# Delving Into Dissipative Quantum Dynamics: From Approximate to Numerically Exact Approaches

Hsing-Ta Chen

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

under the Executive Committee

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

# ABSTRACT

## Delving Into Dissipative Quantum Dynamics:<br>From Approximate to Numerically Exact Approaches

Hsing-Ta Chen

In this thesis, I explore dissipative quantum dynamics of several prototypical model systems via various approaches, ranging from approximate to numerically exact schemes. In particular, in the realm of the approximate I explore the accuracy of Padé–resummed master equations and the fewest switches surface hopping (FSSH) algorithm for the spin–boson model , and non-crossing approximations (NCA) for the Anderson–Holstein model. Next, I develop new and exact Monte Carlo approaches and test them on the spin–boson model. I propose well–defined criteria for assessing the accuracy of Padé-resummed quantum master equations, which correctly demarcate the regions of parameter space where the Padé approximation is reliable. I continue the investigation of spin–boson dynamics by benchmark comparisons of the semiclassical FSSH algorithm to exact dynamics over a wide range of parameters. Despite small deviations from golden-rule scaling in the Marcus regime, standard surface hopping algorithm is found to be accurate over a large portion of parameter space. The inclusion of decoherence corrections via the augmented FSSH algorithm improves the accuracy of dynamical behavior compared to exact simulations, but the effects are generally not dramatic for the cases I consider. Next, I introduce new methods for numerically exact real-time simulation based on real-time diagrammatic QMC and the inchworm algorithm. These methods optimally recycle Monte Carlo information from earlier times to greatly suppress the dynamical sign problem. In the context of the spin–boson model, I formulate the inchworm expansion in two distinct ways: the first with respect to an expansion in the system–bath coupling and the second as an expansion in the diabatic coupling. In addition, a cumulant version of the inchworm Monte Carlo method is motivated by the latter expansion, which allows for further suppression of the growth of the sign error. I provide a comprehensive comparison of the performance of the inchworm Monte Carlo algorithms to other exact methodologies as well as a discussion of the relative advantages and disadvantages of each. Finally, I investigate the dynamical interplay between the electron–electron interaction and the electron–phonon coupling within the Anderson–Holstein model via two complementary NCAs: the first is constructed around the weak-coupling limit and the second around the polaron limit. The influence of phonons on spectral and transport properties is explored in equilibrium, for non-equilibrium steady state and for transient dynamics after a quench. I find the two NCAs disagree in nontrivial ways, indicating that more reliable approaches to the problem are needed. The complementary frameworks used here pave the way for numerically exact methods based on inchworm QMC algorithms capable of treating open systems simultaneously coupled to multiple fermionic and bosonic baths.

This page is intentionally left blank

# Contents

# List of Figures

# Acknowledgements

First and foremost, I would like to thank my advisor, Dave Reichman, for his constant support over these last five years. Dave has shown me many aspects of a respected scientist – never afraid to pursue new ideas, but always check results carefully. He has been a great source of inspiration and discussion. His suggestion and reminder are helpful when I get stuck in research. I am also very grateful for his patience and countless corrections of my writing.

I am extremely fortunate to have an excellent committee: Bruce Berne, Phillip Kim, Andrew Millis, Richard Friesner, and Seogjoo Jang. Their challenging questions and insightful comments help greatly fit my research into a larger scientific context. Bruce has been an important source of advice, both scientifically and professionally, and I am very thankful to have been able to interact with him. Also, I have benefited greatly from working with him assisting his Physical Chemistry course. Andy has had a great impact on my research. Working with Andy on the Anderson–Holstein model became a very important base for the Monte Carlo project.

I am grateful to be part of the group composed by many talented postdocs and graduate students. I had learned so much from working with Guy Cohen. Guy has walked me through all the details of diagrammatic Quantum Monte Carlo method and the inchworm algorithm. I want to thank him for hosting my summer visit in Tel Aviv University and showing me what is good hummus. I also had the opportunity to work with Tim Berkelbach on Pade resummation master equations. Tim inspired my first research project in graduate school. I would like to thank him for always being available for advice. My life in the group would not be as enjoyable without Andrés Montoya-Castillo. It is a luxury for a graduate student to have such a friend who can discuss our research everyday. Apart from works, we have had infinite cups of coffee together and he also teaches (adopts) me to appreciate a very specific type of humor. Andrés has shown me so much outside our science world and shaped part of me now.

Last but not least, I would like to thank my family. I have had unconditional support and encouragement my parents for all of my academic choices. My spouse Aimee Liu has become the anchor of my life. Being with you, I can call anywhere in the world our sweet home.

*For my family.*

# Chapter 1

# Introduction

## 1.1 Motivation

One of the central challenges in chemical physics research is to simulate the dynamics of open quantum systems. At first glance, this task seems to be solvable using modern computer simulation as long as one can write down differential equations that govern the time evolution of the quantum system. However, when trying to implement this in the condensed phase, one soon finds that the common truth of "conservation of difficulty": No matter what approach one takes, the task is nearly always complicated by numerical scaling or stability issues. Despite decades of research and a wide spectrum of methodologies that have been developed, a universal approach to the simulation of quantum dynamics is still difficult to impossible even for model systems. Therefore, understanding the accuracy of existing approaches and developing novel methods are both of great importance.

The challenge of exact real-time quantum dynamics remain substantial for several reasons. First, the size of Hilbert and Fock space scales exponentially with the size of the quantum system, rendering simple numerically exact methods feasible only for small systems. For example, a system of 10 interacting spins formally requires the description of $2^{10}$ many-body states making the direct simulation of such a system unfeasible. In addition, reduced quantities play an important role in dissipative quantum dynamics, and here one needs to keep track of the history of the system and take time non-local dynamics into account. This non-locality in time may manifest itself as strong memory effects which significantly complicate the simulation of quantum dynamics. Lastly, for systems in the condensed phase, multiple distinct types of interactions are presented. The dynamical interplay among these interactions, for example the hybridization of an tagged electron with conduction band electrons and the electron–phonon coupling, may produce myriad subtle physical effects. The existing approaches are often tailored to the needs of a particular system and mechanism at the expense of general accuracy and feasibility.

In this thesis I concentrate on the dynamics of prototypical dissipative quantum systems, describing an electronic subsystem interacting with a simple environment[1–3]. Specifically, I investigate the spin–boson model (a two-level system interacting with an oscillator bath) and the Anderson–Holstein model (a small electronic system interacting with electron and phonon baths). These models are known to provide reasonable descriptions of the essential physics of dissipative quantum systems and the successful simulation of their behavior should be considered a stringent test which any new approach to quantum dynamics must pass.

As mentioned above, the spin–boson model consists of a two-level system coupled linearly to a bosonic bath with a broad band spectrum. It has been used as the basic proxy for dissipative dynamics in many contexts, such as electron and energy transfer in condensed phases and biological systems[4–11], singlet fission dynamics[12–16], and entanglement and decoherence of qubits in condensed media[17–19]. The Anderson–Holstein model[20–22] differs from the spin–boson model in the fact that the electronic system, which takes the form of a correlated impurity site, is coupled to a non-interacting electron reservoir, as well as to phonons. It is a minimal model that embodies the dynamical interplay between electron–electron interactions and electron–phonon coupling, and has implications for the study of superconductivity[23–28], photon-induced metal–insulator transitions in correlated materials[25, 29–32], and the Kondo effect in non-equilibrium nanoscale devices[33–40]. Furthermore, it can be used a basic model for correlated materials with active phonon degrees of freedom. In particular, a strongly-correlated material with active phonon degrees of freedom is precisely mapped onto to the Anderson–Holstein model within the framework of DMFT (dynamical mean-field theory)[41].

## 1.2 Theoretical approaches

### 1.2.1 Spin–boson model

A variety of approaches for solving the dynamics of the spin–boson model have emerged over the past several decades. They can be categorized into three sets: perturbation theories, semiclassical approaches, and numerically exact schemes. In general, perturbation theories are limited to specific regions of parameter space; semiclassical approximations can be efficient and scalable but often involve uncontrolled approximations; numerically exact approaches often scale unfavorably in their computational cost with the simulation time and with the dimensionality of the system. In the following I briefly discuss the relative benefits and drawbacks of the classes of approaches.

### Perturbation theories

Perturbative theories are generally carried out with respect to particular coupling parameters of the model to yield solvable equations of motion. The most widely-used perturbation parameters of the spin–boson model are the system–bath coupling and the electronic (diabatic) coupling. Perturbation expansions carried out to second order in the system–bath coupling yields the so-called Redfield equations[42, 43]. A series of treatments and generalizations[44–47] have been developed based on the Redfield expansion. Expansion in the diabatic coupling yields another family of methods, which includes the noninteracting blip approximation (NIBA) [48, 49] and Padé-resummed master equation approach[50–55]. All such treatments are only reliable within a restricted region of parameter space where the perturbation parameter is sufficiently small, and may lead to unphysical results if the underlying perturbation series does not converge. While extensions of perturbation theories have been proposed[46, 56–58], it is usually unclear how reliable they are in the full parameter space.

### Semiclassical approaches

The semiclassical approaches I will discuss in this thesis rely on the idea that the system must be described quantum mechanically while the bath degrees of freedom can be treated as a swarm of classical trajectories. Many such methods have been developed, for example the linearized semiclassical initial value representation (LSC-IVR) scheme[59–61], the time-dependent self-consistent field (TDSCF) method[62, 63], and the fewest-switches surface hopping (FSSH) approach[64–66]. These methods are in principle inexpensive, non-perturbative, and can be used to treat realistic anharmonic baths and large systems. However, the effectively classical description of the bath degrees of freedom renders them incapable of capturing some quantum mechanical effects, such as nuclear tunneling at low temperatures[67–69]. More importantly, the approximations underlying these methods are uncontrolled, which make it difficult to determine their reliability.

Among these methods, the FSSH approach has proven to be one of the most popular techniques for the calculation of non-adiabatic dynamics in the condensed phase[70–72]. The FSSH provides a superior description of branching processes and microscopic detailed balance. Its wide use can be attributed to the fact that it can conveniently be employed in conjunction with electronic structure calculations in the adiabatic representation. However, aside from the drawbacks associated with its underlying classical dynamics approximation, a long-recognized shortcoming of standard FSSH is the fact that the algorithm does not provide any electronic decoherence mechanism. Recent research efforts have attempted to formulate descriptions of decoherence within FSSH to account for decoherence[73–86].

### Numerically exact schemes

Various numerically exact schemes have been proposed for the spin–boson model based on different quantum mechanical tools, including wavefunction, density matrix, and path integral representations. One of the most successful numerically exact schemes for spin–boson dynamics is the multi–layer multi–configurational time–dependent Hartree (ML–MCTDH) method[87–89], which is a wavefunction-based, fully quantum mechanical approach. ML-MCTDH has been particularly accurate over a quite wide swath of parameter space for equilibrium spin–boson problems, but has difficulty converging for transport problems far from equilibrium[90–92]. The hierarchical equation of motion (HEOM) method[45, 93, 94] is based on the density matrix description. It introduces an infinite hierarchy of auxiliary density matrices and a Matsubara expansion for the bath density matrix. In its standard implementation, the HEOM method relies on the truncation of the infinite coupled differential equations, which make it difficult to converge for the cases of low bath temperatures or strong system–bath coupling.

The path integral description provides a different means of computing exact dynamical properties in the spin–boson model. Two such path integral methods are the quasi–adiabatic propagator path integral (QUAPI) representation[95–98] and Quantum Monte Carlo (QMC) approaches[99–102]. QUAPI is based on the time discretization of the quantum mechanical propagator and a truncation of the temporal range of the influence functional in the path integral representation. However, systematic errors can arise from both the truncation of non-local retarded interactions and the Trotter decomposition that is used to disentangle the system and bath propagators[103]. QMC techniques can provide an unbiased, efficient estimation of the propagator in the path integral description, but the evaluation of dynamical properties requires a summation of complex propagators, which result in a dynamical sign problem. The sign problem causes the computational cost to scale exponentially with increasing time and restricts QMC methods to short times. Recent developments in QMC techniques[101, 104–106] bring the hope for the exact QMC calculation of long-time dynamics closer to reality.

### 1.2.2 Anderson–Holstein model

Despite the importance and simplicity of the Anderson–Holstein model, there is little known about its real-time dynamical properties beyond simple limits where perturbative arguments can be made. Several limits (or simplified forms) of the model, such as the zero and infinite Coulomb interaction cases, have been extensively investigated and provide conceptually essential insight regarding phonon effects[107–113]. However, there are only a handful of approaches that are capable of calculating dynamical properties in the Anderson–Holstein model for generic cases. These approaches largely originate from existing methods for the Anderson impurity model and are tailored to the needs of particular limits.

#### Numerical renormalization group

The numerical renormalization group (NRG) has been developed based on the low energy theory of the Anderson impurity model and can be extended to include electron–phonon interaction. For the Anderson–Holstein model, the NRG results capture interesting features, such as Kondo peak broadening and replication[107, 114–119]. Nevertheless, NRG is generally reliable only for the low energy properties of the system, and remains difficult to apply out of equilibrium.

#### Quantum Monte Carlo

QMC-based methods, such as the auxiliary-field QMC and real-time diagrammatic QMC, are particularly successful in solving the dynamics of the Anderson model, and allow for the exploration of transient dynamics and non-equilibrium transport properties over a wide range of parameters. Their extension to the Anderson–Holstein model poses additional challenges and complications. The auxiliary-field QMC method relies on an analytical continuation to obtain real-time dynamical properties from the density of states calculated under the influence of the phonons in imaginary time[120]. However, the density of states description is only valid for equilibrium and linear response properties and the analytical continuation is uncontrolled and can be problematic for certain parameters[121, 122]. The real-time diagrammatic QMC method evaluates dynamical properties directly by stochastically sampling real-time diagrams[123–128]. In conjunction with partial resummations of the exact diagrammatic series[129] and reduced dynamics techniques[130, 131], it can be used to obtain results up to previously unreachable timescales, at least for the Anderson model[132]. Although it is generically plagued by the dynamical sign problem, recent algorithmic advances, such as the bold-line techniques and the "inchworm" algorithm[106, 121, 122, 129], might allow for an amelioration of the sign problem in real-time QMC simulations for the Anderson–Holstein problem, although such calculations have yet to attempted for this model.

## 1.3 Outline of thesis

The organization of the dissertation is as follows: In Chapter 2, I start with perturbative treatments of spin–boson dynamics. I propose well–defined criteria for assessing the accuracy of quantum master equations whose memory functions are approximated by Padé resummation of the first two moments in the electronic coupling. These criteria partition the parameter space into distinct levels of expected accuracy. Extensive comparison of Padé–resummed master equations with numerically exact results in the context of the spin–boson model demonstrate that the proposed criteria correctly demarcate the regions of parameter space where the Padé approximation is reliable. The applicability analysis I present is not confined to the specifics of the Hamiltonian under consideration and should provide guidelines for other classes of resummation techniques.

In Chapter 3, I continue the investigation of spin–boson dynamics via the semiclassical surface hopping algorithm. I perform extensive benchmark comparisons of surface hopping dynamics with numerically exact calculations for the spin-boson model over a wide range of parameters. FSSH is found to be surprisingly accurate over a large swath of parameter space. The inclusion of decoherence corrections via the augmented FSSH (A-FSSH) algorithm improves the accuracy of dynamical behavior compared to exact simulations, but the effects are generally not dramatic, at least for the case of an environment modeled with the commonly used Debye spectral density.

In Chapter 4, I introduce a set of new methods for the spin–boson model based on real-time diagrammatic QMC and inchworm algorithm. A detailed description of the inchworm Monte Carlo formalism is provided for the exact study of real-time non-adiabatic dynamics. This method optimally recycles Monte Carlo information from earlier times to greatly suppress the dynamical sign problem. Using the example of the spin-boson model, I formulate the inchworm expansion in two distinct ways: the first with respect to an expansion in the system-bath coupling and the second as an expansion in the diabatic coupling. The latter approach motivates the development of a cumulant version of the inchworm Monte Carlo method which has the benefit of further suppression of the growth of the sign error. I provide a comprehensive comparison of the performance of the inchworm Monte Carlo algorithms to other exact methodologies as well as a discussion of the relative advantages and disadvantages of each.

In Chapter 5, I turn attention to the Anderson–Holstein model. I investigate the dynamical interplay between electron–electron interactions and electron–phonon coupling via two complementary non-crossing approximations (NCA). The first NCA is constructed around the weak-coupling limit and the second around the polaron limit. The influence of phonons on spectral and transport properties is explored in equilibrium, for non-equilibrium steady state and for transient dynamics after a quench. Both the particle–hole symmetric and the more generic particle–hole asymmetric cases are studied. In general, the two methods disagree in nontrivial ways, indicating that more reliable approaches to the problem are needed. Importantly, the frameworks used here can form the starting point for numerically exact methods based on bold-line QMC algorithms capable of treating open systems simultaneously coupled to multiple fermionic and bosonic baths.

# Chapter 2

# Padé-Resummed Master Equation Approach to Dissipative Quantum Dynamics*

## 2.1 Introduction

Schemes based on projection operator techniques[133] and generalized quantum master equations (GQMEs) have been used both to design successful approximate approaches and as a platform to develop numerically exact methods[2, 3, 42, 43, 93]. The projection operator

---

*Based on work published in J. Chem. Phys. **144**, 154106 (2016). Copyright 2016, American Institute of Physics.

technique partitions the Hilbert space into system and bath subspaces, leading to the derivation of GQME for the system subspace which accounts for the bath's dynamical influence on the system via a memory kernel. Exact and approximate techniques for the evaluation of the memory kernel have been developed that make use of perturbation theories[45, 46, 48, 49, 56], resummation techniques[50, 51, 54, 55, 134, 135], and self–consistent expansions[130, 131, 136, 137]. Recent progress afforded by these methods has illustrated several advantages of the GQME scheme. First, the memory kernel may decay on a shorter timescale than the system dynamics under study, so that approximate memory kernels may yield more accurate dynamics than would be obtained by direct simulation of the system dynamics using the same level of approximation. Second, the GQME scheme is general enough to treat realistic anharmonic baths[134, 137] and arbitrary system–bath coupling[137]. Finally, the flexibility of different projection operator formulations allows for facile extension to more general situations, such as nonequilibrium initial preparation[52, 53], as well as more complex correlation functions[138]. However, despite these notable results, it remains a difficult task to accurately calculate memory kernels in many regimes of general quantum dissipative systems.

The Padé resummation approach approximates the memory kernel as an infinite resummation based on the kernel's second and fourth moments[50, 51, 139]. At the expense of fourth–order perturbation theory in the electronic coupling, the Padé–resummed GQME is capable of producing an accuracy that exceeds that of simple perturbation theory for the spin–boson model[134], and resummation of higher order kernels provide quantitative corrections[140]. Recently, however, it has been demonstrated that this approach can lead to unphysical, divergent dynamics in the strong electronic coupling regime[135], and the applicability and accuracy of the Padé approximations throughout the entire parameter space is still difficult to evaluate. The aim of the present work is to provide feasible estimates of the applicability based on analysis of the Padé approximation itself.

We propose well–defined criteria in terms of the kernel's second and fourth moments that correspond to conditions leading to "physically reasonable" results within the Padé resummation scheme. To examine the proposed criteria, we perform systematic benchmark comparisons of Padé–resummation with numerically exact results for a prototypical dissipative open quantum system, namely the spin–boson model with a Debye spectral density. The proposed criteria divide the parameter space into subspaces associated with different levels of accuracy, and we confirm that the systematic comparison of population dynamics with exact results clearly demarcate when the approach should provide quantitatively reliable results. It should be noted that the proposed criteria are not limited to the spin–boson model, but are generally applicable for estimating the accuracy of Padé–resummed memory kernels for generic open quantum systems. In addition, the present work may provide guidelines for the applicability of other types of resummation techniques, such as the Landau–Zener resummation[135].

The outline of the chapter is as follows. We present in Sec. 2.2 a brief review of the nonequilibrium Padé–resummed GQME approach to a generic open quantum system. In Sec. 2.3, we analyze the Padé resummation and define the criteria for the validity of the approximation. We apply the proposed criteria to the spin–boson model in Sec. 2.4 and show the correspondence of the different regions of the applicability phase diagrams with exactly computed population dynamics. In Sec. 2.5, we conclude.

## 2.2 The Padé resummed GQME approach

We consider an open quantum system whose Hamiltonian takes the form, $\hat{H} = \hat{H}_s + \hat{H}_b + \hat{V}$, where $\hat{H}_s$ and $\hat{H}_b$ correspond to the system and bath Hamiltonians, respectively, and $\hat{V}$ denotes the system–bath coupling. We denote the quantum states of the system by the kets $|j\rangle$ and the bath density operator by $\hat{\rho}$. It is convenient to adopt the Liouville space notation[50, 51] for the total density operator, $\hat{W} \equiv |W\rangle\rangle$, and define the product $\langle\langle A|B\rangle\rangle \equiv \mathrm{Tr}_s\mathrm{Tr}_b\{A^\dagger B\}$ where $\mathrm{Tr}_s$ and $\mathrm{Tr}_b$ are partial traces over the states of the system and bath, respectively. Time evolution of the density operator is governed by the Liouville–von Neumann equation

$$\frac{d}{dt}|W(t)\rangle\rangle = -i\mathcal{L}|W(t)\rangle\rangle, \tag{2.1}$$

where the Liouville super–operator (the Liouvillian) is defined by $\mathcal{L}|W(t)\rangle\rangle = [\hat{H}, \hat{W}(t)]$ and we set $\hbar = 1$ throughout this chapter. The reduced density matrix of the system can be written as $\sigma_{jk}(t) = \mathrm{Tr}_b\left\{|k\rangle\langle j|\, W(t)\right\} = \langle\langle jk|W(t)\rangle\rangle$ where the Liouville state is given by $|jk\rangle\rangle = |j\rangle\langle k|\otimes\hat{1}$ and $\hat{1}$ is the unit operator for the bath. Then we can denote the population dynamics as

$$P_j(t) = \langle\langle j|W(t)\rangle\rangle, \tag{2.2}$$

where the diagonal elements are expressed as $|jj\rangle\rangle \to |j\rangle\rangle$ for simplicity.

We implement the standard projection operator technique[141] via the super–operator

$$\mathcal{P} = \sum_j |j\rho_j\rangle\rangle\langle\langle j| \tag{2.3}$$

where $|j\rho_j\rangle\rangle = |j\rangle\langle j| \otimes \hat{\rho}_j$ and the bath density operator $\hat{\rho}_j$ is taken to be in equilibrium in the electronic state $|j\rangle$. The projected version of Eq. (2.1) yields the GQME for the population of the $j$-th state,

$$\frac{d}{dt}P_j(t) = \mathcal{I}_j(t) - \sum_k \int_0^t d\tau \mathcal{K}_{jk}(t-\tau)P_k(\tau), \tag{2.4}$$

where the memory kernel matrix is

$$\mathcal{K}_{jk}(t) = \langle\langle j|\mathcal{PL}e^{-i\mathcal{QL}t}\mathcal{QL}|k\rho_k\rangle\rangle, \tag{2.5}$$

and the inhomogeneous terms are given by

$$\mathcal{I}_j(t) = -i\langle\langle j|\mathcal{PL}e^{-i\mathcal{QL}t}\mathcal{Q}|W(0)\rangle\rangle, \tag{2.6}$$

with $\mathcal{Q} = 1 - \mathcal{P}$. The inhomogeneous terms result from the fact that the initial condition for the total density operator will generally satisfy $\mathcal{Q}|W(0)\rangle\rangle \neq 0$. For cases $\mathcal{Q}|W(0)\rangle\rangle = 0$, $\mathcal{I}_j(t) = 0$. In the frequency domain, Eq. (2.4) can be transformed from an integro–differential equation into the algebraic form

$$sp_j(s) = p_j(t=0) + I_j(s) - \sum_k K_{jk}(s)p_k(s) \tag{2.7}$$

with the use of the one–side Laplace transformation, $f(s) = \int_0^\infty e^{-st}F(t)dt$, where $s$ is a complex number. It should be noted that calculation of the memory kernel matrix and the inhomogeneous terms is difficult in part because dynamical evolution involves a projected propagator $e^{-i\mathcal{QL}t}$.

To approximate the projected propagator, one can carry out a perturbation treatment with respect to a perturbation $\hat{H}'$ and an unperturbed Hamiltonian $\hat{H}_0 = \hat{H} - \hat{H}'$. The Liouvillian can be decomposed as $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}'$ and Eq. (2.5) and (2.6) can be expanded in terms of $\mathcal{L}'$. As a result, the memory matrix and the inhomogeneous terms in frequency domain can be expressed as a moment expansion

$$K_{jk}(s) = \sum_{n=1}^{\infty} K_{jk}^{(2n)}(s) \tag{2.8}$$

and

$$I_j(s) = \sum_{n=1}^{\infty} I_j^{(2n)}(s) \tag{2.9}$$

with

$$K_{jk}^{(2n)}(s) = \langle\langle j| \left[\mathcal{L}' G_0(s) \mathcal{L}' G_0(s) \mathcal{Q}\right]^{n-1} \mathcal{L}' G_0(s) \mathcal{L}' |k\rho_k\rangle\rangle, \tag{2.10}$$

and

$$I_j^{(2n)}(s) = \langle\langle j|\mathcal{L}' G_0(s)[\mathcal{Q}\mathcal{L}' G_0(s)]^{2n-1}\mathcal{Q}|W(0)\rangle\rangle, \tag{2.11}$$

where the unperturbed Green's function is $G_0(s) = (s + i\mathcal{L}_0)^{-1}$. In practice, evaluating the $(2n)$-th order moment requires a Laplace transformation for each time variable in a $(2n-1)$-time correlation function. Clearly, the complexity of the terms in the moment expansion grows quickly as the moment order increases.

The memory matrix and inhomogeneous terms may be approximated by a Padé resummation using the second and fourth moments in the frequency domain,

$$K_{jk}(s) \approx \frac{[K_{jk}^{(2)}(s)]^2}{K_{jk}^{(2)}(s) - K_{jk}^{(4)}(s)}, \tag{2.12}$$

$$I_j(s) \approx \frac{[I_j^{(2)}(s)]^2}{I_j^{(2)}(s) - I_j^{(4)}(s)}. \tag{2.13}$$

It should be noted that the Padé resummation is a rational expression that include infinite orders of the perturbation $\hat{H}'$, but the contributions of higher order than the fourth are approximated; for example, $K_{jk}^{(6)} \approx \left[K_{jk}^{(4)}(s)\right]^2 / K_{jk}^{(2)}(s)$. The expressions of this section have been discussed before[55], but a systematic analysis is lacking. We now focus precisely on this issue.

## 2.3 Applicability Analysis and Criteria

The accuracy of the Padé resummation is unknown and depends on the analyticity of an unknown function in the complex plane. Despite this fundamental difficulty, we may estimate its validity via simple convergence properties and physical requirements of the memory kernels. For simplicity below, the criteria are expressed in terms of a single memory kernel element $K(s)$, thereby suppressing the indices associated with memory functions and inhomogeneous terms.

The Padé resummation can be viewed as a complex geometric series which is expected to yield well–behaved results only within the disk of convergence of the Laurent series that represents the expansion in the complex plane. A necessary condition for such convergence is $\|K^{(4)}(s)/K^{(2)}(s)\| < 1$, for all complex number $s$, where $K^{(n)}(s)$ is the $n$-th order expression given in Eq. (2.10). Since the inverse Laplace transformation is performed along the imaginary axis $s = i\omega$, we restrict this condition to

$$\text{(a)} \qquad \|K^{(4)}(i\omega)/K^{(2)}(i\omega)\| < 1 \text{ for real } \omega. \tag{2.14}$$

The above condition is quite strict and may be relaxed by consideration of the physical requirements of a generic memory kernel. Consider the Laplace inversion via the contour integration of the Bromwich integral $K(t) = \frac{1}{2\pi i}\int_{\mathcal{C}} K(s)e^{st}ds$, where $\mathcal{C}$ is the vertical contour in the complex plane chosen to include all singularities of $K(s)$ to the left of it[142, 143]. The asymptotic physical behavior of the memory kernel dictates that the poles of the Padé–resummed approximation *cannot* have a non–negative real part, otherwise the memory function would not be guaranteed to decay to zero as $t \to \infty$. We assume that the distribution of poles changes continuously and smoothly as the parameters of the model changes, allowing us to focus on the imaginary axis $s = i\omega$ and monitor the behavior of $K^{(4)}(i\omega)/K^{(2)}(i\omega)$. In particular, the equality $\mathrm{Re}[K^{(4)}(i\omega^*)/K^{(2)}(i\omega^*)] = 1$ is a necessary (albeit not sufficient) condition for the occurrence of a pole on the imaginary axis at $s = i\omega^*$, which obviate the asymptotic decay of the memory kernel in real–time. We thus propose a second condition

$$\text{(b)} \qquad \mathrm{Re}[K^{(4)}(i\omega)/K^{(2)}(i\omega)] < 1 \text{ for real } \omega, \tag{2.15}$$

which, excepting random occurrences, maintains that all poles are confined to the left of the imaginary axis in the complex plane and that the memory function is well behaved. Note that the first criterion is stricter than the second since it corresponds to the interior of a unit circle in the complex plane while the latter condition corresponds to the entire complex plane to the left of the boundary at $\mathrm{Re}[z] = 1$.

These criteria are indeed crude because they rely on the the limited information of the first two terms of an infinite expansion. We will employ these conditions below as demarcation lines in parameter space to gauge the reliability of the Padé approximation. As will be demonstrated, the criteria provide robust if conservative guidelines for the domain of applicability for Padé–resummed master equations.

## 2.4 Results for the Spin–Boson Model

### 2.4.1 The Spin–Boson Model

In this section, we examine the criteria suggested above via investigation of the population dynamics in the spin–boson model. The spin–boson model is an idealization of an open quantum system which contains most of the important generic features of more complicated dissipative quantum systems while offering the advantage that numerically exact algorithms exist for the calculation of its dynamics over a wide range of parameter space[93, 95, 96, 144]. To produce benchmark results for the spin–boson model in this work, we use the numerically exact hierarchical equations of motion (HEOM) methodology in the Parallel Hierarchy Integrator (`PHI`) [94].

We consider a two–level system with energy bias $\epsilon$ and constant electronic coupling $\Delta$

$$\hat{H}_s = \epsilon\hat{\sigma}_z + \Delta\hat{\sigma}_x,$$

and $\hat{\sigma}_z = |1\rangle\langle 1| - |2\rangle\langle 2|$ and $\hat{\sigma}_x = |1\rangle\langle 2| + |2\rangle\langle 1|$. The two–level system is coupled to a bath consisting of an infinite set of harmonic oscillators

$$\hat{H}_b = \sum_\alpha \frac{\hat{P}_\alpha^2}{2} + \frac{1}{2}\omega_\alpha^2\hat{Q}_\alpha^2. \tag{2.16}$$

Here, the frequency of the $\alpha$-th bath mode is $\omega_\alpha$, while $\hat{P}_\alpha$, $\hat{Q}_\alpha$ refer to the mass–weighted momenta and coordinates of the $\alpha$-th mode. The system–bath coupling is taken to be of the form

$$\hat{V} = \hat{\sigma}_z \sum_\alpha c_\alpha \hat{Q}_\alpha, \tag{2.17}$$

where $c_\alpha$ is the coupling strength between the two–level system and the $\alpha$-th harmonic oscillator. The spectral density compactly describes the influence of the bath on the dynamics of the system, and takes the form

$$J(\omega) = \frac{\pi}{2} \sum_\alpha \frac{c_\alpha^2}{\omega_\alpha} \delta(\omega - \omega_\alpha). \tag{2.18}$$

In our study we choose the commonly used Debye spectral density[144]

$$J(\omega) = \frac{\lambda}{2} \frac{\omega \omega_c}{\omega^2 + \omega_c^2}, \tag{2.19}$$

which is Ohmic at low frequency with a Lorentzian cutoff at high frequency. The Debye spectral density is characterized by two parameters: the characteristic bath frequency $\omega_c$, which represents the average timescale of the bath response, and the reorganization energy $\lambda = \sum_\alpha c_\alpha^2 / 2\omega_\alpha^2$, which is a direct measure of the coupling strength between the system and the bath. In electron–transfer theory, the Debye spectral density is commonly used for the description of a solvent environment with Debye dielectric relaxation (i.e. exponential in time).

Throughout this work, we employ an initial density operator for the bath of the form

$$\hat{\rho}_0 = \frac{e^{-\beta \hat{H}_b}}{\mathrm{Tr}_b\{e^{-\beta \hat{H}_b}\}}, \tag{2.20}$$

where $\beta = 1/k_B T$ is the inverse temperature of the bath. This initial condition corresponds to thermal equilibrium in the reservoir in the absence of the system–bath coupling and is the initial density operator of relevance for the description of an impulsive Franck–Condon excitation.

We implement a commonly used projection operator of the form[50, 51],

$$\mathcal{P} = |1\rho_1\rangle\rangle\langle\langle 1| + |2\rho_2\rangle\rangle\langle\langle 2|, \tag{2.21}$$

where fully–dressed equilibrium bath density operators of the form

$$\hat{\rho}_j = \frac{e^{-\beta\hat{H}_j}}{\mathrm{Tr}_b\{e^{-\beta\hat{H}_j}\}} \tag{2.22}$$

are employed with $\hat{H}_j = \pm(\epsilon + \sum_\alpha c_\alpha \hat{Q}_\alpha) + \hat{H}_b(+$ for 1 and $-$ for 2). Note that with the use of the projector (2.21), factorized initial conditions with an uncorrelated bath (2.20) will necessitate the evolution of inhomogeneous terms (2.13) in the GQME. The second–order moments of the memory kernels ($K^{(2)}_{jk}$) result in an expression equivalent to the noninteracting blip approximation (NIBA) [3]. We carry out the time integrations of the memory kernels and the inhomogeneous terms by the techniques outlined in Ref. 134 and the same Gaussian quadrature subroutine (`DCUTRI`) [145]. The population dynamics of the Padé–resummed GQME is calculated via the accuracy–improved numerical method for Laplace inversion[143].

For this spin–boson model, the Padé–resummed GQME approach has lead to population dynamics in near perfect agreement with numerically exact simulations[54, 55, 134]. On the other hand, Van Voorhis and coworkers have shown the breakdown of the Padé–resummed GQME approach in the strong electronic coupling region[135]. Our goal in the following is to systematically delineate the regime of validity of the approach based on the criteria of Sec 2.3.

### 2.4.2 Parameter Space Diagrams

The model we study in this work can be parametrized by five independent energy scales. We use the electronic coupling $\Delta$ as the unit of energy so that four dimensionless parameters characterize the parameter space. These are: the electronic bias $\epsilon/\Delta$, the reorganization energy $\lambda/\Delta$, the bath's characteristic frequency $\omega_c/\Delta$, and the thermal energy of the bath $k_BT/\Delta$.

To systematically scan parameter space, We consider variation in the scaled $\omega_c$–$\lambda$ plane for different scaled temperature and bias cuts. It is expected that, for a given system–bath coupling $\lambda$, smaller values of $\omega_c/\Delta$ render the Padé approximation less accurate due to the fact that the perturbation series is ordered by $\Delta$. Therefore, we define critical characteristic frequencies, $\omega_c^{(a)}(\lambda)$ and $\omega_c^{(b)}(\lambda)$, as the *lower* bound of scaled $\omega_c$ to satisfy the criteria (a) and (b) for all elements of the memory kernels respectively. The boundaries $\omega_c^{(a)}(\lambda)$ and $\omega_c^{(b)}(\lambda)$ are determined by the conditions that there exists a single imaginary number $i\omega^*$ for which either

$$\text{(a)} \qquad \|K^{(4)}(i\omega^*)/K^{(2)}(i\omega^*)\| = 1,$$

or

$$\text{(b)} \qquad \mathrm{Re}[K^{(4)}(i\omega^*)/K^{(2)}(i\omega^*)] = 1,$$

is satisfied. The critical characteristic frequencies indicate the boundaries of the proposed criteria that partition parameter space into three distinct regions of different levels of accuracy.

Because $K(t)$ and $I(t)$ have similar structure that should decay to zero after a transient time and the Padé approximation takes the same form for both $K(s)$ and $I(s)$, we expect the proposed criteria also apply to the inhomogeneous term. In fact, Refs. 54–134 have shown that the initial preparation of the bath, captured by the inhomogeneous term, is crucial for obtaining the correct dynamics. We only focus here on the memory kernel and expect the inhomogeneous term have similar analytical behaviors.

For illustrative purposes, we show a phase diagram for an unbiased ($\epsilon = 0$), high temperature ($k_B T = 2\Delta$) system in Fig. 2.1 and the corresponding population dynamics of selected points in parameter space calculated by the HEOM, Padé and NIBA approaches in the lower panels. For this example the three regions may be partitioned as:

1. $\omega_c > \omega_c^{(a)}$ (quantitatively accurate):

   In this regime, the results of the Padé approach achieve almost perfect agreement with the numerically exact results. This regime covers the weak electronic coupling (non–adiabatic) limit ($\omega_c/\Delta \gg 1$), where the $\Delta$–perturbation based methods works well. We also find that the Padé–resummed approach provides a significant improvement over NIBA in the large system–bath coupling regime, as can be seen in the upper panels of Fig. 2.1 (b).

2. $\omega_c^{(b)} < \omega_c < \omega_c^{(a)}$ (semi–quantitatively accurate):

   The population dynamics of the Padé approach in this region are not quite as accurate as in the "quantitatively accurate" regime, but the Padé–resummed method still captures most of the important features in a semi–quantitative manner, such as the long–lived oscillations and dissipative relaxation. Since the electronic coupling is considered to be intermediate, the NIBA results become worse while the Padé results remain accurate.

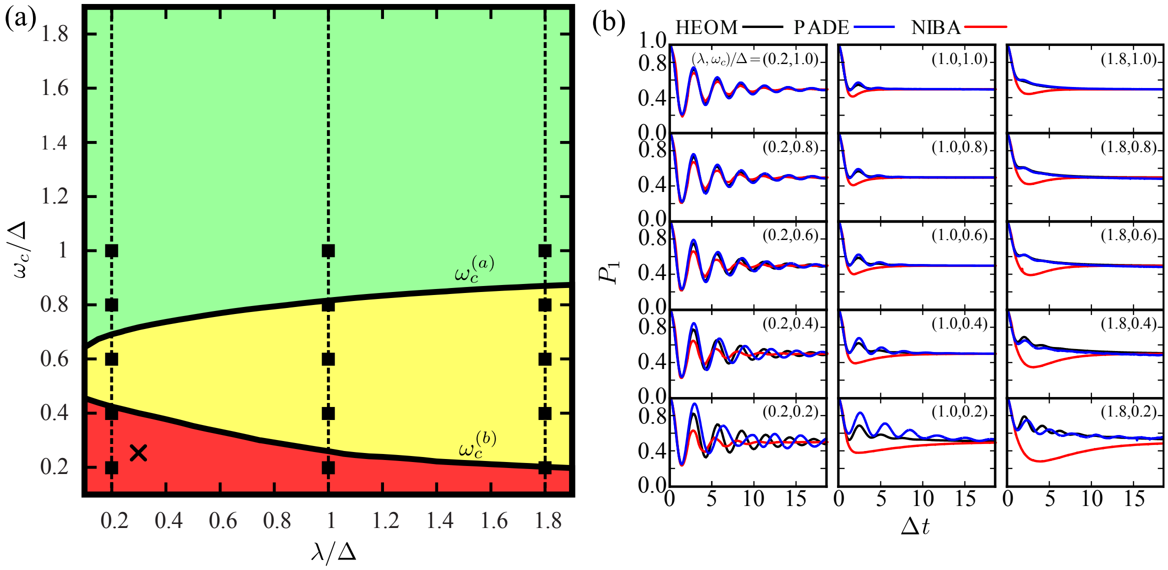


Figure 2.1: Parameter space diagram for the spin–boson model with zero bias ($\epsilon = 0$) and at high temperature ($k_BT = 2\Delta$). The critical frequencies $\omega_c^{(a)}$ and $\omega_c^{(b)}$ are indicated as functions of $\lambda$. The green region ($\omega_c > \omega_c^{(a)}$) is the regime where dynamics are expected to be quantitatively accurate, the yellow region ($\omega_c^{(b)} < \omega_c < \omega_c^{(a)}$) is the regime where dynamics are expected to be semi–quantitatively accurate and the red region ($\omega_c < \omega_c^{(b)}$) is the regime where the Padé–resummed approach is expected to be unreliable or even unstable. The lower panels are the corresponding population dynamics along the vertical cuts (indicated as solid squares connected by dashed lines) calculated by the HEOM approach (red solid lines), Padé–resummed GQME (PADE, green dash lines), and NIBA (blue doted lines). The upper right label in each population dynamics panel denotes the value of $(\lambda, \omega_c)/\Delta$. The symbol × in the phase diagram refers to the parameters corresponding to Fig. 1(d) of Ref. 135.

3. $\omega_c < \omega_c^{(b)}$ (unreliable):

   The discrepancies in the population dynamics between the Padé approach and the HEOM generally become larger in this regime since the large electronic coupling ($\Delta/\omega_c \gg 1$) renders the perturbation theory in $\Delta$ questionable. In this regime, the Padé approach may lead to a shift of the oscillation frequency of the population (see panels labeled by $(\lambda, \omega_c)/\Delta = (0.2, 0.4),\ (0.2, 0.2),\ (1.8, 0.2)$), as well as overly coherent behavior (see the panel labeled by $(\lambda, \omega_c)/\Delta = (1.0, 0.2)$). Extreme cases in the strong electronic coupling (adiabatic) limit may cause the Padé resummation breakdown and result in unphysical population dynamics. Importantly, the parameters of Fig. 1(d) of Ref. 135 lie in the "unreliable" region (labeled by $\times$ in the phase diagram). In this case the Padé–resummed approach yields unphysical population dynamics for the long time behavior.

Despite qualitatively similar behaviors, we notice that our results near the parameters marked by $\times$ appear to be more accurate than those of Ref. 135. One possible reason may be attributed to numerical errors of the FFT–based Laplace inversion method of Honig and Hirdes[146]. Here, we employ a simple improved method proposed by Yonemoto et al.[143]. In addition, we note that Ref. 135 assumes $I_j(s) = 0$ which may yield different population dynamics for short times.

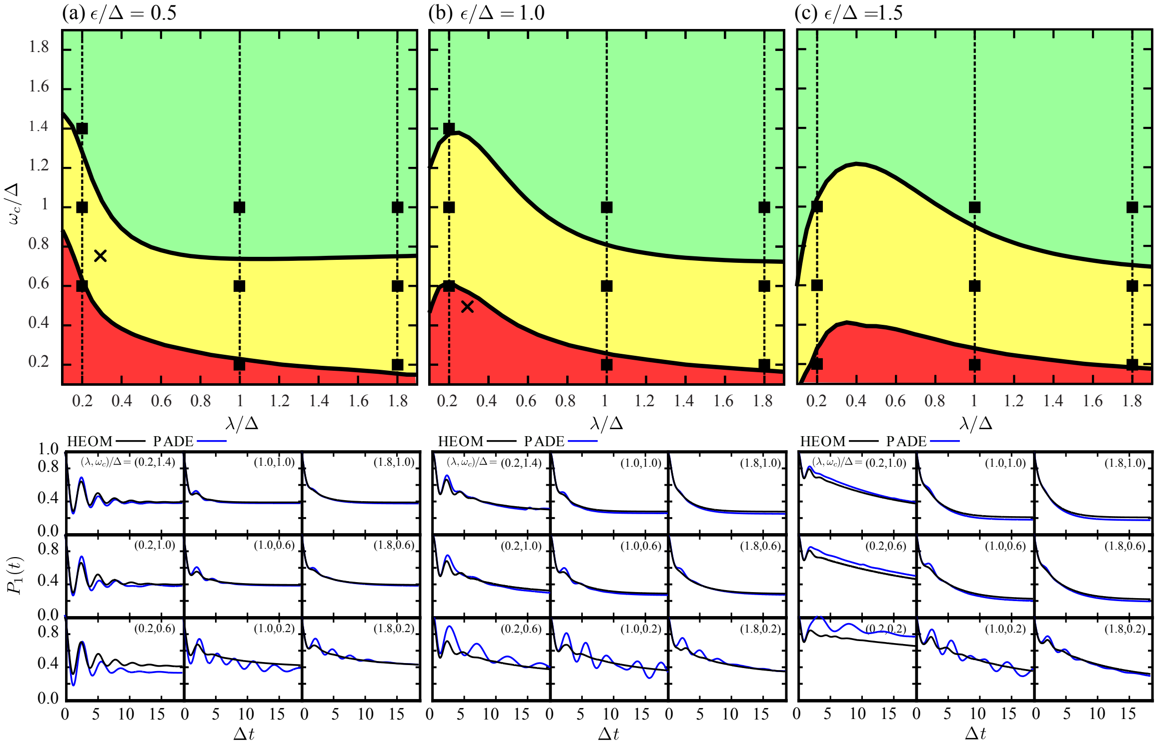


Figure 2.2: Parameter space diagram with increasing bias energies $\epsilon/\Delta = 0.5,\ 1,\ 1.5$ at high temperatures ($k_BT = 2\Delta$) for the spin–boson model. The critical frequencies $\omega_c^{(a)}$ and $\omega_c^{(b)}$ are indicated as functions of $\lambda$ with color regions as in Fig. 2.1. The lower panels are the corresponding population dynamics along the vertical cuts calculated by the HEOM approach (red solid line) and the Padé–resummed GQME (PADE, green dash line). The upper right label in each population dynamics plot denotes the value of $(\lambda, \omega_c)/\Delta$. The symbol × in panel (a) refers to the parameters corresponding to Fig. 3(b) of Ref. 135, while that in panel (b) corresponds to Fig. 4(b) of Ref. 135.

### 2.4.3 Energetic Bias Dependence

The bias dependence of the parameter space phase diagram is shown in Fig. 2.2, as well as the corresponding population dynamics. We find that, as the energetic bias grows, both critical frequencies increase in the low $\lambda$ region. Furthermore, in the region when $\omega_c < \omega_c^{(b)}$, the Padé approach may lead to incorrect steady state population values (see the panels labeled $(\lambda, \omega_c)/\Delta = (0.2, 0.6)$ for $\epsilon = \Delta$ and $(\lambda, \omega_c)/\Delta = (0.2, 0.2)$ for $\epsilon = 3\Delta$) as well as an unphysical "recoherence" behavior (namely the envelope of the population does not decay monotonically) as illustrated in the panels $(\lambda, \omega_c)/\Delta = (1.0, 0.2)$ for all biases. This effect can be attributed to near singularities in the approximate kernels when the Padé resummation does not satisfy the criterion (b). The population dynamics in Fig. 3(b) and Fig. 4(b) of Ref. 135 show qualitatively similar discrepancies from exact calculations as illustrated here. The parameters for these two cases (labeled as (×) in Fig. 2.2) lie in the expected regions of parameter space.

