Metal-insulator transition in the 2D Hubbard model: dual fermion approach with the Lanczos exact diagonalization

Arata Tanaka

1Department of Quantum Matter, ADSM, Hiroshima University, Higashi-hiroshima 739-8530, Japan

(Dated: December 13, 2018)

In this study the metal-insulator transition in the square-lattice Hubbard model at half-filling is revisited in relation to the DOS and spectral functions by means of the ladder dual fermion approximation (LDF). For this purpose, a new expression of the two-body Green’s function in the form of resolvents is proposed, which provides tractable and efficient means to calculate the local vertex function with the Lanczos exact diagonalization (ED) method. This makes it possible to use the Lanczos ED method as a solver of the effective impurity Anderson model for LDF, opening up the way to access low temperatures and obtain accurate DOS and spectral functions on the real frequency axis for these perturbative extensions of the dynamical mean field theory. It is found that for $U \leq 3.5t$, as temperature decreases, the pseudogap formation due to antiferromagnetic correlations in the quasiparticle peak of the spectral function occurs at the X point [${\mathbf{k} = (\pi, 0)}$], spreads through the Fermi surface and ends at the $M_2$ point [${\mathbf{k} = (\pi/2, \pi/2)}$]. The almost simultaneous creation of the pseudogap and the loss of the Fermi liquid feature is consistent with that expected in the Slater regime. Although the pseudogap still appears in the quasiparticle-like single peak for $U \geq 4.0t$, the Fermi-liquid feature is partially lost on the Fermi surface already at higher temperatures as expected in the Mott-Heisenberg regime, in which local spins are preformed at high temperatures.

A sharp crossover from a pseudogap phase to a Mott insulator at finite $U^* \approx 4.7t$ is found to occur below the temperature of the pseudogap formation similar to a previous study with the non-linear $\sigma$ model approach.

PACS numbers:

I. INTRODUCTION

Nowadays the dynamical mean field theory (DMFT) is one of the most powerful methods to describe electronic properties in strongly correlated electron systems. Although DMFT is exact in infinite dimensions, spatial correlation effects, which are absent in DMFT, similar to the ordinary mean field theory, play crucial roles in finite dimensions in some fascinating cases, such as criticality due to thermal or quantum fluctuations, unconventional superconductors and metal-insulator transitions (MIT) in low dimension. In recent years, to include spatial correlations, numerous extensions of DMFT have been developed. In DMFT, the problem of electrons on the lattice is mapped to the effective impurity Anderson model (IAM), where the interaction is explicitly considered on one of the sites (the impurity site) and rest of the sites are replaced by an effective medium. In the cluster extensions of DMFT such as the cellular DMFT (CDMFT) or the dynamical cluster approximation (DCA), the impurity site is replaced by a cluster and spatial correlations within the cluster are considered.

Other attempts to include spatial correlations are perturbative extensions of DMFT and among of them, there is a class of methods in which the local vertex functions instead of the bare interaction are used as the diagrammatic elements of the perturbation. Those in this category are, for example, the pioneering work of Kusunose, the dynamical vertex approximation (DTA), the dual fermion approximation (DFA), the dual boson approximation (DBA), the one-particle irreducible approach (1PI), the DMFT to functional renormalization group approach (DMFT$_c$RG), and the triply irreducible local expansion (TRILEX).

To solve the effective IAM for DMFT, various numerical methods have been developed. In the perturbative extensions of DMFT mentioned above, the two-body Green’s function of the effective IAM is further required to obtain the four-point local vertex function. Efficient schemes to calculate the two-body Green’s function have already been developed and applied to DFA with the continuous-time quantum Monte Carlo (QMC) and the ordinary exact diagonalization (ED) methods. The QMC methods have difficulty in accuracy particularly in low temperatures because of the statical errors. On the other hand, for the ordinary ED technique, the method so far proposed is that with the Lahmann representation and as will be discussed later, it has problem in efficiency and the limitation of the number of the many-body basis functions.

One of the purposes of this paper is to present a new formula for the two-body Green’s function. This renders efficient and accurate means to calculate the local vertex function and makes it possible to use the Lanczos ED technique as a solver of the effective IAM required for DFA and similar perturbative extensions of DMFT.

The MIT in the two-dimensional (2D) Hubbard model at half filling still remains to be a subject of debate even after decades of extensive studies. Although it is well established that the ground state of the half-filled 2D Hubbard model on a square lattice has long-range antiferromagnetic (AFM) order, the difficulty mainly
arises from the fact that long-range AMF order cannot be stable at finite temperatures in two dimension because of the Mermin-Wagner theorem. The MIT has been discussed in relation to the Slater and Mott-Heisenberg mechanisms. In the Slater regime, the gap formation is essentially that of one-body picture, i.e., the Brillouin zone folding caused by the AFM ordering. Hence, the spin and charge degrees of freedom are entwined and both the spin and charge excitations have the same energy scale. In contrast, in the Mott-Heisenberg regime, localized spins preformed at high temperatures and AFM ordering occurs through the exchange coupling between these local spins. Since this picture is essentially based on many-body theories, the energy scale of the charge ordering occurs through the exchange coupling between these operators and also those of the corresponding eigenvector of the Hamiltonian; \( \langle \sigma^+_i \sigma^-_j \rangle \) and \( \langle \sigma^+_i \sigma^-_j \rangle \) are generally different. In a study with the non-linear \( \sigma \) model approach, the pseudogap is formed at finite temperatures. While a clear insulating gap opens in the Mott-Heisenberg model and a Mott gap opens for all \( U > 0 \) in the Slater regime and thus the MIT point \( U_c \approx 4.2t \) is expected to be positioned at the boundary between these two regimes. On the other hand, Anderson has proposed that whole low temperature physics of the 2D Hubbard model is mapped onto the 2D Heisenberg model and a Mott gap opens for all \( U > 0 \). 

DMFT predicts the first-order Mott MIT at finite temperatures with a second-order critical endpoint \( U_c \approx 10t \) at \( T = 0 \) when the paramagnetic (PM) state is assumed, which is essentially the same to the MIT in infinite dimension. CDMFT, the variational cluster approximation (VCA) and the second-order DFA, which are only capable for short-range spatial correlations within the cluster, also find the first-order MIT at finite temperatures similar to DMFT but with substantially smaller critical values \( U_c \approx 6t \). In these theories, however, the AMF insulating state have finite Néel temperatures and the region where the first-order MIT line presence in the \( U-T \) phase diagram is replaced by the AMF insulating phase when the solutions are not constrained to the PM state.

In the studies by means of DFA and extrapolated lattice QMC, the ladder dual fermion approximation (LDFA), and the two-particle self-consistent approximation (TPSC), which incorporate the effects of long-range correlations and fulfill the Mermin-Wagner theorem, the MIT occur at much smaller \( U \) compared to the CDMFT results. In the QMC calculations of finite-size clusters, has also found the pseudogap at least \( U \geq 2.0t \). Schäfer et. al. suggest in their combined study of DFA and lattice QMC that \( U_c = 0 \) for \( T \to 0 \) and thus no MIT occurs at any \( U > 0 \) similar to the 1D Hubbard model. However, the formation of the pseudogap does not necessarily indicate insulating behavior in low temperatures and examination of subtle changes in states inside the gap as a function of temperature is required to verify whether \( U_c \) stays finite or \( U_c = 0 \). For this reason, it is essential to obtain precise information of the DOS and spectral function in the vicinity of the Fermi level to understand the MIT in the 2D Hubbard model. Although there are already several LDFA works and that with the diagrammatic Mote Carlo approach on the 2D Hubbard model at half filling, detailed investigation on the DOS and spectral function with LDFA at low temperatures is still lacking.

In addition to a new formula for the two-body Green’s function, the other purpose of this paper is to investigate the DOS and spectral function of the square-lattice Hubbard model at half filling by means of LDFA to elucidate the behavior and origin of the MIT. In particular, utilizing the newly developed Lanczos ED scheme to calculate the two-body Green’s function as the solver of the effective IAM for LDFA, it is possible to access large \( U \) and low temperature region of the \( U-T \) phase diagram where previous studies still have not reached and obtain results with unprecedented accuracy. It is found that a sharp crossover from a pseudogap phase to a Mott insulator around \( U^* \approx 4.7t \) occurs below the temperature of the pseudogap formation.

The rest of this paper is structured as follows. In Sec. III the new formula of two-body Green’s function is presented. Section IV describes how to calculate the two-body Green’s function approximately with the Lanczos algorithm with the new formula. In Sec. V a brief overview of LDFA is given and some technical points specific to the Lanczos ED method are presented. In Sec. VI a detailed description of the maximum entropy method used in this study is given. In Sec. VII results of LDFA calculations of the 2D Hubbard model are presented. The paper is closed with a discussion in Sec. VIII and a brief summary in Sec. IX. Derivations of the new formula in Sec. III and the update formula of the hybridization function in Sec. IV are referred to two appendices.

II. EXPRESSION OF TWO-BODY GREEN’S FUNCTION IN THE FORM OF RESOLVENTS

As an expression for two-body Green’s function, the Lahmann representation has been used for the ordinary ED technique. However, the structure of the formula is not suitable for the Lanczos exact diagonalization method. The expression has terms like

\[
\sum_{lmnk} \left( e^{-\beta E_l/Z} \langle k | c_2 | l \rangle \langle l | c_1 | m \rangle \langle m | c_3^+ | n \rangle \langle n | c_4^+ | k \rangle \right)
\]

(1)

where \( |l\rangle \) and \( E_l \) are the \( l \)th lowest eigenvalue and corresponding eigenvector of the Hamiltonian; \( c_i \) and \( c_i^+ \) represent creation and annihilation operators of an electron on the impurity site, respectively and index \( i = 1, 2, 3 \) of these operators and also those of \( g_{12} \) and \( \chi_{1234} \) later appeared in this section are shorthand notation of combined index for the spin \( \sigma \) and orbital \( \alpha \) degree of freedom, e.g., \( c_1 \equiv c_{\sigma_1,\alpha_1} \), \( \omega_i \) denote the fermionic Matsubara
frequency: $\omega_i = \pi(2n_i + 1)/\beta$. Since these terms contain the factors in the denominator with the difference of two eigenvalues, e.g., $E_m - E_n$, terms with nearly degenerated two high-energy eigenvectors $E_m$ and $E_n$ can have large contributions. Therefore the precise eigenvalues of whole energy range are necessary. Since the Lanczos ED is accurate only for low-energy eigenvectors, an alternative expression for the two-body Green’s function is desired.

On the other hands, for the Fourier transform of the one-body Green’s function $g_{12}(\tau) \equiv -\langle T[c_1(\tau)c_2^\dagger(0)] \rangle$ the expression with resolvents

$$
g_{12}(i\omega) = \frac{1}{Z} \sum_{l,m} e^{-\beta E_l} \left\{ \frac{\langle l|c_1|m\rangle\langle m|c_1^\dagger|l\rangle}{i\omega + E_l - E_m} + \frac{\langle l|c_2|m\rangle\langle m|c_1|l\rangle}{i\omega + E_m - E_l} \right\}$$

(2)

$$
= \frac{1}{Z} \sum_{l} e^{-\beta E_l} \left\{ \frac{\langle l|c_1|l\rangle}{i\omega + E_l - \mathcal{H}} \frac{1}{c_1^\dagger} - \langle l|c_2|l\rangle \frac{1}{-i\omega + E_l - \mathcal{H}} \right\} c_1^\dagger
$$

(3)

is applicable to the Lanczos ED technique and has already adopted in the DMFT studies. Since eigenvector $|l\rangle$ only within energy range of the thermal excitation contribute due to the presence of the Boltzmann factor $e^{-\beta E_l}$, $|l\rangle$ and $E_l$ can be accurately calculated at low temperatures with the Lanczos ED technique. The resolvents in Eq. (3) can be transformed into continued fractions using the Lanczos algorithm and this continued fraction can be terminated typically several hundreds floors to obtain accurate results even for systems with $\sim 10^7$ basis functions. This technique is called the recursion method. The procedure is also equivalent to replacing $E_m$ and $|m\rangle$ in Eq. (2) by those approximately obtained within a subspace spanned by the Lanczos vectors, i.e., the Krylov subspace.

The expression of the Fourier transform of the two-body Green’s function

$$
\chi_{1234}(\tau_1, \tau_2; \tau_3, \tau_4) \equiv \langle T[c_1(\tau_1)c_2(\tau_2)c_3^\dagger(\tau_3)c_4^\dagger(\tau_4)] \rangle,
$$

(4)

presented here consists of terms each of which has three or two resolvents in the form $1/(i\omega + E_l - \mathcal{H})$ and the factor $e^{-\beta E_l}$ sharing the same eigenvalue $E_l$. Because of this feature, only terms with $E_l$ within the energy range of the thermal excitation contribute similar to Eq. (3) of the one-body Green’s function. Therefore, unlike the Lamm representation, the denominators of these terms are always large with the high-energy eigenvectors of $\mathcal{H}$ and thus no significant contribution of high-energy eigenvectors is expected. This makes it possible to approximate the eigenvectors of $\mathcal{H}$ by those calculated within the Krylov subspace constructed by the Lanczos algorithm, which is less accurate in high-energy eigenvectors.

