

Ferromagnetism in correlated electron systems: Generalization of Nagaoka's theorem

Marcus Kollar,* Rainer Strack, and Dieter Vollhardt†

Institut für Theoretische Physik C, Rheinisch-Westfälische Technische Hochschule Aachen, D-52056 Aachen, Germany

(Received 3 November 1995)

Nagaoka's theorem on ferromagnetism in the Hubbard model with one electron fewer than half filling is generalized to the case where all possible nearest-neighbor Coulomb interactions (the density-density interaction V , bond-charge interaction X , exchange interaction F , and hopping of double occupancies F') are included. It is shown that for ferromagnetic exchange coupling ($F > 0$) ground states with maximum spin are stable already at finite Hubbard interaction $U > U_c$. For nonbipartite lattices this requires a hopping amplitude $t \leq 0$. For vanishing F one obtains $U_c \rightarrow \infty$ as in Nagaoka's theorem. This shows that the exchange interaction F is important for stabilizing ferromagnetism at finite U . Only in the special case $X = t$ is the ferromagnetic state stable even for $F = 0$, provided the lattice allows the hole to move around loops.

I. INTRODUCTION

The single-band Hubbard model was originally introduced as a model for ferromagnetism of itinerant electrons.¹⁻³ The model is given by

$$\hat{H}_{\text{Hubbard}} = -t \sum_{\langle ij \rangle \sigma} (\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \text{H.c.}) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (1)$$

where $\hat{c}_{i\sigma}^\dagger$ ($c_{i\sigma}$) creates (annihilates) an electron with spin $\sigma = \uparrow, \downarrow$, $\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}$ is the number operator, and $\langle ij \rangle$ denotes nearest neighbors. This is the simplest possible correlation model for electrons on a lattice. However, rigorous evidence for itinerant ferromagnetism in this model is very limited. One of the most important results is Nagaoka's theorem,⁴ which states that if the Hubbard repulsion U is infinite, the ground state has maximum total spin on certain lattices in the case of precisely one hole. The physical mechanism behind Nagaoka's theorem is the following. If $U = \infty$, the ground state of (1) is macroscopically degenerate. This degeneracy is lifted by the motion of the hole since it is energetically favorable for it to move in a background of fully aligned spins (provided the lattice allows for motion of the hole around loops⁴). A simpler proof of Nagaoka's theorem was later given by Tasaki,⁵ who also showed that additional density-dependent interactions do not alter this result.

Several other mechanisms leading to ferromagnetism in the Hubbard model have been discussed since then.⁶ Lieb⁷ proved that the ground state is ferromagnetic for bipartite lattices with different numbers of sites in each sublattice. Mielke⁸ and Tasaki⁹ proved the stability of ferromagnetism for special lattices with flat bands.¹⁰ Recently, Müller-Hartmann¹¹ studied ferromagnetism at low particle density in dimension $d = 1$. He included next-nearest-neighbor hopping in such a way that the band has two minima. At low density, the on-site repulsion U generates a ferromagnetic exchange coupling between particles in these two pockets.

Clearly, it is still a long way to a true understanding of itinerant ferromagnetism in solids. It is quite obvious that the single-band Hubbard model is not a *generic* model for ferromagnetism. So far, either the assumption of a special kind of

hopping or of $U = \infty$, or both, was necessary to prove the stability of ferromagnetism. One may therefore ask if there exist other, simple mechanisms leading to itinerant ferromagnetism that are not contained in the Hubbard model. There are two important candidates: (i) band degeneracy, as it exists in 3d-transition metals, and (ii) a nearest-neighbor exchange interaction, which is always present in a fermionic system with a Coulomb interaction. Here we discuss only the latter, since the effect of band degeneracy will be discussed separately.¹²

In solids the exchange part of the Coulomb interaction ("Heisenberg exchange") between orbitals at neighboring sites is usually ferromagnetic. However, since their overlap is small, this interaction may be quite weak. Nonetheless, it will not be strictly zero. This direct exchange interaction, denoted by F below, provides a natural way for stabilizing ferromagnetic states.¹³ Of course, other features of the model, in particular the hopping t , the structure of the lattice, and the electron density, are also important factors concerning the stability of ferromagnetism in the ground state.

In this paper, we are concerned with rigorous criteria for ferromagnetic ground states in the presence of Heisenberg exchange F . To clarify the origin of this exchange term, it is worthwhile to review the steps that originally led to the Hubbard model, and to retain, in a systematic way, Coulomb interaction terms beyond the on-site repulsion U . This is done in Sec. II, and a model Hamiltonian with all nearest-neighbor interactions is derived. In Sec. III we state sufficient stability conditions for ferromagnetic ground states in the case of one hole in a half-filled band. In particular, it turns out that if the direct exchange is ferromagnetic ($F > 0$), and even if $F = 0$ in a special case, the on-site repulsion U need only be larger than a finite value U_c , thereby generalizing Nagaoka's theorem to finite U . The details of the proof, using a method employed previously for the case of half-filling,¹⁴ are deferred to the Appendix. Section IV contains our conclusions.

