Electron-hole correlations in semiconductor quantum dots with tight-binding wave functions

Seungwon Lee, Lars Jönsson, and John W. Wilkins
Department of Physics, Ohio State University, Columbus, Ohio 43210-1106

Garnett W. Bryant
National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8423

Gerhard Klimeck
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109
(Received 14 November 2000; published 27 April 2001)

The electron-hole states of semiconductor quantum dots are investigated within the framework of empirical tight-binding descriptions for Si, as an example of an indirect-gap material, and InAs and CdSe as examples of typical III-V and II-VI direct-gap materials. We significantly improve the energies of the single-particle states by optimizing tight-binding parameters to give the best effective masses. As a result, the calculated excitonic gaps agree within 5% error with recent photoluminescence data for Si and CdSe but they agree less well for InAs. The electron-hole Coulomb interaction is insensitive to different ways of optimizing the tight-binding parameters. However, it is sensitive to the choice of atomic orbitals; this sensitivity decreases with increasing dot size. Quantitatively, tight-binding treatments of Coulomb interactions are reliable for dots with radii larger than 15–20 Å . Further, the effective range of the electron-hole exchange interaction is investigated in detail. In quantum dots of the direct-gap materials InAs and CdSe, the exchange interaction can be long ranged, extending over the whole dot when there is no local (onsite) orthogonality between the electron and hole wave functions. By contrast, for Si quantum dots the extra phase factor due to the indirect gap effectively limits the range to about 15 Å, independent of the dot size.

DOI: 10.1103/PhysRevB.63.195318 PACS number(s): 73.22.−f, 71.35.−y, 71.70.Gm, 71.15.Ap

I. INTRODUCTION

The optical and electronic properties of semiconductor quantum dots have been studied both experimentally[1–10] and theoretically[11–20] for a wide range of sizes, shapes, and materials. This work is stimulated both by a fundamental interest in quantum-confined systems and by the applicability of quantum dots in nanoscale devices. Experimentally, significant recent improvements in both growth techniques[21] and single-dot spectroscopy[22] have enabled detailed studies of the energy spectra of electron-hole complexes (excitons, biexcitons, trions, etc.). Theoretically, the most sophisticated theoretical approaches are multiband effective-mass theory,[4,11] empirical pseudopotential theory,[12–14] tight-binding methods,[15–19] and quasiparticle calculations[20] with the GW approximation.[23]

Quantum dots are intermediate between molecular and bulk systems. This is reflected in the different theoretical approaches; effective-mass theory treats the dot as a confined bulk system whereas pseudopotential theory aims at a detailed atomistic description of the wave functions. Tight-binding theory compromises between these two approaches by including an atomistic description but limiting the local degrees of freedom to a small basis set. Therefore, the computationally less costly tight-binding method can be used to study large quantum dots, up to 25 nm size, without severely restricting atomic-scale variations in the wave functions. However, since the tight-binding matrix elements are empirically optimized without introducing any specific atomic orbitals, there is no direct way to calculate other matrix elements such as Coulomb and exchange matrix elements. Therefore, calculations involving electron-hole interactions require a selection of specific atomic orbitals that cannot be explicitly related to the tight-binding parameters. Hence, a key question that we address in this work is how sensitive the calculated electron-hole properties are to specific choices of orbitals used to calculate the Coulomb matrix elements.

We study spherical semiconductor crystallites centered around an anion atom in a zinc-blende structure. We choose Si, InAs, and CdSe as examples of an indirect-gap material, and typical III-V and II-VI direct-gap materials, respectively. Experimentally, the low-lying exciton energies have been measured for dots up to 20-Å radius in Si,[6] and up to 40-Å radius in InAs (Ref. 4) and CdSe.[3]

We use the empirical nearest-neighbor sp^3s^* tight-binding model[24] for the electron and hole single-particle wave functions. In order to calculate electron-hole Coulomb and exchange matrix elements, we describe the real-space atomic basis orbitals s, p_x, p_y, p_z, and s^* with Slater orbitals as a starting point. Both the Coulomb and exchange interaction are screened by a dielectric function depending on both dot size and the distance between the particles. The energies of the electron-hole states are obtained by diagonalizing the configuration-interaction matrix generated by the lowest-lying electron and hole states with typical convergence of a few meV. We examine the sensitivity of the electron-hole energies on both the choices of atomic basis orbitals and the tight-binding parameters.

Within the tight-binding model, the single-particle Hamiltonian can be improved by either increasing the number of basis orbitals or including interactions between more distant atoms. An s^* orbital was first introduced by Vogl *et al.*[24] to improve the conduction bands near the X point. To some

degree, an s^* orbital can mimic band-structure effects that should be attributed to d bands. However, it is likely that d orbitals need to be added in some materials.[19] An alternative is to add next-nearest-neighbor interactions[15] within the sp^3s^* model, which could improve the band structure without increasing the computational cost of generating Coulomb matrix elements, since the number of orbitals is unchanged. Nevertheless, for the topics we discuss in this work the nearest-neighbor sp^3s^* model is a good starting point. We here focus on how the single-particle energies can be improved by tight-binding parameters specifically optimized to give good effective masses, and on how single-particle wave functions from different parameter sets affect the two-particle Coulomb interactions.

One interesting issue concerning quantum dots is the range of the electron-hole exchange interaction. Franceschetti *et al.*[25] show that due to the lack of local orthogonality between electron and hole wave functions the exchange interaction can extend over the whole dot. We investigate in detail the effective range of the exchange interaction by applying a cutoff range to the Coulomb potential. The origin of the characteristic range of the exchange interaction is revealed by the analysis of the ''exchange charge density'' of the electron-hole pair.

II. THEORY

A. Hamiltonian of an electron-hole pair

The effective Hamiltonian of an electron-hole pair contains a single-particle term with the kinetic and potential energies of an electron and a hole, and a two-particle term with the electron-hole Coulomb interaction.[26] The single-particle term is implicitly defined through the empirical tight-binding matrix elements. The Coulomb interaction between the electron and hole is screened by a dielectric function $\epsilon(|\mathbf{r}'-\mathbf{r}|,R)$ that is assumed to be a function of both the separation $|\mathbf{r}'-\mathbf{r}|$ of the two particles and the dot radius R. Spin-orbit couplings are not included in this work.

We use the nearest-neighbor sp^3s^* tight-binding description. The structure of a quantum dot is modeled as an anion-centered zinc-blende structure.[27] Dangling bonds on the surface are removed by explicitly shifting the energies of the corresponding hybrids well above the highest calculated electron states. This treatment imitates a dot whose surface is efficiently passivated with, for example, hydrogen or ligand molecules.

We obtain an electron-hole basis set $|eh\rangle|j_s,m_s\rangle$ by multiplying the electron and hole eigenstates from the solution of the tight-binding Hamiltonian and their spin states. The spatial part $|eh\rangle$ is the product of an electron and a hole wave function which is expressed as a linear combination of the tight-binding basis orbitals with amplitudes $c_{e;n}$ and $c_{h;n'}$,

$$\langle \mathbf{r}_e,\mathbf{r}_h|eh\rangle \equiv \psi_e(\mathbf{r}_e)\psi_h^*(\mathbf{r}_h)=\sum_{n,n'} c_{e;n}c_{h;n'}^*\phi_n(\mathbf{r}_e)\phi_{n'}^*(\mathbf{r}_h), \tag{1}$$

where the tight-binding orbital index n includes atom-site index i and orbital-type index γ. The spin part $|j_s,m_s\rangle$ is an electron-hole spin state, i.e., either the singlet component $|0,0\rangle$ or one of the triplet components $|1,1\rangle$,$|1,0\rangle$, or $|1,-1\rangle$. This description closely follows that of Leung and Whaley.[17]

The single-particle Hamiltonian can be written in terms of the electron-hole basis set and its eigenvalues,

$$H_{\text{single}}=\sum_{ehj_sm_s}(E_e-E_h)|eh\rangle|j_sm_s\rangle\langle j_sm_s|\langle eh|, \tag{2}$$

where E_e and E_h are the electron and hole energies of the single-particle Hamiltonian.

Projecting the electron-hole Hamiltonian onto the two-particle basis set yields the electron-hole Hamiltonian with a Coulomb interaction J and an exchange interaction K,[17,28]

$$H_{\text{e-}h}=\sum_{j_sm_s}(J+K)|j_sm_s\rangle\langle j_sm_s|, \tag{3}$$

$$J=-\sum_{e'h'eh}|e'h'\rangle\langle eh| \times\int\int d^3r'd^3r\frac{\psi_{e'}^*(\mathbf{r}')\psi_e(\mathbf{r}')\psi_h^*(\mathbf{r})\psi_{h'}(\mathbf{r})}{\epsilon(|\mathbf{r}'-\mathbf{r}|,R)|\mathbf{r}'-\mathbf{r}|}, \tag{4}$$

$$K=2\delta_{j_s}\sum_{e'h'eh}|e'h'\rangle\langle eh| \times\int\int d^3r'd^3r\frac{\psi_{e'}^*(\mathbf{r}')\psi_{h'}(\mathbf{r}')\psi_h^*(\mathbf{r})\psi_e(\mathbf{r})}{\epsilon(|\mathbf{r}'-\mathbf{r}|,R)|\mathbf{r}'-\mathbf{r}|}, \tag{5}$$

where δ_{j_s} is unity for a singlet state and zero for a triplet state. The factor 2 in front of δ_{j_s} in Eq. (5) is due to the fact that the exchange interaction allows two final electron-hole spin states $|\uparrow_e,\downarrow_h\rangle$ and $|\downarrow_e,\uparrow_h\rangle$ for an initial state $|\uparrow_e,\downarrow_h\rangle$ (or $|\downarrow_e,\uparrow_h\rangle$). In contrast, the Coulomb interaction requires that the spin of the electron should be the same between an initial and a final state, as should the spin of the hole. Note that we use atomic units for all the equations in this paper.

