Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 224 (2007) 824–835
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Surface passivation optimization using DIRECT

Peter A. Graf [a,*], Kwiseon Kim [a], Wesley B. Jones [a], Lin-Wang Wang [b]

[a] National Renewable Energy Laboratory (NREL), Golden, CO 80401, United States
[b] Computational Research Division, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA 94720, United States

Received 14 December 2005; received in revised form 6 October 2006; accepted 19 October 2006
Available online 14 December 2006

Abstract

We describe a systematic and efficient method of determining pseudo-atom positions and potentials for use in nanostructure calculations based on bulk empirical pseudopotentials (EPMs). Given a bulk EPM for binary semiconductor X, we produce parameters for pseudo-atoms necessary to passivate a nanostructure of X in preparation for quantum mechanical electronic structure calculations. These passivants are based on the quality of the wave functions of a set of small test structures that include the passivants. Our method is based on the global optimization method DIRECT. It enables and/or streamlines surface passivation for empirical pseudopotential calculations.

Keywords: Global optimization; Electronic structure; Nanostructure; Surface passivation; Pseudopotentials

1. Introduction

For computational nanoscience to become more routine, both in the context of applications and in the exploration of new physics, all the pieces of the process must be automated. Here we present automation of one such piece – surface passivation in the context of the empirical pseudopotential method – using a state-of-the-art global optimization scheme, DIRECT [1], especially suited to this problem.

The empirical pseudopotential method has been very successful in studying the electronic structure of nanostructures [2], and there are mature programs to fit bulk EPMs (throughout, we will use "EPM" to refer to both the method and the actual pseudopotentials). However, one bottleneck in applying the empirical pseudopotential method to nanostructures is taking into account surface effects. Our method will improve this scenario for the investigation of well passivated nanostructures.

In experiments, the dangling bonds at the surface of a semiconductor nanostructure are passivated either by other semiconductors or by organic ligands [3]. Although the surface conditions of the structures vary, photoluminescence (PL) measurements of appropriately passivated nanostructures reveal that the emission originates from interior, "bulklike" states [4]. These observations suggest that an approach to passivating a simulated nanostructure would be to attach "pseudo-atoms" to each dangling bond [5–7].

* Corresponding author.
E-mail address: peter_graf@nrel.gov (P.A. Graf).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2006.10.033

Following the work of [5,6], for each dangling bond, characterized by surface atom element and number of dangling bonds originating from that atom, a three parameter pseudo-atom potential is added to the potential of the nanostructure. For a binary semiconductor compound nanostructure such as CdSe, this results in a 12 real parameter optimization problem. The advantage of the pseudo-atom potential approach is that the potentials are determined once for an ensemble of small test structures and transferable to other structures of the same material.

As the PL data suggest, the surface states near the bandgap are eliminated in a well passivated nanostructure. Thus, the desired effect of the pseudo-atom (i.e., the solution of the 12 parameter optimization problem) is to obtain containment *inside* the nanostructure of wave functions corresponding to eigenvalues in and around the bandgap. In [5] the parameters are chosen by forcing the band edge states to the interior of flat Cd- and Se-terminated surfaces. In [6], the parameters are chosen so that the surface density of states of planar InP surfaces matches that of PL data and local density functional theory calculations.

The optimization of the 12 parameters has heretofore been carried out by trial and error based on physical intuition. Although successful in several circumstances, this approach has three major shortcomings. First, physical intuition of a researcher varies by researcher. Second, it is extremely time consuming; finding acceptable surface passivation parameters requires extensive human intervention, and the parameter space is large. Third, it is not robust; there is no guarantee that the parameter space has been sufficiently sampled and explored. Our method addresses these issues by defining a 12 real parameter optimization problem (an objective function and associated constraints) and solving it using the global optimization method DIRECT. DIRECT has been found effective for problems such as ours in which the search domain consists of real vectors within a bounded region. Our method is a way of enabling *bulk* EPMs for any material to be used to calculate the electronic states of structures having *surfaces*. As such, we offer an improvement in the methodology of surface passivation parameter selection for increased reliability and efficiency, especially in enabling work on new materials.

This work is in the context of electronic structure calculation for crystalline semiconductor nanostructures. The electronic structure theory used here is density functional theory in the local density approximation, solved non-selfconsistently via the empirical pseudopotential method. The specific systems reported on here are tetrahedrally bonded binary semiconductor compounds. The method, however, is not restricted to such systems, and could be applied in the same way to elemental semiconductors (e.g. Si, Ge), alloys (e.g. InGaP), and other crystal structures (e.g. PbSe in the rocksalt crystal structure).

In Section 2 we describe in detail our formulation of surface passivation as an optimization problem and our application of DIRECT to it. In Section 3 we present results for the passivation of two materials. In Appendix A we discuss some of the mathematical underpinnings of the method.

