Precursors of the insulating state in the square lattice Hubbard model

Erik G. C. P. van Loon,1 Hartmut Hafermann,2 and Mikhail I. Katsnelson1

1Radboud University, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands
2Mathematical and Algorithmic Sciences Lab, Paris Research Center, Huawei Technologies France SASU, 92100 Boulogne Billancourt, France

We study the two-dimensional square lattice Hubbard model for small to moderate interaction strengths $1 \leq U/t \leq 4$ by means of the ladder dual fermion approach. The non-local correlations beyond dynamical mean-field theory lower the potential energy, lead to a maximum in the uniform susceptibility and induce a pseudogap in the density of states. While the self-energy exhibits precursors of a possible insulating phase linked to the appearance of long-range fluctuations, the metallic phase persists within the accessible temperature range. Finite-size effects affect results qualitatively. Upper bounds on the crossover temperature are found to be significantly lower than previously reported dynamical vertex approximation results at $U/t = 1$.

PACS numbers: 71.30.+h,71.10.-w,71.10.Fd,

I. INTRODUCTION

The Hubbard1–4 model captures numerous phenomena of strongly correlated electron physics, in particular the Mott metal-insulator transition5. Despite various efforts, the nature of the change from metal to insulator in the two-dimensional (2D) Hubbard model at half-filling and intermediate to small interaction remains elusive.

At large interaction strength $U$, spin and charge degrees of freedom are decoupled and a gap opens in the density of states already at finite temperature. The Hubbard model approximately maps to a Heisenberg model in this regime with an effective exchange coupling $J = -4t^2/U$, where $t$ is the nearest-neighbor hopping. At zero temperature the local moments order due to the superexchange mechanism.

At small interaction, the nature of the phase connecting the weakly correlated Fermi liquid at high temperature with the antiferromagnet (AF) at zero temperature is much less clear. At $U = 0$, the Fermi surface is perfectly nested7 with antiferromagnetic nesting vector $q = (\pi, \pi)$. In one possible scenario, for small finite $U$ a spin-density wave instability develops at $T = 0$ and the MIT is a consequence of the backfolding of the Brillouin-zone due to the magnetic ordering. This is called the Slater transition8. Kyung et al.9 have argued for this scenario based on the two-particle self-consistent approach10–12 (TPSC).

In an alternative scenario propagated by Anderson13, the Hubbard model exhibits strong-coupling behavior for both strong and weak coupling, so that a Mott gap is present at any finite value of $U$ as in 1D14. As the temperature is lowered, local moments develop first because of the MIT and finally order at $T = 0$. In this scenario, AF is a consequence of the MIT, contrary to the Slater scenario. Moukouri et al.15 have argued in favor of this scenario based on the double occupancy and density of states in large cluster DCA calculations. Schäfer et al.16 identified a $U_c(T)$ based on a downturn in the electronic self-energy at the lowest Matsubara frequencies within the dynamical vertex approximation (DFA) and Quantum Monte Carlo (QMC). This downturn has been interpreted as a destruction of the Fermi surface due to scattering of the electrons at Slater paramagnons, fluctuations with a very large correlation length which can extend over thousands of sites. The downturn in self-energy has been found to correlate with a decrease in potential energy. This is consistent with the cellular DMFT results of Fratino et al.17, which show a decrease in potential energy when the finite cluster undergoes a transition to the antiferromagnetically ordered state.

In this work, we contribute to the current physical picture by studying the small interaction, low temperature region using the ladder dual fermion approximation (LDFA)18,19. The LDFA, as well as DFA belong to a class of methods known as diagrammatic extensions of dynamical mean-field theory20. Contrary to cluster methods, they include correlations over length scales covering hundreds of lattice sites. While the self-energy is approximate at any scale, good agreement of the LDFA with benchmarks has been found over a wide parameter range21. Despite similarities in the LDFA and DFA, they differ in how they treat the long-range fluctuations which are essential to respect the Mermin-Wagner theorem22. Because the physical question discussed here is closely related to the presence of these fluctuations, we do not necessarily expect the same results in these methods.