We find that $\omega_c^{(a)}$ and $\omega_c^{(b)}$ become insensitive to the energetic bias in the limit $\lambda \gg \epsilon$. Since the reorganization processes dominate the incoherent decay in this limit, the fluctuations induced by the energetic bias becomes less important here. Hence, the boundaries of accuracy of the Padé–resummed GQME approach do not change when system–bath coupling becomes very large.

### 2.4.4 Temperature Dependence

In general, the Padé–resummed GQME approach becomes less accurate for lower temperature baths. Fig. 2.3 shows that, as the temperature decreases, the critical frequencies increase significantly throughout the entire range of reorganization energies. This may be explained by the fact that the bath degrees of freedom progressively populate lower frequency modes as temperature decreases, rendering $\Delta$ relatively larger with respect to the participating low–frequency modes. However, the Padé approach can still properly capture the dynamical effect of the bath and yield qualitatively reasonable results in the semi–quantitatively accurate region.

In the regions of lower accuracy, the Padé approach tends to overestimate the coherent oscillations. In addition, the coherent oscillations are generally shifted toward lower frequencies. In addition, we observe spurious recoherence in the panel labeled $(\lambda, \omega_c)/\Delta = (1.0, 0.6)$ for $k_B T = 0.2\Delta$. Once again the most sever deviations from exact calculation are found in the region $\omega_c < \omega_c^{(b)}(\lambda)$ as expected.

The value $(\lambda, \omega_c)/\Delta = (1.0, 0.3)$ of Fig. 3(c) of Ref. 135 is labeled $(\odot)$ in panel (c) of Fig. 2.3. However, note that the values of the energetic bias are different in this comparison. As discussed above, we expect both critical frequencies to increase in the low $\lambda$ region as the value of bias grows. Hence, we infer by this trend that the value of $(\lambda, \omega_c)/\Delta$ in the biased case should lie in the region of parameter space where the Padé approach is expected to be unreliable.

## 2.5 Conclusions

In this chapter we provide criteria to estimate the accuracy and applicability of the nonequilibrium Padé–resummed GQME approach to dissipative quantum dynamics. For

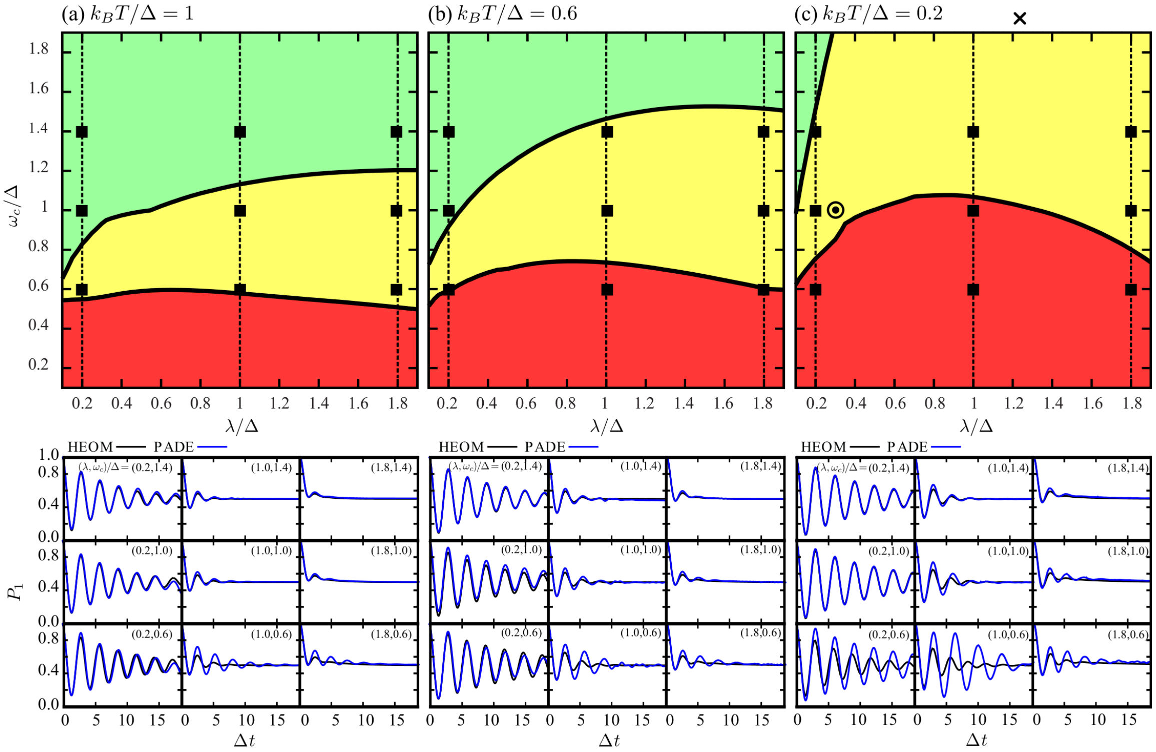


Figure 2.3: Parameter space diagrams with zero bias energy ($\epsilon = 0$) as a function of decreasing temperature $k_BT/\Delta = 1.0,\ 0.6,\ 0.2$ from left to right. The critical frequencies $\omega_c^{(a)}$ and $\omega_c^{(b)}$ are indicated as functions of $\lambda$ with color regions as in Fig. 2.1. The lower panels are the corresponding population dynamics along the vertical cuts calculated by the HEOM approach (red solid line) and Padé–resummed GQME (PADE, green dash line). The upper right label in each population dynamics panel denotes the value of $(\lambda, \omega_c)/\Delta$. The symbol $\times$ in panel (c) refers to the parameters corresponding to Fig. 2(b) of Ref. 135. The symbol $\odot$ indicates the same $(\lambda, \omega_c)$ of Fig. 3(c) of Ref. 135, but with $\epsilon = 0$.

the spin–boson model, the criteria yield critical frequencies $\omega_c^{(a)}(\lambda)$ and $\omega_c^{(b)}(\lambda)$ that partition the parameter space into three distinct regions of expected accuracy. One particularly significant outcome of our analysis is the fact that the difficult intermediate coupling regime, where all energy scales are comparable, falls frequently into a region of parameter space where the Padé approach is expected to be accurate, and indeed we find that the Padé–resummed GQME can still capture significant features of population dynamics within this regime. When $\omega_c < \omega_c^{(b)}(\lambda)$, the Padé–resummed GQME is demonstrated to often exhibit spurious long–time behavior, overestimate oscillations with shifted frequencies, and display unphysical recoherence. Overall, we find that the accuracy of the Padé resummation is relatively insensitive to the system bias and reorganization energy, but becomes worse with decreasing bath frequency and decreasing temperature.

The criteria should be generally applicable in the larger reorganization energy regime than we present here. In fact, the NIBA approach is capable of producing quantitatively accurate results in the "golden rule" regime where the reorganization energy is sufficiently larger than the diabatic coupling ($\lambda \gg \Delta$). In addition, Fig 1 (a) and (b) of Ref. 135 show that Padé GQME approach does capture the dynamics well for large $\lambda/\Delta$. Therefore, we expect the asymptotic behavior of the Padé GQME approach to be as good as or better than the NIBA approach in this regime.

The criteria of accuracy we propose is crude for several reasons. First, it is only based on the analytic properties of the first two moments of an infinite expansion. Second, even with regard to these moments, we merely search for the boundaries in the complex plane where a *single* pole may obviate physical properties required of generic memory functions. In this sense, the boundaries of accuracy are conservative and we expect to see cases where the Padé approach may still yield accurate results even if $\omega_c^{(b)}(\lambda) < \omega_c < \omega_c^{(a)}(\lambda)$ and even occasionally when $\omega_c < \omega_c^{(b)}(\lambda)$ . Indeed, we do find cases where exact calculations demonstrate that the approximate results may be more accurate than expected. However, overall we find that the trends predicted by the criteria of Sec. 2.3 faithfully delineate the trends of accuracy of the Padé–resummed generalized master equation approach.

The proposed criteria should be valid for Padé resummations used to approximate the memory kernels produced by other types of projection operators, and our applicability analysis may provide guidelines for assessing the domain of validity of other resummation techniques. In particular, one can construct applicability phase diagrams for other theories, such as the Landau–Zener resummation, leading to an increased understanding of the domain of validity of complimentary approaches. This line of investigation will be taken up in future work.

# Chapter 3

# Surface Hopping Algorithm for Condensed Phase Non-adiabatic Dynamics *

## 3.1 Introduction

Electronically non-adiabatic transitions lie at the heart of some of the most important dynamical processes in the physical sciences[147]. Phenomena ranging from gas phase atomic and molecular collisions[148] to electron and energy transfer in the condensed phase[2] are often intimately influenced by the coupling between distinct potential energy surfaces that is induced by nuclear motion. Theoretically, the accurate treatment of non-adiabatic dynamics

---

*Based on work published in J. Chem. Phys. **144**, 094104 (2016). Copyright 2016, American Institute of Physics.

is challenging, in particular in condensed phase applications where the interplay between the large number of nuclear degrees of freedom with multiple coupled electronic states greatly increases the complexity of the problem. Exact results may be obtained for specific idealized models such as spin–boson systems where potential energy surfaces are harmonic and linearly displaced[93, 95, 96, 144]. In more realistic condensed phase situations, exact solutions are currently out of reach, despite much recent progress[67].

Among the myriad approximate methods for treating non-adiabatic dynamics, the surface hopping approach[64–66, 149] stands out for several reasons. First, the method is equally applicable to gas phase and condensed phase problems, and can be used to treat realistic anharmonic nuclear motion on potential energy surfaces, albeit in a classical manner[70–72]. Surface hopping has the advantage that it is naturally formulated in the adiabatic picture, so that it can conveniently be employed in conjunction with electronic structure calculations. The method is also inexpensive, non-perturbative, and provides a superior description of branching processes and detailed balance when compared to other approaches, such as the Ehrenfest method[150]. Despite these appealing features, surface hopping naturally suffers from several deficiencies[151–154]. Clearly the description of nuclear motion as classical renders the approach incapable of capturing low temperature effects such as nuclear tunneling on a single potential surface[67–69]. More generally, while surface hopping does not employ perturbation theory in any parameter, as with nearly all mixed quantum-classical approaches to non-adiabatic dynamics, it cannot be *systematically* derived from an exact starting point[155, 156].[†] This fact makes it difficult to evaluate surface hopping's domain

---

[†] Clearly, Ref. 155 and 156 outline steps towards a complete derivation of the FSSH algorithm starting from the exact equation of motion for the density matrix. In each case, however, there is at least one step that needs to be assumed for which the domain of validity is difficult to assess. It is in this strict sense that

of validity.

One long recognized shortcoming of surface hopping is the fact that, in its standard implementation, the algorithm does not provide decoherence for electronic amplitudes. This knowledge has led to the development of important modifications of surface hopping aimed at more accurately describing decoherence[73–86]. In an important recent series of studies, Landry and Subotnik showed that a striking consequence of the neglect of decoherence in surface hopping is the failure to properly capture the golden-rule scaling of the non-adiabatic transfer rate in the Marcus regime[157, 158]. It should be noted, however, that there has been some debate as to just how pervasive this problem is[159–163]. One of the goals of this chapter is to provide an in depth examination of this issue.

More broadly we aim to compare surface hopping, with and without corrections for decoherence, to exact calculations in a model condensed phase system, namely the spin–boson model[48, 164]. Although the spin–boson model is an idealized proxy for a real condensed phase system exhibiting non-adiabatic transitions, it offers the advantage that algorithms now exist that enable the calculation of exact dynamics over a wide swath of the relevant parameter space[93, 95, 96, 144]. While in the past surface hopping was compared to exact benchmark calculations of low dimensional scattering problems[65, 66, 83], we now can provide guidelines for understanding the successes and failures of the surface hopping approach in a broader condensed phase setting. It should be noted, however, that we will restrict our comparison to the "overdamped" case of the coupling to a Debye spectral density, since it is here that facile exact simulations may be performed. While the Debye case represents perhaps the most commonly employed model of a condensed environment in the spin–boson

---

we refer to the lack of a systematic derivation of FSSH.

context, our choice implies that some aspects related to the interplay between surface hopping trajectories and decoherence which are expected to be most dramatic and subtle in the underdamped limit, may not arise[79, 163, 165]. Regardless, this chapter should at least provide a starting point for assessing how surface hopping performs in generic condensed phase settings.

Our chapter is organized as follows: We begin in Sec. 3.2 with a review of the standard surface hopping algorithm for the spin–boson model and various formulations of decoherence corrections. In Sec. 3.3, we present our results for the scaling of the non-adiabatic transfer rate with respect to the electronic coupling in the golden-rule regime. In Sec. 3.4, we explore the full parameter space of spin–boson model. We summarize our results and conclude in Sec. 3.5.

## 3.2 Fewest Switch Surface Hopping (FSSH) and Decoherence

### 3.2.1 Spin-Boson Model

We consider the spin–boson model, $H = H_s + H_b + H_{sb}$, which describes a two-level system with energy bias $\epsilon$ and constant diabatic coupling $\Delta$

$$H_s = \epsilon\sigma_z + \Delta\sigma_x, \tag{3.1}$$

interacting with an infinite set of harmonic oscillators (bath)

$$H_b = \sum_j \frac{1}{2}\left(p_j^2 + \omega_j^2 q_j^2\right), \tag{3.2}$$

where $\omega_j$ is the frequency of the $j$-th bath mode[3]. The isolated electronic system and the bath are coupled bilinearly

$$H_{sb} = \sigma_z \sum_j g_j q_j, \tag{3.3}$$

where $g_j$ is the coupling strength between the two-level system and the $j$-th harmonic oscillator. We adopt the Pauli matrix notation $\sigma_x = |1\rangle\langle 2| + |2\rangle\langle 1|$ and $\sigma_z = |1\rangle\langle 1| - |2\rangle\langle 2|$ where $|i\rangle$ indicate the diabatic states of the system. Throughout the present chapter, we use mass scaled coordinates and momenta for the bath modes, $q_j = \sqrt{M_j} Q_j$ and $p_j = P_j/\sqrt{M_j}$, where $M_j$ are the effective mass of nucleus for the $j$-th harmonic oscillator and set $\hbar = 1$. We denote bold letters $\boldsymbol{q}$, $\boldsymbol{p}$ by the vector of nuclear degrees of freedom.

The influence of the bath on the dynamics of the system can be captured in the compact form of a spectral density,

$$J(\omega) = \frac{\pi}{2} \sum_j \frac{g_j^2}{\omega_j} \delta(\omega - \omega_j). \tag{3.4}$$

In the present chapter, we consider the Debye model of the spectral density[144],

$$J(\omega) = \frac{\lambda}{2} \frac{\omega \omega_c}{\omega^2 + \omega_c^2}, \tag{3.5}$$

which is appropriate for the description of a solvent environment with Debye dielectric relaxation. The Debye spectral density function is characterized by two parameters, the reorganization energy $\lambda$, and the characteristic bath frequency $\omega_c$. In electron-transfer theory, the reorganization energy represents a direct measure of the coupling strength between the system and the bath. The characteristic frequency is related to the relaxation time scale of the bath, $\tau = 1/\omega_c$. The Debye spectral density spans broader frequency than the standard Ohmic ($J(\omega) \propto \omega e^{-\omega/\omega_c}$) and Brownian forms ($J(\omega) \propto \omega/((\omega^2 - \omega_c^2)^2 + \gamma^2\omega^2)$). Following the procedure outlined in Refs. 166 and 87, it is convenient to discretize the

Debye spectral density function via $\omega_j = \tan\left((j-0.5)\tan^{-1}\left(\omega_{\max}/\omega_c\right)/N\right)$ where $\omega_{\max}$ is the largest frequency and $N$ is the number of oscillators employed in the discretization.

The population dynamics of the spin–boson model can be calculated by the numerically exact hierarchical equations of motion (HEOM) methodology[93], implemented in the Parallel Hierarchy Integrator (`PHI`) [94]. The HEOM method is easier to use when the spectral density take the Debye form so that the bath correlation function can be written as a sum of exponentially decaying functions in time[136, 167]. We use the HEOM method to produce all of our benchmark results for the spin–boson model.

We focus on the reduced population dynamics of the system

$$P_i\left(t\right) = \mathrm{Tr}_b\left\{\rho\left(0\right)e^{iHt}\left|i\right\rangle\left\langle i\right|e^{-iHt}\right\} \tag{3.6}$$

where we assume a factorized initial condition $\rho\left(0\right) = \rho_b\left|1\right\rangle\left\langle 1\right|$ and

$$\rho_b = \frac{e^{-\beta H_b}}{\mathrm{Tr}_b\left\{e^{-\beta H_b}\right\}}, \tag{3.7}$$

with the inverse temperature of the bath, $\beta = 1/kT$. The initial condition of the system corresponds to an impulsive Franck-Condon transition with the bath in a state independent of the system with oscillators centered at $q_j = 0$.

### 3.2.2 FSSH and Its Variants

The fewest-switches surface hopping (FSSH) algorithm is a mixed quantum-classical method that treats the bath degrees of freedom classically and the electronic system quantum mechanically[64–66]. A swarm of classical nuclear trajectories evolve on the adiabatic potential energy surfaces associated with the electronic states with each *individual* trajectory evolving on a single *active* surface. Along each trajectory, the electronic wave function propagates according to the Schrodinger equation with the classical nuclear variables evolving as parameters. The essence of FSSH is to simulate the population of the electronic states via the density of trajectories on each surface. For this purpose, a surface-hopping scheme is introduced to allow trajectories to hop among the adiabatic energy surfaces and match the electronic populations. The hopping probability of the classical bath trajectories depends on the electronic wave functions with specific conditions for the acceptance of non-adiabatic transitions. Instead of listing these conditions, we describe them within the context of the spin–boson model.

To implement the FSSH algorithm for the spin–boson model, we transform the model to its adiabatic representation by diagonalizing the Hamiltonian $H|\Phi_i(\boldsymbol{q})\rangle = (\frac{\boldsymbol{p}^2}{2} + V_i(\boldsymbol{q}))|\Phi_i(\boldsymbol{q})\rangle$ where

$$V_i(\boldsymbol{q}) = \frac{1}{2}\sum_j \omega_j^2 q_j^2 + (-1)^i \sqrt{(\boldsymbol{g}\cdot\boldsymbol{q}+\epsilon)^2 + \Delta^2} \tag{3.8}$$

are the adiabatic potential energy surfaces and $\boldsymbol{g}\cdot\boldsymbol{q} = \sum_j g_j q_j$. One may transform the diabatic states to the adiabatic representation via the unitary transformation $|\Phi_i(\boldsymbol{q})\rangle = \sum_j U_{ij}(\boldsymbol{q})|j\rangle$ where

$$U(\boldsymbol{q}) = \begin{pmatrix} \sin\theta(\boldsymbol{q}) & -\cos\theta(\boldsymbol{q}) \\ \cos\theta(\boldsymbol{q}) & \sin\theta(\boldsymbol{q}) \end{pmatrix}. \tag{3.9}$$

The adiabatic-diabatic mixing angle is defined as $\theta(\boldsymbol{q}) = \frac{1}{2}\tan^{-1}(\Delta/(\boldsymbol{g}\cdot\boldsymbol{q}+\epsilon))$ which depends on the bath coordinates. Within the adiabatic representation, the electronic wavefunction can be written as $|\Psi(t)\rangle = c_1(t)|\Phi_1(\boldsymbol{q})\rangle + c_2(t)|\Phi_2(\boldsymbol{q})\rangle$ and the adiabatic amplitudes satisfy an implicit time-dependent Schrodinger equation

$$\frac{d}{dt}c_i\left(t\right) = -iV_i\left(\boldsymbol{q}\right)c_i\left(t\right) - \sum_k \boldsymbol{p}\cdot\boldsymbol{d}_{ik}\left(\boldsymbol{q}\right)c_k\left(t\right), \tag{3.10}$$

where $d_{ik}^j \equiv \left\langle \Phi_i\left(\boldsymbol{q}\right)\right| \frac{d}{dq_j}\left|\Phi_k\left(\boldsymbol{q}\right)\right\rangle$ is the derivative coupling matrix. For the spin–boson model, the derivative coupling matrix elements are $d_{11}^j = d_{22}^j = 0$ and

$$d_{12}^j = -d_{21}^j = \frac{g_j}{2}\frac{\Delta}{\left(\boldsymbol{g}\cdot\boldsymbol{q}+\epsilon\right)^2+\Delta^2}. \tag{3.11}$$

We define the pure state electronic density matrix $\hat{\sigma}$ by $\sigma_{ik} = c_i c_k^*$ and the equivalent equation for the density matrix can be written as

$$\frac{d}{dt}\hat{\sigma}\left(t\right) = -i\left[\hat{V}\left(\boldsymbol{q}\right),\hat{\sigma}\left(t\right)\right] - \left[\boldsymbol{p}\cdot\hat{\boldsymbol{d}}\left(\boldsymbol{q}\right),\hat{\sigma}\left(t\right)\right], \tag{3.12}$$

where the potential energy matrix is $V_{ik}\left(\boldsymbol{q}\right) = \delta_{ik}V_i\left(\boldsymbol{q}\right)$.

The bath in the FSSH algorithm is described via a swarm of trajectories evolving classically on adiabatic potential surfaces. Each individual trajectory propagates on the *active* adiabatic potential surface, $V_a\left(\boldsymbol{q}\right)$, via $\dot{\boldsymbol{q}} = \boldsymbol{p}$ and $\dot{\boldsymbol{p}} = -\partial V_a/\partial \boldsymbol{q}$, and the bath configuration is followed by monitoring the time-dependence of $\left(\boldsymbol{q}^{(n)},\boldsymbol{p}^{(n)},a^{(n)}\right)$ for $n = 1,\cdots,N_{\text{traj}}$. Each trajectory is allowed to switch active surfaces in order to force the relative number of trajectories on each surface to mimic the adiabatic probability calculated by the adiabatic amplitudes. To accomplish this, a minimal switching probability for a hop from surface $a$ (active) to surface $b$ (other) during each time step $dt$ may be employed as[65]

$$\gamma_{ab}^{\text{hop}} = dt\frac{2}{\left|c_a\right|^2}\left[\text{Im}\left(V_{ba}\left(\boldsymbol{q}\right)c_a c_b^*\right) + \text{Re}(\boldsymbol{p}\cdot\boldsymbol{d}_{ab}c_a c_b^*)\right]. \tag{3.13}$$

For the spin–boson model, the hopping probability is determined entirely by the derivative coupling and the adiabatic coherence $c_a c_b^*$. In addition to the hopping probability, trajectories must have enough energy to hop to a new surface and obey energy conservation. If the trajectory switches to a new active surface, the momentum is rescaled in the direction of the derivative coupling by $\boldsymbol{p}' = \boldsymbol{p} + \kappa \boldsymbol{d}_{ab}$ satisfying $|\boldsymbol{p} + \kappa \boldsymbol{d}_{ab}|^2 + 2V_b(\boldsymbol{q}) = |\boldsymbol{p}|^2 + 2V_a(\boldsymbol{q})$.

At time $t = 0$, we require that the initial configuration of the bath mimics the initial electronic density in the adiabatic representation. The initial configuration for the bath modes are sampled from the thermal Wigner distribution, $\rho_b \propto \exp\{-\sum_j \frac{2}{\omega_j} \tanh(\frac{\beta\omega_j}{2})(\frac{1}{2}p_{0j}^2 + \frac{1}{2}\omega_j^2 q_{0j}^2)\}$, with the trace over the bath approximated as $\mathrm{Tr}_b\{\rho_b \cdots\} \approx \frac{1}{N_{\mathrm{traj}}} \sum_{(\boldsymbol{q}_0, \boldsymbol{p}_0)}^{w} \cdots \equiv \langle \cdots \rangle$. In addition, we initialize the active configuration $a^{(n)}$ accordingly by distributing the initial phase terms on surface 1 with the probability $|c_1(0)|^2$ and on surface 2 via probability $|c_2(0)|^2$.

Given that the electronic amplitudes are propagated in the adiabatic representation and the bath trajectories move along adiabatic energy surfaces according to the FSSH algorithm, it is non-trivial to extract diabatic electronic populations. We adopt the interpretation of mixed quantum-classical density matrix[168] for the diabatic population on state $i$, which is given by

$$P_i = \left\langle \sum_j |U_{ij}(\boldsymbol{q})|^2 \delta_{ja} + \sum_{j<k} 2\mathrm{Re}\left[U_{ij}(\boldsymbol{q})\sigma_{jk}U_{ik}^*(\boldsymbol{q})\right] \right\rangle. \tag{3.14}$$

Note that the expression for $P_i$ includes information from the active surface $(a)$ as well as the adiabatic amplitude $(\sigma_{jk})$. For the spin–boson model, we can express the reduced population dynamics of state 1 as

$$\begin{aligned} P_1 = &\left\langle \sin^2\theta(\boldsymbol{q})\,\delta_{1a} + \cos^2\theta(\boldsymbol{q})\,\delta_{2a} \right\rangle \\ &+ \left\langle 2\sin\theta(\boldsymbol{q})\cos\theta(\boldsymbol{q})\,\mathrm{Re}\left[c_1 c_2^*\right]\right\rangle, \end{aligned} \tag{3.15}$$

which is composed of a portion associated with the active surface and a portion contributed by the adiabatic coherence.

### 3.2.3 Decoherence

Within the standard FSSH algorithm, a difficulty arises when a trajectory passes through the coupling region and the electronic wavefunction may bifurcate on different surfaces[65, 74, 169, 170]. Before the bifurcation event, each FSSH trajectory carries a particular electronic amplitude. After the trajectory passes through the coupling region, the wavefunction retains its phase and the density matrix remains pure, even if the trajectories are separated on different surfaces. This failure to incorporate decoherence may lead to an inaccurate description of electronic dynamics.

The augmented FSSH (A-FSSH) [171] has been proposed to resolve this problem by collapsing the electronic state on the inactive surfaces and projecting onto the active surface according to a decoherence rate calculated on the fly. The full procedure of the A-FSSH algorithm is outlined in Ref. 171. Here, for completeness, we briefly review the A-FSSH scheme.

The decoherence rate depends on the matrix of augmented moments of the bath coordinate and momentum $(\delta\hat{\boldsymbol{q}}, \delta\hat{\boldsymbol{p}})$ which provide information regarding the separation of a proxy wave packet in phase space. The augmented moments evolve along a trajectory which follows the equations of motion

$$\frac{d}{dt}\delta\hat{q}_j = \hat{T}_j^q - T_{j,aa}^q\hat{I}, \tag{3.16}$$

$$\frac{d}{dt}\delta\hat{p}_j = \hat{T}_j^p - T_{j,aa}^p\hat{I}, \tag{3.17}$$

where $\hat{T}^q$ and $\hat{T}^p$ are obtained by expanding the full quantum Liouville equation to first order in $\hbar$ (linearized approximation)

$$\hat{T}^q_j \equiv -i\left[\hat{V}, \delta\hat{q}_j\right] + \delta\hat{p}_j - \sum_k p_k \left[\hat{d}_k, \delta\hat{q}_j\right], \tag{3.18}$$

$$\hat{T}^p_j \equiv -i\left[\hat{V}, \delta\hat{p}_j\right] + \frac{1}{2}\left\{\delta\hat{F}_j, \hat{\sigma}\right\} - \sum_k \hat{p}_k \left[\hat{d}_k, \delta\hat{p}_j\right], \tag{3.19}$$

and the matrix of forces is given by $\hat{F}_j \equiv -\partial\hat{V}/\partial q_j|_{\boldsymbol{q}}$ and $\delta\hat{F}_j = \hat{F}_j - F_{j,aa}\hat{I}$. Via the augmented moments, one can derive the off-diagonal correction to the equation of motion for the reduced electronic density matrix,

$$\frac{d}{dt}\hat{\sigma} = -i[\hat{V}, \hat{\sigma}] - \left[\boldsymbol{p}\boldsymbol{\cdot}\hat{\boldsymbol{d}}, \hat{\sigma}\right] + i\left[\hat{\boldsymbol{F}}, \delta\hat{\boldsymbol{q}}\right], \tag{3.20}$$

which incorporates the decoherence mechanism in the last term. The estimated decoherence rate for the separation of wavepackets on the active surface $a$ and the inactive surface $b$ is of the form

$$\gamma^{\mathrm{d}}_{ba} = dt\left\{\frac{(\boldsymbol{F}_{bb} - \boldsymbol{F}_{aa}) \cdot \delta\boldsymbol{q}_{bb}}{2} - 2\left|\boldsymbol{F}_{ab} \cdot \delta\boldsymbol{q}_{bb}\right|\right\}, \tag{3.21}$$

which is obtained by assuming frozen Gaussian wave packets for the bath wavefunction outside of the derivative coupling region ($\hat{\boldsymbol{d}} = 0$) and reducing the decoherence rate for non-zero derivative couplings. The A-FSSH algorithm also permits resetting the augmented moments to avoid the failure of the linearized approximation. The proposed reset rate is given by bifurcate

$$\gamma^{\mathrm{r}}_{ba} = -dt\frac{(\boldsymbol{F}_{bb} - \boldsymbol{F}_{aa}) \cdot \delta\boldsymbol{q}_{bb}}{2}. \tag{3.22}$$

Note that $\gamma^{\mathrm{r}}_{ba}$ is the negative collapsing rate since the moments become invalid when wavepackets aggregate.

A more traditional approach to decoherence corrections within surface hopping consists of damping the coherence of the density matrix via a pure-dephasing-like rate[79, 80, 86].

Within this simpler density-matrix approach, we treat the evolution of the adiabatic coherence outside the derivative coupling region ($\hat{\boldsymbol{d}} = 0$) as pure dephasing in a stochastic formulation[172]. Inside the zero derivative coupling region, the population transfer is excluded and the adiabatic coherences satisfy $\frac{d}{dt}\sigma_{jk} = -i(V_j - V_k)\sigma_{jk}$ and the formal solution is $\sigma_{jk}\left(t+\tau\right) = \sigma_{jk}\left(t\right)\langle\exp\{-i\int_t^{t+\tau} dt'[V_j\left(t'\right) - V_k\left(t'\right)]\}\rangle_w$. The pure-dephasing time within this stochastic formulation is obtained via the energy difference correlation function[172]

$$\frac{1}{T_2^*} = \frac{1}{2}\int_0^\infty \left\langle\left[V_j\left(t'\right) - V_k\left(t'\right)\right]\left[V_j\left(0\right) - V_k\left(0\right)\right]\right\rangle dt'. \tag{3.23}$$

To simulate the decay of the adiabatic coherence within the FSSH algorithm, we introduce a decoherence terms that leads to an exponential decay of the adiabatic coherences. In particular, decoherence is modeled as a Poisson process with the probability that a coherence decay occurs in the time interval $[t, t+dt]$ gives by $\mathrm{Prob}\{N[\sigma_{jk}\left(t\right)] - N[\sigma_{jk}\left(t+dt\right)] = 1\} = e^{-dt/T_2^*}dt/T_2^* \approx dt/T_2^*$ where $N[\sigma_{jk}\left(t\right)]$ is the number of trajectories whose density matrix retains coherence. However, for the spin–boson model, estimation of $T_2^*$ along each trajectory via Eq. (3.23) is not well defined. To circumvent this problem, we assume the decoherence time scale takes a similar form for each trajectory

$$\frac{1}{\tau_{jk}\left(t\right)} = \frac{1}{2}\int_0^t \left(V_j\left(t'\right) - V_k\left(t'\right)\right)\left(V_j\left(0\right) - V_k\left(0\right)\right)dt' \tag{3.24}$$

which gives an estimate of the pure-dephasing time outside of the derivative coupling region. The decoherence rate for the off-diagonal term $\sigma_{jk}$ is then given by

$$\gamma_{jk}^{\mathrm{d}}\left(t\right) = \frac{dt}{\tau_{jk}\left(t\right)}. \tag{3.25}$$

A decoherence factor for the off-diagonal density matrix elements may be defined as $\sigma_{jk} =$

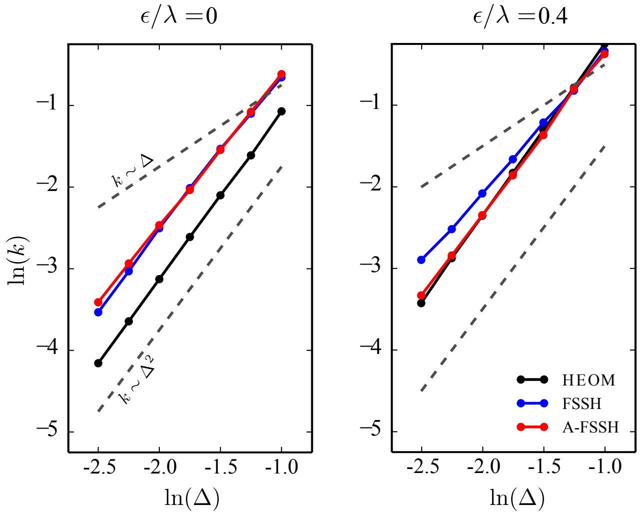


Figure 3.1: Diabatic population transfer rates ($k$) as a function of diabatic coupling $\Delta$ for FSSH (blue), A-FSSH (red), and HEOM (black) in the unbiased $\epsilon/\lambda = 0$ and biased $\epsilon/\lambda = 0.4$ cases. The bath temperature is assumed to be in the classical limit, $T = 300$ K. The reorganization energy is $E_r = 520\ \text{cm}^{-1}$, while the bath frequency scale $\omega_c$ is tuned so that $\Delta/\omega_c \ll 1$. The dashed lines are reference markers of sub-quadratic and quadratic dependence, respectively. The diabatic population transfer rates is extracted from the population dynamics by exponential fitting.

$\eta_{jk} c_j c_k^*$, so that the hopping rate, namely the analogy of Eq. (3.13), becomes

$$\gamma_{ab}^{\text{hop}} = dt \frac{2}{|c_a|^2} \text{Re}(\boldsymbol{p} \cdot \boldsymbol{d}_{ab} \sigma_{ab}). \tag{3.26}$$

For every time step, we calculate the decoherence timescale $\tau_{jk}(t)$ by accumulating energy difference correlations along the trajectory. If a decoherence event occurs, the associated factor $\eta_{jk}$ is set to zero. Then we symmetrize the density matrix and continue the trajectory propagation.

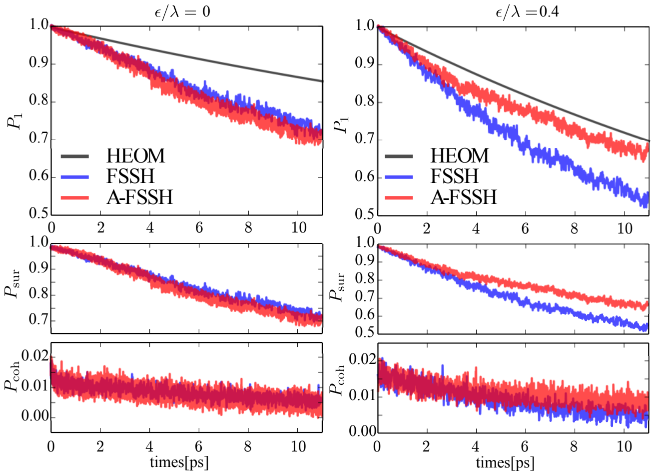


Figure 3.2: Population dynamics of the for FSSH, A-FSSH, and HEOM in the (a) unbiased $\epsilon/\lambda = 0$ and (b) biased $\epsilon/\lambda = 0.4$ cases. The lower panels show the surface and coherence terms separately, as defined in Eq. (3.33). The bath temperature is $T = 300$ K, the reorganization energy is $\lambda = 520$ cm$^{-1}$, and $\omega_c = 85$ cm$^{-1}$.

## 3.3 The Golden-Rule Regime

A surprising feature of the standard FSSH algorithm that has recently been discovered is its failure to capture the quadratic dependence of the (diabatic) transfer rate in the weak coupling regime. The generality of this behavior remains somewhat controversial. Furthermore, the fundamental origin of this apparent failure is unclear. Landry and Subotnik noted an interesting aspect of the simple one-dimensional Landau-Zener (LZ) problem[157]. In the standard treatment of the LZ problem with initial electronic population on one surface only, a single voyage through the crossing region produces population differences in harmony with the expected quadratic coupling dependence of the rate. However, if the system is prepared initially with arbitrary population on *both* diabatic surfaces, then a passage through the crossing point induces a population change that is proportional to *both* the electronic coupling itself as well as its square. It may then be argued that since traversal of the crossing region mixes the populations, multiple crossings will produce a rate with a sub-quadratic coupling dependence. With the addition of decoherence, however, populations are localized after each crossing, such that the rate always retains its proper quadratic golden-rule form. Below we show that while this argument *cannot* explain the deviations from Marcus golden-rule behavior exhibited by FSSH, the notion that decoherence can alter the coupling dependence of the rate in a favorable way is indeed correct.

Let us briefly revisit the simple one dimensional LZ example. As in Ref. 157, let us take an electronic propagator of the form

$$U = \begin{pmatrix} \sqrt{\xi} & \sqrt{1-\xi} \\ -\sqrt{1-\xi} & \sqrt{\xi} \end{pmatrix}, \tag{3.27}$$

where

$$\xi = \exp\left[-\frac{2\pi}{\hbar}\frac{\Delta^2}{|\boldsymbol{v}\cdot(\boldsymbol{F_1}-\boldsymbol{F_2})|}\right] \equiv \exp\left[-\eta\Delta^2\right] \tag{3.28}$$

is the LZ parameter which depends on the crossing velocity $\boldsymbol{v}$ and the difference in the (diabatic) forces, $\boldsymbol{F_1}-\boldsymbol{F_2}$, at the crossing point, and the electronic coupling, $\Delta$. Clearly a pure initial wave packet with amplitude placed entirely on surface $a$, namely

$$P(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3.29}$$

produces a population difference on surface 2 after one crossing that is proportional to $\Delta^2$

for small $\Delta$. On the other hand, if the initial packet has the form

$$P(0) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \tag{3.30}$$

where $\alpha$ and $\beta$ are arbitrary constants satisfying $\alpha^2 + \beta^2 = 1$, then after one passage the population difference on the surface $b$ is given by $(\xi - 1)\beta^2 + (1 - \xi)\alpha^2 - 2\sqrt{\xi(1-\xi)}\alpha\beta \approx (\alpha^2 - \beta^2)\,\eta\Delta^2 - \sqrt{\eta}\Delta\alpha\beta$. The linear term in the electronic coupling heralds an apparent subquadratic dependence of the rate on $\Delta$. Importantly, however, it should be noted that the mixing of populations that occurs during passage through the crossing region depends on $\Delta$. In particular, starting from the "pure" initial state

$$P(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3.31}$$

passage through the crossing region produces populations on each diabatic state that are non-zero, but do depend on $\Delta$ and are thus *not* arbitrary constants. Via consideration of

$$P(n) = U^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3.32}$$

it is straightforward to demonstrate that even in the absence of decoherence, multiple crossings do not generate spurious terms in the $b$-state population that are linear in $\Delta$ within this simple model.

To explore the issue of the behavior predicted by surface hopping in the Marcus regime, we turn to direct simulation. In Fig. 3.1, compare the exact diabatic population transfer rates, numerically extracted from HEOM simulations in the high temperature, weak electronic coupling regime to both the results predicted by FSSH as well as the decoherence based

A-FSSH algorithm. In both cases, we use Eq. (3.15) to extract diabatic quantities. The exact HEOM simulations are not confined to the strict high temperature limit. Thus we expect rates that scale as $\Delta^2$, but do not necessarily conform quantitatively to standard Marcus theory. The results are shown for both an unbiased and strongly biased cases of the spin–boson problem. Several important features should be noted. First, in the symmetric situation, the FSSH approach yields the correct scaling of the rate with $\Delta$ and produces results that are essentially indistinguishable from those of A-FSSH. This is true even as the electronic coupling is varied over a wider range, and for all values of the reorganization energy. On the other hand, when there is a sizable energetic bias, the rate indeed violates Marcus scaling and behaves in a manner qualitatively similar to that described in Ref. 171.[‡] *Importantly, however, the magnitude of the deviations we find are significantly smaller than that expected from the calculations of Ref. 171.* Remarkably, the inclusion of decoherence corrects this failing, producing results in quantitative correspondence with exact numerics. Thus, violations of the expected golden-rule behavior as well[ as the impact of decoherence in the weak-coupling regime appear to depend sensitively on the electronic bias.

To gain a deeper understanding of this surprising result, we decompose the non-adiabatic population into terms that have an explicit dependence on the dynamics on a given surface the coherence between surfaces, respectively. It may be shown that Eq. (3.15) can be recast

---

[‡] The recently published paper, Jain and Subotnik, J. Phys. Chem. Lett. **6** , 4809 (2015), makes a nearly identical observation. We thank Joseph Subotnik for making us aware of this during the writing of this manuscript.

as

$$\begin{aligned} P_1 &= P_{\text{sur}} + P_{\text{coh}} \\ &= \left\langle \frac{1}{2} + \frac{1}{2}\frac{\boldsymbol{g}\cdot\boldsymbol{q}+\epsilon_0}{\sqrt{(\boldsymbol{g}\cdot\boldsymbol{q}+\epsilon)^2+\Delta^2}}(\delta_{2\lambda}-\delta_{1\lambda}) \right\rangle \\ &\quad + \left\langle \frac{\Delta}{\sqrt{(\boldsymbol{g}\cdot\boldsymbol{q}+\epsilon)^2+\Delta^2}}\mathrm{Re}[c_1 c_2^*] \right\rangle , \end{aligned} \tag{3.33}$$

where we have labeled the two relevant terms in Eq. (3.33) as the "surface" term, $P_{\text{sur}}$, and the "coherence" term, $P_{\text{coh}}$. Note that we are using the diabatic interpretation of Ref. 168, so in essence it is the "surface" term that is expected to be most sensitive to decoherence corrections applied in the *adiabatic* basis, not the "coherence" term. Furthermore, note that it is the surface term that has the stronger explicit dependence on the energetic bias, in harmony with the notion that the distinction between FSSH and its decoherence corrected variants will depend on bias as reflected in the way decoherence alters the behavior of the first term of Eq. (3.15). In Fig. 3.2 we show the temporal decay of population in both the unbiased and biased cases, within both FSSH and A-FSSH. We also show separately the surface and coherence terms. For the unbiased case, FSSH and A-FSSH yield essentially identical results, while in the biased case A-FSSH is in near quantitative agreement with the exact result while the standard FSSH result decays too rapidly. The difference between the two results is noticeable only in the surface term, which dominates over the coherence term. Thus, we find that distinction between the unbiased and biased cases reflects the manner in which the bias couples to coherence-sensitive terms as exposed in Eq. (3.33).

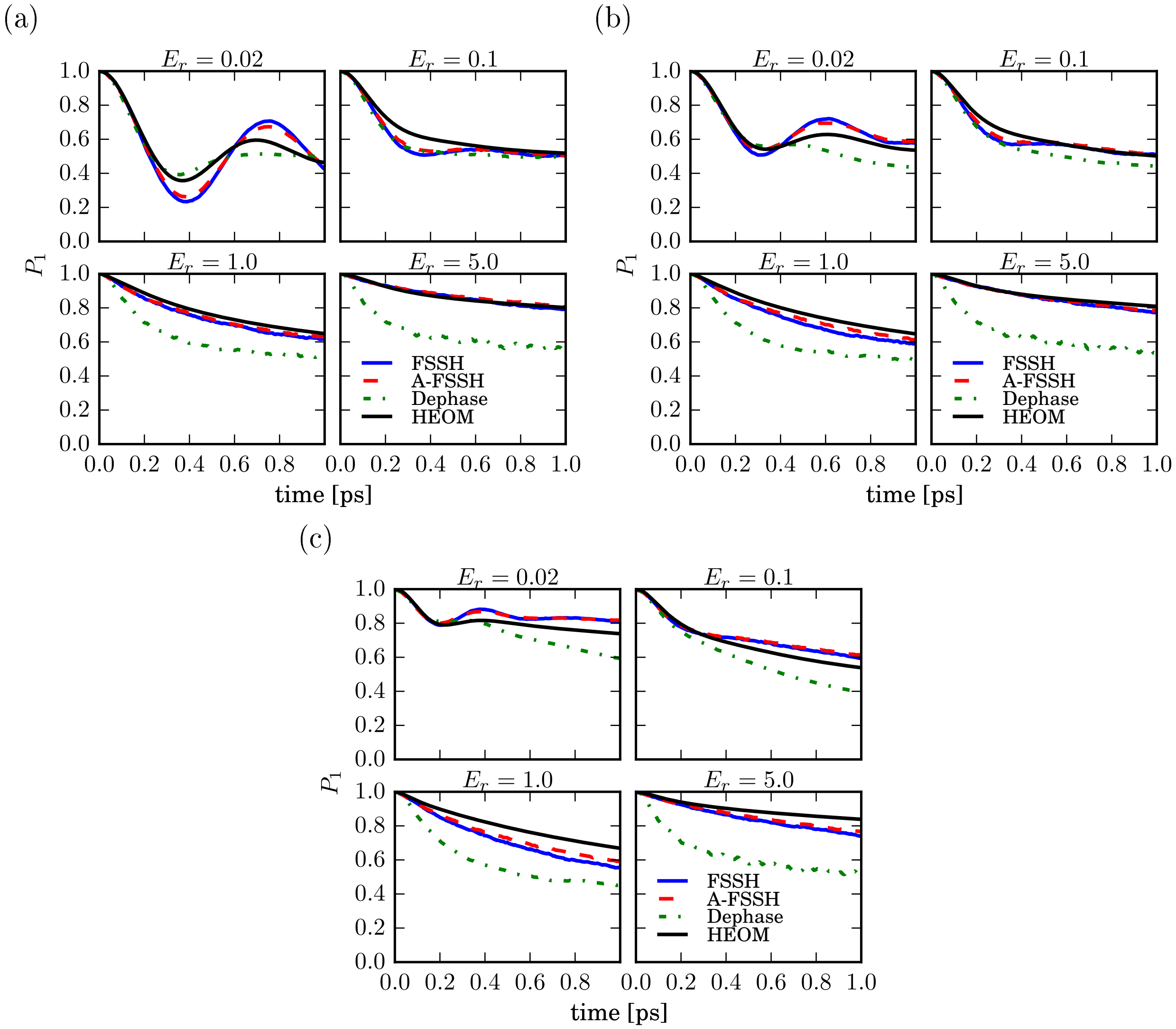


Figure 3.3: High temperature dynamics with intermediate electronic coupling strength. We employ a reference unit of energy of 104 cm$^{-1}$. Parameters are $kT = 2$ ($T = 300$ K), $\Delta = \omega_c = 0.2$ , (a)$\epsilon/\Delta = 0$, (b)$\epsilon/\Delta = 1$, and (c)$\epsilon/\Delta = 2$. Reorganization energies ($E_r = \lambda/\Delta$) are scanned from small to large and are listed on each panel. "Dephase" refers to the use of Eqs. (3.24)-(3.26).

