Inclusion of intersite spatial correlations in the alloy analogy approach to the half-filled ionic Hubbard model

D A Rowlands and Yu-Zhong Zhang

Shanghai Key Laboratory of Special Artificial Microstructure Materials and Technology, School of Physics Science and Engineering, Tongji University, Shanghai 200092, People’s Republic of China

E-mail: yzzhang@tongji.edu.cn

Received 22 September 2013, revised 25 October 2013
Published 17 June 2014

Abstract
Using the nonlocal coherent-potential approximation we study the effect of intersite spatial correlations on the transition from band insulator to metal as well as from metal to Mott insulator in the ‘alloy analogy’ approach to the paramagnetic solution of the half-filled ionic Hubbard model. We find that intersite spatial correlations enhance the metallic phase.

(Some figures may appear in colour only in the online journal)

1. Introduction
The Hubbard model of interacting electrons and their associated correlation effects in narrow d and f energy bands [1] has attracted a huge amount of theoretical and computational effort over a number of decades. However, an early theory due to Hubbard in his third paper on the electron correlation problem remains of significant importance since it was the first to predict a Mott metal–insulator transition (MIT) at a finite critical value of the on-site interaction $U$ [2]. By making an analogy with disordered alloys (the ‘alloy analogy’ (AA)), Hubbard used a Green’s function decoupling method to derive a ‘scattering correction’ to his earlier Hubbard I approximation [1], and his solution has been shown [3] to be equivalent to the coherent-potential approximation (CPA) which was not actually introduced for disordered systems until a few years later [4]. Hubbard also added a ‘resonance broadening’ correction to the AA and this is referred to as the full Hubbard III approximation [2].

Hubbard’s AA improves upon the Hartree–Fock approximation by considering classical or static fluctuations in the potential that an electron sees [5], and remains useful as a computationally simple theory capable of giving a valid approximate description of the MIT. Nevertheless, one drawback of Hubbard’s work is that the scattering correction obtained from his solution to the AA (i.e. the CPA) is a local approximation, meaning that intersite or nonlocal spatial correlations in the required alloy configurational average are neglected. Previous work on addressing this issue has been carried out many years ago by Mors et al [6] who incorporated a crude two-site spatial correlation in order to study nonlocal effects on the MIT, and more recently by Luo et al [7] based on a combination of the full Hubbard III approximation with the spectral density approach [8, 9] and a nonlocal self-energy ansatz. In the present paper we solve the AA using the nonlocal CPA [10–13] which is now an established mean-field approach for describing nonlocal physics and short-range order [14]. Like the CPA it requires the solution of a self-consistent impurity problem but with the added capability of including intersite spatial correlations to arbitrary length scale depending on the size of the impurity cluster used in the calculation.

We study the interesting case of the ionic Hubbard model [15] (IHM) which adds a staggered ionic potential to the conventional Hubbard model on a bipartite lattice i.e. two sublattices with different on-site energies. This enables us to examine nonlocal effects on the expected transition from band insulator to metal as well as the MIT itself across a range of staggered ionic energies. While the two-sublattice framework opens up the possibility of investigating magnetic ordering via incorporating some of the modifications to the AA mentioned later in section 2.1, here we choose to focus on the paramagnetic solution at half-filling.

The IHM has received attention as a potential model for describing the ferroelectric transition in perovskite materials [16] and the neutral–ionic transition in organic mixed-stack donor–acceptor crystals [17]. It has been extensively studied in both one dimension (1D) and two dimensions (2D) but the true nature of its phase diagram at half-filling is still...
undecided. In 1D, Fabrizio et al. [18] performed a study using bosonization and found that the bandgap due to the ionic potential is suppressed at a critical value of $U$ where the system undergoes a transition to an intermediate bond-ordered phase before undergoing a transition to a Mott insulating phase at a second critical value of $U$. Some contradictory results were obtained, for example Wilkens and Martin [19] found no Mott transition using quantum Monte Carlo (QMC). However, other later studies found results similar to Fabrizio et al. [18], for example Batista and Aligia [20], Zhang et al. [21] and Manmana et al. [22], the latter two by using density matrix renormalization group (DMRG) theory.

The situation is far less conclusive in 2D. Using dynamical mean-field theory (DMFT), Garg et al. [23] found a metallic phase sandwiched between the band insulating phase due to the ionic potential and the Mott insulating phase in the strong-coupling regime. Similar results were later obtained by Hoang [24] simply by using the CPA. Craco et al. [25] also used DMFT but found phase transitions of a different nature from metal to Mott insulator and from band insulator to Mott insulator in the phase diagram. Using determinant-QMC, Paris et al. [26] and Bouadim et al. [27] found that in contrast to the DMFT studies, in the absence of the ionic potential the system is an insulator for any value of $U$ due to the inclusion of intersite magnetic fluctuations. The Mott insulating phase was also found to be anti-ferromagnetic. Subsequently, Byczuk et al. [28] re-examined the IHM using DMFT but allowed for anti-ferromagnetic order and found that the system is insulating for all interaction strengths with no metallic phase. Kancharla and Dagotto [29] used cellular-DMFT to incorporate nonlocal spatial correlations to the presence of the staggered ionic energies. Following this reasoning we may approximate the above many-body Hamiltonian by the one-electron Hamiltonian

$$H = \sum_{i \in A, \sigma} E_{A\sigma} n_i^{\sigma} + \sum_{i \in B, \sigma} E_{B\sigma} n_i^{\sigma} - t \sum_{i \in A, j \in B, \sigma} \left[ (c_i^{\dagger})^{\sigma} c_j^{\sigma} + \text{H.c.} \right] + \mu \sum_{i} n_i.$$  (1)

where the disorder potential has been defined to include the chemical potential and is given by

$$E_{\alpha\sigma} = \begin{cases} \epsilon_{\alpha} + U/2 & \text{with probability } \langle n_{\alpha, -\sigma} \rangle \\ \epsilon_{\alpha} - U/2 & \text{with probability } 1 - \langle n_{\alpha, -\sigma} \rangle \end{cases}$$  (3)

at half-filling with $\mu = U/2$. Here $\langle n_{\alpha, \sigma} \rangle$ is the average electron occupancy per site for sublattice $\alpha$ with spin $\sigma$.

