Fundamental Gaps in Finite Systems from Eigenvalues of a Generalized Kohn-Sham Method

Tamar Stein,[1] Helen Eisenberg,[1] Leeor Kronik,[2,*] and Roi Baer[1,†]
[1]*Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, Hebrew University, Jerusalem 91904, Israel*
[2]*Department of Materials and Interfaces, Weizmann Institute of Science, Rehovoth 76100, Israel*

(Received 25 June 2010; published 20 December 2010)

We present a broadly applicable, physically motivated, first-principles approach to determining the fundamental gap of finite systems from single-electron orbital energies. The approach is based on using a range-separated hybrid functional within the generalized Kohn-Sham approach to density functional theory. Its key element is the choice of a range-separation parameter such that Koopmans' theorem for both neutral and anion is obeyed as closely as possible. We demonstrate the validity, accuracy, and advantages of this approach on first, second and third row atoms, the oligoacene family of molecules, and a set of hydrogen-passivated silicon nanocrystals. This extends the quantitative usage of density functional theory to an area long believed to be outside its reach.

DOI: 10.1103/PhysRevLett.105.266802 PACS numbers: 73.22.−f, 71.15.−m

The Kohn-Sham (KS) [1] formulation of density functional theory (DFT) [2] has become the method of choice for ground state electronic-structure calculations across an unusually wide variety of fields [3]. In this approach, the interacting electron system is mapped into an equivalent noninteracting electron system subject to a common local external potential, in the form: [4]

$$\left(-\frac{\nabla^2}{2} + v_{\mathrm{ext}}(r) + v_H(r) + v_{\mathrm{xc}}(r)\right)\varphi_i(r) = \varepsilon_i \varphi_i(r), \quad (1)$$

where $v_{\mathrm{ext}}(r)$ is the ion-electron potential, $v_H([n]; r)$ the Hartree potential, $v_{\mathrm{xc}}([n]; r)$ the exchange-correlation (xc) potential, and ε_i and $\varphi_i(r)$ the KS eigenvalues and orbitals, respectively. This mapping is exact in principle, but in practice $v_{\mathrm{xc}}([n]; r)$ is only known approximately.

A well-known deficiency of the KS approach is that the gap, E_{KS}, defined as the difference between the lowest unoccupied and highest occupied KS eigenvalues, generally differs from the fundamental gap, E_g, defined as the difference between the ionization potential, I, and the electron affinity, A [5]. This is because $E_g = E_{\mathrm{KS}} + \Delta_{\mathrm{xc}}$ [6], where Δ_{xc} is the derivative discontinuity (DD), i.e., the finite "jump" that the xc potential exhibits as the particle number crosses the integer number of particles in the system, N [7,8]. While in some cases Δ_{xc} can be small, extensive numerical investigations show that it is usually sizable [9]. This would hinder the prediction of the fundamental gap from the KS eigenvalues of the N electron system even if the exact form for $v_{\mathrm{xc}}([n]; r)$ were known.

A remedy for this deficiency may lie within the generalized KS (GKS) scheme [5,10,11], in which the interacting electron system is mapped into an interacting model system that can still be represented by a single Slater determinant. This leads to the GKS equation,

$$\left(\hat{O}_s[\{\varphi_i\}] + v_{\mathrm{ext}}(r) + v_R([n]; r)\right)\varphi_j(r) = \varepsilon_j \varphi_j(r), \quad (2)$$

where $\hat{O}_s[\{\varphi_i\}]$ is a nonlocal, orbital-specific operator and $v_R([n]; r)$ is a "remainder" local potential, which includes all Hartree, exchange, correlation, or kinetic energy components not accounted for by $\hat{O}_s[\{\varphi_i\}]$. As in KS theory, this mapping is exact in principle but approximate in practice. It is generally hoped that because $\hat{O}_s[\{\varphi_i\}]$ inherently exhibits a discontinuity as the particle number crosses an integer due to partial occupation of an additional orbital, a judicious choice of $\hat{O}_s[\{\varphi_i\}]$ would greatly diminish the DD in $v_R(r)$, making the GKS gap, E_{GKS}, considerably closer to E_g than E_{KS} [5,10,12,13]. It is further hoped that the nonlocal character of $\hat{O}_s[\{\varphi_i\}]$ would allow it to mimic more efficiently the role played by the self-energy operator in many-body perturbation theory (MBPT) calculations of quasiparticle excitation energies [14].

