12 Introduction to Diagrammatic Approaches

Hartmut Hafermann

Mathematical and Algorithmic Sciences Lab,

Paris Research Center

Huawei Technologies France SASU

92100 Boulogne-Billancourt, France

Contents

1   Introduction

Correlated materials are at the heart of modern solid state physics. These strongly interacting many-body quantum systems can display such diverse phenomena as the Mott transition [1], non-Fermi-liquid and heavy-Fermion behavior [2], and high-temperature superconductivity [3].

At the same time they pose one of most difficult theoretical challenges. The reason is that they are in the middle of two extremes. In solids with broad energy bands electrons can largely avoid each other while moving through the crystal. The properties of such materials are successfully described in a picture of nearly-independent electrons moving freely in an effective potential generated by all other electrons as described by Bloch waves. In such a momentum-space description the interaction is a perturbation. The framework of density-functional theory (DFT) provides a very successful and often even quantitative description of the microscopic properties of such materials, including simple metals like aluminum, or semiconductors and band-insulating materials. It may even be used to design materials with desired properties.

When electrons cannot avoid each other, the process where an electron hops to a neighboring occupied orbital may be energetically so unfavorable that electrons will remain localized. Mott showed that this scenario can be understood in a real-space picture. These Mott insulators are erroneously predicted to be metallic by band theories, with a classic example being NiO.

Strongly correlated materials are in the middle of these two regimes and exhibit a delicate balance between the kinetic and potential energy. Their name emphasizes the fact that the motion of electrons in these materials is correlated in the true sense of the word: they do not move independently. Neither a real-space nor a momentum-space description appears appropriate – a situation which may be as hard to grasp as particle-wave duality.

Materials that fall under this category often have open d- or f-electron shells. Examples are the transition metals V, Fe, Co, Ni, Cu, and their oxides. For example, CuO_2 planes determine the properties of the high-temperature superconductors. Small microscopic changes can have drastic macroscopic effects. These materials are typically very sensitive to externally controllable parameters, such as pressure and doping, and are therefore promising candidates for applications. The development of reliable theoretical tools to calculate their material specific properties therefore remains one of the primary concerns of modern theoretical condensed matter physics.

An exact solution of the many-particle problem is clearly out of reach and we have to resort to approximate methods. Naturally, certain limiting cases are the most accessible. In the opposite regimes of weak and strong interaction, diagrammatic perturbation theory allows us to make quantitative predictions. Importantly, it also provides us with the intuition for the underlying microscopic processes. We will discuss the foundations of diagrammatic perturbation theory below. In correlated materials, however, where the kinetic- is of the order of the potential energy, such a perturbative description necessarily breaks down. In this regime dynamical mean-field theory (DMFT) has provided important insights [4]. It becomes exact in both the localized and noninteracting itinerant regimes and in the limit of infinite dimensions or coordination number [5].

Correlated materials typically are not close to any of these limits. While DMFT describes the important strong local temporal correlations, as a mean-field theory it completely neglects spatial correlations. This shortcoming is well illustrated by the fact that DMFT correctly predicts the temperature dependence of the local moment in the transition metals iron and nickel, but overestimates their Curie temperatures. Phenomena in which the electrons are qualitatively affected by the presence of extended collective excitations cannot be described within DMFT. Because of the complexity of the problem, there is not a single method which gives 'the best' answer. Instead, we are aiming to study the problem using different approximate methods. These will have inherent limitations and exhibit limited parameter ranges of applicability. We may be even unaware of the precise nature of these limitations. It may therefore not be possible to distinguish true physics from artifacts of the method, so that we risk drawing the wrong conclusions. The approach of the theoretical physicist will therefore be to apply different methods to obtain complementary viewpoints and, step by step, draw a complete picture of the underlying physics.

In this lecture, we will concentrate on the so-called diagrammatic extensions of DMFT [6]. These form a relatively recent class of methods which come in different flavors. They nevertheless all build on the common idea to formulate a diagrammatic perturbation theory around DMFT as the starting point, with the main goal to describe the effect of spatial correlations.

In the following we first describe the Hubbard model, which is believed to capture essential properties of correlated materials, and briefly introduce the Green function as a probe for its properties. We then formulate the basic equations of dynamical mean-field theory and sketch the mathematical foundations of diagrammatic perturbation theory. After a brief introduction to diagrammatic extensions of DMFT we develop a diagrammatic perturbation theory for correlated systems by the example of dual fermions. We conclude this chapter with the discussion of some illustrative results and an outlook on further developments.

1.1   Hubbard model

The Hubbard model describes electrons hopping from site to site on a lattice as depicted in Fig. 1, with a probability that is determined by the hopping-integrals t. Whenever two electrons occupy the same site, their mutual repulsion incurs an energy penalty U. Because of its simplicity it is probably the most widely studied model in this context. For concreteness we consider the two-dimensional single-band model whose Hamiltonian is given by

$$H = -t \sum_{\langle ij \rangle \sigma} c_{i\sigma}^{\dagger} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}. \tag{1}$$

Here the indices label the sites on the lattice and the first sum in the kinetic term is over all pairs of sites or bonds.

In the absence of interaction, we can simply diagonalize the kinetic energy term in reciprocal space: $\sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma}$. The second, potential energy term is apparently diagonal in the real-space occupation number basis. In the opposite regimes of large and small interaction, we can either view the kinetic or potential energy as a perturbation and formulate a diagrammatic
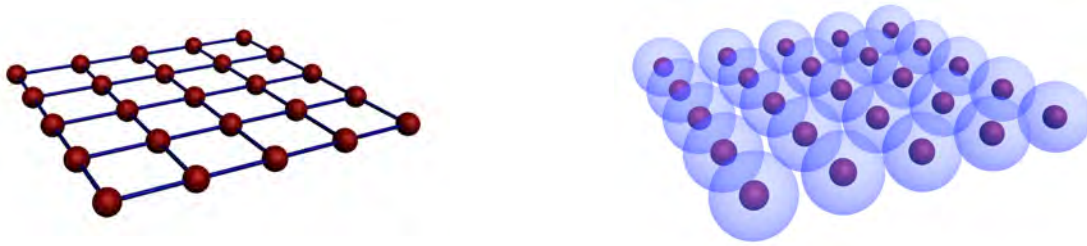
Fig. 1: *Left: Illustration of the Hubbard model showing lattice sites on which electrons interact with Hubbard interaction U (red) connected by bonds through the hopping integrals t. Right: DMFT picture of model showing a collection of decoupled sites exerted to an electronic bath, as indicated by the blue spheres.*

perturbation theory. We will however primarily be interested in the case where the interaction becomes comparable to the bandwidth $U \sim W = 8t$, where such a perturbative description necessarily breaks down.

A complete solution would provide us with the entire spectrum of the Hamiltonian, which would completely specify the equilibrium thermodynamics of the model and include information even on many-particle correlations. A much simpler object is the single-particle Green function. The Matsubara Green function

$$G_{\alpha_1 \alpha_2}(\tau_1 - \tau_2) = -\frac{1}{Z} \text{Tr} \left(e^{-\beta(\hat{H} - \mu \hat{N})} T_\tau c_{\alpha_1}(\tau_1) c_{\alpha_2}^\dagger(\tau_2) \right), \tag{2}$$

where β denotes inverse temperature, completely specifies the thermal single-particle properties. It can be viewed as a probe of the system: A particle inserted into the system in a certain state $\alpha_2 \equiv \{i\sigma\}$ at time τ_2 evolves until we remove it in state α_1 at a later time τ_1. The Green function is the thermal average of this process and tells us how electrons propagate in the system. For this reason, it is sometimes called a propagator. The bosonic (fermionic) Matsubara Green function is a 2β-(anti-)periodic function of imaginary time. As a result, its Fourier representation is a function of odd $\nu_n = (2n + 1)\pi/\beta$ (fermionic) or even $\nu_n = 2m\pi/\beta$ (bosonic) discrete Matsubara frequencies. This simplifies calculations significantly. The retarded Green function is obtained from the Matsubara Green function through analytical continuation $\omega_n \to \omega + i0^+$, which allows to compute the density of states. The latter contains a wealth of information and, for example, tells us whether a system is metallic or insulating.

1.2 Dynamical mean-field theory

DMFT has been an important step towards the understanding of correlated electrons in solids. One of the challenges is to reconcile two vastly different energy scales to describe the redistribution of spectral weight through the interaction. DMFT for the first time allowed a simultaneous description of both the long-lived coherent quasiparticles and the short-lived incoherent high-energy excitations. The former give rise to the 'quasiparticle peak' and the latter lead to broad Hubbard bands in the electronic spectrum.

DMFT fully accounts for the local, time-dependent quantum fluctuations. We can understand this by viewing it as the quantum analog of classical mean-field theories, which provides an intuitive understanding of the approach. The construction of an approximation, in analogy to the classical case, results in a nontrivial mean-field theory. The mean-field can no longer be represented by a single number, but rather by a time or frequency-dependent field, hence the name 'dynamical'.

The electronic self-energy may be expressed as a functional of the Green function. The central approximation of DMFT is to assume that this functional is purely local and a functional of the local Green function only: $\Sigma = \Sigma[G^{\text{loc}}]$. Under this assumption, we may write the DMFT lattice Green function in the form

$$G_{\mathbf{k}\nu}^{\text{DMFT}} = \frac{1}{\imath\nu + \mu - \varepsilon_{\mathbf{k}} - \Sigma[G^{\text{loc}}]} \, . \tag{3}$$

The right-hand side is a functional of the local Green function $G^{\text{loc}} = \frac{1}{N}\sum_{\mathbf{k}} G_{\mathbf{k}\nu}^{\text{DMFT}}$. This is a complicated self-consistent problem. The unknown local Green function determines the self-energy, which in turn fixes the local Green function.

Even if we knew G^{loc}, we still had to sum *all* diagrams for the local self-energy. As shown by Georges and Kotliar [7], we can introduce a local effective quantum impurity model as a tool to accomplish this. Often an Anderson impurity model is used whose action reads

$$S_{\text{imp}}[c^*, c] = -\sum_{\nu\sigma} c_{\nu\sigma}^* \big(\imath\nu + \mu - \Delta_\nu\big)c_{\nu\sigma} + U\sum_{\omega} n_{\omega\uparrow}n_{-\omega\downarrow}\,,$$

where ν are fermionic and ω are bosonic Matsubara frequencies. Here we have introduced the hybridization function Δ_ν, which plays the role of a frequency-dependent Weiss field. In this description the lattice no longer enters explicitly. Instead we can picture the situation as a given lattice site immersed in a structureless 'bath' of conduction electrons. We can imagine that such a description becomes more and more accurate when the coordination number grows large. Electrons can hop from the bath onto the impurity and back. Because of the mutual repulsion when two electrons occupy the same site, the dynamics will strongly depend on the time spent on the impurity. The Weiss field and Green function will therefore still have a non-trivial energy dependence. As we can see in the illustration in Fig. 1, however, different lattice sites are completely decoupled in the DMFT description. The motion of electrons in different parts of the lattice is no longer correlated. We say that DMFT neglects spatial correlations.

By solving the impurity model for a given hybridization we obtain the impurity self-energy and Green function, which are related by a form of Dyson's equation,

$$g_\nu = \frac{1}{\imath\nu + \mu - \Delta_\nu - \Sigma[g]}. \tag{4}$$

We could attempt to do this perturbatively. As for the lattice, this will, however, not work in the strongly correlated regime we are interested in. Since the problem is local, it is nevertheless simpler than that of the lattice. Fortunately, sophisticated and accurate methods exist to solve it efficiently and in particular to compute the impurity Green function $g_\nu := -\langle c_{\nu\sigma}c_{\nu\sigma}^*\rangle_{\text{imp}}$. Notable examples are continuous-time quantum Monte Carlo methods [8].

The hybridization function encodes the dynamics generated by all electrons on average and is a priori unknown. Instead of finding the local lattice Green function, the problem has been reduced to determining the hybridization self-consistently.

If we identify the impurity self-energy with that of the lattice, the hybridization will determine the right-hand side of (3). If we additionally identify the local lattice Green function with the impurity Green function, we obtain the following equation that implicitly determines the hybridization function

$$g_\nu = \frac{1}{N} \sum_{\mathbf{k}} \frac{1}{g_\nu^{-1} + \Delta_\nu - \varepsilon_{\mathbf{k}}} \, . \tag{5}$$

This equation is the only place where the structure of the lattice enters in DMFT, namely in the form of the dispersion $\varepsilon_{\mathbf{k}}$.

In practice, we obtain the self-consistent solution by repeatedly solving the impurity model. Starting from an initial guess for the hybridization, we can iteratively update it based on the solution of the impurity model according to

$$\Delta_\nu \leftarrow \Delta_\nu + \xi\big(g_\nu^{-1} - (G_\nu^{\mathrm{loc}})^{-1}\big), \tag{6}$$

where G^{loc} is calculated from (3) using the impurity self-energy. A solution is found once the impurity Green function equals the local lattice Green function: $g_\nu = G_\nu^{\mathrm{loc}}$. $\xi \in \,]0, 1[$ is a parameter to control the convergence speed. The loop typically converges in a few iterations.