It is well known that the AA itself has many shortcomings. These are discussed in the book by Gebhard [31], and have been found to be mainly due to the arbitrariness in the intuitive way Hubbard performed the decoupling of the higher-order Green’s functions [31]. Consequently many efforts have been made over the years to improve the approximation. Before going on to develop the formalism for this paper, it is appropriate to mention a small selection of these efforts here.

The first forthcoming to be highlighted is the inability of the AA to describe magnetism [32, 33]. For example, in the case of half-filling a transition to an anti-ferromagnetic Mott insulator is not observed in the strong-coupling regime and it is only possible to obtain the paramagnetic solution. One of the better known theories aimed at correcting the AA is the modified alloy analogy (MAA) of Herrmann and Nolting [34], which was derived by combining the spectral density approach [8, 9] (SDA) with the CPA in order to reproduce the exact results of Harris and Lange [35] as $U \to \infty$. The MAA is based upon the notion that the atomic levels and concentrations used in Hubbard’s AA are not the only choice available for constructing an alloy analogy. The MAA still uses the CPA
equation but replaces the atomic levels and concentrations with modified expressions which include a band-shift correction. Potthoff et al. [36] showed that this choice is optimal for reproducing the first four moments of the spectral density in contrast to the first three in Hubbard’s AA. Unlike Hubbard’s AA, the MAA is capable of producing magnetic solutions, albeit over restricted parameter ranges. Another approach capable of capturing the transition to anti-ferromagnetic Mott insulator is the mean-field theory of Janis and Vollhardt [37, 38] who constructed a thermodynamically consistent alloy analogy.

A second deficiency of the AA is that the metallic state is not described as a Fermi liquid. This is a natural consequence of viewing the system in terms of a set of static impurity scattering centres which leads to a self-energy with a finite imaginary part at the Fermi level. Although including the resonance broadening correction of the full Hubbard III approximation will take into account dynamic processes at the impurity scattering centre by relaxing the assumption that the opposite spin electrons are fixed at the lattice sites, this is still not a true description of dynamics in the sense that the effective field seen by the spin σ electron is not time-dependent. Nevertheless, Edwards and Hertz [39, 40] used the limit of infinite dimensions to construct a method for correcting this non-Fermi liquid behaviour in the weak-coupling limit using a modified Green’s function, the method being exact for small U.

The IHM has previously been studied using the CPA by Hoang [24]. This in turn is based on the earlier framework of Gupta et al. [46] who investigated the possibility of magnetic ordering in the Hubbard model using a two-sublattice version of the CPA which we refer to as the ‘coupled-CPA’ method.

In principle, the Green’s function corresponding to the Hamiltonian of equation (2) needs to be averaged over all possible disorder configurations. Such an exact average is not feasible, the CPA introduces a simplified effective medium corresponding to the approximate Hamiltonian

\[ H_{\text{CPA}} = \sum_{i \in A, \sigma} \Sigma_{A\sigma} n_{i\sigma} + \sum_{i \in B, \sigma} \Sigma_{B\sigma} n_{i\sigma} - t \sum_{i \in A, j \in B, \sigma} \left( c_{i\sigma}^\dagger c_{j\sigma} + H.c. \right) \]

(4)

where each sublattice A(B) has its own single-site or local self-energy \( \Sigma_{A\sigma(B\sigma)} \) situated at every sublattice site. In \( k \)-space, the CPA average Green’s function can be written in the two-sublattice matrix form

\[ G_a(k, \omega) = \left( \omega - \Sigma_{A\sigma} - t(k) \right)^{-1} \]

(5)

Carrying out the matrix inversion yields

\[ G_{AA\sigma}(k, \omega) = \frac{1}{\Omega_{\text{BZ}}} \int_{\Omega_{\text{BZ}}} \frac{d\omega}{(\omega - \Sigma_{A\sigma})(\omega - \Sigma_{B\sigma}) - t(k)^2} \]

(6)

\[ G_{BB\sigma}(k, \omega) = \frac{\omega - \Sigma_{B\sigma}}{(\omega - \Sigma_{A\sigma})(\omega - \Sigma_{B\sigma}) - t(k)^2} \]

(7)

for the sites on sublattice A and B respectively, where \( k \) belongs to the first BZ of the sublattice considered. In real space we have

\[ G_{\alpha\sigma}(\omega) = \frac{1}{\Omega_{\text{BZ}}} \int_{\Omega_{\text{BZ}}} \frac{d\omega}{(\omega - \Sigma_{\alpha\sigma})(\omega - \Sigma_{\beta\sigma}) - t(k)^2} \]

(8)

where \( \alpha = A(B) \), \( \beta = B(A) \) and the integral is over the first BZ of the sublattice. Since we are only considering the diagonal terms of the Green’s function matrix, notation has been adopted such that \( \alpha \equiv \alpha \sigma \) for clarity. Using the relation for the bare integrated DOS

\[ \int dE \rho_0(E) = \int dE \frac{1}{N} \sum_k \delta(E - t(k)) \]

