Preconditioned iterative minimization for linear-scaling electronic structure calculations

Arash A. Mostofi,[a] Peter D. Haynes, Chris-Kriton Skylaris, and Mike C. Payne
Theory of Condensed Matter, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom

(Received 24 July 2003; accepted 6 August 2003)

Linear-scaling electronic structure methods are essential for calculations on large systems. Some of these approaches use a *systematic* basis set, the completeness of which may be tuned with an adjustable parameter similar to the energy cut-off of plane-wave techniques. The search for the electronic ground state in such methods suffers from an ill-conditioning which is related to the kinetic contribution to the total energy and which results in unacceptably slow convergence. We present a general preconditioning scheme to overcome this ill-conditioning and implement it within our own first-principles linear-scaling density functional theory method. The scheme may be applied in either real space or reciprocal space with equal success. The rate of convergence is improved by an order of magnitude and is found to be almost independent of the size of the basis. © *2003 American Institute of Physics.* [DOI: 10.1063/1.1613633]

I. INTRODUCTION

Conventional methods for electronic structure calculations, such as the plane-wave pseudopotential approach,[1] have proved to be indispensable tools for the study of condensed matter systems in a diverse range of disciplines.[2] The computational effort required, however, scales asymptotically with the *cube* of the system size, effectively placing a limit on the scientific problems that can be tackled with these approaches. *Linear-scaling* methods,[3,4] which exploit the real-space localization that is inherent in systems with a band gap,[5–8] should make these scientific problems tractable.

Several types of linear-scaling scheme exist and a point of commonality between many of them is the use of localized functions. Some of these approaches use a relatively small basis set of numerical atomic orbitals[9] or Gaussian atomic orbitals[10,11] that have been preoptimized for other environments and transferred to the system under consideration; other approaches[12–16] use much larger localized basis sets of simple functions such as polynomials,[17,18] spherical waves,[19] or bandwidth limited delta functions.[20] Each of these philosophies has its advantages and drawbacks: The former can suffer from transferability problems but is capable of providing good accuracy with modest effort; the latter is computationally more intensive but is capable of giving an accuracy that is *systematically* tunable with a parameter that controls the completeness of the basis set that is being used, akin to the kinetic energy cut-off in plane-wave methods. It is this latter category of method that we discuss here.

The usefulness of any linear-scaling scheme is ultimately determined by its *crossover* point, namely the system size at which the method begins to be faster than conventional cubic-scaling approaches. This crossover depends largely on two factors: First the computational cost per iteration per atom, and second the number of iterations required to reach a given convergence threshold per atom. Even if a method is constructed in which the computational cost per iteration per atom is small and independent of system size, the number of iterations required may be so large that the minimization is prohibitively inefficient. Indeed, it has been observed that methods which use large basis sets suffer from this very problem, known as *ill-conditioning*. We present a discussion of the origin of ill-conditioning and describe a general scheme to overcome it.

We briefly outline the formalism of linear-scaling methods in Sec. II. In Sec. III we discuss the cause of the above-mentioned ill-conditioning, and in Sec. IV, following the work of Bowler and Gillan,[21] we present a general preconditioning scheme for alleviating the problem. In particular, we show that the ''diagonal approximation'' that was invoked in Ref. 21 is unnecessary and we account for the tensorial nature of the nonorthogonal bases correctly. In Sec. V we extend our analysis to the case of an orthogonal basis, and in Sec. VI we use our linear-scaling method[16] as a specific example of the preconditioning scheme. Finally, in Sec. VII we present results that demonstrate the importance of using preconditioning.

II. FORMULATION OF THE PROBLEM

A system of noninteracting particles in a potential V is described by

$$\hat{H}\psi_n(\mathbf{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) \right] \psi_n(\mathbf{r}) = \epsilon_n \psi_n(\mathbf{r}), \tag{1}$$

where $\hat{H}$ is the single-particle Hamiltonian of the system, with energy eigenvalues ϵ_n and corresponding eigenstates $\psi_n(\mathbf{r})$. The eigenstates satisfy the orthogonality constraints given by

[a]Author to whom all correspondence should be addressed. Electronic mail: aam24@cam.ac.uk

$$\int \psi_m^*(\mathbf{r})\psi_n(\mathbf{r})d\mathbf{r}=\delta_{mn}, \tag{2}$$

for all m and n. For instance, within the Kohn–Sham scheme of density-functional theory,[22–24] $\hat{H}$ is the Kohn-Sham Hamiltonian and V is the effective potential.

The total band-structure energy is given by

$$E=\sum_n f_n\epsilon_n=\sum_n f_n\int \psi_n^*(\mathbf{r})\hat{H}\psi_n(\mathbf{r})d\mathbf{r}, \tag{3}$$

where f_n is the occupancy of state $\psi_n(\mathbf{r})$:[25] At the energy minimum, all states below and above the Fermi level have occupancy unity and zero, respectively.