2   Diagrammatic perturbation theory

Perturbation theory can be applied when the problem at hand is close, in some sense, to a solvable reference problem. This implies that it can be described in terms of a small perturbation of this reference. Often the reference problem is the non-interacting one and the perturbation series is a series expansion in powers of the interaction. Wick's theorem allows to express the result in terms of the interaction and products of the known non-interacting Green function. If the coupling, U in case of the Hubbard model, is small, the series will converge to the exact result within any desired accuracy after a finite number of terms. If this is not the case, one may attempt to obtain an approximation by summing an infinite partial series of dominant terms.

The expressions in the expansion become, however, increasingly difficult to handle with increasing order as we will see in a few examples below. In addition the number of terms grows factorially. It helps to visualize the expressions in terms of Feynman diagrams, where one assigns a symbol to the interaction and depicts non-interacting Green functions as lines. This yields a diagrammatic perturbation theory. The usefulness of diagrammatic perturbation theory hinges on the following simplifications.

The linked cluster theorem significantly simplifies the problem, because it allows one to concentrate on connected diagrams. Another simplification is obtained by gathering diagrams with the same structure which yield the same contribution. Finally, one can formulate a set of rules

which allows one to obtain the analytical expression corresponding to any diagram just from its structure, without the need to consider the term in the expansion where it came from.

The diagrammatic perturbation theory is often developed in terms of expectation values of time-ordered second-quantized operators. The theory can equally well be formulated in terms of path integrals. In case of fermions these are coherent state path integrals. Instead of operators we integrate over Grassmann variables to capture the fermion statistics. The expressions are formally very similar and it is often straightforward to translate from one description to the other. The dual fermion approach and the associated perturbation theory are formulated in terms of coherent state path integrals. We therefore briefly introduce them and formulate the perturbation theory in terms of them. An excellent introduction to coherent states and the coherent state path integral can be found in [9].

2.1   Coherent state path integrals

For concreteness, we consider systems of fermions. Starting point for the description of the thermodynamic properties of quantum many-particle systems is the grand canonical partition function $Z = \mathrm{Tr}\, e^{-\beta(\hat{H} - \mu \hat{N})}$, where $\hat{H}$ is the Hamiltonian in second quantized form, $\hat{N}$ is the operator of the total particle number and the chemical potential μ controls the number of particles of the grand canonical ensemble.

Let us first recall the properties of coherent states. Coherent states $|\phi\rangle$ are eigenstates of the annihilation operator c_α with eigenvalue ϕ_α: $c_\alpha |\phi\rangle = \phi_\alpha |\phi\rangle$. For fermions, the eigenvalues are Grassmann numbers. The most important property of the Grassmann algebra is that its generators anticommute: $\phi_\alpha \phi_\beta + \phi_\beta \phi_\alpha = 0$. This implies in particular that $\phi_\alpha^2 = 0$. It is straightforward to show that the state

$$|\phi\rangle = e^{-\sum_\alpha \phi_\alpha c_\alpha^\dagger} |0\rangle = \prod_\alpha (1 - \phi_\alpha c_\alpha^\dagger) |0\rangle \tag{7}$$

has the desired property. The adjoint of the coherent state is $\langle 0| \prod_\alpha (1 + \phi_\alpha^* c_\alpha)$. The overlap of two coherent states follows straightforwardly:

$$\langle \phi | \phi' \rangle = \langle 0| \prod_\alpha (1 + \phi_\alpha^* c_\alpha) \prod_{\alpha'} (1 - \phi_{\alpha'} c_{\alpha'}^\dagger) |0\rangle = \prod_\alpha (1 + \phi_\alpha^* \phi_\alpha) = e^{\sum_\alpha \phi_\alpha^* \phi_\alpha}. \tag{8}$$

The matrix element of a normal-ordered operator $A[c_\alpha^\dagger, c_\alpha]$ therefore is given by

$$\langle \phi | A[c_\alpha^\dagger, c_\alpha] |\phi\rangle = \langle \phi | \phi \rangle \, A[\phi_\alpha^*, \phi_\alpha] = e^{\sum_\alpha \phi_\alpha^* \phi_\alpha} A[\phi_\alpha^*, \phi_\alpha]. \tag{9}$$

The following closure relation is essential for the formulation of the fermionic coherent state path integral:

$$\int \prod_\alpha d\phi_\alpha^* d\phi_\alpha \, e^{-\sum_\alpha \phi_\alpha^* \phi_\alpha} |\phi\rangle \langle \phi| = 1. \tag{10}$$

The coherent states form an over-complete set of states in a generalized Fock space, that is, the set of linear combinations of states in the Fermion Fock space with coefficients in the Grassmann algebra. Any physical Fermion state can be expanded in terms of them.

Given two states $|\psi\rangle$ and $|\phi\rangle$ in the Fock space, it follows from the anticommutation relations of Grassmann numbers that $\langle\psi_i|\phi\rangle\langle\phi|\psi_j\rangle = \langle-\phi|\psi_j\rangle\langle\psi_i|\phi\rangle$. For a complete set of states $\{|n\rangle\}$ in the Fock space, the trace of a second quantized operator A can be written

$$\text{Tr}\, A = \sum_n \langle n|\, A\, |n\rangle = \int \prod_\alpha d\phi_\alpha^* d\phi_\alpha\, e^{-\sum_\alpha \phi_\alpha^* \phi_\alpha} \sum_n \langle n|\phi\rangle\langle\phi|\, A\, |n\rangle$$

$$= \int \prod_\alpha d\phi_\alpha^* d\phi_\alpha\, e^{-\sum_\alpha \phi_\alpha^* \phi_\alpha} \langle-\phi|\, A \sum_n |n\rangle\langle n|\phi\rangle$$

$$= \int \prod_\alpha d\phi_\alpha^* d\phi_\alpha\, e^{-\sum_\alpha \phi_\alpha^* \phi_\alpha} \langle-\phi|\, A\, |\phi\rangle\,. \tag{11}$$

With these prerequisites, the grand canonical partition function can be expressed in the form

$$Z = \text{Tr}\, e^{-\beta(\hat{H}-\mu\hat{N})} = \int \prod_\alpha d\phi_\alpha^* d\phi_\alpha\, e^{-\sum_\alpha \phi_\alpha^* \phi_\alpha} \langle-\phi|\, e^{-\beta(\hat{H}-\mu\hat{N})}\, |\phi\rangle\,, \tag{12}$$

where for fermions, the trace imposes antiperiodic boundary conditions.

We can obtain a coherent state path integral representation of the partition function by viewing the exponential as an imaginary-time evolution operator describing the evolution of the state from time zero to β. While its matrix elements cannot be evaluated directly, we can exploit the fact that the infinitesimal evolution operator can be obtained in normal-ordered form. To this end, we break the time interval $[0,\beta]$ into M time steps of size $\epsilon = \beta/M$, such that $e^{-\beta(\hat{H}-\mu\hat{N})} = \left(e^{-\epsilon(\hat{H}-\mu\hat{N})}\right)^M$.

The second quantized operator $e^{-\epsilon(\hat{H}-\mu\hat{N})}$ is in approximate normal-ordered form, up to a correction of order ϵ^2, $e^{-\epsilon(\hat{H}-\mu\hat{N})} = :e^{-\epsilon(\hat{H}-\mu\hat{N})}: +\mathcal{O}(\epsilon^2)$. Using (9), to evaluate the matrix elements, we can write the partition function in the limit $\epsilon \to 0$ in the form

$$Z = \int \prod_{k=1}^M \prod_\alpha d\phi_{\alpha,k}^* d\phi_{\alpha,k}\, e^{-\sum_{k=1}^M \sum_\alpha \phi_{\alpha,k}^* \phi_{\alpha,k}} \prod_{k=1}^M \langle-\phi_{\alpha,k}|\, :e^{-\epsilon(\hat{H}-\mu\hat{N})}: +\mathcal{O}(\epsilon^2)\, |\phi_{\alpha,k-1}\rangle$$

$$= \int \prod_{k=1}^M \prod_\alpha d\phi_{\alpha,k}^* d\phi_{\alpha,k}\, e^{-\sum_{k=1}^M \sum_\alpha \left(\phi_{\alpha,k}^* \phi_{\alpha,k} - \phi_{\alpha,k}^* \phi_{\alpha,k-1}\right) - \epsilon \sum_{k=1}^M \sum_\alpha \left(H[\phi_{\alpha,k}^*,\phi_{\alpha,k-1}] - \mu\phi_{\alpha,k}^* \phi_{\alpha,k-1}\right)}$$

$$= \int \prod_{k=1}^M \prod_\alpha d\phi_{\alpha,k}^* d\phi_{\alpha,k}\, e^{-S[\phi_{\alpha,k}^*,\phi_{\alpha,k-1}]}\,, \tag{13}$$

where $\phi_{\alpha,0} = -\phi_{\alpha,M}$ because of antiperiodic boundary conditions and we have defined the action

$$S[\phi_\alpha^*,\phi_\alpha] = \epsilon\left(\sum_{k=1}^M \sum_\alpha \phi_{\alpha,k}^* \frac{\phi_{\alpha,k} - \phi_{\alpha,k-1}}{\epsilon} - \mu\phi_{\alpha,k}^* \phi_{\alpha,k-1} + H[\phi_{\alpha,k}^*,\phi_{\alpha,k-1}]\right)\,. \tag{14}$$

In the limit $\epsilon \to 0$, it is convenient to introduce the notation

$$\phi_{\alpha,k}^* \frac{\phi_{\alpha,k} - \phi_{\alpha,k-1}}{\epsilon} \equiv \phi_\alpha^*(\tau)\frac{\partial}{\partial\tau}\phi_\alpha(\tau), \qquad H[\phi_{\alpha,k}^*,\phi_{\alpha,k-1}] \equiv H[\phi_\alpha^*(\tau),\phi_\alpha(\tau)]. \tag{15}$$

Splitting the Hamiltonian into a one-body operator and an interaction, $H = \sum_\alpha \varepsilon_\alpha \phi_\alpha^* \phi_\alpha + V[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]$, the path integral can symbolically be written in trajectory notation as

$$Z = \int\limits_{\phi_\alpha(\beta) = -\phi_\alpha(0)} \mathcal{D}[\phi_\alpha^*(\tau), \phi_\alpha(\tau)] \, e^{-\int_0^\beta d\tau \left(\sum_\alpha \phi_\alpha^*(\tau)(\frac{\partial}{\partial \tau} + \varepsilon_\alpha - \mu)\phi_\alpha(\tau) + V[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]\right)} . \qquad (16)$$

Formally, the problem has been reduced to quadrature. We have to integrate over all possible realizations of paths $\phi_\alpha(\tau)$, $\phi_\alpha^*(\tau)$, weighted by the exponential of the action $S[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]$. In general, it amounts to computing high-dimensional integrals, so that Monte Carlo methods are particularly suitable for this task [9,8]. It must be kept in mind that even though the notation suggests it, it neither implies continuity nor differentiability of the paths and all quantities are defined in terms of the discrete expressions (13) and (14).

2.2   Perturbation theory

In the following, we will develop a perturbation theory based on the path integral formulation of the partition function and the Green function (2),

$$G_{\alpha_1 \alpha_2}(\tau_1 - \tau_2) = -\frac{1}{Z} \int \mathcal{D}[\phi_\alpha^*(\tau), \phi_\alpha(\tau)] \, e^{-S[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]} \phi_{\alpha_1}(\tau_1) \phi_{\alpha_2}^*(\tau_2). \qquad (17)$$

Given that the fermionic path integral always implies antiperiodic trajectories, we omit the indication here and in what follows. Furthermore we need not explicitly indicate the time-ordering, remembering that it is implicit in the construction of the path integral.

