Nonlocal orbital-free kinetic energy density functional for semiconductors

Chen Huang

Department of Physics, Princeton University, Princeton, New Jersey 08544, USA

Emily A. Carter

Department of Mechanical and Aerospace Engineering and Program in Applied and Computational Mathematics,
Princeton University, Princeton, New Jersey 08544-5263, USA

We propose a nonlocal kinetic energy density functional (KEDF) for semiconductors based on the expected asymptotic behavior of its susceptibility function. The KEDF's kernel depends on both the electron density and the reduced density gradient, with an internal parameter formally related to the material's static dielectric constant. We determine the accuracy of the KEDF within orbital-free density functional theory (DFT) by applying it to a variety of common semiconductors. With only two adjustable parameters, the KEDF reproduces quite well the exact noninteracting KEDF (i.e., Kohn-Sham DFT) predictions of bulk moduli, equilibrium volumes, and equilibrium energies. The two parameters in our KEDF are sensitive primarily to changes in the local crystal structure (such as atomic coordination number) and exhibit good transferability between different tetrahedrally-bonded phases. This local crystal structure dependence is rationalized by considering Thomas-Fermi dielectric screening theory.

I. INTRODUCTION

Computer simulation is an increasingly important tool for predicting material properties, providing the underlying rationale for materials behavior, as well as helping to interpret measurements.[1] The most general and accurate methods available for simulating materials are those based on quantum mechanics. At present, the most common quantum mechanics method for studying materials is Kohn-Sham density functional theory (KS-DFT),[2] which is based on the Hohenberg and Kohn theorems.[3] KS-DFT implementations generally exhibit a good balance between accuracy and computational cost. However for large scale material simulations, e.g., thousands of atoms or more, the computational cost of conventional KS-DFT becomes prohibitive. This is partially due to KS-DFT's use of one-electron wave functions (orbitals), which leads to $3N$ degrees of freedom, with N as the number of electrons.

A possible alternative to KS-DFT for large scale materials simulation is orbital-free density functional theory (OF-DFT),[4] in which the total energy functional depends only on the electron density, instead of the N KS orbitals. Consequently, OF-DFT reduces the degrees of freedom in the computation from $3N$ to 3, greatly simplifying the formalism. A recent implementation[5] has made the computational cost of all parts of the OF-DFT calculation scale linearly with respect to system size for all sizes; i.e., there is no crossover between cubic and linear scalings as in linear-scaling KS-DFT.[6] The linear scaling, coupled with parallelization via domain decomposition, now allows an unprecedented number of atoms ($\sim 10^6$) to be treated explicitly with quantum mechanics. Such a fast first-principles method is tantalizing to use to study materials phenomena at the mesoscale, previously unreachable with quantum mechanics methods. However, previous work[7–11] has demonstrated that OF-DFT is only capable of KS-DFT-level accuracy for nearly-free-electron-like main group metals, with only marginal progress made in treating other types of materials.[12,13]

The key element determining the accuracy of OF-DFT is the expression used to evaluate the electron kinetic energy (KE) in terms of the electron density, namely, the kinetic energy density functional (KEDF). In KS-DFT, orbitals are used to evaluate the usual quantum mechanical expectation values of the Laplacian, giving rise to accurate values of the exact noninteracting electron KE. A far greater challenge is posed when evaluating the KE solely from the electron density, since an analogous exact expression for the KEDF is unknown, except in certain idealized limits.

Proposals of new KEDFs have occurred over many decades. These KEDFs generally can be grouped into two classes: (1) local and semilocal KEDFs and (2) nonlocal KEDFs.[4,14–20] Local and semilocal KEDFs are constructed based on the local electron density and its density gradient. Recently, a meta generalized gradient approximation (meta-GGA) KEDF[21] was also proposed, which adds the Laplacian of the electron density into the KEDF. A detailed survey on local and semilocal KEDFs was given recently by García-Aldea and Alvarellos,[22] and an older review was provided by Thakkar.[23] Nonlocal KEDFs typically have nonlocal kernels which relate any two points in space. Most commonly, nonlocal KEDFs have been derived from linear response theory for the perturbed uniform electron gas; i.e., they are based on the Lindhard response function[24] since its form is known analytically exactly in momentum space. However, the Lindhard response function at most can be expected to represent properly nearly-free-electron-like metals, with nearly uniform densities (hence the success mentioned above in applying OF-DFT to main group metals). Consequently, the Lindhard-based KEDFs cannot be expected to treat semiconductors well, since their linear response behavior is far different. In contrast to the uniform electron gas, no exact analytic form exists for the linear response function of

semiconductors, other than some models.[25–27] The purposes of the present work are to examine the basic physics that should be incorporated into a KEDF for semiconductors and then to build and test such a KEDF.

This paper is structured as follows. First we discuss the asymptotic behavior of a semiconductor's susceptibility function $\widetilde{\chi}(\vec{q})$ at the $\vec{q} \to 0$ limit, where $\vec{q}$ is the electron momentum, and then we propose a KEDF based on this behavior. Then we test our KEDF on a variety of binary semiconductors as well as different phases of and defects in silicon. The transferability of the two parameters in our KEDF is also analyzed both formally and numerically. We also propose recommended values for the two parameters of the KEDF that work reasonably well across a broad range of tetrahedrally-bonded semiconductors.

II. KINETIC ENERGY DENSITY FUNCTIONAL FORM

In KS-DFT, the total energy functional is partitioned as

$$E[\rho] = T_S[\rho] + V[\rho], \tag{1}$$

where $T_S[\rho]$ is the KS KEDF of the noninteracting electron gas and $V[\rho]$ contains the Hartree electron-electron repulsion energy, the electron exchange-correlation energy, and the external potential energy (usually the latter is just the ion-electron attraction energy, where the ionic potentials consist of either bare nuclear potentials or pseudopotentials accounting for attraction of the valence electrons to nuclei screened by the core electrons). Within linear response theory, the KS susceptibility function is defined by

$$\delta\rho(\vec{r}) = \int \chi^{\mathrm{KS}}(|\vec{r} - \vec{r}'|)\,\delta v^{\mathrm{KS}}(\vec{r}')d\vec{r}'^{\,3}, \tag{2}$$

where χ^{KS} is simply a function of $|\vec{r} - \vec{r}'|$. $\delta\rho(\vec{r})$ is the change in electron density induced by a perturbation of the KS effective potential $\delta v^{\mathrm{KS}}(\vec{r})$, which includes the electrostatic (i.e., electron-electron plus ion-electron) part $\delta v^{ele}(\vec{r})$ and the exchange-correlation part $\delta v^{\mathrm{XC}}(\vec{r})$,

$$\delta v^{\mathrm{KS}}(\vec{r}) = \delta v^{ele}(\vec{r}) + \delta v^{\mathrm{XC}}(\vec{r}). \tag{3}$$

Equation (2) can be rewritten as

$$\hat{F}\!\left(\frac{\delta v^{\mathrm{KS}}(\vec{r})}{\delta\rho(\vec{r}')}\right) = \hat{F}\!\left(\frac{\delta v^{ele}(\vec{r})}{\delta\rho(\vec{r}')}\right) + \hat{F}\!\left(\frac{\delta v^{\mathrm{XC}}(\vec{r})}{\delta\rho(\vec{r}')}\right) = \frac{1}{\widetilde{\chi}^{\mathrm{KS}}(\vec{q})}, \tag{4}$$

where $\hat{F}$ is the Fourier transform operator.

The general behavior of $\widetilde{\chi}^{\mathrm{KS}}(\vec{q})$ at $q \to 0$ limit has been shown by Pick *et al.*[31] to be

$$\widetilde{\chi}^{\mathrm{KS}}(q \to 0) \to -a \quad \text{(metal)},$$

$$\widetilde{\chi}^{\mathrm{KS}}(q \to 0) \to -bq^2 \quad \text{(insulator)},$$

where a and b are positive numbers. Here we give another proof for the behavior of an insulator's susceptibility function (the second relation above) within the local density approximation (LDA)[28,29] or the GGA[30] for exchange-

correlation. Under an external perturbing potential, the solid undergoes a polarization $\delta\vec{P}(\vec{q})$ (in Fourier space) according to standard electrostatics, expressed as

$$\delta\vec{P}(\vec{q}) = (4\pi)^{-1}[\widetilde{\varepsilon}(\vec{q}) - 1]\delta\vec{E}(\vec{q}), \tag{5}$$

with $\widetilde{\varepsilon}(\vec{q})$ being the dielectric function and with a change in electric field $\delta\vec{E}(\vec{q})$ given by

$$\delta\vec{E}(\vec{q}) = -i\vec{q}\,\delta v^{ele}(\vec{q}). \tag{6}$$

The polarization $\delta\vec{P}(\vec{q})$ is related to the induced density change as

$$-i\vec{q} \cdot \delta P(\vec{q}) = \delta\rho(\vec{q}). \tag{7}$$

By combining Eqs. (5)–(7) we obtain

$$-\frac{\delta\rho(\vec{q})}{q^2} = \frac{1}{4\pi}[\widetilde{\varepsilon}(\vec{q}) - 1]\delta v^{ele}(\vec{q}). \tag{8}$$

Rearranging and substituting Eq. (8) into Eq. (4), we can relate $\widetilde{\chi}^{\mathrm{KS}}(\vec{q})$ to $\widetilde{\varepsilon}(\vec{q})$ as

$$\frac{1}{\widetilde{\chi}^{\mathrm{KS}}(\vec{q})} = -\frac{4\pi}{q^2[\widetilde{\varepsilon}(\vec{q}) - 1]} + \hat{F}\!\left[\frac{\delta v^{\mathrm{XC}}(\vec{r})}{\delta\rho(\vec{r}')}\right]. \tag{9}$$

Up to this point only the linear response approximation has been made. We now invoke the LDA/GGA for exchange-correlation. Under the latter, v^{XC} is a functional only of the density (LDA) and possibly density gradients (GGA), which in turn means that the second term in Eq. (9) is merely a polynomial in q. $\widetilde{\varepsilon}(q=0)$ is just the macroscopic static dielectric constant, which is a finite number greater than 1 for semiconductors. Consequently, the second term in Eq. (9) cannot cancel out the $-1/q^2$ singularity from the first term as $q \to 0$. Thus, for semiconductors treated with the LDA/GGA, $\widetilde{\chi}^{\mathrm{KS}}(\vec{q})$ behaves as $\sim -q^2$ as $q \to 0$.