II. DERIVATION OF THE MODEL

Let us first review the derivation of effective models for metallic ferromagnetism. The general electronic model ex-

pected to describe ferromagnetic phase transitions in transition metals was introduced by Hubbard.¹ It is given by the electronic Hamiltonian¹⁵

$$\hat{H} = \sum_{ij} t_{ij}^{\alpha} \hat{c}_{i\alpha\sigma}^{\dagger} \hat{c}_{j\alpha\sigma} + \frac{1}{2} \sum_{ijmn} v_{ijmn}^{\alpha\beta\mu\nu} \hat{c}_{i\alpha\sigma}^{\dagger} \hat{c}_{j\beta\sigma'}^{\dagger} \hat{c}_{n\nu\sigma'} \hat{c}_{m\mu\sigma}. \quad (2)$$

Here, $\hat{c}_{i\alpha\sigma}^{\dagger}$ ($\hat{c}_{i\alpha\sigma}$) creates (annihilates) an electron with spin σ in a Wannier orbital α localized at site i . The first term describes the hopping between two sites i, j and contains the kinetic energy and the ionic potential $U_{\text{ion}}(\mathbf{r})$. The second term describes the (screened) Coulomb interaction between electrons, $V_{\text{ee}}(\mathbf{r}-\mathbf{r}')$.¹⁶ The matrix elements, expressed in the Wannier basis, are ($\hbar \equiv 1$)

$$t_{ij}^{\alpha} = \langle i\alpha | -\frac{1}{2m} \nabla^2 + U_{\text{ion}}(\mathbf{r}) | j\alpha \rangle, \quad (3a)$$

$$v_{ijmn}^{\alpha\beta\mu\nu} = \langle i\alpha, j\beta | V_{\text{ee}}(\mathbf{r}-\mathbf{r}') | m\mu, n\nu \rangle. \quad (3b)$$

So far no approximation was made. The Hamiltonian (2) contains infinitely many parameters. For simplicity it is therefore often assumed that the essential physics of the problem is captured by a single s band, whereby all other bands are neglected. More precisely, all other bands are *projected* onto one single *effective* s band. This approximation requires the existence of a band gap above the effective band. Then the deviation of the parameters t_{ij} and v_{ijmn} from their multiband values can be determined, in principle, by perturbation theory.

The restriction to a single s band entails considerable simplifications: Orbital indices may be dropped in Eqs. (2) and (3); furthermore, all matrix elements depend only on the separation of the lattice sites (and not on direction). Since the matrix elements are expected to fall off quickly with distance, one usually retains only the first few of them. Thus hopping is restricted to nearest-neighbor sites i and j : $-t \equiv t_{ij}$. It is also natural to assume that $U \equiv v_{iiii}$ is the largest matrix element of the Coulomb interaction. Keeping only t and U one obtains the Hubbard model, Eq. (1).

However, there are other terms that can be of appreciable size.¹ These are the *two-site* terms of the interaction: $V \equiv v_{ijij}$, $X \equiv v_{iijj}$, $F \equiv v_{ijji}$, $F' \equiv v_{iijj}$, where i and j are nearest neighbors. Keeping these terms one obtains the following single-band model:

$$\begin{aligned} \hat{H}_{\text{NN}} = & \hat{H}_{\text{Hubbard}} + V \sum_{\langle ij \rangle} \hat{n}_i \hat{n}_j + X \sum_{\langle ij \rangle \sigma} (\hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma} + \text{H.c.}) \\ & \times (\hat{n}_{i-\sigma} + \hat{n}_{j-\sigma}) + F \sum_{\langle ij \rangle \sigma \sigma'} \hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma'}^{\dagger} \hat{c}_{i\sigma'} \hat{c}_{j\sigma} \\ & + F' \sum_{\langle ij \rangle} (\hat{c}_{i\uparrow}^{\dagger} \hat{c}_{i\downarrow}^{\dagger} \hat{c}_{j\downarrow} \hat{c}_{j\uparrow} + \text{H.c.}), \end{aligned} \quad (4)$$

where \hat{H}_{Hubbard} is defined in (1) and $\hat{n}_i = \sum_{\sigma} \hat{n}_{i\sigma}$. Here V is the density-density interaction between nearest neighbors, X is the bond-charge interaction giving rise to correlated hopping, F is the exchange interaction discussed in the Introduction (ferromagnetic in nature if $F > 0$), and F' repre-

sents hopping of double occupancies. We note that the F term in (4) can be expressed in terms of spin operators as a Heisenberg interaction,

$$F \sum_{\langle ij \rangle \sigma \sigma'} \hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma'}^{\dagger} \hat{c}_{i\sigma'} \hat{c}_{j\sigma} = -2F \sum_{\langle ij \rangle} \left(\hat{\mathbf{S}}_i \hat{\mathbf{S}}_j + \frac{1}{4} \hat{n}_i \hat{n}_j \right), \quad (5)$$

where $\hat{\mathbf{S}}_i = \frac{1}{2} \sum_{\sigma \sigma'} \hat{c}_{i\sigma}^{\dagger} \boldsymbol{\tau}_{\sigma \sigma'} \hat{c}_{i\sigma'}$ and $\boldsymbol{\tau}$ are the Pauli matrices.

While the on-site interaction U usually has the largest numerical value, the other matrix elements are certainly not zero. Hubbard's estimates¹ for transition metals are, for example, $U \approx 10$ eV, $V \approx 2-3$ eV, $X \approx 1$ eV, and $F, F' \approx \frac{1}{40}$ eV, and the hopping amplitude t typically ranges between 0.5 and 1.5 eV. Even if nearest-neighbor interactions are very small, they can be qualitatively important if they have different symmetries than the U term and thus can lift degeneracies.

The model (4) was essentially derived already by Hubbard.¹ Extensions of the actual Hubbard model (1) by some or all of the terms in Eq. (4) have received much attention since then. For example, Campbell, Gammel, and Low¹⁷ presented a detailed investigation of the phase diagram of \hat{H}_{NN} in dimension $d=1$, and discussed the relative magnitude of its parameters for real materials. On a mean-field level, the effect of the terms in (4) on the stability of ferromagnetism was studied by Hirsch.¹⁸ Furthermore, exact solutions are possible in the special case of $X=t$. In this case the number of doubly occupied sites is a conserved quantity, and the exact ground state solution can be obtained in a wide range of parameters.¹⁹⁻²² For $X=t$ and $V=F=F'=0$ the model was recently solved exactly in one dimension,^{23,24} while for $X=t=-V=F=F'$ a solvable supersymmetric model is obtained.^{25,26} The case $X=t$ will play a special role in our analysis, too.

