Reliability of Hybrid Functionals in Predicting Band Gaps

Manish Jain,[1,2] James R. Chelikowsky,[3] and Steven G. Louie[1,2]

[1]*Department of Physics, University of California, Berkeley, California 94720, USA*
[2]*Materials Sciences Division, Lawrence Berkeley, National Laboratory, Berkeley, California 94720, USA*
[3]*Center for Computational Materials, Institute for Computational Engineering and Sciences, Departments of Physics and Chemical Engineering, University of Texas, Austin, Texas 78712, USA*

(Received 28 June 2011; published 18 November 2011)

We show that orbital energies from existing hybrid functionals do not give reliable band gaps. Even if a functional yields a good bulk gap, it in general does not provide accurate gaps in different structural configurations, e.g., surfaces or nanostructures. For example, none of the popular hybrid functionals adequately describe the surface-state gap of the Si(111)-(2 × 1) surface. For graphene nanoribbons, some hybrid functionals give good optical gaps (neglecting strong excitonic effects), but not quasiparticle gaps. In both cases, there are strong variations from different hybrid functionals.

DOI: 10.1103/PhysRevLett.107.216806 PACS numbers: 73.22.−f, 71.15.Mb, 72.80.Rj

Density functional theory (DFT) within the Kohn-Sham formalism [1] has been the method of choice for theoretical predictions of structural and ground-state electronic properties of condensed matter systems. However, for most systems, assigning the Kohn-Sham orbital energy difference as the quasiparticle band gap E_g leads to a dramatic underestimation of the band gap [2]. While in principle E_g is accessible within DFT, the Kohn-Sham gap, however, is not equal to E_g even for the exact functional [3]. This problem is solved by appropriately calculating the quasiparticle energies, e.g., within the *ab initio GW* method [2]. Recently, there has been a new class of DFT exchange-correlation functionals, the hybrid functionals, constructed such that the orbital energies have been reported to give good band gaps in solids [4,5], in particular, semiconductors.

Hybrid functionals go beyond the usual Kohn-Sham formalism and fall within the generalized Kohn-Sham realization of DFT [6]. These functionals mix a fraction b_{HF} of nonlocal single-determinant exchange with conventional DFT exchange-correlation functionals. Popular hybrid functionals—B3LYP [7], PBE0 [8], and HSE [9]—have been constructed to give good structural, thermodynamic, and bonding properties of solids [4,5].

In this Letter, we demonstrate that the orbital energies from existing, popular hybrid functionals are not reliable in predicting the band gaps of materials, either the optical or the quasiparticle gap. Even if a specific functional may give a good value for the bulk band gaps, the same functional in general does not yield accurate gap values for the same material in different configurations such as at its surfaces or in nanostructures.

The quasiparticle gap of an insulator can be conceptually defined in terms of the total energy of the system containing N, $N+1$ and $N-1$ electrons as $E_g = E(N+1) + E(N-1) - 2E(N)$. The optical gap E_{opt} is a different physical quantity and related to E_g as $E_g - E_{\mathrm{exciton}}$, where E_{exciton} is the exciton binding energy. E_{opt} is also, in principle, accessible within time-dependent DFT but the Kohn-Sham eigenvalues can only be used as a rough approximation to the E_{opt} [10].

In hybrid functional calculations [6,11], one solves a self-consistent field equation within the generalized Kohn-Sham formalism:

