Multihole models for deterministically placed acceptor arrays in silicon

Jianhua Zhu,[1,*] Wei Wu,[2,†] and Andrew J. Fisher[1,‡]
[1]*UCL Department of Physics and Astronomy and London Centre for Nanotechnology, University College London, Gower Street, London WC1E 6BT, United Kingdom*
[2]*UCL Institute for Materials Discovery, University College London, Gower Street, London WC1E 6BT, United Kingdom*

(Received 25 May 2021; accepted 16 August 2021; published 8 September 2021)

In this paper, we compute the electronic structure of acceptor clusters in silicon by using three different methods to take into account electron correlations: the full configuration interaction (full CI calculation), the Heitler-London approximation (HL approximation), and the unrestricted Hartree-Fock method (UHF method). We show that both the HL approach and the UHF method are good approximations to the ground state of the full CI calculation for a pair of acceptors and for finite linear chains along [001], [110], and [111]. The total energies for finite linear chains show the formation of a fourfold-degenerate ground state (lying highest in energy), below which there are characteristic low-lying eightfold and fourfold degeneracies, when there is a long (weak) bond at the end of the chain. We present evidence that this is a manifold of topological edge states. We identify a change in the angular momentum composition of the ground state at a critical pattern of bond lengths, and show that it is related to a crossing in the Fock matrix eigenvalues. We also test the symmetry of the self-consistent mean-field UHF solution and compare it to the full CI; the symmetry is broken under almost all the arrangements by the formation of a magnetic state in UHF, and we find further broken symmetries for some particular arrangements related to crossings (or potential crossings) between the Fock-matrix eigenvalues in the [001] direction. We also compute the charge distributions across the acceptors obtained from the eigenvectors of the Fock matrix; we find that, with weak bonds at the chain ends, two holes are localized at either end of the chain while the others have a nearly uniform distribution over the middle; this also implies the existence of the nontrivial edge states. We also apply the UHF method to treat an infinite linear chain with periodic boundary conditions, where the full CI calculation and the HL approximation cannot easily be used. We find the band structures in the UHF approximation, and compute the Zak phases for the occupied Fock-matrix eigenvalues; however, we find they do not correctly predict the topological edge states formed in this interacting system. On the other hand, we find that direct study of the quantum numbers characterizing the edge states, introduced by Turner *et al.* [Phys. Rev. B **83**, 075102 (2011)], provides a better insight into their topological nature.

DOI: 10.1103/PhysRevB.104.125415

I. INTRODUCTION

In the last few decades, studies of defects in semiconducting systems have broadened to include applications to quantum computation and quantum simulation as well as their more traditional role in doping for classical electronics. Donors are especially well studied [1,2], but in materials such as Si having degenerate conduction-band minima they suffer from the disadvantage of intervalley interferences causing rapid oscillations in the wave functions and hence also in hopping or exchange interactions, leading to extreme sensitivity to the precise dopant position. For this reason, and because the spin-orbit coupling present in the valence band provides some additional opportunities to interact with the spin degrees of freedom, acceptors have recently attracted increased attention [3–5].

The electronic structure of a single acceptor can most simply be described by the "spherical model" [6–9], which includes spin-orbit coupling but neglects the cubic anisotropy of the host semiconductor and offers reasonable results for the electronic structure of an isolated acceptor. The interactions between a pair of acceptors in the spherical model have been studied by Durst *et al.* in the framework of a Heitler-London model [10]; later, they found that the interacceptor interactions in the same model are dominated at large distance by electric quadrupole moments [11]. For linear chains of acceptors, an independent-hole model was developed, including the contribution of cubic terms, and the existence of nontrivial single-particle topological edge states was demonstrated for finite chains, and related to band invariants of the corresponding infinite systems [12]. These investigations of pairs and linear arrays of acceptors suggest that the emerging techniques of deterministic doping [13] could lead to interesting results if applied to acceptors [14]. Advances in the experimental characterization of acceptors in silicon include measurements of the optical transitions and spectra of acceptors [15], measurement of the coherence time of the excited state of acceptors [16], and a study of transport properties of

*ucapjhz@ucl.ac.uk
†wei.wu@ucl.ac.uk
‡andrew.fisher@ucl.ac.uk

boron-doped material [17], etc. The readout and control of the spin-orbit state of two coupled acceptors has also been demonstrated experimentally in silicon, suggesting a possible alternative route to quantum computing [18–22]. Finally, there is great potential for applications in the simulation of fermionic strongly correlated many-body systems using acceptors [23]. All these investigations imply that a system of acceptors in well-defined locations could offer some unique properties in its electronic structure.

In this paper, we construct and solve multihole models (including hole-hole Coulomb interactions) for lines of acceptors in silicon with one hole per acceptor, along three high-symmetry directions ([001], [110], and [111]), based on three different methods: full configuration interaction calculation (full CI, in Sec. II A 1), the Heitler-London approximation (HL approximation, in Sec. II A 1) based on the full CI calculation but with a restricted basis, and the unrestricted Hartree-Fock method (UHF method, in Sec. II A 2) which represents the multihole state by a Slater determinant of one-hole states. Some limitations of the full CI calculation and the HL approximation are discussed in Sec. II A 1. We study dimerized chains with staggered bond lengths d_1 and d_2 and concentrate on a "small-separation" case with $d_1 + d_2 = 3a_0$ and a "large-separation" case with $d_1 + d_2 = 6a_0$ where a_0 is the effective Bohr radius; we show that both the HL approach and the UHF method are accurate approximations to the ground state of the fully exact CI calculation for these finite-length linear chains. We investigate the energy spectrum obtained from full CI for a four-acceptor chain and explain the ground state in terms of the formation of edge states; we also relate an anticrossing in the [001] direction for the small-separation case to the behavior of the Fock matrix eigenvalues obtained from the UHF method. We analyze the symmetries of the states produced by symmetry breaking in the UHF solution, and present evidence for the existence of nontrivial many-body edge states in the finite chain system. We point out that the UHF method can be applied to a linear chain with periodic boundary conditions, and calculate the band structure formed by the Fock matrix eigenvalues. We also analyze the topological phases of the system based on two methods: first, a method focusing on the edge states of finite one-dimensional interacting fermionic systems and, second, the Zak phase [24] for an infinite noninteracting system.

II. COMPUTATIONAL DETAILS

A. Multihole models

In our previous paper [12], we developed a one-hole model to describe a pair of acceptors and a linear acceptor chain. Here, we use the same approach to describe the one-hole part of the Hamiltonian, including cubic anisotropy, but only considering the nearest transitions for the chain (see Sec. II A 4). We then combine this one-hole Hamiltonian with two-hole terms representing the interhole Coulomb repulsion, using methods described in Ref. [25]. Our units of energy and length are the effective Rydberg $R_0 = \frac{e^4 m_0}{2\hbar^2\epsilon_0^2\gamma_1}$ and the effective Bohr radius $a_0 = \frac{\hbar^2\epsilon_0^2\gamma_1}{e^4 m_0}$, respectively [7]. We use parameters appropriate for silicon throughout; however, our methods are easily transferable to other cubic semiconductors. With these silicon parameters, $R_0 = 24.8$ meV and $a_0 = 2.55$ nm. In all cases we report our results for lines oriented along the three highest-symmetry directions of the cubic host: [001], [110], and [111].

1. Full configuration interaction calculation (full CI calculation) and Heitler-London approximation (HL approximation)

The configuration interaction calculation (full CI) retains a basis of Slater determinants corresponding to all possible configurations of the holes distributed across basis states on all acceptors, and the Hamiltonian is

$$\hat{H}_{\mathrm{CI}} = \sum_i \hat{H}_i - \sum_{i<j} \frac{2}{r_{ij}}, \tag{1}$$

where $\hat{H}_i$ is the one-hole Hamiltonian from our previous paper [12], $\frac{2}{r_{ij}}$ is the hole-hole interaction in effective Rydberg units, and i, j label the holes. The interaction appears with a minus sign because the Hamiltonian is expressed for electron states. Therefore, throughout this paper, the most favorable states for occupation by holes are those with the *highest* energy: we refer to the highest-energy state as the "ground state." The overlap matrix is also needed and can be written as

$$\hat{S}_{\mathrm{CI}} = \hat{S}_1 \otimes \hat{S}_2 \otimes \cdots \otimes \hat{S}_N, \tag{2}$$

where $\hat{S}_i$ is the overlap matrices for the one-hole model, and N is the number of holes.

The full CI calculation is exact for a given choice of single-particle basis, but scales very badly (superexponentially) with the size of the system. Also, the total energy expression is not extensive so it cannot be implemented under periodic boundary conditions. The first problem is ameliorated by restricting the set of configurations to those with exactly one hole per acceptor; we call this the Heitler-London (HL) approximation because it is in the same spirit as the Heitler-London treatment of the H_2 molecule, and has been used for acceptors pairs in Ref. [10]. The many-particle basis set now grows more slowly (although still exponentially), but the difficulty in treating the infinite system still remains.

2. Unrestricted Hartree-Fock method

To handle the infinite system we employ an unrestricted Hartree-Fock (UHF) method, where the many-hole wave function is optimized over single Slater determinants constructed from a set of one-hole functions, without any restriction on the spin components of each function. The optimization of the one-hole functions results in a self-consistent-field (SCF) approach, where each hole can be understood to experience the average interaction of the others. The one-hole functions are eigenfunctions of the Fock matrix $\hat{F}$, which is given by

$$\hat{F} = \hat{H}^{\mathrm{core}} + \hat{G}, \tag{3}$$

where $\hat{H}^{\mathrm{core}}$ is the Hamiltonian for the one-hole model (including spin-orbit coupling), and $\hat{G}$ is a matrix reflecting the self-consistent influence from other holes. If we expand all quantities in terms of a set of single-hole basis functions $|\phi_\mu\rangle$,

G is given by

$$G_{\mu\nu} = \sum_{\lambda\sigma} P^{\lambda\sigma}[(\mu\nu\|\sigma\lambda) - (\mu\lambda\|\sigma\nu)], \tag{4}$$

where μ, ν, σ, λ are labels running over all basis functions on all acceptors,

$$(\mu\nu\|\sigma\lambda) = \int d\mathbf{x}_1 d\mathbf{x}_2\, \phi^*_\mu(\mathbf{x}_1)\phi^*_\sigma(\mathbf{x}_2)\frac{-2}{|\mathbf{r}_1 - \mathbf{r}_2|}\phi_\nu(\mathbf{x}_1)\phi_\lambda(\mathbf{x}_2) \tag{5}$$

[where $\mathbf{x} = (\mathbf{r}, \tau)$ is a composite coordinate for position $\mathbf{r}$ and intrinsic angular momentum τ, and $(\mu\nu\|\sigma\lambda)$ is the notation used in Ref. [26]] are matrix elements of the Coulomb interaction, and P is the one-hole density matrix which can be constructed as

$$P^{\mu\nu} = \sum_i^N C_i^\mu C^{*\nu}_i, \tag{6}$$

where C_i is an eigenvector of the generalized eigenproblem

$$FC_i = \epsilon_i SC_i, \tag{7}$$

N is the number of holes (hence the number of occupied eigenvectors), and i goes through all eigenvector labels. Once again, because our calculation is describing holes, the single-particle states are occupied according to the Aufbau principle from the highest eigenvalue downwards. The total energy can then be written as

$$E_{\text{tot}} = \frac{1}{2}\sum_{\mu\nu} P^{\nu\mu}\left(H^{\text{core}}_{\mu\nu} + F_{\mu\nu}\right). \tag{8}$$

The self-consistent calculation continues until the output density matrix (6) is similar to the input one used in (4). Further details can be found in Ref. [26]; however, in contrast to the conventional case, our system contains spin-orbit coupling and therefore we cannot separate the single-particle functions into separate sets corresponding to each spin component. So it is necessary to include exchange interactions between *all* pairs of single-hole states, not just those of the same spin.

3. Periodic boundary conditions

Although less accurate than the CI method, the UHF method does not have the limitations mentioned in Sec. II A 1. It scales polynomially, rather than exponentially, as the system size increases, and the total-energy expression (8) is extensive. So it is possible to apply it to a linear chain with periodic boundary conditions. In this case, the Fock matrix $\hat{F}_k$ at a particular Bloch wave vector k will be

$$\hat{F}_k = \sum_X e^{ikX}\hat{F}_X = \sum_X e^{ikX}\left(\hat{H}^{\text{core}}_X + \hat{G}_X\right) = \hat{H}^{\text{core}}_k + \hat{G}_k, \tag{9}$$

where X labels lattice displacements of a single unit cell, $\hat{F}_X$, $\hat{H}^{\text{core}}_X$, and $\hat{G}_X$ are the elements of F, H, and G connecting different cells separated by X, and $\hat{H}^{\text{core}}_k$ and $\hat{G}_k$ are the matrices of $\hat{H}^{\text{core}}$ and $\hat{G}$ in momentum space. The Fock matrix $\hat{F}_k$ can be diagonalized to find a set of eigenvectors C_{ki}, and the corresponding contribution P_k to the the one-particle density matrix is

$$P^{\mu\nu}_k = \sum_i^N C^\mu_{ki} C^{*\nu}_{ki}. \tag{10}$$

The real-space form of $P^{\mu\nu}$ can then be recovered by inverse Fourier transformation, and reinserted into the SCF procedure as previously.