Criteria for the stability of ferromagnetic ground states of the Hamiltonian (4) were recently derived for the case of half-filling (one electron per site).^{14,19-22} The ferromagnetic states are then found to be insulating. To gain insight into the more general problem of *itinerant* ferromagnetism we will now investigate a half-filled band *with one hole*, as in Nagaoka's work.⁴ Thus we consider a finite lattice with L sites and fix the total number of particles at $N=L-1$. The number of nearest neighbors is denoted by Z . We consider lattices with at least Z nearest-neighbor bonds between any subset of lattice sites and the set of remaining sites. For example, all crystal lattices with periodic boundary conditions fulfill this requirement.

III. FERROMAGNETIC GROUND STATES

The Hamiltonian \hat{H}_{NN} commutes with the total spin $\hat{\mathbf{S}} = \sum_i \hat{\mathbf{S}}_i$. The eigenvalues of $\hat{\mathbf{S}}^2$ are denoted by $S(S+1)$. In the following we will be concerned only with saturated ferromagnetic states with largest possible eigenvalue $S_{\text{max}} \equiv N/2 = (L-1)/2$. There are $2S_{\text{max}} + 1 = L$ such states with the same energy eigenvalue.

We are interested in the following question: *Under which circumstances do the ground states of \hat{H}_{NN} have maximum spin?* For the pure Hubbard model (i.e., $V=X=F=F'=0$), Nagaoka's theorem⁴ states that for $U=\infty$, $t < 0$ ($t \neq 0$ if the lattice is bipartite) the ground states have $S=S_{\text{max}}$.

TABLE I. Sufficient conditions for ferromagnetic ground states with one hole.

Case		Condition on U	Condition on lattice	Condition on t	
1a	$F > 0, X \neq t$	$U > U_c^{(1)}$	any lattice	bipartite lattice:	t arbitrary
				nonbipartite lattice:	$t \leq 0$
1b	$F > 0, X \neq t$	$U > U_c^{(2)}$	any lattice	bipartite lattice:	t arbitrary
				nonbipartite lattice:	$t \leq 0$
2	$F > 0, X \neq t$	$U > U_c^{(2)}$	lattice with loops	bipartite lattice:	$t \neq 0$
				nonbipartite lattice:	$t < 0$

This statement can be generalized to arbitrary density-density interaction V .⁵ These results require the lattice to satisfy a certain connectivity condition.²⁷ As pointed out by Tasaki,⁵ this connectivity condition is fulfilled if the lattice has a certain loop structure. Therefore we will refer to such lattices as “lattices with loops.” For example, these include the square, triangular, simple cubic, body-centered-cubic, face-centered-cubic, and hexagonal close-packed lattices, but not the one-dimensional or the Bethe lattice.⁴

The main result of this paper is a generalization of Nagaoka’s theorem to finite values of the Hubbard interaction U . In the Appendix we derive the following conditions for ferromagnetic ground states.

The ground states of \hat{H}_{NN} with one hole (i.e., $N=L-1$) have maximum total spin $S=S_{\text{max}}=(L-1)/2$ and are non-degenerate [apart from $(2S_{\text{max}}+1)$ -fold spin degeneracy] in the following cases.

Case 1. On any lattice, if $F > 0$, $t \leq 0$, and (a) $X \neq t$ and $U > U_c^{(1)}$, or (b) $X = t$ and $U \geq U_c^{(2)}$.

Case 2. On lattices with loops, if $X = t < 0$, $F = 0$, and $U > U_c^{(2)}$.

In both cases $t > 0$ is allowed if the lattice is bipartite.

These results are summarized in Table I. The constants $U_c^{(1)}$ and $U_c^{(2)}$ are given by

$$U_c^{(1)} = Z \left(2|t| + \left| V - F - 2|t| \right| + \frac{(X-t)^2}{F} + \left| F' - \frac{(X-t)^2}{F} \right| \right), \quad (6a)$$

$$U_c^{(2)} = Z \left(2|t| + \left| V - \frac{F}{2} - 2|t| \right| + |F'| \right). \quad (6b)$$

Hence, if $F > 0$, ferromagnetic ground states are stable on any lattice for U larger than a *finite* critical value. For $F \rightarrow 0^+$ we have $U_c^{(1)} \rightarrow \infty$, thus yielding Nagaoka’s condition for the pure Hubbard model. This shows that the Heisenberg interaction F , which is neglected in the Hubbard model, provides an obvious mechanism for stabilizing ferromagnetic ground states at finite U . Note that since X and t are expected to be of the same order of magnitude, the sensitive dependence on F , due to the term $(X-t)^2/F$, may cancel from $U_c^{(1)}$, and values of the order of $U_c \sim 12$ eV are possible. The dependence of U_c on t, V, F is depicted in Fig. 1. The case $X=t$ is special, since in this case the stability of ferromagnetism can be achieved either by $F > 0$ or by $F \geq 0$ and $t < 0$ if the lattice has loops.

The critical couplings $U_c^{(1)}$ and $U_c^{(2)}$ are sums of terms, each of which corresponds to a typical energy scale. This means that the on-site interaction U has to be larger than the

energy describing the paramagnetic state (bandwidth $\sim Z|t|$), as well as the threshold energies for the onset of a charge-density wave or phase separation ($\sim Z|V|$), η -pairing superconductivity²¹ ($\sim Z|F'|$), and a spin-density wave [$\sim (X-t)^2/F$]. Note, however, that these terms do not enter separately, but appear in combinations; i.e., the effects interfere as should be expected.