The Coulomb interaction J describes the scattering of the electron from e to e' and the hole from h to h', whereas the exchange interaction K describes the recombination of a pair e,h at $\mathbf{r}$ and the recreation of a pair e',h' at $\mathbf{r}'$. Since we do not include spin-orbit couplings, the spin states are not coupled to one another in the present Hamiltonian. Therefore, we will use only the spatial part $|eh\rangle$ of the electron-hole basis set from now on. The only constraint that the spin state gives to the Hamiltonian is the spin-selection rule in K.

The matrix elements of the electron-hole interaction Hamiltonian can be rewritten in terms of integrals over the tight-binding basis orbitals by replacing $\psi_e(\mathbf{r}_e)$ with $\Sigma_n c_{e;n}\phi_n(\mathbf{r}_e)$ and $\psi_h(\mathbf{r}_h)$ with $\Sigma_n c_{h;n}\phi_n(\mathbf{r}_h)$ as done in Ref. 17. As a result, the Coulomb and exchange interaction matrix elements are composed of integrals involving four basis orbitals where two orbitals come from the electron wave functions and the other two come from the hole wave functions,

$$\omega(n_1,n_2;n_3,n_4) = \int\int d^3r'd^3r \frac{\phi^*_{n_1}(\mathbf{r}')\phi_{n_2}(\mathbf{r}')\phi^*_{n_3}(\mathbf{r})\phi_{n_4}(\mathbf{r})}{\epsilon(|\mathbf{r}'-\mathbf{r}|,R)|\mathbf{r}'-\mathbf{r}|}. \tag{6}$$

Furthermore, we approximate the Coulomb and exchange interaction matrix elements by considering only terms having at most two distinct basis orbitals following Leung and Whaley.[17] This approximation is reasonable since the integrals involving more than two different orbitals are typically small compared to the kept integrals.[16,17] Integrals in Eq. (6) with $n_1=n_2\equiv n$ and $n_3=n_4\equiv n'$ are *Coulomb integrals* $\omega_{\mathrm{Coul}}(n,n')$, and those with $n_1=n_4\equiv n$ and $n_2=n_3\equiv n'$ or with $n_1=n_3\equiv n$ and $n_2=n_4\equiv n'$ are *exchange integrals* $\omega_{\mathrm{exch}}(n,n')$. The two different kinds of exchange integrals are identical if the tight-binding orbitals are real, as they are in this work.

To make our notations clear, note that the Coulomb and the exchange *integrals* of the basis orbitals are the interactions between the tight-binding basis orbitals. By contrast, the Coulomb and the exchange *interactions* J and K are interactions between the electron and hole wave functions. In fact, the Coulomb interaction has contributions from both the Coulomb and exchange integrals as does the exchange interaction.

Finally, we can describe the electron-hole matrix elements in terms of the Coulomb and exchange integrals,

$$\begin{aligned}\langle e'h'|J|eh\rangle = &-\sum_{n,n'} c^*_{e';n}c_{e;n}c^*_{h;n'}c_{h';n'}\omega_{\mathrm{Coul}}(n,n') \\ &-\sum_{n,n'} c^*_{e';n}c_{e;n'}c^*_{h;n'}c_{h';n'}\omega_{\mathrm{exch}}(n,n') \\ &-\sum_{n,n'} c^*_{e';n}c_{e;n'}c^*_{h;n'}c_{h';n}\omega_{\mathrm{exch}}(n,n'),\end{aligned} \tag{7}$$

$$\begin{aligned}\langle e'h'|K|eh\rangle = &\,2\delta_{j_s}\sum_{n,n'} c^*_{e';n}c_{h';n}c^*_{h;n'}c_{e;n'}\omega_{\mathrm{Coul}}(n,n') \\ &+2\delta_{j_s}\sum_{n,n'} c^*_{e';n}c_{h';n'}c^*_{h;n'}c_{e;n}\omega_{\mathrm{exch}}(n,n') \\ &+2\delta_{j_s}\sum_{n,n'} c^*_{e';n}c_{h';n'}c^*_{h;n}c_{e;n'}\omega_{\mathrm{exch}}(n,n').\end{aligned} \tag{8}$$

Note that Eq. (8) corrects typographical errors in Eq. (22) of Ref. 17.

To evaluate these integrals, we need a real-space description of the tight-binding basis orbitals. To start, we follow Martin *et al.*[16] and model the tight-binding orbitals with atomic Slater orbitals.[29] The Slater orbitals are single-exponential functions with the exponent given by the Slater rules[29] designed to yield a good approximation of the effective radius and the ionization energy. Especially, an s^* orbital is modeled as an excited s orbital by promoting one valence electron to the s orbital of the next shell. The Slater orbitals are an arbitrary choice in the sense that they are not explicitly related to the tight-binding parameters. However, as shown in Sec. III, even in small dots the electron-hole Coulomb interaction is *not* very sensitive to variations in the orbital integrals.

On-site Coulomb and exchange integrals, in which both orbitals are centered on the same atom, are calculated using a Monte Carlo method with importance sampling for the radial integrations. The uncertainty of the Monte Carlo results is within 1%. The angular part is treated exactly by expansion in spherical harmonics. However, the Appendix shows that the integral values must be considered to be uncertain to about 20–30 % due to the arbitrariness of the orbital choice and the effects of orthogonalization.

Off-site exchange integrals, where the two orbitals are centered on two different atom sites, are negligible even for nearest-neighbor integrals. These off site exchange integrals decrease quickly as the distance between atom sites increases, due to the localization and orthogonality of the orbitals. In particular, we show in the Appendix that even nearest-neighbor exchange integrals are negligible as an effect of orthogonalization between off-site hybrids.

Regarding *off-site Coulomb* integrals, Leung and Whaley[17] estimate these integrals with the Ohno formula[30] modified to include screening,

$$\omega_{\mathrm{Coul}}(n,n') \equiv \omega_{\mathrm{Coul}}(i\gamma,i'\gamma') = \frac{1}{\epsilon(|\mathbf{R}_i-\mathbf{R}_{i'}|,R)} \times \frac{1}{\sqrt{\omega^0_{\mathrm{Coul}}(i\gamma,i\gamma')^{-2}+|\mathbf{R}_i-\mathbf{R}_{i'}|^2}}, \tag{9}$$

where $\mathbf{R}_i$ and $\mathbf{R}_{i'}$ are atom site vectors. $\omega^0_{\mathrm{Coul}}(i\gamma,i\gamma')$ is an unscreened on-site Coulomb integral. The superscript 0 designates an unscreened quantity. For the case of binary compounds, $\omega^0_{\mathrm{Coul}}(i\gamma,i\gamma')$ is replaced by the average over the two different atoms. The integrals are screened by the dielectric constant $\epsilon(|\mathbf{R}_i-\mathbf{R}_{i'}|,R)$. This screening is the only modification to the original Ohno formula. From here on, we will refer to this modified Ohno formula in Eq. (9) simply as the Ohno formula.

To test the validity of the Ohno formula in the case that two orbitals are on close atom sites, we calculated the off-site Coulomb integrals with a Monte Carlo method[31] within 1% uncertainty and compared these values with those from the Ohno formula. The Ohno formula severely underesti-

TABLE I. Methods for the computations of the Coulomb and exchange integrals with respect to a site-to-site distance. NN stands for nearest neighbors.

	On site	NN	Beyond NN
Coulomb	Monte Carlo	Monte Carlo	Ohno formula
exchange	Monte Carlo	neglected	neglected

TABLE II. On-site *unscreened* Coulomb integrals and exchange integrals [defined in the paragraph following Eq. (6) in Sec. II A], of the sp^3s^* basis set, in units of eV. Integrals for the sp^3 orbitals are calculated based on the hybridized orbitals along the bonding directions defined in Ref. 17.

Element	Integral	(sp_a^3, sp_a^3)	(sp_a^3, sp_b^3)	(sp_a^3, s^*)	(s^*, s^*)
Si	Coulomb	11.88	8.49	2.87	2.27
Si	exchange	11.88	0.78	0.024	2.27
In	Coulomb	7.82	5.67	2.30	1.36
In	exchange	7.82	0.47	0.024	1.36
As	Coulomb	12.13	9.26	2.34	1.73
As	exchange	12.13	0.47	0.028	1.73
Cd	Coulomb	6.59	5.06	1.77	1.44
Cd	exchange	6.59	0.74	0.017	1.44
Se	Coulomb	12.85	9.70	2.97	2.38
Se	exchange	12.85	0.95	0.030	2.38

mates the off-site integrals as the distance between two atom sites becomes small. For example, the Coulomb integral between the bonding orbitals (see below) from nearest neighbors in a Si quantum dot with radius 18.9 Å given by the Ohno formula is 0.58 eV, while the Monte Carlo calculation gives 2.35 eV. For next-nearest neighbors, the Ohno formula and the Monte Carlo calculation give 0.38 and 0.58 eV, respectively, and for the third-nearest neighbors 0.33 and 0.36 eV.

The reason that we obtain a big discrepancy between the Ohno formula and the Monte Carlo calculation for the bonding orbital integrals is that the effective distance between the bonding orbitals is smaller than the spacing between the nearest-neighbor atom sites. In fact, the spatial overlap of the bonding-orbitals is as big as that of the orbitals on the same atom site. In addition, the spatial dependence of the dielectric function is not fully taken into account in the Ohno formula. This approximation becomes critical when the range of variations in the dielectric function is comparable to the effective distance between orbitals. In that case, the effective dielectric function cannot be represented by $\epsilon(|\mathbf{R}_i-\mathbf{R}_j|,R)$. Therefore, we use the Monte Carlo values for the on-site and the nearest-neighbor integrals and the Ohno formula for the rest of off-site integrals. For clarity, we summarize the methods for the computation of the Coulomb and exchange integrals in Table I.