For solids, this hope is supported, to an extent, by practical schemes: In the screened-exchange approach [10,15], $\hat{O}_s[\{\varphi_i\}]$ corresponds to the sum of the single-particle kinetic energy operator, the Hartree operator, and a Fock-like operator based on a semiclassically screened potential. In the hybrid functional approach (which can be viewed as a special case of the GKS scheme) [5], a weighted mixture of local exchange and nonlocal Fock exchange is used. In many (but not all) cases, either approach leads to meaningful improvement over the KS scheme in the prediction of fundamental gaps [10,15,16].

For finite-sized objects (e.g., atoms, molecules, nanocrystals), the above schemes are not useful because the asymptotic potential, absent in a solid, plays a crucial role in the energetics of electron addition and removal. In the screened-exchange approach, the long-range (LR) exchange term is absent. In conventional hybrid functionals, only a fixed fraction of it remains. And indeed hybrid functional gaps for finite-sized objects are often much smaller than the fundamental gap and arguably in far better agreement with the first excitation energy (known as the optical gap) [5,17].

Range-separated hybrid (RSH) functionals are a class of functionals, in which the exchange energy term is split into LR and short-range (SR) terms, e.g., via $r^{-1} = r^{-1}\mathrm{erf}(\gamma r) + r^{-1}\mathrm{erfc}(\gamma r)$ [18–20]. The SR exchange is represented by a local potential derived from the local-density or the generalized-gradient approximations.The LR part is treated via an "explicit" or "exact" exchange term. In this way, the full LR Fock exchange is obtained, without sacrificing the description of the SR correlation.

If one assumes a system independent γ, its value can be optimized for, e.g., thermochemistry and other desired properties of the system or the functional [19,21–23]. Specifically, Cohen *et al.* have shown that for such a semiempirical RSH functional that they constructed, called MCY3, [23] GKS gaps for several atoms and small molecules were in better agreement with experimental fundamental gaps, with a mean absolute error of 0.7 eV [12].

Assuming a system-independent γ is only an approximation, as γ is itself a functional of the electron density, n [20,21,24]. System-specific studies showed that good prediction of, e.g., the ionization potential is possible, but γ can vary substantially—from 0.3 for Li_2 to 0.7 for HF or O_2 (in atomic units) [21]. In particular, a constant γ is problematic in the study of the quantum size effect because as the system evolves from the molecular to the solid-state limit, the relative importance of the LR and SR exchange must vary. Thus, even with RSH functionals the long-standing question of whether DFT eigenvalues can at all be used as a broadly applicable tool for predicting fundamental gaps of finite-sized systems remains open.

In this Letter, we show that with the aid of a simple, physically motivated, first-principles γ-determining step, the GKS eigenvalues of RSH functionals can be used successfully for quantitative prediction of fundamental gaps of finite-sized objects in general and of quantum size effects in particular. This paves the road towards using DFT as a practical tool in an area dominated by computationally challenging methods such as coupled cluster, quantum Monte Carlo, or MBPT calculations.

In exact KS theory, the DFT version of Koopmans' theorem establishes that the highest-occupied KS eigenvalue is equal and opposite to the ionization potential [7,25]. The same is true in exact GKS theory [12]. This implies that an optimal choice for obtaining the ionization potential from the highest-occupied GKS eigenvalue is to enforce Koopmans' theorem, i.e., to find γ such that

$$-\varepsilon_{\mathrm{H}}^{\gamma} = I^{\gamma}(N) \equiv E_{\mathrm{gs}}(N-1;\gamma) - E_{\mathrm{gs}}(N;\gamma) \qquad (3)$$

where $\varepsilon_{\mathrm{H}}^{\gamma}$ is the highest-occupied molecular orbital (HOMO) per a specific choice of γ and $I^{\gamma}(N)$ is the energy difference between the ground state energies, E_{gs}, of the N and the $N-1$ electron system, per the same γ [26]. While this has been shown to be useful for determining ionization potentials [27], determining the gap requires that we also know the electron affinity. This means that we employ Koopmans' theorem also for I of the $N+1$ electron system which, barring relaxation effects, is the same as the A of the N electron system. Because there is one parameter but two conditions, we seek the γ that minimizes the overall deviation expressed in the target function [28,29]

$$J(\gamma) = |\varepsilon_{\mathrm{H}}^{\gamma}(N) + I^{\gamma}(N)| + |\varepsilon_{\mathrm{H}}^{\gamma}(N+1) + I^{\gamma}(N+1)|. \quad (4)$$