4. Spatial cutoffs

In practice, the sums in Eqs. (4) and (9), as well as the corresponding sums over acceptor cores in $\hat{H}^{\text{core}}$, have to be truncated. For the results in this paper we have performed this truncation after nearest neighbors; for exchange and hopping terms which involve transferring a single hole from site to site, this is justified by the relatively well-localized acceptor wave functions (this means that the relevant matrix elements will decay exponentially with hopping distance). The Coulomb terms (both the hole-hole interaction and the hole-core interaction) decay much more slowly, like $\frac{1}{R}$ (where R is the separation between the charges), but will cancel one another out provided the system is approximately charge neutral at all points. We have checked that the key findings of this paper are reproduced in an extended model which includes all the next-nearest-neighbor transitions but only the largest next-nearest-neighbor hole-hole interactions, for both the finite length chain and periodic boundary case. There, all the key features that we are going to discuss in this paper are kept. Considering introducing the next-nearest hole-hole interactions will more than double the time of the calculations, it is a wise choice to only consider the nearest neighbors.

5. Computing the Zak phase from the unrestricted Hartree-Fock calculation

The Zak phase [24] is a bulk quantity that indicates whether a noninteracting insulator is topologically trivial or nontrivial (supporting edge states): when it is 0 modulo 2π the system should be trivial, when it is $\pm\pi$ modulo 2π the system becomes nontrivial. We can obtain this quantity from the total density matrix, which is available during the SCF procedure of the UHF calculation. We follow a recent paper [27] in calculating the Zak phase for in a general situation is using the formula

$$Z = \arg\left[\text{tr}\left(\prod_k S_k P_k\right)\right], \tag{11}$$

where S_k is the overlap matrix transformed into Fourier space, P_k is the single-particle density matrix as defined above, and k is the wave vector going through the first Brillouin zone. However, we have to remember that Coulomb interactions can change the topological classification [28,29] so we cannot necessarily expect the Zak phase to predict correctly the presence or absence of topological edge states; indeed, we show evidence in Sec. III D that the Zak phases do not correspond to the topological property of the edge states.

B. Single-particle basis

It remains to specify the basis for the single-particle states on each acceptor. As in our previous paper [12], we

TABLE I. The eigenenergy of the fourfold-degenerated ground state (Γ_8^+) obtained from the Gaussian expansion with 21 Gaussian parameters and 5 Gaussian parameters for Si and the difference between them; the energy unit is the effective Rydberg R_0, and the difference is shown in the percentage of the original 21-parameter result.

21-parameter result	5-parameter result	Difference
$1.868314R_0$	$1.854034R_0$	0.7644%

decompose the spatial parts of the acceptor states into linear combinations of Gaussian orbitals. However, as we are interested in the behavior of the low-lying states of the linear chain, we make several changes. First, we consider only the fourfold-degenerate ground-state manifold ($1\Gamma_8^+$) of an isolated acceptor. We expand the radial parts as

$$f_0(r) = r^l \sum_i A_i e^{-\alpha_i r^2}, \tag{12}$$

where l is the orbital angular momentum of the envelope function and α_i is a Gaussian exponent. Second, because we only need to describe the ground state, we use only five Gaussian functions, with exponents $\alpha_i = \{100.0, 25.0, 6.25, 1.5625, 0.390625\}$, rather than 21 as in our previous paper [12]; the single-acceptor ground-state energies in silicon computed with 5 and 21 Gaussians are compared in Table I and found to differ by less than 1%. The reduction in the number of Gaussians saves time in the evaluation of matrix elements for the subsequent calculations.

Finally, we remove the admixture of G-orbital Gaussian components ($l = 4$) in the ground-state manifold, to limit the size of the matrices involved in the calculation and renormalize the remaining parts of the wave function. As an example, we compare the energy of the doubly degenerate ground state for a single hole bound to a pair of acceptors in the [001] direction with and without the G orbitals in Fig. 1. It can be seen that omitting the G orbitals leads to errors in the energy of 1%–2%.

For convenience in the discussion of results in Sec. III B, we assign labels to the states of the fourfold-degenerate ground Γ_8^+ manifold so that we can distinguish them. The main contribution is from the $S_{\frac{3}{2}}$ state with total angular momentum $F = \frac{3}{2}$; we therefore use the values of the angular momentum projections $m_F = \{\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}\}$ to label the different rows of the irreducible representation. (The total angular momentum $\vec{F} = \vec{L} + \vec{I} + \vec{S}$, where $\vec{I}$ is the intrinsic orbital angular momentum of the p states in the valence band, is as defined in Ref. [12].)

III. RESULTS AND DISCUSSION

A. A pair of acceptors

For a pair of acceptors, all the methods and approximation mentioned in Sec. II A can be applied. To show the long-range behavior clearly, we calculate the interaction energy

$$E_{\text{int}} = E_T - 2E_{\text{single}} = E_{\text{tot}} - \frac{2}{R} - 2E_{\text{single}}, \tag{13}$$

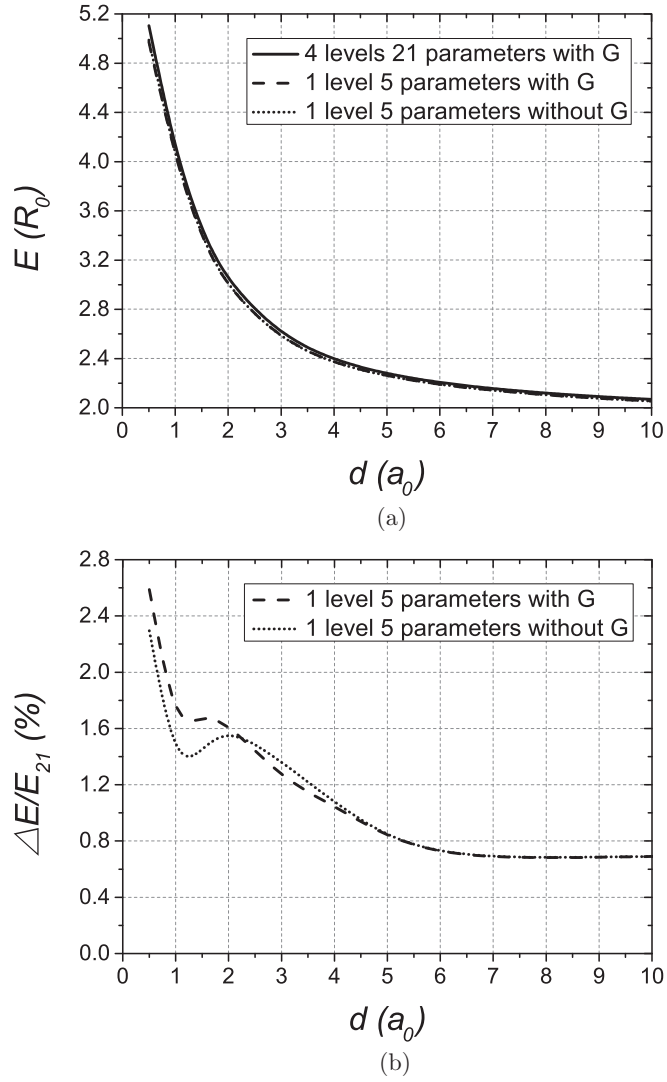


FIG. 1. The behavior of the doubly degenerate ground-state energies with different approximations in the [001] direction for a pair of acceptors in Si under the one-hole model: (a) the ground-state eigenenergies, (b) the differences between the ground-state eigenenergies with different approximations. The solid line is the result of our earlier paper [12] with 21 Gaussians, the dashed lines are for the ground state ($1\Gamma_8^+$) with 5 Gaussian parameters but including G orbitals, the dotted lines are for the ground state ($1\Gamma_8^+$) with 5 Gaussian parameters excluding G orbitals. In (b), energy differences with respect to energy E_{21} [the solid line in (a)] are shown as percentages of the energy E_{21} [the solid line in (a)].

where E_{single} is the single-acceptor energy, E_T is the total energy including the core-core interaction, E_{tot} is the total energy for the holes only [directly obtained from the Hamiltonian (1)], and $\frac{2}{R}$ is the core-core interaction term (appearing with a minus sign to be consistent with our convention for the hole energy). We did not consider the core-core interaction term in our previous paper [12]; we refer to E_{tot} as the "total energy" for the rest of this paper. The interaction energies E_{int} of the ground state from three different models in three high-symmetry directions are shown in the left column of Fig. 2; they appear as the negatives of standard molecular

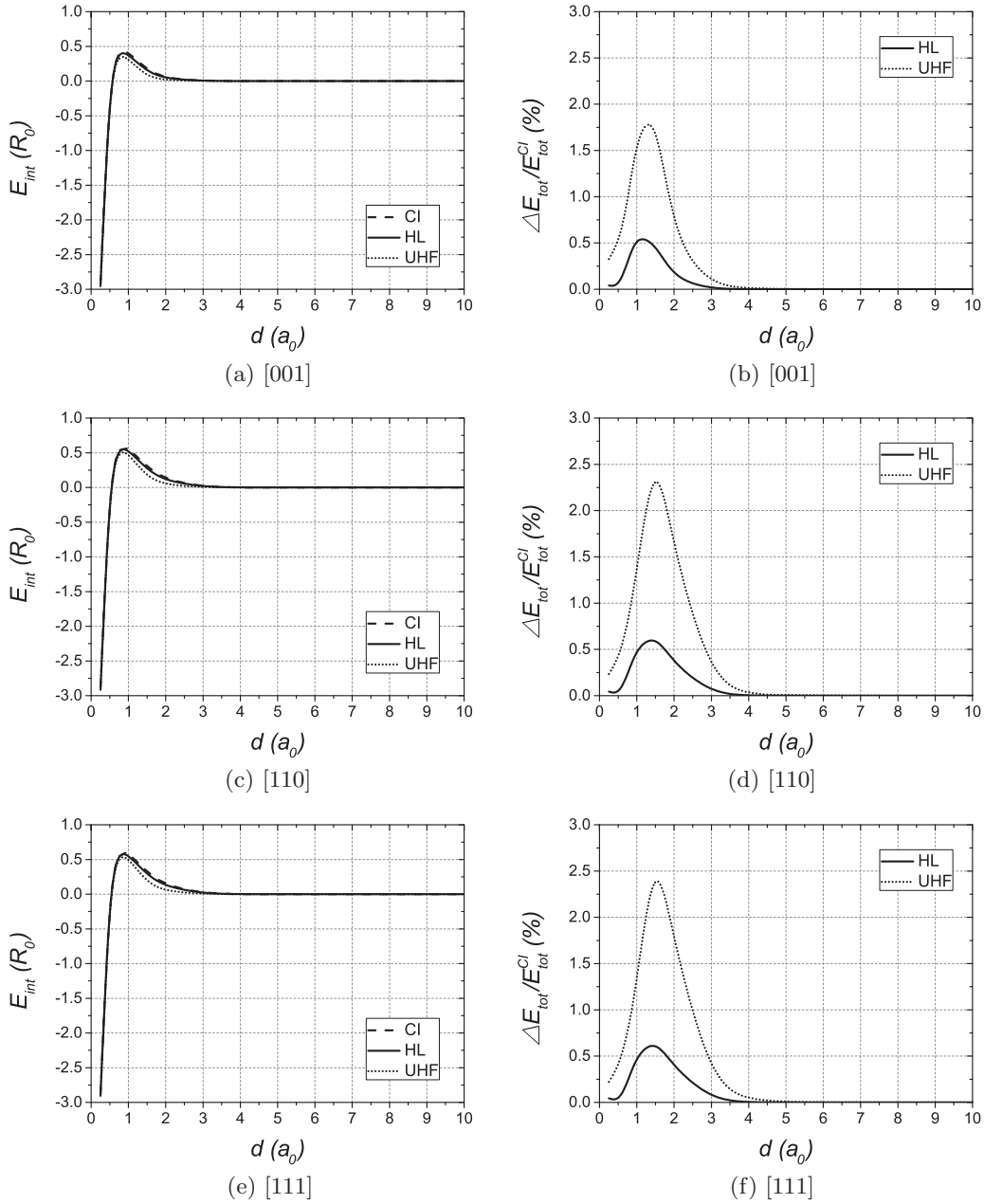
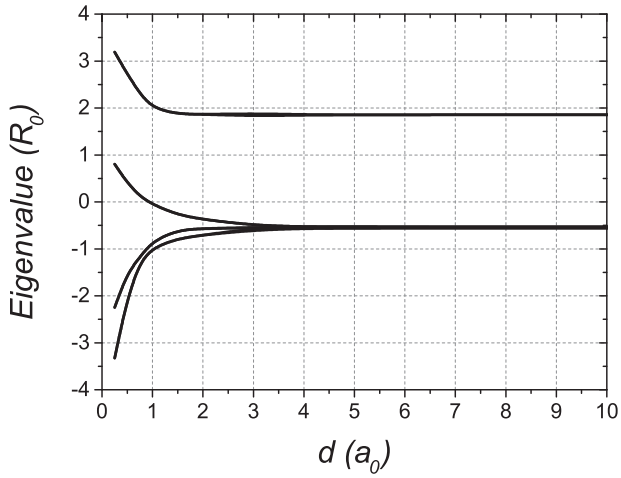


