

Now we turn our attention to the fourth central moment. Since we only want to analyse the behavior of the tails, we will set the Dirac- δ at the center, $y = 0$, so that the odd moments are null and only the $\langle x^4 \rangle$ is necessary to study. Remembering the trick we used, we can write x^4 as a combination of the Chebyshev polynomials, that is

$$x^4 = \frac{T_4(x) + 4T_2(x) + 3}{8}, \quad (2.3.19)$$

and the integral becomes

$$\langle x^4 \rangle = \int_{-1}^1 dx \cdot x^4 \cdot \delta_{\text{KPM}}(x - y) = \frac{g_4^J T_4(y) + 4g_2^J T_2(y) + 3}{8}, \quad (2.3.20)$$

and evaluating the Chebyshev polynomials at $y = 0$ we then have

$$\langle x^4 \rangle = \frac{g_4^J - 4g_2^J + 3}{8}. \quad (2.3.21)$$

We only need to evaluate the corresponding Jackson coefficients, it give us

$$\langle x^4 \rangle = \frac{1}{2} \sin^2 \left(\frac{\pi}{M+1} \right) + \frac{1}{2} \cdot \frac{M-2}{M+1} \cdot \cos \left(\frac{2\pi}{M+1} \right) \sin^2 \left(\frac{\pi}{M+1} \right), \quad (2.3.22)$$

and in large M we can, once again, Taylor expand the expression in first order, getting

$$\langle x^4 \rangle \sim \frac{\pi}{2} \cdot \frac{1}{M+1}.$$

Then, the kurtosis is given by

$$\text{Kurt} \{ \delta_{\text{KPM}}(x) \} = \frac{\langle x^4 \rangle}{\langle x^2 \rangle^2} = \frac{1}{2\pi} \cdot \frac{(M+1)^2}{M} \xrightarrow{M \rightarrow \infty} \frac{M}{2\pi},$$

which grows linearly with M . This is not in agreement with the exponential decrease of a Gaussian distribution, for

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2} \right\},$$

the kurtosis is simply given by

$$\text{Kurt} \{ G(x) \} = 3.$$

This result is of importance because if we want to measure quantities with a Dirac- δ using the Kernel Polynomial Method with the Jackson Kernel then we may fall in the mistake of considering it has the same tails as a regular Gaussian, because of the close resemblance as indicated above. This is not the case as seen in Fig. 2.11, and can contaminate the results greatly.

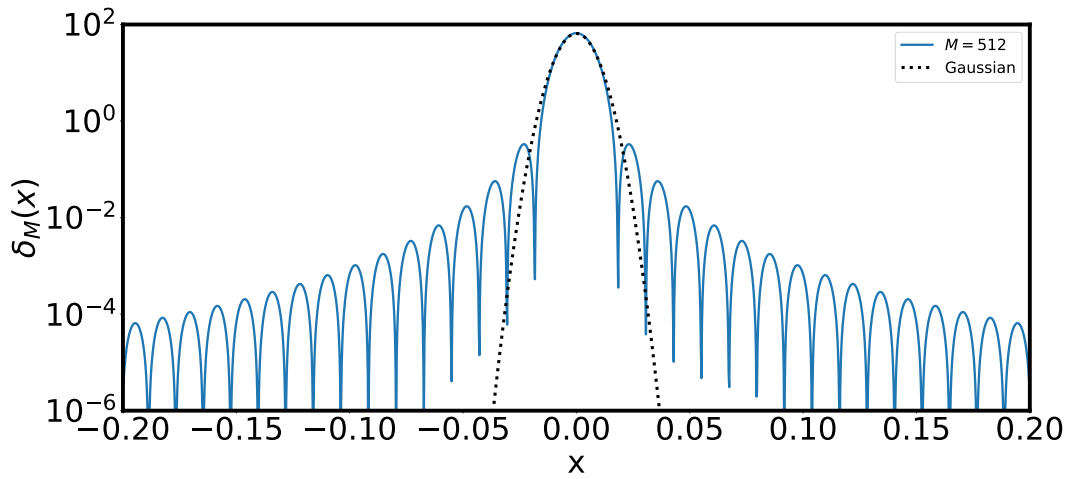


Figure 2.11: Expansion of a Delta-Dirac $\delta(x)$ with 512 moments. This figure shows the remarkable difference in the tails when we expand and compare directly to its Gaussian as discussed.

2.4 Expansion for Operators

2.4.1 Basic Notions

In the above chapters we discussed rather heavily the expansion of functions $f(x)$ and the use of kernels to smooth discontinuous functions. Now we turn to the real application of such method, calculation of quantities with operators. Essentially, we basically make the “upgrade” of

$$x \rightarrow \mathcal{H}, \quad (2.4.1)$$

where x is a scalar quantity and $\mathcal{H}$ an operator, which normally throughout this thesis will be the Hamiltonian of a quantum system. Then, our objective is to expand $f(\mathcal{H})$ in Chebyshev polynomials, which we learned that for a scalar function $f(x)$ is given by

$$f(x) = c_0 + 2 \sum_{m=1}^{\infty} c_m T_m(x), \quad (2.4.2)$$

where c_m are the Chebyshev coefficients given by

$$c_m = \int_{-1}^1 dx \cdot \frac{1}{\sqrt{1-x^2}} f(x) T_m(x). \quad (2.4.3)$$

For a $f(\mathcal{H})$ we then have

$$f(\mathcal{H}) = c_0 + 2 \sum_{m=1}^{\infty} c_m T_m(\mathcal{H}), \quad (2.4.4)$$

We also learned that for scalar functions $f(x)$, x needs to be in the range $x \in [-1, 1]$ for expansion to work, the analogue for operators $\mathcal{H}$ is that we need to guarantee that its eigenvalues are bounded in the same manner, $\epsilon_\alpha \in [-1, 1]$. This can always be achieved, for operators with bounded spectrum, with a rescaling of the

operator $\mathcal{H}$ with parameters b and a ,

$$\mathcal{H} \leftarrow \frac{\mathcal{H} - b}{a}, \quad (2.4.5)$$

where a and b are given by

$$a = \frac{\epsilon_{\max} - \epsilon_{\min}}{2 - \epsilon}, \quad (2.4.6)$$

$$b = \frac{\epsilon_{\max} + \epsilon_{\min}}{2}, \quad (2.4.7)$$

and $\epsilon_{\max}$, $\epsilon_{\min}$ are the maximum and minimum eigenvalues of $\mathcal{H}$, ϵ is just a small cut-off number to avoid divergences and stability problems.

In another light, we can add the closure relation $\sum_{\zeta} |\zeta\rangle\langle\zeta| = \mathbb{I}$ to the expansion $f_{\text{KPM}}(\mathcal{H})$, where $|\zeta\rangle$ are the the eigenvectors of the operator $\mathcal{H}$. Therefore, with $c_0 = 1$,

$$f_{\text{KPM}}(\mathcal{H}) = \mathbb{I} + 2 \sum_{\alpha} \sum_{m=1}^{\infty} c_m T_m(\mathcal{H}) |\zeta\rangle\langle\zeta| \quad (2.4.8)$$

$$= \left(1 + 2 \sum_{\alpha} \sum_{m=1}^{\infty} c_m T_m(\epsilon_{\alpha}) \right) |\zeta\rangle\langle\zeta|, \quad (2.4.9)$$

which basically means we the constructing of the expansion with the eigenvalues of $\mathcal{H}$.

The recursion relation,

$$T_{m+2}(x) = 2xT_{m+1}(x) - T_m(x), \quad (2.4.10)$$

is also valid for operators,

$$T_{m+2}(\mathcal{H}) = 2\mathcal{H} \cdot T_{m+1}(\mathcal{H}) - T_m(\mathcal{H}). \quad (2.4.11)$$

This relation can be proven by simply introducing the closure relations, it goes as follows,

$$\begin{aligned} T_{m+2}(\mathcal{H}) &= \sum_{\alpha} T_{m+2}(\mathcal{H}) |\zeta\rangle\langle\zeta| \\ &= \sum_{\alpha} T_{m+2}(\epsilon_{\alpha}) |\zeta\rangle\langle\zeta|, \end{aligned} \quad (2.4.12)$$

and now since $T_{m+2}(\epsilon_{\alpha})$ is a scalar we can use the previous proven recursion, giving us

$$T_{m+2}(\mathcal{H}) = \sum_{\alpha} \{2\epsilon_{\alpha} T_{m+1}(\epsilon_{\alpha}) - T_m(\epsilon_{\alpha})\} |\zeta\rangle\langle\zeta|, \quad (2.4.13)$$

which is, as we stated

$$T_{m+2}(\mathcal{H}) = 2\mathcal{H} \cdot T_{m+1}(\mathcal{H}) - T_m(\mathcal{H}), \quad (2.4.14)$$

hence proven.

The main advantage of this comes from the fact that for most applications we need to compute are the projections of $f(\mathcal{H})|\alpha\rangle$, where $|\alpha\rangle$ does not need to be in the eigenbasis of the operator $\mathcal{H}$, or scalars $\langle\beta|f(\mathcal{H})|\alpha\rangle$, and that can be done without diagonalizing using this technique.

The advantage is due to the recursion relation, by using Eq. (2.4.4) and projecting onto $|\alpha\rangle$ we have

$$f_{\text{KPM}}(\mathcal{H})|\alpha\rangle = c_0|\alpha\rangle + 2 \sum_{m=1}^{\infty} c_m T_m(\mathcal{H})|\alpha\rangle, \quad (2.4.15)$$

and the ket $T_m(\mathcal{H})|\alpha\rangle$ can be calculated by the use of the recursion relations,

$$T_{m+2}(\mathcal{H})|\alpha\rangle = 2\mathcal{H} \cdot T_{m+1}(\mathcal{H})|\alpha\rangle - T_m(\mathcal{H})|\alpha\rangle. \quad (2.4.16)$$

This can be thought in terms of vectors we use the notation $|\alpha\rangle_m = T_m(\mathcal{H})|\alpha\rangle$, then, the recursion relations is the following

$$|\alpha\rangle_{m+2} = 2\mathcal{H}|\alpha\rangle_{m+1} - |\alpha\rangle_m. \quad (2.4.17)$$

The time consuming part essentially comes from the recursion relation but it is still very fast, because, on one hand, we are dealing with subtraction of vectors, and the other operation, $\mathcal{H}|\alpha\rangle_{m+1}$ is fast because normally $\mathcal{H}$ is sparse, meaning it scales linearly with the dimension of the operator. For $\langle\beta|f(\mathcal{H})|\alpha\rangle$ one only needs to project the vector $|\alpha\rangle_{m+2}$ onto $|\beta\rangle$, giving us a scalar which we denominate μ_{m+2} .

$$\mu_{m+2} = \langle\beta|\alpha\rangle_{m+2}. \quad (2.4.18)$$

Hence we have a method that computes spectral quantities of operators which is linear with the dimension of $\mathcal{H}$ and in the number of moments M used.

2.4.2 A Trick to Compute Moments

We can further reduce the computation time for obtaining the moments μ_{m+2} if $|\beta\rangle = |\alpha\rangle$ [6], that is, they are simply given by

$$g_{m+2} = \langle\alpha|\alpha\rangle_{m+2}. \quad (2.4.19)$$

It is possible because we can exploit another relation between the Chebyshev polynomials, which states that

$$2T_m(x)T_n(x) = T_{m+n}(x) + T_{|m-n|}(x). \quad (2.4.20)$$

For $m = n$ and $m + 1 = n$ we have

$$\begin{aligned} 2T_m^2(x) &= T_{2m}(x) + 1 \\ 2T_m(x)T_{m+1}(x) &= T_{2m+1}(x) + T_1(x), \end{aligned} \quad (2.4.21)$$

if we then upgrade this to operators and project in the $|\alpha\rangle$ basis, in the same manner we did previously, we obtain

$$\begin{aligned} |\alpha\rangle_{2m} &= 2T_m(\mathcal{H})|\alpha\rangle_m - |\alpha\rangle_0 \\ |\alpha\rangle_{2m+1} &= 2T_m(\mathcal{H})|\alpha\rangle_{m+1} - |\alpha\rangle_1, \end{aligned} \quad (2.4.22)$$

projecting further onto $|\alpha\rangle$ we easily see why this reduces the computation time,

$$\begin{aligned} g_{2m} &= 2\langle\alpha_m|\alpha_m\rangle - g_0 \\ g_{2m+1} &= 2\langle\alpha_{m+1}|\alpha_m\rangle - g_1, \end{aligned} \quad (2.4.23)$$

instead of using the recursion relation once again to evaluate $T_m(\mathcal{H})|\alpha\rangle_m$ in Eq. (2.4.22), we use the vector $|\alpha\rangle_m$, which is already computed, and just project it according to the above equations. From one evolution we extract two moments, effectively reducing the time of computation by half.

2.4.3 Error Analysis for the Trace of an Operator

We found in the previous chapter that the moments of the trace of operators can be computed via a stochastic approach, one now questions if it is viable to do it this way, or in more technical words, how well it converges to the desired value, which is a cornerstone aspect to study for such methods. We shall use complex random variables with the same statistics as previously mentioned,

$$\begin{aligned} \langle\xi_i\rangle &= 0 \\ \langle\xi_i\xi_j\rangle &= 0 \\ \langle\xi_i^*\xi_j\rangle &= \delta_{ij}, \end{aligned} \quad (2.4.24)$$

with the random vector defined as

$$|\xi\rangle = \sum_i \xi_i |i\rangle,$$

and consider a general hermitian operator $\mathcal{B}$, defining its elements as $\mathcal{B}_{ij} = \langle i|\mathcal{B}|j\rangle$. With this, we define the random variable θ as

$$\theta = \langle\xi|\mathcal{B}|\xi\rangle. \quad (2.4.25)$$

Its average is the trace of $\mathcal{B}$, which can be trivially seen if we use the definition of the random vectors

$$\langle\theta\rangle = \sum_{i,j} \langle\xi_i^*\xi_j\rangle \mathcal{B}_{ij}, \quad (2.4.26)$$

and using the statistical properties of the random complex numbers we defined above,

$$\langle\theta\rangle = \sum_{i,j} \delta_{ij} \mathcal{B}_{ij} = \sum_i \mathcal{B}_{ii} = \text{Tr}(\mathcal{B}). \quad (2.4.27)$$

But this just proves that, on average, θ will converge to the trace of $\mathcal{B}$. We now study the variance δ of θ , defined as

$$\delta^2 = \langle|\theta|^2\rangle - |\langle\theta\rangle|^2. \quad (2.4.28)$$

With this, it is just a matter of computing $\langle\theta^2\rangle$,

$$\langle|\theta|^2\rangle = \langle\xi|\mathcal{B}|\xi\rangle^* \langle\xi|\mathcal{B}|\xi\rangle = \sum_{i,j} \sum_{k,l} \langle\xi_i\xi_j^*\xi_k^*\xi_l\rangle \mathcal{B}_{ij}^* \mathcal{B}_{kl}, \quad (2.4.29)$$

rearranging and using that $\mathcal{B}$ is hermitian,

$$\langle \theta^2 \rangle = \sum_{i,j} \sum_{k,l} \langle \xi_i^* \xi_j \xi_k \xi_l^* \rangle \mathcal{B}_{ij} \mathcal{B}_{lk}. \quad (2.4.30)$$

The main problem here is the computation of the correlator $\langle \xi_k^* \xi_l \xi_i \xi_j^* \rangle$. This objects are explained in detail in the appendix A.3, here we just state the result

$$g_{ijkl} \langle \xi_i^* \xi_j \xi_k \xi_l^* \rangle = g_{iiii} \cdot \delta_{ij} \delta_{ik} \delta_{il} \langle |\xi_i|^4 \rangle + g_{iikk} \cdot \delta_{ij} \delta_{kl} \bar{\delta}_{ik} + g_{ikik} \cdot \delta_{ik} \delta_{jl} \bar{\delta}_{ij}, \quad (2.4.31)$$

where g is a general function with 4 indexes, δ_{ij} the delta Kronecker and $\bar{\delta}_{ij} = 1 - \delta_{ij}$. Direct application tells us

$$\langle \theta^2 \rangle = \langle |\xi_i|^4 \rangle \sum_i \mathcal{B}_{ii}^2 + \sum_{i,k} \bar{\delta}_{ik} \mathcal{B}_{ii} \mathcal{B}_{kk} + \sum_{i,j} \bar{\delta}_{ij} \mathcal{B}_{ij} \mathcal{B}_{ji} \quad (2.4.32)$$

and applying the $\bar{\delta}_{ij}$ we obtain

$$\langle \theta^2 \rangle = \langle |\xi_i|^4 \rangle \sum_i \mathcal{B}_{ii}^2 + \sum_{i,k} \mathcal{B}_{ii} \mathcal{B}_{kk} - 2 \cdot \sum_i \mathcal{B}_{ii}^2 + \sum_{i,j} \mathcal{B}_{ij} \mathcal{B}_{ji}, \quad (2.4.33)$$

rearranging some terms,

$$\langle \theta^2 \rangle = \left(\langle |\xi_i|^4 \rangle - 2 \right) \sum_i \mathcal{B}_{ii}^2 + \left\{ \sum_i \mathcal{B}_{ii} \right\}^2 + \sum_{i,j} \mathcal{B}_{ij} \mathcal{B}_{ji}, \quad (2.4.34)$$

finally, we just identify once again the trace of the operator $\mathcal{B}$ and directly substract the missing $|\langle \theta \rangle|^2$ from the definition of variance. This procedure results in

$$\delta^2 = \sum_{i,j} \mathcal{B}_{ij} \mathcal{B}_{ji} + \left(\langle |\xi_i|^4 \rangle - 2 \right) \sum_i \mathcal{B}_{ii}^2, \quad (2.4.35)$$

we can further identify the matrix-matrix product in the term $\sum_{i,j} \mathcal{B}_{ij} \mathcal{B}_{ji}$, this will result in the following

$$\delta^2 = \text{Tr}(\mathcal{B}^2) + \left(\langle |\xi_i|^4 \rangle - 2 \right) \sum_i \mathcal{B}_{ii}^2. \quad (2.4.36)$$

The above expression constitutes an interesting results which we can further analyze. The second term can always be removed by means of a clever choice of the probability distribution for ξ , therefore, the main contribution will come from the term $\text{Tr}(\mathcal{B}^2)$. If $\mathcal{B}$ is a sparse matrix, we can infer that $\text{Tr}(\mathcal{B}^2) \sim D$, where D is the dimension of the operator $\mathcal{B}$, therefore, the relative error will go as

$$\frac{\delta}{\langle \theta \rangle} \sim \frac{1}{\sqrt{RD}}, \quad (2.4.37)$$

where we added the term R corresponding to the number of different realizations of ξ , following the central limit theorem. The amazing property from this is that it is self-averaging, meaning, the larger the dimension of the operator, the less its error to the mean is. This is a powerful method of evaluating traces of sparse matrices.

2.5 Self-Averaging of the Density of States

2.5.1 Error Analysis of the Density of States

In this section we will apply the error analysis to the quantity of the Density of States (DoS). From the previous analysis on stochastic evaluation of traces, we learned that the evaluation with random vectors results in a self-averaging behaviour. The calculations of the self-averaging aspect of traces was done for a general sparse operator $\mathcal{B}$. The loss of a clear physical interpretation arises due to the generality surrounding the physical quantity represented by $\mathcal{B}$. Consequently, we try to give some insight of the physical interpretation behind the self-averaging aspect of the DoS.

Let $\mathcal{H}$ be the Hamiltonian of a system and a real-space basis $\{|\mathbf{r}\rangle\}$, the density of states (DOS) is defined by

$$\rho(E) = \frac{1}{D} \sum_{\mathbf{r}} \langle \mathbf{r} | \delta(E - \mathcal{H}) | \mathbf{r} \rangle, \quad (2.5.1)$$

where E is the energy and D the dimension of the Hilbert Space. By using the eigenbasis $\{|\alpha\rangle\}$ of the Hamiltonian and inserting the closure relation identity, $\sum_{\alpha} |\alpha\rangle \langle \alpha| = 1$ in Eq. (2.5.1) we get

$$\begin{aligned} \rho(E) &= \frac{1}{D} \sum_{\mathbf{r}} \sum_{\alpha} \langle \mathbf{r} | \delta(E - \mathcal{H}) | \alpha \rangle \langle \alpha | \mathbf{r} \rangle \\ &= \frac{1}{D} \sum_{\mathbf{r}} \sum_{\alpha} |\psi_{\alpha}(\mathbf{r})|^2 \delta(E - E_{\alpha}), \end{aligned} \quad (2.5.2)$$

where E_{α} is the eigenvalue of H , and using the normalization of the wave function $\sum_{\mathbf{r}} |\psi_{\alpha}(\mathbf{r})|^2 = 1$ we easily see that

$$\rho(E) = \frac{1}{D} \sum_{\alpha} \delta(E - E_{\alpha}), \quad (2.5.3)$$

which is just the common definition of the density of states. We can see that the Density of States is just the trace of an operator

$$\rho(E) = \frac{1}{D} \text{Tr}(\delta(E - \mathcal{H})) \quad (2.5.4)$$

therefore we just do the analysis of the operator $1/D \cdot \delta(E - \mathcal{H})$ in the expression of the error. It will be

$$\delta^2 = \frac{1}{D^2} \left\{ \text{Tr}(\delta(E - \mathcal{H})^2) + (\langle |\xi_i|^4 \rangle - 2) \sum_{\mathbf{r}} [\langle \mathbf{r} | \delta(E - \mathcal{H}) | \mathbf{r} \rangle]^2 \right\} \quad (2.5.5)$$

It is best to expand and trace in the eigenbasis and note that we have the Local Density of States definition in the second term

$$\rho_{\mathbf{r}}(E) = \frac{1}{D} \langle \mathbf{r} | \delta(E - \mathcal{H}) | \mathbf{r} \rangle, \quad (2.5.6)$$

therefore

$$\delta^2 = \frac{1}{D^2} \sum_{\alpha} f_{\alpha}^2 + (\langle |\xi|^4 \rangle - 2) \sum_{\mathbf{r}} \rho_{\mathbf{r}}^2(E). \quad (2.5.7)$$

In Eq. (2.5.7) we get an unconventional object given by $\sum_{\alpha} f_{\alpha}^2 = \sum_{\alpha} \delta(E - E_{\alpha})^2$. The way to deal with such object is with the use of regularized Dirac- δ approximants. It is rather useful to approximate the Dirac- δ with a Gaussian because of the Kernel Polynomial Method, in which f_{α} can be described very closely by a

Gaussian.