(9)

for real energies \( E \), equation (8) can be re-written in the form

\[ G_{\alpha\sigma}(\omega) = (\omega - \Sigma_{\alpha\sigma}) \int dE \frac{\rho_0(E)}{(\omega - \Sigma_{A\sigma})(\omega - \Sigma_{B\sigma}) - E^2} \]

(10)

for which approximations to the bare DOS can be made. To determine the CPA self-energies, Hoang [24] proceeded by employing Hubbard’s semi-elliptic model DOS [1]. This simplifies the above integral and yields a pair of equations for \( \Sigma_{A\sigma} \) and \( \Sigma_{B\sigma} \) which need to be solved self-consistently. Here we do not make any approximations for the DOS; instead we determine the self-energies numerically. We begin by defining the cavity Green’s function \( G_{\alpha\sigma}(\omega) \) through the relation

\[ G_{\alpha\sigma}^{-1}(\omega) = G_{\alpha\bar{\sigma}}^{-1}(\omega) + \Sigma_{\alpha\sigma}(\omega) \]

(11)

for each sublattice \( \alpha \), which describes the medium with the self-energy at some chosen site on each sublattice removed i.e. a cavity. In this definition it is not necessary to consider the off-diagonal terms in the Green’s function matrix, \( \tilde{G}_{AB\alpha} \) and \( \tilde{G}_{BA\alpha} \), since there are no self-energy terms coupling the sublattices. We may now fill the cavity on each sublattice with
some real ‘impurity’ configuration by defining the impurity Green’s functions
\[
G_{\alpha\sigma}^\gamma(\omega) = [G_{\alpha\sigma}(\omega) - E_{\gamma}^{\alpha\sigma}]^{-1}
\]  
with impurity configurations \(E_{\gamma}^{\alpha\sigma} = \epsilon_\alpha \pm U/2\) as defined by equation (3). The CPA demands that
\[
\langle G_{\alpha\sigma}^\gamma(\omega) \rangle = \tilde{G}_{\alpha\sigma}(\omega),
\]  
where the average is taken over the impurity configuration probabilities defined by equation (3), i.e.
\[
(n_{\alpha,-})G_{\alpha\sigma}^{\gamma=}(\omega) + (1 - (n_{\alpha,-}) G_{\alpha\sigma}^{\gamma=-}(\omega) = \tilde{G}_{\alpha\sigma}(\omega).
\]  
Equations (8) and (13) thus need to be solved self-consistently. Since in the IHM electrons generally prefer to be on sublattice B and we have \(n_A + n_B = 2\) at half-filling where
\[
(n_\alpha) = \langle n_{\alpha\uparrow} \rangle + \langle n_{\alpha\downarrow} \rangle = -\frac{1}{\pi} \int_{-\infty}^{0} \text{Im}[\tilde{G}_{\alpha\uparrow} + \tilde{G}_{\alpha\downarrow}] d\omega,
\]  
it is also necessary to ensure the resulting integrated DOS for each sublattice both consistent with the average occupation number probabilities used in equation (13), thus adding an extra layer of self-consistency.

As mentioned in the introduction, there are no antiferromagnetic or ferromagnetic CPA solutions for the (non-degenerate) Hubbard model [32, 33]. This is also true in the presence of staggered ionic energies and so here we are restricted to the paramagnetic solution. Thus the Green’s function is the same for both spin populations so that \(G_{\alpha\uparrow} = G_{\alpha\downarrow}\) and \(\langle n_{\alpha\uparrow} \rangle = \langle n_{\alpha\downarrow} \rangle = n_\alpha/2\). Finally note that although there are no off-diagonal self-energy terms coupling the sublattices, equation (8) shows that the sublattice self-energies are not determined independently of each other.

2.3. Nonlocal CPA for the ionic Hubbard model

The nonlocal CPA introduced by Jarrell and Krishnamurthy [10] systematically includes nonlocal correlations beyond the CPA via the self-consistent embedding of a cluster of sites with periodic boundary conditions imposed. The reader is referred to [10, 14] for full details of the method; here we simply show how to implement for the IHM.

Because of the boundary conditions, the mapping of the cluster to the lattice is achieved in \(k\)-space [10]. The BZ is divided into \(N_c\) regions centred at the set of ‘cluster momenta’ \(\{\mathbf{K}\}\). Correspondingly in real space the lattice is divided into \(N_c\) sublattices, and a site from each sublattice taken together form a real space cluster [14]. Some example real space clusters and corresponding BZ regions for the 2D bipartite lattice are illustrated in figure 1. The main approximation made is to represent the exact lattice self-energy in \(k\)-space \(\Sigma(\mathbf{k}, \omega)\) by the step function \(\Sigma(\mathbf{K}, \omega)\) in the regions centred at the cluster momenta. In this approximation the translational invariance of the underlying lattice is always preserved irrespective of the size of the cluster. In the case of the IHM, the translational invariance of the underlying two-sublattice unit cell must be preserved and hence \(\Sigma(\mathbf{K}, \omega)\) also becomes sublattice-dependent. Therefore an effective medium is introduced corresponding to the approximate \(k\)-space Hamiltonian
\[
H = \sum_{\mathbf{K}, \mathbf{k}'} \Sigma_{\alpha\alpha}(\mathbf{K}, \omega)n_{\alpha, \mathbf{K}+\mathbf{k}'} + \sum_{\mathbf{K}, \mathbf{k}} \Sigma_{\beta\beta}(\mathbf{K}, \omega)n_{\beta, \mathbf{K}+\mathbf{k}'}
\]  
with \(\Sigma_{\alpha\alpha}(\mathbf{K}, \omega) = \frac{1}{2} \bar{U} - t(\mathbf{K}+\mathbf{k}') + \Sigma_{\beta\beta}(\mathbf{K}, \omega) - t(\mathbf{K}+\mathbf{k}')\) and \(\Sigma_{\alpha\beta}(\mathbf{K}, \omega), \Sigma_{\beta\alpha}(\mathbf{K}, \omega)\) being the on-site and nearest-neighbor hopping terms.