In ladder DFA the Mermin–Wagner theorem is a consequence of the so-called Moriya–λ correction23. Its purpose is to ensure the correct leading asymptotic behavior of the self-energy and can be understood as a mass term in the two-particle propagator (susceptibility) which cuts off the divergence that leads to spurious long-range order. In the LDFA, the diagrammatic corrections do not alter the leading term in the asymptotic behavior24. Here the long-range antiferromagnetic fluctuations are included through a self-consistent renormalization procedure which ensures the expected exponential scaling of the susceptibility at low temperature25.

In the following section II we discuss the model and some additional properties of the LDFA.
tical results and their implications for the small interaction and low temperature phase diagram of the two-dimensional Hubbard model are discussed in Sec. III. We conclude in Sec. IV.

II. MODEL AND METHOD

The Hubbard model on the 2D square lattice is described by the Hamiltonian

\[ H = -t \sum_{\langle ij \rangle} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}, \]

where \( t \) is the nearest-neighbor hopping and our unit of energy. We are interested in small to intermediate coupling \( U \) up to half of the bandwidth given by \( W = 8t \).

The LDFA is a particular diagram resummation scheme in the dual fermion (DF) approach. In DF the lattice model is replaced by a lattice of quantum impurity models which interact via auxiliary, so-called dual fermions. The strong local DMFT-like correlations are treated at the level of the impurity model, while non-local correlations are included diagrammatically. For a specific choice of self-consistency condition on the impurities, DMFT corresponds to non-interacting dual fermions. The latter couple to the physical fermions of the same flavor locally. Diagrams in terms of dual fermions can therefore be constructed based on physical considerations. In particular, we expect long-range particle-hole fluctuations to be dominant. The corresponding LDFA self-energy has the form

\[ \Sigma_{k\nu} = -\frac{1}{2N^2} \sum_{k,q} A_r F_{r}^{\nu\nu'} \tilde{G}_{k+q,\nu'+\omega} \tilde{G}_{k+q,\nu} + \left[ F_{\text{lad},r,q}^{\nu\nu'} - \frac{1}{2} F_r^{\nu\nu'} \right], \]

where the ladder diagrams are generated by the Bethe-Salpeter equation (BSE) for the dual vertex \( F_{\text{lad},r,q}^{\nu\nu'} \) of the lattice in the particle-hole channel,

\[ F_{\text{lad},r,q}^{\nu\nu'} = \frac{T}{N} \sum_k F_r^{\nu\nu'} \tilde{G}_{k+q,\nu'} \tilde{G}_{k+q,\nu} F_{\text{lad},r,q}^{\nu\nu'}. \]

Here we have introduced sums over four-momenta \( k = (k,\nu), \quad q = (q,\omega) \) and the spin and charge channels \( r = \text{sp}, \text{ch} \). \( \nu \) and \( \omega \) denote the discrete fermionic and bosonic Matsubara frequencies, respectively. We further have \( A_{\text{ch}} = 1 \), \( A_{\text{sp}} = 3 \), where the latter accounts for the degeneracy of the spin bosonic excitations. \( T \) denotes temperature and \( N \) is the total number of lattice sites. The second term in angular brackets avoids over-counting of the second-order diagram. \( F_r^{\nu\nu'} \) is the exact local reducible vertex of the impurity model. If we approximate the lattice vertex by the local one, \( F_{\text{lad},r,q}^{\nu\nu'} \approx F_r^{\nu\nu'} \), we obtain the second-order approximation \cite{28, 29} DF(2). Note that the LDFA includes diagrams from both the horizontal and vertical particle-hole channels (see for example the discussion in Ref. 21). In the spin channel, these diagrams generate the collective paramagnon excitations. Remarkably, the LDFA reproduces non-mean-field critical exponents. We refer the reader to Refs. 18 and 20 for further details on the method.

Below the DMFT Néel temperature \( T_N^{\text{DMFT}} \) we have to include the long-range fluctuations which destroy the AF order of the underlying mean-field. To this end, the dual Green’s functions \( G_{k\nu} \) in above equations are self-consistently renormalized: The self-energy \( \Sigma_{k\nu} \) is calculated starting from an initial guess for the dual Green’s function (typically the bare dual Green’s function). A new dual Green’s function is obtained via Dyson’s equation, which in turn is inserted into (3) and (2) to calculate a new self-energy (the impurity vertex is fixed). This process is repeated until self-consistency.