Let us define $f_\alpha(\sigma)$ as

$$f_\alpha(\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(E-E_\alpha)^2}{\sigma^2}}, \quad (2.5.8)$$

where σ is the width of the Gaussian. This object has the property of

$$f_\alpha = \lim_{\sigma \rightarrow 0} f_\alpha(\sigma), \quad (2.5.9)$$

and with this, we can express $f_\alpha^2(\sigma)$ as

$$f_\alpha^2(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(E-E_\alpha)^2}{\sigma^2}}, \quad (2.5.10)$$

which can be written in terms of the original $f_\alpha(\sigma)$ as

$$f_\alpha^2(\sigma) = \frac{1}{\sqrt{2\pi}\sigma} f_\alpha(\sigma/\sqrt{2}). \quad (2.5.11)$$

We can confidently say that $\delta(E - E_\alpha)^2$ is, when approximated as a Gaussian of width σ , a new Gaussian of reduced width $\sigma/\sqrt{2}$ and rescaled with an extra normalization factor of $1/\sqrt{2\pi}\sigma$. Then, Eq. (2.5.7) can be written as

$$\delta^2 = \frac{1}{\sqrt{2\pi}D\sigma} \rho(E, \sigma/\sqrt{2}) + (\langle |\xi|^4 \rangle - 2) \sum_r \rho_r^2(E, \sigma), \quad (2.5.12)$$

where we upgraded the DoS and LDoS quantities to approximations with Gaussians. It is quite obvious that the first term $\frac{1}{\sqrt{2\pi}D\sigma} \rho(E, \sigma/\sqrt{2})$ is self-averaging with the system size D and the error increases as the precision σ increases, which agrees with previous numerical knowledge. The second term, $\sum_r \rho_r^2(E, \sigma)$, we could make it disappear by a wise choice of the fourth moment of ξ , $\langle |\xi|^4 \rangle$, but we are interested in the general behavior regardless of the distribution of ξ .

To characterize $\sum_r \rho_r^2(E, \sigma)$ we have to subdivide in two limiting situations: (i) a system having perfectly extended eigenstates and (ii) a localized system.

Extended States

The case for extended states is quite easy because the wavefunctions are Bloch States, $\Psi_k = \frac{1}{\sqrt{D}} e^{ik \cdot r}$ and $|\Psi_k|^2 = \frac{1}{D}$. Therefore, by the definition of the LDoS,

$$\sum_r \rho_r^2(E, \sigma) = \frac{1}{D^2} \sum_r \sum_{\alpha, \beta} |\Psi_\alpha(r)|^2 |\Psi_\beta(r)|^2 f_\alpha f_\beta = \frac{1}{D^3} \sum_{\alpha, \beta} f_\alpha f_\beta = \frac{1}{D} \rho^2(E, \sigma), \quad (2.5.13)$$

we conclude that the second term is just the DoS of the same width σ and weighted by $1/D$. It is trivial to see that for this case the quantity is also self-averaging, decreasing with $1/D$.

Localized States

For localized states the Bloch Theorem no longer applies, and a more careful approach is required to evaluate $\sum_{\mathbf{r}} \rho_{\mathbf{r}}^2(E, \sigma)$. We begin this analysis, as always, using the definition of the LDoS with its wavefunctions.

$$\sum_{\mathbf{r}} \rho_{\mathbf{r}}^2(E, \sigma) = \frac{1}{D^2} \sum_{\mathbf{r}} \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta}, \quad (2.5.14)$$

we can separate the summation in two cases, $\alpha = \beta$ and $\alpha \neq \beta$

$$\sum_{\mathbf{r}} \rho_{\mathbf{r}}^2(E, \sigma) = \frac{1}{D^2} \sum_{\mathbf{r}} \left\{ \sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^4 f_{\alpha}^2 + \sum_{\alpha \neq \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta} \right\}. \quad (2.5.15)$$

We quickly identify the first term in Eq. (2.5.15) as the Inverse Participation Ration of a state α when we sum over $\mathbf{r}$. It would also leave a f_{α}^2 , which we already identified as a rescaled DoS, this gives us

$$\sum_{\mathbf{r}} \rho_{\mathbf{r}}^2(E, \sigma) = \frac{1}{\sqrt{2\pi} D^2 \sigma} \sum_{\alpha} f_{\alpha} \left(\sigma / \sqrt{2} \right) \text{IPR}_{\alpha} + \sum_{\mathbf{r}} \sum_{\alpha \neq \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta}. \quad (2.5.16)$$

To analysis further this object we need to know the general behavior of localized states. It is known the asymptotic behavior is $|\psi_{\alpha}(\mathbf{r})|^2 \sim \frac{1}{4\xi} \exp\left(-\frac{1}{2\xi} |\mathbf{r} - \mathbf{r}_{\alpha}|\right)$, where $\mathbf{r}_{\alpha}$ is the real-space maximum of $|\psi_{\alpha}(\mathbf{r})|^2$ and ξ its localization length. By setting the second summation $\sum_{\alpha \neq \beta}$ and knowing the general behavior of the functions we can argue that, since $\mathbf{r}_{\alpha}$ will be different for each wave-function, $|\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2$ will always suppress itself. We can see that by expressing the product of the wave-functions,

$$|\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 \leq \frac{1}{16\xi^2} e^{-\frac{1}{2\xi} \{|\mathbf{r} - \mathbf{r}_{\alpha}| + |\mathbf{r} - \mathbf{r}_{\beta}|\}}, \quad (2.5.17)$$

and, as you can see, increasing $\mathbf{r}$ and fixed $\mathbf{r}_{\alpha}$ and $\mathbf{r}_{\beta}$ suppresses this quantity and by summing in $\mathbf{r}$ the same happens because of $\mathbf{r}_{\alpha} \neq \mathbf{r}_{\beta}$.

It is left to understand the term IPR_{α} in Eq. (2.5.16). This term will not cause a problem, because the IPR as a very particular behavior for localized states. It will first decay as $1/D$ and then it will reach a plateau because $D > \xi$. Therefore, since we already have a factor $1/D^2$ in term of analysis of Eq. (2.5.16) the IPR_{α} will at most change it to a decay of $1/D^3$ and then return to $1/D^2$ when $D > \xi$. Therefore, even for localized states, this quantity is self-averaging as we expected.

2.5.2 Example of KPM in the DoS

Let us now exemplify with a real application of the Kernel Polynomial Method. The density of states is defined by

$$\rho(E) = \frac{1}{D} \sum_{\mathbf{r}} \langle \mathbf{r} | \delta(E - \mathcal{H}) | \mathbf{r} \rangle, \quad (2.5.18)$$

where E is the energy and D the dimension of the Hilbert Space. According to the Eq. (2.1.12), the moments will be given by

$$\mu_m = \int_{-1}^1 dE \cdot \rho(E) T_m(E) = \frac{1}{D} \text{Tr}(T_m(\mathcal{H})), \quad (2.5.19)$$

that is, the Chebyshev moments of the DoS will be each given by the trace of $T_m(\mathcal{H})$.

To compute the trace of operations we need to know the whole operator or its eigenvalues, which requires the diagonalization. Of course, this breaks our purposes of the efficiency of the KPM. To get around this problem, researchers have devised a stochastic approach to compute traces [6, 54]. The trace will be approximately equal to

$$\text{Tr}(T_m(\mathcal{H})) \approx \langle \xi | T_m(\mathcal{H}) | \xi \rangle, \quad (2.5.20)$$

where $|\xi\rangle$ is a random vector in the basis of $|\mathbf{r}\rangle$, it is defined as

$$|\xi\rangle = \sum_r \xi_r |\mathbf{r}\rangle, \quad (2.5.21)$$

and ξ_r is a complex random number satisfying the following properties

$$\begin{aligned} \langle \xi_r \rangle &= 0 \\ \langle \xi_r \xi_{r'} \rangle &= 0 \\ \langle \xi_r^* \xi_{r'} \rangle &= \delta_{rr'}. \end{aligned} \quad (2.5.22)$$

With this and the lessons learned to compute such objects, $\langle \xi | T_m(\mathcal{H}) | \xi \rangle$, we can therefore calculate the moments μ_m and get the the DoS for a given Hamiltonian $\mathcal{H}$ by performing the following truncated series

$$\rho_{\text{KPM}}(E) = \frac{1}{\pi \sqrt{1-E^2}} \left\{ g_0^J + 2 \sum_{m=1}^M g_m^J \mu_m T_m(E) \right\}, \quad (2.5.23)$$

where the factors g_m^J is the Jackson Kernel. The only additional thing in the number of averages form the moments.

The calculation of objects via a stochastic approach will be the main study in the next chapter.

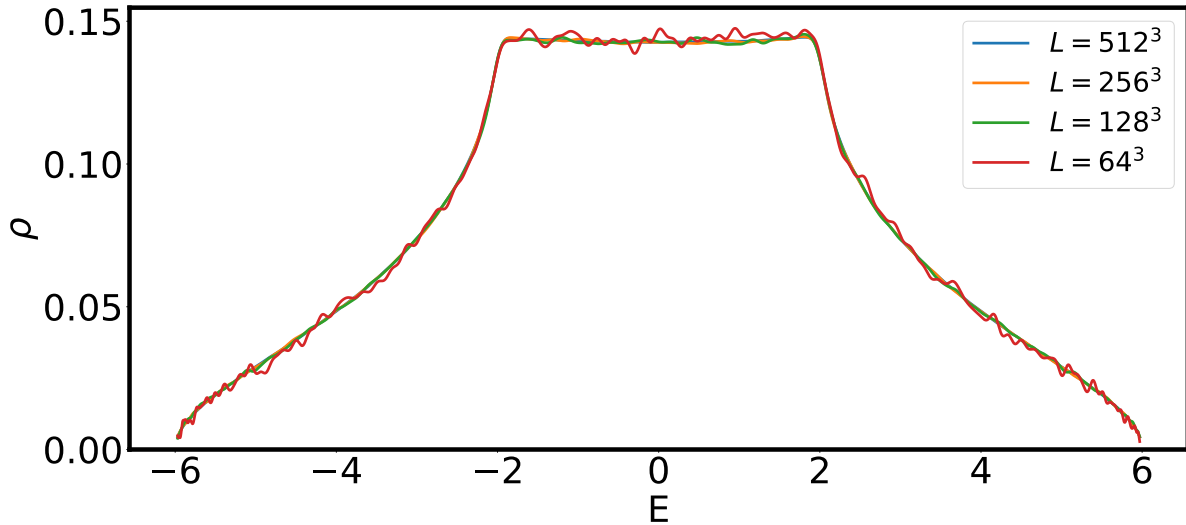


Figure 2.12: Example of a reconstruction of DoS for the clean tight-binding nearest neighbor model for a cubic lattice. There is only a single sample on the random vectors and number of disorders.

Chapter 3

Implementation in Graphical Processing Units

3.1 Basic of GPU Architecture

The GPUs differs from the CPUs. It provides a much higher instruction throughput and memory bandwidth for a similar price. The first remark we want to point out is the number of threads. In a CPU we are accustomed to dozens of process units. However, for a GPU, it is always in the order of thousands or tens of thousands, albeit each unit is much slower and less capable than a CPU one. For simple operations, such as scalar multiplication, this type of computing units offers a great opportunity to process big amounts of data [7, 10, 55, 56].

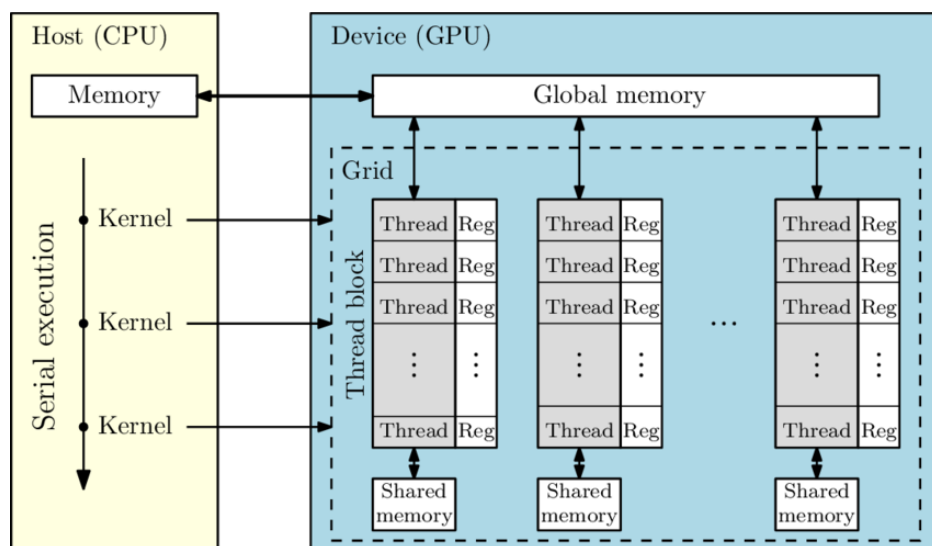


Figure 3.1: Very simple schematic of a GPU architecture and kernel execution. Kernel is just a piece of code that is executed in the GPU [10]. Image taken from [11].

Due to its complexity, a hierarchical structure for the hardware becomes necessary. The basic unit within this hierarchy is referred to as a **block**, which is a collection of threads and always in quantity of a power of two, usually between 128-512 threads per block. Notably, the number of threads can exceed the count of physical CUDA cores, with the hardware automatically assigning threads to the cores. Each **block** possesses access to

a fast memory bank known as **shared memory** which is accessed collectively by threads within the block. It is important to highlight that the **shared memory** of one **block** remains inaccessible to other blocks.

A collection of **blocks** is called a **grid**. This has no particular feature other than a mapping to the GPU. The size of **blocks** and **grids** is determined by software preferences.

The other type of memory is the **global memory/VRAM**. This memory is akin to the **RAM** for a CPU processor. It has moderate speed if accessed by GPU threads during the execution of the code, and is extremely slow if we access it via the CPU. Although various other memory types are available, they are not necessary throughout this chapter. One must always try to avoid the communication of **RAM-VRAM** due to its extreme slowness. This hierarchy is exemplified in Fig. 3.1.

A common ground that one also needs to understand before programming in a GPU is the concept of **warps** and **memory coalescing**. **Warps** is a hardware specification and cannot be altered in code, as it is with the block dimension and grid dimension. It is a group of threads, usually 16-32, that execute and load memory simultaneously.

Memory coalescing is the process of a fast loading from the **global memory** of the adjacent memory elements, that is, if we are accessing the N -th elements, the coalescing will guarantee that we also load the $N \pm m$ -th. Then, if one wants to have an efficient program, the user needs to be sure that this group of threads, **warps**, will be always executing and processing adjacent data, so that it only loads once the required memory from the global memory.

3.2 The Basic Algorithms for the KPM in GPU

3.2.1 Summation of Vectors

The most basic operation one needs to learn for an KPM implementation in GPUs is how effective they are to sum vectors. It is trivially to understand why the type operation is almost made to be exploited in the GPUs. In the CPU case, if we have a vector with N elements, we could break it in N/T chunks, where T is the number of threads, as shown in Fig. 3.3. Each thread would compute its partial sum, meaning a partial calculation that exists in the register of a single thread, and load it to the global memory. This type of operation is one of the simplest to understand and implement, since it requires almost no thought and has no synchronization problems between threads.

C	CUDA C
<pre> void add(int *a, int *b, int *c) { int tid = 0; while (tid < N) { c[tid] = a[tid] + b[tid]; tid += 1; } } </pre>	<pre> __global__ void add(int *a, int *b, int *c) { int tid = threadIdx.x + blockIdx.x * blockDim.x; while (tid < N) { c[tid] = a[tid] + b[tid]; tid += blockDim.x * gridDim.x; } } </pre>

Figure 3.2: Difference between pieces of code that perform the summation of vectors; Left) CPU; Right) GPU, each thread executes it. Notice the jump in the looping variable, for the CPU it is just a skip of 1 while in the GPU it is $\text{blockDim} \times \text{gridDim}$, which is the number of threads. Taken from [10].

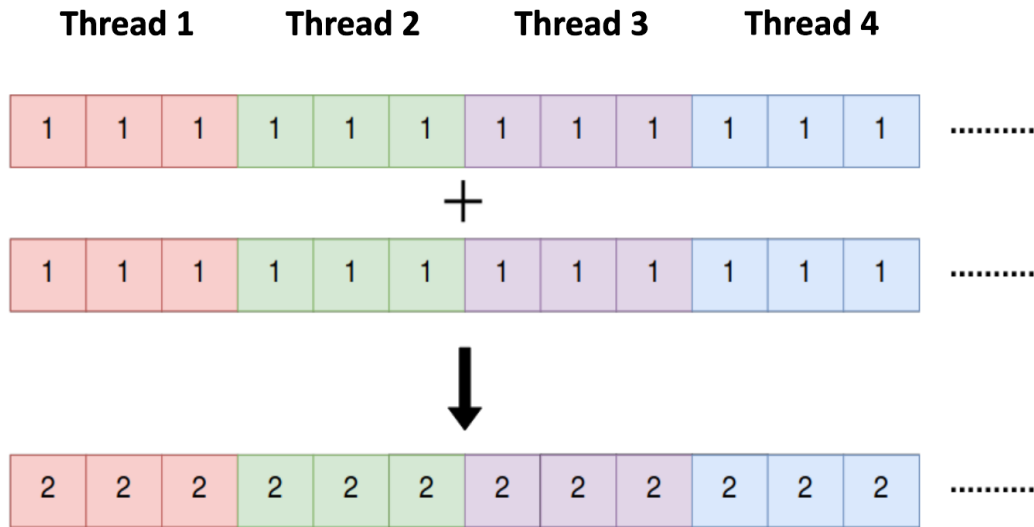


Figure 3.3: Schematics for the parallelization of vector summation in CPUs.

For the GPUs, the same thought process is applied, and basically, since they have a number of threads in the order of the thousands vs the dozens in CPUs, one can understand why the vectors summations are much faster in GPUs. The chunks processed by each thread are much smaller, as shown in Fig. 3.4. A sample code between the difference of the two is shown in Fig. 3.2.

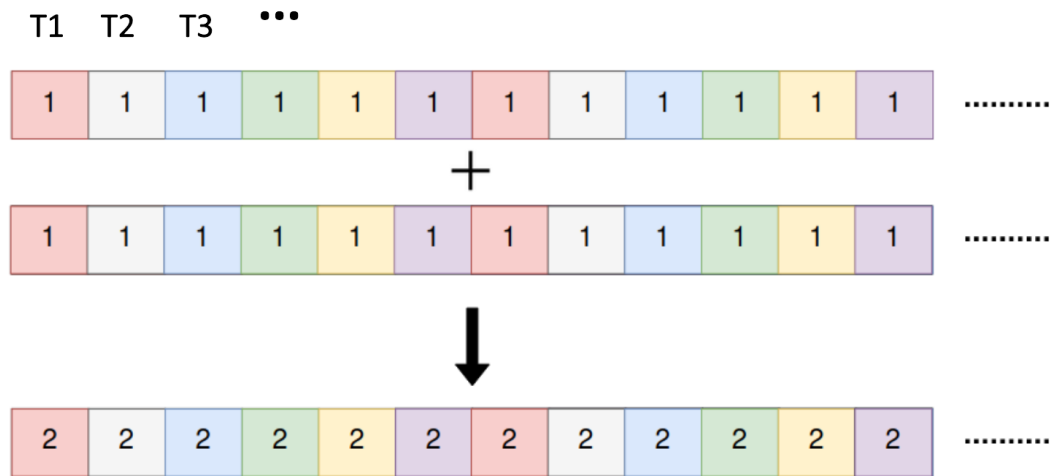


Figure 3.4: Schematics for the parallelization of vector summation in GPUs.

3.2.2 Reduction Algorithm

One of the other fundamental operations in the KPM is the dot product $\langle \alpha | \beta \rangle$ to obtain the moments, for example, for the Density of States. This type of operations are a perfect opportunity to parallelize due to the nature of the dot product, as we saw with the vectors, we can do the same here requiring now only one additional step. This step will be the hard one to understand because it involves the collapse of a higher rank object, vectors, to a single rank, a scalar. We denominate this type of operations as reductions.

In parallel computing, namely in CPUs, one is subject to asynchronization between threads. If the type of

operation does not care at all between parallel segments, as we saw with the summation of vectors, one says it is an embarrassingly parallel algorithm since it only requires the division of work between threads. If however the operation depends on the work of other threads, this is no longer the case and requires some thought to avoid collision problems.