In the case of linear-scaling calculations, the $\mathcal{N}$ lowest extended eigenstates $\psi_n(\mathbf{r})$ $(n\in\{1,\ldots,\mathcal{N}\})$ are expressed in terms of a set of $\mathcal{N}$ localized functions $\phi_\alpha(\mathbf{r})$ $(\alpha\in\{1,\ldots,\mathcal{N}\})$ that are generally nonorthogonal:

$$\psi_n(\mathbf{r})=\sum_\alpha \phi_\alpha(\mathbf{r})M^\alpha{}_n, \tag{4}$$

where $\mathbf{M}$ is a square ($\mathcal{N}$ by $\mathcal{N}$), nonsingular matrix of coefficients, and $\mathcal{N}$ can be equal to or greater than the number of occupied eigenstates. The overlap matrix $S_{\alpha\beta}$ of the localized functions $\phi_\alpha(\mathbf{r})$ is

$$S_{\alpha\beta}=\int \phi_\alpha^*(\mathbf{r})\phi_\beta(\mathbf{r})d\mathbf{r}, \tag{5}$$

and on substitution of Eq. (4) into the orthogonality relation given by Eq. (2) we find that $S_{\alpha\beta}$ satisfies

$$(M^\dagger)_n{}^\alpha S_{\alpha\beta}M^\beta{}_m=\delta_{nm}, \tag{6}$$

where a distinction has been made between contravariant and covariant quantities[26,27] through the use of superscript and subscript Greek suffixes, respectively.

Substituting Eq. (4) into the energy expression of Eq. (3), and defining

$$K^{\alpha\beta}=\sum_n M^\alpha{}_n f_n (M^\dagger)_n{}^\beta, \tag{7}$$

$$H_{\alpha\beta}=\int \phi_\alpha^*(\mathbf{r})\hat{H}\phi_\beta(\mathbf{r})d\mathbf{r}, \tag{8}$$

the band-structure energy becomes

$$E=\sum_{\alpha\beta} H_{\alpha\beta}K^{\beta\alpha}, \tag{9}$$

where $K^{\alpha\beta}$ is referred to as the *density kernel*.[28]

We consider the localized functions $\phi_\alpha(\mathbf{r})$ to be represented in terms of a basis $D_\mu(\mathbf{r})$ as follows:

$$\phi_\alpha(\mathbf{r})=\sum_\mu D_\mu(\mathbf{r})c^\mu{}_\alpha, \tag{10}$$

for some coefficients $c^\mu{}_\alpha$. As the basis functions $D_\mu(\mathbf{r})$ may be in general nonorthogonal, the tensor properties must be taken into account through the use of superscript and subscript Greek suffixes.

Defining

$$h_{\mu\nu}=\int D_\mu^*(\mathbf{r})\hat{H}D_\nu(\mathbf{r})d\mathbf{r}, \tag{11}$$

and using Eqs. (7)–(10), the energy may be written as

$$\begin{aligned}E&=(c^\dagger)_\alpha{}^\mu h_{\mu\nu}c^\nu{}_\beta K^{\beta\alpha}\\&=\sum_n f_n(M^\dagger)_n{}^\alpha(c^\dagger)_\alpha{}^\mu h_{\mu\nu}c^\nu{}_\beta M^\beta{}_n.\end{aligned} \tag{12}$$

Suffixes α and β run over the localized functions $\{\phi\}$, μ and ν run over the basis functions $\{D\}$, and n runs over the extended orthogonal orbitals $\{\psi\}$. We have adopted the Einstein summation convention for all repeated Greek suffixes, and continue to do so from here on.

It is both convenient and physically meaningful to perform the minimization of the energy functional in two nested loops, as in the ensemble density-functional method of Marzari *et al.*:[29] In the inner loop we minimize the energy with respect to the elements of the density kernel $K^{\alpha\beta}$ using one of a number of methods[30–32] to impose the constraint that the ground state density matrix be idempotent and give the correct number of electrons; in the outer loop we optimize the localized functions $\phi_\alpha(\mathbf{r})$ with respect to their coefficients $c^\mu{}_\alpha$ in the basis $D_\mu(\mathbf{r})$.[16]

III. PRINCIPLES OF KINETIC ENERGY ILL-CONDITIONING

The phenomenon of kinetic energy or length-scale ill-conditioning is a familiar one within the plane-wave approach to electronic structure calculations.[1] It is not, however, restricted to this approach and its effects are seen in many methods which use a large basis set.[21,33,34]

The efficiency with which a function can be minimized using iterative techniques such as steepest descents or conjugate gradients is related to the *condition number* $\kappa=\omega_{\max}/\omega_{\min}$, where $\omega_{\max}$ and $\omega_{\min}$ are the extremal curvatures of the function about the minimum.[35] Minimization is most efficient when the condition number is small and the curvatures have a narrow range of values. On the other hand, when the curvatures take a wide range of values, the number of iterations required for convergence can become unacceptably large and the minimization is said to be ill-conditioned.

The curvatures of the total energy functional are determined by the eigenvalues of the Hamiltonian. Hence, the condition number κ depends upon the ratio of the largest and smallest eigenvalues in the basis representation that is being used. With a *large* systematic basis, these eigenvalues span a broad range. As a result, the condition number is large, rendering the problem ill-conditioned. A significant source of this ill-conditioning is associated with the contribution to the total energy due to the kinetic energy E_{kin}, which is given by

$$E_{\mathrm{kin}}=-\frac{\hbar^2}{2m}\sum_n f_n\int \psi_n^*(\mathbf{r})\nabla^2\psi_n(\mathbf{r})d\mathbf{r}. \tag{13}$$

It is clear that high energy eigenstates are dominated by their large kinetic energy. These states contribute little to the total ground state energy, as they are unoccupied, yet they contribute greatly to the broadening of the eigenspectrum. The same argument does not hold, however, for the low-lying states for which the potential and kinetic contributions are

more closely matched. This ill-conditioning may be alleviated, or *preconditioned*, by removing the effect of the kinetic energy operator for the high energy states, making them more degenerate, and hence reducing the width of the eigenspectrum, whilst leaving the low energy states unchanged.