We can view Green function as a thermal average over the interacting system, as symbolized by the following notation: $G_{\alpha_1 \alpha_2}(\tau_1 - \tau_2) = -\langle \phi_{\alpha_1}(\tau_1) \phi_{\alpha_1}^*(\tau_2) \rangle$. We readily obtain a perturbation expansion of the single-particle Green function in powers of the interaction

$$G_{\alpha_1 \alpha_2}(\tau_1 - \tau_2) = -\frac{Z_0}{Z} \left\langle e^{-\int_0^\beta d\tau V[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]} \phi_{\alpha_1}(\tau_1) \phi_{\alpha_2}^*(\tau_2) \right\rangle_0 \qquad (18)$$

$$= -\frac{Z_0}{Z} \sum_{n=0}^\infty \frac{(-1)^n}{n!} \int_0^\beta d\tau_1' \ldots d\tau_n' \left\langle V[\phi_\alpha^*(\tau_1'), \phi_\alpha(\tau_1')] \ldots V[\phi_\alpha^*(\tau_n'), \phi_\alpha(\tau_n')] \phi_{\alpha_1}(\tau_{\alpha_1}) \phi_{\alpha_2}^*(\tau_{\alpha_1}) \right\rangle_0 .$$

The partition function of the non-interacting system, Z_0, is obtained by setting the interacting $V = 0$ in (16). It remains to evaluate the non-interacting average over a product of Grassmann fields arising in above expression. This is accomplished using Wick's theorem, which allows to express it in terms of a product of non-interacting Green functions. Wicks's theorem is usually formulated in terms of second-quantized operators. In its familiar form it states that the non-interacting expectation value of a time-ordered product of operators can be written as the sum over all complete contractions. A contraction of two time-dependent operators is defined as a symbolic pairing of these operators, which evaluates to their non-interacting expectation value. In a complete contraction of a set of an even number of fields, each operator is paired with exactly one other. Having defined thermal expectation values in terms of Grassmann fields, we can equally well write a contraction in terms of them

$$\underline{\phi_{\alpha_1}(\tau_1) \phi_{\alpha_2}^*}(\tau_2) := \left\langle \phi_{\alpha_1}(\tau_1) \phi_{\alpha_2}^*(\tau_2) \right\rangle_0 = -G_{\alpha_1 \alpha_2}^0(\tau_1 - \tau_2) . \qquad (19)$$

In terms of Grassmann numbers, Wick's theorem is based on the following Gaussian integral

$$\frac{\int \mathcal{D}[\phi^*, \phi]\, \phi_{i_1} \phi_{i_2} \dots \phi_{i_n} \phi_{j_n}^* \dots \phi_{j_2}^* \phi_{j_1}^* e^{-\sum_{ij} \phi_i^* M_{ij} \phi_j}}{\int \mathcal{D}[\phi^*, \phi]\, e^{-\sum_{ij} \phi_i^* M_{ij} \phi_j}} = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) M_{i_{\sigma(n)}, j_n}^{-1} \dots M_{i_{\sigma(1)}, j_1}^{-1}. \quad (20)$$

For simplicity we write M_{ij}^{-1} for the elements of the inverse of M. The left-hand side has the form of an expectation value over a product of fields. On the right-hand side we have a sum over the elements of the permutation group S_n, which is the Leibniz formula for the determinant of the inverse of M.

Before relating this expression to the familiar statement of Wick's theorem, we prove it by means of the following generating function

$$G(J^*, J) := \frac{\int \mathcal{D}[\phi^*, \phi]\, e^{-\sum_{ij} \phi_i^* M_{ij} \phi_j + \sum_i \phi_i^* J_i + J_i^* \phi_i}}{\int \mathcal{D}[\phi^*, \phi]\, e^{-\sum_{ij} \phi_i^* M_{ij} \phi_j}} = e^{J_i^* M_{ij}^{-1} J_j}. \quad (21)$$

The name generating function will become apparent below. Here the sources J^*, J are Grassmann numbers and M is a complex matrix with elements M_{ij}. We evaluate the following derivatives

$$\frac{\delta^{2n} G}{\delta J_{i_1}^* \dots \delta J_{i_n}^* \delta J_{j_n} \dots \delta J_{j_1}}\bigg|_{J = J^* = 0} = (-1)^n \frac{\int \mathcal{D}[\phi^*, \phi]\, \phi_{i_1} \dots \phi_{i_n} \phi_{j_n}^* \dots \phi_{j_1}^* e^{-\sum_{ij} \phi_i^* M_{ij} \phi_j}}{\int \mathcal{D}[\phi^*, \phi]\, e^{-\sum_{ij} \phi_i^* M_{ij} \phi_j}}. \quad (22)$$

Here we have used that the derivatives anticommute with Grassmann numbers. Applying the same derivatives to the right-hand side of (21) yields

$$\frac{\delta^{2n}}{\delta J_{i_1}^* \dots \delta J_{i_n}^* \delta J_{j_n} \dots \delta J_{j_1}} \left(e^{\sum_{ij} J_i^* M_{ij} J_j} \right)\bigg|_{J = J^* = 0}$$

$$= (-1)^n \frac{\delta^n}{\delta J_{i_1}^* \dots \delta J_{i_n}^*} \left(\sum_{k_n} J_{k_n}^* M_{k_n, j_n}^{-1} \right) \dots \left(\sum_{k_1} J_{k_1}^* M_{k_1, j_1}^{-1} \right) \left(e^{\sum_{ij} J_i^* M_{ij} J_j} \right)\bigg|_{J = J^* = 0}$$

$$= (-1)^n \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) M_{i_{\sigma(n)}, j_n}^{-1} \dots M_{i_{\sigma(1)}, j_1}^{-1}, \quad (23)$$

which proves the identity (20). It remains to prove the identity for the generating function, Eq. (21). This is readily accomplished by imposing the linear shift transformation whose Jacobian is unity (summation over repeated indices implied),

$$\phi_i^* \to \phi_i^* + J_j^* M_{ji}^{-1}, \qquad \phi_i \to \phi_i + M_{ij}^{-1} J_j, \quad (24)$$

upon which the exponential on the left-hand side of (21) transforms to

$$\exp\left(-\phi_i^* M_{ij} \phi_j + \phi_i^* J_i + J_i^* \phi_i \right) \quad \to \quad \exp\left(-\phi_i^* M_{ij} \phi_j + J_i^* M_{ij}^{-1} J_j \right). \quad (25)$$

The term containing the sources J, J^* can be taken out of the integral. The numerator is seen to cancel the denominator, hence proving the identity.

A first application of Wick's theorem is the evaluation of the noninteracting Green function. By inserting the discrete matrix $M_{ij} = -(\partial_\tau + \varepsilon_\alpha - \mu)_{ij}^{-1}$ defined through (14) into Eq. (20) and replacing ϕ_j by $\phi_{\alpha,k}$ in the matrix elements of the path integral, we obtain

$$G_{\alpha_1\alpha_2}^0(\tau_1 - \tau_2) = -\frac{\int \mathcal{D}[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]\, e^{-\int_0^\beta d\tau \sum_\alpha \phi_\alpha^*(\tau)(\frac{\partial}{\partial\tau} + \varepsilon_\alpha - \mu)\phi_\alpha(\tau)}\phi_{\alpha_1}(\tau_1), \phi_{\alpha_2}^*(\tau_2)}{\int \mathcal{D}[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]\, e^{-\int_0^\beta d\tau \sum_\alpha \phi_\alpha^*(\tau)(\frac{\partial}{\partial\tau} + \epsilon_\alpha - \mu)\phi_\alpha(\tau)}}$$

$$= -(\partial_\tau + \varepsilon_\alpha - \mu)_{\alpha_1\tau_1;\alpha_2\tau_2}^{-1} = G_{\alpha_1}^0(\tau_1 - \tau_2)\, \delta_{\alpha_1\alpha_2}. \tag{26}$$

This shows that the non-interacting Green function equals the inverse of the matrix $-(\frac{\partial}{\partial\tau} + \epsilon_\alpha - \mu)$ and allows us to recast the action in the following form

$$S[\phi_\alpha^*(\tau), \phi_\alpha(\tau)] = \int\limits_0^\beta d\tau \left(-\sum_\alpha \phi_\alpha^*(\tau)\, G_\alpha^{0\,-1}(\tau_{\alpha_1} - \tau_{\alpha_2})\, \phi_\alpha(\tau) + V[\phi_\alpha^*(\tau), \phi_\alpha(\tau)] \right). \tag{27}$$

Similarly, with $M_{ij} = -(\partial_\tau + \epsilon_\alpha - \mu)_{ij}^{-1}$, the left-hand side of Eq. (20) equals the non-interacting average of fields. The right-hand side evaluates to a sum over products of non-interacting Green functions, or, by means of Eq. (19), to the sum over all complete contractions. To see this, we note that the sign of the permutation apparently equals the sign of the permutation that brings the fields in each contraction next to each other in the desired order.

With Wick's theorem at hand, we can evaluate the non-interacting expectation values in each term of the perturbation expansion for the Green function. This allows us to obtain successively more accurate approximations. The result is expressed in terms of the coupling and the known non-interacting Green function. As will be shown below, we can depict these elements with symbols and draw a diagram for each of the resulting expressions to obtain a diagrammatic perturbation theory.

The right-hand side of Eq. (20) contains a sum over all permutations of indices, so that the number of terms grows factorially with order. The construction of perturbation expansions in practice would be hopeless without two major simplifications.

The first simplification is based on the observation that different contractions at a given order can lead to structurally identical diagrams, which give the same contribution to the overall result. Instead of enumerating all of them, their multiplicity can be accounted for in terms of combinatorial factors. Another major simplification is given by the linked-cluster theorem: Any diagram generated by the perturbation expansion (18) can be decomposed into two parts: one is a connected part with an incoming and an outgoing line which stems from the contractions involving $\phi_{\alpha_1}(\tau_1)$ and $\phi_{\alpha_2}(\tau_2)$. The second part is a (not necessarily connected) part consisting of diagrams without external lines contributing to the vacuum amplitude. It is straightforward to verify that the perturbation expansion of the partition function,

$$\frac{Z}{Z_0} = \left\langle e^{-\int_0^\beta d\tau\, V[\phi_\alpha^*(\tau), \phi_\alpha(\tau)]} \right\rangle_0, \tag{28}$$

generates all the disconnected vacuum fluctuation graphs that appear in the expansion of the Green function. The linked-cluster theorem states that the logarithm of the above expectation

value, or $\ln Z/Z_0$, yields the sum over all connected graphs. An important consequence is that this contribution exactly cancels the factor $(Z/Z_0)^{-1}$ in the expansion of Green function, Eq. (18). As a consequence, when evaluating the perturbation expansion of G, we only have to take into account the fully connected diagrams with external lines.

3   Diagrammatic extensions of dynamical mean-field theory

We have seen that diagrammatic perturbation theory is a powerful tool to construct approximations starting from a solvable reference system. The problem we are facing when dealing with strongly correlated systems is that neither the non-interacting, nor the localized limit are appropriate starting points. DMFT, on the other hand, already includes the presumably dominant strong local temporal correlations. It would hence be desirable to include the effects of the presumably weaker spatial correlations beyond DMFT perturbatively.

It turns out that there are different ways to systematically construct a perturbation expansion around DMFT. It is crucial, however, that the DMFT reference system is solvable. Here and in contrast to DMFT solvable means that not only the single-particle Green function, but at least the two-particle and in principle all many-particle Green functions are assumed to be known. Fortunately, reasonable approximations can be constructed from the knowledge of the single- and two-particle Green function only. The need to compute higher-order correlation functions of an impurity model explains why the first works of this kind appeared only around ten years ago. To a large extent these developments were driven by the advent of continuous-time quantum impurity solvers, which allow the efficient and accurate computation of higher-order correlation functions [8].

The first approaches along these lines were the dynamical vertex approximation (DΓA) [10, 11] and the dual fermion approach [12]. A number of different approaches followed [6]. While the construction and the justification of necessary approximation differs for the various approaches, the underlying principle is the same: the propagators and the interaction vertices of the diagrammatic perturbation theory are obtained from the numerical solution of the impurity model. Non-local processes are described in terms of renormalized propagation between sites and a local frequency-dependent interaction between particles. Care has to be taken in the choice of propagators and vertex functions to avoid double counting of contributions that are already contained in DMFT.

The advantage of such methods is that the diagrammatic expressions are relatively inexpensive to evaluate numerically, so that it becomes possible to treat truly long-ranged correlations. The results, however, will be approximate on any scale. On the other hand, the diagrammatic approach provides intuition about the underlying microscopic origin of the observed effects. We will discuss some examples in the result section below.
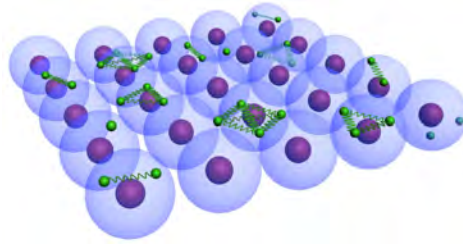
Fig. 2: *Illustration of the dual fermion approach. Spatial correlations neglected in the DMFT description of Fig. 1 are mediated through dual fermions (green). The dual fermions interact locally via n-particle interactions.*

3.1 Dual fermions

In the following we show how to construct a diagrammatic perturbation theory around DMFT, based on the idea of dual fermions. We then study the perturbation expansion in detail and formulate the rules of the diagrammatic perturbation theory.

As in conventional perturbation theory, we separate the problem into a solvable reference system and a perturbation. The underlying idea is to treat the strong local correlations on the level of the single-site impurity model and to embrace the presumably weaker coupling between the sites perturbatively. A diagrammatic extension of DMFT is obtained by setting the hybridization function equal to its DMFT value.