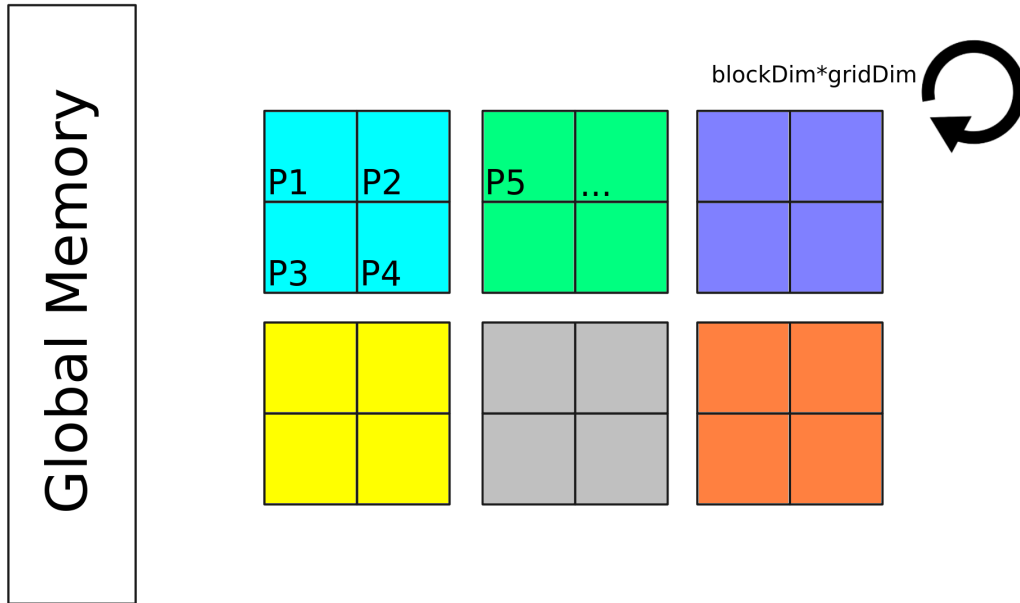


Figure 3.5: Schematics for the operation of the dot product between vectors before the reduction. Each thread will accumulate its partial sum P_n while iterating in the vector in chunks of size $\text{blockDim} \times \text{gridDim}$. Each color represents a block, which has shared memory, of 4 threads, while in real life GPUs its on the order of 128 – 512 threads.

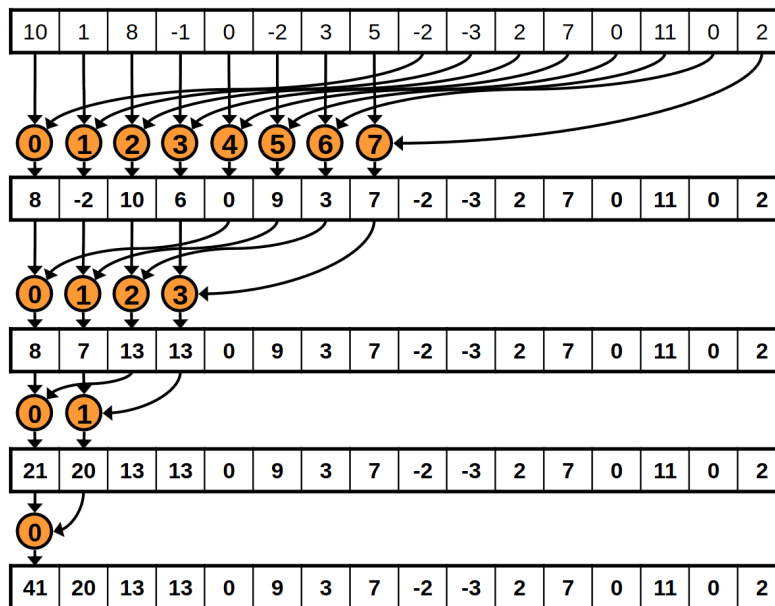


Figure 3.6: Visual representation of thread activation in a block with shared memory while performing the reduction. Taken from [10]

Asynchronization between threads means they are working independently, without knowing the state of other threads. This causes problems if we are trying to access a specific memory/register and changing its value. If two threads access at the same time and add a value, the result will be wrong because the thread did not wait for the other one to complete its task. Therefore, since we are trying to achieve a reduction, that is to say to sum a large number of floats into a single register, we are subject to this problem.

This problem exists in CPUs and is greatly aggravated in GPUs because of the sheer number of threads. Fortunately, in CUDA C there is an instruction that tells the threads to synchronize, by waiting the full completion of the thread works. This is of great importance to understand the Reduction algorithm.

First we perform the same operation we did for the summations of vectors, but each thread will not load its partial sum to the global memory, instead it will load it to the shared memory of its corresponding block, Fig. 3.5 shows the accumulation of the partial sums. The task now is to devise a clever way to retrieve the partial sum of all the blocks while using the maximum number of threads to avoid being idle.

Python

`np.sum(a*b)`

CUDA C

```
__global__ void dot( float *a, float *b, float *c ) {
    __shared__ float cache[threadsPerBlock];

    int tid = threadIdx.x + blockIdx.x * blockDim.x;
    int cacheIndex = threadIdx.x;

    float temp = 0;
    while (tid < N) {
        temp += a[tid] * b[tid];
        tid += blockDim.x * gridDim.x;
    }

    // set the cache values
    cache[cacheIndex] = temp;

    // synchronize threads in this block
    __syncthreads();

    // for reductions, threadsPerBlock must be a power of 2
    // because of the following code
    int i = blockDim.x/2;
    while (i != 0) {
        if (cacheIndex < i)
            cache[cacheIndex] += cache[cacheIndex + i];
        __syncthreads();
        i /= 2;
    }

    if (cacheIndex == 0)
        c[blockIdx.x] = cache[0];
}
```

Figure 3.7: Left: Sane Python code for the dot product. Right: Real CUDA C code showing how to apply the reduction algorithm. Black box: Accumulates the partial sums of the vector to vector elements products; Blue box: Loading to corresponding index of shared memory of the block; Red box: Algorithm collapsing the partial sums into a single number for each block in $\text{blockDim}/2$ loops. Taken from [10].

After completing the partial sums, loading them into shared memory, and synchronizing, the next step involves

summing the memory element $m + \text{blockDim}/2$ to the m , where m is at maximum $\text{blockDim}/2$. This process will collapse the necessary data for the dot product in half while maintaining half the threads of the block occupied. It is best to follow Fig. 3.6 to visualize the algorithm. Next, we basically do the same but reducing the maximum memory element by half, that is, we would now only collapse the summation $n + \text{blockDim}/4$ to the n -th element in the next iteration, where n is now at maximum $\text{blockDim}/4$. We keep repeating this process until we reach the final stage where we have one only element of the shared memory containing the sum of all of the other elements.

However, this approach doesn't entirely resolve the issue because in the end we would have an array of the size of the number of blocks with its final sum. Although it is much smaller than the number of threads and we can just instead load this memory element to the global memory and then sum the smaller array. This process is shown in Fig. 3.7.

This is a heavy thought process to do just a dot product, but it will be worth it in the end.

3.2.3 Memory Optimization

While comprehending algorithms for the problem at hand is crucial, it raises a question: Why invest effort in understanding and implementing these algorithms when readily available libraries such as Numpy in Python or its GPU counterpart, Cupy, can execute the same tasks? The thought process behind this lies in the control we gain over memory management by opting for a manual implementation.

By writing the algorithm ourselves we have unparalleled control over the memory accesses and thus we can improve its performance. Memory management has a bigger part to consider when programming in GPUs than in CPUs, as it should be, as they were developed to process vast amounts of data fast.

While employing simple GPU libraries can yield performance improvements, it might not harness the full potential with a precise memory control. The critical problem is if one frequently accesses the RAM of the computer several times during the code execution. Since the communication between GPU-CPU is the slowest part, by accessing several times we lose the performance gained by the GPU, or even worsens if we only used the CPU, a visual representation of memory speed between components is shown in Fig. 3.8.

The other aspect against libraries is their inherent inflexibility. An example is the dot product, in the Cupy library it only accepts two vectors and nothing more. This restriction means that we always need the vectors prepared before hand, killing some performance.

By crafting the kernels ourselves we get rid of constraints and gain the flexibility to implement customized solutions. Since reducing the memory accesses is ideal we can put the the recursion formula of the KPM,

$$|\alpha\rangle_{m+2} = 2\mathcal{H}|\alpha\rangle_{m+1} - |\alpha\rangle_m, \quad (3.2.1)$$

inside of Reduction algorithm, the black box in Fig. 3.7. This is one of the main reasons this low level type of programming is preferable to pre made libraries if one is considering squeezing the last bit of performance of the hardware.

CPU compilers are rather advanced and they can avoid small mistakes in memory management. Such is not always the case in GPUs, small improvements of how we access the memory are relevant and improve performance. The best example is how we do the recursion formula for our problem, in Algorithm 3.1 is what one could first do, step by step doing the requirements and transformation for the recursion. This, however,

is not the ideal way and the compiler will not correct it. Each small operation requires access to the Global Memory and hinders the execution of the code. By joining several operation, as with Algorithm 3.2, one reduces the Global memory and improves the speed. If one uses some of its elements in several parts of the code it is preferable to create temporary variables to store the memory value in the close registers of the threads. Each of this steps and carefulness seems rather dull but it is necessary since the compiler does not correct it, and every one less access to the memory can mean 2% – 10% improvements.

One now just has to combine this knowledge and craft the best kernel for the operations.

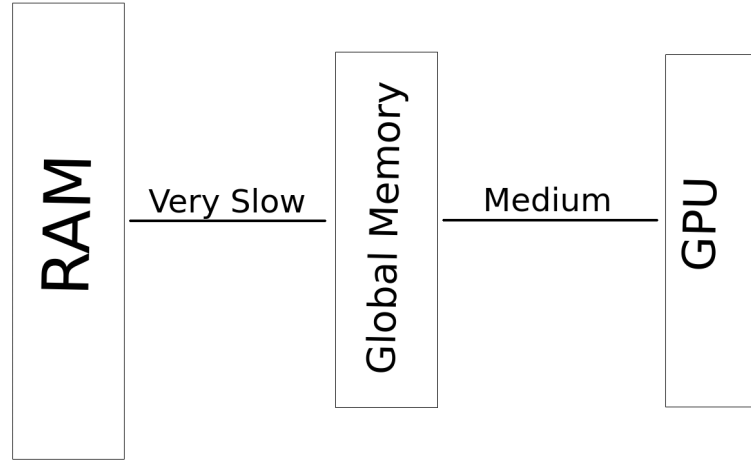


Figure 3.8: Visual representation of memory speed between components in a computer. CPU to GPU (RAM to Global Memory) is the slowest, while the access of the GPU to its global memory is of moderate speed. The fastest memory is not represented, since it is the Shared Memory, that is on chip.

Algorithm 3.1 Naive approach to perform the recursion and get KPM moments.

```

for  $r = 0$  to  $R - 1$  do
   $|v\rangle \leftarrow |\text{random vector}\rangle$ 
  Initialization
  for  $m = 1$  to  $M/2$  do
     $\text{swap}(|w\rangle, |v\rangle)$ 
     $|u\rangle \leftarrow \mathcal{H}|v\rangle$ 
     $|u\rangle \leftarrow |u\rangle - b|v\rangle$ 
     $|w\rangle \leftarrow -|w\rangle$ 
     $|w\rangle \leftarrow |w\rangle + 2a|v\rangle$ 
     $\alpha_{2m} \leftarrow \langle v|v\rangle$ 
     $\alpha_{2m+1} \leftarrow \langle w|v\rangle$ 
  end for
end for
  
```

Algorithm 3.2 A GPU adapted algorithm to obtain the KPM moments.

```

for  $r = 0$  to  $R - 1$  do
   $|v\rangle \leftarrow |\text{random vector}\rangle$ 
  Initialization
  for  $m = 1$  to  $M/2$  do
     $\text{swap}(|w\rangle, |v\rangle)$ 
     $|w\rangle \leftarrow 2a(\mathcal{H} - b)|v\rangle - |w\rangle$ 
     $\alpha_{2m} \leftarrow \langle v|v\rangle$ 
     $\alpha_{2m+1} \leftarrow \langle w|v\rangle$ 
  end for
end for
  
```

3.3 Performance Study

Having studied thoroughly the principles of GPU programming and the application of various methodologies to improve the efficiency, let us see if it was worth all this trouble. We will compare it to one of the most established CPU based programs in regards to spectral calculations, the Quantum Kite [53] .

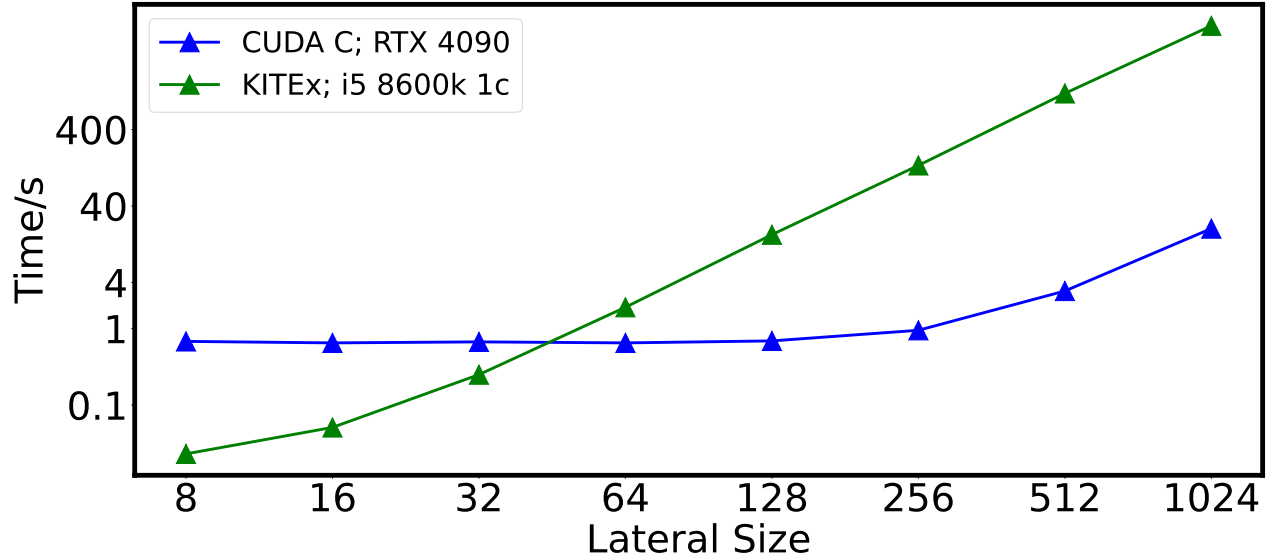


Figure 3.9: Times between different KPM programs. The computation example was of the DoS of 3D tight-binding lattice with 1024 moments.

From the depicted in Fig. 3.9, several noteworthy observations can be drawn. Primarily, the computation time versus the lateral size of the lattice in the initial sizes, $8^3 - 128^3$, of the GPU code, exhibits a zero scaling. This is easily explained by looking at the hardware specification, the RTX 4090 has 16384 Cuda cores. This then means that only sizes exceeding this threshold activate all of the cores, and therefore we would observe no scaling up to this point. However, this alone does not explain the stagnation in scaling up to 128^3 because 16384 cores corresponds to 3D lattice of only 25^3 . One other important explanation is that, up to 128^3 , the GPU utilizes very few memory. At this size each vector uses only 16MB of memory. What we then see is the computation of the DoS of the sizes $25^3 - 128^3$ being faster than the memory overhead, only when we surpass this size, 128^3 , we start to feel the scaling $\mathcal{O}(L^3)$, of the problem.

After that, the vectors starts to occupy hundreds of MB and really justifies the use of GPUs by using the great memory bandwidth it provides. Although, at low sizes, one is better suited with the CPU based programs. Each one has its benefits.

We want to be very cautious when comparing both software's. Although the plot suggests, at 1024^3 , that the GPU based program is about 450x faster than its CPU counterpart, we need to analyze it better. We only used one core on the KITE software, with the intention that this program has a perfect linear scaling in the number of cores used [53] and it is easy to extrapolate. Therefore, this plot essentially implies that the GPU code is equivalent to using 450 single cores of an i5-8600k.

To extrapolate to other kind of CPU processors is hard because it will not only depend on its kind, but also the RAM memory and communication between CPU-RAM, we would have to come up with a single score for

this hardware to really start comparing between other types of hardware.

Needless to say, being equivalent to this number of processors, 450 single cores of an i5 8600k, is an incredible achievement in terms of scientific computation. With only one GPU, which definitely cheaper than hundreds of CPUs, we output incredible compute power.

Chapter 4

Computational Methods for the Local Density of States

4.1 Computational Complexity of the LDoS with KPM

The LDoS, which measures the density of states in a given point, $\mathbf{r}$, of the quantum system is defined by

$$\rho_{\mathbf{r}}(E) = \frac{1}{D} \langle \mathbf{r} | \delta(E - \mathcal{H}) | \mathbf{r} \rangle, \quad (4.1.1)$$

Using the eigenbasis of $\mathcal{H}$,

$$\rho_{\mathbf{r}}(E) = \frac{1}{D} \sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^2 \delta(E - E_{\alpha}), \quad (4.1.2)$$

which is the more familiar expression of the LDoS. With the KPM, as explained in Chapter 2, one would just compute the moments μ_m

$$\mu_m = \int_{-1}^1 dE \cdot \rho_{\mathbf{r}}(E) T_m(E) = \frac{1}{D} \int_{-1}^1 dE \cdot \langle \mathbf{r} | \delta(E - \mathcal{H}) | \mathbf{r} \rangle T_m(E) = \frac{1}{D} \langle \mathbf{r} | T_m(\mathcal{H}) | \mathbf{r} \rangle, \quad (4.1.3)$$

and perform the Chebyshev expansion,

$$\rho_{\mathbf{r},\text{KPM}}(E) = \frac{1}{\pi \sqrt{1 - E^2}} \left\{ g_0^J + 2 \sum_{m=1}^M g_m^J \mu_m T_m(E) \right\}. \quad (4.1.4)$$

Although this approach is very efficient for one lattice point $\mathbf{r}$, being of complexity $\mathcal{O}(D)$, one would be stumped if we cared about the LDoS measurement in the whole lattice. Since we want to do statistics, the problem is very similar. We have to do point by point in the lattice to obtain the data. This effectively increases the complexity of the calculation to $\mathcal{O}(D^2)$.

Aiming to circumvent this important issue, we come up with a new way to compute the LDoS that is less expensive than the bare calculation of the LDoS in every point in space. Previously, there have been some work on the development of alternative method to calculate the LDoS, such as the work by Ju Li et. al. [57], which provided the LDoS for integer multiples of an energy at a time, $\rho_{\mathbf{r}}(n\omega)$. The approach we will implement is inspired by the stochastic trace of operators, explained in some detail in Chapter 2.

4.2 Stochastic LDoS

Having exposed the drawbacks of the normal way of computing the LDoS with the KPM in the previous section, we now attempt to improve the computational complexity, inspired by the stochastic trace evaluation of operators, which was reviewed in Chapter 2. Let ξ_r be a complex random variable drawn from a probability distribution with the same characteristics as the DoS case, that is $\langle \xi_r^* \xi_{r'} \rangle = \delta_{rr'}$ and $\langle \xi_r \rangle = 0$. Then, we can define another random variable as

$$\tilde{\rho}_r(E) = \frac{1}{D} \sum_{r'} \langle \mathbf{r} | \delta(E - H) | \mathbf{r}' \rangle \xi_r^* \xi_{r'} = \frac{1}{D} \xi_r^* \cdot \langle \mathbf{r} | \delta(E - H) | \xi \rangle, \quad (4.2.1)$$

where $|\xi\rangle$ is a random vector. One can immediately tell that the average of this quantity is the same as the exact LDoS,

$$\begin{aligned} \langle \tilde{\rho}_r(E) \rangle &= \frac{1}{D} \sum_{r'} \langle \mathbf{r} | \delta(E - H) | \mathbf{r}' \rangle \langle \xi_r^* \xi_{r'} \rangle = \frac{1}{D} \sum_{r'} \langle \mathbf{r} | \delta(E - H) | \mathbf{r}' \rangle \delta_{rr'} \\ &= \frac{1}{D} \langle \mathbf{r} | \delta(E - H) | \mathbf{r} \rangle = \rho_r(E). \end{aligned} \quad (4.2.2)$$

One quick remark we want to point out is that each evaluation of the above expression only gives the LDoS for a single point. When implementing the software to do the calculations, one can extract all the LDoS points by simple performing the Hadamard product between $\langle \xi |$ and $\delta(E - H) | \xi \rangle$, which essentially is a new vector made up of the product of corresponding elements in $\langle \xi |$ and $\delta(E - H) | \xi \rangle$, we denote this operation by the symbol $\circ$. The random LDoS variable now becomes $\tilde{\rho}_r(E) = \frac{1}{D} \langle \xi | \circ \delta(E - H) | \xi \rangle$. This way we will get all the LDoS values in each position r .

Now that we have all the foundations laid out by promoting the LDoS to a random variable, whose average corresponds to the true LDoS, and obtaining all the lattice points calculations with the stochastic evaluation of the Hadamard product. We question now if this approach is worth it. The answer to this can be found by analyzing the variance of $\tilde{\rho}_r(E)$. The definition of variance, considering the use of complex random vectors, is

$$\sigma_r^2 = \langle \tilde{\rho}_r^*(E) \tilde{\rho}_r(E) \rangle - |\langle \tilde{\rho}_r(E) \rangle|^2, \quad (4.2.3)$$

and before plugging $\tilde{\rho}_r(E)$ in the above expression we first diagonalize it, so that we can work with wavefunctions instead,

$$\tilde{\rho}_r(E) = \frac{1}{D} \sum_{\alpha} \sum_{r'} \Psi_{\alpha}(\mathbf{r}) \Psi_{\alpha}^*(\mathbf{r}') f_{\alpha} \xi_r^* \xi_{r'}, \quad (4.2.4)$$

where $\Psi_{\alpha}(\mathbf{r})$ is the wavefunction with corresponding energy E_{α} and $f_{\alpha} = \delta(E - E_{\alpha})$. For each term in the variance we would then have

$$\langle \tilde{\rho}_r^*(E) \tilde{\rho}_r(E) \rangle = \frac{1}{D^2} \sum_{\alpha, \beta} \sum_{r', r''} \Psi_{\alpha}^*(\mathbf{r}) \Psi_{\alpha}(\mathbf{r}') \Psi_{\beta}(\mathbf{r}) \Psi_{\beta}^*(\mathbf{r}'') f_{\alpha} f_{\beta} \langle |\xi_r|^2 \xi_{r'}^* \xi_{r''} \rangle, \quad (4.2.5)$$

$$|\langle \tilde{\rho}_r(E) \rangle|^2 = \frac{1}{D^2} \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta}. \quad (4.2.6)$$

By looking at the expression we see that main hard part comes from the correlator $\langle |\xi_r|^2 \xi_{r'}^* \xi_{r''} \rangle$. We leave its

calculation and further manipulations in the Appendix A.3 and A.2. The final result of this computation is

$$\sigma_r^2 = \frac{1}{D^2} \left[\sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^2 f_{\alpha}^2 + \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta} \left\{ \langle |\xi|^4 \rangle - 2 \right\} \right], \quad (4.2.7)$$

and we can extract the definition of the LDoS in the second term, resulting in

$$\sigma_r^2 = \frac{1}{D^2} \sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^2 f_{\alpha}^2 + \rho_r(E)^2 \left\{ \langle |\xi|^4 \rangle - 2 \right\}. \quad (4.2.8)$$

We once again encounter the term f_{α}^2 , as we did when analyzing the error in the random evaluation of the DoS, which we saw that can be expressed as a Gaussian with width Δ in the same manner. A reminder that the trick for understanding f_{α}^2 does not stop us from using other Dirac- δ regularizations such that the square gives what we expect, such as the square root of a Lorentzian. This would mean, however, that the original LDoS, $\rho_r(E) = \frac{1}{D} \sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^2 f_{\alpha}$, would have no physical meaning due to f_{α} not being a Dirac- δ in the limit of zero width, only the variance.