We remark that the above conditions are *sufficient* conditions. The occurrence of ground states with maximum spin outside the above parameter region is not ruled out. For example, it may still be possible to find ferromagnetic ground states even for $X \neq t$, $F = 0$, and $U < \infty$. It is interesting to note that the values for U_c appearing in Eq. (6) are the same as those for the case of half-filling (no hole).¹⁴ The bound U_c for half-filling was recently improved by de Boer and Schadschneider.²²

As shown in the Appendix, the ferromagnetic ground states are the same as those discussed by Nagaoka;⁴ i.e., the wave function with $\hat{S}^z = S_{\text{max}}$ corresponds to a band filled with spin-up electrons, with the hole at the top of the band, as illustrated in Fig. 2. If $t < 0$, the band maximum is at the origin, and the corresponding wave function is $|\psi_0\rangle = \hat{a}_{\mathbf{0}\uparrow}|\uparrow\rangle$, where

$$\hat{a}_{\mathbf{k}\sigma} = \frac{1}{\sqrt{L}} \sum_i \exp(i\mathbf{k}\mathbf{R}_i) \hat{c}_{i\sigma},$$

and

$$|\uparrow\rangle = \prod_i \hat{c}_{i\uparrow}^\dagger |0\rangle = \prod_{\mathbf{k}} \hat{a}_{\mathbf{k}\uparrow}^\dagger |0\rangle$$

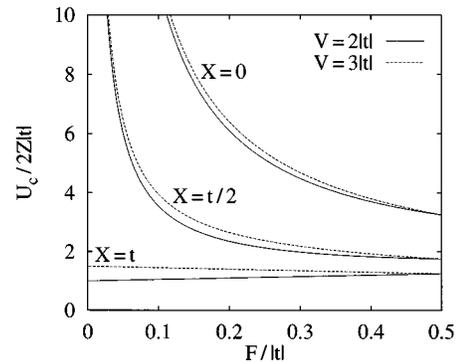


FIG. 1. Critical value U_c vs exchange interaction F for different t, V, X , and $F' = 0$. For $U > U_c$ the ground state is ferromagnetic. See Table I for details.

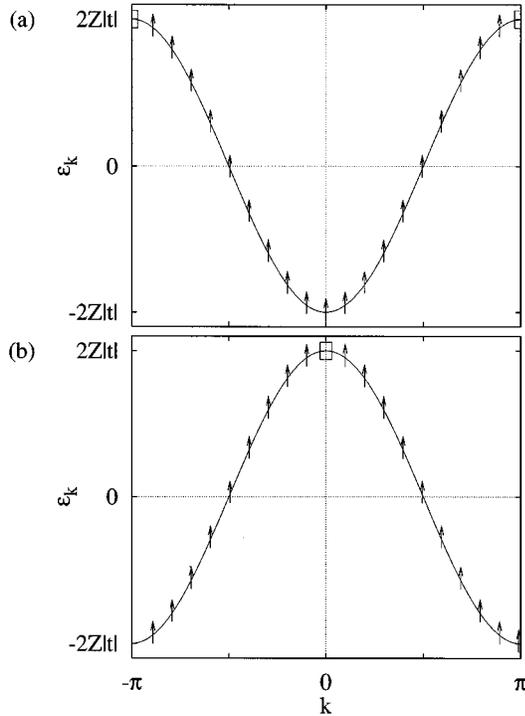


FIG. 2. The figure illustrates (for $d=1$) that the ferromagnetic ground state corresponds to a band filled with spin-up electrons (indicated by arrows), from which a single electron has been removed at the top of the band (open square). (a) $t > 0$ (with a hole at $k = \pi$), (b) $t < 0$ (with a hole at $k = 0$).

is the filled spin-up band. For $t > 0$ a bipartite lattice is required and the band maximum is at wave vector \mathbf{Q} defined by $\epsilon_{\mathbf{k}+\mathbf{Q}} = -\epsilon_{\mathbf{k}}$; e.g., for a hypercubic lattice we have $\mathbf{Q} = (\pi, \pi, \dots, \pi)$. In this case the wave function is $|\psi_0\rangle = \hat{a}_{\mathbf{Q}\uparrow}|\uparrow\rangle$. All $2S_{\max} + 1$ ground states can be obtained from $|\psi_0\rangle$ by global $SU(2)$ rotations; i.e., the state

$$|\psi_M\rangle = (\hat{S}^x - i\hat{S}^y)^M |\psi_0\rangle \quad (7)$$

has $\hat{S}^z = S_{\max} - M$, where $M = 0, \dots, L-1$. The states $\{|\psi_M\rangle\}$ have the same energy eigenvalue.

If $F > 0$, our criteria permit ferromagnetism in any spatial dimension. In particular they hold in dimension $d=1$. This is not in conflict with the Lieb-Mattis theorem²⁸ which rules out ferromagnetism for any continuum model with k^2 dispersion and a symmetric interaction potential in $d=1$. In a Bloch basis a k^2 dispersion corresponds to infinitely many bands as in (2), whereas the Hamiltonian \hat{H}_{NN} in (4) is obtained by a projection onto *one* of these bands, as discussed in Sec. II. Hence the model (4) is quite different from the initial multiband Hamiltonian \hat{H} in (2); in particular there does not exist a potential energy in a continuum model that gives rise to exactly the terms appearing in the truncated Hamiltonian \hat{H}_{NN} . Therefore the Lieb-Mattis theorem does not apply to the single-band Hamiltonian (4) in $d=1$. In the Appendix of Ref. 28 Lieb and Mattis proved a second theorem, using the occupation number formalism, which precludes ferromagnetism in $d=1$ in the case of next-neighbor hopping and purely density-dependent (i.e., site-diagonal) in-

teractions. Therefore this second theorem also does not apply to the Hamiltonian \hat{H}_{NN} , Eq. (4), where the *nondiagonal* exchange interaction $F > 0$ (which is always present as part of the Coulomb interaction) plays a crucial role. The role of nondiagonal interactions for the stability of ferromagnetism was already discussed in Ref. 28.