The detailed derivation of the expression is in Appendix A. The expression can be separated into three components in terms of the similarity to the three scattering channels: those of the horizontal $\chi^{ph}$, and vertical $\chi^{ph}$ particle-hole, and the particle-particle $\chi^{pp}$ types:

$$
\chi_{1234} = \chi^{ph}_{1234} + \chi^{ph}_{1234} + \chi^{pp}_{1234}.
$$

(5)

Each of the components can be given as

$$
\chi^{ph}_{1234}(i\omega_1, i\omega_2; i\omega_3, i\omega_4) =
$$

$$
- \frac{1}{Z} \sum_{l} e^{-\beta E_l} \langle l|\left\{ (c_1|c_4^\dagger)_{E_l+i\omega_1} - (c_4^\dagger|c_1)_{E_l-i\omega_1} \right\} (c_2^\dagger|c_3^\dagger)_{E_l+i\omega_4} - (c_3^\dagger|c_2)_{E_l-i\omega_4} \right\} |l\rangle
$$

$$
- \frac{1}{Z} \sum_{l} e^{-\beta E_l} \langle l|\left\{ (c_2|c_3^\dagger)_{E_l+i\omega_2} - (c_3^\dagger|c_2)_{E_l-i\omega_2} \right\} (c_1|c_4^\dagger)_{E_l+i\omega_4} - (c_4^\dagger|c_1)_{E_l-i\omega_4} \right\} |l\rangle
$$

$$
+ \frac{1}{Z} \sum_{l,m} \delta_{E_l, E_m} e^{-\beta E_l} \langle l|\left\{ (c_1|c_4^\dagger)_{E_l+i\omega_1} - (c_4^\dagger|c_1)_{E_l-i\omega_1} \right\} |m\rangle \langle m|\left\{ (c_2^\dagger|c_3^\dagger)_{E_l+i\omega_4} - (c_3^\dagger|c_2)_{E_l-i\omega_4} \right\} |l\rangle
$$

$$
+ \frac{1}{Z} \sum_{l,m} \delta_{E_l, E_m} e^{-\beta E_l} \langle l|\left\{ (c_1^\dagger|c_4)_{E_l+i\omega_1} - (c_4^\dagger|c_1)_{E_l-i\omega_1} \right\} |m\rangle \langle m|\left\{ (c_2^\dagger|c_3^\dagger)_{E_l+i\omega_4} - (c_3^\dagger|c_2)_{E_l-i\omega_4} \right\} |l\rangle
$$

$$
+ \beta \delta_{\omega_1, \omega_4} \frac{1}{Z} \sum_{l,m} \delta_{E_l, E_m} e^{-\beta E_l} \langle l|\left\{ (c_1^\dagger|c_4)_{E_l+i\omega_1} - (c_4^\dagger|c_1)_{E_l-i\omega_1} \right\} |m\rangle \langle m|\left\{ (c_2^\dagger|c_3)_{E_l+i\omega_4} - (c_3^\dagger|c_2)_{E_l-i\omega_4} \right\} |l\rangle
$$

(6)
\[\chi_{1234}(i\omega_1, i\omega_2; i\omega_3, i\omega_4) =\]
\[+ \frac{1}{Z} \sum_l e^{-\beta E_l} |l\rangle \langle (c_1||c_2)_{E_l+i\omega_1} - (c_2||c_1)_{E_l+i\omega_2} |(c_1||c_3)_{E_l+i\omega_4} - (c_3||c_1)_{E_l+i\omega_3} - (c_2||c_2)_{E_l-i\omega_4} - (c_3||c_2)_{E_l-i\omega_3} - (c_1||c_2)_{E_l-i\omega_2} - (c_2||c_2)_{E_l-i\omega_2} |l\rangle\]
\[+ \frac{1}{Z} \sum_l e^{-\beta E_l} |l\rangle \langle (c_2||c_3)_{E_l+i\omega_2} - (c_2||c_1)_{E_l-i\omega_4} |(c_1||c_3)_{E_l+i\omega_3} - (c_3||c_1)_{E_l-i\omega_3} - (c_2||c_2)_{E_l-i\omega_2} - (c_3||c_2)_{E_l-i\omega_2} |l\rangle\]
\[+ \frac{1}{Z} \delta_{E_l, E_m} e^{-\beta E_l} |l\rangle \langle (c_1 \Delta c_3)_{E_l+i\omega_1} - (c_3 \Delta c_1)_{E_l+i\omega_3} |m\rangle \langle (c_2 \Delta c_4)_{E_l+i\omega_2} - (c_4 \Delta c_2)_{E_l+i\omega_2} |l\rangle\]
\[+ \frac{1}{Z} \delta_{E_l, E_m} e^{-\beta E_l} |l\rangle \langle (c_1 \Delta c_3)_{E_l-i\omega_1} - (c_3 \Delta c_1)_{E_l-i\omega_3} |m\rangle \langle (c_2 \Delta c_4)_{E_l-i\omega_2} - (c_4 \Delta c_2)_{E_l-i\omega_2} |l\rangle\]
\[+ \frac{1}{Z} \delta_{E_l, E_m} e^{-\beta E_l} |l\rangle \langle (c_1 \Delta c_3)_{E_l+i\omega_1} - (c_3 \Delta c_1)_{E_l+i\omega_3} |m\rangle \langle (c_2 \Delta c_4)_{E_l+i\omega_2} - (c_4 \Delta c_2)_{E_l+i\omega_2} |l\rangle\]
\[+ \frac{1}{Z} \delta_{E_l, E_m} e^{-\beta E_l} |l\rangle \langle (c_1 \Delta c_3)_{E_l-i\omega_1} - (c_3 \Delta c_1)_{E_l-i\omega_3} |m\rangle \langle (c_2 \Delta c_4)_{E_l-i\omega_2} - (c_4 \Delta c_2)_{E_l-i\omega_2} |l\rangle\]

where, the operators in the form of \(\langle A||B\rangle\) and \(A \triangle B\) are the abbreviations of those contain one and two resolvents, respectively, and are defined as

\[(A||B) \equiv A \frac{1}{z - \mathcal{H}'},\]
\[(A \triangle B) \equiv A \frac{1}{z - \mathcal{H}} \frac{1}{z' - \mathcal{H}}.\]

The resolvent with the Hamiltonian \(\mathcal{H}'\) in the denominator of Eq. (9) is the same to that of \(\mathcal{H}\) except that all the eigenvectors \(|l\rangle\) whose eigenvalue \(E_l\) are equal to \(z\) are projected out as

\[\langle A||B\rangle \equiv A \frac{1}{z - \mathcal{H}} \frac{1}{z' - \mathcal{H}}.\]

The first two lines of the right-hand side of each of Eqs. (9)-(10) contain 8 terms with three resolvents such as

\[-\sum_l \frac{e^{-\beta E_l}}{Z} |l\rangle \langle (c_1||c_3)_{E_l+i\omega_1} \rangle \langle (c_2||c_4)_{E_l+i\omega_3} \rangle \langle (c_1||c_2)_{E_l-i\omega_4} \rangle \langle (c_3||c_4)_{E_l-i\omega_3} \rangle \langle (c_2||c_2)_{E_l-i\omega_2} \rangle \langle (c_3||c_2)_{E_l-i\omega_2} \rangle \langle (c_1||c_2)_{E_l-i\omega_2} \rangle \langle (c_2||c_2)_{E_l-i\omega_2} |l\rangle\]

In total \(4! = 24\) of terms of this kind exist and we call them the major terms. Each term contains a resolvent with a bosonic Matsubara frequency, e.g., that with \(i(\omega_3 - \omega_2)\) in Eq. (12), and for this resolvent, eigenvectors with eigenvalue \(E_l\) are projected out to avoid divergence at zero frequency. The proper treatment of these special cases further requires 36 counter terms (for details, see Appendix A) and there are two kinds of them: one is those consist of products of two factors containing one
and two resolvents as
\[
\sum_{l,m} \frac{e^{-\beta E_l}}{Z} (l|c_1^l E_{l+i\omega_1} |m) (m|c_2^l E_{l+i\omega_2} |l)
\]
\[
= \sum_{l,m} \frac{e^{-\beta E_l}}{Z} (l|c_1^l E_{l+i\omega_1} |m) (m|c_2^l E_{l+i\omega_2} |l)
\times (m|1_{i\omega_2 + E_l - \mathcal{H}}^l (m) \frac{1}{i\omega_2 + E_l - \mathcal{H}^l |c_1^l |l). \quad (13)
\]
and the other kind of the counter terms have the form
\[
\beta \delta_{\omega_1,\omega_4} \sum_{l,m} \frac{e^{-\beta E_l}}{Z} (l|c_1^l E_{l+i\omega_1} |m) (m|c_2^l E_{l+i\omega_2} |l)
\]
\[
= \beta \delta_{\omega_1,\omega_4} \sum_{l,m} \frac{e^{-\beta E_l}}{Z} (l|c_1^l E_{l+i\omega_1} |m) (m|c_2^l E_{l+i\omega_2} |l)
\times (m|1_{i\omega_2 + E_l - \mathcal{H}^l} (m) \frac{1}{i\omega_2 + E_l - \mathcal{H}^l |c_1^l |l). \quad (14)
\]
Note that if the eigenenergies of wave functions with different electron numbers are degenerated, the counter terms in Eq. (3) can have non-zero values even without finite superconducting order parameter. This can happen if the system possesses the electron-hole symmetry, e.g., half-filled square-lattice Hubbard model with the nearest neighbor hopping in this study.

III. APPROXIMATION OF TWO-BODY GREEN’S FUNCTION WITH LANCZOS ALGORITHM

Having the new expression in hand, in this section we discuss how to calculate the two-body Green’s function approximately with the Lanczos algorithm. The Lanczos algorithm is a unitary transformation, which converts a symmetric or Hermitian matrix \( \mathcal{H} \) into a tridiagonal form:

\[
T^{(n)} = \begin{pmatrix}
  a_1 & b_1 & 0 & \cdots & 0 \\
  b_1 & a_2 & b_2 & \ddots & \vdots \\
  0 & b_2 & a_3 & \cdots & 0 \\
  \vdots & \ddots & \ddots & \ddots & b_{n-1} \\
  0 & \cdots & 0 & b_{n-1} & a_n
\end{pmatrix}.
\quad (15)
\]

Starting from a properly chosen initial vector \( |v_1\rangle \), it creates one of the orthonormal basis vector \( |v_n\rangle \) in every iteration step, and at \( n \)th step these basis vectors span the Krylov subspace \( \mathcal{K}^n (|v_1\rangle, \mathcal{H}) = \text{span} \{ |v_1\rangle, \mathcal{H}|v_1\rangle, \mathcal{H}^2|v_1\rangle, \ldots, \mathcal{H}^{n-1}|v_1\rangle \} \). In practice, because of round off error, the orthogonality of the vectors \( |v_n\rangle \) breaks midway through the iteration. This occurs as soon as the lowest eigenvector converges and one may set the criterion to terminate the iteration when this convergence is reached:

\[
|b_{n-1} s_{1,n}| < \epsilon_{\text{lan}},
\quad (16)
\]

where \( s_{i,j} \) is the \( j \)th component of the eigenvector of \( T^{(n)} \) with the \( i \)th lowest eigenvalue \( \theta_i \), i.e., \( T^{(n)} |s_{i}\rangle = \theta_i |s_{i}\rangle \). In this study, the threshold \( \epsilon_{\text{lan}} = \alpha_{\text{lan}} \sqrt{N_{\text{sys}}} \) is assumed, where \( N_{\text{sys}} \) is the order of matrix \( \mathcal{H} \), \( \epsilon \) denotes the machine accuracy \( \epsilon = 10^{-15} \) and \( \alpha_{\text{lan}} = 10 \). The \( i \)th lowest eigenvalue \( E_i \) and corresponding eigenvector \( |i\rangle \) of \( \mathcal{H} \) can be approximated as

\[
|i\rangle \approx \sum_{j=1}^{n} s_{i,j} |v_j\rangle, \quad E_i \approx \theta_i.
\]

High accuracy (typically more than 10 digits) can be expected for the eigenvector \( |1\rangle \) of the lowest eigenvalue \( E_1 \) and the rest of them are less accurate.