With the Gaussian as a regularization with width Δ this comes about with an amazing result,

$$\sigma_r^2 = \frac{1}{\sqrt{2\pi}D\Delta} \rho_r(E, \Delta/\sqrt{2}) + \rho_r(E, \Delta)^2 \left\{ \langle |\xi|^4 \rangle - 2 \right\}, \quad (4.2.9)$$

in which the variance of $\tilde{\rho}_r(E)$ is a function of $\rho_r(E)$ itself, with different resolutions Δ and the first term is self-averaging with system size D and goes to zero in the thermodynamic limit. In the limit of infinite size and fixed resolution we could say

$$\lim_{D \rightarrow \infty} \sigma_r^2 = \rho_r(E, \Delta)^2 \left\{ \langle |\xi|^4 \rangle - 2 \right\}. \quad (4.2.10)$$

However, we are more interested in the numerical aspect of this problem. By choosing a probability distribution with $\langle |\xi|^4 \rangle = 2$, which is possible by using a combination of two gaussian probability distributions,

$$p(\xi) = \frac{1}{2\sqrt{2\pi}\sigma} \left\{ e^{-\frac{1}{2} \frac{(\xi-\mu)^2}{\sigma^2}} + e^{-\frac{1}{2} \frac{(\xi+\mu)^2}{\sigma^2}} \right\}, \quad (4.2.11)$$

where $\mu = 2^{-\frac{1}{4}}$ and $\sigma = \sqrt{1 - 1/\sqrt{2}}$, we have that

$$\sigma_r^2 = \frac{1}{\sqrt{2\pi}D\Delta} \rho_r(E, \Delta/\sqrt{2}), \quad (4.2.12)$$

and the relative variance is equal to

$$\sigma_{r,\text{relative}}^2 = \frac{1}{\sqrt{2\pi}D\Delta} \frac{1}{\rho_r(E, \Delta/\sqrt{2})}. \quad (4.2.13)$$

From this we take three main conclusions relative to the numerical performance of the method. It is self averaging with system size D , which is good for our purposes of many atom quantum systems, but in contrast we lose precision of our result with increasing resolution Δ of the LDoS. The third conclusion, and the most

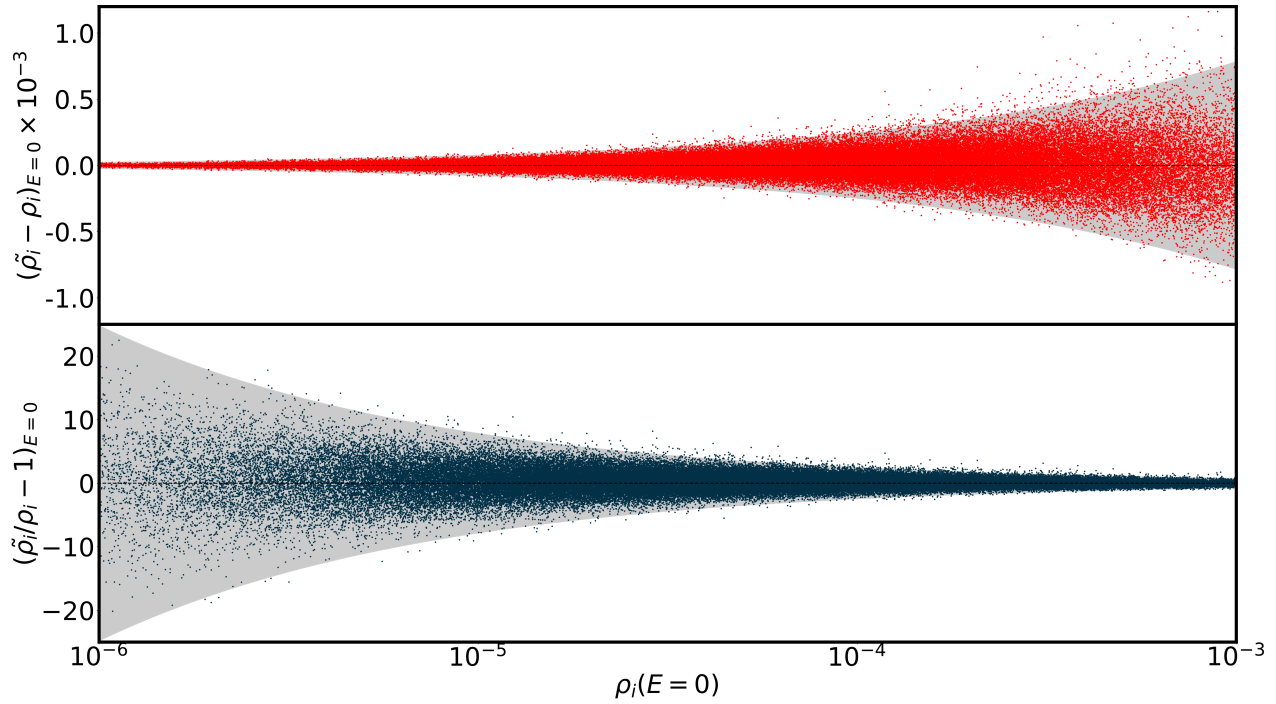


Figure 4.1: Top) Absolute error of the LDoS measured through the sample mean. Bottom) Relative error of the LDoS measured through the sample mean. It is 64x64 Square Lattice Tight-Binding nearest neighbor model. The disorder strength is $W = 6.0$ and 30 disorder realizations with 512 random vectors. The light dark filled region is a 3σ error bar, corresponding to a 99% certainty.

important one, is that the relative error is inversely proportional to $\sqrt{\rho_r(E, \Delta/\sqrt{2})}$, the lesser the LDoS the higher the error is. This comes with its set of problems, since we will be dealing with the LDoS of disordered quantum systems and its probability distributions this is a great disadvantage because we could have LDoS values of the order 10^{-6} but its error could be 2000%, which completely negates the usability of the method for this kind of work. Fig. 4.1 shows an exact agreement of the analytical formula for the variance deduced.

4.3 Measuring the LDoS through the Error Bar

In the previous section we showed that it is possible to measure the LDoS through a stochastic evaluation akin to the DoS. This, however, has a drawback that the error is inversely proportional to $\sqrt{\rho_r(E, \Delta/\sqrt{2})}$. If one wishes to study disordered systems we will encounter low values of the LDoS. This combined with the error bar associated with the previous method defeats the purpose of it. In this section we provide a way to go around this problem.

Normally, in a simulation we compute the LDoS with several random vectors and accumulate its sample average, given by

$$\bar{\rho}_r = \frac{1}{N} \sum_{j=1}^N \tilde{\rho}_{r,j}, \quad (4.3.1)$$

where $\tilde{\rho}_{r,j}$ is the j -th sample of random vectors. The new random variable, $\bar{\rho}_i$, follows, as we know, the central limit theorem. Stating that the variance of $\bar{\rho}_r$ follows σ^2/N , where σ^2 is the variance of $\tilde{\rho}_r$, given by Eq.

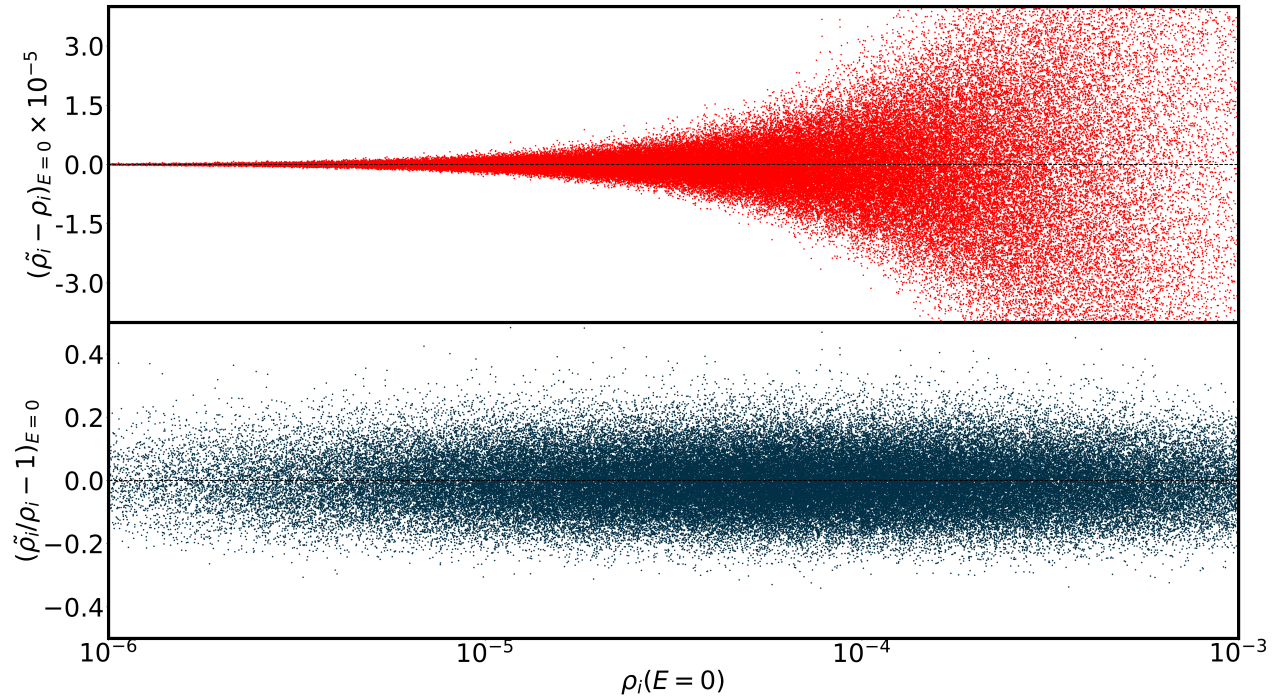


Figure 4.2: Top) Absolute error of the LDoS measured through the sample variance. Bottom) Relative error of the LDoS measured through the sample variance. It is 64x64 Square Lattice Tight-Binding nearest neighbor model. The disorder strenght is $W = 6.0$ and 30 disorder realizations with 512 random vectors.

(4.2.12). This of course allows us to reduce the error with arbitrary precision but it does not solve the relative error problem. We would need millions of average to get it to a reasonable degree, defeating the purpose in compute time.

Instead of writing down the LDoS as a mean value estimator we considered another approach. Since we know the analytical prediction of the variance of $\tilde{\rho}_r$ is given by Eq. (4.2.12), which is directly proportional to ρ_r , we could estimate it by using the sample mean,

$$\tilde{\sigma}_r^2 = \frac{1}{N} \sum_{j=1}^N (\tilde{\rho}_{r,j} - \bar{\tilde{\rho}}_r)^2. \quad (4.3.2)$$

This will be our new estimator for the LDoS, we will denote it by $R(N)$, where N is the number of samples. It is basically measuring the average through its error bar. In Fig. 4.2 we see this, while the absolute error is two orders of magnitude smaller than the measurement through the sample mean, the relative error stays somewhat constant throughout the whole range of LDoS.

In Fig. 4.3 we see the central limit theorem at work, where we measure the fluctuations of our new estimator (fluctuations of the fluctuations). The distributions get peaked around zero and the variance Δ^2 follows the known pattern $1/N$, where N is the number of samples. We did only 8192 random vectors at maximum, corresponding to the relative error of around 2.2% in the whole range of the LDoS. Therefore, this method compensates for any system bigger than 8192 lattice sites. Suffice is to say that we can get the LDoS in the whole lattice in linear time, $\mathcal{O}(L)$.

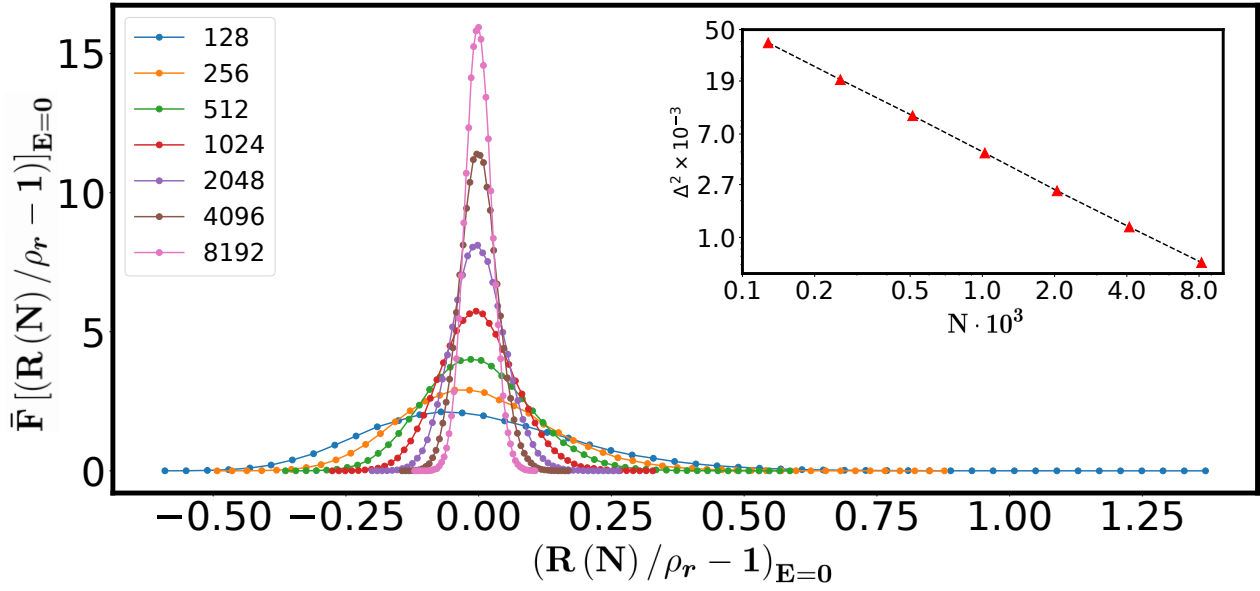


Figure 4.3: Distribution of the relative error of the $R(N)$ estimator for the LDoS with several different number of averages. The inset shows the progression of the width Δ of such distributions. The simulated system is a 64×64 Square Lattice Tight-Binding nearest neighbor model with disorder strength $W = 6.0$ and 50 realizations of disorder.

4.4 Importance of the Twisted Boundary Conditions in Statistics

The numerical aspect of the simulation forbids us to use infinite and continuum system sizes, this in turn forces us to use discretized models. Normally, one would use periodic boundaries conditions (PBC) on the lattice out of convenience. But this approach suffers from a problem if we study the LDoS statistics. In the clean limit, PBC can be imposed on the Hamiltonian $\mathcal{H}$ and be solved with the Bloch theorem. For the sake of an argument, we consider the 2D tight-binding nearest neighbor Hamiltonian, solving with the Bloch theorem, we get the equation for $E = 0$,

$$-2t \{\cos(k_x) + \cos(k_y)\} = 0, \quad (4.4.1)$$

where t is the hopping, k_x and k_y are the x and y momentum, respectfully. The PBC imposes that

$$k_x = \frac{2\pi n_x}{L} \quad k_y = \frac{2\pi n_y}{L}, \quad (4.4.2)$$

where L is the lateral size of the lattice and n_x, n_y integer values. For the equation $E = 0$ to be true one would need that

$$n_x + n_y = \frac{L}{2} \quad \vee \quad n_x - n_y = \frac{L}{2} \quad (4.4.3)$$

This means that L needs to be a multiple of 2 for solutions to exist. System sizes that are not multiple of 2 will not have solutions at $E = 0$. Therefore, the σ of the kernel in KPM will capture states that widely depend on the system size.

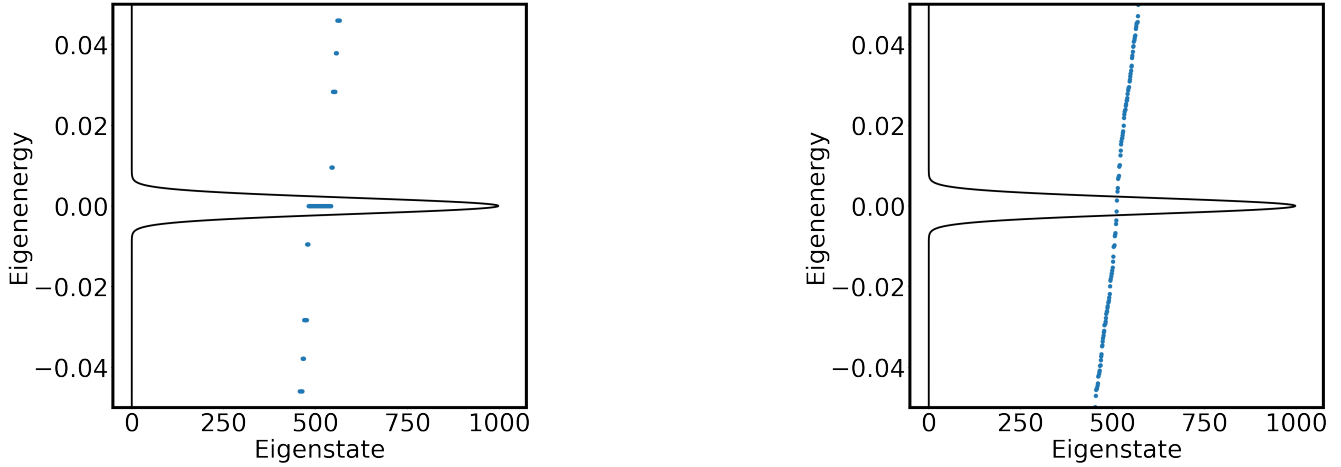


Figure 4.4: Diagonalization of a 2D tight-binding system with Left) PBC Right) TBC; The plotted gaussian represents the width σ of the kernel chosen, the states that fall in between are the selected.

This poses some obvious problems for the statistics to be comparable. One possible solution is the use of the Twisted Boundary Conditions (TBC). TBC are a product of the study of infinite systems where there exists a supercell. When we diagonalize it with the Bloch theorem, the Hamiltonian in momentum space becomes diagonal in blocks, where the block represents the Hamiltonian of the supercell.

Then, the trace of the quantities $\text{Tr}(f(\mathcal{H}))$, where $f(\mathcal{H})$ is a function of the Hamiltonian, can be written as $\text{Tr}(f(\mathcal{H})) = \sum_k \text{Tr}(f(\mathcal{H}_k))$, where $\mathcal{H}_k$ is the Hamiltonian of the block. We can estimate $\sum_k \text{Tr}(f(\mathcal{H}_k))$ by a Monte Carlo approach, where we randomly change the phase of the factor between the supercells, which are a leftover of the Bloch diagonalization. This also corresponds to changing the hoppings at the boundary by a phase in a finite system where the system size is one supercell,

$$-t_{\text{boundary}} \rightarrow -t_{\text{boundary}} \cdot e^{i\phi}, \quad (4.4.4)$$

where ϕ is a random phase between $[0, 2\pi]$.

Another insight is that when we impose the TBC on a finite system, the momentum k given by PBC will always suffer a shift of $\phi/2\pi L$.

$$k' = k + \frac{\phi}{2\pi L} \quad (4.4.5)$$

This is great because if we do many samples of random phases ϕ we “dilute” the momentum space, as shown for example in Fig. 4.4. The width σ of the kernel of the KPM will now see the continuum of the number of states, when the number of samples goes to infinity.

4.5 Parameters and Details of the LDoS Statistics

So far we created a new method that calculates the LDoS on the whole lattice in linear time and that has a constant relative error throughout the whole lattice. Then we reviewed and commented on the Twisted Boundary Conditions (TBC) and their necessity so that in measuring the LDoS we are consistent, guaranteeing the existence of states within the regularization of $\delta(E - E_\alpha)$. In this section, we explain and review the

methods used when computing the LDoS for different system sizes or energies, allowing comparison between results.

For starters, Schubert et al. [5], used the expansion of the LDoS directly from the KPM, meaning, it is a direct calculation of the operator $\rho_r(E) = \frac{1}{D} \langle \mathbf{r} | \delta(E - H) | \mathbf{r} \rangle$ as explained in Chapter 2.4. The direct expansion, using the Jackson Kernel, follows closely a Gaussian with a width σ of

$$\sigma^2 \sim \frac{\pi^2}{(M+1)^3} \left\{ M - (M-1) E^2 \right\}, \quad (4.5.1)$$

where M is the expansion order. In order to compare the statistics of equal relevance, one would need to adjust the expansion order M so that σ is constant throughout. This can be solved with numerical procedures for M .

On the other note, not only, using the above procedure, one would need to have attention to M but also the the number of states, N_k , within the Kernel. Meaning, if one wishes to study the statistics of different lattice sizes, one needs to guarantee there are always an equal number N_k between sizes. We can compute this with the following,

$$N_k(E) = N \cdot \int_{E-\sigma(E)}^{E+\sigma(E)} dE \rho_{\text{med}}(E), \quad (4.5.2)$$

where N is the number of lattice sites in the system and $\rho_{\text{med}}(E)$ the average DoS, the latter can be computed with the KPM with several disorder realizations.