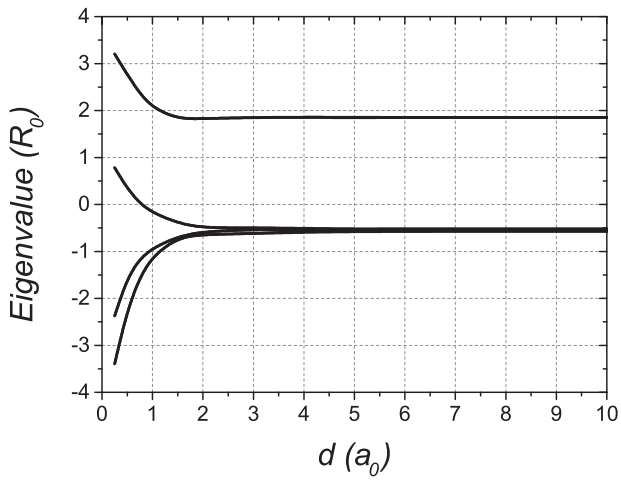
FIG. 2. The interaction energy $E_{\rm int}$ of the ground state and the difference of the total energy $E_{\rm tot}$ towards the full CI calculation in three typical directions for a pair of acceptors: (a) the interaction energy $E_{\rm int}$ in the [001] direction, (b) the difference of the total energy $E_{\rm tot}$ in the [001] direction, (c) the interaction energy $E_{\rm int}$ in the [110] direction, (d) the difference of the total energy $E_{\rm tot}$ in the [110] direction, (e) the interaction energy $E_{\rm int}$ in the [111] direction, (f) the difference of the total energy $E_{\rm tot}$ in the [111] direction. For (a), (c), (e), the dashed line is for the full CI calculation, the solid line is for the HL approximation, the dotted line is for the UHF method. For (b), (d), (f), the solid line is for the HL approximation, the dotted line is for the UHF method, all the differences are in the percentage of the full CI result.

binding-energy curves. We also show the difference in the total energy $E_{\rm tot}$ between the full CI calculation and the other approaches (as a percentage of the full CI result) in the right column. Both the HL approximation and the UHF method are good approximations to full CI for all directions, but the differences are greatest at small separations; the HL approach generally provides a better energy than UHF (since they involve variational approximations to the true wave function, both methods give a lower bound on the true ground-state energy in the hole system). For the [001] direction, the differences reach a maximum around $1.5a_0$ and can be ignored when the separation $d \geqslant 4a_0$; for the [110] and [111] directions, they peak around $1.5a_0$ and could be ignored for $d \geqslant 5a_0$.

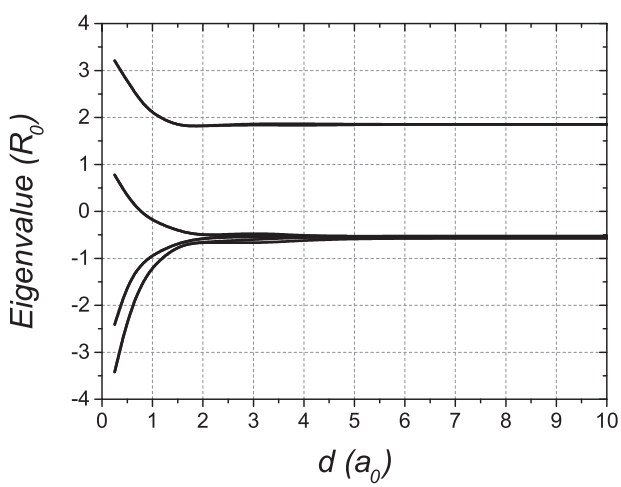
For the convenience of further discussion in Sec. III B, Fig. 3 shows the Fock matrix eigenvalues for pairs oriented along different directions. The ground state appears at the top of the pictures as this is a calculation for acceptors. Each line represents a pair of almost doubly degenerate states; since there are two holes, only the doubly degenerate ground state at the top of the diagram will be filled. There is a large gap between the filled and empty states at all separations; this



(a) [001]



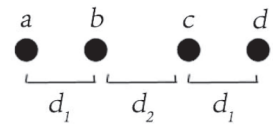
(b) [110]



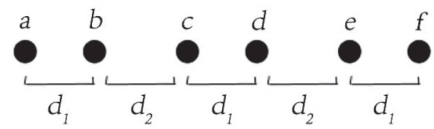
(c) [111]

FIG. 3. The behavior of the Fock matrix eigenvalues in different directions for a pair of acceptors: (a) the [001] direction, (b) the [110] direction, (c) the [111] direction.

is generated by the strong hole-hole repulsion within the self-consistent field. We will see that this feature persists in the calculations on larger systems.



(a) 4-acceptor linear chain



(b) 6-acceptor linear chain

FIG. 4. Schematics of the linear chains studied in this paper. a, b, c, d, e, f are the labels of acceptors; $d_1 < d_2$ is known as the "short-long arrangement," $d_1 > d_2$ is known as the "long-short arrangement."

In the absence of cubic anisotropy, Durst *et al.* [11] argue that the long-range interaction between two acceptors is dominated by quadrupolar effects, which they find favor a doubly degenerate state with total angular momentum $M_F = \pm 2$ about the core axis. This corresponds to partially aligned pairs of holes, with $m_F = \pm(\frac{3}{2}, \frac{1}{2})$ on the two acceptors. However, with the inclusion of significant cubic anisotropy appropriate for Si ($\delta > 0$ and indeed $\delta \sim \mu$) we find that the pair ground state in the quantized direction (the [001] direction) only crosses over to this form for very large separations $d > 5a_0$; for smaller separations, the ground state is nondegenerate and dominated by antiferromagnetically coupled configurations such as $m_F = \pm(\frac{3}{2}, -\frac{3}{2})$ (for $d < 3a_0$) and $m_F = \pm(\frac{1}{2}, -\frac{1}{2})$ for $3a_0 \leqslant d \leqslant 5a_0$.

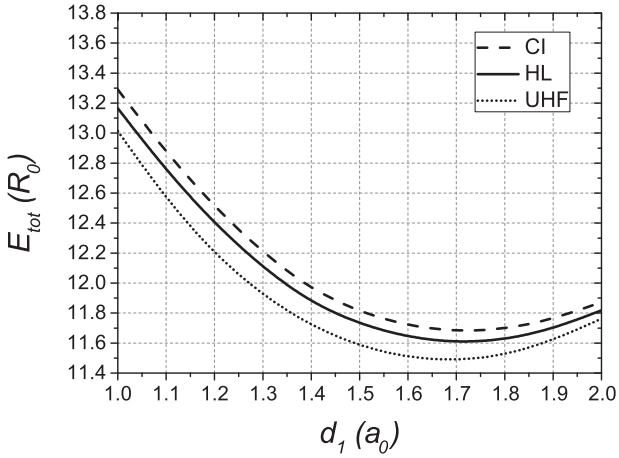
B. Finite dimerized linear chains

We next consider chains of four and six acceptors, with one hole per acceptor and with the separations (d_1, d_2) alternating to form a dimer chain as shown in Fig. 4. When $d_1 < d_2$, we will refer to a "short-long arrangement" throughout the rest of the paper, while when $d_1 > d_2$ we will call it a "long-short arrangement." We investigate two different regimes, each defined by a fixed value of $d_1 + d_2$: a "small-separation" case with $d_1 + d_2 = 3a_0$, and a "large-separation" case with $d_1 + d_2 = 6a_0$.

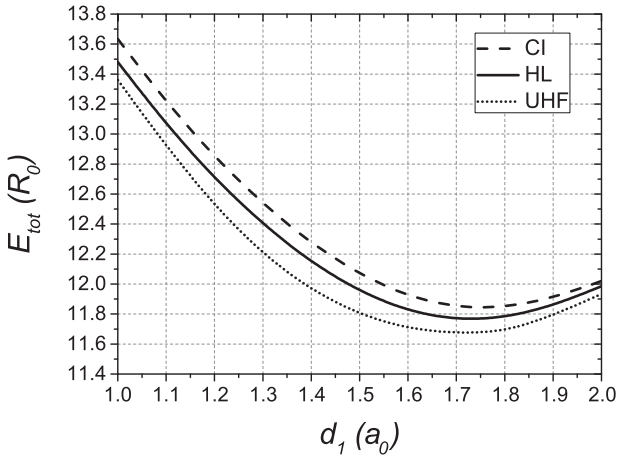
1. Small-separation case ($d_1 + d_2 = 3a_0$)

The hole-hole repulsion term now strongly influences the distribution of the holes: although the parabolic potential due to the negative acceptor cores found in our earlier work [12] is still present, the holes are no longer concentrated in the middle of the chain but are kept apart by their mutual Coulomb repulsion and have a nearly uniform distribution along the chain. This suggests that our small-separation case is already on the insulating side of the Mott transition, so the HL approximation can naturally be applied and may be expected to give good results.

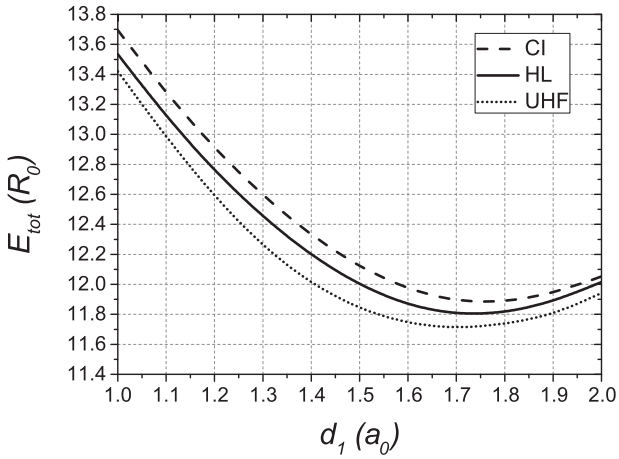
For chains of four acceptors, the ground-state total energy was obtained from all the methods mentioned in Sec. II A along three high-symmetry directions, and is shown in Fig. 5. Both the HL and UHF methods are reasonable approximations



(a) [001]



(b) [110]



(c) [111]

FIG. 5. The behavior of the total energy of the ground state under different arrangements in three typical directions for the small-separation case ($d_1 + d_2 = 3a_0$) of the four-acceptor linear chain: (a) the [001] direction, (b) the [110] direction, (c) the [111] direction. The dashed line is for the full CI calculation, the solid line is for the HL approximation, the dotted line is for the UHF method.

to the full CI result in all directions, with the HL approach offering a better agreement with the full CI calculation. The difference between the full CI and the HL results reduces as the arrangement changes from short-long to long-short; the HL approximation should be more accurate for chains with more acceptors, especially when the average separation between each pair of acceptors is larger. The average separation grows as d_1 increases; the result can then be understood by noting that the accuracy of the HL method for a pair remains roughly constant from $d = 1a_0$ to $d = 1.5a_0$ (see the right column of Fig. 2) but then improves from $d = 1.5a_0$ to $d = 2a_0$. The UHF approximation also becomes more accurate as more acceptors are included, but the significant discrepancies in the energy of a pair with separations around $1.5a_0$ (Fig. 2) are reflected in significant errors in the middle of Fig. 5, where $d_1 \approx d_2 \approx 1.5a_0$. We also computed results for chains of six acceptors, using the HL and UHF methods only; the behavior of the total energies was similar.

We analyze the full CI ground-state eigenvector by looking at the dominant components (those with largest absolute values) in the basis of single-acceptor states described in Sec. II. We can separate the four degenerate states of an isolated acceptor into two groups, those derived from $m_F = \pm\frac{3}{2}$ and those from $m_F = \pm\frac{1}{2}$. We refer to the ground state as "unhybridized" if the dominant components contain either $m_F = \pm\frac{3}{2}$ or $m_F = \pm\frac{1}{2}$ single-acceptor states (but not both), while we refer to it as "hybridized" if they contain both types of single-accpetor states.

In Fig. 6, we show the behavior of the 50 highest-energy (hence, most favorable) states of the full CI calculation under different arrangements of the bonds along three high-symmetry directions. For the [001] direction, the ground state is nondegenerate on the left-hand (short-long) side of the picture, while it joins three other states and forms a fourfold-degenerate state on the right-hand side (long-short arrangement side) which is followed in energy by an eightfold-degenerate state and another fourfold-degenerate state as shown in Fig. 6(b). We observe that among the dominant components, only the states on the acceptors at the end of the chain change between these states; the dimensionality 16 of these highest manifolds comes from the 4 levels on one end multiplied by 4 levels on the other end, implying the existence of a manifold of edge states. The situation is similar for the other directions; we analyze the structure of this manifold in more detail in Sec. III D. It also can be seen that the ground state crosses with the highest excited states between $d_1 = 1.4a_0$ and $1.5a_0$ in the [001] direction; the dominant components of the ground state are unhybridized to the left of the dotted line but become hybridized to the right of it. We will refer to the separation where the crossing (or anticrossing) between the states happens as the "crossing point," and the separation where dominant component of the ground state changes as the "changing point." We see that within the resolution of the step size used ($0.1a_0$), the crossing point and the changing point are the same in the [001] direction.

