Electronic excitations in single-walled GaN nanotubes from first principles: Dark excitons and unconventional diameter dependences

Sohrab Ismail-Beigi
Department of Applied Physics, Yale University, New Haven, Connecticut 06520, USA
(Received 21 August 2007; revised manuscript received 22 October 2007; published 4 January 2008)

We present the first *ab initio* predictions of electronic and excitonic states in gallium nitride nanotubes. Electron-hole interactions dramatically affect optical properties. Low-energy excitons are dark (dipole forbidden) in all nanotubes with key ramifications for applications. We describe an unusual decrease of energy gaps with decreasing nanotube diameter, opposing expectations from quantum confinement. This stems from a combination of nanoscale curvature and gallium chemistry, should apply to other systems, and furnishes an interesting way to reduce excitation energies with decreasing dimension.

DOI: 10.1103/PhysRevB.77.035306 PACS number(s): 78.67.Ch, 73.22.−f, 73.63.Fg

INTRODUCTION

Gallium nitride (GaN) is of great importance due to its optoelectronic properties, high thermal and mechanical stability, and the realization of band-gap engineering via In and Al alloying. GaN is already used in light-emitting diodes and room-temperature laser diodes with long operating lifetimes.[1] Having a large band gap, GaN is a candidate for use in high-temperature, high-power electronics and short-wavelength (UV) detectors.[2] Nanostructures of GaN are of interest for many reasons: they should be semiconducting regardless of precise details so that doping will provide control over conductivity, their optical properties should be tunable through quantum size effects, they can serve as molecular sensors, etc.

Currently, GaN nanowires are fabricated and show promise for applications.[3] In these nanowires, most atoms reside in the crystalline wurtzite wire core. Genuine nanotubular geometries, i.e., hollow core cylinders, may offer physical properties different from the bulk. To date, nanotubular GaN has been synthesized,[4] but the nanotube walls are many nanometers thick which, again, means essentially bulklike atomic layouts. Atomically thin GaN nanotubes have not been yet fabricated, in contrast with carbon, boron nitride, or boron single-walled nanotubes.[5–7] GaN nanotubes are of interest for possibly unique optical and luminescent properties. Here, theory can help predict these properties, how they differ from the bulk, to what extent they can be engineered, and whether they are appropriate for envisaged applications.

Theoretical works have used density functional theory[8,9] (DFT) to study single-walled GaN nanotubes (GaNNTs).[11–15] Because GaN is of primary interest for its electronic, optical, and luminescent properties, the use DFT is a serious drawback: DFT fails badly for predicting these electronic excitations.[16,17] We present results for the electronic and optical properties of GaNNTs by using many-body electronic Green's functions. These *ab initio* methods describe excitations accurately in many materials.[16,18–20] Excitation energies are typically within 0.1–0.2 eV of experiment, sufficient for meaningful comparison and believable prediction. Briefly, we underline the significant effect of electron-hole interactions on optical properties, predict and verify the presence of low-energy dark (dipole forbidden) bound excitons in all GaNNTs, and describe and explain an unconventional *decrease* of excitation energies with decreasing nanotube diameter. We conclude with implications of our findings for applications and related material systems.

I. METHODOLOGY

We describe the electronic ground state using DFT with the local-density approximation (LDA), a plane-wave basis, supercells, and norm-conserving Kleinman-Bylander pseudopotentials.[9,10] The N pseudopotential is generated with the $2s^2 2p^3$ configuration, $r_c = 1.80 a_0$, and the p potential is local. The Ga pseudopotential is generated for $4s^2 4p^{0.5} 4d^{0.5}$ with scalar-relativistic corrections, nonlinear core corrections (NLCC),[21] $(r_c^s, r_c^p, r_c^d) = (2.30, 2.19, 4.48) a_0$, and the d potential is local. A plane-wave cutoff of 50 Ry converges ground-state properties. The NLCC gives accurate ground-state physics without requiring a high plane-wave cutoff or inclusion of semicore Ga $3d$ states.[22,23] This then allows for feasible excited-state computations. (We discuss the $3d$ semicore states below.)

Electronic excitations are calculated with *ab initio* many-electron Green's function methods. Quasiparticle physics is described by the Dyson equation within the *GW* approximation to the electron self-energy.[16] We fix the dielectric function to its LDA random-phase approximation (RPA) value. We use identical LDA and *GW* quasiparticle states: Tests show that off-diagonal self-energy elements have negligible effects. Optical excitations are found by solving the Bethe-Salpeter equation (BSE) for the two-particle Green's function.[19] To study isolated nanotubes, we truncate unwanted Coulomb image interactions between periodic supercells.[24] All other parameters are chosen to converge energy eigenvalue differences (LDA, *GW*, or BSE) to 0.1 eV. For optical transitions, we only consider light polarized parallel to the nanotube axis: Depolarization effects ensure much weaker coupling for the perpendicular polarization.[25] For example, for the (3,3) GaN nanotube below, the perpendicular polarization radiative transition rates are ~15 000 times smaller than the parallel ones.

Due to the localized nature of N orbitals, the *GW* calculations require the inclusion of many high-energy bands and a large screening cutoff. We accelerate convergence by split-

TABLE I. Basic lattice and electronic properties of wurtzite GaN based on DFT-LDA, the *GW* approximation, and experiment. Lattice constants a and c are in Å, and energy gaps are in eV and calculated at the theoretically predicted lattice parameters. All calculations use plane waves and pseudopotentials but differ in the treatment of the 3*d* semicore states. "No 3*d*" means that the 3*d* states are incorporated into the frozen core of the pseudopotential with no further treatment. "Explicit" means that the 3*d* states were included as part of the valence manifold. "NLCC" means that the 3*d* are moved in the frozen core but they effect the valence states via the nonlinear core correction.

Ref.	Method	a	c	E_g (LDA)	E_g (*GW*)
Present	No 3*d*	3.08	5.02	3.29	4.5
28	Explicit 3*d*	3.16	5.14	2.04	
30	Explicit 3*d*	3.15	5.14	2.14	
31	NLCC	3.13	5.12	2.3	3.5
Present	NLCC	3.18	5.19	2.25	3.5
32	Experiment	3.19	5.13	3.50	

ting the *GW* contributions into static and dynamic parts.[26] However, we must still include plane waves in the screening matrix with energies up to 10 Ry as well as unoccupied states with energies up to 10 Ry above the Fermi energy, which makes for taxing nanotube computations.

