Quantum Science and Technology

TOPICAL REVIEW

Beyond quantum cluster theories: multiscale approaches for strongly correlated systems

Herbert F Fotso1,∗, Ka-Ming Tam2,3 and Juana Moreno2,3

1 Department of Physics, University at Albany (SUNY), Albany, New York 12222, United States of America
2 Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803, United States of America
3 Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, United States of America

∗Author to whom any correspondence should be addressed.

E-mail: herbert.fotso@gmail.com

Keywords: multiscale, strongly correlated systems, Hubbard model, computational methods, quantum Monte Carlo, Parquet formalism, quantum materials

Abstract

The degrees of freedom that confer to strongly correlated systems their many intriguing properties also render them fairly intractable through typical perturbative treatments. For this reason, the mechanisms responsible for their technologically promising properties remain mostly elusive. Computational approaches have played a major role in efforts to fill this void. In particular, dynamical mean field theory and its cluster extension, the dynamical cluster approximation have allowed significant progress. However, despite all the insightful results of these embedding schemes, computational constraints, such as the minus sign problem in quantum Monte Carlo (QMC), and the exponential growth of the Hilbert space in exact diagonalization (ED) methods, still limit the length scale within which correlations can be treated exactly in the formalism. A recent advance aiming to overcome these difficulties is the development of multiscale many body approaches whereby this challenge is addressed by introducing an intermediate length scale between the short length scale where correlations are treated exactly using a cluster solver such QMC or ED, and the long length scale where correlations are treated in a mean field manner. At this intermediate length scale correlations can be treated perturbatively. This is the essence of multiscale many-body methods. We will review various implementations of these multiscale many-body approaches, the results they have produced, and the outstanding challenges that should be addressed for further advances.

1. Introduction

Strongly correlated systems include some of the most technologically promising materials of our time. To harness their significant promise, understanding the fundamental mechanisms responsible for their intriguing properties is essential [1–5]. This understanding remains a challenge for the condensed matter community despite several decades of intense effort. For instance, although the discovery of high temperature superconductors dates back to 1987 [6], the underlying superconducting mechanism remains the subject of intense research activity. Following their discovery, the Hubbard model was postulated to contain the ingredients necessary to explain the properties of high temperature superconductors and their low-energy excitations [7]. But despite its simplicity, an exact solution of the Hubbard model beyond one dimension remains elusive [8, 9]. Therefore, numerical methods have played a crucial role. These methods are however constrained by the minus sign problem for quantum Monte Carlo (QMC), or by the exponential scaling of the Hilbert space for exact diagonalization (ED), to relatively small system sizes. Embedding schemes have emerged as an important avenue to treat the problem in the thermodynamic limit. These schemes map the lattice problem onto an impurity, for the case of dynamical mean field theory (DMFT), or onto a cluster for the case of its cluster extensions, dynamical cluster approximation (DCA) or cellular DMFT, embedded into a mean field [10, 11]. Embedding approaches allow the exact treatment of
short length scales at the level of the cluster or the impurity, and the treatment of longer length scales at the mean field level. Multiscale many body (MSMB) methods follow this logic to its natural next step by incorporating between the previous two length scales, an intermediate one at which correlations are treated with diagrammatic perturbation theory [12].

In general, the difficulty in understanding correlated systems lies in the fact that there are no simple theories to explain both the weak interaction limit of the metallic state and the strong interaction limit of the Mott insulating phase [13, 14]. The most successful theory of interacting fermions is the Fermi liquid theory [15–17]. The basic underlying assumption is that the interaction can be turned on adiabatically from the non-interacting free fermions limit. The consequence is that the quantum numbers of the non-interacting fermions remain unchanged. Electrons can be treated as quasi-particles in a rather stable state with a lifetime that becomes very long for those states near the Fermi level.

The Fermi liquid theory is a very efficient description of interacting fermions in a metallic phase [15–17]. It is applicable to almost all metallic phases, except for special circumstances such as the notable exception of one-dimensional systems. The theory owes its simplicity to being an effective renormalized single particle theory. Once the system is beyond the simple single particle description, there is no universal prescription to handle the competition or cooperation between the different degrees of freedom and the interplay between the kinetic and the potential energies. Precisely for this reason, numerical methods are often inevitable for practical calculations.

Widely used mean field methods factorize the interaction terms in the Hamiltonian to reduce the problem to an effective single particle theory in a static potential. The mean field, Hartree–Fock (HF), approximation often provides reasonable results [18–21], but its shortcomings are also obvious, in particular for intermediate interaction strengths where quantum fluctuations are large. The HF approximation quenches the quantum or temporal fluctuations completely. This may be a reasonable assumption if the interactions are overpowered by the kinetic energy terms. However, for many physical realizations of strongly correlated systems, perhaps the most well known one being the cuprate superconductors, the interaction is of the same order of magnitude as the bandwidth. Naïvely factorizing the interaction term to suppress all the quantum fluctuations is questionable at best. Indeed, there is currently no simple mean field theory that can explain most features of the cuprate superconductors.

Understanding the metallic phases beyond Fermi liquid theory is key for understanding broken symmetry phases, such as d-wave superconducting pairing in the cuprates. While one can construct a phase with no explicit broken symmetry and use the mean field method to understand the effective theory, this always involves fractionalized particles and strong constraints such as those of gauge theories [22].

Beyond mean field theory, there exists a plethora of techniques based on weak coupling expansion. They are typically based on low order perturbative methods, such as second order perturbation theory, or on selecting a certain class of diagrams and summing them up to infinite order. A typical example is the random phase approximation (RPA), which selects the class of ladder diagrams and sums them up to infinite order [23–26]. A more sophisticated approach is to sum a large class of diagrams in an iterative way. For example, parquet diagrams are generated when second order diagrams are inserted iteratively into the interaction vertex. This generates a class of diagrams that can only be separated into two disconnected pieces by cutting at least two fermion lines [27–29]. The advantage of the parquet approach compared to second order perturbation theory is in the ability to sum up a large variety of diagrams including those at infinite order. This, in principle, allows the instability towards a broken symmetry to be captured [30–32]. Its main advantage over RPA is in its unbiased sum of diagrams in different scattering channels to enable the study of the competition among different broken symmetries.

Instead of using a diagrammatic expansion approach, the DMFT maps the strongly correlated lattice onto an impurity site embedded in a self-consistently determined effective medium. The interest in the high spatial dimension limit of strongly correlated models in the late 80’s and early 90’s led to the understanding that in this limit, strongly correlated models with local interactions can be greatly simplified. This is due to the fact that an expansion in terms of the hopping amplitude in infinite dimension leads to the vanishing of all diagrams except the local ones, and, for a translationally invariant system, the model loses all spatial dependence. This simplification led to the DMFT.

DMFT remains the subject of active research efforts, particularly because there is no universal quantum impurity solver. Various methods have been proposed over the past few decades. These include semi-analytical methods based on perturbation theory or modified mean field theories. The more well known methods include the iterative perturbation theory and the local moment approximation. Numerical approaches include various kinds of QMC and ED methods. Recently, density matrix renormalization group and matrix product state methods have also been explored. Quantum computing algorithms for solving the quantum impurity problem have been proposed recently [33–36]. After all, solving even a single
impurity is a non-trivial problem, as the mean field hybridization function is not given by a simple form that can allow an analytical solution.

It is worth noting that the DMFT can be viewed as a formal generalization of the coherent potential approximation (CPA) proposed by Soven in the 60’s [37, 38]. The CPA has since been extensively used for studies of random disorder models with negligible interactions, in particular for random alloys. Various extensions of CPA have been proposed over the years. The earliest one that goes beyond the single site approximation is the molecular CPA [39, 40]. It embeds the cluster into an effective medium that possesses the same structure as the cluster itself. It thus generates an effective medium different from that of the CPA. The obvious deficiency of the method is the breaking of translational invariance.

A different scheme of embedding cluster methods is the DCA [41, 42]. By construction, the method naturally preserves the translational invariance of the original model by directly working with both the cluster and the effective medium in momentum space. The method has been extensively employed on the Hubbard model [43–45], Anderson model [46], periodic Anderson model [47], and Falicov–Kimball model [48]. The cluster extension is not just a quantitative improvement on the DMFT. It is necessary to produce important features that are absent in the DMFT results. Perhaps the most important one is the DCA’s ability to capture nonlocal correlations such as that of d-wave pairing, which is obviously absent for approximations that do not consider spatial dependence explicitly [49]. The method has also been considered in the context of multiple scattering theory where it is re-branded as non-linear CPA [50].

The difficulty of solving a cluster impurity (or embedded cluster) problem scales exponentially. Roughly speaking, QMC based methods scale exponentially with the number of impurity sites (cluster size), with inverse temperature, or with the interaction strength [51–54]. The exception is strong-coupling expansion based Monte Carlo methods, but this is usually limited to a rather small number of impurity sites [55, 56]. Another class of impurity/cluster solvers is based on diagonalization of the effective finite size Hamiltonian. For these Hamiltonian-based solvers, such as ED [57–62], the Hilbert space grows exponentially with the cluster size and thus both computing memory and time requirements grow at the same rate. This is also true to a large extent for another Hamiltonian-based approach, the numerical renormalization group method [63–66]. In general, for practical calculations, the maximum number of impurity sites is rather modest (∼10 sites).

Over the last couple of decades various novel methods have been proposed. These include the density matrix renormalization group [67, 68], the related matrix product wavefunction [69] and, even more recently, different forms of machine learning approaches [70, 71]. These more recent methods may have potential for certain applications. For instance, they may be more efficient for calculating real time Green functions in nonequilibrium problems [72]. Approaches based on machine learning could also be more efficient in solving a large set of impurity problems, and this may be useful for applications on random systems that require averaging over random disorder realizations [73]. However, none of these novel impurity solvers are suitable for the calculation of the vertex function which is essential for most methods that are built on an expansion on top of the DMFT solution. Additionally, Monte Carlo sampling of the partition function provides more flexibility for controlling the error as the impurity cluster size is increased. Also, although it has been proven that the single impurity problem does not exhibit a minus sign problem, the absence of minus sign in the Monte Carlo sampling cannot be assumed for a generic impurity problem [74].

Following the logic of embedding a small system into a mean-field host, one can anticipate better accuracy in the result if an intermediate length scale is inserted between the previous two. Since short length scales are appropriately treated with exact solvers and the long length scale by a mean field, this intermediate length scale can be treated reliably with diagrammatic methods. This is the essence of MSMB methods for strongly correlated systems as formulated in the early 2000’s [12] and the subject of continued efforts since then [75].

This review focuses exclusively on the methods for studying systems in equilibrium. It is noteworthy to point that effort has been devoted to the generalization of the MSMB approaches to non-equilibrium problems [76–78]. We refer the reader to the original article for details [76], as we do not discuss these nonequilibrium approaches in the present paper.

The rest of the review is structured as follows. In section 2, we summarize the DMFT method and its connection to the Anderson impurity problem by approaching it from its ‘cavity method’ formulation. In section 3, we discuss two cluster extensions of DMFT, DCA and cellular dynamical mean field theory (CDMFT). We proceed in section 4 with a discussion of the extended dynamical mean field theory (EDMFT) that extends DMFT to the treatment of nonlocal interaction. Section 5 is focused on the 1/d expansion, a systematic expansion of DMFT with respect to the hopping amplitude. In section 6, we describe the original formulation of the MSMB method. The parquet formalism, which encompasses various commonly used diagrammatic approximations, and which is essential for the diagrammatic
treatment of intermediate length scales, is described in section 7. Following this, we briefly describe different implementations to incorporate nonlocal corrections into the DMFT/DCA starting with the dynamical vertex approximation in section 8, and then the dual fermion method in section 9. In section 10, we present the dual boson extension for corrections to EDMFT. In section 11, we discuss efforts to incorporate nonlocal correction into the GW approximation (an approach that obtains the self-energy from the single particle Green function ($G$) and the screened Coulomb interaction ($W$)) in the form of the ‘triply irreducible local expansion’ (TRILEX). In section 12, we discuss functional renormalization group and its usage for nonlocal corrections to DMFT. In section 13, an important computational challenge, the numerical representation of the vertex functions in memory, is discussed. In section 14, we discuss important physical constraints on the methods. In section 15, we summarize results obtained on different models with various implementations of the multiscale many-body approach before ending with our conclusions.

2. Dynamical mean field theory

Following the discovery of high temperature superconductors, it was argued that the Hubbard model (1) captures their low energy properties [7]. This deceptively simple model describes itinerant electrons that can hop between nearest-neighbor sites $\langle i,j \rangle$ on a lattice with a hopping integral $t_{ij}$, and are subject to a Coulomb interaction $U$ when a site is doubly occupied. The model is schematically depicted in figure 1 and defined as:

$$\hat{H} = -\sum_{\langle i,j \rangle, \sigma} t_{ij} c_{i,\sigma}^\dagger c_{j,\sigma} + \sum_i U \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow},$$

where $c_{i,\sigma}$ ($c_{i,\sigma}^\dagger$) is the destruction (creation) operator that destroys (creates) an electron with spin $\sigma$ at site $i$. $\hat{n}_{i,\sigma} = c_{i,\sigma}^\dagger c_{i,\sigma}$ is the number of particles of spin $\sigma$ at site $i$.

The DMFT method in itself is a generalization of the usual mean field theory. But unlike the usual mean field theory, say for the Ising model for example, the mean field here is not an order parameter. Instead it is a function of time or frequency. Thus, the approach captures all temporal fluctuations. Focusing on the translationally invariant paramagnetic phase, there is no explicit order parameter. One may also construct an explicit order parameter to represent the broken symmetry, but this is not done routinely and is not the focus of this discussion.

Within the DMFT solution of a translationally invariant system, the spatial fluctuations are completely suppressed as in a traditional mean field theory. As we will see, a major way to improve the approximation is to systematically incorporate the corrections due to the spatial dependence of the model into the DMFT solution. Also, since the self-consistency condition is valid only at the single particle level, there is no guarantee that high order Green functions, such as the different susceptibilities, are matched between the lattice and impurity models.

When the DMFT was developed in the early 90’s [11, 79–86], solving even a single impurity problem was rather challenging. With the formulation of new numerical algorithms and advances in computing power, a single impurity problem can, in general, be numerically solved quite efficiently. While there is still not a completely satisfactory method that is accurate for a wide range of parameters, particularly for solving interacting problems with random disorder which requires a large ensemble of different impurity.