2. Method

Here we describe in detail our passivation method. First we define the 12 real parameters of the optimization space that describe the pseudo-atoms and the bounds on the domain that is explored. Next we describe the objective function, including its dependence on the electronic states of five test structures, constraints on the eigenvalues of these states, and a measure of quality of these states. Finally, we describe the sequence of steps in the process of finding the parameters for the pseudo-atoms.

2.1. Parameter space

Following the work of [5,6], for each dangling bond a pseudo-atom passivant with a Gaussian potential is added at some length γd along the ideal bond of length d between the surface atom and the missing atom. The potential of the pseudo-atom is given by the expression

$$v(\mathbf{r}) = \alpha e^{-(|\mathbf{r}-\mathbf{R}|/\sigma)^2}. \tag{1}$$

Here $\mathbf{R} = \mathbf{R}(\gamma)$ is the location of the passivant, and α and σ represent the amplitude and width of the Gaussian potential, respectively. All the constants are in atomic units; α is in hartree, and σ is in Bohr (γ is unitless). The set $\{\alpha, \sigma, \gamma\}$ are parameters that are fit to produce desired effects.

Different passivants are required for different surface atom elements and numbers of dangling bonds. We deal exclusively with binary semiconductor compounds on ideal or "relaxed" crystal lattice sites, for example, CdSe and InP. They form tetrahedrally bonded zincblende or wurtzite structures. We index the four passivants as follows:

1. cations with 1 dangling bond ("c1"),
2. cations with 2 dangling bonds ("c2"),
3. anions with 1 dangling bond ("a1"),
4. anions with 2 dangling bonds ("a2").

Atoms with more than two dangling bonds are removed prior to surface passivation. This results in a total of 12 real parameters to fit. Further details of the test structures are discussed in the following section.

The electronic structure calculation is performed in Fourier space, and thus the real space parameters have associated q-space parameters. Transforming Eq. (1) to Fourier space [8], we obtain

$$v(\mathbf{q}) = \alpha\pi^{1.5}\sigma^3 e^{i\mathbf{q}\cdot\mathbf{R}} e^{-(\sigma|\mathbf{q}|/2)^2}. \tag{2}$$

The effect of γ is introduced in real space through the location of the passivant in the input configuration. Now let $a \equiv \alpha\pi^{1.5}\sigma^3$, $b \equiv \sigma/2$, $c \equiv \gamma$. This defines a mapping between the parameters $\{\alpha, \beta, \gamma\}$ in real space and $\{a, b, c\}$ in q-space. Our system seeks the optimal set of 12 parameters $\xi \equiv \{a_i, b_i, c_i\}_{i=1}^4$, where the passivant potential for passivant type i is defined by

$$v_i(\mathbf{q}) \equiv a_i e^{i\mathbf{q}\cdot\mathbf{R}_i} e^{-(b_i|\mathbf{q}|)^2}, \tag{3}$$

with the passivant located a distance c_i along the ideal tetrahedral bond from the passivated atom to the passivant.

2.2. Objective function

To continue we need a precise measure of the quality of passivation represented by a set of parameters, i.e. we must formulate the objective function of our optimization problem. The goal is the passivation of large nanostructures. These may contain many instances of all the passivants, which would be chosen such that when we calculate the near band edge electronic states of the nanostructure, those states would not be surface states. However, the calculation of electronic states of such structures is too computationally intensive to be used in the passivation optimization. Therefore, following [5,6], we consider an ensemble of small structures such that the pseudo-atoms found to passivate the ensemble of small test structures will be used to passivate nanostructures of arbitrary shape. No single structure contains all the passivants, but together all are represented. In addition, small structures have larger quantum confinement, which provides stronger tests for surface passivation. Further, the set of structures allows critical points in the Brillouin zone (e.g. K, L, X, W) to be represented so that the surface states occurring at these point in the Brillouin zone of a real nanostructure can be accurately produced.

For each test structure we require the conduction band minimum (CBM) of the passivated structure to be above that of the corresponding bulk solid. (Knowledge of the bulk band edges allows us to find the CBM; it is the first state above a known "reference energy" inside the bulk bandgap.) Similarly, we need the valence band maximum (VBM) to be below the bulk level. These constraints represent the effect of quantum confinement. Formally, we can write this constraint as $g(\xi) > 0$ for an appropriately defined g.

The measure of quality of a passivation is a measure of the extent to which the CBM and VBM states are *not* on the surface. To make this measure precise, we integrate the squared wave function across the interior ("core") of a structure. Note that this involves explicit consideration of the wave function of the CBM and VBM states, not just the eigenenergies.