Now, let us consider the major term with the form

\[
\sum_{l} \frac{e^{-\beta E_l}}{Z} (l|\mathcal{O}_1 \frac{1}{i\omega + E_l - \mathcal{H}^l} \mathcal{O}_2 \\
\times \frac{1}{i\nu + E_l - \mathcal{H}^l} \mathcal{O}_3 \frac{1}{i\omega' + E_l - \mathcal{H}^l} \mathcal{O}_4 |l), \quad (17)
\]

where each \( \mathcal{O}_i \) \((i = 1, 2, 3, 4)\) is one of \( c_1, c_2, c_3^\dagger \) or \( c_4^\dagger \) and \( \nu \) is the bosonic and \( \omega \) and \( \omega' \) are the fermionic Matsubara frequencies. The eigenvectors \( |l\rangle \) need to be included in the calculations are limited to those of low energies by the Boltzmann factor \( e^{-\beta E_l} \). Note that high accuracy is particularly needed for \( E_l \) and \( |l\rangle \) since it affects later calculations. Therefore, instead of calculating all the \( E_l \) and \( |l\rangle \) required at a time by the Lanczos algorithm described above, it is preferable to use the so called restart Lanczos method, where only one lowest eigenvector is calculated at a time and repeat the same Lanczos procedure except in each step \(|v_n\rangle \) is orthogonalized to all the previously obtained eigenvectors. In this way, all the required \( E_l \) and \( |l\rangle \) can be calculated with high precision.

One of the three resolvents in Eq. (17) on the left can be obtained approximately using \( N_L \) eigenvalues \( E_m^L \) and eigenvectors \( |m^L\rangle \) generated by combined use of the restart Lanczos method for low laying eigenvectors within the reach of the thermal excitation and the ordinary Lanczos method with the initial vector \( \mathcal{O}_1^l |l\rangle \) for rest of high-energy eigenvectors as

\[
\frac{1}{i\omega + E_l - \mathcal{H}} \approx \sum_{m=1}^{N_L} |m^L\rangle \frac{1}{i\omega + E_l - E_m^L} \langle m^L |.
\quad (18)
\]

The same can be done for the resolvent on the right using \( N_R \) eigenvalues \( E_m^R \) and eigenvectors \( |m^R\rangle \) generated by these Lanczos methods with the initial vector \( \mathcal{O}_2^l |l\rangle \). Although the high-energy eigenvectors obtained with the ordinary Lanczos method are less accurate compared to the low-energy eigenvectors calculated with the restart
Lanczos method, the resultant left and right resolvents have proper asymptotic behavior. For instance,

$$\frac{1}{z - H}O_4|l\rangle = \sum_{n=0}^{\infty} \frac{1}{z^{n+1}}H^nO_4|l\rangle$$

(19)

for the right resolvent. Obviously the expansion is correct up to the $N_0$th order, since their coefficients belong to the Krylov space $K^{N_0} (O_4|l\rangle, H)$.

For the right resolvent, the center, expansions through the left and right resolvents with different energy scales are required to take into account. For this purpose, vectors

$$|v_\alpha^L\rangle = \sum_{m=1}^{N_L} O_2^m|m^L\rangle \frac{\langle m^L|O_1^L|l\rangle}{(\Omega_\alpha + E_l - E_m^L)^2 + \gamma_\alpha^2},$$

$$|v_\alpha^R\rangle = \sum_{m=1}^{N_R} O_3|m^R\rangle \frac{\langle m^R|O_4|l\rangle}{(\Omega_\alpha + E_l - E_m^R)^2 + \gamma_\alpha^2}$$

(20)

with different energies $\Omega_\alpha$ are calculated and orthogonalized in each other to use them as the initial vectors of the band Lanczos method\textsuperscript{19}, which is an extension of the Lanczos method with multiple initial vectors. The Lanczos vectors need to be orthogonalized to the eigenvector $|l\rangle$ and, if exist, all the degenerated eigenvectors with the eigenvalue $E_l$ [see Eq. (17)]. It is essential to chose one of $\Omega_\alpha$ to be $\Omega_\alpha \gg E_{\text{max}} - E_l$ to make $\chi$ have proper asymptotic behavior, where $E_{\text{max}}$ is the maximum eigenvalue of $H$.

For example, in the calculations of the two-dimensional Hubbard model in this study we adopted four energy points $\Omega_1 = 0, \Omega_2 = 0.2W, \Omega_3 = 0.4W$ and $\Omega_4 = 4W$ with all $\gamma_\alpha = 0.1W$, where the effective band width $W$ is defined as

$$W \equiv \sqrt{U^2 + 64t^2}.$$ (21)

Only few energy points $\Omega_\alpha$ are required to obtain accurate results in this case and excessive energy points tend to lesser results.

Once $N_C$ eigenvalues $E_n^C$ and eigenvectors $|m^C\rangle$ are obtained with the band Lanczos method, we can approximate the major term in the form of Eq. (17) as

$$\frac{1}{Z} \sum_{i=1}^{N_L} \sum_{m=1}^{N_C} \sum_{n=1}^{N_C} \sum_{m'=1}^{N_R} e^{-\beta E_l} (l|O_4|m^L\rangle \langle m^L|O_2^m|n^C\rangle (n^C|m^C\rangle m^R|m^R\rangle (m^R|O_4|l\rangle \langle i\omega + E_l - E_n^C)(i\nu + E_l - E_n^R)(i\omega' + E_l - E_R^m)$$

(22)

which bears a resemblance to Eq. (11) but all the three factors in the denominator and $e^{-\beta E_l}$ shares the same $E_l$. This makes terms with all the eigenvalues $E_l$, $E_{m'}$, $E_n^C$ and $E_R^m$, being within the energy range of thermal excitation mostly contribute and for these terms, one can use the restart Lanczos method to obtain accurate eigenvectors and eigenvalues. For the rest of terms with high-energy excitations, the combined use of the ordinary and band Lanczos methods as described ensures accurate asymptotic behavior. Since the typical number of the Lanczos vectors, $N_L$, $N_C$, and $N_R$, is several hundreds even for $10^6$ actual basis functions, the method proposed here renders drastic reduction of computational workload over ordinary ED method.

The counter terms, such as Eq. (13) and Eq. (14) can be calculated in the similar way. For instance, the second factor with two resolvents in Eq. (14) can be approximated using $N_R$ eigenvalues $E_n^R$ and eigenvectors $|m^R\rangle$ generated from the band Lanczos algorithm with the initial vectors $|v_\alpha^L\rangle$ and $|v_\alpha^R\rangle$ for all the degenerated eigenvectors $|l\rangle$ and $|m\rangle$ ($E_l = E_m^R$) as

$$\sum_{n=1}^{N_R} \frac{(m|c_2|n^R\rangle \langle n^R|c_1^\dagger|l\rangle)}{(i\omega_2 + E_l - E_n^R)(i\omega_3 + E_l - E_n^R)}.$$ (23)

IV. LADDER DUAL FERMION APPROXIMATION

In this section, a brief overview of LDFA\textsuperscript{8,40} and some technical points specific to the present calculation scheme are provided. The action for the 2D Hubbard model on the square lattice with the nearest neighbor hopping integral $t$ and the on-site Coulomb interaction $U$ is

$$S[\tau, c] = - \sum_{k_\sigma} \bar{c}_{k\sigma}(i\omega + \mu - \epsilon_k)c_{k\sigma} + U \sum_i \int_0^\beta d\tau \bar{c}_{i\uparrow}\bar{c}_{i\sigma}c_{i\sigma}\bar{c}_{i\uparrow},$$ (24)

where $\bar{c}_{k\sigma}$ and $c_{i\sigma}$ are the fermionic Grassmann fields corresponding to the creation (annihilation) operators $c_{k\sigma}^\dagger$ ($c_{k\sigma}$) and $c_{i\sigma}^\dagger$ ($c_{i\sigma}$), respectively; $\omega$ represents the fermionic Matsubara frequency, $\mu$ denotes the chemical potential, and $\epsilon_k = -2t(\cos k_x + \cos k_y)$.

The IAM at site $i$ can be written as

$$S_{\text{imp}}[\tau, c] = - \sum_{\omega \sigma} \bar{c}_{i\omega \sigma}(i\omega + \mu - \Delta_\omega)c_{i\omega \sigma} + U \int_0^\beta d\tau \bar{c}_{i\uparrow\sigma}\bar{c}_{i\uparrow\sigma}c_{i\downarrow\sigma}\bar{c}_{i\downarrow\sigma},$$ (25)
where $\Delta_\omega$ denotes the hybridization function, which is arbitrary at this point. The lattice action in Eq. \[21\] can be represented by the action of the IAM for each site $i$, plus a correction term:

$$S_{[\tau, c]} = \sum_i S_{\text{imp}}[\tau_i, c_i] + \sum_{k, \omega} \tau_{k, \omega} (\epsilon_k - \Delta_\omega)c_{k, \omega}. \tag{26}$$

Instead of directly performing the perturbative calculations with Eq. \[26\], a new fermionic auxiliary field, which is called the dual fermion, $f_{k, \omega}$, is introduced using a Hubbard-Stratonovich transformation. The original action can be mapped onto that of the dual fermion by integrating out the real electron field $c_{k, \omega}$. In this way, one can separate the problem of solving the IMA to obtain the local approximation and the perturbative corrections for the spatial fluctuations, avoiding the double counting of local contributions. The action of $f_{k, \omega}$ within the fourth order is

$$S_d[\bar{f}, f] = -\sum_{k, \omega} \tau_{k, \omega} [G_{k, \omega}^{d, 0}]^{-1} f_{k, \omega} - \frac{1}{4} \sum_{i = 1234} \gamma_{1234}^{(4)} \bar{f}_i f_1 f_2 f_3 f_4, \tag{27}$$

where shorthand notations such as $1 \equiv (\omega_1, \sigma_1)$ are used for the indices; $G_{k, \omega}^{d, 0}$ denotes non-interacting dual-fermion one-body Green’s function, and $\gamma^{(4)}_{1234}$ represents the reducible four-point vertex function of the impurity site for original electrons and they are defined using the local one-body $g_{\omega}$ and two-body $\chi_{1234}$ Green’s functions of the impurity site for original electrons as

$$G_{k, \omega}^{d, 0} = -g_{\omega} + [g_{\omega}^{-1} + \Delta_\omega - \epsilon_k]^{-1}, \tag{28}$$

$$\gamma^{(4)}_{1234} = g_1^{-1} g_2^{-1} \chi_{1234}^{-1} (\beta(\delta_{14} + \delta_{23} - \delta_{13} \delta_{24}) g_1 g_2) g_3^{-1} g_4^{-1}. \tag{29}$$

Note that diagrams containing the six-point vertex function give a negligible contribution. For the sake of convenience, we use notation

$$\gamma^{(\sigma_1, \sigma_2, \sigma_3, \sigma_4)}_{\omega_1, \omega_2} \equiv \gamma^{(4)}_{\omega_1, \sigma_1, \omega_2, \sigma_2, \omega_3, \sigma_3, \omega_4, \sigma_4}. \tag{30}$$

Since we are dealing with the PM state, the system has spin rotational symmetry and the vertex function can be diagonalized with respect to the spin indices and separated into the charge ($S = 0$) and spin ($S = 1$) components:

$$\gamma^{(\text{ch})}_{\omega_1, \omega_2} \equiv \gamma^{\uparrow\uparrow\downarrow\downarrow}_{\omega_1, \omega_2} + \gamma^{\downarrow\uparrow\uparrow\downarrow}_{\omega_1, \omega_2}; \tag{31}$$

$$\gamma^{(\text{sp})}_{\omega_1, \omega_2} \equiv \gamma^{\uparrow\uparrow\downarrow\downarrow}_{\omega_1, \omega_2} - \gamma^{\downarrow\uparrow\uparrow\downarrow}_{\omega_1, \omega_2}; \tag{32}$$

To include effects of long-range spin fluctuations, the ladder diagram of the particle-hole channel is taken into account, which is considered to be the dominant correction to DMFT for the spatial fluctuations in low temperatures at half filling. The Bethe-Salpeter equation of the dual fermion for the charge ($\alpha = \text{ch}$) and spin ($\alpha = \text{sp}$) components are

$$\Gamma^{(\alpha)}_{\omega, \omega'; q_\Omega} = \gamma^{(\alpha)}_{\omega, \omega'; \Omega} + \frac{1}{\beta} \sum_{\omega''} \gamma^{(\alpha)}_{\omega, \omega''; \Omega} \lambda_{q_\Omega \omega''; \Omega} \Gamma^{(\alpha)}_{\omega'', \omega'; \Omega}; \tag{33}$$

where

$$\lambda^{d, 0}_{q_\Omega \omega; \Omega} = -\frac{1}{N} \sum_k G_{k, \omega, \Omega} G_{k + q_\Omega + \omega, \Omega}. \tag{34}$$

We define the effective interaction of each component $\alpha$ as

$$\mathbf{V}_\alpha^{(\alpha)} = \frac{1}{2\beta} \sum_{\omega'} \gamma^{(\alpha)}_{\omega, \omega'; \Omega} \lambda^{d, 0}_{q_\Omega \omega; \Omega} \left( \Gamma^{(\alpha)}_{\omega', \omega; \Omega} - \frac{1}{2} \gamma^{(\alpha)}_{\omega', \omega; \Omega} \right); \tag{35}$$

and the self-energy for the dual fermion can be written as

$$\Sigma^{d}_{k, \omega} = -\frac{1}{\beta N} \sum_{q_\Omega} \gamma^{(\text{ch})}_{\omega, \omega'; \Omega} G^{d}_{q_\Omega} + \frac{1}{\beta N} \sum_{q_\Omega} \left( \mathbf{V}^{(\text{ch})}_{\omega, \omega'} + 3 \mathbf{V}^{(\text{sp})}_{\omega, \omega'} \right) G^{d}_{q_\Omega + q_\Omega + \omega, \Omega}. \tag{36}$$

The Green’s function of the dual fermion is obtained from Dyson’s equation

$$[G^{d}_{k, \omega}]^{-1} = [G^{d, 0}_{k, \omega}]^{-1} - \Sigma^{d}_{k, \omega}. \tag{37}$$

The electron Green’s function can be obtained from its dual counterpart as

$$G_{k, \omega} = [\epsilon_k - \Delta_\omega]^{-1} g_{\omega}^{-1} G_{k, \omega}^{d} g_{\omega}^{-1} [\epsilon_k - \Delta_\omega]^{-1}$$

$$- [\epsilon_k - \Delta_\omega]^{-1}. \tag{38}$$

There is also similar one-to-one correspondence between electrons and their dual counterparts for the higher-order correlation functions. The local corrections can be efficiently included in the impurity problem with a proper choice of $\Delta_\omega$. For the purpose, the condition $\langle G^{d}_{k, \omega} \rangle_k = 0$, with which diagrams containing a local loop vanish, is commonly used, where $\langle \cdots \rangle_k = \langle 1/N \rangle_k \sum \cdots$. This condition is reduced to the self-consistent condition of DMFT for the non-interacting dual fermions and therefore DMFT can be regarded as the lowest order in DFA. The nonlocal corrections can be included in $\Sigma^{d}_{k, \omega}$ by higher orders of dual fermion perturbation theory as already discussed.