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{ion}}(\vec{r}) + v_H([n],\vec{r}) + v_c([n], b_{\mathrm{HF}}, \vec{r}) + b_{\mathrm{HF}}\hat{v}_{\mathrm{HF}}[n] + (1 - b_{\mathrm{HF}})v_x([n],\vec{r})\right]\phi_i(\vec{r}) = \epsilon_i\phi_i(\vec{r}), \quad (1)$$

where v_{ion} is the ionic potential, v_H the Hartree potential, $\hat{v}_{\mathrm{HF}}$ the Fock operator, v_x and v_c the local exchange and correlation potentials within standard Kohn-Sham DFT, respectively. Equation (1) goes beyond the regular Kohn-Sham formalism as the total potential is no longer local. In cases of range-separated hybrid functionals (such as HSE), one replaces the $\hat{v}_{\mathrm{HF}}$ with a short [9] or long [12] range part of the Fock operator. It is noted that different values of b_{HF}, if done properly, would give the same ground-state total energy and density, but give different orbital eigenvalues ϵ_i. As in the original Kohn-Sham formulation, ϵ_i are just Lagrange multipliers in minimizing the total energy and they are *not* quasiparticle excitation energies. The fact that the ϵ_i's can be changed with a different and arbitrary choice of the parameter b_{HF} nicely illustrates this point.

The physical quasiparticle gap is given by the sum of energies needed to create a quasielectron and a quasihole independently in the system. Such quasiparticle energies E_{QP} are given by Dyson's equation:

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{ion}}(\vec{r}) + v_H(\vec{r})\right]\phi_i(\vec{r}) + \int \Sigma(\vec{r},\vec{r}', E_{\mathrm{QP}})\phi_i(\vec{r}')d\vec{r}' = E_{\mathrm{QP}}\phi_i(\vec{r}), \quad (2)$$

where $\Sigma(\vec{r}, \vec{r}', E_{QP})$ is the electron self-energy operator [2]. The Kohn-Sham potential has been shown to be the best local approximation to the self-energy operator [13]. At best, the hybrid functional of Eq. (1) which mixes the Kohn-Sham potential and single-determinant exchange, is a rough approximation to the self-energy Σ which is a nonlocal frequency-dependent operator that is sensitive to many-electron and hence environment effects. From this point of view, the hybrid functional eigenvalue gap from Eq. (1) should be compared to the quasiparticle gap since there is no interaction between the excited electron and hole. The hybrid functional gap has also been related to the quasiparticle gap by comparing the parameter b_{HF} to the effect of an average dielectric screening within the *GW* approximation to Σ [4]. But since b_{HF} is fixed for a given hybrid functional, such screening is fixed and cannot respond to a change in the environment. Band gaps calculated with hybrid functionals have also been associated with E_{opt} [14] without much justification since as mentioned already electron-hole interaction is missing in Eq. (1). To put these issues in concrete terms, we carry out calculations on the Si(111)-(2×1) surface and on armchair graphene nanoribbons using the popular hybrid functionals in the literature.

While there have been a number of studies using hybrid functionals for bulk crystals [15], there are only a few calculations on predicting E_g or E_{opt} in one-dimensional [16,17] and two-dimensional [18] systems. Low-dimensional systems often have optical excitations with large exciton binding energies making the quasiparticle gaps and optical gaps quite different [19–22]. The Si(111)-(2×1) surface is a system with multiple gaps, bulk-state gap and surface-state gap, that have been measured by photoemission and optical experiments. It provides a good test for hybrid functionals because not only are the surface-state wave functions qualitatively different from the bulk wave functions, but also screening at the surface is very different from the bulk.

The π-bonded chain reconstruction of the Si(111)-(2×1) surface [23] has been studied extensively, both theoretically [19,20] and experimentally [24,25]. On an ideal Si (111) surface, the surface Si atoms are bonded to only three atoms rather than the usual four. Owing to the 2×1 reconstruction, the dangling p_z orbitals form π-bonded chains along the $[01\bar{1}]$ direction. These dangling orbitals give rise to two surface-state bands inside the quasiparticle bulk band gap, one occupied and one unoccupied [24]. Because of the quasi-1D nature of surface states localized on the π-bonded chains, the photo-excited surface-state electron and holes form discrete excitonic states with a large binding energy [19,20,25] as compared to the exciton in the bulk. Thus, unlike bulk Si, the quasiparticle gap of the surface states differs significantly from the optical gap.