For the UHF calculations we can understand the overall state most clearly in terms of the behavior of the Fock matrix

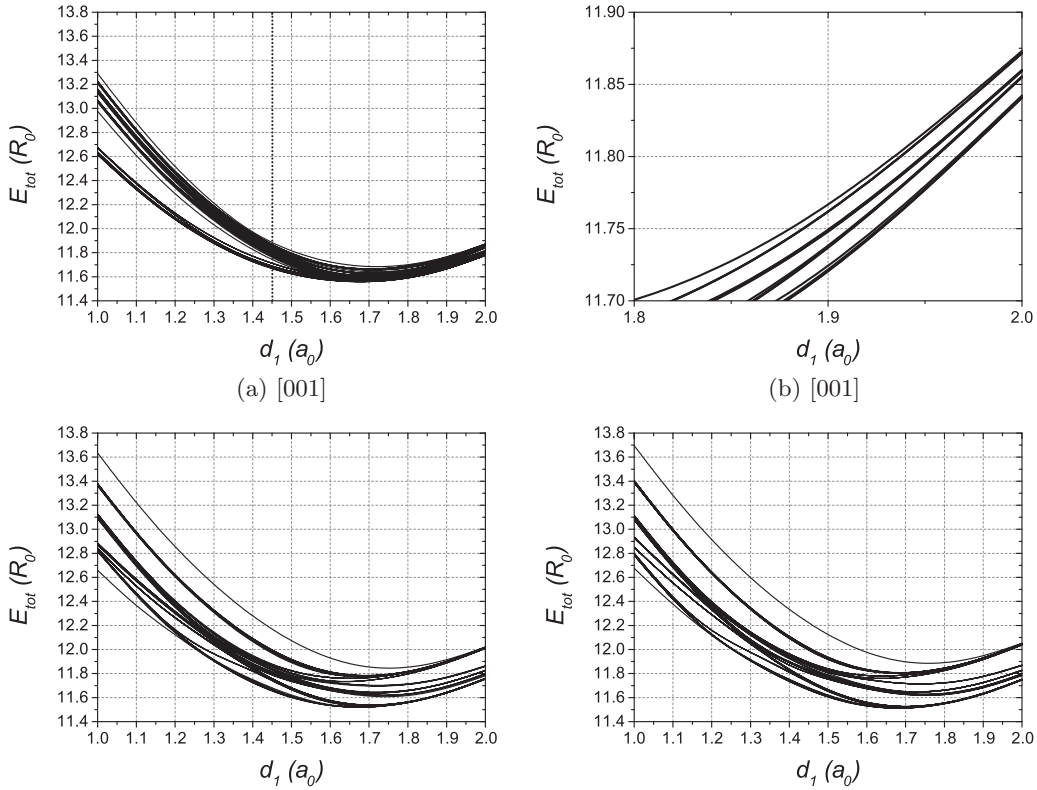


FIG. 6. The behavior of the total energy of the highest 50 energy states of the full CI result under different arrangements in three typical directions for the small-separation case ($d_1 + d_2 = 3a_0$) of the four-acceptor linear chain with the changing point: (a) the [001] direction, (b) details of the long-short arrangement side in the [001] direction, (c) the [110] direction, (d) the [111] direction. In (a), the dotted line is for the changing point.

eigenvalues, shown for different directions in Figs. 7 (four-acceptor chain) and 8 (six-acceptor chain). Here the states are usually doubly degenerate (corresponding to Kramers degeneracy under time-reversal symmetry) but show splittings for certain acceptor arrangements where the symmetry is lower (see Sec. III B 3). The four highest states in Fig. 7, and the six highest in Fig. 8, will be occupied by holes. In all cases there is a large gap between filled and empty states due to the effect of the strong hole-hole repulsion. Compared with Fig. 3 for a dimer, the two significant differences are (i) the splitting of degenerate states, and (ii) the crossing between filled states in the [001] direction in Fig. 7(b). In general we find that the self-consistency cycle in the UHF method breaks the symmetry of the system, with different sets of eigenvectors of the Fock matrix corresponding to the same total energy; we analyze this symmetry breaking further in Sec. III B 3. The crossing occurs close to the changing point identified in the CI calculation, so the change in the single-acceptor energy levels in the dominant component of the CI ground state is related to a change in the ordering of single-electron states in UHF. For the six-acceptor chain, it can be seen from Fig. 8 that another crossing appears around $d_1 = 1.7a_0$, implying another similar crossing between the total-energy ground state and higher excited states which is not able to be shown due to the limit number of acceptors.

The HL approach for the four-acceptor chain (not shown) gives similar results to the CI method, including a fourfold-degenerate ground state when $d_1 > d_2$ and the presence of a changing point where the composition of the ground state changes; however, the changing point now appears between $d_1 = 1.3a_0$ and $1.4a_0$, while the crossing point is still around $d_1 = 1.4a_0$. This suggests that the HL method is a good approximation for both the ground state and low-lying excited states, and preserves some of the main features of the energy spectrum. For the six-acceptor chain there is only one obvious crossing between the ground state and the first excited states, as the degenerate states appear for significantly smaller values of d_1 than before. But we now see two changing points for the eigenvectors: one is between $d_1 = 1.3a_0$ and $1.4a_0$, the other is between $d_1 = 1.6a_0$ and $1.7a_0$.

To understand in more detail the behavior of the energy gap, we show in Fig. 9(a) the difference between the total-energy ground state and first excited state in the [001] direction as a function of d_1. There are two regions of particular interest: the first is the neighborhood of the crossing/changing point where the gap reduces and then increases again ($d_1 = 1.3a_0$ to $1.4a_0$). The minimum gap for four acceptors is around $1.4a_0$ for both the CI case (solid line) and the HL case (dashed line), but shifts to shorter separations for six acceptors (dotted line). To show the details of the crossings among the first few states, a good choice is to show the energy difference between the ground state and excited states as the energies shift dramatically from the short-long arrangement to the long-short arrangement according to Fig. 6(a). In this way, the crossings

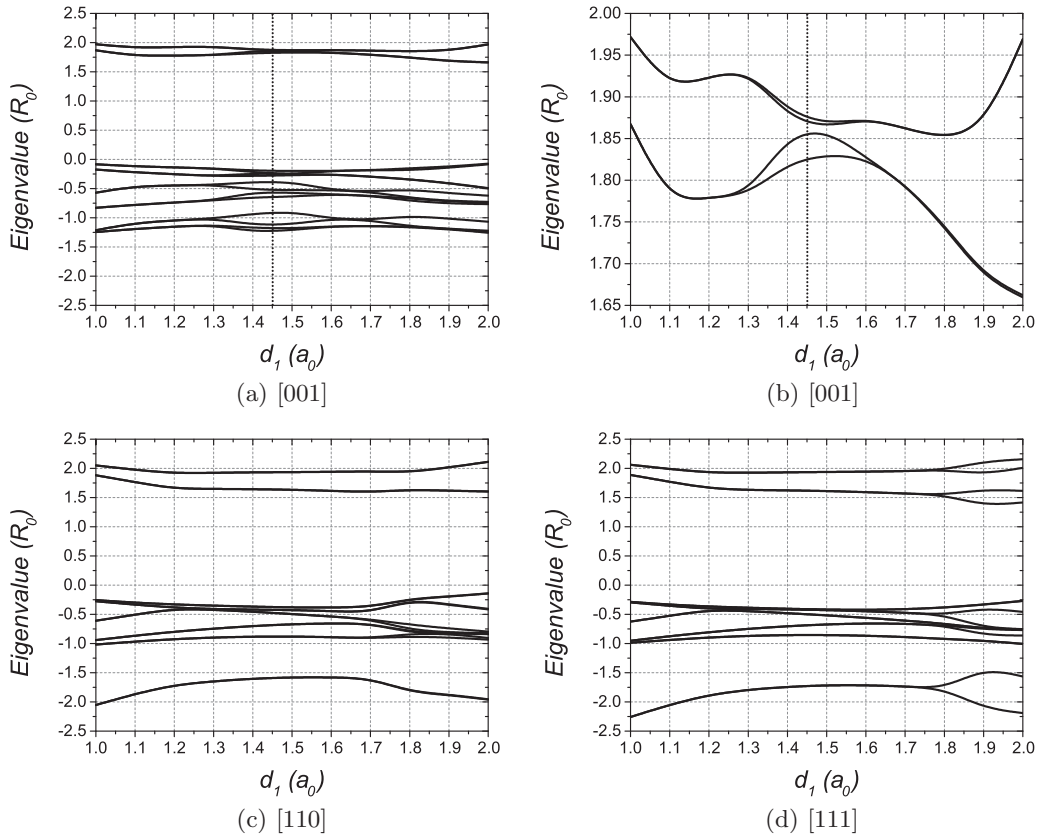


FIG. 7. The behavior of the Fock matrix eigenvalues obtained from the UHF method under different arrangements in three typical directions for the small-separation case ($d_1 + d_2 = 3a_0$) of the four-acceptor linear chain with the changing point: (a) the [001] direction, (b) details of the highest four eigenvalues in the [001] direction, (c) the [110] direction, (d) the [111] direction. For (a) and (b), the dotted lines are for the changing points.

between excited states are shown as usual, while the crossing between the ground state and excited states will be reflected by the value of the difference. Here for the convenience of the further discussion, the ground state before the changing point is called as $|\phi_0\rangle$, while the ground state after the changing point is called as $|\phi_0'\rangle$. In Fig. 9(b), we show the energy difference between the ground state and first 15 excited states for the full CI calculation, where we find a small gap between excited states around $1.4a_0$, which appears to make the "crossings" here anticrossings as $|\phi_0'\rangle$ is found above this gap before the changing point. It is also reasonable to believe the others in the full CI calculation and HL approach are anticrossings. As there is a band of excited states with similar energies in Fig. 9(b), it is helpful to follow the energy difference between the ground state and the excited state that crosses with it, rather than the minimum gap; in Fig. 9(c), we show the energy difference between the previous and new ground states during the anticrossing. This suggests that the true anticrossing is between $d_1 = 1.40a_0$ and $1.41a_0$, a slightly larger value than in the HL approach. The second region of interest is the right-hand side (large d_1), where the fourfold-degenerate manifold of ground states in the six-acceptor system forms for smaller values of d_1 than in the four-acceptor system; alternatively, for a given $d_1 > d_2$, the degeneracy of the ground state becomes better as more acceptors are involved (the same is true for the following eightfold-degenerate and fourfold-degenerate manifolds). This is what would be expected if the degeneracy arises from almost independent sets of localized edge states at either end of the chain (see Sec. III D).

Figure 10 shows that the magnitude of the expectation value of the angular momentum vector on each acceptor in the symmetry-broken UHF solution. At the smallest values of d_1 (the short-long case) the angular momentum is zero everywhere, whereas for large d_1 (the long-short case) it is dominantly located at the ends of the chain. To see if this is related to possible nontrivial edge states, we show the hole distributions from each eigenvector of the Fock matrix for different arrangements in the three high-symmetry directions in Fig. 11. Here short-long refers to $d_1 = 1a_0$, $d_2 = 2a_0$, and long-short to $d_1 = 2a_0$, $d_2 = 1a_0$. The one-hole states do not localize at any particular acceptor under the short-long or uniform arrangements; however, for the long-short case, two states localize at the ends of the chain [the dotted lines in Figs. 11(c), 11(f), and 11(i)], while the others have a nearly uniform distribution across the middle. The states localized at the ends [the dotted lines in Figs. 11(c), 11(f), and 11(i)] are always the lowest (i.e., least favorable) states occupied by holes, which may imply the existence of the nontrivial edge states occurring in the long-short case (since the charge rearrangements we previously identified in the noninteracting case [12] in response to the parabolic potential no longer force the states localized at the end of the chain to be the highest ones and intervene to shift the edge states to the short-long limit).