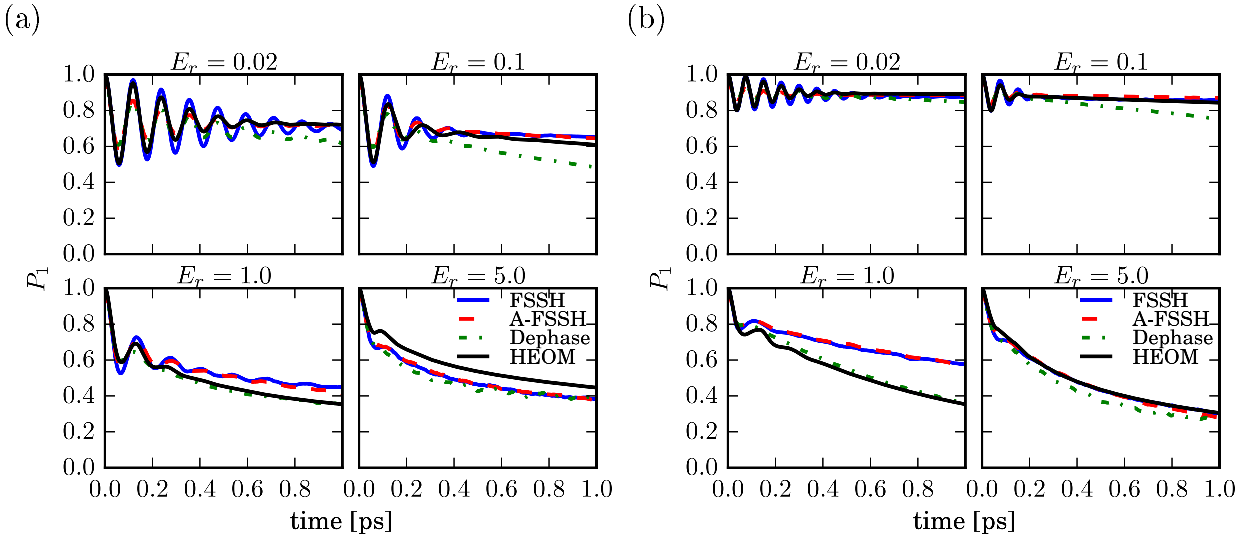


Figure 3.4: High temperature dynamics in the adiabatic regime. We employ a reference unit of energy of 104 $\text{cm}^{-1}$. Parameters are $kT = 2$ ($T = 300$ K), $\Delta = 1$, $\omega_c = 0.2$, (a)$\epsilon/\Delta = 1$ and (b)$\epsilon/\Delta = 2$. Reorganization energies ($E_r = \lambda/\Delta$) are scanned from small to large and are listed on each panel. “Dephase” refers to the use of Eqs. (3.24)-(3.26).

## 3.4 Full Parameter Space

In this section, we explore more broadly the comparison of surface hopping to benchmark calculations of dynamics in the spin–boson model. Fig. 3.2 illustrates that in the golden-rule regime, standard FSSH produces results in qualitative agreement with the exact behavior produced by HEOM calculations. The inclusion of decoherence can lead to improved and even quantitatively accurate results, however the improvement over FSSH will depends sensitively on the parameters of the underlying Hamiltonian, such as the energetic bias. Similar behavior is seen away from the weak coupling limit.

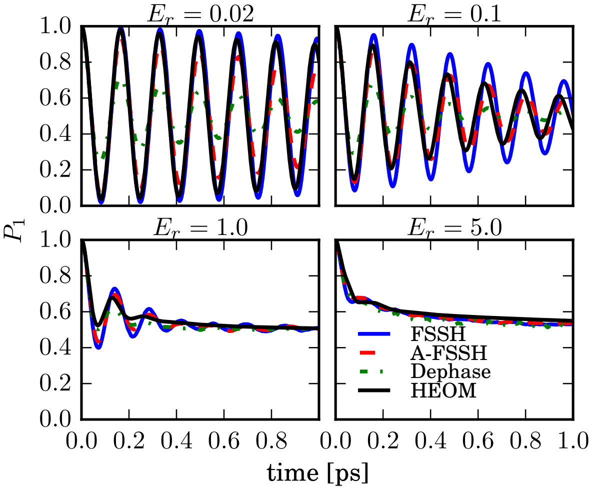


Figure 3.5: High temperature dynamics in the adiabatic regime. We employ a reference unit of energy of 104 cm$^{-1}$. Parameters are $kT = 2$ ($T = 300$ K), $\Delta = 1$, $\omega_c = 0.2$, and $\epsilon = 0$. Reorganization energies ($E_r = \lambda/\Delta$) are scanned from small to large and are listed on each panel. "Dephase" refers to the use of Eqs. (3.24)-(3.26).

### 3.4.1 High Temperature Regime

For intermediate electronic coupling and high temperature ($kT/\omega_c \gg 1$), a regime often difficult to treat via approximate perturbative approaches, we find that FSSH is quite accurate, with an accuracy that is not altered by inclusion of decoherence within the A-FSSH approach. On the other hand, direct decoherence damping with a pure-dephasing-type rate generally leads to less accurate results than FSSH in this regime, especially when the reorganization energy is large. These observations are illustrated in Fig. 3.3. In the adiabatic regime, where the electronic coupling is large, we again find that FSSH is in good agreement with the exact behavior of the simulated non-equilibrium populations at high temperatures, especially for large reorganization energies. Some select examples of this comparison are illustrated in Fig. 3.4. In situations where the reorganization energy is small and the system has no energetic bias, the upper left panel of Fig. 3.4(a) and the upper right panel of Fig. 3.5 illustrate how A-FSSH provides a damping of population oscillations that brings the approximate results into quantitative correspondence with exact simulations. With respect to more phenomenological treatments of decoherence, two new features stand out. First, direct decoherence damping with a pure-dephasing rate generally leads to more accurate results in the strong-coupling regime than it does in situations where the electronic coupling is intermediate or small as compared to other energy scales in the problem. In particular, unlike in the case of intermediate coupling, the more phenomenological treatment of decoherence appears not to lead to gross overestimates of the rate of population decay in energetically biased cases for large electronic couplings. Furthermore, we find, for the first time, examples where a simple "pure dephasing" correction leads to clearly improved accuracy over both FSSH and A-FSSH. We emphasize however that in general we find A-FSSH to be, on average, the most accurate approach across the full parameter space.

### 3.4.2 Low Temperature Regime

Lastly, we turn to situations where the temperature is comparable to, or lower than, the characteristic bath frequency. In such situations we expect any surface hopping approach to be unreliable due to the fact that the dynamics of the nuclei are treated classically. Thus processes such as nuclear quantum tunneling cannot be described. While we find this to be generally the case, there are situations where the surface hopping approaches find some success even in this "quantum bath" regime. In particular, when the electronic coupling

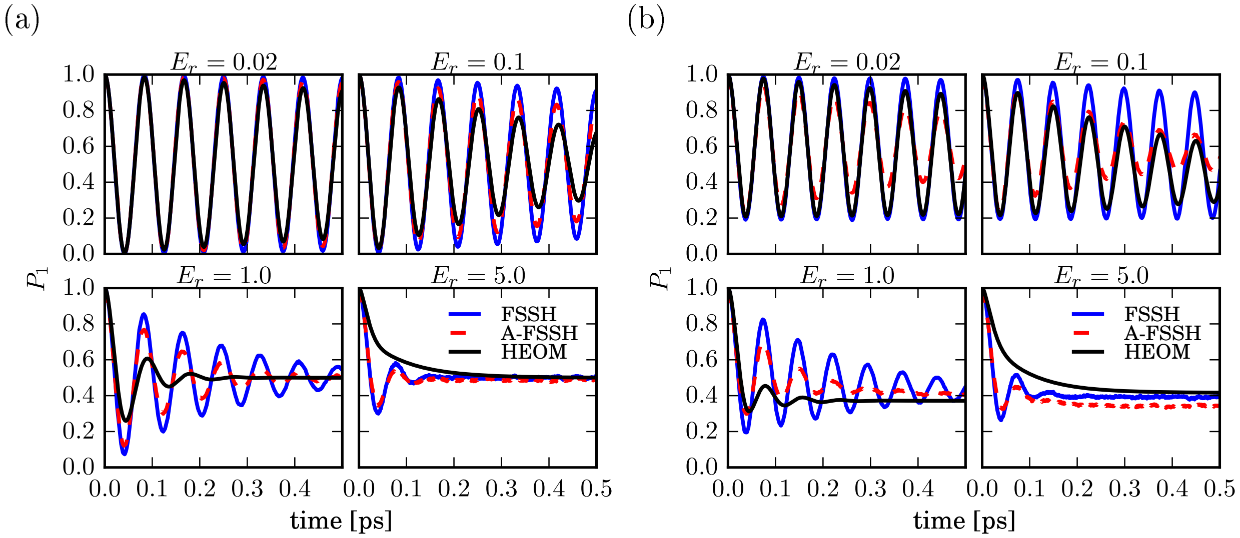


Figure 3.6: Low temperature dynamics in the intermediate regime. We employ a reference unit of energy of 104 cm$^{-1}$. Parameters are $kT = 0.2$ ($T = 30$ K), $\Delta = 2$, $\omega_c = 2$, $kT = 0.2$, (a) $\epsilon/\Delta = 0$, and (b) $\epsilon/\Delta = 1/2$. Reorganization energies ($E_r = \lambda/\Delta$) are scanned from small to large and are listed on each panel.

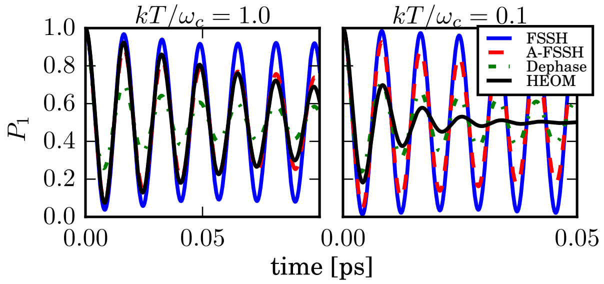


Figure 3.7: Intermediate and low temperature dynamics in the adiabatic regime. We employ a reference unit of energy of 104 cm$^{-1}$. $\epsilon = 0$ and reorganization energy is large $\lambda/\Delta = 5$. Parameters are (left) $\Delta = 10$, $\omega_c = 1$, $kT = 1$ and (right) $\Delta = 20$, $\omega_c = 2$, $kT = 0.2$. "Dephase" refers to the use of Eqs. (3.24)-(3.26).

is strong and the time scale is relatively short, both FSSH and A-FSSH can accurately model the Rabi-like oscillations for several periods of motion as illustrated in the upper left panel of Fig. 3.6 (a) and (b). For intermediate temperatures A-FSSH can accurately correct the decay rate of the amplitude of oscillations, however its accuracy diminishes at lower temperatures as shown in Fig. 3.7. In general, however, surface hopping fails to quantitatively capture population relaxation in these regimes, with some "worst-case" examples illustrated in Fig. 3.6.

## 3.5 Conclusions

In this chapter we have provided, to the best of our knowledge, the first detailed comparison of surface hopping with exact quantum dynamics for an idealized but non-trivial model of condensed phase non-adiabatic dynamics. In particular, we have focused on the role played by decoherence across the entire parameter space in general, and in the incoherent golden-rule regime in particular. Our results provide both an understanding of how decoherence influences behavior in the weak electronic coupling regime as well as general guidelines for the reliability of surface hopping with or without decoherence corrections across all regimes.

With respect to recovery of Marcus golden-rule scaling behavior, we present several novel findings. First, we find that deviations from golden-rule scaling, at least within the confines of the spin–boson model with a standard Debye spectral density, do not occur for symmetric systems and only become apparent in systems with a large energetic bias. In biased cases the inclusion of decoherence appears to correct the errant behavior of the standard FSSH approach. On the other hand, we show that the origins of the inability of FSSH to yield golden-rule behavior are subtle and the departure from the quadratic scaling are smaller than expected from past work. Lastly, we note that while it is clear that the decoherence based A-FSSH algorithm alters the electronic dependence of the transfer rate in the weak coupling limit so that the standard golden-rule is recovered, we have no analytical argument that demonstrates that this should occur, or that it will continue to be true over a wider range of $\Delta$ than we have investigated.

A systematic survey of parameter space provides important guidelines concerning the accuracy of surface hopping and its decoherence-corrected variants. One major conclusion that can be immediately reached is that, in general, the standard FSSH is surprisingly accurate in large portions of parameter space. Furthermore, while the decoherence-based A-FSSH approach often leads to some improvement in the description of the temporal decay of non-equilibrium population, on average the corrections are not dramatic. The largest improvements fostered by the inclusion of decoherence provided within the A-FSSH approach are found in the previously discussed golden-rule regime (c.f. Fig. 3.2) as well as in cases where decoherence damps otherwise oscillatory population decay. Thus, at least with respect condensed phase environments with widely dispersed spectral properties, the standard FSSH approach should generally provide a reasonable description of dynamics.

All of the surface hopping approaches we have employed in this chapter have difficulty in accurately describing low temperature situations, with the exception of symmetric cases where the electronic coupling is so weak that essentially pure Rabi oscillations are observed on short to intermediate time scales. However this breakdown of surface hopping is unsurprising as the approach is incapable of capturing nuclear tunneling effects. Quantitative breakdowns also appear at high and intermediate temperatures not only in the golden-rule limit, but also for intermediate to strong electronic coupling when the coupling to the bath (as given in the reorganization energy) is also sizable. However, even in these regimes failures appear as isolated examples more than generic trends.

We have also investigated decoherence corrections that are perhaps less well justified than that provided by A-FSSH but are simpler conceptually. In particular, we have explored an approach similar to the earliest decoherence corrections which employs a simple damping term given by the pure dephasing rate along a trajectory. In general we find that such an approach decoheres relaxation dynamics too strongly, often worsening agreement between the standard FSSH algorithm and the exact results. Somewhat surprisingly however, the degree of decoherence provided by this approach may be seen to quantitatively correct the failures of both FSSH and A-FSSH in the "isolated" cases where both fail, namely the regimes of sizable electronic and system-bath couplings mentioned above. This coincidence should be investigated further, as it may foster a deeper understanding of the physics associated with these isolated examples, something that we currently have been unable to provide.

# Chapter 4

# Inchworm Quantum Monte Carlo Method for Exact Non-adiabatic Dynamics

## 4.1 Introduction

The description of real-time dynamics in many-body quantum systems continues to provide major challenges for current research. An accurate theoretical understanding of nonequilibrium processes ranging from charge and energy transport in quantum dots and molecular junctions[34] to laser-induced electronic phase transformations[31] is crucial for the interpretation of experimental results and the eventual design of new materials and technologies. Quantum Monte Carlo (QMC) techniques form the basis for the exact description of the *thermodynamics* of systems dominated by quantum fluctuations[173]. In this setting, a variety of QMC methods may be used to *exactly* calculate the properties of lattice and continuum systems, including systems where boson particle statistics induce non-trivial collective phenomena such as the transition to a superfluid state[174]. The inclusion of fermionic statistics within QMC is more difficult, reflecting the NP-hardness of the generic electronic structure problem[174–176]. This difficulty reveals itself in the "fermionic sign problem," where Monte Carlo summands alternate sign, leading to a poor signal-to-noise ratio that can inhibit the accurate calculation of the thermodynamic properties of fermionic assemblies. Despite this difficulty, the umbrella of QMC techniques has essentially solved the problem of the thermodynamics of non-fermionic systems[177], while great progress continues to be made towards the development of accurate QMC approaches for fermions[101, 103, 178–180].

The simulation of real-time quantum dynamics presents another layer of difficulty that is absent when thermodynamics alone is considered. In general, when considering the exact simulation of quantum dynamics, the computational cost scales exponentially with increasing time. This poor scaling manifests in distinct ways in different methodologies[124, 127, 129, 181–185]. Within attempts to extend QMC to the real-time axis, exponentially poor scaling arises from the oscillating phase factors generated by the time evolution operator $e^{-iHt}$. The summation of random phase information leads to a shrinking signal to noise ratio known as the "dynamical sign problem". This afflicts all dynamical QMC simulations, regardless of the nature of the underlying particle statistics.

Modern diagrammatic variants of QMC (dQMC) have proven extremely powerful in the study of thermodynamic properties of impurity models, which consist of a small interacting subsystem coupled to noninteracting fermionic or bosonic baths[127]. The extension of these approaches to real-time dynamics has also met with some success[124, 182–184, 186, 187]. In particular, in conjunction with partial resummations of the exact diagrammatic series[127, 129] and reduced dynamics techniques[130, 131], real-time dQMC has proven capable of the exact simulation of nonequilibrium properties in the paradigmatic Anderson model for nontrivial time scales in select parameter regimes[132]. Despite the aforementioned successes, previous real-time dQMC methods have all been plagued by the dynamical sign problem to differing degrees[124, 182–184, 186, 187]. Very recently, a new dQMC method, dubbed the "inchworm algorithm," has been introduced that largely overcomes the dynamical sign problem[106]. The inchworm algorithm optimally recycles diagrammatic information so that the computational cost scales approximately quadratically, as opposed to exponentially, with time. For the case of the Anderson model, the inchworm algorithm has enabled exact real-time simulation even deep within strongly correlated regions of the parameter space, such as the Kondo and mixed valence regimes.

While progress for the Anderson model has been impressive, it should be noted that the number and range of exact benchmarks for this model are far fewer and less impressive than those available for a simpler impurity model: the spin–boson model. The spin–boson model consists of a two-level system coupled linearly to a bosonic bath, and constitutes the basic proxy for dissipative condensed phase charge and energy transfer problems[2, 3, 48]. Two decades of numerical effort aimed at the spin–boson problem have produced a suite of methodologies capable of long-time simulation of nonequilibrium observables over essentially the entire parameter space of the model[87, 88, 93–96, 99–102, 144]. In this sense, the spin–boson model embodies a stringent test which should be passed by any new numerically exact approach to real-time quantum dynamics.

While the spin–boson model employs seemingly unrealistic features such as linear coupling to a harmonic reservoir, even anharmonic systems may be mapped to this form of environmental interaction within linear response theory[188–192]. This generality explains the wide usage of the spin–boson paradigm in the modeling of systems ranging from charge and energy transfer in condensed phases and biological systems[4–11] to the relaxation of dilute impurities in the solid state[193–195] and in Josephson junction arrays[196]. In the rotating wave approximation, the spin–boson model is reduced to the Jaynes–Cummings model, which is of great importance in quantum optics.[197–199]

In the following work, we use the spin–boson model as a platform to provide the essential details of the inchworm approach and to improve and expand upon the methodology. In particular we describe two diagrammatic expansions (and their resummations within the inchworm framework) rooted in distinct exactly solvable reference systems. We further introduce a new cumulant-based approach[200–203] that reduces the computational cost from quadratic to linear in time. In essence, the use of cumulants allows for the construction of an inchworm expansion for the memory function directly from the moment expansion and without the need for any *a priori* information about the memory kernel itself. We argue that taken together, the distinct inchworm algorithms presented here should essentially cover the relevant parameter space of the spin–boson model.

In this work we compare the results of the inchworm algorithm to those produced by the other methodologies mentioned above in essentially all regimes of interest. Our results allow us to compare and contrast the strengths and weakness of the relative approaches. We demonstrate that the inchworm algorithm is competitive with the most advanced real-time approaches and is capable of producing converged long-time results even in some regimes difficult for several prominent approaches. The success of the inchworm algorithm as outlined in this work paves the way for a host of novel applications, a few of which we enumerate at the conclusion of this chapter.

The organization of the chapter is as follows. In Sec. 4.2, we review the real-time dQMC scheme and the inchworm algorithm in a general formalism. In Sec. 4.3, we formulate the system–bath coupling expansion and its corresponding inchworm expansion. In Sec. 4.4, the diabatic coupling expansion is described. In Sec. 4.5, we introduce cumulant inchworm expansions based on the diabatic coupling expansion. In Sec. 4.6, we provide an analysis of convergence for the system–bath coupling inchworm (SBCI) and the diabatic coupling cumulant inchworm (DCCI) approaches. In Sec. 4.7, the detailed comparison of our new approach to established benchmarks, as well as a discussion of the relative benefits and drawbacks of our approach, are presented. A conclusion is presented in Sec. 4.8.

## 4.2 dQMC Scheme and the Inchworm Algorithm

In this section, we briefly review the real-time dQMC approach[127], the dynamical sign problem and the inchworm algorithm[106] in a general framework.

We consider a generic Hamiltonian of an open quantum system in the form

$$H = H_s + H_b + H_{sb}, \tag{4.1}$$

where $H_s$ and $H_b$ are the Hamiltonian of the system and the bath, respectively, and $H_{sb}$ describes the system–bath coupling. For a given observable $O$, we are interested in its time-dependent expectation value

$$\langle O\left(t\right)\rangle = \mathrm{Tr}\left\{\rho_0 e^{iHt} O e^{-iHt}\right\}. \tag{4.2}$$

Here, $\langle \cdot \rangle = \mathrm{Tr}\left\{\rho_0 \cdot\right\}$ is the trace performed over all degrees of freedom and $\rho_0$ is the initial density matrix of the full system. It should be noted that equilibrium time correlation functions may also be calculated within the framework outlined below[121, 132], however for simplicity we focus on one-time non-equilibrium quantities of the form (4.2).

### 4.2.1 Dyson Series

To evaluate the dynamics of the observable $\langle O\left(t\right)\rangle$, a key needed element is the propagator $e^{-iHt}$, which is difficult to calculate in a computationally useful form. In general, we can expand the propagator in a perturbative fashion by writing the Hamiltonian as

$$H \quad = \quad H_0 + H', \tag{4.3}$$

thus partitioning $H$ into a (solvable) $H_0$ and an interaction Hamiltonian $H'$. In this interaction picture, the dynamics of an operator $O$ is given by

$$e^{iHt} O e^{-iHt} = U^{\dagger}\left(t\right) \widetilde{O}\left(t\right) U\left(t\right), \tag{4.4}$$

where the propagator is $U\left(t\right)$ given by $U\left(t\right) = e^{iH_0 t} e^{-iHt}$. We denote the time-dependent operator in the interaction picture by $\widetilde{O}\left(t\right) = e^{iH_0 t} O e^{-iH_0 t}$. One can expand the propagator using the time-ordered Dyson series ($\hbar = 1$)

$$\begin{aligned} U\left(t\right) = & \sum_{n=0}^{\infty} \left(-i\right)^n \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_{n-1}} \mathrm{d}t_n \\ & \times \widetilde{H}'\left(t_1\right) \widetilde{H}'\left(t_2\right) \cdots \widetilde{H}'\left(t_n\right) \end{aligned} \tag{4.5}$$

which contains a series of interaction operators $\widetilde{H}'\left(t_i\right) = e^{iH_0 t} H' e^{-iH_0 t}$ with the chronological time ordering $t > t_1 > t_2 > \cdots > t_n > 0$. If this expansion is applied to the two interaction picture propagators in Eq. 4.4, the folded Keldysh contour naturally emerges from the sequence of interaction operators generated by the product. The interaction operators arising from $U(t)$ have time arguments denoted as $\left\{t_i^+\right\}$, and are thought of as existing on the forward or + branch of the contour, while those emanating from $U^\dagger(t)$ have time arguments denoted as $\left\{t_i^-\right\}$, and exist on the backward or − branch. This is illustrated in Fig. 4.1a. The contour is folded at $t = t_{\text{max}}$, where the observable operator is applied. Each set of time arguments, $\left\{t_i^+\right\}$ and $\left\{t_i^-\right\}$, is time ordered: $t_{\text{max}} > t_1^\pm > t_2^\pm > \cdots > 0^\pm$, where $0^\pm$ denote the initial time on the ± branch, respectively. Therefore, we can write Eq. 4.4 as

$$\begin{aligned} O(t) = &\sum_{n=0}^{\infty} \int_0^{t_{\text{max}}} \mathrm{d}t_1^+ \int_0^{t_1^+} \mathrm{d}t_2^+ \ldots \int_0^{t_{n-1}^+} \mathrm{d}t_n^+ \times \\ &\sum_{n'=0}^{\infty} \int_0^{t_{\text{max}}} \mathrm{d}t_1^- \int_0^{t_1^-} \mathrm{d}t_2^- \ldots \int_0^{t_{n'-1}^-} \mathrm{d}t_{n'}^- \times \\ &(-i)^n i^{n'} \widetilde{H}'\left(t_{n'}^-\right) \cdots \widetilde{H}'\left(t_1^-\right) \times \\ &\widetilde{O}\left(t_{\text{max}}\right) \widetilde{H}'\left(t_1^+\right) \cdots \widetilde{H}'\left(t_n^+\right). \end{aligned} \tag{4.6}$$

For brevity, it will be convenient to write the two types of time arguments on the two branches of the contour in terms of a single time argument label $s_i$:

$$s_i = \begin{cases} s_i^+ = t_{n-i+1}^+ & i \leq n, \\ s_i^- = t_{i-n}^- & n < i \leq m. \end{cases} \tag{4.7}$$

Here, $m = n + n'$ and $\{s_i\}$ is ordered according to the Keldysh contour causality, $s_1 < \cdots < s_m$ as shown in Fig. 4.1. We define $s_i < s_j$ if $s_i$ occurs before $s_j$ on the Keldysh contour.

Therefore, we can write Eq. 4.4 as an expansion in terms of $s_i$,

$$\begin{aligned} O\left(t\right) = & \sum_{m=0}^{\infty}\sum_{n=0}^{m}\int \mathrm{d}s_m \cdots \int \mathrm{d}s_1 \left(-1\right)^n i^m \times \\ & \widetilde{H}'\left(s_m\right)\cdots\widetilde{H}'\left(s_{n+1}\right)\widetilde{O}\left(t_{\max}\right)\widetilde{H}'\left(s_n\right)\cdots\widetilde{H}'\left(s_1\right), \end{aligned} \tag{4.8}$$

where the integration $\int \mathrm{d}s_m \cdots \int \mathrm{d}s_1$ is taken to represent

$$\begin{aligned} \int \mathrm{d}s_m \cdots \int \mathrm{d}s_1 = & \int_0^{t_{\max}} \mathrm{d}t_1^+ \int_0^{t_1^+} \mathrm{d}t_2^+ ... \int_0^{t_{n-1}^+} \mathrm{d}t_n^+ \\ & \int_0^{t_{\max}} \mathrm{d}t_1^- \int_0^{t_1^-} \mathrm{d}t_2^- ... \int_0^{t_{n'-1}^-} \mathrm{d}t_{n'}^- . \end{aligned} \tag{4.9}$$

Each term in the expansion can be represented by diagrams, in which a vertex or open circle in Fig. 1(a) represents the interactions occurring at the times $\{s_i\}$ and a cross symbol indicates the tip or the folding time $t_{\max}$ of the Keldysh contour where the observable operator acts. For instance, Fig. 4.1b shows the diagrams of the unperturbed term ($m = 0$) and some example diagrams of second order ($m = 2$, two vertices) and of fourth order ($m = 4$, four vertices).

### 4.2.2 Real-time path integral formulation

The dynamical quantities of interest can be expressed in the form of a path integral, or more generally the integral over the contour configuration space

$$\langle O\left(t\right)\rangle = \int \mathrm{d}\boldsymbol{s}\mathcal{O}\left(\boldsymbol{s}\right), \tag{4.10}$$

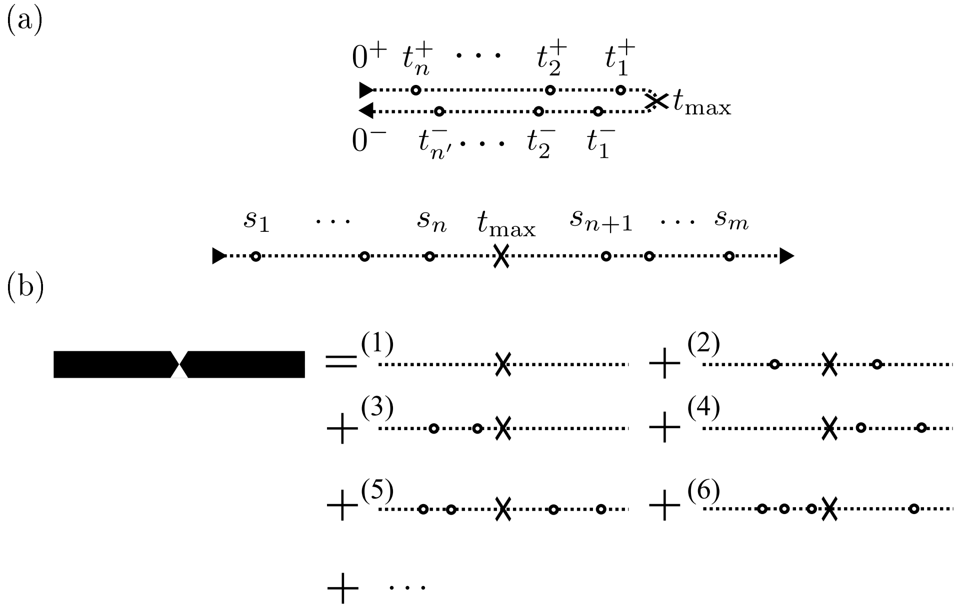


Figure 4.1: (a) A configuration $\boldsymbol{s}$ drawn on the Keldysh contour, with physical times $t_i$ on the forward or + branch and $t_i'$ on the backward or − branch. Below, the configuration is shown on the unfolded contour with contour times $s_i$. The × indicates the tip or fold of the contour and the ticks indicate interaction operators $H'$. (b) General framework of *bare* dQMC. The thin line represents an unperturbed propagator $e^{-iH_0 s}$, while the thick line represents the exact sum over all possible configurations contributing to some observable $\langle O(t) \rangle$. (1) is the zeroth ($m = 0$) order contribution, $\left\langle e^{iH_0 t} O e^{-iH_0 t} \right\rangle$. (2)–(4) are examples of second ($m = 2$) order contributions with (2)$n = 1$, (3)$n = 2$, and (4)$n = 0$. (5) and (6) are examples of fourth ($m = 4$) order configurations.

where we denote $\boldsymbol{s} = \{s_i\}$ as the contour configuration. Note that this expression is implicitly time-ordered and the integration $\int \mathrm{d}\mathbf{s}$ is taken to mean

$$\int \mathrm{d}\boldsymbol{s} = \sum_{m=0}^{\infty} \sum_{n=0}^{m} \int \mathrm{d}s_m \cdots \int \mathrm{d}s_1. \tag{4.11}$$

The contribution of a given configuration $\boldsymbol{s}$ is given by

$$\begin{aligned} \mathcal{O}(\boldsymbol{s}) = (-1)^n i^m \Big\langle \widetilde{H}'(s_m) \cdots \widetilde{H}'(s_{n+1}) \times \\ \widetilde{O}(t_{\max}) \widetilde{H}'(s_n) \cdots \widetilde{H}'(s_1) \Big\rangle. \end{aligned} \tag{4.12}$$

This object is a seemingly complicated multi-time quantity, but in many cases is that it can be efficiently evaluated since it is defined by an interaction picture under the propagation associated with a solvable $H_0$.

The dQMC method provides an unbiased estimator for the infinite-dimensional integral over all configuration parameters, $\int \mathrm{d}\boldsymbol{s}\mathcal{O}(\boldsymbol{s})$, by summing over a set of sample configurations $\boldsymbol{s}_i$ drawn from some normalized probability distribution defined by

$$\mathrm{Prob}(\boldsymbol{s}) = \frac{\mathrm{w}(\boldsymbol{s})}{\int \mathrm{d}\boldsymbol{s}\mathrm{w}(\boldsymbol{s})} \equiv \frac{\mathrm{w}(\boldsymbol{s})}{Z_{\mathrm{w}}}. \tag{4.13}$$

The Metropolis–Hastings algorithm[204, 205] is method for generating a sample set of this type when only $\mathrm{w}(\boldsymbol{s})$ is known. To see how this works, consider that for any prescribed weight function $\mathrm{w}(\boldsymbol{s})$, we have

$$\int \mathrm{d}\boldsymbol{s}\mathcal{O}(\boldsymbol{s}) = Z_{\mathrm{w}} \int \mathrm{d}\boldsymbol{s} \frac{\mathcal{O}(\boldsymbol{s})}{\mathrm{w}(\boldsymbol{s})} \mathrm{Prob}(\boldsymbol{s}). \tag{4.14}$$

Given that the $\{\boldsymbol{s}_i\}$ for $i \in \{1, ..., M\}$ are drawn from $\mathrm{Prob}(\boldsymbol{s})$, in the limit of large $M$ one obtains

$$\int \mathrm{d}\boldsymbol{s}\mathcal{O}(\boldsymbol{s}) \simeq \frac{Z_{\mathrm{w}}}{M} \sum_{i=1}^{M} \frac{\mathcal{O}(\boldsymbol{s}_i)}{\mathrm{w}(\boldsymbol{s}_i)} \equiv Z_{\mathrm{w}} \left\langle \frac{\mathcal{O}}{\mathrm{w}} \right\rangle_{\mathrm{w}}. \tag{4.15}$$

Importantly, we note that $Z_{\mathrm{w}}$ is completely independent of the observable calculated. Therefore, to remove the dependence on $Z_{\mathrm{w}}$, we introduce a "normalizing" observable $N = \int \mathrm{d}\boldsymbol{s}\mathcal{N}(\boldsymbol{s})$ which can be evaluated analytically. Evaluating $N$ via the same Monte Carlo procedure, one obtains

$$N = \int \mathrm{d}\boldsymbol{s}\mathcal{N}(\boldsymbol{s}) \simeq Z_{\mathrm{w}} \left\langle \frac{\mathcal{N}}{\mathrm{w}} \right\rangle_{\mathrm{w}} . \tag{4.16}$$

With this normalization, $Z_{\mathrm{w}}$ cancels out of all final expressions:

$$\int \mathrm{d}\boldsymbol{s}\mathcal{O}(\boldsymbol{s}) = N \frac{\int \mathrm{d}\boldsymbol{s}\mathcal{O}(\boldsymbol{s})}{\int \mathrm{d}\boldsymbol{s}\mathcal{N}(\boldsymbol{s})} \simeq N \frac{Z_{\mathrm{w}} \left\langle \frac{\mathcal{O}}{\mathrm{w}} \right\rangle_{\mathrm{w}}}{Z_{\mathrm{w}} \left\langle \frac{\mathcal{N}}{\mathrm{w}} \right\rangle_{\mathrm{w}}} = N \frac{\left\langle \frac{\mathcal{O}}{\mathrm{w}} \right\rangle_{\mathrm{w}}}{\left\langle \frac{\mathcal{N}}{\mathrm{w}} \right\rangle_{\mathrm{w}}}. \tag{4.17}$$

Since we have complete freedom in the choice of $N$, one is therefore free to choose a quantity which is easy to evaluate in both the Monte Carlo and the analytical calculation. The choice used here is $\mathcal{N}(\boldsymbol{s}) = 1$, such that $N$ is simply the hypervolume of the multidimensional space of interaction times. Since this hypervolume normalization is positive definite, it cannot have a sign problem, and all potential sign problems must appear in the nominator. For $\mathrm{w}(\boldsymbol{s})$, we typically choose the absolute value $|\mathcal{O}(\boldsymbol{s})|$ of the contribution to the observable itself or a closely related property, such that the summation is optimized for summing large contributions to a particular observable. It is currently unknown whether this choice is optimal.

### 4.2.3 Dynamical sign problem and inchworm algorithms

Unfortunately, summing individual contributions to an observable in this manner, the so-called *bare* dQMC algorithm, generally involves a dynamical sign problem. In real-time dQMC, the dynamical sign problem is caused by the oscillatory nature of real-time propagators which results an exponentially growing computational cost as time increases[124, 182, 186, 206]. To circumvent the dynamical sign problem, we employ inchworm expansions[106]. This allows us to efficiently reuse quantities propagated within short time intervals in the calculation of quantities propagated between longer times. Two concrete examples of practical inchworm algorithms for the spin–boson model will be developed below.

We briefly introduce the general concept behind inchworm expansions. Let $s_i < s_\uparrow < s_f$ be three times: an "initial," "inchworm" and "final" time, respectively. Assume some set of properties have been exactly evaluated for all cases where all interaction vertices are restricted to the time interval $[s_i, s_\uparrow]$. Given knowledge of these auxiliary restricted quantities, it is often possible to construct an efficient expansion for the same set of quantities with the vertices restricted to the *longer* interval $[s_i, s_f]$. This describes an *inchworm step*, or the process of *inching*. A series of inchworm steps allows one to start with a set of easily evaluated restricted quantities defined over very short intervals, eventually obtaining the set of unrestricted physical quantities for which interaction vertices span the full length of the Keldysh contour.

The inchworm algorithm has the distinct advantage that much fewer diagrams must be sampled to obtain a converged answer, since each inchworm diagram contains an infinite number of bare diagrams, such that often relatively few low-order inchworm diagrams contain all important contributions from the relevant bare diagrams at all orders. This advantage comes at two important costs. First, as when working with nonequilibrium Green's functions, one is forced to calculate a complete set of two-time properties even if only single-time properties are of interest. Specifically, all propagators between any two points in $[s_i, s_\uparrow]$ are required to obtain a propagator between $s_i$ and $s_f$. Second, Monte Carlo evaluations at long times are no longer independent of short-time evaluations, and errors are carried forward in time during the stepping procedure. This has profound computational implications in that the algorithm is not "embarrassingly parallel" like standard Monte Carlo techniques, since information concerning short-time propagators must be distributed between the various computer nodes performing the calculation. Furthermore, careful error analysis is required in order to take error propagation into account. Essentially, a series of completely independent calculations must be carried out to evaluate the statistical errors, and one must then verify that systematic errors due to the error propagation (in addition to the statistical ones common to all Monte Carlo techniques) are assessed and converged to within the desired accuracy.

Within the formulation of inchworm algorithm, each single inchworm step is numerically exact if Monte Carlo samples are sufficient for the convergence. The sequence/grid of inchworm steps with finite size is formally exact if all propagators of shorter times are smooth enough and well-representative in the discrete sequence/grid. In practice, we have to truncate the maximum order of sampled configuration for each inchworm step and interpolate the discrete propagator data.

### 4.2.4 Spin–boson model

We now specialize the discussion to the case of the spin–boson model. This allows us to give explicit expressions for the terms that emerge in expansions that employ different choices of $H_0$. The form of the Hamiltonian is given by Eq. (4.1). The system Hamiltonian $H_s$ is taken to be a two-level system in the diabatic basis $|\alpha\rangle \in \{|1\rangle, |2\rangle\}$,

$$H_s = \epsilon\hat{\sigma}_z + \Delta\hat{\sigma}_x. \tag{4.18}$$

In this notation, $\hat{\sigma}_z = |1\rangle\langle 1| - |2\rangle\langle 2|$ and $\hat{\sigma}_x = |1\rangle\langle 2| + |2\rangle\langle 1|$. The energetic bias $\epsilon$ is the energy difference between the two diabatic states, and the diabatic coupling $\Delta$ characterizes spin flip processes within the electronic system. The boson bath consists of a set of harmonic oscillators with frequencies $\omega_\ell$ described by the bath Hamiltonian

$$H_b = \sum_\ell \frac{1}{2}\left(p_\ell^2 + \omega_\ell^2 x_\ell^2\right) = \sum_\ell \omega_\ell \left(b_\ell^\dagger b_\ell + \frac{1}{2}\right). \tag{4.19}$$

The system–bath coupling $H_{sb}$ is assumed to be linear in the bath coordinates

$$H_{sb} = \hat{\sigma}_z \sum_\ell c_\ell x_\ell. \tag{4.20}$$

The coupling constants $c_\ell$ describe the strength of the interaction between the harmonic modes and the spin. The system–bath coupling is typically parametrized in compact form by the spectral density

$$J(\omega) = \frac{\pi}{2}\sum_\ell \frac{c_\ell^2}{\omega_\ell}\delta\left(\omega - \omega_\ell\right). \tag{4.21}$$

We specify the system–bath coupling strength by a spectral density that is linear for small $\omega$ and has a Lorentzian cutoff for large $\omega$:

$$J_D(\omega) = \frac{\lambda}{2}\frac{\omega_c \omega}{\omega_c^2 + \omega^2}, \tag{4.22}$$

namely the so-called the Debye spectral density. For this spectral density, we define the reorganization energy as $\lambda = \frac{4}{\pi}\int \frac{J(\omega)}{\omega} d\omega = 2\sum_\ell c_\ell^2/\omega_\ell^2$, which provides a measure of the system–bath coupling strength. The cutoff frequency of the Lorentzian function, $\omega_c$, characterize the frequency of the bath. Therefore, a spin–boson model can be characterized by five parameters (with $\hbar = 1$): the diabatic coupling $\Delta$, the bias energy $\epsilon$ of the electronic system, the cut-off frequency $\omega_c$, the temperature $k_B T = 1/\beta$ of the boson bath, and the reorganization energy $\lambda$ characterizing system–bath coupling.

Throughout this work, we will concentrate on the local dynamics of the spin $\hat{\sigma}_z$ in the diabatic basis

$$\langle \sigma_z (t) \rangle = \mathrm{Tr} \left\{ \rho_0 e^{iHt} \hat{\sigma}_z e^{-iHt} \right\}. \tag{4.23}$$

Here we only address factorized initial conditions corresponding to thermal equilibrium of the bath in the absence of the system–bath coupling, such that the initial density matrix is given by the factorized form $\rho_0 = \rho_s \otimes \rho_b$, with the bath initially in equilibrium $\rho_b = \frac{e^{-\beta H_b}}{\mathrm{Tr}_b \left\{ e^{-\beta H_b} \right\}}$. We specify the initial condition of the spin as $\rho_s = |1\rangle \langle 1|$. Treatment of more general initial conditions is simple but will not be discussed further here.

There are several useful ways of partitioning $H$ into $H_0$ and $H'$ such that the perturbation series of Eq. 4.3 can be carried out, each yielding a different type of expansion. We will discuss two such choices. One treatment takes $H' = H_{sb}$, expanding with respect to the system–bath coupling. Another takes $H' = \Delta \hat{\sigma}_x$, expanding in the diabatic coupling $\Delta$. In the following sections, we discuss these expansions and their inchworm Monte Carlo implementations.

## 4.3 System–Bath Coupling Inchworm (SBCI) Expansion

### 4.3.1 Bare dQMC

We start with the example of the bare dQMC expansion in terms of the system–bath coupling $H' = H_{sb}$. This expansion is the analogous to the hybridization expansion in the Anderson model, for which the first inchworm expansion was formulated. The unperturbed Hamiltonian is taken to be $H_0 = H_s + H_b$ and the initial density matrix is $\rho_0 = |1\rangle \langle 1| \otimes \rho_b$.

To write a dQMC expression for the expectation value of the observable $O = \hat{\sigma}_z$, we must determine the contribution $\mathcal{O}(\boldsymbol{s})$ of any given configuration $\boldsymbol{s}$ to this expectation value in the form of Eq. 4.12:

$$\begin{aligned}\mathcal{O}(\boldsymbol{s}) = (-1)^n\, i^m \Big\langle \widetilde{H}_{sb}(s_m) \cdots \widetilde{H}_{sb}(s_{n+1}) \times \\ \widetilde{\sigma}_z(t)\, \widetilde{H}_{sb}(s_n) \cdots \widetilde{H}_{sb}(s_1) \Big\rangle.\end{aligned} \tag{4.24}$$

In the interaction picture $\widetilde{H}_{sb}(s) = e^{iH_0 s} H_{sb} e^{-iH_0 s}$ can be factorized as

$$\widetilde{H}_{sb}(s) = \widetilde{\sigma}_z(s) \times \sum_{\ell} c_\ell \widetilde{x}_\ell(s), \tag{4.25}$$

and we define the operator of the bath part as

$$\widetilde{B}(s) = \sum_{k} c_\ell \widetilde{x}_\ell(s). \tag{4.26}$$

It turns out that for a linear coupling of the form of Eq. (4.20), one can write Eq. (4.24) as a product of a system influence functional $\mathcal{U}(\boldsymbol{s})$ and a bath influence functional $\mathcal{L}(\boldsymbol{s})$:

$$\mathcal{O}(\boldsymbol{s}) = (-1)^n\, i^m \mathcal{U}(\boldsymbol{s})\, \mathcal{L}(\boldsymbol{s}). \tag{4.27}$$

The system influence functional $\mathcal{U}(\boldsymbol{s})$ for the given initial condition $|1\rangle\langle 1|$ is defined as

$$\begin{aligned}\mathcal{U}(\boldsymbol{s}) = \langle 1|\, \widetilde{\sigma}_z(s_m) \cdots \widetilde{\sigma}_z(s_{n+1}) \times \\ \widetilde{\sigma}_z(t_{\max})\, \widetilde{\sigma}_z(s_n) \cdots \widetilde{\sigma}_z(s_1)\, |1\rangle.\end{aligned} \tag{4.28}$$

For the spin-$\frac{1}{2}$ case, all operators can be written in the form of matrices of rank 2 in the basis of the Hilbert space of the isolated spin. Eq. (4.28) can then be efficiently evaluated as a matrix product of unperturbed system propagators of the form $e^{-iH_s(s_i - s_{i-1})}$, sandwiched between $\hat{\sigma}_z$ operators with $s_i - s_j$ denoting the difference of physical times given by Eq. (4.7).

The bath influence functional is given by an $m$-time interaction picture correlation function of the bath operator $\widetilde{B}\left(s\right)$ in the form of

$$\mathcal{L}\left(\boldsymbol{s}\right)=\left\langle \widetilde{B}\left(s_m\right)\cdots\widetilde{B}\left(s_1\right)\right\rangle_b, \tag{4.29}$$

where we denote $\left\langle\cdot\right\rangle_b=\mathrm{Tr}_b\left\{\rho_b\cdot\right\}$ and $\rho_b$ is the initial bath density matrix. Using Wick's theorem, which is valid for the bath operators within the interaction picture, one can express $\mathcal{L}\left(\mathbf{s}\right)$ as a sum of products of two-time correlation functions by use of the identity

$$\left\langle \widetilde{B}\left(s_m\right)\cdots\widetilde{B}\left(s_1\right)\right\rangle_b=\sum_{q\in\mathcal{Q}_m}\prod_{(j,k)\in q}\left\langle \widetilde{B}\left(s_k\right)\widetilde{B}\left(s_j\right)\right\rangle_b. \tag{4.30}$$

The bath influence functional is zero for odd $m$. $\mathcal{Q}_m$ denotes the set of possible distinct pairings of the integers 1, 2, ..., $m$: each element $q\in\mathcal{Q}_m$ is a set of ordered tuples corresponding to a single pairing. For example, for $m=2$ there is only one pairing, $q=\{(2,1)\}$, and

$$\mathcal{Q}_2=\left\{\left\{(1,2)\right\}\right\}.$$

For $m=4$ there are three possible pairings:

$$\mathcal{Q}_4=\left\{\left\{(1,2),(3,4)\right\},\left\{(1,3),(2,4)\right\},\left\{(1,4),(2,3)\right\}\right\},$$

and so on. With these definitions, the bath influence functional takes the form

$$\mathcal{L}\left(\boldsymbol{s}\right)=\sum_{q\in\mathcal{Q}_m}\mathcal{L}_q\left(\boldsymbol{s}\right), \tag{4.31}$$

where the functional of a given pairing $q$;

$$\mathcal{L}_q\left(\boldsymbol{s}\right)=\prod_{(j,k)\in q}\left\langle \widetilde{B}\left(s_k\right)\widetilde{B}\left(s_j\right)\right\rangle_b, \tag{4.32}$$

corresponding to a particular *diagram* with the coupling lines connecting $s_j$ and $s_k$ on the Keldysh contour (see Fig. 4.3a). Diagram (a.1) is the zeroth order contribution. Diagram (a.2) is the diagram associated with a given $2^{\text{nd}}$ order configuration $(s_1, s_2)$. Diagrams (a.3)–(a.5) are three diagrams (corresponding to three possible pairings of $\mathcal{Q}_4$) associated with a $4^{\text{th}}$ order configuration $(s_1, s_2, s_3, s_4)$.