In the plane-wave approach, the effect of kinetic energy ill-conditioning is reduced by multiplying the steepest descents directions in reciprocal space by a diagonal preconditioning matrix which behaves as the inverse of the kinetic energy operator at high wave vectors and is a constant at low wave vectors.[1] Such a preconditioner, as pointed out in Ref. 21, is qualitatively equivalent to the *exact* preconditioner for a model Hamiltonian $\hat{X}$ given by

$$\hat{X}=1-k_0^{-2}\nabla^2. \tag{14}$$

The preconditioner for this model problem may be derived analytically in any basis, as shown in Sec. IV.

IV. GENERAL FORMALISM FOR KINETIC ENERGY PRECONDITIONING

We introduce a positive-definite model Hamiltonian $\hat{X}$ and write the energy of the system that it describes as

$$E_X=\sum_n \int \psi_n^*(\mathbf{r})\hat{X}\psi_n(\mathbf{r})d\mathbf{r}. \tag{15}$$

We proceed to derive exact expressions for preconditioning the minimization of Eq. (15). For suitable choice of $\hat{X}$, these same expressions may be used to improve the condition number for minimizing the true energy Eq. (3). It is worth noting that all of the occupation numbers f_n for the model system have been set to unity. This amounts to an additional *occupancy preconditioning*, first introduced by Gillan[36] in the context of metallic systems and then by Marzari *et al.*[29] in the general framework of ensemble density-funtional theory.

Following along the same lines as in Sec. II, defining

$$x_{\mu\nu}=\int D_\mu^*(\mathbf{r})\hat{X}D_\nu(\mathbf{r})d\mathbf{r}, \tag{16}$$

and substituting this, Eq. (4) and Eq. (10) into Eq. (15) we obtain

$$E_X=\sum_n (M^\dagger)_n{}^\alpha (c^\dagger)_\alpha{}^\mu x_{\mu\nu}c^\nu{}_\beta M^\beta{}_n. \tag{17}$$

It is at this point that a tensorially incorrect ''diagonal approximation'' is made in Ref. 21. In our notation, this would be given by

$$\sum_n M^\beta{}_n(M^\dagger)_n{}^\alpha=(S^{-1})^{\beta\alpha}\simeq J\delta_{\beta\alpha}, \tag{18}$$

where J is some constant, and the first equality follows from Eq. (6). We do not make this unnecessary approximation.

Formally, as it has been defined to be positive-definite, the matrix $\mathbf{x}$ may be expressed in terms of its unique Cholesky factor $\mathbf{G}$:[37]

$$x_{\mu\nu}=\sum_k G_{\mu k}(G^\dagger)_{k\nu}. \tag{19}$$

Substituting this into Eq. (17) gives

$$E_X=\sum_{kn} |a_{kn}|^2, \tag{20}$$

where the new variables a_{kn} which make the energy surface spherical are given by

$$a_{kn}=(G^\dagger)_{k\nu}c^\nu{}_\beta M^\beta{}_n. \tag{21}$$

In a steepest descents procedure, although the following easily generalizes to the conjugate gradients method, a line minimization is performed along the steepest descents search direction to find the new values of the coefficients a'_{kn}:

$$a'_{kn}=a_{kn}-\lambda\frac{\partial E_X}{\partial a_{kn}^*}, \tag{22}$$

where λ is chosen to minimize the energy. We wish to minimize the energy with respect to the coefficients $c^\mu{}_\alpha$, yet the functional is spherical (and hence preconditioned) in the new coefficients a_{kn}. In order to find the new values $c'^\mu{}_\alpha$ of the coefficients $c^\mu{}_\alpha$ that minimize the energy, we use the chain rule to write

$$\frac{\partial E_X}{\partial a_{kn}^*}=\frac{\partial E_X}{\partial c^{\mu *}_\alpha}\left(\frac{\partial c^\mu{}_\alpha}{\partial a_{kn}}\right)^*, \tag{23}$$

and from this, and Eqs. (21) and (22), it may be shown that

$$c'^\mu{}_\alpha=c^\mu{}_\alpha-\lambda(x^{-1})^{\mu\nu}\frac{\partial E_X}{\partial c^{\nu *}_\beta}S_{\beta\alpha}, \tag{24}$$

where we have used the relations

$$\sum_n (M^{-\dagger})_{\alpha n}(M^{-1})_{n\beta}=S_{\alpha\beta}, \tag{25}$$

and

$$\sum_k (G^{-\dagger})^\mu{}_k(G^{-1})_k{}^\nu=(x^{-1})^{\mu\nu}, \tag{26}$$

obtained from Eqs. (6) and (19), respectively.