We use a size- and distance-dependent dielectric function to screen the Coulomb and exchange interaction of the electron-hole pair. The dielectric function, as a function of the separation r of two particles and the radius R of a quantum dot, is approximated by the Thomas-Fermi model of Resta[32] and the Penn model generalized for quantum dots.[33,34] The separation dependence is given by the Thomas-Fermi model, while the size dependence is given by the Penn model. This way of combining the two models to describe the dielectric function is taken from Ref. 13,

$$\epsilon(r,R)=\begin{cases}\epsilon_\infty^{\mathrm{dot}}(R)\,qr_0/[\sinh q(r_0-r)+qr], & r<r_0\\ \epsilon_\infty^{\mathrm{dot}}(R), & r\geqslant r_0\end{cases} \tag{10}$$

where

$$\epsilon_\infty^{\mathrm{dot}}(R)=1+(\epsilon_\infty^{\mathrm{bulk}}-1)\frac{(E_{\mathrm{gap}}^{\mathrm{bulk}}+\Delta)^2}{[E_{\mathrm{gap}}^{\mathrm{dot}}(R)+\Delta]^2}. \tag{11}$$

The Thomas-Fermi wave vector q is $(4/\pi)^{1/2}(3\pi^2 n_0)^{1/6}$, where the valence electron density $n_0=32/a_0^3$ in a zinc-blende structure. The screening radius r_0 is determined by the condition $\sinh qr_0/qr_0=\epsilon_\infty^{\mathrm{dot}}(R)$. The shift $\Delta=E_2-E_{\mathrm{gap}}^{\mathrm{bulk}}$, where E_2 is the energy of the first pronounced peak in the bulk absorption spectrum. The energies $E_{\mathrm{gap}}^{\mathrm{bulk}}$ and $E_{\mathrm{gap}}^{\mathrm{dot}}(R)$ are the single-particle gaps for bulk and a dot with radius R, respectively. $\epsilon_\infty^{\mathrm{bulk}}$ is the dielectric constant for the bulk material.

The unscreened on-site Coulomb and exchange integrals for the sp^3s^* basis set are listed in Table II, and the screened on-site Coulomb integrals are listed in Table III. The screened nearest-neighbor Coulomb integrals are listed in Table IV. The integrals with s and p orbitals are calculated using the four hybridized orbitals along bonding directions as defined in Ref. 17.

As shown in Tables II and III, the screening effect is significant even for on-site integrals. This is because the screening radius r_0 is 2–4 Å and is similar to the effective radius of the tight-binding basis orbitals. The comparison of

TABLE III. On-site *screened* Coulomb integrals of the sp^3 hybridized orbitals and s^* orbital for the Si dot with radius $R=18.9$ Å, InAs with $R=21.1$ Å, and CdSe with $R=21.1$ Å, in units of eV. The integrals are screened by the dielectric function in Eq. (10), which is a function of electron-hole separation and of the radius of the quantum dot. For comparison, the values in parentheses are the integrals obtained from full screening with the dielectric constant in the long-distance limit $\epsilon_\infty^{\mathrm{dot}}(R)$.

Element	(sp_a^3, sp_a^3)	(sp_a^3, sp_b^3)	(sp_a^3, s^*)	(s^*, s^*)
Si	3.16 (1.26)	1.67 (0.91)	0.32 (0.31)	0.26 (0.24)
In	1.26 (0.80)	0.77 (0.58)	0.24 (0.24)	0.17 (0.14)
As	2.33 (1.24)	1.33 (0.95)	0.28 (0.24)	0.23 (0.18)
Cd	1.86 (1.29)	1.20 (0.99)	0.35 (0.35)	0.28 (0.28)
Se	4.74 (2.51)	2.76 (1.89)	0.58 (0.58)	0.48 (0.46)

TABLE IV. Nearest-neighbor *screened* Coulomb integrals of the sp^3 hybridized orbitals for the Si dot with radius $R=18.9$ Å, InAs with $R=21.1$ Å, and CdSe with $R=21.1$ Å, in units of eV. The integrals are screened by the dielectric function in Eq. (10). The values given by the Ohno formula Eq. (9) are listed within parentheses.

	Si bonding	Si nonbonding
Si bonding	2.35 (0.58)	0.95 (0.53)
Si nonbonding	0.95 (0.53)	0.55 (0.53)
	In bonding	**In nonbonding**
As bonding	1.43 (0.54)	0.61 (0.54)
As nonbonding	0.81 (0.53)	0.42 (0.53)
	Cd bonding	**Cd nonbonding**
Se bonding	2.41 (1.03)	1.03 (1.02)
Se nonbonding	1.62 (1.02)	0.79 (1.00)

the unscreened and screened on-site integrals shows that the effective screening of these integrals is generally about half the long-range screening given by $\epsilon_{\infty}^{\mathrm{dot}}(R)$. Further, based on this observation, we use half of the long-range dielectric constant to screen the on-site exchange integrals that we did not calculate explicitly by the Monte Carlo method.[35]

B. Lowest excitonic states

To obtain the excitonic states near the band edge, we diagonalize the configuration-interaction matrix in the $|eh\rangle$ basis set given by the sum of H_{single} and $H_{e\text{-}h}$ defined by Eqs. (2)–(5). We include sufficient electron and hole states in the configurations to achieve convergence of the first few excitonic states to within a few meV. The typical number of electron and hole states needed is about 10–15 each.

There are two types of Hamiltonians, depending on the total spin of the electron-hole pair. The Hamiltonian for a spin singlet includes both the Coulomb and the exchange interaction, whereas the Hamiltonian for a spin triplet has only the Coulomb interaction. By diagonalizing these two Hamiltonians separately, we obtain a set of spin-singlet and spin-triplet excitonic states. The lowest excitonic state is the lowest triplet state due to the absence of the positive exchange interaction. However, since only spin-singlet states are optically allowed, the *optical* excitonic gap is the energy of the lowest spin-singlet state.

This order of the states can be seen most easily by applying first-order perturbation theory to the electron-hole pair made from the highest hole state and the lowest electron state, which yields

$$E_{\mathrm{singlet}}=E_e-E_h+\langle eh|J|eh\rangle+\langle eh|K|eh\rangle. \tag{12}$$

$$E_{\mathrm{triplet}}=E_e-E_h+\langle eh|J|eh\rangle. \tag{13}$$

The sign of $\langle eh|J|eh\rangle$ is always negative and the sign of $\langle eh|K|eh\rangle$ is always positive. Therefore, the energy of the lowest spin-triplet excitonic state is smaller than the energy of the lowest spin-singlet excitonic state by $\langle eh|K|eh\rangle$ within first-order perturbation theory. The energy difference between these excitonic states is the *exchange splitting*. Further, we denote the difference between the lowest spin-triplet energy and the single-particle energy gap as the *Coulomb shift*, which is $\langle eh|J|eh\rangle$ in first-order perturbation theory. The Coulomb shift is the main correction to the single-particle gap since the Coulomb interaction is roughly one order of magnitude larger than the exchange interaction.

This simple description becomes more complicated when configuration interaction is included due to the correlation of several electron-hole configurations near the band edges. However, the main ideas of Coulomb shift and exchange splitting are still valid. Generally, as we include more electron-hole configurations the Coulomb shift and the exchange splitting increase and converge.

C. Effective range of the exchange interaction

The long-range component of the exchange interaction was investigated by Franceschetti and co-workers[25] with the pseudopotential method. They show that there is a long-range component in the monopole-monopole exchange interaction for several direct-gap semiconductor quantum dots. To verify this long-range exchange interaction with the tight-binding model, we follow their approach using a cutoff potential. With the step function $\Theta(r)$, we replace the Coulomb potential with a cutoff potential $\Theta(r_c-|\mathbf{r}'-\mathbf{r}|)/|\mathbf{r}'-\mathbf{r}|$ to obtain an exchange interaction that is a function of the cutoff distance r_c.

The unscreened exchange interaction with the cutoff potential for the electron-hole state $|eh\rangle$ is

$$\begin{aligned}
&\langle eh|K^0(r_c)|eh\rangle \\
&\quad=\int\int d^3r'\,d^3r \\
&\qquad\times\frac{\psi_e^*(\mathbf{r}')\psi_h(\mathbf{r}')\psi_h^*(\mathbf{r})\psi_e(\mathbf{r})}{|\mathbf{r}'-\mathbf{r}|}\Theta(r_c-|\mathbf{r}'-\mathbf{r}|) \\
&\quad\approx 2\sum_{n_1,n_2} c_{e;n_1}^* c_{h;n_1} c_{h;n_2}^* c_{e;n_2}\omega_{\mathrm{Coul}}^0(n_1,n_2) \\
&\qquad\times\Theta(r_c-|\mathbf{R}_{n_1}-\mathbf{R}_{n_2}|)+4\sum_{n_1,n_2} c_{e;n_1}^* c_{h;n_2} c_{h;n_2}^* c_{e;n_1} \\
&\qquad\times\omega_{\mathrm{exch}}^0(n_1,n_2),
\end{aligned} \tag{14}$$

where the superscript 0 refers to the unscreened interaction. In line with the discrete spatial character of the tight-binding model, we make an approximation that replaces the true cutoff potential with the one based on the site indices $\Theta(r_c-|\mathbf{R}_{n_1}-\mathbf{R}_{n_2}|)$. If there is a long-range exchange interaction, it would stem from the first term of Eq. (14) which includes the Coulomb integrals. The second term, the sum of exchange integrals, has only on-site integrals, since all off-site exchange integrals are negligible as shown in the Appendix.

