

Advances and applications in the FIREBALL *ab initio* tight-binding molecular-dynamics formalism

Review Article

James P. Lewis^{*1}, Pavel Jelínek², José Ortega³, Alexander A. Demkov⁴, Daniel G. Trabada³, Barry Haycock¹, Hao Wang¹, Gary Adams⁵, John K. Tomfohr⁵, Enrique Abad³, Hong Wang¹, and David A. Drabold⁶

¹Department of Physics, West Virginia University, Morgantown, WV 26506-6315, USA

²Institute of Physics, Czech Academy of Sciences, Cukrovarnická 10, 162 53, Praha 6, Czech Republic

³Departamento de Física Teórica de la Materia Condensada, Universidad Autónoma de Madrid, Madrid 28049, Spain

⁴Department of Physics, The University of Texas at Austin, Austin, TX 78712-1081, USA

⁵Department of Physics, Arizona State University, Tempe, AZ 85217-1504, USA

⁶Department of Physics and Astronomy, Ohio University, Athens, OH 45701, USA

Received 23 May 2011, revised 9 July 2011, accepted 11 July 2011

Published online 15 August 2011

Keywords *ab initio* molecular-dynamics

* Corresponding author: e-mail james.lewis@mail.wvu.edu, Phone: +1-304-293-3422 x1409, Fax: +1-304-293-5732

One of the outstanding advancements in electronic-structure density-functional methods is the Sankey–Niklewski (SN) approach [Sankey and Niklewski, Phys. Rev. B **40**, 3979 (1989)]; a method for computing total energies and forces, within an *ab initio* tight-binding formalism. Over the past two decades, several improvements to the method have been proposed and utilized to calculate materials ranging from biomolecules to semiconductors. In particular, the improved

method (called FIREBALL) uses separable pseudopotentials and goes beyond the minimal sp^3 basis set of the SN method, allowing for double numerical (DN) basis sets with the addition of polarization orbitals and d-orbitals to the basis set. Herein, we report a review of the method, some improved theoretical developments, and some recent application to a variety of systems.

© 2011 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

1 Introduction With the increase in computational power, greater efforts have been made by the electronic-structure community to optimize the performance of quantum mechanical methods. Quantum mechanical methods have become increasingly reliable as a complementary tool to experimental research. A variety of methods exist ranging in complexity from semi-empirical methods to density-functional theory (DFT) methods to highly-accurate methods going beyond the one-electron picture. Judicious approximations enable the computational materials science community to more efficiently examine a wider range of materials questions.

Otto F. Sankey was one of the early visionaries by, firstly, demonstrating that molecular-dynamics (MD) simulations can be coupled efficiently with electronic-structure methods to optimize structures and evaluate energetics of materials [1]. Secondly, his judicious approximations in the

difficult exchange-correlation matrix elements using a multi-center approach enables faster evaluation of materials over other approaches [2]; his single greatest innovation was the use of basis functions with compact support. Finally, his immense contributions in predicting and understanding the properties of materials, and continuously developing new theoretical techniques to this end, is demonstrated by his over 225 publications with over 10 000 citations to this point.

One of the first reported approaches using an *ab initio* tight-binding MD formalism was the development by Sankey and Niklewski [2]. Within a tight-binding-like formalism many materials science investigations involving larger systems can be calculated with only a slight decrease in the accuracy. This is particularly useful where a quantum mechanical description is important to the investigated system's fundamental chemistry or atomic dynamics, but where a smaller model system would inadequately describe

the proper physical environment. In addition, there are even larger systems (*i.e.*, enzymes or zeolites) which can only be currently calculated using these approximate methods; using more exact methods would be computationally prohibitive. The original Sankey–Niklewski (SN) method is based on norm-conserving pseudopotentials [3, 4] and solely uses the local density approximation (LDA) limit of DFT, using the Harris–Foulkes functional [5, 6]. The atomic basis set was originally a minimal non-orthogonal local-orbital basis of slightly excited *fireball* orbitals; this was the first introduction of short-range numerical atomic-like orbitals used in electronic-structure methods [2, 7]. The electronic eigenstates are expanded as a linear combination of pseudo-atomic orbitals within a localized minimal sp^3 basis for the atoms.

Several studies on a variety of systems ranging from biomolecules to semiconductors have shown the SN method to be an efficient and successful tool for performing efficient electronic-structure calculations. Some of the earlier seminal work using the SN method is noted in calculations of fullerenes [8], Si surfaces [9], and a-Si [10, 11], the latter reference demonstrated an excellent representation of the Staebler–Wronski effect through simulation. The achievement of the SN method is evident by noting that aspects of the method are utilized in other notable computational methods (*e.g.*, SIESTA) [12–17]. Improvements to the method, now called FIREBALL, include a self-consistent extension originally developed and successfully applied to complex silicas [7, 18–22]. We have applied an expansion of the basis set [23] to correctly obtain self-consistent polarization states to polar molecules such as octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX, an highly energetic explosive molecule) [23–26]. In other earlier work, we have included an overlap-expansion formalism in the SN method for modeling hydrogen-bonding interactions [27, 28]. Implementation of a linear-scaling algorithm [12, 29, 30] in addition to this hydrogen-bonding model provided for the first *ab initio* DFT calculation of a 10-mer deoxyribonucleic acid (DNA) molecule [31].

Some more recent impressive highlights of FIREBALL include two surface studies in high impact journals. The first by Sugimoto et al. reported precise measurements of short-range chemical forces that occur between a scanning probe tip and a surface. FIREBALL was used to sensitively predict how these chemical forces in silicon, tin, and lead (with very similar chemical properties) would change depending on the atom type; thereby providing interpretation of the experimental results [32–34]. The second by Otero et al. reported the formulation of C_{60} and the triazafullerene $C_{57}N_3$ from aromatic precursors using a highly efficient surface-catalyzed cyclodehydrogenation process. FIREBALL was used here to predict the structure of the aromatic precursors and the structures at different stages of the process [35, 36].

In this review, which is by no means exhaustive, we give a summary of the most recent FIREBALL formalism and discuss a survey of recently published results from the past decade. Primarily, we report on recent enhancements to the exchange-correlation multi-center expansion of the SN

method and we summarize more recent FIREBALL results regarding a variety of systems such as DNA, transition-metal oxides, metallic nanostructure systems, disordered systems, as well as calculated properties of surfaces and interfaces. The sampling of results represented in this paper indicated a tribute to the impact that Otto F. Sankey has had in the field of electronic-structure, materials science, physics, and chemistry.

2 Methodology FIREBALL is a real-space local-pseudoatomic-orbital MD implementation of DFT, cast in a tight-binding-like form, that provides a substantial improvement in computational efficiency and high accuracy. At the core of the theory is the replacement of the Kohn–Sham self-consistent density functional by an approximate self-consistent functional based on atomic occupations numbers [18, 37]. We provide a summary of the details of the method in the following sections.

2.1 Localized pseudo-atomic orbitals and basis sets In solving the one-electron Schrödinger equation [see Eq. (2)], we use a set of numerical atomic-like orbitals based on a pseudo-potential formalism. The construction of the pseudopotentials commences with the self-consistent solution of the Kohn–Sham equations for all electrons of the free atom. For the free atom, the exchange-correlation energy $E_{xc}[\rho]$, and respectively for the exchange-correlation potential $V_{xc}[\rho; r]$, various parameterizations of the LDA and of the generalized gradient approximation (GGA) are available [38–45]. Having solved the single particle equations for all electrons of the free atom, the pseudo-potential and pseudo-atomic wavefunctions are generated via the freely available fhiPP code from Fuchs and Scheffler [46] by means of the Hamann-scheme or in the Troullier–Martins form as discussed in detail in Refs. [47] and [48]. The pseudopotentials are then transformed into the fully separable form suggested by Kleinman and Bylander [49]. The absence of unphysical ghost states is checked by examining the bound state spectrum using the analysis of Ref. [50].

Our localized pseudo-atomic orbitals, which we refer to as *fireballs*, are expressed in the form ($\psi_{\text{fireball}}^{\text{atomic}}(r) = f(r)Y_{lm}(\theta, \phi)$), where $f(r)$ is a radial component and $Y_{lm}(\theta, \phi)$ is the angular component which are the spherical harmonic functions. The *fireballs* orbitals are slightly excited due to the boundary condition that they vanish at some radius r_c ($\psi_{\text{fireball}}^{\text{atomic}}(r)|_{r \geq r_c} = 0$) instead of the atomic boundary condition that they vanish at infinity. This boundary condition is equivalent to a “particle in the box” and has the effect of raising the electronic energy levels ($\epsilon_s, \epsilon_p, \epsilon_d, \dots$ atomic eigenvalues). This slight excitation accounts for Fermi compression in solids which apparently gives a better representation of solid-state charge densities [51]. The radial cutoffs (r_c) are chosen such that these electronic eigenvalues remain negative and are mildly perturbed from the free atom. For instance, Fig. 1 shows a plot of the s, p, and d *fireball* orbitals representing oxygen’s free-atomic

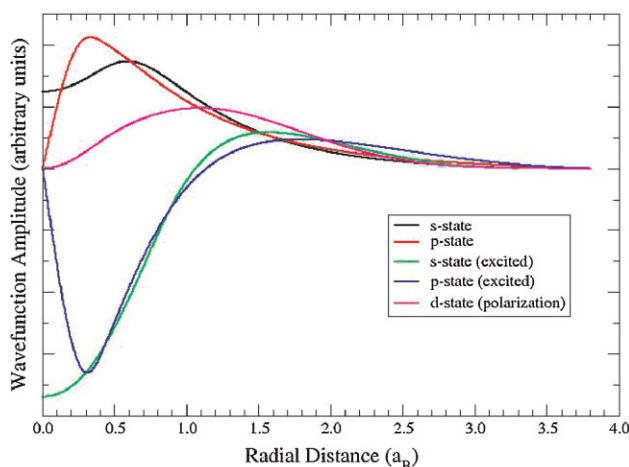


Figure 1 (online color at: www.pss-b.com) Local pseudo-atomic orbitals for oxygen. The basis shown is double numerical in s and p, with a d polarization state.

wavefunctions. It is important that the r_c 's are chosen to preserve the chemical trends of the atoms, *i.e.*, the excitation of the atoms must be done in a manner that preserves the relative ionization energies and relative atomic sizes. A basis for judiciously choosing these r_c 's was discussed in a previous work [7].

We find the use of localized *fireball* orbitals to be computationally advantageous. Given any two atomic orbitals, i and j , beyond some cutoff radius ($r_{ci} + r_{cj}$), the matrix elements H_{ij} and S_{ij} become exactly zero. Therefore, there is only a prescribed interaction range over which the integrals must be evaluated. A multi-center approach is used (see Fig. 2) and all interactions up to three-centers are computed exactly. This inherent sparseness allows one to more readily implement linear-scaling algorithms for

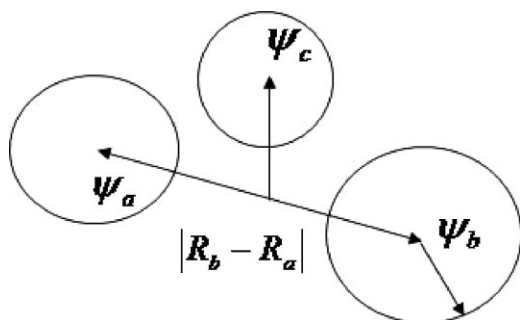


Figure 2 Interactions are pre-computed using a multi-center expansion; a three-center expansion is shown. If the distance between any two centers is larger than the cutoffs of the two wavefunctions on the centers, then the overlap is zero and the interactions between centers will go to zero. For example, if $|R_b - R_a|$ is larger than the cutoffs of the two wavefunctions, r_{ca} and r_{cb} , where r_{ca} is the cutoff of atom a and r_{cb} is the cutoff of atom b , *e.g.*, $|R_b - R_a| > r_{ca} + r_{cb}$ as shown, then the overlap is zero and the interactions between centers will go to zero.

obtaining the band-structure energy. Within the FIREBALL approach integrals are pre-calculated on a numerical grid and the specific values needed are gleaned from the tabulated values via interpolation. We choose an interpolation method that maximizes numerical stability. Because these integral tables depend only on the atom type, their r_c values, and the type of DFT exchange-correlation functional used, the integral tables only need to be generated once, for a given number of atomic species, rather than calculating integrals “on-the-fly” during an MD simulation. This pre-generation process lends itself to parallelization via spreading the integrals out over multiple processors based on integral types [23, 30]. This parallelization is particularly important, since the number of integrals needed for each database grows as order N^3 with the number of atom types in the database.

Our approach allows a flexible choice of basis sets permitting double numerical (DN) or additional polarization sets. Earlier work shows that addition of the DN set yields better results in *ab initio* tight-binding methods similar to FIREBALL [13, 14]. DN basis sets are currently generated by holding the ground state wavefunction fixed and exciting electron density to a higher orthogonal state with the same r_c value. There are other approximations for generating additional DN basis sets as discussed in [13, 14, 52]. Polarizing d functions (unoccupied orbitals in the ground state atom) are generated by exciting the electron density into a d shell with a well chosen r_c value. These radii must be chosen such that the proper chemical trends (such as ionization) are maintained between atomic orbitals. For example, choosing one orbital radius as very large and another orbital radius as too small would artificially encourage charge transfer from the latter to the former. It should be noted that the d shells used in FIREBALL consist of five spherical harmonic d functions as opposed to the six Cartesian d functions used with some Gaussian basis sets. Polarization provides a flexibility in the basis set that was not available in the functions of lower angular momentum, and thus may improve the chemistry. The use of DN basis sets is particularly important in the self-consistent implementation, since charge transference will significantly alter the net potential energy, which alters the kinetic energy via the virial theorem, which in turn effects the curvature of the wavefunction.

