Bachelor Thesis

Physics Faculty

of the

Ludwig-Maximilians-Universität München

Application of Linear Prediciton on the Chebyshev Expansion of Spectral Functions

Anwendung von Linear Prediction auf die Tschebyscheff Entwicklung von Spektralfunktionen

Christophe Pixius

Supervisor:

Prof. Dr. Jan von Delft

July 24, 2015

Contents

Abstract

In this work we want to demonstrate how linear prediction can be used to improve the Chebyshev expansion of spectral functions. The Chebyshev moments of a spectral function can be calculated recursively, using a matrix product state (MPS) approach. They can then be extrapolated with linear prediction, which leads to a drastic increase in precision of the reconstructed spectral function. We are going to look into two recently developed linear prediction setups and apply them to different systems. We will analyze the quality of the obtained results and discuss the influence of the system parameters on the error of the reconstruction. To make this work as self contained as possible we will first give an introduction into the underlying physical concepts. Also we will have a short look at the most important numerical principles that are applied in the framework of this thesis. The linear prediction methods will be tested on an artificial test function and on the Single Impurity Anderson Model (SIAM).

Chapter 1

Introduction

The SIAM is an example of a quantum impurity system. Such systems are characterized by a small subsystem (the impurity) with very few degrees of freedom, coupled to a bath of an infinite amount of degrees of freedom. As an example, we can imagine a magnetic atom residing in a nonmagnetic metal. The infinite bath thus contains the conduction electrons of the host metal.

Systems of this kind represent an interesting branch of modern solid state physics and can be used to describe a large amount of physical phenomena [1–3]. Because of the complexity of these models a purely analytical treatment is out of question in most cases. Also the strong interactions in these systems -e.g. the coulomb interaction of two electrons on the impurity- prevent a successful description based on perturbation theory. Therefore we often have to resort to a numerical treatment.

One of the most established computational methods to treat these quantum impurity models is the numerical renormalization group (NRG), first introduced by Wilson to treat the Kondo problem [4]. It discretizes the continuous bath of conduction electrons and maps it onto a discrete semi-infinite chain. The first site of this chain couples to the impurity and the coupling between the chain sites decreases exponentially. Thus, NRG can resolve the low energy physics of impurity systems with high accuracy. However, the high precision of this method is restricted to low energy scales; in order to avoid an exponential growth of the Hilbert space the high energy states of the system are eliminated in each iteration step.

This implies that the high energy feature of the spectral function can only be poorly resolved.

In order to be able to capture also the high energy physics of these impurity systems, a method with a uniform energy resolution is in order. This can be achieved by expanding the spectral function in terms of Chebyhsev polynomials, which has been introduced in the context of kernel polynomial methods [5]. In 2011 Holzner et al. [6] combined the Chebyshev expansion with matrix-product-state (MPS) techniques to compute the spectral function at zero temperature. To this end one determines the ground state of the system using the density-matrix-renormalization group (DMRG) [7, 8], and then recursively calculates the Chebyhsev moments of the series expansion. The spectral function can then be reconstructed from the calculated moments. Employing this method, we are able to resolve the whole spectral width of the spectral function *uniformly*. However, the calculation of the moments is numerically costly and can be plagued by finite-size effects for higher expansion orders.

Determining only a limited amount of expansion coefficients yields artificial oscillations, so called *Gibb's oscillations* in the reconstructed spectral function. These oscillations can be removed by employing damping factors, which comes at the price of 'smearing out' features of the spectral function itself and thus decreases spectral resolution [6].

Ganahl et al. [9] proposed to extrapolate the Chebyshev moments of spectral functions by applying linear prediction, which removes Gibb's oscillations *and* increases the precision of the reconstructed function considerably. Even more recently, Wolf et al. [10] have proposed an adaptation of linear prediction to further enhance the spectral resolution.

The goal of this thesis is to give an overview over the recently developed linear prediction methods [9,10]

in the context of spectral functions. In Chapter 2 we will first introduce some underlying physical concepts such as Green's functions and spectral functions. In Chapter 3 the basic numerical instruments will be discussed as well as the discretization of the SIAM hamiltonian. Finally we will present the results for various test- and spectral functions in Chapter 4 and give a conclusion in Chapter 5.

Chapter 2

Physical Prerequisites

This chapter focuses on the quantum mechanical background of the systems treated in this work. We are first going to introduce the SIAM Hamiltonian, which is a typical impurity model. Secondly we will have a look at the Green's function, since it is a very powerful and important tool to solve problems in solid state physics. This then leads us to the concept of spectral functions, which will be treated in the last part of this chapter.

2.1   The Single Impurity Anderson Model

The SIAM is a model to describe magnetic impurities in metals. These impurities are ions with a magnetic moment, occupying lattice sites in the metallic host. Because of their magnetic properties, these ions can interact with the conduction electrons of the host metal, leading to a vast number of interesting physical phenomena, such as Kondo correlations.

In order to understand this impurity model, we are first going to study the Hamiltonian and then describe and explain the terms individually. In second quantization, the Hamiltonian reads:

$$H = H_0 + H_U + H_{hyb} \tag{2.1}$$

with

$$H_0 = \sum_{\sigma} \epsilon_d c_{\sigma d}^\dagger c_{\sigma d} + \sum_{\sigma \mathbf{k}} \epsilon_{\mathbf{k}} c_{\sigma \mathbf{k}}^\dagger c_{\sigma \mathbf{k}}, \tag{2.2}$$

$$H_U = U c_{\uparrow d}^\dagger c_{\uparrow d} c_{\downarrow d}^\dagger c_{\downarrow d}, \tag{2.3}$$

$$H_{hyb} = \sum_{\sigma \mathbf{k}} \left(t_{\mathbf{k}} c_{\sigma \mathbf{k}}^\dagger c_{\sigma d} + t_{\mathbf{k}}^* c_{\sigma d}^\dagger c_{\sigma \mathbf{k}} \right). \tag{2.4}$$

where $c^\dagger$ and c are the fermionic creation and annihilation operator, respectively. The physical interpretation of the various constants U, ϵ_d, $\epsilon_{\mathbf{k}}$ and $t_{\mathbf{k}}$ ($t_{\mathbf{k}}^*$) will be discussed later. First, we elaborate on the creation and annihilation operators in more detail, since they are the key to understanding the Hamiltonian.

The fermionic creation operator $c_\nu^\dagger$ creates a fermion, in our case an electron, in the quantum state characterized by the quantum number ν, whereas the annihilation operator c_ν destroys a particle in that state. These operators, although not hermitian, fulfil the following anticommutation relations:

$$\{c_{\nu_i}, c_{\nu_j}^\dagger\} = \delta_{\nu_i, \nu_j} \qquad\qquad \{c_{\nu_i}^{(\dagger)}, c_{\nu_j}^{(\dagger)}\} = 0. \tag{2.5}$$

where the brackets represent the anticommutator $\{A, B\} = AB + BA$. Using $c^\dagger$ and c we can construct another useful operator, namely the number operator, defined as

$$n_\nu = c_\nu^\dagger c_\nu. \tag{2.6}$$

When applied to a quantum state this operator counts the number of particles occupying this particular state. In the fermionic case, this number can be either 0 or 1, since the Pauli Principle prohibits any occupation number larger than 1. Per definition, the number operator is hermitian and thus an observable. In the following, we are going to analyze the individual terms of the Hamiltonian in equation (2.1).

The non-interacting term

$$H_0 = \sum_\sigma \epsilon_d c_{\sigma d}^\dagger c_{\sigma d} + \sum_{\sigma \mathbf{k}} \epsilon_\mathbf{k} c_{\sigma \mathbf{k}}^\dagger c_{\sigma \mathbf{k}} \tag{2.7}$$

contains two sums, the first describing the energy of free electrons in the outer shell of the impurity atom. Since this is often the d-shell the creation and annihilation operators are labeled with the index d. Comparing this term to equation (2.6) we see that it is nothing else but the sum over the two spin configurations of the number operator multiplied with the energy ϵ_d. Thus, the total (non-interacting) energy of the impurity is determined by its occupation number and the constant ϵ_d

The second sum represents the energy of the electrons in the conduction band. Analogous to the energy of the d-electrons, it is given by the number operator multiplied with the energy $\epsilon_\mathbf{k}$ of the occupied state. It turns out that it is generally a good approximation to treat the conduction band as a system of free electrons [3].

The interaction term of the SIAM

$$H_U = U c_{\uparrow d}^\dagger c_{\uparrow d} c_{\downarrow d}^\dagger c_{\downarrow d} \tag{2.8}$$

contains the particle density for spin up as well as spin down electrons. It describes the interaction of two electrons in the d-shell of the impurity ion due to the Coulomb potential. The constant U is the strength of the interaction.

Since the d-shell of the impurity ion and the conduction band of the metal are energetically close, it is possible for electrons to jump from the d-shell to the conduction band and the other way around. This hybridization of the d-orbital is represented by

$$H_{hyb} = \sum_{\sigma \mathbf{k}} \left(t_\mathbf{k} c_{\sigma \mathbf{k}}^\dagger c_{\sigma d} + t_\mathbf{k}^* c_{\sigma d}^\dagger c_{\sigma \mathbf{k}} \right). \tag{2.9}$$

The expression contains on the one hand the combination $c_{\sigma \mathbf{k}}^\dagger c_{\sigma d}$, where we create an electron in the conduction band with momentum $\mathbf{k}$ and at the same time annihilate a d-electron. On the other hand we can also create an electron in the d-shell and annihilate a conduction electron, which is represented by the term $c_{\sigma d}^\dagger c_{\sigma \mathbf{k}}$. Expressions of this form are called hopping terms, since they mean that an electron is 'hopping' from one state to another. The constants $t_\mathbf{k}$ and $t_\mathbf{k}^*$ represent the probability for such a process to occur.

2.1.1 Remark: The Statistical Ensemble

Since we are interested in many-body problems throughout this whole thesis it is important to be precise on which statistical ensemble we are working in. It is favourable to choose the grand canonical ensemble when working with quantum particles. This means that neither energy nor particle number are fixed values, but can only be known in average. In the grand canonical ensemble the expectation value of an observable A is defined as

$$\langle A \rangle = \frac{1}{Z} \text{Tr} \left(A e^{-\beta(H - \mu N)} \right), \tag{2.10}$$

where $Z = \text{Tr}(e^{-\beta(H-\mu N)})$ is the partition function. It is important to note that the time evolution of operators in this ensemble is dependent on $H' = H - \mu N$ instead of H, which means that the particle energies are measured respectively to the chemical potential μ. In this whole thesis the operator H therefore has to be understood as $H - \mu N$ and the eigenenergies $\epsilon_{\mathbf{k}}$ as $\epsilon_{\mathbf{k}} - \mu$, even though we don't write this explicitly.

2.2 Green's -and Spectral Function

2.2.1 The Green's Function

The theoretical analysis of interacting many-body systems is often based on dynamical correlation functions or Green's functions. Analogous to classical electrodynamics, the Green's function $G(\mathbf{r}, \mathbf{r}')$ of a many-body system can be defined for a particular differential operator D. This function fulfills the relation:

$$D\, G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'). \tag{2.11}$$

The one-particle retarded Green's Function $G^R(\mathbf{r}, \mathbf{r}')$ for many-body systems in position space has the following form:

$$G^R(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') \equiv -i\theta(t - t')\langle\{\Psi(\mathbf{r}\sigma, t), \Psi^\dagger(\mathbf{r}'\sigma, t')\}\rangle, \tag{2.12}$$

where $\Psi(\mathbf{r}, t)$ and $\Psi^\dagger(\mathbf{r}', t')$ are so called quantum field operators, which are essentially spatial creation and annihilation operators. Analogously to $c_\nu^\dagger$ and c_ν, $\Psi^\dagger(\mathbf{r}, t)$ and $\Psi(\mathbf{r}, t)$ create, respectively destroy a fermion at a spatial point $\mathbf{r}$. Of course we can make this definition more general, simply using $c_\nu^\dagger$ and c_ν for any set of quantum numbers ν instead of $\Psi^\dagger(\mathbf{r}, t)$ and $\Psi(\mathbf{r}, t)$:

$$G^R(\nu t, \nu' t') \equiv -i\theta(t - t')\langle\{c_\nu(t), c_{\nu'}^\dagger(t')\}\rangle. \tag{2.13}$$

The Green's function is called retarded because it is nonzero only for times $t > t'$, which guarantees causality. Note that this expression holds for bosons as well if we change the anticommutator into a commutator.

In addition to the retarded Green's function, it is common to introduce two other Green's functions, namely $G^>(\nu t, \nu' t')$ (G-greater) and $G^<(\nu t, \nu' t')$ (G-lesser) defined as (again in the fermionic case):

$$G^>(\nu t, \nu' t') \equiv -i\langle c_\nu(t) c_{\nu'}^\dagger(t')\rangle \qquad\qquad G^<(\nu t, \nu' t') \equiv i\langle c_{\nu'}^\dagger(t') c_\nu(t)\rangle. \tag{2.14}$$

The retarded Green's function $G^R(\nu t, \nu' t')$ as defined in (2.12) can be expressed in terms of the greater- and lesser Green's functions by

$$G^R(\nu t, \nu' t') = \theta(t - t')\left(G^>(\nu t, \nu' t') - G^<(\nu t, \nu' t')\right). \tag{2.15}$$

The expectation values are taken with respect to the grand canonical ensemble that we have introduced in 2.1.1. As is shown in Appendix A, the Green's function of free particles in the frequency domain reads:

$$G_0^R(\sigma\mathbf{k}; \omega) = \frac{1}{\omega - \epsilon_{\mathbf{k}} + i\eta}. \tag{2.16}$$

This makes sense, because the Hamiltonian of free particles is diagonal in $c^\dagger c$ (see 2.7). This implies that the Green's function must have the same property, which is evident in (2.16). Physically, this means that a particle with momentum $\mathbf{k}$ will always stay in this state, which is consistent with the assumption that the particles do not interact.

