Analytical and numerical treatment of the Mott–Hubbard insulator in infinite dimensions

M.P. Eastwood[1], F. Gebhard[2,a], E. Kalinowski[2], S. Nishimoto[2], and R.M. Noack[3]

[1] Department of Chemistry and Biochemistry, University of California, La Jolla, CA 92093-0371, USA
[2] Fachbereich Physik, Philipps-Universität Marburg, 35032 Marburg, Germany
[3] Institut für Theoretische Physik III, Universität Stuttgart, 70550 Stuttgart, Germany

Abstract. We calculate the density of states in the half-filled Hubbard model on a Bethe lattice with infinite connectivity. Based on our analytical results to second order in t/U, we propose a new 'Fixed-Energy Exact Diagonalization' scheme for the numerical study of the Dynamical Mean-Field Theory. Corroborated by results from the Random Dispersion Approximation, we find that the gap opens at $U_c = 4.43 \pm 0.05$. Moreover, the density of states near the gap increases algebraically as a function of frequency with an exponent $\alpha = 1/2$ in the insulating phase. We critically examine other analytical and numerical approaches and specify their merits and limitations when applied to the Mott–Hubbard insulator.

PACS. 71.10.Fd Lattice fermion models (Hubbard model, etc.) – 71.27.+a Strongly correlated electron systems; heavy fermions – 71.30.+h Metal-insulator transitions and other electronic transitions

1 Introduction

The Mott–Hubbard metal-insulator transition is a fascinating but very difficult problem in condensed matter theory [1–4]. One hopes that the basic mechanism can be understood with the help of conceptually simple models such as the Hubbard Hamiltonian. This model describes spin-1/2 electrons which move on a lattice (band-width $W = 4t$) and interact locally with strength U. For small interactions, $U \ll W$, the model describes a metal with a finite density of states at the Fermi energy. For large interactions, $U \gg W$, and when there is on average one electron per lattice site (half band-filling), the model describes an insulator because it takes an energy of the order of $\Delta(U \gg W) \approx U - \mathcal{O}(W) > 0$ to generate a charge excitation, i.e., the single-particle gap $\Delta(U \gg W)$ is finite, irrespective of any symmetry breaking [1,2].

These basic insights leave much room for speculations on the properties of the transition itself, for example, on the precise value of the critical interaction strength U_c at which the gap opens. Because the Mott–Hubbard transition is a zero-temperature quantum phase transition [2], it is difficult to access reliably using analytical or numerical techniques; only exact solutions for all interaction strengths can provide conclusive answers. Unfortunately, Hubbard-type models can be solved exactly only in one dimension [5,6], and several features in one dimension are not generic to three-dimensional systems.

$^{\rm a}$ e-mail: `florian.gebhard@physik.uni-marburg.de`

The phase diagram of a model in three dimensions can often be understood from the limit of infinite dimensions. For correlated lattice electrons [7,8], this has led to the formulation of effective single-impurity models whose dynamics must be solved self-consistently (Dynamical Mean-Field Theory, DMFT) [3,9,10]. An alternative approach, the Random Dispersion Approximation (RDA) [2,11], also becomes exact in the limit of infinite dimensions. However, an exact solution of the DMFT is not feasible because single-impurity models with a Hubbard interaction cannot be solved in general.

Using various kinds of approximations, two different scenarios for the Mott–Hubbard transition have been proposed:

Discontinuous Transition.
The gap jumps to a finite value when the density of states at the Fermi energy becomes zero at some critical interaction strength $U_{c,2}$; the gap is preformed above $U_{c,1} < U_{c,2}$, and the co-existing insulating phase is higher in energy than the metallic state.

Continuous Transition.
The gap opens continuously when the density of states at the Fermi energy becomes zero, $U_{c,1} = U_{c,2} \equiv U_c$.

Various approaches to the Mott–Hubbard transition for lattices with infinite coordination number yield results in favor of one or the other of the conflicting scenarios; for a review, see references [2,3]. References [11–18] give more recent treatments. Insightful physical arguments

have helped to sharpen the physical and mathematical implications of the scenario of a preformed gap transition [19–22] but have not been able to resolve the issue.

All analytical approaches [17,23–25] are necessarily approximate in nature, and numerical investigations [12–18,26–28] of the dynamical mean-field equations involve (i) discretization, (ii) numerical diagonalization, (iii) interpolation, (iv) iteration of the self-consistency cycle, and (v) extrapolation to the thermodynamic limit. The Random Dispersion Approximation [11] requires similar steps in its numerical implementation, apart from step (iv). Therefore, the region of applicability of the available numerical techniques is not a priori clear either.

In order to address this issue, two of us (E.K. and F.G.) have developed a strong-coupling perturbation theory at zero temperature [29] which provides a benchmark test for the available analytical and numerical techniques for the Mott–Hubbard insulator. We assess the quality of the Iterated Perturbation Theory (IPT) and the Local Moment Approach (LMA) as well as three numerical methods for the Hubbard model in infinite dimensions: the Random Dispersion Approximation and two Exact Diagonalization (ED) schemes for the Dynamical Mean-Field Theory. As we shall show, this comparison will provide insight as to the limitations of the various methods and will motivate an improved ED scheme, the 'Fixed-Energy Exact Diagonalization' (FE-ED). This scheme successfully reproduces our findings from the t/U expansion even down to the collapse of the Mott–Hubbard insulator.

In this work, after the introduction of some definitions in Section 2, we reanalyze our strong-coupling perturbation theory to second order in t/U for the Bethe lattice with infinite connectivity. In Section 3 we estimate that the Mott–Hubbard insulator is destroyed at $U_{\mathrm{c}}^{\mathrm{sc}}/t = 4.40 \pm 0.09$. In Section 4 we set up the Fixed-Energy Exact Diagonalization scheme for the Dynamical Mean-Field Theory. The FE-ED very well reproduces our analytical findings, and confirms our conjecture for the critical interaction strength, $U_{\mathrm{c}}^{\mathrm{FE-ED}}/t = 4.43 \pm 0.05$. In Section 5 we favorably compare our results with those from the Random Dispersion Approximation (RDA) which provides an independent check on our results from the strong-coupling expansion and the FE-ED. Therefore, we are confident that we accurately describe the Mott–Hubbard insulator for all interaction strengths.

In Section 6 we compare our results from Sections 3 and 4 with two analytical approximations for the Mott–Hubbard insulator, the Iterated Perturbation Theory (IPT) and the Local Moment Approach (LMA). Both of them provide a reasonable description down to $U = 6t$ (IPT) and $U = 5.5t$ (LMA), respectively, where the LMA is generally superior to IPT. However, both approximations underestimate the stability of the Mott–Hubbard insulator, $U_{\mathrm{c},1}^{\mathrm{IPT}} = 5.2t$ and $U_{\mathrm{c},1}^{\mathrm{LMA}} = 4.8t$.

In Section 7, we critically examine two numerical approaches. For finite systems, the Caffarel–Krauth exact-diagonalization scheme [14,18,26] does not provide unique solutions of the self-consistency equations. Therefore, neither the size of the gap nor the shape of the density of states can be extrapolated reliably. The exact diagonalization of the effective single-impurity Anderson model in 'two-chain geometry' [3,27,28] fails because the reconstruction of the density of states from its moments is a numerically too delicate inverse problem. Conclusions, Section 8, and appendices on the IPT and LMA algorithms in the limit of strong coupling close our work.

2 Definitions

In this section, we discuss the basic properties of lattice electrons in the limit of infinite dimensions, and we define the Hubbard Hamiltonian, the single-particle Green function, and the corresponding density of states.

2.1 Hamilton operator

We investigate spin-1/2 electrons on a lattice whose motion is described by the kinetic energy operator

$$\hat{T} = \sum_{\mathbf{i},\mathbf{j};\sigma} t_{\mathbf{i},\mathbf{j}} \hat{c}_{\mathbf{i},\sigma}^{+} \hat{c}_{\mathbf{j},\sigma}, \tag{1}$$

where $\hat{c}_{\mathbf{i},\sigma}^{+}$, $\hat{c}_{\mathbf{i},\sigma}$ are creation and annihilation operators for electrons with spin $\sigma =\uparrow,\downarrow$ on site $\mathbf{i}$. The matrix elements $t_{\mathbf{i},\mathbf{j}}$ are the electron transfer amplitudes between sites $\mathbf{i}$ and $\mathbf{j}$, and $t_{\mathbf{i},\mathbf{i}} = 0$. Since we are interested in the Mott insulating phase, we consider exclusively a half-filled band where the number of electrons N equals the number of lattice sites L.

For lattices with translational symmetry, $t_{\mathbf{i},\mathbf{j}} = t(\mathbf{i} - \mathbf{j})$ and the operator for the kinetic energy is diagonal in momentum space,

$$\hat{T} = \sum_{\mathbf{k};\sigma} \epsilon(\mathbf{k}) \hat{c}_{\mathbf{k},\sigma}^{+} \hat{c}_{\mathbf{k},\sigma}, \tag{2}$$

where

$$\epsilon(\mathbf{k}) = \frac{1}{L} \sum_{\mathbf{i},\mathbf{j}} t(\mathbf{i} - \mathbf{j}) e^{-i(\mathbf{i}-\mathbf{j})\mathbf{k}}. \tag{3}$$

The density of states for non-interacting electrons is then given by

$$\rho(\epsilon) = \frac{1}{L} \sum_{\mathbf{k}} \delta(\epsilon - \epsilon(\mathbf{k})). \tag{4}$$

The mth moment of the density of states is defined by

$$\overline{\epsilon^m} = \int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \epsilon^m \rho(\epsilon), \tag{5}$$

and $\overline{\epsilon} = t(\mathbf{0}) = 0$.

In the limit of infinite lattice dimensions and for translationally invariant systems without nesting, the Hubbard model is characterized by $\rho(\epsilon)$ alone, i.e., higher-order correlation functions in momentum space factorize [30], for

example,

$$\rho_{12}(\epsilon_1, \epsilon_2) \equiv \frac{1}{L} \sum_{\mathbf{k}} \delta(\epsilon_2 - \epsilon(\mathbf{k} + \mathbf{q}_2))\delta(\epsilon_1 - \epsilon(\mathbf{k} + \mathbf{q}_1))$$

$$= \delta_{\mathbf{q}_1, \mathbf{q}_2} \delta(\epsilon_1 - \epsilon_2)\rho(\epsilon_1) + (1 - \delta_{\mathbf{q}_1, \mathbf{q}_2})\rho(\epsilon_1)\rho(\epsilon_2). \quad (6)$$

This observation is the basis for the Random Dispersion Approximation (RDA) which becomes exact in infinite dimensions for paramagnetic systems, *i.e.*, when nesting is ignored [2,11]; see Section 5.

For our explicit calculations we shall later use the semi-circular density of states

$$\rho_0(\omega) = \frac{2}{\pi W} \sqrt{4 - \left(\frac{4\omega}{W}\right)^2}, \quad (|\omega| \leq W/2), \quad (7)$$

$$1 = \int_{-W/2}^{W/2} d\omega \rho_0(\omega), \quad (8)$$

where $W = 4t$ is the band-width. In the following, we take $t \equiv 1$ as our unit of energy. This density of states is realized for non-interacting tight-binding electrons on a Bethe lattice of connectivity $Z \to \infty$ [31]. Specifically, each site is connected to Z neighbors without generating closed loops, and the electron transfer is restricted to nearest-neighbors: $t_{\mathbf{i},\mathbf{j}} = -t/\sqrt{Z}$ when $\mathbf{i}$ and $\mathbf{j}$ are nearest neighbors and zero otherwise. The limit $Z \to \infty$ is implicitly understood henceforth. Later on we shall use either of the two equivalent viewpoints, RDA or tight-binding Bethe lattice, whichever is more convenient for our considerations.

The electrons are taken to interact only locally, and the Hubbard interaction reads

$$U\hat{D} = U \sum_{\mathbf{i}} \left(\hat{n}_{\mathbf{i},\uparrow} - \frac{1}{2}\right)\left(\hat{n}_{\mathbf{i},\downarrow} - \frac{1}{2}\right), \quad (9)$$

where $\hat{n}_{\mathbf{i},\sigma} = \hat{c}_{\mathbf{i},\sigma}^+ \hat{c}_{\mathbf{i},\sigma}$ is the local density operator at site $\mathbf{i}$ for spin σ. This leads us to the Hubbard model [32],

$$\hat{H} = \hat{T} + U\hat{D}. \quad (10)$$

The Hamiltonian explicitly exhibits particle-hole symmetry, *i.e.*, $\hat{H}$ is invariant under the particle-hole transformation

$$\hat{c}_{\mathbf{i},\sigma}^+ \mapsto (-1)^{|\mathbf{i}|}\hat{c}_{\mathbf{i},\sigma}; \quad \hat{c}_{\mathbf{i},\sigma} \mapsto (-1)^{|\mathbf{i}|}\hat{c}_{\mathbf{i},\sigma}^+, \quad (11)$$

where $|\mathbf{i}|$ counts the number of nearest-neighbor steps from the origin of the Bethe lattice to site $\mathbf{i}$. The chemical potential $\mu = 0$ then guarantees a half-filled band for any temperature [2].

For later use we also define the operators for the local electron density, $\hat{n}_{\mathbf{j}} = \hat{n}_{\mathbf{j},\uparrow} + \hat{n}_{\mathbf{j},\downarrow}$. The operators $\hat{S}_{\mathbf{j}}^x = (\hat{c}_{\mathbf{j},\uparrow}^+ \hat{c}_{\mathbf{j},\downarrow} + \hat{c}_{\mathbf{j},\downarrow}^+ \hat{c}_{\mathbf{j},\uparrow})/2$, $\hat{S}_{\mathbf{j}}^y = (\hat{c}_{\mathbf{j},\uparrow}^+ \hat{c}_{\mathbf{j},\downarrow} - \hat{c}_{\mathbf{j},\downarrow}^+ \hat{c}_{\mathbf{j},\uparrow})/(2i)$, and $\hat{S}_{\mathbf{j}}^z = (\hat{n}_{\mathbf{j},\uparrow} - \hat{n}_{\mathbf{j},\downarrow})/2$ are the three components of the spin-1/2 vector operator $\hat{\mathbf{S}}_{\mathbf{j}}$. The operator for the total spin $\hat{\mathbf{S}} = \sum_{\mathbf{j}} \hat{\mathbf{S}}_{\mathbf{j}}$ commutes with the Hamiltonian.

2.2 Green functions

The time-dependent local single-particle Green function at zero temperature is given by [33]

$$G(t) = -i\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{\mathcal{T}} \left[\hat{c}_{\mathbf{i},\sigma}(t)\hat{c}_{\mathbf{i},\sigma}^+ \right] \right\rangle. \quad (12)$$

Here $\hat{\mathcal{T}}$ is the time-ordering operator and $\langle \ldots \rangle$ implies the average over all ground states with energy E_0, and (taking $\hbar \equiv 1$ henceforth)

$$\hat{c}_{\mathbf{i},\sigma}(t) = \exp(i\hat{H}t)\hat{c}_{\mathbf{i},\sigma}\exp(-i\hat{H}t) \quad (13)$$

is the annihilation operator in the Heisenberg picture.

In the insulating phase we can readily identify the contributions from the lower (LHB) and upper (UHB) Hubbard bands to the Fourier transform of the local Green function ($\eta = 0^+$),

$$G(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t} G(t) = G_{\text{LHB}}(\omega) + G_{\text{UHB}}(\omega),$$

$$G_{\text{LHB}}(\omega) = \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^+ \left[\omega + (\hat{H} - E_0) - i\eta \right]^{-1} \hat{c}_{\mathbf{i},\sigma} \right\rangle,$$

$$G_{\text{UHB}}(\omega) = -G_{\text{LHB}}(-\omega). \quad (14)$$

The last equality follows from the particle-hole symmetry, *cf.* (11). Therefore, it is sufficient to evaluate the local Green function for the lower Hubbard band which describes the dynamics of a hole inserted into the system.

The density of states for the lower Hubbard band can be obtained from the imaginary part of the Green function (14) for real arguments *via* [33]

$$D_{\text{LHB}}(\omega) = \frac{1}{\pi}\Im G_{\text{LHB}}(\omega),$$

$$= \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^+ \delta\left(\omega + \hat{H} - E_0\right) \hat{c}_{\mathbf{i},\sigma} \right\rangle, \quad (15)$$

with $\mu_{\text{LHB}}^- \leq \omega \leq \mu_{\text{LHB}}^+ < 0$. Particle-hole symmetry results in $D_{\text{UHB}}(\omega) = D_{\text{LHB}}(-\omega)$ so that the single-particle gap in the Mott–Hubbard insulator is given by

$$\Delta(U) = 2|\mu_{\text{LHB}}^+(U)| > 0. \quad (16)$$

We define the (shifted) moments $M_n(U)$ of the density of states in the lower Hubbard band *via*

$$M_n(U) = \int_{\mu_{\text{LHB}}^-}^{\mu_{\text{LHB}}^+} d\omega \left(\omega + \frac{U}{2}\right)^n D_{\text{LHB}}(\omega). \quad (17)$$

In particular, from (15) we find that [33]

$$M_0(U) = 1, \quad (18)$$

$$M_1(U) = \frac{1}{L}\left(E_0 + U\frac{\partial E_0}{\partial U}\right) + \frac{U}{2} \quad (19)$$

are two useful sum-rules which we shall employ later.