We consider two kinds of structures – spherical quantum dots and oriented slabs. The slabs represent structures with planar terminating surfaces. We define the core of a spherical dot of radius r to be the central sphere of radius $\frac{3}{4}r$. Note that the volume is $(\frac{3}{4})^3 \sim 0.4$ of the dot. For planar slabs, we begin with a supercell having a certain orientation. For example, a [1 1 0] oriented slab has a z-axis in the [1 1 0] direction. The supercell is then

filled halfway along this axis with the material in question, and the rest of the supercell is left as vacuum (i.e. empty). For a slab of thickness d, the supercell has z-axis of length $2d$. The core is defined to be the inner $\frac{1}{2}d$ of the slab. In our system there are two spherical dots of radius ~5.5 Å. "Dot 1" is cation centered and requires passivants c1 and a2 (see Fig. 1). "Dot 2" is anion centered and requires passivants a1 and c2. And there are three slabs of thickness ~20 Å. These slabs, of orientations [0 0 1], [1 1 0], and [1 1 1], have 16, 8, and 22 atomic layers, respectively, and an equal amount of vacuum. They require passivants c2 and a2, c1 and a1, and c1 and a1, respectively. Table 1 lists the atom counts for our test structures.

These structures are created in the following way. First, a large supercell of the appropriate crystal structure – here, zincblende – is created. Now atoms lying outside the dimensions of the structure we are creating (e.g. for a dot centered in the middle of the large supercell, atoms whose distance from the center is greater than the chosen dot radius) are removed. This leaves atoms at the surface of the dot that are not fully bonded (i.e. some of whose coordinating atoms have been removed). These are the surface atoms we must passivate. For each such atom, we count how many dangling bonds it has (i.e. how many of the atoms it was bonded to in the large supercell were removed). If an atom has more than two dangling bonds, we remove it to prevent overly rough surfaces. Removal of an atom can result in a neighbor of that atom having more the two dangling bonds, so we repeat the process until there are no such atoms left. If an atom has one or two dangling bonds, we add the appropriate passivant atom c1, c2, a1, a2 described above, that is, at the distance along the missing bond specified by the c_i parameters. Thus, for example, dot1 requires only passivants c1 and a2 because the atoms that lie on the surface are all either cations with one dangling bond or anions with two dangling bonds. Note, too, that the distances c_i are parameters we are optimizing, thus the test structures are regenerated at

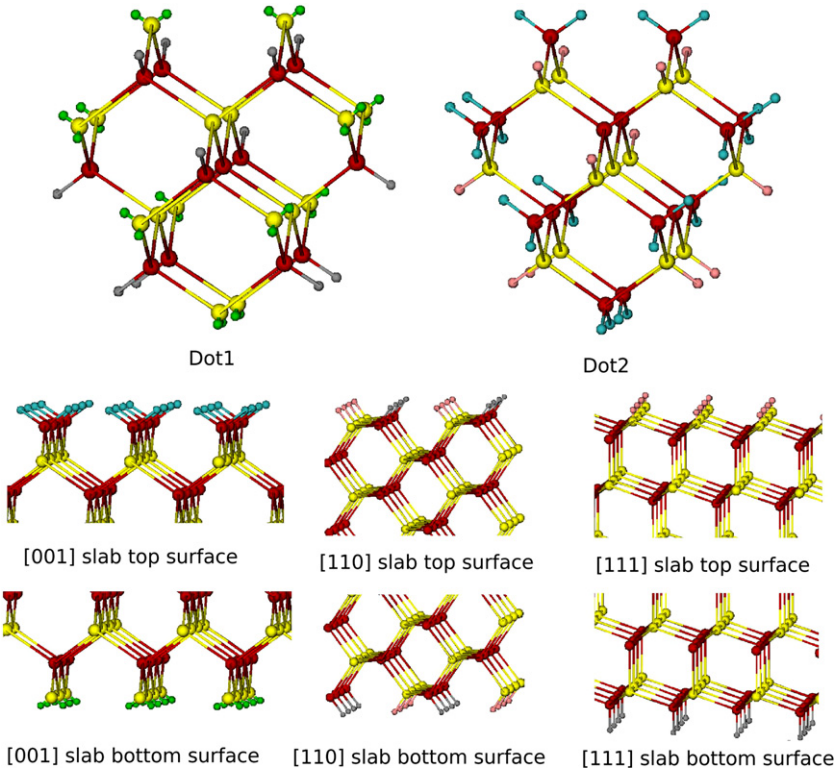

Fig. 1. The test structure surfaces. The large, dark (in color, red) circles are Cd atoms. The large, light (in color, yellow) circles are Se atoms. The small circles are the passivant atoms. (In color, c1 is silver, c2 is cyan, a1 is pink, and a2 is green.) The entire dot1 and dot2 structures are shown. For the slabs, the top and bottom surfaces are shown. The slabs continue infinitely in the plane orthogonal to the surface normal. Note the surface orientations and surface atom types determine the passivants required. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Atom counts of test structures

Test structure	Number of cations	Number of anions	Number of passivants
Dot 1	13	16	36
Dot 2	16	13	36
[0 0 1] slab	16	16	8
[1 1 0] slab	16	16	8
[1 1 1] slab	44	44	8

every iteration of the search. Fig. 1 illustrates the surface environments and resulting passivant atoms (for a particular set of passivant atom distances) used for the five test structures.