Choosing the model Hamiltonian $\hat{X}$ introduced in Eq. (14), and defining

$$s_{\mu\nu}=\int D_\mu^*(\mathbf{r})D_\nu(\mathbf{r})d\mathbf{r}, \tag{27}$$

$$t_{\mu\nu}=-\int D_\mu^*(\mathbf{r})\nabla^2 D_\nu(\mathbf{r})d\mathbf{r}, \tag{28}$$

Eq. (24) becomes

$$c'^\mu{}_\alpha=c^\mu{}_\alpha-\lambda[(s+k_0^{-2}t)^{-1}]^{\mu\nu}\frac{\partial E}{\partial c^{\nu *}_\beta}S_{\beta\alpha}, \tag{29}$$

where, following the discussion in Sec. III, we have replaced the model energy E_X with the true energy E. We see from Eq. (29) that preconditioning is effected by premultiplying the steepest descent gradient by the matrix $(\mathbf{s}+\mathbf{t}/k_0^2)^{-1}$ and postmultiplying it by $\mathbf{S}$.

V. THE CASE OF AN ORTHOGONAL BASIS

In the special case of an orthogonal basis $\{D\}$ there is no distinction between covariant and contravariant quantities with respect to the expansion coefficients of this basis, and as such we use Latin suffixes to denote them: $D_i(\mathbf{r})$. In this case, Eq. (24) becomes

$$c'_{i\alpha}=c_{i\alpha}-\lambda\sum_j x_{ij}^{-1}g_j{}^{\beta}S_{\beta\alpha}, \tag{30}$$

where we have defined $g_j{}^{\beta}\equiv\partial E/\partial c^*_{j\beta}$.

Let $\mathbf{F}$ be that unitary transformation which diagonalizes the (Hermitian) matrix $\mathbf{x}$, i.e., $\mathbf{FxF}^{\dagger}=\tilde{\mathbf{x}}$, where $\tilde{\mathbf{x}}$ is a matrix with eigenvalues ξ_p on its diagonal:

$$\tilde{x}_{pq}=\xi_p\delta_{pq}. \tag{31}$$

Denoting transformed variables by $\tilde{v}_p=\Sigma_j F_{pj}v_j$, we apply $\mathbf{F}$ to Eq. (30) to obtain

$$\tilde{c}'_{p\alpha}=\tilde{c}_{p\alpha}-\lambda\sum_q \tilde{x}_{pq}^{-1}\tilde{g}_q{}^{\beta}S_{\beta\alpha}. \tag{32}$$

From Eq. (31) we see that $\tilde{x}_{pq}^{-1}=\xi_p^{-1}\delta_{pq}$ is diagonal, hence Eq. (32) becomes

$$\tilde{c}'_{p\alpha}=\tilde{c}_{p\alpha}-\lambda\frac{1}{\xi_p}\tilde{g}_p{}^{\beta}S_{\beta\alpha}. \tag{33}$$

In other words, for the case of an orthogonal basis $\{D\}$, the *transformed gradient* is preconditioned by premultiplying by a diagonal matrix of inverse eigenvalues ξ_p^{-1}. Postmultiplication by the overlap matrix $\mathbf{S}$ is still present in order to account for the non-orthogonality of the localized functions $\{\phi\}$.

VI. PRECONDITIONING AND PERIODIC SINC FUNCTIONS

We consider a unit cell (which we shall refer to as the simulation cell) with primitive lattice vectors $\mathbf{A}^{(i)}$ ($i\in\{1,2,3\}$), volume $V=|\mathbf{A}^{(1)}\cdot(\mathbf{A}^{(2)}\times\mathbf{A}^{(3)})|$, and $N_i=2J_i+1$ grid points along direction i, where the J_i are integers. Our basis set is composed of periodic bandwidth-limited delta functions,[20] from here on referred to as periodic sinc or psinc functions, defined as the following finite sum of plane waves:

$$D_{klm}(\mathbf{r})=D(\mathbf{r}-\mathbf{r}_{klm})=\frac{1}{N_1N_2N_3}\times\sum_{p=-J_1}^{J_1}\sum_{q=-J_2}^{J_2}\sum_{s=-J_3}^{J_3}e^{i(p\mathbf{B}^{(1)}+q\mathbf{B}^{(2)}+s\mathbf{B}^{(3)})\cdot(\mathbf{r}-\mathbf{r}_{klm})}, \tag{34}$$

where p, q, and s are integers, and the $\mathbf{B}^{(i)}$ are the reciprocal lattice vectors:

$$\mathbf{B}^{(1)}=\frac{2\pi}{V}(\mathbf{A}^{(2)}\times\mathbf{A}^{(3)}),\quad \text{etc.} \tag{35}$$

and the $\mathbf{r}_{klm}$ are the grid points of the simulation cell,

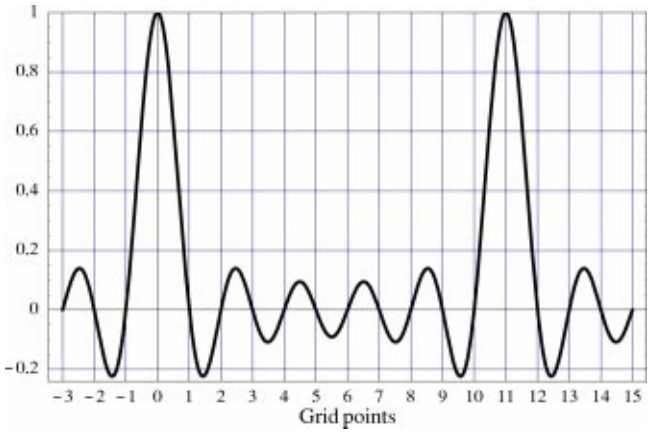

FIG. 1. One-dimensional analogue of a single periodic sinc, or psinc function, centered on the origin. In this example the simulation cell is eleven grid points in length.