3 Strong-coupling expansion

In this section we first summarize the Kato–Takahashi strong-coupling perturbation theory. Next, we prove that the entropy density of the Mott–Hubbard insulator is finite to all orders in perturbation theory. We then derive the density of states to second order in $1/U$; for further details, see reference [29]. Finally, we present explicit results for physical quantities including the gap and an estimate of the critical interaction strength.

3.1 Kato–Takahashi perturbation theory

Based on Kato's degenerate perturbation theory [34], Takahashi developed a perturbation expansion in $1/U$ for the Hubbard model at zero temperature [35]. For large interaction strength U, the set of ground states $|\psi_n\rangle$ of $\hat{H}$ in (10) can be obtained from states $|\phi_n\rangle$ without double occupancy,

$$|\psi_n\rangle = \hat{\Gamma}|\phi_n\rangle, \quad \hat{P}_0|\phi_n\rangle = |\phi_n\rangle, \tag{20}$$

where $\hat{P}_j$ projects onto the subspace with j double occupancies and $\hat{\Gamma}$ is an operator to be determined.

Equation (20) is readily interpreted. In the large-coupling limit, $\hat{T}$ in (10) is considered to be a perturbation to $U\hat{D}$. Addressing the ground states we therefore start from eigenstates with zero double occupancies $|\phi_n\rangle$ into which the operator $\hat{\Gamma}$ successively introduces double occupancies and holes to generate the ground states $|\psi_n\rangle$ of $\hat{H}$. The operator $\hat{\Gamma}$ reduces all operators to the subspace with zero double occupancies. In particular, $\hat{\Gamma}^+\hat{\Gamma} = \hat{P}_0$ so that overlap matrix elements obey [35] $\langle\psi_m|\psi_n\rangle = \langle\phi_m|\phi_n\rangle$.

The Schrödinger equation for the ground states

$$\hat{H}\hat{\Gamma}|\phi_n\rangle = E_0\hat{\Gamma}|\phi_n\rangle \tag{21}$$

leads to

$$\hat{h}|\phi_n\rangle = E_0|\phi_n\rangle, \quad \hat{h} = \hat{\Gamma}^+\hat{H}\hat{\Gamma}, \tag{22}$$

i.e., the eigenvalue does not change under the transformation. The creation and annihilation operators are transformed accordingly,

$$\hat{c}^+_{\mathbf{i},\sigma} \mapsto \widetilde{c}^+_{\mathbf{i},\sigma} = \hat{\Gamma}^+\hat{c}^+_{\mathbf{i},\sigma}\hat{\Gamma}, \quad \hat{c}_{\mathbf{i},\sigma} \mapsto \widetilde{c}_{\mathbf{i},\sigma} = \hat{\Gamma}^+\hat{c}_{\mathbf{i},\sigma}\hat{\Gamma}. \tag{23}$$

The derivation of the explicit expression for $\hat{\Gamma}$ can be found in references [34,35], where it is shown that

$$\hat{\Gamma} = \hat{P}\hat{P}_0\left(\hat{P}_0\hat{P}\hat{P}_0\right)^{-1/2}. \tag{24}$$

Here

$$\hat{P} = \hat{P}_0 - \sum_{m=1}^{\infty}\sum_{\substack{r_1+\ldots+r_{m+1}=m \\ r_i \geq 0}}\hat{S}^{r_1}\hat{T}\hat{S}^{r_2}\cdots\hat{T}\hat{S}^{r_{m+1}},$$

$$\hat{S}^0 = -\hat{P}^0, \tag{25}$$

$$\hat{S}^r = \frac{(-1)^r}{U^r}\sum_{j\neq 0}\frac{\hat{P}_j}{j^r} \equiv \frac{(-1)^r}{U^r}\widetilde{S}^r.$$

The mth order term in $\hat{P}$ contains all possible electron transfers generated by m applications of the perturbation $\hat{T}$; its contribution is proportional to $(1/U)^m$. The square-root factor in (24) guarantees the size-consistency of the expansion, *i.e.*, it eliminates the 'disconnected' diagrams in a diagrammatic formulation of the theory [35]. The square root of an operator is understood in terms of its series expansion, *i.e.*,

$$\left(\hat{P}_0\hat{P}\hat{P}_0\right)^{-1/2} \equiv \hat{P}_0 + \sum_{m=1}^{\infty}\frac{(2m-1)!!}{(2m)!!}\left[\hat{P}_0(\hat{P}_0-\hat{P})\hat{P}_0\right]^m. \tag{26}$$

The Green function for the lower Hubbard band becomes

$$G_{\text{LHB}}(\omega) = \frac{1}{L}\sum_{\mathbf{i},\sigma}\left\langle \widetilde{c}^+_{\mathbf{i},\sigma}\left[\omega + (\hat{h}-E_0) - \mathrm{i}\eta\right]^{-1}\widetilde{c}_{\mathbf{i},\sigma}\right\rangle, \tag{27}$$

where $\langle\ldots\rangle$ now implies the average over all states $|\phi_n\rangle$ with energy E_0.

A straightforward expansion in $1/U$,

$$\widetilde{c}_{\mathbf{i},\sigma} = \sum_{m=0}^{\infty}(-U)^{-m}\hat{P}_0\widetilde{c}^{(m)}_{\mathbf{i},\sigma}\hat{P}_0,$$

$$\hat{h} = \frac{U}{4}(L-2\hat{N}) + \sum_{m=0}^{\infty}U^{-m}\hat{h}_m, \tag{28}$$

gives

$$\hat{h}_0 = \hat{P}_0\hat{T}\hat{P}_0, \tag{29}$$

$$\hat{h}_1 = -\hat{P}_0\hat{T}\widetilde{S}\hat{T}\hat{P}_0, \tag{30}$$

$$\hat{h}_2 = \hat{P}_0\hat{T}\widetilde{S}\hat{T}\widetilde{S}\hat{T}\hat{P}_0 + \hat{h}_1\hat{h}_0 \tag{31}$$

for the Hamiltonian and

$$\widetilde{c}^{(0)}_{\mathbf{i},\sigma} = \hat{c}_{\mathbf{i},\sigma}, \tag{32}$$

$$\widetilde{c}^{(1)}_{\mathbf{i},\sigma} = \hat{c}_{\mathbf{i},\sigma}\widetilde{S}\hat{T}, \tag{33}$$

$$\widetilde{c}^{(2)}_{\mathbf{i},\sigma} = \hat{c}_{\mathbf{i},\sigma}\widetilde{S}\hat{T}\widetilde{S}\hat{T} + \hat{T}\widetilde{S}\hat{c}_{\mathbf{i},\sigma}\widetilde{S}\hat{T} + \frac{1}{2}\hat{c}_{\mathbf{i},\sigma}\hat{h}_1 + \frac{1}{2}\hat{h}_1\hat{c}_{\mathbf{i},\sigma} \tag{34}$$

for the annihilation operators. In the derivation of (29)–(34) we have used the fact that $\widetilde{c}_{\mathbf{i},\sigma}$ in (27) acts on states $|\phi_n\rangle$ with no holes and no double occupancies; see also Section 3.2. We have also used the fact that in the Green function for the lower Hubbard band of a half-filled Bethe lattice we can express the Hamiltonian $\hat{h}$ in terms of a polynomial in the bare hole-hopping operator $\hat{h}_0$ [29].

3.2 Ground states

To leading order in $1/U$, the states $|\phi_n\rangle$ are eigenstates of the Hamiltonian $\hat{h}_1$. Because there are no doubly occupied sites in $|\phi_n\rangle$ and because the system is half filled, all sites are singly occupied ('spin-only states'). Therefore, $\hat{h}_1$

can be expressed in terms of spin operators. For general hopping amplitudes $t(\mathbf{i} - \mathbf{j})$ ($t(\mathbf{0}) = 0$) on a regular lattice, one finds [35,36]

$$\hat{h}_1 = \sum_{\mathbf{i,j}} 2|t(\mathbf{i} - \mathbf{j})|^2 \left(\hat{\mathbf{S}}_{\mathbf{i}} \cdot \hat{\mathbf{S}}_{\mathbf{j}} - \frac{1}{4} \right). \qquad (35)$$

In momentum space this becomes

$$\hat{h}_1 = \sum_{\mathbf{q}} J(\mathbf{q}) \left[\hat{\mathbf{S}}_{\mathbf{q}} \cdot \hat{\mathbf{S}}_{-\mathbf{q}} - \delta_{\mathbf{q,0}} \frac{L}{4} \right],$$

$$\frac{1}{2} J(\mathbf{q}) = \frac{1}{L} \sum_{\mathbf{k}} \epsilon(\mathbf{k}) \epsilon(\mathbf{k} + \mathbf{q})$$

$$= \int_{-\infty}^{\infty} d\epsilon_1 \, \epsilon_1 \int_{-\infty}^{\infty} d\epsilon_2 \, \epsilon_2 \, \rho_{\mathbf{q}}(\epsilon_1, \epsilon_2). \qquad (36)$$

Here $\hat{\mathbf{S}}_{\mathbf{q}} = \sqrt{1/L} \sum_{\mathbf{j}} \exp(-i\mathbf{qj}) \hat{\mathbf{S}}_{\mathbf{j}}$ is the Heisenberg spin operator in momentum space.

Since the joint density of states (6) factorizes, the Heisenberg model (35) in the absence of any nesting reduces to

$$\hat{h}_1 = J(\mathbf{0}) \left[\frac{\hat{\mathbf{S}}^2}{L} - \frac{L}{4} \right] \qquad (37)$$

in infinite dimensions. This shows that all global singlets are ground states with energy $E_0^{(1)} = -J(\mathbf{0})L/(4U) = -\overline{\epsilon^2}L/(2U)$, see (5). For the Bethe lattice with the semi-elliptic density of states (7) we have $\overline{\epsilon^2} = 1$ so that $E_0^{(1)} = -L/(2U)$.

To first order in $1/U$, the degeneracy of the spin-only states is not lifted because, in the thermodynamic limit, all states with $S^z = 0$ are also global singlets, $S = 0$. More precisely, the ground-state entropy density is given by $s = \ln(2) + \mathcal{O}(\ln(L)/L)$ [2].

This degeneracy is not lifted to all orders in perturbation theory. We illustrate the argument in the next non-vanishing order of the perturbation theory. As was shown by Takahashi [35] in the presence of particle-hole symmetry, the Hamilton operator for the ground state to second order in $1/U$ vanishes, and the third-order contribution can be cast into the form

$$\hat{h}_3 = \sum_{\mathbf{a,b,c}} 2|t(\mathbf{a} - \mathbf{b})|^2 |t(\mathbf{b} - \mathbf{c})|^2 \left(\hat{\mathbf{S}}_{\mathbf{a}} \cdot \hat{\mathbf{S}}_{\mathbf{c}} - \frac{1}{4} \right)$$

$$+ \sum_{\mathbf{a,b,c,d; b \neq d, a \neq c}} t(\mathbf{a} - \mathbf{b}) t(\mathbf{b} - \mathbf{c}) t(\mathbf{c} - \mathbf{d}) t(\mathbf{d} - \mathbf{a})$$

$$\left[5 \left(\hat{\mathbf{S}}_{\mathbf{b}} \cdot \hat{\mathbf{S}}_{\mathbf{c}} \right) \left(\hat{\mathbf{S}}_{\mathbf{a}} \cdot \hat{\mathbf{S}}_{\mathbf{d}} \right) + 5 \left(\hat{\mathbf{S}}_{\mathbf{a}} \cdot \hat{\mathbf{S}}_{\mathbf{b}} \right) \left(\hat{\mathbf{S}}_{\mathbf{c}} \cdot \hat{\mathbf{S}}_{\mathbf{d}} \right) \right.$$

$$-5 \left(\hat{\mathbf{S}}_{\mathbf{a}} \cdot \hat{\mathbf{S}}_{\mathbf{c}} \right) \left(\hat{\mathbf{S}}_{\mathbf{b}} \cdot \hat{\mathbf{S}}_{\mathbf{d}} \right) - \frac{1}{4} \left(\hat{\mathbf{S}}_{\mathbf{a}} \cdot \hat{\mathbf{S}}_{\mathbf{b}} \right)$$

$$- \frac{1}{4} \left(\hat{\mathbf{S}}_{\mathbf{b}} \cdot \hat{\mathbf{S}}_{\mathbf{c}} \right) - \frac{1}{4} \left(\hat{\mathbf{S}}_{\mathbf{c}} \cdot \hat{\mathbf{S}}_{\mathbf{d}} \right) - \frac{1}{4} \left(\hat{\mathbf{S}}_{\mathbf{d}} \cdot \hat{\mathbf{S}}_{\mathbf{a}} \right)$$

$$\left. - \frac{1}{4} \left(\hat{\mathbf{S}}_{\mathbf{a}} \cdot \hat{\mathbf{S}}_{\mathbf{c}} \right) - \frac{1}{4} \left(\hat{\mathbf{S}}_{\mathbf{b}} \cdot \hat{\mathbf{S}}_{\mathbf{d}} \right) + \frac{1}{16} \right]. \qquad (38)$$

We have omitted terms where two sites are connected by four hopping processes as they give a vanishing contribution in the limit $Z \to \infty$. By Fourier transforming, we can ensure that the factorization of the correlation functions as in (6) hold. Therefore, in infinite dimensions in the absence of perfect nesting equation (38) reduces to

$$\hat{h}_3 = 2[\overline{\epsilon^2}]^2 \left[\frac{\hat{\mathbf{S}}^2}{L} - \frac{L}{4} \right]$$

$$+ \left(\overline{\epsilon^4} - 2[\overline{\epsilon^2}]^2 \right) \left[\frac{5}{L} \left(\frac{\hat{\mathbf{S}}^2}{L} \right)^2 - \frac{3}{2} \frac{\hat{\mathbf{S}}^2}{L} + \frac{L}{16} \right]. \qquad (39)$$

The second term describes the (usually ferromagnetic) ring exchange whereas the first term favors ground states with spin zero.

Like $\hat{h}_1$ in (37), $\hat{h}_3$ in (39) is solely a function of the operator for the total spin. This argument is readily generalized to higher orders in the perturbation expansion, *i.e.*, $\hat{h}$ is a function of $\hat{\mathbf{S}}^2$ where the leading term favors total spin zero. Therefore, as long as the perturbation series in $1/U$ converges, the degeneracy of the ground state is not lifted. As a consequence of its huge degeneracy, $s = \ln(2)$, each lattice site is equally likely to be occupied by an electron with spin ↑ or ↓, irrespective of the spin on any other lattice site. This is a characteristic feature of the Mott–Hubbard insulator in the limit of high dimensions.

On the Bethe lattice there are no closed loops, and the ring exchange is absent, $\overline{\epsilon^4} = 2\overline{\epsilon^2} = 2$. Thus, the ground-state energy to third order becomes $E_0^{(3)} = -L/(2U^3)$.

3.3 Green function to second order

As shown in reference [29], to second order in $1/U$ the Green function for the lower Hubbard band can be expressed in the form

$$G_{\text{LHB}}(\omega) = \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^+ s(\hat{h}_0, U) \left[z + g(\hat{h}_0, U) \right]^{-1} \hat{c}_{\mathbf{i},\sigma} \right\rangle, \qquad (40)$$

where $z = \omega + U/2 - i\eta$, and $s(\epsilon, U)$ and $g(\epsilon, U)$ are polynomials in ϵ and $1/U$. They are called the shape-correction factor and the gap-renormalization factor, respectively. Up to second order in $1/U$, the general structure of $\hat{h}$ in (29–31) and of $\widetilde{c}_{\mathbf{i},\sigma}$ in (32–34) implies

$$s(\epsilon, U) = 1 + \frac{s_{1,1}\epsilon}{U} + \frac{s_{2,2}\epsilon^2 + s_{2,0}}{U^2},$$

$$g(\epsilon, U) = \epsilon + \frac{g_{1,2}\epsilon^2 + g_{1,0}}{U} + \frac{g_{2,3}\epsilon^3 + g_{2,1}\epsilon}{U^2} \qquad (41)$$

because each electron transfer operator $\hat{T}$ generates one bare hole-hopping operator $\hat{h}_0$ at most.