2.2.2 The Spectral Function

Starting from the Green's function we can define what we call the *spectral function*, or *spectral density*. It consists of the imaginary part of the retarded Green's function $G^R(\nu, \omega)$:

$$A(\nu, \omega) = -\frac{1}{\pi} \text{Im}\, G^R(\nu, \omega). \tag{2.17}$$

This function has a few useful properties, which make it particularly interesting to work with and. Very often solving a model in solid state physics consists in finding exactly this particular function. It can be shown, that the spectral function $A(\nu, \omega)$ satisfies

$$\int_{-\infty}^{+\infty} \mathrm{d}\omega\, A(\nu, \omega) = 1, \tag{2.18}$$

as well as $A(\nu, \omega) \geq 0$. To prove these relations, one needs the Lehmann representation of the Green's Functions, which we are not going to introduce here - for a detailed calculation, see [3]. Nevertheless we can check that they hold for the example of free particles. Using the Green's function in (2.16), we find the relation:

$$A_0(\sigma \mathbf{k}; \omega) = \delta(\omega - \epsilon_{\mathbf{k}}). \tag{2.19}$$

Per definition the Dirac delta is always larger or equal 0 and normalized to 1 and thus A_0 respects the required properties.
The equations (2.18) and $A(\nu, \omega) \geq 0$ suggest that $A(\nu, \omega)$ can be interpreted as a probability density. Thus $A(\nu, \omega)\mathrm{d}\omega$ can be seen as the probability for a particle with quantum number ν to have an energy lying in $[\omega, \omega + \mathrm{d}\omega]$.

Zero temperature spectral functions

The general expression for *zero-temperature* spectral functions can also be derived quite quickly from the Green's functions. As we have seen in Appendix A, the greater and lesser Green's functions can be interpreted as propagators for holes and electrons respectively [3]. From both these propagators, we can formally derive an expression for two spectral functions $A^>(\omega)$ and $A^<(\omega)$. The physical interpretation of the respective Green's functions suggests that $A^>(\omega)$ stand for the unoccupied and $A^<(\omega)$ for the occupied part of the spectrum [9]. We will execute this short calculation for the greater function $G^>(\omega)$ only. We begin by writing down the formal expression of the Green's function in time domain:

$$G^>(t, t') = -i \langle 0|c(t)c^\dagger(t')|0\rangle \tag{2.20}$$

where $|0\rangle$ denotes the energy ground state of the system. We can now explicitly express the time dependence of the operators, which yields

$$G^>(t, t') = -i \langle 0|c\, e^{-i(\hat{H} - E_0)(t - t')}c^\dagger|0\rangle \tag{2.21}$$

with E_0 being the ground state energy. Note that in this step we have explicitly used the $T = 0$ condition. For finite temperatures the system is in general not in an eigenstate of the Hamiltonian, which is the critical requisit of this calculation.
Performing the Fourier transformation like in Appendix A and extracting the imaginary part yields the spectral function $A^>(\omega)$:

$$A^>(\omega) = \langle 0|c\, \delta(\omega - (\hat{H} - E_0))\, c^\dagger|0\rangle. \tag{2.22}$$

Analogously, we can compute $A^<(\omega)$:

$$A^<(\omega) = \langle 0|c^\dagger\, \delta(\omega - (\hat{H} - E_0))\, c|0\rangle. \tag{2.23}$$

The full spectral function $A(\omega)$ is then given by the sum of these two functions:

$$A(\omega) = A^>(\omega) + A^<(-\omega). \tag{2.24}$$

In general, the full spectral function is smooth, whereas $A^>(\omega)$ and $A^<(-\omega)$ have a step at $\omega = 0$. We are going to discuss the implications of this property in the next chapter.

2.2.3 The SIAM Spectral Function

Having introduced the SIAM Hamiltonian and the expression for zero temperature spectral function earlier in this chapter, we can now find an expression for the spectral density of the Anderson model. In general this is however a very difficult task, since the onsite interaction leads to non-diagonal terms in the Hamiltonian. What we can do as a first step is to ignore the Coulomb interaction of the electrons on the impurity by setting U to 0. This limit is known as the *resonating level model* (RLM). In this case, using the equation of motion ansatz (see [3]) we can derive an analytical expression for the spectral function:

$$A(\omega) = -\frac{1}{\pi}\text{Im}\left(\frac{1}{\omega - \epsilon_d + \Sigma(\omega)}\right), \tag{2.25}$$

with

$$\Sigma(\omega) = \Gamma\left(i + \frac{1}{\pi}\ln\left(\frac{1-\omega}{1+\omega}\right)\right). \tag{2.26}$$

Γ is determined by the hybridization: $\Gamma = \pi V^2 \rho(0)$ and gives the width of the spectral function.
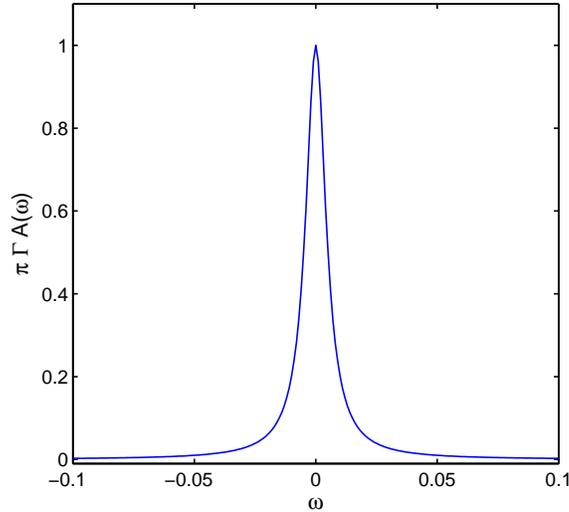
Figure 2.1: The rescaled resonating level model (RLM) spectral function for $\Gamma = 0.005$ and $\epsilon_d = 0$. In this limit we ignore the coulomb interaction of the electrons on the impurity atom.

This expression is structurally not far from the spectral function of free fermions (2.19). Instead of a delta function, which is infinitely narrow and high, we have a peaked curve of finite width Γ, given by the hybridization (see Figure 2.1). The maximum value of $A(\omega)$ is reached for $\omega = 0$ and it's value is given by $A(0) = (\pi\Gamma)^{-1}$. For $\Gamma = 0.005$ this corresponds to $A(0) \approx 63.662$. In the current literature the spectral function is usually plotted in a rescaled form as $\pi\Gamma A(\omega)$ as is shown in Figure 2.1.

As mentioned above it is generally not easily possible to find closed analytical expressions for the spectral densities of complex systems. We therefore have to resort to a numerical treatment. There are different approaches to finding spectral functions for quantum impurity systems systems, e.g. the aforementioned NRG or t-DMRG in combination with MPS. The code that I used is based on the Chebyshev expansion of the spectral function, where the ground state is calculated with MPS. The concepts of this method will be explained in the next chapter.

Chapter 3

Numerical Preparations

In this chapter I want to focus on the tools needed for the calculation of spectral functions of impurity models and of the SIAM in particular. It is possible to reconstruct spectral functions numerically by expanding them in Chebyshev polynomials and determining the ground state with MPS techniques (CheMPS). Since the latter is far beyond the scope of this work, I want to focus on improving the Chebyshev expansion by using linear prediction. To be as complete as possible we will first have a look at the properties of Chebyshev polynomials and the expansion of functions in terms of them. We are also going to treat the expansion of spectral functions and the needed rescaling of the Hamiltonian. In the second section we are going to introduce linear prediction in the context of Chebyshev moments and outline the subtleties of this numerical technique. In the last section of this chapter we are then going to show how the SIAM Hamiltonian has to be discretized in order to be treatable numerically.

3.1   Chebyshev Expansion

The Chebyshev function of first kind T_n are a set of orthonormal polynomials of degree n. They can be defined by the recursive relation

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \qquad\qquad T_0(x) = 1, \qquad\qquad T_1(x) = x, \qquad (3.1)$$

or alternatively by the expression

$$T_n(x) = \cos(n \arccos x). \qquad (3.2)$$

Since they form an orthonormal basis in the interval [-1,1] any smooth and continuous function can be expanded in terms of them:

$$f(x) = \frac{1}{\pi\sqrt{1-x^2}} \left(\mu_0 + 2 \sum_{n=1}^{\infty} \mu_n T_n(x) \right). \qquad (3.3)$$

Here μ_n are the expansion coefficients and the factor $(\pi\sqrt{1-x^2})^{-1}$ is a weight function, which assures that the Chebyshev polynomials are actually orthonormal:

$$\langle T_n, T_m \rangle = \int_{-1}^{1} \frac{1}{\pi\sqrt{1-x^2}} T_n(x) T_m(x) \, \mathrm{d}x = \delta_{n,m}. \qquad (3.4)$$

Combining (3.3) and (3.4) it follows that the coefficients μ_n are given by

$$\mu_n = \int_{-1}^{1} T_n(x) f(x) \, \mathrm{d}x. \qquad (3.5)$$

Having introduced the most important features of Chebyshev expansions, we can now discuss their application for the expansion of the spectral function.
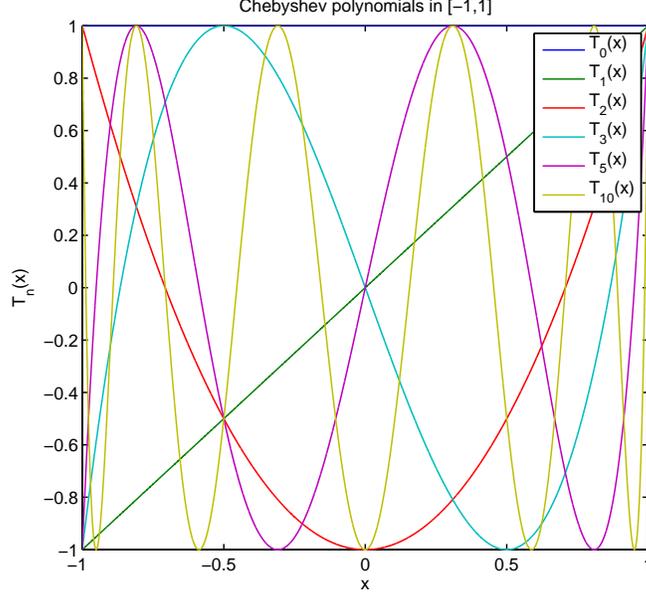
Figure 3.1: A few Chebyshev Polynomials in the interval I=[-1,1]. An important feature of these polynomials is that within I, we have $|T_n(x)| \leq 1$ and that the extremal values of T_n are either 1 or -1 for any n.

3.1.1 Expansion of the spectral function

Our goal is to find a method to numerically determine the spectral function of a system. The approach of CheMPS is to expand the spectral function in Chebyshev polynomials and to recursively determine the expansion coefficient.

Starting with the expression of the one-particle spectral function in Eqs. (2.22-2.24) at zero temperature, we expand the quantity in terms of the orthonormal Chebyhsev polynomials. Using (3.3) and (3.5), as well as the properties of the delta function we find

$$\delta(\omega - H) = \frac{1}{\pi\sqrt{1-x^2}}\left(1 + 2\sum_{n=1}^{\infty} T_n(H)T_n(\omega)\right), \tag{3.6}$$

and thus

$$A^>(\omega) = \frac{1}{\pi\sqrt{1-\omega^2}}\left(\langle 0|cc^\dagger|0\rangle + 2\sum_{n=1}^{\infty}\langle 0|c\,T_n(H)c^\dagger|0\rangle\,T_n(\omega)\right). \tag{3.7}$$

Comparing this equation to the general expression of the Chebyshev expansion (3.3), we can identify the coefficients as $\mu_0^> = \langle 0|cc^\dagger|0\rangle$ and $\mu_n^> = \langle 0|c\,T_n(H)c^\dagger|0\rangle$. These coefficients can be calculated recursively by exploiting relation (3.1). This can be achieved quite efficiently in the MPS framework. Since a detailed discussion of this numerical approach is beyond the scope of this thesis, we refer to Ref. [6,9] for a detailed introduction of CheMPS.

Of course we can calculate the expansion of the negative part of the spectral function $A^<(\omega)$ analogously.

With $A(\omega) = A^>(\omega) + A^<(-\omega)$ the full spectral function then reads:

$$A(\omega) = \frac{1}{\pi\sqrt{1-\omega^2}}\left([\mu_0^> + \mu_0^<] + 2\sum_{n=1}^{\infty}[\mu_n^> T_n(\omega) + \mu_n^< T_n(-\omega)]\right). \tag{3.8}$$

Using the relation $T_n(-\omega) = (-1)^n T_n(\omega)$ we find

$$A(\omega) = \frac{1}{\pi\sqrt{1-\omega^2}} \left(\mu_0 + 2\sum_{n=1}^{\infty} \mu_n T_n(\omega) \right), \tag{3.9}$$

with $\mu_0 = \mu_0^> + \mu_0^<$ and $\mu_n = \mu_n^> + (-1)^n \mu_n^<$.