On other hand, one wishes to find a universally-robust minimal basis set, which would significantly speed up performed simulations, but still keeping desired accuracy. Therefore, we devised an alternative approach to define a single numerical basis set obtained as a linear combination of only two primitive numerical atomic orbitals corresponding to the ground states of the neutral atom and the $2+$ ion, respectively [53]. We applied the method searching for an optimized single numerical basis sets for some sample hydrocarbons and for other systems made of carbon (*e.g.*, graphene, fullerenes, nanotubes). We found that the new optimized basis sets yield an important lowering of the total energy, and bondlengths in very good agreement with the experimental evidence. The approach is a promising tool for

the simulation of complex organic materials, beyond the hydrocarbons, using optimized minimal basis sets.

2.2 Tight-binding density-functional formalism At the core of the FIREBALL method is the initial replacement of the Kohn-Sham energy functional by an approximate self-consistent extension of the Harris–Foulkes functional [5, 6, 18]:

$$E_{\text{tot}}^{\text{Harris}} = E^{\text{BS}} + \{U^{\text{ion-ion}} - U^{\text{ee}}[\rho_{\text{in}}(\mathbf{r})]\} + \{U^{\text{xc}}[\rho_{\text{in}}(\mathbf{r})] - V^{\text{xc}}[\rho_{\text{in}}(\mathbf{r})]\}. \quad (1)$$

The main difference between the Kohn–Sham and Harris–Foulkes functional is that the latter is defined entirely in terms of an input charge density, $\rho_{\text{in}}(\mathbf{r})$; whereas, the former is defined in terms of both an input and output charge density. In Eq. (1), E^{BS} is the band-structure energy ($2 \sum_{i \in \text{occ}} \varepsilon_i$), where ε_i are the eigenvalues of the one-electron Schrödinger equation given by

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right\} \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \quad (2)$$

with

$$V(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + \mu_{\text{xc}}[\rho_{\text{in}}(\mathbf{r})] + \frac{e^2}{2} \int \frac{\rho_{\text{in}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r'. \quad (3)$$

The second term of Eq. (1) is the short-range repulsive interaction which is the ion-ion interaction offset by the overcounting of the Hartree interactions. This term is given by

$$U^{\text{short-range}}(\mathbf{r}) = \{U^{\text{ion-ion}} - U^{\text{ee}}[\rho_{\text{in}}(\mathbf{r})]\}, \quad (4)$$

with

$$U^{\text{ion-ion}} = \frac{e^2}{2} \sum_{i,j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \quad (5)$$

and

$$U^{\text{ee}}[\rho_{\text{in}}(\mathbf{r})] = \frac{e^2}{2} \int \frac{\rho_{\text{in}}(\mathbf{r}) \rho_{\text{in}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r'. \quad (6)$$

The last term of Eq. (1) is a correction of the exchange-correlation, given by

$$\int \rho_{\text{in}}(\mathbf{r}) \{ \varepsilon_{\text{xc}}[\rho_{\text{in}}(\mathbf{r})] - \mu_{\text{xc}}[\rho_{\text{in}}(\mathbf{r})] \}. \quad (7)$$

This term arises because the one-electron Schrödinger equation contains the exchange-correlation potential $\mu_{\text{xc}}[\rho_{\text{in}}(\mathbf{r})]$; however, the correct exchange-correlation energy is obtained from $\varepsilon_{\text{xc}}[\rho_{\text{in}}(\mathbf{r})]$. In general, the calculation of the one-electron Hamiltonian matrix elements as well as the double-counting Hartree and exchange-correlation terms [Eqs. (4) and (7)] in terms of pre-calculated data files as outlined in the original SN paper. Long-range Hartree terms are discussed in Ref. [18].

In evaluating the total energy of the system [Eq. (1)], the input density to be evaluated is a sum of confined spherical atomic-like densities,

$$\rho_{\text{in}}(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r} - \mathbf{R}_i)|^2. \quad (8)$$

The orbitals $\phi_i(\mathbf{r} - \mathbf{R}_i)$ are the slightly excited *fireball* pseudo-atomic wavefunctions which are used as basis functions for solving the one-electron Schrödinger equation [Eq. (2)]. The self-consistent occupation number, n_i , determines the number of electrons occupying each spherically confined atomic-like densities, and is given by

$$n_i = n_i^0 + \delta n_i. \quad (9)$$

As a comparison, in the Harris–Foulkes approximation (implemented within the original SN method), the input density is not determined self-consistently, but rather the occupation numbers are taken as the reference atomic density ($n_i = n_i^0$). It has been shown that the total energy has errors that are only second order in the errors of the input density [2, 5]. Therefore, to the first order in $\delta \rho(\mathbf{r})$ Eq. (1) can be rewritten by replacing the total self-consistent density [Eq. (9)] by the reference density ($n_i = n_i^0$).

2.3 Exchange–correlation interactions We evaluate exchange–correlation (XC) interactions in FIREBALL using the McWEDA method proposed by Jelínek et al. [54], which uses a many-center expansion based on an expansion of the density a site at a time. This method provides advantages over the *nearly uniform density approximation* which was utilized in the SN method [2]. Primarily, this is a higher order approximation than the nearly uniform density approximation and it can be used in principle with gradient corrected functionals. The aim is to calculate the exchange-correlation contributions from pre-calculated integrals stored in data tables, using a three-center approximation. We distinguish two cases (i_μ is the atomic site corresponding to orbital μ and i_ν is the atomic site corresponding to orbital ν): (a) on-site matrix elements ($i_\mu = i_\nu$) and (b) off-site matrix elements ($i_\mu \neq i_\nu$).

In order to improve the SN approach and generalize it beyond the minimal sp^3 basis sets (see Ref. [54] for a discussion), we calculate exchange-correlation matrix elements as the sum of a main (one-center or two-center) term plus a correction. The main term is calculated exactly, while the correction is calculated using the many-center expansion proposed by SN. In order to calculate this correction, we define a generalized SN (GSN) approximation for the exchange-correlation matrix elements as follows

$$\int \phi_{i\mu}(\mathbf{r}) V_{\text{xc}}[\rho; \mathbf{r}] \phi_{i\nu}(\mathbf{r}) d\mathbf{r} = V_{\text{xc}}[\bar{\rho}_{\mu\nu}; \mathbf{r}] S + V'_{\text{xc}}[\bar{\rho}_{\mu\nu}; \mathbf{r}] \left(\int \phi_{i\mu}(\mathbf{r}) \rho(\mathbf{r}) \phi_{i\nu}(\mathbf{r}) d\mathbf{r} - \bar{\rho}_{\mu\nu} S \right), \quad (10)$$

with

$$S = \int \phi_{i\mu}(\mathbf{r})\phi_{iv}(\mathbf{r})d\mathbf{r}. \quad (11)$$

This approximation allows us to calculate correction terms in a practical way. We define the average densities, $\bar{\rho}_{\mu\nu}$, as

$$\bar{\rho}_{\mu\nu} = \frac{\langle w_\mu | \rho | w_\nu \rangle}{\langle w_\mu | w_\nu \rangle}, \quad (12)$$

where the *weighting functions*, w_μ , are taken as a spherical average of the pseudo-atomic orbital, so $w_{il} = |R_{il}(r)|Y_{00}(\Omega)$, where $|R_{il}(r)|$ is the absolute value of the radial function $R_{il}(r)$ associated with the orbital ϕ_{il} . These functions use the idea of importance-sampling to define an average density for each matrix element.

- (a) *On-site matrix elements* ($i_\mu = i_\nu$). Here, we simply add and subtract a contribution associated with the atomic density ρ_i at site i , and write *formally*, the matrix elements as a one-center contribution plus a *correction* consisting of several terms based on the functional derivative of the potential and the average density, $\bar{\rho}$. Using the GSN approximation Eq. (13) to calculate the correction we obtain (see Ref. [54] for further details)

$$\begin{aligned} \int \phi_{i\mu}(\mathbf{r})V_{xc}[\rho; \mathbf{r}]\phi_{iv}(\mathbf{r})d\mathbf{r} &= \int \phi_{i\mu}(\mathbf{r})V_{xc}[\rho_i; \mathbf{r}]\phi_{iv}(\mathbf{r})d\mathbf{r} \\ &+ V'_{xc}[\bar{\rho}_{\mu\nu}; \mathbf{r}] \left(\int \phi_{i\mu}(\mathbf{r})\rho(\mathbf{r})\phi_{iv}(\mathbf{r})d\mathbf{r} - \bar{\rho}_{\mu\nu}S \right) \\ &- V'_{xc}[\bar{\rho}_i; \mathbf{r}] \left(\int \phi_{i\mu}(\mathbf{r})\rho_i(\mathbf{r})\phi_{iv}(\mathbf{r})d\mathbf{r} - \bar{\rho}_i S \right) \\ &+ (V_{xc}[\bar{\rho}_{\mu\nu}; \mathbf{r}] - V_{xc}[\bar{\rho}_i; \mathbf{r}])S, \end{aligned} \quad (13)$$

where S is defined as in Eq. (11), and with

$$\bar{\rho}_i = \frac{\langle w_\mu | \rho_i | w_\nu \rangle}{\langle w_\mu | w_\nu \rangle}. \quad (14)$$

The one-center term (first term on the right-hand side) is calculated exactly and is much larger than the *correction* terms.

- (b) *Off-site matrix elements* ($(i_\mu = i) \neq (i_\nu = j)$). Proceeding in a similar manner as for the on-site matrix elements, we first write the matrix element as a two-center main contributions that we calculate exactly (also using the two-center density, $\rho_{ij} = \rho_j + \rho_j$), and a correction that is evaluated using the GSN approximation in Eq. (10). We

obtain for the McWEDA evaluation of the off-site matrix elements

$$\begin{aligned} \int \phi_{i\mu}(\mathbf{r})V_{xc}[\rho; \mathbf{r}]\phi_{j\nu}(\mathbf{r})d\mathbf{r} &= \int \phi_{i\mu}(\mathbf{r})V_{xc}[\rho_{ij}; \mathbf{r}]\phi_{j\nu}(\mathbf{r})d\mathbf{r} \\ &+ V'_{xc}[\bar{\rho}_{\mu\nu}; \mathbf{r}] \left(\int \phi_{i\mu}(\mathbf{r})\rho(\mathbf{r})\phi_{j\nu}(\mathbf{r})d\mathbf{r} - \bar{\rho}_{\mu\nu}S \right) \\ &- V'_{xc}[\bar{\rho}_{ij}; \mathbf{r}] \left(\int \phi_{i\mu}(\mathbf{r})(\rho_i(\mathbf{r}) + \rho_j(\mathbf{r}))\phi_{j\nu}(\mathbf{r})d\mathbf{r} - \bar{\rho}_{ij}S \right) \\ &+ (V_{xc}[\bar{\rho}_{\mu\nu}; \mathbf{r}] - V_{xc}[\bar{\rho}_{ij}; \mathbf{r}])S, \end{aligned} \quad (15)$$

where S is defined as in Eq. (11), and with

$$\bar{\rho}_{ij} = \frac{\langle w_\mu | \rho_{ij} | w_\nu \rangle}{\langle w_\mu | w_\nu \rangle} \quad (16)$$

(indices μ and ν are omitted here for clarity). In Eqs. (13) and (15), $\bar{\rho}_{\mu\nu}$, which includes all density contributions, is defined using Eq. (12).

Equations (12)–(16) form the basis of the McWEDA approximation for the calculation of the exchange-correlation matrix elements. The initial SN method's multi-center expansion with these subsequent improvements through McWEDA, coupled with *fireball* orbitals and pre-computed data files are what form the important theoretical underpinnings of the FIREBALL methodology. Notice that in this approach the GSN approximation, Eq. (10), is only used to evaluate the correction terms to the dominant one- or two-center contributions. Our approach for determining the exchange-correlation interactions is independent of the type of functional used. Previously, we have considered two types of exchange-correlation density functionals within FIREBALL – LDA and Becke exchange (B88) [41] with Lee–Yang–Parr (LYP) correlation [40]; implementation of other functionals is academic.

2.4 Self-consistency implementation The representation of the input electron density in the form given by Eq. (8) prevents us from achieving self-consistency in the standard way, where $\rho_{in}(\mathbf{r}) = \rho_{out}(\mathbf{r})$ at each point \mathbf{r} . Instead of that, we take a point of view based on a local-orbital formulation of DFT [37], where self-consistency is implemented on the orbital occupancies. Therefore, in our self-consistent procedure [18] the occupation numbers of the input density in Eq. (8) are non-neutral (*i.e.*, $n_i = n_i^0 + \delta n_i$) and the total energy is now a function of the self-consistent occupation numbers n_i , $E_{tot}[\rho_{in}(\mathbf{r})] \equiv E_{tot}[n_i]$ ($n_i \neq n_i^0$). In order to determine the self-consistent n_i 's we use the following approximate method. Firstly, we define the output occupation numbers n_i^{out} as

$$n_i^{out} = 2 \sum_\alpha |\langle \psi_\alpha | \varphi_i \rangle|^2, \quad (17)$$

where ψ_α are the occupied eigenstates of Eq. (2) and φ_i are the atomic-like orthogonal orbitals of Löwdin [55]:

$$\varphi_i = \sum_j (S)_{ij}^{-\frac{1}{2}} \phi_j. \quad (18)$$

Secondly, we determine the occupation factors n_i from the “self-consistency” condition $n_i^{\text{out}} = n_i$ for all i , where n_i is the input occupation numbers from Eq. (8). Equation (17) may be viewed as a way of projecting the *output* electron density obtained from Eq. (2) into a density of the form given by Eq. (8). The problem of finding the self-consistent solution has therefore been simplified from that of matching of the input and output densities on the $M \times M \times M$ grid in real space (M is the number of points on the grid and can be extremely large for fast varying densities) to that of matching $n \times N$ real numbers (N is the number of atoms, and n is the number of valence shells on each atom).

3 Applications of the FIREBALL method

3.1 Doped TiO₂ and TiO₂ Nanostructures The success of the FIREBALL methodology in predicting bulk properties of materials stems from the early inception of the SN method [2], later developments with self-consistency [18, 37], and improvements in the multi-center expansion of the exchange-correlation [54]. One of the important features of FIREBALL is the flexibility of constructing real-space localized basis functions to take advantage of fundamental chemistry in atomic bonding. This allows a substantial improvement in computational efficiency without suffering the loss of accuracy and hundreds of published results to a variety of materials attest to effectiveness of the method. More recently, much success has been demonstrated in the area of transition metal oxides, particularly, a review of bulk and nanostructures of TiO₂ using FIREBALL is presented here.