$$\mathbf{r}_{klm}=\frac{k}{N_1}\mathbf{A}^{(1)}+\frac{l}{N_2}\mathbf{A}^{(2)}+\frac{m}{N_3}\mathbf{A}^{(3)}, \tag{36}$$

where k, l, and m are integers: $k\in\{0,1,\ldots,N_1-1\}$, and similarly for l and m. There is one psinc function centered on each grid point of the simulation cell.

The name ''periodic sinc,'' or psinc, has been chosen to reflect the connection that this function has with the familiar ''cardinal sine'' or sinc function. The sinc function is a continuous integral of plane waves with unit coefficients up to a maximum cut-off frequency. The psinc function differs only in that this continuous integral is replaced by a finite sum over the reciprocal lattice vectors of the simulation cell, as in Eq. (34). As a result, whereas the sinc function decays to zero at infinity, the psinc function is cell-periodic, namely $D(\mathbf{r})=D(\mathbf{r}+\mathbf{R})$, where $\mathbf{R}$ is any lattice vector. Figure 1 shows a one-dimensional analogue of a single psinc function.

From this point onward, for simplicity of notation, we write the psinc functions introduced in Eq. (34) as

$$D_i(\mathbf{r})=\frac{1}{N}\sum_p e^{i\mathbf{k}_p\cdot(\mathbf{r}-\mathbf{r}_i)}, \tag{37}$$

where $\mathbf{k}_p$ denotes a reciprocal lattice point, $\mathbf{r}_i$ denotes a grid point of the simulation cell, and $N=N_1N_2N_3$ is the total number of grid points in the simulation cell.

Using the same model Hamiltonian $\hat{X}$ given by Eq. (14) along with the definitions presented in Eqs. (27) and (28), we write

$$x_{ij}=s_{ij}+k_0^{-2}t_{ij}. \tag{38}$$

As shown in the Appendix, the psinc functions are orthogonal,

$$s_{ij}=w\,\delta_{ij}, \tag{39}$$

and the matrix elements of $-\nabla^2$ in the psinc basis are given by

$$t_{ij}=\frac{w}{N}\sum_p k_p^2 e^{i\mathbf{k}_p\cdot(\mathbf{r}_i-\mathbf{r}_j)}, \tag{40}$$

where $w=V/N$, the grid point weight, and $k_p=|\mathbf{k}_p|$.

The operator **F** which diagonalizes **x** is none other than the discrete Fourier transform:

$$\tilde{b}_p = \sum_j F_{pj} b_j \equiv \frac{1}{\sqrt{N}} \sum_j b_j e^{-i\mathbf{k}_p \cdot \mathbf{r}_j}, \tag{41}$$

$$b_i = \sum_p F^{\dagger}_{ip} \tilde{b}_p \equiv \frac{1}{\sqrt{N}} \sum_p \tilde{b}_p e^{i\mathbf{k}_p \cdot \mathbf{r}_i}, \tag{42}$$

where the b_i are values on the real-space grid and the $\tilde{b}_p$ are values on the reciprocal-space grid. Using these definitions, along with Eqs. (38)–(40) and Eq. (A3), it is a simple matter to show that

$$\tilde{x}_{pq} = \sum_{ij} F_{pi} x_{ij} F^{\dagger}_{jq} = w\left(1 + \frac{k_p^2}{k_0^2}\right)\delta_{pq}. \tag{43}$$

Thus the eigenvalues ξ_p of **x** are given by $\xi_p = w(1 + k_p^2/k_0^2)$. Substituting this into Eq. (33) gives the final expression for our preconditioned line minimization:

$$\tilde{c}'_{p\alpha} = \tilde{c}_{p\alpha} - \frac{\lambda}{w} \frac{k_0^2}{k_0^2 + k_p^2} \tilde{g}_p{}^{\beta} S_{\beta\alpha}. \tag{44}$$

VII. RESULTS

We present some illustrative examples of the importance of kinetic energy preconditioning for the convergence of calculations with our method, described in more detail in Ref. 16. In all test cases we use norm-conserving pseudopotentials in Kleinman–Bylander[38] form, the local-density approximation[39,40] for the exchange and correlation term, and the Γ point only for the k-point sampling.

A silane molecule is placed in a cubic simulation cell of side length 40 a_0, with a grid-spacing 0.5 a_0 (corresponding to a plane-wave cut-off of 537 eV) in each direction. The orbitals are initialized to atom-centered fireballs[41] which are strictly localized within spheres of radius 6.0 a_0. Each orbital is allowed to vary freely within its localization region. There is one orbital on each hydrogen atom and four on the silicon. In Fig. 2 we show the convergence of the total energy as a function of iteration number. The effect of using different fixed values of the kinetic energy preconditioning parameter k_0 may be seen. The limit $k_0 = \infty$ corresponds to the case of no preconditioning. It can be seen that improved performance is achieved for a range of values of k_0.