IV. CONCLUSIONS

We presented a generalization of Nagaoka's theorem to a Hubbard model with all nearest-neighbor interactions. For this model, with one hole in a half-filled band, we derived rigorous, sufficient conditions for the stability of saturated ferromagnetism in the ground state. The ferromagnetic ground state is found to be stable on any lattice provided the next-neighbor exchange F is ferromagnetic and the Hubbard repulsion U is larger than a critical value $U_c < \infty$, with $U_c \sim 1/F$ for $F \rightarrow 0^+$. If $F=0$, only the special case $X=t$ (unlikely to be fulfilled exactly in real materials) yields a ferromagnetic ground state on lattices with loops.

The ferromagnetic ground state is an itinerant state with nonzero kinetic energy. The proof of its stability cannot be easily extended to doping beyond a single hole. Of course, this would be highly desirable since a single hole is irrelevant in the thermodynamic limit. However, the ground states of the model with next-neighbor interactions are difficult to obtain for finite hole densities, since in this case simple eigenstates of the Hamiltonian are not known, such that the present methods can be applied only in special cases.¹²

To be able to explain ferromagnetism in more detail, it is clear that more ingredients are needed than those contained in the single-band Hubbard model with nearest-neighbor hopping. Of greatest interest is the case of band degeneracy where the present methods can be applied, too.¹²

ACKNOWLEDGMENTS

The authors would like to acknowledge valuable discussions with D. Baeriswyl, E. H. Lieb, D. C. Mattis, E. Müller-Hartmann, and A. Schadschneider. M.K. and R.S. gratefully acknowledge financial support of the Studienstiftung des Deutschen Volkes. This work was supported in part by the Sonderforschungsbereich 341 of the Deutsche Forschungsgemeinschaft.

APPENDIX: DETAILS OF THE DERIVATION

To derive the sufficient conditions for ferromagnetic ground states stated in Sec. III we rearrange the Hamiltonian (4) as a sum of positive semidefinite terms. This enables us to construct a lower bound on the ground state energy. If this lower bound coincides with the eigenvalue of a trial state, this state is an exact ground state of (4).^{29,30} Here we consider a wave function of the form

$$|\psi_0\rangle = \sum_j a_j \hat{c}_{j\uparrow} |\uparrow\rangle,$$

where $|\uparrow\rangle = \prod_i \hat{c}_{i\uparrow}^\dagger |0\rangle$ is the state with all sites occupied by spin-up electrons. The coefficients $a_j \neq 0$ will be determined below. $|\psi_0\rangle$ is a state with $N = L - 1$ particles. Since it is an

eigenstate of \hat{S}^z with eigenvalue S_{\max} , it is a representative of the subspace with $S=S_{\max}$. It is sufficient to consider a state of this form, because whenever this state is a ground state, so are its $(2S+1)=L$ global SU(2) rotations.

We introduce the operators

$$\hat{P}_{ij\sigma} = (1 - \hat{n}_{i-\sigma})(\hat{c}_{i\sigma} + \lambda_1 \hat{c}_{j\sigma})(1 - \hat{n}_{j-\sigma}),$$

$$\hat{Q}_{ij\sigma} = \hat{n}_{i-\sigma}(\hat{c}_{i\sigma} + \lambda_1 \hat{c}_{j\sigma})\hat{n}_{j-\sigma},$$

$$\hat{A}_{ij} = \alpha^{-1}(\hat{c}_{i\downarrow}\hat{c}_{i\uparrow} + \hat{c}_{j\downarrow}\hat{c}_{j\uparrow}) + \alpha\lambda_2(\hat{c}_{j\downarrow}\hat{c}_{i\uparrow} + \hat{c}_{i\downarrow}\hat{c}_{j\uparrow}),$$

$$\hat{B}_{ij} = \hat{c}_{i\downarrow}\hat{c}_{i\uparrow} + \lambda_3\hat{c}_{j\downarrow}\hat{c}_{j\uparrow}.$$

Here $\lambda_1 = \text{sgn}(t)$, $\lambda_2 = \text{sgn}(X-t)$, $\lambda_3 = \text{sgn}(F' - |X-t|/\alpha^2)$, and $\alpha \neq 0$ is an arbitrary, real parameter. Furthermore, we introduce

$$\hat{\Omega}_{ij} = \begin{cases} \hat{e}_i\hat{e}_j + 3\hat{d}_i\hat{d}_j + \hat{p}_i\hat{d}_j + \hat{p}_j\hat{d}_i & \text{if } V > 2|t| + (F + \alpha^2|X-t|)/2, \\ \frac{1}{2}(\hat{p}_i - \hat{p}_j)^2 + 2(\hat{d}_i\hat{e}_j + \hat{d}_j\hat{e}_i) & \text{if } V < 2|t| + (F + \alpha^2|X-t|)/2, \end{cases}$$

where $\hat{e}_i = (1 - \hat{n}_{i\uparrow})(1 - \hat{n}_{i\downarrow})$, $\hat{p}_i = (\hat{n}_{i\uparrow} - \hat{n}_{i\downarrow})^2$, and $\hat{d}_i = \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}$ are the projectors onto an empty, singly, and doubly occupied site, respectively. It is straightforward to verify that \hat{H}_{NN} can be written as

$$\hat{H}_{\text{NN}} = \sum_{\langle ij \rangle} \left[|t| \sum_{\sigma} (\hat{P}_{ij\sigma} \hat{P}_{ij\sigma}^{\dagger} + \hat{Q}_{ij\sigma} \hat{Q}_{ij\sigma}^{\dagger}) + |X-t| \hat{A}_{ij}^{\dagger} \hat{A}_{ij} + |\tilde{F}'| \hat{B}_{ij}^{\dagger} \hat{B}_{ij} \right] + |\tilde{V}| \sum_{\langle ij \rangle} \hat{\Omega}_{ij} + \tilde{U} \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} - 2\tilde{F} \sum_{\langle ij \rangle} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j, \quad (\text{A1})$$

where

$$\tilde{F}' = F' - \frac{|X-t|}{\alpha^2}, \quad (\text{A2a})$$

$$\tilde{F} = F - \alpha^2|X-t|, \quad (\text{A2b})$$

$$\tilde{V} = V - \frac{F + \alpha^2|X-t|}{2} - 2|t|, \quad (\text{A2c})$$

$$\tilde{U} = U - Z \left(2|t| + |\tilde{V}| + \frac{|X-t|}{\alpha^2} + |\tilde{F}'| \right), \quad (\text{A2d})$$

and an overall constant was dropped.