The next step is to define the ‘coarse-grained’ \(k\)-space sublattice Green’s functions by integrating over each region of \(k\)-space or ‘tile’ \(n\) centred at the cluster momenta \(\mathbf{K}\):
\[
\bar{G}_{\alpha\beta}(\mathbf{K}_n, \omega) = \frac{N_c}{2} \frac{1}{\Omega_{\text{BZ}}} \int d\mathbf{k}' \bar{G}_{\alpha\beta}(\mathbf{k}', \omega).
\]  

These can be transformed to real space using the Fourier transform
\[
\bar{G}_{\alpha\beta}(\mathbf{k}_n, \omega) = \frac{2}{N_c} \sum_{\mathbf{k}_n} \bar{G}_{\alpha\beta}(\mathbf{k}_n, \omega) e^{i\mathbf{k}_n(\mathbf{R}_I - \mathbf{R}_J)}
\]  
for cluster matrix elements \([I, J]\) where \(N_c\) is the number of sites in the cluster. Together these form the supermatrix
\[
\bar{G}(\omega) = \begin{bmatrix}
\bar{G}_{\alpha\alpha}(\omega) & \bar{G}_{\alpha\beta}(\omega) \\
\bar{G}_{\beta\alpha}(\omega) & \bar{G}_{\beta\beta}(\omega)
\end{bmatrix}
\]  
which has dimension \(N_c\). It can be seen that sites belonging to sublattices A and B are grouped separately. We can similarly Fourier transform the \(k\)-space self-energies to real space to obtain the cluster self-energy matrix \(\Sigma\). To determine the
Green's functions and self-energies, we define the cavity Green's function \( G(\omega) \) through the relation

\[
G^{-1}(\omega) = G^{-1}(\omega) + \Sigma(\omega).
\]  

(23)

Note that unlike the 'coupled-CPA' method of section 2.2, here it is necessary to include the inter-sublattice terms in the Green’s function matrix, \( G_{AB} \) and \( G_{BA} \), when defining the cavity Green’s function due to the presence of the inter-sublattice self-energy terms. The impurity Green’s function is then given by

\[
G^\gamma(\omega) = \left[ G^{-1}(\omega) - E^\gamma \right]^{-1}
\]  

(24)

for some cluster disorder configuration \( \gamma \), where the site component of each configuration is one of \( \epsilon_A \pm U/2 \) or one of \( \epsilon_B \pm U/2 \) as defined by equation (3) for every site on sublattice A and B. The nonlocal CPA demands that

\[
\langle G^\gamma(\omega) \rangle = \bar{G}(\omega)
\]  

(25)

where the average is taken over all \( 2^{N_c} \) possible cluster configurations with the site probability component of each cluster probability defined by equation (3). The nonlocal CPA solution is obtained by solving equations (21) and (25) which strictly speaking are cluster quantities. Again it is necessary to ensure that the resulting integrated DOS for each cluster site is consistent with the relevant site occupation number probability used in the cluster configurations of equation (25), thus adding an extra layer of self-consistency. As a consequence of translational invariance, all sites on sublattice A will have the same average occupation number probability (integrated DOS per site), as will all sites on sublattice B. At half-filling, the integrated DOS summed over all cluster sites should equal \( N_c \).

In the limit where \( \Delta = 0 \), the two sublattices are equivalent. Nevertheless, the system can still be viewed in terms of a two-sublattice unit cell. Then for a cluster of size \( N_c \) it is still only necessary to consider the \( N_c/2 \) cluster momenta associated with the BZ of the two-sublattice unit cell. This is because for \( \Delta = 0 \) the BZ integration with \( N_c \) cluster momenta for the usual single-site unit cell becomes equivalent to two sublattice-BZ integrations, each with \( N_c/2 \) cluster momenta, which in turn are equivalent to the BZ of the two-sublattice unit cell. It is also worth looking at the limit where the cluster size \( N_c = 2 \). In that case there will still be a different self-energy value for each sublattice in \( k \)-space and hence off-diagonal self-energy terms which couple the two sublattices in real space. Hence we obtain an improved version of the coupled-CPA method of section 2.2. It should also be mentioned that the choice made for the set of cluster momenta is not unique, leading to non-unique results for small cluster sizes [47]. A reformulation of the nonlocal CPA which uses multiple sets of cluster momenta has been derived which corrects this problem [47], also naturally providing a lattice self-energy and Green’s function unlike equations (21) and (25) which strictly speaking are cluster quantities. However, in this paper we use only one set of cluster momenta for simplicity.

Finally, we note that a multi-sublattice implementation of the nonlocal CPA has previously been carried out by Marradodo and Staunton for a 1D tight-binding model alloy [48]. Our formulation for the IHM above can be viewed as an important extension to the case of strongly correlated systems and has potential applications to various strongly correlated models like the d–p model originally proposed for describing high-\( T_c \) cuprates [49], the Hubbard model on the honeycomb lattice [50] and the extended Hubbard model on the chequerboard lattice [51], etc.