In the Grassmann path integral formalism, the Hubbard model is described by the action

$$S[c^*, c] = -\sum_{\nu\sigma} c^*_{\nu\sigma}\big(i\nu + \mu - \varepsilon_{\mathbf{k}}\big)c_{\nu\sigma} + U\sum_{\omega} n_{\omega\uparrow}n_{-\omega\downarrow},$$

We can formally add and subtract an arbitrary hybridization function at each lattice site. The lattice action can then be expressed in terms of the impurity action as follows

$$S_{\text{latt}}[c^*, c] = \sum_{i} S_{\text{imp}}[c^*_i, c_i] - \sum_{\mathbf{k}\nu\sigma} c^*_{\mathbf{k}\nu\sigma}\big(\Delta_\nu - \varepsilon_{\mathbf{k}}\big)\, c_{\mathbf{k}\nu\sigma}\,. \qquad (29)$$

In this form, the lattice problem can be viewed as a collection of impurity models at each lattice site, each with their own electronic bath. The sites are spatially coupled by a hybridization- and dispersion-dependent term. In principle it is possible to develop a perturbation theory around the impurity model (and around DMFT for a correspondingly chosen hybridization function) directly from this starting point: We could change the basis to transform the second term to a sum over sites instead of momenta and expand the path integral $Z = \int \mathcal{D}[c^*, c]\, \exp(-S_{\text{latt}})$ in that second term. The integration over the fields c^*, c could formally be performed and would give rise to averages of the form $\big\langle c^*_{i_1} c_{j_1} \ldots c^*_{i_n} c_{j_n} \big\rangle_{\text{imp}}$. There is clearly no Wick theorem in this case, but we could nevertheless compute at least some of these impurity correlation functions numerically and evaluate the perturbation series. It turns out that the dual fermion approximation is a clever resummation of diagrams appearing in this approach [13].

To derive it, we first take a somewhat different route. Here we decouple the impurity models by introducing new fermionic fields f, f^*, the dual fermions. This is achieved by means of

a Hubbard-Stratonovich transformation. Formally this transformation is a Gaussian integral which we can obtain by generalizing (21) using the substitutions

$$\phi_i^* \to f_i^*, \qquad \phi_j \to f_j, \qquad J_i^* \to c_j^* b_{ji}, \qquad J_j \to b_{ij} c_j, \tag{30}$$

which yields

$$\frac{\int \mathcal{D}[f^*, f]\, e^{-\sum_{ij} f_i^* M_{ij} f_j + \sum_{ij} f_i^* b_{ij} c_j + c_i^* b_{ij} f_j}}{\int \mathcal{D}[f^*, f]\, e^{-\sum_{ij} f_i^* M_{ij} f_j}} = e^{\sum_{ijkl} c_i^* b_{ij} M_{jk}^{-1} b_{kl} c_l}. \tag{31}$$

Letting $M_{ij}^{-1} = \Delta_\nu - \varepsilon_{\mathbf{k}}$ and setting the coupling between the physical and dual fermions to the local quantity $b_{ij} = -g_{ij}^{-1}$, we obtain the partition function in the form

$$Z = D_f \int \mathcal{D}[f^*, f]\, e^{-\sum_{\mathbf{k}\nu\sigma} f_{\mathbf{k}\nu\sigma}^* g_{\nu\sigma}^{-1} (\Delta_{\nu\sigma} - \varepsilon_{\mathbf{k}})^{-1} g_{\nu\sigma}^{-1} f_{\mathbf{k}\nu\sigma}} \int \mathcal{D}[c^*, c]\, e^{-\sum_i (S_{\mathrm{imp}}[c_i^*, c_i] + S_{\mathrm{cf}}[c_i^*, c_i; f_i^*, f_i])}. \tag{32}$$

Here D_f is a determinant which arises from the numerator in Eq. (31) and which will be irrelevant for the calculation of expectation values. The term S_{cf} is the local coupling between dual and physical fermions

$$S_{\mathrm{cf}}[c^*, c; f^*, f] = \sum_{\nu\sigma} \left(f_{\nu\sigma}^* g_{\nu\sigma}^{-1} c_{\nu\sigma} + c_{\nu\sigma}^* g_{\nu\sigma}^{-1} f_{\nu\sigma} \right).$$

The goal is to arrive at a representation which depends solely on dual variables. We can formally perform the second integral by expanding the exponential in powers of S_{cf}. We can do this for each site i separately. For a translationally invariant system, the result will be site-independent

$$\int \mathcal{D}[c_i^*, c_i] e^{-S_{\mathrm{imp}}[c_i^*, c_i]} e^{-S_{\mathrm{cf}}[c_i^*, c_i; f_i^*, f_i]} = \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} \frac{(-1)^n}{n!} \left\langle \left(\sum_{\nu\sigma} \left(f_{\nu\sigma}^* g_{\nu\sigma}^{-1} c_{\nu\sigma} + c_{\nu\sigma}^* g_{\nu\sigma}^{-1} f_{\nu\sigma} \right) \right)^n \right\rangle_{\mathrm{imp}}.$$

Only even orders contribute since the Gaussian integral over an odd number of c and c^* must vanish due to particle conservation. Let us exemplify this step for the first non-vanishing term. We have

$$\frac{1}{2} \sum_{\nu\sigma} \sum_{\nu'\sigma'} \left\langle f_{\nu\sigma}^* g_{\nu\sigma}^{-1} c_{\nu\sigma} c_{\nu'\sigma'}^* g_{\nu'\sigma'}^{-1} f_{\nu'\sigma'} + c_{\nu\sigma}^* g_{\nu\sigma}^{-1} f_{\nu\sigma} f_{\nu'\sigma'}^* g_{\nu'\sigma'}^{-1} c_{\nu'\sigma'} \right\rangle_{\mathrm{imp}}$$

$$= \sum_{\nu\sigma} \sum_{\nu'\sigma'} g_{\nu\sigma}^{-1} g_{\nu'\sigma'}^{-1} \left\langle c_{\nu\sigma} c_{\nu'\sigma'}^* \right\rangle_{\mathrm{imp}} f_{\nu\sigma}^* f_{\nu'\sigma'} = -\sum_{\nu\sigma} f_{\nu\sigma}^* g_{\nu\sigma}^{-1} f_{\nu\sigma}, \tag{33}$$

where in the first line we have used the anticommutation relations to bring the two terms into the same form. In the second line we have used the definition of the Green function for the impurity, $\langle c_{\nu\sigma} c_{\nu'\sigma'}^* \rangle_{\mathrm{imp}} = -g_{\nu\sigma} \delta_{\nu\nu'} \delta_{\sigma\sigma'}$, which is diagonal in spin and frequency.

We can apply the same procedure to the higher order terms. For example, the fourth-order term involves averages of the form

$$g_{\nu\nu'\omega}^{\sigma\sigma\sigma'\sigma'} := \left\langle c_{\nu\sigma} c_{\nu+\omega,\sigma}^* c_{\nu'+\omega,\sigma'} c_{\nu'\sigma'}^* \right\rangle_{\mathrm{imp}}. \tag{34}$$

The result of this expansion can be cast into the following form

$$\ln\langle e^{-S_{\mathrm{cf}}[c_i^*, c_i; f_i^*, f_i]}\rangle_{\mathrm{imp}} = -\sum_{\nu\sigma} f_{\nu\sigma}^* \, g_{\nu\sigma}^{-1} \, f_{\nu\sigma} - \tilde{V}[f^*, f]. \tag{35}$$

As one might expect from the linked-cluster theorem, the left hand side generates the *connected* correlation functions of the impurity model coupled to dual variables. To leading order, the resulting dual interaction is given by

$$\tilde{V}[f^*, f] = -\frac{1}{4} \sum_{\nu\nu'\omega\,\sigma_i} \gamma_{\nu\nu'\omega}^{\sigma_1\sigma_2\sigma_3\sigma_4} f_{\nu\sigma_1}^* f_{\nu+\omega,\sigma_2} f_{\nu'+\omega,\sigma_3}^* f_{\nu'\sigma_4} + \cdots, \tag{36}$$

where γ is the reducible two-particle vertex of the impurity model

$$\gamma_{\nu\nu'\omega}^{\sigma\sigma\sigma'\sigma'} := \frac{g_{\nu\nu'\omega}^{\sigma\sigma\sigma'\sigma'} - \beta\, g_{\nu\sigma}\, g_{\nu'\sigma'}\, \delta_\omega + \beta\, g_{\nu\sigma}\, g_{\nu+\omega\sigma}\, \delta_{\nu\nu'}\, \delta_{\sigma\sigma'}}{g_{\nu\sigma}\, g_{\nu+\omega,\sigma}\, g_{\nu'+\omega\sigma'}\, g_{\nu'\sigma'}}. \tag{37}$$

The higher-order terms contain the three-particle (six-leg) and higher-order vertices describing the interaction between a successively larger number of particles. In terms of the original Hubbard interaction, the vertices correspond to the sum of many high-order processes. Combining Eqs. (32) and (35), we see that the action in dual variables is given by

$$\tilde{S}[f^*, f] = -\sum_{\mathbf{k}\nu\sigma} f_{\mathbf{k}\nu\sigma}^* \, \tilde{G}_{\mathbf{k}\nu\sigma}^{0\,-1} \, f_{\mathbf{k}\nu\sigma} + \tilde{V}[f^*, f] \tag{38}$$

and the dual Green function is identified from the bilinear terms in the same two equations as

$$\tilde{G}_{\mathbf{k}\nu\sigma}^0 = \left(g_{\nu\sigma}^{-1} + (\Delta_{\nu\sigma} - \varepsilon_{\mathbf{k}})\right)^{-1} - g_{\nu\sigma}. \tag{39}$$

We have reformulated the problem in terms of a Green function and an interaction which can be computed given the solution of the impurity model. The interaction, however, is rather complicated and frequency-dependent. So what have we gained? To see this, we first establish a connection to DMFT by noting that by means of (3), the bare dual Green function can be written in the form $G_{\mathbf{k}\nu\sigma}^{\mathrm{DMFT}} - g_{\nu\sigma}$. While the hybridization function is arbitrary by construction, see Eq. (29), we can fix it through the self-consistency condition $\sum_{\mathbf{k}} \tilde{G}_{\mathbf{k}\nu\sigma}^0 = 0$. It apparently corresponds to the DMFT self-consistency condition discussed in Sec. 1.2.

At this point we have used the bare dual Green function and have not taken the dual interaction into account. DMFT corresponds to the case of non-interacting dual fermions. This is not surprising, since in the DMFT description sites are decoupled. It means that even the lowest-order corrections based on the action (38) will already introduce corrections beyond DMFT.

Once diagrams are taken into account, the hybridization function can be fixed using a similar condition on the interacting dual Green function, $\sum_{\mathbf{k}} \tilde{G}_{\mathbf{k}\nu\sigma} = 0$. This corresponds to the summation of an infinite partial series: all diagrams with a local propagator are eliminated from the expansion. This property is unique to the dual fermion approach and highlights the fact that the dual Green function is not a physical, but rather an auxiliary quantity. Since we have used an exact transformation to introduce the associated field, we can nevertheless establish exact relations between dual and physical quantities. In particular, the physical self-energy is given in terms of the dual self-energy $\tilde{\Sigma}$ by

$$\Sigma_{\mathbf{k}\nu\sigma} = \Sigma_{\nu\sigma}^{\mathrm{imp}} + \frac{\tilde{\Sigma}_{\mathbf{k}\nu\sigma}}{1 + \tilde{\Sigma}_{\mathbf{k}\nu\sigma} g_{\nu\sigma}}. \tag{40}$$

3.2   Perturbation theory for the dual propagator

We are now in the position to construct the diagrammatic perturbation theory for the dual propagator and the dual self-energy. To this end, we explicitly expand the path integral in powers of the dual interaction and apply Wick's theorem to evaluate the resulting expressions. We will show a couple of example diagrams to illustrate the procedure. We then describe the particularities of the dual perturbation theory, namely the determination of combinatorial prefactors of the diagrams and how to determine the sign of resulting expressions. Finally we arrive at the general Feynman rules for evaluating the expression of any diagram appearing in the diagrammatic perturbation theory.

To have a more condensed notation, we gather frequency-, spin and other possible indices into a single Greek index. To emphasize that the diagrams describe non-local corrections, we write the positions explicitly and use Latin indices for them. To further clearly mark external lines of a diagram, we use the combined index $1 \equiv \{i_1, \alpha_1\}$. Even though the three-particle vertex $\gamma^{(6)}$ is often neglected in actual calculations, we consider it here to illustrate how the theory generalizes to higher-order vertices.