3.3.1 Gap-renormalization factor

We determine the coefficients $g_{i,j}$ first. To this end, we set $s(\epsilon, U) \equiv 1$ in (40) and expand this equation in $1/U$. We

equate the result of this expansion with the corresponding expression from (27). To first order in $1/U$ we find

$$\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \left[z + \hat{h}_0 \right]^{-2} \left[g_{1,2}(\hat{h}_0)^2 + g_{1,0} \right] \hat{c}_{\mathbf{i},\sigma} \right\rangle =$$
$$\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \left[z + \hat{h}_0 \right]^{-1} \left(\hat{h}_1 + \frac{L}{2U} \right) \left[z + \hat{h}_0 \right]^{-1} \hat{c}_{\mathbf{i},\sigma} \right\rangle. \tag{42}$$

As we need to fix two coefficients only, we further expand this expression in $1/z$. Keeping the first two non-trivial orders we arrive at

$$g_{1,2} + g_{1,0} = \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \left[g_{1,2}(\hat{h}_0)^2 + g_{1,0} \right] \hat{c}_{\mathbf{i},\sigma} \right\rangle$$
$$= \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \left(\hat{h}_1 + \frac{L}{2} \right) \hat{c}_{\mathbf{i},\sigma} \right\rangle, \tag{43}$$

and

$$2g_{1,2} + g_{1,0} = \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \left[g_{1,2}(\hat{h}_0)^4 + g_{1,0}(\hat{h}_0)^2 \right] \hat{c}_{\mathbf{i},\sigma} \right\rangle$$
$$= \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \left(\hat{h}_1 + \frac{L}{2} \right) (\hat{h}_0)^2 \hat{c}_{\mathbf{i},\sigma} \right\rangle. \tag{44}$$

It is not difficult to calculate the expectation values in these equations from the definition (30). Recall that $\langle \ldots \rangle$ implies the average over all spin-only states. In equation (43) we note that an intermediate nearest-neighbor pair of hole and double occupancy can be created with probability $1/2$ by $\hat{h}_1$ everywhere but at site $\mathbf{i}$. In the sum over the whole lattice, two contributions are therefore missing. Altogether, $\hat{h}_1$ generates the contribution $-(L-2)/2$ in (43), which, together with the first-order contribution $L/2$ from the ground-state energy, results in $g_{1,2}+g_{1,0} = 1$. In (44) the three-site term in $\hat{h}_1$ contributes with probability $1/2$. Thus, $2g_{1,2} + g_{1,0} = 1 - 1/2 = 1/2$. Therefore, $g_{1,2} = -1/2$ and $g_{1,0} = 3/2$, and $\hat{h}_1^{\text{eff}} \equiv g_{1,2}\hat{h}_0^2 + g_{1,0} = -(\hat{h}_0^2 - 3)/2$.

In second order, we split the contributions arising from $\hat{h}_2$ and those from the correction terms. Let $\hat{C} = 1/(z+\hat{h}_0)$ and $\hat{h}_1' = \hat{h}_1 - E_0^{(1)} - \hat{h}_1^{\text{eff}}$, $\hat{h}_2^{\text{eff}} \equiv g_{2,3}\hat{h}_0^3 + g_{2,1}\hat{h}_0 = \hat{h}_{2a}^{\text{eff}} + \hat{h}_{2b}^{\text{eff}}$, then

$$\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{h}_{2a}^{\text{eff}} \hat{C}^2 \hat{c}_{\mathbf{i},\sigma} \right\rangle = \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{C} \hat{h}_2 \hat{C} \hat{c}_{\mathbf{i},\sigma} \right\rangle, \tag{45}$$

$$\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{h}_{2b}^{\text{eff}} \hat{C}^2 \hat{c}_{\mathbf{i},\sigma} \right\rangle = -\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{C} \hat{h}_1' \hat{C} \hat{h}_1' \hat{C} \hat{c}_{\mathbf{i},\sigma} \right\rangle. \tag{46}$$

Expanding in $1/z$, the corresponding equations for $\hat{h}_{2a}^{\text{eff}}$ read

$$2g_{2,3;a} + g_{2,1;a} = \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{h}_2 \hat{h}_0 \hat{c}_{\mathbf{i},\sigma} \right\rangle, \tag{47}$$

$$g_{2,3;a} = \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{h}_2 [(\hat{h}_0)^3 - 2\hat{h}_0] \hat{c}_{\mathbf{i},\sigma} \right\rangle. \tag{48}$$

Note that in (48) the operator $(\hat{h}_0)^3 - 2\hat{h}_0$ transfers the hole at $\mathbf{i}$ into its third neighbor shell. After some calculations we find $2g_{2,3;a} + g_{2,1;a} = -3/2$ and $g_{2,3;a} = 3/4$ so that $g_{2,1;a} = -3$.

The correction terms were correctly taken into account for the Falicov–Kimball model in reference [29], but were erroneously omitted for the Hubbard model. The operator $\hat{h}_1'$ describes the motion of a hole by two sites whereby a spin-flip is induced. In (46) the hole hops from site $\mathbf{i}$ to some site $\mathbf{j}$ where the hole induces a spin-flip. This needs to be healed after the hole has made an excursion into the lattice so that

$$\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{h}_{2b}^{\text{eff}} \hat{C}^2 \hat{c}_{\mathbf{i},\sigma} \right\rangle = -G_0(z) \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{C} (\hat{h}_1')^2 \hat{C} \hat{c}_{\mathbf{i},\sigma} \right\rangle$$
$$= -\frac{3}{4} G_0(z) \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{C}^2 \hat{c}_{\mathbf{i},\sigma} \right\rangle$$
$$= \frac{3}{8} \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \hat{h}_0 \hat{C}^2 \hat{c}_{\mathbf{i},\sigma} \right\rangle. \tag{49}$$

This leads to $\hat{h}_{2b}^{\text{eff}} = (3/8)\hat{h}_0$, i.e., $g_{2,3;b} = 0$, $g_{2,1;b} = 3/8$.

Altogether, $g_{2,3} = g_{2,3;a} + g_{2,3;b} = 3/4$ and $g_{2,1} = g_{2,1;a} + g_{2,1;b} = -21/8$, and $\hat{h}_2^{\text{eff}} = 3\hat{h}_0[2(\hat{h}_0)^2 - 7]/8$. The gap-renormalization factor becomes

$$g(\epsilon, U) = \epsilon - \frac{\epsilon^2 - 3}{2U} + \frac{3\epsilon (2\epsilon^2 - 7)}{8U^2} \tag{50}$$

through second order in $1/U$.

3.3.2 Shape-correction factor

In order to determine the coefficients $s_{i,j}$ we first employ the sum rules (18) and (19). We use $g(\epsilon, U)$ from the previous subsection and find through terms of order $1/U^3$,

$$1 = \int d\epsilon \rho_0(\epsilon) s(\epsilon, U), \tag{51}$$

$$\mathcal{O}\left(\frac{1}{U^3}\right) = -\int d\epsilon \rho_0(\epsilon) s(\epsilon, U) g(\epsilon, U), \tag{52}$$

because $E_0/L = -U/4 - 1/(2U) - 1/(2U^3) + \mathcal{O}(U^{-5})$. This immediately gives $s_{2,2} = -s_{2,0}$ from (51). From (52) we obtain $s_{1,1} = -1$.

For the other coefficients we expand (40) and (27) in $1/U$ and in $1/z$. The shape-correction coefficients to second order follow from the equation

$$s_{2,2} = 2\frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \hat{c}_{\mathbf{i},\sigma}^{+} \left[(\hat{h}_0)^2 - 1 \right] \widetilde{c}_{\mathbf{i},\sigma}^{(2)} \right\rangle$$
$$+ \frac{1}{L} \sum_{\mathbf{i},\sigma} \left\langle \left(\widetilde{c}_{\mathbf{i},\sigma}^{(1)} \right)^{+} \left[(\hat{h}_0)^2 - 1 \right] \widetilde{c}_{\mathbf{i},\sigma}^{(1)} \right\rangle$$
$$= 2 \left[\frac{3}{4} + \frac{1}{2} - \frac{1}{4} \right] + \frac{1}{4} = \frac{9}{4}. \tag{53}$$

We hereby correct a minor mistake in reference [29] where $s_{2,2} = 5/4$ was reported.

Altogether the shape-correction factor becomes

$$s(\epsilon, U) = 1 - \frac{\epsilon}{U} + \frac{9\left(\epsilon^2 - 1\right)}{4U^2} \qquad (54)$$

up to and including second order in $1/U$.

3.4 Physical quantities

3.4.1 Density of states

The density of states to order $(1/U)^0$ is readily obtained. It results from the motion of a single hole created at site **i** which moves through the Bethe lattice *via* $\hat{h}_0$ and returns to **i**. The hole has not disturbed the spin background after its return to **i**, *i.e.*, its motion appears to be free. Therefore, $\rho_h(\omega) = \rho_0(\omega + U/2)$ [29,37]. Now that we have expressed $G(\omega)$ in (40) in terms of $\hat{h}_0$ we can easily read off the expression for the density of states,

$$D_{\mathrm{LHB}}(\omega) = \int_{-2}^{2} \mathrm{d}\epsilon\, \rho_0(\epsilon) s(\epsilon, U) \delta\left(\omega + U/2 + g(\epsilon, U)\right) . \qquad (55)$$

The second-order expressions for the gap-renormalization factor $g(\epsilon, U)$ and the shape-correction factor $s(\epsilon, U)$ are given in (50) and (54).

To all orders in the $1/U$ expansion, the density of states increases algebraically near the gap,

$$D_{\mathrm{UHB}}(\omega) \sim \left(\omega - \frac{\Delta(U)}{2}\right)^{\alpha_{\mathrm{sc}}}, \quad \omega \to \frac{\Delta(U)}{2}, \qquad (56)$$

$$\alpha_{\mathrm{sc}} = \frac{1}{2}. \qquad (57)$$

This reflects the fact that $\rho_0(\epsilon)$ displays a square-root increase at the band edges. As long as the shape-correction factor $s(\epsilon, U)$ remains finite for $|\epsilon| \to 2$, the shape of the Hubbard bands near the edges remains the same for all $U > U_{\mathrm{c}}$. This behavior is qualitatively and quantitatively the same as in the exactly solvable Falicov–Kimball model [38].

3.4.2 Single-particle gap and width of the Hubbard bands

As seen from (55) the density of states becomes zero at

$$\mu_{\mathrm{LHB}}^{\pm}(U) + \frac{U}{2} + g(\mp 2, U) = 0. \qquad (58)$$

Using (16) we find for the gap

$$\Delta(U) = U + 2g(-2, U) = U - 4 - \frac{1}{U} - \frac{3}{2U^2} + \mathcal{O}\left(U^{-3}\right). \qquad (59)$$

The Hubbard bands have the width

$$W_{\mathrm{LHB}} = W_{\mathrm{UHB}} = \mu_{\mathrm{LHB}}^{+} - \mu_{\mathrm{LHB}}^{-} = 4 + \frac{3}{2U^2} \qquad (60)$$

up to second order in $1/U$. The $1/U$ expansion therefore provides a very good guess for the support for the density of states as a function of frequency in the Mott–Hubbard insulator for all U. We shall profit from this result in our numerical calculations.

The gap appears to converge rapidly as a function of $1/U$. Using the expansion parameter $x = 2/U$, we may write (59) as

$$\Delta(U) = U - 4 - \frac{1}{2}\left[x - \frac{3}{4}x^2 - \lambda x^3 + \mathcal{O}(x^4)\right]. \qquad (61)$$

Preliminary results from our calculations to third order indicate that indeed $\lambda \approx 1$.

The critical interaction strength is determined from $\Delta(U) = 0$. If we denote by $U_{\mathrm{c}}^{(m)}$ the critical interaction strength at which the mth order expression for $\Delta(U)$ vanishes, we find using $\lambda = 1$

$$\begin{aligned}
U_{\mathrm{c}}^{(0)} &= 4, \\
U_{\mathrm{c}}^{(1)} &= 4.236\,[5.9\%], \\
U_{\mathrm{c}}^{(2)} &= 4.313\,[1.8\%], \\
U_{\mathrm{c}}^{(3)} &\approx 4.357\,[\approx 1.0\%]. \qquad (62)
\end{aligned}$$

The numbers in brackets give the percentage change to the result from the previous order. Under the assumptions that all higher-order contributions go in the same direction, and decay similarly fast, we *conjecture* that

$$U_{\mathrm{c}}^{\mathrm{sc}} = 4.40 \pm 0.09, \qquad (63)$$

which allows for another 3% increase of U_{c} with respect to the third-order result. We shall see in Sections 3 and 4 that this estimate is in excellent agreement with our results from numerical calculations.

This analysis may also be carried out for the exactly solvable Falicov–Kimball model [38], where

$$\begin{aligned}
U_{\mathrm{c,FK}}^{(0)} &= 2.828, \\
U_{\mathrm{c,FK}}^{(1)} &= 2.414\,[-14.6\%], \\
U_{\mathrm{c,FK}}^{(2)} &= 2.242\,[-7.15\%], \\
U_{\mathrm{c,FK}}^{(3)} &= 2.167\,[-3.31\%], \\
U_{\mathrm{c,FK}} &= 2\,[-8.35\%]. \qquad (64)
\end{aligned}$$

It appears to us that for the Falicov–Kimball model the convergence of the critical interaction strength is not quite as fast as for the Hubbard model. The relative changes are about a factor of three smaller for the Hubbard case, and correspondingly, we have reason to expect that for the Hubbard model the exact result is within 3% of the third-order result.

3.4.3 Self-energy and momentum distribution

Up to an overall coefficient, the single-particle density of states is identical to the imaginary part of the Green function (15). The Kramers–Kronig relation provides the real

part as

$$\Re G(\omega) = \mathrm{P} \int_{\mu_{\mathrm{LHB}}^-}^{\mu_{\mathrm{LHB}}^+} \mathrm{d}\omega' D_{\mathrm{LHB}}(\omega') \left(\frac{1}{\omega - \omega'} + \frac{1}{\omega + \omega'} \right). \tag{65}$$

The local Green function per spin $G_\sigma(\omega) = G(\omega)/2$ and the single-particle self-energy $\Sigma_\sigma(\omega) = -\Sigma_\sigma(-\omega)$ are related by

$$G_\sigma(\omega) = \int_{-\infty}^\infty \mathrm{d}\epsilon \rho_0(\epsilon) \frac{1}{\omega - \epsilon - \Sigma_\sigma(\omega) + \mathrm{i}\eta \, \mathrm{sgn}(\omega)},$$
$$= G_\sigma^0(\omega - \Sigma_\sigma(\omega)) \tag{66}$$

in infinite dimensions, and

$$G_\sigma^0(z) = \frac{z}{2} \left[1 - \sqrt{1 - \frac{4}{z^2}} \right] \tag{67}$$

holds for the Bethe lattice [31] so that

$$\omega - \Sigma_\sigma(\omega) = G_\sigma(\omega) + \frac{1}{G_\sigma(\omega)}. \tag{68}$$

The real and imaginary part of the self-energy are then readily obtained as

$$\Re\Sigma_\sigma(\omega) = \omega - \Re G_\sigma(\omega) \left[1 + \frac{1}{(\Re G_\sigma(\omega))^2 + (\Im G_\sigma(\omega))^2} \right], \tag{69}$$

$$\Im\Sigma_\sigma(\omega) = -\Im G_\sigma(\omega) \left[1 - \frac{1}{(\Re G_\sigma(\omega))^2 + (\Im G_\sigma(\omega))^2} \right]. \tag{70}$$

Note that this self-energy is not perturbatively linked to the weak-coupling limit. For example, its imaginary part displays a peak $\delta(\omega)$ and its real part diverges proportional to $(-1/\omega)$ as $\omega \to 0$.

The single-particle spectral function $A_\sigma(\epsilon; \omega)$ is defined by

$$A_\sigma(\epsilon; \omega) = -\frac{1}{\pi} \mathrm{sgn}(\omega) \Im \left(\frac{1}{\omega - \epsilon - \Sigma_\sigma(\omega)} \right)$$
$$= -\frac{1}{\pi} \mathrm{sgn}(\omega) \frac{\Im\Sigma_\sigma(\omega)}{(\omega - \epsilon - \Re\Sigma_\sigma(\omega))^2 + (\Im\Sigma_\sigma(\omega))^2}. \tag{71}$$

It does not display a quasi-particle contribution. The momentum distribution

$$n_\sigma(\epsilon) = \int_{\mu_{\mathrm{LHB}}^-}^{\mu_{\mathrm{LHB}}^+} \mathrm{d}\omega A_\sigma(\epsilon; \omega) \tag{72}$$

depends on momentum implicitly via $\epsilon \equiv \epsilon(\mathbf{k})$. In the insulating phase, the momentum distribution is an analytic function of ϵ. Since $n_\sigma(\epsilon) = 1 - n_\sigma(-\epsilon)$ due to particle-hole symmetry, the Taylor expansion around $\epsilon = 0$ has the form

$$n(\epsilon) = \frac{1}{2} - c_1\epsilon - c_3\epsilon^3 \ldots \tag{73}$$

For the Bethe lattice we obtain $c_1 = 1/(2U) + \mathcal{O}(U^{-3})$ [35], and, in general, $c_{2n-1} = \mathcal{O}(U^{1-2n})$.