The ground state $\langle 0|cc^\dagger|0\rangle$ can be calculated in the MPS framework, and from there we can determine the expressions for higher n recursively using (3.1). This recursion is however numerically very costly and thus only a certain number of moments can be calculated until the numerical expense becomes too high. This implies that the reconstruction of the spectral function will stop at a finite value N:

$$A(\omega) = \frac{1}{\pi\sqrt{1-\omega^2}} \left(\mu_0 + 2\sum_{n=1}^{N} \mu_n T_n(\omega) \right). \tag{3.10}$$

Of course this finite expansion order leads to considerable errors in the reconstructed function, which take the form of artificial oscillations, so called *Gibbs oscillations*. One method to avoid these oscillations is to smear them out using broadening kernels. The idea is that a convolution of an oscillating function with a cleverly chosen peaked function - either Gaussian or Lorentzian - will erase the oscillations (see Figure 3.2). A more extensive discussion on the broadening kernels can be found in Ref. [6]. The problem with this method is that the interesting features of the spectral function are smeared out out as well. For this reason the authors of [9] propose to extrapolate the calculated moments by using a linear prediction algorithm. As we will see, this eliminates Gibbs oscillations and leads to a drastic increase in accuracy of the reconstructed spectral function.
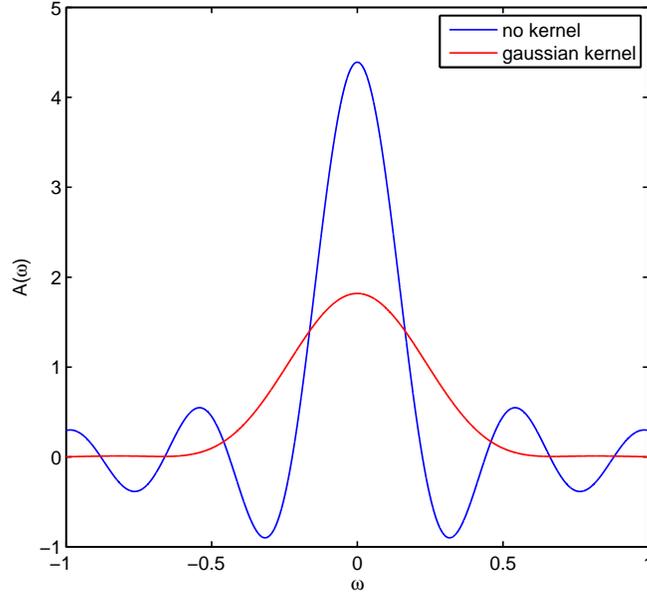
Figure 3.2: The spectral function of the resonating level model for a chain of 25 sites and $\Lambda = 1.3$. The blue curve is obtained by reconstructing the spectral function with 100 moments and no kernel. The red curve shows the reconstructed function when a gaussian kernel is employed. Note that in this plot neither of the functions has been rescaled.

3.1.2 Rescaling of $\hat{H}$

In the previous section we have seen that the expansion of $A(\omega)$ in terms of Chebyshev functions is indeed a valid approach to find the spectral function. We have to keep in mind however that in general the support of the spectral function may reach beyond the interval [-1,1]. This implies that it is generally not possible to expand the spectral function directly, since the Chebyshev polynomials $T_n(x)$ rapidly diverge for $|x| > 1$. We thus have to rescale the support of $A(\omega)$, i.e. the bandwidth, to the interval I=[-1,1]. This rescaling goes as follows [9,10]: Let us assume that the spectral function is nonzero in the interval $\omega \in$ [-W,W]. Then we are searching for a linear function, mapping this interval into I, which can be achieved by defining

$$\tilde{H} = f(H) = \frac{H - E_0}{a} + b \qquad\qquad x = \frac{\omega}{a} + b. \qquad (3.11)$$

If a and b are chosen appropriately, our rescaled energy $x \in$ [-1,1]. Now we can express the new spectral function $\tilde{A}^{>}$ in terms of the rescaled energy x as

$$\tilde{A}^{>}(x) = \langle 0|c\,\delta(x - \tilde{H})\,c^{\dagger}|0\rangle \qquad (3.12)$$

and the two functions $\tilde{A}^{>}(x)$ and $A^{>}(\omega)$ are connected by the relation

$$A^{>}(\omega) = \frac{1}{a}\tilde{A}^{>}\left(\frac{\omega}{a} + b\right). \qquad (3.13)$$

The factor $\frac{1}{a}$ follows from the property $\delta(ax) = \frac{1}{a}\delta(x)$ of the delta function. During the rest of this work we will always be working with the rescaled spectral function if not stated elsewise. I will therefore omit the tilde and replace x by ω, and thus $A(\omega)$ is from now on the *rescaled* spectral function.

3.2 Linear Prediction

Linear prediction is a mathematical method mainly used in statistical signal processing. It can be traced back as far as 1941 and still finds many applications nowadays [11]. In the context of strongly correlated quantum systems it was first introduced by Ref. [12] and [13] for zero- and finite-temperature t-DMRG. Later, in 2014 the authors of [9] proposed to use linear prediction to extrapolate the Chebyshev moments of spectral functions calculated with CheMPS. Up until now there are still improvements to be made to further ameliorate the precision of this numerical method, most recently shown by the authors of [10].

The concept of linear prediction follows a simple idea: Assuming that we have a set of values, for example our Chebyshev moments μ_n ($n \in [0, N]$), with a certain functional decay, then we can make the Ansatz:

$$\mu_k \approx \tilde{\mu}_k = -\sum_{i=1}^{L} a_i \mu_{k-i} \qquad (3.14)$$

for $k > N$ and an, a priori, unknown set of fixed coefficients a_i. This formula states that any μ_k can be approximated by a linear superposition of the previous L moments. Therefore we can predict the evolution of values for larger expansion orders if we know their functional decay, which is contained in the coefficients a_i. Predicting μ_k thus comes down to finding these coefficients.

The Ansatz to determine the values of a_i is to minimize the mean square error

$$F = \sum_{n=T_0+1}^{T} \frac{|\tilde{\mu}_n - \mu_n|^2}{w_n}. \qquad (3.15)$$

Here w_n is a weighing function that we set to 1 and the parameter n runs over the set $\{T_0 + 1, T_0 + 2, ..., T\}$, which is called the fitting, or *training interval*. Minimizing this function F means that we want our reconstructed moments to be as close to the corresponding exact moments as possible. As is shown in Appendix B, the minimization of F leads to the set of linear equations

$$\mathbf{a} = -R^{-1}\mathbf{r}, \tag{3.16}$$

where $\mathbf{a}$ contains the coefficients a_i, and R and $\mathbf{r}$ are defined by

$$R_{ij} = -\sum_{n=T_0+1}^{T} \frac{\mu_{n-i}^* \mu_{n-j}}{w_n} \qquad \text{and} \qquad r_i = -\sum_{n=T_0+1}^{T} \frac{\mu_{n-i}^* \mu_n}{w_n}. \tag{3.17}$$

To determine the coefficients a_i in (3.13), the matrix R has to be inverted. This can cause problems, since it is nearly singular. It is therefore necessary to perform the inversion only up to a cutoff $\delta \approx 10^{-6}$ (pseudo-inverse) instead of calculating the full inverse, as is pointed out in Ref. [9, 10].
We can now write the coefficients into a matrix M:

$$M = \begin{pmatrix} -a_1 & -a_2 & -a_3 & \ldots & -a_L \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}. \tag{3.18}$$

M is called the companion matrix. The reason why we introduce it is because numerical inaccuracies can lead to a divergence of the predicted moments. This occurs when the eigenvalues of M are larger than 1. In order for the extrapolation to converge, these eigenvalues have to be either eliminated or rescaled. To do this we first have to diagonalize M:

$$\Lambda = T^{-1}MT.$$

Now we set all eigenvalues $\lambda_i > 1$ to 0, thus defining a new diagonal matrix Λ', which we can use to reconstruct the companion matrix:

$$M' = T\Lambda'T^{-1}.$$

Linear prediction now consists in multiplying $\mu = (\mu_N, \mu_{N-1}...\mu_{N-L})^T$ with powers of M':

$$\mu_{N+k} = \sum_{j=1}^{L} (M'^k)_{1j} \, \mu_{N-j}. \tag{3.19}$$

For linear prediction to provide precise results is is important that the initial moments show an *exponential decay*. For the Chebyshev moments the decay is exponential if the expanded function is smooth and analytical, as is pointed out in Ref. [9].
We have introduced the parameters L, T and T_0 in this section, but how should they be chosen to optimize the prediction? First of all, there is one condition that has to be fulfilled, namely that $T_0 > L$, otherwise the definition (3.14) of R and $\mathbf{r}$ would allow for negative indices. The choice of L determines how many known moments are used to predict the new μ. As it turns out L should not be chosen too large, because one runs the risk of overfitting. We follow the suggestion of Ref. [10] and use $L = \min((T - T_0)/2, 100)$. For the choice of T and T_0 Ref. [10] proposed $T = N$ and $T_0 = 3N/4$. We have tried a few different intervals and found that it makes little difference, as long as the interval is chosen large enough to caption the functional behaviour and the initial data show no sign of finite size effects. We have therefore settled with the same training interval, i.e. $[3N/4, N]$.

3.3 Discretization of the SIAM Hamiltonian

In Chapter 2 we have introduced the Hamiltonian of the SIAM and explained the individual expressions. The full expression reads

$$H = \sum_\sigma \epsilon_d c_{\sigma d}^\dagger c_{\sigma d} + U c_{\uparrow d}^\dagger c_{\uparrow d} c_{\downarrow d}^\dagger c_{\downarrow d} + \sum_{\sigma \mathbf{k}} \epsilon_\mathbf{k} c_{\sigma \mathbf{k}}^\dagger c_{\sigma \mathbf{k}} + \sum_{\sigma \mathbf{k}} \left(t_\mathbf{k} c_{\sigma \mathbf{k}}^\dagger c_{\sigma d} + H.c. \right). \tag{3.20}$$

where $H.c.$ denotes the hermitian conjugate.

However we can not numerically treat the Hamiltonian in this form, since it contains the infinite sum over all possible momenta $\mathbf{k}$. Simply truncating the sum at a random value $\mathbf{k}'$ leads to wrong results, because there is no physical argument why the momenta beyond this value can be ignored. In order to make the model accessible for numerical calculations, we therefore need to introduce a discretization of the energy spectrum of the continuous band. Following Wilson's NRG approach, we choose a logarithmic discretization with a discretization parameter Λ and map the Hamiltonian onto a chain geometry. This mapping can be done analytically if we assume a constant hybridization $t_\mathbf{k} = V$ and a flat symmetric bath-spectral function $\rho(\omega) = \sum_\mathbf{k} \delta(\omega - \epsilon_\mathbf{k}) = 1/2$ for $\omega \in [-1, 1]$. The transformed Hamiltonian reads:

$$H = \sum_\sigma \epsilon_d c_{\sigma d}^\dagger c_{\sigma d} + U c_{\uparrow d}^\dagger c_{\uparrow d} c_{\downarrow d}^\dagger c_{\downarrow d} + V^2 \sum_\sigma \left(c_{\sigma d}^\dagger c_{\sigma 1} + H.c. \right) + \sum_{\sigma, n=1}^\infty \left(h_n c_{\sigma n}^\dagger c_{\sigma n+1} + H.c. \right). \tag{3.21}$$

In this mapping, the infinite bath of conduction electrons engulfing the impurity is transformed into a discrete chain. The impurity only couples to the first site of this chain and the transition probability between the sites is determined by the hopping parameters h_n, which are given by

$$h_n = \frac{(1 + \Lambda^{-1})(1 - \Lambda^{-n-1})\Lambda^{-n/2}}{2\sqrt{(1 - \Lambda^{-2n-1})(1 - \Lambda^{-2n-3})}}.$$

The energy of a bath site decreases exponentially with n: $E_n = \pm\Lambda^{-n}$.

The term $V^2 \sum_\sigma \left(c_{\sigma d}^\dagger c_{\sigma 1} + H.c. \right)$ denotes the coupling of the impurity to the first bath site, and thus V is a measure for the hybridization strength. Last but not least, we have to cut the chain of bath sites at a finite value N, which corresponds to a cutoff at low energies E_N. Note that the Hamiltonian has already been rescaled, otherwise the bandwidth D would generally not be 1 and $\omega = [-D, D]$.

With this discretization the SIAM is now treatable with numerical methods. One of the most powerful techniques to treat impurity models is the NRG [2, 4]. It is strongly dependent on the logarithmic discretization, and Λ is usually chosen > 2, since lower values considerably increase the computational cost. This implies that we have an excellent resolution around the chemical potential, but a very poor one for high energies. The choice of large Λ is what makes the NRG a very established method to examine low energy properties of a system, but also what prevents its applicability to high energies.

CheMPS and MPS based methods in general do not require the discretization to be logarithmic, because they use a different truncation scheme. Thus a high spectral resolution can be achieved over the whole energy spectrum, making it useful to analyze higher energy regimes of a model. An extensive review of MPS methods is given by Ref. [7].

Chapter 4

Numerical Calculations and Results

In this chapter I want to present two different setups of linear prediction for the computation of spectral functions and compare the respective results. To this end, we will first introduce both setups and test them on an artificial test function. We will then move on to the RLM spectral function and compare the accuracy of the results of the two setups. Finally we are going to treat the SIAM spectral function for nonzero onsite interaction and discuss the results.

4.1   Discussion of the two setups

The idea of extrapolating the Chebyshev moments of spectral functions with linear prediction was first introduced in 2014 by Ganahl et al. [9]. In their paper the authors argued that it is advantageous to predict the moments $\mu_n = \mu_n^> + (-1)^n \mu_n^<$ of the *full spectral function* $A(\omega)$. Since the latter is smooth, its moments decay *exponentially*, which makes linear prediction a precise technique. Thus the obtained results show a high increase in accuracy compared to the kernel polynomial method, which up to then was the method of choice to eliminate Gibbs oscillations [6]. The setup of Ganahl et al. requires the choice of $b = 0$ for the rescaling of the Hamiltonian.

More recently the authors of Ref. [10] proposed a new linear prediction setup to further increase the accuracy of the reconstructed spectral function. They argued that a gain in precision by a factor ~ 6.4 can be achieved if the dominant contribution to $A^>$ and $A^<$ is shifted to the edge of the interval [-1,1] respectively, by choosing $b \approx -1$ for the rescaling. The idea is that the Chebyshev polynomials $T_n(x)$ oscillate faster around $x \approx \pm 1$ than around $x \approx 0$, which can be observed in Figure 3.1. We can also see this by determining the positions of the extrema of the n-th polynomial in the interval [-1,1]:

$$\frac{d}{dx} \cos(n \arccos x_n) = 0. \tag{4.1}$$

After a short calculation we find an expression for the location of the extrema:

$$x_n = \cos\left(\frac{k\pi}{n}\right) \qquad \text{for } k = 0, 1, \dots, n. \tag{4.2}$$

Calculating some values for a given n reveals that most of these extrema are located at high values of x. The faster oscillation of the polynomials at the edges of the interval directly implies that the resolution of a Chebyshev expansion is higher in these areas than around $\omega = 0$. A better resolution implies that the polynomials capture more information of the spectral function.