We start from the definition of the dual propagator

$$\tilde{G}_{12} := - \langle f_1 f_2^* \rangle = - \frac{1}{\tilde{Z}} \int \mathcal{D}[f^*, f] \, f_1 f_2^* \, e^{-\tilde{S}[f^*, f]} \,, \tag{41}$$

where the dual action is given by

$$\tilde{S}[f^*, f] = - \sum_{k, \alpha\beta} f_\alpha^* \tilde{G}_{\alpha\beta}^{0 \, -1} f_\beta + \sum_i V_i[f_i^*, f_i] \,. \tag{42}$$

Formally these equations have the same form as those discussed in the introduction of the diagrammatic perturbation theory. This means in particular that we can make use of Wick's theorem and the linked cluster theorem. We can therefore concentrate on connected diagrams. Before considering the diagrams, let us first derive some explicit expressions. The perturbation series is generated by expanding the exponential appearing under the path integral in the interaction

$$e^{-\sum_i V_i[f_i^*, f_i]} = 1 - \sum_i V_i[f_i^*, f_i] + \frac{1}{2!} \sum_{i,j} V_i[f_i^*, f_i] \, V_j[f_j^*, f_j] - \frac{1}{3!} \sum_{i,j,k} V_i[f_i^*, f_i] \, V_j[f_j^*, f_j] \, V_k[f_k^*, f_k] + \tag{43}$$

The zero-order term yields the bare dual Green function

$$\tilde{G}_{12}^0 = - \frac{1}{\tilde{Z}} \int \mathcal{D}[f^*, f] \, f_1 f_2^* \, e^{-\tilde{S}_0[f^*, f]}, \tag{44}$$

with $\tilde{S}_0[f^*, f] = - \sum_{\alpha\beta} f_\alpha^* \tilde{G}_{\alpha\beta}^{0 \, -1} f_\beta$. The next order gives two local contributions for the Green function $\tilde{G}_{12}$, one from each of the vertices

$$\left(-\frac{1}{4}\right) \sum_i \gamma_{i\alpha\beta\gamma\delta}^{(4)} \int \mathcal{D}[f^*, f] \, f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} \, e^{-\tilde{S}_0[f^*, f]} \,, \tag{45}$$

$$\left(\frac{1}{36}\right) \sum_i \gamma_{i\alpha\beta\gamma\delta\epsilon\zeta}^{(6)} \int \mathcal{D}[f^*, f] \, f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{i\epsilon}^* f_{i\zeta} \, e^{-\tilde{S}_0[f^*, f]} \,. \tag{46}$$

The second-order terms (only those involving $\gamma^{(4)}$ and $\gamma^{(6)}$) are

$$-\frac{1}{2!}\left(-\frac{1}{4}\right)^2 \sum_{i,j} \gamma^{(4)}_{i\alpha\beta\gamma\delta}\, \gamma^{(4)}_{j\kappa\lambda\mu\nu} \int \mathcal{D}[f^*,f]\, f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu}\, e^{-\tilde{S}_0[f^*,f]} \qquad (47)$$

$$-\frac{1}{2!}\left(-\frac{1}{4}\,\frac{1}{36}\right) \sum_{i,j} \gamma^{(4)}_{i\alpha\beta\gamma\delta}\, \gamma^{(6)}_{j\kappa\lambda\mu\nu\epsilon\zeta} \int \mathcal{D}[f^*,f]\, f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu} f_{j\epsilon}^* f_{j\zeta}\, e^{-\tilde{S}_0[f^*,f]} \quad (48)$$

$$-\frac{1}{2!}\left(-\frac{1}{4}\,\frac{1}{36}\right) \sum_{i,j} \gamma^{(6)}_{i\kappa\lambda\mu\nu\epsilon\zeta}\, \gamma^{(4)}_{j\alpha\beta\gamma\delta} \int \mathcal{D}[f^*,f]\, f_1 f_2^* f_{i\kappa}^* f_{i\lambda} f_{i\mu}^* f_{i\nu} f_{i\epsilon}^* f_{i\zeta} f_{j\alpha}^* f_{j\beta} f_{j\gamma}^* f_{j\delta}\, e^{-\tilde{S}_0[f^*,f]} \quad (49)$$

$$-\frac{1}{2!}\left(\frac{1}{36}\right)^2 \sum_{i,j} \gamma^{(6)}_{i\alpha\beta\gamma\delta\epsilon\zeta}\, \gamma^{(6)}_{j\kappa\lambda\mu\nu\rho\eta} \int \mathcal{D}[f^*,f]\, f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{i\epsilon}^* f_{i\zeta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu} f_{j\rho}^* f_{j\eta}\, e^{-\tilde{S}_0[f^*,f]}$$

$$(50)$$

Likewise, the third-order term involving only the two-particle vertex $\gamma^{(4)}$ is the following:

$$+\frac{1}{3!}\left(-\frac{1}{4}\right)^3 \sum_{i,j,k} \gamma^{(4)}_{i\alpha\beta\gamma\delta} \gamma^{(4)}_{j\kappa\lambda\mu\nu} \gamma^{(4)}_{k\epsilon\zeta\rho\eta} \int \mathcal{D}[f^*,f]\, f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu} f_{k\epsilon}^* f_{k\zeta} f_{k\rho}^* f_{k\eta}\, e^{-\tilde{S}_0[f^*,f]}$$

$$(51)$$

The terms quickly grow very complicated and it is hard to see the structure in these expressions. In the following we show how we can represent the relevant contributions in terms of diagrams.

3.3 Self-energy diagrams

In general we can use the Dyson equation to sum infinite partial series of diagrams. We therefore concentrate on one-particle irreducible self-energy diagrams in the following. We obtain these simply from diagrams for the Green function by omitting the contractions corresponding to the external lines (those connecting f_1 and f_2^*). With the exception of diagram a), which we include for illustration, we omit any diagram with a local closed loop on at least one of its vertices. Their contribution vanishes because the local part of the dual Green function is taken to be zero via the self-consistency condition. A contraction in the following is defined as $\underline{f_{i\alpha} f_{j\beta}^*} = -\tilde{G}_{ij\,\alpha\beta}$.

We simply write $\tilde{G}$ instead of $\tilde{G}^0$ because the expressions are also valid when written in terms of the interacting Green functions.
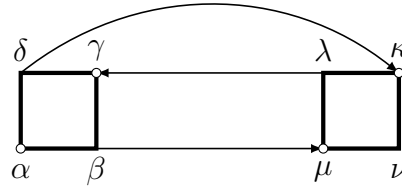
• Diagram (a)
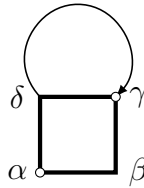
This diagram is derived from (45). The corresponding pairing or contraction is

$$+\underline{f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta}} = -\underline{f_1 f_{i\alpha}^* f_{i\delta} f_{i\gamma}^* f_{i\beta} f_2^*} = (-1)^4\, \tilde{G}_{1\,i\alpha} \tilde{G}_{i\beta\,2} \tilde{G}_{ii\delta\gamma}\,. \qquad (52)$$

The combinatorial prefactor of a diagram is the prefactor of the contribution times the number of pairings that lead to the same (topologically equivalent) diagram. Apparently we obtain

the same diagram if we exchange the 'incoming' points of the vertex (marked by open circles in the figure) or the outgoing ones. This yields 4 pairings which result in the same diagram, which cancels the prefactor $1/4$ of the vertex. The corresponding correction to the self-energy is thus $\Sigma^{(a)}_{ii\,\alpha\beta} = -\gamma^{(4)}_{i\alpha\beta\gamma\delta}\,\tilde{G}_{ii\delta\gamma}$.
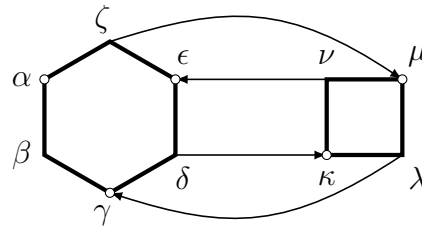
- Diagram (b)

This diagram yields the first non-local correction. The relevant term in the perturbation expansion is (47). The particular pairing corresponding to this diagram is:

$$+f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu} = -f_1 f_{i\alpha}^* f_{i\beta} f_{j\mu}^* f_{j\lambda} f_{i\gamma}^* f_{i\delta} f_{j\kappa}^* f_{j\nu} f_2^*$$

$$= (-1)^6 \, \tilde{G}_{1\,i\alpha} \, \tilde{G}_{ij\beta\mu} \, \tilde{G}_{ji\lambda\gamma} \, \tilde{G}_{ij\delta\kappa} \, \tilde{G}_{i\nu\,2} \,. \tag{53}$$

There are 16 different pairings that correspond to the same diagram. We can count them as follows: Draw two squares corresponding to the two vertices. There are four possibilities to connect an incoming line (two on each vertex). After attaching the incoming line to say, α, there are two possibilities to attach the outgoing line, since it must be connected to the other vertex to yield the desired diagram. Connect this line to ν. Now there are two more possibilities to draw a directed line connecting the two vertices: A line going from the left to the right vertex can only be connected to one point on the left vertex, but to two on the right one. After this line is connected, say from γ to λ, there is only one possibility to connect the remaining two lines. The number of equivalent pairings is thus $4 \cdot 2 \cdot 2 = 16$. The correction to the self-energy is hence given by

$$\Sigma^{(b)}_{ij\,\alpha\nu} = -\frac{1}{2}\gamma^{(4)}_{i\alpha\beta\gamma\delta}\,\gamma^{(4)}_{j\kappa\lambda\mu\nu}\,\tilde{G}_{ij\beta\mu}\,\tilde{G}_{ji\lambda\gamma}\,\tilde{G}_{ij\delta\kappa}\,. \tag{54}$$

- Diagram (c)
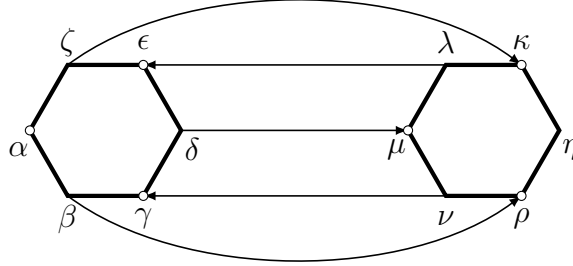

This diagram appears in the two terms (48) and (49), which differ in the order of vertices. The corresponding pairing is

$$f_1 f_2^* f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{i\epsilon}^* f_{i\zeta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu} = +f_1 f_{i\alpha}^* f_{j\lambda} f_{i\gamma}^* f_{i\delta} f_{j\kappa}^* f_{j\nu} f_{i\epsilon}^* f_{i\zeta} f_{j\mu}^* f_{i\beta} f_2^*$$

$$= (-1)^6 \tilde{G}_{1\,i\alpha}\tilde{G}_{ji\lambda\gamma}\tilde{G}_{ij\delta\kappa}\tilde{G}_{ji\nu\epsilon}\tilde{G}_{ij\zeta\mu}\tilde{G}_{i\beta\,2}\,. \tag{55}$$

Counting the number of equivalent pairings as before yields 36. The self-energy correction therefore reads

$$\Sigma_{ii\,\alpha\beta}^{(c)} = \left(\frac{1}{4}\right) \sum_j \gamma_{i\alpha\beta\gamma\delta\epsilon\zeta}^{(6)} \, \gamma_{j\kappa\lambda\mu\nu}^{(4)} \, \tilde{G}_{ji\lambda\gamma} \, \tilde{G}_{ij\delta\kappa} \, \tilde{G}_{ji\nu\epsilon} \, \tilde{G}_{ij\zeta\mu} \, . \tag{56}$$
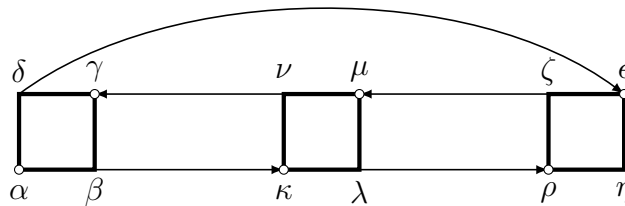
● Diagram (d)

This diagram stems from (50). It corresponds to the pairing

$$f_1 f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{i\epsilon}^* f_{i\zeta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu} f_{j\rho}^* f_{j\eta} f_2^* = f_1 f_{i\alpha}^* f_{i\beta} f_{j\rho}^* f_{j\nu} f_{i\gamma}^* f_{i\delta} f_{j\mu}^* f_{j\lambda} f_{i\epsilon}^* f_{i\zeta} f_{j\kappa}^* f_{j\eta} f_2^*$$

$$= (-1)^7 \tilde{G}_{1\,i\alpha} \tilde{G}_{ij\beta\rho} \tilde{G}_{ji\nu\gamma} \tilde{G}_{ij\delta\mu} \tilde{G}_{ji\lambda\epsilon} \tilde{G}_{ij\zeta\kappa} \tilde{G}_{j\eta\,2}. \tag{57}$$