4 Exact diagonalization with fixed energy interval (FE-ED)

In this section, we first discuss the single-impurity model onto which the Hubbard model can be mapped in the limit of infinite dimensions. Then we propose the Fixed-Energy Exact Diagonalization (FE-ED) as a new scheme for the numerical solution of the DMFT equations. The results of this approach corroborate our findings from Section 3.

4.1 Dynamical mean-field theory (DMFT)

In the limit of infinite dimensions [7] and under the conditions of translational invariance and convergence of perturbation theory in strong and weak coupling, lattice models for correlated electrons can be mapped onto single-impurity models which need to be solved self-consistently [3,9,10]. Unfortunately, these impurity models cannot be solved analytically.

For an approximate numerical treatment, various different implementations are conceivable; see, $e.g.$, reference [39] for a recent implementation. One realization is the single-impurity Anderson model in 'star geometry',

$$\hat{H}_{\mathrm{SIAM}} = \sum_{\ell=1;\sigma}^{n_s-1} \epsilon_\ell \hat{\psi}_{\sigma;\ell}^+ \hat{\psi}_{\sigma;\ell} + U \left(\hat{d}_\uparrow^+ \hat{d}_\uparrow - \frac{1}{2} \right) \left(\hat{d}_\downarrow^+ \hat{d}_\downarrow - \frac{1}{2} \right)$$
$$+ \sum_\sigma \sum_{\ell=1}^{n_s-1} V_\ell \left(\hat{\psi}_{\sigma;\ell}^+ \hat{d}_\sigma + \hat{d}_\sigma^+ \hat{\psi}_{\sigma;\ell} \right), \tag{74}$$

where V_ℓ are real, positive hybridization matrix elements. The model describes the hybridization of an impurity site with Hubbard interaction to $n_s - 1$ bath sites without interaction at energies ϵ_ℓ with $0 < \epsilon_1 < \epsilon_2 < \ldots < \epsilon_{(n_s-1)/2}$. In order to ensure particle-hole symmetry, we have to set $\epsilon_\ell = -\epsilon_{n_s-\ell}$ and $V_\ell = V_{n_s-\ell}$ for $\ell = (n_s+1)/2, \ldots, n_s-1$. Moreover, since we are interested in the Mott–Hubbard insulator, we only use odd n_s so that there is no bath state at $\epsilon = 0$.

For a given set of $(n_s - 1)$ parameters (ϵ_ℓ, V_ℓ) the model (74) defines a many-body problem for which the single-particle Green function

$$G_\sigma^{(n_s)}(t) = -\mathrm{i} \left\langle \hat{\mathcal{T}} \left[\hat{d}_\sigma(t) \hat{d}_\sigma^+ \right] \right\rangle_{\mathrm{SIAM}} \tag{75}$$

can be calculated numerically for $n_s \leq 15$ with the (dynamical) Lanczos technique. In (75) $\langle \ldots \rangle_{\mathrm{SIAM}}$ implies the ground-state expectation value within the single-impurity model. Typically, the imaginary part of the Green function displays n_s peaks with large weight and many other small peaks whose number depends on the number of states n_{L} kept in the Lanczos diagonalization.

Ultimately, we are interested in the limit $n_s \to \infty$ where the hybridization function

$$H^{(n_s)}(\omega) = \sum_{\ell=1}^{n_s-1} \frac{V_\ell^2}{\omega - \epsilon_\ell + \mathrm{i}\eta \, \mathrm{sgn}(\omega)} \tag{76}$$

should smoothly approach the hybridization function of the *continuous* problem,

$$H(\omega) = \lim_{n_s \to \infty} H^{(n_s)}(\omega). \quad (77)$$

Correspondingly, the Green function should fulfill

$$G_\sigma(\omega) = \lim_{n_s \to \infty} G_\sigma^{(n_s)}(\omega). \quad (78)$$

At self-consistency, the Green function of the impurity problem describes the Hubbard model in infinite dimensions. As shown in [3], on the Bethe lattice the hybridization function must obey the simple relation

$$H(\omega) = G_\sigma(\omega). \quad (79)$$

This equation closes the self-consistency cycle for the continuous problem: we have to choose bath energies and hybridizations in such a way that the single-particle Green function and the hybridization function fulfill (79).

4.2 Implementation of the FE-ED

Any numerical scheme for the DMFT faces two problems. First, and foremost, the DMFT requires the Green function $G_\sigma(\omega)$ on the whole, continuous frequency axis but only the information for $(n_s - 1)$ energies ϵ_ℓ is provided. This is in contrast to standard numerical problems in many-body theory where a given energy interval, typically of the order of the band-width W, is resolved to accuracy W/n_s. Concomitantly, in numerical DMFT calculations it is not a priori clear how observables scale as a function of $1/n_s$. Moreover, there can be more than one self-consistent set of parameters (ϵ_ℓ, V_ℓ) for fixed n_s.

Second, the self-consistency condition (79) holds for the continuous $H(\omega)$ and $G_\sigma(\omega)$, not for their discretized counterparts. Therefore, the various ED methods will differ in the way the information is extracted from $G_\sigma^{(n_s)}(\omega)$ in one iteration in order to specify the input parameters (ϵ_ℓ, V_ℓ) for the next iteration. This is another source of ambiguity. Since the scaling as a function of $1/n_s$ is not clear, it cannot be guaranteed that different schemes will ultimately coincide as $n_s \to \infty$.

In this work, we use the results of the $1/U$ expansion to circumvent the first problem. We have seen in Section 3 that the Hubbard bands are distributed symmetrically around $\omega = 0$ with width $W_{\text{LHB}} = W_{\text{UHB}} = W^* \approx W$. Thus, we actually need to resolve a *finite* frequency interval. In practice, we use $W^* = 4.5$ as our maximum bandwidth. Then, we determine the onset of the upper Hubbard band self-consistently, $\Delta(U)/2$; see below. To this end, we start with some input guess $\Delta_{\text{in}}(U)$. In practice, we use the second-order estimate of the gap from (59) to speed up the convergence of this outer self-consistency cycle.

We choose to discretize the Hubbard bands equidistantly, *i.e.*, we fix the energies ϵ_ℓ for $1 \le \ell \le (n_s-1)/2$ by

$$\epsilon_\ell = \frac{\Delta_{\text{in}}(U)}{2} + \left(\ell - \frac{1}{2}\right)\delta W; \quad \delta W = \frac{2W^*}{n_s - 1}. \quad (80)$$

For $n_s \le 15$ and $U > W$ none of the ϵ_ℓ can move outside the Hubbard bands. By fixing the energies at the centers of the equidistant intervals

$$I_\ell = \left[\frac{\Delta_{\text{in}}(U)}{2} + (\ell - 1)\delta W, \frac{\Delta_{\text{in}}(U)}{2} + \ell\delta W\right], \quad (81)$$

$1 \le \ell \le (n_s - 1)/2$, we can be sure that our resolution of the Hubbard bands becomes increasingly better as n_s increases.

We address the second problem in the following way. We apply a constant broadening of width δW to the individual peaks in $\Im G_\sigma^{(n_s)}(\omega_r)$, $r = 1, \ldots, n_{\text{L}}$. Then, we collect the weight into the intervals I_ℓ and assign this weight to $w_\ell = V_\ell^2$. In the Caffarel–Krauth scheme [18,26] a fictitious temperature and a χ^2-fitting procedure is used. We have found that our simple approach is equally suitable.

Typically, the weight of the peaks at energies outside the Hubbard bands is very small. Thus we set

$$w_\ell = \int_{I_\ell} d\omega \sum_{r=1}^{n_{\text{L}}} \Im G_\sigma^{(n_s)}(\omega_r)$$
$$\times \frac{\Theta(\omega - \omega_r + \delta W/2) - \Theta(\omega - \omega_r - \delta W/2)}{\delta W}. \quad (82)$$

In order to generate an educated input guess for $V_\ell = \sqrt{w_\ell}$ we use the results from second-order perturbation theory in (82).

Now that we have determined the input energies ϵ_ℓ and hybridizations V_ℓ, we can start the inner self-consistency cycle. We perform a dynamical Lanczos diagonalization which gives $\Im G_\sigma^{(n_s)}(\omega_r)$ for n_{L} Lanczos frequencies. Most of the peaks carry little weight, only n_s of them are significant. Using (82), the Green function provides the hybridization elements for the next iteration of the inner self-consistency cycle. We have checked that, for fixed n_s, a unique solution for w_l is found for various starting choices.

After we have determined the parameters V_ℓ self-consistently for fixed n_s and a given U, we may calculate the single-particle gap from the difference in the ground-state energy at half band-filling, $E_0^{(n_s)}(N; U)$, and the ground-state energy for one particle more than half band-filling, $E_0^{(n_s)}(N + 1; U)$, as

$$\Delta^{E;(n_s)}(U) = 2\left(E_0^{(n_s)}(N + 1; U) - E_0^{(n_s)}(N; U)\right). \quad (83)$$

After the extrapolation to the thermodynamic limit,

$$\Delta^E(U) = \lim_{n_s \to \infty} \Delta^{E;(n_s)}(U), \quad (84)$$

we obtain a new estimate for the onset of the Hubbard bands, and we could restart the outer self-consistency cycle. Unfortunately, the extrapolation $n_s \to \infty$ has limited accuracy, and it is difficult to obtain a truly converged solution which fulfills $\Delta^E(U) = \Delta_{\text{in}}(U)$ ('energy criterion'), even at very large values, *e.g.*, $U = 14$. An example for $U = 6$ is shown in Figure 1. It is seen that the results around $\Delta_{\text{in}} = 0.88$ are almost but not quite converged.
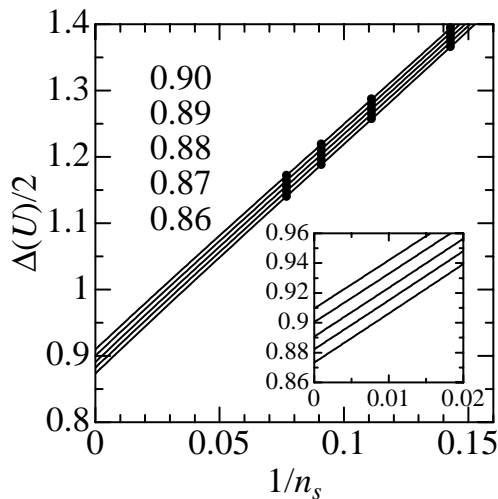
Fig. 1. Gap $\Delta^{E;(n_s)}(U)$ from (83) for $U = 6$ as a function of $1/n_s$ for various Δ_{in}.
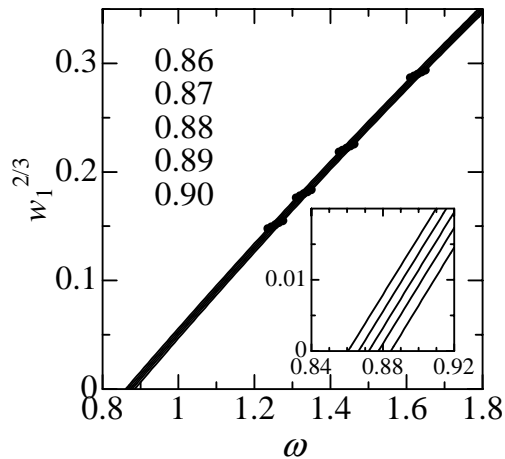

Fig. 2. Weight of the first peak in the upper Hubbard band for $n_s = 7, 9, 11, 13$, $U = 6$, and various Δ_{in}. For a square-root increase of the density of state, $w_1^{2/3}$ should scale linearly with $1/n_s$, see (85). The intersection of the extrapolated curves give $\Delta^w(6)$. Inset: expansion of the region around the zero intercept.

In order to stabilize the outer self-consistency cycle, we need an additional criterion. To this end, we confirm numerically that the density of states near the gap increases algebraically; see (56). The weight $w_1 = V_1^2$ of the peak in $D_{\text{UHB}}(\omega)$ at ϵ_1 represents the integral of the density of states from $\Delta(U)/2$ up to $\epsilon_1 = \Delta(U)/2 + \mathcal{O}(1/n_s)$. Using (56) we should find

$$[w_1]^{1/(\alpha+1)} = \epsilon_1 - \frac{\Delta^w(U)}{2}. \qquad (85)$$

The linear behavior with $\alpha_{\text{FE}-\text{ED}} = 1/2$ is confirmed in Figure 2 for $U = 6$. The plot also provides $\Delta^w(U)$ for a given $\Delta_{\text{in}}(U)$ as the intersection of the extrapolated curves with the frequency axis ('weight criterion').

The two values for the gap from the 'energy criterion' (84) and from the 'weight criterion' (85) do not agree
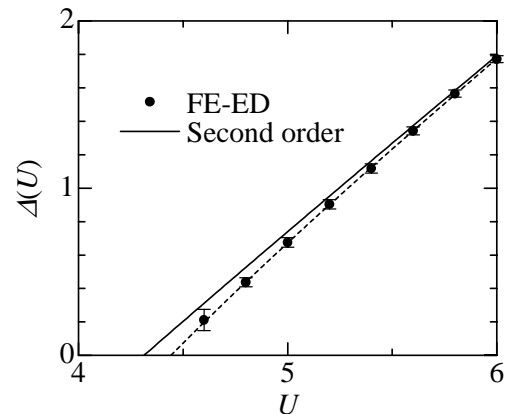
Fig. 3. Mott–Hubbard gap as a function of U in FE-ED in comparison with the $1/U$ expansion to second order (59). The error bars give the two values which enter (86) as obtained from the energy and weight criterion. The dashed line is a cubic polynomial interpolation based on the data points for $U \geq 4.8$.

perfectly. Therefore, we use the self-consistency condition

$$\Delta_{\text{in}}(U) \overset{!}{=} \Delta_{\text{out}}(U) \equiv \frac{\Delta^E(U) + \Delta^w(U)}{2}, \qquad (86)$$

which determines $\Delta^{\text{FE}-\text{ED}}(U)$ after convergence of the outer self-consistency cycle. The difference,

$$\delta\Delta^{\text{FE}-\text{ED}}(U) = \frac{|\Delta^E(U) - \Delta^w(U)|}{2} \qquad (87)$$

is our estimate for the accuracy of $\Delta^{\text{FE}-\text{ED}}(U)$. We call this implementation of the DMFT the 'Fixed-Energy Exact Diagonalization (FE-ED)' scheme.

The converged FE-ED gap as a function of the interaction strength is shown in Figure 3. The agreement with the $1/U$ expansion is excellent for $U \geq 5$ because the two curves agree within the error bars. Unfortunately, the uncertainty in $\delta\Delta^{\text{FE}-\text{ED}}(U)$ increases towards the transition. Using a cubic polynomial extrapolation for data points $U \geq 4.8$, as shown in the figure, leads to $U_c^{\text{FE}-\text{ED}} = 4.44$. As seen from the figure, the spline interpolation also touches the data point for $U = 4.6$, to within its error bars. If this data point were included, the critical value would be $U_c^{\text{FE}-\text{ED}} = 4.42$. Thus, we give

$$U_c^{\text{FE}-\text{ED}} = 4.43 \pm 0.05 \qquad (88)$$

as our best estimate for the location of the closing of the gap.

In Figure 4 we compare the density of states of our FE-ED for $n_s = 15$ with those of the $1/U$ expansion to second order. Both for $U = 5$ and $U = 6$ the agreement is very good. Our FE-ED thus allows us to determine the density of states quantitatively with increasing resolution as n_s increases.

Aside from the fact that the $1/U$ expansion provides an educated guess for the position of the Hubbard bands, the FE-ED is an independent numerical method based on
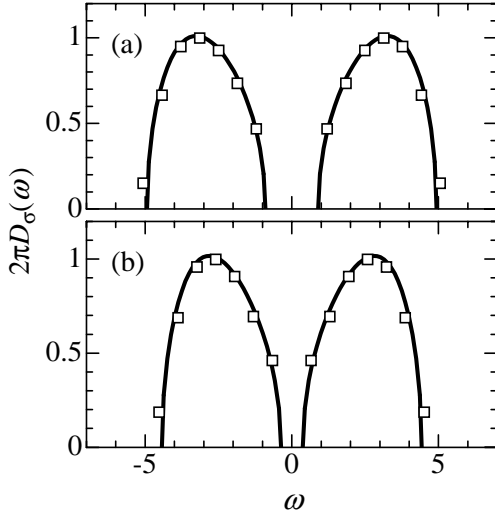
Fig. 4. Density of states, $D_\sigma(\omega) = D(\omega)/2$, as obtained from the $1/U$ expansion, equation (55), and the FE-ED for $n_s = 15$; (a) $U = 6$, (b) $U = 5$.

the DMFT. The good agreement of the results mutually supports the applicability of either method. Therefore, we are confident that neither of the limitations poses a serious problem for our investigation of the Mott–Hubbard insulator, *i.e.*, the problem of convergence of the series expansions in $1/U$ and in $1/n_s$, respectively, are not too serious in our approaches.

5 Random dispersion approximation

In this section, we present numerical results from the Random Dispersion Approximation (RDA). This method is not based on the DMFT and thereby provides yet another independent check for the validity of our results as obtained in Sections 3 and 4.