The two-time correlation function of the harmonic bath in the interaction picture can be evaluated semi-analytically prior to the start of the dQMC calculation as

$$\begin{gathered}\left\langle \widetilde{B}\left(s_k\right)\widetilde{B}\left(s_j\right)\right\rangle_b = \frac{2}{\pi}\int d\omega J\left(\omega\right)\times \\ \left[\coth\left(\frac{\beta\omega}{2}\right)\cos\omega\left(s_k - s_j\right) - i\sin\omega\left(s_k - s_j\right)\right].\end{gathered} \tag{4.33}$$

In practice, an $m$-time path configuration includes $(m-1)!!$ diagrams, and computing each diagram requires $m/2$ evaluations of the bath correlation function. Thus, calculating an $m$–time correlation function requires a total of $(m-1)!!\,(m/2)$ function evaluation, which approaches $\frac{m}{\sqrt{2}}\left(m/e\right)^{m/2}$ in the large $m$ limit. This rapidly becomes a bottleneck for high perturbation order. However, rather than explicitly summing over all diagrams in a configuration, it is possible to sum over the pairings as defined in Eq. (4.31) within the Monte Carlo procedure, thus effectively removing this scaling issue at the cost of an overall increase in the sign problem.

### 4.3.2 Restricted propagators and observable

To facilitate our discussion of the inchworm algorithm, we now define *restricted propagators* on contour subintervals. Propagators are defined with respect to particular physical observables. The *bare* restricted propagator $G^{(0)}_{\alpha\beta}\left(s_f, s_i\right)$ is defined as follows. When the subinterval

$[s_i, s_f]$ is on a single branch of the contour, such that $s_i^+, s_f^+ < t_{\max}$ or $s_i^-, s_f^- > t_{\max}$ , then

$$G_{\alpha\beta}^{(0)}\left(s_f^{\pm}, s_i^{\pm}\right) = \langle\alpha|\, e^{-iH_s\left(s_f^{\pm}-s_i^{\pm}\right)}\, |\beta\rangle\,. \tag{4.34}$$

When the endpoints of the interval are on two different branches, it is defined differently in order to account for the observable at the contour's folding point:

$$G_{\alpha\beta}^{(0)}\left(s_f^{-}, s_i^{+}\right) = \langle\alpha|\, e^{-iH_s\left(s_f^{-}-t_{\max}\right)}\sigma_z e^{-iH_s\left(t_{\max}-s_i^{+}\right)}\, |\beta\rangle\,. \tag{4.35}$$

These restricted propagators are designated by thin solid lines in the diagrammatic representation (see Fig. 4.2). The *full* restricted propagator from $s_i$ to $s_f$ can be defined in terms of an integral over configurations

$$G_{\alpha\beta}\left(s_f, s_i\right) = \int_{\boldsymbol{s}\in\left[s_i,s_f\right]} \mathrm{d}\boldsymbol{s}\mathcal{G}_{\alpha\beta}\left(\boldsymbol{s}\right). \tag{4.36}$$

The notation $\boldsymbol{s} \in [s_i, s_f]$ indicates that the vertex times appearing in the configuration $\boldsymbol{s}$ are restricted to the interval $[s_i, s_f]$. The influence functional then takes the same general form as Eq. (4.27)

$$\mathcal{G}_{\alpha\beta}\left(\boldsymbol{s}\right) = \left(-1\right)^n i^m \mathcal{U}'_{\alpha\beta}\left(\boldsymbol{s}\right)\mathcal{L}\left(\boldsymbol{s}\right), \tag{4.37}$$

namely it is composed of system and bath parts, $\mathcal{U}'_{\alpha\beta}\left(\boldsymbol{s}\right)$ and $\mathcal{L}\left(\boldsymbol{s}\right)$. The bath influence functional is identical to the one given by Eqs. (4.31) and (4.32), and the system influence functional will be discussed immediately below.

The system influence functional, like the bare propagator, takes on different forms for intervals on a single branch as compared to across branches. For a single branch interval, it is

$$\begin{aligned}&\mathcal{U}'_{\alpha\beta}\left(\boldsymbol{s}\in\left[s_i^{\pm}, s_f^{\pm}\right]\right) = \\ &\quad\langle\alpha|\, e^{-iH_s s_f^{\pm}}\widetilde{\sigma}_z\left(s_n^{\pm}\right)\cdots\widetilde{\sigma}_z\left(s_1^{\pm}\right)e^{iH_s s_i^{\pm}}\, |\beta\rangle\,,\end{aligned} \tag{4.38}$$

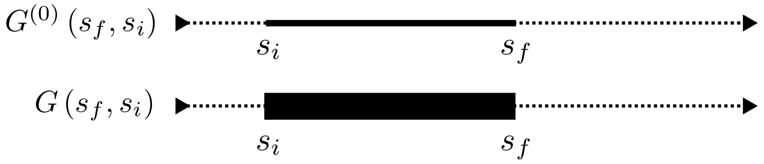


Figure 4.2: Diagrammatic representation of the bare restricted propagator $G^{(0)}$ (thin solid line) and the full restricted propagator $G$ (thick solid line) of the subinterval $[s_i, s_f]$ on an unfolded Keldysh contour.

while for a cross-branch interval it becomes

$$\begin{aligned}&\mathcal{U}'_{\alpha\beta}\left(\boldsymbol{s} \in \left[s_i^+, s_f^-\right]\right) = \\ &\quad\langle\alpha| \, e^{-iH_s s_f^-} \widetilde{\sigma}_z\left(s_m^-\right) \cdots \widetilde{\sigma}_z\left(s_{n+1}^-\right) \times \\ &\quad\widetilde{\sigma}_z(t)\, \widetilde{\sigma}_z\left(s_n^+\right) \cdots \widetilde{\sigma}_z\left(s_1^+\right) e^{iH_s s_i^+} \, |\beta\rangle \,.\end{aligned} \tag{4.39}$$

Note that if $s_i^+ = s_f^+$ or $s_i^- = s_f^-$, with both times on the same branch, the restricted propagator is trivially equal to $G^{(0)}(s_f, s_f) = G(s_f, s_f) = 1$. However, if $s_i^+ = s_f^-$, with the times appearing on opposite branches, $G_{\alpha\beta}\left(s_f^-, s_i^+\right)$ becomes the expectation value of the observable given that the the system density matrix was initially in the state $|\beta\rangle\langle\alpha|$:

$$G_{\alpha\beta}\left(s_f^-, s_f^+\right) = \langle\langle\alpha| \, \sigma_z(t - s_f) \, |\beta\rangle\rangle_b \,. \tag{4.40}$$

In terms of diagrams, the full restricted propagator is represented by a thick segment (see Fig. 4.2).

### 4.3.3 Inchworm algorithm

Suppose that the full set of restricted propagators $G_{\alpha\beta}(s_k, s_j)$ for all $s_i < s_j, s_k < s_\uparrow$ is known, and one wants to evaluate a restricted propagator over a longer interval $[s_i, s_f]$, with

$s_f > s_\uparrow$. It is possible to define an *extended* propagator for the interval $[s_i, s_f]$ by appending the bare propagator to the full propagator:

$$\overline{G}\left(s_k, s_j\right) = \begin{cases} G^{(0)}\left(s_k, s_j\right) & s_j, s_k > s_\uparrow, \\ G\left(s_k, s_j\right) & s_j, s_k < s_\uparrow, \\ G^{(0)}\left(s_k, s_\uparrow\right) G\left(s_\uparrow, s_j\right) & s_j < s_\uparrow < s_k. \end{cases} \tag{4.41}$$

Since the contributions of all configurations $\boldsymbol{s} \in [s_i, s_\uparrow]$ are included in the extended propagator, it is only necessary to sum over configurations in which every inclusion has at least one vertex contained in the interval $[s_\uparrow, s_f]$ . The propagator over the entire interval $[s_i, s_f]$ can then be constructed as a path integral over configurations

$$G\left(s_f, s_i\right) = \int_{\boldsymbol{s}\in\left[s_i, s_f\right]} \mathrm{d}\boldsymbol{s}\overline{\mathcal{G}}_{\alpha\beta}\left(\boldsymbol{s}\right). \tag{4.42}$$

The influence functional $\overline{\mathcal{G}}_{\alpha\beta}$ is defined in terms of extended propagators and a bath influence functional. It takes the form

$$\overline{\mathcal{G}}\left(\boldsymbol{s}\right) = \overline{G}\left(s_f, s_m\right)\cdots\overline{G}\left(s_2, s_1\right)\overline{G}\left(s_1, s_i\right) \times \sum_{q\in\mathcal{Q}'_m} \mathcal{L}_q\left(\boldsymbol{s}\right). \tag{4.43}$$

The bath influence functional $\sum_{q\in\mathcal{Q}'_m} \mathcal{L}_q\left(\boldsymbol{s}\right)$ is similar to that of Eq. (4.30), but summation is only carried out over $\mathcal{Q}'_m \subseteq \mathcal{Q}_m$, a subset of the pairings including only *inchworm proper* pairings.

To define inchworm propriety, we first define two pairs to be *connected* if their interaction lines, which are drawn between the members of each pair, cross each other. As connectedness is clearly an equivalence relation, any pairing can be partitioned into disjoint sets of connected pairs. A pairing or diagram is inchworm proper if there does not exist any such set with all of its vertices contained in $[s_i, s_\uparrow]$. Put differently, to check whether a particular diagram is inchworm proper one should cluster together sets of interaction lines which cross each other. If and only if every cluster includes at least one line with an endpoint in $[s_\uparrow, s_f]$ is the diagram inchworm proper. This is illustrated in Fig. 4.3, where two examples of improper diagrams are crossed out. In Fig. 4.3b, diagram (b.1) is the zeroth order inchworm diagram. Diagram (b.2), (b.4) and (b.5) are all inchworm proper 2nd order diagrams. Diagrams (b.3) is an inchworm improper diagram that is included in diagram (b.1). Diagrams (b.6)–(b.8) are associated with the same 4thorder configuration. Diagrams (b.6) and (b.7) are included in diagrams (b.4) and (b.5), respectively and only diagram (b.8) is inchworm proper.

It is straightforward to prove that any diagram in the bare expansion is accounted for once and only once within the inchworm scheme; therefore, it is formally exact. *However, every inchworm diagram contains an infinite number of bare diagrams, making the expansion substantially more efficient than the bare one.*

This method will be referred to as the System–Bath Coupling Inchworm (SBCI) approach in the following.

## 4.4 Diabatic Coupling Expansion

### 4.4.1 Polaron transformation

We now consider an expansion in terms of the diabatic coupling $H' = \Delta\sigma_x$, i.e. the spin-flip interaction. The unperturbed Hamiltonian is in this case $H_0 = H_b + \sigma_z \left(\epsilon + \sum_k c_k x_k\right)$. Since $H_0$ commutes with $\sigma_z$, its eigenstates maintain the spin quantum number $\sigma = \pm 1$, which partitions them into two subspaces. Within each subspace the Hamiltonian is easily

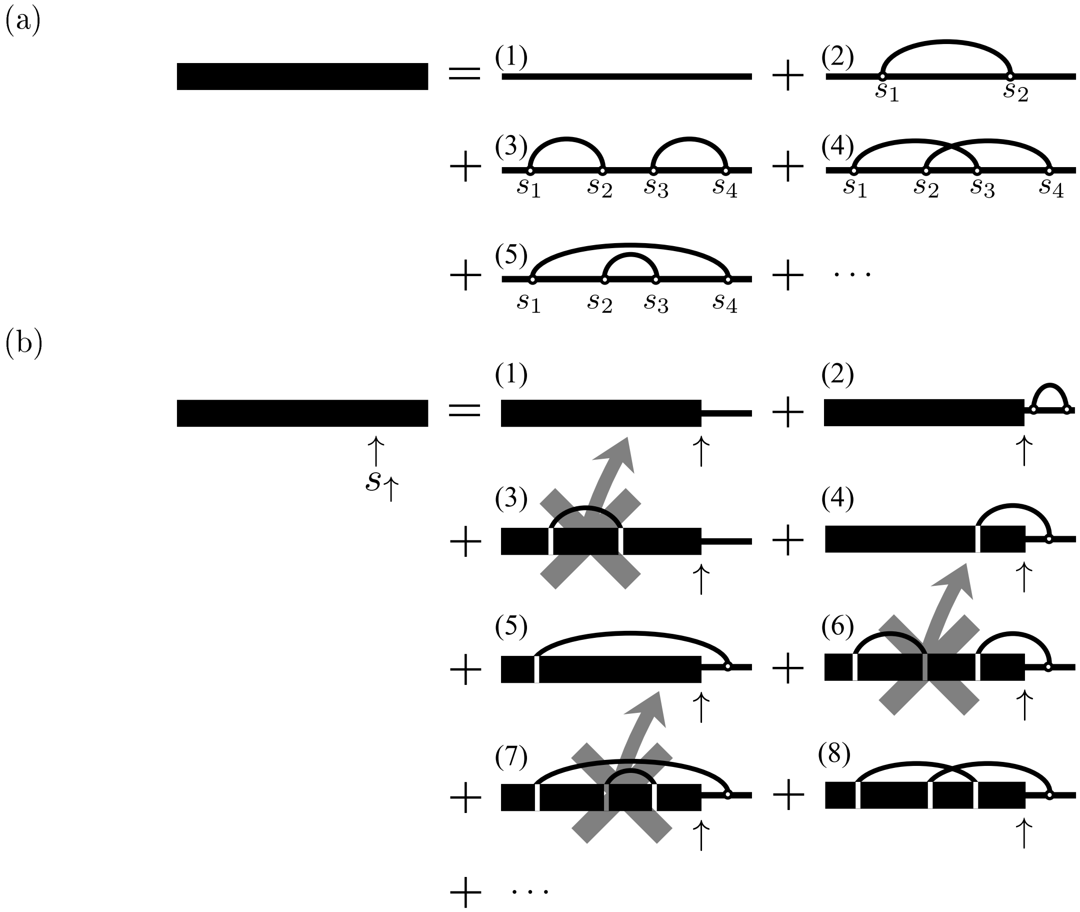


Figure 4.3: (a) The bare dQMC expression for the system–bath coupling expansion. The arched curves connecting pairs of vertices within each configuration describe the coupling interaction. (b) The inchworm algorithm in the system–bath coupling expansion. All the full restricted propagators are assumed to be known for any subinterval to the left of the $s_\uparrow$ time.

diagonalized by a polaron transformation. The effective Hamiltonian for the $\sigma = +1$ and $\sigma = -1$ subspaces, respectively, is

$$H_\sigma = H_b + \sigma \left( \epsilon + \sum_\ell c_\ell x_\ell \right). \tag{4.44}$$

We apply the transformation

$$\mathcal{B}_\sigma H_\sigma \mathcal{B}_\sigma^\dagger = H_b + \sigma\epsilon - \sum_\ell \frac{c_\ell^2}{\omega_\ell^2}, \tag{4.45}$$

where

$$\mathcal{B}_\sigma = e^{\theta_\sigma}, \tag{4.46}$$

$$\theta_\sigma = \sigma \sum_\ell \frac{c_\ell}{\omega_\ell^{3/2}} \left( b_\ell^\dagger - b_\ell \right). \tag{4.47}$$

Since $\theta_+ = -\theta_-$, it is convenient to write $\mathcal{B}_\sigma^\dagger = \mathcal{B}_{\bar{\sigma}}$. We also define $\epsilon_\sigma = \sigma\epsilon - \sum_\ell \frac{c_\ell^2}{\omega_\ell^2}$. With these definitions, the unperturbed propagator can be written in the form

$$e^{-iH_0 t} = \sum_{\sigma=\pm} e^{-i\epsilon_\sigma t} \mathcal{B}_{\bar{\sigma}} e^{-iH_b t} \mathcal{B}_\sigma \left|\sigma\right\rangle \left\langle\sigma\right|. \tag{4.48}$$

In this form the interaction picture time evolution will turn out to be very easy to evaluate, as discussed below.

The natural initial condition for the expansion in the diabatic coupling is $\rho_b = \exp\left[-\beta H_\pm\right]$, and using one of these two choices simplifies the expressions substantially. However, in order to allow for rigorous comparison with the system–bath coupling expansion, we choose to start from a state described by $\rho_b = \exp\left[-\beta H_b\right]$. Unfortunately, this introduces additional complications in the expressions given below, and we will comment on this as we proceed. The choice of initial condition does not otherwise impact the formalism.

### 4.4.2 Bare dQMC

To obtain a dQMC algorithm for the expectation value of $O = \hat{\sigma}_z$, we must write the contribution $\mathcal{O}\left(\boldsymbol{s}\right)$ of a configuration $\mathbf{s}$ in the form of Eq. 4.12. In the interaction picture, $\widetilde{\sigma}_x\left(s\right) = e^{iH_0 s}\sigma_x e^{-iH_0 s}$, we can write

$$\begin{aligned}\mathcal{O}\left(\boldsymbol{s}\right) = \left(-1\right)^n i^m \Delta^m \Big\langle \widetilde{\sigma}_x\left(s_m\right)\cdots\widetilde{\sigma}_x\left(s_{n+1}\right) \times \\ \widetilde{\sigma}_z\left(t\right)\widetilde{\sigma}_x\left(s_n\right)\cdots\widetilde{\sigma}_x\left(s_1\right)\Big\rangle.\end{aligned} \tag{4.49}$$

We designate the state between $[s_{k,}s_{k+1}]$ as $\sigma_{k+1}$ for $k \in \{0,\ldots,m-1\}$, with $s_0 \equiv 0^+$ and $s_{m+1} \equiv 0^-$. The observable $\sigma_z$ at the tip of the contour does not change the state, while every application of $\sigma_x$ flips the state from $\sigma$ to $\bar{\sigma}$. Since the initial condition of the spin is specified $\rho_s = |1\rangle\langle 1| = |+\rangle\langle +|$, we have $\sigma_1 = \sigma_{m+1} = +$. The contribution $\mathcal{O}\left(\boldsymbol{s}\right)$ of a configuration $\boldsymbol{s}$ to the expectation value of $O = \hat{\sigma}_z$ can then be expressed as a product of a system influence functional $\Phi\left(\boldsymbol{s}\right)$ and a bath influence functional $\mathcal{J}\left(\boldsymbol{s}\right)$:

$$\mathcal{O}\left(\boldsymbol{s}\right) = \left(-1\right)^n i^m \Delta^m \Phi\left(\boldsymbol{s}\right)\mathcal{J}\left(\boldsymbol{s}\right). \tag{4.50}$$

The system functional $\Phi\left(\boldsymbol{s}\right)$ handles the influence of propagation within the system,

$$\begin{aligned}\Phi\left(\boldsymbol{s}\right) = &\langle +1|\,\sigma_x^{n'}\sigma_z\sigma_x^{n}\,|+1\rangle \\ &\times \exp\left[-i\sum_{k=1}^{m+1}\epsilon_{\sigma_k}\left(s_k - s_{k-1}\right)\right],\end{aligned} \tag{4.51}$$

whiles the bath functional $\mathcal{J}\left(\boldsymbol{s}\right)$ is a multi-time correlation function of bath operators

$$\mathcal{J}\left(\boldsymbol{s}\right) = \left\langle \widetilde{\mathcal{B}}_-\left(s_{m+1}\right)\prod_{k=1}^{m}\widetilde{\mathcal{B}}^2_{\bar{\sigma}_k}\left(s_k\right)\widetilde{\mathcal{B}}_+\left(s_0\right)\right\rangle_b. \tag{4.52}$$

Here $\langle\cdot\rangle_b = \mathrm{Tr}_b\left\{\rho_b\cdot\right\}$ and $\rho_b$ is the initial bath density matrix. The first and last factors are induced by the initial condition. By a generalized Wick's theorem for polaron shift operator (see Appendix 4.A), we can write $\mathcal{J}(\mathbf{s})$ as a product of two–time correlation functions,

$$\mathcal{J}(\boldsymbol{s}) = \frac{\prod_{(j,k)\in\mathcal{C}_{m+2}^{\mathrm{odd}}} C(s_k, s_j)^{r_k r_j}}{\prod_{(j,k)\in\mathcal{C}_{m+2}^{\mathrm{even}}} C(s_k, s_j)^{r_k r_j}}. \tag{4.53}$$

where $r_i = 1$ if $i = 1, m+1$, otherwise $r_i = 2$. The fact that the powers in the numerator and denominator may differ arises from the initial condition. Here we have defined

$$\mathcal{C}_{m+2}^{\mathrm{even}} = \{(j,k) \in \mathcal{C}_{m+2} |\, |k-j| \text{ even}\}, \tag{4.54}$$

and

$$\mathcal{C}_{m+2}^{\mathrm{odd}} = \{(j,k) \in \mathcal{C}_{m+2} |\, |k-j| \text{ odd}\}. \tag{4.55}$$

which are subsets of all possible pairings of $m+2$ elements. The pairings of $m$ elements, $\mathcal{C}_m$, denotes the set of all ordered tuples composed of different integers between 0 and $m-1$. For example,

$$\mathcal{C}_2^{\mathrm{odd}} = \{(0,1)\}, \ \mathcal{C}_2^{\mathrm{even}} = \{\},$$

where $\{\}$ denotes the empty set and

$$\begin{aligned}\mathcal{C}_4^{\mathrm{odd}} &= \{(0,1), (1,2), (2,3), (0,3)\}, \\ \mathcal{C}_4^{\mathrm{even}} &= \{(0,2), (1,3)\}.\end{aligned}$$

The correlation function $C(s_k, s_j)$ is one of the expressions complicated by the initial condition, and is given by

$$C(s_k, s_j) = \left\langle \widetilde{\mathcal{B}}_-(s_k)\, \widetilde{\mathcal{B}}_+(s_j) \right\rangle_b \tag{4.56}$$

In general, we can write the two-time correlation function of the polaron shift operator as (see Appendix. (4.A))

$$C(s_k, s_j) \;=\; e^{-Q_2(s) - iQ_1(s)}, \tag{4.57}$$

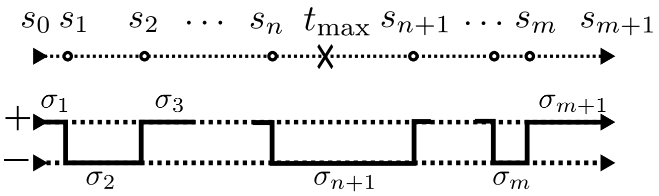


Figure 4.4: A configuration including $s_0 = s_i$ and $s_{m+1} = s_f$ for the diabatic coupling expansion. The state of the system flips at every $s_i$.

with

$$Q_1(s) = \frac{2}{\pi}\int \mathrm{d}\omega \frac{J(\omega)}{\omega^2}\sin\omega s, \tag{4.58}$$

and

$$Q_2(s) = \frac{2}{\pi}\int \mathrm{d}\omega \frac{J(\omega)}{\omega^2}\coth\left(\frac{\beta\omega}{2}\right)(1-\cos\omega s). \tag{4.59}$$

In the diagrammatic representation shown in Fig. 4.4, the two-time correlation function is represented by dashed lines. There exists an extra set of lines due to the initial condition, which connect every vertex to the edges of the diagram. To avoid overcrowding the diagram with information, these are not shown. A dashed line above the contour describes a contribution to the numerator, while one under the contour describes one associated with the denominator. Each vertex is connected by such interaction lines to every other vertex in the configuration, and since only one way to do this exists, each configuration generates exactly one diagram. The bare Monte Carlo implementation based on this expansion is illustrated in Fig. 4.5a.

### 4.4.3 Inchworm algorithm

The process of formulating an inchworm expansion is analogous to that of Sec. 4.3.3, but with the diagrammatic structure of the diabatic coupling expansion. Inchworm proper and improper diagrams are illustrated in Fig. 4.5b. The main difference is that whereas diagrams in the system–bath coupling expansion include interaction lines only between vertices paired within a particular pairing, the diabatic coupling expansion includes an interaction line between every two vertices. Therefore, there is only one "cluster" of vertices in every diagram, and that diagram is required to have at least one vertex in $[t_\uparrow, t_f]$. The only diagram not containing such a cluster is the order zero diagram (shown as (1) in Fig. 4.5b). This is also the only diagram containing an infinite number of bare diagrams: each diagram containing a cluster is completely identical to the one and only bare diagram that it represents.

The main advantages of the inchworm algorithm are therefore lost in the direct diabatic coupling scheme described here, and indeed we have verified that upon implementation of such an algorithm an exponential dynamical sign problem appears (not shown). However, in the remainder of this chapter, it will be shown that this problem can be circumvented by transforming the expansion to a cumulant form. From this perspective, a very useful inchworm algorithm then emerges.

## 4.5 Diabatic Coupling Cumulant Inchworm (DCCI) Expansion

As noted in Sec. 4.4, the diabatic coupling expansion in its direct Keldysh formulation has a peculiar diagrammatic structure in which each interaction vertex is directly connected to every other vertex. As such, this expansion does not significantly benefit from the inchworm algorithm, which relies on the ability to cut diagrams into weakly-connected subgraphs. We now show that by reformulating the diabatic coupling expansion in cumulant form,

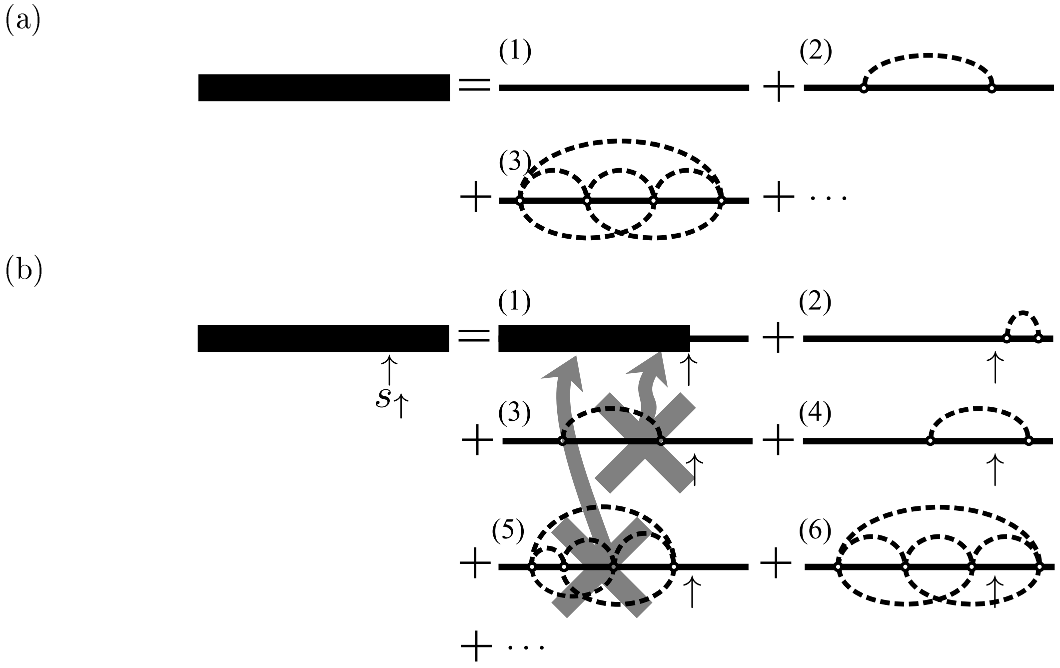


Figure 4.5: (a) Diagrams appearing in bare dQMC. The dashed curve (12) indicates an interaction line in either the numerator (above the contour) or the denominator (below it). Only one diagram corresponds to each configuration. (b) The naive inchworm scheme. Diagrams with no vertices after $s_\uparrow$ (such as (b.3) and (b.5)) are contained in the zeroth order term (b.1) and need not be summed. Other diagrams, such as (b.4) and (b.6), are analogous to those of the bare dQMC.

one obtains a formalism which is much more amenable to inchworm dQMC. The cumulant formalism has the additional advantage of being written in physical (rather than contour) time, such that in the absence of a sign problem the computation scales linearly with time, as will be demonstrated in the following discussion.

Since cumulants are most conveniently defined in terms of moments, the moment form of the expansion will first be presented. Cumulants and the cumulant inchworm algorithm will then be presented.

### 4.5.1 Moments

Consider the evaluation of the dynamics of an observable $O$ in terms of its *moments*, $\mu_m(\tau_1, ..., \tau_m)$. Given that we have Eq. (4.10), such that $\langle O(t) \rangle = \int \mathrm{d}\boldsymbol{s} \mathcal{O}(\boldsymbol{s})$, the observable can be written in terms of a moment expansion

$$\int \mathrm{d}\boldsymbol{s} \mathcal{O}(\boldsymbol{s}) = \sum_{m=0}^{\infty} \int \mathrm{d}\boldsymbol{\tau} \mu_m(\tau_1, \tau_2, \cdots, \tau_m). \tag{4.60}$$

While the integration $\int \mathrm{d}\boldsymbol{s}$ is performed over contour time, the integration $\int \mathrm{d}\boldsymbol{\tau}$ is performed over physical times $\tau_1 > \tau_2 > \cdots > \tau_m$, such that

$$\int \mathrm{d}\boldsymbol{\tau} = \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \cdots \int_0^{\tau_{m-1}} \mathrm{d}\tau_m. \tag{4.61}$$

An $m^{\text{th}}$-order moment $\mu_m(\tau_1, \tau_2, \cdots, \tau_m)$ is defined as

$$\mu_m(\tau_1, ..., \tau_m) = \sum_{\alpha_i \in \{+,-\}} \mathcal{O}\left(\mathcal{T}_c\left[\tau_1^{\alpha_1}, ..., \tau_m^{\alpha_m}\right]\right), \tag{4.62}$$

where $\mathcal{T}_c$ indicates contour time ordering and the $\alpha_i = \pm$ are the Keldysh branch indices. The moments are defined as functions of a set of real times, and the Keldysh branch indices are summed over. This is equivalent to simultaneously collecting the contributions from entire classes of path configurations associated with the real times, $\tau_1, ..., \tau_m$, as illustrated diagrammatically in Fig. 4.6a. Notably, it is exponentially expensive as a function of the order $m$ to evaluate moments in terms of diagrams, as an $m^{\text{th}}$ order moment is the sum of $2^m$ diagrams.

For the population operator $O = \sigma_z$ in the diabatic coupling expansion, the $0^{\text{th}}$ order moment is $\mu_0 = 1$ and odd moments vanish, $\mu_{2n+1} = 0$. The expectation value of $\sigma_z$ can therefore be written in terms of the even moments

$$\begin{aligned}\langle \sigma_z(t) \rangle = 1 &+ \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \mu_2(\tau_1, \tau_2) \\ &+ \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \int_0^{\tau_2} \mathrm{d}\tau_3 \int_0^{\tau_3} \mathrm{d}\tau_4 \\ &\times \mu_4(\tau_1, \tau_2, \tau_3, \tau_4) \\ &+ \dots .\end{aligned} \tag{4.63}$$

With the initial density matrix $|1\rangle \langle 1| e^{-\beta H_b}$ specified earlier, the second population moment simplifies to

$$\mu_2(\tau_1, \tau_2) = -4\Delta^2 \mathrm{Re}\left\{ e^{2i\epsilon(\tau_1 - \tau_2)} \mathcal{J}\left(0^+, \tau_2^+, \tau_1^+, 0^-\right) \right\}, \tag{4.64}$$

and the fourth moment to

$$\begin{aligned}\mu_4(\tau_1, \tau_2, \tau_3, \tau_4) = 4\Delta^2 \times \mathrm{Re}\Bigg\{ & \\ & e^{2i\epsilon(\tau_1 - \tau_2 + \tau_3 - \tau_4)} \mathcal{J}\left(0^+, \tau_4^+, \tau_3^+, \tau_2^+, \tau_1^+, 0^-\right) + \\ & e^{2i\epsilon(\tau_1 - \tau_2 - \tau_3 + \tau_4)} \mathcal{J}\left(0^+, \tau_4^+, \tau_3^+, \tau_1^-, \tau_2^-, 0^-\right) + \\ & e^{-2i\epsilon(\tau_1 - \tau_2 - \tau_3 + \tau_4)} \mathcal{J}\left(0^+, \tau_4^+, \tau_2^+, \tau_1^-, \tau_3^-, 0^-\right) + \\ & e^{-2i\epsilon(\tau_1 - \tau_2 + \tau_3 - \tau_4)} \mathcal{J}\left(0^+, \tau_3^+, \tau_2^+, \tau_1^+, \tau_4^-, 0^-\right) \Bigg\}.\end{aligned} \tag{4.65}$$

The diagrammatic description of moments is shown in Fig. 4.6.

Evaluating the moments within dQMC is therefore an alternative scheme for calculating dynamics. While linear in time (rather than quadratic, like the Keldysh formalism which involves two times), this expansion involves an additional exponential cost in the diagram order, due to the summation over the Keldysh indices. However, bare moment expansions typically converge very slowly if at all, and hold no real advantage over a direct calculation (though they may be of help with sign problems in certain cases[207]). It is therefore often advantageous to resum moments into cumulants. It turns out that a relationship exists between cumulant resummation and the inchworm algorithm, and this will be shown below.

### 4.5.2 Cumulants

Moment expansions can be immediately resummed into cumulant expansions in several ways. For the present purpose, it is advantageous to choose the chronological ordering prescription (COP) cumulant expansion[203], which yields the time-nonlocal equation of motion

$$\begin{aligned}\frac{\mathrm{d}\left\langle \sigma_z\left(t\right)\right\rangle}{\mathrm{d}t} = & \sum_{m=2}^{\infty}\int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1}\mathrm{d}\tau_2 ... \int_0^{\tau_{m-2}}\mathrm{d}\tau_{m-1} \\ & \times \gamma_m\left(t,\tau_1,...,\tau_{m-1}\right)\left\langle \sigma_z\left(\tau_{m-1}\right)\right\rangle .\end{aligned} \tag{4.66}$$

An advantage from the inchworm perspective is immediately apparent: the expression depends on the population at shorter times, such that previously calculated properties can perhaps be reused. The $m$-th order COP cumulant $\gamma_m\left(t,\tau_1,...,\tau_{m-1}\right)$ can be obtained by plugging Eq. (4.63) into both sides of Eq. (4.66) and equating terms of equal order. For

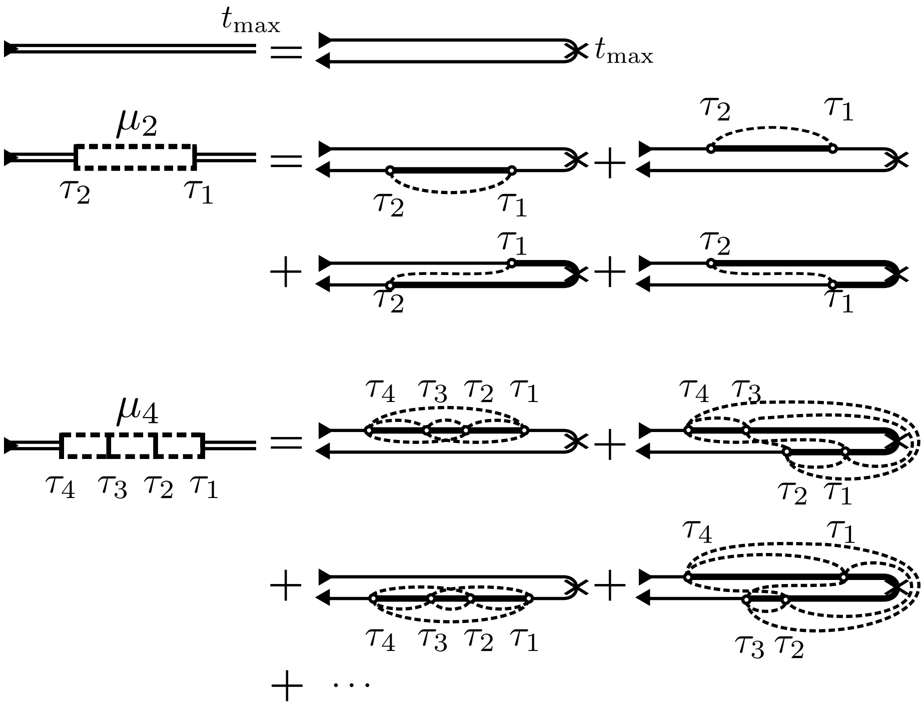


Figure 4.6: The real-time coordinate is represented by the thin double lines. The bare double line segment $[0, t_{\max}]$ corresponds to the bare propagator in the diabatic expansion on the Keldysh contour folded at $t_{\max}$. A $m^{\text{th}}$ order moment of a real time configuration $(\tau_1, \tau_2 \cdots, \tau_m)$ is illustrated as a dashed-edged box from $\tau_1$ to $\tau_m$ with solid vertical ticks at each configuration time. There are 4 distinct diagrams on the Keldysh contour associated with the $2^{\text{nd}}$ moment $\mu_2(\tau_1, \tau_2)$: $\boldsymbol{s} = \left(\tau_1^-, \tau_2^-\right), \left(\tau_2^+, \tau_1^+\right), \left(\tau_1^+, \tau_2^-\right), \left(\tau_2^+, \tau_1^-\right)$. These diagrams are plotted by connecting the vertices with the diabatic interaction lines as in Fig. (4.5). The $4^{\text{th}}$ moment contains $2^4$ diagrams on the contour. Here, we demonstrate only 4 example diagrams.

example,

$$\gamma_2\left(\tau_1,\tau_2\right)=\mu_2\left(\tau_1,\tau_2\right), \tag{4.67}$$

$$\gamma_4\left(\tau_1,\tau_2,\tau_3,\tau_4\right)=\mu_4\left(\tau_1,\tau_2,\tau_3,\tau_4\right)-\mu_2\left(\tau_1,\tau_2\right)\mu_2\left(\tau_3,\tau_4\right), \tag{4.68}$$

and $\gamma_{2n-1}=0$. A general $m$-th order cumulant, $\gamma_m$, can be obtained recursively:

$$\begin{aligned}\gamma_m\left(\tau_1,...,\tau_m\right)=&\sum_{p\in\mathcal{P}_m}-\left(-1\right)^{|p|}\times\\&\prod_{(i_1,i_2,\cdots,i_{2n})\in p}\mu_{2n}\left(\tau_{i_1},\tau_{i_2}\cdots,\tau_{i_{2n}}\right).\end{aligned} \tag{4.69}$$

The set $\mathcal{P}_m$ describes all possible ways of partitioning a sequence of integers $1,2,\ldots,m$ into subsequences of adjacent numbers, each having an even number elements. Each partition $p\in\mathcal{P}_m$ can be represented by a set of ordered tuples $(i_1,i_2,\ldots,i_{2n})$ corresponding to one subsequence, and $|p|$ is the number of subsequences within the partition. For instance,

$$\mathcal{P}_2=\left\{\left\{(1,2)\right\}\right\},$$

$$\mathcal{P}_4=\left\{\left\{(1,2,3,4)\right\},\left\{(1,2),(3,4)\right\}\right\},$$

$$\begin{aligned}\mathcal{P}_6=&\left\{\left\{(1,2,3,4,5,6)\right\},\left\{(1,2),(3,4,5,6)\right\},\right.\\&\left.\left\{(1,2,3,4),(5,6)\right\},\left\{(1,2),(3,4),(5,6)\right\}\right\}.\end{aligned}$$

The diagrammatic description of COP cumulants in terms of moments is shown in Fig. 4.7. A cumulant of any given order can be expressed in terms of moments up to and and including the same order.

### 4.5.3 Naive inchworm algorithm

The dynamics of $\langle\sigma_z\left(t\right)\rangle$ within the COP cumulant expansion can be evaluated by dQMC. To simplify the notation, it is convenient to redefine the times $t,\tau_1,...,\tau_{m-1}$ as $\tau_1,\tau_2...,\tau_m$,

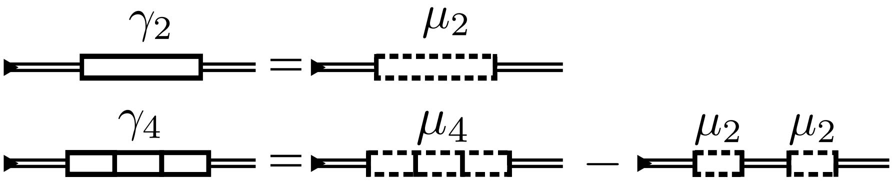


Figure 4.7: The COP cumulants of a real-time configuration $(\tau_1, \tau_2 \cdots, \tau_m)$ are illustrated as a solid-edged box with vertical ticks at each configuration time. Here, we show the diagrammatic representation of Eq. (4.67) and (4.68), which illustrate the 2nd and 4th cumulants in terms of the moments.

respectively; these obey the physical time ordering $\tau_1 > \cdots > \tau_m$. As before, the times will be indicated by the vector quantity $\boldsymbol{\tau}$ when possible. By carrying out the integration $\int_0^t \mathrm{d}\tau_1$ on both sides of Eq. (4.66), an expression for $\langle \sigma_z(t) \rangle$ in terms of itself is obtained:

$$\langle \sigma_z(t) \rangle = 1 + \int_0^t \mathrm{d}\boldsymbol{\tau} \mathcal{K}(\boldsymbol{\tau}), \tag{4.70}$$

$$\mathcal{K}(\boldsymbol{\tau}) = \gamma_m(\tau_1, ..., \tau_m) \langle \sigma_z(\tau_m) \rangle . \tag{4.71}$$

Since $\gamma_{2n-1} = 0$, the path integration $\int_0^t \mathrm{d}\boldsymbol{\tau}$ can be explicitly written as

$$\int_0^t \mathrm{d}\boldsymbol{\tau} = \sum_{m \in \text{even}, \geq 2} \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 ... \int_0^{\tau_{m-1}} \mathrm{d}\tau_m. \tag{4.72}$$

Since the functional $\mathcal{K}(\boldsymbol{\tau})$ depends on $\langle \sigma_z(t_m) \rangle$, the observable is evaluated at the smallest time in the configuration $\boldsymbol{\tau}$. Since this is the quantity being evaluated, it is not known to begin with and there is no bare expansion of the COP type. However, it is straightforward to implement a simple inchworm algorithm: assume $\langle \sigma_z(\tau) \rangle$ is known for all $\tau \in [0, \tau_\uparrow]$. The expectation value at $t > \tau_\uparrow$ can then be expressed as:

$$\langle \sigma_z(t) \rangle = \langle \sigma_z(\tau_\uparrow) \rangle + \int_{\tau_\uparrow}^t \mathrm{d}\boldsymbol{\tau} \mathcal{K}(\boldsymbol{\tau}) . \tag{4.73}$$

Here $\int_{\tau_\uparrow}^{t} \mathrm{d}\boldsymbol{\tau}$ represents the path integral

$$\int_{\tau_\uparrow}^{t} \mathrm{d}\boldsymbol{\tau} = \sum_{m=2}^{\infty} \int_{\tau_\uparrow}^{t} \mathrm{d}\tau_1 \int_{0}^{\tau_1} \mathrm{d}\tau_2 ... \int_{0}^{\tau_{m-1}} \mathrm{d}\tau_m, \tag{4.74}$$

which describes integration over the configuration subspace for which at least one $\tau_1$ is within the interval $[\tau_\uparrow, t]$. This defines a formally exact inchworm step, which appears to leverage knowledge of $\langle \sigma_z(\tau) \rangle$ for times up to $\tau_\uparrow$ in order to obtain the same observable for the final time $t$. Examples of diagrams appearing in this expansion are shown in Fig. 4.8. Diagrams in which the rightmost time index is to the left of $\tau_\uparrow$ (crossed out in the figure) are included in the $0^{\text{th}}$ order contribution (diagram (1) in Fig. 4.8) and need not be summed.

Unfortunately, the inchworm step we have just described cannot be implemented as it stands, and has been introduced chiefly for didactic purposes. This is because it includes contributions where $\langle \sigma_z(\tau) \rangle$ is needed at time argument $\tau > \tau_\uparrow$. Two examples are overlaid with a question mark in Fig. 4.8. Such contributions are unknown and must be dropped from the expansion, leading to an error the magnitude of which can be shown to be linear in $\Delta t = t - \tau_\uparrow$. In practice, this makes convergence of the algorithm to the exact result (by progressively reducing the size of the inching time step $\Delta t$) very hard to achieve. However, as the next subsection shows, this issue can be completely overcome by taking a closer look at the structure of the diagrams.