Since we will be dealing if small σ we can just say that ,

$$N_k(E) = 2N\sigma\rho_{\text{med}}(E), \quad (4.5.3)$$

then, the σ , and consequently the expansion order M , can be known,

$$\sigma = \frac{N_k(E)}{2N\rho_{\text{med}}(E)}. \quad (4.5.4)$$

This would allows us to also judge the compute time given M . It is worth noting that, although in this thesis we do not explore quantum systems where the DoS vanishes at a certain energy E in the thermodynamic limit, we would have to perform a self-consistent evaluation of Eq. (4.5.2) in this case, because the value of the DoS will vary greatly with the expansion order M of the KPM.

The above explanation is how one would perform the study with the regular KPM expansion for the LDoS. However, in this thesis we developed a new stochastic method that allows us to get the LDoS throughout the whole lattice, and the resolution will be independent of the energy E if one uses a specific Kernel. The following results we use a Lorentzian kernel, since we will be first testing the claims by Alt'shuler et al. [3], and be attentive to Eq. (4.5.2) when we study the evolution of the statistics with system size.

Chapter 5

Results

5.1 Local DoS Fluctuations in One-Dimensional Systems

Up to this point we reviewed the Kernel Polynomial method, a numerical method that allow us to efficiently compute spectral quantities but not necessarily ideal to accurately compute local real-space quantities over the entire system. Later, we created a new efficient method to compute the LDoS in the whole lattice by measuring the fluctuations of itself. Furthermore, to improve even further the new LDoS method, we implemented this algorithm in the GPUs. This work will then be used to study the distributions of the LDoS on the numeral aspect to give a better insight.

Schubert et. al. [5] left some key ingredients to understand the mechanism behind the change of the LDoS probability distribution for systems that have a lateral size inferior than the localization length, which obviously raises the question of how it is possible to measure such differences in the distributions. Fortunately, we have analytical results by Al'tshuler et. al. [3], which will take on a center role in our understanding of the LDoS statistics for a one-dimensional chain for the local density of states, defined by

$$\tilde{\rho}_r(E) = \frac{1}{\rho(E)} \cdot \frac{1}{L} \sum_v |\Psi_v(r)|^2 \delta(E - E_v), \quad (5.1.1)$$

where L is length of the system (also the number of orbitals), and the extra factor of the DoS $\rho(E)$ is only there to guarantee that $\sum_r \tilde{\rho}_r(E) = 1$,

If the sample is finite the spectrum is discrete. It does not make much sense to speak of a probability distribution because, by the definition of the LDoS, we would capture only two values, it either exists at E_v or not due to the $\delta(E - E_v)$. Because of this, it is necessary to introduce a regularization function for the Dirac- δ . A Lorentzian is used and the LDoS becomes,

$$\rho_r(\sigma, E) = \frac{1}{\pi \rho(E)} \cdot \frac{1}{L} \sum_v |\Psi_v(r)|^2 \frac{\sigma}{(E - E_v)^2 + \sigma^2}, \quad (5.1.2)$$

where σ is the broadening parameter. Equation (5.1.2) can also be thought as doing a convolution to the probability density $|\Psi_v(r)|^2$ of the system, Eq. (5.1.1), given by,

$$\tilde{\rho}_r(\sigma, E) = \frac{1}{\pi} \int_{-\infty}^{\infty} dE' \tilde{\rho}_r(E') \frac{\sigma}{(E - E')^2 + \sigma^2}. \quad (5.1.3)$$

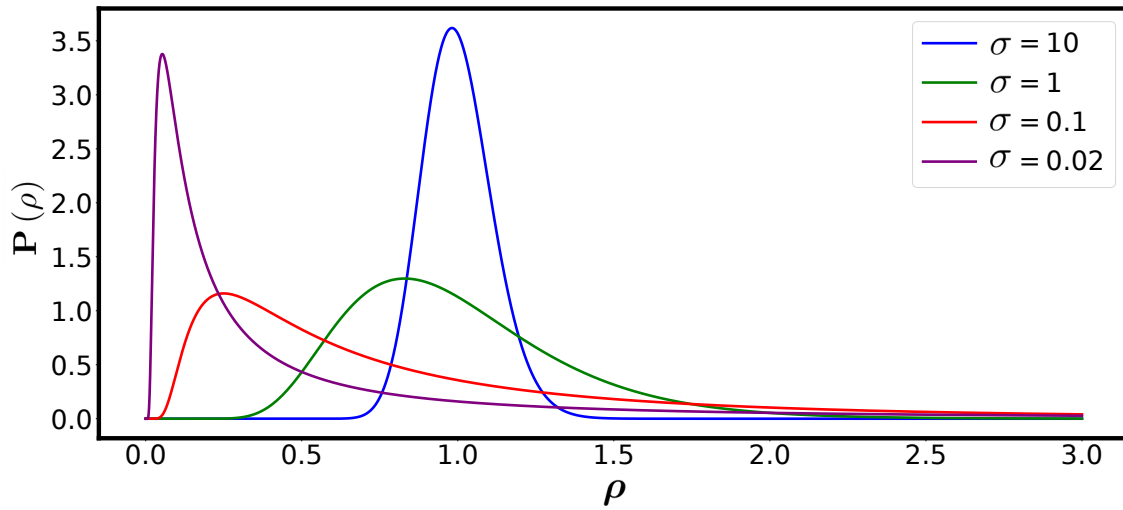


Figure 5.1: Example of the behavior shift of the Inverse-Gaussian distribution with $\tau = 1$ and varying σ .

We know that in one-dimension, all single-electron states are localized at any value of Anderson disorder [20]. The fluctuation moments of the LDoS can be computed for any finite σ by Berezinskii diagrammatic technique [3, 58]. Alt'shuler used this technique to calculate the moments $\langle \rho_r(\sigma, E)^n \rangle$ and reconstructed the probability distribution of the LDoS, having found that for a infinite system it does indeed depend on the broadening σ , and it results in an Inverse-Gaussian distribution,

$$P(\rho) = \sqrt{\frac{4\sigma\tau}{\pi\rho^3}} \exp\left(-4\frac{(\rho-1)^2}{\rho}\sigma\tau\right), \quad (5.1.4)$$

where σ is the broadening and τ the mean free path time¹ of the electron. The mean free path time is the typical time scale associated until an electron finds an obstacle and the mean free path, l , is the typical length scale. Following [3], the localization length, ξ , is related to these quantities and in 1D $\xi = 4l$ and $l = v_e\tau$, where v_e is the electron velocity in a 1D tight-binding chain. This distribution is interesting because it has different asymptotic behaviors depending on the parameter $4\sigma\tau$, as illustrated in Fig. 5.1. The distribution $P(\rho)$ decays exponentially fast for $\rho \rightarrow \infty$. If the parameter $4\sigma\tau \gg 1$ the contributions for the LDoS are wide in energy, capturing many states of the system, and therefore, the distribution will stay centered around 1 and its close to a Gaussian distribution. On the different limit, $4\sigma\tau \ll 1$, the behavior completely shifts. The distribution becomes extremely asymmetric, only capturing a few states of the system, and we have an accumulation near $\rho = 0$, this is expected due to the tails of the wavefunctions, for localized wavefunctions the tails have an asymptotic form $|\Psi_v(r)| = \exp(-|r - r_v|/\xi)$, where r_v is the location of the peak of $|\Psi_v(r)|^2$, where the value is maximum.

So far the work by Alt'shuler et. al. [3] has been purely analytical and the system used to arrive at Eq. (5.1.4) was a 1D semi-infinite one. The distribution should only be valid for systems where the size of the system, L , is many times greater than the mean free path, l , $L \gg l$.

The following results were obtained using a 1D nearest neighbor tight-binding quantum system, described by

¹We used this nomenclature to be consistent with the paper by Alt'shuler. This term is more often known as Elastic Scattering Time.

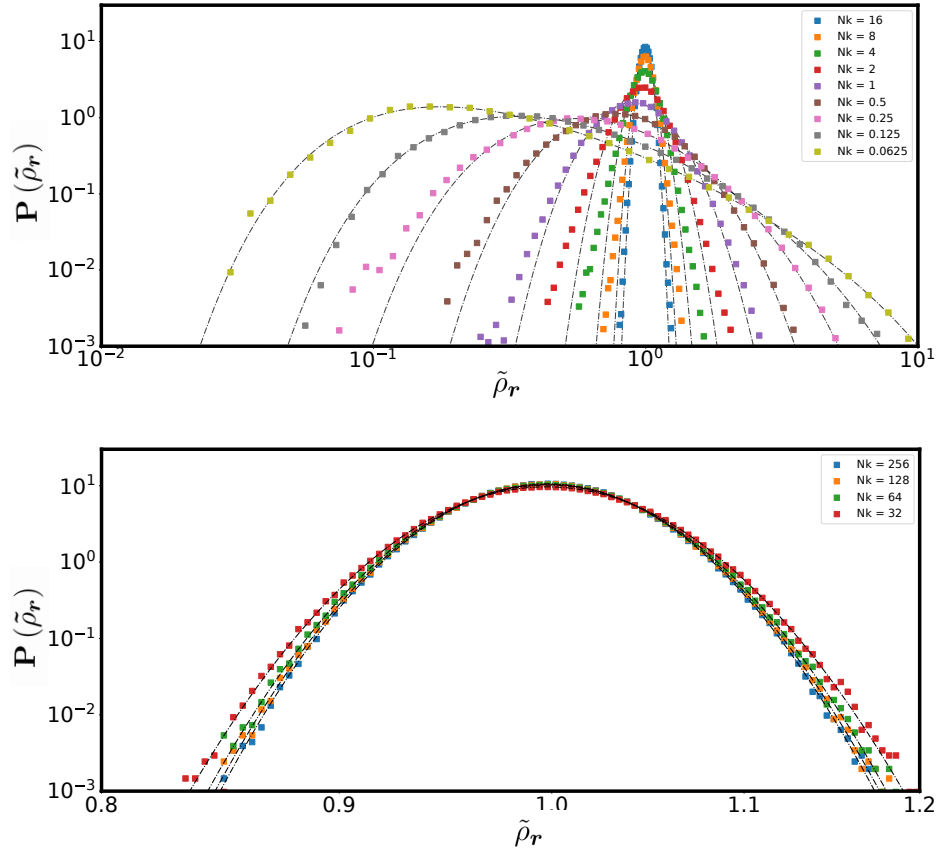


Figure 5.2: Local density of states statistics for a periodic 1D tight-binding chain nearest neighbor of size $L = 512000$ and Anderson disorder $W = 0.3$ with increasing broadening of the Lorentzian. Top) Intermediate and small broadening. Bottom) Large broadening.

the Hamiltonian,

$$\mathcal{H} = \sum_r \epsilon_r c_r^\dagger c_r - t \sum_r \left(c_{r+1}^\dagger c_r + h.c. \right),$$

with twisted boundary conditions to approach the continuum. The random energies ϵ_i are drawn from the Anderson disorder box distribution $[-W/2, W/2]$, where W is the disorder strength. All of the distributions presented below were obtained by means of the new technique introduced in Chapter 4, allowing to calculate the LDoS in the whole system with about 4% of error over all space points.

Although the analytical computations use the semi-infinite system to obtain the Inverse-Gaussian, using a periodic one as a numerical system should not matter in the regime of very large systems and the size L is many times larger than the corresponding localization length ξ .

The results, shown in Figs. 5.2, demonstrates a good agreement for $4\sigma\tau \ll 1$ and $4\sigma\tau \gg 1$. the number of states within the spectral resolution, N_k as the control parameter, instead of σ . This is just a choice of convenience because the two are directly related by

$$N_k(E) = 2N\rho_{\text{med}}(E)\sigma, \quad (5.1.5)$$

where $\rho_{\text{med}}(E)$ is the mean DoS of system, upon many disorder samples and N the system size. For in-

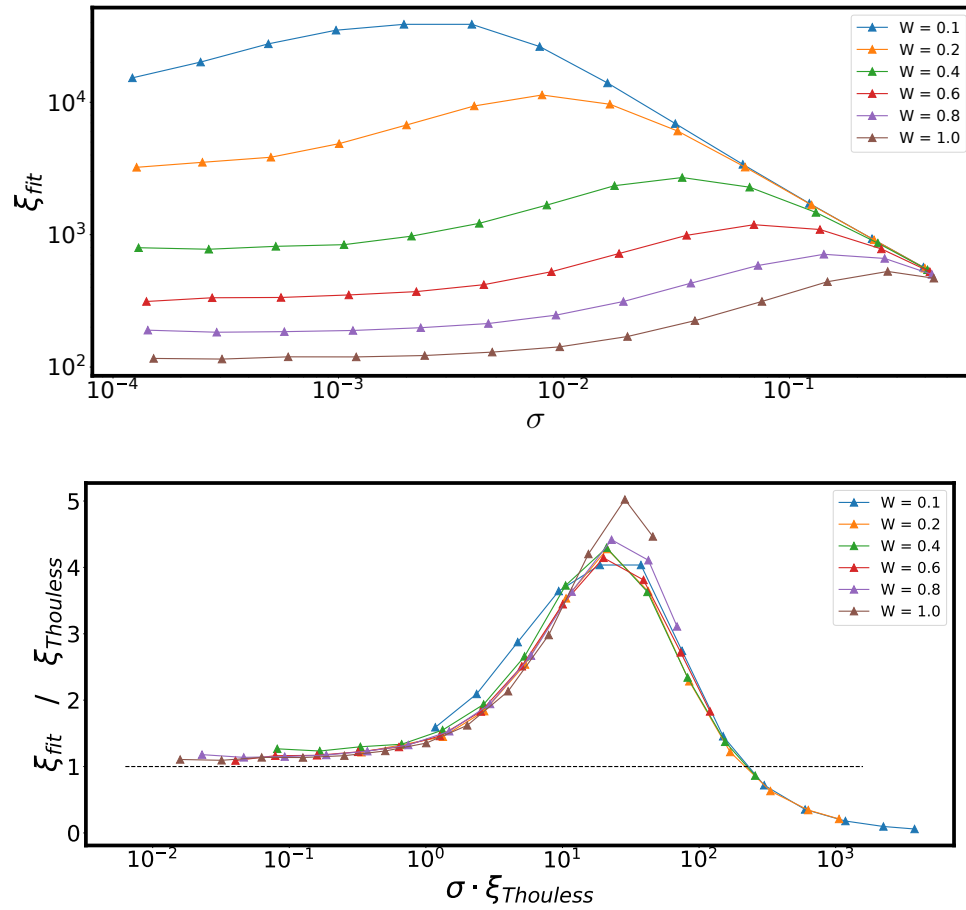


Figure 5.3: Top) Localization length obtained by a fit of the Inverse-Gaussian to the data. Bottom) Collapse of the curves. Localization length rescaled obtained by a fit of the Inverse-Gaussian to the data.

intermediate values the distributions start to deviate greatly in the tails. Alt'shuler et. al. [3] made an exact diagrammatic computation for the limit $4\sigma\tau \ll 1$ and it is argued that it should hold for any $4\sigma\tau$. Our numerical computations shows that there is a very good agreement in $4\sigma\tau \ll 1$ and $4\sigma\tau \gg 1$. The latter, $4\sigma\tau \gg 1$, is trivial because in this limit the distributions becomes a Gaussian. However, for the intermediate values of $4\sigma\tau$ we see a deviation on the tails. This does not contradict the analytical computation by the authors because on the intermediate values of $4\sigma\tau$ the argument that the Inverse-Gaussian should hold is by means of an interpolation.

From the individual distributions it is possible to extract the mean free path time τ from a fit to the Inverse-Gaussian, Eq. (5.1.4), because we always know σ , i.e. the width of the Lorentzian. With this and the definitions the authors use, we can relate the localization length to the mean free path time in the 1D case. This comes to

$$\xi_{fit} = \tau_{fit} \frac{4}{\pi \rho(E)}. \quad (5.1.6)$$

And thus, as expected, the fit of the Inverse-Gaussian to the intermediate cases are off in the tails. Figure 5.3 Bottom) shows this discordance. The ξ_{fit} recovered from the fit of the Inverse-Gaussian to the data, and, as you can clearly see, there is a huge increase of this ξ_{fit} in the intermediate region. As we continue to decrease the width the value seems to converge. It is possible from this results to collapse all the curves.

With a single scaling factor we were able to collapse all the curves, as shown in Fig. 5.3 Bottom). That factor is the exact localization length given by the Thouless formula [59],

$$\frac{1}{\xi(E)} = \int d\mu \rho(\mu) \ln(|E - \mu|). \quad (5.1.7)$$

The plot shows some discordance in the middle section of σ . This is expected because, as we previously mentioned, the analytical calculations by Alt'shuler et. al. [3] is exact in the low limit of $4\sigma\tau$. In the intermediate regime we do not now with certainties what the distribution should look like. All of the curves go to a constant, $\xi_{\text{Fit}}/\xi_{\text{Thouless}} = 1$, at the limit $\sigma \cdot \xi_{\text{Thouless}} \ll 1$.

5.2 Finite Size Effects

The important aspects to take from Fig. 5.3 is the collapse of all the curves with this scaling and that they all go towards a constant in the limit $\sigma \cdot \xi_{\text{Thouless}} \ll 1$. This asks for the creation of a new effective parameter, which we call it effective length,

$$L_{k,\text{eff}} = \frac{L}{N_k \cdot \xi_{\text{Thouless}}} = \frac{1}{2\rho_{\text{med}}(E) \sigma \cdot \xi_{\text{Thouless}}}. \quad (5.2.1)$$

This effective length is also related to the mean free path time σ by Eq. (5.1.1), suggesting it will be enough to describe the statistics but also has a new physical meaning which we shall explore soon.

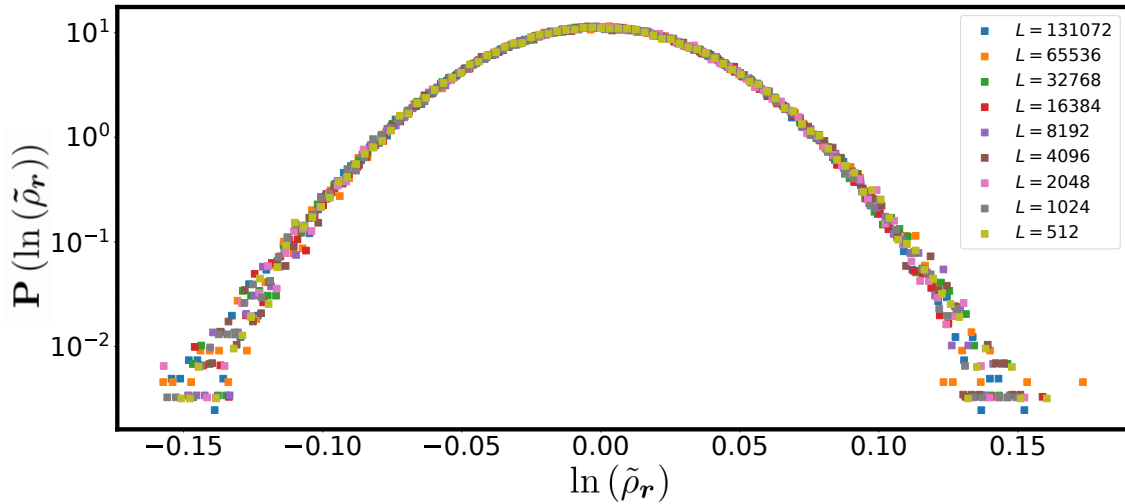


Figure 5.4: Statistics of the LDoS $\ln(\tilde{\rho}_r)$ with increasing system size, L , for a fixed Lorentzian width, $\sigma = 0.05$ and localization length, $\xi_{\text{Thouless}} = 16384$, corresponding to a disorder of $W = 0.076$.

Figure 5.4 shows a study that corroborates Eq. (5.2.1), showing no apparent difference of the LDoS statistics with increasing system size for a fixed σ and ξ_{Thouless} . For a fixed width σ and ξ_{Thouless} we see no difference in the statistics when we increase the system size. Another way to study for the numerical verification of Eq. (5.2.1) is the statistics when we fix the product $\sigma \cdot \xi_{\text{Thouless}} = \text{const}$, this would also allow a fixed $L_{k,\text{eff}}$ and reduces 3 scales, σ , ξ_{Thouless} and L down to 2.

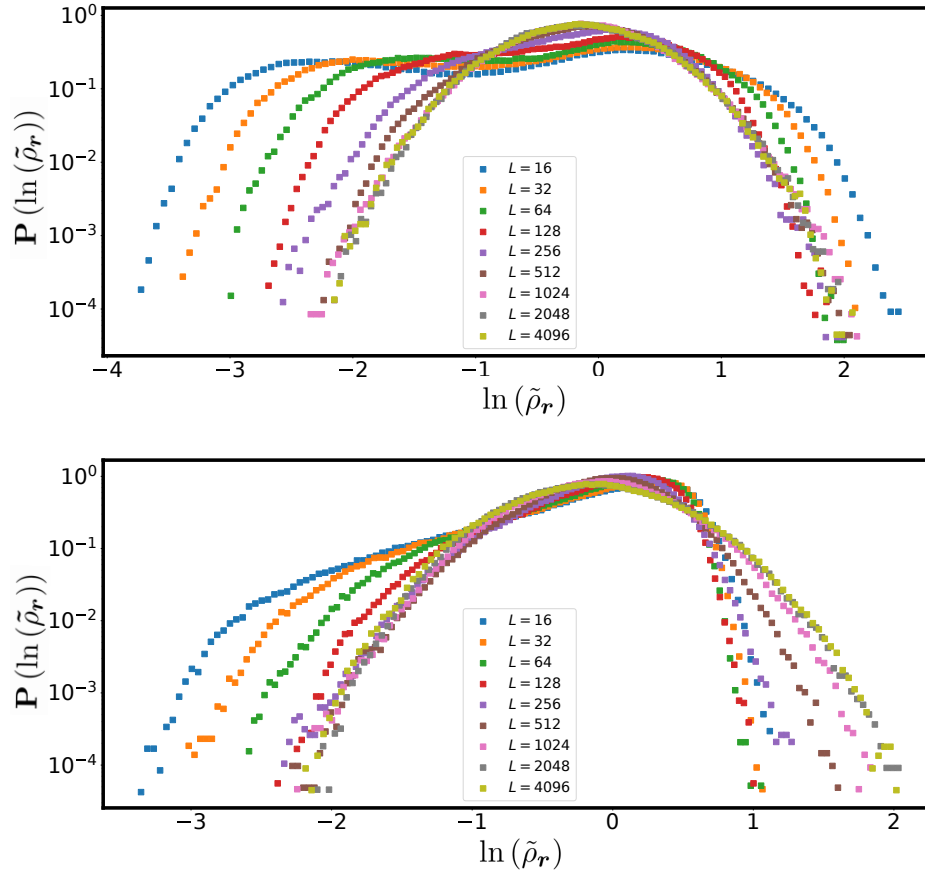


Figure 5.5: Statistics of the LDoS $\ln(\tilde{\rho}_r)$ with increasing system size, L , for a fixed product $\sigma \cdot \xi_{\text{Thouless}} = 1$. Here we do $W = 0.3$, which corresponds to $\xi_{\text{Thouless}} \sim 1024$. Top: Statistics divided by a normalization factor $\langle \rho_r \rangle$. Bottom: Statistics of each system divided by its LDoS mean. This results uses TBC.