Before presenting results, we discuss the role of the Gd 3*d* semicore states on first principles calculations of GaN. At the DFT level, it has been recognized that a Ga pseudopotential that merely moves the 3*d* states into the frozen core with no further ado is quite unsatisfactory;[27,28] predicted lattice constants can be too small, and more importantly, properties such as lattice constants, bulk moduli, and band gaps are highly variable. The NLCC allows one to keep the 3*d* in the frozen core while greatly improving such predictions and bringing them closer to all-electron results (as well as experiments).[27,28] The NLCC improvement is also observed for the energetics and vibrational modes of GaN defects,[22,23] although in some cases, involving Ga vacancies or pairs of Ga atoms as nearest neighbors, the NLCC is not accurate enough for high quality calculations of defect formation energies, and the 3*d* states must be included explicitly.[23,29] Briefly, for DFT-LDA, some accounting of the semicore 3*d* is a necessity; the NLCC accomplishes this in many, but not all, cases.

Moving beyond DFT, the first *GW* calculations of GaN did not include the 3*d* states but used the NLCC.[31] Those authors noted that the NLCC was crucial in producing reasonable results. Studies that attempted to include the semicore *d* electrons in GaN and other materials where cations have semicore *d* states found that for quasiparticle *GW* calculations that start with the LDA, it is necessary to include the entire atomic shell to obtain good results.[33,34] For Ga, this means taking the entire $3s^2 3p^6 3d^{10}$ as part of the valence. Physically, this is because the Fock exchange matrix elements between these orbitals are sizable: LDA describes such interactions poorly, whereas *GW* is supposed to correct them. However, for a LDA pseudopotential that moves part of the shell into the frozen core, the core-valence exchange interactions are not properly described. Recently, *GW* calculations based on the exact exchange method within the optimized effective potential (OEPx) framework have been applied to materials with semicore states.[35] A key feature of this approach is that the pseudopotential is consistently generated within OEPx so that a proper description of core-valence exchange is automatically included into the frozen core. These OEPx pseudopotentials allow one to perform high quality *GW* calculations while only including the 3*d* states of interest.

From a practical standpoint, however, inclusion of the entire semicore shell requires a plane-wave cutoff that is too high for calculations of the size reported here with currently available computers. While OEPx can, in principle, allow for the inclusion of only the 3*d* states, the computational workload of this approach is quite prohibitive even for small unit cells.[35] Therefore, we are forced to employ a Ga pseudopotential which includes no explicit 3*d* states and uses the NLCC. We treat core-valence interactions within DFT-LDA:[16,31,36] when using the *GW* self-energy Σ as a correction to the LDA exchange-correlation potential V_{xc}, the partial core density is not used when calculating this V_{xc}.

To gauge the expected accuracy of the approach, Table I compiles plane-wave pseudopotential results for the basic properties of wurtzite GaN. The first entry shows a typical result when we try to use a Ga pseudopotential with no treatment of the 3*d* states: lattice constants are quite small and band gaps are large. More generally, we have observed that results based on such pseudopotentials are highly *variable* and depend on seemingly irrelevant details of the Ga potential. With the NLCC, results are much closer to those explicitly including the 3*d* states (and also to experiment).

We note that while the LDA results with and without core corrections are quite different, the *GW correction* to the band gap is essentially the same for both to within ~0.1 eV. This effect was seen previously in the case of CdS,[33] where, for energy bands about the valence band maximum and conduction band minimum, *GW* corrections to the LDA were basically the same for pseudopotentials with no valence semicore states and those with the entire semicore shell in the valence. Therefore, a careful choice of Ga NLCC pseudopotential should be able to deliver a good description of the ground-state structure as well as the low-energy electronic excitations based on *GW* corrections and the BSE approach.

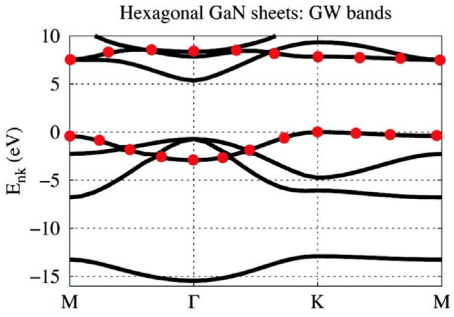

FIG. 1. (Color online) Quasiparticle bands for the hexagonal GaN sheet based on the *GW* approximation. The valence band maximum at K is the zero of energy. The bands outlined with red circles are the π and π^* bands arising from out-of-plane p_z orbitals. The other bands are in-plane σ and σ^* states.

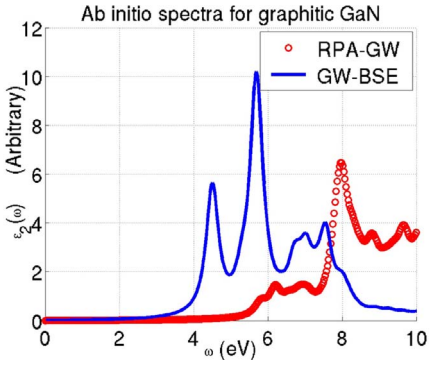

FIG. 2. (Color online) Imaginary part of the frequency-dependent dielectric function $\varepsilon_2(\omega)$ for the hexagonal GaN sheet based on the RPA approximation with *GW* band energies (red circles) and based on the BSE which includes excitonic effects (solid blue). The first BSE peak corresponds to the bright exciton at 4.3 eV.

II. SHEETLIKE (TWO-DIMENSIONAL) FORM OF GaN

To properly orient a study of GaNNTs, we begin with the physics of the two-dimensional sheetlike material from which the nanotubes are constructed. We perform LDA calculations on a variety of possible GaN sheet structures. For example, we consider graphiticlike hexagonal lattices, triangular and square lattices, bilayers of such sheets, flat and buckled sheets, etc. The hexagonal sheet, isostructural to BN sheets,[6] with lattice vector of 3.17 Å has the lowest energy. This sheet is metastable by 0.39 eV/atom when compared to wurtzite GaN. Thus, fabrication of GaNNTs may require kinetic growth techniques, but it is encouraging that AlN nanotubular materials can be fabricated[37] even though their sheet energy is even higher at 0.68 eV/atom above bulk.[38] For use below, we find optical phonon frequencies at Γ for the sheet at 288 cm^{-1} (out of plane) and 758 cm^{-1} (in plane). The quadratic elastic energy increase $\delta\mathcal{E}=\mu\epsilon^2$ for in-plane stretching by a fraction ϵ has modulus $\mu=20.6$ eV/atom.