### 4.5.4 Cumulant inchworm algorithm

It is now necessary to solve the problem raised in the previous subsection, namely the fact that one is unable to evaluate diagrams from configurations having $\tau_m > \tau_\uparrow$ for some $m$. To do so, it is possible to first *unwind* the resumming done implicitly by the cumulant

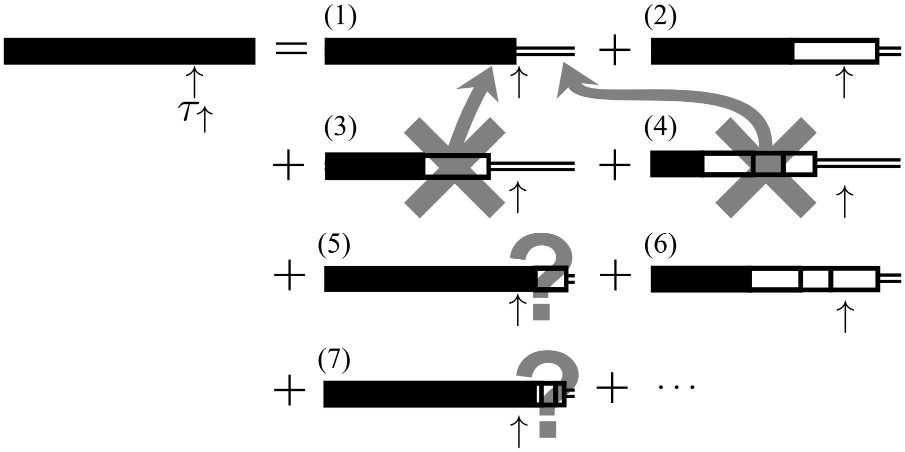


Figure 4.8: Diagrammatic representation of the naive prescription of the inchworm algorithm, Eq. (4.73). The solid-edged boxes with vertical ticks are the COP cumulants as shown in Fig. (4.6). The $\tau_\uparrow$ is indicated as the ↑ on the physical time coordinate. Each configuration corresponds to one single diagram. Diagrams (3) and (4) have all cumulant boxes lying in the known region (to the left of ↑) and are considered been included in diagram (1) for this inchworm step. The cumulant boxes in diagrams (2) and (6) straddle the $\tau_\uparrow$ time and their contribution can be calculated by Eq. (4.71). Diagrams (5) and (7) have all cumulant boxes located to the right of the ↑, are unknown for this inchworm step in the naive version.

expansion, then reintroduce it wherever possible. To see how this works, one inserts the functional Eqs. (4.70) and (4.71). This gives

$$\begin{aligned}\langle\sigma_z(t)\rangle =& 1+\int_0^t \mathrm{d}\boldsymbol{\tau}\gamma_m(\boldsymbol{\tau})\\ &+\int_0^t \mathrm{d}\boldsymbol{\tau}\gamma_m(\boldsymbol{\tau})\int_0^{\tau_m}\mathrm{d}\boldsymbol{\tau}'\mathcal{K}(\boldsymbol{\tau}')\end{aligned}\tag{4.75}$$

and we sample an additional configuration $\boldsymbol{\tau}'$ for the integration $\int_0^{\tau_m}\mathrm{d}\boldsymbol{\tau}'$. This can be iterated any number of times, generating an expansion in terms of multiple cumulants, with the population pushed to increasingly high-order terms. Examples of terms appearing in this unwound cumulant expansion are shown in Fig. 4.9a. Each configuration may yield more than one diagram: 2-time configurations gives one 2nd order diagram (a.2); 4-time configurations yield diagrams (a.3) and (a.4) corresponding to 2 partitions in $\mathcal{P}_4$; 6-time configurations contain diagrams (a.5)–(a.8) corresponding to 4 partitions in $\mathcal{P}_6$. The term "wound / wind" is used to distinguish this procedure from "dressing / dress" used in the context of Dyson equations, and in particular to distinguish "unwound" from "bare."

The unwound expansion can be written as

$$\langle \sigma_z(t) \rangle = 1 + \int \mathrm{d}\boldsymbol{\tau} \Gamma(\boldsymbol{\tau}), \tag{4.76}$$

where the functional $\Gamma$ depends only on the COP cumulants. At a general (even) order $m$, $\Gamma$ contains terms of various partitions $\mathcal{P}_m$, as introduced in Sec. 4.5.2:

$$\Gamma(\tau_1, ..., \tau_m) = \sum_{p \in \mathcal{P}_m} \prod_{(i_1, i_2, \cdots, i_{2n}) \in p} \gamma_{2n}(\tau_{i_1}, \tau_{i_2} \cdots, \tau_{i_{2n}}). \tag{4.77}$$

For instance, the lowest order functional ($m = 2$) is simply

$$\Gamma(\tau_1, \tau_2) = \gamma_2(\tau_1, \tau_2), \tag{4.78}$$

while that with $m = 4$ contains two terms originating from the iteration procedure:

$$\Gamma(\tau_1, \tau_2, \tau_3, \tau_4) = \gamma_4(\tau_1, \tau_2, \tau_3, \tau_4) + \gamma_2(\tau_1, \tau_2)\gamma_2(\tau_3, \tau_4). \tag{4.79}$$

Unlike the original bare expansion in diabatic coupling, each configuration now generates multiple diagrams (corresponding to partitions). For instance, as we show in Fig. 4.9a, a 4th order configuration generates 2 diagrams, (a.3) and (a.4), and a 6th order configuration generates 4 diagrams, (a.5)–(a.8). We note briefly that it is easy to show that the unwound expansion corresponds exactly to the moment expansion, in the sense that $\Gamma_i = \mu_i$. However, the advantages of the unwound representation will immediately become apparent.

The unwinding completely removes the dependence on the population $\langle \sigma_z (\tau) \rangle$, but does so at the cost that the resummation properties of COP expansion are lost. We now *partially rewind* the series wherever this does not interfere with the assumption of the inchworm step, in particular the fact that we only have access to populations for $\tau < \tau_\uparrow$. The inchworm step is performed by stochastically sampling configurations $\boldsymbol{\tau} = (\tau_1, ..., \tau_m) \in [0, t]$, as before

$$\langle \sigma_z (t) \rangle = \langle \sigma_z (\tau_\uparrow) \rangle + \int_{\tau_\uparrow}^{t} \mathrm{d}\boldsymbol{\tau} \mathcal{K}' (\boldsymbol{\tau}) . \tag{4.80}$$

For each configuration, one sums only diagrams stemming from a *proper* subset of the partitions, $\mathcal{P}'_m \subseteq \mathcal{P}_m$, obtained by excluding partitions with subsequences (parts) having all times in $[0, \tau_\uparrow]$. With this, we define

$$\Gamma' (\boldsymbol{\tau}) = \sum_{p \in \mathcal{P}'_m} \prod_{(i_1, i_2, \cdots, i_{2n}) \in p} \gamma_{2n} (\tau_{i_1}, \tau_{i_2} \cdots, \tau_{i_{2n}}) , \tag{4.81}$$

such that the functional to be summed is

$$\mathcal{K}' (\boldsymbol{\tau}) = \begin{cases} \Gamma' (\boldsymbol{\tau}) \langle \sigma_z (\tau_m) \rangle & \text{if } \tau_m < \tau_\uparrow, \\ \Gamma' (\boldsymbol{\tau}) \langle \sigma_z (\tau_\uparrow) \rangle & \text{if } \tau_m > \tau_\uparrow. \end{cases} \tag{4.82}$$

The diagrammatic representation of this cumulant inchworm expansion is illustrated in Fig. 4.9b, where three examples of improper partitions (diagrams) are crossed out. Diagram (b.3) is included in diagram (b.1); diagram (b.4) is included in diagram (b.2); diagram (b.7) is included in diagram (b.5). Note that the contribution of diagram (b.8) takes into account the missing diagrams by the naive inchworm algorithm.

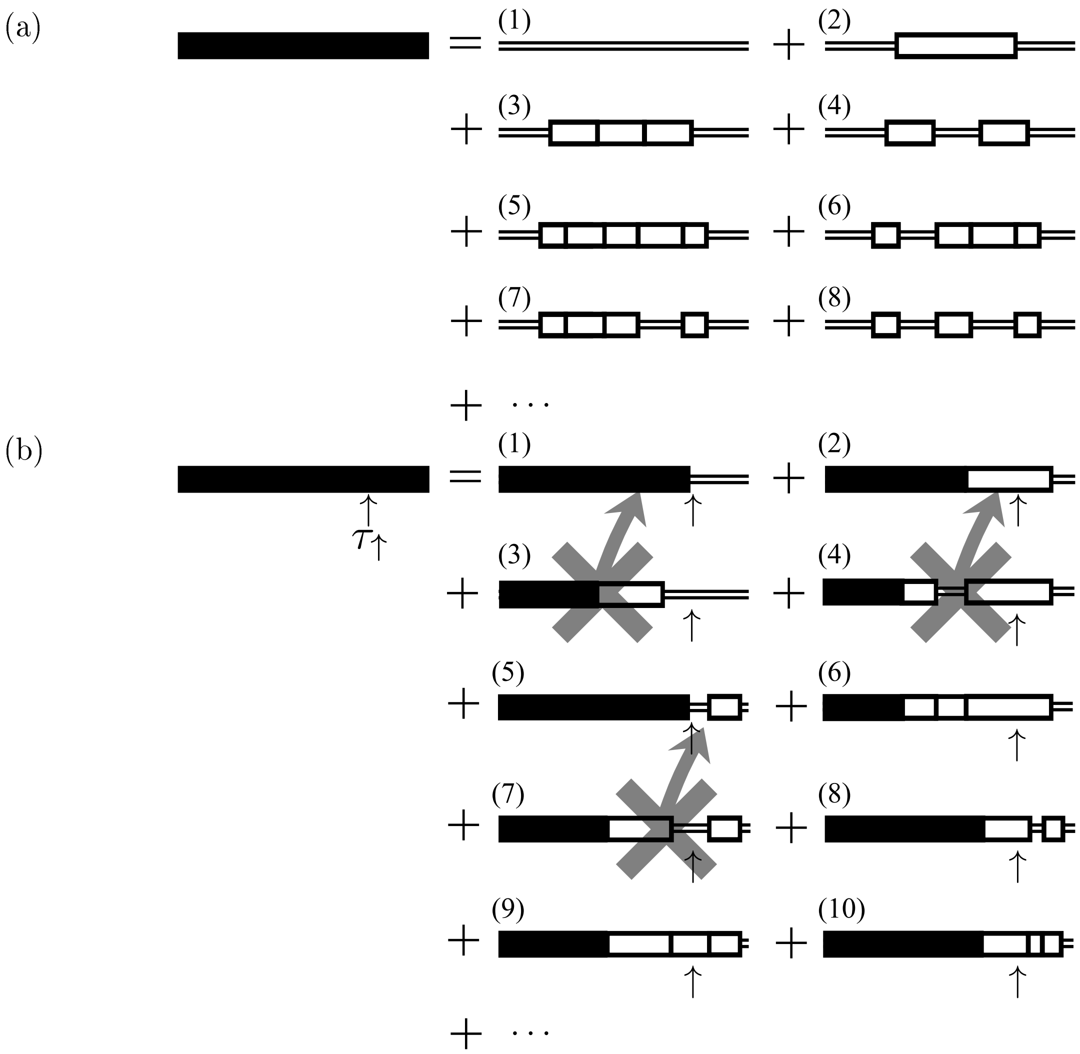


Figure 4.9: (a) The unwound dQMC expression for the full cumulant expansion. The thick solid lines are the exact dynamics of expectation value and the thin double lines are the unperturbed value 1 within the diabatic expansion. The solid-edged boxes with vertical ticks are the COP cumulants as shown in Fig. (4.6). (b) The cumulant inchworm algorithm. Any diagram that has a stand alone part (a cumulant box) to the left of the ↑ has been included in the other diagrams and needs to be neglected in the inchworm step.

To justify that the cumulant inchworm expansion is formally equivalent to the unwound expansion, it must be shown that the two sets of diagrams generated by respective expansions are identical. To do so, we have to prove that (a) these two sets of diagrams contain each other, in the sense that every unwound diagram in one set is *represented* in the other; and (b), each diagram in one set is mapped to only a *single* diagram in the other set (such that the measure is conserved under summation). We will proceed by example, rather than presenting a formal proof.

To argue point (a), we need to show a containment relationship in both directions. First, any cumulant inchworm diagram generates only diagrams contained in the set of unwound diagrams. This is trivial since the thick solid segment in each cumulant inchworm diagram can be considered an infinite sum of unwound diagrams within that segment. In the reverse direction, any unwound diagram can be found in the set of cumulant inchworm diagrams: given an unwound diagram, one can construct a cumulant inchworm diagram containing it by Eqs. (4.81) and (4.82). As an example, we consider the lowest order in Fig. 4.9b with a $2^{\text{nd}}$ order configuration $\boldsymbol{\tau} = (\tau_1, \tau_2)$. The configuration generates one unwound diagram of the (a.2) type. For the same configuration's cumulant inchworm diagram, three cases are possible: $\tau_1 > \tau_\uparrow > \tau_2$, $\tau_\uparrow > \tau_1 > \tau_2$, and $\tau_1 > \tau_2 > \tau_\uparrow$, which correspond to diagrams (b.2), (b.3), and (b.5), respectively. It is clear that diagram (b.3) is improper and has been included in diagram (b.1). Thus, an unwound diagram of the (a.2) type is contained in (b.2), (b.3), or (b.1) depending on its relationship with $\tau_\uparrow$.

Point (b) requires unique correspondence in both directions. One direction is trivial—each cumulant inchworm diagram can be written as an infinite sum of unique unwound diagrams. In the other direction, we need to show that if there exist two cumulant inchworm diagrams which contain the same unwound diagram, one of these two cumulant inchworm diagrams must be eliminated. The propriety of cumulant inchworm diagrams ensures this uniqueness: consider a $4^{\text{th}}$-order unwound diagram of type (a.4) with configuration $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3, \tau_4)$. If $\tau_1 > \tau_\uparrow > \tau_2$, the unwound diagram could in principle be contained in two (not necessarily proper) cumulant inchworm diagrams, (b.2) and (b.4). Diagram (b.4) is then eliminated by the requirement of propriety. Similarly, if $\tau_2 > \tau_\uparrow > \tau_3$, the unwound diagram could be contained in two cumulant inchworm diagrams, (b.5) and (b.7), but (b.7) is improper and therefore can be eliminated. For other cases, the uniqueness is trivial: there is only one (necessarily proper) cumulant inchworm diagram containing the unwound diagram. For example, if $\tau_3 > \tau_\uparrow > \tau_4$, only diagram (b.8) can contain it.

With points (a) and (b) justified, it is clear that an exact correspondence exists between the cumulant inchworm expansion and the unwound expansion. Every diagram in the cumulant inchworm expansion corresponds to an infinite number of unwound diagrams, and while the expansion does not perform resummation over the entire length of the contour like the system–bath coupling expansion, it also has the distinct advantage of scaling linearly in time. It therefore constitutes a highly efficient method which is complementary to the system–bath coupling inchworm approach.

This method will be referred to as the Diabatic Coupling Cumulant Inchworm (DCCI) approach in the following.

## 4.6 Convergence Estimation

To have an estimate of how rapidly the inchworm approaches are expected to converge in different regions of parameter space, we focus on the lowest-order nontrivial contribution in each type of expansion and determine its magnitude as a function of model parameters. We consider the $2^{\text{nd}}$-order term of the SBCI and DCCI expansions, which in both cases can be written in the form

$$G_2(t) = \int_0^t \mathrm{d}t_1 \int_0^{t_2} \mathrm{d}t_2 C(t_1, t_2). \tag{4.83}$$

Here, $C(t_1, t_2)$ is the bath correlation function associated with each expansion. Loosely speaking, one might expect an expansion to converge rapidly as long as the corresponding $G_2(t)$ is not significantly greater than unity. Given the functional form of the Debye spectral density, we can easily estimate $G_2(t)$ in a semi-analytical fashion.

For the SBCI expansion, we can evaluate $G_2(t)$ in the high and low temperature limits and then derive the convergence conditions from the appropriate dimensionless parameters that emerge. This scheme is analogous to one which has been used to determine the limitations of Redfield theory[47]. The bath correlation function in the SBCI expansion is given by $C(t_1, t_2) = \left\langle \widetilde{B}(t_1) \widetilde{B}(t_2) \right\rangle_b$ where $\widetilde{B}(t) = \sum_\ell c_\ell \widetilde{x}_\ell(t)$, and with the definition $\langle \cdot \rangle_b \equiv \mathrm{Tr}_b\{\rho_b \cdot\}$. The integral takes the form $G_2(t) = \xi g(t)$, where $\xi$ is a dimensionless parameter and $g(t)$ is a time-dependent dimensionless function. We expect the expansion to converge rapidly as long as $\xi \lesssim 1$. In the high temperature limit $\beta\omega_c \ll 1$, and we can approximate $\coth\left(\frac{\beta\omega}{2}\right) \approx \frac{2}{\beta\omega}$ and obtain the dimensionless form for $\xi$:

$$\xi = \frac{\lambda}{\beta\omega_c^2} - i\frac{\lambda}{2\omega_c}. \tag{4.84}$$

Thus, in this regime, an estimate for the condition for convergence of the SBCI approach is

$$\frac{\lambda}{\beta\omega_c^2} \lesssim 1. \tag{4.85}$$

In the low temperature limit $\beta\omega_c \gg 1$, we can use $\coth\left(\frac{\beta\omega}{2}\right) \approx 1$, but cannot carry out the integral analytically. In the same spirit, we factor out the dimensionless scale of the integral, $\mathrm{Re}\left[G_2(t)\right] = \frac{\lambda}{\pi\omega_c}\int \mathrm{d}x \frac{1}{x^2+1}\frac{1}{x}\left(1-\cos\omega_c x t\right)$, which yields a convergence condition for the low temperature limit

$$\frac{\lambda}{\pi\omega_c} \lesssim 1. \tag{4.86}$$

It is noteworthy that since $G_2(t)$ is proportional to $\lambda\coth(\beta\omega/2)$, the SBCI expansion becomes more difficult to converge as $\lambda$ increases or $\beta$ decreases.

The explicit form of the bath correlation function in the DCCI expansion is given by $C(t_1, t_2) = e^{-4\mathcal{Q}_2(t_1-t_2) - i4\mathcal{Q}_1(t_1-t_2)}$, where

$$\mathcal{Q}_1(t) = \frac{2}{\pi}\int \mathrm{d}\omega \frac{J_D(\omega)}{\omega^2}\sin\omega t, \tag{4.87}$$

$$\mathcal{Q}_2(t) = \frac{2}{\pi}\int \mathrm{d}\omega \frac{J_D(\omega)}{\omega^2}\coth\left(\frac{\beta\omega}{2}\right)\left(1-\cos\omega t\right). \tag{4.88}$$

Due to the complicated form of these correlation functions, one cannot obtain an analytical expression to extract a dimensionless scale parameter. Therefore, we evaluate the integral numerically at a large enough time for given model parameters. We also note that $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are linearly dependent on $\lambda$, which yields a $1/\lambda^2$ dependence for $G_2(t)$. Therefore, the DCCI expansion becomes easier to converge as $\lambda$ increases.

A two-dimensional "phase diagram" can be drawn as cuts of the full parameter space with varying $\lambda$ and $\omega_c$, shown in Fig. 4.10. Here we limit the discussion to the subspace with zero energy bias $\epsilon = 0$. The horizontal axis is the scaled reorganization energy $(\lambda/\Delta)$ in

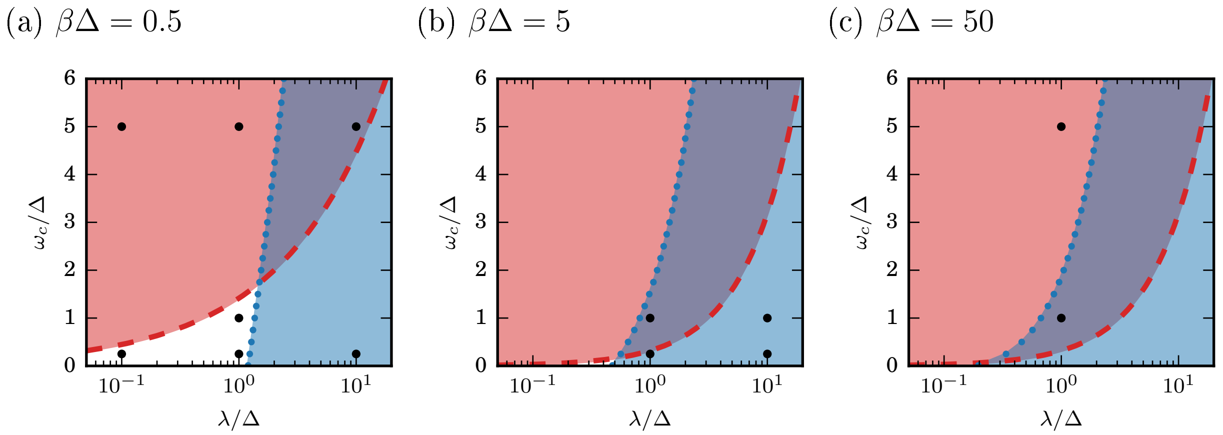


Figure 4.10: Spin–boson model parameter space with zero bias $\epsilon = 0$. The $x$-axis is $\lambda/\Delta$ in log scale and the $y$-axis is $\omega_c/\Delta$ in linear scale. The bath temperatures are (a) $\beta\Delta = 0.5$, (b)$\beta\Delta = 5$, and (c) $\beta\Delta = 50$. In each "phase diagram", the estimated region of rapid convergence for the SBCI approach is to the left of the dashed line (red) and is to the right of the dotted line (blue) for the DCCI approach. Points indicate the parameters for plots presented in this work.

log scale and the vertical axis is the scaled cutoff frequency ($\omega_c/\Delta$). Within this coordinate system, we can demarcate the estimated region of facile convergence for the SBCI and DCCI expansions by the conditions given above. The red regions indicate the subspace satisfying Eq. (4.85) and (4.86), in which the SBCI approach is expected to converge rapidly. The blue regions are obtained by semi-analytical estimation of the analogous condition for the DCCI approach.

Fig. 4.10 exhibits these complementary regions and shows that their combined area covers much of the relevant parameter space. We will briefly point out some important features of the phase diagram. First, for any cutoff frequency $\omega_c$, the SBCI converges better in the small

$\lambda$ direction while the DCCI is expected to work better as $\lambda$ increases. Second, the region of utility for the SBCI expansion shrinks in the adiabatic regime (small $\omega_c$), which is due to the fact that the correlation functions in the SBCI expansion have a longer correlation time when $\omega_c$ is small. Lastly, as the temperature decreases, the regions of rapid convergence of both the SBCI and DCCI approaches expand and cover almost the entire parameter space.

While Fig. 4.10 provides an illustration of applicable regions of our approach, the regions are determined by rough estimation of lowest order contribution. In principle, our inchworm expansions are numerically exact in the entire parameter space, as discussed in previous sections. In the "uncovered" or white region, our approaches should continue yield exact dynamical behavior at least on some time scales, albeit with potentially much greater numerical effort.

## 4.7 Results

In the following, we present a detailed comparison of our new approaches to established benchmarks, as well as a discussion of the relative benefits and drawbacks of our approach in comparison to established methods.

### 4.7.1 Computational Methodology

Each inchworm step is limited to a fixed run time and the order of each individual inchworm diagram is restricted to a maximum order $M$. We use $\mathrm{d}t\Delta = 0.1$ for the size of the inchworm step in the following calculation, unless otherwise specified. One may then check for convergence by systematically increasing $M$, decreasing $\mathrm{d}t$ and increasing the number of

Monte Carlo samples[106]. The SBCI calculation requires the full information contained in two-time restricted propagators, thus for the SBCI propagation to a time $t = N\mathrm{d}t$ requires $N^2$ inchworm steps (in fact, by taking advantage of time-reversal symmetry and the contour ordering of the time arguments, the number of steps needed turns out to be $\sim \frac{1}{4}N^2$). On the other hand, the DCCI expansion is phrased solely in terms of single-time properties, such that it requires only $N$ inchworm steps to reach a simulation time $t = N\mathrm{d}t$. For both approaches, we perform multiple inchworm calculations in order to properly account for error propagation[106].

We compare our calculations with several existing numerically exact methods, including the quasi-adiabatic propagator path integral (QUAPI) approach[95–98], hierarchical equations of motion (HEOM) method[45, 93, 94], and the multi-configuration time-dependent Hartree (MCTDH) approach[87, 88, 144]. QUAPI is based on the discretization of influence functional for reduced propagation on the Keldysh contour. The maximum number of short-time propagators that the path integral spans is determined by a parameter $k_{\max}$, which governs the the memory length. The approach becomes difficult to converge when the memory length is long. The HEOM approach introduces a hierarchy of auxiliary density matrices and employs a Mastsubara expansion for the bath density matrix. The hierarchy truncation level $L$ and number of Matsubara terms $K$ are numerical parameters that are tuned to converge the HEOM calculation. A standard, highly parallel implementation is available[94], known to be accurate in the high temperature limit and for the Debye spectral density. Generically, the HEOM approach has more difficulty for low temperatures and non-Debye spectral densities. The MCTDH approach is based on the expansion of the interacting many-body wave function as a tensor product of wavefunctions defined in a convenient set of orbitals. A highly efficient protocol may then be used to control, in a time-dependent manner, the number of orbitals needed for exact convergence. Exact MCTDH results for the spin–boson model are reported in Ref. 144.

We will be using the benchmarks to investigate accuracy, and will make no attempt to compare numerical efficiency beyond general points having to to with the computational scaling of the algorithms. To provide a general idea, we will say that using our current implementation, most of the (linear scaling in time) DCCI results presented here can be comfortably obtained on a laptop in minutes to hours, whereas the (quadratically scaling in time) SBCI results typically require a small cluster. However, it should be noted that the data below was obtained with a very flexible but not at all optimized code written in the high-level Python programming language. From our experience with similar algorithms for the Anderson impurity model[106], we estimate that 1-2 orders of magnitude in overall runtime could be achieved by writing an efficient code, or simply by switching to a compiled language.

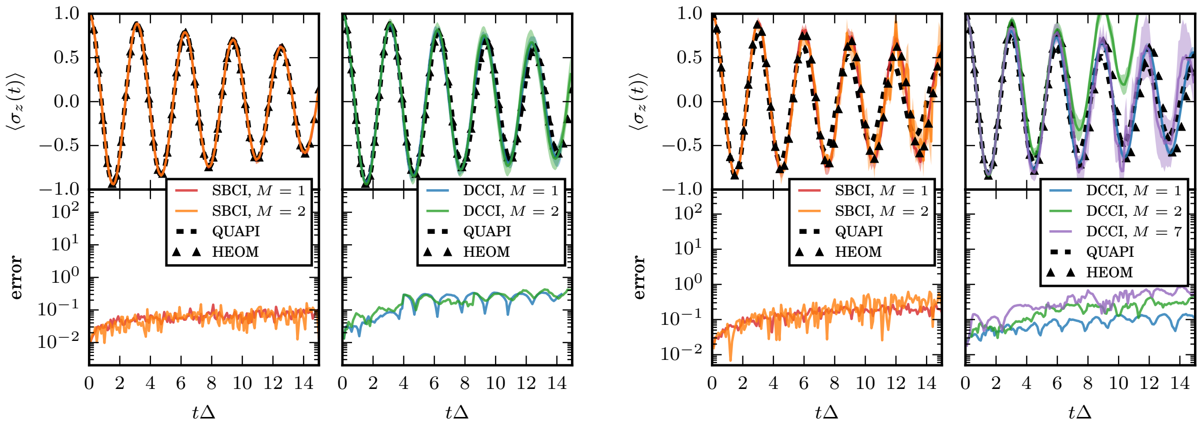
### 4.7.2 High Temperature Regime

We start our comparison of the inchworm approaches with other exact methods in the high temperature regime (Fig. (4.10)(a)), specifically $\beta\Delta = 0.5$ ($k_BT/\Delta = 2$), and consider the vertical cuts at weak coupling $\lambda/\Delta = 0.1$, strong coupling $\lambda/\Delta = 10$, and intermediate coupling $\lambda/\Delta = 1$ in the following.

### Weak coupling

In the weak system–bath coupling regime, we consider cases with scaled reorganization energy ($\lambda/\Delta = 0.1$) where we expect the SBCI expansion to converge more easily than the DCCI expansion. In Fig. 4.11, we find that the lowest order ($M = 1$) results for the SBCI expansion always gives a quantitative account of the dynamics with the error remaining nearly constant over the simulation time. The SBCI result also converges rapidly upon increasing the maximum order $M$ of each inchworm step. We note that a smaller cut-off frequency yields a greater statistical error (see the lower panels of Fig. 4.11(a) and (b)) with the same computational cost. This is due to the long correlation time induced by a small $\omega_c$, which makes it more difficult to converge the SBCI expansion. On the other hand, the DCCI calculation also yields surprisingly accurate results. However, for a small cut-off frequency, it becomes more difficult to converge the DCCI approach, as can be seen in the right panel of Fig.4.11(b). Note that for the DCCI approach, the $M = 2$ case actually yields results that are less accurate than $M = 1$. This lack of convergence is due to the carrying of the large short-time errors to longer times by the inchworm algorithm, and could in principle be overcome by a larger investment of computer time or a faster code.

Thus, in the high temperature, weak coupling regime, both inchworm approaches appear capable of reproducing the results obtained by the HEOM method, which easily converges to the exact answer for the Debye spectral density at high temperatures. The DCCI approach does show some convergence difficulties in this regime for the slow bath case. We could not converge QUAPI in the slow bath regime, and quantitative discrepancies can be found between QUAPI and HEOM here, as seen in Fig. 4.11.



(a) $\lambda/\Delta = 0.1$, $\omega_c/\Delta = 5$

(b) $\lambda/\Delta = 0.1$, $\omega_c/\Delta = 0.25$

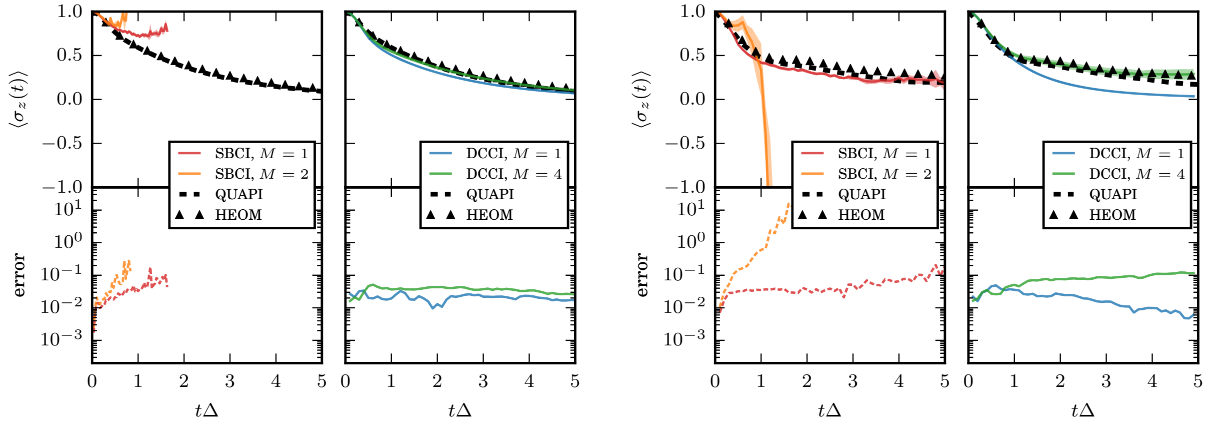
Figure 4.11: Nonequilibrium population difference $\langle\sigma_z(t)\rangle$ (top subplots) and corresponding error estimates (bottom subplots) as a function of time in the weak coupling ($\lambda/\Delta = 0.1$) and high temperature ($\beta\Delta = 0.5$) regime. The bias energy is $\epsilon = 0$. The results calculated by the SBCI (left panels, red and orange) and DCCI (right panels, blue and green) inchworm expansions are plotted for (a) a non-adiabatic (fast) bath with $\omega_c/\Delta = 5$, and (b) an adiabatic (slow) bath with $\omega_c/\Delta = 0.25$. Maximum order for an inchworm step is indicated by $M$. The thickness of the Monte Carlo results results from our error estimates. The dashed lines are the QUAPI results with (a)$\Delta t = 0.1$, $k_{\max} = 6$ and (b)$\Delta t = 0.1$, $k_{\max} = 12$. The triangles indicate the HEOM result with $K = 2$ and $L = 20$.

### Strong coupling

For strong system–bath coupling ($\lambda/\Delta = 10$), we anticipate that the SBCI expansion will be difficult to converge and the DCCI expansion will show rapid convergence. The right panels of Fig. 4.12(a) and (b) show that the DCCI results converge to accurate population dynamics as we increase the maximum order $M$ of each inchworm step, but that at least $M = 4$ is required for convergence. As expected, the SBCI expansion is difficult to converge in this parameter regime. The origins of this convergence issue can be gleaned from the behavior of the error estimate. In particular, the error estimates found in the left panels in Fig. 4.12 show the statistical error for one single SBCI calculation, which indicates the error of the Monte Carlo estimation of the integral within each inchworm step. This is an underestimate of the error margin, as it does not take into account the error propagation from shorter times; other plots in this section show the full error analysis. We note that even the single run error increases exponentially with time, so that it is clear that the origin of the exponential growth in noise to signal ratio is actually the Monte Carlo and not error propagation. Within an inchworm step of finite size $\mathrm{d}t\Delta = 0.1$, the weight of high order configurations to the integral becomes large when $\lambda$ increases. To capture these high order contributions, one may increase $M$, however, as shown in Fig. 4.12, the slope of the statistical error grows unfavorably in this case as we increase $M$, rendering the SBCI expansion difficult to converge.

### Intermediate coupling

Lastly, we focus on the intermediate system–bath coupling regime where the scaled reorganization energy is $\lambda/\Delta = 1$. Fig. 4.13 exhibits a general feature of the inchworm approaches: convergence with respect to the maximum order becomes more difficult to obtain as the cut-off frequency decreases. For a fast bath ($\omega_c/\Delta = 5$), both the SBCI and DCCI expansions yield quite accurate results at lowest order. For $\omega_c/\Delta = 1$, the parameter set $(\lambda/\Delta, \omega_c/\Delta) = (1, 1)$ is located outside of the "safe" regions for the SBCI and DCCI as demarcated in Fig. 4.10(a). Here we observe clear, but small, discrepancies between the SBCI/DCCI results for $M = 1$ and numerically exact dynamics. By systematically in-



(a) $\lambda/\Delta = 10$, $\omega_c/\Delta = 5$

(b) $\lambda/\Delta = 10$, $\omega_c/\Delta = 0.25$

Figure 4.12: Nonequilibrium population difference $\langle\sigma_z(t)\rangle$ (top subplots) and corresponding error estimates (bottom subplots) as a function of time in the strong coupling ($\lambda/\Delta = 10$) and high temperature ($\beta\Delta = 0.5$) regime. The bias energy is $\epsilon = 0$. The results calculated by the SBCI (left panels, red and orange) and DCCI (right panels, blue and green) inchworm expansions are plotted for (a) a non-adiabatic (fast) bath with $\omega_c/\Delta = 5$, and (b) an adiabatic (slow) bath with $\omega_c/\Delta = 0.25$. The error estimate of the SBCI calculation is for one single run. Maximum order for a inchworm step is indicated by $M$. The thickness of the Monte Carlo results results from our error estimates. The dashed lines are the QUAPI results with (a)$\Delta t = 0.1$, $k_{\max} = 6$ and (b)$\Delta t = 0.3$, $k_{\max} = 11$. The triangles indicate the HEOM result with $K = 2$ and $L = 20$.

creasing $M$, the discrepancies can be corrected. When the cut-off frequency is small, the parameter set $(\lambda/\Delta, \omega_c/\Delta) = (1, 0.25)$ is particularly difficult for both SBCI and DCCI expansions, although convergence is still seen for $M = 6$. Lastly, note that here, as in Fig. 4.12, some notable discrepancies exist between the HEOM and QUAPI results. The inchworm expansions always converge to the HEOM results, which are expected to be more accurate in the high temperature regime.

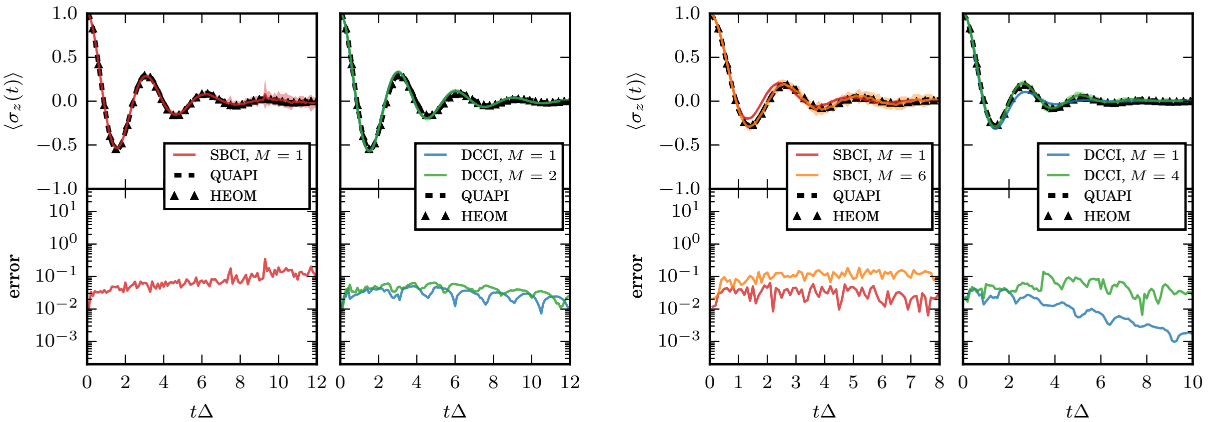
### 4.7.3 Low Temperature Regime

We now turn the attention to the phase diagram in the low temperature regime, specifically $\beta\Delta = 5$ $(k_BT/\Delta = 0.2)$, and concentrate on vertical cuts at intermediate coupling $\lambda/\Delta = 1$ and strong coupling $\lambda/\Delta = 10$, using the more suitable of the two methods in each case. These parameters correspond to Fig. 4.10(b).

#### Intermediate coupling

For intermediate coupling strength $(\lambda/\Delta = 1)$, the SBCI expansion is expected to converge at low temperatures more easily than in the high temperature regime. In particular, Fig. 4.10 shows the region of rapid convergence for the SBCI expansion becomes larger at low temperatures (b) than high temperatures (a). In Fig. 4.14, we find that the SBCI expansion can provide accurate results even at $M = 1$ for the parameter sets $(\lambda/\Delta, \omega_c/\Delta) = (1, 1)$ and $(1, 0.25)$, which would be more difficult to converge in the high temperature regime discussed in Sec. 4.7.2.

(a) $\lambda/\Delta = 1$, $\omega_c/\Delta = 5$ (b) $\lambda/\Delta = 1$, $\omega_c/\Delta = 1$



(c) $\lambda/\Delta = 1$, $\omega_c/\Delta = 0.25$

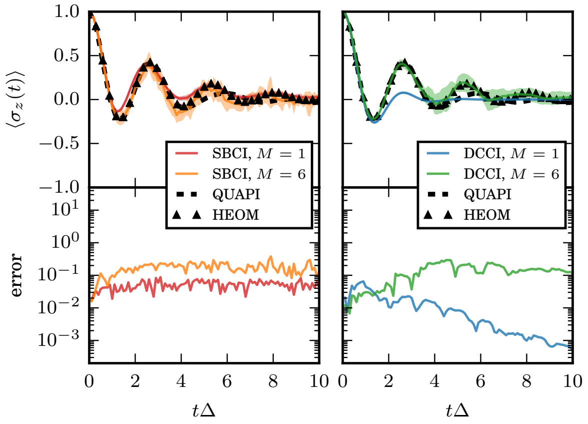


Figure 4.13: Nonequilibrium Population difference $\langle\sigma_z(t)\rangle$ (top subplots) and corresponding error estimates (bottom subplots) as a function of time in the intermediate coupling ($\lambda/\Delta = 1$) and high temperature ($\beta\Delta = 0.5$) regime. The bias energy is $\epsilon = 0$. The results calculated by the SBCI (left panels, red and orange) and DCCI (right panels, blue and green) expansions are plotted for (a) a non-adiabatic (fast) bath with $\omega_c/\Delta = 5$, (b) an intermediate bath with $\omega_c/\Delta = 1$, and (c) an adiabatic (slow) bath with $\omega_c/\Delta = 0.25$. Maximum order for a inchworm step is indicated by $M$. The thickness of the Monte Carlo results results from our error estimates. The dashed line are the QUAPI results with (a) $\Delta t = 0.1$, $k_{\max} = 6$, (b) $\Delta t = 0.2$, $k_{\max} = 10$, and (c) $\Delta t = 0.3$, $k_{\max} = 11$. The triangles indicate the HEOM result with $K = 2$ and $L = 20$.

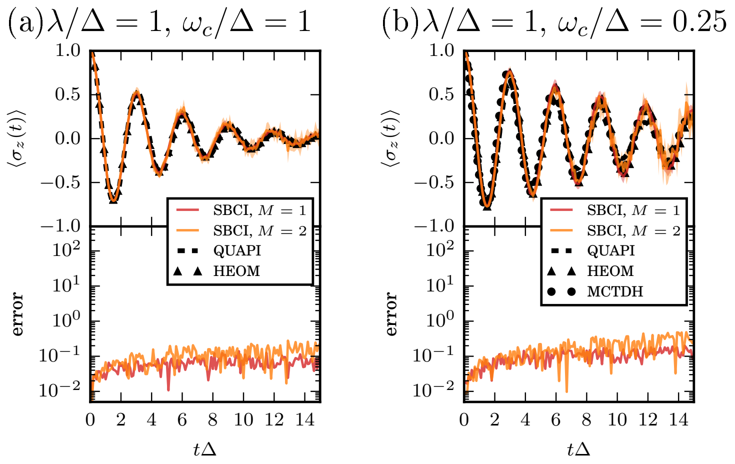


Figure 4.14: Nonequilibrium Population difference $\langle \sigma_z(t) \rangle$ (top subplots) and corresponding error estimates (bottom subplots) as a function of time in the intermediate coupling ($\lambda/\Delta = 1$) and low temperature ($\beta\Delta = 5$) regime. The bias energy is $\epsilon = 0$. The results calculated by the SBCI (red lines) expansions are plotted for (a) an intermediate bath with $\omega_c/\Delta = 1$ and (b) an adiabatic bath with $\omega_c/\Delta = 0.25$. The maximum order for the inchworm step shown is $M = 1$. The thickness of the Monte Carlo results results from our error estimates. The dashed lines are the QUAPI results with (a) $\Delta t = 0.1$, $k_{\max} = 6$ and (b) $\Delta t = 0.1$, $k_{\max} = 10$. The triangles indicate the HEOM result with $K = 2$ and $L = 20$. The MCTDH data is reported in Ref. 144.

### Strong coupling

In the strong coupling regime ($\lambda/\Delta = 10$), the DCCI approach is more rapidly convergent and efficient than the SBCI expansion (see Fig. 4.15). In particular, we show the DCCI results for parameter sets in the the adiabatic and intermediate regime, namely $(\lambda/\Delta, \omega_c/\Delta) = (10, 1)$ and $(10, 0.25)$. In these regimes, the lowest order DCCI results tend to over-estimate the incoherent decay of the population. Including higher order contributions within each inchworm step is necessary, as it provides significant corrections to population dynamics leading to agreement with the HEOM and MCTDH results. At small $\omega_c/\Delta$, one needs to go as far as $M = 8$, which is too difficult to fully converge with our current prototype code without spending a great deal of computer time. In the adiabatic regime (small $\omega_c$), QUAPI also tends to overestimate the decay for the long time behavior. To obtain correct long-time dynamics, one would need to increase the truncation of the memory length $k_{\max}$, which greatly increases the need for memory and makes QUAPI difficult to converge.

### 4.7.4 Very Low Temperature Limit

Finally, we explore the very low temperature limit $\beta\Delta = 50$ ($k_BT/\Delta = 0.02$) corresponding to the phase diagram Fig. 4.10(c). In this limit, the standard HEOM implementation[94] can be computationally expensive to converge. Indeed, the lower the bath temperature, the more Matsubara terms that are needed to capture the bath density matrix and the more hierarchical levels are required to converge the long-time dynamics. We find that the HEOM implementation available to us becomes unfeasible for very low temperatures, though we not that recent advances may ameliorate this problem in at least some instances[208–212]. Fig. 4.10 suggests that the SBCI and DCCI expansions hold an advantage over HEOM (though not MCTDH) at low temperatures, in that the computational cost does not increase with decreasing temperature. However, since at low enough temperatures strong correlation effects may alter the picture, it is not trivial that the simple analysis used to generate this figure should hold.

In Fig. 4.10(c), the combined area of strong convergence for the SBCI and DCCI expansions covers almost the entire parameter space in the very low temperature case. For the

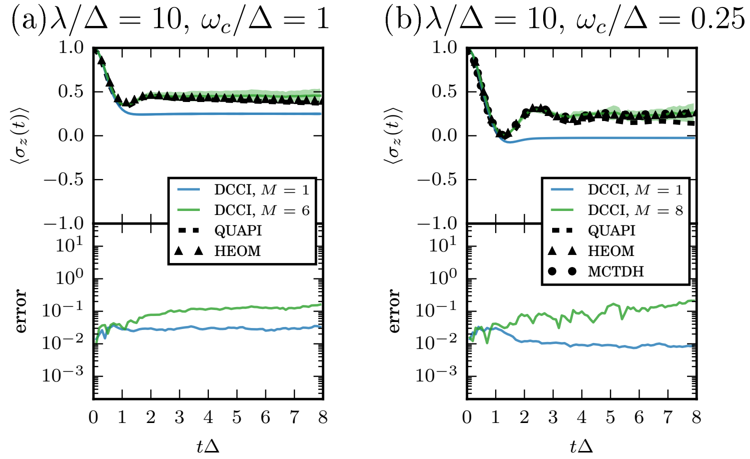


Figure 4.15: Nonequilibrium Population difference $\langle \sigma_z(t) \rangle$ (top subplots) and corresponding error estimates (bottom subplots) as a function of time in the strong coupling ($\lambda/\Delta = 10$) and low temperature ($\beta\Delta = 5$) regime. The bias energy is $\epsilon = 0$. The results calculated by the DCCI (blue and green lines) expansions are plotted for (a) an intermediate bath with $\omega_c/\Delta = 1$ and (b) an adiabatic bath with $\omega_c/\Delta = 0.25$. Maximum order for a inchworm step is indicated by $M$. The thickness of the Monte Carlo results results from our error estimates. The dashed line are the QUAPI results with (a) $\Delta t = 0.2$, $k_{\max} = 11$ and (b) $\Delta t = 0.4$, $k_{\max} = 10$. The triangles indicate the HEOM result with $K = 3$ and $L = 20$. The MCTDH data is reported in Ref. 144.

fast bath case ($\omega_c/\Delta = 5$), the parameter set falls out of the region of facile convergence for the DCCI approach, however we find that the DCCI expansion can still provide accurate population dynamics, and is in fact more efficient than the SBCI expansion. On the other hand, for the intermediate cut-off frequency case ($\omega_c/\Delta = 1$), the SBCI expansion results in the population dynamics that agree perfectly with the MCTDH result, while we note that the DCCI expansion is difficult to converge with respect to the maximum order $M$. This

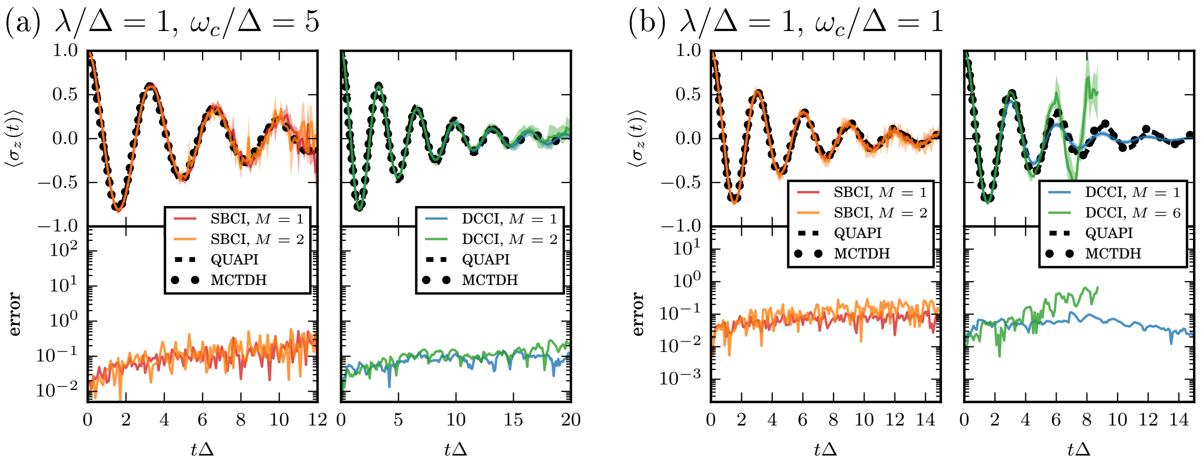


Figure 4.16: Nonequilibrium population difference $\langle\sigma_z(t)\rangle$ (top subplots) and corresponding error estimates (bottom subplots) as a function of time in the intermediate coupling ($\lambda/\Delta = 1$) and very low temperature ($\beta\Delta = 50$) regime. The bias energy is $\epsilon = 0$. The results calculated by the SBCI (left panels) and the DCCI (right panels) expansions are plotted for (a) a non-adiabatic (fast) bath with $\omega_c/\Delta = 5$ and (b) an intermediate bath with $\omega_c/\Delta = 1$. Maximum order for each inchworm step is indicated by $M$. The thickness of the Monte Carlo results results from our error estimates. The dashed line are the QUAPI results with $\Delta t = 0.1$ and $k_{\max} = 10$. The MCTDH data is reported in Ref. 144.

clearly does not agree with our naive estimates for convergence of the DCCI expansion.