Figure 1. Schematic illustration of the Hubbard model on a square lattice. Electrons are allowed to hop between nearest neighboring sites with a hopping amplitude $t$, and the interaction of strength $U$ is local.
realizations for disorder averaging [73, 87, 88], the single impurity problem can mostly be handled at the present time.

As a side note, it is worth mentioning that earlier work on DMFT can be traced back to the study of the transverse-field Sherrington–Kirkpatrick model [89]. In that model the spatial fluctuations are completely suppressed as the model consists of spin couplings in the fully connected network with the transverse field. The physics of the Hubbard model via DMFT was discussed by Georges and Kotliar [83]. The first numerically ’exact’ DMFT solution of fermionic systems in the infinite dimensional external magnetic field, thus it can be mapped to a single site problem [90].

The DMFT formalism can be explained quite transparently from a path integral formulation. Consider the action of the Hubbard model on a lattice,

\[ S = - \sum_{r_i, \tau_j} \int_0^\beta \int_0^\beta d\tau_i d\tau_j \psi^*_\sigma(r_i, \tau_i) G_0^{-1}(r_i, \tau_i; r_i, \tau_j) \psi_\sigma(r_j, \tau_j) + U \sum_{r_i} \int_0^\beta d\tau_i \psi^*_\uparrow(r_i, \tau_i) \psi^*_\downarrow(r_i, \tau_i) \psi_\downarrow(r_i, \tau_i) \psi_\uparrow(r_i, \tau_i), \]  \tag{2}

where \( G_0 \) is the bare Green function, \( \psi^*_\sigma(r, \tau) \) and \( \psi_\sigma(r, \tau) \) are the Grassmann fields for spin \( \sigma \) at location \( r \) and imaginary time \( \tau \).

The first part of the action contains the kinetic energy as characterized by the bare Green function \( G_0 \). It is simply obtained from the bare dispersion of the considered model. The second term includes the interaction characterized by the parameter \( U \), which we assume to be local. Any interaction beyond the local Hubbard term will involve further approximations in the context of the DMFT.

The exact Green function of the above action can be completely characterized by the self-energy \( \Sigma \). If we write the self-energy in the frequency-momentum space, the relation between the bare Green function and the exact Green function \( G \) is given by the Dyson equation,

\[ G(k, \omega) = \frac{1}{G_0(k, \omega) - \Sigma(k, \omega)}. \]  \tag{3}

In the simplified case of a translationally invariant system, the idea of the DMFT is to relate the full lattice problem with spatial dependence to a single site problem. To this end, DMFT reframes (2) into an effective action for a single site with a bare Green’s function \( G \):

\[ S_{\text{eff}} = - \int d\omega \sum_\sigma \psi^*_\sigma(\omega) G^{-1}(\omega) \psi_\sigma(\omega) \]  \tag{4}

\[ + U \int_{\omega_1 + \omega_2 = \omega_3 + \omega_4} d\omega_1 d\omega_2 d\omega_3 d\omega_4 \psi^*_\uparrow(\omega_1) \psi^*_\downarrow(\omega_2) \psi_\downarrow(\omega_3) \psi_\uparrow(\omega_4). \]

Consider the Anderson impurity model characterizing an impurity coupled to a band of conducting electrons and given by the Hamiltonian:

\[ H_{\text{AIM}} = U n_i n_j - \mu \sum_\sigma \epsilon_\sigma c_\sigma^\dagger c_\sigma \]

\[ + \sum_{j \neq i} V_j (f_j^\dagger c_\sigma + c_\sigma f_j) \]  \tag{5}

\[ + \sum_{j \neq i} f_j^\dagger f_j + c_\sigma c_\sigma. \]

where \( c_\sigma^\dagger, c_\sigma \) are creation and destruction operators for the impurity electrons while \( f_j^\dagger, f_j \) are those of the conduction electrons. Its action is:

\[ S_{\text{AIM}} = - \int d\omega \sum_\sigma \psi^*_\sigma(\omega) G_{\text{AIM}}^{-1}(\omega) \psi_\sigma(\omega) \]  \tag{6}

\[ + U \int_{\omega_1 + \omega_2 = \omega_3 + \omega_4} d\omega_1 d\omega_2 d\omega_3 d\omega_4 \psi^*_\uparrow(\omega_1) \psi^*_\downarrow(\omega_2) \psi_\downarrow(\omega_3) \psi_\uparrow(\omega_4), \]

where \( G_{\text{AIM}} \) is the single impurity Anderson model non-interacting Green’s function defined by:

\[ G_{\text{AIM}}^{-1}(\omega_n) = i \omega_n + \mu - \Delta(\omega_n) \]  \tag{7}
with
\[ \Delta(i\omega_n) = \int_{-\infty}^{+\infty} \frac{d\omega}{i\omega_n - \omega} \sum_j V_j^2 \delta(\omega - \epsilon_j). \] (8)

The AIM action \( (6) \) is equivalent to \( (4) \) with \( \mathcal{G} \) playing the role of the non-interacting AIM Green’s function. The construction can be justified via the concept of the ‘cavity method’ in the infinite dimension limit whereby all degrees of freedom are integrated out except for the site labelled by the index 0. In this limit of \( d \to \infty \) for a hypercubic lattice, the hopping amplitude is rescaled as \( t_j = t/\sqrt{2d} \) so that the kinetic energy and the interaction energy remain of the same order. The effective action, \( S_{\text{eff}} \), in this process is defined by:

\[
\frac{1}{Z_{\text{eff}}} \exp(-S_{\text{eff}}[\psi^0_{\sigma,0}, \psi_{\sigma,0}]) = \frac{1}{Z} \int \prod_{j \neq 0, \sigma} \exp(-S[\psi^0_{\sigma,0}, \psi_{\sigma,0}])).
\] (9)

where \( Z \) and \( Z_{\text{eff}} \) are the partition functions associated with \( S \) and \( S_{\text{eff}} \) respectively. The effective action can be written as:

\[
S_{\text{eff}}[\psi^0_{\sigma,0}, \psi_{\sigma,0}] = S_0 + \sum_{\eta_{i,0}} \sum_{j=1}^N \eta_{i,0} \eta_{j,0} \eta_{j,0} \cdots \eta_{i,0} \eta_{j,0} \prod_{j=1}^N \int d\tau dG^{(0)}(\tau) \psi^0_0(\tau) + U \int d\tau \psi^*_0(\tau) \psi_0(\tau) + \int d\tau \psi^*_0(\tau) \psi_0(\tau).
\] (10)

where \( S_0 \) is the local action at site ‘0’:

\[
S_0 = -\int_0^\beta \int_0^\beta \int_0^\beta d\tau d\tau' \int d\tau'' \psi^0_0(\tau) G^{(0)}(\tau, \tau', \tau'').
\] (11)

\( G^{(0)} \) is the Green’s function connecting the cavity to the impurity, \( \eta_{i,0} = t_{i,0} \psi_{0,\sigma} \), with \( t_{i,0} \) the hopping from site \( i \) to \( 0 \).

Only terms of order \( n = 2 \) survive the expansion \( (10) \) in the \( d \to \infty \) limit. Leading to:

\[
S_{\text{eff}}[\psi^0_{\sigma,0}, \psi_{\sigma,0}] = S_0 + \sum_{i,j,1=1}^N t_{i,j} \psi^0_0(\tau) G(i, \tau_0, \tau_1, \tau_j).
\] (12)

Rewriting this in frequency space, gives the AIM action \( (6) \) with \( \mathcal{G}_{\text{AIM}} \) replaced by \( \mathcal{G} \) such that:

\[
\mathcal{G}^{-1}(i\omega_n) = i\omega_n + \mu - \sum_{j,0} t_{i,j} G(i\omega_n).
\] (13)

This relation connects the impurity Green function to the lattice Green function. For the Bethe lattice \([91]\), \( \mathcal{G}^{-1}(i\omega_n) = i\omega_n + \mu - t^2 G(i\omega_n) \). For a general lattice, the connection between the lattice Hubbard model and the single impurity model is established by setting the self-energy,

\[
\Sigma_{\text{lattice}}(k, \omega) = \Sigma_{\text{impurity}}(\omega).
\] (14)

Since the self-energy of the original lattice Hubbard model has spatial dependence while that of the single impurity Anderson model does not, to construct the lattice Green function one has to rely on the coarse-graining process that assumes the self-energy of the lattice model to be the same in the entire Brillouin zone.

\[
G_{\text{lattice}}(k, \omega) = \frac{1}{G_{\text{lattice},0}(k, \omega) - \Sigma(k, \omega)} \approx \frac{1}{G_{\text{lattice},0}(k, \omega) - \Sigma_{\text{lattice}}(\omega)}.
\] (15)

The missing link between the Hubbard model and the Anderson model is to determine the effective bare Green function of the Anderson model. For the \( D \)-dimension case, this is given by:

\[
G_{\text{impurity}}(\omega) = \sum_k \frac{1}{(2\pi)^D} G_{\text{lattice}}(k, \omega) = \sum_k \frac{1}{(2\pi)^D} \frac{1}{G_{\text{lattice},0}(k, \omega) - \Sigma_{\text{lattice}}(\omega)}.
\] (16)

The summation over the momentum can thus be replaced by an integral over the bare density of states to simplify the calculation. The bare density of states for the hypercubic lattice at the limit of infinite spatial dimensions can be exactly calculated \([81, 82]\). We have gathered all the ingredients for the DMFT approximation and the algorithm can be summarized as in figure \( 2 \).
Figure 2. DMFT algorithm: step 1. The impurity solver provides the impurity Green function, $G^{\text{imp}}$. Step 2. The impurity Dyson equation is used to extract the self-energy $\Sigma$. Step 3. The self-energy is coarse-grained over the entire first Brillouin zone to obtain the lattice Green function. Step 4. The bath Green function for the impurity problem is calculated and fed into the impurity solver. Step 5. Repeat steps 1 to 4 until convergence is obtained for the self-energy or Green function.

3. Cluster route for extending the DMFT

A natural and direct avenue to generalize the DMFT is to incorporate nonlocal correlations by including more than a single impurity site, i.e. by formulating the theory around a cluster of multiple sites in a self-consistently determined host. This type of cluster DMFT remains an important method for the study of strongly correlated systems, as it allows, by increasing the cluster size, a systematic correction unlike perturbative expansion methods. Moreover, one can envision that a perturbative expansion on top of the cluster method would produce an even better result, since the bare effective Hamiltonian or action for the perturbative expansion, which corresponds to the DCA or CDMFT solution on the smaller system, is presumably more accurate and already includes a substantial amount of nonlocal correlations.

For the classical spin model, the first attempt of a multiple site mean field theory was the so-called Bethe–Peierls–Weiss approximation using a cluster of $(z + 1)$ sites, with one site at the center surrounded by $z$ sites on the shell [91–94]. The interaction between the center spin and its $z$ nearest neighbor spins is treated explicitly while the interaction between the remaining $z$ spins with the other spins outside their own cluster is treated by a mean field.

Another approach by Oguchi is a more direct generalisation of the mean field method [95]. A cluster of $N_c$ spins is considered. The interaction among these $N_c$ spins is treated explicitly, and the interaction between the spins at the edge of the cluster and spins outside of the cluster is treated by a mean field. Unlike the Bethe–Peierls–Weiss method where pairwise interactions are treated explicitly only for pairs involving the central spin, pairwise interactions among all spins of the $N_c$ spin cluster are treated explicitly in the Oguchi approach.

A cluster of impurities in real space is considered instead of a single one for the cellular dynamical mean field theory [96, 97]. There is a technical problem with using such an approach for the paramagnetic solution of the Hubbard model as the cluster naturally breaks translational invariance. A procedure for restoring the symmetry is needed for a periodic solution [98].

A further approach for a cluster generalization of DMFT is based on the idea of coarse-graining that is central to the DCA [41, 99–101]. A cluster of impurities is used, but after the local cluster is solved the lattice quantities are averaged over different patches of the first Brillouin zone for a coarse-grained quantity. The advantage is that the method is manifestly translationally invariant. This allows a perturbative expansion to be implemented on top of the DCA solution more naturally. As we will discuss, almost all perturbative expansion methods become simpler and less cumbersome to implement when formulated in momentum space.

One can consider that the DMFT impurity bare Green function is the coarse-grained lattice Green function. Since the self-energy which is obtained by solving the impurity model does not have spatial dependence, the coarse-graining procedure for DMFT is done over the entire first Brillouin zone with one single impurity site. Effectively, from a diagrammatic point of view, DMFT neglects momentum conservation at the internal vertices. The DCA systematically restores momentum conservation at these internal vertices. To this end, it divides the Brillouin zone into $N_c$ cells with each cell (of linear size $\Delta k$) represented by a cluster momentum $\mathbf{K}$ in the center of the cell. The DCA then requires that momentum conservation in the internal vertices be respected for momentum transfers between cells (momentum
Figure 3. Illustration of the coarse-graining process in DCA for an eight-site cluster. The momentum $\mathbf{k}$ is mapped onto the nearest cluster point $\mathbf{K}$ so that $\mathbf{\tilde{k}} = \mathbf{k} - \mathbf{K}$ remains inside the cell containing $\mathbf{K}$.

Figure 4. DCA algorithm: 1. The impurity solver provides the impurity Green function, $G_{\text{imp}}(\mathbf{K})$. 2. The Dyson equation is used to extract the self-energy $\Sigma(\mathbf{K})$. 3. The self-energy is coarse-grained over the patches that fill the first Brillouin zone to obtain the lattice Green function, $G(\mathbf{K})$. 4. The bath Green function, $G_b(\mathbf{K})$, for the impurity problem is fed into the impurity solver. 5. Repeat from step 1 until convergence is obtained for the self-energy or Green function.

transfers larger than $\Delta k$), but neglected for momentum transfers within a cell (less than $\Delta k$). In this way momentum conservation is fully recovered in the limit of $N_c \to \infty$, while the DMFT result is obtained for $N_c = 1$.

The DCA coarse-graining process is illustrated by figure 3 for $N_c = 8$. In DCA, the self-energy is no longer momentum-independent. Rather, we have for a lattice momentum $\mathbf{k}$, $\Sigma(\mathbf{k}) = \Sigma(\mathbf{K})$. Where $\mathbf{K}$ is the momentum at the center of the cell containing $\mathbf{k}$. The impurity, or in this case the cluster, Green’s function is here related to the lattice Green’s function by:

$$G_{\text{imp}}(\mathbf{K}, \omega) \approx \sum_k \frac{1}{(2\pi)^D} G_{\text{lattice}}(\mathbf{k}, \omega) = \sum_k \frac{1}{(2\pi)^D} G_{\text{lattice}}^{-1}(\mathbf{k}, \omega) \frac{1}{\Sigma_{\text{lattice}}(\mathbf{K}, \omega)},$$

where the summation over $\mathbf{k}$ is restricted to the patch corresponding to the impurity/cluster site with momentum $\mathbf{K}$. The impurity solver is now a cluster solver for the momentum-dependent self-energy $\Sigma(\mathbf{K})$.

