Efficient iterative method for calculations of dielectric matrices

Hugh F. Wilson,[1] François Gygi,[2] and Giulia Galli[1]
[1]*Department of Chemistry, University of California, Davis, One Shields Avenue, Davis, California 95616, USA*
[2]*Department of Applied Science, University of California, Davis, One Shields Avenue, Davis, California 95616, USA*
(Received 13 May 2008; published 3 September 2008)

A major bottleneck in first-principles calculations of excited and spectroscopic properties of materials is the evaluation of dielectric matrices. We show that by computing a relatively small number of eigenvectors via iterative linear-response calculations, we may construct the dielectric matrix of a system to high numerical accuracy. We demonstrate the procedure on several systems. The proposed method bypasses the need for the calculation of large numbers of excited states required by widely used expressions within the random-phase approximation and opens the way to efficient calculations of excited-state properties in materials and nanostructures.

DOI: 10.1103/PhysRevB.78.113303 PACS number(s): 71.45.Gm, 71.15.Dx, 77.22.Ch

Atomic-scale variation in the dielectric screening properties of materials and nanostructures is central in understanding many phenomena. These include electronic excitations,[1] several forms of spectroscopy,[2,3] gate insulation in devices,[4] and electronic stopping in ion implantation.[5] In the linear regime, the atomic-scale static dielectric screening is expressed by the dielectric matrix $\epsilon(\mathbf{r},\mathbf{r}')$, which relates the external potential applied to a system, V_{ext}, and the screened potential V_{scr} which forms in response:

$$V_{\text{scr}}(\mathbf{r}) = \int \epsilon(\mathbf{r},\mathbf{r}')^{-1} V_{\text{ext}}(\mathbf{r}')d\mathbf{r}'. \quad (1)$$

Static dielectric matrices for bulk insulators and semiconductors have been calculated in works dating back to the 1970s.[6–9] More recently calculations for nanostructures as large as 1 nm Si dots[10] have appeared in the literature. However, computations of dielectric matrices are still limited to systems with relatively few atoms. Their efficient evaluation for larger systems would be particularly desirable, for example, to obtain the electronic self-energy Σ and thus quasiparticle and optical properties[1] within many-body perturbation theory (MBPT) and to evaluate spectroscopic properties of materials and nanostructures. Once the static dielectric matrix is known, in many cases the frequency-dependent response may be determined from the static response via a plasmon-pole model.[11]

In most prior studies,[6–10] dielectric matrices have been calculated via an approach based on perturbation theory under the random-phase approximation (RPA), or an extension thereof to include exchange-correlation effects.[9] Under this formalism and within a reciprocal-space representation, ϵ is given by[12,13]

$$\epsilon(\mathbf{G},\mathbf{G}') = \delta_{\mathbf{G},\mathbf{G}'} + \frac{8\pi e^2}{N\Omega|\mathbf{q}+\mathbf{G}|^2} \sum_{\mathbf{k},v,c} \frac{1}{E_c(\mathbf{k}+\mathbf{q}) - E_v(\mathbf{k})} \langle \mathbf{k} + \mathbf{q}, c | e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | \mathbf{k}, v \rangle \langle \mathbf{k}, v | e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}} | \mathbf{k}+\mathbf{q}, c \rangle, \quad (2)$$

where the conduction- and valence-band wave functions are indexed by c and v, respectively, $\mathbf{k}$ and $\mathbf{q}$ are wave vectors, and $\mathbf{G}$ and $\mathbf{G}'$ denote plane-wave components. The so-called direct methods for the calculation of the dielectric response have also been explored, but so far they have been restricted to special cases ($\mathbf{q}$ points with high symmetry).[14,15]

Calculation of dielectric matrices via RPA is computationally very demanding. First, the perturbation summation requires the evaluation of a large number of empty states $|c,\mathbf{k}+\mathbf{q}\rangle$, which in calculations carried out within density-functional theory (DFT) is usually a time-consuming procedure and the convergence of ϵ with respect to the number of empty states in the sum is slow.[9] Second, the evaluation of Eq. (2) involves the calculation and summation of a large number of terms, with every combination of c, v, and $\mathbf{k}$ treated separately. Third, within the RPA approach, one calculates the dielectric matrix ϵ. However, it is the inverse ϵ^{-1} that is required for most purposes (for example, self-energy calculations within MBPT), necessitating a costly large matrix inversion. For these reasons, calculations of dielectric matrices are possible for only relatively small systems. Consequently MBPT (e.g., GW calculations and solutions of the Bethe-Salpeter equation) has so far been applied to systems containing a few tens of atoms at most.