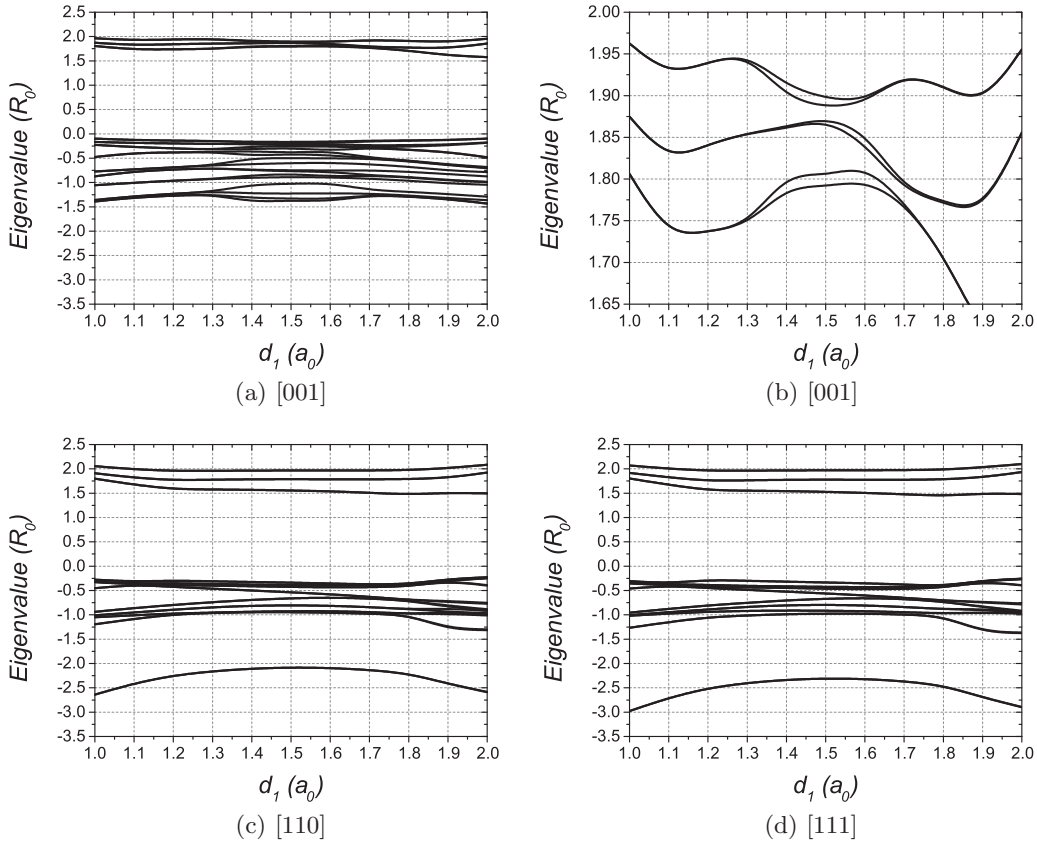


FIG. 8. The behavior of the Fock matrix eigenvalues obtained from the UHF method under different arrangements in three typical directions for the six-acceptor linear chain when $d_1 + d_2 = 3a_0$: (a) the [001] direction, (b) details of the highest six eigenvalues in the [001] direction, (c) the [110] direction, (d) the [111] direction.

2. Large-separation case ($d_1 + d_2 = 6a_0$)

For the large-separation case, we show the behavior of the ground-state total energy obtained from all the methods mentioned in Sec. II A for a four-acceptor chain along high-symmetry directions in Fig. 12. The HL and UHF results are closer to the full CI results than in the small-separation case, consistent with the better agreement found between the methods for larger separations in the case of pairs (Fig. 2). The best agreement is around the uniform chain ($d_1 = d_2 = 3a_0$); once again, the HL approach offers a better approximation than UHF.

The highest 50 energy states from the full CI result are shown in Fig. 13 and and the Fock matrix eigenvalues in Fig. 14. In all three directions the ground state is nondegenerate on short-long side (small d_1), although this is not clearly visible from Fig. 13(a) for the [001] direction; as found for smaller spacings in Sec. III B 1, the ground state joins three other states in each case and forms a fourfold-degenerate manifold on the right-hand side (large d_1). This time there is no change in the character of the ground state and no (anti)crossing visible among the states in Fig. 13 or 14; instead, the Fock eigenvalues show a group of four occupied states strongly separated from the unoccupied ones by the self-consistent potential. There are some small splittings visible among the eigenvalues in Fig. 14 at particular geometries; these are due to the loss of symmetry in the UHF solution, as discussed in Sec. III B 1.

To compare the fourfold-degenerate many-hole ground states obtained in the long-short limit for the small- and large-separation cases, and to understand how they relate to our previous results for noninteracting holes [12], we show in Fig. 15 the energy difference between the ground state and three closest excited states as a function of Coulomb interaction strength for a four-acceptor linear chain in the [001] direction (interpolating between the noninteracting and fully interacting cases). We choose the four-acceptor system because it provides a more straightforward comparison to the one-hole edge states of the noninteracting system, as there will be fewer other states complicating the picture [12]. In both cases, there is a gap in the noninteracting limit because one-hole edge states move apart in the long-short limit to join two different bulk bands as shown in our previous paper [12]; the fourfold-degenerate ground state forms once the interaction strength exceeds a critical value, which is smaller in the large-separation case than in the small-separation case. This can be understood because the energy scale set by the noninteracting part of the Hamiltonian is weaker in the large-separation case, so a smaller hole-hole interaction is sufficient to overcome the parabolic confining potential.

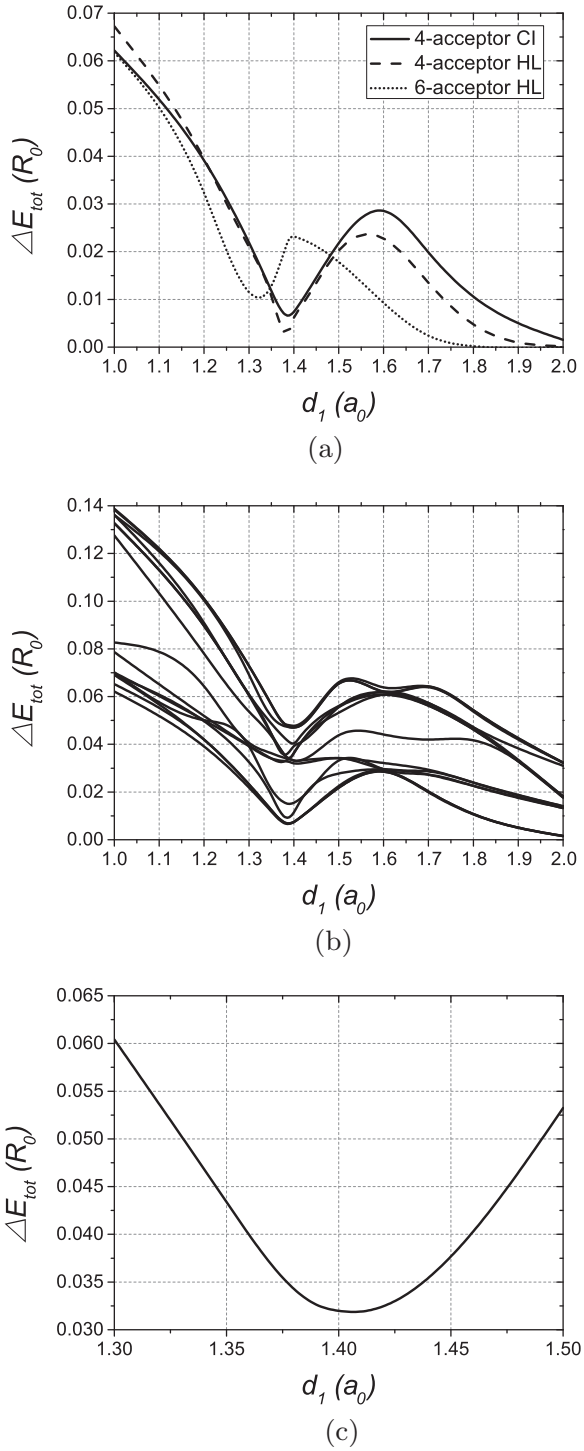


FIG. 9. (a) The differences between the total-energy ground state and the first excited state in different systems under different models in the [001] direction when $d_1 + d_2 = 3a_0$: the solid line is for the four-acceptor full CI calculation, the dashed line is for the four-acceptor HL calculation, the dotted line is for the six-acceptor HL calculation. (b) The differences between the total-energy ground state and the first 15 excited states for the full CI calculation in the [001] direction when $d_1 + d_2 = 3a_0$. (c) The differences between the old and new total-energy ground state during the anticrossing for the full CI calculation in the [001] direction when $d_1 + d_2 = 3a_0$.

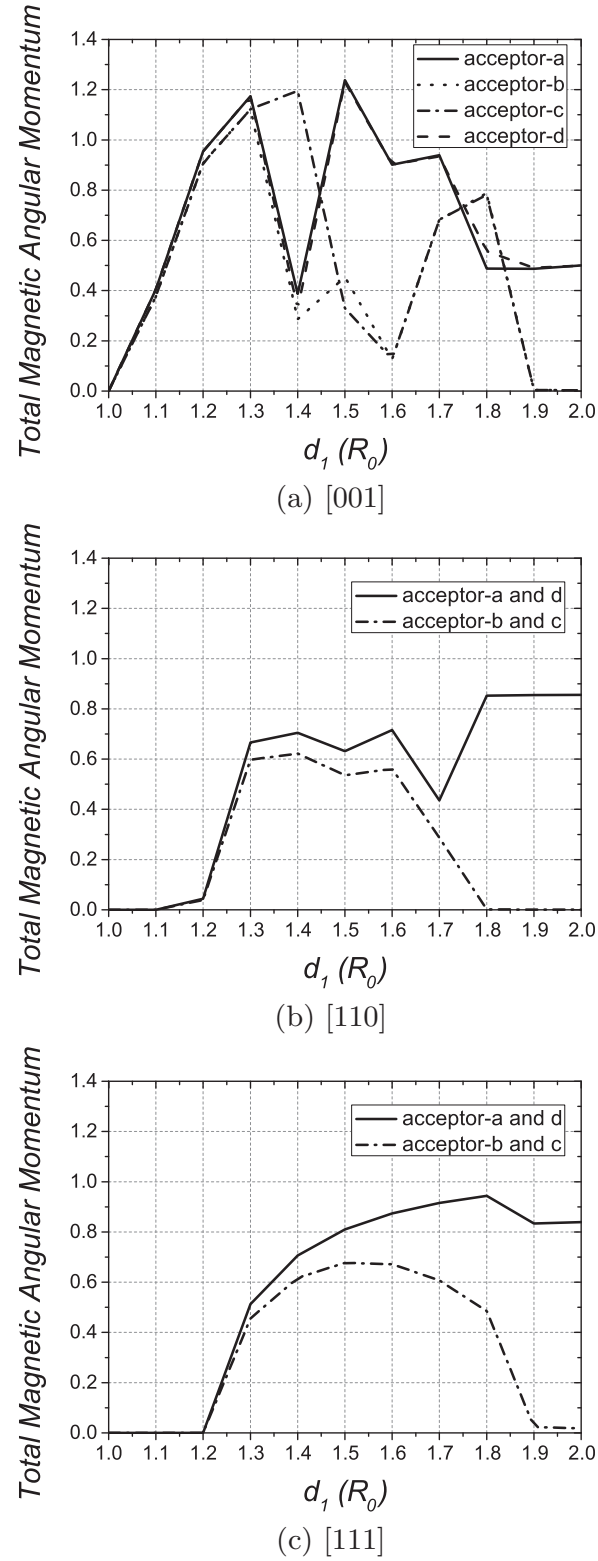


FIG. 10. The behavior of the total magnetic angular momentum obtained from the UHF method under different arrangements in three typical directions for the small-separation case ($d_1 + d_2 = 3a_0$) of the four-acceptor linear chain: (a) the [001] direction, (b) the [110] direction, (c) the [111] direction.

