9   Making Use of Self-Energy Functionals: The Variational Cluster Approximation

Michael Potthoff
I. Institut für Theoretische Physik
Universität Hamburg

Contents

1 Motivation

Self-energy-functional theory (SFT) [1–4] is a general theoretical framework that can be used to construct various approximate approaches by which the thermal properties and the spectrum of one-particle excitations of a certain class of correlated electron systems can be studied. The prototype considered here is the single-band Hubbard model [5–7], but quite generally, the SFT applies to models of strongly correlated fermions on three or lower-dimensional lattices with local interactions.

There are several extensions of the theory, e.g. to systems with non-local interactions [8], to bosonic systems [9, 10] and the Jaynes-Cummings lattice [11, 12], to electron-phonon systems [13], to systems with quenched disorder [14], as well as for the study of the real-time dynamics of systems far from thermal equilibrium [15]. To be concise, those extensions will not be covered here.

The prime example of an approximation that can be constructed within the SFT is the variational cluster approximation (VCA) [2, 16]. Roughly, one of the main ideas of the VCA is to adopt a *divide and conquer* strategy: A tiling of the original lattice into disconnected small clusters is considered, as shown in Fig. 1, for example. While the Hubbard model on the infinite square lattice cannot be solved exactly, there are no serious practical problems in solving the same model for an isolated cluster or for a set of disconnected clusters. The VCA constructs an approximate solution for the infinite lattice from the solution of the individual clusters by means of all-order perturbation theory for those terms in the Hamiltonian that connect the clusters.

This is actually the concept of the so-called cluster perturbation theory (CPT) [17, 18]. However, it is not sufficient in most cases, and we would like to go beyond the CPT. The essential problem becomes apparent, e.g., for a system with spontaneously broken symmetry such as an antiferromagnet. The antiferromagnetic state is characterized by a finite value for the sublattice magnetization which serves as an order parameter. On the other hand, quite generally, the order parameter must be zero for a system of finite size and thus for a small cluster in particular. Coupling finite (and thus necessarily non-magnetic) clusters by means of the CPT, however, one never gets to an antiferromagnetic solution for the infinite lattice. Divide and conquer is not sufficient to describe the emergence of new phases with broken symmetries.
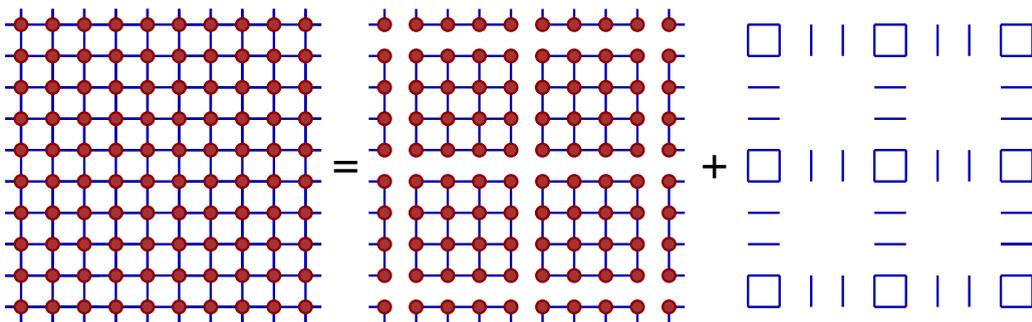

Fig. 1: *Sketch of the decomposition of the original system $H = H_0(\boldsymbol{t}) + H_1$ into a reference system $H' = H_0(\boldsymbol{t}') + H_1$ and the inter-cluster hopping $H_0(\boldsymbol{V})$ for a square lattice and cluster size $L_{\mathrm{c}} = 16$. Blue lines: nearest-neighbor hopping t. Red dots: on-site Hubbard interaction.*
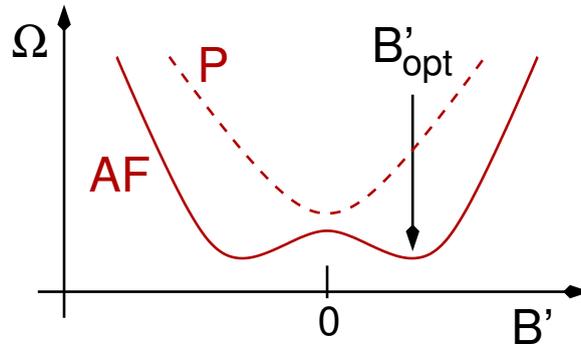
Fig. 2: *Grand potential Ω as a function of a Weiss field B' in the case of a paramagnet (P) and in the case of an antiferromagnet (AF). B' is a fictitious staggered field, the optimal value of which (B'_{opt}) must be determined by minimization of Ω. As there is no physically applied staggered field, i.e. $B = 0$, a finite B'_{opt} indicates* spontaneous *symmetry breaking.*

An obvious way out is to *enforce* a finite antiferromagnetic order parameter within each of the isolated clusters by applying a (staggered) magnetic field B'. Coupling those antiferromagnetic clusters may then result in an antiferromagnetic solution for the entire lattice.

However, what determines the strength of this magnetic field? As we are aiming at a description of *spontaneous* antiferromagnetic order, there is no external *physical* field B that is applied to the original system ($B = 0$). The field B' is actually a *Weiss field*, i.e. a fictitious field or *mean field* that is produced by the system itself. We are seeking for a formalism that allows for the formation of a finite Weiss field if this is favorable, i.e. if a thermodynamical potential can be lowered in this way.

Self-energy-functional theory provides a relation $\Omega(B')$ between the grand potential of the system Ω and the Weiss field B' that can be used to fix the optimal value B'_{opt} of the staggered magnetic field by minimization (see Fig. 2):

$$\left. \frac{\partial \Omega(B')}{\partial B'} \right|_{B'=B'_{\mathrm{opt}}} \overset{!}{=} 0 \tag{1}$$

The purpose of this lecture is to show how this can be achieved in practice. To this end we have to answer the following *how to* questions:

- How can we solve the problem for an isolated cluster?

- With this at hand, how can we construct a solution for the problem on the infinite lattice?

- How can we construct the relation $\Omega(B')$ such that Eq. (1) determines B'_{opt}?

Actually, there is no reason to consider only a staggered magnetic field as a Weiss field. Another goal is therefore to generalize the idea to arbitrary Weiss fields or to an arbitrary set of variational parameters $\boldsymbol{\lambda}'$ that characterize the isolated cluster and that are optimized via $\partial \Omega(\boldsymbol{\lambda}'_{\mathrm{opt}})/\partial \boldsymbol{\lambda}' \overset{!}{=} 0$. Finally, the VCA should be compared with other theories available, and its practical as well as fundamental limitations have to be discussed.

2   The cluster approach

2.1   Tiling the lattice into small clusters

We start with the second question and consider a simple non-interacting system given by

$$H_0 = \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} = H_0(\boldsymbol{t}) \,. \tag{2}$$

Here, $c_{i\sigma}^\dagger$ creates an electron with spin $\sigma = \uparrow, \downarrow$ at the site i of a D-dimensional lattice, and t_{ij} are the (spin-independent) hopping parameters, which are also considered as the elements of the hopping matrix $\boldsymbol{t}$. Furthermore,

$$H_0' = \sum_{ij\sigma} t_{ij}' c_{i\sigma}^\dagger c_{j\sigma} = H_0(\boldsymbol{t}') \,, \tag{3}$$

denotes the Hamiltonian of the system with decoupled clusters (see Fig. 1 and take $H_1 = 0$). If L is the number of lattice sites in the original lattice model H_0 and L_c is the number of sites in an individual cluster, there are L/L_c decoupled clusters. We assume that all clusters are identical. In terms of hopping matrices, we have

$$\boldsymbol{t} = \boldsymbol{t}' + \boldsymbol{V} \tag{4}$$

where $\boldsymbol{V}$ is the inter-cluster hopping.

Consider the resolvent of the hopping matrix, i.e. the Green's function

$$\boldsymbol{G}_0(\omega) = \frac{1}{\omega + \mu - \boldsymbol{t}} \,. \tag{5}$$

Here, ω is a complex frequency (units with $\hbar = 1$ are used). We have also introduced the chemical potential μ (which is not important here but used later). Furthermore, we employ a matrix notation and write ω rather than $\omega \boldsymbol{1}$ etc. for short. Note that $(\cdots)^{-1}$ and $1/(\cdots)$ both mean matrix inversion.

Having the Green's function of the reference system at hand,

$$\boldsymbol{G}_0'(\omega) = \frac{1}{\omega + \mu - \boldsymbol{t}'} \,, \tag{6}$$

how can be get the Green's function of the original model? With some algebra, one easily derives the equation

$$\boldsymbol{G}_0(\omega) = \boldsymbol{G}_0'(\omega) + \boldsymbol{G}_0'(\omega) \boldsymbol{V} \boldsymbol{G}_0(\omega) \tag{7}$$

which is solved by

$$\boldsymbol{G}_0(\omega) = \frac{1}{\boldsymbol{G}_0'(\omega)^{-1} - \boldsymbol{V}} \,. \tag{8}$$

We see that using Green's functions it is formally rather easy to couple a system of isolated clusters.

2.2   Cluster perturbation theory

Actually, we are interested in interacting systems. For the single-orbital model H_0, the only possible local interaction is a Hubbard interaction of the form

$$H_1 = \frac{U}{2} \sum_{i\sigma} n_{i\sigma} n_{i-\sigma} \tag{9}$$

with $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$, where U is the interaction strength. Since H_1 is completely local, the Hamiltonian of the so-called *reference system* $H' = H_0(\boldsymbol{t}') + H_1$ is obtained form the Hamiltonian of the original system $H = H_0(\boldsymbol{t}) + H_1$ by switching off the inter-cluster hopping $\boldsymbol{V}$.

For a small cluster and likewise for a system of disconnected clusters, even for the interacting case, it is comparatively simple to solve the problem exactly (by numerical means if necessary), while for the original lattice model this is a hard problem. One therefore cannot expect a simple relation between the original and the reference system like Eq. (7). Nevertheless, as it is too tempting, we will write down

$$\boldsymbol{G}(\omega) = \boldsymbol{G}'(\omega) + \boldsymbol{G}'(\omega)\boldsymbol{V}\boldsymbol{G}(\omega) \tag{10}$$

where now $\boldsymbol{G}$ and $\boldsymbol{G}'$ are interacting Green's functions. This is an equation that constitutes the cluster-perturbation theory [17, 18]. It must be seen as an approximate way to compute the Green's function of the interacting model from the exact cluster Green's function. In a way the approximation is controlled by the size L_c of the clusters in the reference system since for $L_\mathrm{c} \to \infty$ one can expect the approximation to become exact. In fact, the CPT is not too bad and has been successfully applied in a couple of problems, see Ref. [19] and references therein.

2.3   Green's function and exact diagonalization

Before proceeding with the interpretation of the CPT equation (10), which provides an approximate expression for $\boldsymbol{G}(\omega)$, let us give the *exact* definition of the Green's function for the interacting case. Its elements are defined as

$$G_{ij\sigma}(\omega) = \int_{-\infty}^{\infty} dz\, \frac{A_{ij\sigma}(z)}{\omega - z}\,, \tag{11}$$

where ω is an arbitrary complex frequency and where

$$A_{ij\sigma}(z) = \int_{-\infty}^{\infty} dt\, e^{izt} A_{ij\sigma}(t) \tag{12}$$

is the single-particle spectral density whose Fourier transform

$$A_{ij\sigma}(t) = \frac{1}{2\pi} \langle [c_{i\sigma}(t), c_{j\sigma}^\dagger(0)]_+ \rangle \tag{13}$$

is given as the thermal expectation value of the anti-commutator of the annihilator with the creator in the (grand-canonical) Heisenberg picture, e.g.

$$c_{i\sigma}(t) = e^{i(H-\mu N)t}\, c_{i\sigma}\, e^{-i(H-\mu N)t} \tag{14}$$

with $N = \sum_{i\sigma} n_{i\sigma} = \sum_{i\sigma} c_{i\sigma}^{\dagger} c_{i\sigma}$. The thermal average is a grand-canonical average $\langle \cdots \rangle = Z^{-1}\mathrm{Tr}(e^{-\beta(H-\mu N)} \cdots)$, where $Z = Z(\beta, \mu) = \mathrm{Tr}\, e^{-\beta(H-\mu N)}$ is the partition function at chemical potential μ and inverse temperature β.

In the case of a non-interacting system, one may use the Baker-Campbell-Hausdorff formula to get the simple result ($\boldsymbol{t}$: hopping, t: time):

$$c_{i\sigma}(t) = \sum_{j} \left(e^{-i(\boldsymbol{t}-\mu)t}\right)_{ij} c_{j\sigma} \tag{15}$$

which can be used in Eq. (13), and then via the Fourier transformation (12) and finally the Hilbert transformation (11) one arrives at the result given by Eq. (5) above.

In the interacting case ($U > 0$), one may compute the Green's function from the eigenvalues E_n and eigenstates $|n\rangle$ of the (grand-canonical) Hamiltonian:

$$(H - \mu N)|n\rangle = E_n|n\rangle . \tag{16}$$

Using a resolution of the unity $\boldsymbol{1} = \sum_n |n\rangle\langle n|$ in Eq. (13), one can easily do the calculation and arrives at

$$G_{ij\sigma}(\omega) = \frac{1}{Z} \sum_{mn} \frac{(e^{-\beta E_m} + e^{-\beta E_n})\langle m|c_{i\sigma}|n\rangle\langle n|c_{j\sigma}^{\dagger}|m\rangle}{\omega - (E_n - E_m)} . \tag{17}$$

However, as one must solve the many-body energy eigenvalue problem (16), this way of calculating the Green's function is obviously impossible in practice for the Hamiltonian of the original system – the Hilbert-space dimension exponentially increases with L. On the other hand, for the reference system and if the size of the cluster L_c is not too large, this can be done numerically. For a half-filled system ($N = L_c$), up to $L_c = 8$ sites can be managed in this way easily. At zero temperature, using the Lanczos algorithm [20], the Green's function for somewhat larger clusters can be computed, typically $L_c \leq 12$ at half-filling. This already answers the first question posed in the introduction.