In this Brief Report, we present an *ab initio* method (referred to as the projective dielectric eigenpotential or PDEP method) for the calculation of dielectric matrices, based on an iterative projection of the eigenvectors of ϵ via repeated linear-response computations within density-functional perturbation theory (DFPT).[16] This approach bypasses the need for the costly calculation of and summation over unoccupied states and also avoids the inversion of the ϵ matrix. Thus, it opens the way to calculating dielectric matrices and performing many-body perturbation theory calculations for systems much larger than is presently feasible. We begin by investigating the dielectric eigenvalues of a variety of representative solid and molecular systems and showing that ϵ may be reconstructed to high numerical accuracy from a relatively small number of the most highly screened eigenvectors. We then use an orthogonal iteration procedure within DFPT to calculate these eigenvalues and eigenvectors, and we show that they match those obtained via direct diagonalization within the RPA. Finally, we demonstrate the application of the PDEP method to calculate the ϵ matrix of a realistic silicon nanostructure, and we discuss the efficiency and scaling of the proposed algorithm.

The eigenvectors of the dielectric matrix represent poten-

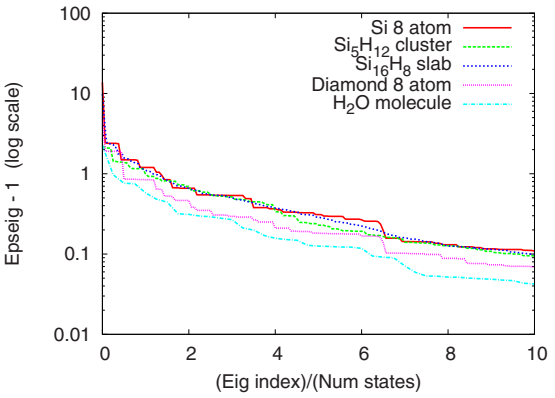

FIG. 1. (Color online) Eigenvalue spectra of the dielectric matrices of bulk Si, two Si nanostructures, bulk eight-atom diamond carbon, and an isolated water molecule, plotted on a logarithmic scale as a function of eigenvalue index. The x axis is normalized for each structure to the number of filled states in the system. All eigenvalues were calculated using the PDEP method described in the text.

tials which, when applied to the system, yield a screened potential of the same spatial form. Although ϵ is not a Hermitian matrix, a Hermitian form denoted $\tilde{\epsilon}$ may be obtained[6] by a similarity transformation, $\tilde{\epsilon}(\mathbf{q},\mathbf{G},\mathbf{G}')=(|\mathbf{q}+\mathbf{G}|/|\mathbf{q}+\mathbf{G}'|)\epsilon(\mathbf{q},\mathbf{G},\mathbf{G}')$. Note that ϵ and $\tilde{\epsilon}$ have the same eigenvalues λ_i, and the corresponding eigenvectors are simply related by a diagonal transformation. The matrix $\tilde{\epsilon}$ may be represented in terms of its eigenvectors $\tilde{\mathbf{v}}_i$ and eigenvalues λ_i as

$$\tilde{\epsilon}=\sum_i \tilde{\mathbf{v}}_i^H \lambda_i \tilde{\mathbf{v}}_i=\sum_i \tilde{\mathbf{v}}_i^H(\lambda_i-1)\tilde{\mathbf{v}}_i+\mathbf{I}. \quad (3)$$

All eigenvalues of the dielectric matrix are real and are known to be greater than or equal to 1.[8] From the form of Eq. (3), we may see that those eigenvalues which are very close to 1 will not have a significant impact on this sum. Thus the full dielectric matrix may be constructed to high accuracy using only those eigenvectors corresponding to eigenvalues significantly greater than 1. In practice we find that the number of such eigenvalues is relatively small: For the various systems examined so far, the eigenvalue spectrum decays to 1 quite rapidly, with a decay constant which appears to be proportional to the number of filled states in the system. This is illustrated in Fig. 1, which shows ϵ eigenvalue spectra (generated using the method described below) for three silicon-based systems (bulk, a 1.2 nm slab, and a Si_5H_12 cluster) as well as bulk diamond and an isolated water molecule. The x axis is scaled by the number of filled states for each system. The decay rates for the eigenvalues of the three silicon-based systems are almost identical, and the decay is slightly faster for the other two. Using a larger plane-wave basis set or a larger vacuum spacing for the molecular systems increases the size of the dielectric matrix and hence the total number of eigenvalues. However this does not significantly change the eigenvalue spectrum—the new eigenvalues thus added are very close to unity.