At the LDA level, this sheet has an indirect band gap of 2.4 eV between the valence band maximum (VBM) at *K* and the conduction band minimum (CBM) at Γ, and the smallest direct gap is 3.1 eV at Γ and is optically allowed. When we include *GW* self-energy effects, the gap remains indirect ($K \rightarrow \Gamma$) but increases to 5.0 eV. The minimum *GW* direct gap is 5.6 eV and at Γ. The VBM at *M* has an energy of 0.4 eV below the VBM at *K*. The *GW* corrections are much larger than those of bulk GaN (see Table I). Thus, bulk *GW* corrections cannot simply be copied over to reduced dimensional systems without serious loss of accuracy.[11] A main physical reason is the dissimilar screening. In a three-dimensional medium, all electric field lines emanating from an electron or hole are screened since there is material in all directions. In reduced dimensions, field lines can easily leave the material, so screening is anisotropic and much reduced. It is the screening behavior of the many-body medium that determines the *GW* corrections to DFT.

Figure 1 shows the sheet band structure. The ionicity difference of Ga and N causes the π and π^* bands derived from the out-of-plane p_z orbitals to split by a large energy. In graphene, these bands are degenerate at the Fermi energy at *K*.[39] Here, a "memory" of this degeneracy remains: the π-π^* gap is smallest at *K* and the sheet VBM remains there.

Figure 2 displays the imaginary part of the dielectric function for the sheet, which is directly related to the optical absorption cross section. The large and qualitative differences between the single-particle RPA-*GW* and two-particle *GW*-BSE methods highlight the crucial role and increased strength of electron-hole attractions in reduced dimensional systems. The lowest-energy *GW*-BSE peak is from an exciton with binding energy of 1.3 eV. This is the sheet's lowest excitation and is composed of electrons and holes about Γ.

The fact that the sheet has an indirect gap has important consequences for the electronic excitations of the nanotubes. Within the zone-folding scheme for nanotubes,[39] we fold the sheet bands along the chiral direction to find the nanotube states. When considering direct interband transitions, "rolling" the sheet turns states about the sheet *K* point into nanotube VBM states at the nanotube Γ for $(n,0)$ nanotubes, whereas the region about the sheet *M* point turns into the nanotube VBM at the nanotube Γ for (n,n) tubes. The CBM is at Γ for both sheets and tubes. Thus, the nanotube's lowest-energy direct transitions will be at its Γ point and will take place between states with differing sheet crystal momenta or equivalently different angular momenta about the tube axis. Optical transitions between them are forbidden by this momentum mismatch, so we expect dipole-forbidden (dark) low-energy excitations in GaNNTs.

To state this in more general terms, when a sheetlike material with an indirect band gap is "rolled up" to create nanotubes, the zone-folding scheme predicts that the lowest-energy nanotube interband transitions will be dark due to a mismatch of crystal momenta. *Ceteris paribus*, we would expect that the lowest-energy excitons would also be dark.

III. GaN NANOTUBES

A. Structure and vibrational properties

We now describe our results for GaN nanotubes. We begin with basic structural and vibrational properties. Table II

TABLE II. The set of GaN nanotubes in this work for which LDA calculations are performed. All lengths are in Å. Listed are the nanotube (n,m) indices, the number of atoms per unit cell along the nanotube length N_{at}, the size of the square base of the unit cell in the x and y directions a_{xy}, the length of the unit cell along the nanotube or z direction a_z, the mean diameter d, the radial buckling Δr, and the direct LDA band gap for the nanotube E_g^{LDA} (in eV). See text for details.

Tube	N_{at}	a_{xy}	a_z	d	Δr	E_g^{LDA}
(4,0)	16	12	5.15	4.41	0.25	1.93
(5,0)	20	13	5.25	5.34	0.21	1.90
(6,0)	24	13	5.43	6.24	0.17	1.88
(7,0)	28	14	5.45	7.20	0.14	2.02
(3,3)	12	13	3.12	5.40	0.21	2.09
(4,4)	16	15	3.11	7.14	0.16	2.31
(5,5)	20	17	3.18	8.78	0.12	2.31
(6,6)	24	19	3.18	10.5	0.10	2.42

describes the nanotubes that we have considered. For each nanotube indexed by (n,m), we perform calculations by placing the nanotube axis along the z direction. This is the physically periodic direction with primitive lattice constant a_z in the table. The choice of (n,m) determines N_{at}, which is the number of atoms per unit cell along z. We choose the supercell to be square in the xy plane. Since we use the Coulomb truncation method,[24] the actual shape or precise dimension of the cell in the xy direction is not relevant once it is large enough to ensure convergence of calculated quantities. The length of the side of the square in the xy plane where convergence is achieved is a_{xy} in the table and is the value we employ.

We relax the nanotube structures within LDA. Our starting guess for the structure is generated by rolling the GaN sheet into the nanotube. As found previously,[11] the N atoms lie at a larger radius than the Ga atoms from the nanotube central axis. Therefore, the diameter d in the table is the average of the two values. The difference in radius, the radial buckling Δr, decays to zero with increasing diameter. We also report the direct LDA band gap at the nanotube Γ point in the table. The $(n,0)$ and (n,n) direct band gaps behave differently and approach different values: for very large tubes, these turn into to the $K \rightarrow \Gamma$ and $M \rightarrow \Gamma$ transitions of the sheet with LDA band gaps of 2.4 and 2.8 eV, respectively.

Previous work has reported on ground-state properties such as structure, strain energies, or mechanical properties,[11,13–15] so we concentrate on other findings. Since Raman spectroscopy is an important tool for experimental identification of nanotubes,[7,40] we describe the basic data for important long-wavelength phonons corresponding to radial breathing modes. We calculate phonon frequencies of a few GaNNTs within LDA, and Table III lists the results. The symmetric mode where Ga and N move radially and in phase has frequency ν_s, while the out-of-phase mode has frequency ν_a. For a large nanotubes where $d \rightarrow \infty$, these modes revert to those of the sheet: the symmetric mode turns into an arbitrarily long wavelength in-plane stretch, and the antisymmetric mode turns into the out-of-plane optical mode at Γ of the sheet, i.e., $\nu_s \rightarrow 0$ and $\nu_a \rightarrow 288$ cm^{-1}.

In Fig. 3, we show the frequencies of both modes versus d. We note that lack of dependence of the frequencies on the class of nanotube, which mirrors the behavior of the strain

TABLE III. Vibrational frequencies in cm^{-1} for radial breathing modes of GaN nanotubes based on LDA calculations. d is the mean nanotube diameter in Å. The frequency ν_s is for the symmetric mode where Ga and N move in phase, and ν_a is for the antisymmetric out-of-phase mode.