2.4 Freedom in the CPT construction

The CPT gives a preliminary answer to the second question. However, it is easily seen that the answer is not unique: consider a *modified* reference system with a Hamiltonian

$$H_0(\widetilde{\boldsymbol{t}'}) + H_1 = H_0(\boldsymbol{t}') + H_0(\Delta\boldsymbol{t}) + H_1 = \sum_{ij\sigma}(t'_{ij} + \Delta t_{ij})c_{i\sigma}^{\dagger} c_{j\sigma} + H_1 , \tag{18}$$

i.e. a reference system where $\boldsymbol{t}' \mapsto \widetilde{\boldsymbol{t}'} = \boldsymbol{t}' + \Delta\boldsymbol{t}'$. The new reference system shall still describe the same set of decoupled clusters but with different intra-cluster hoppings $\widetilde{\boldsymbol{t}'}$. The modified non-interacting Green's function of the reference system is $\widetilde{\boldsymbol{G}}'_0(\omega) = 1/(\omega + \mu - \widetilde{\boldsymbol{t}'})$. Now, the non-interacting Green's function of the original model is obtained from the equation

$$\boldsymbol{G}_0(\omega) = \widetilde{\boldsymbol{G}}'(\omega) + \widetilde{\boldsymbol{G}}'(\omega)\widetilde{\boldsymbol{V}}\boldsymbol{G}_0(\omega) \tag{19}$$

with the modified inter-cluster hopping $\widetilde{V} = t - \widetilde{t}' = V - \Delta t'$. $G_0(\omega)$ can be considered as the limit of a geometrical series that is found by iterating equation (19):

$$G_0(\omega) = \widetilde{G}'(\omega) + \widetilde{G}'(\omega)\widetilde{V}\widetilde{G}'(\omega) + \cdots . \tag{20}$$

We infer that $G_0(\omega)$ can be obtained by (all-order) perturbation theory in $\widetilde{V}$ when expanding around the Green's function of the modified reference system given by the hopping matrix $\widetilde{t}'$. Obviously, the same result is obtained by perturbation theory in V around the Green's function of the modified reference system with hopping matrix t'. This freedom in choosing the starting point for perturbation theory that we have in the non-interacting case turns into a real problem for the interacting case. Namely, since the CPT equation (10) is approximate, we generally have:

$$\widetilde{G}(\omega) \equiv \widetilde{G}'(\omega) + \widetilde{G}'(\omega)\widetilde{V}\widetilde{G}'(\omega) + \cdots \neq G'(\omega) + G'(\omega)VG'(\omega) + \cdots \equiv G(\omega) . \tag{21}$$

Concluding, different starting points t' and $\widetilde{t}'$ for the all-order cluster perturbation theory in V and $\widetilde{V}$ lead to different results $G(\omega)$ and $\widetilde{G}(\omega)$, respectively.

But which is the *right* starting point? The idea is to turn the problem into an advantage by *optimizing* the starting point: this can be done by making use of a variational principle, i.e. by expressing a thermodynamical potential, e.g. the grand potential Ω, as a function of t' and by subsequent minimization. The optimal t'_{opt} shall be obtained by

$$\left.\frac{\partial \Omega(t')}{\partial t'}\right|_{t'=t'_{\text{opt}}} \overset{!}{=} 0 . \tag{22}$$

We see that the set of variational parameters is just the set of hopping parameters of the reference systems or, in the case of multi-orbital models, simply the set of *all* one-particle parameters except for those, of course, that would couple the different clusters. This set also includes a staggered magnetic field

$$H_0(\widetilde{t}') = H_0(t') - B' \sum_i z_i(n_{i\uparrow} - n_{i\downarrow}) , \tag{23}$$

where $z_i = \pm 1$ alternates between the sites of a bipartite lattice.

2.5 The Ritz principle?

The most popular variational principle is the Ritz variational principle. It states that

$$E[|\Psi\rangle] = \langle \Psi|H|\Psi\rangle = \min. \tag{24}$$

for the ground state of H when the search extends over all normalized trial states $\langle \Psi|\Psi\rangle = 1$. Evaluated at the ground state $|\Psi_0\rangle$, the functional yields the ground-state energy $E[|\Psi_0\rangle] = E_0$.

Hence, a straightforward idea that suggests itself is to compute the normalized ground state $|\Psi(\boldsymbol{t}')\rangle$ of a reference system with hopping matrix $\boldsymbol{t}'$ and to use this as a trial state. The trial state can be varied by varying the parameters $\boldsymbol{t}'$, and the optimal parameters are given by

$$\left.\frac{\partial E[|\Psi(\boldsymbol{t}')\rangle]}{\partial \boldsymbol{t}'}\right|_{\boldsymbol{t}'=\boldsymbol{t}'_{\mathrm{opt}}} \stackrel{!}{=} 0 . \tag{25}$$

To test this idea, let

$$|\Psi(\boldsymbol{t}')\rangle = |\Psi_1(\boldsymbol{t}'_1)\rangle \otimes |\Psi_2(\boldsymbol{t}'_2)\rangle \otimes \cdots \otimes |\Psi_{L/L_{\mathrm{c}}}(\boldsymbol{t}'_{L/L_{\mathrm{c}}})\rangle \tag{26}$$

be the ground state of $H' = H_0(\boldsymbol{t}') + H_1$. It is given as a product of the ground states of the L/L_{c} individual clusters where the ground state of the I-th cluster with hopping matrix $\boldsymbol{t}'_I$ is $|\Psi_I(\boldsymbol{t}'_I)\rangle$. Now, if $E_0(\boldsymbol{t}')$ denotes the ground-state energy of the reference system,

$$E[|\Psi(\boldsymbol{t}')\rangle] = \langle\Psi(\boldsymbol{t}')|(H_0(\boldsymbol{t}') + H_0(\boldsymbol{V}) + H_1)|\Psi(\boldsymbol{t}')\rangle = E_0(\boldsymbol{t}') + \langle\Psi(\boldsymbol{t}')|H_0(\boldsymbol{V})|\Psi(\boldsymbol{t}')\rangle . \tag{27}$$

However, the inter-cluster hopping Hamiltonian $H_0(\boldsymbol{V})$ only contains terms like $c_{i\sigma}^\dagger c_{j\sigma}$ where the sites i and j belong to *different* clusters, say I and J. Hence, $\langle\Psi(\boldsymbol{t}')|c_{i\sigma}^\dagger c_{j\sigma}|\Psi(\boldsymbol{t}')\rangle = \langle\Psi_I(\boldsymbol{t}'_I)| \otimes \langle\Psi_J(\boldsymbol{t}'_J)|c_{i\sigma}^\dagger c_{j\sigma}|\Psi_J(\boldsymbol{t}'_J)\rangle \otimes |\Psi_I(\boldsymbol{t}'_I)\rangle = \langle\Psi_I(\boldsymbol{t}'_I)|c_{i\sigma}^\dagger|\Psi_I(\boldsymbol{t}'_I)\rangle\langle\Psi_J(\boldsymbol{t}'_J)|c_{j\sigma}|\Psi_J(\boldsymbol{t}'_J)\rangle = 0$ as enforced by the conservation of the total particle number. This means that we are left with $E[|\Psi(\boldsymbol{t}')\rangle] = E_0(\boldsymbol{t}')$. As this implies that the optimal parameters $\boldsymbol{t}'_{\mathrm{opt}}$ do not at all depend on $\boldsymbol{V}$, the result is trivial and useless, unfortunately. Even worse, the Hellmann-Feynman theorem [21] tells us that

$$\frac{\partial}{\partial \boldsymbol{t}'} E[|\Psi(\boldsymbol{t}')\rangle] = \frac{\partial}{\partial \boldsymbol{t}'} E_0(\boldsymbol{t}') = \frac{\partial}{\partial \boldsymbol{t}'} \langle\Psi(\boldsymbol{t}')|(H_0(\boldsymbol{t}') + H_1)|\Psi(\boldsymbol{t}')\rangle = \langle\Psi(\boldsymbol{t}')|\frac{\partial H_0(\boldsymbol{t}')}{\partial \boldsymbol{t}'}|\Psi(\boldsymbol{t}')\rangle . \tag{28}$$

This means that, using the Ritz principle, the variational parameters should be determined such that all one-particle intra-cluster correlation functions $\langle c^\dagger c\rangle$, in addition to the inter-cluster correlation functions, vanish.

Concluding, optimizing cluster-perturbation theory cannot be done with the help of the Ritz principle. We mention in passing that this also holds for its finite-temperature and mixed state generalization [22,23]

$$\Omega[\rho] = \mathrm{Tr}\Big(\rho(H - \mu N + T \ln \rho)\Big) \stackrel{!}{=} \min. , \tag{29}$$

where the grand potential, expressed as a functional of the density matrix, is at a minimum for the thermal density matrix $\rho = \exp(-\beta(H - \mu N))/\mathrm{Tr}\exp(-\beta(H - \mu N))$. While this is an extremely useful variational principle, it cannot be used here: a trial density matrix $\rho(\boldsymbol{t}')$, defined as the thermal density matrix of a reference system with a hopping matrix $\boldsymbol{t}'$ that describes decoupled clusters, is a simple product of individual cluster density matrices only. As for the standard Ritz principle, this implies that inter-cluster one-particle correlations are neglected altogether.

3 Diagrammatic perturbation theory

3.1 S-matrix and Green's function

As we have already seen, Green's functions, as opposed to wave functions or density matrices, can be used to couple isolated clusters. All-order perturbation theory in the inter-cluster hopping $\boldsymbol{V}$ yields the exact Green's function in the non-interacting ($U = 0$) case and an approximate (CPT) Green's function for $U > 0$. For the necessary optimization of the starting point, i.e. of the intra-cluster one-particle parameters $\boldsymbol{t}'$, we should therefore try to formulate a variational principle based on Green's functions, i.e. a principle of the form $\delta\Omega[\boldsymbol{G}(\omega)]/\delta\boldsymbol{G}(\omega) \stackrel{!}{=} 0$, and try *test Green's functions* $\boldsymbol{G}'(\omega)$ taken from the reference system. In fact, a variational principle of this type can be constructed with the help of all-order perturbation theory in U [24, 25]. *Vice versa*, a systematic and general perturbation theory in U (and also in $\boldsymbol{V}$) requires putting Green's functions at the focus of the theory. Here, only a brief sketch is given, details can be found in Refs. [25–27], for example. Our goal is to use diagrammatic perturbation theory as a language that can be used to formulate a Green's-function-based variational principle.

We decompose the (grand-canonical) Hamiltonian $\mathcal{H} \equiv H - \mu N$ into a *free* part $\mathcal{H}_0 = H_0 - \mu N$ and the interaction $H_1 \equiv \mathcal{H} - \mathcal{H}_0$. Next we define, for $0 \leq \tau, \tau' \leq \beta$, the so-called S-matrix as

$$S(\tau, \tau') = e^{\mathcal{H}_0\tau}\, e^{-\mathcal{H}(\tau-\tau')}\, e^{-\mathcal{H}_0\tau'} \;, \tag{30}$$

One may interpret $\tau = it$ as an *imaginary-time* variable (where t is real). This *Wick rotation* in the complex time plane has the formal advantage that the thermal density matrix, $\propto e^{-\beta\mathcal{H}}$, is just given by the time-evolution operator, $e^{-i\mathcal{H}t} = e^{-\mathcal{H}\tau}$ at $\tau = \beta$.

There are two main purposes of the S-matrix. First, it can be used to rewrite the partition function in the following way:

$$Z = \mathrm{Tr}\, e^{-\beta\mathcal{H}} = \mathrm{Tr}\left(e^{-\beta\mathcal{H}_0} e^{\beta\mathcal{H}_0} e^{-\beta\mathcal{H}}\right) = \mathrm{Tr}\left(e^{-\beta\mathcal{H}_0} S(\beta, 0)\right) = Z_0 \langle S(\beta, 0)\rangle^{(0)} \;. \tag{31}$$

The partition function of the interacting system is thereby given in terms of the partition function of the *free* system, which is known, and a *free* thermal expectation value of the S-matrix. The second main purpose is related to the imaginary-time Green's function which, for $-\beta < \tau < \beta$, is defined via

$$G_{ij\sigma}(\tau) = -\langle \mathcal{T} c_{i\sigma}(\tau) c_{j\sigma}^\dagger(0)\rangle \tag{32}$$

in terms of an annihilator and a creator with imaginary Heisenberg time dependence:

$$c_{i\sigma}(\tau) = e^{\mathcal{H}\tau}\, c_{i\sigma}\, e^{-\mathcal{H}\tau} \;, \qquad c_{j\sigma}^\dagger(\tau) = e^{\mathcal{H}\tau}\, c_{j\sigma}^\dagger\, e^{-\mathcal{H}\tau} \;. \tag{33}$$

Furthermore, $\mathcal{T}$ is the (imaginary) time-ordering operator. With the help of the S-matrix, the interacting time dependence can be transformed into a *free* time dependence, namely:

$$c_{i\sigma}(\tau) = S(0, \tau)\, c_{I,i\sigma}(\tau)\, S(\tau, 0) \;, \qquad c_{j\sigma}^\dagger(\tau) = S(0, \tau)\, c_{I,j\sigma}^\dagger(\tau)\, S(\tau, 0) \;. \tag{34}$$

Here, the index I (*interaction picture*) indicates that the time dependence is due to $\mathcal{H}_0$ only. This time dependence is simple and can, again, be derived with the Baker-Campbell-Hausdorff formula:

$$c_{I,i\sigma}(\tau) = \sum_j \left(e^{-(\boldsymbol{t}-\mu)\tau} \right)_{ij} c_{j\sigma} \,, \qquad c_{I,i\sigma}^\dagger(\tau) = \sum_j \left(e^{+(\boldsymbol{t}-\mu)\tau} \right)_{ij} c_{j\sigma}^\dagger \,. \tag{35}$$

Outside the imaginary-time interval $-\beta < \tau < \beta$, the Green's function is defined as the periodic continuation: $G_{ij\sigma}(\tau + k \cdot 2\beta) = G_{ij\sigma}(\tau)$ for any integer k. This function has a discrete Fourier representation:

$$G_{ij\sigma}(\tau) = \frac{1}{\beta} \sum_{n=-\infty}^{\infty} G_{ij\sigma}(i\omega_n) \, e^{-i\omega_n\tau} \,, \tag{36}$$

where the Fourier coefficients $G_{ij\sigma}(i\omega_n)$ are defined at the so-called fermionic Matsubara frequencies $i\omega_n = i(2n+1)\pi/\beta$ for integer n and can be computed from $G_{ij\sigma}(\tau)$ as

$$G_{ij\sigma}(i\omega_n) = \int_0^\beta d\tau \, G_{ij\sigma}(\tau) \, e^{i\omega_n\tau} \,. \tag{37}$$

The Green's function $G_{ij\sigma}(\tau)$ is just a different representation of the Green's function $G_{ij\sigma}(\omega)$ introduced with Eq. (11) as its Fourier coefficients are given by $G_{ij\sigma}(i\omega_n) = G_{ij\sigma}(\omega)\big|_{\omega=i\omega_n}$. The remaining problem consists in finding a much more suitable representation of the S-matrix. From its definition, one straightforwardly derives the following equation of motion:

$$-\frac{\partial}{\partial\tau}S(\tau,\tau') = H_{1,I}(\tau)S(\tau,\tau') \,. \tag{38}$$

Here, the time dependence of $H_{1,I}(\tau)$ is due to H_0 only. A formal solution of this differential equation with the initial condition $S(\tau,\tau) = 1$ can be derived easily using the time-ordering operator $\mathcal{T}$ again:

$$S(\tau,\tau') = \mathcal{T} \exp\left(-\int_{\tau'}^{\tau} d\tau'' H_{1,I}(\tau'') \right) \,. \tag{39}$$

Note that, if all quantities were commuting, the solution of Eq. (38) would trivially be given by Eq. (39) without $\mathcal{T}$. The appearance of $\mathcal{T}$ can therefore be understood as necessary to enforce commutativity.