To understand the physical origin of the long-range exchange, we expand the exchange interaction K^0 in a multipole expansion,

$$\langle eh|K^0|eh\rangle \approx \sum_{i\neq j} \frac{q(\mathbf{R}_i)^*_{eh} q(\mathbf{R}_j)_{eh}}{|\mathbf{R}_i-\mathbf{R}_j|} + O(|\mathbf{R}_i-\mathbf{R}_j|^{-2}) + \text{on-site interaction.} \quad (15)$$

Here the ''exchange charge density'' $q(\mathbf{R}_i)_{eh}$ at atom site $\mathbf{R}_i$ is a monopole moment defined as

$$\begin{aligned} q(\mathbf{R}_i)_{eh} &\equiv \int d\Omega_i \psi_e(\mathbf{r})\psi_h^*(\mathbf{r}) \\ &= \sum_{\gamma\gamma'} c_{e;\gamma}(\mathbf{R}_i) c^*_{h;\gamma'}(\mathbf{R}_i) \int d^3r \phi_{i\gamma}(\mathbf{r})\phi^*_{i\gamma'}(\mathbf{r}) \\ &= \sum_{\gamma\gamma'} c_{e;\gamma}(\mathbf{R}_i) c^*_{h;\gamma'}(\mathbf{R}_i)\delta_{\gamma\gamma'} \\ &= \sum_{\gamma} c_{e;\gamma}(\mathbf{R}_i) c^*_{h;\gamma}(\mathbf{R}_i), \end{aligned} \quad (16)$$

where $\int d\Omega_i$ is defined to integrate only the orbitals on the atom site $\mathbf{R}_i$. For clarity, the tight-binding orbital index n is specifically replaced by the atom-site index i (or $\mathbf{R}_i$) and the orbital-type index γ. Note that the final expression for $q(\mathbf{R}_i)_{eh}$ has a sum over only one orbital-type index due to the assumed orthogonality of the tight-binding basis orbitals. The distribution of the exchange charge density determines the long-range character of the exchange interaction. If the exchange charge density is zero, that is, the electron and hole states are locally (on-site) orthogonal, there is no exchange interaction beyond the on-site contribution according to Eq. (15). In contrast, if the exchange charge density is nonzero due to the local nonorthogonality of the electron and hole wave functions from site to site, a long-range exchange interaction is caused by the monopole-monopole interaction.

III. RESULTS

A. Real-space description of basis orbitals

The empirical tight-binding model has an inherent difficulty concerning the tight-binding basis orbitals. The real-space description of the basis orbitals is not provided since the tight-binding matrix elements are determined by fitting to the bulk band structure. However, to include electron-hole correlations the electron-hole Coulomb and exchange matrix elements need to be calculated, which requires an explicit choice of real-space basis orbitals. Since this choice is largely arbitrary in the sense that there is no way to connect the chosen basis orbitals to the empirically chosen tight-binding parameters, we need to test to what degree the choice of orbitals affects the electron-hole Coulomb interaction.

We perform this test by scaling the onsite Coulomb and exchange integrals from the values listed in Tables II and III. This scaling scheme is an indirect way of testing the sensitivity on the real-space description. New off-site Coulomb integrals are calculated by replacing the unscreened on-site integrals with the scaled ones in the Ohno formula, Eq. (9). Note that the offsite Coulomb integrals are not directly scaled by the same factor as the on-site integrals, but change only indirectly through the scaled unscreened on-site integrals in the Ohno formula. Therefore, this scheme is in effect changing atomistic details of the basis orbitals.

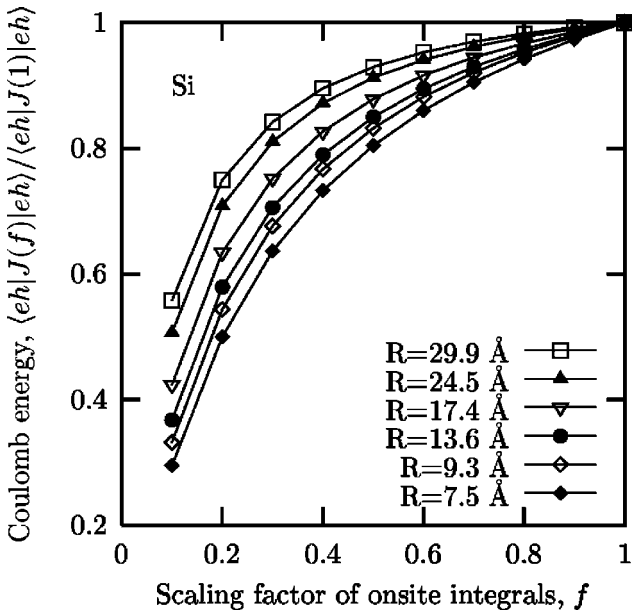

FIG. 1. Coulomb energy $\langle eh|J(f)|eh\rangle$ as a function of scaling factor f of the on-site integrals. The Coulomb energy with the highest hole wave function and the lowest electron wave function $|eh\rangle$ is shown for various radii of Si spherical quantum dots, with the on-site integrals scaled by the factor f. That is $\omega \rightarrow f\omega$ from the values in Tables II and III. The off-site integrals are only indirectly scaled through the on-site integrals in the Ohno formula, Eq. (9). The Coulomb energy is normalized by its value at $f=1$. As the dot size increases, the Coulomb energy becomes significantly less sensitive to variations in the on-site integrals.

Figure 1 shows the variation of the Coulomb interaction between the highest hole state and the lowest electron state with the scaling of the on-site Coulomb and exchange integrals. It shows that as the dot size increases, the sensitivity of the Coulomb interaction to the on-site integrals decreases. For example, if the on-site integrals are reduced by 50%, the reduction in the Coulomb energy varies from 20% in the smallest shown dot to only 5% at 30-Å radius. Since the contribution from the on-site integrals decreases as the dot size increases, the specific model of the real-space functions for the basis orbitals is less critical for larger dots.

We can explain this effect by a closer look at the Ohno formula for the off-site integrals. In the limit of large distances between two atom sites, the off-site integrals become independent of the on-site integral values. The unscreened offsite integral in the limit of large distance between the two atom sites is

$$\omega^0_{\text{Coul}}(i\gamma, i'\gamma') \approx \frac{1}{|\mathbf{R}_i-\mathbf{R}_{i'}|} - \frac{1}{2\omega^0_{\text{Coul}}(i\gamma, i\gamma')^2 |\mathbf{R}_i-\mathbf{R}_{i'}|^3}. \quad (17)$$

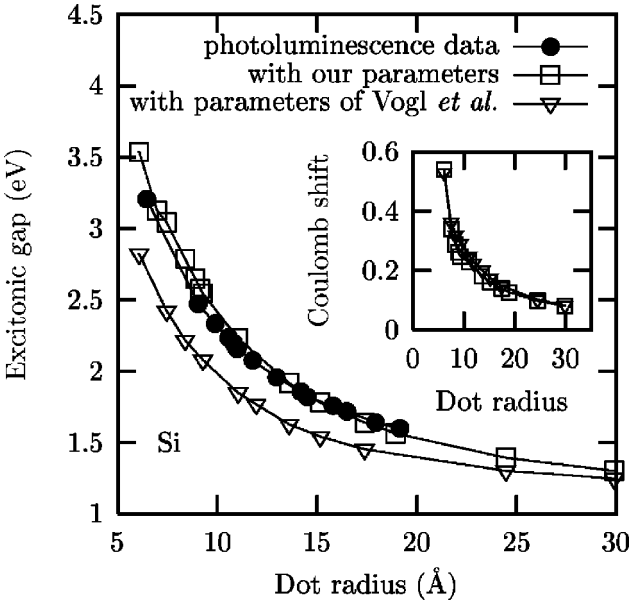

FIG. 2. Excitonic gap of Si spherical quantum dots as a function of the dot radius. The photoluminescence data are taken from Ref. 6. The other two sets of excitonic gaps are calculated with the tight-binding parameters of Vogl *et al.* (Ref. 24) and our parameters in Table VI, respectively. Our parameters give significantly better agreement with experiment than the parameters of Vogl *et al.* (Ref. 24). This good agreement is due to the improved effective masses obtained with our optimized parameters. Inset: Coulomb shift versus the dot radius. The Coulomb shift does not vary much between the parameter sets.

Therefore, the off-site integrals become a point-charge interaction, making the atomic-scale details of the basis orbitals irrelevant in this limit.

The significance of Fig. 1 is that it quantifies this qualitative explanation for decreasing sensitivity with increasing dot sizes. For example, with a targeted 10% accuracy in the Coulomb interaction, only a 50% accuracy for the on-site integrals is needed for dots of radius larger than 30 Å , while for dots of 10-Å radius only a 20% error can be afforded in the on-site integrals. The discussion in the Appendix shows that the dominant integrals must be considered uncertain to about 20–30 %. The tight-binding description of correlation effects can therefore be considered reliable for dots with radii larger than 15–20 Å .

B. Excitonic states near the band edge

We apply the tight-binding configuration-interaction scheme described in Sec. II A to Si, InAs, and CdSe quantum dots in order to calculate the lowest excitonic states near the band edge. We have included sufficient electron and hole states to converge the energies to within a few meV.

For Si, we first used the tight-binding parameters of Vogl *et al.*[24] to determine the single-particle Hamiltonian matrix elements. As shown in Fig. 2, the excitonic gap using their tight-binding parameters gives a discrepancy as large as 0.3 eV compared with experimental data.[6] The tight-binding parameters of Vogl *et al.*[24] necessarily give poor effective masses since the parameters are determined by fitting to the energies of only the Γ and X points in the bulk band structure. The resulting effective masses are listed in Table V. The comparison with the experimental values listed in Table V shows that their parameters fail to produce good effective masses especially for the transverse effective masses at the lowest conduction energy near X.