In the Random Dispersion Approximation, the dispersion relation $\epsilon(\mathbf{k})$ in the kinetic energy is replaced by a random quantity $\epsilon^{\mathrm{RDA}}(\mathbf{k})$ where the bare density of states acts as probability distribution,

$$\rho(\epsilon) = \frac{1}{L} \sum_{\mathbf{k}} \delta(\epsilon - \epsilon^{\mathrm{RDA}}(\mathbf{k})). \tag{89}$$

This is the characteristic property of the dispersion relation in infinite dimensions [2] so that the RDA with the semi-elliptic density of states (7) becomes exact for the Bethe lattice with infinite coordination number in the thermodynamic limit $L \to \infty$.

All correlation functions factorize, analogously to (6). For the mth-order correlation function,

$$\rho_{1...m}(\epsilon_1, \ldots, \epsilon_m) = \frac{1}{L} \sum_{\mathbf{k}} \prod_{l=1}^{m} \delta\left(\epsilon_l - \epsilon(\mathbf{k} + \mathbf{q}_l)\right), \tag{90}$$

we give a recursive algorithm. Let $\rho_l(\epsilon_l) \equiv \rho(\epsilon_l)$ be the bare density of states, for $m = 2$, see (6), and for $m \geq 3$

we define

$$\rho_{1...(n-1);n,...,m}(\epsilon_1, \ldots, \epsilon_m) = \prod_{i,j=n}^{m} (1 - \delta_{\mathbf{q}_i, \mathbf{q}_j})$$
$$\prod_{i=n}^{m} \prod_{j=1}^{n-1} (1 - \delta_{\mathbf{q}_i, \mathbf{q}_j})$$
$$\rho_{1...m}(\epsilon_1, \ldots, \epsilon_m). \tag{91}$$

In this quantity, all $\mathbf{q}_l$ for $l \geq n$ are pairwise different, and they are also different from those $\mathbf{q}_i$ with $i \leq n - 1$. The recursion formula

$$\rho_{1...n;(n+1),...,m}(\epsilon_1, \ldots, \epsilon_m) =$$
$$\rho_{1...(n-1);n,...,m}(\epsilon_1, \ldots, \epsilon_m)$$
$$+ \left[1 - \prod_{l=1}^{n-1} (1 - \delta_{\mathbf{q}_n, \mathbf{q}_l} \delta(\epsilon_n - \epsilon_l)) \right]$$
$$\times \rho_{1...(n-1);(n+1),...,m}(\epsilon_1, \ldots, \epsilon_{n-1}, \epsilon_{n+1}, \ldots, \epsilon_m) \tag{92}$$

starts with $\rho_{1...m}(\epsilon_1, \ldots \epsilon_m) \equiv \rho_{1...m;}(\epsilon_1, \ldots \epsilon_m)$ and terminates at

$$\rho_{1;2,...,m}(\epsilon_1, \ldots \epsilon_m) = \prod_{l=1}^{m} \rho(\epsilon_l) \prod_{i,j=1}^{m} (1 - \delta_{\mathbf{q}_i, \mathbf{q}_j}). \tag{93}$$

For example,

$$\rho_{123}(\epsilon_1, \epsilon_2, \epsilon_3) = \left[1 - \prod_{l=1}^{2} (1 - \delta_{\mathbf{q}_3, \mathbf{q}_l} \delta(\epsilon_3 - \epsilon_l)) \right] \rho_{12}(\epsilon_1, \epsilon_2)$$
$$+ \rho_{12;3}(\epsilon_1, \epsilon_2, \epsilon_3) \tag{94}$$

with

$$\rho_{12;3}(\epsilon_1, \epsilon_2, \epsilon_3) = \delta_{\mathbf{q}_1, \mathbf{q}_2} \delta(\epsilon_1 - \epsilon_2) \rho(\epsilon_1, \epsilon_3)$$
$$+ \rho(\epsilon_1) \rho(\epsilon_2) \rho(\epsilon_3)(1 - \delta_{\mathbf{q}_1, \mathbf{q}_2})(1 - \delta_{\mathbf{q}_1, \mathbf{q}_3})(1 - \delta_{\mathbf{q}_2, \mathbf{q}_3}). \tag{95}$$

In order to put the RDA into practice, we choose a one-dimensional lattice of L sites in momentum space

$$k_\ell = \frac{2\pi}{L} \left(-\frac{L+1}{2} + \ell \right) (\ell = 1, \ldots, L), \tag{96}$$

and determine the dispersion relation $\epsilon(k)$ as the solution of the implicit equation

$$k/2 = (2\epsilon(k)/W) \left[1 - (2\epsilon(k)/W)^2 \right]^{1/2}$$
$$+ \arcsin\left(2\epsilon(k)/W\right). \tag{97}$$

This choice guarantees $\rho(\epsilon) = \rho_0(\epsilon)$ in the thermodynamic limit.

Next, we choose a permutation $\mathcal{Q}_\sigma$ for each spin direction σ which permutes the sequence $\{1, \ldots, L\}$ into $\{\mathcal{Q}_\sigma[1], \ldots, \mathcal{Q}_\sigma[L]\}$; in this way there are $(L!)^2$ independent realizations for a given finite L. A specific realization

of the RDA dispersion is then defined by $\mathcal{Q} = [\mathcal{Q}_\uparrow, \mathcal{Q}_\downarrow]$. Then, the numerical task is the Lanczos diagonalization of the Hamiltonian

$$\hat{H}^{\mathcal{Q}} = \sum_\sigma \sum_{\ell=1}^{L} \epsilon(k_{\mathcal{Q}_\sigma[\ell]})\hat{c}_{k_\ell,\sigma}^+ \hat{c}_{k_\ell,\sigma} + U\hat{D}. \qquad (98)$$

In this way we obtain the ground-state energy $E_0^{\mathcal{Q}}(N; U)$, the single-particle gap

$$\Delta^{\mathcal{Q}}(U) = 2[E_0^{\mathcal{Q}}(N = L + 1; U) - E_0^{\mathcal{Q}}(N = L; U)], \quad (99)$$

and the momentum distribution

$$n^{\mathcal{Q}}(\epsilon; U) = \frac{1}{2}\sum_\sigma \left. \langle \hat{n}_{k_\ell,\sigma}\rangle \right|_{\epsilon(k_\ell)=\epsilon}, \qquad (100)$$

where $\langle \ldots \rangle$ denotes the ground-state expectation value for the realization $\mathcal{Q}$.

As a next step, we obtain all physical quantities for fixed system size L by averaging over the number $N_{\mathcal{Q}}$ of realizations $\mathcal{Q}$. Typically, we choose at least $N_{\mathcal{Q}} = 100$ for $6 \leq L \leq 16$. For the physical quantities we obtain Gaussian-shaped distributions for which we can determine the average values, e.g.,

$$\Delta(U) = \frac{1}{N_{\mathcal{Q}}}\sum_{\mathcal{Q}} \Delta^{\mathcal{Q}}(U), \qquad (101)$$

$$n(\epsilon; U) = \frac{1}{N_{\mathcal{Q}}}\sum_{\mathcal{Q}} n^{\mathcal{Q}}(\epsilon; U) \qquad (102)$$

with accuracy $\mathcal{O}(1/N_{\mathcal{Q}})$.

In order to improve slightly the quality of our distributions, we impose a filter on our randomly chosen permutations. For a truly random dispersion, $L|t^{\mathrm{RDA}}(\ell)|^2 = \overline{\epsilon^2}$ is independent of ℓ; see (36). Therefore, we discard those realizations for which

$$\sum_{\ell=1}^{L-1} \left[|t^{\mathcal{Q}_\sigma}(\ell)|^2 - \overline{\epsilon^2} \right]^2 > d_L \qquad (103)$$

with, $d_L \approx 0.2$ for $L \leq 16$. In this way we admit about every second of the randomly chosen configurations. Note that there are of the order of $(L!)^2$ different realizations so that our filter does not introduce any unwanted bias.

In Figure 5 we show the gap as calculated from (99). Typically, the statistical errors are of the symbol size or smaller. Due to the sizable odd-even effect, we extrapolate the data for odd and even L separately. This gives another check on the influence of finite-size effects. It is seen that the behavior as a function of $1/L$ is quite regular, i.e., the finite-size gap behaves as in generic many-body problems.

In Figure 6 we show the result of an extrapolation of the finite-size data for even and odd system sizes. We consider the agreement with the $1/U$ expansion as quite acceptable for $U > U_c$ which we can estimate as

$$U_c^{\mathrm{RDA}} = 4.0 \pm 0.4. \qquad (104)$$
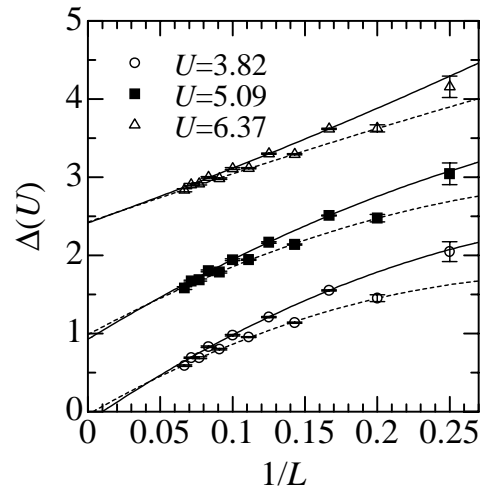
**Fig. 5.** Gap as a function of system size ($L \leq 15$) in the Random Dispersion Approximation for various values of the interaction strength. The extrapolations are done for even and odd system sizes separately.
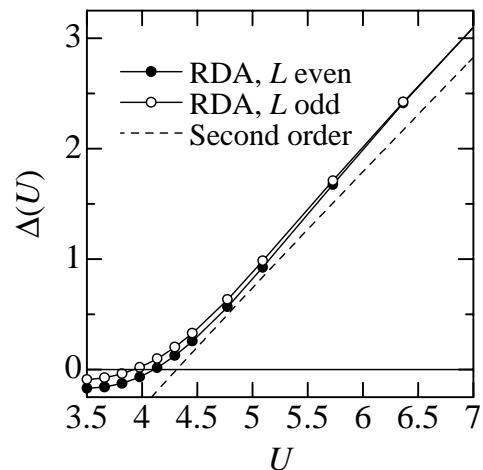
Fig. 6. Gap in RDA as a function of the interaction strength. Filled symbols: extrapolation from even L, open symbols: extrapolation from odd L. Also shown is the result from second-order perturbation theory in $1/U$ (59).

Finite-size effects result in a smearing of the gap at the transition, and a gap exponent cannot be determined. However, the fact that the extrapolated gap turns slightly negative can be taken as an indication that the gap opens at U_c with an exponent not too far from unity.

Lastly, in Figure 7 we show the momentum dispersion within the RDA for $U = 16/\pi \approx 5.1$, together with the result of the $1/U$ expansion. Obviously, there are large finite-size effects which make a finite-size extrapolation difficult. The reason for this behavior is related to the fact that all finite-size systems appear to have a finite discontinuity at the Fermi energy as well as a finite single-particle gap. For insulators, the apparent jump in $n(\epsilon)$ only vanishes in the thermodynamic limit [11]. The agreement between the RDA data and the momentum distribution as
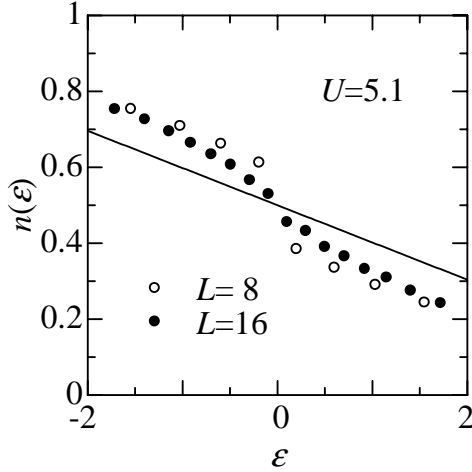
Fig. 7. Momentum distribution in RDA for $U = 16/\pi \approx 5.1$ for system sizes $L = 8$ and $L = 16$. The full line is the result from first-order perturbation theory in $1/U$ (73).

obtained from the $1/U$ expansion is good enough to argue that the RDA will eventually converge to the $1/U$ result for $L \to \infty$.

The RDA becomes exact for lattice electrons in high dimensions. In contrast to the FE-ED it is not based on the DMFT self-consistency equations. Moreover, it does not require the convergence of the $1/U$ expansion. Despite the observed limitations in accuracy, the results from RDA confirm of our findings in Sections 3 and 4: the $1/U$ expansion and the FE-ED provide an accurate description of the Mott–Hubbard insulator whose gap opens at the critical interaction strength $U_c = 4.43 \pm 0.05$.

6 Comparison with analytic approximations

In this section we compare our results with two analytic approximations to the DMFT which become exact for the Bethe lattice in the strong-coupling limit, *i.e.*, they reproduce $D_{\mathrm{LHB}}(\omega)$ to order $(1/U)^0$. The Local Moment Approach (LMA) [25] outlined in Appendix B fulfills this condition by construction, whereas this is a non-trivial result within the Iterated Perturbation Theory (IPT) [40]; a proof [41] is given in Appendix A.

The Hubbard-III approximation [42] also fulfills this criterion but it fails to reproduce qualitatively the density of states for smaller interaction strengths [29]. Therefore, we do not further discuss the Hubbard-III approximation.

6.1 Iterated perturbation theory

For the $Z \to \infty$ Bethe lattice, the self-energy and the local single-particle Green function are related by

$$G_\sigma(\omega) = \frac{1}{\omega - G_\sigma(\omega) - \Sigma_\sigma(\omega)}. \qquad (105)$$

In IPT, the self-energy is approximated by

$$\Sigma_\sigma(\omega) = U^2 \int_{-\infty}^{\infty} \frac{\mathrm{d}\Omega}{2\pi\mathrm{i}} \Pi(\Omega)\mathcal{G}_\sigma(\omega - \Omega). \qquad (106)$$

Here $\mathcal{G}_\sigma(\omega)$ is the host Green function,

$$\mathcal{G}_\sigma(\omega) = \frac{1}{\omega - G_\sigma(\omega)}, \qquad (107)$$

and we have defined the polarization bubble

$$\Pi(\omega) = -\int_{-\infty}^{\infty} \frac{\mathrm{d}\omega_1}{2\pi\mathrm{i}} \mathcal{G}_\sigma(\omega_1)\mathcal{G}_\sigma(\omega_1 - \omega) \qquad (108)$$

in terms of the host Green function.

It is helpful to consider the low-frequency behavior of the IPT equations. First, consider the host Green function (107), whose low-frequency behavior in the insulator is given generally by

$$\mathcal{G}_\sigma(\omega) = \frac{|m_0|}{\omega + \mathrm{i}\eta \, \mathrm{sgn}(\omega)} \qquad \omega \to 0 \,, \qquad (109)$$

where the pole weight $|m_0|$ is given in terms of

$$|\gamma_1| = -\left. \frac{\mathrm{d}G_\sigma(\omega)}{\mathrm{d}\omega} \right|_{\omega=0} \qquad (110)$$

as

$$|m_0| = (1 + |\gamma_1|)^{-1}. \qquad (111)$$

$\mathcal{G}_\sigma(\omega)$ also contains a Hubbard-band contribution, which is strictly separated from the pole at $\omega = 0$ throughout the insulating phase, whence we may write

$$\mathcal{G}_\sigma(\omega) = \mathcal{G}_\sigma^{\mathrm{band}}(\omega) + \frac{|m_0|}{\omega + \mathrm{i}\eta \, \mathrm{sgn}(\omega)}, \qquad (112)$$

which actually defines $\mathcal{G}_\sigma^{\mathrm{band}}(\omega)$. In IPT the low-frequency behavior of $\mathcal{G}_\sigma(\omega)$ directly determines the low-frequency behavior of $\Sigma_\sigma(\omega)$; inserting (112) into (106) and (108) yields

$$\Sigma_\sigma(\omega) = \frac{A}{\omega} \qquad \omega \to 0 \,, \qquad (113)$$

where

$$A = \frac{U^2}{4} |m_0|^3 \,. \qquad (114)$$

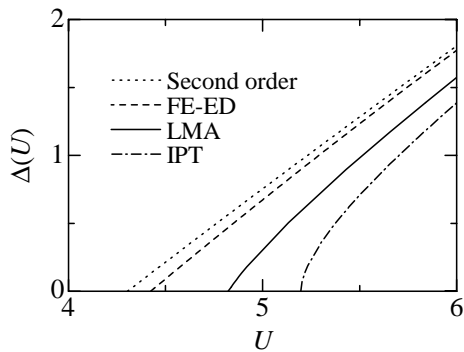To satisfy self-consistency (105) requires

$$A = \frac{1}{|\gamma_1|}, \qquad (115)$$

so that (111), (114) and (115) determine the low-frequency behavior of the problem. In particular, they yield
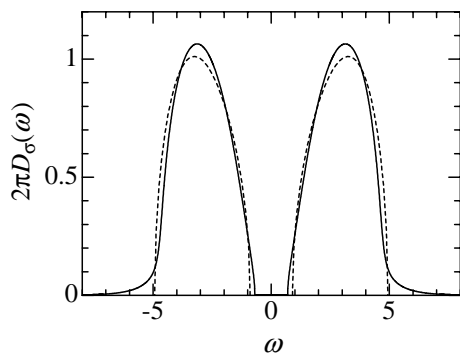
$$\frac{2}{|m_0| \sqrt{1 - |m_0|}} = U. \qquad (116)$$

The denominator $|m_0| \sqrt{1 - |m_0|}$ reaches its maximum value when

$$|m_0| = |m_0^{\mathrm{min}}| = \frac{2}{3}, \qquad (117)$$

Fig. 8. Gap as a function of interaction strength from second-order perturbation theory in $1/U$ (59), the Fixed-Energy Exact Diagonalization (FE-ED), the Local Moment Approach (LMA), and Iterated Perturbation Theory (IPT).