Since below \( T_N^{\text{DMFT}} \) the BSE (3) initially diverges, we cut off the AF fluctuations in the initial iterations of this inner self-consistency by restricting the eigenvalues of the BSE to values strictly smaller than 1. Once the iterations converge the cutoff is removed. When all BSE eigenvalues are smaller than unity in the final iteration the solution is well-defined and independent of the cutoff. The scheme is not guaranteed to converge and the number of iterations may diverge, which ultimately limits the accessible temperature range. The impurity model hybridization function is adjusted in an outer self-consistency loop based on the condition that the lowest-order dual diagram vanishes. At the values of \( U \) we are interested in, its effect is merely a small enhancement of the imaginary part of the impurity self-energy.

The calculations are carried out on lattices of finite size subject to periodic boundary conditions imposed by the discrete Fourier transform. Once the calculation is converged, we compute the physical self-energy \( \Sigma_{k\nu} \) and momentum-resolved susceptibility \cite{28}, \( \chi(\omega, q) = \langle \hat{S}^z \hat{S}^z \rangle_{\omega,q} \) with \( \hat{S}^z = (\hat{n}_\uparrow - \hat{n}_\downarrow) / 2 \).

III. RESULTS

A. Energetics

We first discuss the energetics of the model. Entropy disfavors antiferromagnetism, so the AF transition is driven by the interplay of kinetic and potential energy. In the picture of nearly free electrons at small interaction, AF ordering reduces the double occupancy. Magnetic ordering is hence potential energy driven. This is the Slater regime. At strong interaction on the other hand, double occupancy is largely suppressed. AF ordering promotes hopping processes and the (negative) kinetic energy is lowered. Magnetic ordering is stabilized through the reduction in kinetic energy. This is the Heisenberg regime. It is true even in the Mott insulator, where the exchange
The kinetic energy is measured with respect to the kinetic energy at $T = 0$ and $U = 0$, $E_{\text{kin}}^0 = -16/\pi^2$. Dashed lines and circles are the DMFT results, plus signs are LDFA.

In DMFT it is possible to compute the energy difference between the ordered and unordered states. In the Slater regime, the ordered state has a lower potential energy compared to the un-ordered state at $T = 0$ and $U = 0$, $E_{\text{pot}}^0 = -16/\pi^2$. Dashed lines and circles are the DMFT results, plus signs are LDFA.

DFA includes nonlocal correlations diagrammatically and up to significantly larger length scales compared to the cluster calculations. The correlations increase the kinetic energy at small $U$, while at large $U$ they decrease it. The potential energy was found to be reduced compared to DMFT at all studied values of $U$, both in the Slater and in the Heisenberg regime. However, the ambiguity of the potential energy in DMFT complicates the analysis.

In Figs. 1 and 2 we show the energetics extracted from LDFA calculations as a function of temperature at $U/t = 4$ and $U/t = 1$, respectively. In both cases we find that the potential energy is lowered in LDFA compared to DMFT (albeit only slightly at $U/t = 1$) in accordance with the cluster DMFT and DFA results and as expected in the Slater regime. At $U/t = 4$ the kinetic energy is also somewhat higher than in DMFT, while it seems slightly lower at $U/t = 1$ and low temperature. Here relatively large uncertainties however prevent a definite statement (note the scale compared to $U/t = 4$). The total energy of the system is lower when nonlocal correlations are included.

B. Density of States

The non-local fluctuations have a drastic effect on the local density of states (DOS) as seen in Fig. 3. While the DMFT solution exhibits a quasiparticle peak, the spectral weight is reduced in DFA. The reduction is small in second-order DFA (labeled DF(2)), but increases with order of the ladder diagrams. Remarkably, diagrams at all orders contribute to the pseudogap. The dual Green’s function decays rather rapidly in real space, with exponential decay on a length scale $\xi_G \approx 2.8$. Low-order diagrams hence mediate short-range correlations,
while long-range correlations require high diagram orders. They nevertheless contribute to short-range correlations as well.