Figure 5.5 shows the statistics results for smaller systems sizes, down to $L = 16$ and for two different definitions for the statistics. On top we have the usual LDoS divided by a normalization factor, which is the mean of the LDoS. Conversely, the bottom section showcases a distinct statistical approach, where instead of employing a global normalization factor, each simulated disorder sample is individually normalized by the mean of its respective LDoS. In essence, each system of size L is divided by its mean of the LDoS and the resultant statistics essentially quantify the fluctuations around this mean.

In each case, the curves do not collapse for all system sizes L . But it is clear that the higher the system size is, L , the curves collapse. The use of twisted boundary conditions for the smaller systems was required to simulate an infinite system effectively. A further inspection reveals that the reason for the smaller system sizes to not go under a universal curve is that the twisted boundary conditions do not fully remove the finite size effects. The TBC effectively emulates, given enough random phases, an infinite system with a supercell equal to simulated one with disorder. The DoS of such systems inevitably has gaps, which will subsequently contaminate the statistics.

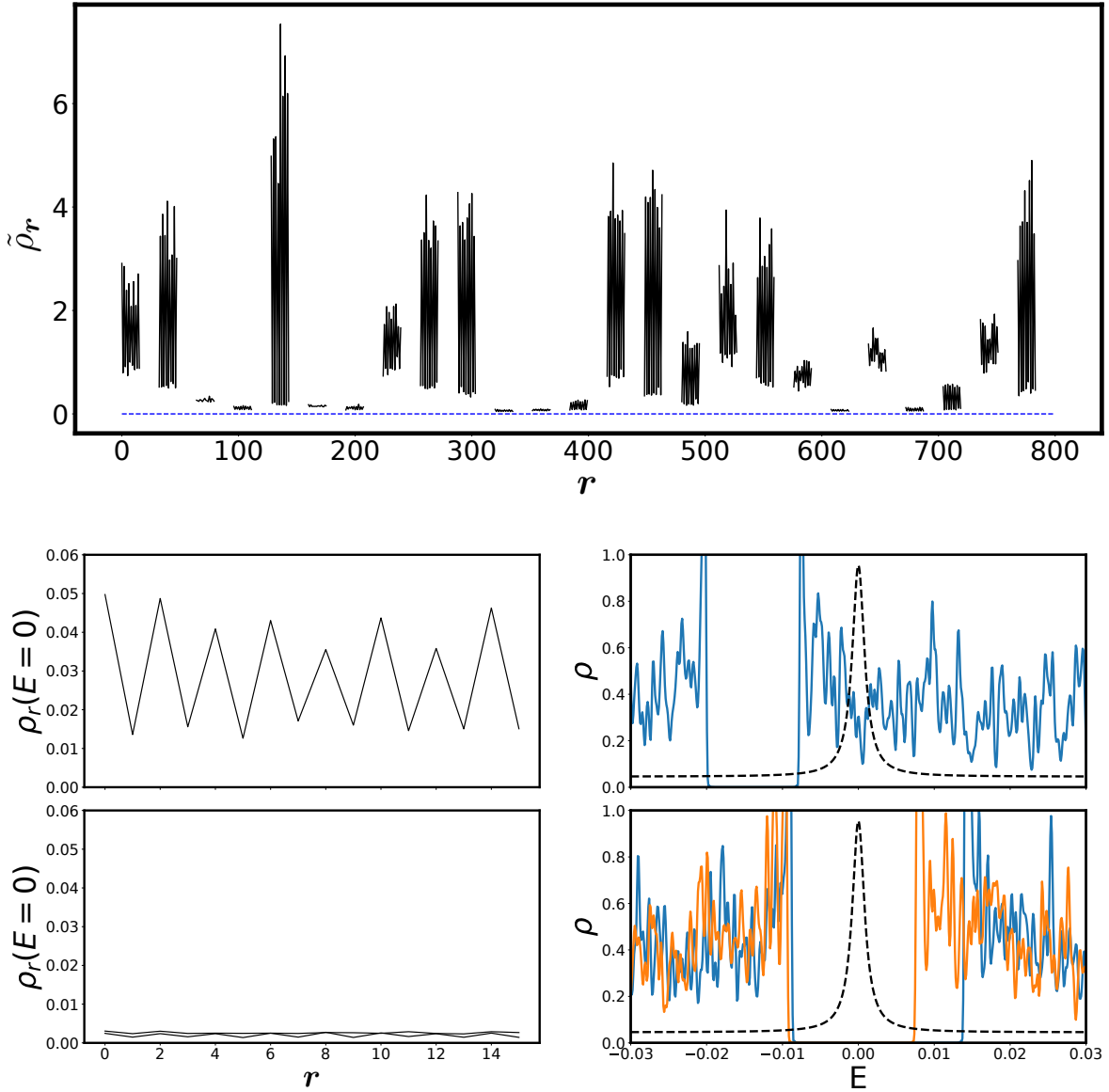


Figure 5.6: Top: Samples of the LDoS in Fig. 5.5 for $L=16$, evenly separated. Bottom: Difference between the low and high LDoS samples and its corresponding DoS, the systems chosen was the first one for the high LDoS and 7th and 11th for the low LDoS.

Figure 5.6 serves as conclusive evidence for the presence of gaps at smaller system sizes in the Local Density of States (LDoS). As the LDoS approaches extremely low values, the associated Lorentzian used for its numerical computation, characterized by a broadening parameter σ , aligns with gaps observed in the Density of States (DoS). Importantly, the DoS computation utilizes the same broadening parameter σ as employed in the LDoS computation, ensuring methodological consistency.

With an increase in the system size L , these gaps in the energy spectrum are reduced. Simultaneously, the Lorentzian profile begins to enclose a greater number of states until it reaches the continuum. Consequently, as the system size increases, the gaps gradually fade away with the chosen broadening parameter. It is important to be aware that adjusting the broadening parameter has a significant impact; reducing it magnifies the effect of the gaps, particularly at lower system sizes. This dynamic interplay between system size, broadening parameter,

and the bandwidth of the gaps in the DoS dictates the statistics.

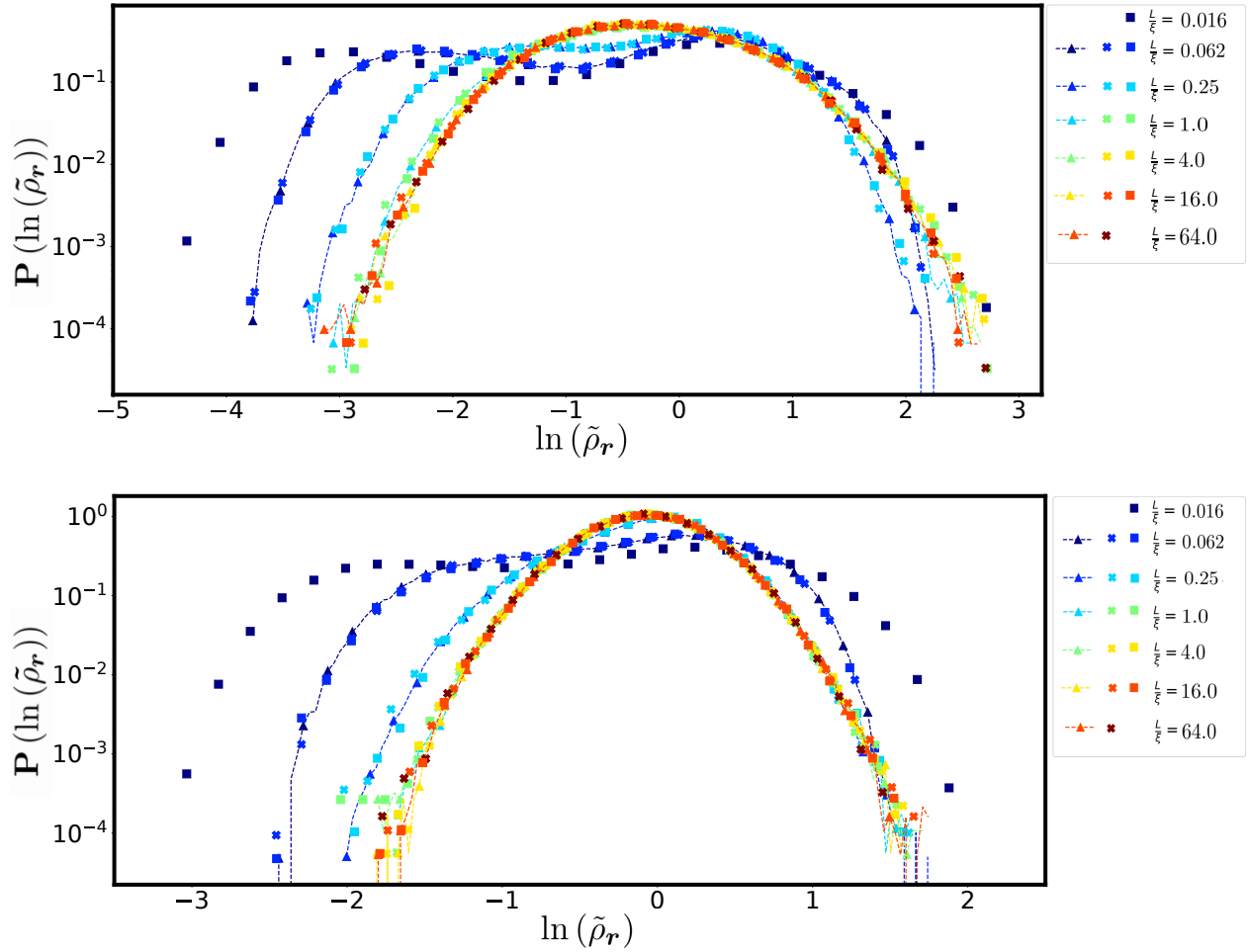


Figure 5.7: Top: Statistics for $\sigma \cdot \xi = 1.0$. Bottom: Statistics for $\sigma \cdot \xi = 4.0$. The sizes are sequentially encoded in the colors, being $L = 2^j$, starting at $L = 16$ and ending in $L = 32768$. While the shape, triangles, crosses and squares, has the localization length information, being 256, 512 and 1024 respectively.

Figure 5.7 shows the statistics for two different products $\sigma \cdot \xi$. While we proposed that for a fixed L_{eff} the statistics should be fixed subsequent observations contradict this assumption, revealing nuanced details on this behavior. By evaluating the ratio L/ξ , certain curves exhibit consistent statistics if L/ξ is constant.

Furthermore, we see that for both cases, the statistics starts to converge if and only if $L/\xi > 1$, meaning the system size is larger than the localization length ξ . This observation is particularly significant, as it introduces an additional criteria, not only we have the L_{eff} parameter we now have $L/\xi > 1$ to guarantee that we are in the thermodynamic limit of the LDoS statistics.

Notice the broadening of the statistics when we increase ξ , for a fixed $L_{\text{k,eff}}$. This goes against our intuition because as we increase ξ , the disorder strength W goes down. Because of this, we thought the statistics should reduce in broadening due to the disorder background having less fluctuations. We know that this phenomenon must be connected to the decrease of the gaps with W , L and how $\sigma \cdot \xi$ stays fixed, this requires more research.

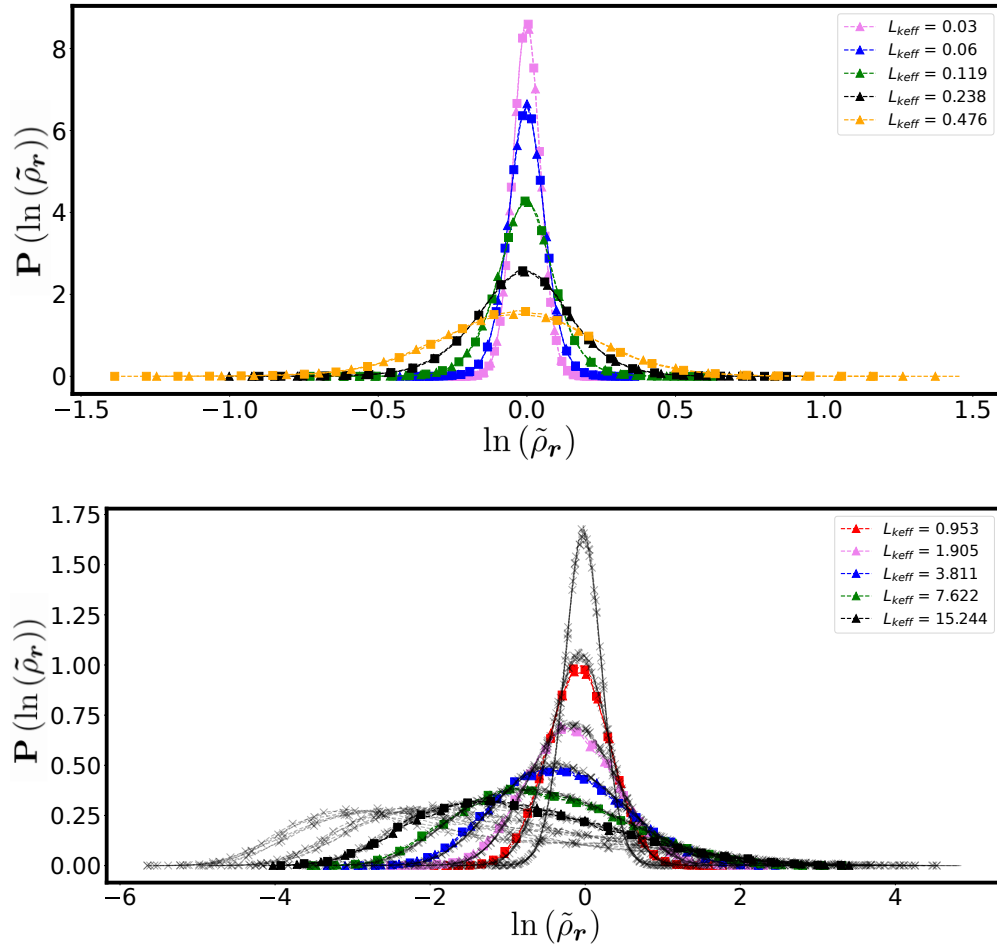


Figure 5.8: Distribution of the local density of states statistics for a one-dimensional tight-binding chain nearest neighbor for several sizes and Anderson disorder $W = 0.3$ and different product $\sigma \cdot \xi$ values, being from $\sigma \cdot \xi = 0.1 - 51.2$, always doubling, for the KPM (Colored curves). The grey curves are obtained from exact diagonalization from sizes ranging from 8 – 1024, localization lengths from 64 – 1024 and different L_{keff} . This shows that the L_{keff} is the important parameter. Top) Small L_{keff} Down) Intermediate-Big L_{keff} . The triangles were simulated with $L = 1024$ while the square with $L = 2048$.

We now shift the focus and study the behavior of the distributions by varying $L_{\text{k,eff}}$. For that, we vary the product parameter $\sigma \cdot \xi$ that is in the definition of $L_{\text{k,eff}}$, defined by Eq. (5.2.1). In Fig. 5.8 the localization length for this disorder, $W = 0.3$, obtained with the Thouless formula, corresponds to localization length of $\xi_{\text{Thouless}} = 1024$. With the aforementioned knowledge that the statistics of the LDoS is well converged when $L/\xi > 1$, we used this criteria for the smallest size used in the KPM (colored curves). The grey curves are obtained from exact diagonalization, for ease of computation due to the small sizes, and also using the criteria $L/\xi > 1$. The system sizes of the grey curves are widely different but maintain $L_{\text{k,eff}}$ and $L/\xi > 1$, and we still obtain a perfect agreement with the KPM data. This shows that the really important parameter when studying the LDoS statistics is $L_{\text{k,eff}}$.

The distribution stays centered at the mean until we reach a threshold of $L_{\text{k,eff}} \sim 1$ when it starts shifting to lower values of $\tilde{f}$, as show in Fig. 5.9. This is expected because when we reach that threshold we basically increasing the separation between the localized states. Therefore, we see more of the lattice with low LDoS, shifting the statistics to the left.

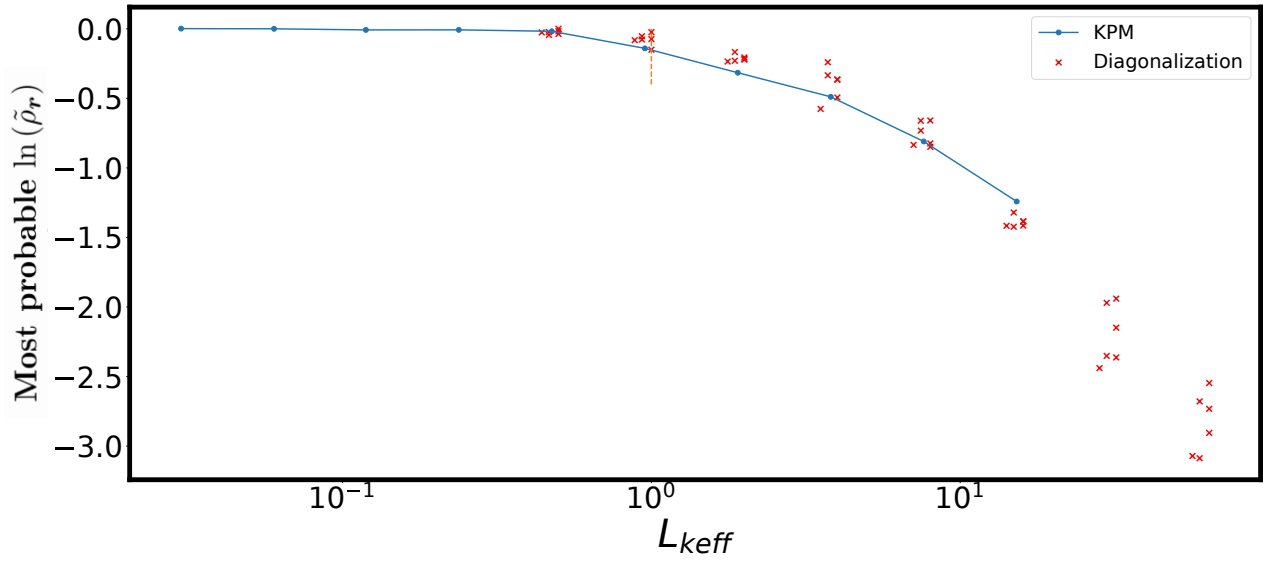


Figure 5.9: Position of the maximum of the distributions in Fig. 5.8. The diagonalization samples corresponds to different disorders realizations.

5.3 Additional Insights on the Dirac- δ Regularization

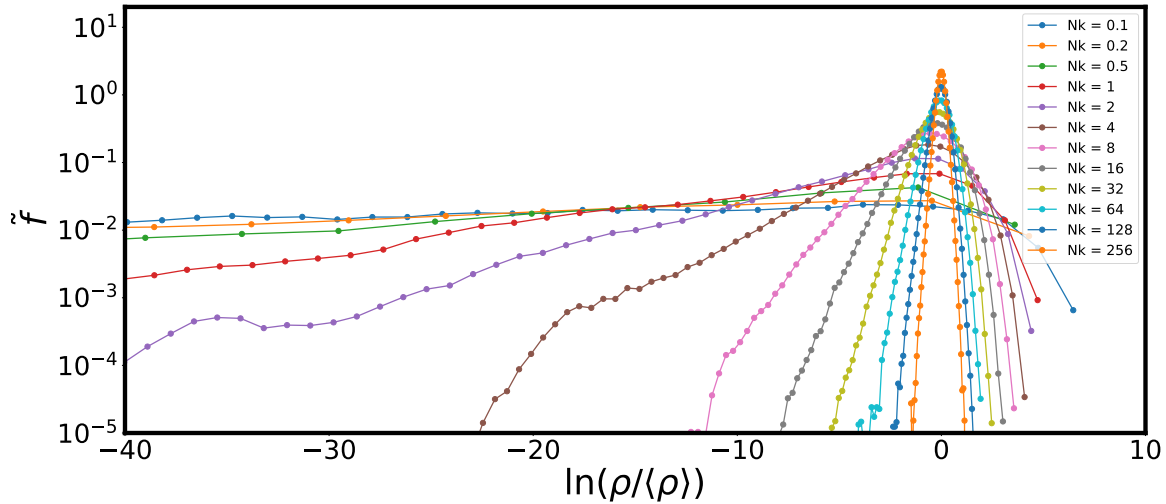


Figure 5.10: LDoS distribution for a system size $L = 16384$ and $W = 0.3$ for small N_k .