Our simple proof above relies on use of the most common exchange-correlation functionals, namely, the LDA/GGA, and therefore does not include the nonlocal response function part of the exchange-correlation functional. However, our conclusion was proved in general by Pick *et al.*[31] Since our purpose here is to derive a KEDF for use with the LDA/GGA for exchange-correlation anyway, it is sufficient for us to prove the asymptotic behavior of $\widetilde{\chi}^{\mathrm{KS}}(\vec{q})$ within the LDA/GGA. Note also that our proof above only works for semiconductors and not for metallic systems where $\widetilde{\varepsilon}(q=0) \to \infty$. [The denominator of the first term in Eq. (9) is then indeterminate as $q \to 0$.] However, we expect that $\widetilde{\chi}^{\mathrm{KS}}(\vec{q})$ for metals should approach a finite value as $q \to 0$ based on the known linear response function of the uniform gas, namely, the Lindhard response function.

The contrasting behavior of $\widetilde{\chi}^{\mathrm{KS}}(q)$ as $q \to 0$ for metals and semiconductors is a key feature that distinguishes their susceptibility functions. From the easily-derived relationship between $\widetilde{\chi}^{\mathrm{KS}}(q)$ and $T_S[\rho]$, namely,[4]

$$\frac{1}{\widetilde{\chi}^{\mathrm{KS}}(\vec{q})} = -\hat{F}\!\left(\frac{\delta^2 T_S[\rho]}{\delta\rho(\vec{r}')\delta\rho(\vec{r})}\right), \tag{10}$$

and the asymptotic behavior of $\widetilde{\chi}^{\mathrm{KS}}(\vec{q})$ above, we obtain the condition that $T_S[\rho]$ needs to satisfy as $q \to 0$

$$\hat{F}\left[\frac{\delta^2 T_S(\vec{r})}{\delta\rho(\vec{r}')\delta\rho(\vec{r})}\right] \sim \frac{1}{q^2[\tilde{\varepsilon}(\vec{q}=0)-1]}, \quad (11)$$

which in turn implies the following asymptotic behavior in real space as $|\vec{r}-\vec{r}'| \to \infty$:

$$\frac{\delta^2 T_S(\vec{r})}{\delta\rho(\vec{r}')\delta\rho(\vec{r})} \sim \frac{1}{|\vec{r}-\vec{r}'|[\tilde{\varepsilon}(\vec{q}=0)-1]}. \quad (12)$$

To our knowledge, the latter condition has not been accounted for in any KEDF to date.

Here we propose to explicitly impose the asymptotic behavior given in Eq. (12) by generalizing a previous KEDF form due to Wang and Teter.[15] Other KEDF forms could be generalized as well, but in the present work we consider only this form. Wang and Teter partitioned the KEDF as

$$T_S[\rho] = T_{\text{TF}} + T_{\text{vW}} + T_{NL}, \quad (13)$$

where $T_{\text{TF}}=C_{\text{TF}}\int \rho^{5/3}dr^3$, with $C_{\text{TF}}=\frac{3}{10}(3\pi^2)^{2/3}$ is the local Thomas-Fermi (TF) KEDF,[32–34] $T_{\text{vW}}=\frac{1}{8}\int(|\nabla\rho|^2/\rho)dr^3$ is the semilocal von Weizsäcker KEDF,[35] and the nonlocal part is generally written as

$$T_{NL} = C \int\int \rho(\vec{r})^\alpha \omega[\xi(\vec{r},\vec{r}')|\vec{r}-\vec{r}'|]\rho(\vec{r}')^\beta dr^3 dr'^3. \quad (14)$$

Here $\xi(\vec{r},\vec{r}')$ is an effective Fermi wave vector that may depend on densities at $\vec{r}$ and $\vec{r}'$, ω is a dimensionless kernel, $C=C_{\text{TF}}8(3\pi^2)$, and α and β are two exponents that satisfy $\alpha+\beta=8/3$ to ensure the correct dimensionality. The kernel ω depends on the distance between $\vec{r}$ and $\vec{r}'$, scaled by $\xi(\vec{r},\vec{r}')$. Clearly in the Wang-Teter form, the nonlocal part is a function of $|\vec{r}-\vec{r}'|$. We generalize the kernel to make it a function of both $|\vec{r}-\vec{r}'|$ and $1/|\vec{r}-\vec{r}'|$ in order to explicitly include the $1/|\vec{r}-\vec{r}'|$ asymptotic behavior of semiconductors given in Eq. (12). Our general kernel thus takes the form

$$\omega(\vec{r},\vec{r}') = \omega\left(G[\rho]|\vec{r}-\vec{r}'| + \frac{Y[\rho]}{|\vec{r}-\vec{r}'|}\right), \quad (15)$$

where $G[\rho]$ and $Y[\rho]$ are unknown functionals. Here we consider a specific kernel

$$\omega(\vec{r},\vec{r}') = \omega[\xi(\vec{r},\vec{r}')|\vec{r}-\vec{r}'|] \quad (16)$$

with

$$\xi(\vec{r},\vec{r}') = k_F(\vec{r})\left\{1 + \lambda\left[\frac{\rho(\vec{r})-\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}\right]^2 \frac{1}{\rho(\vec{r})^{8/3}}\right\}, \quad (17)$$

where $k_F(\vec{r})=[3\pi\rho(\vec{r})]^{1/3}$ is the Fermi wave vector and the parameter λ controls the mixing between the $|\vec{r}-\vec{r}'|$ term and the $1/|\vec{r}-\vec{r}'|$ term in the kernel. If $\lambda=0$, we have a kernel that depends only on $|\vec{r}-\vec{r}'|$ and the density $\rho(\vec{r})$. $\rho(\vec{r})-\rho(\vec{r}')$ in Eq. (17) is introduced to remove the singularity when $|\vec{r}-\vec{r}'|=0$ and the $1/\rho(\vec{r})^{8/3}$ factor preserves correct dimensionality.

By introducing the reduced density gradient $s(\vec{r})$, we approximate the argument of the quadratic term in Eq. (17) by

$$\frac{\rho(\vec{r})-\rho(\vec{r}')}{|\vec{r}-\vec{r}'|} \to \nabla\rho(\vec{r}), \quad (18)$$

leading Eqs. (16) and (17) to simplify to

$$\omega(\vec{r},\vec{r}') = \omega[k_F(\vec{r})(1+\lambda s(\vec{r})^2)|\vec{r}-\vec{r}'|],$$

$$s(\vec{r}) = \frac{|\nabla\rho(\vec{r})|}{\rho(\vec{r})^{4/3}}.$$

Justification of the approximation made in Eq. (18) is given later by considering TF dielectric screening theory.

Our total KEDF in this work finally becomes

$$T_S[\rho] = T_{\text{TF}} + T_{\text{vW}} + T_{NL} \quad (19)$$

with

$$T_{NL} = C \int\int \rho(\vec{r})^{8/3-\beta}\omega[k_F(\vec{r})(1+\lambda s(\vec{r})^2)|\vec{r}-\vec{r}'|]$$
$$\times \rho(\vec{r}')^\beta dr^3 dr'^3. \quad (20)$$

We observe that λ, the coefficient of $s(\vec{r})$, is implicitly linked to the static dielectric constant by Eq. (12); implications of this relationship are discussed in Sec. IV.

To specify the kernel $\omega(\vec{r},\vec{r}')$, we enforce the exact linear response of a uniform electron gas onto our KEDF as one limit we wish to satisfy

$$\hat{F}\left(\left.\frac{\delta^2 T_S[\rho]}{\delta\rho(\vec{r})\delta\rho(\vec{r}')}\right|_{\rho(r)=\rho^*}\right) = -\frac{1}{\tilde{\chi}_{\text{Lindhard}}(q)},$$

$$\tilde{\chi}_{\text{Lindhard}}(\eta) = -\frac{k_F}{\pi^2}\left(\frac{1}{2} + \frac{1-\eta^2}{4\eta}\ln\left|\frac{1+\eta}{1-\eta}\right|\right) = -\frac{k_F}{\pi^2}\frac{1}{F(\eta)}. \quad (21)$$

Here ρ^* is the uniform electron gas density and $\eta=q/(2k_F)$. Imposition of Eq. (21) leads to a first-order ordinary differential equation for the kernel

$$-\beta\eta\tilde{\omega}(\eta)' + (5-3\beta)\beta\tilde{\omega}(\eta) = \frac{5}{3}[F(\eta) - 3\eta^2 - 1], \quad (22)$$

where $\tilde{\omega}(\eta)$ is the Fourier transform of the kernel $\omega(\vec{r},\vec{r}')$ and β arises from the functional derivatives of the density exponents in Eq. (20). The parameter λ does not appear in the kernel equation, which indicates that λ should not have any effect as the system approaches a uniform electron gas. We integrate Eq. (22) from $+\infty$ to zero with the Runge-Kutta method implemented in the RKSUITE[36] code in order to evaluate the kernel.

Although it is an important limit to retain, an inconsistency arises by forcing our KEDF to reproduce the Lindhard response function in Eq. (21). Analysis of Eq. (22) reveals that $\tilde{\omega}(\eta)$ is zero at $q=0$, since $F(\eta)=1$ at $q=0$. Consequently, our kernel does not exhibit the correct behavior for semiconductors as $q \to 0$. Our kernel runs into this problem partially because we selected the Wang-Teter KEDF form. In the latter, T_{NL} has to be zero for a uniform electron gas in order to recover the correct (TF) limit, which makes $\int\int\omega(|\vec{r}-\vec{r}'|)dr^3 dr'^3=0$ and therefore $\tilde{\omega}(\eta=0)$ is zero [since $\tilde{\omega}(0) \sim \int\omega(\vec{r})dr^3$]. Even if the Wang-Teter form is a good

approximation for nearly-free-electron gases (simple metals), it is likely not the optimal starting point to construct KEDFs for semiconductors. However, here we still use the Wang-Teter KEDF form and explore whether our KEDF can be applied successfully to both semiconductors and metals. Thus, our proposed kernel in Eqs. (19) and (20) is admittedly compromised by using the Wang-Teter KEDF form. Imposition of the Lindhard response behavior itself may also be a limitation on accuracy, but unfortunately no more sophisticated analytical form for the response function is available to use for KEDF kernel construction. A KEDF form fully correct for both the nearly-free-electron gas and semiconductors remains unknown.