However it is not possible to predict the moments of the full spectral function in this setup, since its support would be shifted outside of [-1,1]. We therefore work with the positive and negative part of the spectral function seperately, which have a step at $\omega = 0$. Without loss of generality we continue the discussion of the positive part $A^>(\omega)$ only.
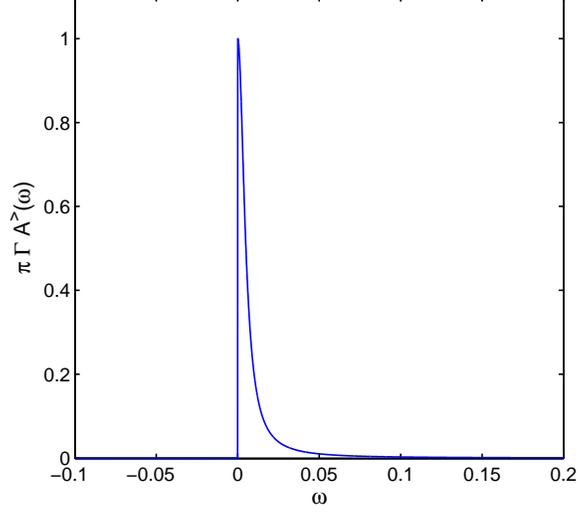
Figure 4.1: The positive part of the rescaled RLM spectral function with $\Gamma = 0.005$. We can shift it very closely to the edges of the interval [-1,1] to enhance the spectral resolution per expansion order, but this implies that the moments do not decay exponentially.

The discontinuity of $A^>(\omega)$ implies that their Chebyshev moments do not decrease exponentially, but rather show an *algebraic* decay, which can not be handled by linear prediction.
Nevertheless, the continuity of the spectral function can be restored by defining

$$A'^>(\omega) = A^>(\omega) - A^>(0), \quad \forall\, \omega \geq b. \tag{4.3}$$

This corresponds to a vertical adjustment of $A^>(\omega)$. The moments of $A'^>(\omega)$ can be determined by inserting (4.3) into (3.5). We find the expression

$$\mu_n'^> = \int_{-1}^{1} d\omega \; T_n(\omega) A^>(\omega) - \int_{-1}^{1} d\omega \; T_n(\omega) A^>(0) \; \Theta(\omega - b), \tag{4.4}$$

leading to

$$\boxed{\mu_n'^> = \mu_n^> - A^>(0)\mu_n^\Theta} \tag{4.5}$$

The coefficients μ_n^Θ are the moments of the step function, given by

$$\mu_n^\Theta = \int_b^1 d\omega \; \arccos(n \cos \omega), \tag{4.6}$$

which yields

$$\begin{cases} \mu_0^\Theta &= 1 - b, \\ \mu_n^\Theta &= \frac{1}{2} \left(\frac{\cos[(n+1)\arccos\omega]}{n+1} - \frac{\cos[(n-1)\arccos\omega]}{n-1} \right)\Big|_b^1, \; \forall n > 0. \end{cases} \tag{4.7}$$

As illustrated below, the moments of the adjusted spectral function $A'^>(\omega)$ do decrease exponentially. The original spectral function is recovered by shifting the reconstructed function back up.
A last problem remains to be faced, namely that the value of $A^>(0)$ in equation (4.5) is *a priori* unknown. It can however be determined recursively: We start with a random value $A_0^>(0)$ and with

equation (4.5) we calculate $\mu_n'^>$, which are then extrapolated with linear prediction. Using these predicted moments we can now reconstruct the spectral function $A^>(\omega)$, which provides us with a new value for $A^>(0)$. This means that with every linear prediction a new $A^>(0)$ is determined as well. This recursion converges monotonically and quickly [10].

In the next section we will show our first results for both these linear prediction setups by applying them to a test function. We refer to the two setups henceforth as the $b = 0$ and the $b = -1$ setup respectively.

4.2 Results for a Test Function

As a first test of the two linear prediction setups introduced in the previous section we use an artificial function, because its moments can easily be calculated by numerical integration.
Spectral functions of strongly correlated interacting systems, e.g. the SIAM with $U \neq 0$, often have a feature around $\omega = 0$ (e.g. Kondo resonance) and an additional peak for $\omega \neq 0$ (e.g. Hubbard satellites). To resemble such features in the test function, we choose a sequence of Lorentzian peaks, given by

$$L(\omega) = \frac{\eta^2}{(\omega + 2)^2 + \eta^2} + \frac{\eta^2}{\omega^2 + \eta^2} + \frac{\eta^2}{(\omega - 2)^2 + \eta^2}, \tag{4.8}$$

with $\eta = 0.05$. The positive part of this function is given by

$$L^>(\omega) = \begin{cases} 0 & \text{for } \omega < 0 \\ \frac{\eta^2}{\omega^2 + \eta^2} + \frac{\eta^2}{(\omega - 2)^2 + \eta^2} & \text{else.} \end{cases} \tag{4.9}$$

This is a good choice to imitate spectral functions in the sense that the moments of a Lorentzian show a similar decay to those of the RLM spectral function, i.e. they decay exponentially.
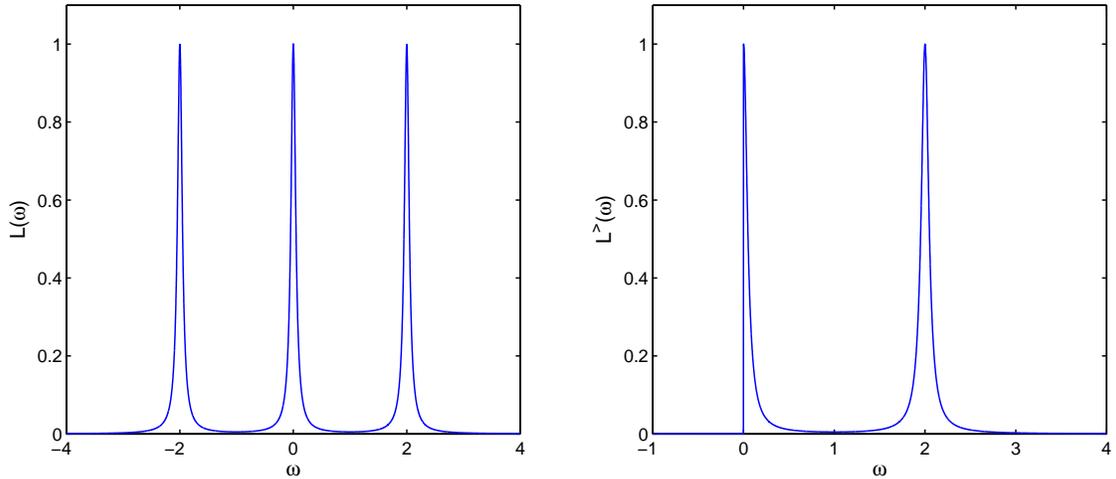
Figure 4.2: Test function defined in (4.8) and (4.9). The left panel shows the full test function $L(\omega)$, which will be used in the $b = 0$ setup. In the right panel we see the positive part of the test function $L^>(\omega)$. Due to the discontinuity at $\omega = 0$ it can be shifted very closely to $\omega = -1$.

The functions are shown in the two panels of Figure 4.2 respectively. Note that in this case we have

not yet rescaled ω to [-1,1].

An explicit example of how to rescale such functions correctly is given in Appendix C. We choose a rescaling constant $a = 100$ to rescaled the test function.

4.2.1 Linear Prediction for the Full Test Function - The $b = 0$ Setup

We want to test the $b = 0$ linear prediction algorithm by applying it to the Chebyshev moments of the full test function and compare the result to the calculated moments.

We find the Chebyshev moments of the full test function by using numerical integration. Figure 4.3 shows the curve of the calculated moments of the full spectral function (green), as well curves obtained with linear prediction (blue and purple).

The moments in the left panel are predicted from 2000 initial moments. For this high amount of initial moments the prediction leads to almost perfect results.

The right panel shows that if we reduce the number of initial moments the predicted curve differs considerably from the precise one. This can be understood by the fact that the training interval shrinks with the number of initial moments. For 1000 initial moments the training interval contains 250 moments in my setup. This does not seem to be enough to capture the precise functional decay of the moments and thus leads to errors.
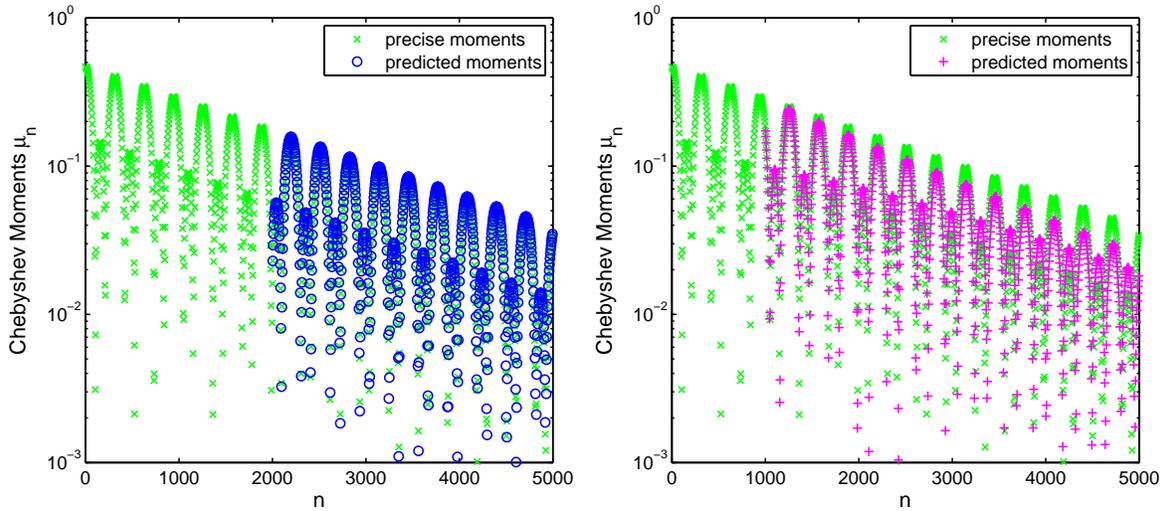
Figure 4.3: Both panels show the exactly calculated as well as the predicted Chebyshev moments of the full test function. **Left Panel**: The blue curve was obtained by applying linear prediction to 2000 initial moments. They are in good agreement with the green curve, which represents the calculated moments. **Right Panel**: The moments obtained when using linear prediction on 1000 initial moments (purple curve). For high expansion orders it deviates considerably from the exact moments.

We can use the extrapolated moments in order to reconstruct the test function. To this end we have predicted the moments up to a convergence of $\sim 10^{-5}$, which in this case correspond to 20 000 moments. The resulting reconstructed function for the case of 2000 initial moments is plotted in the left panel of Figure 4.4.

The error of the reconstruction is shown in the right panel. The blue curve is obtained by using 2000 initial moments, while purple shows the error if we start the prediction with 1000 known moments. In both cases the moments were extrapolated until they converged to $\sim 10^{-5}$. As expected the error is

much higher for a smaller number of initial moments.

Linear prediction thus provides accurate results for the test function in the old setup if we start with enough explicitly computed moments. In the next section we will introduce the $b = -1$ linear prediction setup, i.e. the prediction of the positive spectral function proposed by Wolf et al. in Ref. [10].
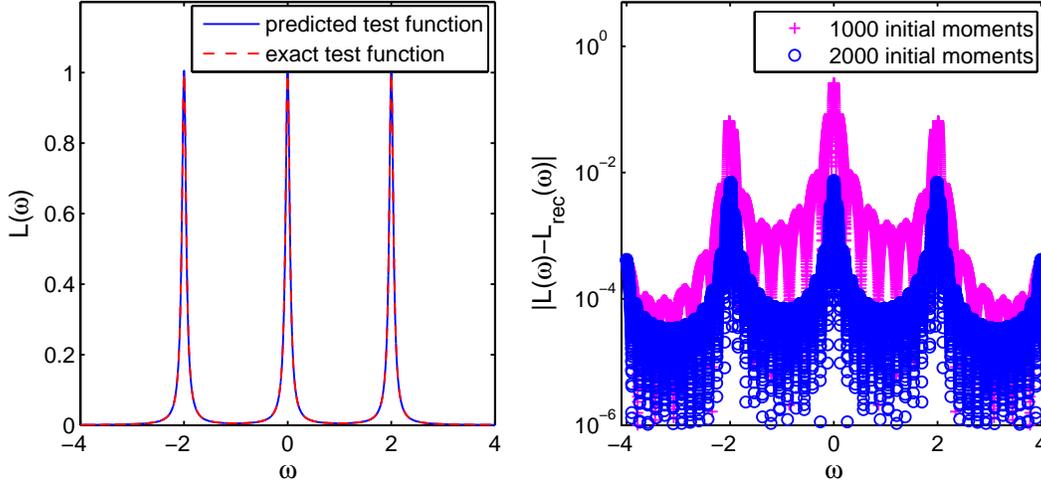
Figure 4.4: **Left Panel:** The reconstructed as well as the exact test function. The predicted function was reconstructed from an initial number of 2000 moments, that were extrapolated to 20 000. The two curves are in good agreement. **Right Panel:** The error of the reconstructed function for different numbers of initial moments. In both cases the moments were extrapolated until they converged to $\sim 10^{-5}$. The precision of the method increases with the number of initial moments.

4.2.2 Linear Prediction for the Positive Test Function - The $b = -1$ Setup

In the $b = -1$ linear prediction setup the test function is shifted to the edge of the interval [-1,1]. To employ this setup we work with the positive part of the function (see Figure 4.1). Because of the higher resolution of the Chebyshev polynomials around $\omega = \pm 1$, the Chebyshev expansion can extract more information from the function for a fixed expansion order. We therefore expect the moments of the test function to decay faster in this setup. The left panel of Figure 4.5 shows the exact moments obtained by numerical integration in different setups. The blue curve again shows the Chebyshev moments of the full test function. The red curve is obtained from the positive test function by setting $b = -0.995$ for the rescaling but without shifting the the function vertically. In this case the step of the function is *not* removed which leads to an algebraic decay of the moments. Finally the black curve shows the moments in the new setup: The positive part of the test function was rescaled with $b = -0.995$ and shifted vertically in order to remove the discontinuity.