### 4.7.5 Biased Systems

We now turn to a discussion of the last dimension of the parameter space of the spin–boson model, namely the bias energy of the spin subsystem. We expect the SBCI and DCCI approaches have similar behavior with respect to convergence within parameter space for non-zero bias energy. However, non-zero bias energy may introduce an additional phase in

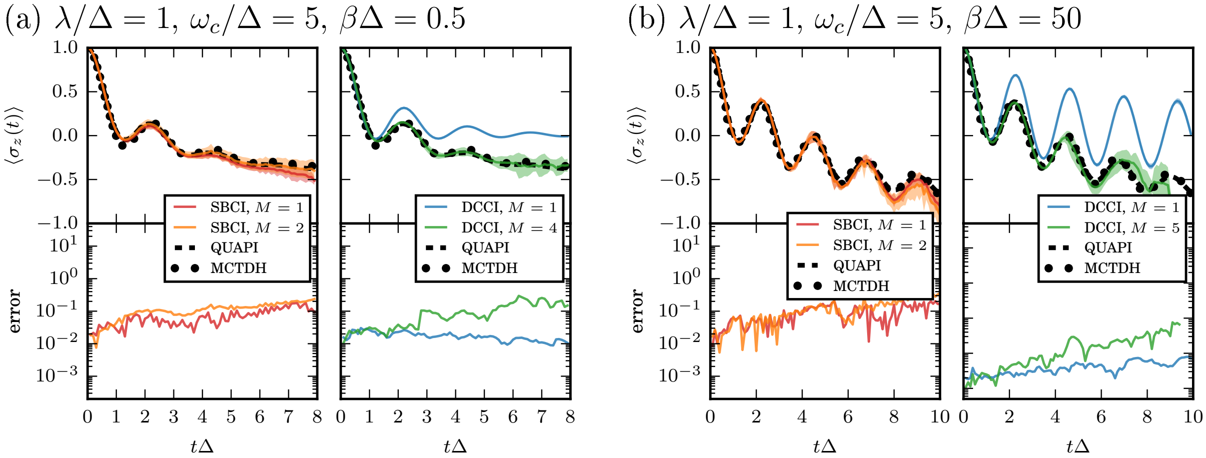


Figure 4.17: Nonequilibrium population difference $\langle\sigma_z(t)\rangle$ (top subplots) and corresponding error estimates (bottom subplots) as a function of time in the intermediate coupling ($\lambda/\Delta = 1$) and non-adiabatic ($\omega_c/\Delta = 5$) regime. The bias energy is $\epsilon = \Delta$. The results calculated by the SBCI (left panels) and the DCCI (right panels) expansions are plotted for (a) high temperature with $\beta\Delta = 0.5$ and (b) very low temperature with $\beta\Delta = 50$. Maximum order for each inchworm step is indicated by $M$. The thickness of the Monte Carlo results results from our error estimates. The dashed line are the QUAPI results with $\Delta t = 0.1$ and $k_{\max} = 10$. The MCTDH data is reported in Ref. 144.

the reduced propagators and cause a more difficult dynamical sign problem.

For the SBCI expansion, the $\epsilon$-dependence is only found within the system influence functional, explicitly in the operators in the interaction picture, $\widetilde{\sigma}_z(t) = e^{iH_s t}\hat{\sigma}_z e^{-iH_s t}$, where $H_s = \epsilon\hat{\sigma}_z + \Delta\hat{\sigma}_x$. The bath influence functional does not depend on the bias energy, so that the inchworm propriety of an individual diagram remains unchanged. Therefore, it is straightforward to account for the bias energy within the SBCI algorithm. On the other hand, the DCCI algorithm only contains $\epsilon$-dependence in the phase influence functional

$\Phi(\boldsymbol{s})$. Note that the phase functional is a real number only if $\epsilon = 0$, while $\epsilon \neq 0$ renders $\Phi(s)$ complex and thus potentially increases the dynamical sign problem of the dQMC method, making the DCCI algorithm somewhat more difficult to converge.

In Fig. 4.17, we show the SBCI and DCCI results for non-zero bias energy $\epsilon = \Delta$ cases at high temperature $\beta\Delta = 0.5$ and the very low temperature case $\beta\Delta = 50$. The system–bath coupling is taken to be $\lambda/\Delta = 1$ and a cut-off frequency of $\omega_c/\Delta = 5$ is used. In general, the error estimate of the SBCI expansion grows more rapidly with time, so that more computational effort to control the error propagation is needed. The DCCI expansion shows a clear convergence with respect to the maximum order $M$. Compared to the same parameter set $\lambda/\Delta = 1$, $\omega_c/\Delta = 5$, and $\beta\Delta = 50$ for zero bias energy, we note that the non-zero bias energy does increase the computational effort, especially for the DCCI approach.

## 4.8 Conclusions

In this chapter we develop two complementary dQMC inchworm approaches for the simulation of exact real-time non-adiabatic dynamics. These approaches are based on generic expansions in either the system–bath coupling or the diabatic coupling, respectively, and thus should be of general utility. For concreteness, as well as to permit benchmarking of the approach, we specialize to the case of the spin–boson model.

Our first approach is based on a system–bath coupling expansion, analogous to the hybridization expansion in the Anderson model. Indeed the scheme is nearly identical to that employed in original inchworm algorithm formulated for the Anderson impurity model.[106] We formally show that proper inchworm diagrams account for any diagram in the bare

Monte Carlo expansion once and only once. The major advantages of the SBCI approach are twofold: there are far fewer proper inchworm diagrams than bare diagrams and an infinite number of bare diagrams are resummed in the inchworm expansion. However, this advantage comes at a cost, namely one has to calculate two-time restricted propagators and perform a more involved error analysis during inchworm propagation.

The second inchworm approach is based on the diabatic coupling expansion and its cumulant form. Due to the fact that diagrams within the diabatic coupling expansion include an interaction line between every two vertices, the main advantages of the inchworm algorithm are lost if one follow the previous scheme. To circumvent this problem, we introduce a cumulant form of the expansion and propose an alternative inchworm approach, the diabatic coupling cumulant inchworm (DCCI) expansion. The DCCI expansion has the notable advantage that only single-time properties are needed, and the simulation scales linearly in time. Since cumulant forms can also be used in other inchworm expansions (namely the SBCI approach), this property should be of general utility. We also note that since the DCCI and SBCI expansions converge differently in distinct parameter regimes, we expect their combined use to cover much, if not all, of the relevant parameter space.

In Sec. 4.7 we have presented benchmark calculations of the inchworm Monte Carlo approach for the real-time nonequilibrium dynamics in the spin–boson model. A rather extensive swath of the full parameter space has been explore and a detailed discussion of the convergence properties of both the SBCI and DCCI has been made. We have compared these inchworm expansions to several prominent, numerically exact schemes such as QUAPI, HEOM, and MCTDH.

In general, we find that at least one of the inchworm expansions appears to converge to the exact result in essentially all tested regions of parameter space. This appears to include regions of parameter space that are difficult for the QUAPI and HEOM methods. On the other hand, at this stage the QUAPI and HEOM algorithms are simpler to employ. In particular, more work needs to be done to fully understand the factors that govern error growth and convergence of the various inchworm approaches so that a general "black-box" implementation may be developed which would render inchworm Monte Carlo as user-friendly as these approaches.

In our view, the MCTDH approach is the most reliable and stable approach for the description of dynamics in the standard spin–boson problem. The inchworm approaches presented here provide results that appear compatible, but not quite as robust, as those produced by MCTDH. Inchworm Monte Carlo is essentially an efficient means to stochastically sample an exact perturbation expansion. This gives hope that the approach may provide a general utility beyond the simplest incarnation of the spin–boson model, in cases where other methods may not be viable. Indeed, inchworm works very well for the Anderson impurity model, where QUAPI appears to suffer memory length issues[105] and MCTDH appears to have trouble in strongly correlated regimes[213].

The biggest potential niche for the suite of inchworm Monte Carlo approaches outlined here appears to be in nonequilibrium setting where transport occurs between two or more reservoirs. In such situations, MCTDH is far more expensive, while diagrammatic Monte Carlo actually becomes easier to converge[124, 129, 132]. A particularly interesting case is nonequilibrium heat transport in the multi-bath spin–boson problem[90, 91, 214–216]. Here, as far as we know, only one exact calculation has been performed[90, 91], but owing to the numerical difficulty of the problem, a systematic study could not be performed. This is just one example of a class of physically important problems that may be probed in far greater detail by the inchworm Monte Carlo methods of this work.

# Appendix

## 4.A Wick's theorem in the diabatic coupling expansion

The multi-time correlation function of polaron shift operators given in Eq. (4.52) can be written as

$$\mathcal{J}\left(\boldsymbol{s}\right)=\left\langle \widetilde{\mathcal{B}}_{-}\left(s_{m+1}\right)\widetilde{\mathcal{B}}_{+}^{2}\left(s_{m}\right)\cdots\widetilde{\mathcal{B}}_{-}^{2}\left(s_{1}\right)\widetilde{\mathcal{B}}_{+}\left(s_{0}\right)\right\rangle _{b}. \tag{4.A.1}$$

The explicit form of the polaron shift operator in the interaction picture is given by

$$\begin{aligned}\widetilde{\mathcal{B}}_{\sigma}\left(s\right)= & e^{\widetilde{\theta}_{\sigma}\left(s\right)}, & (4.A.2)\\
\widetilde{\theta}_{\sigma}\left(s\right)= & \sigma\sum_{\ell}\frac{c_{\ell}}{\omega_{\ell}^{3/2}}\left(e^{i\omega_{\ell}s}b_{\ell}^{\dagger}-e^{-i\omega_{\ell}s}b_{\ell}\right). & (4.A.3)\end{aligned}$$

To simplify the notation, we drop the $\ell$ index for the time being and define $\xi\left(s\right)=\frac{c}{\omega^{3/2}}e^{i\omega s}$, such that

$$\widetilde{\theta}_{\sigma}\left(s\right)=\sigma\left(\xi\left(s\right)b^{\dagger}-\xi\left(s\right)^{*}b\right). \tag{4.A.4}$$

The arguments of the polaron shift operators product can be combined using the identity

$$e^{vb^{\dagger}-v^{*}b}e^{ub^{\dagger}-u^{*}b}=e^{(v+u)b^{\dagger}-(v^{*}+u^{*})b}\times e^{(vu^{*}-v^{*}u)/2} \tag{4.A.5}$$

for boson operators $b$ and $b^{\dagger}$ (as can easily be derived using the Baker–Campbell–Hausdorff formula).

Next, the two-time correlator of polaron shift operators is

$$\begin{aligned}\mathcal{B}_{\sigma'}^{r'}\left(s'\right)\mathcal{B}_{\sigma}^{r}\left(s\right) = &\exp\left\{\left[\sigma' r'\xi\left(s'\right)+\sigma r\xi\left(s\right)\right]b^{\dagger}-\text{c.c.}\right\}\times\\ &\exp\left\{i\sigma'\sigma r' r\,\text{Im}\left[\xi\left(s'\right)\xi\left(s\right)^{*}\right]\right\}.\end{aligned}\tag{4.A.6}$$

We note that the boson operator part of the correlator takes the same form as the polaron shift operator and an additional scalar factor (not a boson operator) emerges. Therefore, we can recursively combine the argument using the above identity and find a general expression for the multi-time correlator

$$\begin{aligned}\prod_{j}\mathcal{B}_{\sigma_j}^{r_j}\left(s_j\right) = &\exp\left\{\sum_{j}\sigma_j r_j\xi\left(s_j\right)b^{\dagger}-\text{c.c.}\right\}\times\\ &\exp\left\{i\sum_{j}\sum_{k<j}\sigma_j\sigma_k r_j r_k\text{Im}\left[\xi\left(s_j\right)\xi\left(s_k\right)^{*}\right]\right\}.\end{aligned}\tag{4.A.7}$$

The scalar factor part can be rewritten in the form $\text{Im}\left[\xi\left(s'\right)\xi\left(s\right)^{*}\right]=\frac{c^2}{\omega^3}\sin\omega\left(s'-s\right)$.

We now focus on the thermal average of the boson operator, $\exp\left\{\sum_j\sigma_j r_j\xi\left(s_j\right)b^{\dagger}-\text{c.c.}\right\}$. The thermal average of free boson operator of the above form can be obtained as

$$\text{Tr}_b\left\{\rho_b e^{\kappa b^{\dagger}-\kappa^{*}b}\right\}=\exp\left\{-\frac{1}{2}\kappa\kappa^{*}\coth\left(\frac{\beta\omega}{2}\right)\right\},\tag{4.A.8}$$

where $\rho_b=e^{-\beta H_b}$. We can take the thermal average of the two-time correlator to obtain

$$\begin{aligned}&\text{Tr}_b\left\{\rho_b\mathcal{B}_{\sigma'}^{r'}\left(s'\right)\mathcal{B}_{\sigma}^{r}\left(s\right)\right\}=\\ &\exp\left\{\sigma'\sigma r' r\frac{c^2}{\omega^3}\left[1-\cos\omega\left(s'-s\right)\right]\coth\left(\frac{\beta\omega}{2}\right)\right\}\times\\ &\exp\left\{i\sigma'\sigma r' r\frac{c^2}{\omega^3}\sin\omega\left(s'-s\right)\right\},\end{aligned}\tag{4.A.9}$$

where a time-independent phase is dropped, since it cancels out when a configuration on the Keldysh contour is considered. By choosing $\sigma'=-1$ and $\sigma=1$ and putting the $\ell$ index back, we can carry out $\sum_\ell$ in terms of spectral density $J\left(\omega\right)$,

$$\begin{aligned}\mathrm{Tr}_b\left\{\rho_b\mathcal{B}_-^{r'}\left(s'\right)\mathcal{B}_+^{r}\left(s\right)\right\} = \\ \exp\left\{r'r\left[-\mathcal{Q}_2\left(s'-s\right)-i\mathcal{Q}_1\left(s'-s\right)\right]\right\},\end{aligned} \tag{4.A.10}$$

where $\mathcal{Q}_1$and $\mathcal{Q}_2$ are defined by Eqs. (4.58) and (4.59). The two-time correlation function, Eq. (4.57), is then given by

$$C\left(s',s\right)^{r'r} \equiv \mathrm{Tr}_b\left\{\rho_b\mathcal{B}_-^{r'}\left(s'\right)\mathcal{B}_+^{r}\left(s\right)\right\} \tag{4.A.11}$$

Finally, we take the thermal average of the multi-time correlator

$$\begin{aligned}\mathrm{Tr}_b\left\{\rho_b\prod_j\mathcal{B}_{\sigma_j}^{r_j}\left(s_j\right)\right\} = \exp\left\{\sum_j\sum_{k<j}\sigma_j\sigma_k r_j r_k\frac{c^2}{\omega^3}\left[1-\cos\omega\left(s_j-s_k\right)\right]\coth\left(\frac{\beta\omega}{2}\right)\right\}\times \\ \exp\left\{i\sum_j\sum_{k<j}\sigma_j\sigma_k r_j r_k\frac{c^2}{\omega^3}\sin\omega\left(s_j-s_k\right)\right\}.\end{aligned} \tag{4.A.12}$$

We can finally carry out $\sum_\ell$ in terms of spectral density $J\left(\omega\right)$, concluding that

$$\mathrm{Tr}_b\left\{\rho_b\prod_j\mathcal{B}_{\sigma_j}^{r_j}\left(s_j\right)\right\} = \prod_j\prod_{k<j}\left(C\left(s_j,s_k\right)^{r_jr_k}\right)^{-\sigma_j\sigma_k}. \tag{4.A.13}$$

Since we have $\sigma_j=1$ for $j$ even and $\sigma_j=-1$ for $j$ odd, the powers are

$$-\sigma_j\sigma_k = \begin{cases}1 & \left|j-k\right| \text{ odd} \\ -1 & \left|j-k\right| \text{ even}\end{cases}. \tag{4.A.14}$$

# Chapter 5

# Two Flavors of the Non–Crossing Approximation for the Anderson–Holstein Model[*]

## 5.1 Introduction

The interaction between electrons and phonons plays an essential role in condensed matter physics: it is for example the fundamental factor responsible for the resistivity of conduction electrons in crystals at relatively high temperatures and the onset of superconductivity at low temperatures[23]. In non-equilibrium molecular electronics experiments[33–35], electron–phonon interactions are ever present and have major implications[38, 39] which can

---

[*] Based on work published in Phys. Rev. B **93**, 174309 (2016).

be exploited in the design of phononic devices[36, 37]. In addition, the interplay between electron–electron interactions (responsible for Coulomb blockade and the Kondo effect) and electron–phonon scattering leads to novel and subtle behaviors[40, 217]. For example, conductance side peaks replicating the Kondo resonance[40, 218, 219] and negative differential resistance at voltages corresponding to the vibrational energy of the molecule[112] have been observed. In a broader sense, explicating the role played by electron–phonon interactions in strongly correlated materials remains a fertile area of research, where recent interest has focused, for example, on the role played by phonons in fulleride[24], cuprate[26] and pnictide superconductors[27, 28] and the control of superconductivity and metal–insulator transitions in correlated materials via strong laser fields[25, 29–32].

A standard model that simultaneously describes both electronic interactions and electron–phonon coupling in nanoscale devices is the Anderson–Holstein model[20, 21, 114]. This model consists of a single interacting site (sometimes called the dot or impurity) coupled to a non-interacting electron reservoir (or reservoirs) and to a set of localized phonon modes. The Anderson–Holstein model can be considered a minimal description of the essential aspects of a correlated electron system interacting with phonon excitations, and has been used to describe vibrational effects in molecular electronics[113–115, 220]. Furthermore, within the framework of dynamical mean-field theory (DMFT) [41], the characterization of a strongly correlated material with active phonon degrees of freedom may be effectively reduced to the Anderson–Holstein model and its variants[126, 220, 221].

Despite the importance of the Anderson–Holstein model, there is surprisingly little known about its real–time dynamical properties outside of simple limits where perturbation arguments can be made. The case of zero on-site electron–electron interactions can describe some phenomena associated with the electron–phonon interaction,including non–equilibrium transient dynamics, inelastic transport, and phonon-induced side peaks[107–111]. This limit has been widely considered in the literature; despite its simplicity, it is non-trivial to solve, especially out of equilibrium. A variety of techniques have been used to analyze this model, including perturbation theory in the electron–phonon coupling[222], a semi-classical treatment[223], and master–equation approaches[222, 224–228]. Semi-analytical approximations within nonequilibrium Keldysh Green's functions (NEGF) [229–235], the equation-of-motion (EOM) approach[236–239], an interpolative ansatz[240], and a recent dressed tunneling approximation[109] have been applied to the model in various limits. Numerically exact methods have also been applied, including real-time Quantum Monte Carlo (QMC) [110, 111, 124, 241], iterative path integral schemes[242–244] and the multi-layer multi-configuration time-dependent Hartree (ML-MCTDH) method[88, 245].

Treatment of the combined effect of electron–electron and electron–phonon interactions is simplest when the on-site Coulomb repulsion is effectively infinite ($U \to \infty$). In this limit, some methods used to treat the non-interacting case can be adopted and generalized, including certain Monte Carlo approaches[113, 246], the equation-of-motion technique[247–250], a decoupling scheme for NEGF[251], and the slave-boson technique[252, 253]. Studies of the infinite–$U$ Anderson–Holstein model predict non-trivial effects, such as the appearance of Kondo replicas above and below the chemical potential and negative differential resistance associated with the destruction of the Kondo resonance[112, 113]. However it remains unclear if these predictions are valid outside of linear response from equilibrium, and in general neither the $U = 0$ nor $U \to \infty$ limits describe the bulk of interesting cases of experimental relevance.

Only a handful of approaches are capable of calculating properties of a generic Anderson–Holstein model outside of the idealized limits discussed above. Approximate methods, such as the master equation approach, can accurately describe transport phenomena at high temperatures and large voltages[254]. The ML-MCTDH method is numerically exact, but has difficulty converging for strong electron–phonon coupling or far from equilibrium[255, 256]. The numerical renormalization group (NRG) can also be extended to include electron–phonon interactions, but remains difficult to apply out of equilibrium and is generally reliable only for the low energy properties of the system[107, 114–119]. The auxiliary-field QMC method has been used to calculate the density of states under the influence of the phonons in imaginary time[120], but application to dynamics involves an uncontrolled analytical continuation which is problematic at certain parameters[121, 122], and the Matsubara formulation is only valid for equilibrium and linear response properties. Real time QMC provides an alternative numerically exact approach which has the ability to describe transient dynamics and non-equilibrium transport properties over a wide range of parameters[123–128]. In combination with reduced dynamics techniques[130–132] it can sometimes be used to obtain results over very long timescales[132]. However, real time QMC is generically plagued by a dynamical sign problem which limits the accessible timescales. Although not the direct focus of this manuscript, we note that the approaches described here can provide a foundation to allow for an amelioration of the sign problem in QMC simulations[121, 122, 129].

The self-consistent resummation of particular classes of interaction terms may allow for an extension of the domain of validity provided by bare perturbation theory. A prominent example is provided by the non-crossing approximation (NCA) [257, 258]. The NCA is a semi-analytical method based on the resummation to all orders of a specific subset of diagrams (those that do not cross temporally on the Keldysh contour) associated with the hybridization between the impurity and the non-interacting leads. It provides a computationally inexpensive approach for solving generic impurity models out of equilibrium[259]. NCA is exact in the atomic limit, and works best in the limit of infinite $U$ and finite $\epsilon$. The approximation does not fully capture low energy properties and does not correctly reproduce the noninteracting limit. But despite the quantitative inaccuracies, the NCA qualitatively predicts the emergence and some properties of the Kondo resonance, and is generally accurate for high-energy features. While the NCA as a "stand alone" approximation may quantitatively fail, higher order approximations (e.g. one–crossing approximation) based on the same principles have been used[260, 261], and recent numerically exact QMC approaches have been formulated that sample corrections to the NCA in a numerically exact way[106, 121, 122, 127, 129].

The NCA has been extended to include the electron–phonon coupling, via the slave-boson technique[252, 253], in nonequilibrium DMFT studies[126, 221], and within a pseudoparticle picture[262]. A first goal of our work is to clearly formulate two complementary NCA-like approximations in the full many-body basis of the impurity, in a form suitable for studying the non-equilibrium behavior of the Anderson–Holstein model, and to compare and contrast the predictions of these distinct self-consistent procedures. A second goal is to clearly delineate the diagrammatic rules associated with each self-consistent resummation on the Keldysh contour so that future exact QMC schemes which sample remaining diagrams may be explicitly formulated. The outline of this chapter is as follows. In Sec. 5.2 we introduce the Anderson–Holstein model and provide the needed formalism. In Sec. 5.3, two distinct types of NCA-like approximation are described. In Sec. 5.4, we present and compare results for transient dynamics, steady state spectral function and differential conductance for a generic Anderson–Holstein model in the Kondo regime. A summary and conclusion are presented in Sec. 5.5.

## 5.2 Coupling Expansion for Anderson–Holstein Model

### 5.2.1 Model and definitions

We consider a single spin-degenerate impurity or quantum dot level with a linear coupling to a phonon bath and to a pair of metallic leads which will be referred to as "left" ($L$) and "right" ($R$). This model is described by the nonequilibrium Anderson–Holstein Hamiltonian[114, 115, 220]

$$H = H_d + H_b + V_b + \sum_{\ell \in L,R} \left( H_\ell + V_\ell \right). \tag{5.1}$$

The electronic part of the dot Hamiltonian, $H_d$, is

$$H_d = \sum_{\sigma=\uparrow,\downarrow} \epsilon_\sigma n_\sigma + U n_\uparrow n_\downarrow, \tag{5.2}$$

where $\epsilon_\sigma$ denotes the energy of singly-occupied states and $U$ is the Coulomb interaction. The operators $d_\sigma^\dagger$ creates an electron of spin $\sigma$ on the dot and the occupation $n_\sigma = d_\sigma^\dagger d_\sigma$.

The local phonon bath Hamiltonian is

$$H_b = \sum_q \omega_q b_q^\dagger b_q. \tag{5.3}$$

Here the $b_q^\dagger$ are phonon creation operators, and $\omega_q$ is the frequency associated with a phonon mode $q$. We will typically assume that the phonons are initially in equilibrium, such that the occupation of the phonon modes is given by the Bose–Einstein distribution $\langle b_q^\dagger b_q \rangle = \frac{1}{e^{\beta_d \omega_q}-1}$, $\beta_d$ being the inverse temperature of the phonon bath. The electron–phonon coupling Hamiltonian $V_b$ is

$$V_b = \sum_q \lambda_q (b_q^\dagger + b_q)\left(n_d - \delta\right), \tag{5.4}$$

where $n_d = \sum_\sigma n_\sigma$ is the total electronic occupation of the dot and $\lambda_q$ the coupling strength between the dot and phonon mode $q$. The parameter $\delta$ is of no physical significance, in the sense that it may be absorbed into a redefinition of the zero point of the oscillator coordinate. However, it is convenient to set $\delta = 1$, so that $\epsilon = 0$ describes the particle–hole symmetric dot, and we will primarily consider this case. We will also investigate the case $\delta = 0$, which provide a more convenient description of a molecular junction in which polaron formation is linked to the presence of extra electrons on the dot. In either case, the electron–phonon coupling is characterized by a spectral density

$$J(\omega) \equiv \frac{\pi}{2} \sum_q \frac{\lambda_q^2}{\omega_q} \delta\left(\omega - \omega_q\right). \tag{5.5}$$

The left and right lead Hamiltonians are

$$H_\ell = \sum_{k\in\ell} \sum_\sigma \epsilon_k c_{k\sigma}^\dagger c_{k\sigma}, \tag{5.6}$$

with $\ell \in \{L, R\}$ and the index $k$ denoting a level within a lead. We assume the leads to be non-interacting, such that they are fully described by the dispersion relation $\epsilon_k$ and the creation operators $c_{k\sigma}^\dagger$. The leads are taken to each be initially isolated and at an equilibrium state with density matrix $\rho_\ell$, and their thermodynamic properties characterized by an inverse temperature $\beta_\ell$ and a chemical potential $\mu_\ell$. The initial density of states is then described by a Fermi–Dirac distribution, $\langle c^\dagger_{k\sigma} c_{k\sigma}\rangle = f_\ell(\epsilon_k) = \frac{1}{e^{\beta_\ell(\epsilon_k-\mu_\ell)}+1}$.The hybridization $V_\ell$ between the dot and lead electrons is described by the dot-lead coupling Hamiltonian

$$V_\ell = \sum_{k\in\ell}\sum_{\sigma}\left[t_k d_\sigma c^\dagger_{k\sigma} + t^*_k d^\dagger_\sigma c_{k\sigma}\right], \tag{5.7}$$

where $t_k$ enumerates the coupling strength between the dot and level $k$ of lead $\ell$. We define a coupling density

$$\Gamma_\ell(\omega) = 2\pi \sum_{k\in\ell} |t_k|^2 \delta\left(\omega - \epsilon_k\right), \tag{5.8}$$

which fully characterizes the $t_k$ within this model.

In steady state the dynamical response of a system is characterized by its spectral function

$$A(\omega) = \frac{i}{2\pi}\mathrm{Tr}\left\{G^r(\omega) - G^a(\omega)\right\}, \tag{5.9}$$

which may be considered a probe of the density of electron and hole excitations as a function of energy. To calculate the spectral function at frequency $\omega'$, we use the auxiliary current method[121, 122] by appending two auxiliary leads to the model, $H \to H + H_A + V_A$, where

$$H_A = \sum_{k\in A}\epsilon_k a^\dagger_k a_k, \tag{5.10}$$

and

$$V_A = \sum_{k\in A}\sum_{\sigma}\left[t_k d_\sigma a^\dagger_k + t^*_k d^\dagger_\sigma a_k\right]. \tag{5.11}$$

These auxiliary leads are coupled to the dot at the single frequency $\omega'$ with a spectral density $\Gamma^{\omega'}_A(\omega) = \eta\delta(\omega - \omega')$. One lead is kept fully occupied, such that $f_{A1}(\omega) = 1$; the other lead is kept empty, such that $f_{A0}(\omega) = 0$. We can calculate the *auxiliary* spectral function $A\left(\omega; t\right)$

at any finite time by the following relation:

$$A\left(\omega;t\right)=\lim_{\eta\to0}-\frac{2h}{e\pi\eta}\left[I_{A1}^{\omega}\left(t\right)-I_{A0}^{\omega}\left(t\right)\right]. \tag{5.12}$$

Here, $I_{A0}^{\omega}\left(t\right)$ and $I_{A1}^{\omega}\left(t\right)$ are the currents flowing out of lead $A0$ and $A1$, respectively, at time $t$. At long times, the auxiliary spectral function approaches the steady state spectral function, Eq. 5.9. While at finite times the auxiliary spectral function does not conform to the standard definition of a spectral function in terms of a Fourier transform of a correlation function, it retains the appealing physical interpretation as a measure of the single-particle excitation density in energies, and could in principle be accessed experimentally by way of three-lead experiments[121, 122, 263, 264].

We shall also be interested in the differential conductance,

$$G(V)=\frac{\mathrm{d}}{\mathrm{d}V}(I_L-I_R). \tag{5.13}$$

which is directly accessible in transport experiments. Here, $V=\mu_L-\mu_R$ is the bias voltage between the two leads. The current $I_\ell(t)$ out of lead $\ell$ is given by $I_\ell(t)=\langle\mathcal{I}_\ell(t)\rangle$, where the current operator for a given lead,

$$\mathcal{I}_\ell=\dot{N}_\ell=i\sum_{k\in\ell}\left(t_kc_{k\sigma}^\dagger d_\sigma-t_k^*c_{k\sigma}d_\sigma^\dagger\right), \tag{5.14}$$

describes the rate at which carriers flow out of that lead. The differential conductance is often interpreted as an estimator for the equilibrium spectral function of the model. However, this interpretation is only valid if the spectral function is independent of the bias voltage. In practice, the two quantities may be qualitatively different[121].

### 5.2.2 Coupling expansion: general formalism

We now formulate a double expansion in the electron–phonon and dot–lead couplings. A brief review will be provided here for completeness; we refer readers interested in a more detailed technical outline of the formalism and algorithm elsewhere[122]. We begin by recasting the Hamiltonian as $H = H_0 + V$. $H_0$ describes the isolated dot and bath subsystems, while $V = V_b + \sum_\ell V_\ell$ describes the coupling Hamiltonian.

The expectation value of an operator $\mathcal{O}$ at time $t$ can be written in the form $\langle \mathcal{O}(t) \rangle = \langle e^{iHt} \mathcal{O} e^{-iHt} \rangle = \langle U^\dagger(t) \mathcal{O}_I(t) U(t) \rangle$, where $U(t) = e^{iH_0 t} e^{-iHt}$ and $\mathcal{O}_I(t) = e^{iH_0 t} \mathcal{O} e^{-iH_0 t}$. The subscript $I$ denotes an operator in the interaction picture. We also define thermal averaging by way of the notation $\langle O \rangle \equiv \mathrm{Tr}\left\{\rho O\right\}$, with the averaging performed with respect to the uncorrelated initial density matrix formed by the product of subsystem density matrices: $\rho = \rho_d \otimes \prod_\ell \rho_\ell \otimes \rho_b$. Thus the dynamics that appear in the following are not in equilibrium and illustrate the approach to equilibrium in the appropriate limits. Other than in some very special cases, a *finite* system coupled to an infinite thermal bath which is allowed to evolve in time is generally found to reproduce the steady state results at long times. Moreover, this is often the only rigorous way to construct the correct nonequilibrium steady state in open quantum systems. Initial correlations allow the system to be thermalized at time zero. Within DMFT[126, 187, 220, 221, 261], one deals with an *infinite* interacting system which is not coupled to a bath, and the role of the initial correlations therefore becomes more important. They are needed to model an initially thermalized system, which might be thought of as a system that had been weakly coupled to a bath and allowed to relax before the beginning of the calculation.

We now describe the details of a Dyson expansion for the reduced propagator on the

Keldysh contour. We can expand $U(t)$ in a Dyson series

$$U(t) = \sum_{n=0}^{\infty} (-i)^n \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n V_I(t_1) V_I(t_2) \cdots V_I(t_n), \tag{5.15}$$

such that the propagator can be expressed as $e^{-iHt} = e^{-iH_0 t} U(t)$. We adopt the many-body atomic states of the isolated dot, $\{|\alpha\rangle\} = \{|00\rangle \equiv |0\rangle, |\uparrow\rangle \equiv |1\rangle, |\downarrow\rangle \equiv |2\rangle, |\uparrow\downarrow\rangle \equiv |3\rangle\}$, as a basis, and define the reduced propagator matrix element $G_{\alpha\beta}(t) \equiv \left\langle \alpha \left| \mathrm{Tr}_B \left\{ \rho e^{-iHt} \right\} \right| \beta \right\rangle$. The trace is taken over the lead and phonon degrees of freedom: $\mathrm{Tr}_B \equiv \mathrm{Tr}_\ell \mathrm{Tr}_b$. The remaining quantity is reduced to the dimensionality of the (many-body) dot subspace. We also define the unperturbed reduced propagator $G^{(0)}_{\alpha\beta}(t) \equiv \left\langle \alpha \left| \mathrm{Tr}_B \left\{ \rho e^{-iH_0 t} \right\} \right| \beta \right\rangle$. $G^{(0)}_{\alpha\beta}$ is diagonal for the model treated here, and takes the form $G^{(0)}_{\alpha\beta}(t) = \Phi(t)\delta_{\alpha\beta} e^{-iE_\alpha t}$. The state energy $E_\alpha$ is evaluated from the isolated dot Hamiltonian. The factor $\Phi(t) = \mathrm{Tr}_B \left\{ \rho e^{-i(H_0 - H_d)t} \right\}$ is related to fluctuations in the noninteracting baths, and is independent of the dot state. It is exactly canceled when considering any quantity defined on the two branch Keldysh contour, and can therefore be safely ignored.

The full, or perturbed, reduced propagator $G_{\alpha\beta}(t)$ is also diagonal. Contributions to it from the coupling Hamiltonian are nonzero only when the creation and annihilation operators occur in pairs, such that only even orders must be accounted for:

$$G_{\alpha\alpha}(t) = G^{(0)}_{\alpha\alpha}(t) - \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \alpha | \mathrm{Tr}_B \left\{ \rho e^{-iH_0 t} V_I(t_1) V_I(t_2) \right\} |\alpha\rangle + \cdots . \tag{5.16}$$

This series can be represented as a summation of diagrams in which the coupling Hamiltonian appears an even number of times. An example diagram is shown Fig. 5.1: in (a), the representation of $G^{(0)}_{\alpha\alpha}$ (thin lines) and $G_{\alpha\alpha}$ (bold lines) in terms of pairs of solid and dashed lines is shown. In (b) a diagram is shown which contains Fermion hybridizations, denoted by wiggly lines which change the dot population, and phonon interactions, denoted by wavy lines with loops which do not change the population (and may appear only within certain dot states, as detailed below).

The reduced propagator satisfies a causal Dyson equation of the form

$$G_{\alpha\alpha}(t) = G^{(0)}_{\alpha\alpha}(t) + \int_0^t dt_1 \int_0^{t_1} dt_2 G^{(0)}_{\alpha\alpha}(t-t_1)\Sigma_{\alpha\alpha}(t_1,t_2)G_{\alpha\alpha}(t_2), \tag{5.17}$$

where all non-trivial aspects of the problem are contained in the (proper) self energy $\Sigma_{\alpha\alpha}(t_1,t_2)$. Solving the Dyson equation self-consistently is in itself an inexpensive computation if the self energy is known. Within the hybridization expansion for the phonon-free version of the model, the simplest approximation to the self energy includes only a single pair of coupling Hamiltonians:

$$\begin{aligned} \Sigma^{\mathrm{2BA}}_{\alpha\alpha}(t_1-t_2) &= -\langle\alpha|\mathrm{Tr}_b\left\{\rho V e^{-iH_0(t_1-t_2)}V\right\}|\alpha\rangle \\ &= \sum_\beta G^{(0)}_{\beta\beta}(t_1-t_2)\times\Delta^\beta_{\alpha\alpha}(t_1-t_2), \end{aligned} \tag{5.18}$$

where the hybridization function is defined as

$$\Delta^\beta_{\alpha\alpha}(t_1-t_2) \equiv -\langle\alpha|\mathrm{Tr}_b\left\{\rho V_I(t_1)|\beta\rangle\langle\beta|V_I(t_2)\right\}|\alpha\rangle. \tag{5.19}$$

This is known as the second-order Born approximation (2BA). The non-crossing approximation (NCA), also known as the self-consistent Born approximation (SCBA), takes the same form, but inserts the full propagator $G$ into the self energy:

$$\Sigma^{\mathrm{NCA}}_{\alpha\alpha}(t_1-t_2) = \sum_\beta G_{\beta\beta}(t_1-t_2)\times\Delta^\beta_{\alpha\alpha}(t_1-t_2). \tag{5.20}$$

With this self energy, we can obtain an approximate propagator containing an infinite, but partial, subset of the diagrams contributing to the reduced propagator, namely all diagrams in which hybridization lines do not cross each other. In the following section, two ways of generalizing this idea to the full Anderson–Holstein model will be described.

So far, in order to simplify the discussion, we have limited our attention to a reduced propagator living on a single branch of the Keldysh contour. To calculate a physical observable, we must consider a two-branch Keldysh contour with the observable operator $\mathcal{O}$ placed at the final time $t$, and take into account diagrams with lines crossing between the two branches. To this end, we define a vertex function of the observable $\mathcal{O}$, with the two time variables $t$ and $t'$ placed on opposite branches of the contour. With $t' \to t$, this object yields the physical expectation value of observable $\mathcal{O}(t)$. In particular, the current out of the lead $\ell$ can be obtained from $I_\ell(t) = \langle \mathcal{I}_\ell(t) \rangle$, where the current operator

$$\mathcal{I}_\ell = \dot{N}_\ell = i \sum_{k \in \ell} \left( t_k c^\dagger_{k\sigma} d_\sigma - t^*_k c_{k\sigma} d^\dagger_\sigma \right) \tag{5.21}$$

and the $c$ and $d$ operators are understood to be at the tip of the Keldysh contour.

Because $\mathcal{I}_\ell$ is composed of the same operators appearing in the dot–bath hybridization Hamiltonian, within the coupling expansion the current can be obtained by summing over diagrams which have a special hybridization line placing the current operator at the final time of the Keldysh contour. An example of such a diagram is given in Fig. 5.1 (c).

## 5.3 Two Types of NCA for Electron–phonon Coupling

In this section we lay out the construction of NCA-like approximations in two limits: First, a *bare NCA* based on self-consistently resummed *second order* perturbation theory for the electron–phonon and dot-lead and electron–phonon Hamiltonians. Second, a *dressed NCA* in which the Hamiltonian is modified by a Lang-Firsov transformation so that the coupling

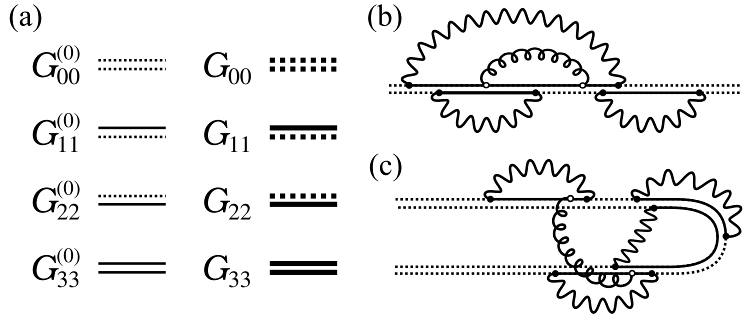


Figure 5.1: (a) The elements of the unperturbed propagator $G_{\alpha\alpha}^{(0)}$ (left column, thin lines) and of the NCA propagator $G_{\alpha\alpha}$ (right column, bold lines). The upper line represents spin up occupation and the lower line spin down occupation. The dotted line signifies that the spin level is unoccupied, while a solid line marks it as occupied. (b) An example of a diagram included in the reduced propagator $G_{00}$. Electronic hybridization lines are shown as wiggly lines, and phonon interaction lines as gluon lines. (c) An example of a diagram on the Keldysh contour with inter-branch lines and a special hybridization line ending at the final time, corresponding to a contribution to the current.

Hamiltonian becomes a phonon-dressed dot-lead coupling, and includes non-crossing diagrams composed of phonon-dressed hybridization lines[126, 187, 220]. Both approximations can be extended to higher orders, or used as the preliminary step withing a numerically exact bold-line QMC algorithm. We initially formulate these two types of NCA for the symmetric Anderson–Holstein model in the following two subsections, then discuss the asymmetric case.

### 5.3.1 Weak coupling perturbation theory

The *bare NCA* approximation is specified by the following equations

$$\mathbf{G}^{-1} = \mathbf{G}_0^{-1} - \mathbf{\Sigma}^{\ell} - \mathbf{\Sigma}^{b}, \tag{5.22}$$

with $\mathbf{G}$, $\mathbf{G}_0$ and $\mathbf{\Sigma}$ matrices (diagonal, in the cases of interest here) in the Hilbert space of the decoupled dot, and the lead ($\ell$) and phonon ($b$) self energies$\mathbf{\Sigma}$ given by:

$$\Sigma^{\ell}_{\alpha\alpha}(t_1, t_2) = \sum_{\beta} G_{\beta\beta}(t_1, t_2) \times \Delta^{\beta}_{\alpha\alpha}(t_1, t_2) \tag{5.23}$$

$$\Sigma^{b}_{\alpha\alpha}(t_1, t_2) = G_{\alpha\alpha}(t_1, t_2) \times \Lambda_{\alpha\alpha}(t_1, t_2) \tag{5.24}$$

with the lead hybridization function

$$\begin{aligned} \Delta^{\beta}_{\alpha\alpha}(t_1, t_2) = & \sum_{\sigma} \langle\alpha|d_{\sigma}|\beta\rangle\langle\beta|d^{\dagger}_{\sigma}|\alpha\rangle \sum_{k\in\ell} |t_k|^2 \mathrm{Tr}_{\ell}\left[\rho_{\ell} c^{\dagger}_{k\sigma}(t_1) c_{k\sigma}(t_2)\right] \\ & + \sum_{\sigma} \langle\alpha|d^{\dagger}_{\sigma}|\beta\rangle\langle\beta|d_{\sigma}|\alpha\rangle \sum_{k\in\ell} |t_k|^2 \mathrm{Tr}_{\ell}\left[\rho_{\ell} c_{k\sigma}(t_1) c^{\dagger}_{k\sigma}(t_2)\right]. \end{aligned} \tag{5.25}$$

We also define the lesser and greater hybridization functions $\Delta^{<,>}_{\ell}(\tau_1, \tau_2) = \sum_{k\in\ell} |t_k|^2 \mathrm{Tr}_{\ell}\left[\rho c^{\dagger}_{k\sigma}(\tau_1) c_{k\sigma}(\tau_2)\right]$ for each lead $\ell$ and times $\tau_1$, $\tau_2$ on the Keldysh contour. $\Delta^{>}_{\ell}$ is used when $\tau_1$ precedes $\tau_2$, and $\Delta^{<}_{\ell}$ is used otherwise. The dot–lead hybridization function for each lead can be expressed in terms of the coupling densities $\Gamma_{\ell}(\omega)$ and the initial occupation of that lead:

$$\Delta^{>}_{\ell}(t_1, t_2) = i \int_{-\infty}^{\infty} \frac{d\omega}{\pi} e^{-i\omega(t_1 - t_2)} \Gamma_{\ell}(\omega) \left[1 - f_{\ell}(\omega - \mu_{\ell})\right], \tag{5.26}$$

$$\Delta^{<}_{\ell}(t_1, t_2) = -i \int_{-\infty}^{\infty} \frac{d\omega}{\pi} e^{-i\omega(t_1 - t_2)} \Gamma_{\ell}(\omega) f_{\ell}(\omega - \mu_{\ell}). \tag{5.27}$$

We similarly define the phonon hybridization function

$$\Lambda_{\alpha\alpha}(t_1, t_2) = \langle\alpha|\left(n_d(t_1) - \delta\right)\left(n_d(t_2) - \delta\right)|\alpha\rangle \times \\ \sum_q \lambda_q^2 \mathrm{Tr}_b\left[\rho_b\left(b_q^\dagger(t_1) + b_q(t_1)\right)\left(b_q^\dagger(t_2) + b_q(t_2)\right)\right]. \tag{5.28}$$

This is analogous (but not identical) to the pseudoparticle NCA approximation of ref. 262. Since the the electron–phonon coupling $V_b$ does not modify the electronic state of the dot, one can write

$$\langle\alpha|\left(n_d(t_1) - \delta\right)\left(n_d(t_2) - \delta\right)|\alpha\rangle = \left(n_d^{(\alpha)} - \delta\right)^2. \tag{5.29}$$

We also define the bath correlation function,

$$B_q(t_1, t_2) = \mathrm{Tr}_b\left\{\rho_b\left(b_q^\dagger(t_1) + b_q(t_1)\right)\left(b_q^\dagger(t_2) + b_q(t_2)\right)\right\}. \tag{5.30}$$

It can be expressed in terms of the frequency $\omega_q$ and the inverse temperature $\beta$ of the local phonon modes, $B_q(t) = \coth(\beta\omega_q/2)\cos(\omega_q t) - i\sin(\omega_q t)$, if we consider a bath initially composed of free harmonic phonon modes. Thus, it is possible to recast the phonon hybridization function as $\Lambda_{\alpha\alpha}(t_1 - t_2) = \left(n_d^{(\alpha)} - \delta\right)^2 \times \Lambda_b(t_1 - t_2)$, where

$$\Lambda_b(t_1 - t_2) = \sum_q \lambda_q^2 B_q(t_1 - t_2). \tag{5.31}$$

Just as the electronic hybridization function is described by $\Gamma_\ell(\omega)$, the phonon bath is usually characterized by its spectral density, $J(\omega) = \frac{\pi}{2}\sum_q \left(\lambda_q^2/\omega_q\right)\delta(\omega - \omega_q)$. In particular,

$$\Lambda_b(t_1 - t_2) = \frac{2}{\pi}\int d\omega J(\omega)\omega B_\omega(t_1 - t_2). \tag{5.32}$$

Fig. 5.2 and Fig. 5.3 illustrate the diagrams included in the self energy of the *bare NCA* approach (for the symmetric case $\delta = 1$). The wiggly lines in Fig. 5.2 denote the dot–lead

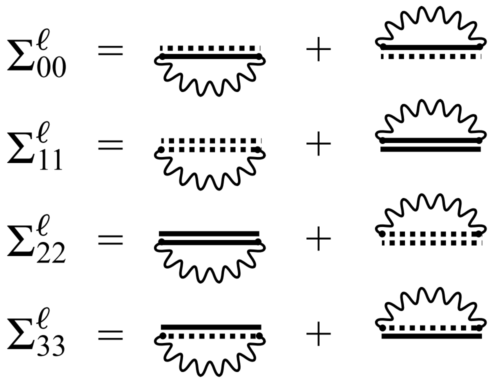


Figure 5.2: The electron hybridization diagrams included in the bare NCA self energy, where the wiggly lines denote electronic dot–lead hybridization lines. The pairs of straight lines represent the dot's electronic state, with the two lines standing for the two possible spins: a solid line represents an occupied spin level, whereas dashed lines stand for empty spin levels.

hybridization $\Delta^{\beta}_{\alpha\alpha}(t_1 - t_2)$, while the phonon lines of Fig. 5.3 symbolize the phonon coupling $\Lambda_{\alpha\alpha}(t_1 - t_2)$. The computation of the Green's function from the Dyson equation using this approximate self energy embodies a self-consistent perturbative expansion including the lowest order skeleton diagrams in both the dot–lead hybridization and electron–phonon coupling. We expect this bare NCA approach to be more applicable in the regime where both $\lambda$ and $\Gamma$ are small. Additionally, the Green's function resulting from the bare NCA does not contain certain multi–phonon excitations, related to crossing diagrams, which might be expected to affect the dot electron if the phonon relaxation is slow. This implies that the bare NCA is more accurate in the limit of the fast phonon bath.