In the calculations of the dual fermionic quantities in Eqs. \[35\]-\[37\], the frequency cutoff $-N_\omega \leq n \leq N_\omega$ is set for the variables with the fermionic Matsubara frequency $\omega_n = (2n - 1)\pi/\beta$ and $|m| \leq 2N_\omega - 1$ for those with the bosonic Matsubara frequency $\Omega_m = 2m\pi/\beta$. Choosing $N_\omega \approx W/\beta$ is found to be sufficient to obtain accurate results, where $W$ is the effective band width in Eq. \[21\].

In the ED method, the conduction band of the effective IAM is replaced by discretized $N_0$ energy levels (or bath
Effective IAM to fulfill the condition of the self-consistency loop of $\Sigma$ via a modified Broyden's method is applied to accelerate the convergence of the fluctuation exchange approximation in the particle-hole channel using Eqs. (28)-(37). The one-body Green's function at the impurity site can be calculated using Eqs. (4)-(8) and the prescription in Sec. III. Once the new hybridization function is obtained, the parameter set $\{\varepsilon_l^b, V_l\}$ for the next iteration is determined by minimizing the distant function defined as

$$d = \sum_{p=1}^{2N_\omega} \frac{1}{|z_p|} \left| \Delta_{\text{new}}(z_p) - \Delta(z; \{\varepsilon_l^b, V_l\}) \right|^2,$$

where the set of $2N_\omega$ points $\{z_p\}$ on the complex plane with $\text{Im} \, z_p > 0$ consists of $N_\omega$ points on the imaginary axis at the fermionic Matsubara frequencies $z_n = i(2n - 1)\pi/\beta$ ($n = 1, 2, \cdots, N_\omega$) depicted as the crosses on the imaginary axis, the equally spaced points represented as the dots on the curve consisting of two quarter circle arcs with the radius $R$ connected by the straight line with the length $W$ in Eq. (21) are included in this study.

The detailed derivation of Eq. (41) can be found in Appendix B. Once the new hybridization function is obtained, the parameter set $\{\varepsilon_l^b, V_l\}$ for the next iteration is determined by minimizing the distant function defined as

$$d = \sum_{p=1}^{2N_\omega} \frac{1}{|z_p|} \left| \Delta_{\text{new}}(z_p) - \Delta(z; \{\varepsilon_l^b, V_l\}) \right|^2,$$

where the set of $2N_\omega$ points $\{z_p\}$ on the complex plane with $\text{Im} \, z_p > 0$ consists of $N_\omega$ points on the imaginary axis at the fermionic Matsubara frequencies $z_n = i(2n - 1)\pi/\beta$ ($n = 1, 2, \cdots, N_\omega$) depicted as the crosses on the imaginary axis, the equally spaced points represented as the dots on the curve consisting of two quarter circle arcs with the radius $R = \pi(2N_\omega + 5)/\beta$. This choice of data points alleviates the problems in the accuracy of the results and the stability of the convergence encountered in the above mentioned iteration process. The similar technique is also used in the analytic continuation with MEM in Sec. V. To find the optimal solution of the parameter set $\{\varepsilon_l^b, V_l\}$ is not a straightforward task because of the presence of numerous solutions with nearly the same distance. A genetic algorithm is applied in combination with the conjugate gradient method to improve the slow convergence of the solution.

V. MAXIMUM ENTROPY METHOD

Since the perturbative calculations are performed with the Matsubara frequency in the present formalism for LDFA, the analytic continuation is required to convert the results as functions of the real frequency. For the spectral function $A(\omega) = -(1/\pi)\text{Im} \, G(\omega)$, the relation to
the Green’s function at an arbitrary complex number $z$

\[
G(z) = \int_{-\infty}^{\infty} \frac{1}{z - \omega} A(\omega) \, d\omega
\]  

(43)
can be utilized. If a set of $N_G$ data of the Green’s function $G \equiv \{G(z_1), G(z_2), \ldots\}$ is given as input, one may obtain approximately a set of $N_A$ discretized data of spectral function $A \equiv \{A(\omega_1) \Delta \omega, A(\omega_2) \Delta \omega, \ldots\}$ by solving the linear equation

\[
G = K A,
\]

(44)
where $K$ denotes the $N_G \times N_A$ matrix with $K_{ij} \equiv 1/(z_i - \omega_j)$. Solving Eq. (44) is known to be severely ill-posed problem and the effective number of constraints imposed to $A$ by the equation is far less than $N_G$ within the practical numerical precision. To extract the limited information about $A$ properly, the maximum entropy method based on Bayesian inference is employed\[12\], where the entropic prior as a means of the regularization is introduced to circumvent the problem of the over fitting.

The joint probability of $A$ and hyperparameters $\alpha_{\chi}$ and $\alpha_S$ for given $G$ is described by using Bayes’ theorem as

\[
P(A, \alpha_{\chi}, \alpha_S | G) = P(G | A, \alpha_{\chi}) P(A, \alpha_S) P(\alpha_{\chi}) P(\alpha_S) / P(G).
\]

(45)
The distribution of the sum of squared relative errors

\[
\chi^2 \equiv \sum_{i=1}^{N_G} \left| G_i - \sum_j K_{ij} A_j \right|^2
\]

(46)
is assumed to be represented as a Gaussian function

\[
P(G | A, \alpha_{\chi}) = \frac{\alpha_{\chi}^{N_G/2}}{(2\pi)^{N_G/2} \prod_i |G_i|^2} \exp(-\alpha_{\chi} \chi^2/2).
\]

(47)
Here, $\alpha_{\chi}$ is also optimized as a hyperparameter, since the standard deviation $\sigma = 1/\sqrt{\alpha_{\chi}}$ for the spectral function $A_{\mathbf{k}}(\omega)$ inferred by MEM from the LDFA Green’s function of the 2D Hubbard model has strong momentum $\mathbf{k}$ dependence: while the values of $\sigma$ at $\mathbf{k}$ points on the Fermi surface, i.e., the $X$–$M_2$ line in Fig. 8(d), range from $\sigma = 5 \times 10^{-2}$ to $2 \times 10^{-6}$, those at $\mathbf{k} = (0, 0)$ are from $\sigma = 3 \times 10^{-5}$ to $4 \times 10^{-4}$. The entropic prior is given as

\[
P(A | \alpha_S) \approx \frac{\alpha_S^{N_A/2}}{(2\pi)^{N_A/2} \prod_i m_i^{1/2}} \exp(\alpha_S S),
\]

(48)
where $S$ is the relative entropy between $A$ and the default model $m \equiv \{m_1, m_2, \ldots\}$, and is defined as

\[
S = \sum_{i=1}^{N_A} [A_i - m_i - A_i \log(A_i/m_i)].
\]

(49)
The uniform distribution is adopted for the default model: $m_i = 1/N_A$. The prior probabilities for hyperparameters $P(\alpha_{\chi})$ and $P(\alpha_S)$ are assumed to be constants and $P(G)$ is the normalization factor.

The joint probability of the hyperparameters $\alpha_{\chi}$ and $\alpha_S$ can be obtained inserting Eqs. (46)–(49) into Eq. (45) and integrated it over $A$ within the Gaussian approximation:

\[
\log P(\alpha_{\chi}, \alpha_S | G) \approx \text{const.} + \frac{N_G}{2} \log \alpha_{\chi} + \frac{1}{2} \sum_{i=1}^{N_A} \log \alpha_S A_i^{\text{max}} + \frac{\alpha_{\chi}}{2} \chi^2(A^{\text{max}}) + \alpha_S S(A^{\text{max}}),
\]

(50)
where $A^{\text{max}}$ denotes $A$ at which $P(A, \alpha_{\chi}, \alpha_S | G)$ is the maximum with given values of $\alpha_{\chi}$ and $\alpha_S$. $\lambda_i$ represents the eigenvalue of the matrix $\Lambda$:

\[
\Lambda_{ij} = \sum_{l=1}^{N_G} (A_i^{\text{max}})^{1/2} K_{il} K_{lj} (A_j^{\text{max}})^{1/2} / |G_l|^2.
\]

(51)
Our problem of the analytic continuation is now reduced to find the set of $A$, $\alpha_{\chi}$ and $\alpha_S$ which is at the maximum of $P(A, \alpha_{\chi}, \alpha_S | G)$. To do so, we first set guess values of $\alpha_{\chi}$ and $\alpha_S$ and maximize $P(A, \alpha_{\chi}, \alpha_S | G)$ with respect to $A$. This is equivalent to minimizing $Q(A) \equiv \alpha_{\chi} \chi^2(A)/2 - \alpha_S S(A)$, which can be achieved using the Newton-Raphson method. The calculations are repeated with different values of $\alpha_S$ to find the maximum of $P(\alpha_{\chi}, \alpha_S | G)$ in Eq. (50) with respect to $\alpha_S$ with fixed value of $\alpha_{\chi}$. The golden section search method is applied for this optimization of $\alpha_S$. To find the maximum of $P(\alpha_{\chi}, \alpha_S | G)$ in Eq. (50) with respect to $\alpha_{\chi}$,

\[
\frac{\partial}{\partial \alpha_{\chi}} P(\alpha_{\chi}, \alpha_S | G) = 0
\]

(52)
is calculated assuming the $\alpha_{\chi}$ dependence of $A^{\text{max}}$ is negligible. The resultant equation for the optimal $\alpha_{\chi}$ is

\[
\alpha_{\chi} = \frac{1}{\chi^2(A^{\text{max}})} \left( N_G \sum_{i=1}^{N_A} \frac{\lambda_i}{\lambda_i + \alpha_S / \alpha_{\chi}} \right).
\]

(53)
This can be solved iteratively by inserting previously obtained value of $\alpha_{\chi}$ repeatedly on the right-hand side of the equation. After solving Eq. (53), the optimization of $A$ and $\alpha_S$ follows with this new value of $\alpha_{\chi}$ and again solving Eq. (50). This process is repeated until convergence is reached.

In the LDFA calculations, the $N_z$ data of $G(z_n)$ at the fermionic Matsubara frequencies, i.e. $z_n = i(2n - 1)\pi/\beta$ ($n = 1, 2, \ldots, N_z$) on the imaginary axis are adopted. For the rest of the data, instead of taking them on the Matsubara frequencies, a set of data points placed at the same distance $R$ from the nearest pole of $G(z)$ on the real axis is chosen: equally spaced points on the curve consisting of two quarter circle arcs with the radius $R$ connected by a straight line with the length $W$ in Eq. (21).
As shown in Fig. 2, this choice of data points ameliorates the difficulty of solving Eq. (44) and the accuracy of $A(\omega)$ obtained is improved, particularly in structures away from the Fermi level. The value of $R$ adopted in this study is $R = 2.7 W$ except for $U = 3.0$, for which $R = 1.8 W$ is used and the number of these additional points is $N_R \approx R\beta$.

Note that altering the default model other than the uniform distribution scarcely affects the results. What is even more important is the model for $\chi^2$ in Eq. (46), where the square sum of the relative errors is assumed instead of the absolute errors as in the previous studies and the data points are chosen as in Fig. 2. This is probably due to the different quality of the data we obtained here, where the main source of errors comes from arithmetic operations and statistical errors are absence unlike previous studies with QMC.