Let us consider the terms in (A1) one by one. The first term is positive semidefinite. Since $\hat{A}_{ij}|\psi_0\rangle = 0$, $\hat{B}_{ij}|\psi_0\rangle = 0$, and $\hat{Q}_{ij\sigma}|\psi_0\rangle = 0$, the action of this term on $|\psi_0\rangle$ is given by

$$\begin{aligned} |t| \sum_{\langle ij \rangle \sigma} \hat{P}_{ij\sigma} \hat{P}_{ij\sigma}^{\dagger} |\psi_0\rangle &= |t| \sum_{\langle ij \rangle m \sigma} (1 - \hat{n}_{i-\sigma})(1 - \hat{n}_{j-\sigma})(\hat{c}_{i\sigma} + \lambda_1 \hat{c}_{j\sigma})(\hat{c}_{i\sigma}^{\dagger} + \lambda_1 \hat{c}_{j\sigma}^{\dagger}) a_m \hat{c}_{m\uparrow} |\uparrow\rangle \\ &= |t| \sum_{\langle ij \rangle \sigma} (\hat{c}_{i\uparrow} + \lambda_1 \hat{c}_{j\uparrow})(\hat{c}_{i\uparrow}^{\dagger} + \lambda_1 \hat{c}_{j\uparrow}^{\dagger})(a_i \hat{c}_{i\uparrow}^{\dagger} + a_j \hat{c}_{j\uparrow}^{\dagger}) |\uparrow\rangle \\ &= |t| \sum_{\langle ij \rangle \sigma} (a_i + \lambda_1 a_j)(\hat{c}_{i\uparrow} + \lambda_1 \hat{c}_{j\uparrow}) |\uparrow\rangle. \end{aligned}$$

Thus, if $a_i = -\text{sgn}(t)a_j$ for all nearest neighbors $\langle ij \rangle$, we find that $|\psi_0\rangle$ has zero eigenvalue and hence is a ground state of this term. Let us consider the case $t < 0$ first. Then we must have $a_i = a_j$, and the resulting wave function is

$$|\psi_0\rangle = \frac{1}{\sqrt{L}} \sum_i \hat{c}_{i\uparrow} |\uparrow\rangle = \hat{a}_{\mathbf{k}=\mathbf{0}\uparrow} |\uparrow\rangle,$$

where

$$\hat{a}_{\mathbf{k}\sigma} = \frac{1}{\sqrt{L}} \sum_i \exp(i\mathbf{k}\mathbf{R}_i) \hat{c}_{i\sigma}.$$

Note that the filled spin-up band can be expressed as $|\uparrow\rangle = \prod_{\mathbf{k}} \hat{a}_{\mathbf{k}\uparrow} |0\rangle$. The (noninteracting) band structure is given by $\epsilon_{\mathbf{k}} = -t \sum_{\langle ij \rangle} \exp[i\mathbf{k}(\mathbf{R}_i - \mathbf{R}_j)]$. Thus for $t < 0$ the hole at $k=0$ is created at the *band maximum*, where $\epsilon_{\mathbf{k}=\mathbf{0}} = Z|t|$. This is shown in Fig. 2(a).

Next consider $t > 0$. Then $a_i = -a_j$ must hold for all nearest neighbors $\langle ij \rangle$. This is only possible if the lattice is bipartite, i.e., if any two nearest neighbors are located on different sublattices \mathcal{A} and \mathcal{B} . In this case the wave function becomes

$$|\psi_0\rangle = \frac{1}{\sqrt{L}} \left(\sum_{i \in \mathcal{A}} \hat{c}_{i\uparrow} - \sum_{i \in \mathcal{B}} \hat{c}_{i\uparrow} \right) |\uparrow\rangle = \hat{a}_{\mathbf{k}=\mathbf{Q}\uparrow} |\uparrow\rangle. \quad (\text{A3})$$

Here \mathbf{Q} is defined by $\exp[i\mathbf{Q}(\mathbf{R}_i - \mathbf{R}_j)] = -1$ for nearest neighbors $\langle ij \rangle$ or, equivalently, by $\epsilon_{\mathbf{k}+\mathbf{Q}} = -\epsilon_{\mathbf{k}}$. This implies that the band maximum is at $\mathbf{k}=\mathbf{Q}$ for $t>0$, and this is the state in which the hole is created, as shown in Fig. 2(b). The two cases are related by a gauge transformation $\hat{c}_{i\sigma} \rightarrow -\hat{c}_{i\sigma}$ on one of the sublattices, which effectively reverses the sign of t .

From now on we choose the coefficients a_i as just described, $|\psi_0\rangle$ thus being a ground state of the first term in (A1). If the lattice is not bipartite, this requires $t \leq 0$. (The trivial case $t=0$ can also be included here.)

Turning to the positive semidefinite Ω term in (A1), we observe that for $\tilde{V} \geq 0$ it annihilates $|\psi_0\rangle$, which in this case is a ground state of this term. Now consider $\tilde{V} < 0$, in which case $|\psi_0\rangle$ is an eigenstate of $\sum_{\langle ij \rangle} \hat{\Omega}_{ij}$ with eigenvalue $Z/2$. We now show that this is the lowest possible eigenvalue for $N=L-1$. Consider states with a fixed configuration $\{p_i=0,1\}$ of P singly occupied sites, $P = \sum_i p_i$. The possible (integer) values of P are $0 \leq P \leq L$.