Here the number of equivalent pairings already becomes quite large: 216. The prefactor is hence given by $(1/2)(1/36)^2 \times 216 = 1/12$. The self-energy correction reads

$$\Sigma_{ij\,\alpha\eta}^{(d)} = \left(\frac{1}{12}\right) \gamma_{i\alpha\beta\gamma\delta\epsilon\zeta}^{(6)} \, \gamma_{j\kappa\lambda\mu\nu\rho\eta}^{(6)} \, \tilde{G}_{ij\beta\rho} \, \tilde{G}_{ji\nu\gamma} \, \tilde{G}_{ij\delta\mu} \, \tilde{G}_{ji\lambda\epsilon} \, \tilde{G}_{ij\zeta\kappa} \, . \tag{58}$$
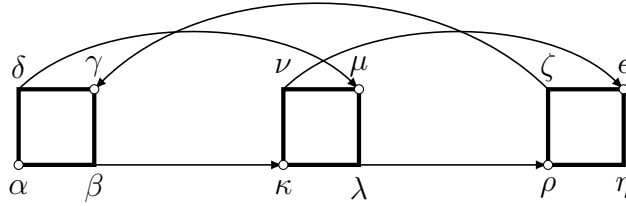
● Diagram (e)

The diagram stems from (51). The corresponding pairing is

$$f_1 f_{i\alpha}^* f_{i\beta} f_{i\gamma}^* f_{i\delta} f_{j\kappa}^* f_{j\lambda} f_{j\mu}^* f_{j\nu} f_{k\epsilon}^* f_{k\zeta} f_{k\rho}^* f_{k\eta} f_2^* = -f_1 f_{i\alpha}^* f_{i\beta} f_{j\kappa}^* f_{j\nu} f_{i\gamma}^* f_{i\delta} f_{k\epsilon}^* f_{j\lambda} f_{k\rho}^* f_{k\zeta} f_{j\mu}^* f_{k\eta} f_2^*$$

$$= (-1)^8 \tilde{G}_{1\alpha} \, \tilde{G}_{ij\beta\kappa} \, \tilde{G}_{ji\nu\gamma} \, \tilde{G}_{ik\delta\epsilon} \, \tilde{G}_{jk\lambda\rho} \, \tilde{G}_{kj\zeta\mu} \, \tilde{G}_{\eta 2} \tag{59}$$

and the number of equivalent pairings is: 384. This gives the self-energy correction

$$\Sigma_{ik\,\alpha\eta}^{(e)} = (-1) \sum_j \gamma_{i\alpha\beta\gamma\delta}^{(4)} \, \gamma_{j\kappa\lambda\mu\nu}^{(4)} \, \gamma_{k\epsilon\zeta\rho\eta}^{(4)} \, \tilde{G}_{ij\beta\kappa} \, \tilde{G}_{ji\nu\gamma} \, \tilde{G}_{ik\delta\epsilon} \, \tilde{G}_{jk\lambda\rho} \, \tilde{G}_{kj\zeta\mu} \, . \tag{60}$$

• Diagram (f)


The diagram stems from the same term as diagram (e). The following pairing:

$$f_1 f^*_{i\alpha} f_{i\beta} f^*_{i\gamma} f_{i\delta} f^*_{j\kappa} f_{j\lambda} f^*_{j\mu} f_{j\nu} f^*_{k\epsilon} f_{k\zeta} f^*_{k\rho} f_{k\eta} f^*_2 = -f_1 f^*_{i\alpha} f_{i\beta} f^*_{j\kappa} f_{k\zeta} f^*_{i\gamma} f_{i\delta} f^*_{j\mu} f_{j\lambda} f^*_{k\rho} f_{j\nu} f^*_{k\epsilon} f_{k\eta} f^*_2$$

$$= (-1)^8 \tilde{G}_{1 i\alpha} \tilde{G}_{ij\beta\kappa} \tilde{G}_{ki\zeta\gamma} \tilde{G}_{ij\delta\mu} \tilde{G}_{jk\lambda\rho} \tilde{G}_{jk\nu\epsilon} \tilde{G}_{k\eta 2} \,,$$

(61)

however, leads to a topologically inequivalent diagram, as seen in the figure. The number of equivalent pairings is also different from diagram (e) and equal to 96. The resulting self-energy correction reads

$$\Sigma^{(\mathrm{f})}_{ik\,\alpha\eta} = \left(-\frac{1}{4}\right) \sum_j \gamma^{(4)}_{i\alpha\beta\gamma\delta} \gamma^{(4)}_{j\kappa\lambda\mu\nu} \gamma^{(4)}_{k\epsilon\zeta\rho\eta} \, \tilde{G}_{ij\beta\kappa} \, \tilde{G}_{ki\zeta\gamma} \, \tilde{G}_{ij\delta\mu} \, \tilde{G}_{jk\lambda\rho} \, \tilde{G}_{jk\nu\epsilon} \,. \qquad (62)$$

We can deform this diagram as shown below by drawing the vertex in a different way. We see that we have two parallel arrows connecting neighboring vertices. This diagram therefore describes renormalization of the self-energy by scattering of particle-particle pairs. Diagram (e) by contrast describes renormalization through particle-hole scattering.
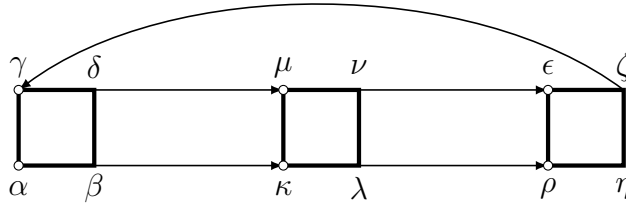
3.3.1 Determination of combinatorial prefactors

As we have seen, the determination of the combinatorial prefactors by counting the number of equivalent pairings becomes cumbersome already for diagrams at moderate orders of the perturbation theory. It is therefore desirable to have general rules to obtain these prefactors simply by looking at the structure of a diagram. Fortunately we can derive such rules by the following analysis:

First consider diagrams which contain no equivalent lines. Equivalent lines are equally directed lines connecting to the same or same two vertices. This is the case for diagrams (a) and (e). The prefactor of such a diagram is unity at any order of the perturbation theory. In order to see this, recall that the prefactor of each vertex is $1/[(n/2)!]^2$, where n is the number of edges. This applies to higher-order vertices as well. It is exactly the number of possibilities to permute equivalent endpoints of each vertex among themselves. For example, the three-particle vertex

has 3 incoming endpoints, and there are 3! permutations of them. We have the same number for the outgoing ones, which yields $(3!)^2$ permutations.

An additional factor $1/m!$ arises from the expansion of the exponential, where m is the perturbation order. Attaching a label to each vertex of one sort (e.g. two-particle vertices) to make them distinguishable, one sees that all $m!$ permutations appear in a complete contraction. If vertices of different sorts are present in a diagram, the factor corresponding to the permutation of these vertices among themselves explicitly appears in the expansion, for example the 2! terms (48) and (49) contributing to diagram (c). Hence a diagram corresponds to the sum of $1/[(n/2)!]^2 m!$ contributions with the same value, so that the prefactor exactly cancels.

This only holds if all ways of attaching the lines or permuting the vertices yield a different, *distinguishable* diagram. If for example two vertices are connected by k equivalent lines, this reduces the number of distinguishable diagrams (pairings) by the number of permutations of these lines, since a permutation yields the identical, distinguishable diagram. Hence the prefactor is cancelled only up to a factor $1/k!$ for each set of k equivalent lines connecting the same two vertices. For example, there are two equivalent lines going from left to right and three parallel lines from right to left in diagram (d). Hence the prefactor is $1/2! \cdot 1/3! = 1/12$. Diagram (f) has two sets of two equivalent lines, while diagram (e) does not. This yields a prefactor of $(1/2!)^2 = 1/4$ for diagram (f) instead of unity for diagram (e).

For vacuum fluctuation diagrams, that is, those with no external lines and no unpaired endpoints, additional symmetry factors arise. An example is the generic n-th order ring diagram shown below. In this diagram, $2n$ cyclic permutations of the sequences $(1, 2, \ldots n)$ and $(n, \ldots, 2, 1)$ correspond to the same distinguishable diagram. Hence the symmetry factor of this diagram is $1/(2n)$. The symmetry factor is obviously unity for self-energy diagrams. The diagrammatic rules for the dual propagator are similar to those for Hugenholtz diagrams [14].
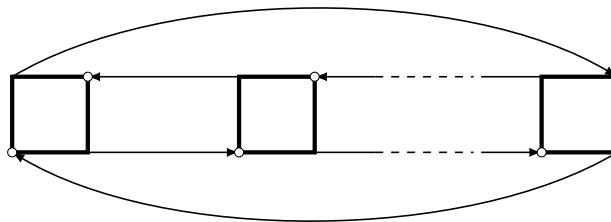
3.3.2 Determination of the sign

Finally we need rules to determine the sign of a given diagram. For the case where the dual potential is truncated after the two-particle interaction term, they can be obtained as follows:
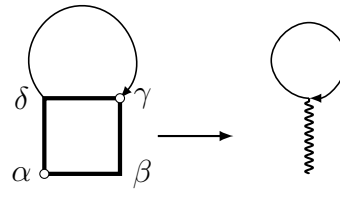
A priori, i.e., regardless of the particular pairing, the contribution to a diagram for Green function or the self-energy has positive sign. A sign of $(-1)^n$, where n is the order or the number of vertices, arises from the expansion of the exponential. This sign cancels at any order due to the negative prefactor of $-1/4$ of each vertex. In each contribution, the Grassmann numbers appear as $-f f^* f^* f \ldots f^* f$, where the sign is due to the definition of Green function, $\tilde{G} = -\langle f f^* \rangle$. Reversing the order of pairs to form a complete contraction and recalling that a contraction of two Grassmann numbers is defined as $f f^* = -\tilde{G}$, one sees that a diagram has positive sign. An

overall sign of $(-1)^{n_L}$ of a diagram arises due to n_L closed fermion loops, as in standard per-
turbation theory. It is, however, not obvious how to count the loops for a given diagram in the
present antisymmetrized technique, where the interaction is fully antisymmetric with respect
to permutation of its endpoints. This can be resolved by comparing with the *un*symmetrized
technique, where the interaction has the form $U_{\alpha\beta\gamma\delta}\, f_\alpha^* f_\beta f_\gamma^* f_\delta = U\, \delta_{\alpha\beta}\, \delta_{\gamma\delta}\, f_\alpha^* f_\beta f_\gamma^* f_\delta$, which can
be represented by a wiggly line as in

$$\tag{63}$$

Since the order of Grassmann variables is the same for the interaction $\gamma_{\alpha\beta\gamma\delta} f_\alpha^* f_\beta f_\gamma^* f_\delta$, the sign
of a diagram is obtained by replacing the square by a wiggly line as in

$$\tag{64}$$

and counting the number of closed loops (a single loop in this example). The relative orientation
of the line and the square must be kept fixed, e.g., both the line and the square in (64) have been
rotated counterclockwise by $\pi/2$ with respect to (63).

3.4 Dual perturbation theory in momentum space

With the above prerequisites, we are now in the position to formulate the dual perturbation
theory in momentum space. The rules to evaluate the expression for a given diagram are:

- draw all topologically distinct connected diagrams involving any n-body interaction $\gamma^{(2n)}$
  depicted by regular polygons with $2n$ edges or endpoints, whereof n are outgoing (in-
  coming) endpoints, where a directed line originates (terminates)

- connect the vertices with directed lines, compliant with ingoing and outgoing endpoints

- with each line associate a dual Green function $\tilde{G}_{\mathbf{k}\nu}$

- assign a frequency, momentum, orbital and spin label to each endpoint

- sum / integrate over all internal variables taking into account energy-, momentum-, and
  spin-conservation at each vertex

- for each tuple of n equivalent lines, associate a factor $1/n!$

- multiply the expression by $(T/N)^m S^{-1} \times s$, where m counts independent frequency /
  momentum summations and S and s are the symmetry factor and sign described above.

4   Numerical results

In the following, we describe two important diagrammatic approximations to the dual self-energy and show results to illustrate their physical content.