The algorithm for DCA is summarized in figure 4.
4. Extended dynamical mean field theory (EDMFT)

The earliest attempt to include the effect of nonlocal interactions was motivated by the competition between the local interaction and the RKKY interaction in heavy fermion materials [102]. The basic idea was to include the density-density and/or spin–spin interaction term and scale the interaction strength with respect to the spatial dimensionality so that its fluctuations are non-zero in the high dimension limit [102–107]. This procedure leads to an impurity embedded in a self-consistent fermionic bath and, at the same time, a self-consistent bosonic bath due to the nonlocal interactions.

In DMFT, the hopping is considered as a function of spatial dimension \( d \), \( t_{ij} \sim 1/\sqrt{d^{j-\beta}} \) in the infinite dimension limit. The EDMFT includes the nonlocal interaction in a similar manner: \( V_{ij} \sim 1/\sqrt{d^{j-\beta}} \) [102–104]. Here, we only consider the density–density interaction. A generic two-body interaction can include spin–spin, correlated hopping, and pair hopping. For example, if the nonlocal interaction is on-site Hubbard interaction. We note that the decoupling of the interaction by Hubbard–Stratonovich fields is not unique: one can decouple the coupling in different ways [108]. The decoupling presented above only decouples the density–density interaction, the \( V \) term in the extended Hubbard model.

The algorithm is thus similar to that of the DMFT, except that at each iteration both the mean fields for the fermionic and for the bosonic baths have to be updated. The algorithm can be summarized as in figure 5.

5. 1/\( d \) expansion

One of the earliest attempts to consider the effect of finite dimension and move away from the infinite dimension limit is the systematic expansion with respect to the hopping amplitude. This approach has been studied in other contexts, especially in the closely related CPA for disordered systems [109–113]. Since the hopping term carries the factor of \( 1/\sqrt{d} \), an expansion with respect to the hopping regains the dimensional dependence at the expense of including multiple sites. Work by Schiller and Ingersent studied the case of two impurities for the Falicov–Kimball model [114]. The approach has also been applied to the half-filled Hubbard model [115].

The 1/\( d \) corrections can be recovered by considering a multi-impurity problem. For a single impurity \(( n = 1)\) or a two-impurity \(( n = 2)\) problem,

\[
S^{(n)} = \sum_{\sigma} \sum_{\alpha_i=1}^{n} \int d\tau \int d\tau' \sum_{\sigma} \psi^*_{\sigma \alpha_i}(\tau) G_{\sigma \alpha_i}^{-1}(\tau) \psi_{\sigma \alpha_i}(\tau') - U \int d\tau n_{\sigma \alpha_i}(\tau) n_{\sigma \alpha_i}(\tau) + \sum_{\tau} U \int d\tau n_{\alpha \beta}(\tau) n_{\alpha \beta}(\tau),
\]

(21)
Figure 5. EDMFT algorithm: 1. The impurity solver provides the impurity Green function for the electrons, $G_{\text{imp}}$, and the impurity Green function for the Bosonic field, $D_{\text{imp}}$. 2. Dyson equation is used to extract the self-energy $\Sigma$ and the polarization, $\Pi$. 3. The self-energy and the polarization is coarse-grained over the entire first Brillouin zone to obtain the lattice Green functions. 4. The bath Green function for the electrons and the Bosonic field of the impurity problem is feed into the impurity solver. 5. Repeat from step 1 until convergence is obtained for the self-energy and polarization or Green functions.

where $\alpha$ and $\beta$ label the sites for $n = 2$ [114, 115]. The mean fields $G^{(1)}$ and $G^{(2)}$ are chosen in such a way that the impurity Green functions $G^{(1)}$ and $G^{(2)}$ coincide with the full on-site and nearest neighbor lattice propagators, $G_{\text{lattice}}^{(1)}$ and $G_{\text{lattice}}^{(2)}$:

$$G^{(1)} = G_{11}^{(2)} = G_{\text{lattice}}^{(1)}$$

(22)

For skeleton diagrams (diagrams without self-energy dressing or vertex correction) of order $O(1/d)$, the impurity self energies $\Sigma^{(1)}$ and $\Sigma^{(2)}_{\alpha \beta}$ and the diagonal and off-diagonal lattice self energies, $\Sigma_{0}$ and $\Sigma_{1}$ are related by [114, 115].

$$\Sigma_{\text{lattice}}^{(1)} = \Sigma^{(1)} + 2d(\Sigma_{11}^{(2)} - \Sigma^{(1)}_{11})$$

(23)

$$\Sigma_{\text{lattice}}^{(2)} = \Sigma_{12}^{(2)}$$

(24)

The lattice Green function can be obtained from the lattice self-energy:

$$G_{\text{lattice}}^{(lm)}(i\omega) = \frac{1}{1 + \sqrt{d} \Sigma_{\text{lattice}}^{(1)} G_{\text{lattice}}^{(lm)}(i\omega)}$$

(25)

where $G_{\text{lattice}}^{(lm)}(z)$ and $G_{\text{lattice}}^{(lm)}(z)$ denote the unperturbed and dressed lattice propagators between sites $l$ and $m$, respectively [114, 115].

Among the shortcomings of the $1/d$ expansion are its limitations in the description of long range fluctuations. Perhaps more importantly, the truncation at finite order may lead to non-analytic properties of some dynamical quantities [116]. The method is an example of the nested cluster scheme and is also related to the recently proposed self-energy embedding theory [98, 117, 118].

6. Multiscale many body

Given the challenge of solving a large cluster by numerical methods and the lack of accurate analytical approaches, perturbation theory is a possible route for the exploration of physics beyond local theories. Initial efforts aimed to use perturbative methods as solvers for the cluster impurity problem. Notably, the fluctuation exchange method was used as an impurity cluster solver [119]. The fluctuation-exchange or FLEX is one of the simplest methods to incorporate correlations among different channels. It is, therefore, a conceptually appealing approach for systems in which particle–hole or particle–particle vertex fluctuations are not small. If correlations of the two particle fluctuations are ignored, one obtains the second order perturbation theory for which the irreducible vertex is replaced by the bare vertex. FLEX allows vertex contributions from different channels to be correlated, thus the name of fluctuation exchange approximation.

The general scheme of the MSMB method as envisioned by Jarrell and collaborators is to construct a theory which allows for the treatment of different length scales by different approaches [12, 120]. The short length scale is addressed by some highly accurate numerical approach, such as QMC. The long length scale
Figure 6. Conceptual idea of the multiscale many body approach. The original lattice model is mapped onto a small cluster (yellow–orange–red colors), embedded in a larger cluster (rainbow colors) embedded in a mean field (orange color). The information passed between the clusters and the effective medium is composed of irreducible quantities like the self-energy and the fully irreducible vertex function.

Figure 7. Algorithm for the multiscale many body method. 1. The fully irreducible vertex ($\Lambda$) and optionally the self-energy ($\Sigma$) are obtained from the DMFT or DCA solver. 2. The parquet method or some other perturbative method is used to calculate the full vertex to restore spatial dependence or enhance spatial resolutions in the full vertex ($F$). 3. The full vertex is fed into Schwinger Dyson equation and Dyson equation to update the self-energy and the Green function respectively. 4. The Green function can be coarse-grained to be fed back to the DMFT or DCA solver as the bath Green function for self-consistent calculations.

is treated at a mean field level, while the intermediate length scale is treated by some form of perturbative technique.

An early proposition was to supplement the QMC calculation on small cluster sizes with FLEX on larger clusters via a self-energy self-consistency scheme between the two methods [121]. It was generalized in 2006 to address short and long length scales within the DCA formalism while the intermediate length scale would be treated by the parquet formalism [12]. In this approach, the connection between the two methods is established by the fully irreducible vertex from the DCA calculation on the cluster. The promise of the approach stems from the better scaling of the parquet formalism compared to that of QMC as a cluster solver (figure 6).

The general construction of the MSMB method, as summarized in figure 7, is obviously rather general. There is in fact plenty of freedom in the choice of a solver for the intermediate length scale. Indeed, this area of research has been the subject of significant activity over the past decade or so. Nevertheless, most of the developed methods are based on some simplification of the original proposal, either in picking a certain subclass of diagrams from the parquet formalism or in simplifying the solution by numerical techniques.

7. Diagrammatic methods and the parquet formalism

A major route for perturbative expansions around the DMFT solution is based on the parquet diagrams. This approach encapsulates many of the approximations that have emerged in this field, including the dynamical vertex approximation. For this reason, we review the parquet method in rather self-contained detail in this section.

Standard diagrammatic perturbative expansions attempt to describe all the scattering processes as single or two-particle Feynman diagrams. In the single-particle formulation the self-energy describes the many-body processes that renormalize the motion of a particle in the interacting background of all the other particles. In the two-particle context, one is able to probe the interactions between particles using the...
so-called vertex functions, which are matrices describing two particle scattering processes. For example, the reducible (full) two-particle vertex $F^{\text{ph}}(12; 34)$ describes the scattering amplitude of a particle–hole pair from its initial state $|3, 4\rangle$ into the final state $|1, 2\rangle$. Here, $i = 1, 2, 3, 4$ represents a set of indices which combines the momentum $k_i$, the Matsubara frequency $i\omega_n$, and, if needed, the spin $\sigma_i$ and band index $m_i$. Since the total momentum and energy of the vertex are conserved, it is convenient to adopt the notation $F^{\text{ph}}(2–4)_{1,3}$ for the numerical implementation on the single band Hubbard model. Other representations are also possible [122].

In general, depending on how particles or holes are involved in the scattering processes, one can define three different two-particle scattering channels. These are the particle–hole (p–h) horizontal channel, the p–h vertical channel and the particle–particle (p–p) channel. The parquet formalism is, in essence, a method for summing up diagrams that characterize scattering processes at the two-particle level. From another perspective, the diagrams are generated by inserting the one loop, second order, diagrams repeatedly into itself. Without channel mixing, this is equivalent to the RPA. With the mixing of three channels, this becomes the parquet formalism.

The vertices can be categorized by extending the notion of diagram reducibility to the two-particle level as illustrated by figure 8. At the one particle level, a diagram is said to be reducible if it can be split in two disconnected parts by breaking a single Green’s function line. A two-particle diagram will be said to be irreducible if it cannot be separated in two disconnected parts by breaking two Green function lines in the same channel [30]. It will be said to be fully irreducible if it cannot be separated in two disconnected parts by breaking two-Green’s function lines in any channel. In the single particle formalism, the Green function is related to the self-energy containing all single-particle irreducible diagrams by the Dyson equation. A connection is made between the single-particle and the two-particle diagrams by the Schwinger–Dyson equation that connects the self-energy to the (full/reducible) vertex $F$ containing all allowed diagrams in a given channel. The subset of all two-level diagrams in the full vertex that are irreducible in the same channel is known as the irreducible vertex $\Gamma$. The subset of irreducible vertices that are irreducible in any channel is called the fully irreducible vertex $\Lambda$.

It is worth noting that the above idea for the decomposition is not unique. One can devise other possible decompositions. For example, a recent attempt is to decompose diagrams in terms of the fermion–boson vertex. We will not explore this direction in detail in this review. The above decomposition is the most natural one in the sense that the method can be easily understood in terms of an iterative process. The higher order diagrams are all generated by iteratively replacing the vertex function at a lower order approximation.

Furthermore, we are mostly interested in models that preserve the $SU(2)$ spin rotation symmetry. Since this symmetry is always obeyed for the two-dimensional calculations at non-zero temperature, it is convenient to preserve this symmetry. This is accomplished by decomposing the vertices in the so-called spin-diagonalized representation. In this representation, the spin degrees of freedom decompose the particle–hole channel into the density and the magnetic channels, and the particle–particle channel into the spin singlet and the spin triplet channels which we denote as $d$-channel, $m$-channel, $s$-channel, and
$t$-channel respectively [30]. They are defined for the irreducible vertex as follows,

$$\Gamma_d = \Gamma_{d,1}^{\text{ph}} + \Gamma_{d,1}^{\text{ph}};$$

$$\Gamma_m = \Gamma_{m,1}^{\text{ph}} - \Gamma_{m,1}^{\text{ph}};$$

$$\Gamma_s = \Gamma_{s,1}^{\text{ph}} - \Gamma_{s,1}^{\text{ph}};$$

$$\Gamma_t = \Gamma_{t,1}^{\text{ph}} + \Gamma_{t,1}^{\text{ph}},$$

and similarly for $F$ and $\Lambda$.

The formalism is completed by equations that connect the different types of vertices. The full vertex is related to the irreducible vertex by the Bethe–Salpeter equation and the irreducible vertex is in turn related to the fully irreducible vertex by the parquet equation. We reproduce the full set of equations for the parquet formulation in the spin diagonalized representation in the following.

The Schwinger–Dyson equation that connects the vertex to the self-energy is

$$\Sigma(P) = -\frac{UT^2}{4N} \sum_{P',Q} \{ G(P') G(P' + Q) G(P - Q) (F_d(Q)_{P' - Q,P'} - F_m(Q)_{P' - Q,P'}) + G(-P') G(P' + Q) G(-P + Q) (F_s(Q)_{P' - Q,P'} + F_t(Q)_{P' - Q,P'}) \},$$

where $G$ is the single-particle Green function, which itself can be calculated from the self-energy using the Dyson equation,

$$G^{-1}(P) = G_0^{-1}(P) - \Sigma(P),$$

where $G_0$ is the bare Green function. Here, the indices $P, P'$ and $Q$ combine momentum $k$ and Matsubara frequency $i\omega_n$, i.e. $P = (k, i\omega_n)$.