Tube	d	ν_s	ν_a
(4,0)	4.41	193	439
(5,0)	5.34	168	406
(6,0)	6.24	151	370
(7,0)	7.20	135	353
(3,3)	5.40	166	399
(4,4)	7.14	137	358
(5,5)	8.78	116	335

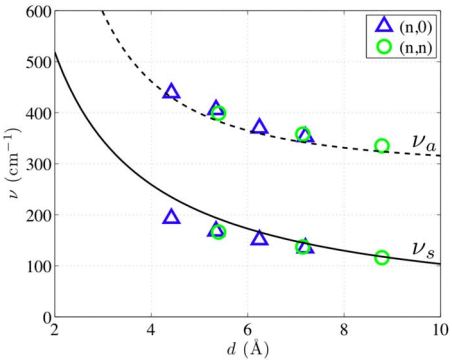

FIG. 3. (Color online) Radial breathing mode frequencies in cm^{-1} versus nanotube diameter d in Å. The upper set of data is for the antisymmetric mode ν_a, and the lower set of data is for the symmetric mode ν_s. Open (blue) triangles are for $(n,0)$ nanotubes, while (green) circles are for (n,n) nanotubes. The upper continuous dashed curve is $\nu_a = 288 + 2070/d^2$, and the lower solid continuous curve is $\nu_s = 1040/d$.

TABLE IV. Lowest-energy exciton energies in eV for single-walled GaNNTs. The lowest excitations are dark (dipole forbidden); the lowest bright (optically allowed) exciton is also listed. d is the nanotube diameter, Ω the exciton energy, E_{inter} the dominant interband transition for that exciton, and Δ the exciton binding energy. The sheet data provide the relevant limiting case for large diameter (Ref. 45).

		Dark excitons			Bright excitons		
Tube	d (Å)	Ω	E_{inter}	Δ	Ω	E_{inter}	Δ
(4,0)	4.4	3.2	4.7	1.5	3.5	4.9	1.5
(5,0)	5.4	3.3	4.7	1.4	3.8	5.2	1.4
Sheet	∞		5.0		4.3	5.6	1.3
(3,3)	5.4	2.5	4.9	2.4	3.2	5.4	2.2
(4,4)	7.1	3.0	5.1	2.1	3.5	5.4	1.9
Sheet	∞		5.4		4.3	5.6	1.3

energies:[11] the results for $(n,0)$ and (n,n) lie on the same curve versus diameter. We have only calculated results for the achiral $(n,0)$ and (n,n) tubes. However, since these two are the extreme cases for the choice of chiral angle,[39] we expect that results for other chiral nanotubes should lie between the two extremes. Under this assumption, we expect no chirality dependence to these vibrational frequencies in GaNNTs.

For the symmetric mode, one can make a simple analytical model to extract ν_s versus d using the data from the GaN sheet alone. In a continuum elastic picture, the symmetric radial mode corresponds to in-plane stretching of a GaN sheet. As noted above, such stretches cost a potential energy $\delta\mathcal{E}=\mu\epsilon^2$, when stretching by a fraction ϵ. A radial displacement δr means $\epsilon=\delta r/r$. The kinetic energy per atom is $\bar{m}\delta\dot{r}^2/2$, where $\bar{m}$ is the average atomic mass of Ga and N. The frequency of this quadratic system is $\nu_s=\sqrt{\mu/\bar{m}}/(\pi d)=1040/d$, with d in Å and ν_s in cm^{-1}. We plot this curve in Fig. 3: there is some deviation at small d, but for $d\approx 9$ Å, the it is already less than 2% and will only improve for larger d.

The behavior of the antisymmetric mode ν_a is more difficult to model analytically. To understand the scaling of ν_a versus d, we implement a recent tight-binding model for GaN.[41] Although the model was designed for bulk systems, we have verified that it describes well the geometries of GaN sheets and nanotubes. We use this model to calculate phonon frequencies for a large sample of nanotubes. For the symmetric mode, we verify again that the behavior for large d is $\nu_s\propto d^{-1}$. For the antisymmetric mode, the behavior is $\nu_a=A+Bd^{-2}$ for constants A and B. We fit our *ab initio* data to this mathematical form: A must equal 288 cm^{-1} as explained above, and the best fit for B yields the final form $\nu_a=288+2070/d^2$. By $d\approx 9$ Å, the error is less than 4%. We emphasize that we have used the tight-binding model only to find the mathematical *form* of the function but not any of the actual constants.

Finally, we note that there exists a tangential mode where Ga and N atoms move out of phase along the surface of the nanotube. When $d\to\infty$, this mode turns into the in-plane optical phonon of the sheet at Γ with frequency $\nu_t=758$ cm^{-1}. The tight-binding model predicts the behavior $\nu_t=C+Dd^{-2}$, where $D<0$. We have not calculated any *ab initio* data for this mode, so precise determination of D is not possible. However, based on a crude scaling of the tight-binding data that sets $C=758$ cm^{-1}, we roughly estimate that $\nu_t\approx 758-2190/d^2$.

B. Electronic and optical properties

We now consider the electronic states of GaN nanotubes. Table IV summarizes our findings for the low-energy optical properties of four small-diameter achiral $(n,0)$ and (n,n) nanotubes. Figure 4 shows the imaginary part of the dielectric function for each of these nanotubes. We now discuss salient features of these results. First, exciton binding energies are large and $\sim 1-2$ eV. Due to further dimensional reduction, binding energies are larger in the nanotubes than for the sheet. (The exciton binding energy is the difference between the excitation energy Ω and the lowest interband transition E_{inter} for which the exciton wave function has a sizable amplitude.) Figure 5 shows a representative probability distribution for the lowest exciton of the (3,3) nanotube. Due to its large binding energy, the exciton is rather localized both around and along the nanotube axis. Second, as anticipated, the lowest-energy excitons are dipole forbidden (dark). There is a significant $\sim 0.5-1.0$ eV difference between the lowest dark and first bright exciton energies. Third, the bright excitons in $(n,0)$ and (n,n) tubes both approach the same limit: the sheet's bright exciton. However, the dark excitons approach different limits because, as per band folding, the lowest direct transition for $(n,0)$ tubes corresponds to $K\to\Gamma$ in the sheet, whereas for (n,n), we have $M\to\Gamma$. Based on the $(n,0)$ data, we estimate that the dark $K\to\Gamma$ excitons in large GaNNTs are at ≈ 3.6 eV with binding energy of ≈ 1.4 eV.