As expected, the moments in the $b = -1$ setup show a much faster decay than the moments of the full spectral function. However, around $n = 3500$ the black curve begins to flatten, deviating from the initial exponential devolution. As we will see this feature of the curve can *not* be captured by linear prediction. Nevertheless we expect precise results, since the moments have already dropped under 10^{-4} at that point. In the $b = 0$ setup, this value is only reached for expansion orders around 20 000. Note again that this plot only contains moments that have been calculated directly with numerical integration.

We can apply the $b = -1$ linear prediction setup to the moments of the $b = -0.995$ *unshifted* test function (red curve). The result is shown in the right panel of Figure 4.5, where we have plotted the predicted moments which were obtained with 1000 initial moments (black curve). For $n < 3000$ the decay is almost identical to the precise moments in the left panel. Only when the precise curve begins
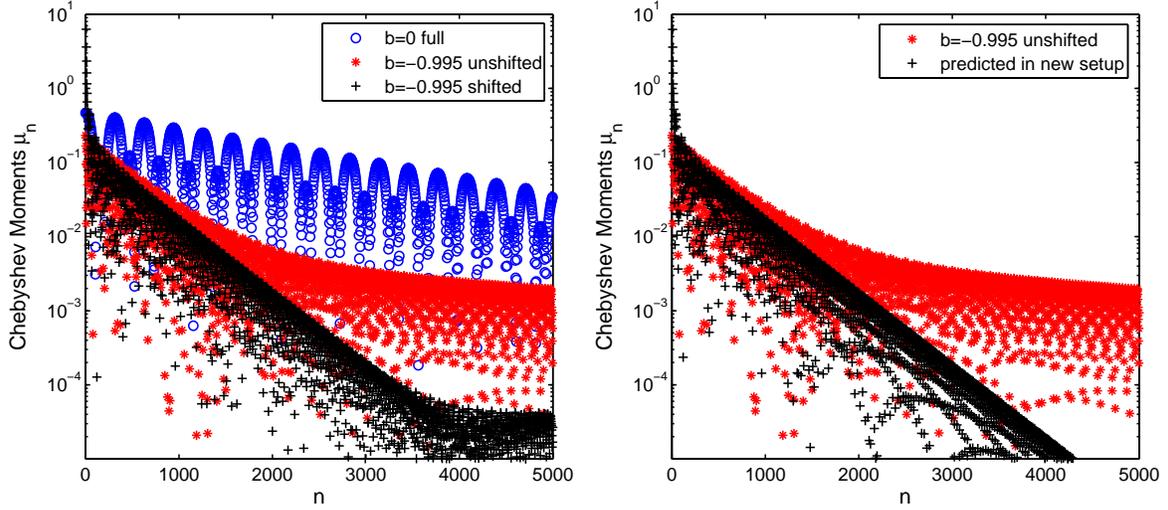
Figure 4.5: **Left Panel:** The moments of the test function in different setups. The blue curve again represents the moments of the full test function. The red curve shows the moments of the positive test function in the $b = -1$ setup. However in this case the function has *not* been shifted vertically and thus still contains a discontinuity. Therefore the decay of these moments is only algebraic. The black curve contains the moments of the positive, shifted spectral function in the $b = -1$ setup. They decay very swiftly to $\sim 10^{-4}$ until they deviate from the initial exponential decay.
Right Panel: Application of the $b = -1$ linear prediction to the moments of the positive test function (red curve). By starting with 1000 initial moments, we obtain the black curve which coincides very well with the precise moments in the right panel. Of course, we see that the deviation from the exponential decay around 3500 can not be reproduced, which is expected.

to flatten, do the predicted moments deviate considerably. Since the moments have dropped under 10^{-4} at this point we do not expect a large error.

Using the predicted moments we can reconstruct the test function and evaluate the error. In Figure 4.6 (left panel) we have plotted the initial as well as the reconstructed test function. In this setup the reconstruction shows only small errors, even though we only used 1000 initial moments. This shows the advantage of the new method: Since the moments of the shifted function decay much faster than in the $b = 0$ setup we need fewer moments to start with.

This fact is reflected in the right panel, where we have plotted the error of the reconstructed function for the $b = 0$ and the $b = -1$ setup respectively. Both curves were obtained by using 1000 initial moments. They show a large gain in precision for the new method.

As a last comparison between the two setups we have plotted the error of the reconstruction as a function of the number of initial moments, shown in Figure 4.7. In both cases, the error decreases as the number of initial moments N goes up. However, we observe a faster decrease in the case of the $b = -1$ setup, especially for small N.

From this example we can conclude that the $b = -1$ setup proposed by Wolf et al. does increase the precision of the reconstructed function. However, we want to stress that this rise in accuracy is only reached under certain conditions. By varying the parameters of the system, e.g. the width of the curve or the rescaling parameter a we found that the $b = -1$ method does not always provide better results. One should be aware that the $b = -1$ method requires a high value for the rescaling parameter a, because it needs all the contributions to the positive function to be shifted close to $\omega = -1$. If a is
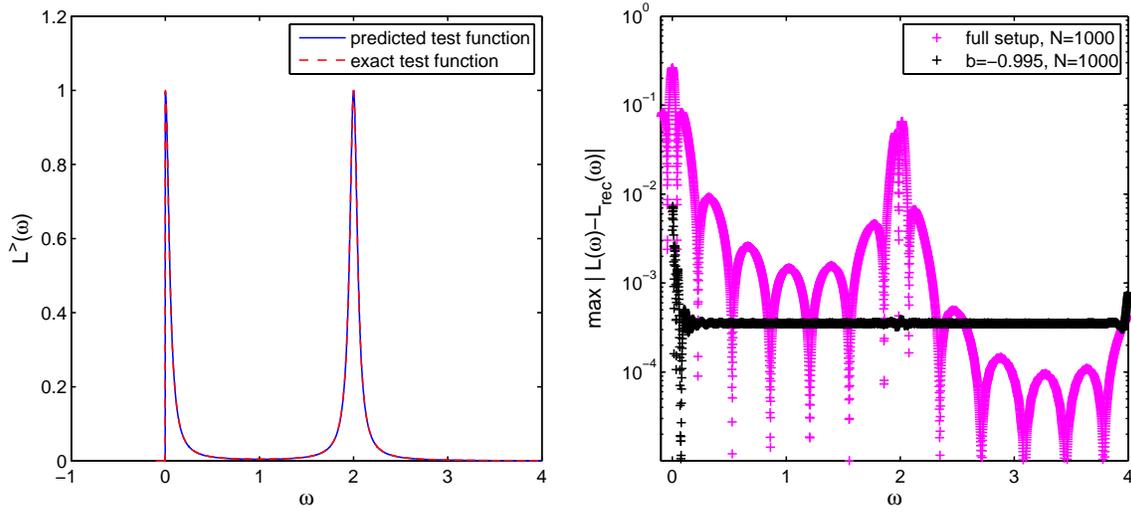
Figure 4.6: **Left Panel:** Plot of the exact and the reconstructed test function. We applied linear prediction on 1000 initial moments in this case. **Right Panel:** Comparison of the error of the reconstructed test function for the two setups. In both setups we used 1000 initial moments to start with. The $b = -1$ setup shows a maximum error of $\sim 10^{-2}$, which is considerably lower than for the $b = 0$ method $(\sim 5 \times 10^{-1})$
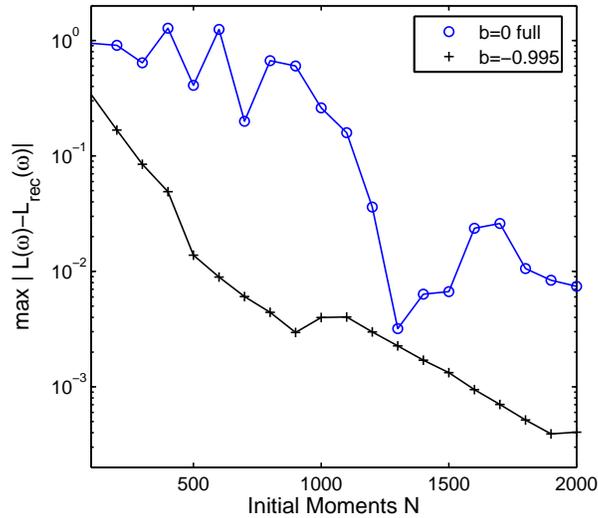
Figure 4.7: Error of the reconstructed test function for $a = 100$. In this case the $b = -1$ setup provides preciser results for any number of initial moments.

chosen too small, the high energy peak of the expanded function can not be shifted close enough to the edge of [-1,1] to fully profit from the increased resolution of the new method.

On the other hand a not only rescales the support of the function but also its width. This implies that for high values of the rescaling parameter the peaks of the function become narrower, and thus the Chebyshev moments display a slower decay.

An example of how the rescaling parameter can influence the error of the reconstructed function for
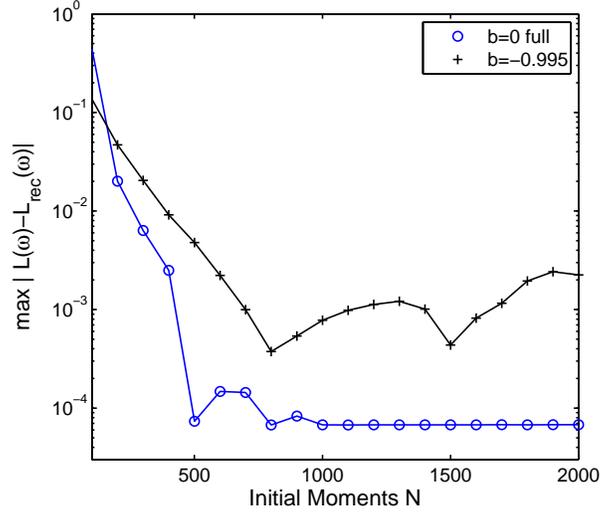
Figure 4.8: Error of the reconstructed test function with width $\eta = 0.02$ and for a rescaling parameter $a = 20$. In this case the prediction of the full spectral function provides better results, showing that the gain in precision of the new setup is not always guaranteed.

the respective setups is shown in Figure 4.8. Here we used the same test function as before, but with a width of $\eta = 0.02$. The rescaling was done with $a = 20$. For this choice of parameters, the error of the $b = 0$ setup is lower throughout the whole interval of initial moments, dropping swiftly to 10^{-4} for only $N = 500$ moments. Therefore, the gain in precision of the new setup can not be regarded as absolute, but depends on the chosen parameters.

In the next section we will have a look at some results for the resonating level model spectral function and compare the two setups for this case.

4.3 Results for the Resonating Level Model

As a first application of linear prediction to physical systems, we consider the reconstruction of the RLM spectral function. This model is well suited as a benchmark for the two setups, since it has an analytical solution. Before presenting the results, we first want to discuss the role of finite size effects, since we found that they can pose major difficulties for the calculation of Chebyshev moments with CheMPS.

4.3.1 Finite-size effects

As the name suggests, finite-size effects arise due to the fact that we approximate an infinite system -the conduction electrons- by a chain of finite length. Recalling the section about the discretization of the Hamiltonian we know that the defining parameter of this approximation is the *discretization parameter* Λ. Two different finite size effects can emerge due to the discretization:

1. Choosing Λ too small compared to the length of the chain of bath sites can lead to reflections at the end of the chain. To fully understand this effect we would have to explain the method that is used to calculate the Chebyshev moments using MPS, which would go far beyond the scope of this work. We have to be content with the notion that, for the calculation of the moments, we send a signal through the chain of bath sites. This signal is damped by Λ and thus becomes

smaller as it is propagating through the chain. If Λ is not large enough however, the signal has not been damped sufficiently when it reached the end of the bath chain. In this case the reflections of the signal at the end of the chain can not be ignored and interfere with the incoming signal. This causes fluctuation in the calculated Chebyshev moments, which can not be handled by linear prediction.

2. For large Λ, finite-size effects manifest themselves because the δ-peaks of the discrete bath spectral function are resolved at the edge of the bandwidth. To understand this effect we have to recall what we did when we discretized the Hamiltonian.
We assumed a flat bath- spectral density $\rho(\omega) = \sum_{\mathbf{k}} \delta(\omega - \epsilon_{\mathbf{k}})$, which is transformed to $\rho(\omega) = \sum_n \delta(\omega - E_n)$, with $E_n = \pm\Lambda^{-n}$ by the discretization. For large Λ this means that the first few delta peaks are quite far apart. If we reach high expansion orders the Chebyshev expansion will therefore inevitably start to resolve these individual delta peaks, which leads to deviations in the smoothly decaying moments.
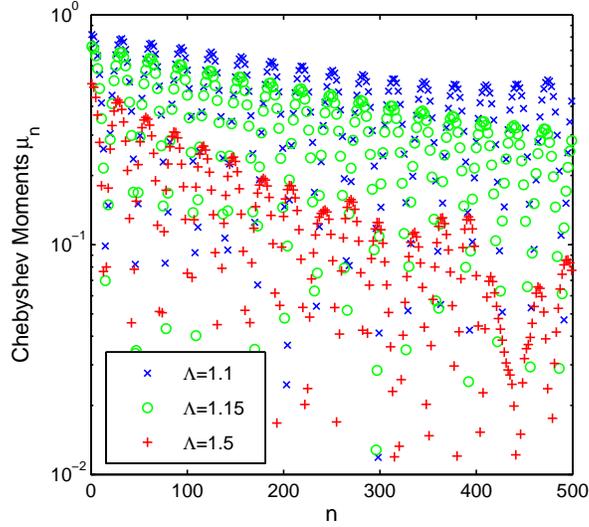
Figure 4.9: The moments of the positive spectral function of the RLM for different Λ. The chain length is set to 51 sites. The plot shows the two aforementioned finite size effects.