The reducible and the irreducible vertices in a given channel are related by the Bethe–Salpeter equation,

$$F_t(Q)_{P',P} = \Gamma_t(Q)_{P',P} + \Phi_t(Q)_{P',P'},$$

$$F_s(Q)_{P',P} = \Gamma_s(Q)_{P',P} + \Psi_s(Q)_{P',P'},$$

where $r = d$ or $m$ for the density and magnetic channels and $r' = s$ or $t$ for the spin singlet and spin triplet channels. The vertex ladders are defined as

$$\Phi_t(Q)_{P',P'} \equiv \sum_{P''} F_t(Q)_{P'',P'} \chi_0^{\text{ph}}(Q)_{P''} \Gamma_t(Q)_{P'',P'},$$

$$\Psi_s(Q)_{P',P'} \equiv \sum_{P''} F_s(Q)_{P'',P'} \chi_0^{\text{ph}}(Q)_{P''} \Gamma_s(Q)_{P'',P'},$$

where $\chi_0$, the bare susceptibility, is the product of two single-particle Green functions.

The parquet equations in the spin diagonalized representation are

$$\Gamma_d(Q)_{P,P'} = \Lambda_d(Q)_{P,P'} - \frac{1}{2} \Phi_d(P' - P)_{P,P+Q} - \frac{3}{2} \Phi_m(P' - P)_{P,P+Q} + \frac{1}{2} \Phi_t(P + P' + Q)_{P,P+Q} + \frac{1}{2} \Psi_s(P + P' + Q)_{P,P+Q} + \frac{1}{2} \Psi_t(P + P' + Q)_{P,P+Q},$$

$$\Gamma_m(Q)_{P,P'} = \Lambda_m(Q)_{P,P'} - \frac{1}{2} \Phi_d(P' - P)_{P,P+Q} + \frac{1}{2} \Phi_m(P' - P)_{P,P+Q} - \frac{1}{2} \Phi_t(P + P' + Q)_{P,P+Q} - \frac{1}{2} \Psi_s(P + P' + Q)_{P,P+Q} - \frac{1}{2} \Psi_t(P + P' + Q)_{P,P+Q},$$

$$\Gamma_s(Q)_{P,P'} = \Lambda_s(Q)_{P,P'} + \frac{1}{2} \Phi_d(P' - P)_{P,P+Q} - \frac{3}{2} \Phi_m(P' - P)_{P,P+Q} + \frac{1}{2} \Phi_t(P + P' + Q)_{P,P+Q} + \frac{1}{2} \Psi_s(P + P' + Q)_{P,P+Q} + \frac{1}{2} \Psi_t(P + P' + Q)_{P,P+Q},$$

$$\Gamma_t(Q)_{P,P'} = \Lambda_t(Q)_{P,P'} + \frac{1}{2} \Phi_d(P' - P)_{P,P+Q} + \frac{1}{2} \Phi_m(P' - P)_{P,P+Q} - \frac{1}{2} \Phi_t(P + P' + Q)_{P,P+Q} - \frac{1}{2} \Psi_s(P + P' + Q)_{P,P+Q} - \frac{1}{2} \Psi_t(P + P' + Q)_{P,P+Q}.$$
It is important to note that if we substitute the irreducible vertices $\Gamma$ (equations (36)–(39)) into the Bethe–Salpeter equation (equations (32) and (33)) the crossing symmetries (symmetry relations of the vertex that are a consequence of the Pauli exclusion principle for identical fermionic particles) in the full vertex $\Phi$ is automatically satisfied regardless of the numerical values of the vertex ladders $\Psi$ and $\Phi$, assuming the fully irreducible vertices, $\Lambda$, obey the crossing symmetries. We write all the full vertices explicitly in the following, using only the vertex ladders, $\Phi$, $\Psi$, and the fully irreducible vertices, $\Lambda$.

\[
F_d(Q)_{P,P'} = \Lambda_d(Q)_{PP'} - \frac{1}{2} \Phi_d(P' - P)_{PP'+Q} - \frac{3}{2} \Phi_m(P' - P)_{PP'+Q} + \frac{1}{2} \Psi_t(P + P' + Q)_{-P, -Q, -P} + \frac{3}{2} \Psi_t(P + P' + Q)_{-P, -Q, -P} \Phi_d(Q)_{PP'};
\]  

\[
F_m(Q)_{P,P'} = \Lambda_m(Q)_{PP'} - \frac{1}{2} \Phi_m(P' - P)_{PP'+Q} + \frac{1}{2} \Phi_m(P' - P)_{PP'+Q} - \frac{1}{2} \Psi_s(P + P' + Q)_{-P, -Q, -P} + \frac{1}{2} \Psi_s(P + P' + Q)_{-P, -Q, -P} \Phi_m(Q)_{PP'};
\]  

\[
F_s(Q)_{P,PP'} = \Lambda_s(Q)_{PP'} + \frac{1}{2} \Phi_s(P' - P)_{PP'+Q} - \frac{3}{2} \Phi_m(P' - P)_{PP'+Q} + \frac{1}{2} \Phi_s(P + P' + Q)_{-P, -Q, -P} + \frac{3}{2} \Phi_m(P + P' + Q)_{-P, -Q, -P} \Phi_s(Q)_{PP'};
\]  

\[
F_t(Q)_{P,PP'} = \Lambda_t(Q)_{PP'} + \frac{1}{2} \Phi_t(P' - P)_{PP'+Q} + \frac{1}{2} \Phi_t(P' - P)_{PP'+Q} - \frac{1}{2} \Phi_t(P + P' + Q)_{-P, -Q, -P} - \frac{1}{2} \Phi_t(P + P' + Q)_{-P, -Q, -P} \Phi_t(Q)_{PP'}.
\]  

The full parquet formalism encompasses a variety of approximations that are widely used in condensed matter physics and materials science. The hierarchy of these different approximations is neatly summarized in figure 9.

- Hartree–Fock and second order perturbation theory: at the highest level, we might make the approximation on the two-particle Green function (analogous to the conventional single particle HF perturbation) such that the four-point correlation function can be factorized as a product of two two-point correlation functions. It is equivalent to ignoring the contribution from the full vertex functions. From the Schwinger–Dyson equation, the self-energy has two contributions a HF term and a second order perturbation theory term.

- Self-consistent second order perturbation theory: substituting the bare vertex for the full vertex in the Schwinger–Dyson equation and solving for the self-energy self-consistently results in the self-consistent second order perturbation theory.

- Random-phase approximation (RPA): the irreducible vertex in the longitudinal charge channel is approximated by the bare Coulomb interaction. And then the Bethe–Salpeter equation is used to sum up over all the ring-type diagrams [23, 24].

- T-matrix approximation (TMA): similar to RPA, the irreducible vertex in the transverse particle–hole channel or particle–particle channel, instead of longitudinal particle–hole channel, is approximated by the bare Coulomb interaction. And then the Bethe–Salpeter equation is used to sum all the ladder-type (instead of ring-type in RPA) diagrams [124].

- Fluctuation exchange approximation (FLEX): a combination of RPA and TMA, such that the fluctuations in different channels are treated equally [125].

The parquet formalism dates back to the 50’s, but as can be seen from the analytical form of the governing equations, a general analytical solution is rather hard to track [126, 127]. Over the years various approximations to simplify the equations have been proposed to solve a set of problems ranging from the Anderson impurity to nuclear structure [30, 31, 128–154]. In addition, a numerical solution is computationally demanding. This difficulty is in general common to theories that involve the two-particle vertex functions. The computational difficulty arises mostly from the memory requirements to store the vertex functions as they are four-legged objects unlike the single particle quantities that only have two legs. If momentum and energy conservation are implemented, the single particle quantities scale linearly with the size of the space-time grid, on the other hand the vertex function scales as the third power of the size of the space-time grid. This challenge is not insurmountable and may be overcome with appropriate
parallelization. In fact, the advent of petascale computing enabled the first solution for the two-dimensional problem [123, 155].

The algorithm for the numerical solution of the parquet formalism is summarized in figure 10.

Perhaps the most challenging problem from the numerical point of view is the difference in the nature of the single particle and the two-particle functions. The single particle quantities in the Matsubara frequency space do not diverge. On the other hand, the two particle vertex has strong divergences when the metallic phase is unstable at the momentum and energy corresponding to an instability. For instance, in the Hubbard model, the particle–hole vertex at the verge of the antiferromagnetic instability has a strong divergence for momentum transfer ($\pi, \pi$). For this reason, the vertex is represented by numbers spanning a wide range of values over many orders of magnitude. This clearly is a recipe for possible numerical instabilities in the iterative solution for the vertex functions at low temperature and on the verge of long range order phase transitions. As we will discuss in the rest of this review, various methods have been proposed to improve the stability of the numerical solutions. These include simplifying the equations, or abandoning the self-consistent approach for the vertex functions.

8. Dynamical vertex approximation (DΓA)

The original dynamical vertex approximation (DΓA) is a simplification of the multiscale many-body method with a pragmatic mindset that limits the calculation to the local fully irreducible vertex using perturbation theory [156, 157]. As discussed in section 7 above, a full parquet solution with space-time resolution is very challenging. A natural scheme to sidestep the difficulty is, in the spirit of DMFT, to consider the vertex to be only time or frequency dependent. The local fully irreducible vertex from a DMFT calculation can then be used as input for the parquet formalism. The method subsequently follows the procedure discussed above for the parquet method, therefore, this aspect will not be repeated in this section.
Since this approach is relatively transparent and numerically practical with modest computational costs [158], besides the two-dimensional Hubbard model, it has been applied to several problems including the three-dimensional Hubbard model, the attractive Hubbard model [159], nanoscopic quantum junction systems [160], and it has also been recently combined with ab initio calculations [161]. A more elaborate calculation based on parquet method has been performed for the one-dimensional Hubbard model [162].

There are a few variations of the dynamical vertex approximation. The full parquet formalism can be solved self-consistently with the fully irreducible vertex from the DMFT solution used as input. One can also consider self-consistency at the level of both the parquet equations and the DMFT equations: this involves finding the fully irreducible vertex from the DMFT solution, then using this to solve the parquet equations. The parquet formalism provides both the full vertex and the dressed Green function. The dressed Green function can in turn be treated as input for the DMFT equation to obtain self-consistency in both the DMFT loop and the parquet equations loop.

Initial applications on the half-filled Hubbard model motivated a further simplification by decoupling the particle–hole channels from the particle–particle channels. This simplification can be justified by the physics of the systems of interest. For example, in the Hubbard model near half-filling, the density wave is driven by the nesting in the particle–hole channels and one can argue for choosing the particle–hole ladder summations. On the other hand, when the system is driven by s-wave pairing such as that in the attractive Hubbard model, one can keep only the particle–particle channel. There is no systematic universal argument on which channel should be dominant. For most interesting regimes, such as the d-wave pairing in the Hubbard model, presumably all channels could contribute and one may have no choice but to try to tackle the full set of equations of the parquet formalism.

The algorithm is summarized in figure 11.

9. Dual fermions

Another path to the multiscale treatment of correlations in a fermionic system is that of the dual fermions approach that was built on previous analogous methods for bosonic systems. This approach systematically incorporates nonlocal correlations into the DMFT solution. The method is distinguished from others in that it maps a strongly correlated fermionic lattice onto weakly correlated delocalized fermions. This allows a perturbative treatment of nonlocal correlations using some subsets of allowed diagrams to produce satisfactory corrections on top of the short-length scale correlations that are addressed by an exact solver.

The dual fermion formalism is an extension of the theory by Sarker for strongly correlated system [163]. He proposed a strong coupling expansion of the solution from the atomic limit that predates widespread usage of DMFT. Similar ideas have also been proposed for the study of one dimensional system [164]. In this theory, the Hubbard model is mapped onto another interacting fermionic model in which the multi-particle hopping-exchange processes appear explicitly. The formulation is equivalent to the dual fermion formalism as currently known [165–167].
Starting from the action of itinerant electrons on a lattice that can be written as:

\[ S[c^*, c] = \sum_{w, k, \sigma} c^*_{w, k, \sigma} [i\omega + \mu - h_k] c_{w, k, \sigma} + \sum_i S_{\text{loc}}[c^*, c], \tag{44} \]

where \( \mu \) is the chemical potential, \( h_k \) is the hoping term in momentum space, \( c^*_{w, k, \sigma}(c_{w, k, \sigma}) \) are the Grassmann variables corresponding to the creation (annihilation) operator, and \( S_{\text{loc}} \) is the local part of the action. The lattice problem can be reframed into that of a set of impurities and an additional term to account for the remaining contributions:

\[ S[c^*, c] = \sum_i S_{\text{imp}}[c^*_i, c_i] - \sum_{w, k, \sigma} c^*_{w, k, \sigma} [\Delta(i\omega) - h_k] c_{w, k, \sigma}. \tag{45} \]

Revisiting the expression of the partition function, a Hubbard–Stratonovich transformation can be applied on the second term, introducing new fermionic degrees of freedom. The action can then be expressed as:

\[ S[c^*, c; f^*, f] = \sum_i S_{\text{restr}, i}[c^*, c; f^*, f] + \sum_{w, k, \sigma} \frac{f^*_{w, k, \sigma} f_{w, k, \sigma}}{\Delta(i\omega) - h_k}. \tag{46} \]

where \( g \) is the single particle DMFT Green function and \( S_{\text{restr}, i} \) is the action restricted to site \( i \), and is defined by:

\[ S_{\text{restr}, i}[c^*, c; f^*, f] = S_{\text{imp}}[c^*, c] + \sum_{\omega, \sigma} f^*_{\omega, \sigma} g^{-1}(i\omega) c_{\omega, \sigma} + \text{h.c.}. \tag{47} \]

The lattice fermionic degrees of freedom can then be integrated out of the action restricted to the respective sites following:

\[ \int \exp \left( -S_{\text{restr}, i} [c^*, c; f^*, f] \right) D [c^*, c] = Z_{\text{imp}} \exp \left( -\sum_{\omega, \sigma} f^*_{\omega, \sigma} g^{-1}(i\omega) f_{\omega, \sigma} + V [f^*, f] \right). \tag{48} \]

This last expression introduces the dual potential \( V[f^*, f] \) in terms of the new fermionic degrees of freedom. It is shown to include all \( 2n \)-vertices for the impurity with \( n = 2, 3, 4, \ldots \).