Importantly, using Eq. (4) to choose the optimal γ, denoted γ^*, does not require any empirical input and contains no adjustable parameters. Furthermore, two figures of merit can serve to evaluate if the result is expected to yield a usefully accurate fundamental gap. First, $J(\gamma^*)$ should be substantially smaller than the desired accuracy. Second, one could define a similar condition involving the lowest unoccupied molecular orbital (LUMO) $\epsilon_{\mathrm{L}}^{\gamma}$ [29]:

$$J'(\gamma') = |\varepsilon_{\mathrm{H}}^{\gamma'}(N) + I^{\gamma'}(N)| + |\varepsilon_{\mathrm{L}}^{\gamma'}(N) + I^{\gamma'}(N+1)|. \quad (5)$$

Unlike Eq. (4), Eq. (5) has no rigorous basis because there is no formal equivalent to Koopmans' theorem involving the LUMO, owing to the DD. However, if indeed the residual DD is small, as hoped for above, then γ^* should be close to γ'^* and $J(\gamma^*)$ should be close to $J'(\gamma'^*)$.

To examine how well the GKS eigenvalues, obtained from such an optimally tuned RSH functional, can predict fundamental gaps in practice, we performed extensive optimally tuned calculations based on the Baer, Neuhauser, and Livshits (BNL) RSH functional, [21,30] as implemented in version 3.2 of Q-CHEM [31].

Consider a series of atoms from the first three rows of the periodic table. A comparison of computed and experimental gaps for these atoms is given in Fig. 1. Several observations are drawn from these results. First, the tuning procedure, based on Eq. (4), produces consistently excellent gap prediction, with a mean deviation of −0.01 eV, a mean absolute deviation of 0.1 eV, and a maximal absolute deviation of 0.3 eV (for P). Second, different atoms require different range parameters (see Fig. 1): For alkali metals γ is relatively high because these atoms bear more resemblance to one-electron systems where Hartree-Fock theory is exact. With the exception of Be, the general trend along a given row in the periodic table is that higher values of the gap require higher values of γ. First row atoms require values of γ between 0.5 and 0.65, whereas second and third row atoms require lower γ values, between 0.37 and 0.47. Third, for a given atom the gap is a very sensitive function of the range parameter. For example, for F and O the gap changes by as much as 8 eV when γ changes from 0.3 to 1. Fourth, the inset to Fig. 1 compares the gaps from our BNL* functional to those from the semiempirical, fixed-γ MCY3 [12] functional. Although the latter presents a huge improvement over previous attempts, the optimal tuning leads to higher accuracy and consistency. Taken together, these four observations underline the importance and nontriviality of our first-principles tuning procedure.

To further test our approach, we studied the evolution of the fundamental gap in two systems known to exhibit a significant quantum size effect: the oligoacene molecules

$C_{2+4n}H_{4+2n}(n = 1 - 6)$ [32–35] and the hydrogen-passivated spherical Si nanocrystals [36,37]. For both systems, ground state properties are well described by conventional functionals, but the fundamental gap is seriously underestimated. For the oligoacenes, even gap values obtained from total energy differences using B3LYP were found to be smaller by an average of 0.5 eV with respect to experiment, with a maximal deviation for hexacene that exceeds 0.7 eV [35]. Eigenvalue-based BNL* gaps, compared with experiment, are given in Fig. 2 (left). The mean absolute deviation is 0.3 eV, with a maximal deviation of 0.5 eV. The tuned γ values decrease consistently with system size, from 0.3 to 0.19, because electron delocalization increases with system size, rendering the necessary weight of exact exchange smaller. Again an optimally tuned γ outdoes a fixed one.