3. Results

The 2D phase diagram for the system at zero temperature has previously been calculated using the CPA by Hoang [24] who compared the CPA results with those given by other approaches. Since our main motivation here is to study the effect of intersite correlations using the nonlocal CPA, we calculate the 2D phase diagram for the system using the same parameters as Hoang [24] so that direct comparisons can be made. Thus the energy unit is set by \( W = 4t = 1 \) where \( W \) is the half-width of the band. In order to determine the phase boundaries, the CPA or nonlocal CPA medium and site occupation number probabilities were first determined self-consistently using energies with an imaginary part \( \delta = 0.01 \). Using the same number probabilities, the DOS at the Fermi level \( \rho(0) \) was then calculated by determining the medium self-consistently as a function of \( \delta \). Extrapolation via
polynomial fitting was used in order to obtain the value with \( \delta \rightarrow 0^+ \). In the case of a finite gap, \( \rho(0) \) extrapolates to zero. In contrast \( \rho(0) \) remains finite in the metallic region [25]. Taking into account tolerances used in the calculations, \( \rho(0) \geq 0.01 \) was considered to be a finite value. Figure 2 shows example CPA calculations for \( \rho(0) \) as a function of \( \delta \). It is quite clear that the graphs for \( U = 0.85 \) and 1.00 will not extrapolate to zero via polynomial fitting and hence indicate metallic phases, whilst the graphs for \( U = 0.1 \) and 1.45 indicate insulating phases.

The phase diagram for the system calculated using both the CPA and nonlocal CPA with a cluster size \( N_c = 4 \) is shown in figure 3, with the CPA phase boundaries denoted by dotted lines and nonlocal CPA by solid lines. The CPA phase diagram is similar to that of Hoang [24], though the metallic phase is very slightly wider. The small difference in shape in the weak-coupling regime is likely to be due to the fact that a model DOS has not been used in our calculations. It can be seen that the phase boundary between the band insulator and metallic phases is unchanged in the nonlocal CPA calculations compared to the CPA and approaches the line \( U = 2\Delta \). On the other hand, the phase boundary for the MIT is significantly raised from the CPA value in the weak-coupling region, but again approaches the strong-coupling line \( U = 2\Delta \) as \( \Delta \) is increased. Thus intersite spatial correlations appear to enhance the metallic phase compared to the CPA, particularly in the weak-coupling regime. We note that for \( \Delta = 0 \) this behaviour is in agreement with the results of both Mors et al [6] and Luo et al [7] for the conventional Hubbard model, where according to their methods it was found that intersite spatial correlations shift the critical value of \( U \) for the MIT to a higher value. Also present in figure 3 is a small insulating region within the metallic sector between \( \Delta = 0.0 \) and 0.25 which we believe to be unphysical and resulting from a shortcoming of the nonlocal CPA method for small cluster sizes (see \( N_c = 8 \) calculation below).

To identify the origin of the differences between the CPA and nonlocal CPA phase diagrams, we now examine some example DOS results. First we consider the CPA. Figures 4(a)–(d) show plots with ionic energy \( \Delta = 0.1 \) for a selection of values of \( U \). The general features of the plots are similar to those of Hoang [24], though the curves are of more realistic shape since here the use of a model DOS has been avoided. For \( U = 0.1 \) the system is a band insulator with charge gap slightly smaller than the value \( 2\Delta \) for the non-interacting (\( U = 0 \)) limit. The results for values \( U = 0.50 \) and \( U = 1.30 \) clearly lie within the metallic and Mott insulating regimes respectively. The critical value for the metal–insulator transition is found to be very close to \( U = 1 \).

The nonlocal CPA results for the same parameters with \( N_c = 4 \) are shown in figures 4(e)–(h). The curve for \( U = 0.1 \) is similar to that of the CPA, indicating that for relatively small values of \( U \) the effect of intersite correlations is small. On the other hand, immediately apparent is the emergence of a peak at the Fermi level in the metallic phase as seen in the example plot for \( U = 1.30 \). As the Coulomb repulsion is increased further, the peak gradually diminishes. However, the transition from metal to Mott insulator is significantly delayed resulting in a larger critical value for \( U \). A similar peak for the conventional Hubbard model (i.e. \( \Delta = 0 \)) was found by Luo et al [7] using their method involving combination of the SDA with the full Hubbard III approximation and making a nonlocal self-energy ansatz. In their work it was interpreted as a quasiparticle peak resulting from the nonlocal nature of the self-energy i.e. a different mechanism to the well-known peak observed in DMFT calculations due to the dynamical though local nature of the self-energy. However Luo et al [7] do not indicate the nature of the imaginary part of the self-energy at the Fermi level. In fact a finite value was observed in DMFT calculations [25] which implies a finite lifetime for the state. Craco et al [25] identified this as being a consequence of a pseudo-gap feature at the Fermi level in the DOS on the B sublattice, meaning that the one-electron DOS in the metallic phase is not attached to the unperturbed (\( U = \Delta = 0 \)) value. In our work we also obtain a finite value but we regard this as being in agreement with the notion that viewing the system in terms of a set of static impurity scattering centres in the AA approach rules out the possibility of a Fermi liquid. Hence we consider the presence of the peak to be simply due to DOS at the band edges arising from cluster contributions as is the case.
in conventional disordered alloy calculations in the split-band regime.