The sufficiency of the objective function described here – geared toward preventing band edge states from being surface states – lies in the results of its use rather than in complete theoretical justification; this basic approach has a successful history in the literature [5,6]. Furthermore, explicit modeling of the organic molecules at the surface of a nanostructure is outside the scope of the electronic structure method used here. We reiterate that our main contribution is automation of this approach rather than justification of it.

In summary, the objective function $f(\xi)$ is the geometric average of the core wave function densities of CBM and VBM over all the test structures. It is a number between zero and one representing the proportion of the wave function contained, on average, in the core.

2.3. Domain of ξ

Recall that $\xi \equiv \{a_i, b_i, c_i\}_{i=1}^{4}$ defines a certain passivation. To further specify the optimization problem, we must state the domain of ξ. The distances c_i are proportions of the ideal bond distance. It is reasonable to assume that the passivant is no farther than the missing atom would be, and that it is not extremely close to the atom it passivates. Therefore we constrain the c_i to the interval [0.1, 1.0]. By considering that the Gaussian potential must decay to zero as *both* $|q|$ and $|r|$ become large, we arrive at lower and upper bounds for the width parameter b_i of 0.1 and 1.5, respectively. The amplitude parameter differs between cation and anion passivant, with the cation having positive amplitude and the anion negative. This choice is motivated by the physics. Cations have lower electronegativity than anions, so a positive ligand passivation potential will ensure that they will lose electrons, whereas the anions have higher electronegativity, so a negative passivation potential will ensure that they gain extra electrons. Further, the maximum amplitude is determined by the magnitudes of the corresponding bulk pseudopotentials. For convenience, we rescale the amplitudes of all the passivants for all the EPMs studied by a constant value of 12.028 (which is $v(q=0)$ for the existing CdSe passivation), transforming the amplitude parameters a_i into "weights" w_i having $|w_i| \sim 1$. Together the domain of ξ is a hyperrectangle we refer to as $\mathscr{D}$. Specifically, we have chosen the following limits:

$$\begin{aligned}
0.1 &\leqslant \text{ cation weight } w_i \leqslant 2.0, \quad i \in \text{cation},\\
-2.0 &\leqslant \text{ anion weight } w_i \leqslant -0.1 \quad i \in \text{anion},\\
0.1 &\leqslant \text{ width } b_i \leqslant 1.5,\\
0.1 &\leqslant \text{ distance } c_i \leqslant 1.0.
\end{aligned}$$

2.4. The DIRECT optimization method

We have thus formulated our surface passivant optimization problem as

$$\max_{\xi \in \mathscr{D}} \quad f(\xi) \quad s.t. \quad g(\xi) > 0.$$

This is a nonlinear constrained global optimization problem.

However, the structure of $\mathscr{D}$ – a hyperrectangle of real vectors – makes this problem especially suited to the global optimization method DIRECT. DIRECT is a deterministic algorithm for global search introduced in

1993 [1]. The name is a mnemonic for 'divide' and 'rectangle' because it involves systematically dividing rectangles in its search for a global optimum. In particular, the method is founded on an explicit recognition of the tradeoff between global and local search. In DIRECT, the search begins at the center of the hyperrectangle $\mathscr{D}$. Subsequent search is directed toward larger unexplored rectangles (representing global search) as well as areas that have promising objective function values (representing local search). The simultaneous exploration of the entire space of tradeoffs between global and local search is a powerful innovation built into the method. It is related to finding and exploring the entire "Pareto Front" of the global/local tradeoff [9]. This connection, and further details of the method, are pursued further in Appendix A. We emphasize that DIRECT is a true *global* optimization method; it does not require an initial estimate of the solution to find the global optimum. Thus our method can start with minimal input by the physicist.

DIRECT is one of many optimization methods. For our purposes it has the advantages of being (i) global and (ii) deterministic; beyond a good solution, we also get a robust, comprehensive *scan* of the parameter space. This is important, as the objective function landscape for this problem is largely unknown. The dimensionality of our problem is also well suited to DIRECT. DIRECT appears to be a good candidate for problems of dimension up to approximately 20–30. This size space would encompass most passivation search problems, such as those arising from considering a different crystal structure with coordination number greater than four.