In Fig. 2 (right) the BNL* HOMO and LUMO energies for nanocrystaline Si systems are compared with experimental [38] ionization potentials and GW-computed ionization potentials and electron affinities [36]. The three small-diameter systems do not bind an electron. The mean deviation of the BNL* eigenvalue gaps from GW gaps is 0.1 eV and the mean absolute deviation is 0.2 eV with a maximal deviation of 0.45 eV. Moreover, I and A values are very well-reproduced separately, at a fraction of the computational cost of a GW calculation. Here too γ^* decreases steadily as the size increases, and for the same physical reason.

We note that for the molecular and nanocrystaline systems, the remaining difference between our results and the reference values may also reflect limitations of the reference. For molecules, vertical electron affinities are hard to come by owing to structural relaxation effects. Furthermore, GW calculations of molecules may exhibit some deviation from experiment, especially for the electron affinity [39]. For example, the experimental ionization potential of Si_5H_{12} is closer to the BNL* value than to the GW one (deviations of -0.01 and $+0.15$ eV, respectively).

For both oligoacenes and Si nanocrystals, the tuning procedure results in $J(\gamma^*)$ and $J'(\gamma'^*)$ that are close to zero (~ 0.02 eV on average), indicating a negligible DD. Likely, this excellent performance arises from the fact that addition of an electron to the system does not change its chemical nature, ergo γ^* for neutral and anion is similar. For atoms, addition of a single electron does change their nature appreciably. Coupled with the larger gaps in general, we expect larger deviations in this case. Indeed, optimal tuning based only on A leads to an average root mean square deviation from experiment of 0.15 eV, but tuning according to Eq. (4) leads to a deviation of 0.35 eV. Furthermore, tuning based on Eq. (4) and (5) yields an average $J(\gamma^*)$ or $J'(\gamma^{*\prime})$ of 0.65 or 0.4 eV, respectively. The difference indicates a non-negligible, but still small, DD. But $J(\gamma^*)$ in this case is too strict a criterion. The remaining error in I and A is of the same sign, resulting in the excellent gaps of Fig. 1. Thus, even in this worst-case scenario, the method still yields quantitatively useful fundamental gaps. Partly, this is because we mimic the "straightline between integers" dependence of the total energy on the electron number, expected for the exact functional [7,13]: By construction, Eq. (3) demands that the slope approaching N from below (given by the HOMO eigenvalue due to Janak's theorem) remains the same as the average slope between $N - 1$ and N (given by the energy difference), with an obvious generalization to $N + 1$ in Eq. (4). This imposes severe constraints on the deviation of the eigenvalues from their correct value, even in the presence of nonlinearity.

Finally, we comment on the importance of our approach. For finite systems, in principle one can always compute the

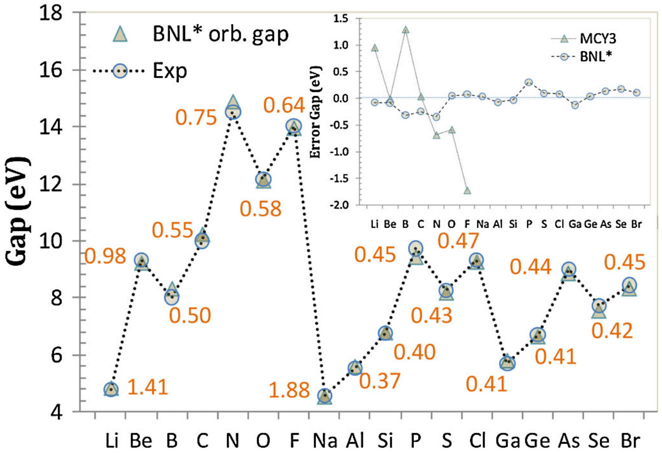

FIG. 1 (color online). BNL* HOMO-LUMO gaps (computed using the aug-cc-pVTZ basis set), compared with experimental fundamental gaps [43]. The value of γ, determined by minimizing J, is indicated near each point. Inset: the deviation from experiment of GKS HOMO-LUMO gaps based on BNL* (this work) and MCY3 [12].