III. NUMERICAL IMPLEMENTATION

Our KEDF is implemented in the PROFESS code,[37] which is a plane-wave-based OF-DFT code that imposes periodic boundary conditions. KS-DFT benchmark calculations are performed with the ABINIT code.[38] All calculations use the LDA for electron exchange-correlation.[28,29]

Various phases of silicon are studied in this work. The structures and unit cells used for cubic diamond (CD), hexagonal diamond (HD), complex body-centered-cubic (cbcc), β-tin, body-centered-tetragonal 5 (bct5), simple-cubic (sc), hexagonal-close-packed (hcp), body-centered-cubic (bcc), and face-centered-cubic (fcc) structures are all as given in our previous work.[11] We also examine the (cubic) zincblende (ZB) and (hexagonal) wurtzite (WZ) structures of the III-V semiconductors GaP, GaAs, GaSb, InP, InAs, and InSb.

In KS-DFT calculations with nonlocal pseudopotentials (NLPSs), we use Troullier-Martins[39] NLPSs generated in the FHI98 code[40] using the default cutoff radii. For all local pseudopotential KS-DFT and all OF-DFT calculations, we employ bulk-derived local pseudopotentials (BLPSs).[11,41] For silicon, we use the one previously reported and tested in Ref. 11, whereas for other binary semiconductors, we use new BLPSs built with the same method.[11] The details of constructing and the qualities of the BLPSs for Ga, In, P, As, and Sb are summarized in Appendix A.

In all KS-DFT and OF-DFT calculations, the number of plane waves (i.e., the kinetic energy cutoff) is increased until the total energy is converged to within 1 meV/cell (for defect formation energy calculations, the total energy is converged to 5 meV/cell). k-point meshes for the KS-DFT calculations are generated with the Monkhorst-Pack method.[42] The kinetic energy cutoff, the number of k points, and the number of atoms in each periodic cell are listed in Table I. In KS-DFT calculations, Fermi-Dirac smearing is used (smearing width of 0.1 eV) for metallic phases, with no smearing for semiconductors. In OF-DFT, the kinetic energy cutoff used is 1600 eV for all structures, which converges the total energies to within 1 meV/cell.

For the OF-DFT calculations, we describe in Appendix B how to render the cost of computing T_{NL} with our KEDF and its potential $\delta T_{NL}/\delta\rho$ linear scaling with respect to system size. Other details of the implementation of the OF-DFT calculations can be found in Ref. 37.

A variety of properties was calculated to test our KEDF, including equilibrium volumes, bulk moduli, phase energy

TABLE I. Kinetic energy cutoffs (E_{cut} in eV) and k-point meshes used for various KS-DFT calculations in this work. The numbers of atoms per unit cell are given in parentheses next to each phase (sc is simple cubic, ZB is zincblende, WZ is wurtzite, CD is cubic diamond, HD is hexagonal diamond, and cbcc is complex bcc).

	E_{cut}	k-point mesh
fcc(1), hcp(2), bcc(1), and sc(1) silicon	1000	$20\times20\times20$
β-tin(2) and bct5(2) silicon	800	$20\times20\times20$
ZB(2) and WZ(4) III-V semiconductors	800	$12\times12\times12$
CD(2), HD(4), and cbcc(8) silicon		
Point defects in CD Si (63: vacancy; 65: self-interstitial)	760	$8\times8\times8$

differences, and defect formation energies. The KS-DFT equilibrium structures were determined by relaxing each structure with a force threshold of 5×10^{-5} hartree/bohr and a stress threshold of 5×10^{-7} hartree/bohr^3. The OF-DFT equilibrium structures were obtained by relaxing the internal (atomic) coordinates to within a force threshold of 2×10^{-4} hartree/bohr, while the lattice parameters, including the c/a ratio, were optimized manually (as opposed to minimizing the stress tensor). We have not yet derived the stress tensor expression for this KEDF, hence the manual optimization of the lattice vectors. OF-DFT predictions of equilibrium volumes and bulk moduli were then calculated by expanding and compressing the OF-DFT equilibrium unit cell structure by up to 5% to obtain eight energy versus volume points, which are then fit to Murnaghan's equation of state.[43] Phase energy differences are simply the differences in total energy (per atom or formula unit) between different phases at their equilibrium volumes. The phase transition pressures were calculated using the common tangent rule,

$$\left.\frac{dE}{dV}\right|_{\text{phase 1}} = \left.\frac{dE}{dV}\right|_{\text{phase 2}} = -P_{\text{trans}}.$$

A vacancy in CD Si was modeled by putting eight cubic unit cells together in a $2\times2\times2$ fashion and then removing one Si atom from the corner. With this $2\times2\times2$ supercell, a self-interstitial defect was constructed by inserting an extra Si atom at a tetrahedral interstitial site. For the vacancy and self-interstitial defects, spin-restricted DFT was used and the Si atoms were not structurally relaxed during these point defect benchmark calculations. We are not attempting to model the actual physical defect accurately, which would be better described by spin-polarized DFT and structural relaxation; we are only interested in testing the KEDF in comparison to KS-DFT within a given KS-DFT model—hence the use of spin-restricted DFT and unrelaxed structures for simplicity. Point defect energies are calculated according to Gillan's expression,[44]

$$E_{\text{defect}} = E\left(N\pm1, 1, \frac{N\pm1}{N}\Omega\right) - \frac{N\pm1}{N}E(N,0,\Omega),$$

TABLE II. Optimal nonlocal KEDF parameters λ and β fitted to reproduce KS-DFT/BLPS equilibrium volumes (V_0 in Å^3) and total energies (E_0 in eV) per unit cell for CD Si and various ZB semiconductors. OF-DFT/BLPS and KS-DFT/BLPS (in parentheses) predictions of bulk moduli (B in GPa) as well as best fits to V_0 and E_0 are also given.

	B	V_0	E_0	$\lambda(\times 10^{-2})$	β
Silicon	97	39.926	−219.248	1.00	0.650
	(98)	(39.562)	(−219.258)		
AlP	91	40.423	−240.199	1.20	0.845
	(90)	(40.637)	(−240.182)		
AlAs	76	44.700	−232.912	1.25	0.825
	(80)	(43.616)	(−232.908)		
AlSb	61	55.996	−206.588	1.20	0.750
	(60)	(56.607)	(−206.606)		
GaP	87	37.788	−243.057	1.00	0.791
	(80)	(37.646)	(−243.079)		
GaAs	81	40.798	−235.782	1.30	0.783
	(75)	(40.634)	(−235.799)		
GaSb	58	52.686	−209.739	1.00	0.720
	(56)	(52.488)	(−209.697)		
InP	66	45.854	−235.696	1.20	0.885
	(73)	(46.040)	(−235.722)		
InAs	63	49.500	−228.523	1.42	0.875
	(65)	(49.123)	(−228.537)		
InSb	49	62.847	−202.381	1.20	0.810
	(50)	(62.908)	(−202.387)		

where $E(N,m,\Omega)$ is the total energy for a cell of volume Ω, with N atoms and m defects. The "+" sign is for the self-interstitial defect and the "−" sign is for the vacancy defect.

IV. RESULTS AND DISCUSSION

A. Bulk properties

To test the quality of our KEDF, we focus on the bulk modulus, the equilibrium volume, and the equilibrium total energy of various semiconductors. If our KEDF is a good model of the KS T_S, we should at least be able to reproduce the above three properties for each semiconductor by adjusting the only two parameters in our KEDF: λ and β. As a point of reference, corresponding values of the exponents α and β used in the Wang-Teter KEDF are $\alpha=\beta=5/6$,[15] in Perrot's version of the Wang-Teter KEDF $\alpha=\beta=1$,[20] in Smargiassi and Madden's version of the Wang-Teter KEDF $\alpha=\beta=1/2$,[7] and in the Wang-Govind-Carter 1998 (WGC98) version of the Wang-Teter KEDF $\alpha,\beta=(5\pm\sqrt{5})/6$.[17] In this work, $\alpha+\beta=8/3$, which is different from the $\alpha+\beta=5/3$ in the Wang-Teter and WGC98 KEDFs, because our kernel as defined in Eq. (20) is dimensionless whereas the Wang-Teter and WGC98 KEDF kernels have dimensions of the electron density.

Table II lists the optimal λ and β for the ground state phase of each semiconductor, which were fitted to the KS-DFT equilibrium energies and volumes of each semiconductor ground state. The bulk moduli were not part of the fit and

therefore represent a verification test of the KEDF. These bulk properties are mainly controlled by the parameter λ that scales the reduced density gradient, whereas the parameter β has a much smaller effect. We adjusted β only to refine the final OF-DFT equilibrium total energy. Table II shows that with optimal parameters, our KEDF yields very good bulk properties for all these semiconductors. To our knowledge, this is the only KEDF model able to reproduce KS-DFT bulk moduli, equilibrium energies, and equilibrium volumes well for this large set of semiconductors, with only two parameters.

Figures 1 and 2 further illustrate the close correspondence between KS and OF-DFT predictions, via total energy versus isotropic volume curves for CD and β-tin silicon, as well as ZB GaAs. Optimal λ and β were used for CD silicon ($\lambda=0.01$ and $\beta=0.65$) and for ZB GaAs ($\lambda=0.013$ and $\beta=0.783$), producing truly excellent agreement. By contrast, the OF-DFT total energy versus volume curve for β-tin silicon deviates from the KS-DFT curve due to use of a nonoptimal β. In principle, if one wanted to model solely β-tin silicon within OF-DFT, it would be best to optimize β as well as λ to obtain higher accuracy in the KEDF.