Fig. 5.10 shows the distribution of the LDoS for small parameters, that is, for a disorder $W = 0.3$, which corresponds to a localization length to about $\xi = 1024$ units and a system size $L = 16384$, this system size allows us to see the tails of the localized wavefunctions. For small N_k we see the distribution to flatten. This can be explained by a simple analysis of the distribution of the tails of the wavefunctions, which are described by $|\Psi_v(r)| = \exp(-|r - r_v|/\xi)$. Remember that for small N_k what we really are doing is selecting just one state for the LDoS, therefore, we expect to see the statistics of this state.

We can compute the logarithm distribution with the following,

$$\tilde{f}(y) = \int dx \delta(y - \ln(|\Psi_v(r)|)), \quad (5.3.1)$$

therefore, by using asymptotic behavior of the localized wavefunction, we have the integral

$$\tilde{f}(y) = \int d\mathbf{r} \delta\left(y + \frac{\sqrt{(\mathbf{r} - \mathbf{r}_\alpha)^2}}{\xi}\right), \quad (5.3.2)$$

which in spherical coordinates is

$$\tilde{f}(y) = \Omega_d \int dr r^{d-1} \delta\left(y + \frac{r}{\xi}\right), \quad (5.3.3)$$

where Ω_d is the surface area of a d -sphere. Finally, the integral comes about

$$\tilde{f}(y) = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma\left(\frac{d+1}{2}\right)} (-y\xi)^{d-1}. \quad (5.3.4)$$

For 1D, the distribution is constant, which is what we see in the plot of Fig. 5.10. This also means that the distribution for small N_k depends on the dimension of the problem.

This study always used, of course, the Lorentzian as a regularization for the LDoS, one now questions that, since many researchers use the Gaussian kernel, if this affects the results considerably or we can use Gaussian as a good regularization function for the distributions.

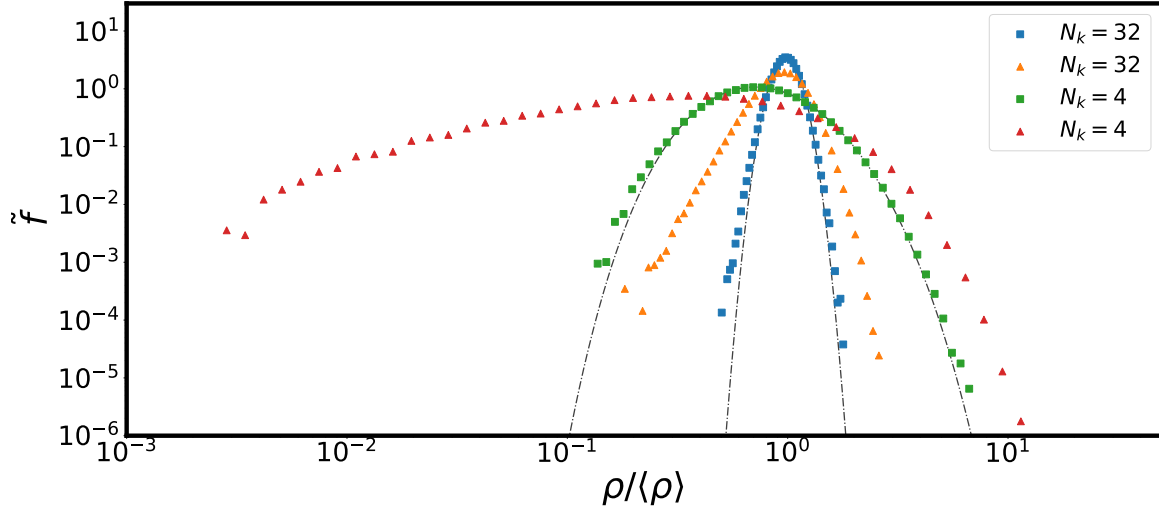


Figure 5.11: Example of the difference between the statistics of the LDoS for a system size $L = 512000$ and $W = 1.0$ using; squares: Lorentzian; triangles: Gaussian; regularization of the Dirac- δ .

In Fig. (5.11), this question is trivially answered and easily understood. Due to the Gaussian having exponentially suppressed tails, it will immediately kill the states that are far beyond the selected energy E of study, adding no value to the LDoS and, thus, having a greater contribution in the small range of the LDoS. On the other hand, the Lorentzian will always add contribution, shifting the LDoS to greater values and a better uniformly spread LDoS on the whole range. This conclusion is of importance because it heavily impacts the study of this subject. It is important to be aware of this fact.

5.4 Basic Scales

In this chapter we learned about the probability distribution of the Local Density of States in the finite disordered one-dimensional Anderson model. Upon a detailed exploration of all relevant parameters in the system, we have demonstrated that this distribution is controlled by three basic length scales (or, equivalently, energy scales). The parameters of this problem comprise two natural length scales -- the simulated system size L and the localization length ξ -- and three natural energy scales -- the broadening introduced by the Dirac- δ regularization, σ , the disorder strength W and the energy E . It is worth noting that ξ and W are intimately related and, in fact, $\xi \propto 1/W^2$ in one-dimensional systems. In our numerical study, we have found that in the thermodynamic limit, the statistical fluctuations of the LDoS are controlled by a single dimensionless parameter: the effective length, $L_{k,\text{eff}} = 1/(2\rho_{\text{med}}(E)\sigma\xi)$. This effective length essentially measures the typical distance that exists between localization centers for electrons having an energy within the spectral resolution considered. If the localized states are close in energy their respective centers tend to be separate in space, and $L_{k,\text{eff}}$ will quantify this spacing. As we defined it, $L_{k,\text{eff}}$ is measured in units of the localization length. This means that if $L_{k,\text{eff}}$ is much larger than 1, the states that contribute to the LDoS at that energy are typically separated by more than their localization length. In this case, the LDoS fluctuations in space will display a behavior typical of a localized system. While the effective length controls the impact of single-particle localization in the results of a finite resolution LDoS calculation, it is not allowed to exceed the actual size of the simulated system. As $L_{k,\text{eff}}$ gets closer to 1, finite size effects become apparent because the calculation is picking up too few states, and the discrete nature of the spectrum becomes apparent. Similarly, if the boundary conditions affect the individual quantum states themselves, finite-size effects may still matter even in cases where $L_{k,\text{eff}} \ll 1$. In other words, the tails of the localized wavefunctions extend beyond the boundaries of the system and experience self-interference if ξ approaches or exceeds L . By adjusting all the physical parameters, we were able to confirm the existence of each of these regimes in the results that are shown in this chapter.

Chapter 6

Conclusions

On this thesis, in face of the computational challenge, we developed a powerful method to compute the Local Density of States (LDoS). We introduced a novel method that uses random vectors, resulting in a highly efficient algorithm capable of computation of the LDoS throughout the whole lattice in linear complexity and with controllable error. What makes this method particularly intriguing is not the clever use of the random vectors but rather the way we derive the average LDoS through the known variance.

We have further revisited the utilization of the Graphical Processing Units (GPU) on the Kernel Polynomial Method (KPM). The KPM relies heavily on sparse matrix operations, which makes it a perfect candidate for the hardware of GPUs. We developed a CUDA C-based KPM software, devoid of known GPU libraries, affording us superior memory control and yielding performance gains of several orders of magnitude compared to CPU-based implementations. This conjoined with the new algorithm of the LDoS computation in whole lattice it allows us to amass an impressive amount of data that allows one to conduct sampling-intensive studies of fluctuations in local properties of disordered systems.

Building on this foundation, we investigated the statistics of LDoS in disordered systems. Having found that, for the 1D systems, the statistics of the LDoS perfectly agrees with the analytical results by Alt'shuler, with minor deviations in the tails at intermediate Lorentzian widths. However, these deviations at the intermediate level does not contradict the analytical results. After this study, we were able to extract the localization length with a fit to the data and managed to collapse all curves for different disorders. This in turn revealed a new effective parameter, $L_{k,eff}$, that measures the maximum spacial distance between localized centers that are close in energy, consolidating system characteristics.

Subsequently, we probed this $L_{k,eff}$ with increasing L , starting at 512 and showed that the statistics does not change, initially confirming our hypothesis for $L_{k,eff}$. Then, with smaller system sizes we found that such is not case, and it revealed that a good parameter for the statistics convergence is that $L/\xi > 1$, this is essentially how bigger L is compared to the localization length ξ . Statistics with the same ratio L/ξ are equal.

We also found that for the lower systems sizes the existing of gaps occurred, consequence of the twisted boundary conditions. These gaps essentially meant that we were not on the thermodynamic limit, and saw the elongation of the LDoS statistics due to the Lorentzian broadening σ capturing the gaps. This also opened a question of why the LDoS elongates, maintaining $L_{k,eff}$, and increasing ξ .

In addition, we investigated the case where the width of the Lorentzian was extremely small, basically selecting only a state to contribute to the LDoS. This gave us insight about the statistics of the wavefunctions assuming

the known behavior of the tails when localized. This approach is very simple and concluded that the statistics depends on the dimension of the problem. For 1D the statistics assumes a constant and it is what we observed in the limit of low width.

Lastly, we performed a simple study by changing the Dirac- δ representation, which had been the standard Lorentzian, to a Gaussian, commonly employed by many researchers. This indeed changed the statistics in a meaningful manner, notably the broadening the distribution of LDoS. The tails of the Gaussian are heavy and effectively suppresses rapidly the contributions of the wavefunctions outside the energy we investigate. This permits a larger variation of the LDoS values, and thus, a broader distribution.

To sum up, this thesis not only devised innovative solutions to computational challenges regarding the LDoS but also made substantial contributions to the understanding of LDoS statistics in disordered systems, offering valuable insights into dimension-dependent behaviors. These findings provide a robust foundation for future research in this field such as the original problem that initiated the work presented here, the study of the unconventional disorder-induced phase transition from a dirty semi-metal to a metal phase in 3D point-node Weyl systems.

Appendix A

Calculation of the Variance of the DoS and LDoS

A.1 Density of States

In this Appendix we show to compute the variance of the DoS, given in Eq. ???. As stated before, the main object in

$$\sigma^2 = \frac{1}{D^2} \left\{ \sum_{r,r_1,r_2,r_3} \sum_{\alpha,\beta} \Psi_\alpha^*(r) \Psi_\alpha(r_1) \Psi_\beta(r_2) \Psi_\beta^*(r_3) f_\alpha f_\beta \langle \xi_r^* \xi_{r_1} \xi_{r_2} \xi_{r_3}^* \rangle \right\} - \rho^2(E), \quad (\text{A.1.1})$$

is the computation of the correlator $\langle \xi_r \xi_{r'} \xi_{r''} \xi_{r'''} \rangle$. This is already done in Appendix A.3, stating the result we have

$$g_{ijkl} \langle \xi_i^* \xi_j \xi_k \xi_l^* \rangle = g_{iiii} \cdot \delta_{ij} \delta_{ik} \delta_{il} \langle |\xi_i|^4 \rangle + g_{iikk} \cdot \delta_{ij} \delta_{kl} \bar{\delta}_{ik} + g_{ikik} \cdot \delta_{ik} \delta_{jl} \bar{\delta}_{ij} \quad (\text{A.1.2})$$

Performing the summation in Eq. (4.2.5) and plugging the result of the correlator we quickly find that

$$\begin{aligned} \sigma^2 = & \frac{1}{D^2} \left\{ \langle |\xi|^4 \rangle \sum_r \sum_{\alpha,\beta} |\Psi_\alpha(r)|^2 |\Psi_\beta(r)|^2 f_\alpha f_\beta + \sum_{r \neq r_2, r} |\Psi_\alpha(r)|^2 |\Psi_\beta(r_2)|^2 f_\alpha f_\beta \right\} \\ & + \frac{1}{D^2} \left\{ \sum_{r \neq r_2, r} \sum_{\alpha,\beta} \Psi_\alpha^*(r) \Psi_\alpha(r_2) \Psi_\beta(r) \Psi_\beta^*(r_2) f_\alpha f_\beta \right\} - \rho^2(E), \end{aligned} \quad (\text{A.1.3})$$

We can easily see the definition of the LDoS appearing in the first term, and for the second and third term we separate the summation as $\sum_{r \neq r_2} = \sum_r - \sum_{r=r_2}$, resulting in

$$\begin{aligned} \sigma^2 = & \langle \xi^4 \rangle \sum_r \rho_r^2(E) + \frac{1}{D^2} \sum_{r,r_2} \sum_{\alpha,\beta} |\Psi_\alpha(r)|^2 |\Psi_\beta(r_2)|^2 f_\alpha f_\beta - \frac{1}{D^2} \sum_r \sum_{\alpha,\beta} |\Psi_\alpha(r)|^2 |\Psi_\beta(r)|^2 f_\alpha f_\beta \\ & + \frac{1}{D^2} \sum_{r,r_2} \sum_{\alpha,\beta} \Psi_\alpha^*(r) \Psi_\alpha(r_2) \Psi_\beta(r) \Psi_\beta^*(r_2) f_\alpha f_\beta - \frac{1}{D^2} \sum_{r_2} \sum_{\alpha,\beta} |\Psi_\alpha(r_2)|^2 |\Psi_\beta(r_2)|^2 f_\alpha f_\beta \\ & - \rho^2(E), \end{aligned} \quad (\text{A.1.4})$$

we then proceed to use the normalization of the wave-functions $\sum_r |\Psi_\alpha(r)|^2 = 1$ in the second term, immediately seeing the definition of the DoS, which cancels with the last term. The third and fifth term is again just the the LDoS definition. Having said this, we get for the expression

$$\sigma^2 = \left\{ \frac{1}{D^2} \sum_{r,r_2} \sum_{\alpha,\beta} \Psi_\alpha^*(r) \Psi_\alpha(r_2) \Psi_\beta(r) \Psi_\beta^*(r_2) f_\alpha f_\beta + \left(\langle |\xi|^4 \rangle - 2 \right) \sum_r \rho_r^2(E) \right\}. \quad (\text{A.1.5})$$

The final step comes from the orthogonality relation, $\sum_r \Psi_\alpha^*(\mathbf{r}) \Psi_\beta(\mathbf{r}) = \delta_{\alpha\beta}$, and again the normalization of ψ , providing us with the result

$$\sigma^2 = \left\{ \frac{1}{D^2} \sum_\alpha f_\alpha^2 + \left(\langle |\xi|^4 \rangle - 2 \right) \sum_r \rho_r^2(E) \right\}, \quad (\text{A.1.6})$$

thus concluding the calculation.

A.2 Local Density of States

In this appendix we provide a detailed calculation of the variance of the LDOS, starting where we left out at 4.2.5.

$$\langle \tilde{\rho}_{\mathbf{r}}(E) \tilde{\rho}_{\mathbf{r}}^*(E) \rangle = \frac{1}{D^2} \sum_{\alpha, \beta} \sum_{\mathbf{r}', \mathbf{r}''} \Psi_{\alpha}^*(\mathbf{r}) \Psi_{\alpha}(\mathbf{r}') \Psi_{\beta}(\mathbf{r}) \Psi_{\beta}^*(\mathbf{r}'') f_{\alpha} f_{\beta} \langle |\xi_{\mathbf{r}}|^2 \xi_{\mathbf{r}'} \xi_{\mathbf{r}''}^* \rangle. \quad (\text{A.2.1})$$

The calculation of correlators is a different technique altogether and is explained in the Appendix A.3, the relevant result for us is that

$$f_{ijk} \langle |\xi_i|^2 \xi_j \xi_k^* \rangle = f_{iii} \cdot \delta_{ij} \delta_{ik} \langle |\xi|^4 \rangle + f_{ijj} \cdot \delta_{jk} \bar{\delta}_{ij}. \quad (\text{A.2.2})$$

By plugging this result in A.2.1 we have

$$\begin{aligned} \langle \tilde{\rho}_{\mathbf{r}}(E) \tilde{\rho}_{\mathbf{r}}^*(E) \rangle &= \frac{1}{D^2} \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta} \langle |\xi|^4 \rangle \\ &+ \frac{1}{D^2} \sum_{\alpha, \beta} \sum_{\mathbf{r}' \neq \mathbf{r}} \Psi_{\alpha}^*(\mathbf{r}) \Psi_{\alpha}(\mathbf{r}') \Psi_{\beta}(\mathbf{r}) \Psi_{\beta}^*(\mathbf{r}') f_{\alpha} f_{\beta}. \end{aligned} \quad (\text{A.2.3})$$

We can extend the summation $\sum_{\mathbf{r}' \neq \mathbf{r}}$ to a full one and subtract the left over, $\sum_{\mathbf{r}' \neq \mathbf{r}} = \sum_{\mathbf{r}'} - \sum_{\mathbf{r}' = \mathbf{r}}$, giving us

$$\begin{aligned} \langle \tilde{\rho}_{\mathbf{r}}(E) \tilde{\rho}_{\mathbf{r}}^*(E) \rangle &= \frac{1}{D^2} \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta} \langle |\xi|^4 \rangle \\ &+ \frac{1}{D^2} \sum_{\alpha, \beta} \sum_{\mathbf{r}'} \Psi_{\alpha}^*(\mathbf{r}) \Psi_{\alpha}(\mathbf{r}') \Psi_{\beta}(\mathbf{r}) \Psi_{\beta}^*(\mathbf{r}') f_{\alpha} f_{\beta} \\ &- \frac{1}{D^2} \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta}. \end{aligned} \quad (\text{A.2.4})$$

Using using the orthogonality relation $\sum_{\mathbf{r}} \Psi_{\alpha}^*(\mathbf{r}) \Psi_{\beta}(\mathbf{r}) = \delta_{\alpha\beta}$ we get

$$\langle \tilde{\rho}_{\mathbf{r}}(E) \tilde{\rho}_{\mathbf{r}}^*(E) \rangle = \frac{1}{D^2} \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta} \left\{ \langle |\xi|^4 \rangle - 1 \right\} + \frac{1}{D^2} \sum_{\alpha, \beta} \Psi_{\alpha}^*(\mathbf{r}) \Psi_{\beta}(\mathbf{r}) \delta_{\alpha\beta} f_{\alpha} f_{\beta}, \quad (\text{A.2.5})$$

$$\langle \tilde{\rho}_{\mathbf{r}}(E) \tilde{\rho}_{\mathbf{r}}^*(E) \rangle = \frac{1}{D^2} \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta} \left\{ \langle |\xi|^4 \rangle - 1 \right\} + \frac{1}{D^2} \sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^2 f_{\alpha}^2. \quad (\text{A.2.6})$$

It only have to subtract the term 4.2.6 to obtain σ^2 . Giving us the intended result

$$\sigma^2 = \frac{1}{D^2} \left[\sum_{\alpha} |\Psi_{\alpha}(\mathbf{r})|^2 f_{\alpha}^2 + \sum_{\alpha, \beta} |\Psi_{\alpha}(\mathbf{r})|^2 |\Psi_{\beta}(\mathbf{r})|^2 f_{\alpha} f_{\beta} \left\{ \langle |\xi|^4 \rangle - 2 \right\} \right]. \quad (\text{A.2.7})$$

A.3 Correlators

During the calculation of statistical properties of the approach of the random local density of states we encounter quite frequently the correlators. That is, averages of objects of the form $\langle \xi_i \xi_j \xi_k \dots \xi_n \rangle$, where ξ are random variables drawn from the same probability distribution. In this appendix we explain how to compute such objects.

The way to deal with such correlators is to exhaust all possibilities of the indexes. This is rather vague and it is best shown with an example, we shall use a simple one, such as

$$C_{ijk} = \langle \xi_i^2 \xi_j \xi_k \rangle. \quad (\text{A.3.1})$$

To go through all combinations of the indexes we shall use the notation δ_{ij} which is the kronecker delta meaning that we set $i = j$ and $\bar{\delta}_{ij} = 1 - \delta_{ij}$ for the case $i \neq j$. When we start this process both δ_{ij} and $\bar{\delta}_{ij}$ will accompany each other. For example,

$$C_{ijk} = \delta_{ij} \langle \xi_i^3 \xi_k \rangle + \bar{\delta}_{ij} \langle \xi_i^2 \xi_j \xi_k \rangle, \quad (\text{A.3.2})$$

in the first term we immediately know that $i = j$ and so it already appears a cubic term. Such is not the case in the second term. We continue this process of index combinations,

$$C_{ijk} = \delta_{ij} \langle \xi_i^3 \xi_k \rangle + \bar{\delta}_{ij} \langle \xi_i^2 \xi_j \xi_k \rangle \quad (\text{A.3.3})$$

$$= \delta_{ij} \delta_{ik} \langle \xi_i^4 \rangle + \delta_{ij} \bar{\delta}_{ik} \langle \xi_i^3 \rangle \langle \xi_k \rangle + \bar{\delta}_{ij} \langle \xi_i^2 \xi_j \xi_k \rangle, \quad (\text{A.3.4})$$

as you can see, the correlator starts to unfold quite in a predictably manner. The new second term we were able to separate the average because, as indicated in the $\bar{\delta}_{ik}$, $i \neq k$ and there are no more combinations. We are left to deal with the last term. Continuing we get

$$C_{ijk} = \delta_{ij} \delta_{ik} \langle \xi_i^4 \rangle + \delta_{ij} \bar{\delta}_{ik} \langle \xi_i^3 \rangle \langle \xi_k \rangle + \bar{\delta}_{ij} \delta_{jk} \langle \xi_i^2 \rangle \langle \xi_j \rangle + \bar{\delta}_{ij} \bar{\delta}_{jk} \langle \xi_i^2 \xi_j \xi_k \rangle, \quad (\text{A.3.5})$$

The last term is not so obvious, because we know that $i \neq j$ and $j \neq k$ but it does not mean that $i \neq k$, so, we get another possibility for the last term,

$$C_{ijk} = \delta_{ij} \delta_{ik} \langle \xi_i^4 \rangle + \delta_{ij} \bar{\delta}_{ik} \langle \xi_i^3 \rangle \langle \xi_k \rangle + \bar{\delta}_{ij} \delta_{jk} \langle \xi_i^2 \rangle \langle \xi_j \rangle + \bar{\delta}_{ij} \bar{\delta}_{jk} \delta_{ik} \langle \xi_i^3 \rangle \langle \xi_j \rangle + \bar{\delta}_{ij} \bar{\delta}_{jk} \bar{\delta}_{ik} \langle \xi_i^2 \rangle \langle \xi_j \rangle \langle \xi_k \rangle, \quad (\text{A.3.6})$$

and finally, we exhausted all possibilities.