An explicit expression for the dual potential is obtained by expanding both sides of this equation and comparing the resulting expressions order by order. The dual potential to lowest order reads

\[ V[f^*, f] = \frac{1}{4} \sum_{\omega, \nu, \nu', \nu'', \nu'''} \gamma \gamma' \gamma'' \gamma''' \sum_{\sigma, \sigma', \sigma'', \sigma'''} \int_{\nu, \nu'} L_{\nu, \nu'} f^*_{\nu, \nu'} f_{\nu', \nu''} f^*_{\nu', \nu''} f_{\nu''}, \ldots \]

where \( \gamma \) is the DMFT reducible or full vertex. Thus, nonlocal correlations are addressed by solving the many-body problem with bare Green function \( g \) and interaction potential \( V \). The lattice fermions can finally be integrated out to produce an action that only depends on the dual fermions:

\[ S_D[f^*, f] = -\sum_{w, k, \sigma} f^*_{w, k, \sigma} G^D_{\sigma}(i\omega, k) f_{w, k, \sigma} + \sum_i V[f^*_i, f_i]. \tag{50} \]

With the bare dual fermion Green function defined by:

\[ G^D_{\sigma}(i\omega, k) = \frac{g^2(i\omega)}{[g(i\omega) + (\Delta(i\omega) - h_k)]^{-1}} \tag{51} \]

The action of equation (50) is the tool to account for nonlocal correlations. It can be treated using diagrammatic perturbation theory. In this context, the interaction potential is usually truncated to the four-point vertex. For most practical calculations, higher order terms of the dual potential are truncated [168], though they may have non-negligible effect [169, 170]. However, the formalism is shown, by construction, to be convergent both in the strong coupling and in the weak coupling regime. The process for solving the formalism follows a typical diagrammatic procedure. The impurity Green function and the vertex are obtained from the DMFT calculation. The impurity Green function is used to evaluate the dual fermions non-interacting Green function. The vertex and the non-interacting Green’s function are then used for a self-consistent diagrammatic solution with a given subset of all the allowed diagrams. This solution produces the dressed dual fermion Green’s function. The dual Green function can subsequently be used to evaluate the lattice Green function. Alternatively, phase transitions can be studied directly using the dual fermion diagrams since the instability identified in the dual fermions space is found to be equivalent to that of the lattice Green function. In general, upon solving the dual fermion problem, a new expression for
the impurity self-energy or the impurity hybridization can be extracted and fed back into the impurity solver and the entire procedure repeated iteratively until convergence. The algorithm of the dual fermion method is summarized in figure 12.

The dual fermion formalism, as discussed above, was initially introduced to add nonlocal correlations to the DMFT result. It was subsequently extended to the DCA [171, 172]. In this context, it serves as a way to address intermediate length scales beyond the short ones that are treated by the cluster solver of the DCA formalism.

Because the dual fermions concept is rather general, it has been the subject of numerous developments [173–178]. A notable generalization is the treatment of disorder with this formalism [179, 180]. The effect of disorder in correlated systems has long been an important outstanding problem in condensed matter physics, particularly for the two dimensional case. While experimentally available systems such as semi-conductors usually involve long range Coulomb interactions, the problem of a perhaps simpler Anderson–Hubbard model which have both interaction and disorder at local sites still represents an outstanding challenge which has attracted a lot of attention.

Another important development is the treatment of nonlocal interactions, such as that of the extended Hubbard model with nearest neighbor interaction. The inclusion of nonlocal interaction opens up the possibility for interesting physics such as charge density wave and the more exotic bond order wave. The dual fermions method has also been generalized for the EDMFT to the dual bosons theory as we discuss in the following section [181].

The dual fermion method has additionally been used to generalize the real space DMFT to the real space dual fermion method which allows the study of systems with open boundary condition [182].

10. Dual bosons, extension of EDMFT

One can apply ideas similar to those of the dual fermions for DMFT to EDMFT [181]. As discussed in section 4, the nonlocal part of the EDMFT is not limited to the hopping term, but also includes interaction terms [102]. For this problem, the action can be written as:

$$S [c^*, c] = \sum_{\omega, \k, \sigma} [\epsilon_{\omega, \k} + \mu - h_{\k}] c^{\dagger}_{\omega, \k, \sigma} c_{\omega, \k, \sigma} + \frac{1}{2} \sum_{\omega, \k} V(k)n_{\omega, \k} n_{\omega, -\k} + \sum_{\omega} S_{\text{local}}[c^*, c].$$

(52)

The next step is to introduce a Hubbard–Stratonovich transformation as in the dual fermion formulation for the nonlocal bilinear term of the kinetic energy. In addition, the nonlocal density-density potential energy term is decoupled by another Hubbard–Stratonovich transformation for the bosonic charge density. In parallel with the dual fermion method, the original fermionic degrees of freedom, and the bosonic charge density can formally be integrated out exactly, yielding a new effective potential given by the
Figure 13. Algorithm for the dual boson method. This algorithm is parallel to the dual fermion method, except that both the bosonic Green function and the dual potential with the bosonic degrees of freedom are needed. 1. The DMFT solver provides the Green function and the reducible vertex functions. 2. The Green function and the reducible vertex functions for both fermion–fermion vertex $\left( F^\text{ff}_{\text{imp}} \right)$ and fermion–boson vertex $\left( F^\text{fb}_{\text{imp}} \right)$ from the DMFT solver are used to construct the Green functions ($G^\text{dual}$ and $D^\text{dual}$) and the dual potential ($V^\text{dual}$). 3. Perturbative methods can be used to calculate the self-energy and the Green function of the dual fermions and dual bosons. 4. Inverse transformation of the Green function from dual fermions and dual bosons to physical fermionic and bosonic degrees of freedom. The results can be read out after this step. For the full self-consistent calculation 5. The Green function can be coarse-grained to obtain the bath Green function of the DMFT impurity problem.

vertices of the impurity model. The action becomes

$$S_0 [\phi^*, \phi] = -\sum_{\omega, k, \sigma} f^{\text{LO}}_{\omega, k, \sigma} G^\text{imp}_{\omega, k, \sigma} - \sum_{\omega, k} \phi^{*}_{\omega, k, \sigma} &D^{\text{dual}}_{\omega, k, \sigma} \phi_{\omega, k, \sigma} + \sum_{i} V[f^*_i, f_i, \phi_i].$$ (53)

Similar to the dual fermions, the potential of the dual variables can be expanded and truncated at finite order of the vertex functions of the impurity problem. Perturbative methods can be used to solve the effective problem expressed in terms of the dual variables. The method has been applied to problems with nonlocal interaction, such as the extended Hubbard model and models with long range Coulomb coupling [183–187].

The algorithm of the dual bosons method is summarized in figure 13.

11. GW approximation with three-particle irreducible vertex

A conventional scheme to address quantum fluctuations beyond HF is the so-called GW approximation. It was developed for solutions of the electron gas problem in the 50’s. Due to its simplicity, it has been widely adapted in density functional theory calculations and it can be derived from many-body perturbation theory. The form of the self-energy in the GW approximation is kept as that of the HF approximation, but the interaction, originally just the Coulomb term, is dynamically screened [188–192].

Recent proposals have utilized the properties of the GW approximation and extended it by introducing a dynamical three point vertex function. The method can be derived from the partial bosonization of the electron–electron interaction via the Hubbard–Stratonovich transformation in different channels. The effective action becomes that of an electron–boson coupling problem. The bosonic fields from the decoupling of the electron–electron interaction can be considered as the electrons coupled to the charge and spin fluctuations. This is equivalent to solving the self-energy in the Hedin’s equation with the vertex correction [191]. Instead of solving the electron–boson vertex with spatial dependence, the vertex is calculated via an effective impurity problem similar to that of DMFT. The spatial dependence of both the fermion and the boson self-energies is generated by the fermion ($G$) and the boson ($D$) Green functions. This generalization was introduced by Aryal and Parcollet as the TRILEX [193, 194]. The method can be formally derived from a functional of the vertex given by three-particle irreducible diagrams [195, 196].

The key step for incorporating nonlocal correlations is to consider the polarization and the electron self-energy with the local boson-fermion vertex from the numerical solution of the impurity problem, see figure 14 for the diagrams. The rest of the algorithm is parallel to that of the EDMFT. The updated self-energy and the polarizability are coarse-grained and fed back into the EDMFT effective impurity problem.

The algorithm for the TRILEX method is summarized in figure 15.
Figure 14. Diagrams for generating momentum dependence for the self-energy of the single particle function and the polarization of the two particle function.

Figure 15. Algorithm for the TRILEX method. 1. The DMFT solver provides the boson-fermion vertex. 2. The Hartree diagrams of the GW approximation are calculated with the vertex from the DMFT for both the self-energy and the polarization to generate spatial dependence. 3. The Dyson equation is employed to calculate the Green function for both the fermions ($G$) and the bosons ($D$). 4. The coarse-grained Green functions can be fed back to the effective DMFT impurity problem. Repeat from step 1 until convergence is attained.

12. Functional renormalization group

The functional renormalization group for fermionic systems evolved from approaches used in the one dimensional Luttinger liquid [197–200]. The full vertex in this context is usually denoted as ‘$g$’, for this reason, the method is often referred as g-ology [201, 202]. From the renormalization group perspective, the frequency dependence of the vertex function is not the most relevant contribution. Thus for systems that do not have explicit frequency dependence in the bare vertex, such as the electron–phonon coupling, the frequency or time dependence is often neglected [203].

The generalization of g-ology beyond one dimension gained increased attention after the work by Shankar [203, 204]. Traditional implementations of Wilsonian renormalization group only consider the flow of a handful of coupling constants, which is physically the case for one dimensional systems as there are only two Fermi points instead of a surface. The functional formalism for all coupling constants is considered when the method is generalized to high spatial dimensions.

The idea of performing renormalization group on a system with an extended Fermi surface was suggested by Anderson [205]. Benfatto and Gallavotti [206], and Feldman and Trubowitz [207, 208] studied the stability of the Fermi liquid against perturbations. Effective theories based on the idea of renormalization group for Fermi liquid were derived by Polchinski [209], and for superconductors by Weinberg [210].

Most studies in the late 90’s and early 00’s focused on the two dimensional Hubbard model [211–217]. Two important developments paved the way for using FRG to improve DMFT solutions. First, originally, the conventional renormalization group uses an energy cutoff [203]. This is clearly not a unique choice, and other cutoffs such as temperature, interaction, and even an hybridization cutoff for the impurity problem have been proposed and implemented [218–220]. The second development is the implementation of the
frequency dependent vertex. This is largely motivated by the interest in studying more complicated electron–phonon coupling models [221, 222]. When only electron–phonon coupling is considered, the frequency dependence, even at the single particle level, is not considered. That is, the self-energy is not renormalized. The situation changes when explicitly retarded interactions are considered in models with electron–phonon coupling or local disorder, such as the Anderson impurity model [122, 223–225].

It was realized early in the study of the frequency dependent vertex that the method can be used to study mean field fluctuations. Attempts have been made to use it for the study of Gaussian fluctuations of the slave boson mean field solution for the t-J model [226]. Progress in this research direction has been limited because the conventional assumption that the bandwidth should be larger than the interaction is not met for a wide range of parameters of the slave boson t-J model.

Related ideas have been revived in the past few years to consider mean field fluctuations. Instead of a static mean field solution, the solution of the DMFT is considered [227]. The technique developed as a result can be directly applied to build up the spatial dependence from the DMFT or DCA solution. The main idea is that the bare vertex is replaced by the full vertex from the DMFT or DCA, and the self-energy is set to that of the DMFT or DCA for the initial conditions of the renormalization flows.

From a general perspective, the renormalization group can be viewed as summing up diagrams that are often referred to as forming the leading divergence [228]. In particular if second order perturbation is used in the renormalization group calculation, the diagrams generated have the same topology as those of the parquet method discussed above.

There is however a subtlety. The integration over the internal frequency and momentum are not the same as those of the parquet method. The difference depends on the different cutoff schemes being used, but the summation is never performed over the full range of frequency and momentum, instead it is done on a shell of the energy range, and iteratively approaches the desired low energy or low temperature manifold.

It has long been believed that the leading divergences of the one-loop FRG and the parquet should be equivalent [202, 229, 230]. A recent proposal of a multi-loop flow equation for the four-point vertex framework showed that the FRG flow consisting of successive one-loop calculations is equivalent to a solution of the parquet equations. This further supports the idea of FRG as a possible alternative to solving the parquet equations [231, 232].

Given the same topology of the diagrams, one might expect that the leading divergence among these two methods should be the same even outside of the multi-loop setting. Therefore, for the purpose of looking for the instability from metallic to ordered phase, these two approaches should be expected to give the same result. On the other hand, for the metallic phase with no proximity to an instability, the results are not naively equivalent.

The clear advantage of the functional renormalization group is the ease of the numerical calculations. Unlike, the parquet or even the simplified dynamical vertex approximation, the functional renormalization group works directly on the full reducible vertex. The numerical solution does not involve solving self-consistent equations. It is given by the flow of the full vertex as the cutoff is lowered to the desired energy or temperature. The formulation is represented in terms of ordinary differential equations.

The main challenge of the parquet method is the instability of the numerical solution. Even if we assume the existence of a unique solution, a robust method to attain this solution remains highly non-trivial. This is ultimately the main issue with methods that seek self-consistent solutions. Unlike the DMFT in which only single particle quantities are involved, the two-particle methods, full parquet or simplified forms, involve solving for the vertex through self-consistent equations. Divergence in the vertex function is expected to occur as the temperature is lowered. Therefore, equations with variables spanning a large range of values over many orders of magnitude have to be solved. This is clearly a non-trivial task from the perspective of numerical simulations. Although one can solve the equations at high temperature, it is not always clear how far the solution can be pushed down in temperature. Contrary to this, the functional renormalization group method sums up the same set of diagrams without the necessity of solving relevant equations self-consistently. The divergence is approached step by step rather than via shooting as is done in the self-consistent solution.

While the advantage of functional renormalization group from the point of view of numerical stability is clear, the justification of its usage for nonlocal corrections with the vertex function from the DMFT solution is not obvious. The conventional wisdom of renormalization group calculations is that the bandwidth should be large and the interaction is a small parameter. The full vertex from DMFT is not necessarily small compared to the effective bandwidth. Therefore, the conventional wisdom of justifying the low order expansion is not incontrovertibly fulfilled. Moreover, these couplings generate self-energy corrections, which have been shown to be important for studying systems with retardation effects, leading to the renormalization of the Fermi velocity and quasiparticle lifetime. Thus, these couplings have to be kept even
though they are often ignored in the non-retarded systems. In brief, the effective system being solved is retarded although the original Hubbard model is not.