**VI. 2D HUBBARD MODEL**

**A. Accuracy of the results and bench mark**

Before discussing the LDFA results of the 2D Hubbard Model obtained with the new Lanczos ED method in detail, here, we evaluate the accuracy of the results and make comparisons with results of other methods. We first check the accuracy of the four-point vertex function in Eq. (29). Since we are dealing with the PM state, the spin rotational symmetry can be verified. To do this, the difference between the horizontal-spin $\gamma^{\uparrow \downarrow \uparrow \downarrow}$ and vertical-spin $\gamma^{\uparrow \uparrow \downarrow \downarrow}$ components, which should be zero for the exact calculations, are examined with respect to $\Omega$ as

$$
\Delta \gamma(\Omega) = \frac{\sum_{\omega, \omega'} |\gamma^{\uparrow \downarrow \uparrow \downarrow}(\omega) - \gamma^{\uparrow \uparrow \downarrow \downarrow}(\omega)|}{\sum_{\omega, \omega'} \left| |\gamma^{\uparrow \downarrow \uparrow \downarrow}(\omega) + |\gamma^{\uparrow \uparrow \downarrow \downarrow}(\omega)\right|}.
$$

It is found that $\Delta \gamma(\Omega) \sim 10^{-5}$ for the lowest $\Omega \approx 0$ and the highest $\Omega = 4N_\omega \sim 4W$ frequencies and less accurate $\Delta \gamma(\Omega) \sim 10^{-4}$ for intermediate frequencies $\Omega \sim 2W$. This is expected from the approximation made in the Lanczos ED method, which is accurate in low-energy excitations and asymptotic behavior as mentioned in Sec. III.

Similarly, the accuracy of the LDFA results can be assessed by calculating the same quantity with different directions of spin for the PM state. The accuracy is typically about 6 digits for values such as the double occupancy $D \equiv \langle n_{i\uparrow} n_{i\downarrow}\rangle$ calculated using the Migdal-Galitskii formula. At low temperatures, however, because of the divergent property of the spin susceptibility $\chi_{sp} \sim e^{\Delta/T}$ as $T \to 0$, the Bethe-Salpeter equation in Eq. (33) is unstable when $\chi_{sp}^{-1} \propto 1 - \lambda_{sb} \approx e^{-\Delta/T}$ is too small $1 - \lambda_{sb} \lesssim 10^{-3}$, where $\lambda_{sb}$ is the maximum eigenvalue of a $2N_\omega \times 2N_\omega$ matrix

$$M_{\omega, \omega'} \equiv \frac{1}{\beta} \langle \chi_{sp}^{(sp)}(\omega') \chi_{sp,0}(\pi, \omega') \rangle.$$

Although the problem can be partly avoided by using technique in Ref. [42] obtained results are less accurate.

As mention before, in the ED method the conduction branch is replaced by a finite number of the bath energy levels in the effective IAM. While the accuracy of the results is expected to increase as the number of the bath levels $N_b$ increases, numerical errors introduce by the Lanczos algorithm, where high-energy excitations are omitted, would increase as the number of basis function increases. To check the convergence of the results as a function of $N_b$, the values of $D$ obtained with $N_b = 3, 5, \text{and} 7$ have been compared. Note that one of the bath level is required to be placed at the Fermi level to describe a metallic state and thus $N_b$ needs to be an odd number for the half-filled square-lattice Hubbard model because of the electron-hole symmetry. Whereas considerable difference $|D_{N_b=3} - D_{N_b=5}|/D_{N_b=5} \sim 10^{-3}$ between those obtained with $N_b = 3$ and $N_b = 5$ is found for $U = 4.0$ at low temperatures, the discrepancy between those obtained with $N_b = 5$ and $N_b = 7$ is already within $2 \times 10^{-5}$.

The deviations from the hybridization sum rule, which relates the hybridization strength of the effective IAM and the lattice hopping integrals, have also been examined. The sum rule for the square-lattice Hubbard...
model with the nearest-neighbor hopping is

\[ \sum_{i=1}^{N_b} V_i^2 = 4t^2. \] (56)

The deviation is rapidly reduced with increasing \( N_b \): the relative errors \( \varepsilon = |(\sum V_i^2)^{1/2} - 2t/2t| \) are utmost \( \varepsilon = 0.15, 6 \times 10^{-3}, \) and \( 6 \times 10^{-5} \) for \( N_b = 3, 5, \) and 7, respectively. These findings indicate that well converged values can be obtained with \( N_b = 7 \). Hence, all results presented in the rest of the paper were calculated with \( N_b = 7 \). The number of discretized momentum points in the Brillouin zone of the square lattice used in the calculations is \( 64 \times 64 \). For simplicity, we set the value of the hopping integral \( t = 1 \).

In Fig. 4, the imaginary part of the self-energy \( \text{Im} \Sigma_k(i\omega_n) \) on the imaginary axis at momenta \( k = (\pi/2, \pi/2) \) and \( k = (\pi, 0) \) for \( U = 8.0 \) and \( \beta = 2.0 \) obtained with various methods are compared. The differences are mainly found at the lowest Matsubara frequency, i.e., \( \omega_1 \equiv \pi/\beta \). The LDFA values of \( \text{Im} \Sigma_k(i\omega_n) \) are substantially reduced from that of DMFT at \( \omega_1 \) and are in good agreements with the DCA result for \( k = (\pi, 0) \) in Ref. 60. Our LDFA results are also in good agreement with the previous LDFA results with the diagrammatic QMC method in Ref. 37. On the other hand, the DΓA values are placed between those of DMFT and LDFA. The DΓA values at \( \omega_1 \) considerably deviate from that of DCA.

**B. Energetics**

It has been a long-standing debate over where and how the crossover or transition from the Slater to Mott-Heisenberg regime occurs as \( U \) increases in the 2D Hubbard model.24,28,34,35,41,52–55. In the Slater regime electrons are delocalized at high temperatures. In the weak coupling limit, the potential energy decreases in the presence of AFM correlations and therefore the stabilization of AFM order is expected to be mainly driven by the potential energy in the Slater regime. On the other hand, in the Mott-Heisenberg regime the localization of electrons already occurs at high temperatures and local spins are formed. In the strong coupling limit, the AFM coupling between the localized spins can be regarded as a virtual process through the electron hopping and the stabilization of AFM order is, therefore, expected to be mainly driven by the kinetic energy in the Mott-Heisenberg regime. In this subsection, we discuss the temperature and \( U \)-dependence of the double occupancy and the kinetic energy obtained with LDFA.

Figure 4 shows temperature dependence of the double occupancy \( D \equiv \langle n_{\uparrow} n_{\downarrow} \rangle \) for various values of \( U \) calculated using the Migdal-Galitskii formula.56 For \( U \leq 5.0 \) at high temperature, \( D \) increases with decreasing temperature, reaches its local maximum, and then decreases. The local maximum is positioned just above the DMFT \( T_N^\text{DMFT} \) indicated by the vertical arrows. Although no long-range AFM order is presence in LDFA in finite temperatures, which abides by the Mermin-Wagner theorem, the connection between the local maximum of \( D \) found at temperature near \( T_N^\text{DMFT} \) for \( U \leq 5.0 \) and AFM correlations is apparent. The presence of the local maximum is also found in DCA study for \( U = 4.0 \) in Ref. 60 as indicated by the diamonds in Fig. 4(c). The temperature dependence of \( D \) for \( U \leq 5.0 \) is consistent with what expected in the Slater regime: high-temperature metallic state, because of the Fermi degeneracy, \( D \) increases with decreasing temperature and \( D \) decreases as AFM correlations develop below \( T_N^\text{DMFT} \), where electrons efficiently avoid each other. These findings are also consistent with the previous LDFA study, where the reduction of the potential energy due to nonlocal AFM correlations found in a range \( 1.0 \leq U \leq 4.0 \).

To elucidate the relation between the AFM correlation and the temperature dependence of \( D \) more in detail, in Fig. 5 the imaginary part of the self-energy \( \text{Im} \Sigma_k(i\omega_n) \)
for $U = 4.0$ at momenta on the Fermi surface $\mathbf{k} = (\pi, 0)$ (the X point) and $\mathbf{k} = (\pi/2, \pi/2)$ (the M$_2$ point) are shown for various values of $\beta$ in panels (a) and (b), respectively. For a non-correlated metal, $\text{Im} \Sigma_k(i\pi/\beta)$ the local maximum temperature of $\text{Im} \Sigma_k$ at temperatures close to the onset of the short-range AFM order and the temperature of AFM correlations, i.e., $T_{\text{DMFT}}^\text{N}$ coincides with $T_{\text{DMFT}}^\text{N}$ at momenta on the Fermi surface $\mathbf{k} = (\pi, 0)$ (the X point) and indicates non-Fermi-liquid or “bad metallic” behavior at high temperatures. The fact is consistent with the Mott-Heisenberg regime, where local spins are expected to be preformed above the onset temperature of AFM correlations, i.e., $T_{\text{DMFT}}^\text{N}$.

To clarify where the crossover from the Slater to Mott-Heisenberg regime occurs, in Fig. 5(a), $D$ as a function of $U$ are plotted for $\beta = 5.0$ and $\beta = 8.0$. As can be seen in the $D$ curve of $\beta = 5.0$, $D$ decreases linearly with increasing $U$ at high temperatures. However, at low temperatures, the decrease of $D$ does not evenly happen as can be observed in the $D$ curve of $\beta = 8.0$: the faster (slower) reduction of $D$ for $U < 4.7$ ($U > 4.7$) causes the formation of the concave (convex) in the $D$ curve. This variation of $D$ at low temperatures is in accordance with the crossover from the Slater to Mott-Heisenberg regime with increasing $U$ and thus one can regard the inflection point of the $D$ curve as a boundary between the two regimes. To examine the inflection point of $D$ curve more in detail, in Fig. 6(b), $\partial^2 D/\partial U^2$ as a function of $U$ is shown for various values of $\beta$. For $\beta = 4.0$, there is no inflection point in the $D$ curve in a range from $U = 4.1$ to $U = 5.5$. As temperature decreases, however, $\partial^2 D/\partial U^2$ as a function of $U$ rapidly converges into a curve, which intersects the $\partial^2 D/\partial U^2 = 0$ line at $U \approx 4.7$, and changes its sign from positive to negative at $U \approx 4.7$. These results indicates that the crossover from the Slater to Mott-Heisenberg regime occurs around $U^* \approx 4.7$. Similar crossover from the Slater to Mott-Heisenberg regime has been also found in the AMF phase in the CDMFT and variational QMC studies.

Figure 7 shows temperature dependence of the kinetic energy $E_K$ for various values of $U$. For small $U$, the inclination of $E_K$ is reduced with decreasing temperature and no clear change of this tendency at $T_{\text{DMFT}}^\text{N}$ is found. In contrast, for $U > 4.5$ the inclination of $E_K$ increases with decreasing temperature particularly $T < T_{\text{DMFT}}^\text{N}$ resulting in steep precipitation of $E_K$ below $T_{\text{DMFT}}^\text{N}$. This tendency becomes more clear as $U$ increases. To make a rough estimation of the lowering of $E_K$ caused by the AFM correlation, we assume $E_K$ without the AFM correlation can be approximated by a linear function of $T$.

FIG. 5: The imaginary part of the self-energy $\text{Im} \Sigma_k(i\omega_n)$ for $U = 4.0$ with various values of $\beta$ at momenta on the Fermi surface $\mathbf{k} = (\pi, 0)$ (the X point) (a) and $\mathbf{k} = (\pi/2, \pi/2)$ (the M$_2$ point) (b).

FIG. 6: (a) Comparison between the $U$ dependence of the double occupancy $D$ at $\beta = 5.0$ (open tip-up triangles) and $\beta = 8.0$ (closed tip-down triangles). The inset shows the values of $D$ at $\beta = 5.0$ subtracted from that of $\beta = 8.0$ $D_{\beta=8.0} - D_{\beta=5.0}$ in a range from $U = 4.0$ to $U = 5.6$. (b) $\partial^2 D/\partial U^2$ as a function of $U$ for various values of $\beta$. 
below \( T_{N}^{\text{DMFT}} \) and subtract this approximated value from \( E_k \) at the lowest temperature \( T_L \) available as

\[
\Delta E_k = E_k(T_L) - E_k(T_{N}^{\text{DMFT}}) + \frac{\partial E_k}{\partial T} \bigg|_{T=T_{N}^{\text{DMFT}}} (T_L - T_{N}^{\text{DMFT}}). \tag{57}
\]

The result is presented in the inset of Fig. 7. Whereas \( \Delta E_k \) remains small \( |\Delta E_k| < 10^{-2} \) for \( U \leq 4.5 \), \( \Delta E_k \) rapidly decreases with increasing \( U \) for \( U > 5.0 \). This fact is consistent with the crossover behavior of \( D \) from the Slater to Mott-Heisenberg regime around \( U^* \approx 4.7 \), since in the latter the stabilization of the short-range AFM order is expected to be driven by the kinetic energy.