We perform DFT calculations using *ab initio* pseudopotentials and plane wave formalism as implemented in PARATEC [26]. The Si ionic pseudopotential was generated using the Troullier-Martins scheme [27] in the PBE approximation [28] to the exchange-correlation functional. The use of pseudopotentials generated within PBE for hybrid functional calculations can in principle lead to some errors. We explicitly check the bulk Si band gap calculated with PBE pseudopotential and a Hartree-Fock pseudopotential and found the differences in the band gaps for all the functionals studied here to be less than 50 meV [29]. Thus, any possible error from the pseudopotentials is negligible. We use a 12-layer slab to simulate the (111) surface in a 24 atom centrosymmetric supercell with $\sim$20 Å of vacuum in between the slabs. The Brillouin zone was sampled with a 4×8 $\vec{k}$-point grid along the surface directions. We used a plane wave cutoff of 35 Ry for the wave functions. Because of the unscreened Coulomb interaction in the Fock operator being long ranged, we employed a slab truncation scheme for the Coulomb interaction [30] for all the hybrid functional calculations, except HSE where the Coulomb interaction is now short ranged due to the error-function complement in the expression [9]. The structure was relaxed within the PBE approximation (forces ≤ 0.02 eV/Å) and was kept fixed for the hybrid functional calculations. Because the forces were small, relaxing the structure within each hybrid functional is not expected to change the electronic structure. Explicitly, within HSE, the relaxation of the atoms moved them by <0.1 Å.

Figure 1 shows our calculated band structure for the Si(111)-(2×1) surface using semilocal (PBE) and hybrid functionals. Also shown is the projected bulk Si band

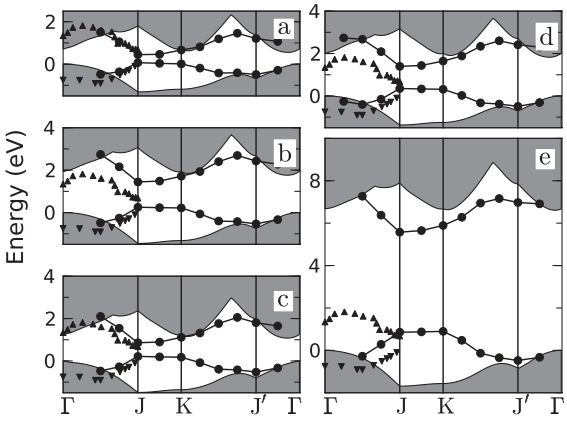

FIG. 1. Surface bands (black dots) of the Si(111)-(2×1) surface calculated with (a) PBE, (b) PBE0, (c) HSE, (d) B3LYP, and (e) Hartree-Fock approximations. Up and down filled triangles are experimental data (from direct and inverse photoemission) of Ref. [24]. The shaded gray is the projected bulk band structure. Energy (in eV) of the top of the bulk valence band is set to zero in each panel.

structure. To calculate the projected bulk band structure with a refined grid, we interpolated the bands using Wannier functions constructed from a bulk calculation with a $8 \times 8 \times 8$ $\vec{k}$-point mesh. We aligned the bulk and slab band structures by setting the top of the bulk valence band at the Γ point to be at the same energy in both calculations. Based on quantum confinement energy estimated for a particle in a one-dimensional well, we estimate the uncertainty in alignment in the two band structures to be about 0.2 eV. It is important to point out that, the uncertainty in alignment does not affect any of the surface-state band gaps in Fig. 1 since there is no slab confinement effect on the surface states. As is evident from Fig. 1, different hybrid functionals give quite different band structures. Besides the PBE band structure, which qualitatively resembles the direct and inverse photoemission experiment reasonably well, the HSE band structure also looks similar to experiment. Experimentally, the top of the occupied surface band at J is known to be below the top of the bulk valence band at Γ by 0.1 eV [24]. However, in all the hybrid functional and Hartree-Fock calculations, this was found not to be the case, even after accounting for a possible 0.2 eV shift in the alignment. This effect is clearly seen in the Hartree-Fock results where the surface band is 0.7 eV (after shifting by 0.2 eV) above the top of the valence band at Γ. All hybrid functionals show similar incorrect behavior, with the surface band at J nearly at the same energy as that of the top of the bulk valence band at Γ for HSE and PBE0 and 0.14 eV above for B3LYP. The GW results (not shown in Fig. 1) from Ref. [20], on the other hand, are in excellent agreement with experiment.