To leading second order expansion, the renormalization group equations for the scale ($\Lambda$) dependent full vertex function, $F(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) [211, 212]$, and the self-energy, $\Sigma(\mathbf{k})$ for a spin rotationally invariant two-body interacting system, are given by:

$$
\frac{\partial}{\partial \Lambda} F_{\mathbf{A}}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = - \int dp \partial \Lambda[G_{\mathbf{A}}(p)G_{\mathbf{A}}(\mathbf{k})] F_{\mathbf{A}}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) F_{\mathbf{A}}(p, \mathbf{k}, \mathbf{k}_3)
$$

$$
- \int dp \partial \Lambda[G_{\mathbf{A}}(p)G_{\mathbf{A}}(\mathbf{q}_1)] F_{\mathbf{A}}(\mathbf{p}, \mathbf{k}_2, \mathbf{q}_1) g_{\mathbf{A}}(\mathbf{k}_1, \mathbf{q}_1, \mathbf{k}_3)
$$

$$
- \int dp \partial \Lambda[G_{\mathbf{A}}(p)G_{\mathbf{A}}(\mathbf{q}_2)] \left[-2g_{\mathbf{A}}(\mathbf{k}_1, \mathbf{p}, \mathbf{q}_2)F_{\mathbf{A}}(\mathbf{q}_2, \mathbf{k}_2, \mathbf{k}_3) + F_{\mathbf{A}}(\mathbf{p}, \mathbf{k}_1, \mathbf{q}_2)F_{\mathbf{A}}(\mathbf{q}_2, \mathbf{k}_2, \mathbf{k}_3) + F_{\mathbf{A}}(\mathbf{k}_1, \mathbf{p}, \mathbf{q}_2)F_{\mathbf{A}}(\mathbf{q}_2, \mathbf{k}_2, \mathbf{k}_3)\right],
$$

$$
\frac{\partial}{\partial \Lambda} \Sigma_{\mathbf{k}}(\mathbf{k}) = - \int dp \partial \Lambda[G_{\mathbf{A}}(p)] [2F_{\mathbf{A}}(p, \mathbf{k}) - F_{\mathbf{A}}(p, \mathbf{k}, \mathbf{k})],
$$

where $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{p}, \mathbf{q}_1 = \mathbf{p} + \mathbf{k}_3 - \mathbf{k}_1, \mathbf{q}_2 = \mathbf{p} + \mathbf{k}_3 - \mathbf{k}_2, \int dp = \int dp \sum_{\omega} 1/(2\pi\beta)$, and $G_{\mathbf{A}}$ is the self-energy corrected propagator at cutoff $\Lambda$ (figure 16).

The RG equation can be presented in terms of diagrams. For a non-retarded system, the low energy instability can be obtained from the renormalization flow of the couplings, and different phases can be identified by the fixed points corresponding to the relevant spin and charge modes. One can explicitly construct the flows of the susceptibilities of different order parameters. For example the pairing susceptibility, $\chi^\Delta(\mathbf{k}, \omega)$, is defined by:

$$
\chi^\Delta(\mathbf{k}, \omega) = \int dp dp_2 dp_3 c_{\mathbf{p}_1} c_{\mathbf{p}_1'} c_{\mathbf{p}_2} c_{\mathbf{p}_2'} \Lambda.
$$

The RG equations are:

$$
\frac{\partial}{\partial \Lambda} \chi^\Delta(0, 0) = \int dp \partial \Lambda[G_{\mathbf{A}}(p)G_{\mathbf{A}}(-p)](Z^\Delta(p))^2,
$$

$$
\frac{\partial}{\partial \Lambda} Z^\Delta(p) = - \int dp' \partial \Lambda[G_{\mathbf{A}}(p')G_{\mathbf{A}}(-p')] Z^\Delta(p') g_{\mathbf{A}}(-p' - p, -p, p).
$$

The function $Z^\Delta(p)$ is the effective vertex in the definition for the susceptibility $\chi^\Delta$. The RG equations for susceptibilities are solved with initial condition $\chi^\Delta(0, 0) = 0$. The dominant instability in the ground state is given by the most divergent susceptibility by solving the renormalization equations numerically. Similar equations can be derived for other susceptibilities.

For the FRG boosted DMFT approach, DMF$^2$RG [227], the initial condition for the full vertex functions and the self-energy are both given by the DMFT [227]. The scale dependent bare propagator is defined as an interpolation between the DMFT propagator and the bare lattice propagator as:

$$
G_{\mathbf{A}}^0(\mathbf{k}, i\omega) = \Lambda G_{\text{imp}}^0(i\omega)^{-1} + (1 - \Lambda) G_{\text{lat}}(\mathbf{k}, i\omega)^{-1}.
$$

Since one can treat the functional renormalization group as a vehicle for summing diagrams, it can in principle be applied on any effective fermionic interacting system. For example, it has recently been used as
Figure 17. Algorithm for the functional renormalization group boosted DMFT. 1. The full reducible vertex and the Green function are obtained from the DMFT solver. 2. They are fed into the FRG method as the initial conditions. 3. Solve the flow equations for the FRG to obtain the Green function, full vertex, and susceptibilities.

The DMF2RG algorithm is summarized in figure 17. For the derivation of the FRG formulation in the context of strongly correlated systems we refer the readers to references [203, 212, 219, 236]. Details of the implementation and approximations of the DMF2RG can be found in reference [227].

13. Numerical methods to represent vertex functions

The various methods described above in principle provide nonlocal corrections to the conventional dynamical mean field theory or additional nonlocal corrections in the case of the DCA or the cellular dynamical mean field theory. The ultimate goal is to attain a better approximation to the exact solution. The physically most interesting regimes are often in the intermediate coupling away from the trivial limits that could serve as a good basis for perturbative treatments. After all, all considered methods are based on some truncation with respect to the interaction or more complicated objects as those in the ‘dual’ variables approach. Moreover, the actual numerical implementation is sometimes rather challenging. Unlike DMFT where only single particle quantities are involved in the self-consistent equations, the storage requirement of two particle vertex functions is increased by two powers of the space-time grid size. Simply storing those vertices is in itself a rather difficult task. Of course, that largely depends on the interaction and the temperature range. One can naively expect that more fine resolution in the space-time grid is needed for intermediate interactions and low temperatures. For high temperature and very weak or very strong coupling, a rather sparse grid for discretizing the space-time functions could be sufficient.

The above storage problem can be resolved to an extent in today’s clusters with tens of thousands of computing nodes. How to efficiently manipulate an object with such large memory requirement, and involving all-to-all memory swaps, is an active research problem in computer science [155, 237]. Altogether, with the improvement of computer hardware and better implementations, the storage problem can be mitigated.

The main issue that is ultimately common to most of these methods is the convergence to a solution. There is in general no guarantee that a self-consistent solution exists. While even in the single particle self-consistent method, there is no guarantee that there is a uniform convergence to the solution. The situation is more acute in theories that require two particle self-consistency. The vertex function is a measure of the instability towards an ordered phase, therefore it should become singular as the instability is approached. One can often work in the range of interactions and temperatures where singularities are far from being reached. However, this may altogether defeat the purpose of the methods as the regimes near the instabilities are usually those where corrections to the DMFT are of most interest. The challenge of finding a stable numerical solution of the parquet formalism was already discussed in the early days when the formalism was applied to the single impurity Anderson model [132, 238]. It is fair to say that a universally reliable approach for a solution has yet to be found.

For DMFT, besides the simple iterative method, the more powerful Broyden method which utilizes the gradient of the hybridization function in the impurity problem has been implemented and is occasionally used in cases where convergence is difficult to achieve [239]. For the two-particle theories, given the complexity of the equations being solved, more sophisticated methods have been attempted. One of them is the homotopy method, in which a known convergent solution is relaxed to hopefully lead to a solution for
another temperature or interaction strength. Practically, the methods are not in general easy to apply and convergence is not guaranteed.

The above difficulties in solving for the two-particle vertex function self-consistently may give an advantage to FRG based methods in which no self-consistent solution is sought. The solution is obtained not by solving non-linear integral self-consistent equations, but rather by solving differential equations with initial conditions. This allows more flexibility in the numerical solution.

For substantial progress to be accomplished, it is essential that numerical methods perform well while, at the same time, requiring reasonable computational resources. Although storing the full vertex is a daunting task, the amount of information it actually contains is, in practice, not very large. At least for the weak coupling case, the vertex functions contain very small entropy in the sense that it can be compressed numerically to a large extent and still retain most of the information. It has been suggested for a long time that, a possible route to storing the vertex function is to use the spectral representation [240, 241]. A clear advantage in the spectral representation is that the high frequency information is built in the representation. The spectral representation has been further explored in recent studies [242, 243].

Various efforts have also been devoted to understanding the frequency structure of the vertex function. These may help with new ideas on approximation schemes for the vertex function [244–250]. The latest proposal is to use a tensor network representation [251].

Yet another intuitive scheme is to consider an inhomogeneous grid to represent momentum-frequency space indices. Generically, the frequency or momentum dependence is described by an interpolation scheme. This can be justified specifically for the frequency indices because the low frequency information should be more important, contains most of the information and thus requires higher resolution. Moreover the high frequency contribution can be well fitted by simple functions for convenient storage. Generally, such methods which are based on interpolations, can be seen as approximating the vertex as follows:

\[
F(\omega_k, \omega_{k'}, \omega_q) = S(\tilde{F}(\tilde{\omega}_k, \tilde{\omega}_{k'}, \tilde{\omega}_q), \omega_k, \omega_{k'}, \omega_q),
\]

where \( F \) is the vertex in the uniform frequency grid \( \omega \) and \( \tilde{F} \) is the actual data stored in the grid of some arbitrary basis in \( \tilde{\omega} \). \( S \) is the interpolating function or the basis function that maps \( \tilde{F} \) to \( F \).

Note that there is no inverse for such an interpolation or basis expansion. Similar ideas can also be extended to the space grid. This has been studied recently in the context of FRG and also the parquet method [252–256]. An extreme case of only retaining the transfer frequency and momentum has been proposed as a simplification of the parquet equations for analytical solutions for single impurity problems [133–135].

The lack of an inverse transformation to the original vertex means that the self-consistent equations are, in principle, altered by the transformation. The representation is however a controlled approximation in the sense that the larger number of basis functions or the larger number of grid points can reduce the acquired error. On the other hand, strictly speaking, the crossing symmetries of the vertex, that are a manifestation of the Pauli exclusion principle, are also broken by such interpolation or basis expansion schemes. These symmetries are one of the key features of the parquet algorithm.

A scheme that has been studied in the context of FRG is to factorize the vertex function approximately [227]. In the representation of the vertex function with two fermionic frequencies and one bosonic transfer frequency, the frequency transfer has the most dominant contribution. Thus, one can argue physically that the frequency or momentum transfer part of the vertex function can be factorized. This drastically reduces the computational effort and, in particular, the storage requirement. A less drastic approximation is the so-called two-level approximation [246], for which the low frequency part of the vertex is calculated exactly while only the frequency transfer dependence is kept for the high frequency part of the vertex. Another recent proposal is to consider the momentum dependence by using an expansion in terms of, say, spherical harmonics. This can help reduce the storage for the spatial dimension and thus allow studies of larger cluster sizes [252].

It is important to point out out that the corrections on top of the DMFT solution, while they are indeed quite meaningful, are often not large for most momentum-frequency points, particularly for weak or very strong interactions. It is, in this sense, appropriate to inquire what are the effects of numerical errors on a method that would, in principle, be capable of significant corrections to DMFT. Might the approximate numerical solutions overwhelm the expected solution? Of course, one could expect that corrections in the intermediate regime are larger, but the convergence problem in this regime may prevent a satisfactory solution.

Overall, a number of challenges still remain for general numerical solutions. Algorithmic breakthroughs, advances in the representation of the vertex function and in the solvers, are needed for general converging solutions over a wider range of parameters.
14. Exact constraints

14.1. Conserving approximation
A milestone in the theory of condensed matter physics is the formulation of the so-called conserving approximations by Baym and Kadanoff [257, 258]. The idea of constructing approximations based on a functional of the Green function was pioneered by Luttinger and Ward, \( \Phi(G) \) [259]. It was shown that the functional derivative with respect to the Green function can be used to define the self-energy \( \delta \Sigma = \text{tr} \delta G \). Putting it together with the Dyson equation \( G^{-1} = G_0^{-1} - \Sigma \), the correlation functions generated by the functional obey conservation laws and thus the Ward identities.

Conventional methods that are also conserving approximations include self-consistent solutions of the Hatree–Fock, TMA, and GW approximations. Another well known method that is also a conserving approximation is the fluctuation exchange (FLEX) approximation.

Dynamical mean field theory and its cluster extensions, CDMFT and DCA, are both conserving approximations although the Ward identity on the lattice is strictly speaking not satisfied due to the coarse-graining procedure. Though ‘conserving approximation’ does not necessary imply ‘better approximation’, conserving approximations may be more important for broken symmetry cases, that are not routinely studied by the DMFT.

It has been argued that both the dual fermion and the dual boson approximations give conserving approximations for ladder summation [181]. That should also be the case for the dynamical vertex approximation, at least for the ladder approximation.

14.2. Mermin–Wagner–Coleman theorem
This theorem forbids broken continuous symmetry for a system in dimensions two or less at a finite temperature [260–262]. A particularly interesting case is that of the spin rotation symmetry in the two dimensional Hubbard model that should not be broken and thus no true phase transition should occur.

We are not aware of a proof that any of the methods presented in this review satisfies this constraint. There is plenty of evidence that the DMFT and its cluster extensions do show finite transition temperatures through measurements of the susceptibilities.

A recent study by the parquet method seems to provide numerical evidence that Mermin–Wagner theorem could be satisfied [252]. The dual boson theory has also been argued to fulfill the Mermin–Wagner theorem by restricting the Hubbard model double occupancy to be the same as that of the impurity model which is solved numerically in an exact manner [187].

14.3. Causality
DMFT, CDMFT, and DCA can be shown to be causal. We are not aware of a proof of causality for dual fermion, dual boson, dynamical vertex approximation, parquet, or FRG [41, 42]. However, in practice, that does not seem to be a serious problem.

14.4. Crossing symmetry, Pauli principle
An important feature of the parquet method is that the crossing symmetries, as illustrated for the particle–particle vertex in figure 18, expressing the identity:

\[
F_{pp}(12, 34) = -F_{pp}(21, 34),
\]

are satisfied. The full FRG also satisfies the crossing symmetries. However, in practical implementations, the crossing symmetries may sometimes be broken due to the approximations made on the vertex functions.

15. Applications
In this section we discuss the major applications of the MSMB methods discussed above. The list of applications here is not intended to be exhaustive, in particular we focus exclusively on studies of lattice
models. Efforts to combine density functional theory with these post-DMFT/DCA methods are the subject of current active research. Readers interested in these efforts will find useful discussions in references [161, 263–272].