The importance of TiO₂ is mostly evident from noting that in 1972, Fujishima and Honda demonstrated that the photoexcitation presented in TiO₂ is promising to chemically split water to form H₂ and O₂; thus the potentiality of solar energy conversion through semiconducting materials was conceived [56]. The only issue which restricts the broad range of applications for TiO₂ in solar energy conversion is the intrinsic large band gap of TiO₂. Only 3% of the solar spectrum (ultraviolet) can be utilized, which dramatically decreases the efficiency of TiO₂ in light absorbance. Therefore, considerable efforts have been devoted to decrease the band gap and increase the light absorbance. Impurity doping is one of the typical approaches to extend the spectral response of a wide band gap semiconductor to visible light. Recently, it was shown that anionic non-metal doping of TiO₂ are ideal for optimal engineering of the valence band states by introducing deep-level states and thereby decreasing the band gap [57–62].

3.1.1 Results for bulk TiO₂ Two phases of TiO₂ are considered of primary interest in photoactivity – rutile and

anatase; both are tetragonal. The tetragonal rutile (anatase) structure belongs to the space group $P4_2/mnm$ (D_{4h}^{14}) ($I4_1/amd$ (D_{4h}^{19})), containing 6 (12) atoms per unit cell. Structural parameters for both rutile and anatase have been determined to a high degree of accuracy from the neutron diffraction experiments performed by Burdett et al. [63]. We have calculated the optimal structures by minimizing the total energy of the rutile and anatase structures with respect to the lattice parameters a and c , and the internal parameter u . Our results of these structural parameters in both rutile and anatase are within 1% of the experimental results; bulk moduli data are within 5% of the experimental data [54, 64]. From the equilibrium lattice parameters, we have calculated the self-consistent electronic band structures for rutile and anatase along the high-symmetry directions of the irreducible Brillouin zone. Band-width results using FIREBALL are consistent with other calculations [54, 64]. The traditional use of the LDA and the Kohn–Sham approximation generally underestimates (compared to experiment) the band gap for insulators and semiconductors; however, finite local orbital basis sets tend to overestimate band gaps. Thus, the Γ point band gap of 3.05 eV (3.26 eV) for rutile (anatase) is in agreement with the reported experimental gap of 3.06 eV (3.20 eV) [64].

One approach to developing next-generation TiO₂ photocatalysts is to form alloys by substitutional “doping” with anionic species such as C, S, or N as metal oxide materials like TiO₂ exhibit wide band gaps (3.0 eV for TiO₂); resulting in photoactivity which utilizes only a small percentage of the solar spectrum. This approach is quite promising because such doping will contribute p-states near the valence band much like any other deep donor level in semiconductors. Using FIREBALL, Wang and Lewis [65] found that carbon produces states above the reference TiO₂ valence band edge and seen in Fig. 3.

In their work, two supercells for both rutile and anatase were constructed based on our theoretically determined equilibrium structures. The larger supercells contains 64 primitive unit cells (384 atoms in total) and randomly substituted carbons were placed at oxygen sites, yielding carbon concentrations of 0.26%. In smaller supercells, which contain 16 primitive unit cells (96 atoms in total), we randomly substitute 5 carbons yielding a 5.2% concentration. The band gap narrows as a result of the carbon and gap energies are 2.30 and 2.45 eV for 5.2 and 0.26% concentration of carbon, respectively. Similar results are observed experimentally as several studies report a red shift in the absorption spectra [57–62]. Irie et al. [60] and Khan et al. [58] both show energy levels just above the valence energy which corresponds to *substitutional* carbon only. They also conclude that the substitutional carbon atoms are responsible for the visible light absorption. The experiments by Khan et al. [58] show appreciable absorption from carbon at wavelengths less than 535 nm in samples of ~5% carbon. This corresponds to a band gap energy of 2.32 eV and is comparable to our calculated band gap for 5.2% carbon concentration.

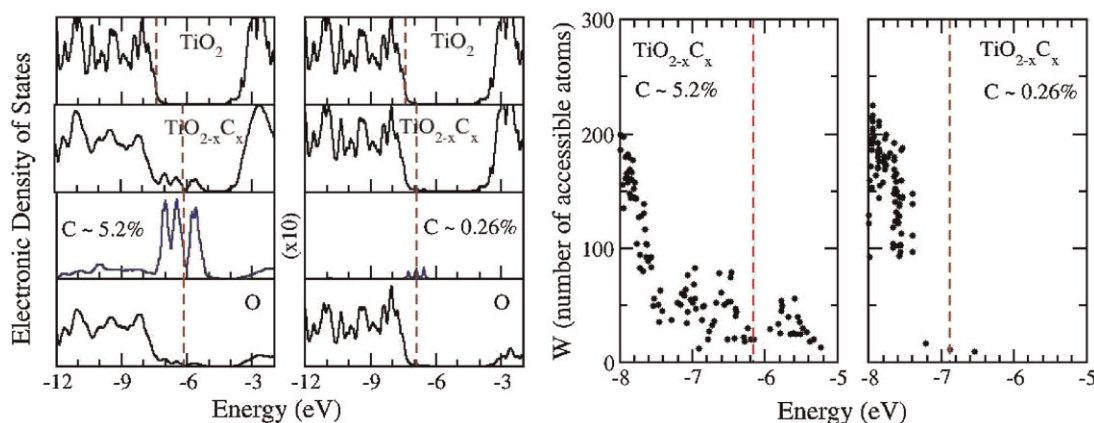


Figure 3 (online color at: www.pss-b.com) Plot of the density of states (*left* plots) and the degree of localization (*right* plots), both are functions of the electronic energy, for two concentrations of C doping (5.2 and 0.26%) in TiO₂. The localization is measured by the entropic quantity *W* (from Refs. [64, 65]) which is the number of accessible atoms for the given electronic energy and which describes the spatial extent of the electronic state.

An interpretation given by Asahi et al. [57], which we support is that the states in the gap should overlap sufficiently with the bands of TiO₂ to transfer photoexcited carriers to reactive sites at the catalyst surface within their lifetime. Our DOS calculations in Fig. 3 show that there is significant overlap with the carbon states and TiO₂ valence bands for 5.2% carbon concentration. Contrarily, the three energy levels in the DOS for the lower concentration are quite distinct and highly localized on the single carbon dopant. As seen in the right plots of Fig. 3, the states are highly localized and there is no significant overlap for the 0.26% carbon concentration; hence, the states do not mix with Ti or O and as a result, the charges will not easily migrate to the surface where they are required for catalysis [64, 65]. Instead, these highly localized states indicate that lower-percentage dopings produce impurities in the material which will act as recombination centers. Not until 5.2% carbon concentrations are considered, do the dopant states mix with the Ti and O states and thereby produce sufficient charge injection to Ti/O reaction sites. Accordingly, our prediction is that 5.2% carbon concentration will exhibit much more photocatalytic efficiency than the 0.26% carbon concentration. Experimentally, there is an obvious correlation between carbon content and photocatalytic efficiency. A concentration of 5% yields an improved efficiency in splitting water [58]; but a much lower efficiency (~0.2%) for the decomposition of gaseous 2-propanol (IPA) albeit photocatalytic in the visible.

3.1.2 Nanostructured TiO₂ In recent decades, as improvements have been made to synthesize nanostructures, nanoscale titanium dioxide materials have been exploited extensively in many applications including photocatalysis, solar cells, biomaterials, memory devices, and as environmental catalysts (see Ref. [66]). Among TiO₂ polymorphs, anatase is of paramount importance since it exhibits higher reactivity in many cases [67], especially as the system size

decreases to the nanoscale; TiO₂ nanocrystals appear to prefer the metastable anatase phase rather than the rutile phase due to the lower surface energy of anatase particles [68]. Generally, it is believed that the behavior of TiO₂ nanoparticles are greatly different from either bulk or atomic scale clusters due to the formation of disordered curvature on the nanoparticle facets. The surface energies of finite size particles are different from the surface energies of planar surfaces, as the curvature of the finite size particles must be accounted for in addition to their surface crystal direction.

Calculations using FIREBALL have demonstrated the atomic structures and electronic properties of medium size (from 182–774 atoms) anatase nanoparticles, as well as interpret the underlying mechanism for chemisorption of water molecules at nanoparticle sites [66]. These results have established that the atomic and electronic properties of nanoparticles are very different from the bulk anatase. The reason why the behavior of TiO₂ nanoparticles exhibits a great difference from bulk and atomic scale clusters is the formation of corner (or edged) Ti atoms which can act as *hot spots* to interact with the species existing around nanoparticles. The edged Ti atoms are juncture connected between two different surfaces exposed in nanoparticles (see Fig. 4). Small peak states were found near the Fermi energy levels in the band gap of calculated nanoparticles. The highly localized states and the symmetric distribution on the edged Ti atoms, these small peak states are surface states which interact with electrons acceptor species around nanoparticles. The underlying mechanism is also shown in Fig. 4; these gap state peaks on the nanoparticles surfaces act as reaction or chemisorption centers. The concept of nanoparticle reactivity and the role of edges is explored in several joint experimental and theoretical investigations such as the decomposition of the methyl orange in the visible irradiation for doped TiO₂ nanobelts and shape-enhanced nanoparticles [69–71]. These results give insights into the development

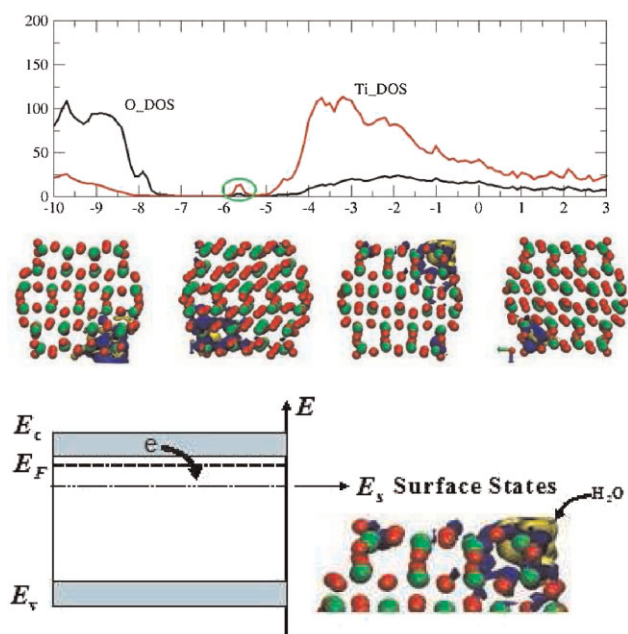


Figure 4 (online color at: www.pss-b.com) The electronic properties of anatase nanoparticle with 228 atoms. The small peaks existing in the band gap are surface states distributed in the edged Ti atoms. The surface states act as possible relay trap of electrons and hot spots to interact with species, such as water molecules.

of visible light photocatalysts based on nitrogen-doped nanostructures with large specific surface area and unique morphology, which are important ingredients to improve photocatalytic, electronic, or gas sensing performance.

3.2 Surfaces and interfaces The FIREBALL approach has been applied extensively to study the atomic, electronic and dynamical properties of surfaces and interfaces [72–75]. The advantage of this *ab initio* MD technique is that, due to the excellent balance between accuracy and computational performance, it makes an *exploratory tool* to search for unknown surface atomic structures [76, 77], study their electronic structure, and investigate atomic motions on the surface [78, 79]. One of the first applications of the FIREBALL code to surface systems was the analysis of the atomic and electronic structure of the Si(111)- 5×5 dimer-adatom-stacking (DAS) fault reconstruction, in what was the largest first-principles MD relaxation to that date [9]. These calculations show that the adatom dangling-bond (DB) states control the electronic structure around the Fermi energy, the six DB states (corresponding to the six adatoms) being filled with only three electrons. Similar results have been found for the Si(111)- 7×7 DAS reconstruction, where the twelve adatom DBs share five electrons [80]. FIREBALL calculations have been performed as a first-step to analyze the electron correlation effects on these surfaces [80, 81].

A primary example of the use of FIREBALL as a first-principles exploratory tool is the theoretical investigation of

reversible phase transitions at clean and adatom decorated semiconductor surfaces [82–84]. In these transitions the translational symmetry of the surface changes reversibly as a function of the temperature [82]; this structural two-dimensional transition is often accompanied by an electronic insulator-metal (IM) transition. Typical examples are the $\sqrt{3} \times \sqrt{3} \leftrightarrow 3 \times 3$ transitions in Sn or Pb adatom decorated Ge(111) or Si(111) surfaces [82, 84, 85], the $4 \times 1 \leftrightarrow 8 \times 2$ transition in In/Si(111) [86], and the $2 \times 1 \leftrightarrow c(4 \times 2)$ order–disorder transitions in Si(100), Ge(100), and SiC(100) [83]. The atomic mechanisms responsible for these 2D reversible transitions have been under intense debate. From a fundamental point of view, two questions must be solved: (1) what is the origin and precise atomic structure of the low T phase? (2) which are the atomic dynamical processes that give rise to the structural (and eventually IM) transitions? The first question has already proven a difficult theoretical/computational problem due to (i) the delicate energy differences between candidate atomic structures, (ii) the inaccuracies present in all first-principles methods, and (iii) the complex energy landscape that has to be explored in order to find the relevant atomic geometries which usually involve not obvious distortions from ideal structures. First-principles MD simulations may be performed in an attempt to answer the second question.