Next, excitons in $(n,0)$ and (n,n) tubes approach the sheet limit at different rates. Given a diameter, the $(n,0)$ are much closer to the sheet limit. The difference is large; for example, for exciton binding energies, at $d\approx 5$ Å, the (5,0) energy is within 0.1 eV of the sheet values, while the (3,3) is 0.9 eV off. The difference is also visible in the plots of Fig. 4; comparing the *GW*-BSE spectra for (5,0) and (3,3) to the sheet's, we see that the (5,0) results are closer to the sheet for the first and second low-energy peaks. We have not found the reason for the difference. For example, a key structural difference between the sheet and the nanotubes is the radial

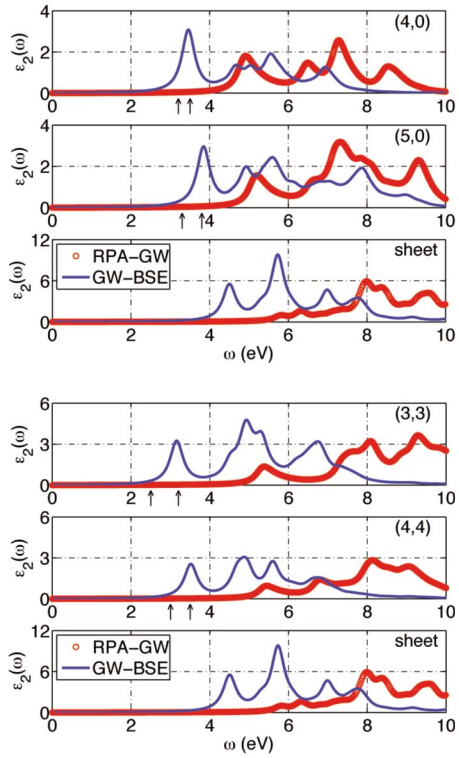

FIG. 4. (Color online) Imaginary part of the dielectric function $\varepsilon_2(\omega)$ for four GaN nanotubes and the GaN sheet (from Fig. 2). The top panel has plots for (4,0), (5,0), and the sheet; the bottom panel has plots for (3,3), (4,4), and the sheet. Thick red circles are results based on the RPA with *GW* band energies, and the thinner solid blue lines are BSE results including excitonic effects. The vertical scale in each plot is arbitrary. For each nanotube, two arrows mark the energies of lowest dark and bright excitons from Table IV. The leftmost arrow is for the dark exciton, and the other is for the bright exciton and necessarily coincides with the peak in the BSE spectrum.

buckling, but the buckling amplitude for $(n,0)$ and (n,n) are equal for the same diameters (see Table II).

Finally, a very unusual feature in Table IV is the decrease in excitation energies with decreasing diameter. This trend was seen at the LDA level,[11] and it holds at the more sophisticated *GW*-BSE level for interband and exciton energies. This effect is opposite to that expected from quantum confinement; e.g., in carbon nanotubes, band gaps increase with decreasing diameter.[39]

C. Unusual diameter dependence of excitation energies

A refined examination of the states at the nanotube Γ point shows that the unusual decrease of excitation energies with decreasing nanotube diameter is due mainly to the systematic variation of the CBM energy. Figure 6 shows the behavior of the *GW* quasiparticle VBM and CBM energies at Γ versus diameter. In Fig. 6, the VBM energies show no significant pattern, while the CBM energies display a clear decrease with decreasing diameter that is also independent of the class of nanotube, i.e., $(n,0)$ versus (n,n).

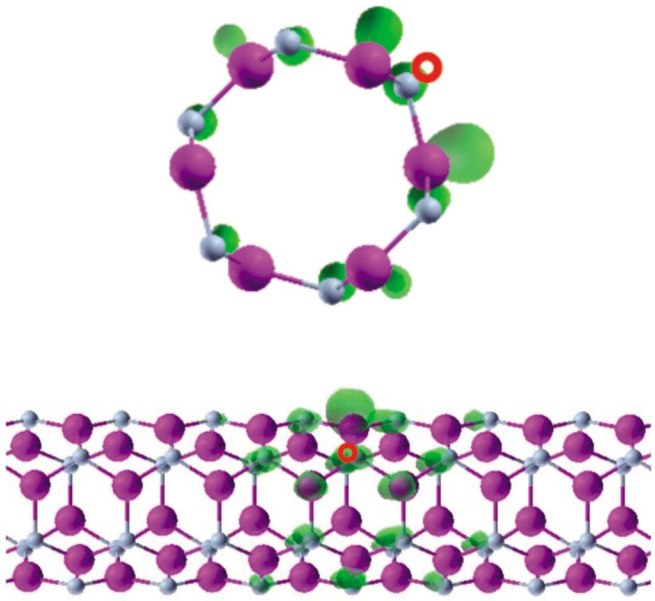

FIG. 5. (Color online) Isosurfaces of probability density (in green/medium gray) of finding an electron when the hole is placed at the position of the red open circle (close to a N atom) for the lowest dark exciton of the (3,3) nanotube. The isosurface is at 25% of its maximum value. The top view is down the nanotube axis and shows the distribution around the nanotube circumference. The bottom view shows the distribution along the nanotube axis. Larger purple/dark gray balls are Ga, and smaller cyan/light gray balls are N atoms.

To obtain the data in Fig. 6, we had to resolve two technical obstacles. First, for the smaller nanotubes listed in Tables IV and V, we have the calculated *GW* band energies on hand, but for the larger nanotubes where *GW*-BSE calculations are prohibitive and have not been performed. For the larger tubes, we have interpolated the *GW* correction to the LDA. Specifically, we already have LDA data on the nanotubes in Table II. As shown in Table V, *GW* corrections to the LDA change little with diameter. Hence, for the larger nanotubes, we interpolate the *GW* correction between the available nanotube and sheet extremes. For example, the in-

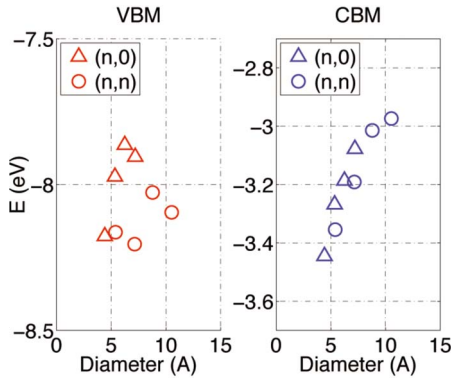

FIG. 6. (Color online) Valence band maximum (left) and conduction band minimum (right) energies, both at Γ, for single-walled GaNNTs versus diameter and with respect to the vacuum. Triangles are for $(n,0)$, and circles for (n,n) tubes.