The pseudogap develops in a regime where magnetic fluctuations are very strong and is linked to a reduction in potential energy, in accordance with DCA. The temperature here is slightly below the DMFT Néel temperature. The AF correlation length is of the order of $\xi_{AF} \sim 4$ (see Sec. III D). Extended AF fluctuations are hence present in the system and mediate correlations between electrons. The correlations may be singlet-like, which can lead to a strong suppression of the density of states and ultimately open the gap. These results are consistent with those of Ref. 15, where a gap fully opens at $U/t = 4$ and $T/t = 0.125$. Unfortunately we cannot reach such low temperatures at this value of $U$. On the other hand, the opening of the gap can be interpreted as a consequence of the scattering of quasiparticles off the AF fluctuations, which we include by construction in the LDFA. This interpretation and our numerical results are consistent with the DGA study of Ref. 16, which pointed out the importance of Slater paramagnons. The qualitative DOS predicted by TPSC at crossover temperatures is consistent with the LDFA results.

C. Uniform susceptibility

Figure 4 shows the temperature dependence of the uniform spin susceptibility $\chi^{sp}$ at $U/t = 4$. We find a decreasing susceptibility at high temperatures, even though Curie’s law $\chi \propto 1/T$ only sets in at higher temperatures, $T \approx W = 8t$. While the results of different methods agree at high temperature, they are qualitatively different at low $T$. In DMFT, $\chi$ continues to increase up to the point where the antiferromagnetic susceptibility diverges due to the second-order transition to the mean-field antiferromagnetic state at $T_N^{DMFT}$ (not shown). In LDFA the susceptibility decreases at low temperature and exhibits a maximum. AF correlations that build up reduce the uniform susceptibility. This occurs at the energy scale of the effective exchange interaction $J$ between neighboring sites. DMFT does not include such nonlocal correlations, so that the maximum is absent. Note that the maximum occurs at a slightly lower temperature than $T_N^{DMFT}$. At $T_N^{DMFT}$ the ladder diagram series (3) diverges, which in turn causes large effects in the self-consistent renormalization of the Green’s functions. The result are strong magnetic fluctuations which destroy the mean-field long-range order in the underlying DMFT solution. At the same time they lead to spin correlations between sites.

The LDFA susceptibility is in excellent agreement with lattice QMC results for the same lattice size. The comparison with results for a larger lattice reveals that finite size effects play a role. The magnitude of the susceptibility is significantly reduced (by about 10%) in the larger system. The position of the maximum however is not affected by finite size effects, which is consistent with its interpretation as an effective exchange energy scale.

The uniform spin susceptibility for $U/t = 1$ is shown in Fig. 5. As for $U/t = 4$, the nonlocal correlations reduce the susceptibility compared to DMFT, especially at lower temperatures. In LDFA, the magnitude of the susceptibility is reduced on the larger lattice, even more strongly than for $U/t = 4$. However we do not find a maximum in the susceptibility at the accessible temperatures down to $T/t = 1/70 \approx 0.014$. At half-filling and $U = 0$, there is a Van Hove singularity in the electron density of states exactly at the Fermi level and the uniform susceptibility shows a logarithmic divergence as $T \to 0$. Based on our results, it is not possible to distinguish between a maximum at a finite but very small temperature, or at $T = 0$. The DMFT Néel temperature is between $T/t = 0.025$ and $T/t = 0.022$, showing that it is not necessarily linked to the occurrence of the maximum as for $U/t = 4$. 

![Figure 4](image1.png)

**FIG. 4.** Uniform spin susceptibility as a function of temperature compute with different methods, all at $U/t = 4$. In QMC and LDFA, $\chi^{sp}$ exhibits a maximum at the effective exchange energy scale. The QMC results are taken from Ref. 38 for an $8 \times 8$ lattice.

![Figure 5](image2.png)

**FIG. 5.** Uniform spin susceptibility as a function of temperature, at $U/t = 1$. 

---

**RAW TEXT END**
In Fig. 6 we show results for several values of $U$. The location of the maxima for $U/t = 2, 3$ and $4$ agree within error bars with the QMC results of Ref. 40, where the location of the maximum has been used to find the energy scale at which magnetic correlations become important in optical lattice experiments. The location of the maximum in the susceptibility decreases with decreasing $U$, showing that the physics is clearly not Heisenberg-like. For the $U$ values where we obtain a maximum, it is close to the DMFT Néel temperature.