3.2.1 Dynamical fluctuations on the Sn/Ge(111) surface

The Sn/Ge(111)- $\sqrt{3} \times \sqrt{3} \leftrightarrow 3 \times 3$ transition [85] has been extensively studied (*e.g.*, see Ref. [84] and references therein). The ground state atomic geometry was found using FIREBALL and an MD annealing technique [87, 88]. In this 3×3 structure one Sn adatom (per unit cell) is displaced upwards and the other two Sn adatoms are displaced downwards, see Fig. 5, with a total vertical distortion ~ 0.3 Å; notice that there are three equivalent 3×3 ground state structures, since the up Sn-atom can be any of the three Sn atoms in the unit cell. This distortion has been shown to be related with a surface soft phonon [89]. In the ideal $\sqrt{3} \times \sqrt{3}$ surface there is one electron in each Sn DB but in the 3×3 structure the DB of the upper Sn atom is completely filled with two electrons while the two down Sn DBs share the other electron. These results indicate a strong coupling of 2D lattice vibrations with electrons in DB states close to the Fermi energy (or frontier orbitals).

The observed transition to a $\sqrt{3} \times \sqrt{3}$ symmetry has been studied using FIREBALL MD simulations [88, 90], see Fig. 6. These first-principles MD show how the system develops a chaotic motion above the transition temperature T_c , jumping between configurations associated with the different (but equivalent) 3×3 geometries. The time-scale of these fluctuations is on the order of 10^{-12} s; these *dynamical fluctuations* yield a $\sqrt{3} \times \sqrt{3}$ symmetry on average, explaining the apparent contradiction between electronic properties (time-scale of photoemission $\sim 10^{-15}$ s) and the structural symmetry observed (*e.g.*, time-scale of STM or LEED $\sim 10^{-3}$ s).

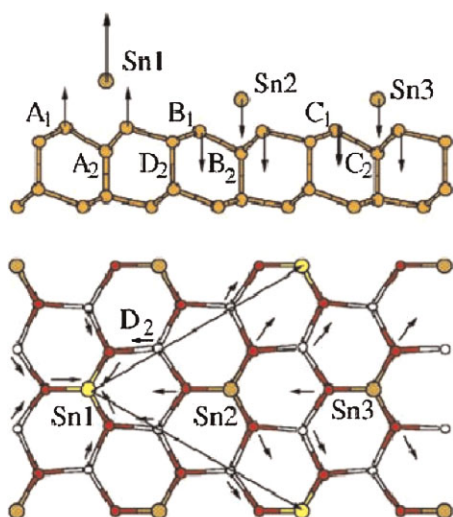


Figure 5 (online color at: www.pss-b.com) The Sn/Ge(111)- 3×3 atomic structure (side view and top view). The short arrows indicate the soft phonon distortions linking the ideal 3×3 structure to the 3×3 ground state. The long arrows (top view) indicate the 3×3 lattice vectors. Notice that there are three equivalent 3×3 ground states, corresponding to Sn atoms 1, 2, or 3 in the up-position (and the other 2 in down-position).

3.2.2 Dynamical fluctuations on the In/Si(111) and SiC (100) surfaces A similar investigation has been carried out for the In/Si(111)- $4 \times 1 \leftrightarrow 8 \times 2$ [91–93] and β -SiC(100)- $2 \times 1 \leftrightarrow c(4 \times 2)$ [94] reversible transitions. In these systems, due to the underlying soft phonons, at high temperature the system frequently visits atomic configurations that are metallic while jumping between ground states that are insulating, explaining the observed IM transitions [94].

3.2.3 Metal–organic interfaces FIREBALL has also been used as the basic computational tool to analyze different metal–organic interfaces, such as C_6H_6 /Au(111) and C_{60} /Au(111) [95], as well as TCNQ/Au(111) [96]. Metal–organic systems represent a challenge for standard DFT simulations not only because these are typically large

systems, but also for two important problems: the poor description of van der Waals forces [97, 98], resulting in many cases in a non-reliable molecule–surface adsorption geometry, and the organic energy gap problem [99, 100]. To correct these deficiencies, FIREBALL DFT-LDA is combined with a calculation of the charging energy U of the organic molecule on the metal surface [100, 101] to properly describe the transport gap and the energy level alignment at the interface; also, van der Waals interactions are taken into account to obtain a reliable molecule–surface adsorption distance. FIREBALL has also been used in combination with other DFT techniques to investigate the formation of fullerenes and triazafullerenes on the Pt(111) surface through the catalytic dehydrogenation of aromatic molecules [35], and the hydrogenation of β -SiC(100) semiconductor surfaces [102].

3.3 Transport properties of nanostructures The Local-orbital formulation of DFT offers certain technical advantages in describing molecular conductance [103–105]. The formulation in terms of local orbitals has an added value, as the transport properties can be easily calculated from the resulting first-principles tight-binding Hamiltonian using Green’s function techniques. In particular, Sankey has developed a Green’s function method based on the FIREBALL Hamiltonian and used it to describe molecular conductance in simple alkene chains, ply-alanine, DNA, and other molecular systems [104, 106–110]. More recently, a non-perturbative scattering theory framework for transport calculations was introduced inside the FIREBALL Hamiltonian to describe ballistic conductance in nanostructures using the Lippmann–Schwinger (LS) scattering theory [111, 112]; the local-orbital DFT Hamiltonian can be readily used within a self-consistent NEGF formalism [113, 114] to describe the inelastic tunneling spectroscopy (e.g., C_{60} [115]) and electron–phonon interactions in quatro-alanine bridges [116].

The methodology of using the FIREBALL-like Hamiltonian with its local orbital basis, and a non-perturbative scattering theory framework for transport calculations, was introduced in [112]. In that work, the system under consideration is two semi-infinite leads with a

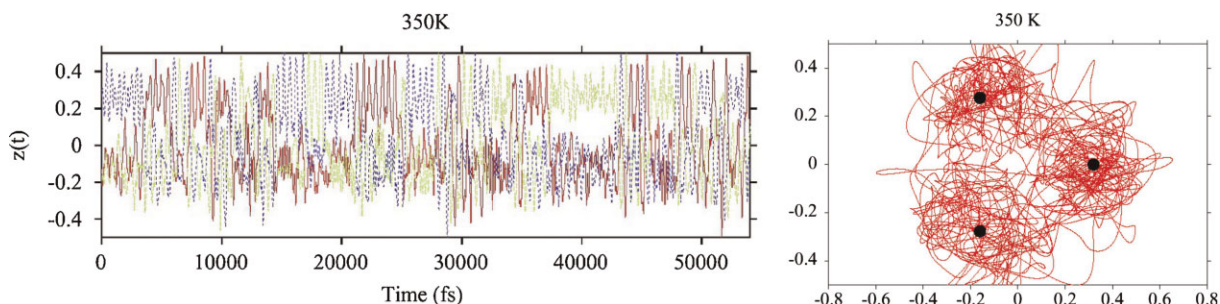


Figure 6 (online color at: www.pss-b.com) (left) MD simulation at 350 K showing the vertical position (in Å) of the 3 Sn atoms in the unit cell as a function of time, $z(t)$. The first-principles MD simulation shows how the system at room T is jumping between the three different 3×3 ground states in a chaotic fashion, giving rise on average to a $\sqrt{3} \times \sqrt{3}$ symmetry. (right) The 3 Sn atom motion on the phase space ($\gamma(t)$ versus $\beta(t)$) of the normal coordinates $\beta(t) = [2z_1(t) - z_2(t) - z_3(t)]/\sqrt{6}$; $\gamma(t) = [z_2(t) - z_3(t)]/\sqrt{2}$. The black dots in this figure indicate the position of the three different 3×3 ground states.

nano-device sandwiched between them. A structure like this could describe a molecule connected to two wires or a leaky capacitor with a very thin dielectric layer, or a scanning tunneling microscopy (STM) experiment. The formalism developed by Zhang et al. is extremely flexible and can work with any local orbital DFT code with minor modifications.

The goal of the theory is to describe the current flowing from one lead to another through a nano-size “device” region. Using the Fermi golden rule the total current is given by the usual expression:

$$I = \frac{8\pi^2 e}{h} \int_{-\infty}^{\infty} dE \sum_{l,r} |T_{lr}(E)|^2 \times \delta(E-E_l) \delta(E-E_r) \times [f(E-\mu_l) - f(E-\mu_r)], \quad (19)$$

where μ_l and μ_r refer to the electrochemical potential of the left and right leads, respectively, and a multiplicative factor of two has been included to account for spin degeneracy. The probability of scattering from left to right is given by the square of the corresponding matrix element, which we obtain by means of the elastic scattering theory. The transmission matrix element T_{lr} is the matrix element of the scattering operator, which is given by the LS equation [117],

$$\hat{T} = \hat{V} + \hat{G}\hat{V}\hat{V}. \quad (20)$$

Here, \hat{V} is the scattering potential describing the interaction of the Bloch electrons coming from the lead with the “device structure” and \hat{G} is the full Green’s function of the problem. If only the first term on the right hand side of Eq. (20) is retained, the theory reduces to the first Born approximation. However, the full \hat{T} operator is accurate to all orders.

At low temperature, the Fermi functions in Eq. (19) can be approximated by step functions resulting in the final expression for the total current:

$$I = \frac{2e}{h} \int_{\mu_l}^{\mu_r} dE T(E), \quad (21)$$

where the transmission function $T(E)$ is given by:

$$T(E) = 4\pi^2 \sum |T_{lr}(E)|^2 \delta(E-E_l) \delta(E-E_r). \quad (22)$$

Therefore, the problem of computing the current is reduced to that of calculating the transmission function $T(E)$. The current–voltage characteristic is readily obtained by integrating the transmission function over the energy window defined by the electrochemical potentials of the leads. If $T(E)$ is assumed to vary little, then this reduces to the Landauer formula $I = \frac{2e^2}{h} T_0 (\mu_r - \mu_l) = \frac{2e^2}{h} T_0 V$ with the quantum of conductance $2e^2/h = 12.9 \text{ k}\Omega$ [118].

Starting with Eq. (20), one can directly obtain the following compact expression for the transmission:

$$T(E) = \text{Tr} [\hat{\Gamma}_r \hat{G}_d \hat{\Gamma}_l \hat{G}_d^*], \quad (23)$$

where the indices l , r , and d stand for the left, right, and “device” (i.e., the scattering region) portions of the modeled system, respectively. Equation (23) is widely used to calculate the transmission function for a molecule sandwiched between two contacts. Operators Γ and G describe the leads and the device region of the problem, respectively. We describe both regions microscopically by a DFT FIREBALL Hamiltonian within LDA. The Hamiltonian describing our system is simply:

$$\hat{H} = \hat{H}_l + \hat{H}_r + \hat{H}_d + \hat{V}, \quad (24)$$

where \hat{H}_l and \hat{H}_r describe uncoupled left and right leads, e.g., semi-infinite metallic wires, respectively, \hat{H}_d describes the scattering region (also uncoupled), and \hat{V} describes the coupling between the leads and the scattering region (the “device”).

It is worth mentioning that Hamiltonians in Eq. (24) operate only in their own mutually orthogonal Hilbert spaces, while the operator \hat{V} provides the coupling between these spaces. In principle, the direct vacuum tunneling between the leads still can be described, but a device structure with a vacuum gap in the middle needs to be constructed. Thus, the device region can be constructed large enough to ensure that there is no direct interaction (non-zero matrix elements) between left and right leads, and the matrix elements of the T -matrix are given by:

$$\begin{aligned} \langle r | \hat{T} | l \rangle &= \langle r | \hat{V} + \hat{V} \hat{G} \hat{V} | l \rangle = \langle r | \hat{V}_{dr} + \hat{V}_{dr} \hat{G} \hat{V}_{dl} | l \rangle \\ &= \langle r | \hat{V}_{dr} \hat{G} \hat{V}_{dl} | l \rangle, \end{aligned} \quad (25)$$

where \hat{V}_{dl} and \hat{V}_{dr} are the couplings between the device and the left/right leads, respectively, and the interaction operator \hat{V} is a sum of these two couplings $\hat{V} = \hat{V}_{dl} + \hat{V}_{dr}$. Using the definitions of Eq. (25) the transmission function can, after some algebra, be put in the form given by Eq. (23) (for more details see [112]).

This formalism has been applied to the leakage problem in a Si–SiO₂–Si capacitor [21]. The atomic scale capacitor structures were generated using FIREBALL molecular-dynamical simulations, and the quantum leakage current was calculated using the LS formalism, results shown in Fig. 7 were in good agreement with experiment [119].

The periodic boundary conditions are applied in two lateral directions, and the two-dimensional Brillouin zone integrals are evaluated using the special k -points technique. We use the real space block-recursion scheme to describe semi-infinite leads in the direction of propagation. Complex self-energy operators take into account the open boundary condition. The presented theory suffers three major limitations: (i) it is not self-consistent in the applied field, (ii) the difficulties of the DFT with the excited states lead to incorrect tunneling barriers, and (iii) this is a scattering theory and its applicability to the non-equilibrium problem is not clear. The first and the third problems can be overcome in the non-equilibrium Green’s function formulation. However, the band gap problem is a serious limitation for

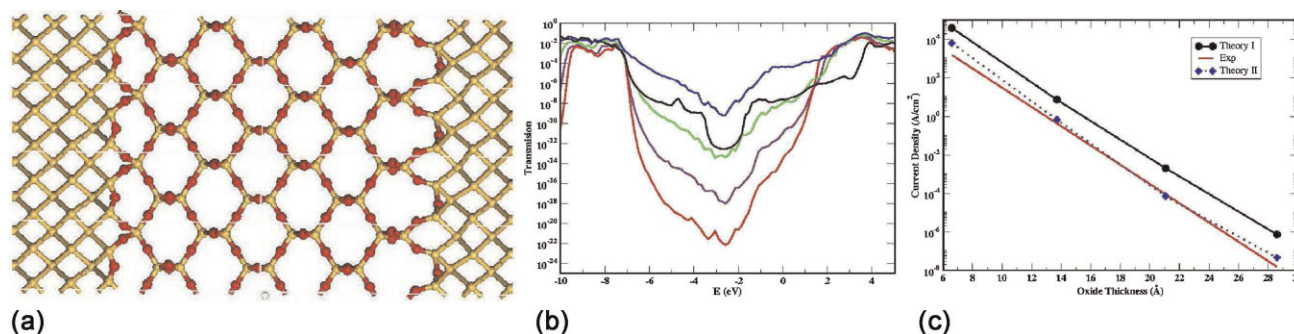


Figure 7 (online color at: www.pss-b.com) The quantum leakage current through an ultra thin silicon dioxide layer calculated within the scattering theory. Plots are (a) atomic level model of the Si-SiO₂-Si structure, (b) the transmission through the structures of different thickness, and (c) the theoretical tunneling current compared with experiment (Ref. [119]).

the methodology if applied to device simulations, and can be overcome only outside LDA. The recent development of hybrid potentials [120] may offer a practical solution of this difficulty.