TABLE V. Lowest-energy direct band gaps at Γ for GaN nanotubes. All energies are in eV, and d is the nanotube diameter in Å. Listed are the DFT-LDA gaps E_g^{LDA}, the GW gaps E_g^{GW}, and the difference between the two ΔE_g (i.e., the GW correction to LDA). The sheet data provide the relevant limiting case for large diameter.

Tube	d (Å)	E_g^{LDA}	E_g^{GW}	ΔE_g
(4,0)	4.4	1.9	4.7	2.8
(5,0)	5.3	1.9	4.7	2.8
Sheet	∞	2.4	5.0	2.6
(3,3)	5.4	2.1	4.9	2.8
(4,4)	7.1	2.3	5.1	2.8
Sheet	∞	2.8	5.4	2.6

terpolated GW correction for the VBM-CBM gap is taken to be 2.7 eV for the larger nanotubes. We expect this procedure to yield energies with $\lesssim$0.1 eV of error.

Second, more important than this interpolation issue is the problem of a common zero of energy: We are comparing electronic states of different nanotubes that are calculated in different-sized supercells. It is well known that electronic energies from plane-wave calculations can have an arbitrary energy shifts depending on the supercell size.[9] We reference energies with respect to the vacuum and remove arbitrary energy shifts from the calculations by performing a few DFT-LDA calculations at different cell sizes and extrapolating to infinite volume.[42] The initial ambiguity of the zero of energy stems from the long range of the Hartree interaction, which is handled by DFT. GW corrections stem from shorter range effects of exchange and correlation[16] and do not suffer from this problem (we have explicitly checked this for the GaN sheet and the smaller nanotubes). We estimate that our extrapolation procedure has an uncertainty of ≈0.1 eV, mainly stemming from the limited number and relatively small values of volumes used in the extrapolation. Part, but not all, of the variation of the VBM energy in Fig. 6 is due to this uncertainty.

To gain further insight into the differing behavior of the VBM and CBM, we describe the physical character and properties of these states from a linear combination of atomic orbital (LCAO) viewpoint. Projection of the LDA wave functions onto the pseudoatomic orbitals shows that (a) the CBM at Γ for the sheet and the nanotubes is dominated by atomic Ga s orbitals with a smaller out-of-phase N s component (i.e., antibonding σ^* states, see Fig. 7) and (b) the nanotube VBM at Γ and the sheet VBM at K or M are composed essentially of out-of-plane N p orbitals (i.e., bonding π states). Our LCAO projection method is described in the Appendix.

Figure 7 shows a curved segment of a GaN sheet that might be part of a $(n,0)$ GaNNT. Curvature reduces interatomic distances in the circumferential direction. In particular, second- and higher-neighbor distances are more strongly modified than nearest-neighbor ones on a curved surface. Since N atomic orbitals are spatially localized, their second- and higher-neighbor hopping elements are small and not greatly modified by these changes, and this explains why the VBM energies are not much affected by curvature. On the other hand, the CBM is dominated by Ga s states that are spatially extended, so the second- and higher-neighbor interactions are greatly affected, and this, in turn, modifies the energies systematically. The large extent of Ga orbitals is physically significant: e.g., one has a very poor tight-binding description of GaN if one neglects second-neighbor Ga-Ga interactions.[41]

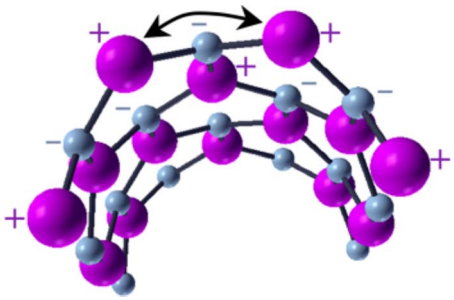

FIG. 7. (Color online) Curved segment of a $(n,0)$ GaNNT. Large purple spheres are Ga, smaller cyan are N, and black lines are bonds. Signs indicate phases of s orbitals for the antibonding CBM state. The arrow highlights the second-neighbor Ga s-s interaction discussed in the text.

Quantitatively, Tables VI and VII present LCAO Hamiltonian and overlap matrix elements between pseudoatomic s orbitals for the GaN sheet and the (4,0) nanotube, respectively. The most important entries in the tables are the Ga-Ga interactions since the CBM state is dominated by the Ga s component. Numerically, the Ga-N interaction turn out to play a small role in determining the CBM energy, while the N-N elements can be ignored for that purpose.

Comparison of the two tables shows how the reduction in distances magnifies the magnitudes of Ga-Ga interactions, whereas changes in Ga-N and N-N matrix elements are smaller. Since s-s interaction energies are always negative, the values become larger in magnitude when the sheet is curved to form the nanotube. The effect is most significant for the second-neighbor Ga-Ga interactions but holds for higher-neighbor interactions. Since the CBM states has all

TABLE VI. LCAO matrix elements of the LDA Hamiltonian H and the overlap operator S between atomic s states for the GaN sheet. Hamiltonian matrix elements are in meV, and overlap elements have been multiplied by 1000. Neighbor shells are identified by their separation in units of the lattice constant of the sheet (a =3.19 Å). We only tabulate matrix elements where H is larger than 10 meV in magnitude.

	Shell	H (meV)	$S\times 1000$
Ga-Ga	1	−2320	121
	$\sqrt{3}$	−110	5
	2	−30	1
Ga-N	$\sqrt{1/3}$	−6780	299
	$\sqrt{4/3}$	−577	23
	$\sqrt{7/3}$	−90	3
N-N	1	−420	14

TABLE VII. LCAO matrix elements of the LDA Hamiltonian H and the overlap operator S between atomic s states for the (4,0) nanotube. Hamiltonian matrix elements are in meV, and overlap elements have been multiplied by 1000. r_{ij}/a is the interatomic distance in units of the sheet lattice constant a. The column "Sheet shell" identifies the sheet neighbor shell to which the interaction belongs in the nomenclature of Table VI. We only tabulate matrix elements where H is larger than 100 meV in magnitude.

	r_{ij}/a	Sheet shell	H (meV)	$S\times 1000$
Ga-Ga	0.942	1	−2920	150
	0.967	1	−2670	140
	1.33	2	−591	31
	1.48	$\sqrt{3}$	−337	16
	1.64	$\sqrt{3}$	−180	7
Ga-N	0.584	$\sqrt{1/3}$	−6760	293
	0.596	$\sqrt{1/3}$	−6490	280
	1.07	$\sqrt{4/3}$	−878	34
	1.15	$\sqrt{4/3}$	−614	22
	1.33	$\sqrt{7/3}$	−261	9
	1.46	$\sqrt{7/3}$	−133	4
	1.50	$\sqrt{7/3}$	−103	4
N-N	1.00	1	−444	13
	1.05	1	−328	9

Ga s states in phase (see Fig. 7), the enhanced Ga s-s matrix elements between all the Ga atoms add together and *lower* the CBM energy. Obviously, this result cannot be explained by a simple appeal to a nearest-neighbor bonding picture: reduced distances would lead us to expect larger bonding-antibonding splittings and the opposite trend versus diameter.