In Figs. 3 and 4 we show the convergence of the total energy as a function of iteration number for different grid spacings and localization radii, respectively. For the calculations presented in these two figures we used a kinetic energy preconditioning parameter $k_0 = 4.0\ a_0^{-1}$. As the grid spacing is reduced, or the localization radius increased, the size of the basis set and the number of variational parameters in the minimization increases. From Figs. 3 and 4 it is clear that the preconditioning scheme is working well as the number of iterations required to reach a given accuracy does not vary a great deal with the size of the problem. For instance, in Fig. 3, we see that the calculation with a grid spacing of 1.0 a_0 (134 eV) reaches an energy convergence of 10^{-6} hartree after 11 iterations, whilst with a grid spacing of 0.4 a_0 (839 eV), i.e., almost 16 times as many basis functions, the same level of convergence is achieved in just 14 iterations.

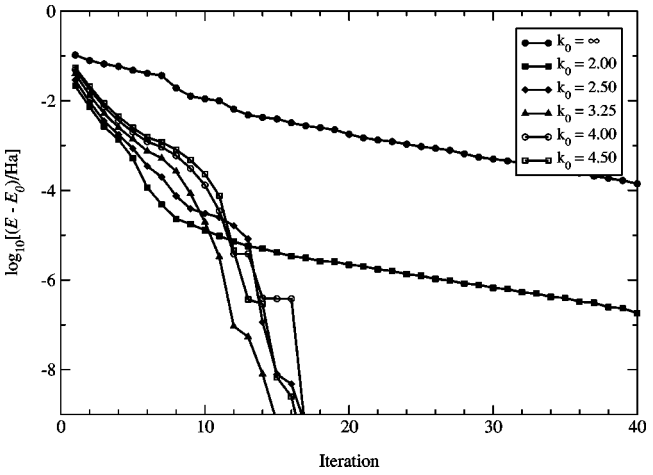

FIG. 2. Convergence of the total energy as a function of the iteration number for the calculation on a silane molecule. The grid spacing was 0.5 a_0 and the localized orbitals were confined within spheres of radius 6.0 a_0. E_0 is the converged value of the total energy for each run, and k_0 is given in units of a_0^{-1}.

Finally, as an alternative to preconditioning the Fourier transformed gradient $\tilde{g}_p{}^{\beta}$ by multiplying it with the preconditioner ξ_p in reciprocal space, in accord with Eq. (44), we have developed a real-space implementation of the preconditioning scheme. In this we *convolve* the real-space gradient $g_i{}^{\alpha}$ with the inverse fast Fourier transform (FFT) of ξ_p. Of course, a full convolution would be costly: If the gradient and preconditioner are both of size N_{grad}, then the computational effort required to perform a full convolution scales as N_{grad}^2. Thus, we truncate the preconditioner in real space at a radial cut-off R_0 so that it is nonzero over only a small number of points $N_{\mathrm{prec}} \ll N_{\mathrm{grad}}$. The computational cost of performing a convolution between the gradient $g_i{}^{\alpha}$ and this truncated preconditioner is much more favorable and scales

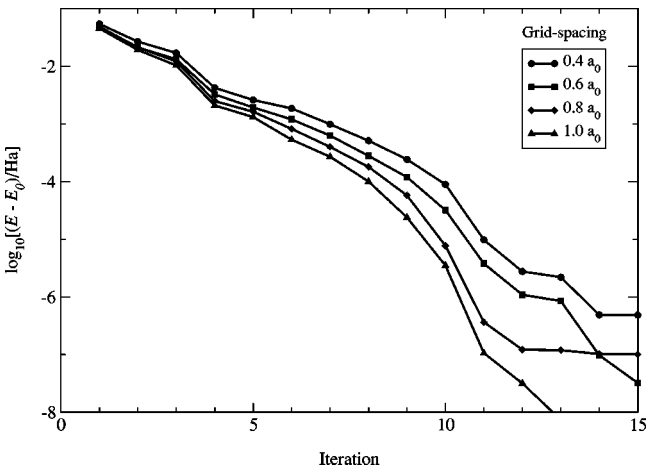

FIG. 3. Convergence of the total energy as a function of the iteration number for the calculation on a silane molecule. The localized orbitals were confined within spheres of radius 6.0 a_0 and kinetic energy preconditioning with $k_0 = 4.0\ a_0^{-1}$ was used. E_0 is the converged value of the total energy for each run.

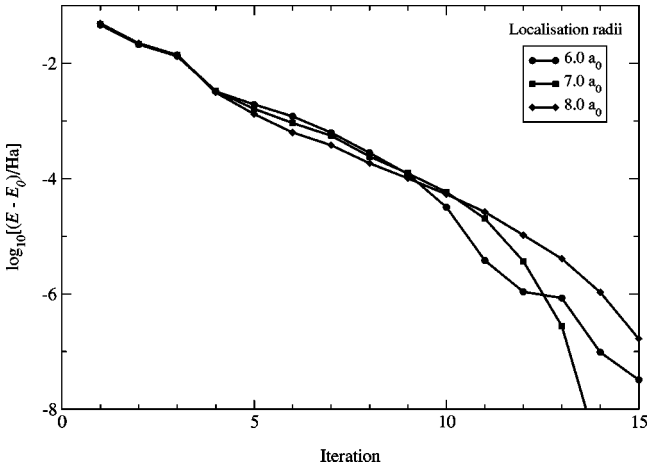

FIG. 4. Convergence of the total energy as a function of the iteration number for the calculation on a silane molecule. The grid spacing was 0.5 a_0 and kinetic energy preconditioning with $k_0=4.0\ a_0^{-1}$ was used. E_0 is the converged value of the total energy for each run.

as $N_{\mathrm{prec}}N_{\mathrm{grad}}$. For typical values of N_{grad} and N_{prec}, this is comparable to the cost of preconditioning in reciprocal-space.