Experimental and theoretical studies of the atomic, electronic, and transport properties of nanostructures and their mechanical stability have received a lot of attention in the last decade, due to its potential application in the electronic industry. Computer simulations play a significant role in exploration of mechanical and transport properties of nanostructures, but due to the complexity of these materials and processes, addressing them with fully-converged DFT methods is computationally very demanding. Efficient DFT methods, such as FIREBALL, with very favorable accuracy/efficiency balance are probably the best available tools for such calculations. Similar to the approach of Zhang et al. described above, approaches to the electron transport based on Green's function technique were also developed within the FIREBALL code by others [121–123].

The methodology was successfully applied to diverse problems such as the electron transport in atomic contacts or molecular junctions. For example, it has been used to simulate different metallic (Al, Pd, or Au) nanowires, included up to ~150 atoms, submitted to mechanical load up to the final breaking point [54, 124–127]. One of the most interesting outcome of these studies is the enhanced chemical activity found in mono-atomic gold nanowires [126, 127]. In addition, the role of irradiation induced defects and temperature in the conducting properties of single-walled carbon nanotubes has been analyzed by means of the FIREBALL code [128, 129]. Theoretical simulations pointed out strong influence of small number of vacancies on the electron transport regime, where strong Anderson localization effect appears and a seemingly universal curve exists for the resistance as a function of the number of defects.

This methodology has also been successfully applied to study the electron transfer in fully relaxed molecular junctions [130]. The electron transport through molecular junction is strongly affected by the Fermi level alignment and charge transfer on molecule/metal interfaces.

Understanding the mechanism responsible for the barrier height formation between a metal and an organic molecule or monolayer has been the subject of intensive debate. As mentioned above, the simulation of metal-organic systems using DFT is not a trivial task; in particular, the DFT energy band gap of the organic material is usually significantly underestimated. Recently, the transport gap and energy level alignment in the nanogap organic molecular junctions formed by a C₆₀ molecule in between two Au-tips [101], or in between an Au-tip and a Au(111) surface [131] have been analyzed using FIREBALL. In these calculations the charging energy of the molecule on the metal surface is calculated and used to properly determine the transport gap.

3.4 Simulations of scanning probe microscopy

Scanning Probe techniques, such as Atomic Force Microscope (AFM) and Scanning Tunneling Microscope (STM), have already achieved outstanding performance in characterization and modification of surfaces and nanostructures. However, further proliferation of the Scanning Probe methods is closely tied with detail understanding of undergoing processes during imaging or manipulation at atomic scale.

In particular, a proper interpretation of experimental STM measurements requires often sophisticated theoretical approach including several effects, such as influence of electronic and atomic structure of tip, the electron transfer going beyond the first order approximation or image potential effects. An important tool to achieve this goal is computationally efficient total energy DFT calculations using the FIREBALL code combined with the electron transport calculations based on Green's function technique (for detail description of methodology see [121, 132]). The theoretical approach is based on the calculation of the independent tip and sample electronic structures, and the coupling of both sides of the interface using Green's function formalism. These simulations included realistic tip models and electron transfer going beyond the first-order Bardeen approximation. In addition, an extended basis formalism, based on a local basis formalism, can also be used to calculate the tip

and sample structure [121]. In this approach, the tip sample interaction is obtained calculating the hopping integrals between the orbitals of both sides. The approach allows us to treat on the same foot both the tunneling and the contact regimes, where commonly used perturbation methods give incorrect results [133].

In this approach the Hamiltonian is written as the sum of three terms: $\hat{H} = \hat{H}^s + \hat{H}^t + \hat{H}^{t-s}$, with one term associated with the sample (s), one with the tip (t), and finally a term that takes into account the interaction between tip and sample ($t-s$). Then, we can write down an expression for the electrical current flowing between tip and sample is, that only involves the calculation of the Green's functions for the isolated tip and sample and the hopping matrix that couples both parts of the system

$$I = \frac{4\pi e}{h} \int_{-\infty}^{+\infty} dE \times [f_t(E) - f_s(E)] \times \text{Tr} \left[\hat{T}_{ts} \hat{\rho}_{ss}(E) \hat{D}_{ss}^r(E) \hat{T}_{st} \hat{\rho}_{tt}(E) \hat{D}_{tt}^a(E) \right], \quad (26)$$

where $\hat{T}_{st}, \hat{T}_{ts}$ are the hopping matrices, $f_i(E)$ ($i = t, s$) is the corresponding Fermi–Dirac distribution function, $\hat{\rho}_{tt}$ and $\hat{\rho}_{ss}$ are the density of states matrices associated with the tip and sample, while

$$\hat{D}_{ss}^r = [\hat{1} - \hat{T}_{ts} \hat{g}_{ss}^r(E) \hat{T}_{st} \hat{g}_{ss}^r(E)]^{-1}$$

and

$$\hat{D}_{tt}^a = [\hat{1} - \hat{T}_{ts} \hat{g}_{ss}^a(E) \hat{T}_{st} \hat{g}_{tt}^a(E)]^{-1}$$

where \hat{g}_{tt}^r and \hat{g}_{ss}^a mean the retarded and advanced Green's function for the tip and sample, respectively. Notice that the denominators \hat{D}_{ss}^r and \hat{D}_{tt}^a take into account the multiple scattering effects via the summation up to all orders of an expansion of the electron scattering across the tunneling barrier between tip and sample. Here, we should stress here these terms are responsible for the saturation of the tunneling current for small tip–sample distances (see *e.g.*, [122]).

3.4.1 STM surfaces The FIREBALL tool has been successfully adopted to analyze atomic scale STM images of semiconductor [134–136] and metal surfaces [36, 122]. In particular, the FIREBALL approach allows to perform complex studies of dependence of atomic scale contrast on atomic and chemical structure of probe [133, 137].

3.4.2 Ga/Si(112) The combination of MD capabilities of FIREBALL and the accurate simulation of STM images (based on the local-orbital Hamiltonian provided by FIREBALL) has proven very useful to unravel the atomic geometry for the quasi-one dimensional surface reconstruction of Ga on Si(112) [138, 139]. Recent progress in scanning probe instrumentation tends to combination of both AFM and STM into one tool, where the tunneling current and force is recorded simultaneously. Therefore, computer simulations analyzing both quantities from tunneling to contact

regime are necessary. The FIREBALL tool has been already used to analyze such dependence of the tunneling current and force along tip approach on nanotubes [140] and graphene [141], semiconductor [142], or metal [122, 143] surfaces.

3.4.3 AFM atomic and chemical resolution

Computer simulations of AFM consists in description of tip–sample interaction as function of many parameters, such as tip–sample distance or chemical and atomic structure of probe. What more, in AFM experiments probes are used to approach to distances where the chemical bond is established between outermost atoms of sample and probe. Therefore, a quantum description of the problem is mandatory. To fulfill these requirements large set of DFT simulations with different parameters/configurations is necessary. Hence, the FIREBALL technique seems to be an optimal choice for such task. Indeed the FIREBALL technique has been successfully used to understand AFM atomic scale contrast on semiconductor [144] or metal-oxides [145]. The computer simulations significantly contributed to a new method allowing the single-atom chemical identification on surfaces using force site spectroscopy [33].

3.4.4 AFM manipulation Understanding different mechanisms leading to manipulation of single atoms on surfaces call for DFT simulations exploring complex phase space of atomic configurations under different tip–surface interactions. After an optimal pathway between selected atomic configuration can be detected, functions of tip–sample distance, temperature or the chemical and atomic structure of different tips can be determined. The FIREBALL code has been successfully applied for description of the lateral single-atom manipulation on the Si(111)-(7 × 7) surface at room temperature performed in the attractive interaction regime [32] and the vertical atomic exchange between tip and sample realized in the strong interaction regime [34]. One of the main obstacles of SPM technique is knowledge of reliable atomic structure of probes. Recently, FIREBALL has been used as a tool for complex study, where reliable atomic structures of Si-based probes have been predicted [146].

3.4.5 AFM dissipation and Kelvin probe force microscopy (KPFM) In addition, DFT simulations performed with the FIREBALL code have shed more light into the energy dissipation at atomic scale [147–149] observed in AFM or the atomic contrast recently observed in KPFM [150]. In this case, simulations using the FIREBALL code provided more insight into the origin of the atomic resolution in KPFM due the electron density redistribution induced by a formation of the chemical bond between outermost atoms of tip and sample (see Fig. 8).

3.5 Disordered bulk systems One of the main frontiers of material science is the physics of disorder, namely answering the questions: “What is the nature of topological disorder in amorphous solids?” and “How does

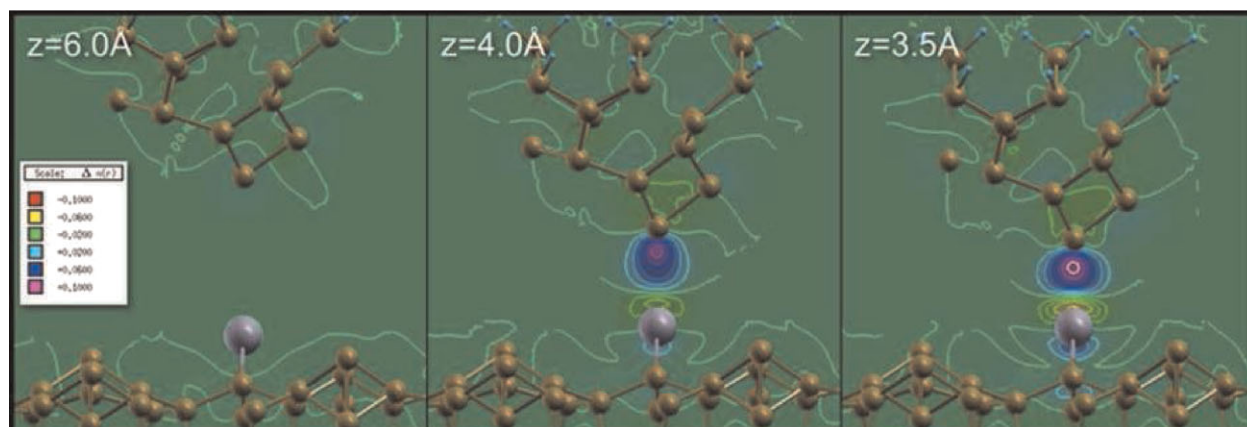


Figure 8 (online color at: www.pss-b.com) Calculated differential electron density ($\text{eV } \text{\AA}^3$) shows formation of the chemical bond between Si-based tip and Pb adatom on the Si(111)-(7 \times 7) surface during a tip approach (from left to right).

topological disorder influence physical observables?” The observables in question involve structural, vibrational, electronic, and optical properties. It is worth reflecting briefly on the state of this field when FIREBALL appeared. Empirical potentials reigned supreme: the Stillinger–Weber [151] and Tersoff [152] potentials were the prime tools for simulating the liquid and amorphous states. With 25 years between the 1985 papers and today, we can state that *ab initio* methods led to a genuinely revolutionary change in materials theory.

Amorphous semiconductors were a hot topic starting in the fifties and impressive strides were made with analytic theory in the early days, and leading scientists debated the most basic features of amorphous materials: Were they nanocrystalline?, Why did amorphous semiconductors possess a band gap (when it was usually stated that this required long-range order)?, What was the character of electronic states in a topologically disordered system?, and: Was the “continuous random network” picture of the topology of glasses due to Zacharisen [153] correct [154]? We now have fairly secure answers to these and other questions in part because the availability of accurate and efficient codes like FIREBALL and its application to these materials.

The lack of translational invariance (long-range order) in amorphous systems means that the standard apparatus of solid state physics based upon Bloch's theorem is useless. The advent of *ab initio* MD allowed for the creation of realistic computer models of materials that could be directly compared to experiment. The basic idea, borrowed from classical MD [155], was to make a model by superficially mimicking the glass transition: equilibrating a liquid and then slowly quenching it until atomic motion was arrested. To eliminate surface effects (but introduce other less destructive artifacts!), periodic boundary conditions were used, and the Brillouin zone of the supercell was usually ignored, sampling k integrals only at Γ . FIREBALL proved to be a great tool to create realistic computer models of amorphous materials and it enabled an answer to the most basic question

of all: “Where are the atoms?”, the overarching challenge to understanding disordered systems.¹

3.5.1 Amorphous silicon By current standards the first few FIREBALL calculations were rather primitive: involving typically 64 atoms, or more impressively (at the time!), 216 atoms. The first application of FIREBALL to a bulk amorphous material [158] compared the FIREBALL forces on each atom in a tiny 32-atom model of a-Si to those estimated by standard empirical potentials. The discrepancies were large, and as expected, most severe for highly strained configurations. This was perhaps the first quantitative measure of the inadequacy of the empirical potentials for representing a topologically disordered network. A useful fringe benefit of the calculation was an indication that computing the electronic structure only at the Γ point was seriously deficient for such a small system, a point that is now generally understood.