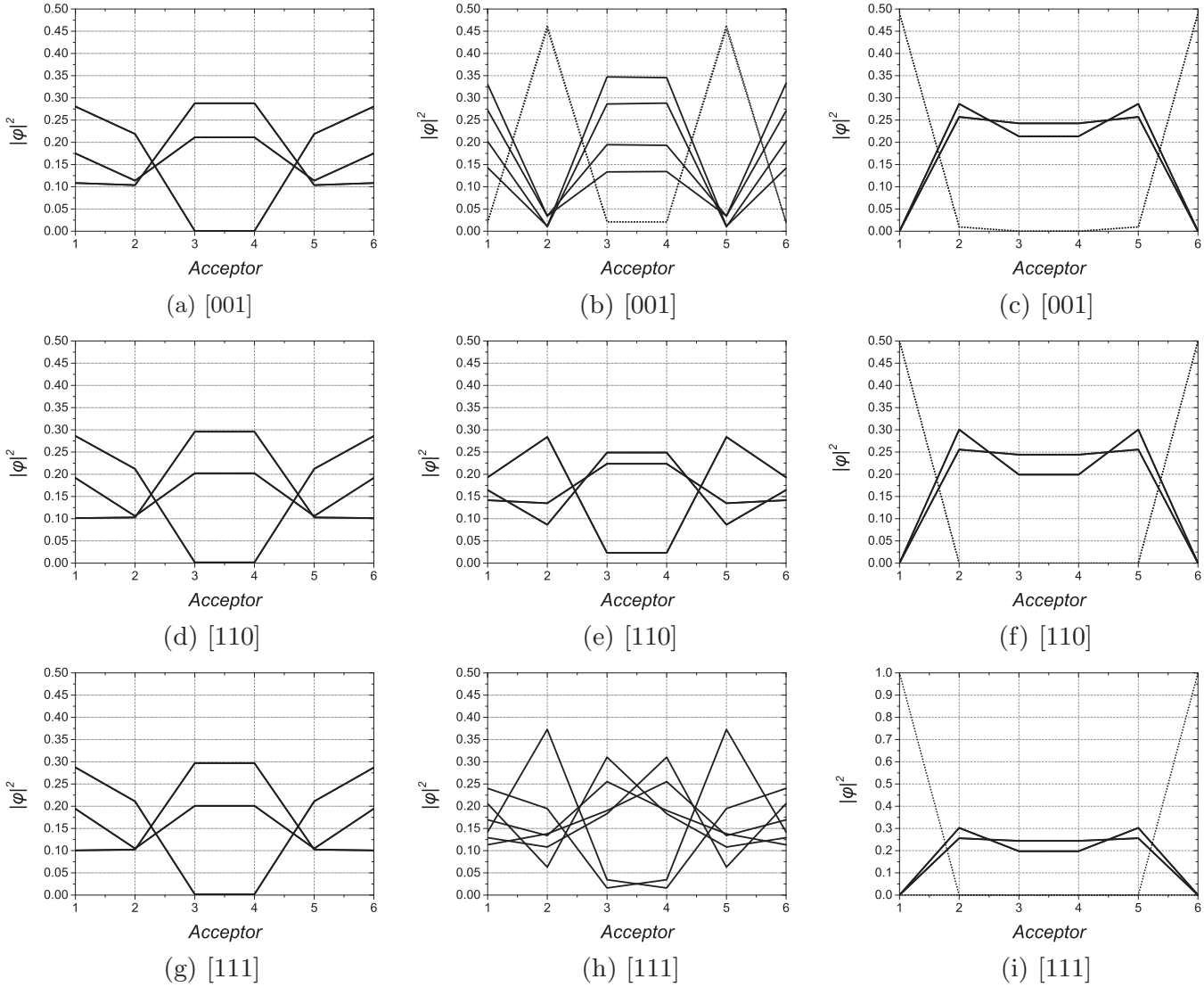
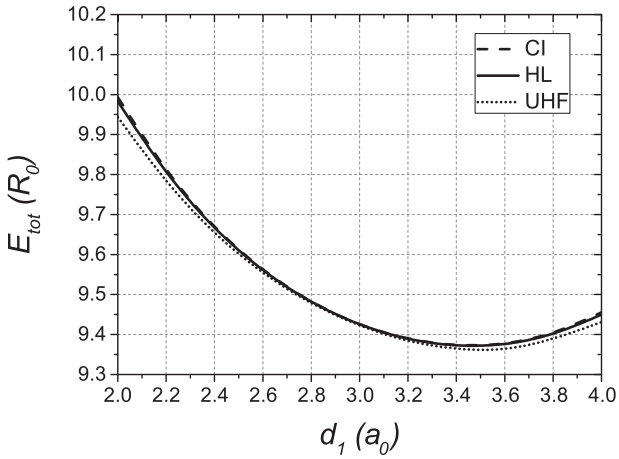


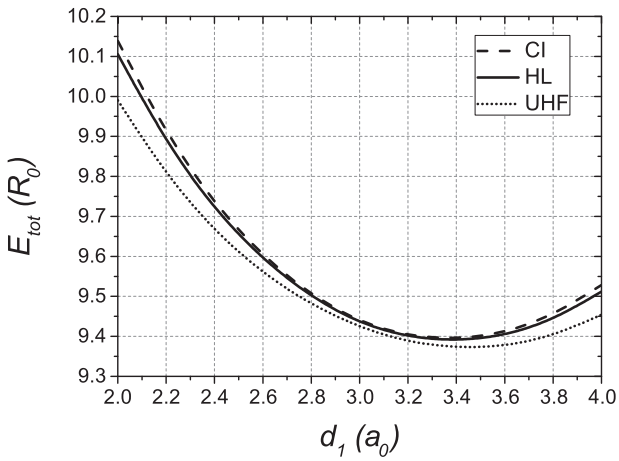
FIG. 11. The total charge distribution among acceptors under different arrangements in three typical directions for the six-acceptor linear chain when $d_1 + d_2 = 3a_0$: (a) the short-long arrangement in the [001] direction, (b) the uniform chain case in the [001] direction, (c) the long-short arrangement in the [001] direction, (d) the short-long arrangement in the [110] direction, (e) the uniform chain case in the [110] direction, (f) the long-short arrangement in the [110] direction, (g) the short-long arrangement in the [111] direction, (h) the uniform chain case in the [111] direction, (i) the long-short arrangement in the [111] direction. For (c), (f), and (i), the dotted lines are for the states localized at the acceptors at the end of the chain, which are always the highest states among the states involved to form the total-energy ground state here. All the lines in (a), (c), (d), (e), (f), (g), the dotted line in (b), and the solid lines in (i) are doubly degenerate.

3. Symmetry breaking in the UHF calculation

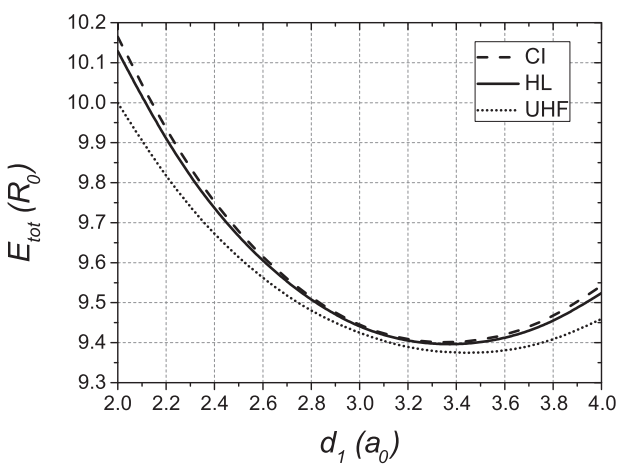
To investigate the symmetry breaking, we determined the symmetry of the one-hole reduced density matrices, both in the full CI case and after the convergence of the UHF calculation; the results are shown using the Hermann-Mauguin notation for magnetic point groups in the upper part of Table II for the small-separation case, and in the lower part for the large-separation case. We observe that for small separations, the UHF solution always begins (for small d_1) with the same symmetry as the CI calculation (and the core Hamiltonian). This is a "gray" magnetic group that contains the time-reversal operation $1'$, meaning that no magnetic moment has developed. The group then loses some symmetry elements as d_1 increases, as magnetic moments develop; it would be more accurate to describe these missing symmetry operations as "hidden" rather than "lost" because they map different members of a manifold of degenerate self-consistent solutions to the UHF equations, each individually having lower symmetry, onto one another. At the points in the [001] direction where the symmetry is lowest ($d_1 = 1.4a_0$, $1.5a_0$, and $1.8a_0$), the convergence of the SCF procedure is poorer than for other separations; Figs. 7(a) and 7(b) show that these correspond to the location of the crossing points between different many-body ground-state compositions. Bearing in mind that the crossing points for six acceptors [Figs. 8(a) and 8(b)] are in slightly different positions than for four acceptors, and the slight change of the crossing point location shown in Fig. 9, it is reasonable to believe that $d_1 = 1.8a_0$ is the true location of



(a) [001]

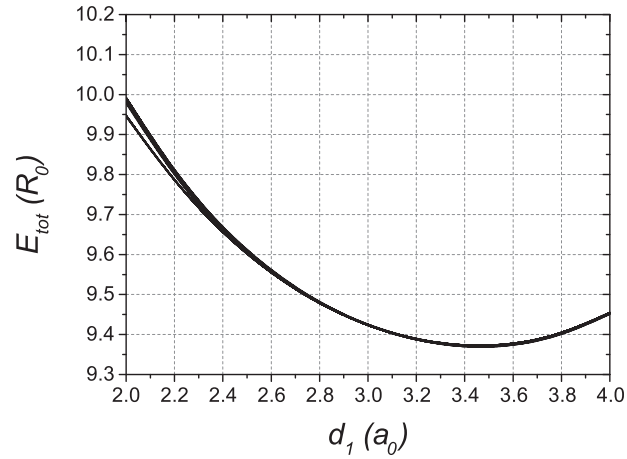


(b) [110]

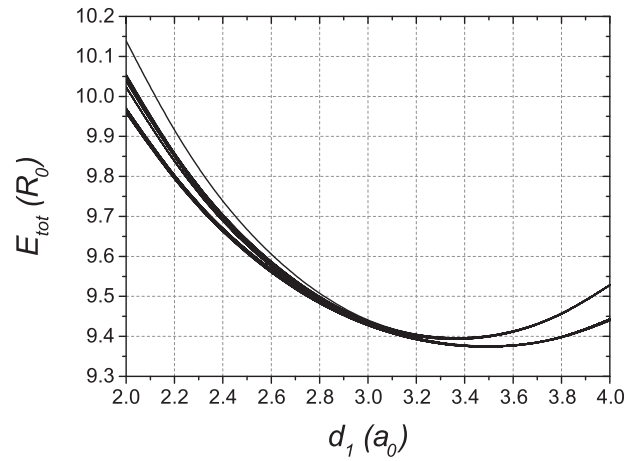


(c) [111]

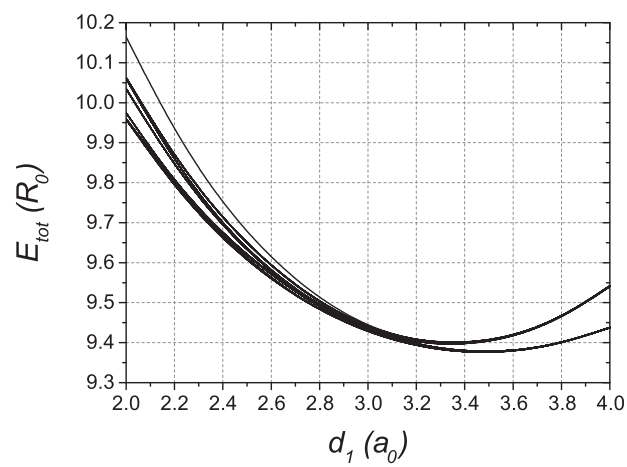
FIG. 12. The total energy of the ground state in three high-symmetry directions for the large-separation case ($d_1 + d_2 = 6a_0$) of the four-acceptor linear chain: (a) the [001] direction, (b) the [110] direction, (c) the [111] direction. The dashed line is for the full CI calculation, the solid line is for the HL approximation, the dotted line is for the UHF method.



(a) [001]

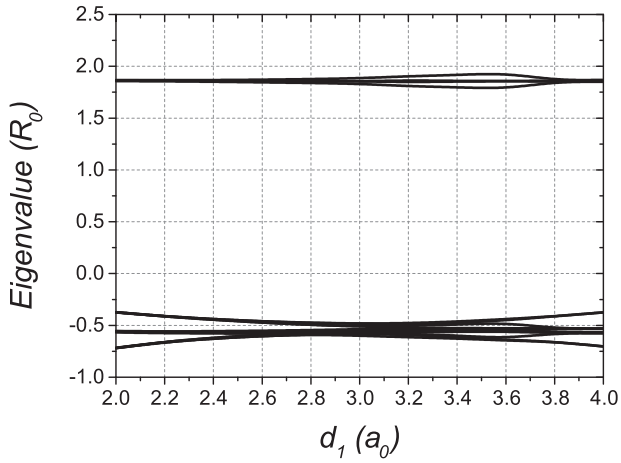


(b) [110]

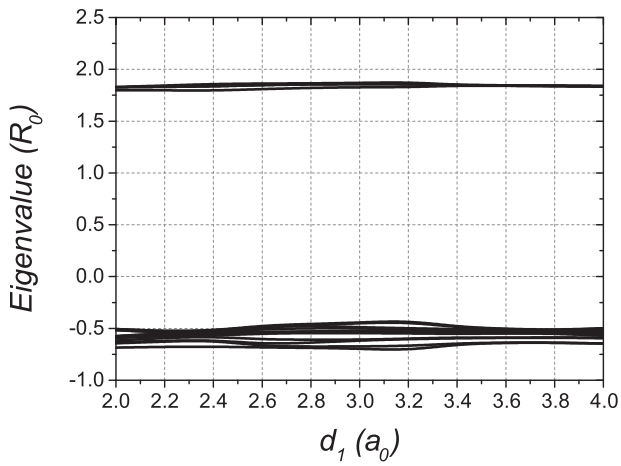


(c) [111]

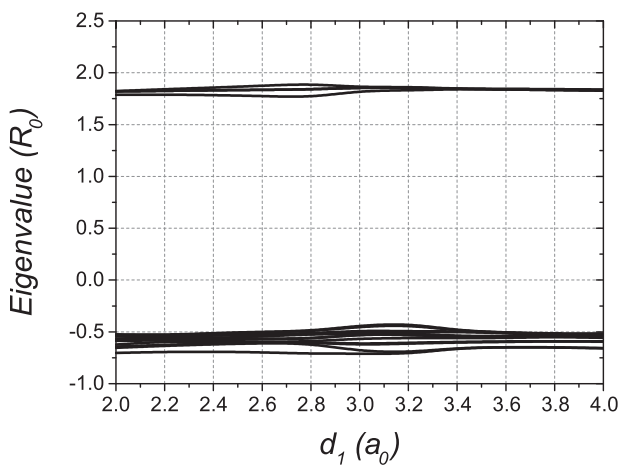
FIG. 13. The total energies of the highest 50 states of the full CI result in three high-symmetry directions for the large-separation case ($d_1 + d_2 = 6a_0$) of the four-acceptor linear chain: (a) the [001] direction, (b) the [110] direction, (c) the [111] direction.