In order to efficiently perform the summation in Eq. (3), it is desirable to find a way of computing the few most significant dielectric eigenvectors without the diagonalization of the full matrix. We note that the eigenvectors of ϵ are the same as those of ϵ^{-1}, and computing the latter is equivalent to computing the response to an applied field to a given system using the DFPT (Ref. 16) linear-response formalism. By applying the method of orthogonal iteration (with Ritz acceleration as described in Ref. 17) with repeated applications of the operator $(I-\tilde{\epsilon}^{-1})$ to an initial set of orthogonal potentials, we may extract the most screened eigenvectors of ϵ^{-1} and use them to reconstruct the full matrix via Eq. (3). This procedure has been implemented in a code based on the linear-response DFPT subroutines of the QUANTUM-ESPRESSO (Ref. 18) package. In order to obtain results directly comparable to RPA results, the exchange-correlation part of the potential in the DFPT formalism has been disabled. However it may be easily included in this method.

As an initial demonstration of this scheme, we apply it to the calculation of the highest dielectric eigenstates of an eight-atom bulk silicon unit cell and compare them to the dielectric eigenstates obtained by diagonalization of the RPA ϵ matrix using DFT wave functions. Both calculations used norm-conserving pseudopotentials, a regular grid of 64 k points, and a $\mathbf{q}$ value of 0.01 in the $\mathbf{z}$ direction; the DFT-RPA calculations used 304 empty states. The PDEP iterations started from an orthogonal set of plane-wave potentials with a small random component added. These were considered to be converged when the difference between eigenvalues calculated at successive iterations was less than 0.01%. The top 20 eigenvalues found by the two methods are compared in Table I and are found to be identical to within 0.1%. Lower-lying eigenvalues are also found to be in close agreement, although the lower eigenvalues obtained within the RPA are highly sensitive to the number of empty states used in the summation.

Since the eigenstates of the ϵ matrix may have high degeneracy, it is not sensible to compare the eigenvectors obtained within RPA and PDEP directly. Instead we compare properties computed from the ϵ matrices reconstructed from finite sets of RPA and PDEP eigenpotentials. We denote these matrices as ϵ_M, where M is the number of eigenvectors included in the summation of Eq. (3). In Fig. 2 we plot the xy-planar-averaged dielectric response (the local ratio between the screened and applied fields) of bulk Si to a constant field in the z direction, as obtained from reconstructed ϵ_M matrices including the highest 1 (upper panel) and 256 (lower panel) eigenvectors, from both PDEP and RPA methods. The same quantity as obtained from the unreconstructed DFT-RPA ϵ calculation is also shown. The peaks in the graphs correspond to the planes which cut through the highly screening Si-Si bonds, while the troughs correspond to the planes which cut through the atomic cores. Figure 2 shows that the computed responses from ϵ_M matrices reconstructed using PDEP or RPA procedures with the same M are basically identical. Moreover, the lower panel shows that as the number of eigenvectors is increased, the ϵ_M matrix converges to the full RPA ϵ.

In Fig. 3 we demonstrate the application of the PDEP scheme to a larger nanostructured system, consisting of a 1.2 nm hydrogen-terminated $Si_{16}H_8$ slab with 1 nm of vacuum spacing. Again, the PDEP procedure was iterated until the

TABLE I. Top 20 eigenvalues of the dielectric matrix for the eight-atom silicon cubic cell calculated via the RPA and PDEP methods, using 64 k points.

Index	RPA	PDEP
1	14.7432	14.7538
2	3.4231	3.4237
3	3.3908	3.3914
4	3.3908	3.3914
5	3.3908	3.3914
6	3.3908	3.3914
7	3.3589	3.3596
8	2.4910	2.4925
9	2.4910	2.4925
10	2.4910	2.4925
11	2.4910	2.4925
12	2.4905	2.4920
13	2.4905	2.4920
14	2.4716	2.4721
15	2.1964	2.1972
16	2.1960	2.1968
17	2.1959	2.1966
18	2.1959	2.1966
19	2.1958	2.1966
20	2.1958	2.1966

difference between eigenvalues at successive iterations was less than 0.01%. As in the bulk silicon case, it was found that, e.g., the highest 20 eigenvalues matched those from an RPA calculation to within 0.5%. Figure 3 shows the xy-planar-averaged dielectric response to a uniform field in the z direction (normal to the slab) as computed from an ϵ_M matrix reconstructed from 100, 250, and 600 eigenvectors, compared to the response calculated from the full RPA ϵ matrix. It is apparent that calculation of the lowest 600 eigenstates for this system is sufficient to reproduce all the features of the full ϵ (with dimensionality of $\approx 3000 \times 3000$) to a high accuracy, and that inclusion of 250 eigenstates may for many applications be sufficient to reproduce the features of the response to an acceptable tolerance. The small remaining discrepancy between the 600-eigenvector PDEP and RPA curves must be attributed partially to incomplete convergence in the number of eigenstates in the PDEP calculation and partially to incomplete convergence in the number of empty states in the RPA calculation.