These early structural models provided much insight into the defect states (which are all important for electronic applications). Since the seminal work of Thorpe and Weaire [159] it was believed that a four-coordinated Si atom would produce a site-projected density of states with an optical gap. Drabold et al. [160] showed that sufficient bond angle strain led to midgap states. With these early models, and taking advantage of an excellent 216-atom model made by Djordjevic et al. [161], they next studied the dynamics of the a-Si models, with an emphasis on both the vibrational density of states, and the thermally induced fluctuations in the Kohn–Sham eigenvalues [10]. Experiments on the temperature dependence of the electronic band tails had just

¹ Since this paper is primarily a review of FIREBALL and its applications, we do not comment extensively on contemporary calculations. Still, for completeness, we should note that a number of important independent studies on disordered systems began appearing about the same time using plane-wave LDA, initially from the school of Car and coworkers [156]. Other workers were developing MD more in the spirit of the Sankey–Allen paper [1], performing *tight-binding MD* simulations [157].

appeared [162], and FIREBALL allowed us to make a direct link to the experiments, working out the physical differences between the broadening mechanisms intrinsic to the valence and conduction band tails. Their work noted for the first time that the electron–phonon coupling was strongly correlated with the spatial localization of the energy eigenstates, a result that proved to be far more general [163]. They also showed that a prevailing explanation for the existence of exponential band tails (the so-called Urbach edges) was incorrect. The inverse participation ratio was computed to characterize the spatial localization of the electron states and its energy dependence hinted at the nature of the mobility edge in a real material, an idea which led to subsequent insights on the Anderson transition [164]. The work has subsequently developed and led to new insight into origin of the Urbach edge in disordered systems [79, 165].

Based on a clever idea of Fedders [11], a first *ab initio* interpretation of the Staebler–Wronski effect, which is the experimental observation that light exposure creates structural (and electronic) defects in a-Si:H that act as carrier traps and reduce the efficiency of photovoltaics, was evaluated. The idea was to take the Kohn–Sham states as “genuine” single-particle eigenstates, and to treat low-lying unoccupied states as excited states. Thus, the approach by Fedders simply transferred an electron (or pair of electrons) from the HOMO state to the LUMO or adjacent levels by changing the occupations and the time evolution of the system in this artificially excited state was studied. The Hellmann–Feynman forces derived from occupying a well-localized LUMO state led to significant and highly non-local rearrangements, in agreement with experiments on the material [11]. While this *occupation change* model has serious shortcomings and ambiguities, it has the virtue that a realistic estimate of the electron–phonon coupling is included. It is probably still the most realistic scheme available, short of a proper non-adiabatic approach which at present is computationally difficult.

3.5.2 Amorphous chalcogenides Chalcogenide glasses (materials involving S, Se, or Te in elemental and alloyed forms) are of great interest for applications (examples include phase change and conducting bridge memory devices), and of scientific interest because for wide composition ranges, these materials are excellent glass formers and have the interesting feature that the composition may be continuously varied within the glass-forming window. Since the connectivity changes with changing composition (*e.g.*, as more highly coordinated atoms are added) the network becomes more rigid, this has led to fascinating work on vibrational phase transitions (the floppy to rigid transition [166]), and the more recent idea of the intermediate phase [167]. FIREBALL has been particularly successful with glassy chalcogenide materials, offering a nearly ideal combination of efficiency and accuracy.

The first realistic model of the classic glass GeSe₂ was made with FIREBALL in 1995 [168] in collaboration with experimentalists studying the vibrational density of states of

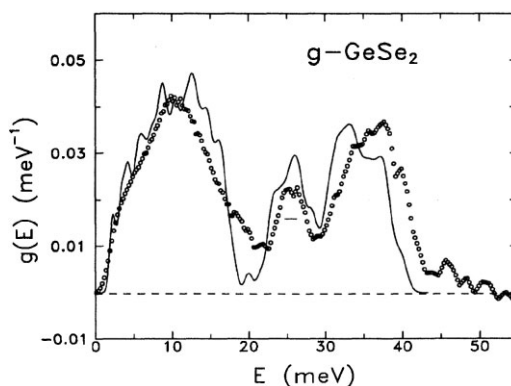


Figure 9 Vibrational density of states, from Ref. [168], experiment (due to Cappelletti and Kamitakahara) (open circles) and theory (solid line) for 63-atom model of glassy GeSe₂. The secondary *companion* mode of the central (A1) band is determined to be due to edge-sharing Se tetrahedra around Ge atoms.

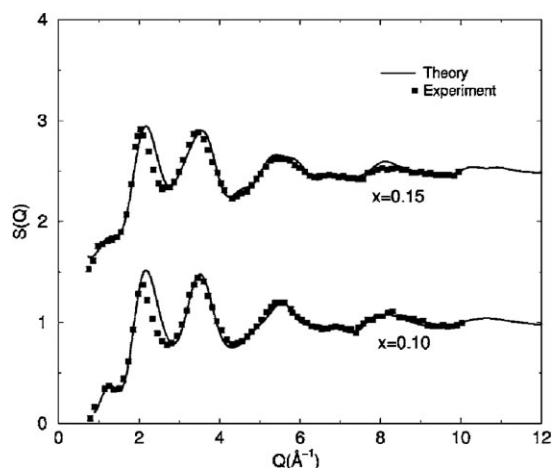


Figure 10 Calculated neutron structure factor $S(Q)$, from Ref. [170], of Ag-doped glasses compared to experiment.

the material. The amalgam of FIREBALL, the structural model, and Sankey’s scheme for computing the dynamical matrix from first principles² easily led to the solution of a much debated puzzle; a split A1 (tetrahedral breathing) band seen in Raman experiments, and its atomistic origin (shown in Fig. 9). By direct calculation with a fairly realistic model, we showed that the two humps in the central and were due to separate contributions from corner and edge sharing GeSe tetrahedra (Fig. 10). This was an early example of the success of FIREBALL as a “mathematical microscope” [169].

A technologically important and scientifically fascinating class of materials are the amorphous solid electrolytes: in our case, insulating glasses heavily doped with transition metals (most commonly Ag or Cu). Based on the pioneering experimental work of Kozicki and Mitkova (summarized in Ref. [164]), the compositions [GeSe₃]_{1-x}:Ag_x were studied,

² Historical aside: FORTRAN comments in this code included Beach Boys lyrics, that the reader can easily guess.

with $x = 0.1$ and 0.15 . FIREBALL was used to create the first computer models of this class of materials and two interesting facts emerged [170]. It is worth noting that this is a very complex material: it has an odd stoichiometry, with a Ge concentration inadequate to allow for a chemically ordered network, and more importantly populated with transition metal ions. *A priori*, the preferred positions for the Ag ions, and even the basic topology of the chalcogenide glass host were unknown.

Even with the application of a mindless “quench from the melt” simulation, 240-atom models were found to be in frankly impressive agreement with experiment (the neutron structure factors from experiment and theory are reproduced in Fig. 10). In the structure factors, every essential feature is present, even the shoulders near $Q = 1^{-1}$, a harbinger of the so-called “first sharp diffraction peak,” well known in studies in many glasses (and weak in this system because of the presence of the Ag and low concentration of Ge relative to the stoichiometric material, GeSe_2).

To comprehend the fundamental processes of ion conduction, and to forge links to experiments, the motion of the Ag is the key. To our initial surprise, we observed that the hopping of Ag ions could be directly simulated with FIREBALL employing using time scales that are easily accessible (albeit, at a somewhat exaggerated temperature). This is of course in stark contrast to the usual experience in MD that diffusive events are very hard to capture by simply integrating the equations of motion (*e.g.*, solving $\mathbf{F} = m\mathbf{a}$). The Ag is weakly enough bound to the host that in simulations under 100 ps, several trap-release events are observed (see Fig. 11).

The situation is somewhat similar to the case of proton hopping in Si [171, 172], where interesting hopping events occur on MD-accessible time scales.

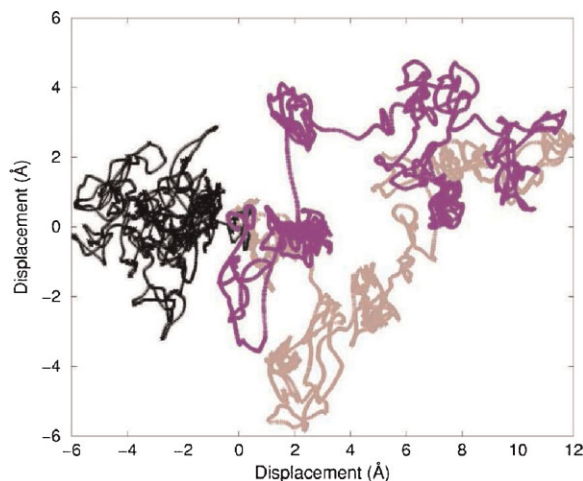


Figure 11 (online color at: www.pss-b.com) Two-dimensional projection of trajectories of the most diffusive (indigo and brown) and least diffusive (black) Ag atoms for GeSe_3 doped with 15 Ag, simulated at $T = 1000$ K, from Ref. [170]. The total simulation time was 62.5 ps.

3.6 Disordered molecular systems: Dynamically amorphous DNA Combining several theoretical techniques, we have developed an electronic-structure-based approach for performing large molecular-dynamical simulations of large biomolecular systems and we were the first to apply electronic-structure techniques to a 10-mer DNA molecule. For large systems such as the 10-mer DNA (644 atoms), we have implemented a variational linear-scaling technique to solve for the total energies and forces from the sparse Hamiltonian and overlap matrices [31]. Furthermore, we have developed a massively-parallel algorithm using message passing interface (MPI) to manipulate extremely large sparse matrices required for linear-scaling algorithms and have exhibited simulations of up to 6000 atoms [30]. The use of local-orbitals in the FIREBALL method yields a very sparse Hamiltonian matrix, which facilitates using a linear-scaling algorithm to obtain the electronic band-structure energy. Evaluation of forces is also written in a massively parallel procedure so that linear-scaling is achieved for MD simulations.

The mechanisms of electron or hole transport in DNA has been examined intensely over the past two decades as DNA has potential nanotechnological applications [173]. Our computational methodological developments in FIREBALL and its application to DNA have undoubtedly contributed to a better understanding of charge mechanisms in the molecule.

In one of our prominent DNA investigations [24], a state-of-the-art empirical force field MD simulations were coupled with FIREBALL to examine the effects of DNA structural changes on the charge transport. As demonstrated in a model periodic structure of poly(dA)-poly(dT) DNA (using an Arnott B-DNA model where each base pair is rotated by 36° and translated by 3.38 Å; thereby, 10 base pairs complete one full pitch of the double helix), the HOMO–LUMO electronic states are ideal periodic Bloch-like states (100 s of atoms participate in the state) which extend greatly throughout the molecule [174]. However, when thermally-driven dynamical fluctuations (no large-scale distortions) of the poly(dA)-poly(dT) DNA molecule are considered, strong Anderson localization results (the state is accessible to only about 30 atoms) in the electronic states near the band edge of the DNA molecule which we attribute to the off-diagonal disorder due to structural changes promoted by the thermal fluctuations. This is a similar localization that is observed in a-Si as discussed in the previous section.

Additionally, these band-edge states, which are predominantly located on a single adenine or thymine base, dynamically trading locations in time from one base to another base. Figure 12 shows the localized HOMO state relocating to different sites in the molecule as the structure of the molecule changes. This disorder results from dynamical variations of DNA with its environment which may very well suggest an important mechanism of charge transport along the DNA molecule. These simulations support experimental data (see Ref. [173] for a review) suggesting that these structural dynamical variations are the predominant

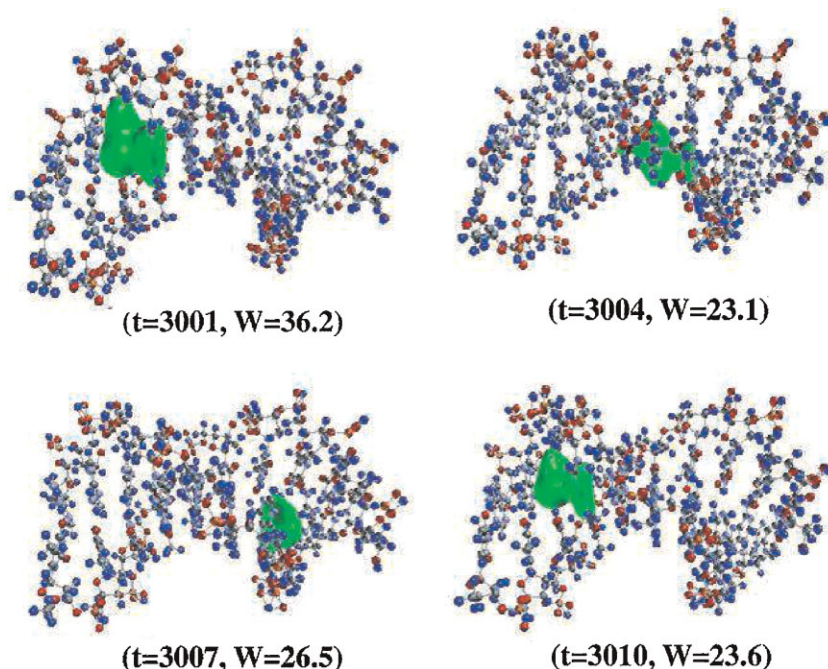


Figure 12 (online color at: www.pss-b.com) Population density plots for the localized HOMO state as a function of time. The time between snapshots is 1.5 ps.

mechanisms for charge transfer in DNA. This is strongly supported also from calculations of the band-gap tunneling states in DNA; we show that the tunneling decay is more rapid than many other organic molecules and that DNA is a poor tunneling conductor [175].

4 Summary FIREBALL has contributed to the physics of condensed matter directly, being the prime tool used in hundreds of papers from groups around the world, and also as a conceptual breakthrough that has motivated the creation of other powerful codes, especially SIESTA. It is worth reflecting on how condensed matter theory has changed since the advent of *ab initio* molecular dynamics in 1985. Many materials science and chemistry problems that were formerly untouchable by theory can now be routinely modeled in highly realistic ways. This is particularly true of complex systems with a structure that cannot be directly inferred from experiment. The legacy of FIREBALL and associated methods is a strong new thread between experiment and theory; Otto F. Sankey was one of the key *spiders* forging the link. The future is very bright indeed, with remarkable new links to new materials and experiments. The authors of this paper have been privileged to be working on these problems at an auspicious time of growth and development. We expect that the work will continue for many more academic generations.