Table I gives the relevant band gaps, calculated as a difference between the highest occupied and lowest unoccupied bulk or surface states, for the different functionals. Also presented are the GW [19,20,31] and the GW plus Bethe-Salpeter equation (GW-BSE) [20,31] and experimental results. As can be seen from Table I, all the hybrid functionals significantly overestimate the optical surface-state gap. Moreover, different hybrid functionals give gaps that differ by up to 0.5 eV. Our calculated values of the bulk gap for PBE0, B3LYP, HSE, and Hartree-Fock functionals agree well with the previous calculations for Si [34]. It is also interesting to note that the values of the bulk gap are not simply related to the fraction of exact exchange in the hybrid functional. This is not surprising given that the band gaps depend on the detailed character of the wave functions and screening environment. The actual self-energy operator is complex and cannot be determined by just one parameter—the fraction of single-determinant exchange. Range separation in HSE may be viewed as a different "effective" screening environment and results in a different gap. Hartree-Fock substantially overestimates both the bulk and the surface-state gap. This is primarily because of the lack of screening in Hartree-Fock. While PBE0 and B3LYP also overestimate all the gaps, HSE gives a bulk gap that is close to the experimental value. However, when compared to the experimental optical surface-state gap, the HSE surface-state gap is also too large. It should be pointed out that, in this case, the HSE surface-state gap is similar to the experimental quasiparticle surface-state gap. This does not agree with the claim in the literature that the HSE gap is expected to agree with the optical gap and it is a good estimate of the quasiparticle gap only when the exciton binding energy is small [14]. In the present case, the bulk Si exciton binding energy is $\sim$15 meV while the surface-state exciton binding energy is 0.28 eV. On the other hand, gaps calculated within the GW and the GW-BSE approaches [20], for the quasiparticle gap and the optical gap, respectively, match the corresponding experimental values well.

TABLE I. Bulk and surface-state gaps of Si(111)-(2×1) calculated using various functionals and with the GW method. The experiment (PE) refers to direct and inverse photoemission values while (opt) refers to optical gap. All values are given in eV.

	Bulk gap	Surface-state gap
Generalized Kohn-Sham gap		
PBE	0.55	0.38
PBE0	1.71	1.19
HSE	1.14	0.65
B3LYP	1.83	1.04
Hartree-Fock	6.63	4.67
Quasiparticle gap		
GW [20,31]	1.23	0.69
Experiment (PE) [24,32]	1.17	0.75
Optical gap		
GW-BSE [20,31]	1.23	0.43
Experiment (opt) [25,33]	1.16	0.47

For the 1D system, we studied armchair graphene nanoribbons (AGNRs) of different widths. The AGNRs studied have armchair-shaped edges with the dangling σ bonds at the edges passivated by hydrogen atoms. Following the conventional notation, a AGNR-N is specified by the number of dimer lines N along the ribbon forming the width. In this study, we employ hybrid functionals to calculate the electronic properties of three AGNR-Ns ($N = 5$, 6, and 7), which cover the distinct three families [35] ($N = 3p - 1$, $3p$, and $3p + 1$, where p is an integer) of AGNRs. These nanoribbons have been previously studied theoretically [17,22,35].