In order to study the effect of longer-ranged intersite correlations, the nonlocal CPA phase diagram calculated using a larger cluster size, $N_c = 8$, is shown in figure 3 together with the earlier CPA calculation. The lower phase boundary between band insulator and metal shows similar behaviour to the $N_c = 4$ calculation. As expected, the small spurious insulating region present in the $N_c = 4$ phase diagram for values of $U$ just below the upper CPA phase boundary has disappeared. On the other hand, as $U$ is increased above the upper CPA phase boundary the system surprisingly appears to switch back and forth between metal and Mott insulator so that the MIT occurs three times. We believe the reason for this lies in a shortcoming of the conventional nonlocal CPA method. The medium self-energy should be a smooth function in $k$-space for all cluster sizes. However, for the purposes of producing an analytic Green’s function, the nonlocal CPA approximates the medium self-energy as a step function in $k$-space [10]. It has previously been shown that such an approximation can lead to spurious DOS results for small cluster sizes where the jumps between the steps are large [52]. In particular, artificial gaps appear at certain energies. In the present context, this could cause artificial gaps to appear at the Fermi level for certain values of $U$, resulting in the extra insulating phases denoted by the shaded regions in the phase diagrams of figures 3 and 5. This can be seen in the example DOS plots of figure 6 for cluster size $N_c = 8$, where after undergoing a MIT between $U = 1.30$ and $1.45$, a small peak appears again at the Fermi level around $U = 1.65$ which subsequently separates into two very small peaks as shown for $U = 1.95$ where they are well separated. These small peaks are contributions to the DOS at the very edges of the bands from cluster configurations. We therefore believe that the gap seen at the Fermi level for $U = 1.45$ is an artificial gap and the final MIT just after $U = 1.65$ is the true one because only the smaller critical $U$ values can result from overlap of bands with the spurious gaps. This is also in agreement with a systematic enhancement of the metallic phase with increasing cluster size.

We point out that a reformulation of the nonlocal CPA has been derived which systematically corrects the self-energy step function approximation by producing a smoother function in $k$-space for any given cluster size [47]. This is achieved by utilizing the fact that the choice of cluster momenta for a given cluster size is not unique. By using multiple sets of cluster momenta, a smoother and more realistic $k$-space self-energy can be self-consistently determined by reducing the jumps between the steps. It was found that implementing this method removes the spurious gaps in the DOS for disordered alloy calculations [47]. It should be stressed that this reformulation simply produces the correct self-energy curve and DOS for a given cluster size and hence correlation length; information beyond the correlation length can only be introduced by using a larger cluster size. In the present context of the IHM, we believe that implementing the method would remove the spurious extra
phase boundaries present in figures 3 and 5. This is left as future work.

In order to help identify the nature of the phase transitions in figure 5, the staggered average occupation number (charge density) $n_B - n_A$ as a function of $U$ for different values of $\Delta$ using the nonlocal CPA with $N_c = 8$ is shown in figure 7. A charge density wave is present throughout the whole parameter regime and gradually softens with increasing $U$. We find that the difference between the nonlocal CPA and the CPA results are negligible, of order $5 \times 10^{-3}$, and the phase transitions are seen to be continuous in all cases. Our results are in excellent agreement with the CPA results of Hoang [24], which in turn are in good agreement with those obtained by single-site DMFT [28]. Particularly noticeable for all three $\Delta$ values in our results are the change in slopes just after $U = 2\Delta$. These indicate the band insulator to metal transition which is dominated by the charge density wave. On the other hand, the MIT is governed by the Hubbard $U$ and so the curves smoothly change with no inflection point at the MIT since they have no relation to the charge density wave.

Since cellular-DMFT calculations for the IHM in 2D indicated an intermediate bond-ordered phase [29], it should be checked that the intermediate phase found in our calculations is indeed metallic. A particularly advantageous feature of the nonlocal CPA in the context of conventional alloy calculations is that the medium preserves the translational invariance of the underlying lattice. Unfortunately, in the present context this rules out the possibility of a long-range bond-ordered phase by construction. Therefore in order to determine whether intersite spatial correlations are able to produce a long-range bond-ordered phase we also implemented the molecular CPA [53, 54] for the IHM which naturally describes long-range order up to the length scale of a ‘molecule’ or cluster of sites. Indeed, cellular-DMFT is the analog of the molecular CPA for correlated electron systems. The implementation of the molecular CPA for the IHM is described in the appendix. The bonds are given by the matrix elements $\{\bar{G}^{IJ}\}$ for nearest-neighbour sites $I, J$ within the cluster. We performed extensive calculations in 1D and found no evidence of a long-range bond-ordered phase within the AA approach.

Finally, the present calculations could be improved by using some of the modifications described in section 2.1 such as the interpolating alloy analogy (IAA) [36]. For the paramagnetic solution at half-filling, the IAA reduces to the conventional AA in the strong-coupling regime so it would only be necessary to correct the non-Fermi liquid behaviour in the weak-coupling regime using the Edwards–Hertz approximation. A further improvement would be to include Hubbard’s resonance broadening correction and hence generalise the full Hubbard III approximation.
4. Conclusions

We have obtained the paramagnetic solution of the nonlocal CPA for the two-dimensional IHM at half-filling and zero temperature. We found that inter-site (nonlocal) spatial correlations enhance the intermediate metallic phase throughout the range of parameters, although the difference decreases with increasing $\Delta$. While the critical value $U_{c2}$ for the transition from band insulator to metal is very similar to that of the CPA, a peak is developed in the DOS in the metallic phase as $U$ is increased which delays the transition from metal to Mott insulator, significantly raising the critical value $U_{c2}$ for the MIT. To confirm the metallic nature of the intermediate phase, we also implemented the molecular CPA for the IHM and found no evidence of an intermediate bond-ordered phase. These findings further support the view [23–25] that for the paramagnetic solution of the IHM the presence of an intermediate metallic phase between the band insulator and Mott insulating phases is a real feature of the system.