TABLE V. Effective masses of Si with the tight-binding parameters of Vogl *et al.* (Ref. 24) and our parameters in Table VI, in units of the free-electron mass. m_{cl} and m_{ct} denote the longitudinal and transverse effective masses at the lowest conduction energy near X. m_{vl} and m_{vh} are the effective masses at Γ of the two highest valence bands with a light mass and a heavy mass, respectively. The hole masses are averages of the three directions given in Ref. 38. The cyclotron resonance data are taken from Ref. 36.

	m_{cl}	m_{ct}	m_{vl}	m_{vh}
Vogl *et al.*	0.73	1.61	0.18	0.39
Our parameters	0.91	0.30	0.15	0.55
Cyclotron resonance	0.92	0.19	0.15	0.54

To improve the effective masses, we replace the parameter set of Vogl *et al.*[24] with our parameter set listed in Table VI. Our parameter set is optimized with a genetic algorithm by fitting the effective masses as well as the energies at high symmetry points of the bulk band structure.[38] The resulting effective masses are listed in Table V. One important note is that we use two different parameter sets to separately optimize the electron and hole single-particle states. Good effective masses are impossible to obtain simultaneously for both the conduction and the valence band of Si with one set of parameters within the nearest-neighbor sp^3s^* tight-binding model (see Ref. 38). Consequently, the electron and hole single-particle wave functions, being generated from different Hamiltonians, are not orthogonal. However, even though the orthogonality has not been enforced, the overlaps between the different electron and hole wave functions are at most 0.001. Thus, we can use these two different parameter sets to verify how important a role the effective masses play in the electronic properties of the quantum dots.

Figure 2 shows the improved excitonic gaps with our parameter set. To further examine the effect of changing pa-

TABLE VI. Tight-binding parameters for electron and hole states of Si in units of eV. The notation (Ref. 39) of Vogl *et al.* (Ref. 24) is used.

	$E(s)$	$E(p)$	$E(s^*)$	$V(s,s)$
Electron	−3.060	1.675	4.756	−8.114
Hole	−4.777	1.674	8.697	−8.465
Vogl *et al.*	−4.200	1.715	6.685	−8.300
	$V(x,x)$	$V(x,y)$	$V(s,p)$	$V(s^*,p)$
Electron	1.675	21.838	8.236	5.994
Hole	1.674	4.919	5.724	6.133
Vogl *et al.*	1.715	4.575	5.729	5.375

TABLE VII. Tight-binding parameters for InAs in units of eV. The notation (Ref. 39) of Vogl *et al.* (Ref. 24) is used. Indices *a* and *c* refer to anion and cation, respectively.

$E(s,a)$	$E(p,a)$	$E(s,c)$	$E(p,c)$	
−8.419	0.096	−2.244	0.096	
$E(s^*,a)$	$E(s^*,c)$	$V(s,s)$	$V(x,x)$	$V(x,y)$
12.147	7.485	−4.267	1.427	5.356
$V(sa,pc)$	$V(sc,pa)$	$V(s^*a,pc)$	$V(s^*c,pa)$	
4.409	5.326	5.846	4.594	

rameters, we can compare the electron-hole interaction energies with our parameters to those energies obtained with the parameters of Vogl *et al.*[24] In particular, we compare the Coulomb shift, the energy difference between the single-particle gap, and the lowest triplet excitonic energy. As shown in the inset in Fig. 2, the Coulomb shifts from the two parameter sets are very similar. This insensitivity indicates that the better description of the excitonic gap with our parameter set is mainly due to the better single-particle eigenvalues and not from a change in the Coulomb matrix elements.

To study direct-gap semiconductors, we choose InAs and CdSe spherical quantum dots. The InAs tight-binding parameters are generated using the genetic algorithm approach,[38] fitting band gaps and effective masses at Γ as well as possible, but neglecting spin-orbit coupling. These parameters are listed in Table VII. The resulting effective masses with these parameters are $m_c = 0.024$, $m_{vl} = 0.025$, and $m_{vh} = 0.405$, where m_c is the effective mass of the lowest conduction band at Γ, and m_{vl} and m_{vh} are defined as in Table V. The tight-binding parameters for CdSe are taken from Ref. 40. Their parameters give good effective masses as well as good energies in high symmetry points of the bulk material.

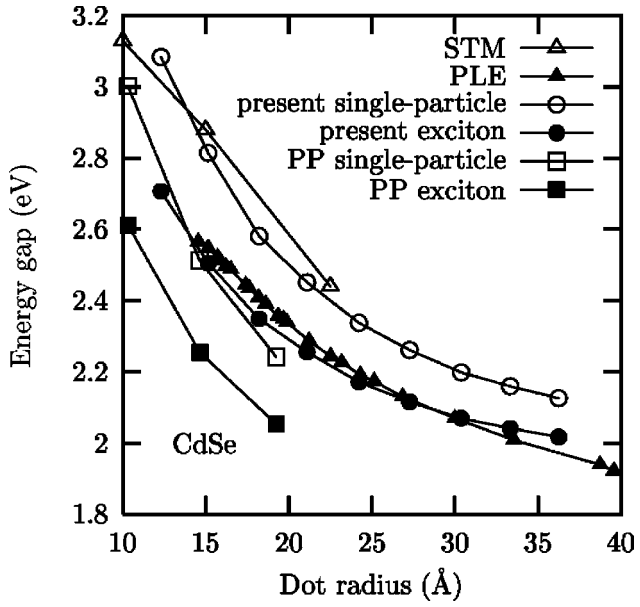

FIG. 3. Excitonic gap and single-particle gap of CdSe spherical quantum dots as a function of the dot radius. The photoluminescence excitation (PLE) gaps are taken from Ref. 3. The scanning tunneling spectroscopy (STM) gaps are obtained from recent STM tunneling dI/dV spectra (Ref. 7 and 37). The excitonic gaps of the pseudopotential (PP) calculations (Ref. 13) are about 0.15 eV lower than the PLE gaps. Our excitonic gaps are in good agreement with the PLE gaps. The small difference between our single-particle gaps and the STM quasiparticle gaps indicates that the quasiparticle polarization energy is small for these dots.

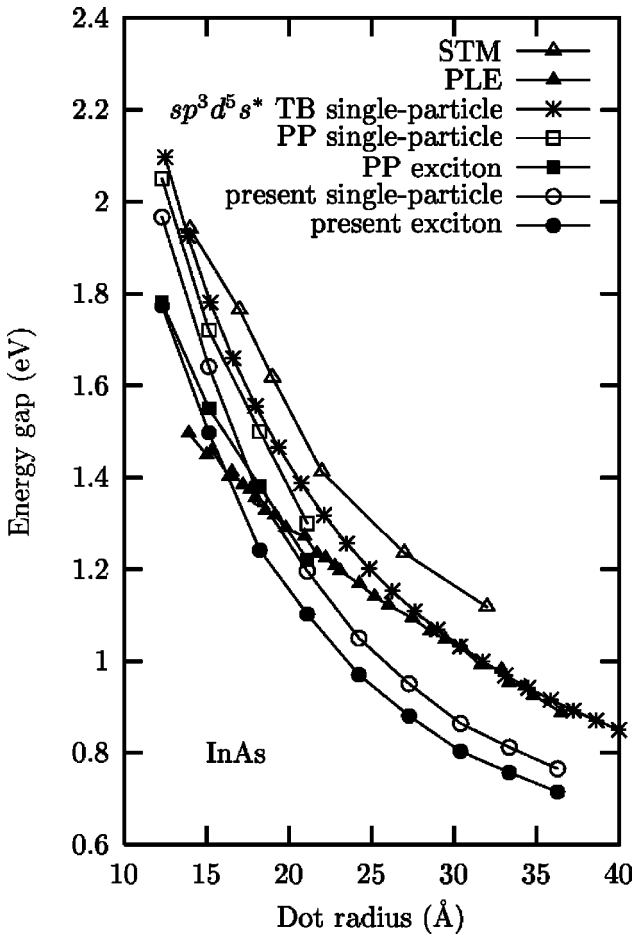

FIG. 4. Excitonic gap and single-particle gap of InAs spherical quantum dots as a function of the dot radius. The measured PLE gaps are taken from Ref. 4. The STM gaps are obtained from the tunneling spectra of Millo *et al.* (Refs. 10 and 37) The pseudopotential gaps (PP) are from Ref. 14. The $sp^3d^5s^*$ tight-binding (TB) single-particle gaps are plotted using the fitting parameters of Allan *et al.* (Ref. 19). The inclusion of *d* orbitals and spin-orbit coupling raises the gaps as much as 0.2 eV in comparison with our sp^3s^* model. It is not understood why the experimental curve is so much flatter than the theoretical curves.

Figures 3 and 4 show the resulting excitonic gaps versus the dot radius. For CdSe quantum dots, our excitonic gaps are in good agreement with optical gaps measured by photoluminescence excitation[3] (PLE). We also plot the energy gaps measured by scanning tunneling spectroscopy (STM) on a single quantum dot.[7] The STM gaps are obtained from

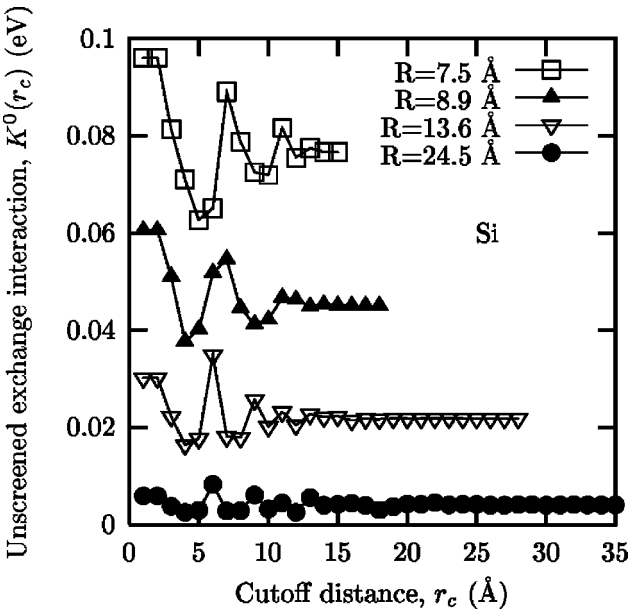

FIG. 5. Unscreened exchange energy, Eq. (14), as a function of the cutoff distance, with the Coulomb potential replaced by a cutoff potential for various radii of Si spherical quantum dots. The energies are for the highest hole wave function and the lowest electron wave function. The curves show that there is an oscillation region for small cutoff distances followed by a saturation region beyond 15 Å. This saturation suggests that the effective range of the exchange interaction in Si quantum dots is around 15 Å regardless of the dot radius.