For the calculations of the AGNRs, we used the Troullier-Martins scheme [27] to generate the C and H pseudopotentials within PBE functional. Similar to the Si case, we checked that the bulk diamond band gap calculated with the PBE pseudopotential was within 50 meV of the band gap calculated with a Hartree-Fock pseudopotential for all the hybrid functionals [29]. The wave functions were expanded in plane waves with a cutoff

energy of 65 Ry. The Brillouin zone was sampled with $1 \times 1 \times 16$ $\vec{k}$ points. The structures of the AGNRs were fully relaxed within PBE. As in the Si surface case, the structures were not relaxed further using hybrid functionals. The details of the *GW* and *GW*-BSE calculations, performed with the BerkeleyGW package [36], were as per Ref. [22]. To avoid nanoribbon-nanoribbon interaction, we employed a wire truncation scheme [30] of the Coulomb potential in the Fock operator in all cases, except for the HSE functional as for the reason explained previously.

Table II shows the calculated band gaps using various methods. In the absence of experiment, the *GW* gaps are expected to be close to the actual quasiparticle gaps of the AGNRs and the *GW*-BSE gaps are close to the actual optical gaps. Our calculations are in good agreement with previous calculations [17,22]. The exciton binding energies were found to be large as seen in the difference between the quasiparticle and optical gap values. All the functionals show the expected family behavior of the gap in the AGNRs [35]. However, in these systems as in the case of Si(111)-(2×1), gaps calculated with different hybrid functionals differ by up to 0.5 eV. Hartree-Fock overestimates all the gaps, while PBE underestimates the optical gap. PBE0 and B3LYP hybrid functionals underestimate the quasiparticle gap and overestimate the optical gap. HSE gaps are in good agreement with optical gaps from the *GW*-BSE calculations, but not in agreement with the quasiparticle gap. This is the opposite behavior from the Si(111)-(2×1) case where the HSE band gaps were in agreement with the quasiparticle gaps. It is also worth pointing out that the *GW*-BSE calculation gives discrete peaks in the absorption spectrum that are related to bound excitons whose excitation energy corresponds to the optical gap. While the band gap from the HSE functional is close to the optical band gap from *GW*-BSE, the HSE absorption onset corresponds to a continuum onset rather than discrete bound states. This is a qualitatively different behavior.

TABLE II. Band gap of armchair graphene nanoribbons calculated using various functionals and the *GW* and *GW*-BSE approaches. All values are in eV.

	AGNR-5	AGNR-6	AGNR-7
Generalized Kohn-Sham gap			
PBE	0.39	1.12	1.56
PBE0	1.02	2.08	2.59
HSE	0.59	1.50	2.02
B3LYP	0.89	1.88	2.39
Hartree-Fock	3.98	5.12	5.42
Quasiparticle gap			
GW	1.32	2.86	3.44
Optical gap			
GW-BSE	0.48	1.51	1.88

These results show that gaps calculated using existing hybrid functionals are not reliable and have large variability depending on the functional and system studied. It is, moreover, not justified which gaps the results from the popular hybrid functionals should correspond to—optical or quasiparticle. Importantly, the reliability of band gaps obtained with hybrid functionals is not dependent just on the materials, but also on the environment. Owing to this, hybrid functionals cannot be assumed to give good optical or quasiparticle gaps. Further, one cannot even treat b_{HF} as an empirical parameter for cases with multiple gaps such as the Si surface.

In conclusion, we have studied materials in different configurations with hybrid functionals. For the Si(111)-(2×1) surface, none of the popular hybrid functionals give the correct surface-state optical gap, while the HSE functional gives a good quasiparticle gap. For AGNR's, none of the popular hybrid functionals give the correct quasiparticle gap. PBE0 and B3LYP overestimate the optical gap as compared to *GW*-BSE results, while HSE gives good agreement with the *GW*-BSE optical gaps although electron-hole interaction is not explicitly included. Overall, there is a large variability in the results from different hybrid functionals, and in their current forms hybrid functionals may not be relied upon to predict band gaps.