The presence of either of these two effects makes linear prediction hardly applicable. How they manifest themselves in the calculated Chebyshev moments is shown in Figure 4.9. The three curves represent the moments of a chain with 51 sites for different choices of the discretization parameter Λ.
The green curve consists of the moments for $\Lambda = 1.15$. It shows the expected exponential decay, meaning that it contains no finite size effects. The blue curve ($\Lambda = 1.1$) shows what happens if Λ is chosen too small for the chain. At first it decays exponentially, but for large n the moments start to rise again, showing an oscillatory pattern. Since this effect is caused by reflections at the end of the chain it can be avoided by simply choosing the chain to be longer. Indeed for $\Lambda = 1.1$ and a chain of 151 sites the moments do not show any sign of these oscillations and can be extrapolated with linear prediction. For the red curve we chose $\Lambda = 1.5$. We see that for low expansion orders the decay of the moments is exponential as well. This can be explained by the fact that for small n the resolution of the Chebyshev functions is not yet high enough to capture the delta peaks of the bath spectral density. For larger n however the trend of the curve deviates strongly from an exponential decay, as we can see in the figure. This effect can not be compensated for by adapting the chain length or any other parameter. A large value for Λ implies that the energy of the bath sites diminishes rapidly, meaning that the last sites of the chain are of negligible energy. Adding more sites in this case hardly changes the bandwidth at all.

We also want to mention that the decay of the Chebyshev moments is influenced by the discretiztion parameter Λ. The latter determines the bandwidth W of the system and thus the rescaling that has to be used for CheMPS. In the case of a small Λ ($\Lambda \to 1$) the bandwidth is larger than for large Λ, which means that we have to rescale a larger interval to [-1,1]. Thus for small Λ the moments decay slower. Of course it is favourable to work with long chains and small discretization parameters, since we want to describe a system in the thermodynamical limes. However we will see in the next section that linear prediction favours a faster decay of the moments and is thus more precise for small systems. Also the numerical cost of CheMPS grows quickly for larger systems.

4.3.2 The $b = 0$ Setup

The $b = 0$ setup is tested on the full RLM spectral function. We chose a system of 151 bath sites with the discretization parameter $\Lambda = 1.1$. The width of the spectral function is chosen to be $\Gamma = 0.005$. The moments of the full spectral function are plotted in Figure 4.10. It is important to note that in these plots the x-axis scales logarithmically.
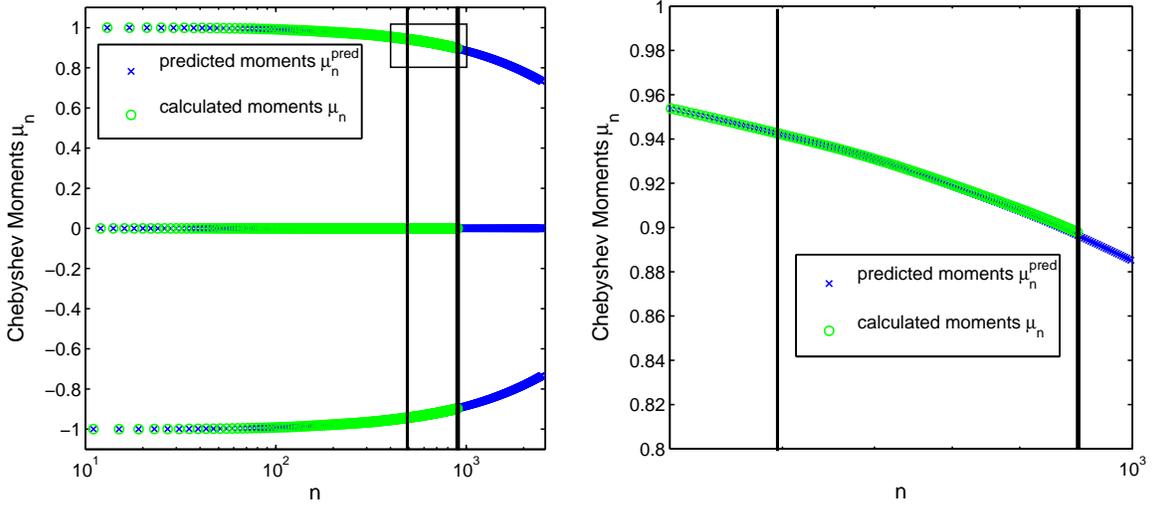
Figure 4.10: **Left Panel:** Moments of the RLM for a system with 151 bath sites and a discretization parameter $\Lambda = 1.1$. The green curve represents the Chebyshev moments calculated with CheMPS. The moments obtained with linear prediction make up the blue curve. The two vertical black lines mark the interval where calculated and predicted moments are plotted simultaneously. **Right Panel:** Zoom into the boxed section of the left panel. Between the two black lines we can compare the CheMPS and the predicted moments. They are in good agreement.

The green curve shows 900 moments obtained with CheMPS. We used the first 500 of them to predict 2000 moments, which make up the blue curve. In the interval $500 \leq n \leq 900$ marked by the vertical black lines the CheMPS and predicted moments are displayed showing little to no discrepancy. The boxed section of the left panel is shown in Figure the right panel.

The decay of the moments is exponential, which is why linear prediction gives quite accurate results. However, we can observe in the right panel that the predicted moments do not coincide *perfectly* with the others. Similar to the case of the test function, 500 moments are not enough to capture the functional decay of the moments. We therefore expect a visible error of the reconstructed spectral function.
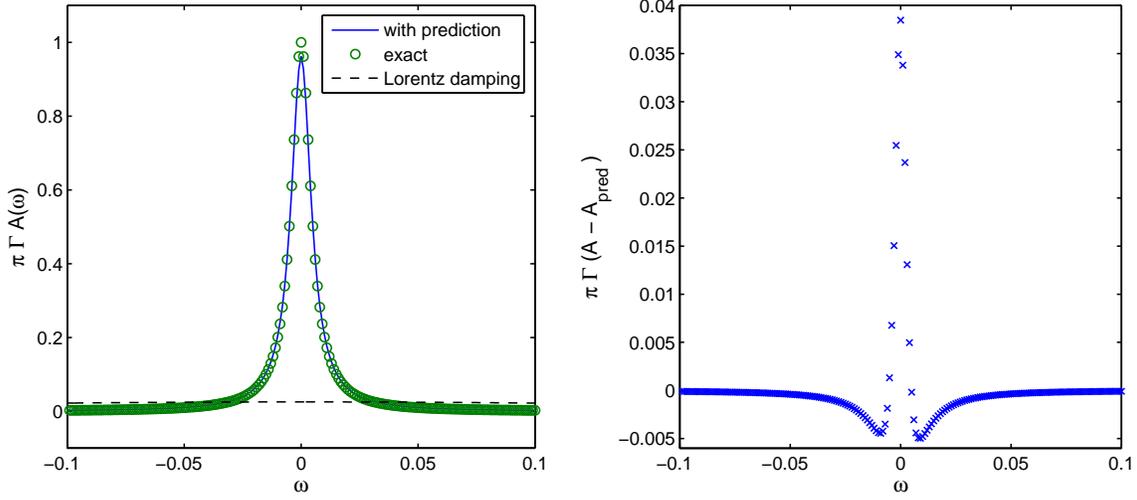
Figure 4.11: **Left Panel:** Plot of the reconstructed spectral function and comparison to the exact curve. To show the extreme improvement of linear prediction compared to the polynomial kernel method I have also plotted the spectral function obtained with 500 moments and a Lorentz kernel (black curve). **Right Panel:** The error of the RLM spectral function for a system of 151 sites and a discretization parameter $\Lambda = 1.1$. The spectral density was reconstructed with 500 initial moments.

In order to reconstruct the spectral function we extrapolate the Chebyshev moments until they converged to $\sim 10^{-5}$, which corresponds to about 50 000 moments. We have plotted the reconstructed spectral density in Figure 4.11 (left panel), together with the exact curve and a curve obtained with a Lorentz kernel. All the curves have been rescaled with $\pi\Gamma$. The predicted and exact function coincide quite well in this case, although we observe a notable discrepancy around the maximum.

The error of the rescaled function is plotted in the right panel. As expected its maximum lies around the peak of the spectral function itself. If we compare the value of the error to the results obtained by the authors of Ref. [9] we realize that our error is quite high. As we mentioned in the previous section, the decay of the moments is strongly influenced by the rescaling parameter a. In our code this rescaling is always identical to the bandwidth W, which in this case is $W = 41.9649$. Comparing this to the value $a = 5$ chosen in Ref. [9] we understand the difference in accuracy.

To illustrate how the accuracy of the reconstruction can be improved by choosing different CheMPS parameters, we calculated the moments of a systen with 25 sites and a $\Lambda = 1.7$. The bandwidth of this system is $W = 7.5599$, and thus we expect the moments to decay faster. Since the discretization parameter is quite large we have to be wary to avoid finite-size effects, which would occur for large expansion orders. For this reason we calculated only a mere 100 moments.

The left panel of Figure 4.12 shows the moments of this system. The blue curve is obtained by applying linear prediction to the 100 initial moments, represented by the green circles. Comparing this plot to Figure 4.10 we can see that the moments decay much faster. This is also reflected in the error of the reconstructed spectral function, plotted in the right panel. Although we have only used 100 initial moments as opposed to the 500 in the previous example, the reconstruction is more precise in this case.
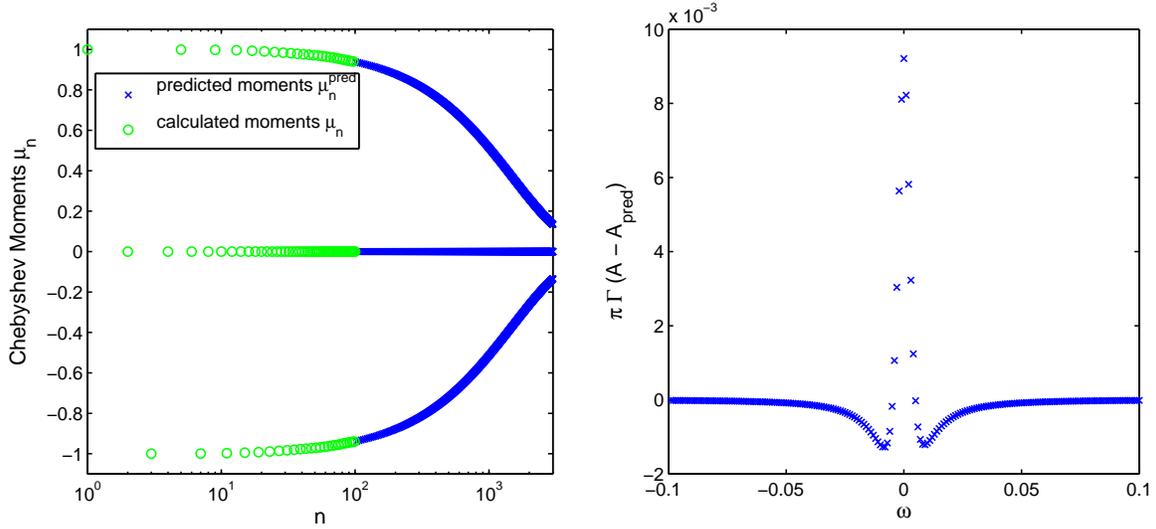
Figure 4.12: **Left Panel:** Moments of the RLM for a system of 25 sits and $\Lambda = 1.7$. This plot shows that the moments decay much faster for higher values of Λ in this setup, leading to an increase in precision of the reconstructed spectral function. **Right Panel:** Error of the reconstructed spectral function for a system of 25 sites and $\Lambda = 1.7$. I started with 100 initial moments and extrapolated them until they converged to 10^{-5}. Although I used fewer initial moments than for the plot in Figure 4.18 the error is smaller.

4.3.3 The $b = -1$ Setup

In this section we apply the $b = -1$ linear prediction setup to the RLM. In this case we only treat the positive part of the spectral function. In order to compare the results to the $b = 0$ setup, we used the same system, i.e. a chain with 151 sites and a discretization parameter $\Lambda = 1.1$. We calculated the first 500 moments of the system with CheMPS and used the new linear prediction to extrapolate the moments to higher orders.

Figure 4.13 shows the first 500 calculated as well as the predicted moments. Unlike in the $b = 0$ setup the predicted and the initial moments do *not* coincide, since the new linear prediction shifts the moments in order to remove the discontinuity in the spectral function. The blue curve, which consists of the predicted moments displays a fast exponential decay for $n > 300$.

The reconstructed spectral function is shown in the left panel of Figure 4.14. Similar to the $b = 0$ setup the error, plotted in the right panel has its maximum around the peak of the spectral function, which is expected. It is surprising however that in this particular case the $b = -1$ setup leads to higher errors than the $b = 0$ method. Ref. [10] mentions that linear prediction consistently overestimates the spectral function. A combination of linear prediction and the kernel polynomial method could therefore lead to more precise results. There is a large amount of different parameters which are involved in the calculation of the moments and the linear prediction itself. This makes it difficult to determine the source of this discrepancy.

We showed in the previous section that there are systems where the $b = 0$ setup leads to smaller errors, which has been confirmed by the calculation in the RLM.