(a) [001]

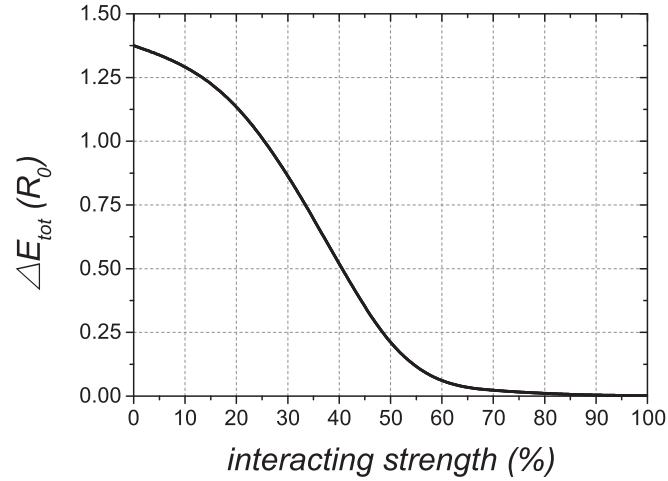


(b) [110]

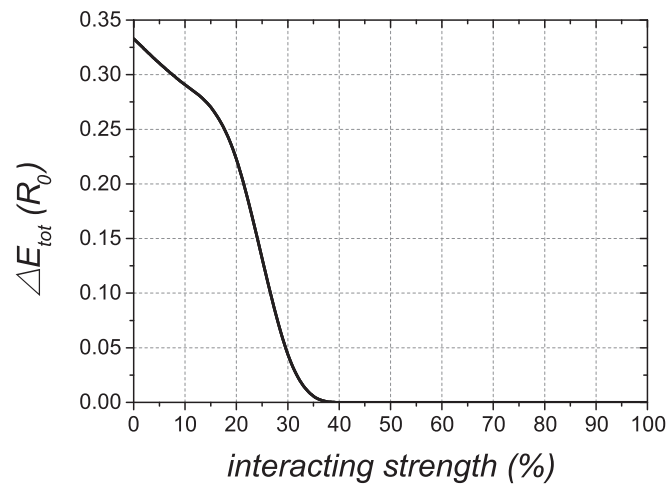


(c) [111]

FIG. 14. The Fock matrix eigenvalues obtained from the UHF method in three high-symmetry directions for the large-separation case ($d_1 + d_2 = 6a_0$) of the four-acceptor linear chain: (a) the [001] direction, (b) the [110] direction, (c) the [111] direction.



(a) $d_1 = 2a_0, d_2 = 1a_0$



(b) $d_1 = 4a_0, d_2 = 2a_0$

FIG. 15. The energy difference between the ground state and three closest excited states as a function of Coulomb interaction strength (expressed as a percentage) in the long-short limit of the four-acceptor linear chain in the [001] direction: (a) the small-separation case ($d_1 = 2a_0,\ d_2 = 1a_0$), (b) the large-separation case ($d_1 = 4a_0,\ d_2 = 2a_0$).

a second crossing point which is not shown in Figs. 7(a) and 7(b) due to the limited number of acceptors. This suggests that the further reduction in symmetry near these points may also be related to the crossings (or potential crossings) between the occupied eigenvalues of the Fock matrix. We also show the behavior of the total magnetic angular momentum for each acceptor in the different chain orientations in Fig. 10; the breaking of symmetry is reflected by splitting into two or (at the lowest-symmetry arrangements the [001] direction) four different inequivalent sets. The magnetization pattern shows that nonzero magnetization becomes increasingly concentrated at the ends of the chain as d_1 increases, which is also true in the large-separation case. The six-acceptor system

TABLE II. The magnetic symmetry groups of the UHF ground states in different arrangements for the three high-symmetry directions in Hermann-Mauguin notation. Here the prime denotes operations that are only symmetries when accompanied by time reversal; the symbols m and m' are abbreviations for $\frac{1}{m}$ and $\frac{1}{m'}$, respectively.

d_1	[001]	[110]	[111]
Hamiltonian	$\frac{4}{m}\frac{2}{m}\frac{2}{m}1'$	$\frac{2}{m}\frac{2}{m}\frac{2}{m}1'$	$\bar{3}\frac{2}{m}1'$
Small-separation case ($d_1 + d_2 = 3a_0$)			
$1.0a_0$	$\frac{4}{m}\frac{2}{m}\frac{2}{m}1'$	$\frac{2}{m}\frac{2}{m}\frac{2}{m}1'$	$\bar{3}\frac{2}{m}1'$
$1.1a_0$	$\frac{4}{m'}\frac{2}{m'}\frac{2}{m'}$	$\frac{2}{m}\frac{2}{m}\frac{2}{m}1'$	$\bar{3}\frac{2}{m}1'$
$1.2a_0$	$\frac{4}{m'}\frac{2}{m'}\frac{2}{m'}$	$\frac{2'}{m}\frac{2'}{m}\frac{2}{m'}$	$\bar{3}\frac{2}{m}1'$
$1.3a_0$	$\frac{4}{m'}\frac{2}{m'}\frac{2}{m'}$	$\frac{2'}{m}\frac{2'}{m}\frac{2}{m'}$	$\frac{2}{m'}$
$1.4a_0$-$1.5a_0$	$4m'm'$	$\frac{2'}{m}\frac{2'}{m}\frac{2}{m'}$	$\frac{2}{m'}$
$1.6a_0$-$1.7a_0$	$\frac{4}{m'}\frac{2}{m'}\frac{2}{m'}$	$\frac{2'}{m}\frac{2'}{m}\frac{2}{m'}$	$\frac{2}{m'}$
$1.8a_0$	$4m'm'$	$\frac{2}{m}\frac{2'}{m'}\frac{2'}{m'}$	$\frac{2}{m'}$
$1.9a_0$-$2.0a_0$	$\frac{4}{m}\frac{2'}{m'}\frac{2'}{m'}$	$\frac{2}{m}\frac{2'}{m'}\frac{2'}{m'}$	$\frac{2'}{m'}$
Large-separation case ($d_1 + d_2 = 6a_0$)			
$2.0a_0$-$3.2a_0$	$\frac{4}{m}\frac{2'}{m'}\frac{2'}{m'}$	$\frac{2}{m}\frac{2'}{m'}\frac{2'}{m'}$	$\frac{2'}{m'}$
$3.4a_0$-$3.6a_0$	$\frac{4}{m}\frac{2'}{m'}\frac{2'}{m'}$	$2'mm'$	m'
$3.8a_0$-$4.0a_0$	$4m'm'$	$2'mm'$	m'

behaves similarly to the four-acceptor system, so we do not show the results here.

In the large-separation case, the symmetry is broken with respect to the underlying Hamiltonian at *all* separations. As previously, the broken symmetries are not really lost, but now map different solutions within the manifold of degenerate states (all having nonzero magnetic moments) into one another. For values of d_1 greater than some critical value (which depends on the direction), the symmetry is further reduced; comparing with Fig. 13, we see this further reduction occurs when the fourfold-degenerate ground states in the full CI calculation show very small energy differences between each other so they are hard to distinguish in the UHF calculation. By checking each data point, the switching from ferromagnetically aligned case to antiferromagnetically aligned case is found across the central (d_2) bond as it shortens.

C. Linear chain with periodic boundary conditions

We now turn to periodic boundary conditions. A schematic of the system is shown in Fig. 16; (a, b, c, d) label four adjacent acceptors, with b, c in the same unit cell, and d_1, d_2 are the separations. (We have swapped the separation labels relative to the convention used in in our previous paper [12].) Approaches based on full diagonalization (full CI calculation and the HL approach) are not extensive and hence not useful with periodic boundary conditions as discussed in Sec. II A 1,

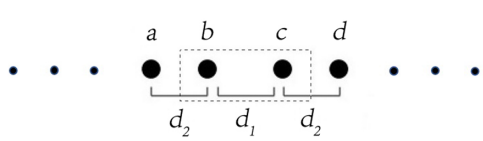


FIG. 16. The schematic of the linear chain with periodic boundary conditions. a, b, c, d are the labels of acceptors, $d_1 < d_2$ is known as the "short-long arrangement," $d_1 > d_2$ is known as the "long-short arrangement."

but the UHF method is still suitable. Since the behavior of finite chains is found to be quite similar in the small- and large-separation cases, we report results for infinite chains only for smaller ($d_1 + d_2 = 3a_0$).

Figure 17 shows the band structures of the Fock matrix eigenvalues. We only show the results for the short-long arrangement ($d_1 < d_2$) in each high-symmetry direction, along with the uniform chains ($d_1 = d_2$), as the short-long arrangements are equivalent to long-short under periodic boundary conditions. All the single-hole states are doubly degenerate, so the two states at the top of each picture will be filled (two holes per cell). There is a large gap between the filled and empty bands in the short-long dimerized arrangement; for uniform chains, the bands move closer but this gap does not close, showing the existence of a cell-doubling perturbation from the self-consistent field. This is related to the broken symmetries found in the corresponding finite chain calculations: as shown in Table II, we found the inversion symmetry is broken (becomes hidden) for some uniform-chain cases. It is reasonable that this also occurs under periodic boundary conditions, leading to an inequivalence of the two atoms in the cell even for a uniform chain and implying that the band structure of the two-atom cell cannot be obtained by simply folding the bands for the one-atom cell.

D. Structure of the edge states

In order to understand the nature of the edge states, we examine the many-hole states from the full CI calculation and compare them to the UHF single-particle states, for both small-separation and large-separation cases in the four-acceptor finite chain. Both methods show edge states localized at the acceptors at the end of the chain in the long-short arrangement ($d_1 > d_2$); however, the signatures are different. The CI method shows a manifold of almost degenerate states spanned by a basis of the form

$$|\psi_{m,n}\rangle = \left|\psi_m^A\right\rangle \otimes |\psi^{\text{bulk}}\rangle \otimes \left|\psi_n^B\right\rangle, \tag{14}$$

where A labels the left end of the chain [acceptor a in Fig. 4 (a)], B labels the right end [acceptor d in Fig. 4(a)], and $|\psi^{\text{bulk}}\rangle$ is a common state residing in the interior of the chain [acceptors b and c in Fig. 4(a)]. The indices m and n label different states of the ends, and the pair (m, n) together label a member of the almost degenerate manifold. The transformation from state $|\Psi_{m,n}\rangle$ to $|\Psi_{m',n'}\rangle$ can therefore be carried out by a unitary operator

$$\hat{U} = \hat{U}^A \otimes \hat{1}^{\text{bulk}} \otimes \hat{U}^B$$

with

$$\hat{U}^A = \left|\psi_{m'}^A\right\rangle\left\langle\psi_m^A\right|,$$
$$\hat{U}^B = \left|\psi_{n'}^B\right\rangle\left\langle\psi_m^B\right|. \tag{15}$$

For finite chains, the eigenstates are particular linear combinations of the $|\psi_{m,n}\rangle$ which are almost (but not quite) degenerate; the splittings decay to zero as d_1 is increased, or as the chain becomes longer (see Fig. 9). It is therefore important to look at the whole space spanned by the $|\psi_{m,n}\rangle$, especially when the splittings become very small. The UHF method instead picks out a single symmetry-broken many-hole ground state