Fig. 9. Density of states for $U = 6$ in Iterated Perturbation Theory (solid line) and in second-order perturbation theory in $1/U$ (55) (dashed line).

thus locating the destruction of the Mott–Hubbard insulator at

$$U_{\mathrm{c},1}^{\mathrm{IPT}} = 3\sqrt{3} \approx 5.2, \qquad (118)$$

as earlier reported by Rozenberg *et al.* [23]. The IPT prediction lies considerably above our best estimate $U_{\mathrm{c}} = 4.43 \pm 0.05$.

This discrepancy can already be seen for larger interaction strength where

$$\Delta^{\mathrm{IPT}}(U) = U - 4 - \frac{5}{2U} - \dots \qquad (119)$$

The first-order correction can be obtained analytically, as shown in Appendix A. It is more than twice as large as our exact coefficient. Therefore, the agreement is good only for $U > 6$. The gap in IPT is compared to our FE-ED and second-order results in Figure 8. As seen from the figure, and shown analytically in [41], the gap closes continuously but with an exponent of $\gamma_{\mathrm{IPT}} = 1/2$.

Despite the deviations in the gap values, the IPT correctly reproduces the overall shape of the density of states for $U > 6$, apart from the artificial high-energy tails of the IPT Hubbard bands. The threshold exponent for the density of states is correct, $\alpha_{\mathrm{IPT}} = 1/2$ in the Mott–Hubbard insulator. Nevertheless, the deviations from the second-order result (55) are quite noticeable at $U = 6$ which is
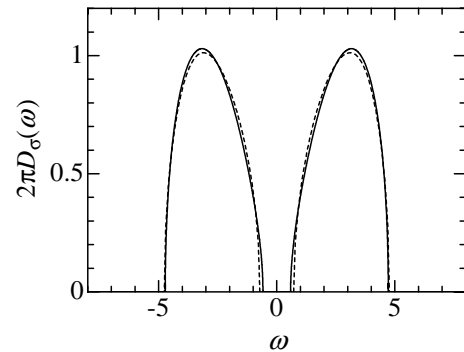
Fig. 10. Density of states for $U = 4\sqrt{2} \approx 5.66$ from the Local Moment Approach (full line) and second-order perturbation theory in $1/U$ (55) (dashed line). On the scale of the figure, the gap from FE-ED and from second-order perturbation theory are almost the same.

shown in Figure 9. We conclude that the IPT provides a quantitatively correct description of the Mott–Hubbard insulator down to $U \approx 6$. However, IPT seriously underestimates the stability of the Mott–Hubbard insulator.

6.2 Local moment approach

Details of the LMA are given in reference [25]. As shown in Appendix B, the LMA gap to first order in $1/U$ is given by

$$\Delta^{\mathrm{LMA}}(U) = U - 4 - \frac{3}{2U}. \qquad (120)$$

The first-order correction is close to the exact one, and, correspondingly, the agreement with our strong-coupling result (59) is very good down to $U \approx 5.5$. For smaller interaction strengths, the deviations become noticeable and the LMA predicts the collapse of the insulator to occur at $U_{\mathrm{c}}^{\mathrm{LMA}} = 4.82$ [25] which is larger than our best estimate $U_{\mathrm{c}} = 4.43 \pm 0.05$. The LMA yields a critical gap exponent $\gamma_{\mathrm{LMA}} = 1$ [41], in contrast to IPT where $\gamma_{\mathrm{IPT}} = 1/2$. The gap in LMA and our second-order and FE-ED results are shown in Figure 8.

It is seen that the LMA gap is more accurate than the gap from IPT. In particular, this holds true for the density of states as the LMA very well reproduces our second-order result (55). The comparison for $U = 4\sqrt{2} \approx 5.66$ is shown in Figure 10. The deviations are largest around the gap. The overall agreement, however, is excellent. In particular, the threshold exponent is $\alpha_{\mathrm{LMA}} = 1/2$.

The LMA thus provides a qualitatively correct description of the Mott–Hubbard insulator. It is a quantitatively reliable approximation down to $U \approx 5.5$.

7 Other numerical approximations to the DMFT

In this section we discuss two other exact-diagonalization approaches. In both of them a discretized version of the single-impurity Anderson model (SIAM), Section 4.1, is

investigated in which the site energies ϵ_ℓ are determined self-consistently. This is the decisive difference to our 'fixed-energy' algorithm as presented in Section 4.

In this work, we do not discuss the Numerical Renormalization Group (NRG) approach to the Mott–Hubbard insulator. The NRG is custom-tailored for the description of the metallic phase in the sense that it tries to resolve structures around $\omega = 0$ whereas features at high energies are broadened on a logarithmic scale [15]. Therefore, the Mott–Hubbard insulator does not display a clear gap. Consequently, as can be seen from Figure 2 in reference [15], the density of states is finite in the gap, and the NRG Hubbard bands display sizable high-energy tails. Therefore, the gap and the width of the Hubbard bands as a function of the interaction strength cannot be easily deduced from this approach.

7.1 Exact diagonalization in 'two-chain geometry'

In contrast to Section 4.1, an alternative formulation of the Hamilton operator for the SIAM describes two chains of length $(n_s - 1)/2$ which represent the upper and lower Hubbard bands. They are coupled to the impurity site at the respective chain origins [3,27]

$$
\begin{aligned}
\hat{H}_{\text{SIAM-tc}} = & \sum_{\ell=1;\sigma}^{(n_s-1)/2} \epsilon_\ell \left(\hat{u}_{\sigma;\ell}^+ \hat{u}_{\sigma;\ell} - \hat{l}_{\sigma;\ell}^+ \hat{l}_{\sigma;\ell} \right) \\
& + U \left(\hat{d}_\uparrow^+ \hat{d}_\uparrow - \frac{1}{2} \right) \left(\hat{d}_\downarrow^+ \hat{d}_\downarrow - \frac{1}{2} \right) \\
& + \sum_\sigma \sqrt{\frac{1}{2}} \left(\hat{u}_{\sigma;1}^+ \hat{d}_\sigma + \hat{l}_{\sigma;1}^+ \hat{d}_\sigma + \text{h.c.} \right) \\
& + \sum_{\ell=1}^{(n_s-3)/2} V_\ell \left(\hat{u}_{\sigma;\ell+1}^+ \hat{u}_{\sigma;\ell} + \hat{l}_{\sigma;\ell+1}^+ \hat{l}_{\sigma;\ell} + \text{h.c.} \right).
\end{aligned}
$$
(121)

Here particle-hole symmetry results in $\epsilon_\ell > 0$ for $\ell = 1, \ldots, (n_s - 1)/2$, and $V_\ell > 0$ for $\ell = 1, \ldots, (n_s - 3)/2$.

The main advantage of the two-chain geometry lies in the fact that on the Bethe lattice the self-consistency condition (79) can be implemented exactly even for finite n_s. In the continued-fraction expansion of the Green function for the Hubbard bands (14), the local energies ϵ_ℓ and the electron transfer amplitudes V_ℓ appear as diagonal and off-diagonal coefficients,

$$
G_{\text{LHB};\sigma}(\omega) = \frac{1/2}{\omega + U/2 - (V_1)^2/[\omega - \epsilon_1 - \ldots]}.
$$
(122)

The (dynamical) Lanczos procedure directly provides the continued-fraction coefficients so that the output of the exact diagonalization procedure yields the input for the next iteration of the self-consistency cycle without a collection of weights as, *e.g.*, in (82).

For this reason, the two-chain ED is rather appealing. Moreover, using the results for $n_s = 3, 5, 7$ and some unspecified extrapolation in $1/n_s$, the gap in [3] was
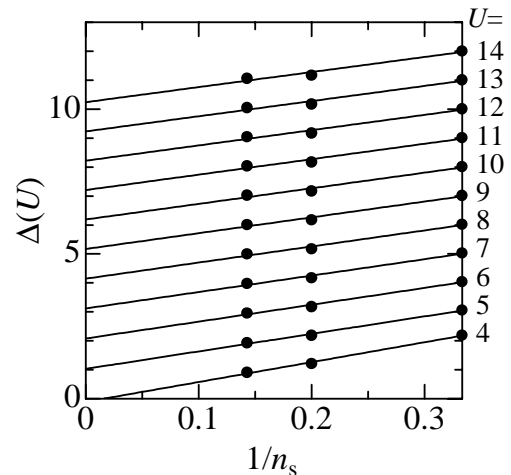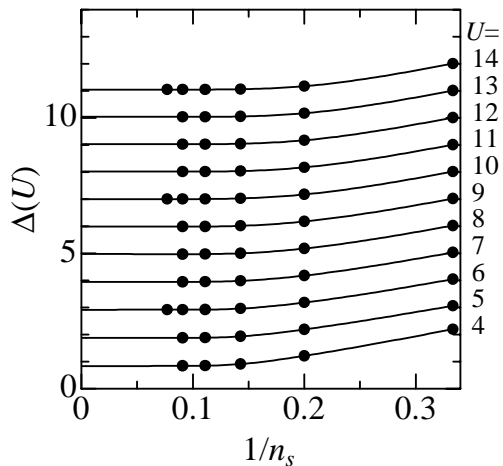

Fig. 11. Gap in exact diagonalization of the single-impurity model in two-chain geometry for $n_s = 3, 5, 7$ for various interactions U. The lines are a linear extrapolation to $1/n_s \to 0$.
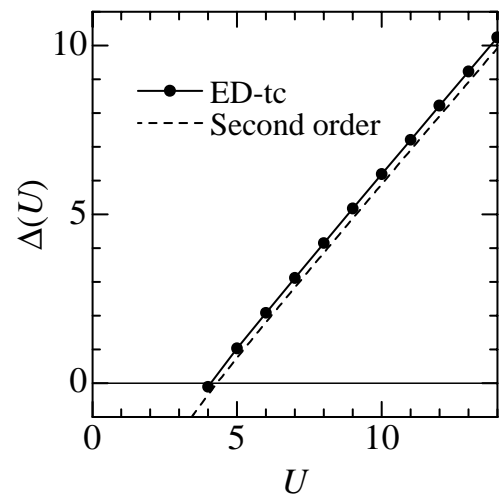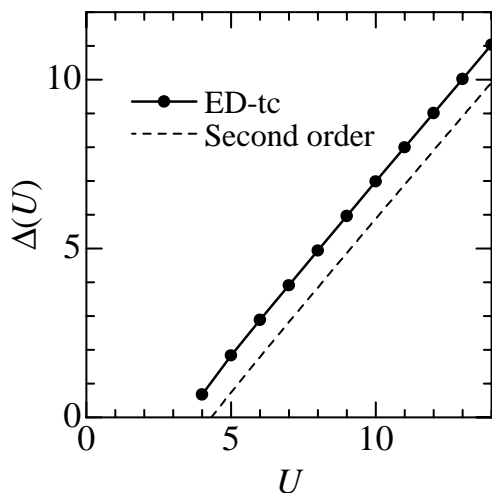

Fig. 12. Extrapolated gap (see Fig. 11) in the exact diagonalization of the single-impurity model in two-chain geometry as a function of the interaction strength, as compared to second-order perturbation theory in $1/U$ (59).

found to be almost linear as a function of the interaction strength above $\tilde{U}_c = 4.30$, in surprisingly good agreement with our best estimate (88). The corresponding results for $n_s = 3, 5, 7$ are shown in Figure 11. If we take only those three data points into account and assume a linear scaling in $1/n_s$, we obtain the gap as shown in Figure 12 which is very similar to the result reported in [3], and is in reasonable agreement with the results from the $1/U$ expansion. Thus, one would be tempted to conclude that the method is quite reliable. Unfortunately, such a conclusion is not warranted.

First, bigger system sizes need to be analyzed because $n_s = 3, 5, 7$ correspond to chain lengths of $L_c = 1, 2, 3$. When we follow the original algorithm as described in [27], we find that the insulating phase becomes unstable at $U = 8$ for $n_s = 13$ ($U = 6.4, 5.0, 3.5$ for $n_s = 11, 9, 7$). Obviously, this is a numerical artifact which is not related

Fig. 13. Gap in exact diagonalization of the single-impurity model in two-chain geometry for $n_s \leq 13$ for various interactions U. The lines are a cubic spline extrapolation to $1/n_s \to 0$.

Fig. 14. Extrapolated gap (see Fig. 13) in the exact diagonalization of the single-impurity model in two-chain geometry as a function of the interaction strength, as compared to second-order perturbation theory in $1/U$ (59).

to the metal-insulator transition. We have found that this problem can be cured by incorporating the particle-hole symmetry into $\hat{H}_{\mathrm{SIAM-tc}}$ (121) right from the start; this was done only 'on average' in reference [27]. In the following, we show results obtained using this particle-hole symmetry.

When we include bigger system sizes, we obtain a behavior quite different from the one seen in Figure 12. The result for the gap as a function of $1/n_s$ for $n \leq 13$ is shown in Figure 13. The behavior of the gap as a function of $1/n_s$ can no longer be approximated by a linear fit. The gap is not regular as a function of $1/n_s$ but seems to saturate. When we interpolate using a cubic spline, we obtain the gap as a function of the interaction strength as shown in Figure 14. Obviously, the results are rather poor as the
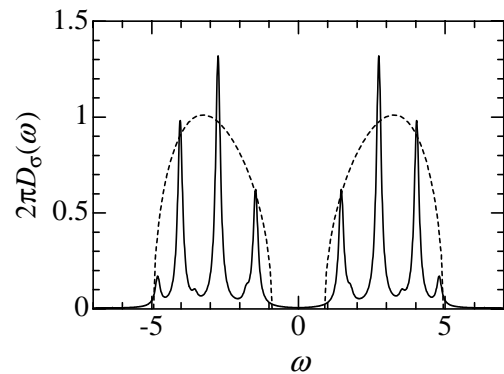
Fig. 15. Density of states as obtained from the two-chain exact diagonalization for $U = 6$ and $n_s = 13$ (full line), as compared to second-order perturbation theory in $1/U$ (55) (dashed line).

large-U limit cannot be reproduced properly and there is nowhere agreement with the results from the $1/U$ expansion. It is thus seen that the agreement between the data reported in [3,28] and ours is fortuitous.

The poor quality of the two-chain ED is also seen in the density of states, as shown in Figure 15 for $n_s = 13$ and $U = 6$. For all $n_s \geq 7$ we obtain a four-peak structure which does not allow the recovery of the true density of states.

The reason for the failure of the method lies in the setup of the self-consistency scheme. In the continued-fraction expansion one optimizes the moments of the density of states. However, the reconstruction of the density of states from its moments is a numerically very delicate inverse problem. In our case, the moments of the upper Hubbard band can be approximated very well by three main peaks and many small peaks at (much) higher energy. For this reason, the structure of the density of states is essentially unchanged as a function of n_s. The convergence as a function of system size saturates, and a reliable extrapolation scheme is not evident. We thus conclude that the two-chain ED is not a suitable tool for the investigation of the Mott–Hubbard insulator.

7.2 Caffarel–Krauth exact diagonalization in 'star geometry'

As an alternative to the two-chain discretization of the previous section, Caffarel and Krauth used the discretization of the single-impurity Anderson model in 'star geometry' (74). In contrast to our FE-ED, both the hybridization parameters V_ℓ and the site energies ϵ_ℓ are determined self-consistently in the Caffarel–Krauth scheme (CK-ED).