Nevertheless, the reason that the intermediate metallic state is stabilized in the nonlocal CPA with respect to its local version when applied to strongly correlated systems from an alloy analogy point of view may be due to the fact that such theories of disorder fail to capture the localized state and hence always favour a metallic solution (see, for example, Ekuma et al [55] and references therein). Indeed, for $\Delta = 0$ we would have expected the inclusion of intersite spatial correlations to actually lower the critical value $U_{c2}$ for the MIT in accordance with calculations for the paramagnetic solution of the conventional 2D Hubbard model using the variational cluster approximation [56], dynamical cluster approximation (DCA) [57] and cellular-DMFT [58, 59], the latter two being the dynamical analogs of the nonlocal CPA and molecular CPA, respectively, for interacting electron systems. In the case of the DCA, it was found that in the weak-coupling region a pseudo-gap at the Fermi level was systematically developed with increasing cluster size and its formation was attributed to the increasing length of intersite anti-ferromagnetic correlations [57, 60]. Eventually, a true gap opens at zero temperature as the correlation length goes to infinity [61]. In the case of cellular-DMFT, it was reported that the MIT appears earlier than that obtained from DMFT due to the inclusion of short-range anti-ferromagnetic correlations [58].

In conclusion, since our calculations are within the alloy analogy approach (i.e. purely static), we have been able to isolate the effects of nonlocal spatial correlations from the dynamics and found the opposite behaviour to that obtained with dynamical methods when nonlocal spatial correlations are included, i.e. systematic enhancement of the metallic phase as opposed to systematic destruction. This important finding indicates that there must be some interplay between nonlocal spatial correlations and dynamical correlations which account for these opposing trends. This is a very interesting issue to investigate as future work.

Acknowledgments

It is an honour to produce this work in tribute to the late Professor Balazs Gyorffy. This work is supported by National Natural Science Foundation of China (No. 11174219), Shanghai Pujiang Program (No. 11PJ1409900), Research Fund for the Doctoral Program of Higher Education of China (No. 20110072110044) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning as well as the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

Appendix. Molecular CPA for the ionic Hubbard model

After dividing the lattice into clusters or ‘molecules’, the Green’s function matrix in $k$-space for any given cluster can be expressed in the form

$$
\bar{G}(q, \omega) = \begin{pmatrix}
\bar{G}_{AA}(q) & \bar{G}_{AB}(q) \\
\bar{G}_{BA}(q) & \bar{G}_{BB}(q)
\end{pmatrix}
$$

$$
= \begin{pmatrix}
\omega - (\Sigma_{AA} + \bar{t}^') - \bar{t}'(q) & -(\Sigma_{AB} + \bar{t}^') - \bar{t}'(q) \\
-(\Sigma_{BA} + \bar{t}^') - \bar{t}'(q) & \omega - (\Sigma_{BB} + \bar{t}^') - \bar{t}'(q)
\end{pmatrix}^{-1}
$$

(A.1)

where for convenience the cluster sites have been labeled by grouping into sublattice components. Here $q$ has been used to denote the wavevector restricted to the BZ of the cluster superlattice and the intra-cluster hopping terms $\bar{t}$ and $\bar{t}'$ have been separated out from the definition of the cluster self-energy. The intra-sublattice hopping terms have been denoted by $\bar{t}$ since these are in general different from the inter-sublattice terms $\bar{t}'$. For nearest-neighbour hopping on the bipartite lattice, the intra-sublattice hopping terms are zero. The Green’s function matrix can be expressed in real space by carrying out the BZ integration

$$
\bar{G}(\omega) = \frac{1}{\Omega_BZ} \int_{\Omega_BZ} \bar{G}(q, \omega) \, dq
$$

(A.2)

where the integration is over the BZ of the cluster superlattice. Note that the inter-cluster hopping is expressed through the $q$-dependence of the hopping terms defined by

$$
\bar{t}(q) = \frac{N_{c}}{N} \sum_{C \neq C} t e^{-i(qR_{C} - R_{C}')}
$$

(A.3)

where the distance vector implies the distance between the centres of clusters $C$ and $C'$, and for the case of nearest-neighbour hopping the appropriate terms in the matrix must be set to zero.

One may now proceed as in the nonlocal CPA by following the cavity matrix construction procedure expressed by equation (23)–(25). Equation (A.2) and the analogue of equation (25) need to be solved self-consistently. It will be necessary to ensure that the resulting integrated DOS for each cluster site is consistent with the relevant site occupation number probability used in the cluster configurations of equation (25), thus adding an extra layer of self-consistency. In contrast to the nonlocal CPA, the real space diagonal terms of $\Sigma_{\alpha\beta}$ and $\bar{G}_{\alpha\beta}$ when $\alpha = \beta$ do not have to be equivalent, and
hence the DOS results will in general differ depending upon the position of the measured site even on the same sublattice.

For the cluster size $N_c = 2$ the above formulation reduces to a generalisation of the coupled-CPA method of section 2.2, albeit a different one to that of the nonlocal CPA due to the different boundary conditions imposed on the cluster.