Using this S-matrix representation, the partition function and the Green's function can be written as:

$$\frac{Z}{Z_0} = \left\langle \mathcal{T} \exp\left(-\int_0^\beta d\tau'' H_{1,I}(\tau'') \right) \right\rangle^{(0)} \tag{40}$$

and

$$G_{ij\sigma}(\tau) = -\frac{\left\langle \mathcal{T} \exp\left(-\int_0^\beta d\tau\, H_{1,I}(\tau) \right) c_{I,i\sigma}(\tau) c_{I,j\sigma}^\dagger(0) \right\rangle^{(0)}}{\left\langle \mathcal{T} \exp\left(-\int_0^\beta d\tau\, H_{1,I}(\tau) \right) \right\rangle^{(0)}} \,. \tag{41}$$

The important point is that the expectation values and time dependencies appearing here are free and thus known. Therefore, expanding the exponentials in Eq. (40) and Eq. (41) provides
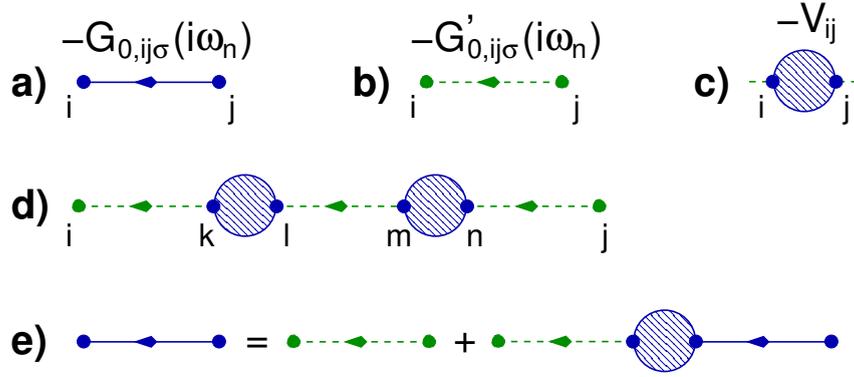
Fig. 3: *Diagrams for potential scattering. See text for discussion.*

an expansion of the partition function and of the Green's function in powers of the interaction strength. The coefficients of this expansion are given as free expectation values of time-ordered products of annihilators and creators with free time dependencies. In kth order, this is a k-particle free correlation function that can be simplified by using Wick's theorem. This is the central theorem of diagrammatic perturbation theory and applies to *free* higher-order correlation functions.

Consider the case of the partition function, as an example. At k-th order, the coefficient is given by a sum of $(2k)!$ terms, each of which factorizes into an k-fold product of terms of the form $\langle \mathcal{T} c_{i\sigma}(\tau) c_{j\sigma}^\dagger(\tau') \rangle^{(0)}$ called propagators. Apart from a sign, a propagator is nothing but the free Green's function. The summation of the $(2k)!$ terms is organized by means of a diagrammatic technique where *vertices* are linked via *propagators*. Wick's theorem and the details of the technique can be found in Refs. [25–27], for example.

3.2   Scattering at the inter-cluster potential, diagrammatically

Here, it is sufficient to illustrate the technique. To this end, we first consider the simple and exactly solvable system that is given by the Hamiltonian $\mathcal{H}_0(\boldsymbol{t}) = H_0(\boldsymbol{t}) - \mu N$ (see Eq. (2)). We decompose the Hamiltonian into a *free* part $\mathcal{H}_0(\boldsymbol{t}') = H_0(\boldsymbol{t}') - \mu N$ (see Eq. (3)) and an "interaction" $H_1 \equiv H_0(\boldsymbol{V})$ (see Eq. (4)). The "fully interacting" propagator, which we are interested in, is $-G_{0,ij\sigma}(i\omega_n)$ and is represented by an oriented line which starts at site j where the electron is created ($c_{j\sigma}^\dagger$) and ends at site i (see Fig. 3a). The *free* propagator $-G'_{0,ij\sigma}(i\omega_n)$ is represented by a dashed line (see Fig. 3b). Propagators carry a frequency $i\omega_n$ and a spin σ. A circle with two links, one for an incoming and one for an outgoing propagator, is called a vertex and stands for the "interaction" $-V_{ij}$ itself (see Fig. 3c). According to Wick's theorem, the contribution of order k to $-G_{0,ij\sigma}$ is obtained by drawing all topologically different diagrams where all links at k vertices are connected by free propagators, except for two external links at the sites i and j (see Fig. 3d). This contribution is calculated by performing the sums over *internal* variables (such as k, l, m, n in Fig. 3d) and respecting frequency and spin conservation at each vertex. These diagram rules can be derived strictly by expanding Eq. (41) and applying Wick's theo-

rem. Together with Eq. (40) this also leads to the important so-called linked-cluster theorem which allows us to concentrate on *connected* diagrams only. The disconnected diagrams for the Green's function (i.e. with external links) exactly cancel diagrammatic contributions from the denominator in Eq. (41). As concerns closed diagrams (no external links) contributing to the partition function, Eq. (40), the sum of only the connected closed diagrams yields, apart from a constant, $\ln Z$, i.e. the grand potential. For the simple case of scattering at the inter-cluster potential discussed at the moment, there is a single connected diagram at each order k only, and thus the "interacting" Green's function is given by

$$- G_{0,ij\sigma}(i\omega_n) = -G'_{0,ij\sigma}(i\omega_n) + \sum_{kl} [-G'_{0,ik\sigma}(i\omega_n)]\,[-V_{kl}]\,[-G'_{0,kj\sigma}(i\omega_n)] + \cdots \quad (42)$$

or, using a matrix formulation and after elimination of the signs,

$$\boldsymbol{G}_0 = \boldsymbol{G}'_0 + \boldsymbol{G}'_0\boldsymbol{V}\boldsymbol{G}'_0 + \boldsymbol{G}'_0\boldsymbol{V}\boldsymbol{G}'_0\boldsymbol{V}\boldsymbol{G}'_0 + \cdots$$
$$= \boldsymbol{G}'_0 + \boldsymbol{G}'_0\boldsymbol{V}(\boldsymbol{G}'_0 + \boldsymbol{G}'_0\boldsymbol{V}\boldsymbol{G}'_0 + \cdots) = \boldsymbol{G}'_0 + \boldsymbol{G}'_0\boldsymbol{V}\boldsymbol{G}_0\,. \quad (43)$$

In this way we have simply re-derived Eq. (7) diagrammatically. This is not yet the CPT equation as the Hubbard interaction has been disregarded.

3.3   Diagram language for systems with Coulomb interaction

Next, let us consider the system given by the Hamiltonian $H_0(\boldsymbol{t}) + H_1$, see Eqs. (2) and (9), and treat the Hubbard (or Coulomb) term H_1 as the interaction, as usual. Also in this case, the free propagator is given by $-\boldsymbol{G}_0$ (Fig. 3a). To represent the interaction $-U$, we need a symbol (red dotted line) with four links, two for *outgoing* and two for *incoming* propagators (Fig. 4a) corresponding to the two creators and the two annihilators in the Hubbard interaction term. Note that the interaction is local and labeled by a site index and that there is energy and spin conservation at a vertex. A diagram contributing to the *interacting propagator* $-G_{ij\sigma}(i\omega_n)$ at order k consists of $2k+1$ propagators fully connecting the k vertices among each other and with the two external links at the sites i and j. Opposed to the potential-scattering problem discussed above, there are many more diagrams at a given order k, namely $(2k+1)!$, one of which, for $k = 3$, is shown in Fig. 3b. The exact Green's function $\boldsymbol{G}(i\omega_n)$ is obtained by summing the algebraic expressions corresponding to those diagrams and summing over all k. The detailed rules necessary for the evaluation of diagrams (see Refs. [25–27]) are not needed here as we do *not* intend to construct a diagrammatically defined approximation by summing a certain subclass of diagrams. While this would be the standard procedure of many-body perturbation theory, here we just want to *speak diagrammatically*.

One can identify so-called self-energy insertions in the diagrammatic series, i.e. parts of diagrams that have links to two external propagators. Examples are given in Fig. 4c where we also distinguish between reducible and irreducible self-energy insertions. The reducible ones can be split into two disconnected parts by removal of a single propagator line. The self-energy
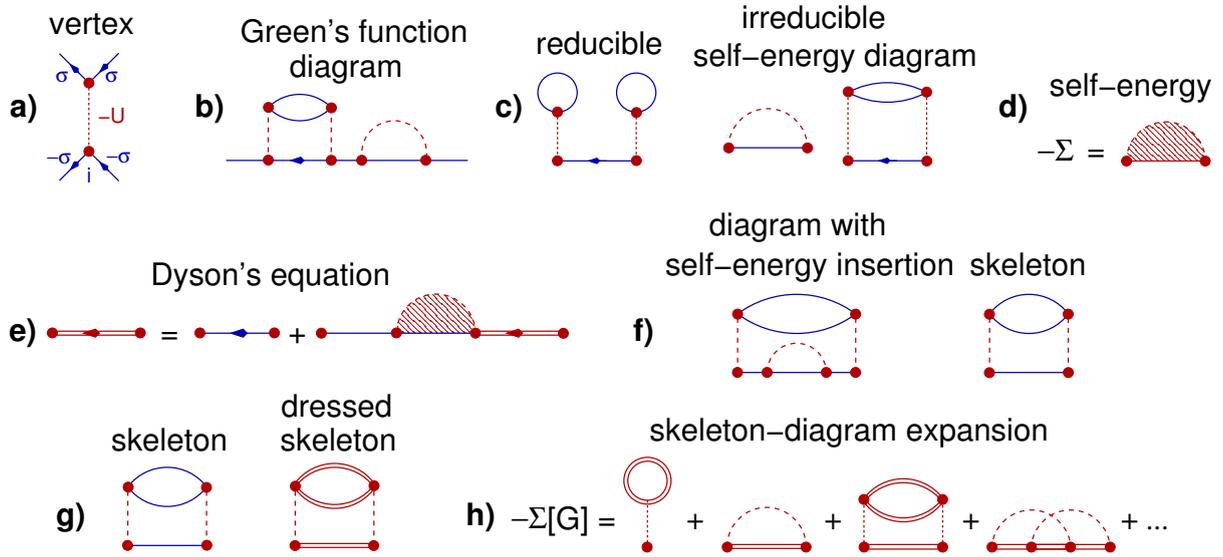
Fig. 4: *Diagram language for systems with Hubbard interaction. See text for discussion.*

$\Sigma_{ij\sigma}(i\omega_n)$ is then defined diagrammatically as the sum over all irreducible self-energy insertions, see Fig. 4d. With this we can derive Dyson's equation

$$\boldsymbol{G}(i\omega_n) = \boldsymbol{G}_0(i\omega_n) + \boldsymbol{G}_0(i\omega_n)\boldsymbol{\Sigma}(i\omega_n)\boldsymbol{G}(i\omega_n) \tag{44}$$

corresponding to Fig. 4e. The double line stands for the interacting propagator $-\boldsymbol{G}(i\omega_n)$. Note that the self-energy plays the same role for the Coulomb interacting system as $\boldsymbol{V}$ does for the scattering problem.

As the first diagram in Fig. 4f shows, there are irreducible self-energy diagrams that contain self-energy insertions. Diagrams without any self-energy insertion are called skeleton diagrams. Skeleton diagrams can be *dressed* by replacing in the diagram the free propagators with interacting propagators (double lines), see Fig. 4g. It is easy to see that the self-energy is given by the sum of the skeleton diagrams only, provided that these are dressed, see Fig. 4h. Therewith, the self-energy is given in terms of the interacting Green's function, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}[\boldsymbol{G}]$. It is only through diagrammatic language that this very important functional relationship, called skeleton-diagram expansion, can be defined rigorously. If combined with Dyson's equation (44), it provides us with a closed equation

$$\boldsymbol{G}(i\omega_n) = \frac{1}{\boldsymbol{G}_0(i\omega_n)^{-1} - \boldsymbol{\Sigma}[\boldsymbol{G}](i\omega_n)} \tag{45}$$

the solution of which is the *exact* Green's function. It is clear, however, that the functional $\boldsymbol{\Sigma}[\boldsymbol{G}]$ is extremely complicated and actually cannot be given in an explicit form, even for the most simple models such as the Hubbard model, and even in cases like small isolated Hubbard cluster, where a numerical computation of the self-energy and the Green's function is easily possible.

3.4   Diagrammatic derivation of the CPT

Equipped with the diagrammatic language, let us come back to the central topic. We have $H = H_0(\boldsymbol{t}') + H_0(\boldsymbol{V}) + H_1$ where the reference system $H' = H_0(\boldsymbol{t}') + H_1$ is easily solvable since it consists of decoupled small clusters, and where $H_0(\boldsymbol{V})$ is the inter-cluster hopping. Ideally, one would start from the solution of H' and perform a perturbative treatment of $H_0(\boldsymbol{V})$. This, however, is not possible (within the above-described standard perturbation theory) as the starting point H' is an interacting system and, therefore, Wick's theorem does not apply. On the other hand, nothing prevents us from starting with $H_0(\boldsymbol{t}')$ and treating both the inter-cluster hopping and the Hubbard interaction, $H_0(\boldsymbol{V})$ and H_1, as the perturbation.

There are two ways to do this: (i) we start from the free ($U = V = 0$) propagator $\boldsymbol{G}'_0$ of $H_0(\boldsymbol{t}')$ and, in a first step, sum the diagrams of all orders in V but for $U = 0$ (see the first line in Fig. 5a). One must merely sum a geometrical series, which can be done exactly. This step has been discussed already in Sec. 3.2. In a subsequent step, the resulting propagator $\boldsymbol{G}_0$ is dressed by taking into account the Hubbard interaction to all orders (see second line in Fig. 5a). This summation would yield the full Green's function $\boldsymbol{G}$ but obviously cannot be done in practice.