In order to close the self-consistency cycle, a χ^2-fitting procedure [18,26] is used. Let T be a fictitious small temperature, typically $T = 1/200$, and $\omega_n = (2n + 1)\pi T$ are the corresponding Matsubara frequencies. For a given parameter set (ϵ_ℓ, V_ℓ) we solve the single-impurity model for the single-particle Green function $G_\sigma^{(n_s)}(\omega)$. Then, the self-consistency equation (79) allows us to deduce the parameter set for the next iteration from the
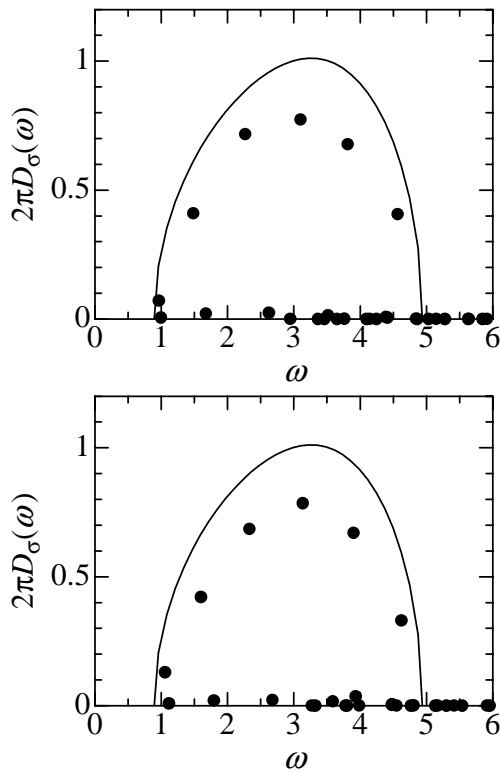
Fig. 16. Density of states for $U = 6$ as obtained from the $1/U$ expansion (solid line), in comparison with two converged solutions of the Caffarel–Krauth Exact Diagonalization scheme (CK-ED) for $n_s = 11$ sites (solid points).

minimization of

$$\chi^2(\epsilon_\ell^{\text{new}}, V_\ell^{\text{new}}) = \sum_{n=0}^{n_M} \Big| H^{(n_s)}(\mathrm{i}\omega_n; \epsilon_\ell^{\text{new}}, V_\ell^{\text{new}})$$
$$- G_\sigma^{(n_s)}(\mathrm{i}\omega_n; \epsilon_\ell, V_\ell) \Big|^2. \quad (123)$$

Here n_M is a (large) upper cut-off.

In Figure 16 we show two typical converged results for the density of states for $U = 6$, starting from two different initial values for the set (V_ℓ, ϵ_ℓ). Note that the solutions are not unique, *i.e.*, there are many self-consistent solutions for the insulator. At first glance, the two solutions appear to be very similar. The main peaks have almost the same weights V_ℓ^2 and the same positions ϵ_ℓ. This can be expected because the main peaks dominate χ^2 in (123). Moreover, the overall distribution of weights and energies by and large follows the true density of states from the analytical calculation.

A closer look into the figures reveals a fundamental problem. In Figure 16a the first non-vanishing peak appears at $\Delta_a^{(11)}(6)/2 = 0.966$ whereas in Figure 16b the value for half the gap is estimated as $\Delta_b^{(11)}(6)/2 = 1.06$. The same results follow if we use the difference of the ground-state energies (83) for the definition of the gap. Obviously, we cannot give a unique value for $\Delta^{\text{CK–ED};(n_s)}$, and an extrapolation $n_s \to \infty$ is impossible. The reason for this problem lies in the fact that the peak at the onset

of the density of states is rather small so that its precise position and weight do not matter much in the χ^2-fit (123). Thus, the onset energy may easily fluctuate by 10% or more, as seen in Figure 16. Therefore, the CK-ED approach cannot provide a reliable estimate for the gap.

The failure of the self-consistency can be traced back to the fact that, for general filling, the Green function $G^{(n_s)}(\omega)$ contains only n_s poles with substantial weight whereas $H^{(n_s)}(\omega)$ is characterized by $(2n_s)$ parameters $(\epsilon_\ell, V_\ell^2)$ in the CK-ED. Too much flexibility in the hybridization function leads to non-unique solutions of the self-consistency equations. Our Fixed-Energy Exact Diagonalization cures this problem because only n_s parameters V_ℓ^2 are to be determined.

8 Conclusions

In this work, we have studied the insulating phase of the half-filled Hubbard model on a Bethe lattice with infinite connectivity (band-width $W = 4t$). As long as the strong-coupling perturbation theory converges, the ground state of the Mott–Hubbard insulator has a finite entropy density, $s = \ln(2)$, and the density of states of the lower and upper Hubbard bands increases as a function of frequency with the edge exponent $\alpha = 1/2$, see (57). Furthermore, we find that the high-order corrections to the gap as a function of $1/U$ are fairly small. Thus, the gap opens at the critical interaction strength $U_c^{\text{sc}} = 4.40 \pm 0.09$. Our explicit results to second order in $1/U$ (55) provide a benchmark test for analytical and numerical approaches to the Mott–Hubbard insulator in infinite dimensions.

We have used our results from the $1/U$ expansion to set up a new numerical scheme for the effective single-impurity Anderson model in the Dynamical Mean-Field Theory. In our Fixed-Energy Exact Diagonalization (FE-ED), we start an outer self-consistency cycle with the position and width of the Hubbard bands from perturbation theory which we resolve equidistantly to accuracy $9t/(n_s - 1)$ for $n_s \leq 15$. The extrapolation $n_s \to \infty$ thereby becomes systematic. In an inner self-consistency cycle we determine the hybridization strengths for given n_s using a dynamical Lanczos procedure. From this we determine the gap after extrapolation to the thermodynamic limit. We merge the extrapolated gap from the 'energy criterion' and the 'weight criterion' in order to obtain a new estimate for the onset of the Hubbard bands. In this way, we iterate to convergence the outer self-consistency cycle.

Our FE-ED very well reproduces the gap, the density of states, and the exponent $\alpha = 1/2$ from perturbation theory. In this way we confirm our estimate for the transition to $U_c^{\text{FE–ED}} = 4.43 \pm 0.05$. As a last check, we favorably compare our results with those from the Random Dispersion Approximation which is an independent approach to the limit of infinite dimensions. Thus, we are confident that the strong-coupling perturbation theory, our Fixed-Energy Exact Diagonalization scheme for the Dynamical Mean-Field Theory, and the Random Dispersion Approximation provide a consistent and accurate description of the Mott–Hubbard insulator.

Other analytical and numerical techniques for the solution of the Dynamical Mean-Field Theory meet our benchmark test with limited success. The best analytical approximation is the Local Moment Approach which is quantitatively reliable for the density of states down to $U = 5.5$; it reproduces the correct exponent $\alpha = 1/2$ but overestimates the critical interaction strength, $U_c^{\mathrm{LMA}} = 4.8$. Iterated Perturbation Theory reproduces the correct threshold exponent $\alpha = 1/2$ but it becomes quantitatively and qualitatively unreliable below $U = 6$ and underestimates the stability of the Mott–Hubbard insulator, $U_c^{\mathrm{IPT}} = 5.2$.

Of the two exact-diagonalization (ED) schemes for the effective single-impurity problem existing in the literature, the two-chain approach fails because it leads to an ill-conditioned inverse problem, namely, the recovery of the density of states from its moments. The Caffarel–Krauth exact diagonalization approach (CK-ED) in star geometry gives a reasonable guess for the shape of the density of states, but it fails to provide a reliable estimate for the gap. For fixed system size, the fitting procedure results in different solutions with noticeable differences for the onset of the density of states. Therefore, it is not possible to set up a sensible extrapolation scheme for $n_s \to \infty$.

In this work we have presented a qualitatively and quantitatively reliable analysis of the Mott–Hubbard insulator. It will provide a solid basis for further studies of the Mott–Hubbard metal-insulator transition.

Appendix A: Iterated Perturbation Theory (IPT) in strong coupling

We here write the IPT self-energy, *cf.* (105–108), in a form amenable to numerical calculations and additionally find its strong coupling form. We start by rearranging (112) as follows

$$
\begin{aligned}
\mathcal{G}_\sigma(\omega) &= \mathcal{G}_\sigma^{\mathrm{band}}(\omega) + \frac{|m_0|}{\omega + \mathrm{i}\eta\,\mathrm{sgn}(\omega)} \\
&= \mathcal{G}_{\sigma,+}^{\mathrm{band}}(\omega) + \mathcal{G}_{\sigma,-}^{\mathrm{band}}(\omega) + \frac{|m_0|/2}{\omega - \mathrm{i}\eta} + \frac{|m_0|/2}{\omega + \mathrm{i}\eta} \\
&\quad + \frac{\mathrm{i}\pi}{2}|m_0|\left(\delta(\omega - 0^-) - \delta(\omega - 0^+)\right), \quad (124)
\end{aligned}
$$

where $\mathcal{G}_{\sigma,\pm}^{\mathrm{band}}(\omega)$ are the retarded/advanced components of $\mathcal{G}_\sigma^{\mathrm{band}}(\omega)$. Inserting (124) into (108), and performing a simple contour integration, yields for $\Pi(\omega)$

$$
\begin{aligned}
\Pi(\omega) &= \frac{(|m_0|/2)^2}{\omega - \mathrm{i}\eta} - \frac{(|m_0|/2)^2}{\omega + \mathrm{i}\eta} \\
&\quad + |m_0|\left\{\mathcal{G}_{\sigma,-}^{\mathrm{band}}(\omega) - \mathcal{G}_{\sigma,+}^{\mathrm{band}}(\omega)\right\} + I_1(\omega), \quad (125)
\end{aligned}
$$

where

$$
I_1(\omega) = -\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega_1}{2\pi\mathrm{i}}\mathcal{G}_\sigma^{\mathrm{band}}(\omega_1)\mathcal{G}_\sigma^{\mathrm{band}}(\omega_1 - \omega). \quad (126)
$$

Note that $\Im\Pi(\omega)$ has a pole contribution at $\omega = 0$, specifically

$$
\Pi(\omega) = \frac{\mathrm{i}\pi}{2}|m_0|^2\,\delta(\omega) + \Pi^{\mathrm{band}}(\omega), \quad (127)
$$

with

$$
\Pi^{\mathrm{band}}(\omega) = |m_0|\left\{\mathcal{G}_{\sigma,-}^{\mathrm{band}}(\omega) - \mathcal{G}_{\sigma,+}^{\mathrm{band}}(\omega)\right\} + I_1(\omega). \quad (128)
$$

Insertion of (124) and (125) into (106) yields for the self-energy

$$
\begin{aligned}
\Sigma_\sigma(\omega) &= U^2\left(\frac{|m_0|}{2}\right)^2\left\{3\mathcal{G}^{\mathrm{band}}(\omega) + \frac{|m_0|}{\omega + \mathrm{i}\eta\,\mathrm{sgn}(\omega)}\right\} \\
&\quad + I_2(\omega) + \frac{|m_0|}{2}U^2\left\{I_1^-(\omega) - I_1^+(\omega)\right\}, \quad (129)
\end{aligned}
$$

where

$$
I_2(\omega) = U^2\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega_1}{2\pi\mathrm{i}}\mathcal{G}_\sigma^{\mathrm{band}}(\omega - \omega_1)\Pi^{\mathrm{band}}(\omega_1) \quad (130)
$$

and

$$
I_1^\pm(\omega) = \pm\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega_1\Im I_1(\omega_1)}{\omega_1 - \omega \mp \mathrm{i}\eta}. \quad (131)
$$

$\Im I_1(\omega)$ may be found using a convolution involving only $\Im\mathcal{G}_\sigma^{\mathrm{band}}(\omega)$, and likewise $\Im I_2(\omega)$ may be found from a convolution of $\Im\mathcal{G}_\sigma^{\mathrm{band}}(\omega)$ and $\Im\Pi^{\mathrm{band}}(\omega)$. For $\omega > 0$,

$$
\Im I_1(\omega) = \frac{1}{\pi}\int_0^\omega \mathrm{d}\omega_1\Im\mathcal{G}_\sigma^{\mathrm{band}}(\omega_1)\Im\mathcal{G}_\sigma^{\mathrm{band}}(\omega - \omega_1), \quad (132)
$$

$$
\Im I_2(\omega) = \frac{U^2}{\pi}\int_0^\omega \mathrm{d}\omega_1\Im\mathcal{G}_\sigma^{\mathrm{band}}(\omega_1)\Im\Pi^{\mathrm{band}}(\omega - \omega_1). \tag{133}
$$

In order to calculate $\Im\Sigma_\sigma(\omega)$ for $\omega > 0$, it is thus necessary, see (129), to perform only the two convolutions (132) and (133), $\Im\Sigma_\sigma(\omega)$ for $\omega < 0$ follows immediately by symmetry, whence $\Re\Sigma_\sigma(\omega)$ follows by Hilbert transform.

The above is what we implement numerically. However, when $U \to \infty$ equation (129) simplifies. In this limit (116) gives $|m_0| \to 1$; note that $|m_0| \to 0$ is not a self-consistent possibility, since $|m_0| \geq 2/3$, *cf.* (117). Specifically,

$$
|m_0| = 1 - \frac{4}{U^2}, \quad U \to \infty. \quad (134)
$$

Since $|m_0|$ is the weight of the pole in $\mathcal{G}_\sigma(\omega)$, (112), and $\int_{-\infty}^{\infty}|\Im\mathcal{G}_\sigma(\omega)|\,\mathrm{d}\omega = \pi$, we have

$$
\Im\mathcal{G}_\sigma^{\mathrm{band}}(\omega) \sim \mathcal{O}(U^{-2}), \quad U \to \infty. \quad (135)
$$

Now, equations (132) and (133) result in $I_1(\omega) \sim \mathcal{O}(U^{-4})$ and $I_2(\omega) \sim \mathcal{O}(U^{-2})$, whereby (129) reduces to

$$
\Sigma_\sigma(\omega) = U^2\left(\frac{|m_0|}{2}\right)^2\left\{3\mathcal{G}_\sigma^{\mathrm{band}}(\omega) + \frac{|m_0|}{\omega + \mathrm{i}\eta\,\mathrm{sgn}(\omega)}\right\}, \tag{136}
$$

with corrections $\mathcal{O}(U^{-2})$. Using (112) and (107) this may be rewritten as

$$\Sigma_\sigma(\omega) = \frac{-3\,|m_0|^2\,U^2/4}{U/2 - \omega' + G_\sigma(\omega)} + \frac{U^2\,|m_0|^3/2}{U/2 - \omega'} + \mathcal{O}(U^{-2}), \tag{137}$$

where $\omega' = \omega + U/2$. In the region of the lower Hubbard band only ($\omega' \sim \mathcal{O}(1)$) equation (137) may be expanded to give

$$\Sigma_\sigma(\omega) = -\frac{U}{2} - \omega' + 3G_\sigma(\omega)$$
$$- \frac{2}{U}\left[\omega'^2 - 6\omega' G_\sigma(\omega) + 3G_\sigma(\omega)^2\right] + \mathcal{O}(U^{-2}). \tag{138}$$

If self-energy terms of $\mathcal{O}(U^{-1})$ and above are ignored, then $G_\sigma(\omega)$ in the region of the lower Hubbard band reduces to

$$G_\sigma(\omega) = \frac{1/2}{\omega' - 2G_\sigma(\omega)}, \tag{139}$$

and the single-particle density of states becomes

$$D_{\text{LHB},\sigma}(\omega) = \frac{1}{4\pi}\sqrt{4 - \omega'^2}, \tag{140}$$

i.e., the correct strong coupling spectrum of a semi-ellipse with the full noninteracting band-width. Retaining the $\mathcal{O}(U^{-1})$ terms in equation (138) gives the $\mathcal{O}(U^{-1})$ corrections to the position and shape of the LHB. The resultant spectrum is a distorted semi-ellipse of the form

$$D_{\text{LHB},\sigma}(\omega) = \sqrt{4 - \left(\omega' - \frac{a_{\text{IPT}}}{U}\right)^2 + \mathcal{O}(U^{-2})}$$
$$\times \frac{1 - (a_{\text{IPT}}/U)\omega' + \mathcal{O}(U^{-2})}{4\pi}, \tag{141}$$

where $a_{\text{IPT}} = 5/4$ and the LHB has band edges at $\omega' = (a_{\text{IPT}}/U) \pm 2$. The IPT result may be compared with the exact result to this order which also has the form of equation (141) but with a different coefficient of $a = 1/2$.

Appendix B: Local Moment Approach (LMA) in strong coupling

The LMA of Logan *et al.* is described in detail in reference [25], where it is proved that the LMA recovers the strong coupling single-particle spectrum up to and including terms of $\mathcal{O}(U^0)$. Unlike IPT, the LMA also recovers the strong coupling spectrum in the antiferromagnetic phase [25,43]. The success of the LMA in strong coupling is unsurprising since the LMA is explicitly motivated by ideas of hole motion in a spin background. In this appendix we extend the previous analysis to recover the $\mathcal{O}(U^{-1})$ corrections to the LMA self-energy and single-particle spectrum in the strong-coupling limit. This is done in three stages. First the necessary equations of the LMA are given. Secondly, it is shown that only the coupling of hole motion to zero-frequency spin-flips survive, with the coupling

of hole motion to higher energy particle-hole excitations only entering at $\mathcal{O}(U^{-2})$. Finally, the resultant self-energy is expanded to $\mathcal{O}(U^{-1})$ and the corresponding density of states is obtained.