**C. Structures of DOS and the spectral function**

Figure 8 shows LDFA results of DOSs and the spectral functions \( A_k(\omega) \) with momenta \( k \) along the symmetry lines \( \Gamma-X, X-M \) and \( M-\Gamma \) for \( U = 4.0 \) with \( \beta = 5.0 \) in panel (a) and with \( \beta = 8.0 \) in panel (b). For comparison those obtained with the cluster perturbation theory (CPT)\(^{61,62}\) with a \( 4 \times 4 \) cluster are also presented in panel (c). Although the pseudogap already exists in DOS with \( \beta = 5.0 \), which is also found in the previous LDFA studies\(^{12,31}\), one can see the prominent development of the pseudogap DOS for \( \beta = 8.0 \). Such drastic development in the pseudogap with decreasing temperature is consistent with exponential growth of the AFM correlation length with decreasing temperature discussed in the DGA\(^{32}\) and LDFA\(^{42}\) studies. The formation of the pseudogap in DOS is found to start at the temperature close to \( T_{N}^{\text{DMFT}} \).

The formation of the pseudogap is also found in the quasiparticle peaks at the Fermi level in \( A_k(\omega) \) with \( k = (\pi, 0) \) (the X point) and \( k = (\pi/2, \pi/2) \) (the M\(_2\) point) for \( \beta = 8.0 \). The pseudogap opening along the \( \Gamma-X \) and \( X-M \) lines in the vicinity of the X point is also seen along with the peaks corresponding to the shadow band (indicated by the arrows), which is the reminiscence of the Brillouin zone folding caused by the long-range AFM order. On the other hand, for \( \beta = 5.0 \), although the incipience of pseudogap formation, e.g., the flattening of the quasiparticle peak top, can be observed at the X point, still no such indication found in the peak at the M\(_2\) point. The formation of the pseudogap occurs first at the X point, spreads through the Fermi surface and ends at the M\(_2\) point with decreasing temperature. The same trends in the temperature and momentum dependence in the pseudogap formation is found in the previous works with TPSC\(^{65}\), DGA\(^{61}\) and the QMC calculations of finite-size clusters\(^{38}\). More details of the momentum dependence of the pseudogap formation will be discussed in Sec. VII E. Although the structures of the LDFA spectral functions of the momenta near the \( \Gamma \) point are blurred because of inaccuracy of the data and the size of the pseudogap in the vicinity of the Fermi level is substantially smaller compared to those of CPT, reasonable agreements can be found between those obtained with LDFA in panel (a) and CPT in panel (c).

**D. Gap formation in DOS**

Although the investigation on spin susceptibility and spin correlation length in the previous works with DGA\(^{61,62}\) and LDFA\(^{42,43}\) have already revealed that the low temperature behavior of the spin fluctuations of the half-filled Hubbard model on a square lattice is consistent with the non-linear \( \sigma \) model, the connection between the spin fluctuation and the pseudogap formation is still not well understood. In the previous study with the non-linear \( \sigma \) model approach\(^{29}\), it is argued that there is a finite critical value of \( U \) \( (U_c \approx 4.25) \) which separates a pseudogap phase and a Mott insulating phase. In the pseudogap phase, finite \( \rho(\omega = 0) \) lingers at finite temperature, whereas clear gap opening occurs in the Mott insulating phase. The purpose of this subsection is to verify whether such abrupt change in the temperature dependence of \( \rho(\omega = 0) \) occurs at a finite \( U \) or not.

To see the temperature dependence of DOS, in Fig. 9 those with various values of \( \beta \) for \( U = 4.5 \) and \( U = 5.0 \) depicted in panels (a) and (b), respectively. Each of these DOSs consists of a peak at the Fermi level flanked by the shoulder structures corresponding to the lower- and upper-Hubbard bands at high temperatures. The pseudogap appears at a temperature close to \( T_{N}^{\text{DMFT}} \) and further develops as temperature decreases. While finite DOS in the vicinity of the Fermi level persists for \( U = 4.5 \)
even at the lowest temperature indicated in the figure, clear opening of the gap can be seen at \( \beta = 9.0 \) for \( U = 5.0 \). From these results it is expected that a pseudogap phase to Mott-Hubbard insulator transition or crossover exists in between \( U = 4.5 \) and \( U = 5.0 \).

To examine how the temperature dependence of \( \rho(\omega = 0) \) varies as \( U \) changes, in Fig. 10 the logarithmic plots of \( \rho(\omega = 0) \) as a function of \( \beta \) for various values of \( U \) are presented. The value of \( \rho(\omega = 0) \) is reduce as temperature decreases and the reduction becomes steeper as \( U \) increases particularly at low temperatures; Whereas the value of \( \rho(\omega = 0) \) for \( U = 4.9 \) is rapidly reduced with decreasing temperature and the value is less than \( 10^{-19} \) at \( \beta = 10.0 \), the reduction is moderate for \( U = 4.3 \) and the value is only reduced to about 47% from \( \beta = 3.0 \) to \( \beta = 10.0 \).

In the non-linear \( \sigma \) model approach\(^{29}\), the temperature dependence of \( \rho(\omega = 0) \) for the weak coupling limit at low temperatures can be approximated by

\[
\rho(\omega = 0) \propto \exp(-\beta\Delta_0),
\]

where \( \Delta_0 \) is half the size of the gap of the AFM state at \( T = 0 \). To verify whether the low-temperature behavior of our results are consistent with the pseudogap phase in the non-linear \( \sigma \) model approach for small \( U \), the linear least squares fit is made for \( \log_{10} \rho(\omega = 0) \) as a function of \( \beta \) to determine \( \Delta_0 \). For example, the results of the linear least squares fit for \( U = 4.5 \) is shown in Fig. 11. The fitting is made within a range from \( \beta = 6.2 \) to 8.2. A reasonably good agreement can be obtained within the range. However, the substantial deviation from the linear approximation are found for \( \beta \geq 9.0 \) and the values of \( \rho(\omega = 0) \) are larger than those expected from Eq. 58. This is probably caused by inaccuracy due to small 1 - \( \lambda_{SP} \approx 5 \times 10^{-4} \) for \( \beta > 8.5 \) as mentioned in Sec. VI B. The similar tendency is found for \( U \leq 4.7 \) at low temperatures. The estimated \( \Delta_0 \)'s for various values of \( U \) are depicted in the inset. The obtained 2\( \Delta_0 \) is, indeed, about the peak to peak distance of the pseudogap structure of DOS for \( U \leq 4.3 \). However, \( \Delta_0 \) rapidly deviates from the actual size of the pseudogap of DOS for \( U > 4.3 \) and the possible range of \( \beta \) for the linear fitting is reduced as \( U \) increases, indicating rapid disappearance of states inside the gap. No reasonable fitting is available for \( U \geq 4.8 \).

These results show that a sharp crossover from the pseudogap phase to the Mott insulator takes place at \( U^* \approx 4.7 \). Indeed, the value of \( U^* \) coincides with the boundary between the Slater and Mott-Heisenberg regimes defined by the inflection point of \( D \) curve as a function of \( U \), i.e., \( (\partial D/\partial U^2)_T = 0 \), discussed in Sec. VII B.

The robustness of the results obtained here has been checked by altering the part of the procedure of MEM described in Sec. VII in several different ways, e.g., by taking...
E. Pseudogap in the spectral function

In the previous study with DΓA\textsuperscript{35,64}, the MIT in the half-filled Hubbard model on a square lattice has been discussed, where the transition temperature is related to the variation in the \(\omega_n\) dependence of \(\text{Im} \Sigma_k(i\omega_n)\). If the system is a Fermi liquid or at least a Fermi liquid like, it is expected that \(|\text{Im} \Sigma_k(i\omega_n)|\) should decrease as \(\omega\) decreases. However, the loss of the Fermi-liquid feature is not necessarily indicate insulating behavior of the system. The purpose of this subsection is to clarify the relation between the pseudogap formation in the spectral function \(A_k(\omega)\) and the change in \(\text{Im} \Sigma_k(i\omega_n)\).

As discussed in Sec. VI C, the pseudogap formation in \(A_k(\omega)\) has \(\mathbf{k}\) dependence; it initiates at the X point and spread through the Fermi surface and terminates at the M\(_2\) point as temperature decreases. Figure 12 shows \(A_k(\omega)\) in the vicinity of the Fermi level for \(U = 3.0\) at temperatures across the pseudogap formation at the X and M\(_2\) points together with the corresponding \(\text{Im} \Sigma_k(i\omega_n)\). It is clearly seen in panels (b) and (d) that the \(\omega_n\) dependence of \(\text{Im} \Sigma_k(i\omega_n)\) changes upturn to downturn in \(\text{Im} \Sigma_k(i\omega_n)\) with decreasing temperature around the pseudogap formation temperature. These results are in good agreement with those of DΓA calculations for \(U = 3.0\) in Ref. 64. This indicates that for small \(U\) the Fermi-liquid feature is gradually lost accompanied by the pseudogap formation in \(A_k(\omega)\) at the Fermi level. It is also found in panels (b) and (d) that the variation mainly occurs in the lowest Matsubara frequency \(\omega_1 = \pi/\beta\), where \(|\text{Im} \Sigma_k(i\pi/\beta)|\) is increases with decreasing temperature. This predominant increase in \(|\text{Im} \Sigma_k(i\pi/\beta)|\), as was already pointed out in Ref. 22, is the main cause of the double peak structure in \(A_k(\omega)\),
i.e., the formation of the pseudogap. Since the low-temperature magnetic excitation of the half-filled Hubbard model on a square lattice is considered to be described by the two-dimensional non-linear $\sigma$ model in the renormalized classical regime, the magnetic scattering at the lowest Matsubara frequency is expected to be the dominant process, which is consistent with our results.

The relation, however, becomes less clear as $U$ increases, although the pseudogap formation still occurs below $T^\text{DMFT}_N$. As shown in Fig. 13 for $U = 5.0$, $|\text{Im } \Sigma_k(i\omega_n)|$ at the X point monotonically increases with decreasing $\omega_n$ already at high temperatures above $T^\text{DMFT}_N$ and therefore no upturn to downturn change occurs in $\text{Im } \Sigma_k(i\omega_n)$ at $\beta = 4.75$ where the pseudogap appears. This non-Fermi-liquid feature in $\text{Im } \Sigma_k(i\omega_n)$ above $T^\text{DMFT}_N$ spreads through the Fermi surface from the X point to the $M_2$ point as $U$ increases from $U = 4.0$ and at $U = 5.5$ the whole Fermi surface is lost already above the temperature where the pseudogap appears, resulting in no upturn to downturn change in $\text{Im } \Sigma_k(i\omega_n)$ at all.

Furthermore, one can see the enhancement of the intensity of shoulder structures around $\omega \approx \pm 2.5$ in $A_k(\omega)$ and the increase of $|\text{Im } \Sigma_k(i\omega_n)|$ up to $\omega_n \sim 5$ with decreasing temperature. The appearance of the high-energy structure $\omega \approx \pm U/2$ is hallmark of the Mott physics and cannot be explained by the magnetic scattering within the energy scale of $T^\text{DMFT}_N \sim 0.3$. These facts can be contrasted with $A_k(\omega)$ and $\text{Im } \Sigma_k(i\omega_n)$ for $U = 3.0$ in Fig. 12 where their temperature effects are mainly found within the energy scale of $T^\text{DMFT}_N \sim 0.2$ around the Fermi level and are in accordance with the Slater mechanism of gap formation due to the short-range AMF ordering. Nevertheless, a quasiparticle-like broad single peak still exists at the Fermi level at high temperatures in $A_k(\omega)$ for $U = 5.0$ in Fig. 13 and predominant increase in $|\text{Im } \Sigma_k(i\pi/\beta)|$ owing to the magnetic scattering leads to the pseudogap formation.

**F. $U$-$T$ phase diagram**

To conclude this section, here we consider how the electronic state changes on the $U$-$T$ parameter space. In Fig. 14 characteristic temperatures so far discussed are summarized. The $D$ local maximum temperature is the onset temperature of AFM correlation in the high-temperature metallic state, which is signaled by increase in $|\text{Im } \Sigma_k(i\pi/\beta)|$ with decreasing temperature around the X point for $U < 5.0$. For $U \geq 5.5$, on the other hand,
strong correlations already develop at high-temperatures and no D local maximum temperature found below $T < 0.5$. The AFM correlations further develop with decreasing temperature and at $T_{\mathrm{N}}^{{\mathrm{D}}M{\mathrm{F}}T}$, $|\text{Im} \Sigma_k(i\pi/\beta)|$ with momenta $k$ on whole Fermi surface start to increase with decreasing temperature.

For $U \leq 3.5$, the Fermi-liquid behavior is gradually lost below $T_{\mathrm{N}}^{{\mathrm{D}}M{\mathrm{F}}T}$ line as the AFM correlations develops with decreasing temperature. This starts at the X point of the Fermi surface and ends at the $M_2$ point with concomitant formation of the pseudogap in the corresponding quasiparticle peak at the Fermi level. The Fermi surface is totally lost and the formation of the pseudogap at the Fermi level is complete at the $M_2$ pseudogap line. The difference between the pseudogap temperatures at the X and $M_2$ points rapidly decreases as $U$ is reduced and is only about $2 \times 10^{-3}$ at $U = 2.0$.