2.5. Implementation

Our implementation is built on a parallel Fortran implementation of DIRECT available at http://www4.ncsu.edu/eos/users/c/ctkelley/www/optimization_codes.html. Our system involves many components. Specifically, it is built around the electronic structure program PESCAN [10,11]. The method of evaluating the objective function for a given set of passivation parameters involves the following steps: Generate atomic configurations of dots of chosen radii and slabs with chosen thicknesses with the passivation distances and passivant weights specified. Generate four pseudopotential files, one for each passivant, using the width parameters. For each structure, find the eigenvalues and eigenstates for the CBM and VBM. Check whether the constraint that the CBM and VBM straddle the bulk band edges is satisfied. If the constraint is not met, set the objective function to 0. Note we are thus enforcing our constraint through a penalty function. This is really our only choice, because the evaluation of the constraint function is a complicated, expensive, nonlinear function of the passivant parameters. Integrate the charge density of these states over the "core" of the structure. If the core density is below a user specified threshold, we have found a surface state; reject this passivation by setting the objective function to 0. This is a penalty function enforcing the constraint that the CBM and VBM should not be a surface state for *any* of the test structures. Finally, return the geometric average of the two core integrations over all the test structures to DIRECT as the objective function value.

To use this system to generate a passivation scheme for a bulk EPM of a given material, users need only supply the necessary components of the bulk EPM, along with a short input file that describes the particular settings used by the electronic structure solver for this EPM.

3. Results

Here we discuss the results obtained for CdSe and InP. We divide the discussion into two parts corresponding to the optimization process itself and the actual passivations obtained.

3.1. Optimization

DIRECT is a nearly parameterless global optimization method. Aside from the upper and lower bounds of the search variables, discussed above, there are few parameters to work with, consisting mainly of a set of tolerance and stopping criteria parameters. Because the search quickly plateaus to small fluctuations in the objective function near the optimum value, we generally turn off all stopping criteria.

Fig. 2 shows the typical progress of the optimization: a short period of rapid improvement followed by a longer plateau of slow improvement. This is typical of global optimization methods [12,13].

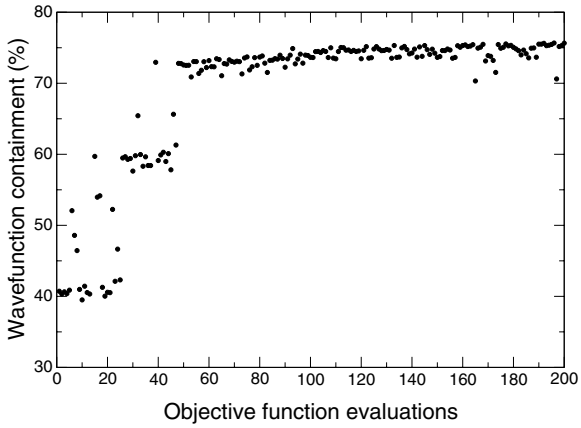

Fig. 2. Typical progress of an optimization run. Each dot represents one set of the 12 passivation parameters. The x-axis represents the progress of an optimization run. The y-axis is the objective function value for each set of parameters. It is the average containment of the squared wave function of the conduction band minimum and valance band maximum in the core region of the five test structures. Typically one observes initial rapid improvement followed by gradual refinement of good solutions. This suggests that a wide range of acceptable parameters exists, as has been noted in [6]. Also, it suggests that we may add additional criteria to our objective function if we need a passivation with properties beyond band edge wave function containment.

3.2. Passivation

We consider two materials, CdSe and InP. First, we have compared our CdSe results with the passivation parameters of [5]. Our optimization scheme produces a set of parameters that, according to our objective function, is slightly better. Specifically, using their parameters, the wave functions are 72% contained in the "core", and using the parameters found by DIRECT, the wave functions are 78% contained. The small improvement is not really the point. More important is the fact that our parameters were found by a systematic method involving no human intervention. As also shown in Fig. 2, there are multiple solutions of similar quality. This situation is acceptable because as long as the surface is well passivated, the electronic structure inside the nanostructure is independent of the details of the passivation [6]. Fig. 3 shows a 447 atom (including passivants)

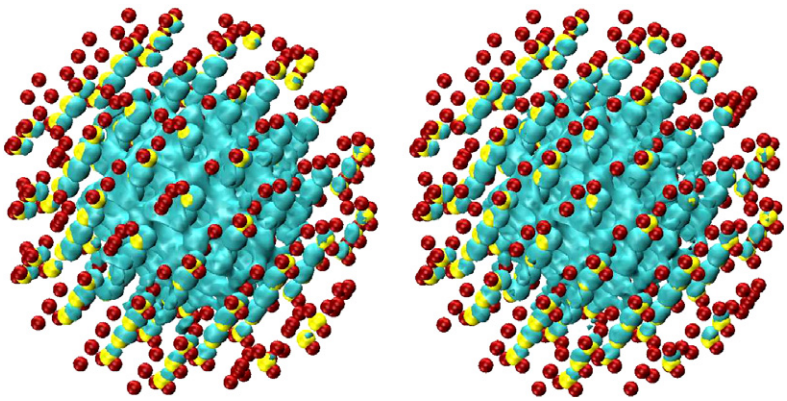

Fig. 3. Four hundred and forty-seven atom CdSe dots passivated using our results (left) and those of Ref. [5] (right). The small dark grey (red) and light grey (yellow) spheres are Cd and Se atoms, respectively. For clarity, the passivation atoms are not shown. The grey (aqua) surfaces are isosurfaces of the conduction band minimum wave function squared. Shown is the isosurface corresponding to the isovalue above which is contained 90% of the wave function. The contours for higher isovalues are *inside* the surface shown. Note the similar, adequate containment of the wave functions in both cases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

CdSe dot passivated with the parameters we found and those of [5], respectively. Fig. 4 shows an example of the passivation achieved for the [1 1 1] oriented slab.