Next we examine briefly the sensitivity of bulk properties to the choice of λ. We focus on CD silicon with β fixed to 0.65 (its optimal value) and vary $\lambda \pm 40\%$ around its optimal value, from 0.006 to 0.014, which is the range of optimal values found across all semiconductors examined thus far. Figure 3 reveals that the bulk modulus is the strongest function of λ, varying $\pm 17\%$ around the optimal value of λ,
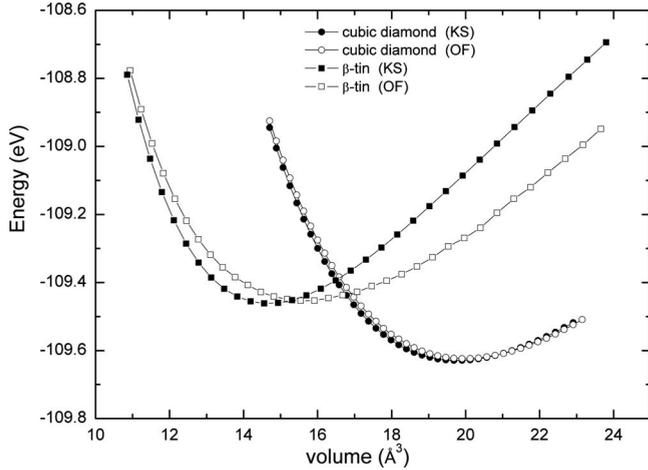
FIG. 1. OF-DFT and KS-DFT total energy versus volume curves for CD and β-tin silicon. For CD silicon, the optimal KEDF parameters are used ($\lambda = 0.01$ and $\beta = 0.65$). For β-tin silicon, the optimal $\lambda = 0.0055$ is used with $\beta = 0.65$ (optimal for CD silicon).

while the equilibrium volume and the total energy per atom change more modestly with λ [$\pm 4\%$ and $\pm 0.04\%$ (<0.1 eV), respectively]. As λ increases, the bulk modulus increases, while the total energy and equilibrium volume tend to decrease. As mentioned earlier, the strength of the reduced density gradient term in the nonlocal KEDF kernel is determined by λ; we see that subtle corrections to the physics are provided by this term, given that the magnitude of λ is quite small.

The physics contained in our KEDF offers a significant improvement over an earlier attempt to develop a KEDF for covalent materials.[13] In that case, a nonlocal KEDF based on uniform-gas linear response theory was proposed, again with only two tunable parameters: (1) the average valence density ρ_* used in a Taylor expansion of the KEDF (used to achieve algorithmic linear scaling) and (2) the γ exponent defining the two-body Fermi wave vector in the density-dependent kernel of the WGC 1999 (WGC99) KEDF.[17] These two pa-
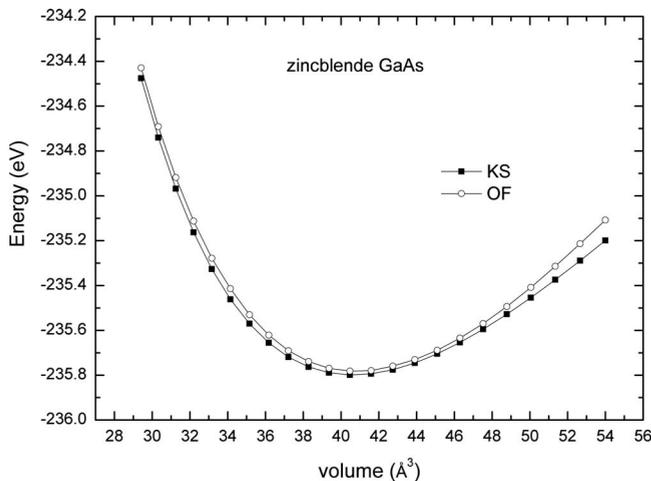

FIG. 2. OF-DFT and KS-DFT total energy versus volume curves for ZB GaAs. The OF-DFT curve is calculated with optimal $\lambda = 0.013$ and $\beta = 0.783$.
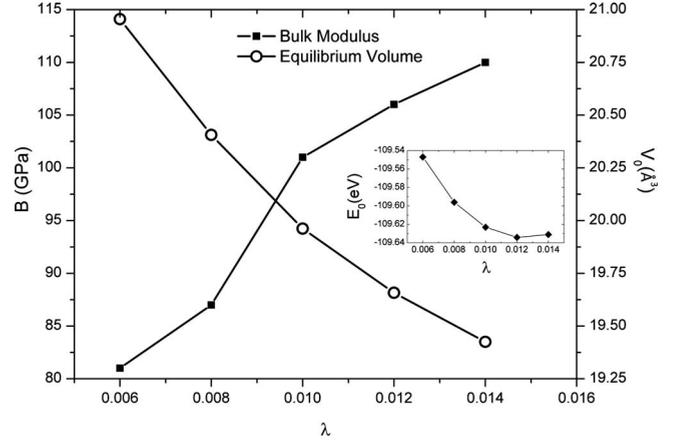

FIG. 3. Variation in CD silicon bulk modulus, equilibrium volume per atom, and total energy per atom with different λ.

rameters were optimized, but the best predicted bulk modulus of CD silicon was in error by 34%. By contrast, our bulk modulus of CD silicon lies within 2% of the KS-DFT result. Moreover, although this earlier KEDF was able to obtain CD Si as the ground state in an OF-DFT calculation, the three basic structural properties (bulk modulus, equilibrium volume, and equilibrium energy) and the equation of state (energy versus volume curve) of CD silicon could not be simultaneously reproduced well by that KEDF, in contrast to the KEDF proposed here.

B. Electron density

To evaluate the performance of our KEDF with respect to reproducing electron densities, we compute the self-consistent valence electron density for CD Si and ZB GaAs using OF-DFT with our nonlocal KEDF, where we set λ either to zero or a nonzero value to see how λ affects the density distribution. The latter case considers the inhomogeneity of the electron distribution, while the former case reduces our KEDF to a WGC99-like KEDF that should work well for metallic phases. (The WGC99 KEDF is double-density dependent,[17] whereas our KEDF is single-density dependent; WGC99 gives quantitative accuracy for nearly free-electron-like metals.) We then compare the resulting electron densities with benchmark KS-DFT electron densities. Figures 4 and 5 reveal that our KEDF with nonzero λ produces a density closer to the KS-DFT density in the core and lower density regions (both are highly inhomogeneous regions); however the density in the bond region (left-hand side of plots) is worse than the $\lambda = 0$ case.

C. Validity of the semilocal approximation

One key assumption made is the introduction of the reduced density gradient in Eq. (18), which approximates the nonlocal part of the kernel with a semilocal term. The question arises as to whether it is physically justified to make this replacement. We now show that this replacement appears valid based on Resta's[45] TF dielectric screening theory and our numerical tests.
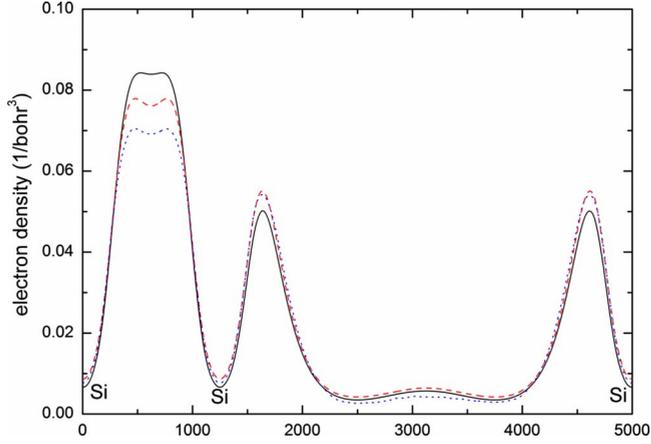
FIG. 4. (Color online) The electron density of CD silicon along the [111] direction. Black solid line: KS-DFT. Red dashed line: OF-DFT with $\lambda=0.0$. Blue dotted line: OF-DFT with $\lambda=0.01$ and $\beta=0.65$. Vertical axis is electron density in $1/\text{bohr}^3$; horizontal axis represents the grid.

Let us take CD silicon as an example. Consider the screening length for each silicon ion in the crystal. If the screening length is long over several nearest-neighbor distances, then it is invalid to make the approximation given in Eq. (18). However, according to Resta's theory,[45] the screening length in CD silicon is roughly equal to the nearest-neighbor distance. Resta extended the TF dielectric screening theory to semiconductors; previously the theory had been used exclusively for metals. With the screening length denoted as R and the ionic charge of each silicon ion as Z, the screened electrostatic potential of each single silicon ion beyond a certain distance R is modeled as

$$V(r) = -\frac{Z}{\varepsilon r}, \quad r \geq R,$$

where ε is the static dielectric constant in CD silicon and r is the distance from the silicon ion. The unscreened electro-
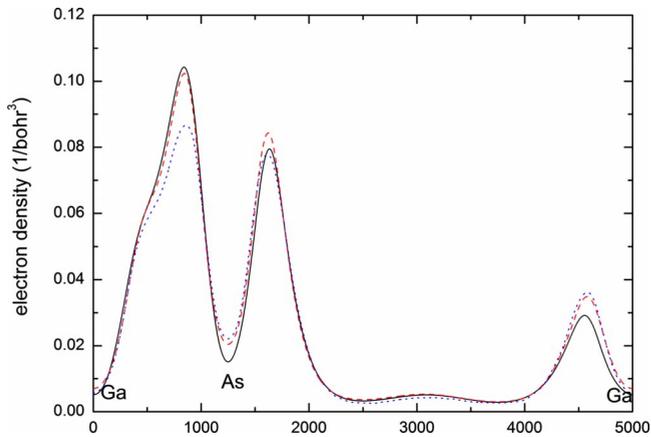

FIG. 5. (Color online) The electron density of ZB GaAs along the [111] direction. Black solid line: KS-DFT. Red dashed line: OF-DFT with $\lambda=0.0$. Blue dotted line: OF-DFT with $\lambda=0.013$ and $\beta=0.783$. Vertical axis is electron density in $1/\text{bohr}^3$; horizontal axis represents the grid.

static potential inside R is assumed to obey the TF theory and is determined from the TF equation, with the general solution of

$$V(r) = -\frac{Z}{r}(\alpha e^{qr} + \beta e^{-qr}) + A, \quad r < R,$$

where α, β, and A are parameters determined by the boundary conditions of $V(r)$ at $r=0$ and $r=R$. Here $q=\sqrt{4k_F/\pi}$ with $k_F=(3\pi^2\rho)^{1/3}$. Resta finally obtained an equation for R as

$$\sinh(qR)/qR = \varepsilon. \tag{23}$$

This equation for R is easily solved using silicon's average valence density ρ and its static dielectric constant ε as input. The screening length R for CD silicon obtained from this equation is very close to the nearest-neighbor distance, which indicates that only nearest-neighbor silicon atoms participate in screening. Similar results were found for CD germanium and carbon by Resta.[45]

Thus, based on this TF dielectric screening theory, we argue that the nonlocality of the KEDF is weak outside first nearest neighbors and therefore it is valid to make the approximation given in Eq. (18). This feature was also exploited implicitly in Cortona's work,[46] in which he successfully calculated various bulk properties of many semiconductors using his embedding theory. In Cortona's approach, he divides the bulk into atoms and solves KS-DFT equations for each atom with an atom-centered Gaussian basis, with the atoms considered to be embedded in the bulk. He numerically showed that for CD Si and many other semiconductors, the "embedding potential" due to the kinetic energy interactions between atoms is approximated reasonably well with the TF KEDF (a local KEDF), and his quite good results provide support for our argument: the nonlocality of the KEDF in semiconductors is weak beyond first nearest neighbors.