(i) $P=L$. This case is impossible for $N=L-1$.

(ii) $1 \leq P \leq L-1$. In this case we have

$$\Omega \equiv \left\langle \sum_{\langle ij \rangle} \hat{\Omega}_{ij} \right\rangle \geq \frac{1}{2} \sum_{\langle ij \rangle} (p_i - p_j)^2 = \frac{1}{2} \sum_{\langle ij \rangle} (1 - \delta_{p_i p_j});$$

i.e., 2Ω is bound from below by the number of bonds $\langle ij \rangle$ with $p_i \neq p_j$. Consider the set of sites with $p_i=1$. By our definition of a lattice at the end of Sec. II, these sites are connected to at least Z sites with $p_j=0$. Hence $\Omega \geq Z/2$.

(iii) $P=0$. Then, since $N=L-1$, the number of lattice sites L must necessarily be odd, and there must be $(L+1)/2$ empty and $(L-1)/2$ doubly occupied sites. (There are no singly occupied sites.) Consider states with a fixed configuration of doubly occupied sites $\{d_i=0,1\}$. We have

$$\Omega \geq 2 \sum_{\langle ij \rangle} [d_i(1-d_j) + d_j(1-d_i)] = 2 \sum_{\langle ij \rangle} (1 - \delta_{d_i d_j}).$$

Except for a factor of 4, this is the same problem as above, with $\{p_i\}$ replaced by $\{d_i\}$. Therefore in this case $\Omega \geq 2Z > Z/2$.

Thus, summarizing the cases (i)–(iii), we obtain $\Omega \geq Z/2$. Hence $|\psi_0\rangle$ is a ground state of the Ω term for any \tilde{V} , since it is always an eigenstate with the lowest possible eigenvalue.

Finally, $|\psi_0\rangle$ is clearly a ground state of the remaining terms in (A1) if $\tilde{U} \geq 0$ and $\tilde{F} \geq 0$, since it has no doubly occupied sites and maximum spin.

So far we proved that $|\psi_0\rangle$ and its global $SU(2)$ rotations are among the ground states of \hat{H}_{NN} if $\tilde{U} \geq 0$ and $\tilde{F} \geq 0$ (and $t \leq 0$ if the lattice is not bipartite). Clearly, if $\tilde{F} > 0$ (and $\tilde{U} \geq 0$), these are the only ground states, since then only states with maximum spin minimize the \tilde{F} term. This will be used below, when we prove that *all* ground states of \hat{H}_{NN} have $S=S_{\text{max}}$ in the cases listed in Table I.

Case 1a. $X \neq t$, $F > 0$, $U > U_c^{(1)}$. In this case we choose $\alpha^2 \equiv (1-\epsilon)F/|X-t| > 0$ with $0 < \epsilon \leq \frac{1}{2}$ to be specified later. Then \tilde{U} in Eq. (A2) becomes

$$\tilde{U} = U - Z \left[2|t| + \left| V - F - 2|t| + \frac{\epsilon F}{2} \right| + \max \left(F', \frac{2(X-t)^2}{(1-\epsilon)F} - F' \right) \right].$$

Using $(1-\epsilon)^{-1} \leq 1 + 2\epsilon$ we obtain the bound

$$\tilde{U} \geq U - Z \left[2|t| + \left| V - F - 2|t| \right| + \max \left(F', \frac{2(X-t)^2}{F} - F' \right) + \epsilon \left(\frac{F}{2} + \frac{4(X-t)^2}{F} \right) \right] = U - U_c^{(1)} - \frac{\epsilon Z}{2F} [F^2 + 8(X-t)^2], \quad (\text{A4})$$

where $U_c^{(1)}$ is defined in Eq. (6). Since $U - U_c^{(1)} > 0$ by assumption, it can be seen from Eq. (A4) that \tilde{U} is positive if ϵ is chosen small enough. For example,

$$\epsilon \equiv \min \left(\frac{F(U - U_c^{(1)})}{Z[F^2 + 8(X-t)^2]}, \frac{1}{2} \right)$$

indeed yields $\tilde{U} > 0$. Furthermore, from Eq. (A2), $\tilde{F} = \epsilon F > 0$. Hence the present choice of α^2 yields $\tilde{U} > 0$ and $\tilde{F} > 0$. Therefore only states with $S=S_{\text{max}}$ are ground states of \hat{H}_{NN} .

Case 1b. $X=t$, $F > 0$, $U \geq U_c^{(2)}$, with $U_c^{(2)}$ as defined in Eq. (6). In this case $\tilde{U} = U - U_c^{(2)} \geq 0$ and $\tilde{F} = F > 0$. Again, only states with $S=S_{\text{max}}$ are ground states of \hat{H}_{NN} .

Case 2. $X=t < 0$, $F=0$, $U > U_c^{(2)}$, for a lattice with loops. (In the case of a bipartite lattice the above-mentioned gauge transformation may be applied to include the case $t > 0$.)

We know that $|\psi_0\rangle$ is a ground state of \hat{H}_{NN} for $\tilde{U} > 0$ and that it is lower in energy than any state with doubly occupied sites. Hence no ground state of \hat{H}_{NN} can have doubly occupied sites. Among the states without double occupancies those with $S=S_{\text{max}}$ are the lowest in energy; this follows from the P term in (A1) by a proof completely analogous to that of Tasaki.⁵ In Nagaoka's basis,²⁷ all off-diagonal matrix elements of the P term are negative. Furthermore, it is represented by an irreducible matrix since the lattice has loops; i.e., it satisfies the connectivity condition defined in Ref. 27. It follows directly from the Perron-Frobenius theorem³¹ that, in every sector with fixed \hat{S}^z , the ground state is unique and is given by a linear combination with strictly positive coefficients of the basis vectors. These ground states are just the states $\{|\psi_M\rangle\}$ of Eq. (7) and all have $S=S_{\text{max}}$.