We therefore consider an alternative and reverse the order of the two steps: first, the free ($U = V = 0$) propagator $\boldsymbol{G}'_0$ is renormalized by the electron-electron interaction U to all orders but at $V = 0$ (first line in Fig. 5b). This yields the fully interacting cluster Green's function $\boldsymbol{G}'$. While, of course, $\boldsymbol{G}'$ cannot be computed by the extremely complicated summation of individual U diagrams, it is nevertheless easily accessible via a direct (numerical) calculation if the cluster size is sufficiently small (see Sec. 2.3). In the second step, the $V = 0$ propagator $\boldsymbol{G}'$ is renormalized due to inter-cluster potential scattering. Again, this is easily done by summing a geometrical series but only yields an approximation $\boldsymbol{G}_{\mathrm{CPT}}$ to the exact Green's function $\boldsymbol{G}$. In fact, as the second line in Fig. 5b demonstrates, this is just the cluster-perturbation theory, see Eq. (10).

Note that the CPT equation (10) has been introduced in an *ad hoc* way. In contrast, the diagram approach enables understanding of the CPT as an approximation that is given by summing a certain subclass of diagrams. Fig. 5c displays a low-order self-energy diagram that is neglected in this summation. This clearly shows that the CPT cannot be exact and suggests two different routes for improvement, namely (i) taking into account missing diagrams and (ii) using the freedom in the CPT construction to optimize the starting point. The first idea is related to the attempt to perform a systematic perturbative expansion around the disconnected-cluster limit and is notoriously complicated (as Wick's theorem does not apply) [29].

Another *ad hoc* way to derive the CPT follows the idea of the so-called Hubbard-I approximation [5]: The main idea is to employ the Dyson equation (44) *of the reference system* to compute the reference system's self-energy,

$$\boldsymbol{\Sigma}'(i\omega_n) = \boldsymbol{G}'_0(i\omega_n)^{-1} - \boldsymbol{G}'(i\omega_n)^{-1} \,, \tag{46}$$

and to consider this as an approximation for the self-energy of the original system: $\boldsymbol{\Sigma}(i\omega_n) \approx \boldsymbol{\Sigma}'(i\omega_n)$. The motivation for this step is that the self-energy, as opposed to the Green's function, is a much more local object, as is well known at least for the weak-coupling regime from
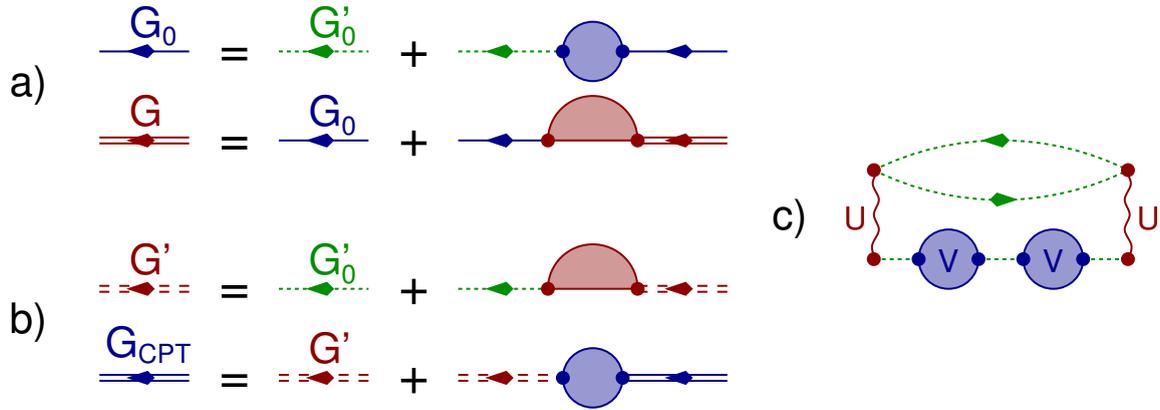
Fig. 5: *Diagrammatic derivation of the CPT, see text for discussion and Ref. [28].*

standard perturbation theory [30–32], and by the fact that the self-energy becomes purely local in the limit of lattices with infinite spatial dimensions [33,34]. Furthermore, the idea is reminiscent of dynamical mean-field theory (DMFT) [33, 35, 36] where the self-energy of an impurity model approximates the self-energy of the lattice model. Using Eq. (46) in Dyson's equation for the original model, we find

$$
\begin{aligned}
\boldsymbol{G}(i\omega_n) &= \frac{1}{\boldsymbol{G}_0(i\omega_n)^{-1} - (\boldsymbol{G}_0'(i\omega_n)^{-1} - \boldsymbol{G}'(i\omega_n)^{-1})} \\
&= \frac{1}{\boldsymbol{G}'(i\omega_n)^{-1} - \boldsymbol{V}} \ ,
\end{aligned}
\tag{47}
$$

i.e., the CPT equation (10) is recovered. We note in passing that the Hubbard-I approach is obtained is for $L_c = 1$ and with some *ad hoc* element of self-consistency [5].

This way to construct the CPT suggests to use, rather than the Ritz principle, a variational principle of the form

$$
\frac{\delta\Omega[\boldsymbol{\Sigma}]}{\delta\boldsymbol{\Sigma}(i\omega_n)} = 0 \ ,
\tag{48}
$$

where the trial self-energy is taken from the reference system and varied by varying the parameters of the reference system. Ideally, this *self-energy-functional approach* should also cure the different defects of the CPT, i.e. besides the arbitrariness of the CPT construction, the non-self-consistent nature of the approach, and its inability to describe spontaneous symmetry breaking, as well as different thermodynamical inconsistencies that show up in the computation of a thermodynamical potential from the Green's function [25–27]. Furthermore, one may ask whether both the CPT and the DMFT can be understood in a single unifying theoretical framework.

4 Self-energy functional theory

4.1 Luttinger-Ward generating functional

The construction of a variational principle based on the self-energy is in fact possible with the help of the so-called Luttinger-Ward functional [24] (see [37] for a pedagogical introduction). This is a scalar functional Φ of the Green's function $\boldsymbol{G}$ that was originally defined by all-order perturbation theory (a construction that uses the path integral can be found in Ref. [38]). More specifically, $\Phi[\boldsymbol{G}]$ is defined as the sum of all closed, connected, and fully dressed skeleton diagrams of any order k. Fig. 6 shows the lowest-order diagrams. Closed diagrams without links to external propagators are diagrams contributing to the partition function, see Eq. (40). The Luttinger-Ward series is given by dressed skeleton diagrams, i.e. diagrams without self-energy insertions where the free propagators are replaced by the fully interacting ones. One easily verifies that, due to dressing of the diagrams, some diagrams in the expansion of Z/Z_0 are counted twice or more. This is done on purpose. The most important property of the Luttinger-Ward functional constructed in this way is that its functional derivative just yields the skeleton-diagram expansion of the self-energy:

$$\frac{\delta\Phi[\boldsymbol{G}]}{\delta\boldsymbol{G}(i\omega_n)} = \frac{1}{\beta}\boldsymbol{\Sigma}[\boldsymbol{G}](i\omega_n) \ . \tag{49}$$

This can be verified, diagram by diagram: the functional derivative of a dressed skeleton just corresponds to the removal of a dressed propagator and results in a dressed skeleton diagram with two links for external propagators that contributes to the self-energy. When carefully taking into account the coefficients of the two different expansions Eq. (40) and Eq. (41), one easily derives Eq. (49). The equation is remarkable as it shows that the different components of the self-energy $\Sigma_{ij\sigma}(i\omega)$ can be obtained from the *scalar* functional. In fact, the existence of the Luttinger-Ward functional can also be proven by verifying a vanishing-curl condition as has been done by Baym and Kadanoff [39, 40].

The value Φ of the Luttinger-Ward functional has no direct physical meaning. Summing all closed diagrams (not only connected skeletons) yields, by construction, the partition function Z/Z_0. Summing connected diagrams only yields $\ln Z$ as is known from the linked-cluster theorem [24, 25]. The sum of dressed connected skeletons, however, cannot provide the grand potential $\propto \ln Z$ because of the above-mentioned double counting.
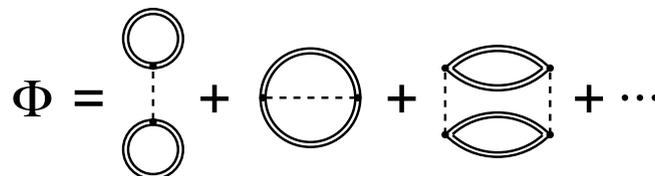

Fig. 6: *Diagrammatic construction of the Luttinger-Ward functional $\Phi[\boldsymbol{G}]$. Double lines stand for fully interacting propagators, dashed lines for the Hubbard interaction.*

4.2 Self-energy functional

We will make use of $\Phi[\boldsymbol{G}]$ by defining the following functional of the self-energy:

$$\Omega[\boldsymbol{\Sigma}] = \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}_0^{-1}-\boldsymbol{\Sigma}} + \Phi[\boldsymbol{G}[\boldsymbol{\Sigma}]] - \mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{G}[\boldsymbol{\Sigma}])\,. \tag{50}$$

Here, the frequency dependencies are suppressed in the notation and

$$\mathrm{Tr}\,\boldsymbol{A} \equiv \frac{1}{\beta}\sum_n\sum_{i\sigma}e^{i\omega_n 0^+}A_{ii\sigma}(i\omega_n) \tag{51}$$

is used where 0^+ is a positive infinitesimal. Furthermore, $\boldsymbol{G}[\boldsymbol{\Sigma}]$ is the inverse of the functional $\boldsymbol{\Sigma}[\boldsymbol{G}]$, i.e. $\boldsymbol{G}[\boldsymbol{\Sigma}[\boldsymbol{G}]] = \boldsymbol{G}$. We assume that this inverse of the skeleton-diagram expansion of the self-energy exists at least *locally*. The second part of the self-energy functional,

$$F[\boldsymbol{\Sigma}] \equiv \Phi[\boldsymbol{G}[\boldsymbol{\Sigma}]] - \mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{G}[\boldsymbol{\Sigma}])\,, \tag{52}$$

is just the Legendre transform of the Luttinger-Ward functional. With $\boldsymbol{\Sigma}[\boldsymbol{G}[\boldsymbol{\Sigma}]] = \boldsymbol{\Sigma}$ and Eq. (49) we immediately have

$$\frac{\delta F[\boldsymbol{\Sigma}]}{\delta\boldsymbol{\Sigma}} = -\frac{1}{\beta}\boldsymbol{G}[\boldsymbol{\Sigma}]\,. \tag{53}$$

Therewith, we can also calculate the functional derivative of $\Omega[\boldsymbol{\Sigma}]$:

$$\frac{\delta\Omega[\boldsymbol{\Sigma}]}{\delta\boldsymbol{\Sigma}} = \frac{1}{\beta}\left(\frac{1}{\boldsymbol{G}_0^{-1}-\boldsymbol{\Sigma}} - \boldsymbol{G}[\boldsymbol{\Sigma}]\right)\,. \tag{54}$$

The equation

$$\boldsymbol{G}[\boldsymbol{\Sigma}] = \frac{1}{\boldsymbol{G}_0^{-1}-\boldsymbol{\Sigma}} \tag{55}$$

is a (highly non-linear) conditional equation for the self-energy of the system $H = H_0(\boldsymbol{t}) + H_1$. Inserting $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}[\boldsymbol{G}]$ shows that it is (locally) equivalent to Eq. (45). It is satisfied by the exact self-energy of the system. Therefore, solving Eq. (55) is equivalent to a search for the stationary point of the self-energy functional:

$$\frac{\delta\Omega[\boldsymbol{\Sigma}]}{\delta\boldsymbol{\Sigma}} = 0\,. \tag{56}$$

This represents the dynamical variational principle we have been looking for. The exact self-energy of the system makes the self-energy functional $\Omega[\boldsymbol{\Sigma}]$, Eq. (50), stationary.

The definition of the self-energy functional given with Eq. (50) is a formal one only. The argument of the ln is not dimensionless, and furthermore, since $\boldsymbol{G}(i\omega_n) \propto 1/\omega_n \propto 1/(2n+1)$ for large n, the sum over the Matsubara frequencies $\sum_n\ln(2n+1)$ does not converge. This problem can be solved, however, by replacing $\Omega[\boldsymbol{\Sigma}] \mapsto \Omega[\boldsymbol{\Sigma}] - \mathrm{Tr}\ln\boldsymbol{G}_{\mathrm{reg}}$ with $G_{\mathrm{reg},ij\sigma}^{-1}(i\omega_n) = \delta_{ij}(i\omega_n - \varepsilon_{\mathrm{reg}})$ and taking the limit $\varepsilon_{\mathrm{reg}} \to \infty$ *after* all calculations are done. As the constant $\mathrm{Tr}\ln\boldsymbol{G}_{\mathrm{reg}}$ does not depend on $\boldsymbol{\Sigma}$, the variational principle is unaffected but now the Matsubara sum over both logarithms is well defined and convergent. One can show [24, 25, 3] that, if evaluated at the physical (exact) self-energy, the regularized $\Omega[\boldsymbol{\Sigma}] - \mathrm{Tr}\ln\boldsymbol{G}_{\mathrm{reg}}$ is just the

grand potential of the system. This provides us with a physical interpretation of the self-energy functional. In the following this regularization is always implicit.

As a remark, we note that at $U = 0$ the self-energy functional reduces to the expression $\Omega_0 \equiv \text{Tr} \ln \boldsymbol{G}_0$ as becomes obvious from the diagrammatic definition of $\Phi[\boldsymbol{G}]$ and of $\boldsymbol{\Sigma}$ since there are simply no diagrams left at zeroth order in the interaction strength:

$$\Phi[\boldsymbol{G}] \equiv 0 \,, \quad \boldsymbol{\Sigma}(i\omega_n) = 0 \qquad \text{for } U = 0 \,. \tag{57}$$

If regularized properly, $\Omega_0 \mapsto \Omega_0 - \text{Tr} \ln \boldsymbol{G}_{\text{reg}}$, this exactly yields the grand potential of the non-interacting system.