We now turn to a discussion of the efficiency of the PDEP method in comparison with that of the RPA. In an RPA calculation, the dominant term is the summation in Eq. (1), which takes a time proportional to $N_{pw\epsilon}^2 N_v N_c$, where $N_{pw\epsilon}$ is the size of the basis set used to represent the ϵ matrix and N_v and N_c are the number of filled and empty states in the sum, respectively. In the PDEP calculation, the most time-consuming part of the calculation consists of the repeated application of the potentials to the system via DFPT. Each linear-response calculation takes a time proportional to $N_{pw\psi} N_v^2$, with $N_{pw\psi}$ as the number of plane waves in the basis

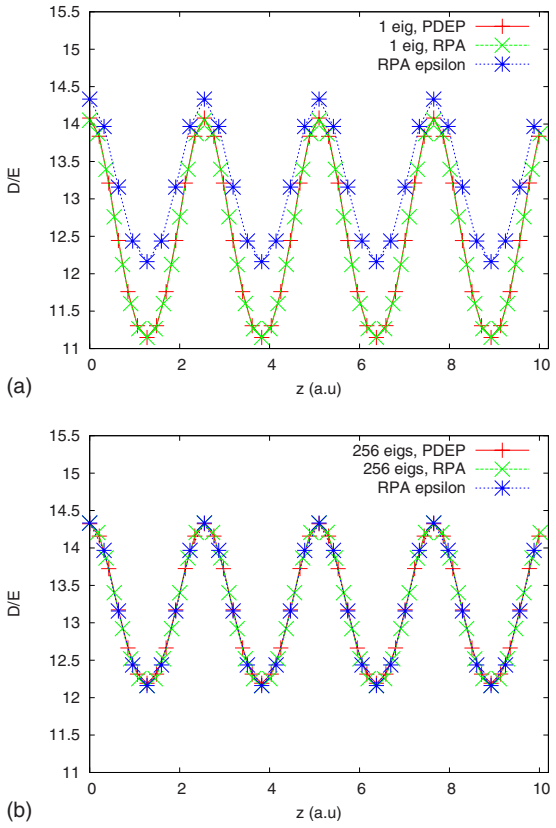

FIG. 2. (Color online) Planar-averaged D/E ratio of the induced (D) and applied (E) electric field magnitudes for an eight-atom Si cell subjected to a constant electric field along the **z** axis, calculated using ϵ_M reconstructed from 1 (upper panel) and 256 (lower panel) eigenvectors determined via the RPA and PDEP methods. The response calculated from an RPA-DFT calculation is shown for comparison.

set used to represent the wave functions. The total number of linear-response calculations equals the desired number of eigenstates (N_{eig}) times the number of iterations (N_{iter}) required to reach convergence. The latter of these factors is independent of system size, while the former is proportional

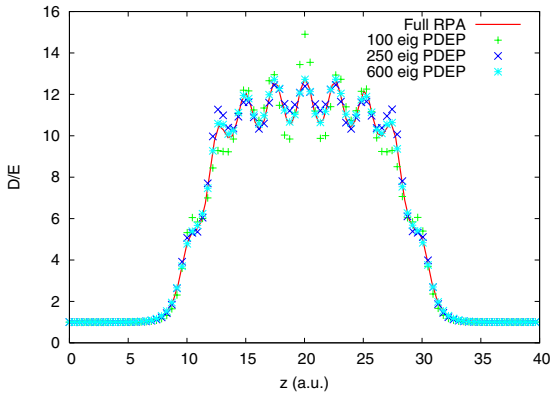

FIG. 3. (Color online) xy-planar-averaged dielectric response to a constant normally oriented field for a 1.2 nm H-terminated Si slab computed using the PDEP method (see text) with 100, 250, and 600 dielectric eigenstates. Also plotted is the same quantity computed from the RPA ϵ using 584 empty electronic states.

to it, and so the total scaling of the PDEP method is proportional to $N_{pw\psi}N_v^2N_{\mathrm{eig}}$. Thus, both methods scale with the fourth power of the size of the system.