Our conclusion to these calculations is that both linear prediction setups lead to a drastic increase in precision compared to the kernel polynomial method. For the considered system of 151 bath sites and $\Lambda = 1.1$ both methods provide accurate spectral functions, with errors around $\sim (10^{-2} - 10^{-1})$. Surprisingly the $b = 0$ linear prediction method seems to provide more accurate results in the considered
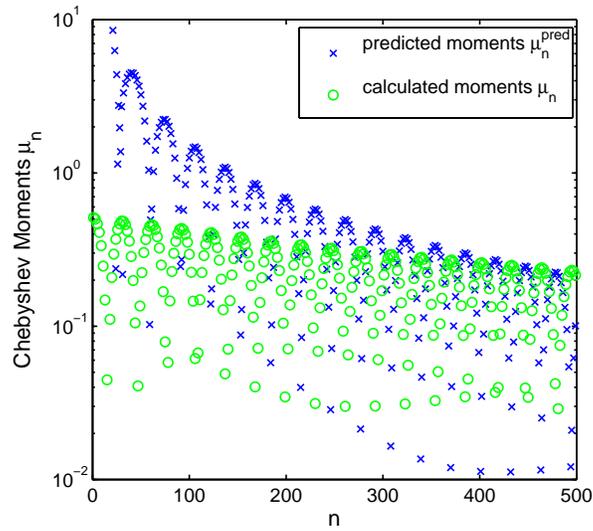
Figure 4.13: The moments of the positive RLM spectral function. The system consists of a chain of 151 sites with a discretization $\Lambda = 1.1$. The green curve shows the calculated moments. The $b = -1$ linear prediction shifts these moments in order to remove the step in the spectral function, which yields the blue curve.
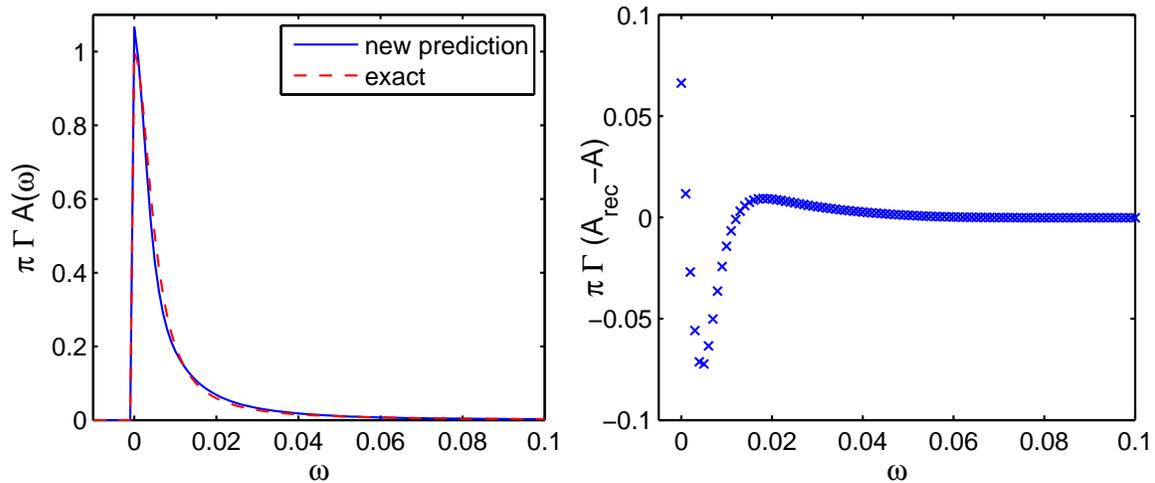
Figure 4.14: **Left Panel:** The reconstructed as well as the exact RLM spectral function for a system of 151 bath sites and $\Lambda = 1.1$. **Right Panel:** The error of the RLM spectral predicted function in the $b = -1$ setup. The error is roughly two times larger than for the $b = 0$ setup.

RLM system.

In the next section we are going to have a short look at the results for the interacting SIAM.

4.4 Results for the SIAM

Up until this point we have only considered non-interacting systems. In this case the energy of the electrons on the impurity atom is simply given by the fermi energy, which we have set to zero. For this reason we observe the peak of the RLM spectral density around $\omega = 0$.

In the case of finite onsite interaction $U \neq 0$ the electrons on the impurity do feel one another's presence due to the coulomb interaction. This has an impact on the energy of the electrons, and thus on the form of the spectral function. We therefore expect to find features for energies $\omega \sim U/2$ in the spectral density. These peaks at finite energies are called *Hubbard satellites*. In addition to these peaks the spectral function always has a feature at $\omega = 0$ (Kondo resonance), the height of which is determined by the hybridization Γ. For any SIAM spectral function the relation $\pi \Gamma A(0) = 1$ is fulfilled, which provides us with a good method to test the quality of our numerical results.

All the presented spectral densities in this chapter correspond to a system with 125 sites, $\Lambda = 1.1$ and a *flat* bath spectral density.
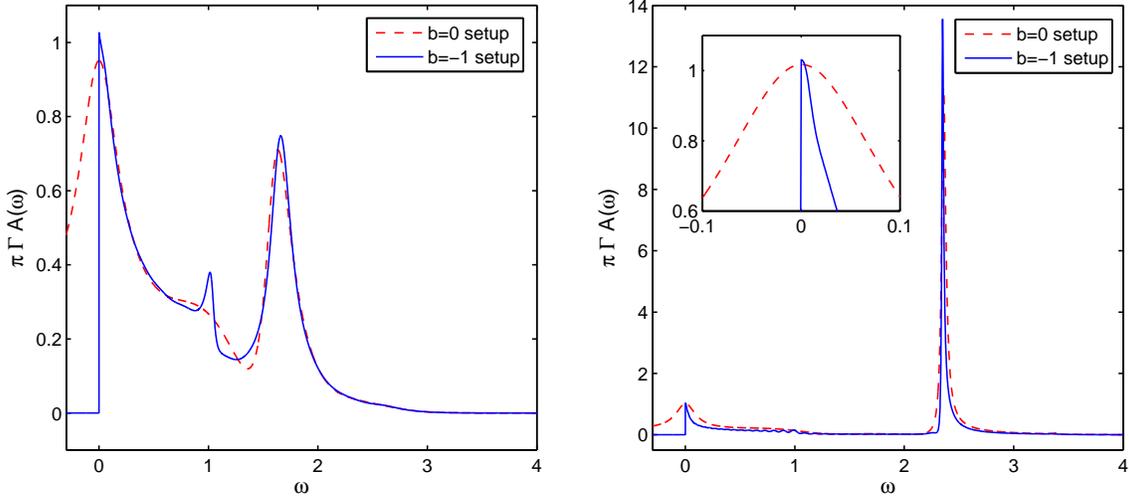
Figure 4.15: **Left Panel:** The SIAM spectral function for $\Gamma = 0.5$ and an onsite interaction $U/\Gamma = 4$. The most obvious difference between the two setups is the small peak that is only shown by the $b = -1$ curve. The origin of this feature may be physical or a numerical inaccuracy. Due to a lack of time, we unfortunately could not make further investigations. The curve of the $b = 0$ setup was obtained from 500 initial moments. For the $b = -1$ setup we started with 150 initial moments. **Right Panel:** The SIAM spectral function for $\Gamma = 0.5$ and an onsite interaction $U/\Gamma = 8$. This plot shows the very prominent Hubbard satellite at high energies. Here it is evident that the $b = -1$ setup provides more precise results, with the central peak being narrower and the Hubbard satellite higher. The $b = -1$ result was obtained by extrapolating 200, the $b = 0$ result for 500 initial moments.

The left panel of Figure 4.15 displays the spectral function of a system with $\Gamma = 0.5$ and an onsite interaction $U/\Gamma = 4$. The red curve was obtained by applying the $b = 0$ setup, whereas the blue curve is the result of the $b = -1$ method. First of all we note that for both setups the sum rule $\pi \Gamma A(0) = 1$ is satisfied. Both curves show roughly the same features, except for the small peak displayed by the $b = -1$ function around $\omega = 1$. Wether it is obtained due to numerical inaccuracies or has a physical origin is unclear. Unfortunately, due to time pressure we could not further investigate this feature.

In the right panel of Figure 4.15 the spectral function is plotted for $\Gamma = 0.5$ and $U/\Gamma = 8$. In both setups the Hubbard satellite is very pronounced and clearly visible. The $b = -1$ setup provides better

results for this system, which is proved by the height of the Hubbard satellite, as well as the fact that the Kondo resonance is narrower. Also for this system the sum rule is satisfied quite well.

A comparison of both panels shows that the position, as well as the height and width of the Hubbard satellite is dependent on the ratio U/Γ. As the onsite energy grows larger the peak moves to higher energies. We also have to take into account fact that of the bath density of states, which only reaches to $\omega = -1$ has an effect on the shape of the peak.
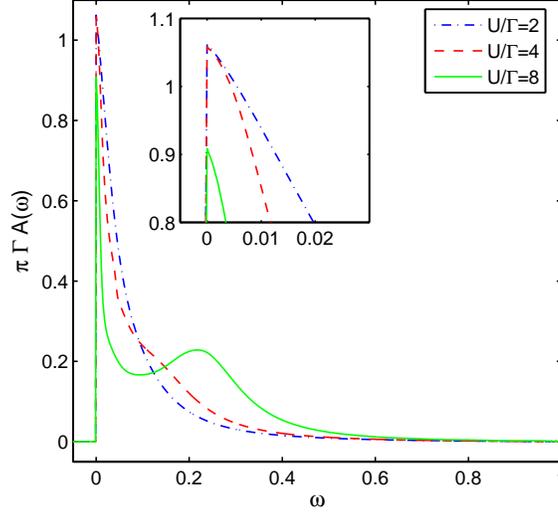
Figure 4.16: SIAM spectral functions for $\Gamma = 0.05$ and three different interaction strength. The three curves were obtained by starting with 500 initial moments and predicting them in the $b = -1$ setup.

Figure 4.16 shows a spectral function of a system with $\Gamma = 0.05$ and different interaction strengths. In all the cases the interaction is small compared to the maximum bath energy. This has a visible effect on the shape of the spectral function. The Hubbard satellites are less pronounced and the Kondo resonance sharper for this set of parameters. Although the sum rule is roughly satisfied in this case as well, we note that for large U/Γ the height of the peak decreases. This is due to the fact that the central peak becomes very sharp; thus the Chebyshev moments decay slower and linear prediction becomes less precise. For even larger onsite interactions we therefore expect linear prediction to provide higher errors. All the spectral functions in Figure 4.16 were calculated in the $b = -1$ setup.

We conclude that linear prediction provides accurate results for the interacting SIAM. For all the plotted spectral functions the sum rule holds quite well. In general we can confirm with these results that the $b = -1$ setup does lead to preciser results for SIAM spectral functions, compared to the $b = 0$ method. If the small peak at $\omega = 1$ for the $\Gamma = 0.5$, $U/\Gamma = 4$ spectral function is of physical origin, this is a strong argument for the $b = -1$ setup.

Chapter 5

Conclusion

We started by giving a short summary of the numerical methods that are currently used to calculate spectral functions of impurity models. The need for an alternative to NRG was motivated by the its low resolution for high energy scales.

As an introduction to the physical properties of quantum impurity systems, the SIAM Hamiltonian was presented and explained in great detail in Chapter 2. We then moved on to the treatment of Green's- and spectral functions, which play key roles in the analysis of models in condensed matter physics. As an example of zero-temperature spectral functions we introduced the RLM spectral function, which we later used as a benchmark for linear prediction.

In Chapter 3 we treated the numerical methods that are necessary to calculate spectral functions with CheMPS. We started by introducing the Chebyshev polynomials of first kind and discussing their properties. Since they only form an orthonormal basis in the interval [-1,1], we had to rescale the Hamiltonian to this interval. We then moved on to the concept of linear prediction and the ideal choice of parameters. Finally, we showed how to discretize the SIAM Hamiltonian in order to make it treatable with numerical methods.

Chapter 4 contained the practical part of the work. We started by explaining the idea of linear prediction in the context of CheMPS and introduced the $b = 0$ and $b = -1$ setup respectively. We started the calculations by applying the two linear prediction setups to a test function. The first result showed that the number of initial moments is a crucial factor for the precision of the reconstructed function. For the function that we considered the $b = -1$ setup leads to more accurate results for any number of initial moments. Nevertheless, there are cases where the $b = 0$ method seems to be more reliable. In general, this example taught us that CheMPS in combination with linear prediction is very dependent on the system parameters. We can get a wide range of accuracies, depending on how well we choose the initial parameters. We have also seen that CheMPS is quite susceptible to finite-size effects. In fact, there is only a small margin for the discretization parameter where the method leads to reliable results. As we have seen, for large Λ only a limited amount of moments can be computed until we start to resolve the δ-peaks of the bath spectral function. For small Λ we have to work with long chains to avoid reflections at the end of the chain. This again increases the computational effort. Last but not least, we have found that the results for finite a onsite interaction $U \neq 0$ were very promising, especially for the $b = -1$ setup. For all the treated interacting systems we obtained a higher accuracy for less initial moments. Thus, we can conclude that, if carefully implemented, linear prediction increases the precision of CheMPS-computed spectral functions enormously. The resolution could be enhanced even further by combining linear prediction with the kernel polynomial method to compensate for the systematic overestimation of the height of the central peak. This makes CheMPS a viable option to calculate zero-temperature spectral functions of quantum impurity models.