In summary, the unusual behavior versus diameter is due to the chemistry of Ga combined with curvature effects accessible only at the nanoscale. The effect follows from the two rather generic facts that Ga atomic orbitals are spatially extended and that the CBM state is dominated by Ga atomic orbitals. Thus, this behavior and explanation should not be specific to the LDA or other approximations employed here. Moreover, this mechanism can apply in other materials containing cations with large atomic orbitals and provides another way to *reduce* band gaps and excitation energies with decreasing size.

CONCLUSIONS

We have offered state-of-the-art *ab initio* theoretical predictions of the basic vibrational, electronic, optical, and excitonic properties of gallium nitride nanotubes as well as the two-dimensional sheet form of GaN which gives rise to the nanotubes. We have shown that electron-hole interactions lead to strongly bound excitons and dramatically modify the optical spectra when compared to simple interband predictions.

The GaN sheet is shown to be an indirect band-gap material which makes for nanotubes whose lowest-energy direct transitions are optically forbidden (dark). This carries over to the excitons; the lowest-energy excitons in the nanotubes are also found to be dark. Finally, we have observed and explained an unusual decrease in excitation energies with decreasing diameter that opposes expectations based on the usual picture of quantum confinement in nanostructures. This unconventional behavior is due to a combination of nanoscopic curvature and the chemistry of gallium, can apply to other materials, and provides another method to reduce excitation energies with decreasing size.

We end with some implications and speculations. As luminescent materials, GaNNTs may fare poorly due to the unavoidable low-energy dark excitons; once excitons are thermalized in the nanotubes, the excitons have relatively weak radiative strength. This conclusion should also apply to InN nanotubes because their sheet forms have indirect band gaps as well.[43]

Conversely, this very property makes these materials useful when long-lived excitons are needed. For example, in current candidates for low-cost photovoltaic materials, photogenerated excitons must first diffuse through the material in which they are created to reach an interface with another material. There, the exciton can dissociate across the interface into free carriers which then travel in opposite directions to create a current.[44] Clearly, this entire process is more effective and efficient if the radiative lifetime of the excitons is not a bottleneck during the (relatively slow) diffusion to the interface.

Our results show that the optical absorption of GaNNTs begins at ~3.5–4.0 eV, which is beyond the visible spectrum but can be useful for UV applications. To reduce this energy into the visible, InN or alloyed Ga/InN nanotubes may be considered. As explained above, the unusual band-gap behavior versus diameter should apply also to InN nanotubes because it stems from the large size of the cation atomic orbitals, a property shared by In and Ga. This class of materials likely offers strong and tunable absorption combined with very long exciton lifetimes.

ACKNOWLEDGMENTS

We thank Lisa Pfefferle and Jung Han for sharing their interest and knowledge which inspired this work. This work was supported primarily by the National Science Foundation under Contract No. MRSEC DMR 0520495. The Bulldog parallel clusters of the Yale High Performance Computing center provided computing resources.

APPENDIX

Here, we describe the simple LCAO projection method employed in the analysis, discussion, and tables of Sec. III C. During the generation of pseudopotentials for Ga and N, we obtain the atomic radial functions $R_{inl}(r)$, where i labels the atom and n, l are standard atomic quantum numbers. The atomic pseudo-wave-function is

$$\phi_{inlm}(\vec{r}) = \langle \vec{r} \,|\, \phi_{inlm} \rangle = R_{inl}(\|\vec{r} - \vec{r}_i\|) Y_{lm}(\theta_i, \phi_i),$$

where $\vec{r}_i$ is the position of atom i and the Y_{lm} are spherical harmonics centered at position $\vec{r}_i$. We wish to find LCAO

matrix elements of the Hamiltonian (H) and overlap (S) between two atomic orbitals,

$$H_{inlm,i'n'l'm'} = \langle \phi_{inlm} | \hat{H} | \phi_{i'n'l'm'} \rangle,$$

$$S_{inlm,i'n'l'm'} = \langle \phi_{inlm} | \phi_{i'n'l'm'} \rangle.$$

Above, $\hat{H}$ is the DFT Hamiltonian operator.

In principle, a variety of approaches can be used to compute H and S. We opt for the simplest one. Performing the GW calculations requires a large number of single-particle orthonormal Bloch eigenstates $\psi_{j\vec{k}}$ of $\hat{H}$ along with their energies $E_{j\vec{k}}$, where j is the band index and $\vec{k}$ is the crystal momentum. We employ completeness to write H and S as

$$H_{inlm,i'n'l'm'} = \sum_{j,\vec{k}} \langle \phi_{inlm} | \psi_{j\vec{k}} \rangle E_{j\vec{k}} \langle \psi_{j\vec{k}} | \phi_{i'n'l'm'} \rangle,$$

$$S_{inlm,i'n'l'm'} = \sum_{j,\vec{k}} \langle \phi_{inlm} | \psi_{j\vec{k}} \rangle \langle \psi_{j\vec{k}} | \phi_{i'n'l'm'} \rangle.$$

The precision is limited by the cutoff in the sum over j. Convergence is monitored easily: the diagonals of H and S approach convergence from below, the diagonals of S approach unity, and off-diagonal elements of S between different orbitals on the same atom approach zero. Computing inner products such as $\langle \phi_{inlm} | \psi_{j,\vec{k}} \rangle$ is standard for DFT calculations with nonlocal pseudopotentials;[9] one uses the plane-wave expansion of $\psi_{j,\vec{k}}$ together with the standard textbook expansion of a plane wave into spherical waves about an arbitrary point ($\vec{r}_i$ or $\vec{r}_{i'}$).

[1] S. Nakamura, T. Mukai, and M. Senoh, Appl. Phys. Lett. **64**, 1687 (1994); F. A. Ponce and D. P. Bour, Nature (London) **386**, 351 (1997).

[2] M. Asif Khan, Q. Chen, Michael S. Shur, B. T. Dermott, J. A. Higgins, J. Burm, W. J. Schaff, and L. F. Eastman, Solid-State Electron. **41**, 1555 (1997).