The same gradual pseudogap formation in the peak at the Fermi level occurs for $U \geq 4.0$ below the $T_{\mathrm{N}}^{{\mathrm{D}}M{\mathrm{F}}T}$ line. In this case, however, $|\text{Im} \Sigma_k(i\pi/\beta)|$ around the X point is large even at high temperatures and the system is already non-Fermi-liquid like above the pseudogap formation temperature. This tendency is strengthen as $U$ increases and for $U \geq 5.5$, the whole Fermi surface is already non-Fermi-liquid like above the pseudogap formation temperature, where $A_k(\omega)$ has a structure with a broad single peak at the Fermi level accompanied by prominent shoulder structures at $\omega \approx \pm U/2$. In this regards, the formation of the pseudogap at low temperatures only heralds the development of AFM correlations for $U \geq 5.5$ and local spins are considered to be preformed at higher temperatures as expected in the Mott-Heisenberg regime. This is contrasted with the simultaneous formation of the pseudogap and loss of the Fermi-liquid feature found for $U \leq 3.5$, which is expected in the Slater regime. Note that the $D$ local maximum and the X and $M_2$ pseudogap lines are only those of crossover and no anomaly which indicates a true transition is found in $D$, DOS or the spectral functions.

Further lowering the temperature, the development of the pseudogap is distinctly different depending on whether $U$ is larger or smaller than $U^* \approx 4.7$. For $U < U^*$, $\rho(\omega = 0)$ is reduced with decreasing temperature but persists even at low temperatures. In contrast, for $U > U^*$, the reduction is much rapid and clear gap opening occurs at certain temperature. The value of $U^* \approx 4.7$ coincides with the boundary between the Slater and Mott-Heisenberg regimes defined by the inflexion point of $D$ curve as a function of $U$.

**VII. DISCUSSIONS**

The MIT in the half-filled Hubbard model on the square lattice has been discussed using DFA and lattice QMC\textsuperscript{35}. Although the difference between the pseudogap formation temperatures at the X and $M_2$ points is substantially larger compared to the present study, the behavior of the formation found in this DFA study for $U \leq 4.0$ is essentially the same to the present study (see Fig. 14): the pseudogap formation starts at the X point and ends at the $M_2$ point with decreasing temperature. From this results it is suggested that $U_c = 0$ for $T \to 0$ and thus no MIT occurs at any $U > 0$ similar to the one-dimensional Hubbard model in this DFA study. Although we found a sharp crossover from the pseudogap phase to Mott insulator around $U^* \approx 4.7$, we do not I found any anomaly, e.g., a discontinuity or kink, in $D$ up to the second derivative with respect to $U$ at $U^*$. Hence, it is unlikely that $U^* \approx 4.7$ is the true MIT point. For this reason, in the strict sense, the true MIT point is considered to be presence at $U_c = 0$. This weakness of the variation from the metallic to insulating phase at $U^*$ as compared to that in infinite dimension is not surprising. Since the pseudogap owing to the AFM correlation already exists at higher temperatures, the crossover we have discussed here amounts to a subtle change in the states in the gap. This is contrasted with the MIT in the Hubbard model in infinite dimension, where the abrupt destruction of the coherent peak at the Fermi level causes the first-order MIT in finite temperature.

In the DFA study, there is no detail investigation on DOS at the Fermi level below the pseudogap formation temperature in particular for $U > 4.0$. For this reason, there is a possibility that a similar sharp crossover from the Slater to Mott-Heisenberg regime is found at finite $U$.
in DΓA and it would be interesting to investigate DOS at the Fermi level below the pseudogap formation temperature with DΓA to clarify this point.

VIII. CONCLUSIONS

A new formula for the two-body Green’s function combined with the Lanczos ED technique proposed in this paper provides efficient and accurate means to calculate the local vertex function of the effective impurity Anderson model required for DΓA and similar perturbative extensions of DMFT. The formula is applicable to not only the impurity Anderson model but cluster models including those with multiple orbitals. Utilizing this new scheme, the double occupancy, kinetic energy, spectral function and DOS of the Hubbard model on the square lattice at the half-filling are calculated by means of LDFA with an unprecedented accuracy at low temperatures and $U$’s ranging from $U_2$ to $U_5$ to clarify this point.

It is found that the pseudogap is first formed in the quasiparticle peak of the spectral function at the $X$ point as temperature decreases. The formation spreads through the Fermi surface and ends at the M point as temperature decreases. The formation spreads in DΓA and it would be interesting to investigate DOS at the Fermi level below the pseudogap formation temperature with DΓA to clarify this point.

IX. ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Numbers 25400377, 18K03541, 18H03683. The author also would like to thank A. Lichtenstein for fruitful discussions.

Appendix A: Derivation of the new expression for two-body Green’s function

The purpose of this section is to provide detailed description of a derivation of the expression for the two-body Green’s function given in Eqs. (8)- (3).

\[
\chi_{1234}(iw_1, iw_2; iw_3, iw_4) = \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \int_0^\beta d\tau_3 \langle T [c_1(\tau_1)c_2(\tau_2)c_3^\dagger(\tau_3)c_4^\dagger(0)] \rangle e^{i(w_1\tau_1 + w_2\tau_2 - w_3\tau_3)}
\]

\[
= \frac{1}{Z} \sum_{lmnk} \langle k|c_1|l\rangle \langle l|c_2|m\rangle \langle m|c_3^n|n\rangle \langle n|c_4^\dagger|k\rangle \phi_{lmnk}(w_1, w_2, -w_3) + \frac{1}{Z} \sum_{lmnk} \langle k|c_3^\dagger|l\rangle \langle l|c_1|m\rangle \langle m|c_2^n|n\rangle \langle n|c_4^\dagger|k\rangle + \frac{1}{Z} \sum_{lmnk} \langle k|c_2|l\rangle \langle l|c_3^\dagger|m\rangle \langle m|c_1^n|n\rangle \langle n|c_4^\dagger|k\rangle \phi_{lmnk}(-w_3, w_1, w_2) - \frac{1}{Z} \sum_{lmnk} \langle k|c_3^\dagger|l\rangle \langle l|c_2|m\rangle \langle m|c_1^n|n\rangle \langle n|c_4^\dagger|k\rangle \phi_{lmnk}(-w_3, w_2, w_1) - \frac{1}{Z} \sum_{lmnk} \langle k|c_1|l\rangle \langle l|c_3^\dagger|m\rangle \langle m|c_2^n|n\rangle \langle n|c_4^\dagger|k\rangle \phi_{lmnk}(w_1, -w_3, w_2) - \frac{1}{Z} \sum_{lmnk} \langle k|c_2|l\rangle \langle l|c_3^\dagger|m\rangle \langle m|c_1^n|n\rangle \langle n|c_4^\dagger|k\rangle \phi_{lmnk}(w_2, w_1, -w_3)
\]  

(A1)
where
\[\phi_{l, mnk}(\omega_1, \omega_2, \omega_3) = e^{-\beta E_k} \int_0^\beta d\tau_1 \int_0^{\tau_1} d\tau_2 \int_0^{\tau_2} d\tau_3 e^{(E_k - E_l + i\omega_1)\tau_1} e^{(E_l - E_m + i\omega_2)\tau_2} e^{(E_m - E_n + i\omega_3)\tau_3}\]
(A2)

If both the condition \(E_k \neq E_m\) or \(\omega_1 + \omega_2 \neq 0\) and the condition \(E_l \neq E_n\) or \(\omega_2 + \omega_3 \neq 0\) are satisfied, we obtain
\[\phi_{l, mnk}(\omega_1, \omega_2, \omega_3) = -e^{-\beta E_n} e^{-\beta E_k} \frac{(E_m - E_n + i\omega_3)(E_l - E_n + i(\omega_2 + \omega_1))(E_k - E_n + i(\omega_1 + \omega_2 + \omega_3))}{e^{-\beta E_k} - e^{-\beta E_k}} + \frac{(E_m - E_n + i\omega_3)(E_l - E_n + i(\omega_2 + \omega_1))(E_k - E_l + i\omega_1)}{e^{-\beta E_k} - e^{-\beta E_k}} - \frac{(E_m - E_n + i\omega_3)(E_l - E_m + i\omega_2)(E_k - E_m + i(\omega_1 + \omega_2))}{e^{-\beta E_k} - e^{-\beta E_k}} - \frac{(E_m - E_n + i\omega_3)(E_l - E_m + i\omega_2)(E_k - E_l + i\omega_1)}{e^{-\beta E_k} - e^{-\beta E_k}}.\]
(A3)

Now we rewrite the right-hand side of Eq. (A3) in such a way that the three factors in denominator and the argument of the exponential function of each term contain the same eigenvalue of \(H\). To do this, we use the identity
\[\frac{1}{z_1 z_2} = \frac{1}{z_1} \left( \frac{1}{z_2} \right) = \left( \frac{1}{z_1} \pm \frac{1}{z_2} \right) \left( \frac{1}{z_1} \pm \frac{1}{z_2} \right).\]
(A4)

For instance, the term with the factor \(e^{-\beta E_l}\) on the second line of Eq. (A3) can be transformed into
\[e^{-\beta E_l} \left( \frac{1}{E_m - E_n + i\omega_3} - \frac{1}{E_l - E_n + i(\omega_2 + \omega_3)} \right) \frac{1}{(E_l - E_m + i\omega_2)(E_k - E_l + i\omega_1)} \left( \frac{1}{E_l - E_n + i(\omega_2 + \omega_3)} \right)\]
(A5)

Together with the term with the factor \(e^{-\beta E_l}\) on the fourth line of Eq. (A3), we get
\[-e^{-\beta E_l} \frac{1}{(E_l - E_n + i(\omega_2 + \omega_3))(E_l - E_m + i\omega_2)(E_k - E_l + i\omega_1)}\]
(A6)

where all the factors in the denominator and \(e^{-\beta E_l}\) contain the same eigenvalue \(E_l\). We can cast all the terms of Eq. (A3) in this form by further repeated use of Eq. (A4) to the terms with \(e^{-\beta E_k}\):
\[\phi_{l, mnk}(\omega_1, \omega_2, \omega_3) = -e^{-\beta E_k} \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i(\omega_1 + \omega_2))(E_k - E_n + i(\omega_1 + \omega_2 + \omega_3))} - e^{-\beta E_m} \frac{1}{(E_m - E_n + i\omega_3)(E_m - E_k - i(\omega_1 + \omega_2))(E_m - E_l - i\omega_2)} + e^{-\beta E_l} \frac{1}{(E_l - E_m + i\omega_2)(E_l - E_n + i(\omega_2 + \omega_3))(E_l - E_k - i\omega_1)} + e^{-\beta E_n} \frac{1}{(E_n - E_k - i(\omega_1 + \omega_2 + \omega_3))(E_n - E_l - i(\omega_2 + \omega_3))(E_n - E_m - i\omega_3)}.\]
(A7)

These terms correspond to the major terms discussed in Sec. II such as Eq. (12).

We now deal with the situations with \(E_k = E_m\) or \(E_l = E_n\), where some of denominators in Eq. (A7) are zero if \(\omega_1 + \omega_2 = 0\) or \(\omega_2 + \omega_3 = 0\) is satisfied. When \(E_k = E_m\) and \(\omega_1 + \omega_2 \neq 0\), using Eq. (A4) the first two lines on the right-hand side of Eq. (A7) can be written as
\[-e^{-\beta E_k} \left\{ \frac{1}{(E_k - E_l + i\omega_1)(E_l - E_n + i(\omega_1 + \omega_2 + \omega_3))} - \frac{1}{(E_k - E_n + i\omega_3)(E_k - E_l - i\omega_2)} \right\} + \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i\omega_3)} \frac{1}{(E_k - E_n + i(\omega_1 + \omega_2))(E_k - E_l - i\omega_2)} + \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i\omega_3)} \frac{1}{(E_k - E_n + i\omega_3)(E_k - E_l - i\omega_2)} + \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i\omega_3)} \frac{1}{(E_k - E_n + i\omega_3)(E_k - E_l - i\omega_2)}\]
(A8)
Similarly, when $E_l = E_n$ and $\omega_2 + \omega_3 \neq 0$, the last two lines on the right-hand side of Eq. (A7) can be written as

\[
e^{-\beta E_l} \left\{ \frac{1}{(E_l - E_m + i\omega_2)(E_l - E_k - i\omega_1)} - \frac{1}{(E_l - E_k - i(\omega_1 + \omega_2 + \omega_3))(E_l - E_m - i\omega_3)} \right\}
\]

\[
= -e^{-\beta E_l} \left\{ \frac{1}{(E_l - E_k - i(\omega_1 + \omega_2 + \omega_3))(E_l - E_m + i\omega_2)} + \frac{1}{E_l - E_k - i\omega_1} \right\}.
\]  

To complete the calculation of $\phi$, we further need to know special cases of the integration of Eq. (A3). If $E_k = E_m$ and $\omega_1 + \omega_2 = 0$, we obtain

\[
\phi_{lmnk}(\omega_1, \omega_2, \omega_3) = e^{-\beta E_k} \left\{ \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_n + i\omega_3)} + \frac{1}{E_k - E_l + i\omega_1} \right\}
\]

\[
+ e^{-\beta E_l} \left\{ \frac{1}{(E_l - E_k - i\omega_1)(E_l - E_m + i\omega_2)} + \frac{1}{E_l - E_k - i\omega_1} \right\}.
\]  