Appendices

Appendix A

Green's Function of Free Particles

To show how the rather abstract concepts we have introduced in Chapter 2, i.e. the Hamiltonian in second quantization and the various Green's functions, are used in a concrete case let us calculate the Green's function of free particles. Remembering the expression we have already used for the Hamiltonian of the SIAM, we can write down the energy of free fermions:

$$H_0 = \sum_{\sigma\mathbf{k}} \epsilon_\mathbf{k} c_{\sigma\mathbf{k}}^\dagger c_{\sigma\mathbf{k}}. \tag{A.1}$$

Since this Hamiltonian is diagonal in $c_{\sigma\mathbf{k}}^\dagger c_{\sigma\mathbf{k}}$ so will be the Green's functions, so we can formally write down $G_0^>(\sigma\mathbf{k}; t, t')$ and $G_0^<(\sigma\mathbf{k}; t, t')$:

$$G_0^>(\sigma\mathbf{k}; t, t') = -i\langle c_{\sigma\mathbf{k}}(t) c_{\sigma\mathbf{k}}^\dagger(t') \rangle \qquad\qquad G_0^<(\sigma\mathbf{k}; t, t') = i\langle c_{\sigma\mathbf{k}}^\dagger(t') c_{\sigma\mathbf{k}}(t) \rangle. \tag{A.2}$$

The equations (2.14) can be solved by extracting the time dependency from the operators. In order to do that, let us explicitly write the time dependency of $c_{\sigma\mathbf{k}}^\dagger$ and $c_{\sigma\mathbf{k}}$:

$$c_{\sigma\mathbf{k}}^{(\dagger)}(t) = e^{iHt} c_{\sigma\mathbf{k}}^{(\dagger)} e^{-iHt}. \tag{A.3}$$

Taking the time derivative of this expression and using the operator algebra (2.5) leads us to a differential equation that can be integrated and yields:

$$c_{\sigma\mathbf{k}}^\dagger(t) = e^{i\epsilon_\mathbf{k} t} c_{\sigma\mathbf{k}}^\dagger \qquad\qquad c_{\sigma\mathbf{k}}(t) = e^{-i\epsilon_\mathbf{k} t} c_{\sigma\mathbf{k}}. \tag{A.4}$$

Thus, we have found an expression for the time dependency of our operators which we can reinsert into (2.14). Remembering the expression (2.6) for the number operator we find:

$$G_0^>(\sigma\mathbf{k}; t, t') = -i e^{-i\epsilon_\mathbf{k}(t-t')} (1 - \langle n_{\sigma\mathbf{k}} \rangle) \tag{A.5}$$

$$G_0^<(\sigma\mathbf{k}; t, t') = i e^{-i\epsilon_\mathbf{k}(t-t')} \langle n_{\sigma\mathbf{k}} \rangle. \tag{A.6}$$

Combining these two equations and using the relation (2.13) we can now write down an explicit expression for the retarded Green's function:

$$G^R(\sigma\mathbf{k}; t, t') = -i\theta(t - t') e^{-i\epsilon_\mathbf{k}(t-t')}. \tag{A.7}$$

In the time domain, the Green's function has the form of a plane wave. The energy of the particle is constant, which makes perfect sense, since there are no interactions involved and the Green's function was diagonal in $c^\dagger c$ from the start. We can interpret equation (A.5) by identifying that $G_0^>(\sigma\mathbf{k}; t, t')$ needs an empty state and therefore describes the propagation of electrons, whereas (A.6) is proportional to the to the electron number and consequently $G_0^<(\sigma\mathbf{k}; t, t')$ gives the propagation of holes.

We see that all the propagators we have found in our example only depend on the difference $t - t'$ and not on t and t' individually. This temporal translation invariance implies that it makes sense to represent the Green's function in the frequency domain, which can be achieved by Fourier transformation:

$$G(\omega) = \int_{-\infty}^{+\infty} \mathrm{d}t e^{i(\omega + i\eta)(t - t')} G(t - t'). \tag{A.8}$$

The imaginary part η of the frequency is an infinitesimal positiv number, needed to make the integral convergent. If we insert the expression (A.7) for the retarded Green's function into (A.8), we find the simple expression

$$G_0^R(\sigma \mathbf{k}; \omega) = \frac{1}{\omega - \epsilon_{\mathbf{k}} + i\eta}. \tag{A.9}$$

A.1 Equation of motion Ansatz

The same result can also be obtained via the equation of motion theory. We find the equation of motion for the Green's function in momentum representation by applying $i\partial_t$ to it:

$$i\partial_t G_0^R(\mathbf{k}\sigma t, \mathbf{k}'\sigma't') = \partial_t \left(\theta(t - t')\right) \langle \{c_{\mathbf{k}\sigma}(t), c_{\mathbf{k}'\sigma'}^\dagger(t')\}\rangle - i\theta(t - t')\langle\{i\partial_t c_{\mathbf{k}\sigma}(t), c_{\mathbf{k}'\sigma'}^\dagger(t')\}\rangle \tag{A.10}$$

The first of these two terms can be reduced to $\delta(t - t')\delta_{\mathbf{k}\sigma, \mathbf{k}'\sigma'}$, using the algebra of creation and annihilation operators (2.5) and the fact that the derivative of the step function is a delta function. The second term contains the time derivative of an operator, which can be written as $i\partial_t c_{\sigma\mathbf{k}}(t) = -[H, c_{\sigma\mathbf{k}}](t)$ using the Hamilton equations of motion. For the free particle case this commutator can be easily calculated, since the Hamiltonian H_0 is diagonal:

$$-[H, c_{\mathbf{k}}](t) = -\sum_{\sigma\mathbf{k}''} \epsilon_{\mathbf{k}''}[c_{\sigma''\mathbf{k}''}^\dagger c_{\sigma''\mathbf{k}''}, c_{\mathbf{k}\sigma}](t) \tag{A.11}$$

Using the relation $[AB, C] = A\{B, C\} - \{C, A\}B$ as well as the equations (2.5) this can be rewritten as

$$-[H, c_{\mathbf{k}}](t) = -\sum_{\sigma''\mathbf{k}''} \epsilon_{\mathbf{k}''}\delta_{\sigma''\mathbf{k}'', \sigma\mathbf{k}} c_{\sigma''\mathbf{k}''} = -\epsilon_{\mathbf{k}} c_{\sigma\mathbf{k}} \tag{A.12}$$

and inserting this back into (A.10) yields

$$(i\partial_t - \epsilon_{\mathbf{k}})G_0^R(\mathbf{k}\sigma t, \mathbf{k}'\sigma't') = \delta(t - t')\delta_{\mathbf{k}\sigma, \mathbf{k}'\sigma'} \tag{A.13}$$

At this point it becomes clear why we are calling G^R a Green's Function, since (A.13) has the same form as (2.11). As is the usual procedure for solving equations of this form, let us consider the Fourier transform $G^R(\mathbf{k}\sigma, \mathbf{k}'\sigma'; \omega)$. This Ansatz makes sense because the right side of the equation indicates that G^R only depends on the difference $t - t'$, and not on t and t' seperately. Hence we can write the Fourier transform of G^R as:

$$G_0^R(t - t') = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} e^{-i(\omega + i\eta)(t - t')} G^R(\omega) \tag{A.14}$$

Inserting (A.11) into (A.12) yields:

$$G_0^R(\mathbf{k}\sigma, \mathbf{k}'\sigma'; \omega) = \frac{1}{\omega - \epsilon_{\mathbf{k}} + i\eta}\delta_{\mathbf{k}\sigma, \mathbf{k}'\sigma'} \tag{A.15}$$

η is an infinitesimally small positive number. It is introduced in order to make the Fourier integral convergent.

In quantum field theory Green's functions $G(\nu, \nu')$ are also called propagators, because they give the probability amplitude for a particle to evolve from the state ν to the state ν'. If ν and ν' are spatial coordinates, the term 'propagation' becomes clear.

This physical interpretation of the Green's function makes sense in our case as well: Since we are dealing with a system of free, non-interacting particles, the momentum of a single particlee remains unchanged and hence we have the Kronecker Delta on the right side of (A.13).

Appendix B

Minimizing the mean squared error

The goal of this section is to derive the system of linear equations (3.17), starting with the idea that the mean squared error needs to be minimal. In order to do this efficiently we are going to use a compact notation for the mean squared error and explain the connection to the expressions (3.15) at the end. We start with our previous ansatz, namely that any value in our set of data x_i can be expressed as a linear combination $\tilde{x}$ of the m previous values:

$$\tilde{x} = -a_1^* x_1 - a_2^* x_2 - ... - a_m^* x_m. \tag{B.1}$$

The estimated error of this reconstructed value is given by

$$e = x - \tilde{x} = x + a_1^* x_1 + a_2^* x_2 + ... + a_m^* x_m \tag{B.2}$$

and in our compact notation we express the mean squared error by $E[|e|^2]$. As we will see in the course of this section, this object $E[\ ]$ essentially has the attributes of a scalar product. What is important to keep in mind is that x is fixed value and that the only variables in these calculation are the coefficients a_i.

So which estimated error e will minimize the mean squared error? If we think geometrically, we can answer this question without any calculations. Suppose we want to approximate a vector $\vec{s}$ in three dimensions with a vector $\vec{s'}$ which is confined to the x-y-plane. The best approximation will be given by the vector $\vec{s'}$ that points into the same x-y direction, and at the same time has the same length as the x-y component of $\vec{s}$. These two requirements can be combined by stating that the error $\vec{e} = \vec{s} - \vec{s'}$ has to be orthogonal to $\vec{s'}$, or $\langle \vec{e}, \vec{s'} \rangle = 0$. As we have already mentioned, the mean squared error can be interpreted as a scalar product, and we can generalize the vectorial case to any type of data. The result of this generalization is the *Orthogonality Principle*:

The estimate $\tilde{x}$ results in a minimal mean squared error if and only if the estimation error is orthogonal to the m variables x_i:

$$E[e x_i^*] = 0, 1 \leq i \leq m. \tag{B.3}$$

A formal proof of this theorem can be found in [11].

Now that we have found this condition we have to use it to derive an expression for the coefficients a_i. To condense our notation even further, let us define the two vectors $\mathbf{a} = (a_1, a_2, ..., a_m)^T$ and $\mathbf{x} = (x_1, x_2, ..., x_m)^T$, which leads us to the compact expression

$$e = x + \mathbf{a}^\dagger \mathbf{x} \tag{B.4}$$

for the error. The orthogonality condition can now be written as

$$E[x \mathbf{x}^\dagger] + \mathbf{a}^\dagger E[\mathbf{x} \mathbf{x}^\dagger] = 0. \tag{B.5}$$

In a way this is already the equation we were looking for, we only have to make the following definitions:

$$R = E[\mathbf{x}\mathbf{x}^{\dagger}] \qquad \text{and} \qquad \mathbf{r} = E[\mathbf{x}x^{*}]. \qquad (B.6)$$

As can be seen from the definition R is a hermitian $m \times m$ Matrix and $\mathbf{r}$ is an m dimensional vector. If we insert these definitions into (B.5), we obtain the equation

$$R\mathbf{a} = -\mathbf{r}, \qquad (B.7)$$

which is the relation we want to find.

Keeping in mind that (B.1) holds for *any* value of the data set, we find the expressions (3.17) for R and $\mathbf{r}$.

Appendix C

Rescaling of the Lorentzian

As an example I will show the rescaling explicitly for a general Lorentzian, since it is directly transferable to any other function. We have

$$f(\omega) = \frac{\eta}{(\omega - \omega_0)^2 + \eta^2} \tag{C.1}$$

with $\text{supp}(f) = [m, n] \nsubseteq [-1, 1]$. In order to expand this function in Chebyshev polynomials, we have to rescale it, using the linear transformation:

$$\omega \to x = \frac{\omega}{a} + b, \tag{C.2}$$

such that $x(m) = -1$ and $x(n) = 1$. Without changing the function f we can generally write

$$f(\omega) = \frac{\eta}{a^2 \left(\frac{\omega}{a} + b - \frac{\omega_0}{a} - b\right)^2 + \eta^2}, \tag{C.3}$$

where we have inserted two identities. The Chebyshev moments are then calculated as follows:

$$\mu_n = \int_m^n \mathrm{d}\omega \, f(\omega) T\left(\frac{\omega}{a} + b\right) \tag{C.4}$$

$$= \int_m^n \mathrm{d}\omega \, \frac{\eta}{a^2 \left(\frac{\omega}{a} + b - \frac{\omega_0}{a} - b\right)^2 + \eta^2} T(\frac{\omega}{a} + b). \tag{C.5}$$

We can make the substitution $x = \frac{\omega}{a} + b$, with $\mathrm{d}\omega = a \mathrm{d}x$:

$$\mu_n = \int_{-1}^1 \mathrm{d}x \, \frac{a\eta}{a^2(x - x_0)^2 + \eta^2} T(x) \tag{C.6}$$

$$= \int_{-1}^1 \mathrm{d}x \, f_{resc}(x) T(x). \tag{C.7}$$

In the last line we have introduced the rescaled function

$$f_{resc}(x) = \frac{a\eta}{a^2(x - x_0)^2 + \eta^2} = a \, f(\omega), \tag{C.8}$$

which thus leads us to the transformation rule (3.13) that we have also found for the spectral function.

Acknowledgements

Bibliography

[1] Kenneth G. Wilson. The renormalization group: Critical phenomena and the Kondo problem. *Rev. Mod. Phys.*, 47:773–840, Oct 1975.

[2] Ralf Bulla, Theo A. Costi, and Thomas Pruschke. Numerical renormalization group method for quantum impurity systems. *Rev. Mod. Phys.*, 80:395–450, Apr 2008.

[3] Karsten Flensberg Henrik Bruus. *Many-body quantum theory in condensed matter physics.* August 2002.

[4] Kenneth G. Wilson. Renormalization Group and Critical Phenomena. I. Renormalization Group and the Kadanoff Scaling Picture. *Phys. Rev. B*, 4:3174–3183, Nov 1971.

[5] Alexander Weiße, Gerhard Wellein, Andreas Alvermann, and Holger Fehske. The kernel polynomial method. *Rev. Mod. Phys.*, 78:275–306, Mar 2006.

[6] Andreas Holzner, Andreas Weichselbaum, Ian P. McCulloch, Ulrich Schollwöck, and Jan von Delft. Chebyshev matrix product state approach for spectral functions. *Phys. Rev. B*, 83:195115, May 2011.

[7] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96–192, 2011.

[8] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[9] Martin Ganahl, Patrik Thunström, Frank Verstraete, Karsten Held, and Hans Gerd Evertz. Chebyshev expansion for impurity models using matrix product states. *Phys. Rev. B*, 90:045144, Jul 2014.

[10] F. Alexander Wolf, Jorge A. Justiniano, Ian P. McCulloch, and Ulrich Schollwöck. Spectral functions and time evolution from the Chebyshev recursion. *Phys. Rev. B*, 91:115144, Mar 2015.

[11] P.P. Vaidyanathan. *The Theory of Linear Prediction.* Morgan and Claypool, 2008.

[12] Steven R White and Ian Affleck. Spectral function for the S= 1 Heisenberg antiferromagetic chain. *Physical Review B*, 77(13):134437, 2008.

[13] Thomas Barthel, Ulrich Schollwöck, and Steven R. White. Spectral functions in one-dimensional quantum systems at finite temperature using the density matrix renormalization group. *Phys. Rev. B*, 79:245101, Jun 2009.

Statutory Declaration

I hereby declare on oath that I completed this work on my own and that information which has been taken directly or indirectly from other sources has been noted as such. Neither this, nor a similar work has been published or presented to an examination committee.

....

place, date   signature