4.1   Second-order approximation

We start with the second-order approximation given by diagram (b). We refer to it as $DF^{(2)}$. Applying the above rules of the perturbation theory, we obtain the following expression

$$\tilde{\Sigma}_{\mathbf{k}\nu\sigma} = -\frac{1}{2}\frac{T^2}{N^2}\sum_{\mathbf{k}'\mathbf{q}}\sum_{\nu'\omega}\sum_{\sigma'}\gamma_{\nu\nu'\omega}^{\sigma\sigma\sigma'\sigma'}\,\tilde{G}_{\mathbf{k}+\mathbf{q}\nu+\omega\sigma}\,\tilde{G}_{\mathbf{k}'+\mathbf{q}\nu'+\omega\sigma'}\,\tilde{G}_{\mathbf{k}'\nu'\sigma'}\,\gamma_{\nu'\nu\omega}^{\sigma'\sigma'\sigma\sigma}$$

$$-\frac{1}{2}\frac{T^2}{N^2}\sum_{\mathbf{k}'\mathbf{q}}\sum_{\nu'\omega}\gamma_{\nu\nu'\omega}^{\sigma\bar{\sigma}\bar{\sigma}\sigma}\tilde{G}_{\mathbf{k}+\mathbf{q}\nu+\omega\bar{\sigma}}\,\tilde{G}_{\mathbf{k}'+\mathbf{q}\nu'+\omega\bar{\sigma}}\,\tilde{G}_{\mathbf{k}'\nu'\sigma}\,\gamma_{\nu'\nu\omega}^{\sigma\bar{\sigma}\bar{\sigma}\sigma}. \qquad (65)$$

Here we have left out the first-order contribution. Similarly to DMFT, we repeatedly solve the impurity model. Contrary to DMFT the hybridization is updated to make the local part of the interacting (instead of non-interacting) dual Green function and the first-order diagram zero. The hybridization is therefore different from DMFT. In addition, we evaluate the self-energy in each iteration and renormalize the Green function self-consistently using Dyson's equation.

Fig. 3 shows results for the spectral function obtained in $DF^{(2)}$ and DMFT just above the critical interaction at which the model becomes insulating, which is $U_c = 9.35t$ in DMFT, while it is significantly reduced by spatial correlations to $6.64t$ in $DF^{(2)}$. The $DF^{(2)}$ value agrees much better with the cellular DMFT value of $U_c = 6.05t$, which takes nearest-neighbor correlations into account [15]. For better comparability the energy axis has been scaled by U_c. We observe a significantly richer structure in the $DF^{(2)}$ spectral function. While most spectral weight follows the non-interacting dispersion (red line), we observe relatively broad 'shadow bands' marked by arrows roughly following the dispersion shifted by the antiferromagnetic nesting vector $\mathbf{q} = (\pi, \pi)$ (blue line). These bands originate from short-range dynamical antiferromagnetic correlations included through the self-energy correction. While the solution is paramagnetic, spins on neighboring sites favor to align antiferromagnetically on short time scales.
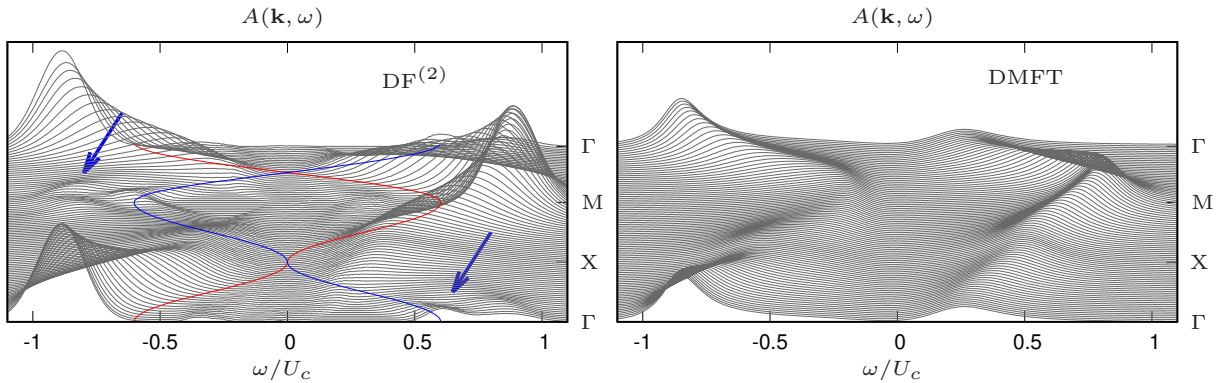

Fig. 3: *Momentum resolved spectral function in $DF^{(2)}$ and DMFT along high-symmetry lines in the Brillouin zone in the paramagnetic insulator. The $DF^{(2)}$ spectral function exhibits shadow bands due to short-range antiferromagnetic correlations.*

4.2   Ladder approximation

A more sophisticated approximation is the so-called ladder dual fermion approximation (LDFA)

$$\Sigma^{\text{LDFA}} = - \; \Box \; - \frac{1}{2} \; \Box \; \Box \; - \; \Box \; \Box \; \Box \; - \cdots \quad (66)$$

The diagrams are given with with corresponding signs and prefactors. All higher-order terms in the expansion have a prefactor of unity. The diagrams look like ladders, with an additional Green function connecting the left and rightmost vertices to close the diagram. Here the rules of the diagrammatic perturbation theory are crucial, because they allow us to derive expressions for every diagram in this infinite partial series without actually carrying out the expansion. To sum the infinite series in a practical way, we take a detour and first sum all the ladder diagrams without the closing Green function (see Fig. 4). Applying the rules of perturbation theory and taking into account all possible spin configurations, we can obtain the following Bethe-Salpeter equation

$$\tilde{\Gamma}^{\sigma\sigma\sigma'\sigma'}_{\mathbf{q}\nu\nu'\omega} = \gamma^{\sigma\sigma\sigma'\sigma'}_{\nu\nu'\omega} - \frac{T}{N} \sum_{\mathbf{k}''\nu''\sigma''} \gamma^{\sigma\sigma\sigma''\sigma''}_{\nu\nu''\omega} \tilde{G}_{\mathbf{k}''+\mathbf{q}\nu''+\omega} \tilde{G}_{\mathbf{k}''\nu''} \tilde{\Gamma}^{\sigma''\sigma''\sigma'\sigma'}_{\mathbf{q}\nu''\nu'\omega}, \quad (67)$$

$$\tilde{\Gamma}^{\sigma\bar{\sigma}\bar{\sigma}\sigma}_{\mathbf{q}\nu\nu'\omega} = \gamma^{\sigma\bar{\sigma}\bar{\sigma}\sigma}_{\nu\nu'\omega} - \frac{T}{N} \sum_{\mathbf{k}''\nu''} \gamma^{\sigma\bar{\sigma}\bar{\sigma}\sigma}_{\nu\nu''\omega} \tilde{G}_{\mathbf{k}''+\mathbf{q}\nu''+\omega} \tilde{G}_{\mathbf{k}''\nu''} \tilde{\Gamma}^{\sigma\bar{\sigma}\bar{\sigma}\sigma}_{\mathbf{q}\nu''\nu'\omega}. \quad (68)$$

Here we have omitted spin labels on the Green functions to emphasize that we consider the paramagnetic state. The negative sign arises because we have a closed loop in the diagram. It is easy to see that by repeatedly inserting the left-hand side into the right, we successively generate the sum over all ladder diagrams. This is illustrated in the left of Fig. 4.

Spin seems to play a particular role here. We have two equations which differ only in the spin labels. The first equation (67) actually corresponds to two coupled equations, one for each of the signs of $\sigma = \pm 1/2$. Apparently it mixes different spin components of the vertex (there is a sum over spins), while the second does not. To understand this, we observe that each vertex γ has an incoming and outgoing line on the left, as well as on the right. If we interpret an arrow pointing to the right as a propagating particle, the arrow to the left is a propagating hole. When we read the diagram from left to right, γ hence describes scattering of particle-hole pairs (the same is true for processes from top to bottom). $\gamma^{\sigma\sigma\sigma'\sigma'}_{\nu\nu'\omega}$ hence describes a particle with spin σ
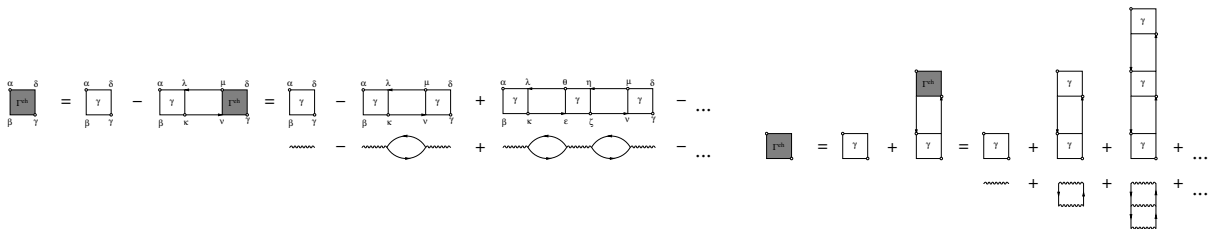

Fig. 4: *Left: Bethe-Salpeter equation in the horizontal electron-hole channel and the first terms in the infinite series it generates. Right: Vertical channel. Diagrams obtained when the vertex γ is replaced by the Hubbard interaction U (depicted by a wiggly line) are also indicated.*

and a hole with spin $\bar{\sigma} = -\sigma$ entering from the left which scatter into a particle with spin σ' and a hole with spin $-\sigma'$. The energy of the pair is ω and is conserved in the scattering process. So is the z-component of the spin: $\sigma - \sigma = \sigma' - \sigma' = 0$. However the total spin is not. By forming the linear combinations $\gamma_{\nu\nu'\omega}^{d(m)} = \gamma_{\nu\nu'\omega}^{\uparrow\uparrow\uparrow\uparrow} \overset{+}{(-)} \gamma_{\nu\nu'\omega}^{\uparrow\uparrow\downarrow\downarrow}$, the two equations decouple. We may say the scattering processes occur in different 'channels' which each correspond to a definite value of the total spin S of the particle hole pair. The plus sign corresponds to the density channel, which can be made plausible by forming the expectation value $\langle n_\sigma n_\uparrow \rangle + \langle n_\sigma n_\downarrow \rangle = \langle n_\sigma n \rangle$, while the minus signs relates to the spin channel, as can be seen from $\langle n_\sigma n_\uparrow \rangle - \langle n_\sigma n_\downarrow \rangle = \langle n_\sigma S_z \rangle$. The latter therefore corresponds to a total spin of $S = 1$, while the former has $S = 0$. With the same reasoning as above we see that the second equation (68) describes scattering of particle hole pairs with projections $S_z = \sigma + \sigma = \pm 1$, so that $S = 1$. We have found all spin states of the particle hole pair: $S = 0, S_z = 0$ and $S = 1, S_z = 0, \pm 1$. In the paramagnetic case, the results cannot depend on the spin projection and we have found the useful identity $\gamma_{\nu\nu'\omega}^{\uparrow\uparrow\uparrow\uparrow} - \gamma_{\nu\nu'\omega}^{\uparrow\uparrow\downarrow\downarrow} = \gamma_{\nu\nu'\omega}^{\uparrow\downarrow\downarrow\uparrow}$. We have just performed a spin-diagonalization, after which the equations can be written in the form

$$\tilde{\Gamma}_{\mathbf{q}\nu\nu'\omega}^{\alpha} = \gamma_{\nu\nu'\omega}^{\alpha} - \frac{T}{N} \sum_{\mathbf{k}''\nu''} \gamma_{\nu\nu''\omega}^{\alpha} \tilde{G}_{\mathbf{k}''+\mathbf{q}\nu''+\omega} \tilde{G}_{\mathbf{k}''\nu''} \tilde{\Gamma}_{\mathbf{q}\nu''\nu'\omega}^{\alpha}, \tag{69}$$

where $\alpha = \mathrm{d}(\mathrm{m})$ denotes the density ($S = 0$) or spin ($S = 1$) channel. The physical content of this equation is the repeated scattering of particle-hole pairs which describes collective charge or spin excitations. The self-energy in the ladder approximation describes effects due to scattering of electrons and these bosonic excitations.

We can solve this equation simply by matrix inversion when viewing the convolution of Green functions as a matrix in ν and ν'

$$[\tilde{\Gamma}_{\mathbf{q}\omega}^{\alpha}]_{\nu\nu'}^{-1} = [\gamma_{\omega}^{\alpha}]_{\nu\nu'}^{-1} + (T/N) \sum_{\mathbf{k}} \tilde{G}_{\mathbf{k}+\mathbf{q}\nu+\omega} \tilde{G}_{\mathbf{k}\nu} \delta_{\nu\nu'}. \tag{70}$$

We can do this independently for fixed $\mathbf{q}, \omega$ and S, which are precisely the quantum numbers of the particle-hole pair and which are conserved in scattering processes.

From the left part of Fig. 4 it would seem that we can obtain the self-energy by simply adding a closing line on the vertex $\tilde{\Gamma}^{eh}$. However we observe a problem here: the second-order contribution would miss the factor $1/2$ expected from the diagrammatic rules.

A more systematic way to obtain the ladder self-energy is the Schwinger-Dyson equation (SDE)

$$\Sigma^{\mathrm{LDFA}} = - \quad\square\quad - \frac{1}{2} \; \boxed{\gamma} \quad \boxed{\Gamma} \tag{71}$$

It connects the self-energy to the exact two-particle vertex. Here we construct an approximation to the vertex via the Bethe-Salpeter equations. However, approximating it as $\tilde{\Gamma} \approx \tilde{\Gamma}^{eh}$ with the horizontal series shown in the left of Fig. 4, we would miss the entire series of vertical diagrams

shown on the right. If the interaction were of Hubbard type, it would be obvious from the figure that the corresponding diagrams are all valid contributions to the vertex which should be taken into account. Both series are in fact the diagrams generated by the well-known fluctuation exchange approximation (FLEX) [16].