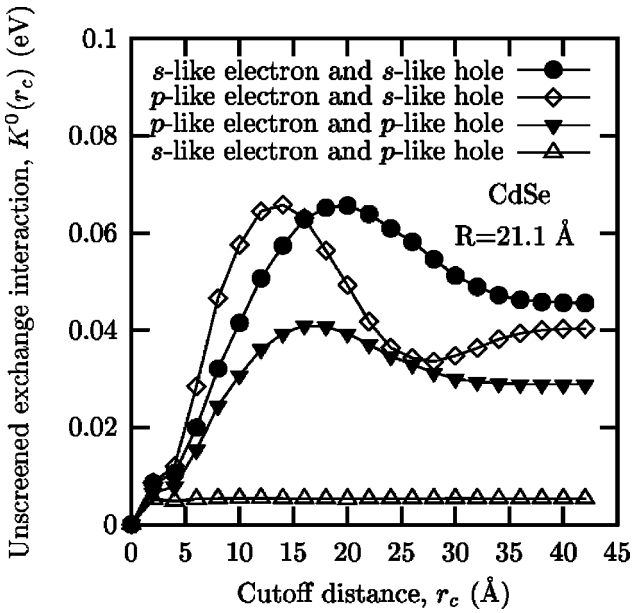

FIG. 6. Unscreened exchange energy, Eq. (14), as a function of the cutoff distance, with the Coulomb potential replaced by a cutoff potential for the CdSe spherical quantum dot of radius $R=21.1$ Å . The unscreened exchange energy of four different types of electron-hole configurations is shown. The electron and hole configurations are labeled by the dominant angular-momentum component of their envelope functions (Ref. 42). Except for the *s*-like electron and *p*-like hole configuration, the variation of the exchange interaction extends over the whole dot.

the difference between the first prominent peaks of the tunneling dI/dV spectra with positive and negative bias voltages, respectively. Since the STM experiment applies bias voltages to add or subtract electrons from the quantum dots, this experiment measures quasiparticle energies. For a finite system, these quasiparticle energies include the (positive) polarization energy between the particle and the image charges on the surface. The polarization energy is roughly $2(1/\epsilon_{\text{out}}-1/\epsilon_{\text{in}})/R$. The small difference between our single-particle gaps and the STM gaps therefore suggests that the dielectric constant ϵ_{out} of the surrounding material is relatively close to the dielectric constant ϵ_{in} in the dots. The results of pseudopotential calculations[13] are also plotted in the figures for comparison.

For InAs quantum dots, there is a significant discrepancy as large as 0.2 eV between our sp^3s^* tight-binding excitonic gaps and PLE gaps.[4] Eight-band effective-mass calculations[4] and pseudopotential calculations[14] also fail to describe the experimental data, especially the lack of significant curvature. The recent results of Allan *et al.*[19] show that the inclusion of *d* orbitals and spin-orbit interaction raises the sp^3s^* results by almost the needed 0.2 eV. However, their results do not include the Coulomb shift and should therefore be shifted down by 200–50 meV as the dot size increases. Recent STM measurements are also plotted in Fig. 4. It is consistent with the larger dielectric constant of InAs that the STM results in this case are well above the other curves by an amount similar to the Coulomb shift.

C. Effective range of the exchange interaction

One of the interesting issues related to the exchange interaction of the electron-hole pair is its effective range. Motivated by the work of Franceschetti *et al.*,[25] we calculate the unscreened exchange interaction defined in Eq. (14) as a function of the cutoff distance r_c to determine the effective range of the exchange interaction. As the cutoff distance increases for a given electron-hole configuration, the exchange interaction eventually saturates to a final value. If this saturation occurs over just a few atomic sites, we call it short ranged, while long ranged exchange implies that the saturation occurs over distances comparable to the dot size. For Si, Fig. 5 shows that for the configuration with the highest hole state and the lowest electron state there is a region of strong oscillations below a cutoff distance of 15 Å . The strong oscillations are due to the phase difference between the electron and hole states stemming from their different locations in *k* space for an indirect-gap material. The oscillations die out beyond a cutoff distance of about 15 Å, suggesting that the effective range of the exchange interaction in Si quantum dots is around 15 Å, regardless of dot size. This short-ranged and oscillatory behavior is universal within the configurations near the band edges.

For the direct-gap InAs and CdSe quantum dots, we calculate the unscreened exchange interaction for several of the lowest electron-hole configurations. We label the electron and hole states by the dominant angular-momentum character of their ''envelope functions.'' Here, the envelope function is defined to be the coefficient of the dominant basis

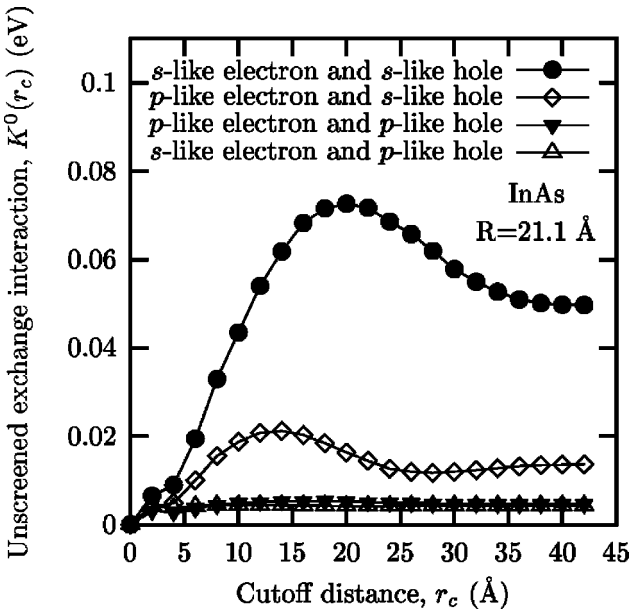

FIG. 7. Unscreened exchange energy, Eq. (14), as a function of the cutoff distance, with the Coulomb potential replaced by a cutoff potential for the InAs spherical quantum dot of radius $R=21.1$ Å . The unscreened exchange energy of four different types of electron-hole configurations is shown. The electron and hole configurations are labeled by the dominant angular-momentum component of their envelope functions (Ref. 42). Long-range exchange interactions appear for the *s*-like hole with both the *s*-like electron and the *p*-like electron.

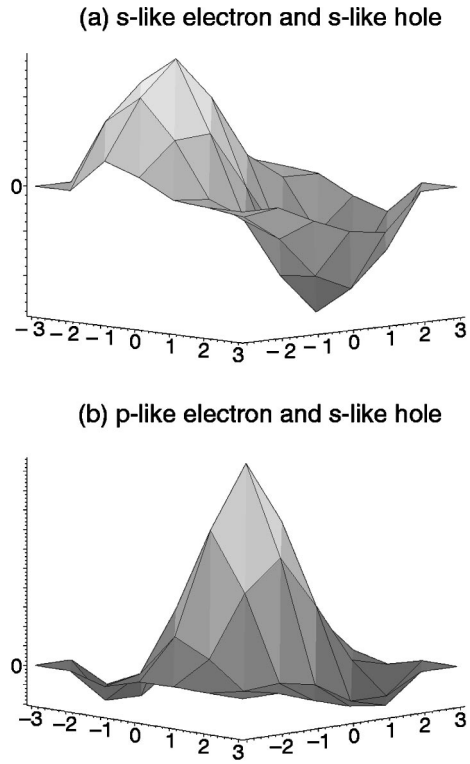

FIG. 8. Exchange charge density $q(\vec{R})_{eh}$ from Eq. (16) of (a) the *s*-like electron and *s*-like hole configuration, and of (b) the *p*-like electron and *s*-like hole configuration for the CdSe quantum dot with radius 21.1 Å. The exchange charge density is plotted in a plane through the center of the dot. The unit of the horizontal axes is the lattice constant of CdSe. The plots show that the orthogonality between the electron and hole wave functions is global not local, with a *p*-like oscillation or a $2s$-like oscillation, respectively. These global oscillations of the exchange charge density lead to the long-range variation of the exchange interactions in Fig. 6.

orbital. The *s* and *p* basis orbitals are typically dominant in the electron and hole states, respectively. In our calculations, a *p*-like hole[41,42] is the highest hole state and an *s*-like hole is the second-highest hole state. This order is opposite that of pseudopotential theory.[13] However, it is possible that the spin-orbit coupling, which is not included in this work, can affect the order of these hole states.

As shown in Figs. 6 and 7, direct-gap quantum dots show a qualitatively different behavior of the exchange interaction with respect to the cutoff distance from the behavior for Si. First, since there is no overall phase difference between the electron and hole states, there is no region with oscillations for small cutoff distances. Second, the exchange interaction for a particular electron-hole pair can grow continuously up to the dot radius. The figures show that the exchange interaction of direct-gap materials is generally long ranged, extending over the whole dot.

To understand why some electron-hole configurations have a slowly varying long-range exchange interaction, we analyze the long-range component by a multipole expansion as written in Eq. (15). The leading term of the long-range exchange interaction is the monopole-monopole interaction. Therefore, the distribution of the monopole moment, or the ''exchange charge density'' defined in Eq. (16), determines the range of the long-range exchange interaction.