15.1. Parquet

The parquet formalism is quite demanding from the point of view of computational resources and, for this reason, it has not been as extensively pursued as other diagrammatic methods despite its promise. We highlight two earlier studies using this method. Recent studies include applying it to the dual fermion diagrams and dynamical vertex approximation, or simplifying the calculation by different representations of the vertex functions as discussed in previous subsections.

Yang et al solved the full parquet equation for the half-filled $4 \times 4$ cluster [123]. Results were compared to those obtained from the determinant quantum Monte Carlo (DQMC), fluctuation exchange (FLEX), and self-consistent second-order approximation methods. This comparison, illustrated in figure 19, shows satisfactory agreement with DQMC and a significant improvement over the FLEX or the self-consistent second-order approximation.

When exploring the origin of instabilities identified in the susceptibility, the parquet formalism enables the identification of separate channel contributions. Indeed, a key feature of the parquet equations is that they express the contribution to a given scattering channel by processes from other channels. This is essential in understanding the mechanisms for key processes and phase transitions. Yang et al analyzed the pairing vertex as a function of temperature and doping. To this end, they wrote it in the form of its contributions from the charge and spin channels: $\Gamma = \Lambda + \Phi_c + \Phi_s$, and applied the $d$-wave projection to this equation to get the expression in terms of the different components: $V_d = V_{d\Lambda} + V_{d\Phi} + V_{d\Phi}$. This analysis indicated that the dominant contribution to $V_d$ originates from the spin channel [273].

The relations (40)–(43) can also allow us to restore the crossing symmetries for the full vertices during the iterative solution of the self-consistent equations. This procedure was found to enhance the stability of the self-consistent solution. With two-particle diagrams, physical instabilities can be identified by examining the divergence of the susceptibility in the associated scattering channel. This divergence is also manifested through the leading eigenvalues of the pairing matrix in the channel $r$, $M_r = \Gamma_r \times \chi_0$, becoming equal to 1. As shown in figure 20, the leading eigenvalues of the pairing matrix typically diverge prematurely in the self-consistent solution of the parquet formalism and thus lead to a breakdown of the numerical solution at moderate to large values of the interaction. However, when the crossing symmetries are enforced...
Kusunose solved parquet equations for both the impurity Anderson model and the Hubbard model on a square lattice mainly for the particle–hole symmetric or the half-filled case \[141\]. He argued that in both models the vertex renormalization in the spin channel eliminates the magnetic instabilities of the mean-field theory to ensure satisfaction of the Mermin–Wagner theorem. The parquet method gives the same critical exponents as the self-consistent renormalization Moriya theory in the quantum critical region.

Pudleiner et al studied the Pariser–Parr–Pople (PPP) model or Hubbard model with nonlocal interaction for the conjugated π bonds in benzene \[274\]. They found that quasiparticle renormalization is much weaker in the PPP than in the Hubbard model, but the static part of the self-energy enhances the band gap of the PPP model. In addition, the vertex corrections to the optical conductivity are much more important in the PPP model.

15.2. Dynamical vertex approximation

Most of the applications of the dynamical vertex approximation are focused on the Hubbard model in two dimensions. The one dimensional Hubbard model has also been studied by the full parquet dynamical vertex approximation \[275\]. For the three dimensional Hubbard model, it has been found that the antiferromagnetic phase develops incommensurate magnetic ordering as the doping increases \[162\]. Interesting results on the critical exponent have also been obtained \[276\].

The main gain of including more spatial fluctuations is the suppression of the tendency towards ordering. The dynamical vertex approximation has been shown to reduce the transition temperature of the antiferromagnetic ordering in the half-filled Hubbard model in both two and three dimensions. As we discussed in the previous section, neither DMFT nor dynamical vertex approximation fulfill the Mermin–Wagner theorem, thus broken symmetry is allowed at a finite temperature even for the two-dimensional case. We summarize the major results from the dynamical vertex approximation in the following.
15.2.1. Hubbard model

Valli et al used the parquet dynamical vertex approximation to study the electronic self-energies and the spectral properties of the finite-size one-dimensional Hubbard model with periodic boundary conditions. In this model the Fermi liquid theory is invalid, and that should present a rather challenging case for any perturbative expansion on top of the DMFT solution [275]. Valli et al suggested that for a non-degenerate bare dispersion, the parquet dynamical vertex approximation quantitatively reproduces the exact many-body solution of the system. This is illustrated in figure 21 that compares the exact self-energy to that of the parquet DΓA. Given that the system should be a Luttinger liquid which cannot be adiabatically tuned into a Fermi liquid, this is a very encouraging result.

Schäfer et al have studied the two-dimensional Hubbard model on a square lattice [278]. They defined two transition lines in the phase diagram: one for the gap cutting across the nodal direction and the other for a gap throughout the Fermi surface. The self-energy data shows that the evolution between the two regimes occurs in a gradual way, not through a phase transition, and also that at low enough temperatures the whole Fermi surface is always gapped.

Schäfer et al also showed that the electron self-energy is well separable into a local dynamical part and a static nonlocal contributions for the three dimensional Hubbard model [279]. The quasiparticle weight remains essentially momentum independent for different fillings, including in the presence of overall large nonlocal corrections to the self-energy.

Pudleiner et al computed the self-energy for the half-filled Hubbard model on a square lattice using lattice QMC simulations and the dynamical vertex approximation [280]. The self-energy is strongly momentum-dependent, but it can be parameterized via the noninteracting energy–momentum dispersion ϵk, except for some pseudogap features right at the Fermi edge.

In reference [281], Schäfer et al studied the two dimensional Hubbard model by combining dynamical vertex approximation, lattice QMC, and variational cluster approximation (VCA). They demonstrated that scattering at long-range fluctuations due to paramagnons opens a spectral gap at weak-to-intermediate couplings, irrespective of the preformed localized or short-ranged magnetic moments. They argued that the two-dimensional Hubbard model has a paramagnetic phase which is insulating at low enough temperatures for any finite interaction and no Mott–Hubbard transition is observed.

Schäfer et al found that the antiferromagnetic phase transition of the Hubbard model in three dimensions is in contradiction with the conventional Hertz–Millis–Moriya theory [162]. They argued that the quantum critical behavior is driven by the Kohn anomalies of the Fermi surface, even when electronic correlations become strong.

Rohringer et al studied the three dimensional half-filled Hubard model [276]. They found the Neel temperature is lowered from that of the DMFT as expected. More interestingly, they found the critical exponents to be the same as those of the three dimensional Heisenberg antiferromagnet in contrast to mean
field exponents. This demonstrates that non-mean-field behavior can indeed be obtained by these systematic nonlocal corrections.

Rohringer and Toschi studied several spectral and thermodynamic properties of the Hubbard model in two and three dimensions [282]. Specifically, by evaluating the electronic scattering rate and the quasiparticle mass renormalization in the low energy regime, they characterized the gradual deterioration by nonlocal correlations of the Fermi liquid physics as a function of the interaction strength. They found that the kinetic energy either increases or decreases compared to that of the DMFT depending on the interaction strength being weak or strong, respectively. They argued that these results correspond to the evolution of the ground state from a nesting-driven (Slater) to a superexchange-driven (Heisenberg) antiferromagnet.

15.2.2. Attractive Hubbard model

Lorenzo Del Re et al studied the attractive Hubbard model in three dimensions for the pairing or charge density wave ordering [159]. They found that the fitted critical exponents from the ladder DΓ results were larger not only than the DMFT ones, but also larger than the exact ones belonging to the corresponding universality class.

15.2.3. Periodic Anderson model

Schäfer et al studied the phase diagram and quantum critical region of the periodic Anderson model [283]. They found a phase transition between a zero-temperature antiferromagnetic insulator and a Kondo insulator. In the quantum critical region, they determined a critical exponent $\gamma = 2$ for the antiferromagnetic susceptibility. This becomes $\gamma = 1$ at high temperature.

15.3. Dual fermions

The dual fermion method has been used to investigate models from the Falicov–Kimball to the Hubbard model in both two and three dimensions. The Anderson model for random disorder has also been studied and so far, this is the only post-DMFT method for the study of disorder. Perhaps a more interesting study is that of the Anderson–Hubbard model, which investigates the long standing problem of the competition between Mott insulator and Anderson insulator. We summarize the major results from the dual fermion method in the following.

15.3.1. Hubbard model

Hafermann et al used the ladder diagrams for the dual fermion and found that the critical Néel temperature of the mean-field solution is suppressed in the ladder approximation of the two-dimensional Hubbard [284].

Rubtsov et al found that the antiferromagnetic pseudogap, the Fermi-arcs formation, and the non-Fermi-liquid effects due to the Van Hove singularity are correctly reproduced by the lowest-order diagrams for the two-dimensional Hubbard model [285].

Otsuki et al obtained the phase diagram for the two-dimensional Hubbard model [286]. This features a phase separation region in the low-doping regime around the Mott insulator.

Astretsov et al mapped out the phase diagram of the 2D Hubbard model as a function of temperature and doping. They identified an antiferromagnetic region at low doping and a superconducting dome at higher doping [177]. Their results support the role of the van Hove singularity as an important ingredient for the high value of $T_c$ at optimal doping. At small doping, the destruction of antiferromagnetism is accompanied by an increase of the charge fluctuations supporting the scenario of a phase-separated state driven by quantum critical fluctuations.

Tanaka studied the square-lattice Hubbard model at half-filling using the ladder dual fermion approximation. He found that the almost simultaneous creation of the pseudogap and the loss of the Fermi liquid feature is consistent with what is expected in the Slater regime [287]. Although the pseudogap still appears in the quasi-particle-like single peak for $U \leq 4$, the Fermi-liquid feature is partially lost on the Fermi surface already at higher temperatures as expected in the Mott–Heisenberg regime, where local spins are performed at high temperatures. A sharp crossover from a pseudogap phase to a Mott insulator at finite $U \approx 4.7t$ was found to occur below the temperature of the pseudogap formation.

van Loon et al applied the dual fermion approach with a second-order approximation to the self-energy for the Mott transition in the two-dimensional Hubbard model [288]. A strong reduction of the critical interaction and an inversion of the slope of the transition lines with respect to single-site dynamical mean-field theory was observed.

Katanin et al showed the suppression of the quasiparticle weight in the three-dimensional Hubbard model [289]. With an additional correction in the susceptibility to fulfill the Mermin–Wagner theorem [290], they also found a dramatically stronger impact of spin fluctuations in two dimensions where the
pseudogap is formed at low enough temperatures. They proposed that the origin of the pseudogap at weak-to-intermediate coupling is in the splitting of the quasiparticle peak.

Hirschmeier et al studied the three dimensional Hubbard model and they reported that in the weak-coupling regime, spin-flip excitations across the Fermi surface are important while the strong-coupling regime is described by Heisenberg physics [291]. For intermediate interaction, aspects of both local and nonlocal correlations appear. They also found that the critical exponents of the transition in the strong-coupling regime are consistent with the Heisenberg model down to an interaction of $U = 10t$. Again the identification of non mean-field exponents is an interesting finding.

Antipov et al demonstrated that diagrammatic multiscale methods anchored around local approximations are indeed capable of capturing the non-mean-field nature of the critical point of lattice models [292]. This is an interesting result as the mean field theory describes the longest length scale in the problem.

van Loon et al studied the two-dimensional square-lattice for small to moderate interaction strengths [293]. The nonlocal correlations beyond dynamical mean-field theory induce a pseudogap in the density of states. The upper bounds on the crossover temperature are found to be significantly lower than previously reported dynamical vertex approximation results at $U = t$.

As mentioned previously, the methods presented here can be applied to quantum cluster theories such as DCA and CDMFT to perturbatively capture nonlocal correlations beyond the length scale of the initial cluster size. Figure 22 shows an analysis of the leading eigenvalue of the antiferromagnetic pairing matrix for different approximate methods within the dual fermion approach. The figure indicates how the unphysical phase transition obtained from the mean field result (DCA with one site) is suppressed with different approximations of the dual fermion solution [171].

Figure 23 shows a compilation of results obtained using different methods and illustrates the systematic corrections to the transition temperatures through the incorporation of nonlocal correlations in the two-dimensional Hubbard model at half-filling. The figure depicts, at large $U$ and at low $T$, the DMFT paramagnetic metal solution indicating the first-order Mott metal insulator transition with a low temperature Mott paramagnetic insulator. The first-order transition terminates at a critical value of $U_c = 10t$. Including the short range antiferromagnetic correlations such as in CDMFT, VCA or second-order dual fermions (DF(2)), modifies the critical interaction value and the shape of the coexistence region. Including longer-range antiferromagnetic fluctuations through ladder $D\Gamma A$ or the two-particle self-consistent method leads to further modifications eventually transforming the MIT into a crossover at small $U$ that is consistent with the $U_c \to 0$ for $T \to 0$ limit.

15.3.2. Hubbard model on a triangular lattice

Yudin et al studied the Hubbard model on a triangular lattice [294]. They showed that the band flattening is driven by correlations and is well pronounced even at sufficiently high temperatures, of the order of 0.1–0.2 times the hopping parameter.
Lee et al studied the Hubbard model on the triangular lattice at half filling. They determined the metal–insulator transition and the hysteresis associated with a first-order transition in the double-occupancy and nearest-neighbor spin-correlation functions as functions of temperature [295]. By calculating the spin susceptibility, an enhancement of antiferromagnetic correlations and evidence for magnetically ordered phases were found.

Antipov et al studied the half-filled Hubbard model on an isotropic triangular lattice with a spin polarized extension of the dual fermion approach [296]. They found that the dual fermion corrections drastically decrease the energy of a spin liquid state while leaving the non-collinear magnetic states almost non-affected. This makes the spin liquid become a preferable state in a certain interval of interaction strength of the same order of magnitude than the bandwidth.

Li et al studied both the half-filled and the doped Hubbard model on a triangular lattice and produced its phase diagram [297].

15.3.3. Hubbard model on the honeycomb lattice

Hirschmeier et al studied the Hubbard model on the honeycomb lattice in the vicinity of the quantum critical point by means of a multiband formulation of the dual fermion approach. They found that the critical interaction strength of the quantum phase transition from a paramagnetic semimetal to an antiferromagnetic insulator is in good agreement with other numerical methods [298]. They also argued that the Hubbard model on the honeycomb lattice behaves like a quantum nonlinear $\sigma$ model, while displaying signs of non-Fermi-liquid behavior.

15.3.4. Falicov–Kimball model

Astleithner et al studied the Falicov–Kimball model. Using the full parquet dynamical vertex approximation, they argued that weak localization corrections in the particle–particle channel are not the dominant vertex corrections to the optical conductivity [299].