Truncating the preconditioner in real space is not simply a matter of improving the computational efficiency, for it also makes physical sense: The reason behind preconditioning is to smear out large kinetic energy variations over short distances, thus a convolution that is *localized* in real-space over just a few grid points is all that should be required. This is demonstrated by the results presented in Fig. 5, which shows the convergence of the total energy with this real-space scheme for the above-introduced silane molecule. The different curves correspond to various radial cut-offs R_0 for the inverse FFT of the preconditioning function ξ_p. Comparing Figs. 2 and 5 we see that preconditioning via local convolution in real space is as successful as the conventional reciprocal-space approach, and that there is little sensitivity to the choice of the cut-off radius R_0, which may be as small as 1.0 a_0.

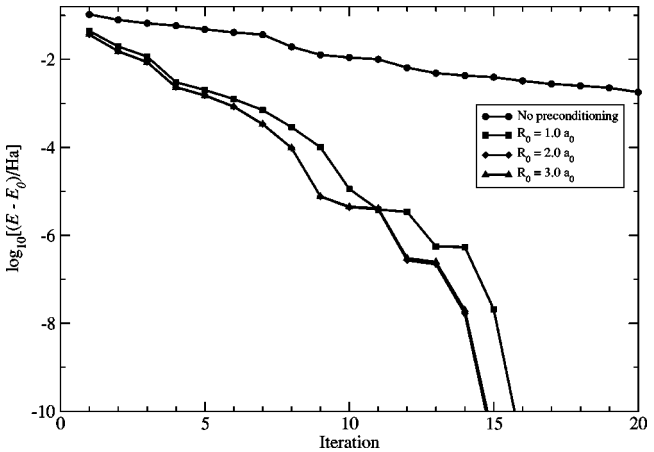

FIG. 5. Convergence of the total energy as a function of the iteration number for the calculation on a silane molecule. The grid spacing was 0.5 a_0 and the localized orbitals were confined within spheres of radius 6.0 a_0. The top curve is for the case of no preconditioning ($k_0=\infty$), while for the others $k_0=3.0\ a_0^{-1}$. R_0 is the convolution radius in real-space. E_0 is the converged value of the total energy for each run.

VIII. CONCLUSIONS

We have presented a preconditioning scheme to improve the convergence of iterative steepest descents or conjugate gradients total energy minimizations. We have derived a general expression for this preconditioning scheme for nonorthogonal basis sets. For the special case of orthogonal basis sets, we have showed that a unitary transformation may be made to a representation in which the preconditioning function is diagonal. In our linear-scaling density functional theory method, which uses an orthogonal basis set of periodic sinc (psinc) functions, this representation is accessed via discrete Fourier transformation: In other words, the preconditioning function is diagonal in reciprocal space. We have also developed an efficient and physically motivated preconditioning scheme which uses a localized convolution directly in real space, with no need for fast Fourier transforms. Both of these approaches (reciprocal space and real space) significantly improve the rate of convergence, and this improvement is found to be almost independent of the size of the basis set.

ACKNOWLEDGMENTS

One of the authors (A.A.M.) acknowledges the EPSRC for a Ph.D. studentship, and Christ's College, Cambridge for its support. P.D.H. acknowledges Sidney Sussex College, Cambridge for a Research Fellowship. C.-K.S. acknowledges the EPSRC for postdoctoral research funding (Grant No. GR/M75525). We thank Dr. Hadrian Green for reading the manuscript.

APPENDIX: THE PSINC BASIS

The overlap matrix **s** of the psinc functions defined in Eq. (37) is given by

$$\begin{aligned} s_{ij} &= \int D_i^*(\mathbf{r}) D_j(\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{N^2} \sum_{pq} e^{i\mathbf{k}_p\cdot\mathbf{r}_i - i\mathbf{k}_q\cdot\mathbf{r}_j} \int e^{i(\mathbf{k}_q-\mathbf{k}_p)\cdot\mathbf{r}} d\mathbf{r} \\ &= \frac{V}{N^2} \sum_{pq} e^{i\mathbf{k}_p\cdot\mathbf{r}_i - i\mathbf{k}_q\cdot\mathbf{r}_j} \delta_{pq} \\ &= \frac{V}{N^2} \sum_{p} e^{i\mathbf{k}_p\cdot(\mathbf{r}_i-\mathbf{r}_j)} = w\,\delta_{ij}, \end{aligned} \tag{A1}$$

where $w=V/N$ is the grid point weight, and we have used the relations

$$\int e^{i(\mathbf{k}_p-\mathbf{k}_q)\cdot\mathbf{r}} d\mathbf{r} = V\delta_{pq}, \tag{A2}$$

and

$$\sum_p e^{i\mathbf{k}_p\cdot(\mathbf{r}_i-\mathbf{r}_j)} = N\delta_{ij}. \tag{A3}$$