4.3   Evaluation of the self-energy functional

The diagrammatic definition of the Luttinger-Ward functional (Fig. 6) uncovers another remarkable property: since any diagram contributing to Φ consists of vertices and dressed propagators only, the functional relation $\Phi[\cdots]$ is completely determined by the interaction U but does not depend on $\boldsymbol{t}$. Clearly, this universality also holds for its Legendre transform $F[\boldsymbol{\Sigma}]$: two systems (at the same chemical potential μ and inverse temperature β) with the same interaction H_1 but different one-particle parameters $\boldsymbol{t}$ and $\boldsymbol{t}'$ are described by the same functional $F[\boldsymbol{\Sigma}]$. In contrast, the first part of the self-energy functional,

$$\Omega[\boldsymbol{\Sigma}] = \text{Tr} \ln \frac{1}{\boldsymbol{G}_0^{-1} - \boldsymbol{\Sigma}} + F[\boldsymbol{\Sigma}] \tag{58}$$

does depend on the hopping, namely via $\boldsymbol{G}_0^{-1}(i\omega_n) = i\omega_n + \mu - \boldsymbol{t}$, but not on the interaction strength U.

The universality property of $F[\boldsymbol{\Sigma}]$ is more important, however, as this functional is basically unknown. The central idea of self-energy-functional theory is to compare the self-energy functional of two systems, the original system with $H = H_0(t) + H_1$ and the reference system with $H' = H_0(\boldsymbol{t}') + H_1$, i.e. Eq. (58) and

$$\Omega'[\boldsymbol{\Sigma}] = \text{Tr} \ln \frac{1}{\boldsymbol{G}_0'^{-1} - \boldsymbol{\Sigma}} + F[\boldsymbol{\Sigma}] \,. \tag{59}$$

Due to its universality, one can eliminate the unknown functional $F[\boldsymbol{\Sigma}]$ by combining both equations:

$$\Omega[\boldsymbol{\Sigma}] = \Omega'[\boldsymbol{\Sigma}] + \text{Tr} \ln \frac{1}{\boldsymbol{G}_0^{-1} - \boldsymbol{\Sigma}} - \text{Tr} \ln \frac{1}{\boldsymbol{G}_0'^{-1} - \boldsymbol{\Sigma}} \,. \tag{60}$$

This equation is still exact. Since the functional dependence of $\Omega'[\boldsymbol{\Sigma}]$ is also unknown in the case of a simple reference system with decoupled clusters, it appears that this step amounts to a mere shift of the problem. The great advantage of Eq. (60) becomes manifest, however, when inserting the exact self-energy of the reference system $\boldsymbol{\Sigma}_{\boldsymbol{t}'}(i\omega_n)$ as a trial self-energy. (We use the notation $\boldsymbol{\Sigma}_{\boldsymbol{t}'}$ for the exact self-energy of the system with hopping parameters $\boldsymbol{t}'$ and interaction H_1.) In this case, the first term on the right-hand side of Eq. (60) just reduces to the
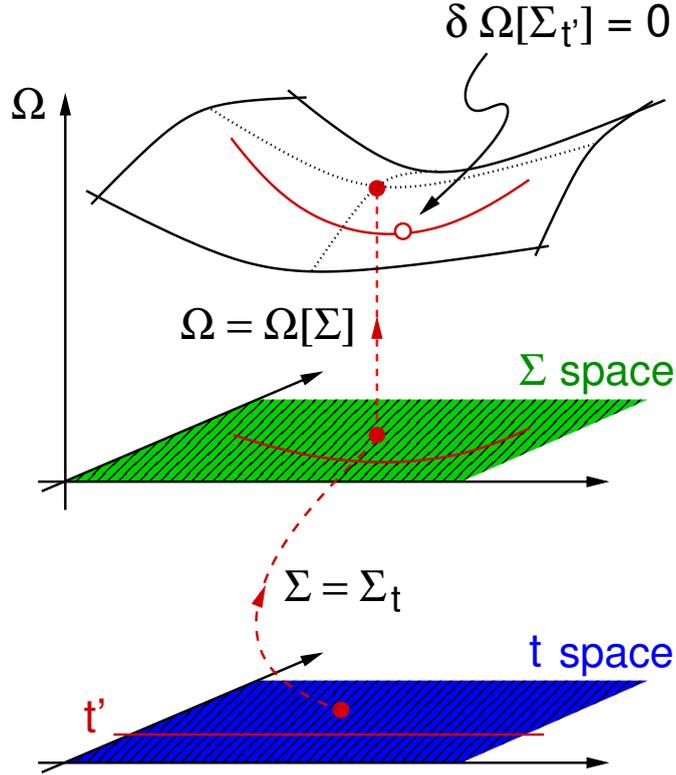
Fig. 7: *The construction of consistent approximations within the self-energy-functional theory. The grand potential is considered as a functional of the self-energy that is stationary at the physical (exact) self-energy $\boldsymbol{\Sigma}_t$ (filled red circles). The functional dependence of $\Omega[\boldsymbol{\Sigma}]$ is not accessible on the entire space of self-energies (Σ space). However, $\Omega[\boldsymbol{\Sigma}]$ can be evaluated exactly on a restricted subspace of trial self-energies $\boldsymbol{\Sigma}_{t'}$ parametrized by a subset of one-particle parameters $\boldsymbol{t}'$ (solid red lines). These $\boldsymbol{t}'$ define an exactly solvable reference system, i.e. a manifold of systems with the same interaction part but a modified one-particle part given by $\boldsymbol{t}'$. Typically, the reference system consists of a set of decoupled clusters. A self-energy at which the grand potential is stationary on this sub-manifold represents the approximate self-energy of the original system and the grand potential at this self-energy represents the approximate grand potential (open circle).*

grand potential of the reference system Ω', which can be computed easily if, as we assume, the reference is amenable to an exact numerical solution. The same holds for the second and the third term. We find:

$$\Omega[\boldsymbol{\Sigma}_{t'}] = \Omega' + \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}_0^{-1} - \boldsymbol{\Sigma}_{t'}} - \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}_0'^{-1} - \boldsymbol{\Sigma}_{t'}} \ . \tag{61}$$

This is a remarkable result. It shows that an *exact* evaluation of the self-energy functional of a non-trivial interacting system is possible, at least for trial self-energies that are taken from an exactly solvable reference system with the same interaction part (see Fig. 7).

4.4   The variational cluster approximation

We recall that the cluster perturbation theory approximates the self-energy of the original lattice-fermion model by the self-energy of a reference system of disconnected clusters. As one may choose the intra-cluster parameters of the reference system different from the corresponding parameters of the original system, there is a certain arbitrariness in the CPT construction. Usually, one simply assumes that e.g. the intra-cluster nearest-neighbor hopping of the reference system is the same as the physical hopping. There are, however, good reasons not to do so. Symmetry breaking Weiss fields are one example, as already mentioned above. Another one becomes obvious from Fig. 5c, where the CPT is seen to neglect the effect of the scattering at the inter-cluster potential on the self-energy. Therefore, an enhanced *intra*-cluster hopping could, at least partially, compensate for the missing feedback of the *inter*-cluster hopping on the approximate self-energy.

With the self-energy-functional framework at hand, we can now remove the arbitrariness of the CPT approach and determine the *optimal* self-energy from Eq. (61). This optimal self-energy is the exact self-energy of an optimized reference system that is specified by a set of one-particle (intra-cluster) parameters $\boldsymbol{t}'$. Note that to derive Eq. (61) it was necessary to assume that the interaction part H_1 of the reference system cannot be optimized and must be the same as the interaction of the original system. Therefore, the role of the reference system is to generate a manifold of trial self-energies $\boldsymbol{\Sigma}_{\boldsymbol{t}'}$ that are parameterized by the one-particle parameters $\boldsymbol{t}'$. As the self-energy functional Eq. (58) can be evaluated *exactly* on this manifold via Eq. (61), the optimal self-energy $\boldsymbol{\Sigma}_{\boldsymbol{t}'_{\mathrm{opt}}}$ is given as the solution of the SFT Euler equation

$$\left. \frac{\partial \Omega[\boldsymbol{\Sigma}_{\boldsymbol{t}'}]}{\partial \boldsymbol{t}'} \right|_{\boldsymbol{t}'=\boldsymbol{t}'_{\mathrm{opt}}} = 0 \, . \tag{62}$$

For a cluster reference system, this constitutes the variational cluster approximation (VCA).

The VCA represents an approximation as it provides the stationary point of the self-energy functional on a restricted manifold of trial self-energies only rather than on the entire *self-energy space* (see Fig. 7). The latter could be defined as the set of the self-energies of all models with the interaction part fixed at H_1 but with a completely arbitrary one-particle part (that also may connect the clusters). This space, of course, contains the exact self-energy $\boldsymbol{\Sigma}_{\boldsymbol{t}}$ of our lattice model $H_0(\boldsymbol{t}) + H_1$ while the optimized VCA self-energy $\boldsymbol{\Sigma}_{\boldsymbol{t}'_{\mathrm{opt}}}$ is constrained to the manifold of cluster trial self-energies. Approximations generated in this way have a number of advantageous properties: first of all, although we have employed the language of perturbation theory, the VCA is non-perturbative. Formally, the diagram series has not been cut at any level, no subclass of diagrams is neglected, etc. The approximation rather results from a restricted domain of a self-energy functional. Second, the VCA is an internally consistent approximation in the sense that all observables derive from an explicit (though approximate) expression for a thermodynamical potential, namely from the self-energy functional evaluated at the optimal self-energy $\Omega[\boldsymbol{\Sigma}_{\boldsymbol{t}'_{\mathrm{opt}}}]$. Third, the VCA can be improved in a systematic way by increasing the cluster size L_{c}, as one has to approach the exact solution for $L_{\mathrm{c}} \to \infty$. Here, the reference system is basically identical

with the original system, and t is basically *within* the space of the variational parameters t'. One may interpret $1/L_c$ as the small parameter that controls the quality of the approximation. It is clear, however, that the numerical effort to solve Eq. (62) also increases with L_c. This increase is even exponential if an exact-diagonalization solver is used to compute the cluster self-energy, Green's function, and grand potential that enter Eq. (61). Unfortunately, one cannot *a priori* give a distance by which the optimal VCA self-energy differs from the exact self-energy. In practice, the quality of the approximation must therefore be controlled by comparing the results obtained for different cluster sizes L_c. For small clusters, also the cluster geometry and the imposed cluster boundary conditions matter, and must be checked.

Although the VCA derives from a general variational principle, it is not variational in the sense that the approximate VCA grand potential $\Omega[\boldsymbol{\Sigma}_{t'_{\text{opt}}}]$ must always be larger than the exact grand potential Ω. As opposed to the Ritz principle and the state functional $E[|\Psi\rangle]$ or, at finite temperatures, the density-matrix functional $\Omega[\rho]$, the self-energy functional $\Omega[\boldsymbol{\Sigma}]$ is not convex and hence there is no reason to assume that the VCA provides an upper bound to Ω.

Concluding, the VCA must be seen as a cluster mean-field approximation that focusses on one-particle correlations, the one-particle excitation spectrum (e.g. photoemission) and thermodynamics. It treats short-range one-particle correlations within the cluster in an explicit way while inter-cluster one-particle correlations are accounted for via Dyson's equation. The feedback of local and intra-cluster *two*-particle (and even higher) correlations on the one-particle self-energy is explicitly and non-perturbatively taken into account while the feedback of non-local two-particle, e.g. magnetic, correlations on the one-particle spectrum is neglected altogether. This is typical for cluster mean-field theories [41] and should be kept in mind when studying, e.g., systems close to a second-order phase transition, where non-local correlations play an important role.

5 Implementation of the variational cluster approximation

Q-matrices

The bottleneck of a practical VCA calculation consists in the computation of the Green's function of the reference system. Using an exact-diagonalization technique, the Green's function for an individual cluster can be obtained in its Lehmann representation, see Eq. (17). Let $\alpha = (i, \sigma)$ be an index referring to the elements of the localized orbitals forming an orthonormal basis of the one-particle Hilbert space. Therewith, the elements of the cluster Green's function can be written in the form

$$G'_{\alpha\beta}(\omega) = \sum_m Q'_{\alpha m} \frac{1}{\omega - \omega'_m} Q'^{\dagger}_{m\beta} . \tag{63}$$

Here, $m = (r, s)$ refers to a single-particle excitation between two energy eigenstates $|s\rangle$ and $|r\rangle$ of the (grand-canonical) Hamiltonian of the reference system $H' - \mu N$, and $\omega'_m = E'_r - E'_s$ is the excitation energy. $Q'_{\alpha m}$ are the elements of the so-called Q'-matrix [42], which is a rectangular matrix with a small number of rows but a large number of columns (dimension of

the one-particle Hilbert space $\times$ number of many-body excitations):

$$Q'_{\alpha m} = \langle r|c_\alpha|s\rangle \sqrt{\frac{\exp(-\beta E'_r) + \exp(-\beta E'_s)}{Z'}} \, , \quad Z' = \sum_r e^{-\beta E'_r} \, , \tag{64}$$

as is readily read off from the Lehmann representation Eq. (17). One also verifies that $Q'Q'^\dagger = 1 \neq Q'^\dagger Q'$. Using the Q'-matrix, we can write the reference system's Green's function in a compact form as

$$G'(\omega) = Q' \frac{1}{\omega - \Lambda'} Q'^\dagger \, , \tag{65}$$

where Λ' is the diagonal matrix with elements $\Lambda'_{mn} = \omega'_m \delta_{mn}$. With $V = t - t'$, the Green's function of the original system is obtained as:

$$G(\omega) = \frac{1}{G'(\omega)^{-1} - V} = G'(\omega) + G'(\omega)VG'(\omega) + \cdots$$

$$= Q' \left(\frac{1}{\omega - \Lambda'} + \frac{1}{\omega - \Lambda'} Q'^\dagger V Q' \frac{1}{\omega - \Lambda'} + \cdots \right) Q'^\dagger = Q' \frac{1}{\omega - M} Q'^\dagger \, , \tag{66}$$

where $M = \Lambda' + Q'^\dagger V Q'$ is a (large) square Hermitian matrix that can be diagonalized by a unitary transformation, $M = S \Lambda S^\dagger$. Here, $\Lambda_{mn} = \omega_m \delta_{mn}$ with the poles ω_m of $G(\omega)$. We find

$$G(\omega) = Q \frac{1}{\omega - \Lambda} Q^\dagger \tag{67}$$

with $Q = Q'S$. The representations Eq. (65) and Eq. (67) are particularly useful to evaluate the self-energy functional Eq. (61) in practice. The trace contains a Matsubara-frequency summation which can be carried out analytically [43] such that one is left with a simple algebraic expression [42],

$$\mathrm{Tr}\ln \frac{1}{G_0^{-1} - \Sigma_{t'}} - \mathrm{Tr}\ln \frac{1}{G_0'^{-1} - \Sigma_{t'}} = -\sum_m \frac{1}{\beta} \ln(1 + e^{-\beta \omega_m}) + \sum_m \frac{1}{\beta} \ln(1 + e^{-\beta \omega'_m}) \, , \tag{68}$$

which involves the poles of $G(\omega)$ and $G'(\omega)$ only. Finally, the grand potential of the reference system in Eq. (61) is easily computed as $\Omega' = -(1/\beta) \ln \sum_r e^{-\beta E'_r}$.