[3] W. Han, S. Fan, Q. Li, and Y. Hu, Science **277**, 1287 (1997); X. F. Duan and C. M. Lieber, J. Am. Chem. Soc. **122**, 188 (2000); J. C. Johnson, H.-J. Choi, K. P. Knutsen, R. D. Schaller, P. Yang, and R. Saykally, Nat. Mater. **1**, 106 (2002); V. Dobrokhotov, D. N. Mcllroy, M. G. Norton, A. Abuzir, W. J. Yeh, I. Stevenson, R. Pouy, J. Bochenek, M. Cartwright, L. Wang, J. Dawson, M. Beaux, and C. Berven, J. Appl. Phys. **99**, 104302 (2006).

[4] Z. Liliental-Weber, Y. Chen, S. Ruvimov, and J. Washburn, Phys. Rev. Lett. **79**, 2835 (1997); J. Goldberger, R. He, Y. Zhang, S. Lee, H. Yan, H.-J. Choi, and P. Yang, Nature (London) **422**, 599 (2003).

[5] S. Iijima, Nature (London) **354**, 56 (1991).

[6] N. G. Chopra, R. J. Luyken, K. Cherrey, V. H. Crespi, M. L Cohen, S. G. Louie, and A. Zettl, Science **269**, 966 (1995); A. Rubio, J. L. Corkill, and M. L. Cohen, Phys. Rev. B **49**, R5081 (1994).

[7] D. Ciuparu, R. F. Klie, Y. Zhu, and L. Pfefferle, J. Phys. Chem. B **108**, 3967 (2004).

[8] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).

[9] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[10] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[11] S. M. Lee, Y. H. Lee, Y. G. Hwang, J. Elsner, D. Porezag, and T. Frauenheim, Phys. Rev. B **60**, 7788 (1999).

[12] S. Hao, G. Zhou, J. Wu, W. Duan, and B.-L. Gu, Phys. Rev. B **69**, 113403 (2004).

[13] E. Durgun, S. Tongay, and S. Ciraci, Phys. Rev. B **72**, 075420 (2005).

[14] M. Zhang, Z.-M. Su, L.-K. Yana, Y.-Q. Qiu, G.-H. Chen, and R.-S. Wang, Chem. Phys. Lett. **408**, 145 (2005).

[15] B. Xu and B. C. Pan, Phys. Rev. B **74**, 245402 (2006).

[16] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[17] *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum, New York 1983), and references therein.

[18] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. **80**, 4510 (1998); L. X. Benedict, E. L. Shirley, and R. B. Bohn, *ibid.* **80**, 4514 (1998); M. Rohlfing and S. G. Louie, *ibid.* **81**, 2312 (1998).

[19] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).

[20] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[21] S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).

[22] S. Limpijumnong, J. E. Northrup, and C. G. Van de Walle, Phys. Rev. B **68**, 075206 (2003), and references therein.

[23] S. Limpijumnong and C. G. Van de Walle, Phys. Rev. B **69**, 035207 (2004), and references therein.

[24] S. Ismail-Beigi, Phys. Rev. B **73**, 233103 (2006).

[25] H. Ajiki and T. Ando, Physica B **201**, 349 (1994).

[26] F. Bechstedt, R. Del Sole, G. Cappellini, and L. Reining, Solid State Commun. **84**, 765 (1992).

[27] V. Fiorentini, M. Methfessel, and M. Scheffler, Phys. Rev. B **47**, 13353 (1993).

[28] A. F. Wright and J. S. Nelson, Phys. Rev. B **50**, 2159 (1994).

[29] J. Neugebauer, T. Zywietz, M. Scheffler, J. E. Northrup, and C. G. Van de Walle, Phys. Rev. Lett. **80**, 3097 (1998).

[30] A. Janotti, D. Segev, and C. G. Van de Walle, Phys. Rev. B **74**, 045202 (2006).

[31] A. Rubio, J. L. Corkill, M. L. Cohen, E. L. Shirley, and S. G. Louie, Phys. Rev. B **48**, 11810 (1993).

[32] *Semiconductors—Basic Data*, 2nd revised ed., edited by O. Madelung (Springer, Berlin, 1996).

[33] M. Rohlfing, P. Kruger, and J. Pollmann, Phys. Rev. Lett. **75**, 3489 (1995).

[34] M. Rohlfing, P. Kruger, and J. Pollmann, Phys. Rev. B **57**, 6485 (1998).

[35] P. Rinke, A. Qtiesh, J. Neugebauer, C. Freysoldt, and M. Scheffler, New J. Phys. **7**, 126 (2005).

[36] M. Palummo, L. Reining, R. W. Godby, C. M. Bertonni, and N.

Bornsen, Europhys. Lett. **26**, 607 (1994).
[37] C. Balasubramanian, S. Belluccib, P. Castrucci, M. De Crescenzi, and S. V. Bhoraskar, Chem. Phys. Lett. **383**, 188 (2004).
[38] M. Zhao, Y. Xia, D. Zhang, and L. Mei, Phys. Rev. B **68**, 235415 (2003).
[39] N. Hamada, S.-I. Sawada, and A. Oshiyama, Phys. Rev. Lett. **68**, 1579 (1992); R. Saito, C. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College, Press, London, 1998).
[40] A. M. Rao, E. Richter, S. Bandow, B. Chase, P. C. Eklund, K. A. Williams, S. Fang, K. R. Subbaswamy, M. Menon, A. Thess, R. E. Smalley, G. Dresselhaus, and M. S. Dresselhaus, Science **275**, 187 (1997); M. S. Dresselhaus, G. Dresselhaus, A. Jorio, A. G. Souza Filho, and R. Saito, Carbon **40**, 2043 (2002).
[41] D. E. Boucher, G. G. DeLeo, and W. B. Fowler, Phys. Rev. B **59**, 10064 (1999).
[42] S. Ismail-Beigi and S. G. Louie, Phys. Rev. Lett. **90**, 076401 (2003).
[43] Z. Qian, S. Hou, J. Zhang, R. Li, Z. Shen, X. Zhao, and Z. Xue, Physica E (Amsterdam) **30**, 81 (2005).
[44] See discussions and references in W. U. Huynh, J. D. Dittmer, and A. P. Alivisatos, Science **295**, 2425 (2002); I. Gur, N. A. Fromer, M. L. Geier, and A. P. Alivisatos, *ibid.* **310**, 462 (2005); B. A. Gregg, MRS Bull. **30**, 20 (2005).
[45] We use the dipole approximation which is valid only for diameters smaller than the wavelength of light. Hence, the ∞ diameter in Table IV really means "quite large."