Furthermore, the matrix elements of $-\nabla^2$ in this basis are given by

$$
\begin{aligned}
t_{ij} &= -\int D_i^*(\mathbf{r})\nabla^2 D_j(\mathbf{r})d\mathbf{r} \\
&= -\frac{1}{N^2}\sum_{pq} e^{i\mathbf{k}_p\cdot\mathbf{r}_i - i\mathbf{k}_q\cdot\mathbf{r}_j}\int e^{-i\mathbf{k}_p\cdot\mathbf{r}}\nabla^2 e^{i\mathbf{k}_q\cdot\mathbf{r}}d\mathbf{r} \\
&= \frac{1}{N^2}\sum_{pq} e^{i\mathbf{k}_p\cdot\mathbf{r}_i - i\mathbf{k}_q\cdot\mathbf{r}_j}|\mathbf{k}_q|^2\int e^{i(\mathbf{k}_q-\mathbf{k}_p)\cdot\mathbf{r}}d\mathbf{r} \\
&= \frac{V}{N^2}\sum_{pq} e^{i\mathbf{k}_p\cdot\mathbf{r}_i - i\mathbf{k}_q\cdot\mathbf{r}_j}|\mathbf{k}_q|^2\delta_{pq} \\
&= \frac{w}{N}\sum_{p} |\mathbf{k}_p|^2 e^{i\mathbf{k}_p\cdot(\mathbf{r}_i-\mathbf{r}_j)}. \qquad \text{(A4)}
\end{aligned}
$$

[1] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[2] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Condens. Matter **14**, 2717 (2002).
[3] G. Galli, Curr. Opin. Solid State Mater. Sci. **1**, 864 (1996).
[4] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[5] W. Kohn, Phys. Rev. **115**, 809 (1959).
[6] J. des Cloizeaux, Phys. Rev. **135**, 685 (1964).
[7] S. Ismail-Beigi and T. A. Arias, Phys. Rev. Lett. **82**, 2127 (1999).
[8] L. He and D. Vanderbilt, Phys. Rev. Lett. **86**, 5341 (2001).
[9] J. M. Soler, E. Artacho, J. D. Gale, A. García, P. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).
[10] C. E. White, B. G. Johnson, P. M. W. Gill, and M. Head-Gordon, Chem. Phys. Lett. **253**, 268 (1996).
[11] G. E. Scuseria, J. Phys. Chem. A **103**, 4782 (1999).
[12] E. Hernández, M. J. Gillan, and C. M. Goringe, Phys. Rev. B **53**, 7147 (1996).
[13] J. E. Pask, B. M. Klein, P. A. Sterne, and C. Y. Fong, Comput. Phys. Commun. **135**, 1 (2001).
[14] J.-L. Fattebert and J. Bernholc, Phys. Rev. B **62**, 1713 (2000).
[15] J. R. Chelikowsky, N. Troullier, and Y. Saad, Phys. Rev. Lett. **72**, 1240 (1994).
[16] C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Diéguez, and M. C. Payne, Phys. Rev. B **66**, 035119 (2002).
[17] E. Hernández, M. J. Gillan, and C. M. Goringe, Phys. Rev. B **55**, 13485 (1997).
[18] J. E. Pask, B. M. Klein, C. Y. Fong, and P. A. Sterne, Phys. Rev. B **59**, 12352 (1999).
[19] P. D. Haynes and M. C. Payne, Comput. Phys. Commun. **102**, 17 (1997).
[20] A. A. Mostofi, C.-K. Skylaris, P. D. Haynes, and M. C. Payne, Comput. Phys. Commun. **147**, 788 (2002). Also known as periodic sinc or psinc functions: See Sec. VI.
[21] D. R. Bowler and M. J. Gillan, Comput. Phys. Commun. **112**, 103 (1998).
[22] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964).
[23] W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133 (1965).
[24] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
[25] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).
[26] E. Artacho and L. Miláns del Bosch, Phys. Rev. A **43**, 5770 (1991).
[27] C. A. White, P. Maslen, M. S. Lee, and M. Head-Gordon, Chem. Phys. Lett. **276**, 133 (1997).
[28] R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[29] N. Marzari, D. Vanderbilt, and M. C. Payne, Phys. Rev. Lett. **79**, 1337 (1997).
[30] X.-P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).
[31] J. M. Millam and G. E. Scuseria, J. Chem. Phys. **106**, 5569 (1997).
[32] P. D. Haynes and M. C. Payne, Phys. Rev. B **59**, 12173 (1999).
[33] C. K. Gan, P. D. Haynes, and M. C. Payne, Comput. Phys. Commun. **134**, 33 (2001).
[34] E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **52**, R5471 (1995).
[35] Y. Saad, *Iterative Methods for Sparse Linear Systems* (PWS, Boston, 1996).
[36] M. J. Gillan, J. Phys.: Condens. Matter **1**, 689 (1989).
[37] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. (The Johns Hopkins University Press, Baltimore, MD, 1996).
[38] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[39] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[40] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[41] O. F. Sankey and D. J. Niklewski, Phys. Rev. B **40**, 3979 (1989).