References

[1] Hubbard J 1963 Proc. R. Soc. Lond. A 276 238
[2] Hubbard J 1964 Proc. R. Soc. Lond. A 276 401
[3] Velecky B, Kirkpatrick S and Ehrenreich H 1968 Phys. Rev. 175 747
[4] Soven P 1967 Phys. Rev. 156 809
[5] Györffy B L, Barbiell A, Staunton J B, Shelton W A and Stocks G M 1991 Physica B 172 35
[6] Mors P M, Iglesias J R and Anda E V 1985 Solid State Commun. 55 827
[7] Luo H-G, Jia C-L, Wang S-J and Zuo W 2002 Phys. Rev. B 65 075108
[8] Geipel G and Nolting W 1988 Phys. Rev. B 38 2608
[9] Nolting W and Borgiel W 1989 Phys. Rev. B 39 6962
[10] Jarrell M and Krishnamurthy H R 2001 Phys. Rev. B 63 125102
[11] Moradian R, Györffy B L and Annett J F 2002 Phys. Rev. Lett. 89 287002
[12] Rowlands D A, Staunton J B and Györffy B L 2003 Phys. Rev. B 67 115109
[13] Rowlands D A, Ernst A, Györffy B L and Staunton J B 2006 Phys. Rev. B 73 165122
[14] Rowlands D A 2009 Rep. Prog. Phys. 72 086501
[15] Hubbard J and Torrance J B 1981 Phys. Rev. Lett. 47 1750
[16] Egami T, Ishihara S and Tachiki M 1993 Science 261 1307
[17] Nagaosa N and Takimoto J 1986 J. Phys. Soc. Japan 55 2735
[18] Fabrizio M, Gogolin A O and Nersesyan A A 1999 Phys. Rev. Lett. 83 2014
[19] Wilkens T and Martin R M 2001 Phys. Rev. B 63 235108
[20] Batista C D and Aligia A A 2004 Phys. Rev. Lett. 92 246405
[21] Zhang Y-Z, Wu C-Q and Lin H-Q 2003 Phys. Rev. B 67 205109
[22] Mannama S R, Meden V, Noack R M and Schonhammer K 2004 Phys. Rev. B 70 155115
[23] Garg A, Krishnamurthy H R and Randeria M 2006 Phys. Rev. Lett. 97 046403
[24] Hoang A T 2010 J. Phys.: Condens. Matter 22 095602
[25] Craco L, Lombardo P, Hayn R, Japaridze G I and Müller-Hartmann E 2008 Phys. Rev. B 78 075121
[26] Paris N, Bouadim K, Hebert F, Batrouni G G and Scalettar R T 2007 Phys. Rev. Lett. 98 046403
[27] Bouadim K, Paris N, Hebert F H, Batrouni G G and Scalettar R T 2007 Phys. Rev. B 76 085112
[28] Byczuk K, Sekania M, Hofstetter W and Kampf A P 2009 Phys. Rev. B 79 121103(R)
[29] Kancharla S S and Dagotto E 2007 Phys. Rev. Lett. 98 016402
[30] Chen H-M, Zhao H, Lin H-Q and Wu C-Q 2010 New J. Phys. 12 093021
[31] Gebhard F 1997 The Mott Metal–Insulator Transition: Models and Methods (Springer Series in Solid-State Sciences) (Berlin: Springer)
[32] Brouers F, Ducastelle F, Treglia G, Brauwers M and Giner J 1976 J. Physique 37 C4–207
[33] Brouers F, Ducastelle F and Giner J 1976 J. Physique 37 1427
[34] Herrmann T and Nolting W 1996 Phys. Rev. B 53 10579
[35] Harris A B and Lange R V 1967 Phys. Rev. 157 295
[36] Potthoff M, Hermann T and Nolting W 1998 Eur. Phys. J. B 4 485
[37] Janis V and Vollhardt D 1993 Z. Phys. B 91 317
[38] Janis V, Masek J and Vollhardt D 1993 Z. Phys. B 91 325
[39] Edwards D M and Hertz J A 1990 Physica B 163 527
[40] Edwards D M 1993 J. Phys.: Condens. Matter 5 161
[41] Corrias M 1990 Mod. Phys. Lett. B 4 549
[42] Corrias M 1991 ‘The Hubbard Model—Recent Results’ (Series on Advances in Statistical Mechanics vol 7) ed M Rasetti (Singapore: World Scientific)
[43] Luo H-G and Wang S-J 2000 Phys. Rev. B 61 5158
[44] Gorski G and Mizia J 2009 Phys. Rev. B 79 064414
[45] Gorski G and Mizia J 2011 Phys. Rev. B 83 064410
[46] Gupta A K, Edwards D M and Hewson A C 1975 J. Phys. C: Solid State Phys. 8 3207
[47] Rowlands D A, Zhang X-G and Gonis A 2008 Phys. Rev. B 78 115119
[48] Marmodoro A and Staunton J B 2011 J. Phys.: Condens. Matter 23 104613
[49] Emery V J 1987 Phys. Rev. Lett. 58 2794
[50] Meng Z Y, Lang T C, Wessel S, Assaad F F and Muramatsu A 2010 Nature 464 847
[51] Zhang Y-Z, Tran M-T, Yushankhai V and Thalmeier P 2005 Eur. Phys. J. B 44 265
[52] Rowlands D A 2006 J. Phys.: Condens. Matter 18 3179
[53] Tsukada M 1972 J. Phys. Soc. Japan 32 1475
[54] Ducastelle F 1974 Solid State Phys. 7 1795
[55] Ekuma C E, Terletska H, Meng Z Y, Moreno J, Jarrell M, Mahmoudian S and Dobrosavljevic V 2013 arXiv:1306.5712
[56] Balzer M, Kyung B, Senechal D, Tremblay A-M S and Potthoff M 2009 Europhys. Lett. 85 17002
[57] Moukouri S and Jarrell M 2001 Phys. Rev. Lett. 87 167010
[58] Zhang Y-Z and Imada M 2007 Phys. Rev. B 76 045108
[59] Park H, Haule K and Kotliar G 2008 Phys. Rev. Lett. 101 186403
[60] Kyung B, Landry J S, Poulin D and Tremblay A-M S 2003 Phys. Rev. Lett. 90 097002
[61] Imada M, Fujimori A and Tokura Y 1998 Rev. Mod. Phys. 70 1039