The exchange charge density of an electron-hole pair has zero total charge due to the orthogonality between the electron and hole wave functions. There are two ways to satisfy this condition: the electron and hole states are either locally orthogonal, which is enforced in effective mass theory due to the orthogonality between the Bloch functions of the valence and conduction bands, or globally orthogonal, which is possible in the atomistic pseudopotential and tight-binding theories. If the former is true, the exchange charge density would be zero at each site and there would be no monopole-monopole interaction. That would make the exchange interaction of the electron-hole pair short ranged. By contrast, without on-site orthogonality the exchange charge density has nonzero values, causing monopole-monopole interactions that lead to significant long-range exchange interactions.

To show that the character of the orthogonality of the electron-hole configuration determines the long-range behavior of the exchange interaction, we plot in Fig. 8 the exchange charge density of two electron-hole configurations in CdSe that have a long-range exchange interaction in Fig. 6. Figure 8 shows the exchange charge density of (a) the *s*-like electron and *s*-like hole configuration, and of (b) the *p*-like electron and *s*-like hole configuration in a plane going through the center of the dot for CdSe with radius 20.1 Å. This figure shows that there is no local orthogonality be-

TABLE VIII. On-site unscreened Coulomb and exchange integrals with Löwdin orthogonalized Gaussian-type hybrids (O-GTO); nonorthogonal Gaussian-type hybrids (NO-GTO); and nonorthogonal Slater orbitals (NO-SO). The GTO integrals were calculated with the MOLPRO (Ref. 43) package using the atomic pseudopotentials from the Los Alamos group (Ref. 44). The SO integrals are from our Monte Carlo calculations. The hybrids a and b are the ones defined as sp_a^3 and sp_b^3 in Ref. 17.

	O-GTO	NO-GTO	NO-SO
$\omega^0_{\mathrm{Coul}}(a,a)$ of Si	11.95	11.65	11.91
$\omega^0_{\mathrm{Coul}}(a,b)$ of Si	9.44	8.85	9.00
$\omega^0_{\mathrm{exch}}(a,b)$ of Si	1.06	0.91	0.73
$\omega^0_{\mathrm{Coul}}(a,a)$ of In	7.90	8.52	7.82
$\omega^0_{\mathrm{Coul}}(a,b)$ of In	6.73	6.54	5.67
$\omega^0_{\mathrm{exch}}(a,b)$ of In	0.77	0.67	0.47
$\omega^0_{\mathrm{Coul}}(a,a)$ of As	12.99	12.57	12.13
$\omega^0_{\mathrm{Coul}}(a,b)$ of As	10.00	9.54	9.26
$\omega^0_{\mathrm{exch}}(a,b)$ of As	1.08	0.99	0.47
$\omega^0_{\mathrm{Coul}}(a,a)$ of Cd	7.09	7.81	6.59
$\omega^0_{\mathrm{Coul}}(a,b)$ of Cd	6.08	5.98	5.06
$\omega^0_{\mathrm{exch}}(a,b)$ of Cd	0.70	0.61	0.74
$\omega^0_{\mathrm{Coul}}(a,a)$ of Se	14.14	13.73	12.85
$\omega^0_{\mathrm{Coul}}(a,b)$ of Se	10.80	10.39	9.70
$\omega^0_{\mathrm{exch}}(a,b)$ of Se	1.15	1.08	0.90

tween the electron and hole wave functions. The orthogonality of the electron and hole wave functions are instead satisfied by a p-like global oscillation [case (a)] or a $2s$-like global oscillation [case (b)]. These shapes of the global oscillations explain why the exchange interaction has growing and decaying regions over global distances as shown in Figs. 6 and 7. Those electron-hole configurations that do not have a long-range exchange interaction have a much smaller exchange charge density than those configurations that do have the long-range exchange interaction. These results show that local nonorthogonality of the electron and hole wave functions leads to a strong monopole-monopole interaction, and that the global variations in the exchange interaction depend on the particular way in which the exchange charge density globally sums to zero for a specific electron-hole configuration.

IV. SUMMARY

We use tight-binding wave functions to calculate electron-hole states near the band edge for both direct-gap and indirect-gap quantum dots. First, we examined to what degree the model of the real-space atomic basis orbitals affects the electron-hole Coulomb interaction. We find that the sensitivity of the Coulomb interaction to the real-space description of the basis orbitals decreases quickly as the dot size increases. Our results shows that tight-binding descriptions of electron-hole Coulomb interactions in quantum dots should be reliable for dots larger than about 15–20-Å radius even for simple models for the basis orbitals. More detailed calculations of basis orbitals are required for smaller dots.

TABLE IX. Nearest-neighbor unscreened Coulomb and exchange integrals with Löwdin orthogonalized Gaussian-type hybrids (O-GTO); nonorthogonal Gaussian-type hybrids (NO-GTO); and nonorthogonal Slater orbitals (NO-SO). The GTO integrals were calculated with the MOLPRO (Ref. 43) package using the pseudopotentials from the Los Alamos group (Ref. 44) for a two-atom molecule with a bond length given by the bulk value. The SO integrals are from our Monte Carlo calculations. The indices B and N designate the bonding and nonbonding sp^3 hybrids, respectively.

Si	O-GTO	NO-GTO	NO-SO
$\omega^0_{\mathrm{Coul}}(B,B)$	8.04	10.01	10.60
$\omega^0_{\mathrm{Coul}}(B,N)$	5.96	6.65	6.78
$\omega^0_{\mathrm{Coul}}(N,N)$	4.64	4.67	4.89
$\omega^0_{\mathrm{exch}}(B,B)$	0.27	6.20	
$\omega^0_{\mathrm{exch}}(B,N)$	0.11	0.43	
$\omega^0_{\mathrm{exch}}(N,N)$	0.04	0.32	
InAs	**O-GTO**	**NO-GTO**	**NO-SO**
$\omega^0_{\mathrm{Coul}}(B,B)$	6.94	8.77	9.06
$\omega^0_{\mathrm{Coul}}(B,N)$	5.50	6.39	6.59
$\omega^0_{\mathrm{Coul}}(N,B)$	5.02	5.42	5.43
$\omega^0_{\mathrm{Coul}}(N,N)$	4.12	4.18	4.24
$\omega^0_{\mathrm{exch}}(B,B)$	0.28	4.90	
$\omega^0_{\mathrm{exch}}(B,N)$	0.16	0.59	
$\omega^0_{\mathrm{exch}}(N,B)$	0.04	0.18	
$\omega^0_{\mathrm{exch}}(N,N)$	0.04	0.29	
CdSe	**O-GTO**	**NO-GTO**	**NO-SO**
$\omega^0_{\mathrm{Coul}}(B,B)$	6.94	8.77	9.06
$\omega^0_{\mathrm{Coul}}(B,B)$	6.89	8.66	8.74
$\omega^0_{\mathrm{Coul}}(B,N)$	5.66	6.62	6.84
$\omega^0_{\mathrm{Coul}}(N,B)$	4.85	5.16	5.01
$\omega^0_{\mathrm{Coul}}(N,N)$	4.06	4.11	4.13
$\omega^0_{\mathrm{exch}}(B,B)$	0.27	4.35	
$\omega^0_{\mathrm{exch}}(B,N)$	0.19	0.69	
$\omega^0_{\mathrm{exch}}(N,B)$	0.03	0.13	
$\omega^0_{\mathrm{exch}}(N,N)$	0.04	0.24	

For excitonic gaps, we obtained good agreement with recent experiments for both Si and CdSe quantum dots. However, the gaps for InAs quantum dots agree less well with experiment. Especially for Si, we improved the agreement with experimental data by optimizing the tight-binding parameters to give better effective masses compared to the parameters of Vogl *et al.*[24] We also showed that, in contrast to the electron and hole single-particle energies, the electron-hole Coulomb interaction is not very sensitive to the choice of parameters.

Finally, we studied the effective range of the exchange interaction. Replacing the Coulomb potential with a cutoff potential, we explored the dependence of the exchange interaction on the cutoff radius. For direct-gap materials, the lack of on-site orthogonality causes the exchange interaction to be long ranged. For an indirect material Si, the calculations

show that the exchange interaction is oscillatory and has a range of about 15 Å .

ACKNOWLEDGMENTS

We thank Jeongnim Kim for many helpful discussions. This work was supported by NSF (Grant No. PHY-9722127) and by the NCSA and OSC supercomputer centers. Part of the work described in this paper was carried out by the Jet Propulsion Laboratory (JPL), California Institute of Technology under a contract with the National Aeronautics and Space Administration. The supercomputer used in this investigation at JPL was provided by funding from the NASA Offices of Earth Science, Aeronautics, and Space Science.

APPENDIX

Since we use empirical tight-binding wave functions, the choice of specific atomistic orbitals for matrix-element calculations is largely arbitrary. The results presented in the main text were based on Coulomb and exchange integrals calculated with Slater's atomic orbitals, obtained from Slater's rules.[29] In addition, we neglected all off-site exchange integrals. An alternative for unscreened integrals is to use one of the standard quantum chemistry Gaussian-based commercial packages. However, screened matrix elements cannot be obtained in this way, since there is no way in these codes to include a spatially varying screening function.

Although we cannot obtain screened integrals, two important questions can be answered by a comparison between integrals from Gaussian-type orbitals (GTO) and Slater's orbitals (SO): what is the typical variation in the integral values for two reasonable choices of orbitals; and what is the effect of using nonorthogonal bond hybrids rather than properly orthogonalized hybrids? The underlying assumption in the tight-binding approach is that the orbitals on different sites are orthogonal.

Table VIII shows a comparison between orthogonal GTO (O-GTO), nonorthogonal GTO (NO-GTO), and nonorthogonal SO (NO-SO) for on-site Coulomb and exchange integrals. Typically, the NO-SO and NO-GTO Coulomb integrals differ by 10%, whereas the (an order of magnitude smaller) exchange integrals differ by 20–50%. Orthogonalization generally gives an additional 10% change. The use of nonorthogonal Slater orbitals can therefore be estimated to imply 20% overall uncertainty in the on-site integrals.