Second, we have optimized passivants for the InP semiempirical pseudopotential of [6]. As Table 2 shows, our passivation is again only slightly better *on average* than the hand tuned one. However, it is clear based both on the core density and visualization of the squared wave function that the [1 1 1] slab CBM state for the passivation from [6] is a surface state. By contrast, our optimization scheme ensures that the optimum passivation will not result in CBM or VBM states which are surface states for *any* of the test structures by explicitly penalizing passivations that generate such states.

Also, we have verified that the surface states were forced far enough out of the bandgap that there are several states above the CBM and below the VBM that are *also* not surface states. In fact, our objective function considers as many states above and below the gap as the user desires.

Table 2 summarizes our results. We note some regularity:

- The optimal anion passivants are typically closer to the atoms they passivate than are the cation passivants.
- Potentials of optimal passivants of anions with *two* dangling bonds are typically more spread out in q space (small width parameter) with low amplitude. They are usually especially close to the anions they passivate.

This regularity is not expected a priori; it is the outcome of our optimization and depends on the material properties and the bulk pseudopotentials that describe the material. Passivations of other materials may or may not have these characteristics.

Finally, consider Fig. 5, where we have plotted the bandgap as a function of radius for a series of quantum dots. We see the expected confinement-induced increase in bandgap.

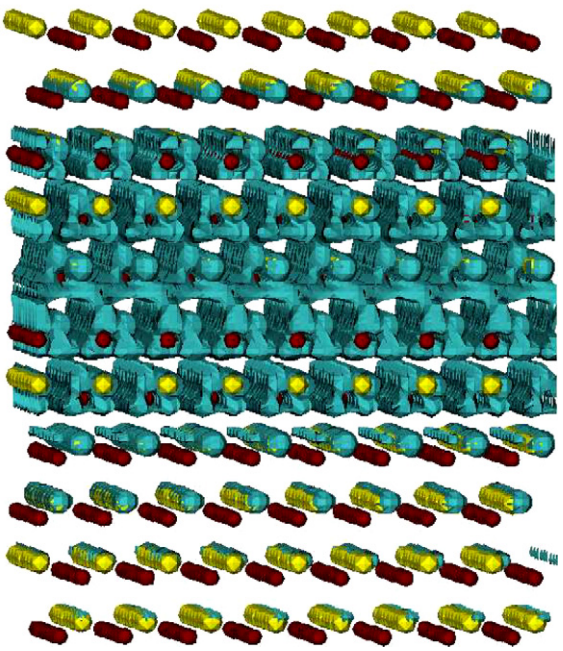

Fig. 4. [1 1 1] oriented CdSe slab passivated by our method, and the conduction band minimum wave function squared. The colors are as in Fig. 3. Again, the passivant atoms are *not* shown. We see that at this, the 85% isovalue, the isovalue surface is fully contained in the slab. Thus more than 85% of the wave function is contained in the interior of the slab. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Passivant parameters and objective function values found by our method and those used in [5 and 6]

	CdSe passivants		InP passivants	
	Current	From [5]	Current	From [6]
w_1	1.89	1.00	1.68	1.18
w_2	1.26	1.00	1.68	1.90
w_3	−1.26	−0.60	−1.68	−0.93
w_4	−0.84	−0.60	−1.05	−0.78
b_1	0.64	0.75	0.80	0.45
b_2	0.64	0.75	0.64	0.70
b_3	0.33	0.75	0.33	0.50
b_4	0.33	0.75	0.80	0.75
c_1	0.45	0.55	0.35	0.25
c_2	0.55	0.55	0.25	0.50
c_3	0.25	0.25	0.25	0.25
c_4	0.28	0.30	0.25	0.25
Function	77.6	71.2	71.8	65.1

The objective function ('Function') is the average percentage of the VBM and CBM wave function contained in the core of the five test structures. w_i are the weights (amplitudes), b_i are the widths of the Gaussians, and c_i are the distances along the ideal bond. The index "i" denotes the surface atom types defined in Section 2. Though the *average* core density is similar for our passivation and the hand tuned passivation of [6], at least one of the band edge states for the passivation in [6] is a surface state.

