Corrected penalty-functional method for linear-scaling calculations within density-functional theory

P. D. Haynes and M. C. Payne
Theory of Condensed Matter, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom
(October 6, 1998)

We present a new method for the calculation of ground-state total energies within density-functional theory, based upon the single-particle density-matrix formulation, which requires a computational effort which scales only linearly with system-size. The difficult idempotency constraint is imposed approximately using a penalty-functional constructed to allow efficient minimization. The resulting error in the total energy due to the violation of idempotency is removed by an analytic correction. The results for a system comprising 216 atoms of crystalline silicon are compared with those from a standard plane-wave code. Linear scaling to 512 atoms is also demonstrated on a workstation.

Within density-functional theory (DFT), the complexity of the problem of calculating the ground-state energy of the inhomogeneous electron gas in an external potential [1] scales linearly with system-size N (i.e. the number of electrons). However, 'traditional' methods based upon the Kohn-Sham (KS) formulation of DFT [2] require a computational effort which scales asymptotically as N^3, either because of the cost of diagonalizing the Hamiltonian, or as a result of the orthogonality requirement for the extended KS single-particle wave-functions. This $\mathcal{O}(N^3)$ scaling restricts the size of systems which can currently be treated. Much interest has therefore been shown in using the single-particle density-matrix (DM) to calculate the total energy [3]. Since the DM is short-ranged in real-space [4] and free from orthogonality constraints, it provides the basis of a linear-scaling method for KS-DFT [5–7] (see [8] for a review of some of these methods). However, most linear-scaling schemes have so far been applied only in the context of tight-binding or restricted basis-set calculations. In contrast, the method presented here has been applied to fully self-consistent DFT calculations with an arbitrarily complete basis-set, a task so far attempted by only one other group [6].

In terms of the KS wave-functions $\{\psi_i\}$ and occupancies $\{f_{0,i}\}$, the ground-state DM $\rho_0(\mathbf{r},\mathbf{r}')$ is

$$\rho_0(\mathbf{r},\mathbf{r}') = \sum_i f_{0,i}\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}') \tag{1}$$

where the occupancies equal zero or unity for wave-functions with KS eigenvalues above or below the chemical potential respectively. Defining the energy functional by

$$E[\rho] = -\int \mathrm{d}\mathbf{r}' \left[\nabla^2_{\mathbf{r}}\rho(\mathbf{r},\mathbf{r}')\right]_{\mathbf{r}=\mathbf{r}'} + E_{\mathrm{Hxc}}[n] + \int \mathrm{d}\mathbf{r}\ V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r}), \tag{2}$$

where $E_{\mathrm{Hxc}}[n]$ is the sum of the Hartree and exchange-correlation energies which depend only on the electronic density $n(\mathbf{r}) = 2\rho(\mathbf{r},\mathbf{r})$ (the factor of two arising from spin degeneracy), and V_{ext} is the external potential arising from the ions. The ground-state energy is found by minimizing this functional with respect to all Hermitian DMs satisfying

$$2\int \mathrm{d}\mathbf{r}\ \rho(\mathbf{r},\mathbf{r}) = N, \tag{3}$$

and the idempotency condition

$$\rho^2(\mathbf{r},\mathbf{r}') = \int \mathrm{d}\bar{\mathbf{r}}\ \rho(\mathbf{r},\bar{\mathbf{r}})\rho(\bar{\mathbf{r}},\mathbf{r}') = \rho(\mathbf{r},\mathbf{r}'). \tag{4}$$

This can be achieved in $\mathcal{O}(N)$ operations by exploiting the short-ranged nature of the DM and truncating it beyond some spatial cut-off r_{cut} i.e. imposing $\rho(\mathbf{r},\mathbf{r}') = 0$ for $|\mathbf{r}-\mathbf{r}'| > r_{\mathrm{cut}}$.

The idempotency condition is required to ensure that the occupancies are all zero or unity. It is difficult to enforce since it is a non-linear constraint. One approach has been to use a *purifying transformation* [9] to drive the DM towards idempotency during minimization [3,6]. The approach most closely related to that introduced in this work is due to Kohn who, in a recent Letter [10], added a *penalty-functional* to the KS energy functional so that the minimum value of the total functional is achieved for the idempotent ground-state DM ρ_0. In terms of a trial Hermitian DM ρ and its eigenvalues $\{f_i\}$, Kohn's penalty-functional is

$$P[\rho] = \left\{\int \mathrm{d}\mathbf{r}\ \left[\rho^2(1-\rho)^2\right]_{\mathbf{r}'=\mathbf{r}}\right\}^{\frac{1}{2}} = \left[\sum_i f_i^2(1-f_i)^2\right]^{\frac{1}{2}}. \tag{5}$$

The square-root is required to force the total functional to take its minimum value at the ground-state, because the energy is not variational with respect to changes in the occupancies. However, the branch point introduced means that the gradient is undefined at the ground-state and the functional cannot be minimized using standard techniques such as the conjugate gradients algorithm, since information from local gradients cannot be used to build up a picture of the global behaviour of the total functional [12].

In this Letter, we introduce a generalized functional $Q[\rho]$ defined by

$$Q[\rho] = E[\rho] + \alpha P^2[\rho]. \tag{6}$$

The Taylor series for this functional in terms of the occupancies is well-defined at all points, so that it may be minimized efficiently. Although the total functional Q may possess multiple minima for certain values of α, there is only one minimum where the normalization constraint (3) is also satisfied [11], and since the search over trial DMs may be straightforwardly constrained to normalized DMs, multiple minima do not cause problems in practice.