However, the PDEP method has many significant advantages for many classes of system. Due to the iterative nature of the method, eigenstate convergence is reached rapidly if the initial guess of the dielectric eigenpotentials is close to the true eigenpotentials. This means that one may save considerably on CPU time by first finding the eigenpotentials for the same system using looser convergence parameters (e.g., basis set and k-point sampling) and then using these eigenpotentials as input into more refined calculations. As an illustrative example, in the above eight-atom Si calculations, it was found that fully converged eigenvalues could be obtained for the system with 64 k points in only two or three iterations if the eigenpotentials from a fully converged gamma-point calculation were used as the initial set of potentials. This makes the PDEP calculation significantly faster than the RPA calculation in our implementation. Furthermore, since the eigenstates of $\epsilon(\mathbf{G},\mathbf{G}',\mathbf{q})$ are similar to those of $\epsilon(\mathbf{G},\mathbf{G}',\mathbf{q}+\Delta\mathbf{q})$ for small $\Delta\mathbf{q}$, the method allows us to easily calculate ϵ matrices at a closely spaced set of $\mathbf{q}$ vectors, as required for several important applications.[1] In a similar fashion, the method allows the efficient calculation of dielectric properties of a series of steps in a molecular-dynamics trajectory. The PDEP method is extremely efficient in situations where only the few most highly screened eigenpotentials are required, such as, for instance, in the calculation of dielectric band structures.[6,8,19] Since the method allows for the construction of either ϵ or ϵ^{-1} with equal ease, the need for the time-consuming and difficult-to-parallelize matrix inversion step is bypassed, which is particularly advantageous for large systems. We note also that the PDEP procedure may be very efficiently parallelized, especially when the number of eigenstates to be calculated is greater than or equal to the number of CPUs. Finally, since the full ϵ matrix need never be calculated or stored, the PDEP method is efficient in memory use and disk space, again a particular advantage for large systems.

In summary, we have presented an iterative method for the calculation of static dielectric matrices of materials and nanostructures, based on the projection of the most highly screened dielectric eigenmodes. The method allows for the construction of either ϵ or ϵ^{-1} matrices at arbitrary $\mathbf{q}$ while bypassing the need for the calculation of empty electronic states and the inversion of ϵ. This opens the way to calculations of optical and excitonic properties of systems much bigger than are presently treatable using RPA-based techniques.

We acknowledge financial support from Intel and the Department of Energy's SciDAC program. We are grateful for useful discussions with Sebastien Hamel, Andrew Williamson, Ed Ratner, Dan Wack, and Deyu Lu.

[1] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[2] L. K. Dash, N. Vast, P. Baranek, M.-C. Cheynet, and L. Reining, Phys. Rev. B **70**, 245116 (2004).
[3] A. Diebold, *Handbook of Silicon Semiconductor Metrology* (CRC, Boca Raton, FL, 2001).
[4] F. Giustino and A. Pasquarello, Phys. Rev. B **71**, 144104 (2005).
[5] T. M. H. E. Tielens, G. E. W. Bauer, and T. H. Stoof, Phys. Rev. B **49**, 5741 (1994).
[6] A. Baldereschi and E. Tosatti, Phys. Rev. B **17**, 4710 (1978).
[7] A. Baldereschi and E. Tosatti, Solid State Commun. **29**, 131 (1979).
[8] R. Car, E. Tosatti, S. Baroni, and S. Leelaprute, Phys. Rev. B **24**, 985 (1981).
[9] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **35**, 5585 (1987).
[10] S. Ogut, R. Burdick, Y. Saad, and J. R. Chelikowsky, Phys. Rev. Lett. **90**, 127401 (2003).
[11] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).
[12] S. L. Adler, Phys. Rev. **126**, 413 (1962).
[13] N. Wiser, Phys. Rev. **129**, 62 (1963).
[14] K. Kunc and E. Tosatti, Phys. Rev. B **29**, 7045 (1984).
[15] A. Fleszar and R. Resta, Phys. Rev. B **31**, 5305 (1985).
[16] S. Baroni, S. de Gironcoli, A. D. Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).
[17] G. H. Golub and C. F. V. Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1983).
[18] S. Baroni *et al.*, http://www.pwscf.org/
[19] D. Lu, F. Gygi, and G. Galli, Phys. Rev. Lett. **100**, 147601 (2008).