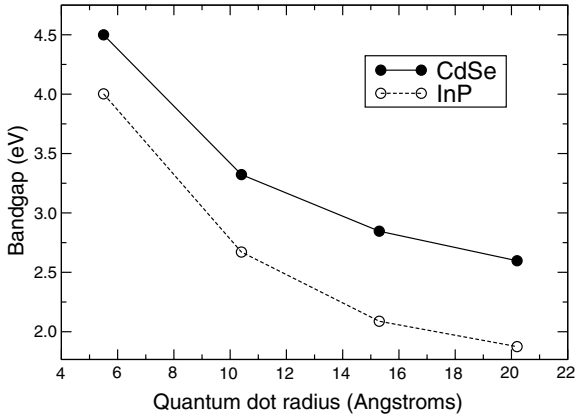

Fig. 5. Bandgap versus radius for a series of quantum dots passivated by our method. Note the expected quantum-confinement induced increase in bandgap with reduced dot size.

4. Conclusion

In this paper we have presented an automated methodology for generating surface passivations for bulk empirical pseudopotentials. Incorporated into a larger computational nanoscience infrastructure, our work represents a much needed improvement in the usability of the empirical pseudopotential method for electronic structure calculations of nanostructures.

Acknowledgments

We gratefully acknowledge discussions with A. Zunger, A. Canning, G. Bester, A. Franceschetti, S. Dudiy, G. Trimarchi, J. Langou. The three-dimensional images were produced with the molecular visualization program Visual Molecular Dynamics (VMD) [14]. This work used computing resources at the NREL Computational Sciences Center and is supported by the US DOE-SC-ASCR-Mathematical, Information and

Computational Sciences program under Lab03-17 Theory Modeling in Nanoscience initiative under Contract No. DE-AC36-99GO10337, and for LBNL under Contract No. DE-AC03-76SF00098.

Appendix A. DIRECT and the optimal tradeoff of local and global search

Here we further pursue the connection between DIRECT and the concept of a Pareto Front. To our knowledge this connection has not been explicitly made before. First, however, we further clarify the search mechanism of the DIRECT algorithm. As described in [1], at each iteration of the search, there is a set of rectangles of varying sizes that have been sampled (the objective function value at the center of the rectangle is known). An iteration then involves the following steps:

1. Identify a set of "potentially optimal" rectangles (see definition below).
2. Evaluate the objective function at specific points in these rectangles.
3. Based on the objective function values, divide these rectangles for further search.

These steps lead to a new set of rectangles to consider for the next iteration. Fig. A.1 illustrates the division of the domain $\mathscr{D}$ into rectangle of varying sizes for a particular cross section of variables for our CdSe passivation optimization.

We now consider the connection between DIRECT, multicomponent optimization, and the optimal tradeoff of global and local search. In multicomponent optimization, with simultaneous objective functions that we cannot combine into a single objective function by, for example, a linear combination, the best one can do is the so called *Pareto front*. This is defined to be the set of objective function values representing points x in the space of search variables that "cannot be improved on" in the following sense: Given n objective functions $\{f_i\}_{i=1}^n$, all of which we seek to maximize, the *Pareto optimal* set is defined to be those points x for which there is *no* y such that both

(i) $f_i(y) \geqslant f_i(x)$ for all i, and
(ii) $f_j(y) > f_j(x)$ for some j.

The *Pareto front* is the set of objective function values of the members of the Pareto optimal set.

The connection with DIRECT is through the method of choosing which rectangles to subdivide. At each stage, the entire set of points at which the objective function has been evaluated, and the rectangles they

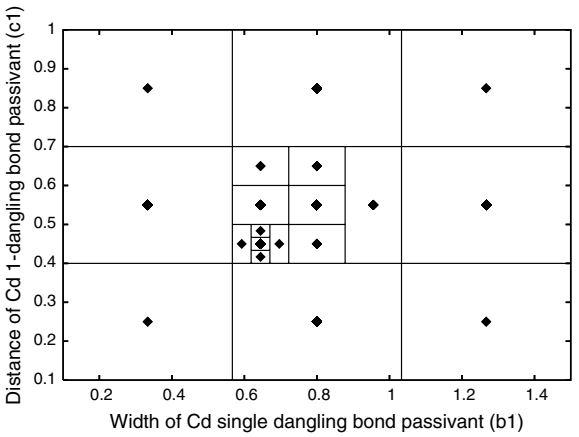

Fig. A.1. A two-dimensional cross section of DIRECT's subdivision of the search domain $\mathscr{D}$. The rectangles and diamonds at their centers represent regions and sampled points at which we have evaluated the objective function. The algorithm proceeds by preferentially subdividing rectangles in which higher quality solutions have been found. Thus the smallest rectangle is where the best solution has been found. At the next iteration, the algorithm will subdivide the *potentially optimal* rectangles; which are, roughly, the rectangles of each size with the best objective function value (see text).