Recipe for practical calculations

A typical VCA calculation is carried out as follows:

- Construct a reference system by tiling the original lattice into identical clusters.

- Choose a set of one-particle parameters t' of the reference system and compute $V = t - t'$.

- Solve the problem for the reference system (U is fixed), i.e. compute the Green's function G' and find the poles ω'_m and the Q'-matrix.

- Get the poles ω_m of the approximate Green's function of the original system by diagonalization of the matrix $M = \Lambda' + Q'^\dagger V Q$.

- Calculate the value of the SFT grand potential via Eq. (61) and Eq. (68) and by calculating the grand potential of the reference system Ω' from the eigenvalues of H'.

- Iterate this scheme for different $\boldsymbol{t}'$, such that one can solve

$$
\left. \frac{\partial \Omega[\boldsymbol{\Sigma}_{\boldsymbol{t}'}]}{\partial \boldsymbol{t}'} \right|_{\boldsymbol{t}'=\boldsymbol{t}'_{\mathrm{opt}}} \overset{!}{=} 0 \tag{69}
$$

  for $\boldsymbol{t}'_{\mathrm{opt}}$.

- Evaluate observables, such as $\Omega[\boldsymbol{\Sigma}_{\boldsymbol{t}'_{\mathrm{opt}}}]$, $\boldsymbol{G}(\omega)$ and static expectation values derived from the SFT grand potential by differentiation, at the stationary point $\boldsymbol{t}'_{\mathrm{opt}}$.

- Redo the calculations for different parameters of the *original* system, e.g. a different U, filling or β to scan the interesting parameter space.

Tips and tricks

For a given topology of the reference system, i.e. for a given cluster geometry, one may in principle consider all one-particle parameters $\boldsymbol{t}'$ as variational parameters. However, besides an exponentially increasing Hilbert-space dimension, a larger cluster also implies an increasing numerical complexity for the search of the stationary point since $\Omega[\boldsymbol{\Sigma}_{\boldsymbol{t}'}]$ is a function of a multi-component variable $\boldsymbol{t}'$. It is therefore advisable to restrict the search to a small number of physically important parameters. In most cases, a few variational parameters suggest themselves.

An overall shift $\Delta\varepsilon'$ of the on-site energies in the cluster (like the chemical potential), $t'_{ii} \mapsto t'_{ii} + \Delta\varepsilon'$, should be among the variational parameters to ensure thermodynamical consistency with respect to the total particle number as has been pointed out in Ref. [44]. This ensures that both ways to compute the total particle number, $\langle N \rangle = -\partial\Omega/\partial\mu$ and $\langle N \rangle = \sum_{i\sigma} \int dz\, A_{ii\sigma}(z)/(e^{\beta z}+1)$, must yield the same result. Analogously, in case of a (ferro- or antiferro-) magnetic system, one should include a (homogeneous or staggered) Weiss field B' in the set of variational parameters. For a paramagnetic system and for a system with manifest particle-hole symmetry, however, symmetry considerations *a priori* fix those variational parameters to $B' = 0$ and $\Delta\varepsilon' = 0$, i.e. $t'_{ii} = t_{ii}$. This can also be verified by a practical VCA calculation.

For the setup of self-energy-functional theory it is inevitable that the original and the reference system have the same interaction H_1. Conversely, the one-particle part of the reference system can be designed at will. One very interesting option in this context is to add additional fictitious sites to the cluster. These *bath sites* have to be non-interacting ($U = 0$), contrary to the *correlated sites* ($U > 0$), which correspond to the physical sites (with the same U) of the original system. Adding the bath sites does not change the interaction part H_1 of the Hamiltonian and therefore leaves the Luttinger-Ward functional as well as its Legendre transform $F[\boldsymbol{\Sigma}]$ unaffected. Bath sites can be coupled via one-particle hopping terms to the correlated sites in the

cluster. This construction has the appealing advantage of increasing the space of variational parameters and thereby improving the quality of the approximation *locally*. Adding bath sites and optimizing the additional associated parameters will improve the description of local temporal correlations while increasing the cluster improves the theory with respect to non-local spatial correlations.

If one decides to consider a reference system with bath sites, it is advisable to formally include the same bath sites also in the original system. Here, of course, they are completely decoupled from the correlated sites (the respective parts of the hopping matrix t have to be set to zero) such that all physical quantities remain unchanged. The advantage of this trick is that t and t' have the same matrix dimension, and that the Hamiltonians H and H' operate on the same Hilbert space. The *inter-cluster hopping* $V = t - t'$ includes the hopping terms between correlated and bath sites in the reference system only.

Rather than employing the above-mentioned Q-matrix technique, one may also perform the traces in Eq. (61), i.e. the trace of the spatial and orbital degrees of freedom and the implicit Matsubara-frequency summation, numerically. This is recommended if the dimension of M, given by the number of poles of G' with non-vanishing spectral weight, becomes too large.

If a full diagonalization of the cluster problem is not feasible and Krylov-space methods shall be applied, one has to make sure that the different elements $G'_{ij\sigma}(\omega)$ have the same set of poles: ω'_m should be independent of i, j. This can be achieved by the band Lanczos method [45]. The dimension of the matrix M is given by the number of iteration steps in the Lanczos procedure. Typically, about 100 steps are sufficient for reasonably well converged results. This should be checked regularly.

The SFT grand potential may exhibit more than a single stationary point. A minimal grand potential among the grand potentials at the different stationary points distinguishes the thermodynamically stable phase in most situations [46]. Often, the occurrence of several stationary points is welcome from a physical point of view. For example, scanning a physical parameter, e.g. U, a second-order magnetic phase transition is characterized by a bifurcation of a non-magnetic solution into a non-magnetic and a magnetic one (or even more magnetic ones).

There are different numerical strategies to determine a stationary point of the self-energy functional, see Ref. [47] for examples. If there is only a single variational parameter to be optimized, iterative bracketing of maxima and minima can be employed efficiently. In the case of more than one variational parameter, the SFT grand potential usually exhibits a saddle point rather than a minimum or maximum. A strategy that has been found to be useful for two or three parameters is to assume (and verify) a certain characteristic of the saddle point and to apply iterated one-dimensional optimizations. The downhill simplex method can be used for higher-dimensional parameter spaces to find the local minima of $|\partial\Omega[\boldsymbol{\Sigma}_{t'}]/\partial t'|^2$. If there is more than one, only those must be retained for which $\Omega[\boldsymbol{\Sigma}(t')]$ has a vanishing gradient.

6 Selected results

The VCA is not restricted to the single-band Hubbard model but has also been applied to a variety of multi-orbital systems. The necessary generalization to the multi-orbital case is straightforward. In this way, the VCA has contributed to the study of the correlated electronic structure of real materials such as NiO [48], CoO and MnO [49], CrO_2 [50], $LaCoO_3$ [51], TiOCl [52], TiN [53], and NiMnSb [54]. Here, however, we will focus on the single-band model and discuss a few and very simple examples to illustrate the theory.

6.1 One-dimensional Hubbard model

In the first example [55], we will consider the one-dimensional Hubbard model at zero temperature and half-filling with hopping $t = 1$ between nearest neighbors, see Fig. 8a. A tiling of the one-dimensional lattice into "clusters" is particularly simple: Each cluster is a finite chain of L_c sites. We treat the intra-cluster nearest-neighbor hopping t' as the only variational parameter. This is the most obvious choice. Nevertheless, one may numerically check that the optimal on-site hopping $t'_{ii,\text{opt}} = t_{ii} = 0$. The same holds for the hopping between second nearest neighbors: $t'_{2-\text{nd,opt}} = 0$. Again this is predicted by particle-hole symmetry. On the other hand, if a third nearest-neighbor hopping is introduced as a variational parameter, it acquires a small finite value at the stationary point. Interestingly, one also finds $t'_{\text{pbc,opt}} = 0$ [2], where t'_{pbc} is a hopping parameter that links the two edge sites of the cluster with each other. $t'_{\text{pbc}} = t'$ would be a realization of periodic boundary conditions, but the calculation shows that open boundaries, $t'_{\text{pbc,opt}} = 0$, are preferred. Furthermore, one may also relax the constraint that the hopping t' be the same for all pairs of nearest neighbors. In this case one finds the strongest deviations close to the edges of the reference systems [55].

Fig. 8b shows the dependence of the SFT grand potential $\Omega[\Sigma_{t'}]$ on the *single* variational parameter t'. Actually, $(\Omega + \mu\langle N\rangle)/L$ is plotted. At zero temperature and at the stationary point, this is the (approximate) ground-state energy of the Hubbard model per site. There is a stationary point, a minimum in this case, with the optimal value for the intra-cluster hopping t'_{opt} being close to but different from the physical value $t = 1$ for strong Coulomb interaction U. Note that the CPT is given by $t' = t$ and that there is a gain in binding energy due to the optimization of t', namely $\Omega(t'_{\text{opt}}) < \Omega(t)$ which implies that the VCA improves on the CPT result.

It is physically reasonable that in the case of a stronger interaction and thus more localized electrons, switching off the inter-cluster hopping is less significant and must therefore be offset to a lesser degree by an increase of the intra-cluster hopping. A considerably large deviation from the physical hopping, $t'_{\text{opt}} > t$, is only found for the weakly interacting system. However, even a "strong" approximation of the self-energy (measured as a strong deviation of t'_{opt} from t) becomes irrelevant in the weak-coupling limit as the self-energy becomes small. With decreasing U, the self-energy functional becomes flatter and flatter until at $U = 0$ the t' dependence is completely irrelevant. Note that not only the non-interacting limit but also the atomic limit ($t = 0$) is exactly reproduced by the VCA. In the latter case, the reference system becomes
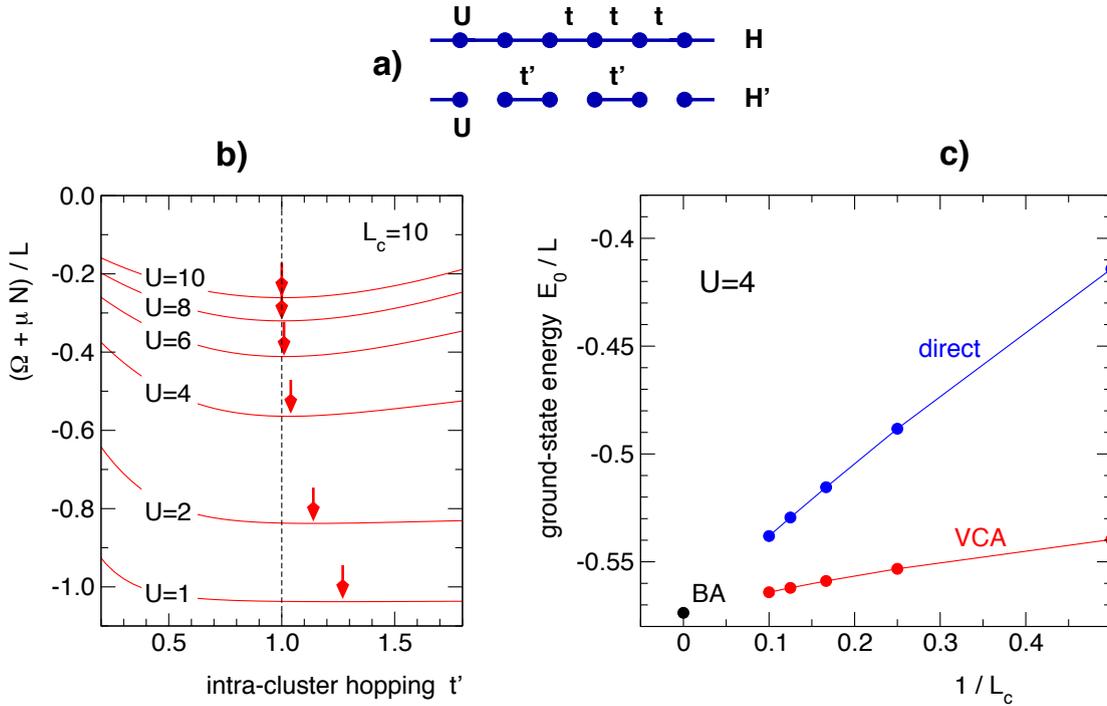
Fig. 8: *a) Original system: one-dimensional Hubbard model. Reference system: decoupled clusters (for $L_c = 2$). b) SFT grand potential per site and shifted by μN as a function of the intra-cluster nearest-neighbor hopping t'. VCA calculation (with $L_c = 10$) for the one-dimensional Hubbard model at zero temperature, half-filling and different U as indicated. The nearest-neighbor hopping $t = 1$ sets the energy scale. Arrows mark the stationary points. c) VCA ground-state energy per site as a function of $1/L_c$ for $U = 4$ at the respective stationary points compared with the corresponding results for an isolated cluster and the exact results known from the Bethe ansatz (BA) [56]. (adapted from Ref. [55])*

identical to the original system at $t' = 0$.

Fig. 8c shows the VCA ground-state energy (per site) at a fixed interaction strength $U = 4$ as a function of the inverse cluster size $1/L_c$. By extrapolation to $1/L_c = 0$ one recovers the exact Bethe-Ansatz result (BA) [56]. Furthermore, the VCA is seen to improve the ground-state energy as compared to calculations done for an isolated Hubbard chain with open boundaries. Convergence to the BA result is clearly faster within the VCA. Note that, as opposed to the VCA, the direct cluster approach is not exact for $U = 0$.