D. Transferability and crystal structure dependence of λ and β

Given that the screening length R derived from Eq. (23) is short ranged for semiconductors, then we may infer that the local bonding environment in the crystal controls most of the dielectric screening. We therefore would expect the parameters λ and β to be transferable between solids of similar local bonding motifs. We acknowledge that this hypothesis ignores the fact that the dielectric function is a global property.

To test this hypothesis, we investigate the transferability of λ and β optimized for CD Si for predicting properties of HD and cbcc Si, all of which are tetrahedrally-bonded (Table III). Likewise, we test parameter transferability for the tetrahedrally-bonded ZB and WZ structures of various binary semiconductors (Table IV). In other words, λ and β were first fit to reproduce the equilibrium volume and energy of CD Si, and then these parameters were used to calculate the bulk properties of HD and cbcc Si. For binary semiconductors, λ and β were first fit to the ZB structures and then were applied to WZ structures. Tables III and IV clearly show that λ and β are transferable between CD, HD, and cbcc silicon,

TABLE III. Bulk properties of silicon in its CD, HD, and cbcc phases, as predicted by KS-DFT/BLPS (in parentheses) and OF-DFT/BLPS using our nonlocal KEDF with $\lambda=1\times10^{-2}$ and β $=0.65$ (parameters optimized for CD silicon only). Bulk moduli (B) are in GPa and equilibrium volumes per atom (V_0) are in Å^3. The equilibrium total energy per atom (E_0) for CD silicon and the energies of other structures relative to the CD phase are in eV.

Silicon structure	B	V_0	E_0
CD	97	19.962	−109.624
	(98)	(19.781)	(−109.629)
HD	98	19.875	0.007
	(99)	(19.642)	(0.015)
cbcc	105	18.419	0.141
	(102)	(17.517)	(0.122)

as well as between the ZB and WZ structures for each binary semiconductor, in terms of predicting the small KS-DFT energy differences between various phases (to within 25 meV). The transferability of λ and β is also evident in the reproduction of KS-DFT bulk moduli (aside from the 15% deviation for WZ GaP) and equilibrium volumes for HD and cbcc silicon, as well as for the WZ structures.

Although we have just argued and then numerically demonstrated good transferability as long as the local bonding environment is similar, once the coordination-number

TABLE IV. KS-DFT/BLPS (in parentheses) and OF-DFT/BLPS bulk moduli (B_{WZ} in GPa) and equilibrium volumes per formula unit (V_{0WZ} in Å^3) for WZ structures of various binary semiconductors. Energy differences per formula unit ($E_{WZ}-E_{ZB}$, in meV) between WZ and ZB structures are also listed. c/a ratios optimized in OF-DFT are compared to KS-DFT ratios (in parentheses). OF-DFT/BLPS results are calculated using the λ and β listed in Table II (fitted for each ZB semiconductor).

	B_{WZ}	V_{0WZ}	$E_{WZ}-E_{ZB}$	c/a
AlP	92	40.667	22	1.64
	(90)	(40.608)	(9)	(1.64)
AlAs	79	44.937	25	1.66
	(80)	(43.621)	(11)	(1.65)
AlSb	59	56.063	16	1.65
	(58)	(56.548)	(13)	(1.65)
GaP	76	37.972	26	1.65
	(88)	(37.625)	(18)	(1.65)
GaAs	76	40.923	25	1.65
	(76)	(40.611)	(19)	(1.65)
GaSb	59	52.826	15	1.64
	(57)	(52.397)	(16)	(1.65)
InP	73	46.085	19	1.64
	(73)	(46.037)	(3)	(1.64)
InAs	66	49.726	32	1.65
	(65)	(49.129)	(7)	(1.64)
InSb	48	63.068	17	1.64
	(50)	(62.884)	(11)	(1.65)

changes, the optimal λ is no longer transferable. To illustrate this, we consider the trend in optimal λ for various phases of Si, with β fixed (since as mentioned earlier λ is found to be much more important than β in determining bulk properties). We optimize λ for each Si phase to yield the best equilibrium energy and volume, with β fixed to 0.65, which is optimal for CD silicon. Table V reveals the general trend that the optimal λ becomes smaller for structures with higher coordination numbers (more metallic); this trend for λ makes complete sense when one considers the relationship between λ and static dielectric constant pointed out in Sec. II. $\lambda=0$ corresponds to the infinite static dielectric constant case, i.e., a metal, whereas finite lambda corresponds to a finite static dielectric constant, i.e., a semiconductor. If we let $\lambda=0$, our KEDF becomes physically similar to the Wang-Teter[15] and WGC99[17] KEDFs, in which the former's kernel has no density dependence, while the latter's kernel has a double-density dependence. Our KEDF's kernel has single-density dependence. Like the Wang-Teter and WGC99 KEDFs that describe nearly-free-electron-like metals well, our KEDF with $\lambda=0$ gives a good description for metallic structures of Si, aside from predicting the hcp phase to be less stable than it should be (however, the energy differences are very small and certainly within the typical uncertainty of KS-DFT). With these optimal λ, the equilibrium volumes are all fairly well reproduced by our KEDF, except for the hypothetical bct5 phase (14% deviation). OF-DFT bulk moduli of the β-tin through the fcc structures are all uniformly shifted downward by 20–30 GPa from the KS-DFT predictions, suggesting that if these metallic phases are of interest, optimization of the β parameter may be critical. Earlier OF-DFT studies of these Si phases using a reparametrization of the WGC99 KEDF by Zhou *et al.*[13] had trouble reproducing the small energy difference between CD and HD silicon whereas our KEDF captures this small energy difference quite well. The notable crystal structure dependence of λ exhibited here suggests that a potential future avenue of research could be to parametrize λ as a function of coordination number.

We also tested our KEDF's transferability by calculating the transition pressure for Si transforming from the CD to the β-tin structure. When the KEDF parameters optimized for CD Si ($\lambda=1\times10^{-2}$ and $\beta=0.65$) are used for both phases, the predicted transition pressure is −2.3 GPa compared to KS-DFT transition pressures of 5.4 (using the BLPS) and 7.4 GPa (using the NLPS). This unphysical OF-DFT result is undoubtedly due to the nonoptimal λ for the β-tin structure, which is metallic and has a larger coordination number than the CD structure. As discussed earlier and shown in Fig. 3, when λ is too large it produces too low a total energy. In this case, the too-large value of λ overly stabilizes the β-tin structure so its energy is below the CD structure of silicon. As a result, the transition pressure becomes negative. If instead, the optimal, smaller $\lambda=0.55\times10^{-2}$ is used for β-tin silicon (see Fig. 1, where $\beta=0.65$ for both phases), the OF-DFT transition pressure is predicted to be a physically reasonable 6.4 GPa, quite close to the KS-DFT/BLPS value of 5.4 GPa. From these results, the need to let λ vary as the coordination-number changes seems clear.

In earlier work, Zhou *et al.*[13] obtained an OF-DFT/BLPS CD to β-tin transition pressure for silicon of 12.0 GPa com-

TABLE V. Bulk moduli B_0 (in GPa), equilibrium volumes (V_0) (in Å^3/atom), and equilibrium total energies (E_0) (in eV/atom) for various silicon phases (abbreviations defined in the text), calculated using KS-DFT/BLPS (in parentheses) and OF-DFT/BLPS with the optimal λ (value listed should be multiplied by 10^{-2}) for each structure and β fixed at 0.65 (optimal for CD Si). "c.n." stands for coordination number.

Si	CD	HD	cbcc	β-tin	bct5	sc	hcp	bcc	fcc
λ	1	1	1	0.55	0.26	0.3	0	0	0
B_0	97	98	105	83	66	81	71	82	64
	(98)	(99)	(102)	(99)	(96)	(112)	(91)	(98)	(83)
V_0	19.962	19.875	18.419	15.662	19.289	16.082	14.445	14.565	14.412
	(19.781)	(19.642)	(17.517)	(14.660)	(16.905)	(15.484)	(14.157)	(14.602)	(14.372)
E_0	−109.624	0.007	0.141	0.170	0.119	0.226	0.353	0.334	0.351
	(−109.629)	(0.015)	(0.122)	(0.168)	(0.215)	(0.229)	(0.340)	(0.351)	(0.381)
c.n.	4	4	4	6	5	6	12	8	12

pared to a KS-DFT/BLPS transition pressure of 10.2 GPa using a different BLPS and the reparametrized WGC99 KEDF. Given that the equations of state obtained for these two phases in that work exhibited large errors, the good agreement of the transition pressures was likely fortuitous.

Using λ and β fitted to perfect CD silicon, we further tested transferability by calculating the vacancy and self-interstitial defect formation energies in CD silicon. The vacancy formation energy is only off by 0.35 eV: KS-DFT/BLPS predicts 3.04 eV while OF-DFT/BLPS yields 2.69 eV. However the self-interstitial formation energy from OF-DFT with our KEDF is again the wrong sign (KS: 3.29 eV versus OF: −1.91 eV), which again is likely due to an improper λ used for the inserted Si atom at the interstitial position. The interstitial Si atom and its neighbors now have higher coordination numbers, which would be better described by a smaller value of λ (see trend in Table V). As mentioned above for β-tin, we find that too large a λ overstabilizes close-packed structures. Therefore the inappropriately large λ used for the interstitial Si atom produces too low energy, which in turn results in the negative self-interstitial formation energy. Again, this artifact could very well disappear if we were to parametrize λ based on the local environment instead of using a constant λ throughout the cell.

E. Recommended KEDF parameters for tetrahedrally-bonded semiconductors

Although the interrelationship of λ and the static dielectric constant formally precludes a single optimum value of λ for all semiconductors, for practical calculations it would be preferable to have one set of parameters to use for any semiconducting material. Consequently, we tested the transferability of the *average* values of λ and β, where we take a simple average of the optimal values shown in Table II. These average parameter values ($\lambda = 0.011\,77$ and $\beta = 0.7143$) are used to define the KEDF. As Table VI shows, KS bulk moduli are reproduced to within 5–10 % for most semiconductors (except for GaP, which is off by ~20%) and KS equilibrium volumes are reproduced to within 4%. The total energy is more sensitive to changes in λ and β, with maximum error of about 1% (errors of <2.3 eV). The total energy error incurred using OF-DFT is fairly uniform for all

the semiconductors, as evidenced by examining the energy differences between CD silicon and other ZB binary semiconductors (Fig. 6). The KS energy ordering trends, i.e., relative stability, among these semiconductors, are well reproduced by OF-DFT with this single set of λ and β. Thus, we conclude that these averaged values of λ and β will be a good first choice for modeling most tetrahedrally-bonded semiconductors.