M. J. would like to thank Brad Malone and Noa Marom for fruitful discussions. This work was supported by National Science Foundation Grant No. DMR10-1006184, the U.S. Department of Energy under Contracts No. DE-AC02-05CH11231 and No. DE-SC0001878. Computational resources have been provided by NSF through TeraGrid resources at NICS. M. J. was supported by the DOE. Part of the simulations were carried out with electronic structure and QP codes developed under NSF support.

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[2] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[3] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983); L. J. Sham and M. Schlüter, *ibid.* **51**, 1888 (1983).
[4] J. Paier *et al.*, J. Chem. Phys. **124**, 154709 (2006).
[5] M. Marsman *et al.*, J. Phys. Condens. Matter **20**, 064201 (2008).
[6] A. Seidl *et al.*, Phys. Rev. B **53**, 3764 (1996).
[7] P. J. Stephens *et al.*, J. Phys. Chem. **98**, 11623 (1994).
[8] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996); C. Adamo and V. Barone, *ibid.* **110**, 6158 (1999).
[9] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003); **124**, 219906 (2006).
[10] A. Görling, Phys. Rev. A **54**, 3912 (1996).
[11] T. Körzdörfer and S. Kümmel, Phys. Rev. B **82**, 155206 (2010).

[12] T. Stein, H. Eisenberg, L. Kronik, and R. Baer, Phys. Rev. Lett. **105**, 266802 (2010).
[13] M. E. Casida, Phys. Rev. A **51**, 2005 (1995).
[14] T. M. Henderson, J. Paier, and G. E. Scuseria, Phys. Status Solidi (b)**248**, 767 (2011).
[15] S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).
[16] V. Barone, J. E. Peralta, and G. E. Scuseria, Nano Lett. **5**, 1830 (2005); V. Barone *et al.*, *ibid.* **5**, 1621 (2005).
[17] V. Barone, O. Hod, and G. E. Scuseria, Nano Lett. **6**, 2748 (2006).
[18] A. Stroppa and G. Kresse, New J. Phys. **10**, 063020 (2008).
[19] J. E. Northrup, M. S. Hybertsen, and S. G. Louie, Phys. Rev. Lett. **66**, 500 (1991); L. Reining and R. DelSole, *ibid.* **67**, 3816 (1991).
[20] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **83**, 856 (1999); Phys. Status Solidi A **175**, 17 (1999).
[21] C. D. Spataru *et al.*, Phys. Rev. Lett. **92**, 077402 (2004).
[22] L. Yang, M. L. Cohen, and S. G. Louie, Nano Lett. **7**, 3112 (2007); L. Yang *et al.*, Phys. Rev. Lett. **99**, 186801 (2007); D. Prezzi *et al.*, Phys. Rev. B **77**, 041404 (2008).
[23] K. C. Pandey, Phys. Rev. Lett. **47**, 1913 (1981).
[24] R. I. G. Uhrberg *et al.*, Phys. Rev. Lett. **48**, 1032 (1982); P. Perfetti, J. M. Nicholls, and B. Reihl, Phys. Rev. B **36**, 6160 (1987).
[25] F. Ciccacci *et al.*, Phys. Rev. Lett. **56**, 2411 (1986).
[26] http://www.nersc.gov/projects/paratec.
[27] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[28] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[29] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.107.216806 for details.
[30] S. Ismail-Beigi, Phys. Rev. B **73**, 233103 (2006).
[31] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **52**, 1905 (1995).
[32] C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, Inc., New York, 1996).
[33] G. G. Macfarlane *et al.*, Phys. Rev. **111**, 1245 (1958).
[34] F. Fuchs *et al.*, Phys. Rev. B **76**, 115109 (2007).
[35] Y.-W. Son, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **97**, 216803 (2006).
[36] http://www.berkeleygw.org.