D. Finite-size effects

Figure 7 shows the leading eigenvalue of the BSE at the AF wave vector for $U/t = 1$ and $U/t = 4$ respectively, for different lattice sizes. At low temperature the antiferromagnetic susceptibility is expected to exhibit an exponential scaling $\chi_{AF} \propto e^{\Delta/T}$. It follows that the leading eigenvalue behaves as $(\chi_{AF})^{-1} \sim 1 - \lambda \sim \exp(-\Delta/T)$. $1 - \lambda$ remains non-zero for finite $T$ as required by the Mermin-Wagner theorem. We observe the scaling for $U/t = 4$ for the larger lattice, and not at all for $U/t = 1$.

Figure 8 shows the spin correlation function in real space. At $U/t = 4$ and $T/t = 0.2$ the AF correlation length is short, $\xi_{AF} \approx 4$. This is consistent with the fact that no finite-size effects are visible at this temperature in Fig. 7. At $U/t = 1$ and $T/t = 0.02$, the correlation length is significantly longer, $\xi_{AF} \approx 17$, but unlikely to explain the absence of the scaling for the $128 \times 128$ lattice.

In addition to the length scale associated with two-particle fluctuations, we can introduce a scale related to single-particle properties obtained from the Green’s function at $\tau = \beta/2$, $\xi_G$. For $U/t = 4$ and $T/t = 0.2$, we find $\xi_G \approx 2.8 < \xi_{AF} \approx 4$, while for $U/t = 1$ and $T/t = 0.02$ we obtain $\xi_G \approx 35$, which is larger than $\xi_{AF}$ and the smaller lattice and may explain why finite-size effects act in opposite directions for $U = 4$ and $U = 1$. Even though we cannot rule out that the exponential scaling of $1 - \lambda$ is obscured by finite size effects, we expect that it sets in at even lower, inaccessible temperatures.

Figs. 10 and 11 show the local Green’s function for the same $U$ values. For $U/t = 4$, finite-size effects are clearly absent for large lattices. For $U/t = 1$ small finite-size effects are visible even on the largest lattice. Note that the finite-size effects are also visible in DMFT. This underlines that they are related to the momentum discretization. Fig. 9 illustrates this for the noninteracting DOS, which requires of the order of $64 \times 64$ points at $T = 0.02$ to be accurately represented.
E. Self-energy

In Fig. 12, we plot the momentum-resolved difference of the self-energy at the two lowest Matsubara frequencies at $U/t = 4$ and $T/t = 0.2$. The presence of positive values near the point $k = (\pi, 0)^{14}$ indicates a downturn in the self-energy and signals a breakdown of Fermi liquid theory. This feature is robust with respect to the lattice size and consistent with the presence of the pseudogap. At the studied temperature, the downturn only occurs close to $k = (\pi, 0)$. By lowering the temperature, the momentum space region with positive values of the self-energy difference grows. DFA results indicate that there should be a downturn along the entire Fermi surface at $T/t \approx 0.11$ which is outside the accessible temperature range $T/t \geq 0.196$. We do not find a divergence of the self-energy and hence no opening of a gap at finite temperature.

In Fig. 13, we show corresponding results for $U/t = 1$ and $T/t = 0.02$. We do not observe a downturn, but values close to zero (a flattening of the self-energy at low frequencies) in a very narrow strip along the Fermi surface (the diagonal from top left to bottom right). Compared to Fig. 12, the tendency towards a downturn occurs in a much narrower part of the Brillouin Zone. This suggests that the phenomenon is related to AF fluctuations on very long length scales.

Figure 14 shows that finite-size effects can change behavior qualitatively. A downturn in the self-energy occurs on a $32 \times 32$ lattice, while it is absent on the larger $128 \times 128$ lattice. If the scattering of electrons at magnetic fluctuations induces the downturn, it is conceivable that the downturn occurs because the antiferromagnetic susceptibility is overestimated, as visible in Figs. 5 and 7. The downturn, and ultimately an opening of the gap may still appear at lower temperatures. However we do not observe a downturn in the accessible temperature range. This excludes a crossover down to $T/t = 1/69$, in contrast to the crossover temperature of $T/t \approx 1/38$ obtained in DFA$^{16}$. The discrepancy between both methods may lie in the way they treat the magnetic fluctuations.