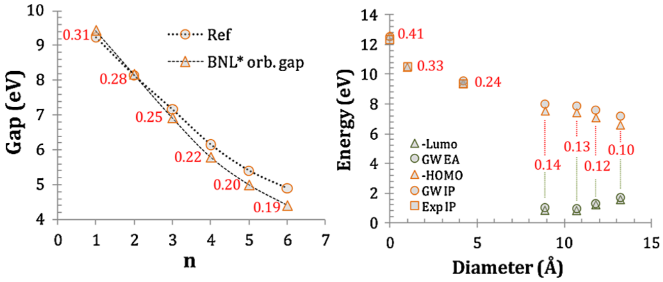

FIG. 2 (color online). Left: BNL* HOMO-LUMO gaps, compared with gaps from experimental (vertical) ionization potentials (IP) [32] and best estimates of vertical electron affinities [33], for the oligoacenes $C_{2+4n}H_{4+2n}$, $n = 1$ (benzene) to 6 (hexacene). The value of γ, determined by minimizing J, is indicated near each point. Right: BNL* HOMO and LUMO energies compared to GW [36] and experimental [38] IP and EA of hydrogen terminated Si nanocrystals, as a function of diameter. The values of the tuned range parameter are shown in red. In both systems the cc-pVTZ basis set was used. Geometries were obtained from a B3LYP calculation for the oligoacenes and from Ref. [44] for the Si nanocrystals.

fundamental gap from the total energy of the N, $N-1$, and $N+1$ electron systems. Depending on the system and the approximate functional, such calculations may be of insufficient accuracy, e.g., for semilocal functionals applied to large systems [40]. But an accurate *eigenvalue-based* DFT gap is also essential in other contexts. Three typical examples are: (1) The (image-charge renormalized) incorrect DFT gap results in gross overestimations of the calculated conductance, e.g., of single-molecule oligophenyldiamine−gold junctions [41]. (2) A more accurate DFT gap provides a much better starting point for DFT-based MBPT calculations that yield the overall quasiparticle spectrum [42]. (3) Accurate DFT-level prediction of the fundamental gap is essential for accurate prediction of optically-induced charge transfer excitations [28,29].

In conclusion, we presented a physically motivated first-principles approach to determining the fundamental gap of finite systems from single-electron orbital energies. It is based on using a RSH functional within the GKS scheme, choosing the range-separation parameter such that Koopmans' theorem for both neutral and anion is obeyed, as closely as possible. We demonstrated the validity and accuracy of this approach on first, second, and third row atoms, oligoacenes, and hydrogen-passivated silicon nanocrystals. This extends the quantitative usage of DFT to an area long believed to be outside its reach.

Work was supported by the Israel Science Foundation.