TABLE VI. Bulk moduli (B), equilibrium volumes (V_0), and equilibrium total energies (E_0) per unit cell for CD silicon and ZB semiconductors calculated by OF-DFT with $\lambda = 0.011\,77$ and $\beta = 0.7143$ (averaged from Table II). KS-DFT values are given in parentheses.

|	B (GPa)	V_0 (Å^3)	E_0 (eV)
Silicon	100	39.567	−219.801
	(98)	(39.562)	(−219.258)
AlP	89	40.290	−238.612
	(90)	(40.637)	(−240.182)
AlAs	76	44.746	−231.702
	(80)	(43.616)	(−232.908)
AlSb	61	55.917	−206.309
	(60)	(56.607)	(−206.606)
GaP	94	36.795	−242.113
	(80)	(37.646)	(−243.079)
GaAs	78	41.214	−235.086
	(75)	(40.634)	(−235.799)
GaSb	62	51.779	−209.686
	(56)	(52.488)	(−209.697)
InP	68	45.930	−233.497
	(73)	(46.040)	(−235.722)
InAs	61	50.596	−226.775
	(65)	(49.123)	(−228.537)
InSb	49	62.461	−201.572
	(50)	(62.908)	(−202.387)
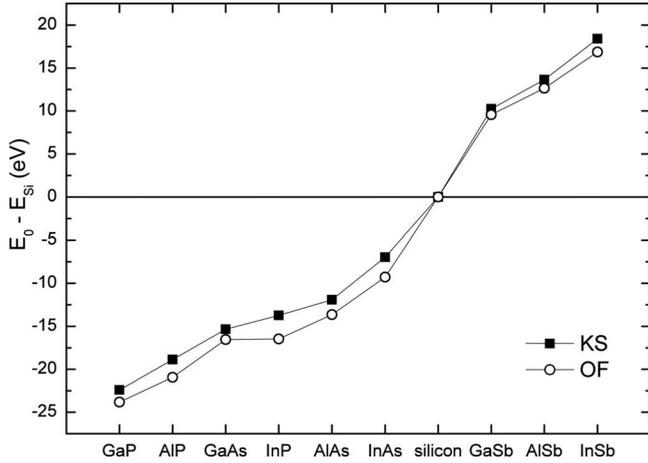
FIG. 6. OF-DFT and KS-DFT relative energy differences between CD silicon and ZB semiconductors (per primitive cell). OF-DFT results using an averaged $\lambda = 0.011\ 77$ and $\beta = 0.7143$ closely match the energy ordering from KS-DFT (see Table VI).

V. SUMMARY

In this work, we discussed the $\sim q^2$ asymptotic behavior at the $q \to 0$ limit of the susceptibility function in semiconductors. We pointed out that any KEDF designed for semiconductors therefore should behave as $1/|\vec{r} - \vec{r}'|$ as $|\vec{r} - \vec{r}'| \to \infty$. Based on this requirement, we proposed a general KEDF form whose kernel explicitly contains $1/|\vec{r} - \vec{r}'|$ and that has only two adjustable parameters. One of the parameters (λ) was shown to be directly related to the static dielectric constant. We tested our KEDF on properties of various binary semiconductors and a variety of phases of silicon. Since each semiconductor has a different dielectric function, a universal value for λ cannot exist. However, the two parameters λ and β in our KEDF can be adjusted to simultaneously reproduce three bulk properties of each semiconductor: KS-DFT bulk moduli, equilibrium volumes, and equilibrium energies. The parameters λ and β in our KEDF are observed to depend on

FIG. 8. Indium BLPS in real space. Coulombic tail is enforced beyond 4.0 bohr.

coordination number; our numerical results can be explained by appealing to Resta's TF dielectric screening theory, which demonstrates that in semiconductors the effective screening length is essentially a bond length. Consequently, the KEDF parameters are quite transferable within phases possessing the same local coordination number. We also determined a trend in the optimal λ for different silicon phases, with larger λ preferred for small coordination numbers and smaller λ for large coordination numbers, which can be understood based on the trend in dielectric constants for semiconducting versus metallic phases. As more severe tests, we calculated the transition pressure for the phase transition from CD Si to β-tin Si, as well as point defect formation energies in CD Si. From these latter calculations, a coordination-number-dependent parametrization of λ, instead of a constant λ throughout the crystal, appears to be required. However, for all tetrahedral-bonded semiconductors considered in this work, a single pair of λ and β in our KEDF, used within OF-DFT, is able to reproduce quite well KS-DFT predictions of basic bulk properties and relative energy orderings among various material
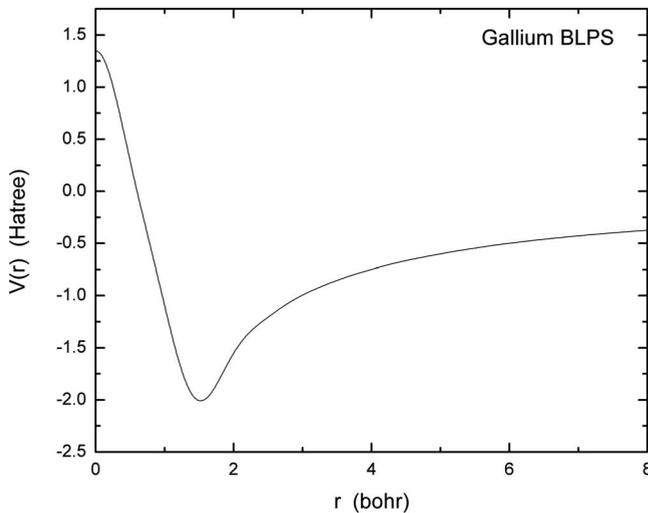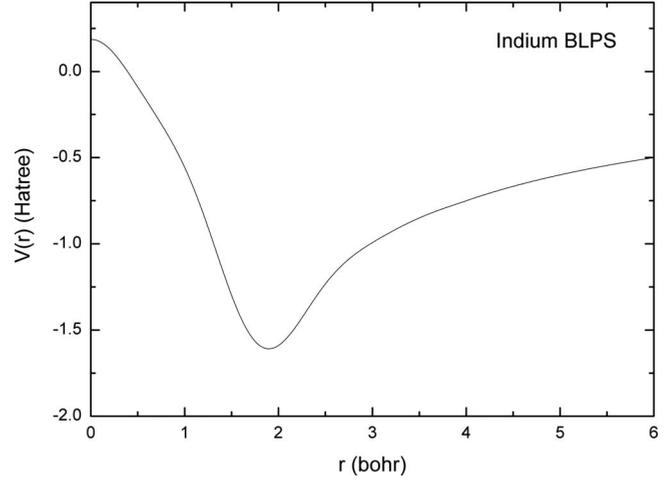

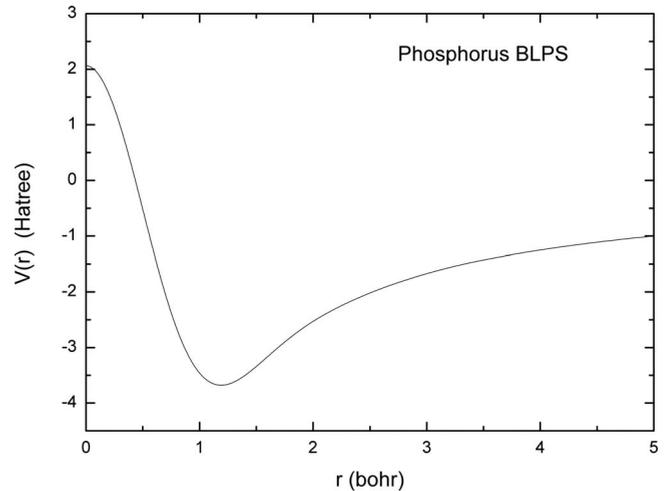FIG. 7. Gallium BLPS in real space. Coulombic tail is enforced beyond 3.5 bohr.

FIG. 9. Phosphorus BLPS in real space. Coulombic tail is enforced beyond 3.5 bohr.

FIG. 10. Arsenic BLPS in real space. Coulombic tail is enforced beyond 4.0 bohr.
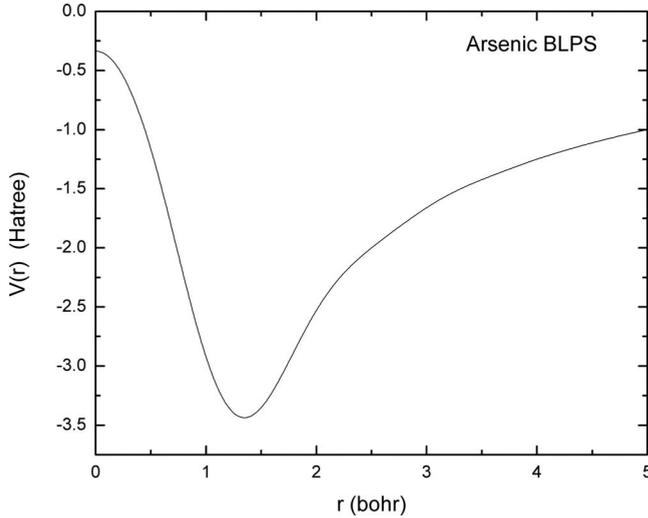

FIG. 11. Antimony BLPS in real space. Coulombic tail is enforced beyond 4.0 bohr.

phases. We believe that this work provides a new direction for future developments of KEDFs for semiconductors.

ACKNOWLEDGMENT

APPENDIX A

For all local pseudopotential KS-DFT and OF-DFT calculations in this work, we employ BLPSs[11,41] and the LDA.[28,29] For silicon, we use the BLPS previously reported and tested in Ref. 11, whereas for other binary semiconductors, we use new BLPSs built with the same method.[11] These BLPSs for Ga, In, P, As, and Sb are shown in Figs. 7–11.