Weak-coupling approximations can put these results into perspective. The Hartree-Fock gap$^{45,46} \Delta E = 32t \exp(-2\pi \sqrt{U/t})$ is exponentially small in $U$, and predicts a typical energy scale $0.06t$ at $U/t = 1$ due to the significant prefactor. The temperatures studied here are actually below this scale. Renormalization-group analysis$^{47}$, on the other hand, suggests a critical energy scale of approximately $0.02t$ for the formation of bound particle-hole pairs, which is exactly in the range studied here.

IV. CONCLUSIONS

We have studied the half-filled 2D Hubbard model on the square lattice in the interaction range $1 \leq U/t \leq 4$ using the LDFA. In accordance with DFA$^{43}$ and cellular DMFT$^{17,32}$ results, non-local AF correlations reduce the potential energy as expected in the Slater regime. Strong AF fluctuations develop in the vicinity of the DMFT Neél temperature. Scattering of electrons off these fluctuations leads to (presumably singlet-like) correlations and
FIG. 12. Difference in self-energy at the lowest two Matsubara frequencies, \( \text{Im} \Sigma_{\nu_1} - \text{Im} \Sigma_{\nu_0} \), at \( U/t = 4 \) and \( T/t = 0.2 \) on a \( 256 \times 256 \) lattice. Positive values indicate a downturn. Only a quarter of the Brillouin Zone is shown. The black contours indicate where the self-energy difference is zero.

FIG. 13. Difference in self-energy at the lowest two Matsubara frequencies, \( \text{Im} \Sigma_{\nu_1} - \text{Im} \Sigma_{\nu_0} \) at \( U/t = 1, T/t = 0.02 \), on a \( 256 \times 256 \) lattice. Note the extremely sharp feature at the Fermi surface. However, it does not become positive, which would signal a downturn.

to a pseudogap in the density of states. A concomitant downturn is observed in the self-energy at \( U/t = 4 \). We further find a maximum in the uniform susceptibility in good agreement to Monte Carlo results. It can be explained through the buildup of AF correlations on an effective exchange energy scale \( J \). This scale gets smaller when \( U \) decreases.

We found that finite-size effects can alter the results qualitatively. While they play a minor role at \( U/t = 4 \) due to relatively short correlations lengths, they lead to a spurious downturn in the self-energy at \( U/t = 1 \). Nevertheless a tendency to a downturn in the self-energy is observed for \( U/t = 1 \) in a very narrow region in the Brillouin zone along the Fermi surface, which suggests that a possible crossover would be associated with long-range AF fluctuations.

We do not find a crossover to an insulator in the accessible temperature range for \( 1 \leq U/t \leq 4 \). From our calculations we obtain an upper bound \( T_{\text{downturn}}/t < 0.0145 \) on the crossover temperature for \( U/t = 1 \), which is significantly lower than the value reported in Ref. 16. The origin of this discrepancy between LDFA and DΓA might lie in the way the methods treat the mean-field divergence of the AF susceptibility.

ACKNOWLEDGMENTS

We thank Emanuel Gull, Georg Rohringer, Thomas Schäfer and Alessandro Toschi for useful discussions. E.G.C.P. v. L. and M.I.K. acknowledge support from ERC Advanced Grant 338957 FEMTO/NANO. Our implementation is based on the ALPS framework and on the impurity solver of Ref. 49.

FIG. 14. Finite-size effects in the self-energy, at \( U/t = 1 \) and \( T/t = 0.02 \). The self-energy is shown at the \( X = (\pi, 0) \) and \( M' = (\pi/2, \pi/2) \) points.

7 We consider the Hubbard model with only nearest-neighbor hopping.
8 J. C. Slater, Phys. Rev. 82, 538 (1951).


23 This is because the dual Green function decays as \(1/(iv)^2\).


34 The LDFA energetics are obtained through frequency- and momentum summations of single-particle quantities as in 33 including tail corrections, which are susceptible to numerical errors. In DMFT, the kinetic energy is obtained from the average perturbation order in the CTHYB solver\(^30\) and the potential energy from the Migdal-Galitskii formula\(^31\). The Matsubara frequency summations are most difficult at low temperatures, where very many frequencies are needed.

35 This scale is determined by an exponential fit to \(G(x, y, \tau = \beta/2)\) along the diagonal \(x = y\). The decay is highly anisotropic and the decay length is longest along the diagonal.