The bare propagators within the LMA are unrestricted Hartree Fock (UHF) propagators. Lattice sites are taken to have up-spin or down-spin permanent local moments at random, and labeled 'A-type' or 'B-type' respectively. We focus solely on an A-type site (Green functions for B-type sites follow by symmetry) and suppress the 'A' labels in what follows. For the $Z = \infty$ Bethe lattice the UHF equations are simply

$$G_\sigma^0(\omega) = -G_{-\sigma}^0(-\omega) = \left[\omega + \frac{\sigma}{2}U\,|\mu| - G^0(\omega)\right]^{-1}, \tag{142}$$

where $\sigma = \pm 1$ for $\uparrow$ / $\downarrow$ respectively, and the symmetric single-particle Green function obeys $G^0(\omega) = [G_\uparrow^0(\omega) + G_\downarrow^0(\omega)]/2$. The local moment is derived from

$$\mu = \int_{-\infty}^0 \mathrm{d}\omega \left[D_\uparrow^0(\omega) - D_\downarrow^0(\omega)\right], \tag{143}$$

$$D_\sigma^0(\omega) = -\frac{1}{\pi}\text{sgn}(\omega)\Im G_\sigma^0(\omega). \tag{144}$$

The RPA on-site particle-hole propagator

$$\Pi^{+-}(\omega) = \frac{\Pi_0^{+-}(\omega)}{1 - U\Pi_0^{+-}(\omega)}, \tag{145}$$

with

$$\Pi_0^{+-}(\omega) = -\int_{-\infty}^\infty \frac{\mathrm{d}\Omega}{2\pi\mathrm{i}}G_\downarrow^0(\Omega)G_\uparrow^0(\Omega - \omega), \tag{146}$$

has a zero-frequency pole [25] (as well as higher-energy Stoner bands) reflecting the presence of zero-frequency spin-flip excitations in the system. In the full LMA single-particle Green function,

$$G_\sigma(\omega) = \left[\omega + \frac{\sigma}{2}U\,|\mu| - G(\omega) - \Sigma_\sigma(\omega)\right]^{-1}, \tag{147}$$

hole motion and accompanying spin-flips are dynamically coupled *via* the self-energy

$$\Sigma_\uparrow(\omega) = -\Sigma_\downarrow(-\omega) = U^2 \int \frac{\mathrm{d}\Omega}{2\pi\mathrm{i}}\Pi^{+-}(\Omega)\mathcal{G}_\downarrow(\omega + \Omega). \tag{148}$$

The LMA equations must be solved self-consistently, since the host Green function reads

$$\mathcal{G}_\sigma(\omega) = \left[\omega + \frac{\sigma}{2}U\,|\mu| - G(\omega)\right]^{-1} \tag{149}$$

with $G(\omega) = [G_\uparrow(\omega) + G_\downarrow(\omega)]/2$.

To calculate the LMA self-energy to $\mathcal{O}(U^{-1})$, we proceed in two stages. First we sketch the proof that only the zero-frequency spin-flip pole in $\Pi^{+-}(\omega)$ need be retained and the self-energy simplifies to

$$\Sigma_\uparrow(\omega) = U^2 \mathcal{G}_\downarrow^-(\omega) + \mathcal{O}(U^{-2}) \tag{150}$$

$$= U^2 \int_{-\infty}^0 \frac{\mathrm{d}\omega_1}{\pi}\frac{\Im\mathcal{G}_\downarrow(\omega_1)}{\omega - \omega_1 - \mathrm{i}\eta} + \mathcal{O}(U^{-2}). \tag{151}$$

Subsequently we analyze $\mathcal{G}_\downarrow^-(\omega)$ to obtain the spectrum to $\mathcal{O}(U^{-1})$.

It has been shown [25] that the RPA particle-hole propagator may be written

$$\Pi^{+-}(\omega) = \frac{Q}{-\omega - \mathrm{i}\eta} + \Pi_>^{+-}(\omega) + \Pi_<^{+-}(\omega), \qquad (152)$$

where $\Pi_>^{+-}(\omega)$ and $\Pi_<^{+-}(\omega)$ are the Stoner-like contributions whose imaginary parts are bands centered at $\pm U$ respectively. The full self-energy is thus

$$\Sigma_\uparrow(\omega) = QU^2\mathcal{G}_\downarrow^-(\omega) + \Sigma_\uparrow^{\mathrm{S}}(\omega), \qquad (153)$$

where

$$\Im\Sigma_\uparrow^{\mathrm{S}}(\omega) = \begin{cases} -U^2 \int_0^\omega \mathrm{d}\Omega \mathcal{D}_\downarrow(\Omega)\Im\Pi_<^{+-}(\Omega - \omega) & \text{for } \omega > 0 \\ U^2 \int_\omega^0 \mathrm{d}\Omega \mathcal{D}_\downarrow(\Omega)\Im\Pi_>^{+-}(\Omega - \omega) & \text{for } \omega < 0 \end{cases}. \tag{154}$$

To show that $\Sigma_\uparrow^{\mathrm{S}}(\omega) \sim \mathcal{O}(U^{-2})$, it is necessary to examine $\Pi^{+-}(\omega)$ and hence $\Pi_0^{+-}(\omega)$ and $G_\sigma^0(\omega)$. By expanding $G_\downarrow^0(\omega)$ in the region of the LHB it quickly follows that

$$|\mu| = 1 - 1/U^2 + \mathcal{O}(U^{-3}). \qquad (155)$$

Then $\Im\Pi_0^{+-}(\omega)$ may be shown from (146) to have a minority band centered around $\omega \approx -U$ with spectral weight $[(1 - |\mu|)/2]^2 = 1/(4U^4) + \mathcal{O}(U^{-5})$, and a majority band centered at $\omega \approx U$ with spectral weight $[(1 + |\mu|)/2]^2 = 1 - 1/(U^2) + \mathcal{O}(U^{-3})$. The bands of the RPA propagator which follow from (145),

$$\Im\Pi^{+-}(\omega) = \frac{\Im\Pi_0^{+-}(\omega)}{\left|1 - U\Pi_0^{+-}(\omega)\right|^2}, \qquad (156)$$

are also centered around $\omega = \pm U$. Since $\Pi_0^{+-}(\omega) \sim \mathcal{O}(1)$ in the region of the upper band it follows that the spectral weight in $\Im\Pi_>^{+-}(\omega)$ is $\mathcal{O}(U^{-2})$. In the region of the lower Hubbard band ($\omega \approx -U/2$) Kramers-Kronig relations result in $\Re\Pi_0^{+-}(\omega) = 1/(2U) + \mathcal{O}(U^{-2})$, whence $\Im\Pi^{+-}(\omega) = 4\Im\Pi_0^{+-}(\omega) + \mathcal{O}(U^{-5})$ and the spectral weight of $\Im\Pi_<^{+-}(\omega)$ is $1/U^4$. The remainder of the spectral weight of $\Im\Pi^{+-}(\omega)$ is contained in the pole weight Q. Equations (145), (146) may be used to show that

$$\Pi^{+-}(\omega \to \infty) = \Pi_0^{+-}(\omega \to \infty) = -\frac{|\mu|}{\omega}, \qquad (157)$$

whence Kramers-Kronig relations give the missing weight as $Q = 1 - \mathcal{O}(U^{-2})$. Finally, since $\mathcal{G}_\downarrow(\omega)$ (in common with $G_\downarrow^0(\omega)$) has minority and majority spectral weights that approach $1/(2U^2)$ and $1 - 1/(2U^2)$ respectively, equation (154) shows the contributions to $\Im\Sigma_\uparrow^{\mathrm{S}}(\omega)$ are $\mathcal{O}(U^{-2})$. A more detailed analysis shows that in the region of the LHB $\Sigma_\uparrow^{\mathrm{S}}(\omega)$ is pure real and $\mathcal{O}(U^{-4})$, and thus neglecting $\Sigma_\uparrow^{\mathrm{S}}(\omega)$ barely changes the LMA results throughout the insulating phase. Nevertheless the analysis outlined here is sufficient to arrive at (150).

Substituting (150) together with $|\mu| = 1 - 1/U^2 + \mathcal{O}(U^{-3})$ into (147) yields in the region of the LHB

$$G_\uparrow(\omega) = \left[\omega' - \frac{1}{2}G_\uparrow(\omega) - U^2\mathcal{G}_\downarrow^-(\omega) + \mathcal{O}(U^{-2})\right]^{-1}, \tag{158}$$

where we have defined

$$\omega' = \omega + U/2, \qquad (159)$$

and we have also used the LHB expansion

$$G(\omega) = \frac{1}{2}G_\uparrow(\omega) - \frac{1}{2U} + \mathcal{O}(U^{-2}). \qquad (160)$$

All that remains is to calculate $\mathcal{G}_\downarrow^-(\omega)$ to $\mathcal{O}(U^{-3})$. We focus exclusively on the LHB. First we expand $\mathcal{G}_\downarrow(\omega)$ in powers of $1/U$,

$$\begin{aligned} \mathcal{G}_\downarrow(\omega) = &-\frac{1}{U} - \frac{1}{U^2}\left[\omega' + \frac{1}{2U} - G(\omega)\right] \\ &- \frac{1}{U^3}\left[\omega' + \frac{1}{2U} - G(\omega)\right]^2 + \mathcal{O}(U^{-4}). \end{aligned} \tag{161}$$

Since in the region of the LHB $\Im\mathcal{G}_\downarrow^-(\omega) = \Im\mathcal{G}_\downarrow(\omega)$ by definition, we have

$$\begin{aligned} U^2\Im\mathcal{G}_\downarrow^-(\omega) &= \Im G(\omega) - \frac{1}{U}\Im\left\{[\omega' - G(\omega)]^2\right\} + \mathcal{O}(U^{-2}) \\ &= \frac{1}{2}\Im G_\uparrow(\omega) - \frac{3}{4U}\Im\left\{\omega' G_\uparrow(\omega)\right\} + \mathcal{O}(U^{-2}), \end{aligned} \tag{162}$$

where we have used (160) and the exact form of the single-particle Green function in strong coupling,

$$G_\uparrow^2(\omega) = \omega' G_\uparrow(\omega) - 1 + \mathcal{O}(U^{-1}), \qquad (163)$$

which is correctly obtained within the LMA. The Hilbert transform (151) then yields

$$\begin{aligned} U^2\mathcal{G}_\downarrow^-(\omega) &= \frac{1}{2}G_\uparrow(\omega) + \frac{3\omega'}{4U}\mathrm{i}\Im G_\uparrow(\omega) \\ &\quad + \frac{3}{4U}\mathrm{P}\int_{\mathrm{LHB}} \mathrm{d}\omega_1 \frac{\omega_1 D_\uparrow(\omega_1)}{\omega - \omega_1} + \mathcal{O}(U^{-2}) \\ &= \frac{1}{2}G_\uparrow(\omega) + \frac{3\omega'}{4U}G_\uparrow(\omega) - \frac{3}{4U} + \mathcal{O}(U^{-2}), \end{aligned} \tag{164}$$

where we have used $2\pi D_\uparrow(\omega) = \sqrt{4 - \omega'^2 + \mathcal{O}(U^{-1})}$ to simplify the $\mathcal{O}(U^{-1})$ correction. Equation (164) is only valid in the region of LHB spectral density. Substituting (164) into (158) gives the single-particle Green function in this region,

$$G_\uparrow(\omega) = \left[\omega' + \frac{3}{4U} - \left\{1 + \frac{3\omega'}{4U}\right\}G_\uparrow(\omega) + \mathcal{O}(U^{-2})\right]^{-1}. \tag{165}$$

The corresponding single-particle spectrum in strong coupling including $1/U$ corrections is

$$D_\uparrow(\omega) = \sqrt{4 - \left(\omega' - \frac{a_{\mathrm{LMA}}}{U}\right)^2 + \mathcal{O}(U^{-2})}$$

$$\times \frac{1 - (a_{\mathrm{LMA}}/U)\omega' + \mathcal{O}(U^{-2})}{2\pi} \qquad (166)$$

with the LMA coefficient $a_{\mathrm{LMA}} = 3/4$ instead of $a = 1/2$ from the strong-coupling expansion.

References

1. N.F. Mott, *Metal–Insulator Transitions*, 2nd edn. (Taylor and Francis, London, 1990)
2. F. Gebhard, *The Mott Metal–Insulator Transition* (Springer, Berlin, 1997)
3. A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
4. M. Imada, A. Fujimori, Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
5. E.H. Lieb, F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)
6. F. Gebhard, A.E. Ruckenstein, Phys. Rev. Lett. **68**, 244 (1992); F. Gebhard, A. Girndt, A.E. Ruckenstein, Phys. Rev. B **49**, 10926 (1994)
7. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
8. E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)
9. U. Brandt, C. Mielsch, Z. Phys. B **75**, 365 (1989), *ibid.* **79**, 295 (1990)
10. M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
11. R.M. Noack, F. Gebhard, Phys. Rev. Lett. **82**, 1915 (1999)
12. J. Schlipf, M. Jarrell, P.G.J. van Dongen, N. Blümer, S. Kehrein, T. Pruschke, D. Vollhardt, Phys. Rev. Lett. **82**, 4890 (1999)
13. M.J. Rozenberg, R. Chitra, G. Kotliar, Phys. Rev. Lett. **83**, 3498 (1999)
14. W. Krauth, Phys. Rev. B **62**, 6860 (2000)
15. R. Bulla, Phys. Rev. Lett. **83**, 136 (1999)
16. R. Bulla, T.A. Costi, D. Vollhardt, Phys. Rev. B **64**, 045103 (2001)
17. R. Bulla, M. Potthoff, Eur. Phys. J. B **13**, 257 (2000)
18. Y. Ōno, R. Bulla, A.C. Hewson, M. Potthoff, Eur. Phys. J. B **22**, 283 (2001)
19. G. Moeller, Q. Si, G. Kotliar, M.J. Rozenberg, D.S. Fisher, Phys. Rev. Lett. **74**, 2082 (1995)
20. D.E. Logan, P. Nozières, Phil. Trans. R. Soc. London A **356**, 249 (1998); P. Nozières, Eur. Phys. J. B **6**, 447 (1998)
21. S. Kehrein, Phys. Rev. Lett. **81**, 3912 (1998)
22. A. Georges, G. Kotliar, Phys. Rev. Lett. **84**, 3500 (2000); S. Kehrein, Phys. Rev. Lett. **84**, 3501 (2000)
23. M.J. Rozenberg, G. Kotliar, X.Y. Zhang, Phys. Rev. B **49**, 10181 (1994)
24. T. Pruschke, D.L. Cox, M. Jarrell, Europhys. Lett. **21**, 593 (1993); Phys. Rev. B **47**, 3553 (1993); M. Jarrell, J.K. Freericks, T. Pruschke, Phys. Rev. B **51**, 11704 (1995)
25. D.E. Logan, M.P. Eastwood, M.A. Tusch, J. Phys. Condensed Matter **9**, 4211 (1997)
26. M. Caffarel, W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)
27. M.J. Rozenberg, G. Möller, G. Kotliar, Mod. Phys. Lett. B **8**, 535 (1994)
28. See also Q. Si, M.J. Rozenberg, G. Kotliar, A.E. Ruckenstein, Phys. Rev. Lett. **72**, 2761 (1994); M.J. Rozenberg, G. Kotliar, H. Kajüter, G.A. Thomas, D.H. Rapkine, J.M. Honig, P. Metcalf, Phys. Rev. Lett. **75**, 105 (1995)
29. E. Kalinowski, F. Gebhard, J. Low Temp. Phys. **126**, 979 (2002)
30. P.G.J. van Dongen, F. Gebhard, D. Vollhardt, Z. Phys. B **76**, 199 (1989)
31. E. Economou, *Green's Functions in Quantum Physics*, 2nd edn. (Springer, Berlin, 1983)
32. J. Hubbard, Proc. Roy. Soc. London Ser. A **276**, 238 (1963); *ibid.* **277**, 237 (1963)
33. A.L. Fetter, J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw–Hill, New York, 1971)
34. T. Kato, Prog. Theor. Phys. **4**, 154 (1949); A. Messiah, *Quantum Mechanics*, Vol. 2 (North-Holland, Amsterdam, 1962), Chap. 16
35. M. Takahashi, J. Phys. C **10**, 1289 (1977)
36. P.W. Anderson, Phys. Rev. **115**, 2 (1959)
37. W. Metzner, P. Schmit, D. Vollhardt, Phys. Rev. B **45**, 2237 (1992); W.F. Brinkman, T.M. Rice, Phys. Rev. B **2**, 1324 (1970)
38. P.G.J. van Dongen, D. Vollhardt, Phys. Rev. Lett. **65**, 1663 (1990); P.G.J. van Dongen, Phys. Rev. B **45**, 2267 (1992)
39. M. Potthoff, preprint `cond-mat/0301137` (2003), unpublished
40. A. Georges, G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
41. M.P. Eastwood, Ph.D. thesis, Oxford University (1998), unpublished
42. J. Hubbard, Proc. Roy. Soc. London **281**, 401 (1964)
43. D.E. Logan, M.P. Eastwood, M.A. Tusch, Phys. Rev. Lett. **76**, 4785 (1997)