The terms on the first line of Eq. (A10) are identical to Eq. (A9) with $E_k = E_m$ and $\omega_1 + \omega_2 = 0$ and the same holds for the terms on the second and third lines of Eq. (A10), which correspond to the last two lines of Eq. (A7). However the term on the last line of Eq. (A10) is newly appeared. Similarly, if $E_l = E_n$ and $\omega_2 + \omega_3 = 0$, we obtain

\[
\phi_{lmnk}(\omega_1, \omega_2, \omega_3) = -e^{-\beta E_k} \left\{ \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i\omega_2)} + \frac{1}{E_k - E_l + i\omega_1} \right\}
\]

\[
- e^{-\beta E_m} \left\{ \frac{1}{(E_m - E_l - i\omega_2)(E_m - E_k - i(\omega_1 + \omega_2))} + \frac{1}{E_l - E_m + i\omega_2} \right\}
\]

\[
- e^{-\beta E_l} \left\{ \frac{1}{(E_l - E_k - i\omega_1)(E_l - E_m + i\omega_2)} + \frac{1}{E_l - E_k - i\omega_1} \right\}
\]

\[
- e^{-\beta E_k} \left\{ \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i\omega_2)} + \frac{1}{E_k - E_l + i\omega_1} \right\}.
\]  

We find the terms on the third line of Eq. (A11) are identical to Eq. (A9) with $E_l = E_n$ and $\omega_2 + \omega_3 = 0$ and the terms on the first two lines of Eq. (A11) correspond to those on the first two lines of Eq. (A7). Again, we see the additional term on the last line of Eq. (A11). From Eqs. (A7)-(A11), we obtain

\[
\phi_{lmnk}(\omega_1, \omega_2, \omega_3) =
\]

\[
- (1 - \delta_{E_k, E_n}) e^{-\beta E_k} \left\{ \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i(\omega_1 + \omega_2)))} + \frac{1}{E_k - E_l + i\omega_1} \right\}
\]

\[
- (1 - \delta_{E_k, E_m}) e^{-\beta E_m} \left\{ \frac{1}{E_m - E_l - i\omega_3)(E_m - E_k - i(\omega_1 + \omega_2))} \right\}
\]

\[
+ (1 - \delta_{E_l, E_n}) e^{-\beta E_l} \left\{ \frac{1}{E_l - E_m + i\omega_2)(E_l - E_n + i(\omega_2 + \omega_3))} \right\}
\]

\[
+ (1 - \delta_{E_l, E_m}) e^{-\beta E_m} \left\{ \frac{1}{(E_n - E_k - i(\omega_1 + \omega_2 + \omega_3))(E_n - E_l - i(\omega_2 + \omega_3))} \right\}
\]

\[
+ \delta_{E_k, E_m} e^{-\beta E_k} \left\{ \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i\omega_3)} \right\}
\]

\[
- \delta_{E_k, E_l} e^{-\beta E_l} \left\{ \frac{1}{(E_k - E_l + i\omega_1)(E_k - E_m + i\omega_3)} \right\}
\]

\[
+ \delta_{E_k, E_n} e^{-\beta E_k} \left\{ \frac{1}{(E_k - E_n + i\omega_3)(E_k - E_l - i\omega_2)} \right\}.
\]  

Finally, inserting Eq. (A12) into Eq. (A11), we obtain the expression for the two-body Green’s function in Eqs. (5)-(11).
Appendix B: Derivation of update formula of hybridization function

Here, we show the derivation of the update formula of $\Delta_\omega$ for DFA in Eq. (11). For the DMFT calculation, one can choose new $\Delta_\omega$ and $g_\omega$ for next iteration

$$g_\omega^{\text{new}} = \langle g_{k\omega} \rangle_k, \quad (B1)$$
$$\Delta_\omega^{\text{new}} = g_\omega^{-1} - [g_\omega^{\text{new}}]^{-1} + \Delta_\omega, \quad (B2)$$

where $g_{k\omega} = [g_\omega^{-1} + \Delta_\omega - \varepsilon_k]^{-1}$ is the DMFT lattice Green's function. These update formulae render robust and rapid convergence of $\Delta_\omega$ for DMFT. A formula similar to Eq. (11) can be derived for DFA. The condition of the convergence of $\Delta_\omega$ adopted in this study is $\langle G_{k\omega} \rangle_k = 0$. We rewrite this equation as

$$\langle G_{k\omega}^d \rangle_k = \langle G_{k\omega}^d[G_{k\omega}^d]^{-1}G_{k\omega}^d \rangle_k$$
$$= \langle G_{k\omega}^d[G_{k\omega}^d]^{-1}(-g_\omega + g_{k\omega}) \rangle_k$$
$$= -\langle G_{k\omega}^d[G_{k\omega}^d]^{-1} \rangle_k g_\omega + \langle G_{k\omega}^d[G_{k\omega}^d]^{-1}g_{k\omega} \rangle_k = 0$$

(B3)

From this one may employ

$$g_\omega^{\text{new}} = \left( G_{k\omega}^d[G_{k\omega}^d]^{-1} \right)^{-1} \left( -g_\omega^d + g_{k\omega}^d \right)$$

(B4)

as an update for $g_\omega$ of DFA. Indeed, if $\Sigma^d_\omega = 0$, this equation is reduced to Eq. (11). With combined use of Eq. (B2), one can obtain new hybridization function for DFA. Finally, substitution of Eq. (B4) into Eq. (B2) and use of the relation $g_{k\omega}^{-1} = g_\omega^{-1} + \Delta_\omega - \varepsilon_k$ results in Eq. (11). Similar formula can be obtained for DMFT:

$$\Delta_\omega^{\text{new}} = g_\omega^{-1}(g_{k\omega}\varepsilon_k)$$

Note that instead of directly calculating $G_{k\omega}^d[G_{k\omega}^d]^{-1}$ in Eq. (11), it is preferable to use

$$G_{k\omega}^d[G_{k\omega}^d]^{-1} = [1 - \Sigma^d_\omega G_{k\omega}^d]^{-1}$$

(B5)

to avoid loss of significant digits with small interaction.

1. A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. 68, 13 (1996).
2. W. Metzner and D. Vollhardt, Phys. Rev. Lett. 62, 324 (1989).
3. T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, Rev. Mod. Phys. 77, 1027 (2005).
4. G. Rohringer, H. Hafermann, A. Toschi, A. A. Katanin, A. E. Antipov, M. I. Katsnelson, A. I. Lichtenstein, A. N. Rubtsov, and K. Held, Rev. Mod. Phys. 90, 025003 (2018).
5. H. Kusunose, J. Phys. Soc. Jpn. 75, 054713 (2006).
6. A. Toschi, A. A. Katanin, and K. Held, Phys. Rev. B 75, 045118 (2007).
7. A. N. Rubtsov, M. I. Katsnelson, and A. I. Lichtenstein, Phys. Rev. B 77, 033101 (2008).
8. A. N. Rubtsov, M. I. Katsnelson, A. I. Lichtenstein, and A. Georges, Phys. Rev. B 79, 045133 (2009).
9. A. N. Rubtsov, M. I. Katsnelson, A. I. Lichtenstein, Ann. Phys. 327, 1320 (2012).
10. G. Rohringer, A. Toschi, H. Hafermann, K. Held, V. I. Anisimov, and A. A. Katanin, Phys. Rev. B 88, 115112 (2013).
11. C. Taranto, S. Andergassen, J. Bauer, K. Held, A. Katanin, W. Metzner, G. Rohringer, and A. Toschi, Phys. Rev. Lett. 112, 196402 (2014).
12. T. Ayral, and O. Parcollet, Phys. Rev. B 92, 115109 (2015).
13. G. Li, H. Lee, and H. Monien, Phys. Rev. B 78, 195105 (2008).
14. H. Hafermann, K. R. Patton, and P. Werner, Phys. Rev. B 85, 205106 (2012).
15. H. Hafermann, Phys. Rev. B 89, 235128 (2014).
16. H. Shimaoka, J. Otsuki, K. Haule, M. Wallerberger, E. Gull, K. Yoshimi, and M. Ohzeki, Phys. Rev. B 97, 205111 (2018).
17. H. Hafermann, C. Jung, S. Brener, M. I. Katsnelson, A. N. Rubtsov and A. I. Lichtenstein, EPL 85, 27007 (2009).
18. H. Hafermann, Numerical Approaches to Spatial Correlations in Strongly Interacting Fermion Systems (Cuvillier Verlag, Göttingen, 2010).
19. Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors, Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide (SIAM, Philadelphia, 2000).
20. J. R. Schrieffer, X. G. Wen, and S. C. Zhang, Phys. Rev. B 39, 11663 (1989).
21. M. Vekić, and S. R. White, Phys. Rev. B 47, 1160 (1993).
22. Y. M. Vil’k, and A.-M. S. Tremblay, Europhys. Lett. 33, 159 (1996).
23. Y. M. Vil’k, and A.-M. S. Tremblay, J. Phys. I 7, 1309 (1997).
24. P. W. Anderson, The Theory of Superconductivity in the High-Tc Cuprates, Princeton Series in Physics (Princeton University Press, NJ, 1997).
25. S. Moukouri, S. Allen, F. Lemay, B. Kyung, and D. Poulin, Y. M. Vil’k, A.-M. S. Tremblay, Phys. Rev. B 61, 7887 (2000).
26 F. Mancini, EPL (Europhysics Letters) **50**, 229 (2000).
27 A. Avella, F. Mancini, and R. Münzner, Phys. Rev. B **63**, 245117 (2001).
28 K. Borejsza, and N. Dupuis, Europhys. Lett. **63**, 722 (2003).
29 K. Borejsza, and N. Dupuis, Phys. Rev. B **69**, 085119 (2004).
30 J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
31 S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).
32 N. D. Mermin and H. Wagner, Phys. Rev. Lett. **63**, 722 (2003).
33 K. Borejsza, and N. Dupuis, Phys. Rev. B **69**, 085119 (2004).
34 J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
35 S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).
36 N. D. Mermin and H. Wagner, Phys. Rev. Lett. **63**, 722 (2003).
37 K. Borejsza, and N. Dupuis, Phys. Rev. B **69**, 085119 (2004).
38 J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
39 S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).
40 N. D. Mermin and H. Wagner, Phys. Rev. Lett. **63**, 722 (2003).
41 K. Borejsza, and N. Dupuis, Phys. Rev. B **69**, 085119 (2004).
42 J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
43 S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).
44 N. D. Mermin and H. Wagner, Phys. Rev. Lett. **63**, 722 (2003).
45 K. Borejsza, and N. Dupuis, Phys. Rev. B **69**, 085119 (2004).
46 J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
47 V. Heine in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz and D. Turnbull (Academic Press, New York, 1980) Vol. 35, p.87.
48 S. Brener, H. Hafermann, A. N. Rubtsov, M. I. Katsnelson, and A. I. Lichtenstein, Phys. Rev. B **77**, 195105 (2008).
49 N. E. Bickers, and S. R. White, Phys. Rev. B **43**, 8044 (1991).
50 R. Žitko, Phys. Rev. B **80**, 125125 (2009).
51 M. Jarrell, and J. E. Gubernatis, Phys. Rep. **269**, 133 (1996).
52 S. Moukouri, and M. Jarrell, Phys. Rev. Lett. **87**, 167010 (2001).
53 B. Kyung, J. S. Landry, D. Poulin, and A.-M. S. Tremblay, Phys. Rev. Lett. **90**, 099702 (2003).
54 E. Gull, P. Werner, X. Wang, M. Troyer, and A. J. Millis, Eur. Phys. Lett. **84**, 37009 (2008).
55 L. F. Tocchio, F. Becca, and S. Sorella, Phys. Rev. B **94**, 205126 (2016).
56 V. M. Galitskii and A. B. Migdal, Zh. Eksp. Teor. Fiz. **34**, 139 (1958) [JETP 7, 96 (1958)].
57 S. Chakravarty, B. I. Halperin, and D. R. Nelson, Phys. Rev. Lett. **60**, 1057 (1988).
58 P. Hasenfratz and F. Niedermayer, Phys. Lett. B **268**, 231 (1991).
59 E. Koch, G. Sangiovanni, and O. Gunnarsson, Phys. Rev. B **78**, 115102 (2008).
60 J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bullik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull, Phys. Rev. X **5**, 041041 (2015).
61 D. Sénéchal, D. Perez, and M. Pioro-Ladrière, Phys. Rev. Lett. **84**, 522 (2000).
62 D. Sénéchal, D. Perez, and D. Plouffe, Phys. Rev. B **66**, 075129 (2002).
63 A. A. Katanin, A. Toschi, and K. Held, Phys. Rev. B **80**, 075104 (2009).
64 G. Rohringer, and A. Toschi, Phys. Rev. B **94**, 125144 (2016).