The result in Eq. (A.3.6) can be further simplified by noting that the random variables are, as previously said, drawn from the same probability distribution. Therefore, we can drop the indices in ξ , which gives,

$$C_{ijk} = \delta_{ij} \delta_{ik} \langle \xi^4 \rangle + \left\{ \delta_{ij} \bar{\delta}_{ik} + \bar{\delta}_{ij} \bar{\delta}_{jk} \delta_{ik} \right\} \langle \xi^3 \rangle \langle \xi \rangle + \bar{\delta}_{ij} \delta_{jk} \langle \xi^2 \rangle^2 + \bar{\delta}_{ij} \bar{\delta}_{jk} \bar{\delta}_{ik} \langle \xi^2 \rangle \langle \xi \rangle^2, \quad (\text{A.3.7})$$

Doing this type of computations by hand is quite painful. Fortunately there is a way to develop some sort of algorithm to perform this combinations automatically. We can use Eq. (A.3.1) as an example again.

This process begins by noting that, of course, $\delta_{ij} + \bar{\delta}_{ij} = 1$. With this we can express product of all the possible combination as follows,

$$C_{ijk} = (\delta_{ij} + \bar{\delta}_{ij}) (\delta_{ik} + \bar{\delta}_{ik}) (\delta_{jk} + \bar{\delta}_{jk}) C_{ijk}, \quad (\text{A.3.8})$$

then we proceed to expand the main expression, obtaining

$$\begin{aligned} & \delta_{ij}\delta_{ik}\delta_{jk} + \delta_{ij}\delta_{ik}\bar{\delta}_{jk} + \delta_{ij}\bar{\delta}_{ik}\delta_{jk} + \delta_{ij}\bar{\delta}_{ik}\bar{\delta}_{jk} \\ & + \bar{\delta}_{ij}\delta_{ik}\delta_{jk} + \bar{\delta}_{ij}\delta_{ik}\bar{\delta}_{jk} + \bar{\delta}_{ij}\bar{\delta}_{ik}\delta_{jk} + \bar{\delta}_{ij}\bar{\delta}_{ik}\bar{\delta}_{jk}. \end{aligned} \quad (\text{A.3.9})$$

After the expansion we begin the process of elimination, that is, we get rid of the of the impossible combinations. For example, the term $\delta_{ij}\delta_{ik}\bar{\delta}_{jk}$ is impossible because it says that $i = j$ and $i = k$ therefore $j = k$ but the last product contradicts this logic, saying that $j \neq k$, therefore, this term does not exist. The same type of conclusion follows for the term $\delta_{ij}\bar{\delta}_{ik}\delta_{jk}$ and $\bar{\delta}_{ij}\delta_{ik}\delta_{jk}$. What we are left with is the terms

$$\delta_{ij}\delta_{ik}\delta_{jk} + \delta_{ij}\bar{\delta}_{ik}\bar{\delta}_{jk} + \bar{\delta}_{ij}\delta_{ik}\bar{\delta}_{jk} + \bar{\delta}_{ij}\bar{\delta}_{ik}\delta_{jk} + \bar{\delta}_{ij}\bar{\delta}_{ik}\bar{\delta}_{jk}. \quad (\text{A.3.10})$$

We then just join C_{ijk} and arrive at the intended result. To not that the terms are not identical to the ones found in Eq. (A.3.6) because in that computation we already suppressed obvious terms such as δ_{jk} in $\delta_{ij}\delta_{ik}\delta_{jk}$. This is an easier way to compute all the possible combinations in the correlators. One quick remark is that doing the whole product before hand is wasteful, because it grows exponentially with the number of correlators. For a three-correlator C_{ijk} this is not a problem because we only have $2^{2+1} = 8$ terms. But for higher order n -correlators this number grows as $2^{\sum_{i=1}^{n-1} i}$, for a 8-correlator we would have 2^{28} terms. The way to circumvent this annoyance is doing a type of recursion algorithm, we compute the first two product and detect the impossible terms, then use this result to do the next product. This allows us to quickly find the possible terms even in high order correlators. With this at hand we can quickly compute the relevant correlators.

After we learned how to compute the correlators, there is only one more caveat. Normally in this computation, there is always a pre-function f of n parameters with the n correlator, $f_{i,j,k,\dots} \cdot \langle \xi_i \xi_j \xi_k \dots \rangle$, and this really complicate matters if such function f has little symmetry in the parameters, one has to take this into account. This being said, we can now express the results for the correlators used in computation of the statistical properties of the DoS and LDoS using the relevant moments of the probability distributions, $\langle |\xi_i|^{2n-1} \rangle = 0$ for $n \in \mathbb{N}$, $\langle \xi_i^* \xi_j \rangle = \delta_{ij}$ and $\langle \xi_i \xi_j \rangle = \langle \xi_i^* \xi_j^* \rangle = 0$.

$$f_{ijk} F_{ijk} = f_{ijk} \langle |\xi_i|^2 \xi_j \xi_k^* \rangle = f_{iii} \cdot \delta_{ij} \delta_{ik} \langle |\xi_i|^4 \rangle + f_{ijj} \cdot \delta_{jk} \bar{\delta}_{ij}, \quad (\text{A.3.11})$$

$$g_{ijkl} G_{ijkl} = g_{ijkl} \langle \xi_i^* \xi_j \xi_k \xi_l^* \rangle = g_{iiii} \cdot \delta_{ij} \delta_{ik} \delta_{il} \langle |\xi_i|^4 \rangle + g_{iikk} \cdot \delta_{ij} \delta_{kl} \bar{\delta}_{ik} + g_{ikik} \cdot \delta_{ik} \delta_{jl} \bar{\delta}_{ij} \quad (\text{A.3.12})$$

For the next two we only need the real valued correlators, meaning that the necessary conditions is $\langle \xi_i^{2n-1} \rangle = 0$ for $n \in \mathbb{N}$, $\langle \xi_i \xi_j \rangle = \delta_{ij}$

$$C_{ijkl} = \langle \xi_i^3 \xi_j \xi_k \xi_l \rangle = \delta_{ij} \delta_{ik} \delta_{il} \langle \xi_i^6 \rangle + 3 \cdot \delta_{ij} \delta_{kl} \bar{\delta}_{ik} \langle \xi_i^4 \rangle \quad (\text{A.3.13})$$

$$\begin{aligned}
u_{ijklm}U_{ijklm} &= \langle |\xi_i|^4 \xi_j \xi_k \xi_l \xi_m \rangle = \delta_{ij} \delta_{ik} \delta_{il} \delta_{im} \langle \xi_i^8 \rangle + 6 \cdot \delta_{ij} \delta_{ik} \delta_{lm} \bar{\delta}_{il} \langle \xi_i^6 \rangle \\
&\quad + \delta_{jk} \delta_{jl} \delta_{jm} \bar{\delta}_{ij} \langle \xi_i^4 \rangle^2 + 3 \cdot \delta_{jk} \delta_{lm} \bar{\delta}_{ij} \bar{\delta}_{il} \bar{\delta}_{kl} \langle \xi_i^4 \rangle
\end{aligned} \tag{A.3.14}$$

Bibliography

- [1] P. W. Anderson. Absence of Diffusion in Certain Random Lattices. *Physical Review*, 109(5):1492–1505, March 1958. Publisher: American Physical Society.
- [2] J. Tersoff and D. R. Hamann. Theory of the scanning tunneling microscope. *Physical Review B*, 31(2):805–813, January 1985.
- [3] BI Altshuler and Vn Prigodin. Distribution of Local Density of States and Shape of Nmr Line in a One-Dimensional Disordered Conductor. *ZHURNAL EKSPERIMENTALNOI I TEORETICHESKOI FIZIKI*, 95(1):348–370, January 1989. Num Pages: 23 Place: Moscow Publisher: Mezhdunarodnaya Kniga Web of Science ID: WOS:A1989T174900034.
- [4] Alexander D. Mirlin and Yan V. Fyodorov. Distribution of local densities of states, order parameter function, and critical behavior near the Anderson transition. *Physical Review Letters*, 72(4):526–529, January 1994.
- [5] Gerald Schubert, Jens Schleede, Krzysztof Byczuk, Holger Fehske, and Dieter Vollhardt. Distribution of the local density of states as a criterion for Anderson localization: Numerically exact results for various lattices in two and three dimensions. *Physical Review B*, 81(15):155106, April 2010. arXiv:1002.2895 [cond-mat].
- [6] Alexander Weisse, Gerhard Wellein, Andreas Alvermann, and Holger Fehske. The Kernel Polynomial Method. *Reviews of Modern Physics*, 78(1):275–306, March 2006. arXiv:cond-mat/0504627.
- [7] Moritz Kreutzner, Georg Hager, Gerhard Wellein, Andreas Pieper, Andreas Alvermann, and Holger Fehske. Performance Engineering of the Kernel Polynomial Method on Large-Scale CPU-GPU Systems. In *2015 IEEE International Parallel and Distributed Processing Symposium*, pages 417–426, May 2015. arXiv:1410.5242 [cond-mat, physics:physics].
- [8] Gerald Schubert, Alexander Weiße, Gerhard Wellein, and Holger Fehske. HQS@HPC: Comparative numerical study of Anderson localisation in disordered electron systems. In Arndt Bode and Franz Durst, editors, *High Performance Computing in Science and Engineering, Garching 2004*, pages 237–249, Berlin, Heidelberg, 2005. Springer.
- [9] Andreas Alvermann and Holger Fehske. Local Distribution Approach. In H. Fehske, R. Schneider, and A. Weiße, editors, *Computational Many-Particle Physics*, Lecture Notes in Physics, pages 505–526. Springer, Berlin, Heidelberg, 2008.

- [10] Jason Sanders and Edward Kandrot. *CUDA by Example: An Introduction to General-Purpose GPU Programming*. Addison-Wesley Professional, Upper Saddle River, NJ, 1st edition edition, July 2010.
- [11] An Li, Robert Maunder, Bashir Al-Hashimi, and L. Hanzo. Implementation of a Fully-Parallel Turbo Decoder on a General-Purpose Graphics Processing Unit. *IEEE Access*, 4:1–1, June 2016.
- [12] H. P. Veiga, S. M. João, J. M. Alendouro Pinho, J. P. Santos Pires, and J. M. Viana Parente Lopes. Unambiguous Simulation of Diffusive Charge Transport in Disordered Nanoribbons, November 2023. arXiv:2311.03983 [cond-mat].
- [13] N. F. Mott. Conduction in non-crystalline materials. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 19(160):835–852, April 1969. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/14786436908216338>.
- [14] Patrick A. Lee and T. V. Ramakrishnan. Disordered electronic systems. *Reviews of Modern Physics*, 57(2):287–337, April 1985. Publisher: American Physical Society.
- [15] N.F. Mott and W.D. Twose. The theory of impurity conduction. *Advances in Physics*, 10(38):107–163, April 1961. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/00018736100101271>.
- [16] R. E. Borland and John Anthony Pople. The nature of the electronic states in disordered one-dimensional systems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 274(1359):529–545, January 1997. Publisher: Royal Society.
- [17] M. A. Paalanen, T. F. Rosenbaum, G. A. Thomas, and R. N. Bhatt. Critical Scaling of the Conductance in a Disordered Insulator. *Physical Review Letters*, 51(20):1896–1899, November 1983. Publisher: American Physical Society.
- [18] Shingo Katsumoto, Fumio Komori, Naokatsu Sano, and Shun-ichi Kobayashi. Fine Tuning of Metal-Insulator Transition in $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ Using Persistent Photoconductivity. *Journal of the Physical Society of Japan*, 56(7):2259–2262, July 1987. Publisher: The Physical Society of Japan.
- [19] B. Kramer and A. MacKinnon. Localization: theory and experiment. *Reports on Progress in Physics*, 56(12):1469, December 1993.
- [20] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan. Scaling Theory of Localization: Absence of Quantum Diffusion in Two Dimensions. *Physical Review Letters*, 42(10):673–676, March 1979. Publisher: American Physical Society.
- [21] A. MacKinnon and B. Kramer. One-Parameter Scaling of Localization Length and Conductance in Disordered Systems. *Physical Review Letters*, 47(21):1546–1549, November 1981. Publisher: American Physical Society.
- [22] A. V. Malyshev, V. A. Malyshev, and F. Dominguez-Adame. Monitoring the localization-delocalization transition within a 1D model with non-random long-range interaction. *Physical Review B*, 70(17):172202, November 2004. arXiv:cond-mat/0303092.

- [23] F. Wegner. Inverse participation ratio in 2+ dimensions. *Zeitschrift für Physik B Condensed Matter*, 36(3):209–214, September 1980.
- [24] Alexander D. Mirlin, Yan V. Fyodorov, Frank-Michael Dittes, Javier Quezada, and Thomas H. Seligman. Transition from localized to extended eigenstates in the ensemble of power-law random banded matrices. *Physical Review E*, 54(4):3221–3230, October 1996. Publisher: American Physical Society.
- [25] L. S. Levitov. Absence of Localization of Vibrational Modes Due to Dipole-Dipole Interaction. *Europhysics Letters*, 9(1):83, May 1989.
- [26] V Dobrosavljević, A. A Pastor, and B. K Nikolić. Typical medium theory of Anderson localization: A local order parameter approach to strong-disorder effects. *Europhysics Letters (EPL)*, 62(1):76–82, April 2003.
- [27] T. R. Albrecht, H. A. Mizes, J. Nogami, Sang-il Park, and C. F. Quate. Observation of tilt boundaries in graphite by scanning tunneling microscopy and associated multiple tip effects. *Applied Physics Letters*, 52(5):362–364, February 1988.
- [28] Guohong Li, A. Luican, J. M. B. Lopes dos Santos, A. H. Castro Neto, A. Reina, J. Kong, and E. Y. Andrei. Observation of Van Hove singularities in twisted graphene layers. *Nature Physics*, 6(2):109–113, February 2010. Number: 2 Publisher: Nature Publishing Group.
- [29] Ad Mirlin and Yv Fyodorov. The Statistics of Eigenvector Components of Random Band Matrices - Analytical Results. *JOURNAL OF PHYSICS A-MATHEMATICAL AND GENERAL*, 26(12):L551–L558, June 1993. Num Pages: 8 Place: Bristol Publisher: IoP Publishing Ltd Web of Science ID: WOS:A1993LN95000012.
- [30] A. D. Mirlin. Distribution of local density of states in disordered metallic samples: logarithmically normal asymptotics. *Physical Review B*, 53(3):1186–1192, January 1996. arXiv:cond-mat/9507082.
- [31] A. D Mirlin and Y. V Fyodorov. Statistical Properties of the Local Density of States and the Knight Shift in Disordered Wires. *Europhysics Letters (EPL)*, 25(9):669–674, March 1994.
- [32] A Mirlin. Statistics of energy levels and eigenfunctions in disordered systems. *Physics Reports*, 326(5-6):259–382, March 2000.
- [33] BI Altshuler, Ve Kravtsov, and Iv Lerner. Statistics of Mesoscopic Fluctuations and Instability of One-Parameter Scaling. *ZHURNAL EKSPERIMENTALNOI I TEORETICHESKOI FIZIKI*, 91(6):2276–2302, December 1986. Num Pages: 27 Place: Moscow Publisher: Mezhdunarodnaya Kniga Web of Science ID: WOS:A1986F487200029.
- [34] BI Altshuler, Ve Kravtsov, and Iv Lerner. Current Relaxation and Mesoscopic Fluctuations in Disordered Conductors. *ZHURNAL EKSPERIMENTALNOI I TEORETICHESKOI FIZIKI*, 94(4):258–275, April 1988. Num Pages: 18 Place: Moscow Publisher: Mezhdunarodnaya Kniga Web of Science ID: WOS:A1988N151000024.
- [35] M H Zare, Mohsen Amini, Farhad Shahbazi, and S A Jafari. Anderson transition in disordered bilayer graphene. *Journal of Physics: Condensed Matter*, 22(25):255503, June 2010.

- [36] Franz X. Bronold and Holger Fehske. Anderson localization of polaron states. *Physical Review B*, 66(7):073102, August 2002. Publisher: American Physical Society.
- [37] Krzysztof Byczuk, Walter Hofstetter, and Dieter Vollhardt. Mott–Hubbard transition vs. Anderson localization of correlated, disordered electrons. *Physical Review Letters*, 94(5):056404, February 2005. arXiv:cond-mat/0403765.
- [38] Gerald Schubert, Alexander Weisse, and Holger Fehske. Delocalisation transition in chains with correlated disorder. *Physica B: Condensed Matter*, 359-361:801–803, April 2005. arXiv:cond-mat/0406212.
- [39] Gerald Schubert, Alexander Weisse, and Holger Fehske. Unusual localisation effects in quantum percolation. *Physical Review B*, 71(4):045126, January 2005. arXiv:cond-mat/0406750.
- [40] A. Alvermann, G. Schubert, A. Weisse, F. X. Bronold, and H. Fehske. Characterisation of Anderson localisation using distributions. *Physica B: Condensed Matter*, 359-361:789–791, April 2005. arXiv:cond-mat/0406051.
- [41] N.P. Armitage, E.J. Mele, and Ashvin Vishwanath. Weyl and Dirac semimetals in three-dimensional solids. *Reviews of Modern Physics*, 90(1):015001, January 2018. Publisher: American Physical Society.
- [42] E. Fradkin. Critical behavior of disordered degenerate semiconductors. II. Spectrum and transport properties in mean-field theory. *Physical Review. B, Condensed Matter*, 33(5):3263–3268, March 1986.
- [43] J. H. Pixley, Pallab Goswami, and S. Das Sarma. Anderson localization and the quantum phase diagram of three dimensional disordered Dirac semimetals. *Physical Review Letters*, 115(7):076601, August 2015. arXiv:1502.07778 [cond-mat].
- [44] J. Klier, I. V. Gornyi, and A. D. Mirlin. From weak to strong disorder in Weyl semimetals: Self-consistent Born approximation. *Physical Review B*, 100(12):125160, September 2019. arXiv:1907.03018 [cond-mat].
- [45] Pallab Goswami and Sudip Chakravarty. Quantum criticality between topological and band insulators in $(3+1)$ -dimensions. *Physical Review Letters*, 107(19):196803, November 2011. arXiv:1101.2210 [cond-mat].
- [46] J. H. Pixley, David A. Huse, and S. Das Sarma. Rare region induced avoided quantum criticality in disordered three-dimensional Dirac and Weyl semimetals. *Physical Review X*, 6(2):021042, June 2016. arXiv:1602.02742 [cond-mat].
- [47] Justin H. Wilson, David A. Huse, S. Das Sarma, and J. H. Pixley. Avoided quantum criticality in exact numerical simulations of a single disordered Weyl cone. *Physical Review B*, 102(10):100201, September 2020. arXiv:2002.11722 [cond-mat].
- [48] Rahul Nandkishore, David A. Huse, and S. L. Sondhi. Rare region effects dominate weakly disordered three-dimensional Dirac points. *Physical Review B*, 89(24):245110, June 2014. Publisher: American Physical Society.

- [49] J. P. Santos Pires, B. Amorim, Aires Ferreira, İnanç Adagideli, Eduardo R. Mucciolo, and J. M. Viana Parente Lopes. Breakdown of universality in three-dimensional Dirac semimetals with random impurities. *Physical Review Research*, 3(1):013183, February 2021. arXiv:2010.04998 [cond-mat, physics:physics, physics:quant-ph].
- [50] J. P. Santos Pires. Quantum Effects of Impurities and Lattice Defects in Topological Semimetals, December 2022. arXiv:2212.11384 [cond-mat, physics:physics, physics:quant-ph].
- [51] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45(4):255, October 1950.
- [52] H. Tal-Ezer and R. Kosloff. An accurate and efficient scheme for propagating the time dependent Schrödinger equation. *The Journal of Chemical Physics*, 81(9):3967–3971, November 1984.
- [53] Simão M. João, Miša Anđelković, Lucian Covaci, Tatiana Rappoport, João M. V. P. Lopes, and Aires Ferreira. KITE: high-performance accurate modelling of electronic structure and response functions of large molecules, disordered crystals and heterostructures. *Royal Society Open Science*, 7(2):191809, February 2020. arXiv:1910.05194 [cond-mat, physics:physics].
- [54] R. N. Silver and H. Röder. Calculation of densities of states and spectral functions by Chebyshev recursion and maximum entropy. *Physical Review E*, 56(4):4822–4829, October 1997. Publisher: American Physical Society.
- [55] Shixun Zhang, Shinichi Yamagiwa, Masahiko Okumura, and Seiji Yunoki. Kernel Polynomial Method on GPU. *International Journal of Parallel Programming*, 41(1):59–88, February 2013.
- [56] Shixun Zhang, Shinichi Yamagiwa, Masahiko Okumura, and Seiji Yunoki. Performance Acceleration of Kernel Polynomial Method Applying Graphics Processing Units, May 2011. arXiv:1105.5481 [cond-mat, physics:physics].
- [57] Ju Li and Sidney Yip. Order-n method to calculate the local density of states. *Physical Review B*, 56(7):3524–3527, August 1997. Publisher: American Physical Society.
- [58] VI Berezinskii. Kinetics of a Quantum Particle in a One-Dimensional Random Potential. *ZHURNAL EKSPERIMENTALNOI I TEORETICHESKOI FIZIKI*, 65(3):1251–1266, 1973. Num Pages: 16 Place: Moscow Publisher: Mezhdunarodnaya Kniga Web of Science ID: WOS:A1973Q822300039.
- [59] D. J. Thouless. A relation between the density of states and range of localization for one dimensional random systems. *Journal of Physics C: Solid State Physics*, 5(1):77, January 1972.