*Electronic address: mkollar@physik.rwth-aachen.de.

†Electronic address: vollha@physik.rwth-aachen.de.

¹J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963).

²M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).

³J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963).

⁴Y. Nagaoka, Solid State Commun. **3**, 409 (1965); Phys. Rev. **147**, 392 (1966). A related mechanism is discussed by D. J. Thouless, Proc. Phys. Soc. London **86**, 893 (1965).

- ⁵H. Tasaki, Phys. Rev. B **40**, 9192 (1989).
- ⁶For a review see E. H. Lieb, in *The Hubbard Model: Its Physics and Mathematical Physics*, edited by D. Baeriswyl *et al.* (Plenum, New York, 1995).
- ⁷E. H. Lieb, Phys. Rev. Lett. **62**, 1201 (1989).
- ⁸A. Mielke, J. Phys. A **24**, 3311 (1991); **25**, 4335 (1992); Phys. Lett. A **174**, 443 (1993).
- ⁹H. Tasaki, Phys. Rev. Lett. **69**, 1608 (1992); **73**, 1158 (1994).
- ¹⁰For a review see A. Mielke and H. Tasaki, Commun. Math. Phys. **158**, 341 (1993).
- ¹¹E. Müller-Hartmann, J. Low Temp. Phys. **99**, 349 (1995).
- ¹²M. Kollar, R. Strack, and D. Vollhardt (unpublished).
- ¹³This exchange interaction between orbitals at different sites may be enhanced, or reversed in sign, by superexchange effects; see J. B. Goodenough, *Magnetism and the Chemical Bond* (Krieger, New York, 1976).
- ¹⁴R. Strack and D. Vollhardt, Phys. Rev. Lett. **72**, 3425 (1994); for a review see R. Strack and D. Vollhardt, J. Low Temp. Phys. **99**, 385 (1995).
- ¹⁵This Hamiltonian is obtained as follows. First one constructs a basis of Bloch functions which diagonalizes the kinetic energy together with the ionic potential, resulting in an infinite number of energy bands. Next one switches to the occupation number formalism by introducing field operators. Using the Bloch functions, one transforms to operators $\hat{a}_{\mathbf{k}\alpha\sigma}^\dagger$ that create an electron in a Bloch state. Finally, one performs a Fourier transform to arrive at the Wannier basis. Equation (2) is obtained by expressing the one-body kinetic energy, the ionic potential, and the two-body Coulomb interaction in this basis.
- ¹⁶The Coulomb interaction is, of course, a spin-independent two-body interaction. The fact that the creation and annihilation operators in (2) do contain spin indices is merely a consequence of the occupation-number formalism. Note that the Bloch basis has been chosen such that the one-body term is diagonal in k space, and therefore the interaction will *in general* not be diagonal in site representation; i.e., it cannot be expressed in terms of density operators \hat{n}_i only.
- ¹⁷D. K. Campbell, J. T. Gammel, and E. Y. Low, Jr., Phys. Rev. B **38**, 12043 (1988); **42**, 475 (1990).
- ¹⁸J. E. Hirsch, Phys. Rev. B **40**, 2354 (1989); **40**, 9061 (1989), **43**, 705 (1991); S. Tang and J. E. Hirsch, *ibid.* **42**, 771 (1990).
- ¹⁹R. Strack and D. Vollhardt, Phys. Rev. Lett. **70**, 2637 (1993).
- ²⁰A. A. Ovchinnikov, Mod. Phys. Lett. B **7**, 1397 (1993).
- ²¹J. de Boer, V. E. Korepin, and A. Schadschneider, Phys. Rev. Lett. **74**, 789 (1995).
- ²²J. de Boer and A. Schadschneider, Phys. Rev. Lett. **75**, 4298 (1995).
- ²³L. Arrachea and A. A. Aligia, Phys. Rev. Lett. **73**, 2240 (1994).
- ²⁴A. Schadschneider, Phys. Rev. B **73**, 10 386 (1995).
- ²⁵F. H. L. Essler, V. E. Korepin, and K. Schoutens, Phys. Rev. Lett. **68**, 2960 (1992); **70**, 73 (1993).
- ²⁶It should be noted that the model considered in Ref. 25 is related to a model for the metal-insulator transition in systems like V_2O_3 [C. Castellani, C. Di Castro, and M. Grilli, Phys. Rev. Lett. **72**, C3626 (1994)].
- ²⁷This connectivity condition may be stated as follows: In Nagao-ka's basis for singly occupied states (Refs. 4,5),
- $$|i, (\sigma_1 \cdots \sigma_{i-1} \sigma_{i+1} \cdots \sigma_L)\rangle$$
- $$= (-1)^j \hat{c}_{1,\sigma_1}^\dagger \cdots \hat{c}_{i-1,\sigma_{i-1}}^\dagger \hat{c}_{i+1,\sigma_{i+1}}^\dagger \cdots \hat{c}_{L,\sigma_L}^\dagger |0\rangle,$$
- all states with the same eigenvalue of \hat{S}^z must be obtainable by repeated application of the hopping operator, $\sum_{\langle ij \rangle \sigma} (\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \text{H.c.})$; i.e., the hopping operator is represented by an irreducible matrix in this basis.
- ²⁸E. Lieb and D. Mattis, Phys. Rev. **125**, 164 (1962); J. Math. Phys. **3**, 749 (1962).
- ²⁹U. Brandt and A. Giesekeus, Phys. Rev. Lett. **68**, 2648 (1992).
- ³⁰R. Strack, Phys. Rev. Lett. **70**, 833 (1993).
- ³¹See, e.g., F. R. Gantmacher, *Theory of Matrices* (Chelsea, New York 1971), Vol. II.