Acknowledgements Several funding agencies have contributed to the ongoing development of FIREBALL and the reported results using FIREBALL. Otto F. Sankey held sustained support from George Wright at ONR through the years of development of FIREBALL and subsequently from NSF. Both Lewis and Drabold are currently funded by the National Science

Foundation through NSF DMR 09-03225. Lewis is also funded by the US Department of Energy under grant DE-FG02-10ER16164. Drabold would also like to thank the Army Research Office for support under MURI W91NF-06-2-0026. Haycock would like to acknowledge the support of the Irish Research Council for Science, Education and Technology under RS/2005/50. Jelínek acknowledges support of the GACR projects no. 202/09/0545 and 204/10/0952 and of the GAAV projects no. M100100904 and KAN400100701. Ortega acknowledges support from the Spanish MICIIN, project FIS2010-16046, and the Comunidad de Madrid, project S2009/MAT-1467. Demkov's work is supported by the National Science Foundation under grant DMR-0548182 and the US Department of Energy under grant DE-SC0001878.

References

- [1] O. F. Sankey and R. E. Allen, *Phys. Rev. B* **33**(10), 7164–7171 (1986).
- [2] O. F. Sankey and D. J. Niklewski, *Phys. Rev. B* **40**(6), 3979–3995 (1989).
- [3] D. R. Hamann, M. Schlüter, and C. Chiang, *Phys. Rev. Lett.* **43**(20), 1494–1497 (1979).
- [4] G. B. Bachelet, D. R. Hamann, and M. Schlüter, *Phys. Rev. B* **26**(8), 4199–4228 (1982).
- [5] J. Harris, *Phys. Rev. B* **31**(4), 1770–1779 (1985).
- [6] W. M. C. Foulkes and R. Haydock, *Phys. Rev. B* **39**(17), 12520–12536 (1989).
- [7] O. Sankey, A. Demkov, and T. Lenosky, *Phys. Rev. B* **57**, 15129–15139 (1998).
- [8] G. B. Adams, O. F. Sankey, J. B. Page, M. O’Keeffe, and D. A. Drabold, *Science* **256**(5065), 1792–1795 (1992).
- [9] G. B. Adams and O. F. Sankey, *Phys. Rev. Lett.* **67**(7), 867–870 (1991).
- [10] D. A. Drabold, P. A. Fedders, S. Klemm, and O. F. Sankey, *Phys. Rev. Lett.* **67**(16), 2179–2182 (1991).

- [11] P. A. Fedders, Y. Fu, and D. A. Drabold, *Phys. Rev. Lett.* **68**(12), 1888–1891 (1992).
- [12] P. Ordejón, E. Artacho, and J. M. Soler, *Phys. Rev. B* **53**(16), R10441–R10444 (1996).
- [13] A. Horsfield, *Phys. Rev. B* **56**(11), 6594–6602 (1997).
- [14] D. Sanchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler, *Int. J. Quantum Chem.* **65**(5), 453–461 (1997).
- [15] J. Soler, E. Artacho, J. Gale, A. Garcia, J. Junquera, P. Ordejón, and D. Sanchez-Portal, *J. Phys.: Condens. Matter* **14**(11), 2745–2779 (2002).
- [16] E. Artacho, E. Anglada, O. Diagnez, J. D. Gale, A. Garcia, J. Junquera, R. M. Martin, P. Ordejón, J. M. Pruneda, D. Sanchez-Portal, and J. M. Soler, *J. Phys.: Condens. Matter* **20**(6), 064208 (2008).
- [17] S. Boker, M. Neale, H. Maes, M. Wilde, M. Spiegel, T. Brick, J. Spies, R. Estabrook, S. Kenny, T. Bates, P. Mehta, and J. Fox, *Psychometrika* **76**, 306–317 (2011), DOI: 10.1007/s11336-010-9200-6.
- [18] A. A. Demkov, J. Ortega, O. F. Sankey, and M. P. Grumbach, *Phys. Rev. B* **52**(3), 1618–1630 (1995).
- [19] A. Demkov and O. Sankey, *Chem. Mater.* **8**(8), 1793–1806 (1996).
- [20] A. Demkov and O. Sankey, *Phys. Rev. B* **56**(16), 10497–10504 (1997).
- [21] A. Demkov and O. Sankey, *J. Phys.: Condens. Matter* **13**(46), 10433–10457 (2001).
- [22] P. Treesukol, J. P. Lewis, J. Limtrakul, and T. N. Truong, *Chem. Phys. Lett.* **350**(1–2), 128–134 (2001).
- [23] J. P. Lewis, K. R. Glaesemann, G. A. Voth, J. Fritsch, A. A. Demkov, J. Ortega, and O. F. Sankey, *Phys. Rev. B* **64**(19), 195103 (2001).
- [24] J. Lewis, *Chem. Phys. Lett.* **371**(5–6), 588–593 (2003).
- [25] J. Lewis, T. Sewell, R. Evans, and G. Voth, *J. Phys. Chem. B* **104**(5), 1009–1013 (2000).
- [26] H. Wang, J. Stalnakier, H. Chevreau, and J. P. Lewis, *Chem. Phys. Lett.* **457**(1–3), 26–30 (2008).
- [27] J. Ortega, J. Lewis, and O. Sankey, *Phys. Rev. B* **50**(15), 10516–10530 (1994).
- [28] J. Ortega, J. P. Lewis, and O. F. Sankey, *J. Chem. Phys.* **106**(9), 3696–3702 (1997).
- [29] P. Ordejón, *Comput. Mater. Sci.* **12**(3), 157–191 (1998).
- [30] S. Shellman, J. Lewis, K. Glaesemann, K. Sikorski, and G. Voth, *J. Comput. Phys.* **188**(1), 1–15 (2003).
- [31] J. Lewis, P. Ordejón, and O. Sankey, *Phys. Rev. B* **55**, 6880–6887 (1997).
- [32] Y. Sugimoto, P. Jelnek, P. Pou, M. Abe, S. Morita, R. Perez, and O. Custance, *Phys. Rev. Lett.* **98**(10), 106104 (2007).
- [33] Y. Sugimoto, P. Pou, M. Abe, P. Jelínek, R. Perez, S. Morita, and O. Custance, *Nature* **446**(7131), 64–67 (2007).
- [34] Y. Sugimoto, P. Pou, O. Custance, P. Jelínek, M. Abe, R. Perez, and S. Morita, *Science* **322**(5900), 413–417 (2008).
- [35] G. Otero, G. Biddau, C. Sanchez-Sanchez, R. Caillard, M. F. Lopez, C. Rogero, F. J. Palomares, N. Cabello, M. A. Basanta, J. Ortega, J. Mendez, A. M. Echavarren, R. Perez, B. Gomez-Lor, and J. A. Martin-Gago, *Nature* **454**(7206), 865–868 (2008).
- [36] G. Otero, C. González, A. L. Pinardi, P. Merino, S. Gardonio, S. Lizzit, M. Blanco-Rey, K. Van de Ruit, C. F. J. Flipse, J. Méndez, P. L. de Andrés, and J. A. Martín-Gago, *Phys. Rev. Lett.* **105**(21), 216102 (2010).
- [37] F. J. García-Vidal, J. Merino, R. Pérez, R. Rincón, J. Ortega, and F. Flores, *Phys. Rev. B* **50**(15), 10537–10547 (1994).
- [38] J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**(10), 5048–5079 (1981).
- [39] J. P. Perdew, *Phys. Rev. B* **34**(10), 7406 (1986).
- [40] C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**(2), 785–7789 (1988).
- [41] A. D. Becke, *Phys. Rev. A* **38**(6), 3098–3100 (1988).
- [42] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiollhais, *Phys. Rev. B* **46**(11), 6671–6687 (1992).
- [43] J. P. Perdew and Y. Wang, *Phys. Rev. B* **45**(23), 13244–13249 (1992).
- [44] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**(18), 3865–3868 (1996).
- [45] J. P. Perdew, K. Burke, and Y. Wang, *Phys. Rev. B* **54**(23), 16533–16539 (1996).
- [46] M. Fuchs and M. Scheffler, *Comput. Phys. Commun.* **119**, 67–69 (1999).
- [47] D. R. Hamann, *Phys. Rev. B* **40**(5), 2980–2987 (1989).
- [48] N. Troullier and J. L. Martins, *Phys. Rev. B* **43**(3), 1993–2006 (1991).
- [49] L. Kleinman and D. M. Bylander, *Phys. Rev. Lett.* **48**(20), 1425–1428 (1982).
- [50] X. Gonze, R. Stumpf, and M. Scheffler, *Phys. Rev. B* **44**(16), 8503–8513 (1991).
- [51] M. W. Finnis, *J. Phys.: Condens. Matter* **2**(2), 331 (1990).
- [52] B. Delley, *J. Chem. Phys.* **92**(1), 508–517 (1990).
- [53] M. Basanta, Y. Dappe, P. Jelínek, and J. Ortega, *Comput. Mater. Sci.* **39**(4), 759–766 (2007).
- [54] P. Jelínek, H. Wang, J. Lewis, O. Sankey, and J. Ortega, *Phys. Rev. B* **71**(23), (2005).
- [55] B. C. Carlson and J. M. Keller, *Phys. Rev. B* **105**(1), 102–103 (1957).
- [56] A. Fujishima and K. Honda, *Nature* **238**, 37–38 (1972).
- [57] R. Asahi, T. Morikawa, T. Ohwaki, K. Aoki, and Y. Taga, *Science* **293**, 269–271 (2001).
- [58] S. Khan, M. Al-Shahry, and J. W. B. Ingler, *Science* **297**, 2243–2245 (2002).
- [59] S. Sakthivel and H. Kisch, *Angew. Chem., Int. Ed.* **42**, 4908–4911 (2003).
- [60] H. Irie, Y. Watanabe, and K. Hashimoto, *Chem. Lett.* **32**, 772–773 (2003).
- [61] H. Irie, Y. Watanabe, and K. Hashimoto, *J. Phys. Chem. B* **107**, 5483–5486 (2003).
- [62] K. Noworyta and J. Augustynski, *Electrochem. Solid-State Lett.* **7**, E31–E33 (2004).
- [63] J. Burdett, T. Hughbanks, G. Miller, J. J. W. Richardson, and J. Smith, *J. Am. Chem. Soc.* **109**, 3639–3656 (1987).
- [64] H. Wang and J. P. Lewis, *J. Phys.: Condens. Matter* **18**, 421–434 (2005).
- [65] H. Wang and J. P. Lewis, *J. Phys.: Condens. Matter* **17**, L209–L213 (2005).
- [66] H. Wang and J. P. Lewis, *J. Phys. Chem. C* **113**(38), 16631–16637 (2009).
- [67] A. Selloni, *Nature Mater.* **7**(8), 613–615 (2008).
- [68] H. Zhang, B. Gilbert, F. Huang, and J. Banfield, *Nature* **424**, 1025–1029 (2003).
- [69] J. Wang, D. N. Tafen, J. P. Lewis, Z. Hong, A. Manivannan, M. Zhi, M. Li, and N. Wu, *J. Am. Chem. Soc.* **131**(34), 12290–12297 (2009), PMID: 19705915.
- [70] D. N. Tafen, J. Wang, N. Wu, and J. P. Lewis, *Appl. Phys. Lett.* **94**(9), 093101 (2009).