A similar comparison for nearest-neighbor integrals is shown in Table IX. Here the difference between NO-GTO and NO-SO is less than 10%, but orthogonalization can yield a lowering of up to 30% in the Coulomb integrals between bonding orbitals. The most dramatic effect, however, is that the exchange integrals essentially become negligible when orthogonalized hybrids are used. Notably, nonorthogonal hybrids cannot be used for the bonding-bonding off-site exchange integrals, since these integrals are quite large without orthogonalization but are reduced by a factor of 20–30 after orthogonalization.

In conclusion, the dominant Coulomb integrals obtained from the Slater orbitals can be considered accurate only to 20–30 % due to the sensitivity to different functional representations and to effects of orthogonalization. Further, proper orthogonalization reveals that all offsite exchange integrals can be neglected, including those between bonding hybrids.

[1] P. D. J. Calcott, K. J. Nash, L. T. Canham, M. J. Kane, and D. Brumhead, J. Phys.: Condens. Matter **5**, L91 (1993).

[2] M. Chamarro, C. Gourdon, P. Lavallard, O. Lublinskaya, and A. I. Ekimov, Phys. Rev. B **53**, 1336 (1996).

[3] D. J. Norris and M. G. Bawendi, Phys. Rev. B **53**, 16 338 (1996).

[4] U. Banin, C. J. Lee, A. A. Guzelian, A. V. Kadavanich, A. P. Alivisatos, W. Jaskolski, G. W. Bryant, Al. L. Efros, and M. Rosen, J. Chem. Phys. **109**, 2306 (1998).

[5] E. Dekel, D. Gershoni, E. Ehrenfreund, D. Spektor, J. M. Garcia, and P. M. Petroff, Phys. Rev. Lett. **80**, 4991 (1998).

[6] M. V. Wolkin, J. Jorne, P. M. Fauchet, G. Allan, and C. Delerue, Phys. Rev. Lett. **82**, 197 (1999).

[7] B. Alperson, I. Rubinstein, G. Hodes, D. Porath, and O. Millo, Appl. Phys. Lett. **75**, 1751 (1999).

[8] U. Banin, Y. Cao, D. Katz, and O. Millo, Nature (London) **400**, 542 (1999).

[9] H. Petterson, C. Pryor, L. Landin, M.-E. Pistol, N. Carlsson, W. Seifert, and L. Samuelson, Phys. Rev. B **61**, 4795 (2000).

[10] O. Millo, D. Katz, Y. Cao, and U. Banin, Phys. Rev. B **61**, 16 773 (2000).

[11] Al. L. Efros, M. Rosen, M. Kuno, M. Nirmal, D. J. Norris, and M. Bawendi, Phys. Rev. B **54**, 4843 (1996).

[12] S. Öğüt, J. R. Chelikowsky, and S. G. Louie, Phys. Rev. Lett. **79**, 1770 (1997).

[13] A. Franceschetti, H. Fu, L. W. Wang, and A. Zunger, Phys. Rev. B **60**, 1819 (1999).

[14] A. J. Williamson and A. Zunger, Phys. Rev. B **61**, 1978 (2000).

[15] C. Delerue, G. Allan, and M. Lannoo, Phys. Rev. B **48**, 11 024 (1993).

[16] E. Martin, C. Delerue, G. Allan, and M. Lannoo, Phys. Rev. B **50**, 18 258 (1994).

[17] K. Leung and K. B. Whaley, Phys. Rev. B **56**, 7455 (1997).

[18] K. Leung, S. Pokrant, and K. B. Whaley, Phys. Rev. B **57**, 12 291 (1998).

[19] G. Allan, Y. M. Niquet, and C. Delerue, Appl. Phys. Lett. **77**, 639 (2000).

[20] C. Delerue, M. Lannoo, and G. Allan, Phys. Rev. Lett. **84**, 2457 (2000).

[21] See, e.g., D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. Den-Baars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993); C. B. Murray, D. J. Norris, and M. G. Bawendi, J. Am. Chem. Soc. **115**, 8706 (1993); J. E. B. Katari, V. L. Colvin, and A. P. Alivisatos, J. Phys. Chem. **98**, 4109 (1994); T. Vossmeyer, L. Katsikas, M. Giersig, I. G. Popovic, and H. Weller, *ibid.* **98**, 7665 (1994).

[22] For example, see N. H. Bonadeo, J. Erland, D. Gammon, D. Park, D. S. Katzer, and D. G. Steel, Science **282**, 1473 (1998); L.

Landin, M. S. Miller, M.-E. Pistol, C. E. Pryor, and L. Samuelson, *ibid.* **280**, 262 (1998); Y. Toda, T. Sugimoto, M. Nishioka, and Y. Arakawa, Appl. Phys. Lett. **76**, 3887 (2000).

[23] L. Hedin and S. Lundqvist, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1969), Vol. 23, p. 1.

[24] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).

[25] A. Franceschetti, L. W. Wang, H. Fu, and A. Zunger, Phys. Rev. B **58**, 13 367 (1998).

[26] See, for example, H. Haken, *Quantum Field Theory of Solids* (North-Holland, Amsterdam, 1976); R. S. Knox, *Solid State Physics* (Academic, New York, 1963), Vol. 5.

[27] Recent experiments for CdSe are based on wurtzite CdSe, but we model CdSe as a zinc-blende material in this study since the small value of the crystal-field splitting in CdSe (25 meV) makes it reasonable to approximate this semiconductor as a zinc-blende material.

[28] R. S. Knox, *Solid State Physics* (Academic, New York, 1963), Vol. 5, p. 7.

[29] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[30] K. Ohno, Theor. Chim. Acta **2**, 219 (1964).

[31] Specifically, a Monte Carlo integration method with biased random walk is used. A sequence of six-dimensional coordinates $(\mathbf{r},\mathbf{r}')$ is generated with the distribution of the charge density $|\phi_{n_1}(\mathbf{r})\phi_{n_2}(\mathbf{r}')|^2$ using a biased random walk which starts at some plausible coordinates and in which the size of the attempted steps is uniformly distributed between preset limits. The criterion to accept a new coordinate set $(\mathbf{r}_{i+1},\mathbf{r}'_{i+1})$ is (i) if the probability $|\phi_{n_1}(\mathbf{r}_{i+1})\phi_{n_2}(\mathbf{r}'_{i+1})|^2$ of the new coordinate set is larger than the probability $|\phi_{n_1}(\mathbf{r}_i)\phi_{n_2}(\mathbf{r}'_i)|^2$ of the previous set, it is always accepted; (ii) otherwise the new set is accepted with a probability of the ratio of the new probability to the previous probability. The integration is done by averaging the Coulomb potential over a set of $(\mathbf{r},\mathbf{r}')$ generated by this process. For a general review, see, for example, K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction* (Springer, New York, 1997).

[32] R. Resta, Phys. Rev. B **16**, 2717 (1977).

[33] D. R. Penn, Phys. Rev. **128**, 2093 (1962).

[34] R. Tsu, L. Ioriatti, J. Harvey, H. Shen, and R. Lux, in *Microcrystalline Semiconductors: Materials Science and Devices*, edited by P. M. Fanchet, C. C. Tsai, L. T. Conham, I. Shimizu, and Y. Aoyagi, Mater. Res. Soc. Symp. Proc. **283** (Materials Research Society, Pittsburgh, 1993), p. 437.

[35] The Monte Carlo integration method with a biased random walk needs a non-negative density function that can be used as the probability of accepting the random walks. Since the exchange integrals involve a density function that is not positive definite, this method is not straightforward to use.

[36] *Semiconductors—Basic Data*, edited by Otfried Madelung, 2nd revised ed. (Springer, New York, 1996).

[37] We have used the difference between the first prominent peaks for positive- and negative-bias voltage to calculate the STM gaps from the published dI/dV curves. No adjustment regarding charging energy or polarization energy is made in the STM gaps.

[38] G. Klimeck, R. C. Bowen, T. B. Boykin, C. Salazar-Lazaro, T. A. Cwik, and A. Stoica, Superlattices Microstruct. **27**, 77 (2000).

[39] $E(\gamma)$ denotes an on-site Hamiltonian matrix element for the orbital γ, and $V(\gamma,\gamma')$ denotes a nearest-neighbor Hamiltonian matrix element between the two orbitals γ and γ'.

[40] P. E. Lippens and M. Lannoo, Phys. Rev. B **41**, 6079 (1990).

[41] A. Franceschetti and A. Zunger, Phys. Rev. B **62**, 2614 (2000).

[42] To be clear about this notation, ''*s*-like hole'' means that the coefficients of the *p* basis orbital has *s* symmetry in real space over the whole dot, whereas ''*p*-like hole'' means that the coefficients of the *p* basis orbital has *p* symmetry. For electron states, the coefficients of the *s* basis orbital are used to label the wave functions. For example, ''*s*-like electron'' means that the coefficients of the *s* basis orbital has *s* symmetry, while ''*p*-like electron'' has *p* symmetry in the coefficients of the *s* orbital.

[43] MOLPRO is a package of *ab initio* programs written by H.-J. Werner, and P. J. Knowles, with contributions from R.D. Amos, A. Bernhardsson, A. Berning, P. Celani, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, C. Hampel, G.Hetzer, T.Korona, R. Lindh, A.W. Lloyd, S.J. McNicholas, F.R. Manby, W. Meyer, M.E. Mura, A. Nicklass, P.Palmieri, R. Pitzer, G. Rauhut, M. Schütz, H. Stoll, A.J. Stone, R. Tarroni, and T. Thorsteinsson. The integral evaluation (SEWARD) is described in R. Lindh, U. Ryu, and B. Liu, J. Chem. Phys. **95**, 5889 (1991). Use of this package does not imply an endorsement or specific validation by NIST.

[44] P. J. Hay and W. R. Wadt, J. Chem. Phys. **82**, 270 (1985).