6.2   Antiferromagnetism

With the second example [16], we return to our original motivation, see Fig. 2: one of the main drawbacks of the CPT consists in its inability to describe spontaneous symmetry breaking. Consider SU(2) transformations in spin space and antiferromagnetic order, for example. As the exact solution of a finite Hubbard cluster is necessarily spin-symmetric, i.e., invariant under SU(2) transformations, and as the CPT equation proliferates this symmetry, the antiferromagnetic order parameter, the staggered magnetization m, must always be zero if there is
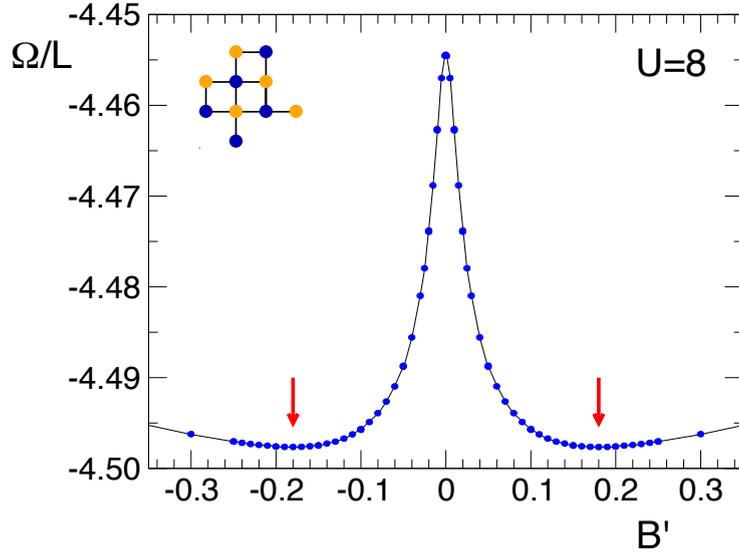
Fig. 9: *SFT grand potential per site as a function of the strength of a fictitious staggered magnetic field B'. Calculation for the two-dimensional half-filled Hubbard model on the square lattice at zero temperature and $U = 8$, $t = 1$. The reference system consists of disconnected clusters with $L_c = 10$ sites each, see the inset for the cluster geometry. Arrows mark the two equivalent stationary points. (adapted from Ref. [16])*

no physically applied staggered magnetic field that would explicitly break the symmetry, i.e. if $B = 0$. The VCA, on the other hand, in principle allows for a spontaneous SU(2) symmetry breaking. Namely, treating an intra-cluster fictitious staggered magnetic field of strength B' as a variational parameter offers the possibility for a symmetry-broken stationary point with $B'_{\rm opt} \neq 0$. The reference-system Hamiltonian is given by

$$H' = H'\big|_{B'=0} - B' \sum_{i\sigma} z_i(n_{i\uparrow} - n_{i\downarrow}) \,, \tag{70}$$

where $z_i = +1$ for sites on sublattice A and $z_i = -1$ for sublattice B (yellow and blue sites in the inset of Fig. 9)

The main part of Fig. 9 displays VCA results for the half-filled two-dimensional Hubbard model on the square lattice at zero temperature. Decoupled clusters with $L_c = 10$ sites are considered as a reference system, and the staggered field B' is the only variational parameter considered. There is a stationary point at $B' = 0$ that corresponds to the paramagnetic phase and to the CPT. In addition, however, there are two equivalent stationary points at *finite* B' corresponding to a phase with antiferromagnetic order. Comparing the ground-state energies of both the antiferromagnetic and the paramagnetic phase shows that the former is thermodynamically stable.

One should be aware, however, that the VCA, like any cluster mean-field approach, tends to overestimate the tendency towards magnetic order. Furthermore, a finite-temperature calculation is expected to produce a finite order parameter for the two-dimensional but also for the one-dimensional case, which would be at variance with the Mermin-Wagner theorem [57]. What is missing physically in the VCA is the effect of long wavelength spin excitations. The VCA is therefore restricted to cases where the physical properties are dominated by short-range correlations on a scale accessible by an exactly solvable finite cluster.

6.3   Mott metal-insulator transition

The third example [58] addresses a first-order (discontinuous) phase transition. This type of phase transition can be studied conveniently within the SFT framework as there is an explicit expression for a thermodynamical potential available from the very beginning. We again consider the two-dimensional Hubbard model on the square lattice at half-filling and zero temperature but disregard the antiferromagnetic phase and enforce a paramagnetic state by choosing $B' = 0$. The paramagnetic system is expected to undergo a transition from a correlated metal at weak U to a Mott insulator at strong U. This Mott transition is first of all interesting from a fundamental point of view as it is driven by electronic correlations opposed to other types of metal-insulator transitions [59]. The Mott insulator is characterized by a gap of the order of U in the single-particle excitation spectrum which is only weakly dependent on temperature. One therefore expects that a possible metal-insulator quantum phase transition at zero temperature is of relevance for the high-temperature state of the system as well, where it should give rise to a smooth crossover between a more metallic and a more insulating state. The crossover takes place at temperatures that may be well above the Néel temperature where the system is paramagnetic. This is the motivation to ignore the magnetic phase for the zero-temperature calculation.

Fig. 10 (right) shows the building block of the reference system. This is a cluster with $L_c = 4$ correlated sites (filled blue dots) but with four bath sites (open red dots) in addition, i.e. there are $n_s = 2$ *local* degrees of freedom (one additional bath site per correlated site). As mentioned above, including bath sites in a VCA calculation improves the description of local correlations. This is an important ingredient in understanding the Mott transition: a paradigmatic picture of the Mott transition could be worked out in the limit of infinite spatial dimensions with the help of the dynamical mean-field theory [33, 35]. In this limit the self-energy becomes a completely local quantity, $\Sigma_{ij\sigma}(i\omega_n) = \delta_{ij}\Sigma_{i\sigma}(i\omega_n)$ [34], and therefore the local temporal degrees of freedom (n_s) dominate the spatial ones (L_c). For two dimensions, the considered reference system with $L_c = 4$ and $n_s = 2$ is expected to represent a good compromise between the importance of local and non-local correlations and to result in a reasonable approximation.

VCA calculations with the full set of variational parameters indicated in Fig. 10 (right) have shown [58] that the hopping between the correlated and the bath sites V is the most important parameter to be optimized while $t'_{\mathrm{opt}} \approx t$ and $t''_{\mathrm{opt}} \approx 0$ can safely be ignored, i.e. set to the *a priori* plausible values $t'_{\mathrm{opt}} = t$ and $t''_{\mathrm{opt}} = 0$. The on-site energies of the correlated and the bath sites are fixed by particle-hole symmetry anyway. This drastically simplifies the study, as the SFT grand potential $\Omega[\boldsymbol{\Sigma}_V]$ can be regarded as a function of a single variational parameter V.

Fig. 10 (left) displays the SFT grand potential, shifted by μN, per correlated site as a function of V for different U. For weak interactions, $U < U_{c2} \approx 6.35$, there is a stationary point (a minimum) at a comparatively large V_{opt} that describes a metallic phase (blue dots). The metallic character of the phase can be inferred from the finite value of the imaginary part of the local Green's function $\mathrm{Im}\, G_{ii\sigma}(i\omega)$ for $\omega \to 0$ (see Ref. [58]). Above the critical value U_{c2}, no metallic solution can be found. There is, however, an insulating phase for strong U (red
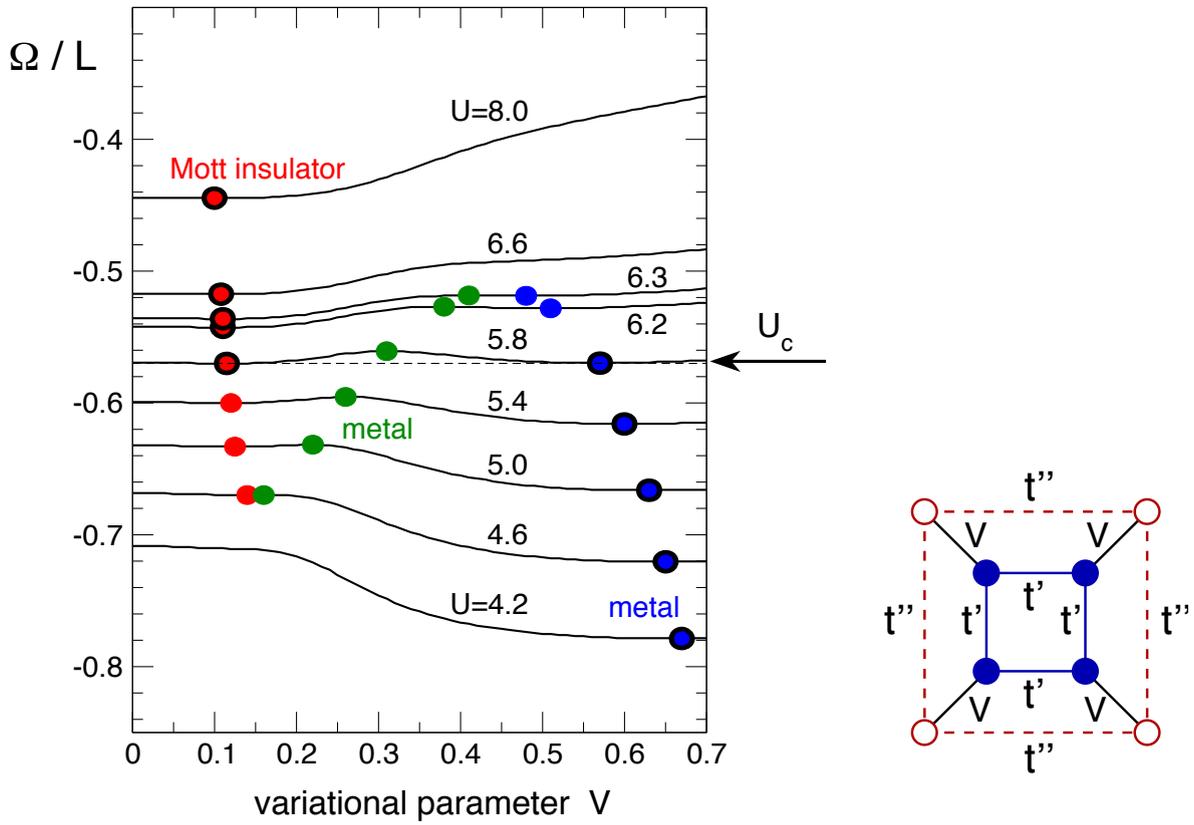
Fig. 10: Left: *SFT grand potential per site as a function of the variational parameter V. VCA calculation for the two-dimensional Hubbard model on the square lattice at zero temperature, half-filling, and different U as indicated. The nearest-neighbor hopping $t = 1$ sets the energy scale. Symbols: stationary points. Red: Mott insulator, green and blue: metal. Fat symbols: thermodynamically stable phase. The first-order Mott transition is marked by an arrow.* Right: *Sketch of the building block of the reference system. Blue filled dots: correlated sites with $U > 0$. Red open dots: bath sites with $U = 0$. Calculations for $t' = t$, $t'' = 0$, arbitrary V. (adapted from Ref. [58]).*

dots). The respective stationary point (a minimum) of the SFT grand potential is found at a comparatively low value of V_{opt} and can be traced with decreasing U down to another critical value $U_{c1} \approx 4.6$. For $U < U_{c1}$, there is no insulating phase.

It is interesting to observe that in the regime $U_{c1} < U < U_{c2}$ the metallic and the insulating phase are coexisting, i.e. that there are two stationary points of the grand potential. There is actually a third stationary point in the coexistence region, indicated by the green dots, where the SFT grand potential is at a maximum. Note that any stationary point, minimum, maximum or saddle point (in higher-dimensional parameter spaces), must be considered as an admissible solution within the SFT. However, the grand potential of the third phase is always higher than the grand potentials of the other phases. It therefore describes a physically irrelevant metastable phase but mathematically explains why the other two phases cease to exist above (below) a certain critical interaction U_{c2} (U_{c1}).

For a given U, the phase with the lowest grand potential is thermodynamically stable (see fat symbols). This means that the system is a correlated metal for $U < U_c$ and a Mott insulator for $U > U_c$ where the critical value for the Mott transition $U_c \approx 5.8$ is given by the interaction strength for which the metal and insulator have the same grand potential (the same ground-state energy at zero temperature), see the arrow in Fig. 10. Therefore, at U_c the optimal hopping parameter $V_{\rm opt}$ jumps between the large metallic and the small insulating value. Consequently, the ground state and thus the self-energy of the reference system changes abruptly at U_c. This leads to a discontinuous change of the SFT Green's function as well as of all observables that are computed as derivatives of the (optimized) SFT grand potential. The phase transition is of *first order* or *discontinuous*. This is interesting since the Mott transition in the Hubbard model on an infinite-dimensional, e.g. hyper-cubic, lattice is known [35] to be of second order or continuous. The VCA calculation discussed here actually corrects a mean-field artifact that is due to the neglect of non-local short-range antiferromagnetic correlations (see Ref. [58] for an extended discussion). For a discussion of more recent developments see Ref. [60], for example.

7   Relation to other methods and conclusions

Concluding, it is an appealing idea to divide a correlated lattice-fermion problem into small isolated clusters for which the problem can be solved easily and in a second step to employ the decoupled-cluster solution to construct the solution for the original lattice model. This construction must be approximate and has the spirit of a cluster mean-field theory, where the intra-cluster correlations are treated in a much better and more explicit way than the inter-cluster correlations. We have learned that this construction cannot be based on a technique that uses many-body wave-functions, rather, one has to employ Green's functions, namely the single-particle Green's function or, equivalently, the self-energy, as is done with the cluster-perturbation theory.

The self-energy-functional theory conceptually improves the CPT in several respects: first, it removes the arbitrariness that is inherent to the CPT regarding the choice of the cluster parameters. Second, it introduces an element of self-consistency or variational character by which it becomes possible to study phases that possess a symmetry different from the symmetry of the isolated cluster, i.e. one can address spontaneously symmetry-broken phases (collective magnetism, superconductivity etc.). Third, the SFT provides us with an explicit, though approximate, expression for a thermodynamical potential from which all observables have to be derived. This ensures that the approach is consistent in itself and obeys general thermodynamical relations, an important point that is missing in the plain CPT as well.

The self-energy-functional theory should actually be seen as a theoretical frame that allows the construction of different approximations. Each approximation is characterized by the choice of a corresponding reference system. Typically, this consists of decoupled clusters with $L_{\rm c}$ correlated sites each and an additional $n_{\rm s} - 1$ uncorrelated bath sites. Large clusters are necessary in order to describe short-range correlations as well as possible, and a large number of local degrees of freedom $n_{\rm s}$ is recommended to improve the description of local temporal correlations.
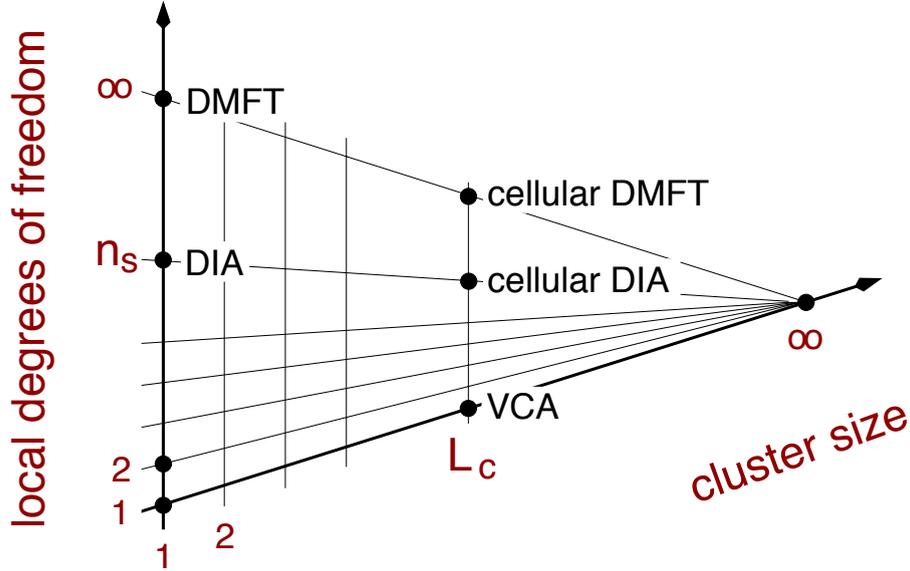
Fig. 11: *Schematic picture for the relation between different approximations that can be constructed within the self-energy-functional theory. See text for discussion.*

There is the ubiquitous tradeoff between the quality of the approximation on the one hand and the numerical effort on the other as the problem must be exactly solved for the isolated cluster. Using an exact-diagonalization solver the effort increases roughly exponentially with L_c and n_s. The space of possible approximations spanned by L_c and n_s is sketched in Fig. 11.