On the other hand, if we replace the fully antisymmetric box by wiggly interaction lines,

$$\gamma_{1234} \rightarrow -U(\delta_{12}\delta_{34} - \delta_{14}\delta_{32})\,, \qquad \boxed{\begin{matrix} 1 & 4 \\ \gamma \\ 2 & 3 \end{matrix}} = \Big\} \quad - \quad \text{wwww}\,,$$

we see that already the series in the left of Fig. 4 generates all FLEX diagrams. It is hence plausible that $\tilde{\Gamma}^{\text{eh}}$ and $\tilde{\Gamma}^{\text{v}}$ give the same contribution to the self-energy. We approximate the vertex as $\tilde{\Gamma} \approx \tilde{\Gamma}^{\text{eh}} + \tilde{\Gamma}^{\text{v}} - \gamma$, where we subtract γ once because it appears in both series for $\tilde{\Gamma}^{\text{eh}}$ and $\tilde{\Gamma}^{\text{v}}$. Inserting $\tilde{\Gamma}$ into the SDE yields the LDFA self-energy

$$\tilde{\Sigma}_{\mathbf{k}\nu} = -\frac{T^2}{N^2} \sum_{\mathbf{k'q}} \sum_{\nu'\omega} A_\alpha\, \gamma^\alpha_{\nu\nu'\omega}\, \tilde{G}_{\mathbf{k+q}\nu+\omega}\, \tilde{G}_{\mathbf{k'+q}\nu'+\omega}\, \tilde{G}_{\mathbf{k'}\nu'} \left(\tilde{\Gamma}^{\text{h},\alpha}_{\nu'\nu\omega} - \frac{1}{2}\gamma^\alpha_{\nu'\nu\omega} \right). \tag{72}$$

Here $A_{\text{d}} = 1$, and $A_{\text{m}} = 3$ accounts for the degeneracy ($S_z = 0, \pm 1$) of the magnetic $S = 1$ channel. The two equal contributions from $\tilde{\Gamma}^{\text{eh}}$ and $\tilde{\Gamma}^{\text{v}}$ have canceled the prefactor $1/2$ of the second diagram in the SDE, Eq. (71). The double counting correction γ in $\tilde{\Gamma}$ attains a factor $1/2$ and provides the correct prefactor of the second-order contribution in Eq. (66).

Fig. 5 shows results obtained within the ladder approximation for the two-dimensional Hubbard model. Fig. 5 (a) shows the static ($\omega = 0$), homogeneous ($\mathbf{q} = 0$) spin susceptibility computed from the vertex. In DMFT the response to a homogeneous field increases as moments form with decreasing temperature. After a similar increase in LDFA, neighboring spins start to couple antiferromagnetically. Hence the temperature of the downturn marks the effective exchange energy scale. We can see that the ladder approximation agrees well with Quantum Monte Carlo (QMC) results within error bars. We can therefore be confident that we have included the dominant contributions. Conversely, the entire series of diagrams made up by terms of the form of diagram (f) describing particle-particle scattering that we have neglected apparently plays only a minor role in this regime. We can further see a significant size dependence, indicating that the self-energy corrections are truly long-ranged.

In DMFT the susceptibility diverges at the antiferromagnetic wave vector $\mathbf{q} = (\pi, \pi)$ (not shown). We can view the Bethe-Salpeter equation as a generalization of a geometric series $\sum_{n=0}^{\infty} q^n = 1/(1 - q)$, where the matrix $(T/N) \sum_{\mathbf{k}\nu} \gamma^\alpha_{\nu\nu'\omega}\, \tilde{G}_{\mathbf{k+q}\nu+\omega}\, \tilde{G}_{\mathbf{k}\nu}\, \delta_{\nu\nu'}$ plays the role of q. The divergence hence appears when the leading eigenvalue of this matrix approaches 1. Here it has physical significance and indicates a second-order transition to the antiferromagnetic Néel state as indicated by the vertical line. It is clearly an artifact of the mean-field approximation, because the Mermin-Wagner theorem forbids breaking of a continuous symmetry in two dimensions [17].

In LDFA, we account for the long-range fluctuations that are essential to the proof of the theorem and which destroy the spurious long-range order. To obtain a finite result in LDFA even
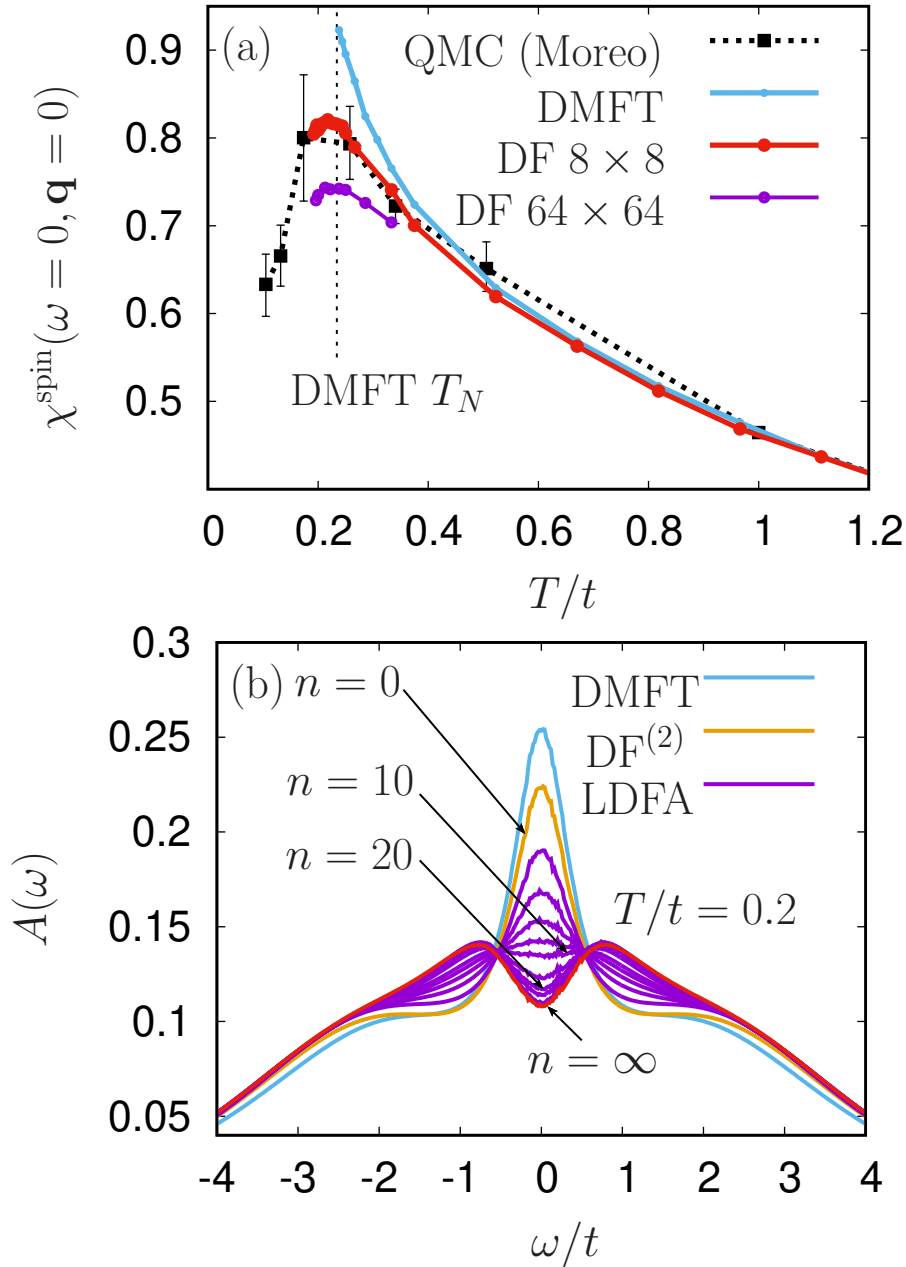
Fig. 5: *(a) Temperature dependence of the static homogeneous susceptibility for $U/t = 4$ and different temperatures. (b) Local density of states in DMFT, $DF^{(2)}$ and different orders of the ladder approximation. All orders contribute to the formation of a pseudogap.*

though the series diverges in DMFT is possible by computing the self-energy from Green functions which are self-consistently renormalized by the self-energy itself. Below the DMFT Néel temperature, this requires a regularization procedure [18].

Fig. 5 (b) shows the local density of states. In DMFT we see the quasi-particle peak and broad Hubbard bands. In the second-order approximation, the spectral weight at the Fermi level is reduced. By iterating the Bethe-Salpeter equation, we can compute the results to successively higher orders. Apparently all orders of the perturbation series contribute to the result. Instead of a quasi-particle peak, the LDFA spectral function exhibits a pseudogap.

5   Summary and outlook

In just two examples we have seen that diagrammatic extensions of DMFT can introduce highly non-trivial effects beyond DMFT and remedy artifacts of the mean-field approximation in particular in low dimensions. They can give qualitatively different and even quantitative results. At the same time, the diagrammatic approach allows us to isolate the dominant contributions and provides us with intuition on the underlying microscopic processes.

In the last ten years of their development, diagrammatic extensions of DMFT have been applied to a variety of systems and scenarios. They have been used to describe unconventional superconductivity in the Kondo lattice model, magnetism in frustrated systems, disordered, inhomogeneous and non-equilibrium systems and even quantum critical behavior in the Hubbard and Falicov-Kimball models. First simple applications to materials have emerged. New methods based on different functionals and the fermion-boson vertices and extensions to cluster- and extended DMFT have been developed. The diagram series has been sampled using diagrammatic Monte Carlo techniques [19, 20] and even more advanced diagrammatic approximations like the parquet equations have been considered. All these developments are summarized in Ref. [6]. For many of the applications the unique ability of the methods to simultaneously describe the strong local dynamical correlations and extended critical fluctuations is crucial. Diagrammatic extensions of DMFT provide a complementary viewpoint to results obtained within other approaches. Most notably within cluster DMFT approaches, in which all diagrams to the self-energy are summed within the range of the comparatively small clusters [21].

The approaches continue to be developed. Recently more fundamental questions are being investigated. For example how to construct conserving approximations when the approximations are made two-particle self-consistent [22], the role of the self-consistency condition in dual fermion [23, 24], or the role of three-particle vertices [25] in dual fermion and DΓA.

Hopefully this introduction will inspire work in two important research directions: i) the combination of diagrammatic extensions with the functional renormalization group and ii) a merger with density-functional theory to arrive at a quantitative theory of correlated materials.

References

[1] N.F. Mott, Rev. Mod. Phys. **40**, 677 (1968)

[2] G.R. Stewart, Rev. Mod. Phys. **56**, 755 (1984)

[3] E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994)

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

[5] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)

[6] G. Rohringer, H. Hafermann, A. Toschi, A.A. Katanin, A.E. Antipov, M.I. Katsnelson, A.I. Lichtenstein, A.N. Rubtsov, and K. Held, Rev. Mod. Phys. **90**, 025003 (2018)

[7] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)

[8] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011)

[9] J.W. Negele and H. Orland: *Quantum Many-Particle Systems* (Perseus books, 1998)

[10] A. Toschi, A.A. Katanin, and K. Held, Phys. Rev. B **75**, 045118 (2007)

[11] H. Kusunose, J. Phys. Soc. Jpn. **75**, 054713 (2006)

[12] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008)

[13] G. Li, Phys. Rev. B **91**, 165134 (2015)

[14] N. Hugenholtz, Physica **23**, 481 (1957)

[15] H. Park, K. Haule, and G. Kotliar, Phys. Rev. Lett. **101**, 186403 (2008)

[16] N.E. Bickers, D.J. Scalapino, and S.R. White, Phys. Rev. Lett. **62**, 961 (1989)

[17] N.D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966)

[18] J. Otsuki, H. Hafermann, and A.I. Lichtenstein, Phys. Rev. B **90**, 235132 (2014)

[19] S. Iskakov, A.E. Antipov, and E. Gull, Phys. Rev. B **94**, 035102 (2016)

[20] J. Gukelberger, E. Kozik, and H. Hafermann, Phys. Rev. B **96**, 035152 (2017)

[21] T. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)

[22] F. Krien, E.G.C.P. van Loon, H. Hafermann, J. Otsuki, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **96**, 075155 (2017)

[23] E.G.C.P. van Loon, M.I. Katsnelson, preprint arXiv:1805.08572 (2018)

[24] T. Ribic, P. Gunacker, preprint arXiv:1805.10996 (2018)

[25] T. Ribic, P. Gunacker, S. Iskakov, M. Wallerberger, G. Rohringer, A.N. Rubtsov, E. Gull, and K. Held, Phys. Rev. B **96**, 235127 (2017)