- [71] N. Wu, J. Wang, D. N. Tafen, H. Wang, J. G. Zheng, J. P. Lewis, X. Liu, S. S. Leonard, and A. Manivannan, *J. Am. Chem. Soc.* **132**(19), 6679–6685 (2010).
- [72] D. R. Alfonso, D. A. Drabold, and S. E. Ulloa, *Phys. Rev. B* **51**(20), 14669–14685 (1995).
- [73] D. R. Alfonso, C. Noguez, D. A. Drabold, and S. E. Ulloa, *Phys. Rev. B* **54**(11), 8028–8032 (1996).
- [74] D. R. Alfonso, S. H. Yang, and D. A. Drabold, *Phys. Rev. B* **50**(20), 15369–15380 (1994).
- [75] K. A. Kilian, D. A. Drabold, and J. B. Adams, *Phys. Rev. B* **48**(23), 17393–17399 (1993).
- [76] J. Dong and D. A. Drabold, *Phys. Rev. B* **57**(24), 15591–15598 (1998).
- [77] S. H. Yang, D. A. Drabold, and J. B. Adams, *Phys. Rev. B* **48**(8), 5261–5264 (1993).
- [78] X. Zhang and D. A. Drabold, *Phys. Rev. B* **62**(23), 15695–15701 (2000).
- [79] F. Inam and D. Drabold, *J. Non-Cryst. Solids* **354**(19–25), 2495–2499 (2008), Amorphous and Nanocrystalline Semiconductors, 22nd International Conference on Amorphous and Nanocrystalline Semiconductors – Science and Technology.
- [80] J. Ortega, F. Flores, and A. L. Yeyati, *Phys. Rev. B* **58**(8), 4584–4588 (1998).
- [81] F. Flores, J. Ortega, and R. Pérez, *Surf. Rev. Lett.* **6**, 411–433 (1999).
- [82] J. M. Carpinelli, H. H. Weitering, E. W. Plummer, and R. Stumpf, *Nature* **381**(6581), 398–400 (1996).
- [83] A. Mascaraque and E. G. Michel, *J. Phys.: Condens. Matter* **14**(24), 6005 (2002).
- [84] J. Ortega, R. Pérez, and F. Flores, *J. Phys.: Condens. Matter* **14**(24), 5979–6004 (2002), and references therein.
- [85] J. M. Carpinelli, H. H. Weitering, M. Bartkowiak, R. Stumpf, and E. W. Plummer, *Phys. Rev. Lett.* **79**(15), 2859–2862 (1997).
- [86] H. W. Yeom, S. Takeda, E. Rotenberg, I. Matsuda, K. Horikoshi, J. Schaefer, C. M. Lee, S. D. Kevan, T. Ohta, T. Nagao, and S. Hasegawa, *Phys. Rev. Lett.* **82**(24), 4898–4901 (1999).
- [87] J. Ortega, R. Perez, and F. Flores, *J. Phys.: Condens. Matter* **12**(1), L21–L27 (2000).
- [88] J. Avila, A. Mascaraque, E. G. Michel, M. Asensio, G. Lelay, J. Ortega, R. Perez, and F. Flores, *Phys. Rev. Lett.* **82**, 442–445 (1999).
- [89] R. Pérez, J. Ortega, and F. Flores, *Phys. Rev. Lett.* **86**(21), 4891–4894 (2001).
- [90] D. Farias, W. Kaminski, J. Lobo, J. Ortega, E. Hulpke, R. Perez, F. Flores, and E. G. Michel, *Phys. Rev. Lett.* **91**(1), 016103 (2003).
- [91] C. González, F. Flores, and J. Ortega, *Phys. Rev. Lett.* **96**(13), 136101 (2006).
- [92] C. González, J. Ortega, and F. Flores, *New J. Phys.* **7**, 100 (2005).
- [93] C. González, J. Guo, J. Ortega, F. Flores, and H. H. Weitering, *Phys. Rev. Lett.* **102**(11), 115501 (2009).
- [94] D. G. Trabada and J. Ortega, *J. Phys.: Condens. Matter* **21**(18), 182003 (2009).
- [95] E. Abad, Y. Dappe, J. Martinez, F. Flores, and J. Ortega, *J. Chem. Phys.* **134**(4), 044701 (2011).
- [96] J. I. Martínez, E. Abad, F. Flores, and J. Ortega, *Phys. Status Solidi B* **248**, 2044 (2011), this issue.
- [97] Y. Dappe, M. Basanta, F. Flores, and J. Ortega, *Phys. Rev. B* **74**(20), 205434 (2006).
- [98] Y. Dappe, J. Ortega, and F. Flores, *Phys. Rev. B* **79**(16), 165409 (2009).
- [99] F. Flores, J. Ortega, and H. Vazquez, *Phys. Chem. Chem. Phys.* **11**(39), 8658–8675 (2009).
- [100] J. D. Sau, J. B. Neaton, H. J. Choi, S. G. Louie, and M. L. Cohen, *Phys. Rev. Lett.* **101**, 026804 (2008).
- [101] E. Abad, C. González, J. Ortega, and F. Flores, *Org. Electron.* **11**(2), 332–337 (2010).
- [102] D. G. Trabada, F. Flores, and J. Ortega, *Phys. Rev. B* **80**(7), 075307 (2009).
- [103] M. Brandbyge, N. Kobayashi, and M. Tsukada, *Phys. Rev. B* **60**(24), 17064–17070 (1999).
- [104] J. K. Tomfohr and O. F. Sankey, *Phys. Status Solidi B* **226**(1), 115–123 (2001).
- [105] M. Brandbyge, J. L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, *Phys. Rev. B* **65**(16), 165401 (2002).
- [106] J. Tomfohr and O. F. Sankey, *J. Chem. Phys.* **120**(3), 1542–1554 (2004).
- [107] M. H. Lee, G. Speyer, and O. F. Sankey, *Phys. Status Solidi B* **243**(9), 2021–2029 (2006).
- [108] M. H. Lee, G. Speyer, and O. F. Sankey, *J. Phys.: Condens. Matter* **19**(21), 215204 (2007).
- [109] M. H. Lee and O. F. Sankey, *Phys. Rev. E* **79**(5), 051911 (2009).
- [110] M. H. Lee and O. F. Sankey, *J. Phys.: Condens. Matter* **21**(3), 035110 (2009).
- [111] A. A. Demkov, X. Zhang, and D. A. Drabold, *Phys. Rev. B* **64**(12), 125306 (2001).
- [112] X. Zhang, L. Fonseca, and A. Demkov, *Phys. Status Solidi B* **233**(1), 70–82 (2002).
- [113] N. A. Bruque, R. R. Pandey, and R. K. Lake, *Phys. Rev. B* **76**(20), 205322 (2007).
- [114] N. A. Bruque, M. K. Ashraf, G. J. O. Beran, T. R. Helander, and R. K. Lake, *Phys. Rev. B* **80**(15), 155455 (2009).
- [115] N. Sergueev, A. A. Demkov, and H. Guo, *Phys. Rev. B* **75**(23), 233418 (2007).
- [116] N. Sergueev and A. A. Demkov, *Phys. Rev. B* **81**(4), 045112 (2010).
- [117] W. Glockle, *The Quantum Mechanical Few-Body Problem* (Springer-Verlag, Paris, France, 1983).
- [118] R. Landauer, *IBM J. Res. Dev.* **1**, 223–231 (1957).
- [119] M. L. Green, T. W. Sorsch, G. L. Timp, D. A. Muller, B. E. Weir, P. J. Silverman, S. V. Moccio, and Y. O. Kim, *Microelectron. Eng.* **48**(1–4), 25–30 (1999), Insulating Films on Semiconductors.
- [120] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, *J. Chem. Phys.* **123**(17), 174101 (2005).
- [121] J. M. Blanco, F. Flores, and R. Pérez, *Prog. Surf. Sci.* **81**(10–12), 403–443 (2006).
- [122] J. M. Blanco, C. González, P. Jelínek, J. Ortega, F. Flores, and R. Pérez, *Phys. Rev. B* **70**(8), 085405 (2004).
- [123] P. Jelínek, R. Perez, J. Ortega, and F. Flores, *Surf. Sci.* **566–568**(Part 1), 13–23 (2004), Proceedings of the 22nd European Conference on Surface Science.
- [124] P. Jelínek, R. Pérez, J. Ortega, and F. Flores, *Phys. Rev. B* **68**(8), 085403 (2003).
- [125] B. Pieczyrak, C. González, P. Jelínek, R. Perez, J. Ortega, and F. Flores, *Nanotechnology* **19**(33), 335711 (2008).

- [126] P. Jelínek, R. Pérez, J. Ortega, and F. Flores, *Phys. Rev. B* **77**(11), 115447 (2008).
- [127] P. Jelínek, R. Pérez, J. Ortega, and F. Flores, *Phys. Rev. Lett.* **96**(4), 046803 (2006).
- [128] C. Gomez-Navarro, P. J. D. Pablo, J. Gomez-Herrero, B. Biel, F. J. Garcia-Vidal, A. Rubio, and F. Flores, *Nature Mater.* **4**(7), 534–539 (2005).
- [129] B. Biel, F. J. García-Vidal, A. Rubio, and F. Flores, *Phys. Rev. Lett.* **95**(26), 266801 (2005).
- [130] W. Kaminski, V. Rozsival, and P. Jelínek, *J. Phys.: Condens. Matter* **22**(4), 045003 (2010).
- [131] E. Abad, J. I. Martínez, J. Ortega, and F. Flores, *J. Phys.: Condens. Matter* **22**(30), 304007 (2010).
- [132] N. Mingo, L. Jurczyszyn, F. J. Garcia-Vidal, R. Saiz-Pardo, P. L. de Andres, F. Flores, S. Y. Wu, and W. More, *Phys. Rev. B* **54**(3), 2225–2235 (1996).
- [133] J. M. Blanco, C. González, P. Jelínek, J. Ortega, F. Flores, R. Pérez, M. Rose, M. Salmeron, J. Méndez, J. Wintterlin, and G. Ertl, *Phys. Rev. B* **71**(11), 113402 (2005).
- [134] W. Kaminski, P. Jelínek, R. Pérez, F. Flores, and J. Ortega, *Appl. Surf. Sci.* **234**(1–4), 286–291 (2004), The Ninth International Conference on the Formation of Semiconductor Interfaces.
- [135] B. Czech and B. Stankiewicz, *Appl. Surf. Sci.* **254**(14), 4279–4285 (2008), *Proc. 3rd Int. Workshop on Surface Physics – IWSP-2007*.
- [136] M. Švec, P. Jelínek, P. Shukryna, C. González, V. Cháb, and V. Drchal, *Phys. Rev. B* **77**(12), 125104 (2008).
- [137] C. Sanchez-Sanchez, C. González, P. Jelínek, J. Mendez, P. L. de Andres, J. A. Martin-Gago, and M. F. Lopez, *Nanotechnology* **21**(40), 405702 (2010).
- [138] C. González, P. C. Snijders, J. Ortega, R. Pérez, F. Flores, S. Rogge, and H. H. Weitering, *Phys. Rev. Lett.* **93**(12), 126106 (2004).
- [139] P. C. Snijders, E. J. Moon, C. González, S. Rogge, J. Ortega, F. Flores, and H. H. Weitering, *Phys. Rev. Lett.* **99**(11), 116102 (2007).
- [140] C. González, J. Ortega, F. Flores, D. Martínez-Martín, and J. Gómez-Herrero, *Phys. Rev. Lett.* **102**(10), 106801 (2009).
- [141] M. Ondráček, P. Pou, V. Rozsival, C. González, P. Jelínek, and R. Pérez, *Phys. Rev. Lett.* **106**(17), 176101 (2011).
- [142] P. Jelínek, M. Švec, P. Pou, R. Perez, and V. Cháb, *Phys. Rev. Lett.* **101**(17), 176101 (2008).
- [143] M. Ternes, C. González, C. P. Lutz, P. Hapala, F. J. Giessibl, P. Jelínek, and A. J. Heinrich, *Phys. Rev. Lett.* **106**(1), 016802 (2011).
- [144] Y. Sugimoto, P. Pou, O. Custance, P. Jelínek, S. Morita, R. Pérez, and M. Abe, *Phys. Rev. B* **73**(20), 205329 (2006).
- [145] R. Bechstein, C. González, J. Schutte, P. Jelínek, R. Perez, and A. Kohnle, *Nanotechnology* **20**(50), 505703 (2009).
- [146] P. Pou, S. A. Ghasemi, P. Jelínek, T. Lenosky, S. Goedecker, and R. Perez, *Nanotechnology* **20**(26), 264015 (2009).
- [147] N. Oyabu, P. Pou, Y. Sugimoto, P. Jelínek, M. Abe, S. Morita, R. Pérez, and O. Custance, *Phys. Rev. Lett.* **96**(10), 106101 (2006).
- [148] N. F. Martinez, W. Kaminski, C. J. Gomez, C. Albonetti, F. Biscarini, R. Perez, and R. Garcia, *Nanotechnology* **20**(43), 434021 (2009).
- [149] W. Kaminski and R. Perez, *Tribol. Lett.* **39**, 295–309(15) (2010).
- [150] S. Sadewasser, P. Jelínek, C. K. Fang, O. Custance, Y. Yamada, Y. Sugimoto, M. Abe, and S. Morita, *Phys. Rev. Lett.* **103**(26), 266103 (2009).
- [151] F. H. Stillinger and T. A. Weber, *Phys. Rev. B* **31**(8), 5262–5271 (1985).
- [152] J. Tersoff, *Phys. Rev. B* **38**(14), 9902–9905 (1988).
- [153] W. H. Zachariasen, *J. Am. Chem. Soc.* **54**(10), 3841–3851 (1932).
- [154] J. Ziman, *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems* (Cambridge University Press, Cambridge, 1979).
- [155] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1989).
- [156] I. Stich, R. Car, and M. Parrinello, *Phys. Rev. B* **44**(20), 11092–11104 (1991).
- [157] L. J. Lewis and N. Mousseau, *Comput. Mater. Sci.* **12**, 210 (1998).
- [158] D. A. Drabold, J. D. Dow, P. A. Fedders, A. E. Carlsson, and O. F. Sankey, *Phys. Rev. B* **42**(8), 5345–5348 (1990).
- [159] M. F. Thorpe and D. Weaire, *Phys. Rev. Lett.* **27**(23), 1581–1584 (1971).
- [160] D. A. Drabold, P. A. Fedders, O. F. Sankey, and J. D. Dow, *Phys. Rev. B* **42**(8), 5135–5141 (1990).
- [161] B. R. Djordjevic, M. F. Thorpe, and F. Wooten, *Phys. Rev. B* **52**(8), 5685–5689 (1995).
- [162] S. Aljishi, J. D. Cohen, S. Jin, and L. Ley, *Phys. Rev. Lett.* **64**(23), 2811–2814 (1990).
- [163] R. Atta-Fynn, P. Biswas, and D. A. Drabold, *Phys. Rev. B* **69**(24), 245204 (2004).
- [164] D. Drabold, *Eur. Phys. J. B* **68**, 1–21 (2009).
- [165] Y. Pan, F. Inam, M. Zhang, and D. A. Drabold, *Phys. Rev. Lett.* **100**(20), 206403 (2008).
- [166] J. C. Phillips and M. F. Thorpe, *Phase Transitions and Self-organization in Electronic and Molecular Networks* (Kluwer, Utrecht, 2001).
- [167] P. Boolchand, D. G. Georgiev, and B. Goodman, *J. Optoelectron. Adv. Mater.* **3**, 703–720 (2001).
- [168] R. L. Cappelletti, M. Cobb, D. A. Drabold, and W. A. Kamitakahara, *Phys. Rev. B* **52**(13), 9133–9136 (1995).
- [169] R. Godby, R. Needs, and M. Payne, *Phys. World* **39** (1990).
- [170] D. N. Tafen, D. A. Drabold, and M. Mitkova, *Phys. Rev. B* **72**(5), 054206 (2005).
- [171] F. Buda, G. L. Chiarotti, R. Car, and M. Parrinello, *Phys. Rev. Lett.* **63**(3), 294–297 (1989).
- [172] T. A. Abtew, F. Inam, and D. A. Drabold, *Europhys. Lett.* **79**(3), 36001 (2007).
- [173] J. C. Genereux and J. K. Barton, *Chem. Rev.* **110**(3), 1642–11662 (2010), PMID: 20214403
- [174] J. Lewis, J. Pikus, T. Cheatham, E. Starikov, H. Wang, J. Tomfohr, and O. Sankey, *Phys. Status Solidi B* **233**(1), 90–100 (2002).
- [175] H. Wang, J. Lewis, and O. Sankey, *Phys. Rev. Lett.* **93**, 016401 (2004).