---

*leeor.kronik@weizmann.ac.il
†roi.baer@huji.ac.il

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[2] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[3] D. S. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction* (Wiley, Hoboken, N.J., 2009).
[4] For notational simplicity, the equation is given in a spin-unpolarized form. All calculations presented below were spin-polarized whenever necessary.
[5] S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).
[6] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983); L. J. Sham and M. Schluter, Phys. Rev. Lett. **51**, 1888 (1983).
[7] J. P. Perdew *et al.*, Phys. Rev. Lett. **49**, 1691 (1982).
[8] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).
[9] O. Gunnarsson and K. Schonhammer, Phys. Rev. Lett. **56**, 1968 (1986); R. W. Godby *et al.*, Phys. Rev. Lett. **56**, 2415 (1986); G. K.-L. Chan, J. Chem. Phys. **110**, 4710 (1999); M. J. Allen and D. J. Tozer, Mol. Phys. **100**, 433 (2002); M. Gruning *et al.*, J. Chem. Phys. **124**, 154108 (2006).
[10] A. Seidl *et al.*, Phys. Rev. B **53**, 3764 (1996).
[11] R. Baer *et al.*, Annu. Rev. Phys. Chem. **61**, 85 (2010).
[12] A. J. Cohen *et al.*, Phys. Rev. B **77**, 115123 (2008).
[13] P. Mori-Sanchez *et al.*, Phys. Rev. Lett. **100**, 146401 (2008).
[14] P. Sanchez-Friera and R. W. Godby, Phys. Rev. Lett. **85**, 5611 (2000); M. Gruning *et al.*, Phys. Rev. B **74**, 161103 (2006).
[15] D. M. Bylander and L. Kleinman, Phys. Rev. B **41**, 7868 (1990).
[16] C. B. Geller *et al.*, Appl. Phys. Lett. **79**, 368 (2001); C. Stampfl *et al.*, Phys. Rev. B **63**, 155106 (2001); J. Robertson *et al.*, Phys. Status Solidi B **243**, 2054 (2006); B. Lee and L. W. Wang, Phys. Rev. B **73**, 153309 (2006); J. Muscat *et al.*, Chem. Phys. Lett. **342**, 397 (2001); J. Heyd *et al.*, J. Chem. Phys. **123**, 174101 (2005); J. Paier *et al.*, J. Chem. Phys. **124**, 154709 (2006); X. Wu *et al.*, Phys. Rev. B **80**, 115201 (2009).
[17] U. Salzner *et al.*, J. Comput. Chem. **18**, 1943 (1997); V. Barone *et al.*, Nano Lett. **5**, 1621 (2005); N. Dori *et al.*, Phys. Rev. B **73**, 195208 (2006); C. G. Zhan *et al.*, J. Phys. Chem. A **107**, 4184 (2003).
[18] A. Savin, in *Recent Advances in Density Functional Methods Part I*, edited by D. P. Chong (World Scientific, Singapore, 1995), p. 129; T. Leininger *et al.*, Chem. Phys. Lett. **275**, 151 (1997).
[19] H. Iikura *et al.*, J. Chem. Phys. **115**, 3540 (2001); T. Yanai *et al.*, Chem. Phys. Lett. **393**, 51 (2004).
[20] R. Baer and D. Neuhauser, Phys. Rev. Lett. **94**, 043002 (2005).
[21] E. Livshits and R. Baer, Phys. Chem. Chem. Phys. **9**, 2932 (2007).
[22] M. J. G. Peach *et al.*, Phys. Chem. Chem. Phys. **8**, 558 (2006); J. D. Chai and M. Head-Gordon, J. Chem. Phys. **128**, 084106 (2008); O. A. Vydrov and G. E. Scuseria, J. Chem. Phys. **125**, 234109 (2006); I. C. Gerber and J. G. Angyan, Chem. Phys. Lett. **415**, 100 (2005).
[23] A. J. Cohen *et al.*, J. Chem. Phys. **126**, 191109 (2007).
[24] R. Baer *et al.*, Chem. Phys. **329**, 266 (2006).
[25] C.-O. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985).
[26] A different strategy, based on per-orbital corrections following a generalized Koopmans' approach, is discussed in I. Dabo *et al.*, Phys. Rev. B **82**, 115121 (2010); S. Lany and A. Zunger, Phys. Rev. B **80**, 085202 (2009).
[27] U. Salzner and R. Baer, J. Chem. Phys. **131**, 231101 (2009).
[28] T. Stein *et al.*, J. Am. Chem. Soc. **131**, 2818 (2009).
[29] T. Stein *et al.*, J. Chem. Phys. **131**, 244119 (2009).
[30] In the original, fixed-γ BNL, part of the short-range exchange energy was subtracted from the correlation energy. This subtraction was not employed here, because it departs from the spirit of the RSH functional in principle and was found to be detrimental in practice.
[31] Y. Shao *et al.*, Phys. Chem. Chem. Phys. **8**, 3172 (2006).
[32] W. Schmidt, J. Chem. Phys. **66**, 828 (1977).
[33] B. Hajgato *et al.*, J. Chem. Phys. **129**, 084308 (2008).
[34] M. Bendikov *et al.*, J. Am. Chem. Soc. **126**, 7416 (2004).
[35] E. S. Kadantsev *et al.*, J. Chem. Phys. **124**, 134901 (2006).
[36] M. L. Tiago and J. R. Chelikowsky, Phys. Rev. B **73**, 205334 (2006).
[37] S. Ogut *et al.*, Phys. Rev. Lett. **79**, 1770 (1997).
[38] U. Itoh *et al.*, J. Chem. Phys. **85**, 4867 (1986).
[39] N. Sai *et al.*, Phys. Rev. B **77**, 161306(R)(2008).
[40] R. W. Godby and I. D. White, Phys. Rev. Lett. **80**, 3161 (1998).
[41] S. Y. Quek *et al.*, Nano Lett. **9**, 3949 (2009).
[42] P. Rinke *et al.*, New J. Phys. **7**, 126 (2005).
[43] NIST, Reference Database No. 69 (2008).
[44] I. Vasiliev *et al.*, Phys. Rev. Lett. **86**, 1813 (2001).