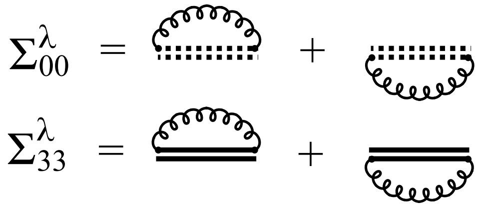


Figure 5.3: The phonon interaction diagrams for the bare NCA self energy in the symmetric case $\delta = 1$. The curly lines denote phonon interaction lines, and straight lines are as in Fig. 5.2.

### 5.3.2 Strong coupling perturbation theory

In this section, we present a version of the non-crossing approximation more suitable to strong coupling between the dot and the phonon bath to the propagator formulation. This approach, which we will refer to as the *dressed NCA*, has previously been employed within a standard Green's function formulation in Ref. 220, 126.

We begin by performing the unitary Lang–Firsov transformation $\tilde{H} = SHS^{-1}$ with $S = e^{\frac{\lambda}{\omega_0}(b^\dagger - b)n_d}$, which eliminates the explicit electron–phonon coupling in the Hamiltonian. We set the unperturbed Hamiltonian to be $H_0 = H_d + H_b + V_b$. After the transformation, this becomes

$$\tilde{H}_0 = \tilde{\epsilon}_d \tilde{n}_d + \tilde{U} \tilde{n}_\uparrow \tilde{n}_\downarrow, \tag{5.33}$$

$$\tilde{V}_\ell = \sum_{k\in\ell} \sum_{\sigma} \left[ t_k \tilde{d}_\sigma c^\dagger_{k\sigma} + t^*_k \tilde{d}^\dagger_\sigma c_{k\sigma} \right]. \tag{5.34}$$

In the above expressions, the bare dot energy $\epsilon$ and the Coulomb interaction strength $U$ are replaced by the renormalized quantities

$$\tilde{\epsilon} = \epsilon + (2\delta - 1)\lambda^2/\omega_0, \tag{5.35}$$

$$\tilde{U} = U - 2\lambda^2/\omega_0. \tag{5.36}$$

Also, the dot electron creation and annihilation operators become

$$\tilde{d_\sigma} = e^{-\frac{\lambda}{\omega_o}(b^\dagger - b)} d_\sigma, \tag{5.37}$$

$$\tilde{d_\sigma^\dagger} = e^{\frac{\lambda}{\omega_o}(b^\dagger - b)} d_\sigma^\dagger. \tag{5.38}$$

All pairs of hybridization events are therefore connected by an infinite set of phonon hybridization lines generated by these exponential phonon displacement operators.

Within the dressed NCA approximation for the self energy, we consider only the dressed phonon lines appearing along the noncrossing fermionic hybridization lines, as illustrated in Fig. 5.4. With this assumption, the effect of the electron–phonon interaction is simply to reweigh each fermionic hybridization line with a phonon-dependent factor, such that the NCA self energy takes the form

$$\begin{aligned} \tilde{\Sigma}^{\ell}_{\alpha\alpha}(t_1 - t_2) &= w(t_1 - t_2) \\ &\quad \times \sum_{\beta} \Delta^{\beta}_{\alpha\alpha}(t_1 - t_2) G^{(0)}_{\beta\beta}(t_1 - t_2). \end{aligned} \tag{5.39}$$

The phonon weight $w(t_1 - t_2)$ is given by

$$w(t) = \exp\left\{ -\sum_q \left(\frac{\lambda_q}{\omega_q}\right)^2 \times \left[(1 - \cos\omega_q t)\coth(\beta\omega_q/2) + i\sin\omega_q t\right] \right\} \tag{5.40}$$

In terms of the bath spectral density $J(\omega)$, this can be written as

$$w(t) = \exp\left\{-Q_2(t) - iQ_1(t)\right\},$$

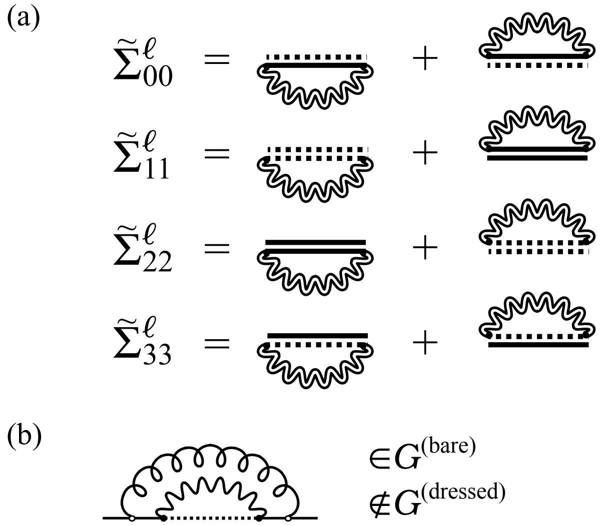


Figure 5.4: (a) The diagrams representing the different matrix elements of the dressed NCA self energy. The wiggly double lines denote electron hybridization lines dressed by phonon interactions. (b) An example of a bare NCA diagram of the lowest order is not included in the dressed NCA diagrams.

where

$$Q_1(t) = \frac{2}{\pi}\int d\omega \frac{J(\omega)}{\omega} \sin\omega t, \tag{5.41}$$

$$Q_2(t) = \frac{2}{\pi}\int d\omega \frac{J(\omega)}{\omega} (1-\cos\omega t)\coth(\beta\omega/2). \tag{5.42}$$

The dressed NCA self energy includes many phonon interactions not included in the bare NCA. The self energy diagrams composed of the transformed dot operators $\tilde{d}_\sigma$ and $\tilde{d}^\dagger_\sigma$ can be expanded in terms of the bare dot operators and effectively contain all the hybridization diagram within the wiggly double lines. Also, the polaron shift of $U$ and $\epsilon$ is explicitly included within the dressed NCA, but not the bare NCA. One might expect it to be a more appropriate approximation in the polaron limit. On the other hand, it also misses some contributions that are included in the bare NCA (see Fig. 5.4 (b)) and over-emphasizes others, and at weak coupling to the phonons it might be expected to be less accurate. The two approximations are therefore somewhat complementary, if in a non-rigorous sense; it is reasonable to assume that conclusions supported by both may be robust to the nature of the approximations, while conclusions supported by only are suspect and should be investigated further.

### 5.3.3 NCA for asymmetric model

We now briefly discuss the structure of the non–crossing approximation for the case of an asymmetric Anderson–Holstein model in which the counter term is not included (i.e. $\delta = 0$ in Eq. (5.4)). The phonon can then only be created or destroyed in the single electron state or the doubly occupied state, not in the empty state. Such a model might be considered a more physically realistic description of a quantum junction, where one is interested in vibrational states coupled to electrons.

In the bare NCA calculation, the phonon coupling lines only connect points with occupied electron states. The interaction diagrams for the bare NCA self energy therefore no longer have the symmetric structure of Fig. (5.4), but rather include a different number of phonon inclusions for each of the matrix elements. This is illustrated in Fig. (5.5).

For the dressed NCA, the same Lang-Firsov transformation is performed to eliminate the explicit electron–phonon coupling. The dressed coupling Hamiltonian then remains the

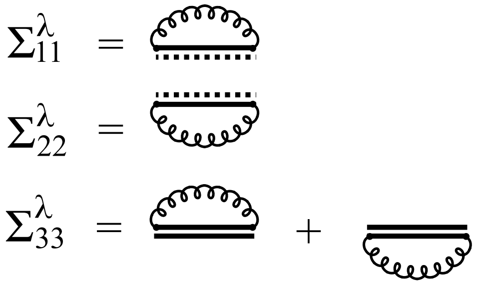


Figure 5.5: The phonon interaction diagrams for asymmetrical model.

same as in the symmetric case. However, the renormalized energy becomes

$$\tilde{\epsilon} = \epsilon - \frac{\lambda^2}{\omega_0}, \tag{5.43}$$

while the renormalized interaction remains the same as Eq. (5.36).

With this coupling, $\epsilon = 0$ does not correspond to a particle–hole symmetric point. In the absence of dot–lead coupling, the charge transfer bands are centered around $\omega_+ = \frac{U}{2} + \frac{\lambda^2}{\omega_0}$ and $\omega_- = -\frac{U}{2} + 3\frac{\lambda^2}{\omega_0}$.

## 5.4 Results

We now discuss the application of the two NCA approaches described above to the Anderson–Holstein impurity model, focusing on a case where the dot has degenerate spin levels ($\epsilon_\uparrow = \epsilon_\downarrow = \epsilon_d$) and obeys particle–hole symmetry ($\epsilon_d = -\frac{U}{2}$) in the absence of phonons. The leads are assumed to be flat with a soft cutoff: $\Gamma_\ell(\omega) = \frac{\Gamma_\ell}{(1+e^{\nu(\omega-\Omega_c)})(1+e^{-\nu(\omega+\Omega_c)})}$, where $\Omega_c = 10$ and $\nu = 10$. We consider only symmetrical couplings to the left and right leads,

$\Gamma_L = \Gamma_R = 0.5\Gamma$, and apply the bias voltage $V$ symmetrically such that the chemical potentials are given by $\mu_L = -\mu_R = 0.5V$.

The methods we have described are suitable for the exploration of systems containing multiple electron and phonon baths with complicated densities of states, but we focus on a phonon bath with single mode, $H_b = \omega_0 b^\dagger b$. The electron–phonon coupling Hamiltonian becomes $V_b = \lambda(b^\dagger + b)\,(n_d - \delta)$ and the strength is characterized by by the parameter $\lambda$. We assume that all baths are initially at the same inverse temperature $\beta = 10/\Gamma$.

To calculate the spectral function $A(\omega)$ by the double probe scheme, we attach a pair of auxiliary leads to the system and measure the corresponding auxiliary currents. The spectral density of the auxiliary leads is a Gaussian delta function $\Gamma_a(\omega, \omega') = \frac{\eta}{\delta_a\sqrt{\pi}} e^{-[(\omega-\omega')/\delta_a]^2}$ where $\eta = 10^{-4}\Gamma$ and $\delta_a = 10^{-2}\Gamma$. The dot is assumed to be initially empty, and the coupling to the thermally equilibrated leads and phonon bath is turned on at time $t = 0$. The auxiliary spectral function exhibits some transient behavior, and approaches the physical steady state spectral function at sufficiently long time, as discussed in Ref. 122.

### 5.4.1 Symmetric Model

We first consider the system which includes the counter term, $\delta = 1$. For this case, the electron–phonon coupling does not break particle–hole symmetry and the spectral function remains symmetric.

#### Transient dynamics

The left panels of Fig. 5.6 and Fig. 5.7 show the transient evolution of the spectral function $A(\omega; t)$. The corresponding right panels display single frequency cuts through this data, highlighting the time evolution of the central peak ($\omega = 0$) and the charge transfer (CT) peak ($\omega/U = 0.5$). We observe an overshooting of the spectral function at short time due to the instantaneous coupling between the dot and the leads. The bare NCA results (Fig. 5.6) exhibit oscillatory behavior in the amplitude of the central peak. We observe that this is composed of a slower oscillation with a period of $2\pi/\omega_0$, which is associated with the phonon frequency; and a rapid oscillation with a period of $2\pi/U$, which comes from the static energetics of the system. However, in the dressed NCA results (Fig. 5.7), oscillatory behavior consistent with the phonon frequency is not apparent. The oscillatory behavior predicted by the bare NCA is consistent with predictions made for the Anderson–Holstein model in the spinless $U = 0$[241] and $U = \infty$ cases[252], where the local density of states at $\omega = 0$ approaches the steady state in an oscillating manner with the periodicity of the phonon mode. Here, the time-evolution of the entire frequency dependent auxiliary spectral function additionally reveals the transient effect of electron–phonon coupling on the charge transfer peaks.

At long times, the bare NCA exhibits a strong suppression of the CT peaks when the phonon frequency is small. However, this suppression of the CT peaks is not nearly as evident in the dressed NCA results. Conversely, the dressed NCA shows a strong enhancement of the central peak at low phonon frequencies, which is not present in the bare NCA results.

### Equilibrium steady state spectral function

We next explore the equilibrium spectral function $A(\omega)$ of the system in the limit of long times, where the system has reached its steady or equilibrium state. We consider two types of cuts through the parameter space: the first is the dependence on the phonon frequency $\omega_0$

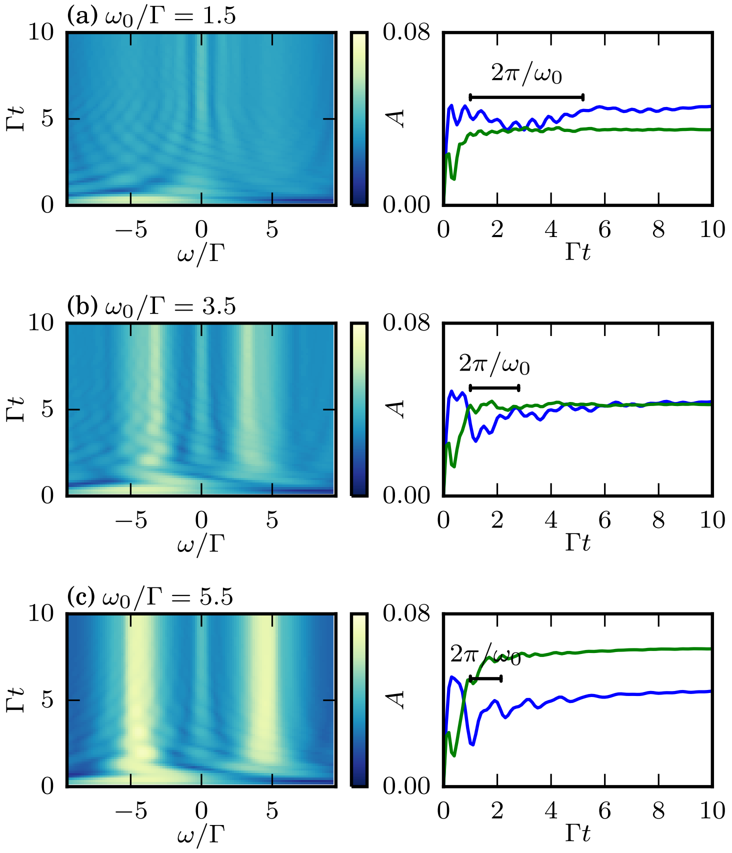


Figure 5.6: (left panels) The time evolution of the spectral function $A(\omega;t)$ within the **bare NCA** is shown for different phonon frequencies. (right panels) Time dependence of cuts at $\omega = 0$ (blue) and $\omega = U/2$ (green). The time scale $2\pi/\omega_0$ related to the phonon frequency is also plotted for comparison. A symmetric dot with $U = -2\epsilon = 10\Gamma$ is considered at equilibrium $V = 0$. The phonon coupling is set to $\lambda = 1.5\Gamma$ and the counter term is symmetric ($\delta = 1$). The inverse temperature of all baths is $\beta = 10/\Gamma$.

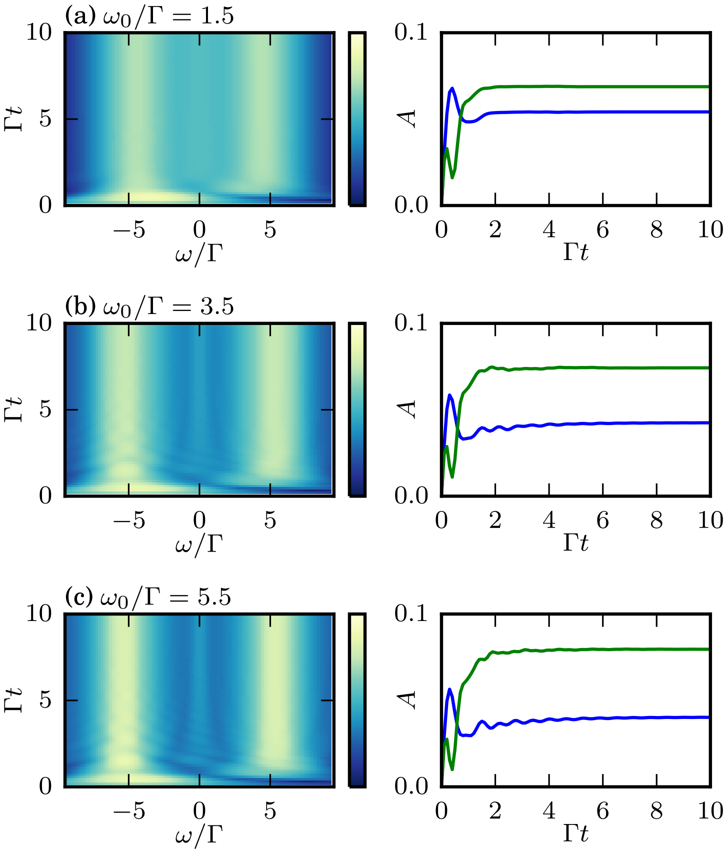


Figure 5.7: The same as Fig. 5.6 within the **dressed NCA**. A symmetric dot with $U = -2\epsilon = 10\Gamma$ is considered at equilibrium $V = 0$. The phonon coupling is symmetric with $\lambda = 1.5\Gamma$ and the inverse temperature of all baths is $\beta = 10/\Gamma$.

at constant dot-phonon coupling strength $\lambda$, and the second is the $\lambda$ dependence at constant $\omega_0$. Here, too, the bare and dressed NCA predict qualitatively different behaviors.

In Fig. 5.8, $A(\omega)$ is shown for a range of phonon frequencies at intermediate electron–phonon coupling $\lambda = 1.5$. Within bare NCA, shown in panel (a), a set of features at $\omega = \pm n\omega_0$ with $n \in \{1, 2, 3\}$ is visible at low frequencies. These features, corresponding to Kondo replicas or sidebands[107–109, 111, 114, 119], appear as a sequence of positive peaks at $\omega = \pm(2n+1)\omega_0$ and negative peaks at $\omega = \pm 2n\omega_0$, and are related to interference effects. In the literature, the Anderson–Holstein impurity model is mostly assumed to be spinless ($U = 0$), and one observes multiple positive side bands due to a resonance with the phonon. For a generic Anderson–Holstein model, negative peaks have previously been predicted in the $T \sim 0$ regime by perturbation theory, but not are exhibited within numerical renormalization group calculation[114, 119]. However, our calculation shows both positive and negative side peaks exist at a finite temperature for generic Anderson–Holstein model. In the high-frequency regime, the Kondo replicas die out and the CT peaks appear. The CT peaks are suppressed by coupling to a low frequency phonon mode, which implies that phonon-induced tunneling dominates the single particle excitation spectrum in this regime.

Replica-like features can also be observed at $\omega = \pm\omega_0$ in the dressed NCA, which is plotted in Fig. 5.8 (b). However, these side peaks are substantially weaker than those observed in the bare NCA calculation. In the dressed NCA the CT peaks are shifted by the reorganization energy, such that their central frequencies are located at $\omega_\pm = \pm\left(\epsilon + \frac{\lambda^2}{\omega_0}\right)$ (as illustrated by the dashed line). A significant enhancement in $A(\omega)$ occurs when the two renormalized CT peaks cross each other. In the low frequency regime $\omega_0 \leq \frac{\lambda^2}{|\epsilon|}$, the two CT peaks merge and form a wide central peak which is clearly unrelated to the Kondo effect.

The Kondo peak only develops in the high frequency regime, and in general it is strongly suppressed for a wide range of parameters.

The $\omega_0$ dependence of the central peak $A(\omega = 0)$ exhibits consistent behavior for the two flavors of NCA only at high frequencies (Fig. 5.8 (c)). At low frequencies, both approximations exhibit enhancement of the central peak, but the context and perhaps the mechanism of the enhancement is different between the two cases. In the bare NCA, the amplitude of the Kondo peak is enhanced as $\omega_0$ decreases because the replicas of the Kondo peak merge when the phonon quasi-states become nearly-degenerate as $\omega_0$ decreases. In the dressed NCA, on the other hand, the enhancement is maximal where the two CT peaks merge at $\omega_0^* = \lambda^2/\epsilon$. The contrast with the bare case is even more notable when one considers that in the bare NCA the CT peaks are almost entirely suppressed at low frequencies.

In Fig. 5.9 We repeat the previous analysis in a different plane of the parameter space, by taking a cut at a constant (low) phonon frequency $\omega_0$ and a range of $\lambda$ values. In the bare NCA (Fig 5.9 (a)), the CT peaks are suppressed as $\lambda$ increases. One can observe a set of ridge-like features developing along with a strong enhancement of the central Kondo peak. In the large $\lambda$ regime, the developed side peaks shifted linearly with $\lambda$ with a spacing of approximately $\omega_0$ between peaks in frequency. These features resemble Kondo replicas[107, 109, 111, 114, 119], but a closer inspection reveals behavior more complicated than simply side peaks generated at the phonon frequency $|\omega| = n\omega_0$. A sharp Kondo peak is only apparent before the crossing point of the ridges. It is significantly enhanced at the crossing point, and is either completely suppressed or split beyond this point.

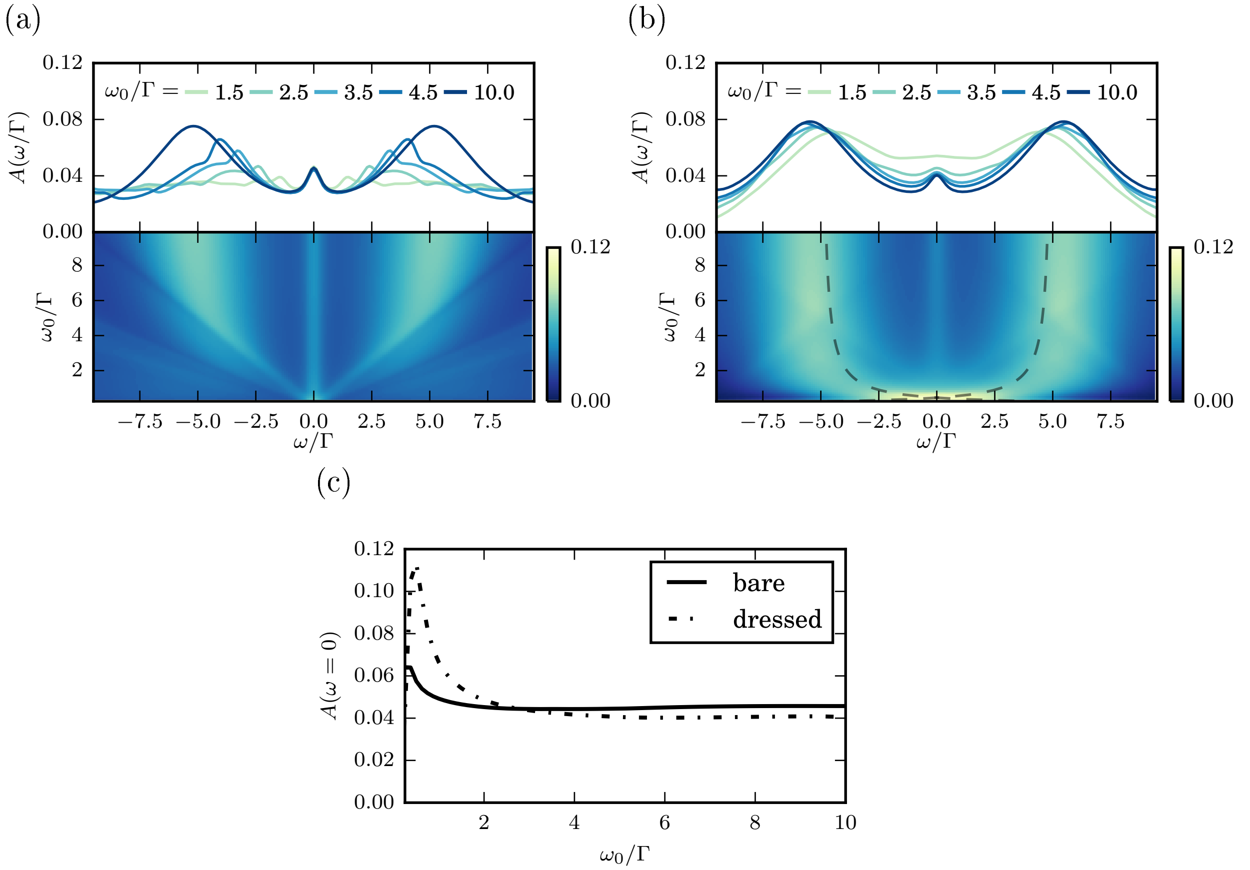


Figure 5.8: The $\omega_0$ dependence of the spectral function $A(\omega)$ is calculated by (a) **bare NCA** and (b) **dressed NCA** for a symmetric dot at equilibrium $V = 0$ with $U = -2\epsilon = 10\Gamma$. The phonon coupling is $\lambda = 1.5\Gamma$ and the counter term is symmetric ($\delta = 1$). All baths at the same inverse temperature $\beta = 10/\Gamma$. The dashed lines indicate the renormalized charge transfer peak at $\omega_\pm = \pm\left(\epsilon + \frac{\lambda^2}{\omega_0}\right)$. The $\omega_0$-dependence of the central peak at $\omega = 0$ is plotted in (c).

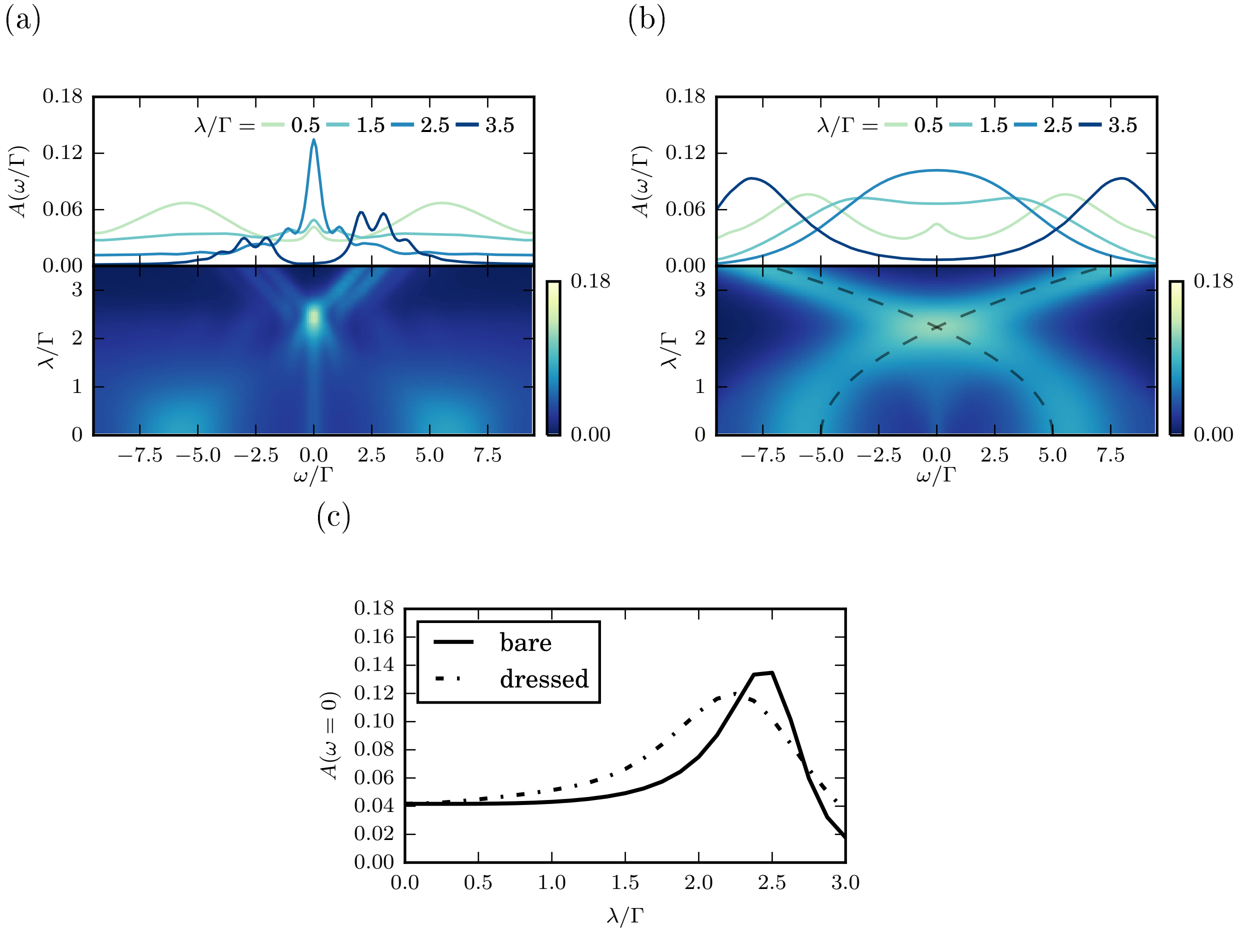


Figure 5.9: The $\lambda$ dependence of the spectral function $A(\omega)$ as calculated within the (a) **bare NCA** and (b) **dressed NCA** for a symmetric dot with $U = -2\epsilon = 10\Gamma$ at equilibrium $V = 0$. The phonon coupling is $\omega_0 = 1.0\Gamma$ and the counter term is symmetric ($\delta = 1$). All baths at the same inverse temperature $\beta = 10/\Gamma$. The dashed lines indicate the renormalized charge transfer peak at $\omega_{\pm} = \pm\left(\epsilon + \frac{\lambda^2}{\omega_0}\right)$. The $\lambda$-dependence of the central peak at $\omega = 0$ is plotted in (c).

No Kondo replicas are observed within the dressed NCA(Fig. 5.9 (b)). The CT peaks are again renormalized, and appear centered at $\omega_{\pm} \approx \pm\left(\epsilon + \frac{\lambda^2}{\omega_0}\right)$ as illustrated by the dashed lines. The crossing at $\lambda^* = \sqrt{\epsilon\omega_0}$ leads to a strong enhancement near $\omega = 0$. The Kondo peak is only observable for $\lambda < \lambda^*$, and is widened beyond the point where it can be distinguished from the CT bands before the crossing point is reached. This widening effect is not observed in the bare NCA. Past the crossing point, no central feature is visible, in agreement with the bare NCA.

While the striking non-monotonic enhancement of the $\omega = 0$ spectral function is predicted by both approximations, it occurs at a different value of $\lambda$ in each case (see Fig. 5.9 (c)). The peak in the dressed NCA occurs precisely at the value of $\lambda$ for which the effective, dressed $\tilde{U}$ change sign. In this regard, the result is reminiscent of the NRG prediction of Hewson and Meyer[114], where the negative–$\tilde{U}$ Anderson–Holstein model flows to the $U = 0$ behavior. Within the bare NCA, the peak value of $A(\omega = 0)$ occurs for a slightly larger value of $\lambda$. Here, the self–consistency of the perturbation theory presumably captures, in an approximate manner, the terms leading to negative–$\tilde{U}$ behavior as well. Lastly, it should be mentioned that this non-monotonic behavior is consistent with the prediction of Ref. 115. We return to this point later in the manuscript.

### Nonequilibrium steady state spectral function

We now consider a nonequilibrium system driven by a bias voltage $V = 2\Gamma$. The $\omega_0$ dependence of $A(\omega)$ is plotted in Fig. 5.10. The voltage splitting of Kondo peak[121, 265] can be observed in both approximations. The central Kondo peak splits into two peaks at $\omega = \pm V/2$ independently of the phonon frequency. Kondo replicas are not clearly distin-

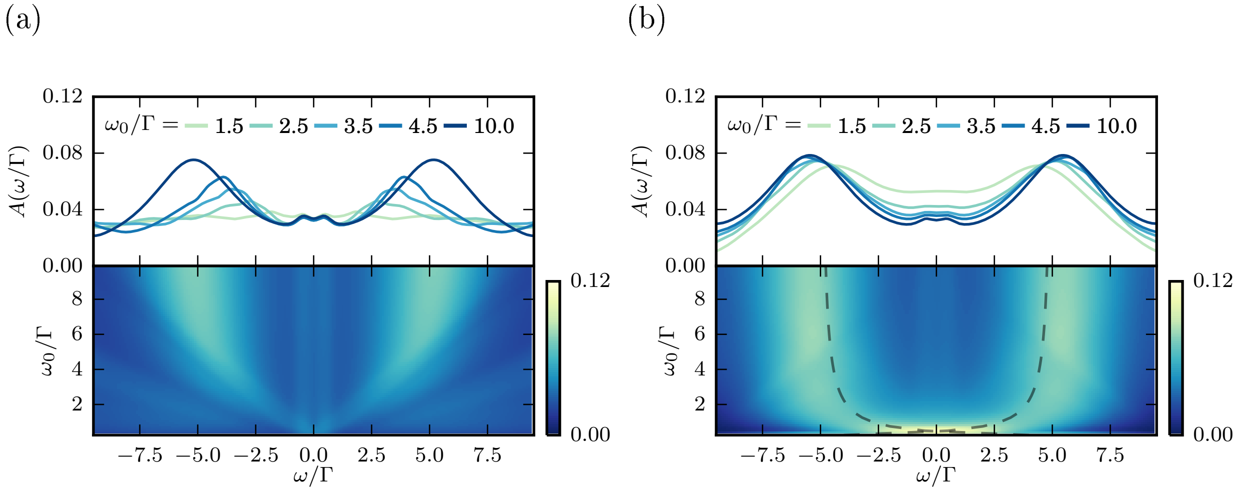


Figure 5.10: The $\omega_0$-dependence of the spectral function $A(\omega)$ for a symmetric dot with $U = -2\epsilon = 10\Gamma$ under a nonequilibrium symmetrically applied bias voltage $V = 2\Gamma$ within the (a) **bare NCA** and (b) **dressed NCA**. The phonon coupling is $\lambda = 1.5\Gamma$ and the counter term is symmetric ($\delta = 1$). All baths at the same inverse temperature $\beta = 10/\Gamma$.

guishable, since the splitting smears out the associated features. However, a set of linearly dependent signatures remains visible.

### 5.4.2 Asymmetric Model

In the following subsection, we consider an Anderson–Holstein model without a counter term, i.e. $\delta = 0$ in Eq. (5.4). While the isolated dot Hamiltonian is still assumed to remain particle-hole symmetric, the electron–phonon coupling breaks the particle–hole symmetry of the system and results in an asymmetric spectral function. The two NCA formulations we employ take this asymmetry into account in different ways, as pointed out in sec. 5.3.3. In addition to the spectral function, we study the effects of the symmetry breaking on transport properties. This is of particular interest, because under a symmetrically applied bias the differential conductance is a symmetric function of frequency even without particle–hole symmetry. Additionally, one may not be able to observe the replicas directly in a transport experiment, due to the nonequilibrium shifting or suppression of the Kondo peak, which would also affect the replicas. We show that an indirect experimental signal of the replica effect may remain.

### Transient dynamics

Within the bare NCA, the CT peaks and Kondo peak oscillate at the phonon frequency $\omega_0$, but the oscillations are manifested in different ways (Fig. 5.11, left panels). In particular, the CT peaks oscillate in *frequency*, while the Kondo peak oscillates in *amplitude*. At short times and in the adiabatic limit, the CT peak oscillations can be explained by oscillating energy levels ($\tilde{\epsilon}_\sigma = \epsilon_\sigma + \frac{2\lambda}{\omega_0}\sin(\omega_0 t + \phi_0)$) with some unknown initial phase. This is illustrated by the black dashed lines in the left panels of Fig. 5.11. All these features are washed out in the dressed NCA.

### Steady state spectral function

To explore the effects of phonons on the equilibrium spectral function, we once again plot first the $\omega_0$ dependence at constant $\lambda$, and then the $\lambda$ dependence at constant $\omega_0$. Within the bare NCA, the Kondo replica features can clearly be seen in Fig. 5.13 (a), but harder to distinguish in the cuts. They are mixed with a variety of other effect including the low-frequency smearing of the Kondo resonance and the suppression of the positive CT peak. The replica effect and the above-mentioned CT suppression are both stronger at positive

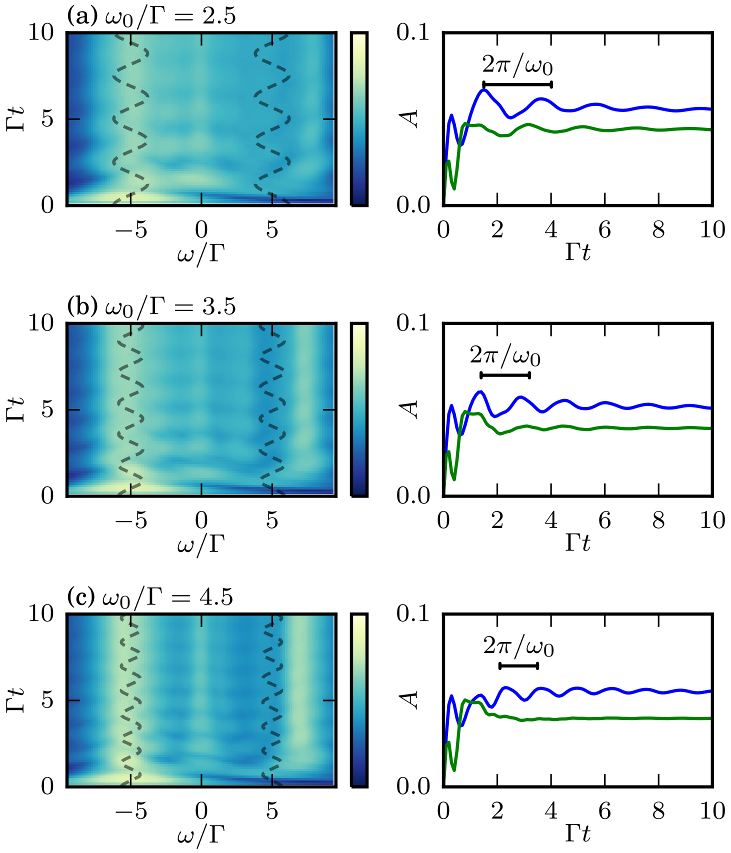


Figure 5.11: (left panels) The time evolution of the spectral function $A(\omega;t)$ within the **bare NCA** is shown for different phonon frequencies. The frequency oscillations of the CT peaks along with an illustration of the expected energy oscillations in the adiabatic limit (dash lines) are also exhibited. (right panels) Time dependence of cuts at $\omega = 0$ (blue) and $\omega = U/2$ (green). The time scale $2\pi/\omega_0$ related to the phonon frequency is also plotted for comparison. The dot is symmetric with $U = -2\epsilon = 10\Gamma$ at equilibrium $V = 0$. The phonon coupling is $\lambda = 1.5\Gamma$ and the counter term is asymmetric ($\delta = 0$). The inverse temperature of all baths is $\beta = 10/\Gamma$.

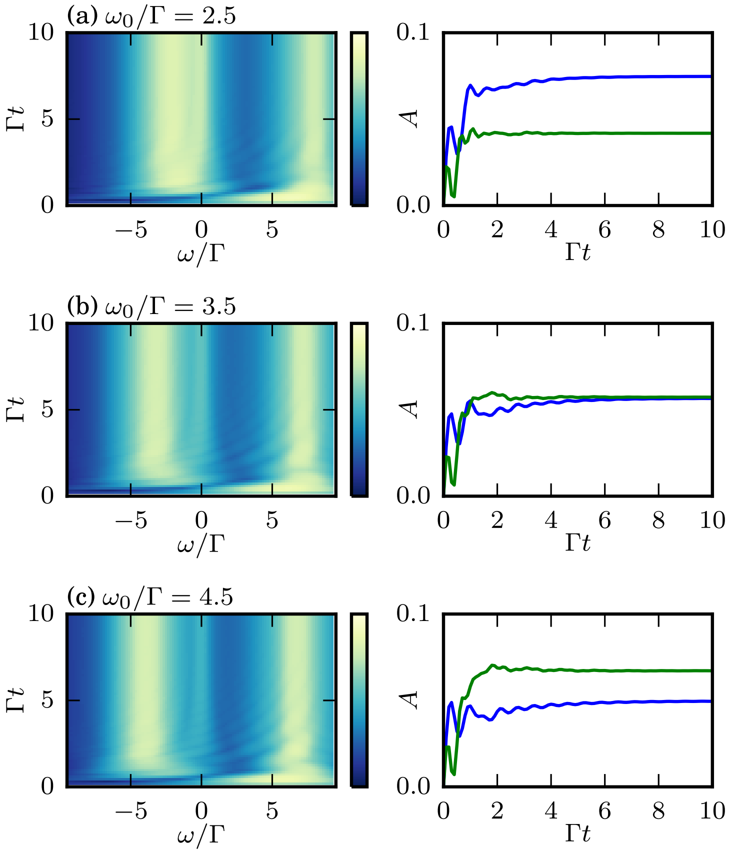


Figure 5.12: The same as Fig. 5.11 within the **dressed NCA**. The dot is symmetric with $U = -2\epsilon = 10\Gamma$ at equilibrium $V = 0$. The phonon coupling is asymmetric ($\delta = 0$) with $\lambda = 1.5\Gamma$ and the inverse temperature is $\beta = 10/\Gamma$.

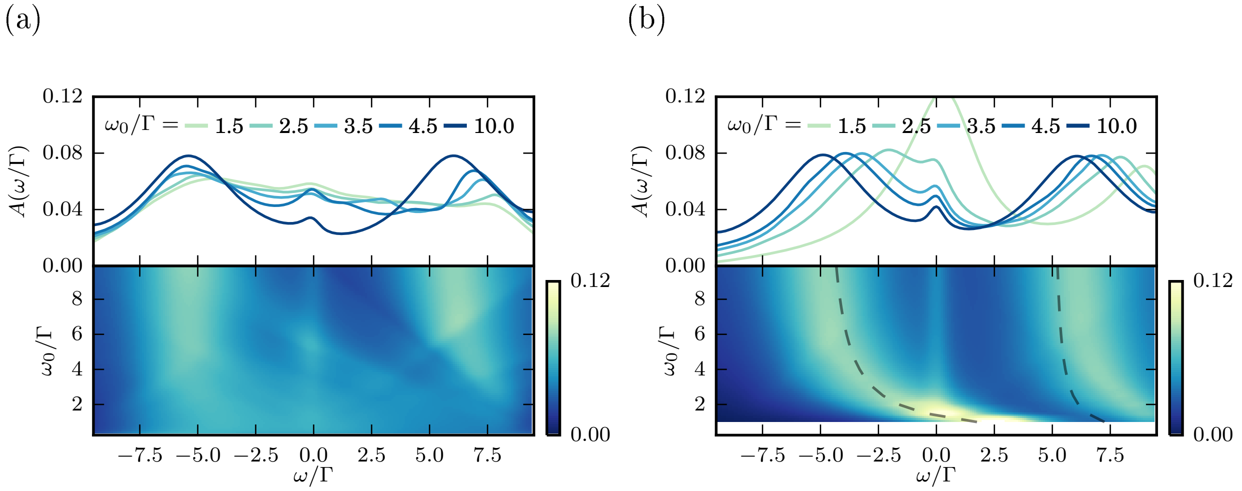


Figure 5.13: The $\omega_0$-dependence of the spectral function $A(\omega)$ for a dot in equilibrium as calculated within the (a) **bare NCA** and (b) **dressed NCA**. The electron–phonon coupling is asymmetric ($\delta = 0$) and the coupling strength is $\lambda = 1.5\Gamma$. The dot is symmetric with $U = -2\epsilon = 10\Gamma$. All baths at the same inverse temperature $\beta = 10/\Gamma$.

frequencies. At small phonon frequencies, the Kondo resonance merges with the negative CT peak.

At the intermediate phonon frequency $\omega_0 = |\epsilon_\sigma - U|$ where the replicas are aligned with the CT peaks, a non-monotonic enhancement of the central peak is evident, and is especially strong at large $\lambda$. This can be seen more clearly in the cut shown in Fig. 5.15 (c). We believe this is due to a phonon-assisted process which is similar to the Kondo spin-flip process, and which becomes possible for electrons with energies closed to the chemical potential[22, 115]. The effects described here are largely washed out in the dressed NCA.