The two parameters required to build these BLPSs, i.e., (1) the value of the non-Coulombic part of the Fourier-transformed BLPS at $q=0$ and (2) the position in real space beyond which a Coulomb tail is enforced, are chosen to reproduce the KS-DFT/NLPS energy ordering for fcc, sc, bcc, and CD phases of these elements. These structures are cho-
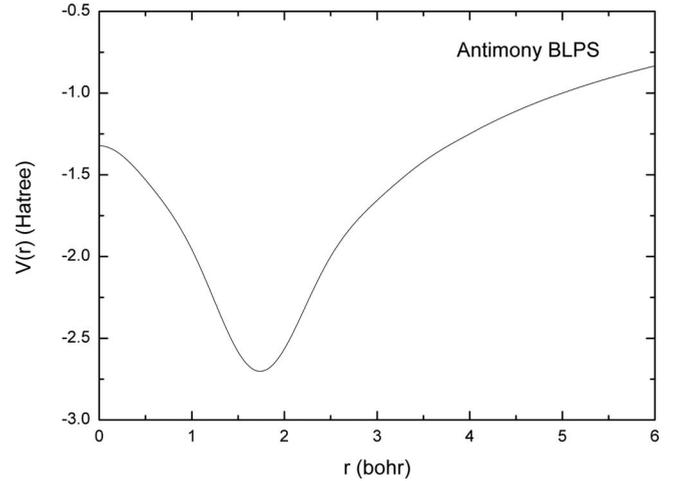
sen because they span a wide range of coordination-number environments in a solid. In building the BLPSs, Troullier-Martins[39] nonlocal pseudopotentials are generated with the FHI98 code[40] using default cutoff radii. Fermi-Dirac smearing with a smearing width of 0.1 eV is used all through the process of building of BLPSs, and the plane-wave basis kinetic energy cutoff is 2000 eV for all cases to obtain accurate target KS-DFT/NLPS electron densities used in the BLPS construction scheme. The k-point mesh used during BLPS construction is $20 \times 20 \times 20$ for fcc, sc, and bcc and $12 \times 12 \times 12$ for CD. The number of atoms per unit cell used is as given in Table I of the main text for elemental Si.

The BLPSs are then tested on the known ground states of these elements, i.e., α-gallium (Ga),[47] α-arsenic (As),[48] and α-antimony (Sb),[49] bct indium (In),[50] and A17 phosphorus (P),[51] and their properties are compared to other phases. In these tests, plane-wave basis kinetic energy cutoffs of 800 eV and $20 \times 20 \times 20$ k-point meshes are used to converge the total energy per cell to within 1 meV. No Fermi smearing is used for insulators. For metallic solids, a Fermi-Dirac smearing width of 0.1 eV is used. The number of atoms used per unit cell for the ground state structures is as follows: two atoms each for α-As and α-Sb and eight atoms each for α-Ga and A17 P.

TABLE VII. Comparison of NLPS and BLPS KS-LDA bulk properties of the α, fcc, hcp, bcc, sc, and CD phases of gallium. The bulk modulus (B) is in GPa, the equilibrium volume per atom (V_0) is in Å^3, and the equilibrium total energy (E_0) for α-Ga is in eV/atom. The total energies (eV/atom) of other Ga structures are listed relative to α-Ga's equilibrium total energy. α-Ga is the experimental ground state at low temperature (Ref. 47). The same convention and units for B, V_0, and E_0 are used in all subsequent tables.

	Gallium	α	fcc	hcp	bcc	sc	CD
B	NLPS	71	74	73	74	69	51
	BLPS	60	51	52	53	50	37
V_0	NLPS	17.488	15.668	15.793	15.990	16.906	21.693
	BLPS	17.232	15.359	15.545	15.927	16.924	21.501
E_0	NLPS	-61.637	-0.018	0.001	0.014	0.058	0.342
	BLPS	-60.871	-0.015	0.000	0.016	0.076	0.335

TABLE VIII. Comparison of NLPS and BLPS KS-LDA bulk properties of the bct, fcc, bcc, sc, and CD phases of indium. bct is the experimental ground state structure (Ref. 50).

	Indium	bct	fcc	bcc	sc	CD
B	NLPS	68	68	67	57	38
	BLPS	64	64	58	49	32
V_0	NLPS	20.237	20.219	20.551	22.350	29.857
	BLPS	20.052	20.050	20.621	23.858	32.274
E_0	NLPS	−56.620	0.000	0.038	0.171	0.576
	BLPS	−53.543	0.000	0.095	0.323	0.716

The KS-LDA/NLPS energy orderings, equilibrium volumes, and bulk moduli of different phases are mostly qualitatively if not quantitatively reproduced with these BLPSs, as shown in Tables VII–XI. Only two energy orderings are inverted for two elements—the bcc versus CD phases of phosphorus and arsenic—but the energy differences between these two phases in both cases are quite small (≤ 25 meV), certainly within the uncertainty of KS-LDA overall. The phosphorus energy differences between phases show the largest errors, but at least the correct ground state is obtained. α-gallium is not predicted to be the ground state structure with either the NLPS or the BLPS, which might be due to the lack of a nonlinear core correction[52] in the Ga NLPS or due to the LDA description of exchange-correlation. The equilibrium volumes of all phases of all elements are quite well reproduced by the BLPSs; errors in bulk moduli in some cases are significantly larger.

Table XII provides a transferability test of the bulk properties produced by these BLPSs for ZB binary semiconductors, which we see agree quite well with the NLPS predictions (except for the bulk modulus of InP). The number of atoms used per unit cell is as given in Table I of the main text. The absolute total energies are included for completeness, but there is no reason that different PSs should give the same absolute total energies, so deviations in the last column are not meaningful. As we are interested primarily in using these BLPSs for studying such binary compounds rather than the pure elements, these results are encouraging.

In addition to verifying electronic structural properties, we tested the electronic structure transferability of these BLPSs by comparing predicted KS-LDA/NLPS band gaps (KS eigenvalue gaps) with KS-LDA/BLPS band gaps for the ZB and WZ structures of each binary semiconductor, as well as for CD and hexagonal diamond (HD) silicon (Table XIII). Again, the number of atoms used per unit cell is as given in Table I of the main text. Generally, our BLPSs give comparable band gaps for most semiconductors except for HD Si (zero band gap) and ZB and WZ InAs and InSb (gaps too large). More difficult is the correct prediction of the nature of the gaps, namely, whether they are direct or indirect. However, since the main purpose of this work is to test our kinetic energy density functional on semiconductors, as long as our BLPSs predict these materials to be semiconductors within KS-DFT (which they do, except for HD silicon), we should be on solid ground as there is no band structure within OF-DFT anyway.

APPENDIX B

In order to evaluate $T_{NL}[\rho]$ and its potential $\delta T_{NL}[\rho]/\delta\rho$ with a plane-wave basis under periodic boundary conditions in a linear-scaling way, we need to efficiently calculate two types of integrals,

$$P(\vec{r}) = \int \Omega[\xi(\vec{r})|\vec{r} - \vec{r}'|]f(\vec{r}')dr'^3, \quad \text{(B1)}$$

$$Q(\vec{r}) = \int \Omega[\xi(\vec{r}')|\vec{r} - \vec{r}'|]f(\vec{r}')dr'^3. \quad \text{(B2)}$$

They are more complicated than the standard convolution,

TABLE IX. Comparison of NLPS and BLPS KS-LDA bulk properties of the A17, sc, CD, bcc, and fcc phases of phosphorus. The A17 structure, i.e., black phosphorus, is the ground structure under ambient conditions (Ref. 51).

	Phosphorus	A17	sc	CD	bcc	fcc
B	NLPS	94	130	50	111	105
	BLPS	133	125	68	116	113
V_0	NLPS	13.973	13.988	18.939	13.534	13.838
	BLPS	13.957	14.700	20.149	13.642	13.768
E_0	NLPS	−180.465	0.002	0.552	0.572	0.742
	BLPS	−181.956	0.513	0.914	0.889	1.005

TABLE X. Comparison of NLPS and BLPS KS-LDA bulk properties of the α, sc, bcc, CD, and fcc phases of arsenic. α-As is the experimental ground state at low temperature (Ref. 48).

	Arsenic	α	sc	bcc	CD	fcc
B	NLPS	75	93	97	54	96
	BLPS	77	82	92	47	89
V_0	NLPS	20.282	18.981	17.764	26.288	18.015
	BLPS	20.033	18.936	18.039	25.713	18.281
E_0	NLPS	−174.015	0.183	0.499	0.511	0.612
	BLPS	−174.013	0.090	0.405	0.382	0.503

$$Y(\vec{r}) = \int \Omega(|\vec{r} - \vec{r}'|)f(\vec{r}')dr'^3, \tag{B3}$$

whose kernel only depends on the relative distance between $\vec{r}$ and $\vec{r}'$ and can be calculated efficiently with fast Fourier transforms (FFTs),

$$\widetilde{Y}(\vec{q}) = \widetilde{\Omega}(\vec{q})\widetilde{f}(\vec{q}), \tag{B4}$$

where $\widetilde{Y}(\vec{q})$, $\widetilde{\Omega}(\vec{q})$, and $\widetilde{f}(\vec{q})$ are the Fourier transforms of Y, Ω, and f. If an FFT can be used, the computational cost is near linear scaling, i.e., O[$N \ln(N)$]. However, the kernels Ω in the integrals shown in Eqs. (B1) and (B2) depend on information at either $\vec{r}$ or $\vec{r}'$, respectively, and therefore the integrals cannot be evaluated with FFTs.

To make the integral in Eq. (B1) linear scaling, we use an interpolation technique.[53] We first evaluate $P(\vec{r})$ for selected values of $\{\xi_i\}$, with $\xi_{i+1}/\xi_i = \kappa$. For each ξ_i, Eq. (B1) reduces to a standard convolution and then an FFT can be applied. After that, we interpolate $P(\vec{r})$ over the space for the actual distribution of $\xi(r)$. The accuracy of this interpolation technique is controlled by the ratio κ. The smaller the κ is, the more accurate our interpolation will be. In our cases, we find that $\kappa = 1.2$ is enough to converge total energy to better than 1 meV/atom. Using this interpolation technique, the computational cost for evaluating Eq. (B1) becomes linear scaling, O[$mN \ln(N)$], where m is the size of the set $\{\xi_i\}$ and is determined by the difference between the maximum and minimum of $\xi(r)$. N is the number of plane waves in the OF-DFT calculation. With $\kappa = 1.2$, typical values of m are around 100 for semiconducting CD Si due to large electron density fluctuations in real space; for metallic fcc Al, m is around 20.