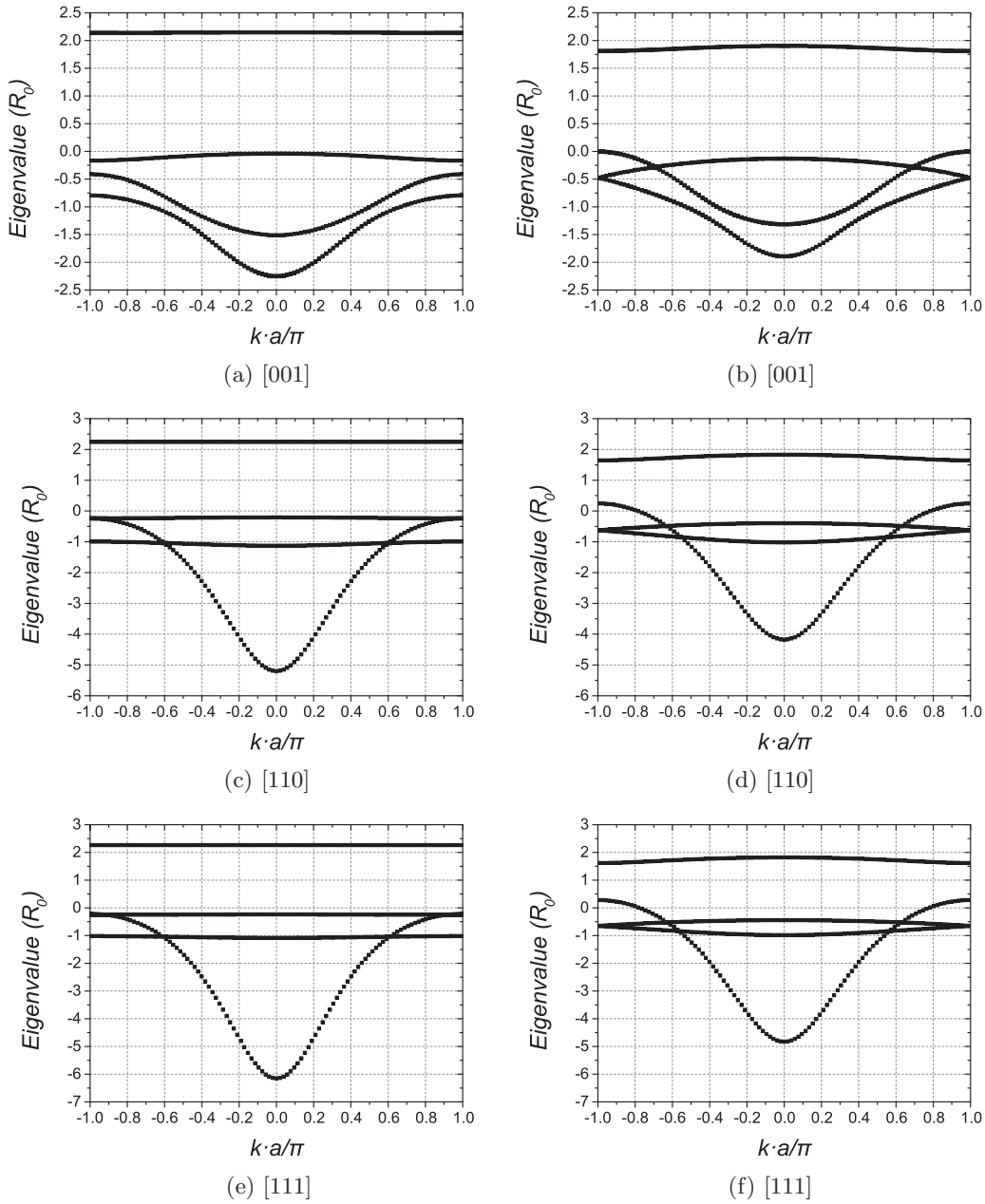


FIG. 17. The band structures of the Fock matrix eigenvalues under different arrangements in three typical directions when $d_1 + d_2 = 3a_0$: (a) the short-long arrangement in the [001] direction, (b) the uniform chain case in the [001] direction, (c) the short-long arrangement in the [110] direction, (d) the uniform chain case in the [110] direction, (e) the short-long arrangement in the [111] direction, (f) the uniform chain case in the [111] direction.

in which one pair of occupied single-particle states is localized at the chain ends [acceptors a and d in Fig. 4(a)] while the other pair is spread over the interior [acceptors b and c in Fig. 4(a)]. The single-hole edge states can be written as linear combinations of particular one-hole kets $|\phi^A\rangle$ and $|\phi^B\rangle$ localized at either end.

We can also examine the symmetries of the edge states $|\psi_m^A\rangle, |\psi_n^B\rangle$ in the light of the classification of the topological phases of one-dimensional interacting fermions proposed by Ref. [29]; in the long-short limit we find the characteristic phases are $(\mu = 0, \phi = 0, \kappa = \pi)$, hence, the state is topologically nontrivial with fourfold degeneracy, while in the short-long arrangement they are $(\mu = 0, \phi = 0, \kappa = 0)$ (topologically trivial, nondegenerate). However, we find some differences between the small- and large-separation cases. For the four-acceptor chain, when $d_1 + d_2 = 3a_0$, $|\psi_m^A\rangle, |\psi_n^B\rangle$ involve only $m_F = \pm\frac{1}{2}$ states in the [001] direction, while $|\psi^{\text{bulk}}\rangle$ includes only $m_F = \pm\frac{3}{2}$ states. This is because in the long-short arrangement case, the system can be considered as two single acceptors at the chain ends and a closely coupled pair of acceptors between them. In that case, $|\psi^{\text{bulk}}\rangle$ is dominated by the central pair, while $|\psi_m^A\rangle, |\psi_n^B\rangle$ are dominated by the single-acceptor ends. Since the doubly degenerate occupied bands at the top of Figs. 17(a) and 17(b) in the [001] direction are always formed predominantly from linear combinations of the $\pm\frac{3}{2}$ states on the two acceptors in the cell, and

a single acceptor perturbed by another acceptor always has a ground state of $m_F = \pm\frac{1}{2}$ symmetry, it is reasonable that $|\psi^{\text{bulk}}\rangle$ and $|\psi_m^A\rangle$, $|\psi_n^B\rangle$ only involve $m_F = \pm\frac{3}{2}$ and $\pm\frac{1}{2}$ states, respectively. When $d_1 + d_2 = 6a_0$, although $|\psi^{\text{bulk}}\rangle$ is similar, the $\{|\psi_m^A\rangle, |\psi_n^B\rangle\}$ involve also the superpositions $\pm\{|\frac{3}{2}\rangle, |\frac{1}{2}\rangle\}$ in two of the four degenerate states. In the large-separation case the degeneracy is more nearly exact, so the properties of individual eigenstates are not clearly defined and we should consider the space spanned by all four degenerate states together. For the six-acceptor system (which we treat in the HL approximation), we always find the edge states composed purely of $m_F = \pm\frac{1}{2}$ states at the end of the chain (as for four acceptors).

We can also calculate the Zak phase for the occupied UHF states in the infinite system by using (11) in Sec. II A 5. We find that Zak phase is 0 for all arrangements in all directions, even though we find the edge states in the finite chains have nontrivial symmetries; this is consistent with the preservation of a gap in the single-particle UHF energy spectrum for all arrangements. The Zak phase is calculated by using the single-hole UHF eigenvectors, and its correspondence with the topological properties of an interacting system is still unclear; it is not surprising that it fails to describe the topological properties of the interacting system in the same way, as was previously noted for the bosonic case [30]. In the absence of a rigorously defined topological quantum number for an infinite system with interactions, the direct study of the quantum numbers characterizing the edge states of the finite system, introduced by Turner *et al.* [29], provides a better insight into their topological nature.

IV. CONCLUSION

In this paper, we constructed multihole models for neutral, one-dimensional multiacceptor chains based on three different methods: full configuration interaction, the Heitler-London approximation, and the unrestricted Hartree-Fock method. The HL approximation solves some of the problems with the CI method, but only the UHF method is able to cope with infinite chains under periodic boundary conditions.

From reference calculations on a pair of acceptors, we found that both the HL approach and the UHF method give good approximations to the ground state of the full CI calculation, with the HL approach offering a better result in the regimes studied (which are on the insulating side of the Mott transition). The UHF method is less useful for the calculation of excited states, so we use the HL approximation to simplify the calculation of low-lying excitations when interactions are strong. The converged UHF state has a large gap between the filled and empty states, due to the self-consistent potential generated by the hole-hole interactions.

For finite chains, the CI ground state is nondegenerate in the short-long arrangement in all directions, but joins three other states to form a fourfold-degenerate manifold in the long-short arrangement, which is followed in energy by an eightfold-degenerate state and another fourfold-degenerate state. By checking the dominant components of these 16 states, we found that only the levels on the acceptors at the end of the chain change between different members of the manifold; the overall 16-fold degeneracy comes from the product of separate sets of four levels on each end acceptor. The topological nature of these edge states is confirmed by the presence of nontrivial phases in the classification of one-dimensional fermion edge states by Turner *et al.* In the small-separation case where $d_1 + d_2 = 3a_0$, an anticrossing occurs between the ground state and the next excited states in the [001] direction, resulting in a switch from an unhybridized ground state dominated by $m_F = \pm\frac{3}{2}$ states to a hybridized state where $m_F = \pm\frac{1}{2}$ states are also present; this transition is related to the crossing between the filled UHF single-particle states. The UHF solution loses part of the symmetry of the underlying Hamiltonian; for particular arrangements, we found the further broken symmetries related to the crossing (or potential crossing) of Fock matrix eigenstates in the [001] direction. The loss of symmetry corresponds to the emergence of static moments on each acceptor in the UHF approach.

We obtained the UHF band structures of the Fock matrix eigenvalues. We found there is a large gap between the filled and empty states in a dimerized chain, which does not fully close in the uniform chain, showing the existence of a period-doubling perturbation. Since a gap is maintained throughout the transition from short-long to long-short arrangements, the Zak phase is constant (and equal to zero), despite the observation of nontrivial many-body edge states in the long-short case. Hence, this method does not capture the formation of edge states, while the previous method introduced by Turner *et al.* can well characterize their topological properties. The nature of the bulk-edge correspondence in such interacting systems requires further investigation.

ACKNOWLEDGMENTS

We wish to acknowledge the support of the Engineering and Physical Sciences Research Council and UK Research and Innovation under the ADDRFSS programme (Grant No. EP/M009564/1). We thank N. Le, B. Murdin, N. Curson, and G. Aeppli for helpful and inspiring discussions.

[1] B. Koiller, X. Hu, and S. Das Sarma, Phys. Rev. Lett. **88**, 027903 (2001).
[2] S. T. Pantelides and C. T. Sah, Phys. Rev. B **10**, 621 (1974).
[3] J. Salfi, M. Tong, S. Rogge, and D. Culcer, Nanotechnology **27**, 244001 (2016).
[4] N. W. Hendrickx, W. I. L. Lawrie, M. Russ, F. van Riggelen, S. L. de Snoo, R. N. Schouten, A. Sammak, G. Scappucci, and M. Veldhorst, Nature (London) **591**, 580 (2021).
[5] Z. Wang, E. Marcellina, A. R. Hamilton, J. H. Cullen, S. Rogge, J. Salfi, and D. Culcer, npj Quantum Inf. **7**, 54 (2021).
[6] J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).
[7] A. Baldereschi and N. O. Lipari, Phys. Rev. B **8**, 2697 (1973).
[8] A. Baldereschi and N. O. Lipari, Phys. Rev. B **9**, 1525 (1974).

[9] N. Lipari and A. Baldereschi, Solid State Commun. **25**, 665 (1978).
[10] A. C. Durst, K. E. Castoria, and R. N. Bhatt, Phys. Rev. B **96**, 155208 (2017).
[11] A. C. Durst, G. Yang-Mejia, and R. N. Bhatt, Phys. Rev. B **101**, 035202 (2020).
[12] J. Zhu, W. Wu, and A. J. Fisher, Phys. Rev. B **101**, 085303 (2020).
[13] S. R. Schofield, N. J. Curson, M. Y. Simmons, F. J. J Rueß, T. Hallam, L. Oberbeck, and R. G. Clark, Phys. Rev. Lett. **91**, 136104 (2003).
[14] K. J. Dwyer, S. Baek, A. Farzaneh, M. Dreyer, J. R. Williams, and R. E. Butera, ACS Appl. Mater. Interfaces **13**, 41275 (2021).
[15] P. Clauws, J. Broeckx, E. Rotsaert, and J. Vennik, Phys. Rev. B **38**, 12377 (1988).
[16] N. Q. Vinh, B. Redlich, A. F. G. van der Meer, C. R. Pidgeon, P. T. Greenland, S. A. Lynch, G. Aeppli, and B. N. Murdin, Phys. Rev. X **3**, 011019 (2013).
[17] P. Dai, Y. Zhang, and M. P. Sarachik, Phys. Rev. B **45**, 3984 (1992).
[18] K. Litvinenko, E. Bowyer, P. Greenland, N. Stavrias, J. Li, R. Gwilliam, B. Villis, G. Matmon, M. Pang, B. Redlich *et al.*, Nat. Commun. **6**, 6549 (2015).
[19] J. van der Heijden, T. Kobayashi, M. G. House, J. Salfi, S. Barraud, R. Laviéville, M. Y. Simmons, and S. Rogge, Sci. Adv. **4**, eaat9199 (2018).
[20] A. Corna, L. Bourdet, R. Maurand, A. Crippa, D. Kotekar-Patil, H. Bohuslavskyi, R. Lavieville, L. Hutin, S. Barraud, X. Jehl *et al.*, npj Quantum Inf. **4**, 6 (2018).
[21] A. Crippa, R. Maurand, L. Bourdet, D. Kotekar-Patil, A. Amisse, X. Jehl, M. Sanquer, R. Laviéville, H. Bohuslavskyi, L. Hutin *et al.*, Phys. Rev. Lett. **120**, 137702 (2018).
[22] A. Crippa, R. Ezzouch, A. Aprá, A. Amisse, R. Lavieville, L. Hutin, B. Bertrand, M. Vinet, M. Urdampilleta, T. Meunier *et al.*, Nat. Commun. **10**, 2776 (2019).
[23] J. Salfi, J. Mol, R. Rahman, G. Klimeck, M. Simmons, L. C. L. Hollenberg, and S. Rogge, Nat. Commun. **7**, 11342 (2016).
[24] J. Zak, Phys. Rev. Lett. **62**, 2747 (1989).
[25] E. Clementi and D. Davis, J. Comput. Phys. **1**, 223 (1966).
[26] A. Szabo, A. Szabó, and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Macmillan, New York, 1982).
[27] N. H. Le, A. J. Fisher, N. J. Curson, and E. Ginossar, npj Quantum Inf. **6**, 24 (2020).
[28] L. Fidkowski and A. Kitaev, Phys. Rev. B **81**, 134509 (2010).
[29] A. M. Turner, F. Pollmann, and E. Berg, Phys. Rev. B **83**, 075102 (2011).
[30] F. Grusdt, M. Höning, and M. Fleischhauer, Phys. Rev. Lett. **110**, 260405 (2013).