15.3.5. Kondo lattice model

Otsuki studied the Kondo lattice model to explore possible superconductivity emerging from the critical antiferromagnetic fluctuations [300]. The d-wave pairing is found to be the leading instability only in the weak-coupling regime. As the coupling is increased, a change of the pairing symmetry into a p-wave spin-singlet pairing was found.

15.3.6. s–d exchange model

Sweep et al studied the critical values of the s–d exchange coupling constant [301]. They reported a difference between the DMFT and dual fermion results that is more than a factor of two for the square lattice and spin one-half localized electrons.

15.3.7. Anderson disorder model

Terletska et al generalized the dual fermion approach to disordered systems using the replica method [179, 302]. The developed method utilizes the exact mapping to the dual fermion variables, and includes inter-site scattering via diagrammatic perturbation theory in the dual variables. As shown in figure 24 nonlocal effects that are missed in the CPA are captured.
15.3.8. Anderson Falicov–Kimball model
Yang et al generalized the dual-fermion formalism for disordered fermionic systems to include the effect of interactions. The phase diagram for the two-dimensional Anderson–Falicov–Kimball model was obtained [303].

15.3.9. Anderson–Hubbard model
Haase et al studied the three-dimensional Anderson Hubbard model. They report that the dual-fermion approach leads to quantitative as well as qualitative improvement of the dynamical mean-field results. This is shown in the phase diagrams of figure 25 obtained with DMFT (a), dual fermions with second order diagrams (b), and dual fermions with FLEX diagrams (c). The systematic improvement of the solution first with the incorporation of non-local corrections and then in terms of the level of the diagrammatic treatment is shown through the expected suppression of the DMFT critical temperatures for the antiferromagnetic phase. These solutions allowed the authors to calculate the hysteresis in the double occupancy in three dimensions, taking into account nonlocal correlations [180].

Otsuki studied the Kondo lattice model [300], and found that different superconductivity pairing symmetries emerge from the critical antiferromagnetic fluctuations. He found the d-wave pairing to be the leading instability only in the weak-coupling regime. As the coupling is increased, a change of the pairing symmetry into a p-wave spin-singlet pairing is observed. The competing superconductivities are ascribed to a crossover between small and large Fermi surfaces, which occurs with the formation of heavy quasiparticles.

15.4. Dual bosons
The dual boson method has been applied on models with nonlocal interactions. These include the truncated long range Coulomb coupling, the nearest neighbor interaction in the extended Hubbard model, and the anisotropic dipolar coupling for cold atoms. A key difference is that those nonlocal density–density type couplings can lead to the competition between charge fluctuations and spin fluctuations which results in charge density wave ordering and possibly bond wave ordering [304].
15.4.1. Extended Hubbard model

Vandelli et al proposed to use QMC to sample diagrams from the dual boson theory for the extended Hubbard model [186]. They proposed that the single-particle Green function allows one to estimate the transition point to the charge density wave phase.

15.4.2. Hubbard model with dipolar coupling

van Loon et al studied the Hubbard model with long-range dipole–dipole interactions [305]. This is an interesting model in the context of experiments with cold atoms on optical lattice [306]. Besides the stripe phase and the checkerboard phase, based on their dual boson calculation they suggest that there is a novel phase with ‘ultralong-range’ density correlations at distances of tens of lattice sites.

15.4.3. Hubbard model with Coulomb coupling

Hafermann et al and van Loon et al studied the polarization for the two dimensional Hubbard model with long range Coulomb coupling [183, 307]. They found that plasmon spectra are qualitatively different from those of the random-phase approximation: they exhibit a spectral density transfer and a renormalized dispersion with enhanced deviation from the canonical behavior.

15.5. TRILEX

Applications of TRILEX are mostly on the two dimensional Hubbard model.

Aryal et al found that the local vertex, for strong interactions, gains a strong frequency dependence, driving the system to a Mott transition for the half-filled Hubbard model on a square lattice [193]. At low enough temperatures, large spin fluctuations lead to an enhancement of the momentum dependence of the self-energy. Upon doping, they find a Fermi arc in the spectral function.

Vučičević et al studied the dependence of the superconducting temperature on the bare dispersion at weak coupling, which shows a clear link between strong antiferromagnetic correlations and the onset of superconductivity [194]. They identified a combination of hopping amplitudes particularly favorable to superconductivity at intermediate doping.

15.6. FRG

Functional renormalization group has been used extensively for over two decades. Recent applications to improve the DMFT solution have so far been mostly limited to the Hubbard model.

Tranto et al first proposed to use FRG to expand the DMFT solution [227]. They studied the half-filled square lattice Hubbard model and found that the method provides more prominent momentum dependence than the conventional FRG method.

Figure 25. $U-T$ phase diagram of the 3D Anderson–Hubbard model for different values of the disorder strength $V$ obtained with DMFT (a), dual fermions with second order diagrams (b), and dual fermions with FLEX diagrams (c). The shaded regions correspond to the antiferromagnetic phase. $W$ is the bandwidth. Note that the authors do not calculate the critical temperatures in the region of $U/W < 1/3$. Reprinted figure with permission from [180], Copyright (2017) by the American Physical Society.
Vilardi et al studied the doped two dimensional Hubbard model [308]. They found strong antiferromagnetic correlations from half-filling to 18% hole doping at low temperature, and a sizable d-wave pairing interaction driven by magnetic correlations at the edge of the antiferromagnetic region.

16. Conclusion

We have reviewed MSMB numerical methods to address strongly correlated systems by appropriately treating the short length scale, the long length scale and the intermediate length scale. The different methods implemented to date have produced promising results despite being hindered by a variety of numerical challenges. Since short length scales are treated exactly, diagrammatic methods arise as a suitable approach to deal with the intermediate length scales by systematically evaluating appropriate subsets of possible diagrams. In this context the parquet formalism is the most natural toolkit. We have reviewed the construction of the parquet formalism and the different diagrammatic approximations that it encompasses as well as algorithms for their numerical solutions. We have not discussed in this review efforts to extend the methods into ab initio calculations. These represent an important next step for appropriate treatments of real materials. In general, MSMB methods to incorporate nonlocal corrections into the DMFT solution represent an active area of research and new implementations are actively being developed to overcome previous shortcomings.

Some of the latest ideas have not been discussed in the present review. These include but are not limited to the parquet method for the vertex in the boson-fermion representation [309–312], the atomic approximation of the four-particle irreducible functional method [313], one-particle irreducible functional method [314], nonlocal expansion method [315], and FLEX + DMFT approach [316, 317].

Another important topic is that of the solvers for the vertex functions. While there are many different numerical solvers for the impurity/cluster problem, most of them are not suitable for the calculation of the vertex function which is essential for perturbative expansions around the DMFT solution. While many solvers may be generalized for the calculation of vertex function, at present the practical methods are exact diagonalization and QMC.

For the ED method, the calculation of the vertex function is usually down to brute force calculation in the Källén–Lehmann spectral representation. Unlike the calculation of the single particle quantity, the method based on expansion in terms of a continued fraction is not applicable for the calculation of the vertex functions [11, 59, 318, 319]. Thus the calculation is limited to a rather small number of bath sites or orbitals [157, 287].

For the QMC approach, besides the minus sign problem [320, 321], the main challenge for calculating the vertex is the noise in the measurements, especially at high frequency. This is particularly acute for the hybridization expansion approach. Significant progress has been made to reduce the noise by measuring in a basis of orthogonal polynomials [322–325]. There is continuous improvement on the sampling efficiency and on the ability to attain ergodicity [326–330]. For further discussion of these approaches, we refer the interested readers to the comprehensive review by Gull et al [53].

As we have seen, the different implementations of MSMB approaches have produced very significant results that validate the motivation of the approach. Indeed, more appropriate treatments of nonlocal correlations improve the results both qualitatively and quantitatively. To improve the robustness of the approach and to extend the methods to broader ranges of parameters, further developments are needed to overcome the computational challenges. This may involve new insights on the physics, leading to modified algorithms, or the development of new numerical techniques.

While we discuss several executions of the MSMB approach, an omission in this paper is a definite guideline with pros and cons of the respective methods. In particular, it is desirable to answer the question of which method provides the best results with the least numerical effort. Presently, there are various reasons why it is rather difficult to address this question. First, many of the methods have not been fully investigated, some of them may not even have been optimally implemented. Second, the question of ‘best results’ needs qualification, it is unlikely that there is one method which holds a clear advantage over the others in terms of getting the best results. This can be understood from the point of view that all the methods discussed are based on some form of perturbative expansion on top of effective interacting models. The range of parameters is an important factor in deciding the quality of different expansions. Third, the implementation of a given method also affects the quality of the results. All the methods require the handling of different types of vertex functions. The procedure for storing and approximating the vertex functions can be a non negligible factor in the final results. While there is intense activity on the MSMB approaches, the field is still rather young. We have painted a detailed picture of the landscape in our discussions of the different methods, the nuances within the methods, and possible subtleties across the
numerical approximations involved. In time, we believe the community will push these different implementations to the point of producing a fuller picture; allowing for more transparent comparisons.

Acknowledgments

We are deeply indebted to Mark Jarrell for introducing us to the computational studies of strongly correlated systems. We thank our collaborators over the years, Shuxiang Yang, Hanna Terletska, N S Vidhyadhiraja, Yi Zhang, Cyrill Slezak, Zi-Yang Meng, Harmut Hafermann, Muhammad Aziz Majid, Karlis Mikelsons, Ehsan Khatami, Peng Zhang, Peter Reis, Kuang-Shing Chen, Chinedu Ekuma, Ryky Nelson, Patrick Haase, Nagamalleswararao Dasari, Sheng Feng, Samuel Kellar, Brian Moritz, Paul Kent, Unjong Yu, Dimitris Galanakis, Hartmut Kaiser, Karen Tomko, Thomas Maier and Thomas Pruschke. We thank Eric Dohner for proofreading our manuscript. HF is partially supported by NSF PHY-2014023. KMT and JM are partially supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0017861. KMT is also supported by NSF DMR-1728457 and OAC-1931445.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Herbert F Fotso  https://orcid.org/0000-0001-7952-6256

References

[1] Wolf S A, Awschalom D D, Buhrman R A, Daughton J M, von Molnár S, Roukes M L, Chtchelkanova A Y and Tregear D M 2001 Science 294 1488
[2] Morosan E, Néel, D, and Néel, S 2012 Adv. Mater. 24 4896–923
[3] Dagotto E and Tokura Y 2008 MRS Bull. 33 1037–45
[4] Tokura Y and Nagosa N 2000 Science 288 462
[5] Dagotto E 2005 Science 309 257
[6] Bednorz J G and Müller K A 1986 Z. Phys. B 64 189
[7] Zhang F C and Rice T M 1988 Phys. Rev. B 37 3759
[8] Lieb E H and Wu F Y 2003 Physica A 321 1
[9] Lieb E H and Wu F Y 1968 Phys. Rev. Lett. 20 1445
[10] Maier T, Jarrell M, Pruschke T and Hettler M H 2005 Rev. Mod. Phys. 77 1027
[11] Georges A, Kotliar G, Krauth W and Rozenberg M J 1996 Rev. Mod. Phys. 68 13
[12] Slezak C, Jarrell M, Maier T and Deisj J 2009 J. Phys.: Condens. Matter 21 435604
[13] Mott N F 1968 Rev. Mod. Phys. 40 677
[14] Mott N 1982 Proc. R. Soc. A 382 1
[15] Landau L D 1959 JETP 3 920
[16] Landau L D 1957 JETP 5 101
[17] Landau L D 1959 JETP 8 70
[18] Hartree D R 1928 Math. Proc. Camb. Phil. Soc. 24 111–32
[19] Fock V 1930 Z. Phys. 61 126
[20] Fock V 1930 Z. Phys. 62 795
[21] Slater J C 1930 Phys. Rev. 35 210
[22] Lee P A, Nagaosa N and Wen X-G 2006 Rev. Mod. Phys. 78 17
[23] Bohm D and Pines D 1951 Phys. Rev. 82 625
[24] Pines D and Bohm D 1952 Phys. Rev. 85 338
[25] Bohm D and Pines D 1953 Phys. Rev. 92 609
[26] Gell-Mann M and Brueckner K A 1957 Phys. Rev. 106 364
[27] Landau L D, Abrikosov A A and Khalatnikov I M 1954 Dokl. Akad. Nauk 95 497
[28] Landau L D, Abrikosov A A and Khalatnikov I M 1954 Dokl. Akad. Nauk 95 773
[29] Landau L D, Abrikosov A A and Khalatnikov I M 1954 Dokl. Akad. Nauk 95 1177
[30] Bickers N E 1998 Numerical Methods for Lattice Quantum Many-Body Problems ed D J Scalapino (Reading, MA: Addison-Wesley) pp 237–96
[31] Bickers N E and White S R 1991 Phys. Rev. B 43 8044
[32] Bickers N E and Scalapino D J 1993 Ann. Phys. 193 206
[33] Bauer B, Wecker D, Millis A J, Hastings M B and Troyer M 2016 Phys. Rev. X 6 031045
[34] Rungger I et al 2020 arXiv:1910.04735 [quant-ph]
[35] Keen T, Maier T, Johnston S and Lougovski P 2019 arXiv:1910.09512 [quant-ph]
[36] McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 New J. Phys. 18 023023
[37] Soven P 1967 Phys. Rev. 156 809
[38] Shiba H 1971 Prog. Theor. Phys. 46 77
[322] Hafermann H, Patton K R and Werner P 2012 Phys. Rev. B 85 205106
[323] Hafermann H 2014 Phys. Rev. B 89 235128
[324] Gull E, Iskakov S, Krivenko I, Rusakov A A and Zgid D 2018 Phys. Rev. B 98 075127
[325] Huang I. 2016 Chin. Phys. B 25 117101
[326] Sémon P, Yee C-H, Haule K and Tremblay A-M S 2014 Phys. Rev. B 90 075149
[327] Sémon P, Sordi G and Tremblay A-M S 2014 Phys. Rev. B 89 165113
[328] Gunacker P, Wallerberger M, Gull E, Hausoel A, Sangiovanni G and Held K 2015 Phys. Rev. B 92 155102
[329] Melnick C, Sémon P, Yu K, D’Imperio N, Tremblay A-M and Kotliar G 2020 arXiv:2010.08482
[330] Gunacker P, Wallerberger M, Ribic T, Hausoel A, Sangiovanni G and Held K 2016 Phys. Rev. B 94 125153