The most simple approximation is given by $L_c = 1$ and $n_s = 1$. Here, one approximates the self-energy of the lattice model by the self-energy of the atomic problem. This is in the spirit of the Hubbard-I approximation [5]. For $L_c > 1$ we find the variational cluster approximation that we have discussed at length here. Obviously, one would recover the exact solution of the lattice model in the limit $L_c \to \infty$.

Choosing a "cluster" with a single correlated site only, $L_c = 1$, but introducing a number of bath sites $n_s - 1 \geq 1$ specifies another approximation which is called dynamical impurity approximation (DIA). This is a true mean-field approximation as all non-local two-particle spatial correlations are neglected in this case.

Obviously, an ideal embedding of a single site into the lattice or an ideal mean-field theory is realized with an infinite number of bath sites $n_s \to \infty$. In this case, all local, temporal correlations are treated exactly – as opposed to a *static* mean-field theory like the Hartree-Fock approach. This optimal mean-field theory turns out to be identical with the well-known dynamical mean-field theory (DMFT) [33, 35]. In fact, one immediately recognizes that the reference system in this case is just the single-impurity Anderson model (if one starts out from a single-band Hubbard model, for example). Within DMFT the parameters of this impurity model are fixed by imposing the so-called DMFT self-consistency condition, stating that the impurity Green's function of the impurity model should be equal to the local Green's function of the lattice model. This is realized by setting up a self-consistency scheme that requires an iterated solution of the impurity model. Within the SFT, on the other hand, the parameters of

the reference impurity system are fixed by the SFT Euler equation which is the condition for the stationarity of the SFT grand potential. In fact both the DMFT self-consistency equation and the SFT Euler equation are basically equivalent in this case. From Eq. (58), we have

$$\frac{\partial}{\partial \boldsymbol{t}'} \, \Omega[\boldsymbol{\Sigma}_{\boldsymbol{t}'}] = \frac{1}{\beta} \sum_{n,ij\sigma} \left(\frac{1}{\boldsymbol{G}_0^{-1}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n)} - \boldsymbol{G}'(i\omega_n) \right)_{ij\sigma} \frac{\partial \Sigma_{ji\sigma}(i\omega_n)}{\partial \boldsymbol{t}'} = 0 \, , \qquad (71)$$

where $\boldsymbol{G}'$ denotes the Green's function of the reference system. As the self-energy of a single-impurity Anderson model is local, $\Sigma_{ij\sigma}(i\omega_n) = \delta_{ij}\Sigma_{i\sigma}(i\omega_n)$ and non-zero on the correlated impurity site only. This SFT Euler equation is satisfied if the *impurity* Green's function $G'_{\mathrm{imp}}(i\omega_n)$ is equal to the (approximate) *local* Green's function of the lattice model, i.e. if the local elements of the bracket vanish. This is the DMFT self-consistency equation.

What happens if $n_{\mathrm{s}} < \infty$? For a small number of bath sites, an exact solution of the reference system by means of exact-diagonalization techniques becomes feasible in practice. The resulting approximation, the DIA, differs from DMFT as the bracket in Eq. (71) will never vanish in this case. This is easily seen by noting that the Green's function of an $n_{\mathrm{s}} < \infty$ impurity model has a finite number of poles on the real frequency axis while the approximate lattice Green's function inherits its analytical structure from the non-interacting lattice Green's function $\boldsymbol{G}_0$, which, for an infinite lattice, may exhibit isolated poles but must have branch cuts as well. The DIA is different and actually inferior compared with the DMFT but does not need an advanced solver (note that solving the single-impurity Anderson model with $n_{\mathrm{s}} = \infty$ is still a demanding many-body problem). It has turned out, however, that with a few bath sites only, one often has a rather reliable approach to study the thermodynamics of a lattice model (see Ref. [43], for an example). This is also known from the exact-diagonalization approach to DMFT [61]. DMFT-ED considers the impurity model with a small n_{s}, as is done in the DIA, but employs another, actually more *ad hoc* condition to fix the parameters of the impurity model, namely one minimizes a suitably defined distance between the two local elements of the Green's function in the DMFT self-consistency condition to satisfy this at least approximately. The approach is able to yield similar results as the DIA in practice and is more easily implemented numerically but lacks internal consistency.

The third example discussed in the previous section has shown that one can favorably make use of approximations where a small number of bath degrees of freedom $n_{\mathrm{s}} > 1$ are combined with a cluster approach $L_{\mathrm{c}} > 1$. This approach is a VCA with additional bath sites and may be termed "cellular DIA" (see Fig. 11) since it is related to the cellular DMFT [62, 63] in the same way as the DIA is related to the DMFT. Note that with increasing cluster size $L_{\mathrm{c}} \to \infty$, all approaches, the VCA, the cellular DIA as well as the cellular DMFT must recover the exact solution of the lattice problem in principle.

Another prominent and widely used cluster mean-field theory is the dynamical cluster approximation (DCA) [64]. As compared to the cellular DMFT, this is a cluster extension of dynamical mean-field theory that avoids one of the main drawbacks of various cluster approaches, namely the artificial breaking of the translational symmetries. Already the CPT yields an approximate Green's function that merely reflects the translational symmetries of the "superlattice," which

periodically repeats the basic cluster, rather than the symmetries of the underlying physical lattice. In contrast, the DCA provides a Green's function and a self-energy with the correct symmetries but, on the other hand, must tolerate that the self-energy is discontinuous as a function of k in reciprocal space. Here, we briefly mention that it is possible to re-derive the DCA within in the framework of the SFT as well. This is carried out in detail in Ref. [14]. It is based on the idea that, with a proper modification of the hopping parameters of the *original* lattice model, $t \mapsto \widetilde{t}$, which becomes irrelevant for $L_c \to \infty$, the DCA becomes equivalent to the cellular DMFT. In a similar way [9] another cluster-mean-field variant, the periodized cellular DMFT [65] can be re-derived within the SFT.

To summarize, the self-energy-functional approach not only recovers a number of well-known mean-field and cluster mean-field concepts and provides a unified theoretical framework to classify the different approaches but has also initiated the construction of new non-perturbative and consistent approximations, the most prominent example of which is the variational cluster approximation. The challenges for future developments are manifold; let us mention only two directions here. The first consists in the generalization of the SFT to many-body lattice models far away from thermal equilibrium [15]. This requires a reformulation of the theory in terms of non-equilibrium Green's functions but offers the exciting perspective of studying the real-time dynamics of strongly correlated systems in a non-perturbative and consistent way. Another equally important direction of future work consists in an extension of the theory to correlated lattice models with non-local and even long-ranged interactions. First promising steps have already been made [8]. The restriction to local Hubbard-type interactions, inherent to all of the approaches mentioned here (see Fig. 11), represents an eventually unacceptable model assumption that must be abandoned.

Acknowledgement

References

[1] M. Potthoff, Euro. Phys. J. B **32**, 429 (2003)

[2] M. Potthoff, M. Aichhorn, and C. Dahnken, Phys. Rev. Lett. **91**, 206402 (2003)

[3] M. Potthoff, AIP Conf. Proc. **1419**, 199 (2011)

[4] M. Potthoff in *Strongly Correlated Systems: Theoretical Methods*, p. 303, ed. by A. Avella
   and F. Mancini, Springer Series in Solid-State Sciences, Vol. 171 (Springer, Berlin, 2012)

[5] J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963)

[6] M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)

[7] J. Kanamori, Prog. Theor. Phys. (Kyoto) **30**, 275 (1963)

[8] N.-H. Tong, Phys. Rev. B **72**, 115104 (2005)

[9] W. Koller and N. Dupuis, J. Phys.: Condens. Matter **18**, 9525 (2005)

[10] M. Knap, E. Arrigoni, and W. von der Linden, Phys. Rev. B **81**, 235122 (2010)

[11] M. Aichhorn, M. Hohenadler, C. Tahan, and P. B. Littlewood,
    Phys. Rev. Lett. **100**, 216401 (2008)

[12] M. Knap, E. Arrigoni, and W. von der Linden, Phys. Rev. B **81**, 104303 (2010)

[13] W. Koller, D. Meyer, Y. Ono, and A. C. Hewson, Europhys. Lett. **66**, 559 (2004)

[14] M. Potthoff and M. Balzer, Phys. Rev. B **75**, 125112 (2007)

[15] F. Hofmann, M. Eckstein, E. Arrigoni, and M. Potthoff, Phys. Rev. B **88**, 165124 (2013)

[16] C. Dahnken, M. Aichhorn, W. Hanke, E. Arrigoni, and M. Potthoff,
    Phys. Rev. B **70**, 245110 (2004)

[17] C. Gros and R. Valenti, Phys. Rev. B **48**, 418 (1993)

[18] D. Sénéchal, D. Pérez, and M. Pioro-Ladrière, Phys. Rev. Lett. **84**, 522 (2000)

[19] D. Sénéchal in *Strongly Correlated Systems: Theoretical Methods*, p. 237, ed. by A. Avella
    and F. Mancini, Springer Series in Solid-State Sciences, Vol. 171 (Springer, Berlin, 2012)

[20] E. Koch: *The Lanczos Method* in
    E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein (eds.):
    The LDA+DMFT approach to strongly correlated materials
    Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011)
    http://www.cond-mat.de/events/correl11

[21] R.P. Feynman, Phys. Rev. **56**, 340 (1939)

[22] J.W. Gibbs: in: *The collected works of J. Willard Gibbs*, Vol. II, Chap. XI, pp. 129
(Yale University Press, New Haven, 1948)

[23] R.P. Feynman, Phys. Rev. **97**, 660 (1955)

[24] J.M. Luttinger and J. C. Ward, Phys. Rev. **118**, 1417 (1960)

[25] A.A. Abrikosow, L.P. Gorkov, and I.E. Dzyaloshinski: *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, New Jersey, 1964)

[26] A.L. Fetter and J.D. Walecka: *Quantum Theory of Many-Particle Systems*
(McGraw-Hill, New York, 1971)

[27] J.W. Negele and H. Orland: *Quantum Many-Particle Systems*
(Addison-Wesley, Redwood City, 1988)

[28] M. Balzer and M. Potthoff, Phys. Rev. B **83**, 195132 (2011)

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

[30] H. Schweitzer and G. Czycholl, Solid State Commun. **74**, 735 (1990)

[31] H. Schweitzer and G. Czycholl, Z. Phys. B **83**, 93 (1991)

[32] M. Potthoff and W. Nolting, Z. Phys. B **104**, 265 (1997)

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

[34] E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)

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

[36] G. Kotliar and D. Vollhardt, Physics Today, March 2004, p. 53

[37] R. Eder: *The Variational Cluster Approximation* in
E. Pavarini, E. Koch, and U. Schollwöck (eds.):
Emergent Phenomena in Correlated Matter
Reihe Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
`http://www.cond-mat.de/events/correl13`

[38] M. Potthoff, Condens. Mat. Phys. **9**, 557 (2006)

[39] G. Baym and L.P. Kadanoff, Phys. Rev. **124**, 287 (1961)

[40] G. Baym, Phys. Rev. **127**, 1391 (1962)

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

[42] M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, Phys. Rev. B **74**, 235117 (2006)

[43] M. Potthoff, Euro. Phys. J. B **36**, 335 (2003)

[44] M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, Phys. Rev. B **74**, 024508 (2006)

[45] R. Freund: *Band Lanczos method,* in *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, ed. by Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst (SIAM, Philadelphia, 2000)

[46] M. Potthoff in *Effective models for low-dimensional strongly correlated systems*, ed. by G. Batrouni and D. Poilblanc (AIP Conf. Proc., Melville, 2006)

[47] W. Press, S.A. Teukolsky, W.T. Vetterling, and B. Flannery: *Numerical Recipes* (Cambridge University Press, 3rd. ed., 2007)

[48] R. Eder, Phys. Rev. B **76**, 241103(R) (2007)

[49] R. Eder, Phys. Rev. B **78**, 115111 (2008)

[50] L. Chioncel, H. Allmaier, E. Arrigoni, A. Yamasaki, M. Daghofer, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **75**, 140406 (2007)

[51] R. Eder, Phys. Rev. B **81**, 035101 (2010)

[52] M. Aichhorn, T. Saha-Dasgupta, R. Valenti, S. Glawion, M. Sing, and R. Claessen, Phys. Rev. B **80**, 115129 (2009)

[53] H. Allmaier, L. Chioncel, and E. Arrigoni, Phys. Rev. B **79**, 235126 (2009)

[54] H. Allmaier, L. Chioncel, E. Arrigoni, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **81**, 054422 (2010)

[55] M. Balzer, W. Hanke, and M. Potthoff, Phys. Rev. B **77**, 045133 (2008)

[56] E.H. Lieb and F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)

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

[58] M. Balzer, B. Kyung, D. Sénéchal, A.-M.S. Tremblay, and M. Potthoff, Europhys. Lett. **85**, 17002 (2009)

[59] F. Gebhard: *The Mott Metal-Insulator Transition* (Springer, Berlin, 1997)

[60] T. Schäfer, F. Geles, D. Rost, G. Rohringer, E. Arrigoni, K. Held, N. Blümer, M. Aichhorn, and A. Toschi, arXiv:1405.7250

[61] M. Caffarel and W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)

[62] G. Kotliar, S.Y. Savrasov, G. Pálsson, and G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)

[63] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **62**, R9283 (2000)

[64] M.H. Hettler, A.N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H.R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998)

[65] G. Biroli, O. Parcollet, and G. Kotliar, Phys. Rev. B **69**, 205108 (2004)