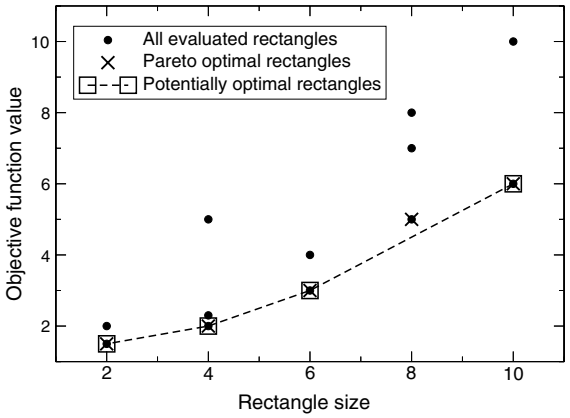

Fig. A.2. Illustration (for simulated data) of how DIRECT chooses which rectangles to further subdivide, and connection to concept of Pareto front. The small circles represent all the rectangles that have been searched. The squares are *potentially optimal* rectangles on the convex hull (dashed line), which will be further searched in the next iteration. The crosses are the rectangles in the *Pareto optimal* set for the multicomponent optimization problem of minimizing functional value and maximizing rectangle size. Note the rectangle in the Pareto optimal set (symbol: X) that is *not* potentially optimal (symbol: square).

represent, are considered. Let $f(x)$ be a function we are trying to minimize (DIRECT is stated and implemented in terms of minimization, so we switch our orientation here), and let z_i be the center of rectangle i of volume d_i. A rectangle j is *potentially optimal* if there is some K such that

(i) $f(z_j) - Kd_j < f(z_i) - Kd_i$ for all i, and
(ii) $f(z_j) - Kd_j < f_{\min} - \epsilon|f_{\min}|$.

Here, $f_{\min}$ is the best function value found so far, and ϵ is a parameter that can be tuned to prevent the system from subdividing extremely small rectangles. Graphically, the meaning of this definition is that the potentially optimal rectangles are those that lie on the lower right portion of the convex hull of all rectangles searched, considered in a coordinate system where the x-axis is the rectangle size and the y-axis is the objective function value at the center of the rectangle (see Fig. A.2).

Now consider the family of functions $f_K(z_j) = -Kd_j$, minimized for large rectangles. Then the potentially optimal rectangles are (modulo a technical distinction – the version of DIRECT explored in [15] is *exactly* searching the Pareto front at each iteration) just those that lie on the Pareto front of the multicomponent optimization problem

$$\min f \quad \text{and} \quad \min f_K.$$

And these problems are precisely the classic tradeoff between local refinement (f is minimized for small rectangles where a good function value is already known) and global exploration (f_K is minimized for large rectangles within which known solutions may be of lower quality, but which may yield high quality solutions on further exploration). Fig. A.2 illustrates the connection. The search performed by DIRECT is thus seen – in the precise sense of Pareto optimality – to be optimally balanced between local refinement and global exploration.

References

[1] D.R. Jones, C.D. Perttunen, B.E. Stuckman, J. Optimiz. Theory Appl. 79 (1) (1993) 157.
[2] A. Zunger, Phys. Stat. Sol. (b) 224 (2001) 727.
[3] P. Castrillo, D. Hessman, M.E. Pistol, S. Anand, N. Carlsson, W. Seifert, L. Samuelson, Appl. Phys. Lett. 67 (1995) 1905; O.I. Micic, C.J. Curtis, K.M. Jones, J.R. Sprogue, A.J. Nozik, J. Phys. Chem. 98 (1994) 4966.
[4] O.I. Micic, J.R. Sprogue, Z.H. Lu, A.J. Nozik, Appl. Phys. Lett. 68 (1996) 3152.
[5] L.W. Wang, A. Zunger, Phys. Rev. B 53 (15) (1996) 9579.
[6] H. Fu, A. Zunger, Phys. Rev. B 55 (3) (1997) 1642.

[7] X. Huang, E. Lindgren, J.R. Chelikowsky, Phys. Rev. B 71 (2005) 165328.
[8] We use the definition $v(q) \equiv \int v(r) \exp(-iq \cdot \mathbf{r})\, d\mathbf{r}$. Note this is a three dimensional integration.
[9] K. Deb, Evol. Comput. 7 (3) (1999) 205.
[10] L.-W. Wang, A. Zunger, J. Chem. Phys. 100 (1994) 2394.
[11] A. Canning, L.-W. Wang, A.J. Williamson, A. Zunger, J. Comput. Phys. 160 (2000) 29.
[12] K. Kim, P.A. Graf, W.B. Jones, J. Comput. Phys. 208 (2005) 735.
[13] A.E. Eiben, J.E. Smith, Introduction to Evolutionary Computing, Springer, 2003.
[14] J. He, L. Watson, N. Ramakrishnan, C. Shaffer, S. Verstak, J. Jian, K. Bae, W. Tranter, Comput. Optimiz. Appl. 23 (2002) 5.
[15] W. Humphrey, A. Dalke, K. Schulten, J. Molec. Graphics 14 (1) (1996) 33.