To make the integral in Eq. (B2) linear scaling, we note that our kernel oscillates over long distance. Therefore we have to first Fourier transform Eq. (B2), obtaining

$$\widetilde{Q}(\vec{G}) = \frac{1}{\Omega} \int f(\vec{r}')e^{-i\vec{G}\cdot\vec{r}'}\widetilde{\Omega}[\vec{G}, \xi(\vec{r}')]dr'^3 \tag{B5}$$

with

$$\widetilde{\Omega}(\vec{G}, \xi(\vec{r'})) = \int \Omega[\xi(\vec{r}')|\vec{r} - \vec{r}'|]e^{-i\vec{G}\cdot(\vec{r}-\vec{r}')}dr^3.$$

We find that $\widetilde{\Omega}$ does not oscillate in reciprocal space; therefore it can be efficiently splined at the beginning of the computation. We use the cubic Hermite spline, which only requires the value and the first derivative of $\widetilde{\Omega}$. The spline is expressed as

$$\widetilde{\Omega}(\vec{G}, \xi(\vec{r}')) = \sum_{i=1}^{n-1} \theta[\xi(\vec{r}') - \xi_i]\theta[\xi_{i+1} - \xi(\vec{r}')]\Bigg[h_{00}(t_i)\widetilde{\Omega}_i$$
$$+ h_{10}(t_i)h_i \frac{\partial\widetilde{\Omega}}{\partial\xi}\Bigg|_{\xi_i} + h_{01}(t_i)\widetilde{\Omega}_{i+1}$$
$$+ h_{11}(t_i)h_i \frac{\partial\widetilde{\Omega}}{\partial\xi}\Bigg|_{\xi_{i+1}} \Bigg], \tag{B6}$$

where we have n nodes $\{\xi_i\}$, $\theta(r)$ is the Heaviside function, and h_{00}, h_{01}, h_{11}, and h_{10} are the standard Hermite basis functions,

$$h_{00}(t) = 2t^3 - 3t^2 + 1,$$

TABLE XI. Comparison of NLPS and BLPS KS-LDA bulk properties of the α, sc, bcc, fcc, and CD phases of antimony. α-Sb is the experimental ground state at low temperature (Ref. 49).

	Antimony	α	sc	bcc	fcc	CD
B	NLPS	63	67	72	71	38
	BLPS	63	63	69	68	30
V_0	NLPS	28.231	27.234	24.606	24.846	38.165
	BLPS	26.816	26.801	25.174	24.811	37.598
E_0	NLPS	−153.364	0.018	0.182	0.277	0.414
	BLPS	−148.483	0.000	0.212	0.310	0.541

TABLE XII. KS-LDA/BLPS predictions of the bulk properties for zincblende binary semiconductors: bulk modulus (B), equilibrium volume (V_0), and equilibrium total energy (E_0) per formula unit. NLPS results are in parentheses for comparison.

	B (GPa)	V_0 (Å³)	E_0 (eV)
AlP	90(89)	40.637(39.577)	−240.182(−238.867)
AlAs	80(75)	43.616(43.708)	−232.908(−232.169)
AlSb	60(57)	56.607(55.917)	−206.606(−210.839)
GaP	80(90)	37.646(37.575)	−243.079(−243.175)
GaAs	75(79)	40.634(42.169)	−235.799(−236.381)
GaSb	56(60)	52.488(52.855)	−209.697(−215.334)
InP	73(88)	46.040(44.001)	−235.722(−238.177)
InAs	65(73)	49.123(48.794)	−228.537(−231.425)
InSb	50(55)	62.908(60.890)	−202.387(−210.306)

$$h_{10}(t) = t^3 - 2t^2 + t,$$

$$h_{01}(t) = -2t^3 + 3t^2,$$

$$h_{11}(t) = t^3 - t^2,$$

with $h_i = \xi_{i+1} - \xi_i$, $t_i = (\xi(\vec{r}') - \xi_i)/h_i$, and $\widetilde{\Omega}_i$ is short for $\widetilde{\Omega}(\xi_i|\vec{r} - \vec{r}'|)$.

After inserting Eq. (B6) into Eq. (B5) and moving the summation in front of the integral, we are able to make use of an FFT to do the standard convolution first and then do the summation. With this spline technique, the computational cost of evaluating integral [Eq. (B2)] becomes $\sim O[mN \ln(N)]$, where m is the number of nodes used in the spline, which is again determined by the distribution of $\xi(r)$ in real space. Again, N is the number of plane waves.

TABLE XIII. Comparison of NLPS and BLPS KS-LDA eigenvalue band gaps (in eV) of various semiconductors (ZB: zincblende; WZ: wurtzite). "(I)" indicates an indirect band gap and "(D)" a direct band gap. HD silicon has zero band gap using the silicon BLPS.

		NLPS	BLPS
Silicon	CD	0.46 (I)	0.29 (I)
	HD	0.28 (I)	0
AlP	ZB	1.41 (I)	1.16 (I)
	WZ	1.79 (I)	1.42 (I)
AlAs	ZB	1.18 (I)	1.17 (I)
	WZ	1.51 (I)	1.33(D)
AlSb	ZB	1.12 (I)	0.74 (I)
	WZ	1.01(D)	0.66 (I)
GaP	ZB	1.38 (I)	1.08 (I)
	WZ	1.36 (I)	1.10(D)
GaAs	ZB	0.90(D)	1.13 (I)
	WZ	0.83(D)	0.99(D)
GaSb	ZB	0.44(D)	0.47 (I)
	WZ	0.19(D)	0.42 (I)
InP	ZB	1.51 (I)	1.23 (I)
	WZ	1.61(D)	1.43(D)
InAs	ZB	0.45(D)	1.26 (I)
	WZ	0.57(D)	1.55(D)
InSb	ZB	0.20(D)	1.06 (I)
	WZ	0.42(D)	1.04 (I)

In this way, we have made the evaluation of two unconventional convolutions defined in Eqs. (B1) and (B2) both almost linear scaling, which makes the evaluation of our KEDF and its potential very efficient, albeit with an extra prefactor of m. However, that prefactor can be eliminated by implementing these two numerical techniques in parallel.

[1] E. A. Carter, Science **321**, 800 (2008).

[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[4] Y. A. Wang and E. A. Carter, in *Theoretical Methods in Condensed Phase Chemistry*, edited by S. D. Schwartz (Kluwer, Dordrecht, 2000), p. 117.

[5] L. Hung and E. A. Carter, Chem. Phys. Lett. **475**, 163 (2009).

[6] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).

[7] E. Smargiassi and P. A. Madden, Phys. Rev. B **51**, 117 (1995).

[8] K. M. Carling and E. A. Carter, Modell. Simul. Mater. Sci. Eng. **11**, 339 (2003).

[9] G. Ho, M. T. Ong, K. J. Caspersen, and E. A. Carter, Phys. Chem. Chem. Phys. **9**, 4951 (2007).

[10] B. J. Jesson and P. A. Madden, J. Chem. Phys. **113**, 5924 (2000).

[11] C. Huang and E. A. Carter, Phys. Chem. Chem. Phys. **10**, 7109 (2008).

[12] B. Zhou and E. A. Carter, J. Chem. Phys. **122**, 184108 (2005).

[13] B. Zhou, V. L. Ligneres, and E. A. Carter, J. Chem. Phys. **122**, 044103 (2005).

[14] E. Chacón, J. E. Alvarellos, and P. Tarazona, Phys. Rev. B **32**, 7868 (1985).

[15] L.-W. Wang and M. P. Teter, Phys. Rev. B **45**, 13196 (1992).

[16] P. García-González, J. E. Alvarellos, and E. Chacón, Phys. Rev. B **53**, 9509 (1996).

[17] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **60**, 16350 (1999); **64**, 089903(E) (2001).

[18] J.-D. Chai and J. D. Weeks, Phys. Rev. B **75**, 205122 (2007).

[19] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **58**, 13465 (1998); **64**, 129901(E) (2001).

[20] F. Perrot, J. Phys.: Condens. Matter **6**, 431 (1994).

[21] J. P. Perdew and L. A. Constantin, Phys. Rev. B **75**, 155109 (2007).

[22] D. Garcia-Aldea and J. E. Alvarellos, J. Chem. Phys. **127**, 144109 (2007).

[23] A. J. Thakkar, Phys. Rev. A **46**, 6920 (1992).

[24] J. Lindhard, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **28**, 8

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

[26] G. Srinivasan, Phys. Rev. **178**, 1244 (1969).

[27] Z. H. Levine and S. G. Louie, Phys. Rev. B **25**, 6310 (1982).

[28] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[29] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[30] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[31] R. M. Pick, M. H. Cohen, and R. M. Martin, Phys. Rev. B **1**, 910 (1970).

[32] L. H. Thomas, Proc. Cambridge Philos. Soc. **23**, 542 (1927).

[33] E. Fermi, Rend. Accad. Naz. Lincei **6**, 602 (1927).

[34] E. Fermi, Z. Phys. **48**, 73 (1928).

[35] C. F. v. Weizsäcker, Z. Phys. **96**, 431 (1935).

[36] R. W. Brankin, I. Gladwell, and L. F. Shampine, SOFTREPORT 92 (1992).

[37] G. Ho, V. L. Ligneres, and E. A. Carter, Comput. Phys. Commun. **179**, 839 (2008).

[38] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, and D. C. Allan, Comput. Mater. Sci. **25**, 478 (2002).

[39] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[40] M. Fuchs and M. Scheffler, Comput. Phys. Commun. **119**, 67 (1999).

[41] B. Zhou, Y. A. Wang, and E. A. Carter, Phys. Rev. B **69**, 125109 (2004).

[42] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[43] F. D. Murnaghan, Proc. Natl. Acad. Sci. U.S.A. **30**, 244 (1944).

[44] M. J. Gillan, J. Phys.: Condens. Matter **1**, 689 (1989).

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

[46] P. Cortona, Phys. Rev. B **44**, 8454 (1991).

[47] M. Bernasconi, G. L. Chiarotti, and E. Tosatti, Phys. Rev. B **52**, 9988 (1995).

[48] L. F. Mattheiss, D. R. Hamann, and W. Weber, Phys. Rev. B **34**, 2190 (1986).

[49] L. M. Falicov and P. J. Lin, Phys. Rev. **141**, 562 (1966).

[50] *Smithells Metals Reference Book*, edited by E. A. Brandes and G. B. Brook, 7th ed. (Elsevier, New York, 1998).

[51] A. Brown and S. Rundqvist, Acta Crystallogr. **19**, 684 (1965).

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

[53] D. J. Singh, Phys. Rev. B **48**, 14099 (1993).