We continue to investigate the $\lambda$ dependence at constant $\omega_0$. Here, we plot the results for both approximations at a relatively large $\omega_0$ (Fig. 5.14). The bare NCA (panel (a)) shows

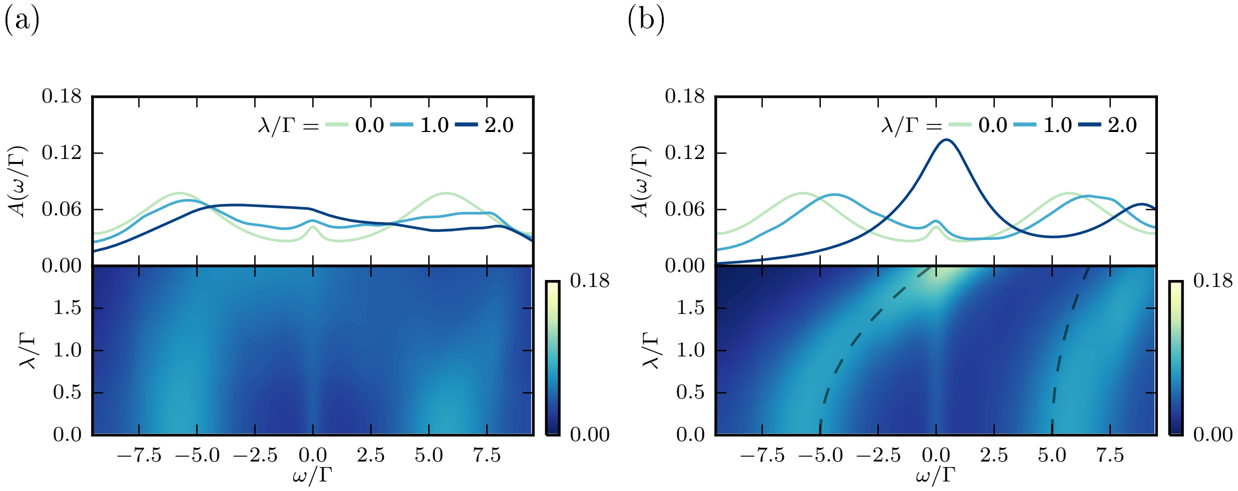
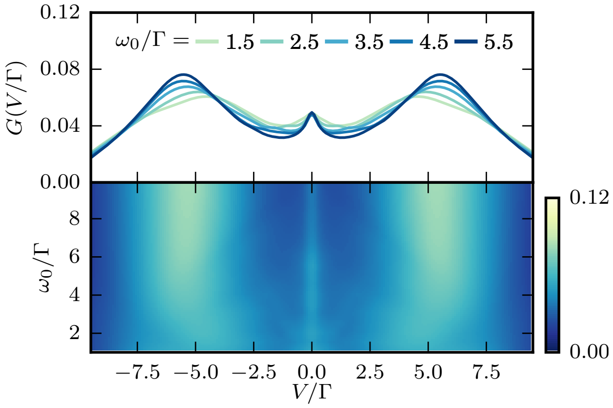


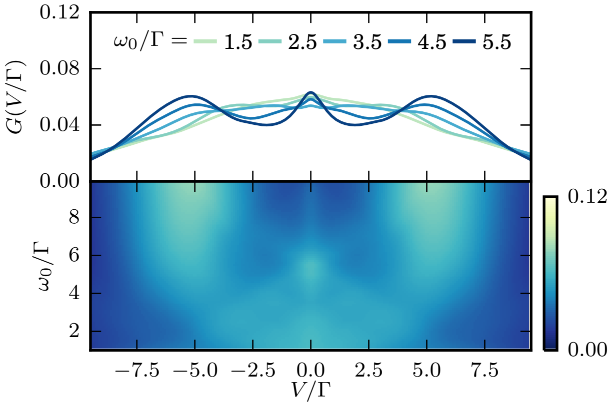
Figure 5.14: The $\lambda$-dependence of the spectral function $A(\omega)$ as calculated within the (a) **bare NCA** and the (b) **dressed NCA** for an equilibrium symmetric dot with $U = -2\epsilon = 10\Gamma$. The phonon frequency is $\omega_0/\Gamma = 2.5$. The dashed lines indicate the center of the CT peaks as estimated by the energy renormalization at the anti-adiabatic limit $\omega_{CT}^{+}/\Gamma = -\epsilon + \frac{\lambda^2}{\omega_0}$ and $\omega_{CT}^{-}/\Gamma = \epsilon + 3\frac{\lambda^2}{\omega_0}$. All baths at the same inverse temperature $\beta = 10/\Gamma$.

a suppression of the charge transfer bands and a widening of the Kondo peak. The dressed NCA (panel (b)) shows an asymmetric shift of the CT peaks to approximately $\omega_+ = \frac{U}{2} + \frac{\lambda^2}{\omega_0}$ and $\omega_- = -\frac{U}{2} + 3\frac{\lambda^2}{\omega_0}$, as might be expected in the anti-adiabatic limit. Some deviation from this occurs, especially for the positive CT band. More interestingly, as the CT peak merges with the Kondo peak at $\lambda = \sqrt{\frac{U\omega_0}{6}}$, a strong enhancement occurs. This enhancement is not observed in the bare NCA.

(a) $\lambda/\Gamma = 1$



(b) $\lambda/\Gamma = 2$



(c)

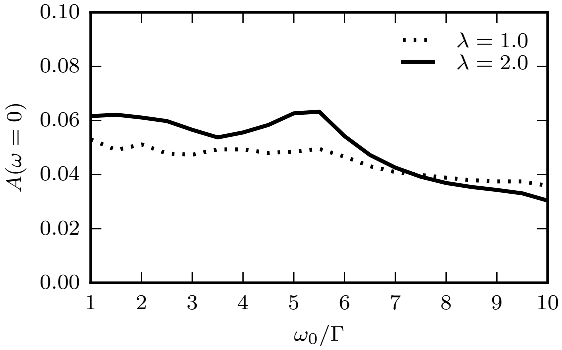


Figure 5.15: The conductance $G(V)$ as calculated within the **bare NCA** for different electron–phonon coupling (a)$\lambda/\Gamma = 1$ and (b)$\lambda/\Gamma = 2$ with a symmetrically applied bias $\mu_L = \mu_R = V$. The dot is also symmetric with $U = -2\epsilon = 10\Gamma$. Panel (c) shows the $\omega_0$-dependence of the central peak at $\omega = 0$. All baths at the same inverse temperature $\beta = 10/\Gamma$.

### Steady state conductance

Despite the symmetry breaking of the spectral function, the differential conductance $G(V) \equiv \frac{dI}{dV}(V)$ under a symmetrically applied bias ($\mu_L = -\mu_R = V/2$) remains a symmetric function of frequency even without the counter term. The replica effect and the non-monotonic enhancement, as visible in, *e.g.,* Fig. 5.13, appears in the spectral function, which could in principle be accessible in spectroscopic experiments. However, spectroscopic studies of single molecules in junctions and mesoscopic quantum dots are difficult to perform, and transport experiments are far more common. It is interesting to consider whether these effects are observable in the differential conductance as well as the spectral function; outside of linear response these quantities may differ qualitatively[121]. Fig. 5.15 shows the differential conductance as it varies under the effect of the phonon frequency $\omega_0$ at two different phonon coupling strengths $\lambda$. The non-monotonic enhancement remains clearly visible, while the side peaks are substantially weaker than their counterparts in the spectral function. The bare NCA therefore predicts that the non-monotonicity could be observed in transport experiments. Since it is related to the side bands merging with the charge transfer bands, an experimental observation of it could also be considered an indirect confirmation of the replica effect. We note that the dressed NCA also predicts a non-monotonicity, but one which does not appear related to the replica effect. It will take a more sophisticated theoretical treatment to determine whether this effect is real or an artifact of the two NCA approaches, and to understand more deeply the mechanism that lies behind it.

In Ref. 115, a non-monotonic effective Kondo temperature and zero-bias conductance has been predicted in the Anderson–Holstein model via the consideration of two limiting cases. In particular, for weak electron–phonon coupling $2\lambda^2/\omega_0 \ll U$, the low–energy excitations of the Anderson–Holstein model can be approximated by an isotropic Kondo Hamiltonian with the coupling to phonons leading to an increase in the effective Kondo temperature. On the other hand, for strong electron–phonon coupling $2\lambda^2/\omega_0 \gg U$, the low-energy excitations can be approximated by an anisotropic Kondo Hamiltonian in which the effective Kondo temperature decreases with increasing $\lambda$. This crossover behavior is observed in both NCAs, though the implied maximum in the spectral function occurs at a different $\lambda$ (see also Fig. 5.9c). Interestingly, when examining the spectral function at all energies simultaneously, a set of higher energy features which appear to be shifted replicas of the maximum is also revealed.

## 5.5 Conclusions

In this chapter we formulate and compare two distinct non-crossing approximations for the study of the Anderson–Holstein model. The first approximation, which we call the bare NCA, is a self-consistent resummation based on a self energy which contains the electron–phonon coupling and hybridization with the leads to lowest order. Within the second approximation, which we term the dressed NCA, a Lang–Firsov transformation is first applied, and the resulting transformed set of interactions are then included in a self-consistent, lowest order self energy. We focus on the predictions of both approximations with regard to transient dynamics as well as the non-equilibrium steady state behavior of the spectral function. In general, it should be expected that any flavor of NCA will be inaccurate for low–frequency properties. For example, NCA predicts a broadened and suppressed Kondo resonance when compared with exact numerics[122]. Due to the paucity of exact and global information related to the dynamical properties of the model, a detailed assessment of the success and failure of the respective methods is not possible even for higher frequency features. On the other hand, we believe it is plausible to favor the bare NCA when the electron–phonon coupling is weak, the dressed NCA when it is strong, and both approaches when they produce consistent results in the intermediate coupling regime. Since the two approximation are based on disparate limits of the electron–phonon portion of the problem, we focus on the intermediate coupling regime in an attempt to assess the validity of the two approximations.

We find that several features appear to be robust within both flavors of NCA. First, the Kondo peak is enhanced in particular regimes, but is universally suppressed in the large electron–phonon coupling regime. Second, low energy tunneling occurs and charge transfer peaks are suppressed when phonon frequency is small compared to other relevant energy scales. Lastly, the voltage splitting of the Kondo peak robustly occurs in the non-equilibrium regime. We expect these features to be real and experimentally reproducible behaviors in the Anderson–Holstein model.

Conversely, several striking dynamical properties appear only within one type of NCA approximation. In particular, the oscillatory transient behavior exhibited in Fig. 5.11 and the replication of the Kondo peak is only observed within the bare NCA, while polaronic shifts of the charge transfer peaks occur only in the dressed NCA approximation. It is important to note that these observations do not necessarily imply that such behaviors are artifacts. In particular, since the bare NCA is expected to capture accurately the weak electron–phonon situation, it is plausible that the features revealed in Fig. 5.9 and 5.11 are real properties of the model in this regime. The dressed NCA may not predict this behavior due to the fact that several low order diagrams associated with the interplay between hybridization and electron–phonon coupling are absent. On the other hand, polaronic effects may only be captured within the dressed NCA, and thus strong coupling shifts of the charge transfer peaks should be expected once the coupling to phonons is sizable.

Perhaps the most important aspect of the work presented here is that it lays the foundation for exact real-time QMC approaches based on expansion around the NCA approximation. These “bold-line” approaches have been successful in the treatment of the simpler Anderson model, and have enabled the simulation of relatively long real time information before the dynamical sign problem becomes problematic. Convergence of these approaches depends crucially on having a reasonably accurate partial summation of diagrams from the outset. With respect to the work presented here, we expect that the bare and dressed NCA approximations should provide a good starting point in the weak and strong electron–phonon coupling regimes, respectively. In addition to validating or falsifying the predictions made by the individual NCA approximations of this chapter, real-time QMC approaches that make use of the bare and dressed NCA techniques should allow for the exact simulation of the Anderson–Holstein model in regimes that are currently inaccessible.

# Appendix

## 5.A Comparison with DMFT-based Monte Carlo results

The top panel of Fig. 3 of Ref. 187 illustrates the behavior of the spectral function of an Anderson–Holstein problem computed via analytical continuation of *exact* imaginary-time quantum Monte Carlo as a function of increasing electron–phonon coupling, and is analogous to our Fig. 5.9. While it is difficult to make a direct comparison between these results and the results presented in our work due to the fact that the previous results were obtained self-consistently in the context of dynamical mean field theory, we have computed the spectral function for the same model and parameters within the NCA approaches outlined in this chapter. In this sense, the results of Fig. 16 represent a type of non-iterated NCA impurity solution in the DMFT context. The electron–phonon coupling parameters used in Fig. 3 of Ref. 187 are sufficiently large to render the bare NCA unstable. On the other hand, the dressed NCA is in qualitative agreement with the analytically continued results.

Quantitatively, the dressed NCA produces peaks in positions similar to those obtained by Monte Carlo for large $\lambda$, but the $\omega = 0$ and low frequency peaks are broadened and suppressed when compared to those of the analytically continued exact data. This broadening and suppression appears to be a general feature of NCA[122]. While the behavior of the gap closing feature can be observed in both the NCA and the analytically continued Monte Carlo data, it is still unclear to what degree the differences in the spectral functions are due to the effects of analytical continuation and the self-consistency of the DMFT calculation.

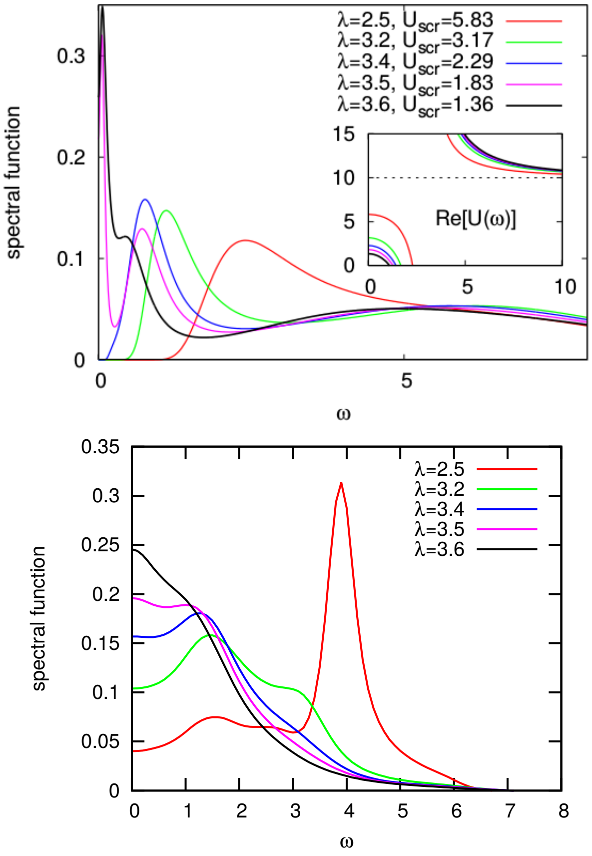


Figure 5.16: Upper: Evolution of the spectral function across the metal–insulator transition (gap closing) by increasing the phonon coupling. Lower: The spectral function $A(\omega)$ in the **strong** coupling regime is calculated within the dressed NCA for a symmetric dot with $U = -2\epsilon = 10\Gamma$ at equilibrium $V = 0$. The density of state is of the semi-circular form $\Gamma(\omega) = \sqrt{4t^2 - \omega^2}$ with $t = 1$. The phonon coupling is $\omega_0 = 3.0\Gamma$ and the counter term is symmetric ($\delta = 1$). The baths are maintained at a temperature $\beta\Gamma = 50$.

# Bibliography

[1] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*, en, Google-Books-ID: 0Yx5VzaMYm8C (Oxford University Press, 2002).

[2] A. Nitzan, *Chemical Dynamics in Condensed Phases: Relaxation, Transfer, and Reactions in Condensed Molecular Systems* (Oxford University Press, New York, 2006).

[3] U. Weiss, *Quantum dissipative systems*, Vol. 10 (World Scientific Publishing Company Incorporated, 1999).

[4] A. Garg, J. N. Onuchic, and V. Ambegaokar, J. Chem. Phys. **83**, 4491 (1985).

[5] Y. Georgievskii, C.-P. Hsu, and R. A. Marcus, J. Chem. Phys. **110**, 5307 (1999).

[6] J. Adolphs and T. Renger, Biophys. J. **91**, 2778 (2006).

[7] G. S. Engel, T. R. Calhoun, E. L. Read, T. K. Ahn, T. Mancal, Y. C. Cheng, R. E. Blankenship, and G. R. Fleming, Nature **446**, 782 (2007).

[8] H. Lee, Y. C. Cheng, and G. R. Fleming, Science **316**, 1462 (2007).

[9] G. Panitchayangkoon, D. Hayes, K. A. Fransted, J. R. Caram, E. Harel, J. Wen, R. E. Blankenship, and G. S. Engel, Proc. Natl. Acad. Sci. U. S. A. **107**, 12766 (2010).

[10] E Collini and G. D. Scholes, Science **323**, 369 (2009).

[11] J. L. Brédas and R. Silbey, Science **323**, 348 (2009).

[12] M. B. Smith and J. Michl, Chem. Rev. **110**, 6891 (2010).

[13] P. E. Teichen and J. D. Eaves, J. Phys. Chem. B **116**, 11473 (2012).

[14] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, J. Chem. Phys. **138**, 114102 (2013).

[15] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, J. Chem. Phys. **138**, 114103 (2013).

[16] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, J. Chem. Phys. **141**, – (2014).

[17] M. Thorwart and P. Hänggi, Phys. Rev. A **65**, 012309 (2001).

[18] T. A. Costi and R. H. McKenzie, Phys. Rev. A **68**, 034301 (2003).

[19] M. J. Storcz and F. K. Wilhelm, Phys. Rev. A **67**, 042319 (2003).

[20] P. W. Anderson, Phys. Rev. **124**, 41 (1961).

[21] T Holstein, Ann. Phys. (N. Y). **8**, 325 (1959).

[22] A. Hewson, *The Kondo Problem to Heavy Fermions - Academic and Professional Books - Cambridge University Press* (1993).

[23] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, 1976).

[24] M. Capone, M. Fabrizio, C. Castellani, and E. Tosatti, Rev. Mod. Phys. **81**, 943 (2009).

[25] M. Capone, C. Castellani, and M. Grilli, Adv. Condens. Matter Phys. **2010**, 920860 (2010).

[26] S. Dal Conte, C. Giannetti, G. Coslovich, F. Cilento, D. Bossini, T. Abebaw, F. Banfi, G. Ferrini, H. Eisaki, M. Greven, A. Damascelli, D. van der Marel, and F. Parmigiani, Science **335**, 1600 (2012).

[27] C. Gadermaier, A. S. Alexandrov, V. V. Kabanov, P. Kusar, T. Mertelj, X. Yao, C. Manzoni, D. Brida, G. Cerullo, and D. Mihailovic, Phys. Rev. Lett. **105**, 257001 (2010).

[28] C. Gadermaier, V. V. Kabanov, A. S. Alexandrov, L. Stojchevska, T. Mertelj, C. Manzoni, G. Cerullo, N. D. Zhigadlo, J. Karpinski, Y. Q. Cai, X. Yao, Y. Toda, M. Oda, S. Sugai, and D. Mihailovic, Phys. Rev. X **4**, 011056 (2014).

[29] L. Perfetti, P. A. Loukakos, M. Lisowski, U. Bovensiepen, H. Berger, S. Biermann, P. S. Cornaglia, A. Georges, and M. Wolf, Phys. Rev. Lett. **97**, 067402 (2006).

[30] L Perfetti, P. A. Loukakos, M Lisowski, U Bovensiepen, M Wolf, H Berger, S Biermann, and A Georges, New J. Phys. **10**, 053019 (2008).

[31] D Fausti, R. I. Tobey, N Dean, S Kaiser, A Dienst, M. C. Hoffmann, S Pyon, T Takayama, H Takagi, and A Cavalleri, Science **331**, 189 (2011).

[32] S Kaiser, S. R. Clark, D Nicoletti, G Cotugno, R. I. Tobey, N Dean, S Lupi, H Okamoto, T Hasegawa, D Jaksch, and A Cavalleri, en, Sci. Rep. **4**, 3823 (2014).

[33] S. V. Aradhya and L. Venkataraman, Nat. Nanotechnol. **8**, 399 (2013).

[34] A. Nitzan and M. A. Ratner, Science **300**, 1384 (2003).

[35] X. H. Qiu, G. V. Nazin, and W Ho, Science **299**, 542 (2003).

[36] N. Li, J. Ren, L. Wang, G. Zhang, P. Hänggi, and B. Li, Rev. Mod. Phys. **84**, 1045 (2012).

[37] Y. Dubi and M. Di Ventra, Rev. Mod. Phys. **83**, 131 (2011).

[38] C. Joachim and M. A. Ratner, Proc. Natl. Acad. Sci. U. S. A. **102**, 8801 (2005).

[39] Y. C. Chen, M. Zwolak, and M. Di Ventra, Nano Lett. **3**, 1691 (2003).

[40] G. D. Scott and D. Natelson, ACS Nano **4**, 3560 (2010).

[41] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[42] F. Bloch, Phys. Rev. **105**, 1206 (1957).

[43] A. Redfield, Adv. Magn. Opt. Reson. **1**, 1 (1965).

[44] J. M. Jean, R. A. Friesner, and G. R. Fleming, J. Chem. Phys., 5827 (1992).

[45] A. Ishizaki and G. R. Fleming, J. Chem. Phys. **130**, 234111 (2009).

[46] M. Yang and G. R. Fleming, Chem. Phys. **275**, 355 (2002).

[47] A. Montoya-Castillo, T. C. Berkelbach, and D. R. Reichman, J. Chem. Phys. **143**, 194108 (2015).

[48] A. J. Leggett, S. Chakravarty, A. T. A. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger, Rev. Mod. Phys. **59**, 1 (1987).

[49] A. Würger, Phys. Rev. Lett. **78**, 1759 (1997).

[50] M. Sparpaglione and S. Mukamel, J. Chem. Phys. **88**, 3263 (1988).

[51] M. Sparpaglione and S. Mukamel, J. Chem. Phys. **88**, 4300 (1988).

[52] D. G. Evans and R. D. Coalson, J. Chem. Phys. **102**, 5658 (1995).

[53] R. D. Coalson, D. G. Evans, and A. Nitzan, J. Chem. Phys. **101**, 436 (1994).

[54] A. A. Golosov and D. R. Reichman, J. Chem. Phys. **115**, 9862 (2001).

[55] A. A. Golosov and D. R. Reichman, J. Chem. Phys. **115**, 9848 (2001).

[56] W. M. Zhang, T. Meier, V. Chernyak, and S. Mukamel, J. Chem. Phys. **108**, 7763 (1998).

[57] A. Ishizaki and Y. Tanimura, Chem. Phys. **347**, 185 (2008).

[58] A. Ishizaki and G. R. Fleming, J. Chem. Phys. **130**, 234110 (2009).

[59] H. Wang, X. Sun, and W. H. Miller, J. Chem. Phys. **108**, 9726 (1998).

[60] X. Sun, H. Wang, and W. H. Miller, J. Chem. Phys. **109**, 7064 (1998).

[61] Q. Shi and E. Geva, J. Chem. Phys. **118**, 8173 (2003).

[62] R. B. Gerber, V. Buch, and M. A. Ratner, J. Chem. Phys. **77**, 3022 (1982).

[63] G. Stock, J. Chem. Phys. **103**, 1561 (1995).

[64] J. C. Tully, J. Chem. Phys. **55**, 562 (1971).

[65] J. C. Tully, J. Chem. Phys. **07974**, 1061 (1990).

[66] J. C. Tully, Faraday Discuss. **110**, 407 (1998).

[67] M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).

[68] S. R. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal, and S. Hammes-Schiffer, J. Chem. Phys. **114**, 6925 (2001).

[69] P. K. Agarwal, S. R. Billeter, and S. Hammes-Schiffer, J. Phys. Chem. B **106**, 3283 (2002).

[70] S. Hammes-Schiffer and J. C. Tully, J. Chem. Phys. **101**, 4657 (1994).

[71] S. Hammes-Schiffer and J. C. Tully, J. Chem. Phys. **103**, 8528 (1995).

[72] S. Y. Kim and S. Hammes-Schiffer, J. Chem. Phys. **124**, 244102 (2006).

[73] F. Webster, E. T. Wang, P. J. Rossky, and R. A. Friesner, J. Chem. Phys. **100**, 4835 (1994).

[74] B. J. Schwartz, E. R. Bittner, O. V. Prezhdo, and P. J. Rossky, J. Chem. Phys. **104**, 5942 (1996).

[75] K. F. Wong and P. J. Rossky, J. Chem. Phys. **116**, 8418 (2002).

[76] K. F. Wong and P. J. Rossky, J. Chem. Phys. **116**, 8429 (2002).

[77] M. J. Bedard-Hearn, R. E. Larsen, and B. J. Schwartz, J. Chem. Phys. **123**, 234106 (2005).

[78] R. E. Larsen, M. J. Bedard-Hearn, and B. J. Schwartz, J. Phys. Chem. B **110**, 20055 (2006).

[79] C. Zhu, S. Nangia, A. W. Jasper, and D. G. Truhlar, J. Chem. Phys. **121**, 7658 (2004).

[80] C. Zhu, A. W. Jasper, and D. G. Truhlar, J. Chem. Phys. **120**, 5543 (2004).

[81] A. W. Jasper and D. G. Truhlar, J. Chem. Phys. **123**, 64103 (2005).

[82] I. Horenko, C. Salzmann, B. Schmidt, and C. Schutte, J. Chem. Phys. **117**, 11075 (2002).

[83] N. Shenvi, J. E. Subotnik, and W. Yang, J. Chem. Phys. **134**, 144102 (2011).

[84] J. Y. Fang and S. Hammes-Schiffer, J. Chem. Phys. **110**, 11166 (1999).

[85] J. Y. Fang and S. Hammes-Schiffer, J. Phys. Chem. A **103**, 9399 (1999).

[86] O. Prezhdo and P. Rossky, Phys. Rev. Lett. **81**, 5294 (1998).

[87] H. Wang, M. Thoss, and W. H. Miller, J. Chem. Phys. **115**, 2979 (2001).

[88] H. Wang and M. Thoss, J. Chem. Phys. **119**, 1289 (2003).

[89] H. Wang and M. Thoss, J. Phys. Chem. A **107**, 2126 (2003).

[90] K. A. Velizhanin, H. Wang, and M. Thoss, Chem. Phys. Lett. **460**, 325 (2008).

[91] K. A. Velizhanin, M. Thoss, and H. Wang, J. Chem. Phys. **133**, 084503 (2010).

[92] H. Wang and M. Thoss, J. Phys. Chem. A **117**, 7431 (2013).

[93] Y Tanimura and R Kubo, J. Phys. Soc. Japan (1989).

[94] J. Strümpfer and K. Schulten, J. Chem. Theory Comput. **8**, 2808 (2012).

[95] D. E. Makarov and N. Makri, Chem. Phys. Lett. **221**, 482 (1994).

[96] N. Makri, J. Math. Phys. **36**, 2430 (1995).

[97] N. Makri and D. E. Makarov, J. Chem. Phys. **102**, 4600 (1995).

[98] N Makri, E Sim, D. E. Makarov, and M Topaler, Proc. Natl. Acad. Sci. U. S. A. **93**, 3926 (1996).

[99] R. Egger and U. Weiss, Zeitschrift für Phys. B Condens. Matter **89**, 97 (1992).

[100] R. Egger and C. H. Mak, Phys. Rev. B **50**, 210 (1994).

[101] R. Egger, L. Mühlbacher, and C. H. Mak, Phys. Rev. E **61**, 5961 (2000).

[102] C. H. Mak and R. Egger, *Monte Carlo Methods for Real-Time Path Integration*, Vol. XCIII (John Wiley & Sons, Inc., 2007), pp. 39–76.

[103] J. E. Hirsch, Phys. Rev. B **38**, 12023 (1988).

[104] L. Mühlbacher and R. Egger, J. Chem. Phys. **118**, 179 (2003).

[105] D. Segal, A. J. Millis, and D. R. Reichman, Phys. Rev. B **82**, 205323 (2010).

[106] G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, Phys. Rev. Lett. **115**, 266802 (2015).

[107] J. Paaske and K. Flensberg, Phys. Rev. Lett. **94**, 176801 (2005).

[108] A. Jovchev and F. B. Anders, Phys. Rev. B **87**, 195112 (2013).

[109] R. Seoane Souto, A. L. Yeyati, A. Martín-Rodero, and R. C. Monreal, Phys. Rev. B **89**, 085412 (2014).

[110] K. F. Albrecht, A. Martin-Rodero, R. C. Monreal, L. Mühlbacher, and A. Levy Yeyati, Phys. Rev. B **87**, 085127 (2013).

[111] K. F. Albrecht, A. Martin-Rodero, J. Schachenmayer, and L. Mühlbacher, Phys. Rev. B **91**, 064305 (2015).

[112] J. Gaudioso, L. J. Lauhon, and W. Ho, Phys. Rev. Lett. **85**, 1918 (2000).

[113] J. E. Han, Phys. Rev. B **81**, 113106 (2010).

[114] A. Hewson and D. Meyer, J. Phys. Condens. Matter **14**, 23 (2001).

[115] P. S. Cornaglia, H. Ness, and D. R. Grempel, Phys. Rev. Lett. **93**, 147201 (2004).

[116] P. S. Cornaglia, D. R. Grempel, and H. Ness, Phys. Rev. B **71**, 075320 (2005).

[117] P. S. Cornaglia, G. Usaj, and C. A. Balseiro, Phys. Rev. B **76**, 241403 (2007).

[118] E. Eidelstein, D. Goberman, and A. Schiller, Phys. Rev. B **87**, 075319 (2013).

[119] M. A. Laakso, D. M. Kennes, S. G. Jakobs, and V. Meden, en, New J. Phys. **16**, 023007 (2014).

[120] L. Arrachea and M. J. Rozenberg, Phys. Rev. B **72**, 41301 (2005).

[121] G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, Phys. Rev. Lett. **112**, 146802 (2014).

[122] G. Cohen, D. R. Reichman, A. J. Millis, and E. Gull, Phys. Rev. B **89**, 115139 (2014).

[123] J. König, H. Schoeller, and G. Schön, Phys. Rev. Lett. **76**, 1715 (1996).

[124] L. Mühlbacher and E. Rabani, Phys. Rev. Lett. **100**, 176403 (2008).

[125] M. Schiró and M. Fabrizio, Phys. Rev. B **79**, 153302 (2009).

[126] P. Werner and M. Eckstein, Phys. Rev. B **88**, 165108 (2013).

[127] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011).

[128] E. Gull, D. R. Reichman, and A. J. Millis, Phys. Rev. B **84**, 085134 (2011).

[129] E. Gull, D. R. Reichman, and A. J. Millis, Phys. Rev. B **82**, 075109 (2010).

[130] G. Cohen and E. Rabani, Phys. Rev. B **84**, 075150 (2011).

[131] G. Cohen, E. Y. Wilner, and E. Rabani, New Journal of Physics **15**, 073018 (2013).

[132] G. Cohen, E. Gull, D. R. Reichman, A. J. Millis, and E. Rabani, Phys. Rev. B **87**, 195108 (2013).

[133] H Grabert, *Projection Operator Techniques in Nonequillibrium Statistical Mechanics*, Vol. 95, Springer Tracts in Modern Physics (Springer-Verlag, 1982).

[134] A. A. Golosov and D. R. Reichman, Chem. Phys. **296**, 129 (2004).

[135] M. G. Mavros and T. Van Voorhis, J. Chem. Phys. **141**, 054112 (2014).

[136] Q. Shi and E. Geva, J. Chem. Phys. **119**, 12063 (2003).

[137] Q. Shi and E. Geva, J. Chem. Phys. **120**, 10647 (2004).

[138] A. Montoya-Castillo and D. R. Reichman, J. Chem. Phys. **144**, 184104 (2016).

[139] J. L. Basdevant, Fortschritte der Phys. **20**, 283 (1972).

[140] Z. Gong, Z. Tang, S. Mukamel, J. Cao, and J. Wu, J. Chem. Phys. **142**, 084103 (2015).

[141] R Zwanzig, Phys. Rev. **124**, 983 (1961).

[142] G. B. Arfken, *Mathematical methods for physicists* (Academic press, 2013).

[143] A Yonemoto, T Hisakado, and K Okumura, IEE Proceedings- Circuits, Devices Syst. **150**, 399 (2003).

[144] M. Thoss, H. Wang, and W. H. Miller, J. Chem. Phys. **115**, 2991 (2001).

[145] J. Berntsen, T. O. Espelid, and A. Genz, ACM Trans. Math. Softw. **17**, 437 (1991).

[146] G Honig and U Hirdes, J. Comput. Appl. Math. **10**, 113 (1984).

[147] J. C. Tully, J. Chem. Phys. **137**, 22A301 (2012).

[148] W. H. Miller, *Classical limit quantum mechanics and the theory of molecular collisions*, Vol. 25 (1974), pp. 69–177.

[149] M. Barbatti, Wiley Interdiscip. Rev. Comput. Mol. Sci. **1**, 620 (2011).

[150] C.-Y. Hsieh and R. Kapral, en, Entropy **16**, 200 (2013).

[151] U. Müller and G. Stock, J. Chem. Phys. **107**, 6230 (1997).

[152] M. S. Topaler, T. C. Allison, D. W. Schwenke, and D. G. Truhlar, J. Phys. Chem. A **102**, 1666 (1998).

[153] M. D. Hack, A. M. Wensmann, D. G. Truhlar, M. Ben-Nun, and T. J. Martinez, J. Chem. Phys. **115**, 1172 (2001).

[154] J. E. Subotnik and N. Shenvi, J. Chem. Phys. **134**, 244114 (2011).

[155] R. Kapral and G. Ciccotti, J. Chem. Phys. **110**, 8919 (1999).

[156] J. E. Subotnik, W. Ouyang, and B. R. Landry, J. Chem. Phys. **139**, 214107 (2013).

[157] B. R. Landry and J. E. Subotnik, J. Chem. Phys. **135**, 1 (2011).

[158] J. E. Subotnik, J. Phys. Chem. A **115**, 12083 (2011).

[159] R. Jiang and E. L. Sibert, J. Chem. Phys. **136**, 224104 (2012).

[160] L. Wang and D. Beljonne, J. Chem. Phys. **139**, 064316 (2013).

[161] C. A. Schwerdtfeger, A. V. Soudackov, and S. Hammes-Schiffer, J. Chem. Phys. **140**, 034113 (2014).

[162] M. J. Falk, B. R. Landry, and J. E. Subotnik, J. Phys. Chem. B **118**, 8108 (2014).

[163] A. Jain, M. F. Herman, W. Ouyang, and J. E. Subotnik, J. Chem. Phys. **143**, 134106 (2015).

[164] D. Mac Kernan, G. Ciccotti, and R. Kapral, J. Chem. Phys. **116**, 2346 (2002).

[165] A. W. Jasper, C. Zhu, S. Nangia, and D. G. Truhlar, en, Faraday Discuss. **127**, 1 (2004).

[166] N. Rekik, C.-Y. Hsieh, H. Freedman, and G. Hanna, J. Chem. Phys. **138**, 144106 (2013).

[167] V. May and O. Kuhn, *Charge and energy transfer dynamics in molecular systems* (John Wiley & Sons, Inc., 2008).

[168] B. R. Landry, M. J. Falk, and J. E. Subotnik, J. Chem. Phys. **139**, 211101 (2013).

[169] E. R. Bittner and P. J. Rossky, J. Chem. Phys. **103**, 8130 (1995).

[170] O. Prezhdo and P. Rossky, J. Chem. Phys. **107**, 5863 (1997).

[171] B. R. Landry and J. E. Subotnik, J. Chem. Phys. **137**, 0 (2012).

[172] J. Skinner and D Hsu, J. Phys. Chem. **47**, 4931 (1986).

[173] J. W. Negele and H. Orland, *Quantum many-particle systems*, Vol. 200 (Addison-Wesley New York, 1988).

[174] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).

[175] S. Trotzky, L. Pollet, F. Gerbier, U. Schnorrberger, I. Bloch, N. V. Prokof'ev, B. Svistunov, and M. Troyer, en, Nature Physics **6**, 998 (2010).

[176] V. A. Kashurnikov, N. V. Prokof'ev, and B. V. Svistunov, Physical Review A **66**, 031601 (2002).

[177] N. Prokof'ev and B. Svistunov, Physical Review Letters **87**, 160601 (2001).

[178] C. H. Mak and D. Chandler, Physical Review A **41**, 5709 (1990).

[179] E. Burovski, N. Prokof'ev, B. Svistunov, and M. Troyer, Physical Review Letters **96**, 160402 (2006).

[180] J. E. Hirsch, R. L. Sugar, D. J. Scalapino, and R. Blankenbecler, Physical Review B **26**, 5033 (1982).

[181] S. Rombouts, K. Heyde, and N. Jachowicz, Phys. Lett. A **242**, 271 (1998).

[182] M. Schiró, Phys. Rev. B **81**, 085126 (2010).

[183] P. Werner, A. Comanac, L. De' Medici, M. Troyer, A. J. Millis, L. De Medici, M. Troyer, and A. J. Millis, Phys. Rev. Lett. **97**, 076405 (2006).

[184] P. Werner and A. J. Millis, Phys. Rev. B **74**, 155107 (2006).

[185] E. Gull, P. Werner, O. Parcollet, and M. Troyer, en, EPL (Europhysics Letters) **82**, 57003 (2008).

[186] P. Werner, T. Oka, and A. J. Millis, Phys. Rev. B **79**, 035320 (2009).

[187] P. Werner and A. J. Millis, Phys. Rev. Lett. **104**, 146401 (2010).

[188] R. A. Kuharski, J. S. Bader, D. Chandler, M. Sprik, M. L. Klein, and R. W. Impey, The Journal of chemical physics **89**, 3248 (1988).

[189] J. S. Bader, R. A. Kuharski, and D. Chandler, The Journal of Chemical Physics **93**, 230 (1990).

[190] A. Warshel, Z. T. Chu, and W. W. Parson, en, Science **246**, 112 (1989).

[191] A. Warshel and J.-K. Hwang, The Journal of chemical physics **84**, 4938 (1986).

[192] N. Makri and K. Thompson, Chemical Physics Letters **291**, 101 (1998).

[193] M. P. A. Fisher and W. Zwerger, Physical Review B **32**, 6190 (1985).

[194] N. V. Prokof'ev and P. C. E. Stamp, J. Low Temp. Phys. **104**, 143 (1996).

[195] N. V. Prokof'ev and P. C. E. Stamp, Physical review letters **80**, 5794 (1998).

[196] M. Goldstein, M. H. Devoret, M. Houzet, and L. I. Glazman, Phys. Rev. Lett. **110**, 017002 (2013).

[197] J. M. Fink, M. Göppl, M. Baur, R. Bianchetti, P. J. Leek, A. Blais, and A. Wallraff, en, Nature **454**, 315 (2008).

[198] E. Solano, G. S. Agarwal, and H. Walther, Phys. Rev. Lett. **90**, 027903 (2003).

[199] J. Casanova, G. Romero, I. Lizuain, J. J. García-Ripoll, and E. Solano, Phys. Rev. Lett. **105**, 263603 (2010).

[200] N. Van Kampen, Physica **74**, 239 (1974).

[201] B Yoon, J. Deutch, and J. H. Freed, J. Chem. Phys. **62**, 4687 (1975).

[202] S. Mukamel, Chem. Phys. **37**, 33 (1979).

[203] D. R. Reichman, F. L. H. Brown, and P. Neu, Phys. Rev. E **55**, 2328 (1997).

[204] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[205] W. K. Hastings, en, Biometrika **57**, 97 (1970).

[206] A. E. Antipov, Q. Dong, and E. Gull, Phys. Rev. Lett. **116**, 036801 (2016).

[207] R. E. Profumo, C. Groth, L. Messio, O. Parcollet, and X. Waintal, Phys. Rev. B **91**, 245154 (2015).

[208] J. Hu, R.-X. Xu, and Y. Yan, J. Chem. Phys. **133**, 101106 (2010).

[209] J. Hu, M. Luo, F. Jiang, R.-X. Xu, and Y. Yan, J. Chem. Phys. **134**, 244106 (2011).

[210] Y. Yan, J. Chem. Phys. **140**, 054105 (2014).

[211] J. M. Moix and J. Cao, J. Chem. Phys. **139**, 134106 (2013).

[212] Z. Tang, X. Ouyang, Z. Gong, H. Wang, and J. Wu, J. Chem. Phys. **143**, 224112 (2015).

[213] H. Wang and M. Thoss, The Journal of chemical physics **138**, 134704 (2013).

[214] D. Segal and A. Nitzan, Phys. Rev. Lett. **94**, 034301 (2005).

[215] L. Nicolin and D. Segal, J. Chem. Phys. **135**, 164106 (2011).

[216] K. Saito and T. Kato, Phys. Rev. Lett. **111**, 214301 (2013).

[217] N. A. Zimbovskaya and M. R. Pederson, Phys. Rep. **509**, 1 (2011).

[218] L. H. Yu, Z. K. Keane, J. W. Ciszek, L. Cheng, M. P. Stewart, J. M. Tour, and D. Natelson, Phys. Rev. Lett. **93**, 266802 (2004).

[219] D. Rakhmilevitch, R. Korytár, A. Bagrets, F. Evers, and O. Tal, Phys. Rev. Lett. **113**, 236603 (2014).

[220] P. Werner and A. J. Millis, Phys. Rev. Lett. **99**, 146404 (2007).

[221] D. Golež, M. Eckstein, and P. Werner, Phys. Rev. B **92**, 195123 (2015).

[222] A. Mitra, I. Aleiner, and A. J. Millis, Phys. Rev. B **69**, 245302 (2004).

[223] A. Mitra, I. Aleiner, and A. J. Millis, Phys. Rev. Lett. **94**, 076404 (2005).

[224] R. Härtle, C. Benesch, and M. Thoss, Phys. Rev. Lett. **102**, 146801 (2009).

[225] M. G. Schultz and F. von Oppen, Phys. Rev. B **80**, 033302 (2009).

[226] M. Esposito and M. Galperin, Phys. Rev. B **79**, 205303 (2009).

[227] M. Esposito and M. Galperin, J. Phys. Chem. C **114**, 20362 (2010).

[228] W. Dou, A. Nitzan, and J. E. Subotnik, J. Chem. Phys. **142**, 084110 (2015).

[229] S. Tikhodeev, M. Natario, K. Makoshi, T. Mii, and H. Ueba, Surf. Sci. **493**, 63 (2001).

[230] T. Mii, S. Tikhodeev, and H. Ueba, Surf. Sci. **502-503**, 26 (2002).

[231] M. Galperin, M. A. Ratner, and A. Nitzan, J. Chem. Phys. **121**, 11965 (2004).

[232] A. Ueda and M. Eto, Phys. Rev. B **73**, 235353 (2006).

[233] L. K. Dash, H. Ness, and R. W. Godby, J. Chem. Phys. **132**, 104113 (2010).

[234] L. K. Dash, H. Ness, and R. W. Godby, Phys. Rev. B **84**, 085433 (2011).

[235] B. Dong, G. H. Ding, and X. L. Lei, Phys. Rev. B **88**, 075414 (2013).

[236] M. Galperin, A. Nitzan, and M. A. Ratner, Phys. Rev. B **73**, 45314 (2006).

[237] M. Galperin, A. Nitzan, and M. A. Ratner, Phys. Rev. B **76**, 035301 (2007).

[238] R. C. Monreal, F. Flores, and A. Martin-Rodero, Phys. Rev. B **82**, 235412 (2010).

[239] S. Sayyad and M. Eckstein, Phys. Rev. B **91**, 104301 (2015).

[240] A. Martin-Rodero, A. Levy Yeyati, F. Flores, and R. C. Monreal, Phys. Rev. B **78**, 235112 (2008).

[241] J. Klatt, L. Mühlbacher, and A. Komnik, Phys. Rev. B **91**, 155306 (2015).

[242] R. Hützen, S. Weiss, M. Thorwart, and R. Egger, Phys. Rev. B **85**, 121408 (2012).

[243] L. Simine and D. Segal, J. Chem. Phys. **138**, 214111 (2013).

[244] L. Simine and D. Segal, J. Chem. Phys. **141**, 014704 (2014).

[245] V. V. Albert, Phys. Rev. Lett. **108**, 180401 (2012).

[246] K. F. Albrecht, H. Soller, L. Mühlbacher, and A. Komnik, Phys. E **54**, 15 (2013).

[247] O. Entin-Wohlman, A. Aharony, and Y. Meir, Phys. Rev. B **71**, 035333 (2005).

[248] J. Koch and F. Von Oppen, Phys. Rev. Lett. **94**, 206804 (2005).

[249] J. Koch, M. Semmelhack, F. Von Oppen, and A. Nitzan, Phys. Rev. B **73**, 155306 (2006).

[250] R. C. Monreal and A. Martin-Rodero, Phys. Rev. B **79**, 115140 (2009).

[251] S Maekawa, J Ko, G Scho, J Martinek, Y Utsumi, H Imamura, and J Barnas, Phys. B **91**, 20 (2003).

[252] A. Goker and B. Uyanik, Phys. Lett. A **376**, 2735 (2012).

[253] P. Roura-Bas, L. Tosi, and A. A. Aligia, Phys. Rev. B **87**, 195136 (2013).

[254] R. Härtle and M. Thoss, Phys. Rev. B **83**, 115414 (2011).

[255] E. Y. Wilner, H. Wang, G. Cohen, M. Thoss, and E. Rabani, Phys. Rev. B **88**, 045137 (2013).

[256] E. Y. Wilner, H. Wang, M. Thoss, and E. Rabani, Phys. Rev. B **89**, 205129 (2014).

[257] N. E. Bickers, Rev. Mod. Phys. **59**, 845 (1987).

[258] T. Pruschke and N. Grewe, Zeitschrift für Phys. B Condens. Matter **74**, 439 (1989).

[259] N. S. Wingreen and Y. Meir, Phys. Rev. B **49**, 40 (1994).

[260] K. Haule, S. Kirchner, J. Kroha, and P. Wölfle, Phys. Rev. B **64**, 155111 (2001).

[261] M. Eckstein and P. Werner, Phys. Rev. B **82**, 115115 (2010).

[262] A. J. White and M. Galperin, en, Phys. Chem. Chem. Phys. **14**, 13809 (2012).

[263] E. Lebanon and A. Schiller, Phys. Rev. B **65**, 035308 (2001).

[264] Q. F. Sun and H. Guo, Phys. Rev. B **64**, 153306 (2001).

[265] Y. Meir, N. S. Wingreen, and P. A. Lee, Phys. Rev. Lett. **70**, 2601 (1993).