Although the penalty-functional P^2 will prevent 'run-away' solutions (such as macroscopic occupation of the low-energy bands at the expense of negatively filling higher bands) so that the minimization process is stable (which is not guaranteed by other schemes), the DM $\bar{\rho}$ which minimizes the functional will not be exactly idempotent. In general, the occupancies $\{\bar{f}_i\}$ corresponding to occupied bands will exceed unity, and those corresponding to unoccupied bands will be negative. We estimate the errors in the occupancies $\{\delta f_i\}$ defined by $\bar{f}_i = f_{0,i} + \delta f_i$ in the following manner: since the functional Q is minimized by $\bar{\rho}$, it is a minimum with respect to all changes in the occupancies which maintain the normalization constraint (3), imposed by introducing a Lagrange multiplier λ:

$$\frac{\partial}{\partial f_i}\left\{Q[\rho] - \lambda\left[2\int d\mathbf{r}\ \rho(\mathbf{r},\mathbf{r}) - N\right]\right\}_{f_i=\bar{f}_i} = 0. \quad (7)$$

Using Janak's theorem [13] for the derivative of the energy functional E in terms of the KS eigenvalues $\{\bar{\varepsilon}_i\}$ at the minimum of the functional, we obtain:

$$\bar{\varepsilon}_i + \alpha\bar{f}_i\left(1-\bar{f}_i\right)\left(1-2\bar{f}_i\right) - \lambda = 0 \quad (8)$$

so assuming that the $\{\delta f_i\}$ are small,

$$\delta f_i = -(\bar{\varepsilon}_i - \lambda)/\alpha. \quad (9)$$

Application of the normalization constraint (3) requires $\sum_i \delta f_i = 0$ from which $\lambda = 2\sum_i \bar{\varepsilon}_i/N$ is obtained. The variance of the occupancy errors is related to the variance of the energy eigenvalues, scaled by the parameter α. As α is increased, the errors in the eigenvalues at the minimum decrease as $\delta f_i \propto \alpha^{-1}$. For small deviations from idempotency, the penalty-functional $P^2 \approx \sum_i \delta f_i^2$ so that the penalty term αP^2 also decreases as α^{-1}, and the energy approaches the true ground-state energy. This behaviour is illustrated in Fig. 1 where the occupancy errors are plotted for four bands as a function of α. This α^{-1} convergence is unsatisfactory for practical applications, since it requires large values of α to obtain accurate estimates of the ground-state energy and at large values of α the penalty term dominates the total functional and hinders efficient minimization of the energy term. To overcome this problem we present a method for obtaining accurate values for the ground-state energy from the results of calculations performed at much smaller values of α. At the minimum of the total functional, $f_i = \bar{f}_i$,

$$\left.\frac{\partial Q[\rho]}{\partial f_i}\right|_{\bar{f}_i} = 0 = \left.\frac{\partial E[\rho]}{\partial f_i}\right|_{\bar{f}_i} + 2\alpha\bar{f}_i(1-\bar{f}_i)(1-2\bar{f}_i). \quad (10)$$

This expression can be used to construct a first-order Taylor expansion for the total energy with respect to the occupancies. We can provide an estimate for the ground-state energy E_0 as

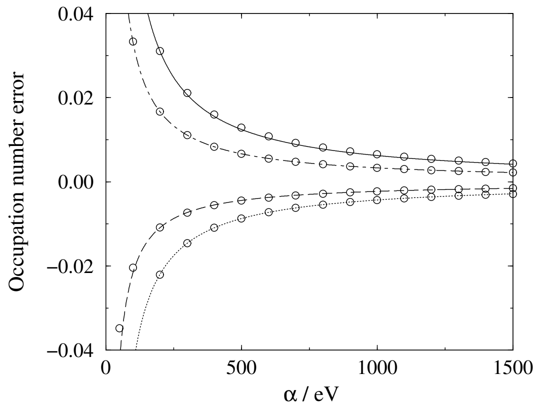

FIG. 1. Variation of occupancy errors with α. Dotted lines show the best fit to α^{-1} behaviour.

$$E_0 = E[\rho_0] \approx E[\bar{\rho}] + 2\alpha\sum_i \bar{f}_i(1-\bar{f}_i)(1-2\bar{f}_i)\delta f_i. \quad (11)$$

For occupied bands, $\delta f_i = \bar{f}_i - 1$ whereas for unoccupied bands $\delta f_i = \bar{f}_i$ so that

$$E_0 \approx E[\bar{\rho}] - 2\alpha\sum_i^{\text{all}} \bar{f}_i(1-\bar{f}_i)^2(1-2\bar{f}_i) + 2\alpha\sum_i^{\text{unocc}} \bar{f}_i(1-\bar{f}_i)(1-2\bar{f}_i). \quad (12)$$

The first term of the correction has been written as a sum over all bands so that it can be expressed in terms of a trace $\mathrm{Tr}[\bar{\rho}(1-\bar{\rho})^2(1-2\bar{\rho})]$, which can always be evaluated in $\mathcal{O}(N)$ operations. The second term only contributes when unoccupied bands are included in the calculation, which is not necessary for insulators. A single eigenvalue of the (sparse) DM can always be obtained in $\mathcal{O}(N)$ operations and so it is possible to evaluate the correction for a small number of unoccupied bands without spoiling the $\mathcal{O}(N)$ scaling. Even so, since the correction is only applied when the minimum of the total functional has been found, a single $\mathcal{O}(N^2)$ step to obtain all of the eigenvalues of the DM is still an insignificant fraction of the total computational effort. Importantly, this analytic correction also permits the calculation of forces consistent with the *corrected* energy. Fig. 2 shows the total energy, the total functional and the corrected energy as a function of α. Even at $\alpha = 50$ eV, the corrected energy agrees with the $\alpha \to \infty$ limits of both the total energy and total functional to better than 10^{-4} eV per atom.

The lowest-order term omitted from the Taylor expansion involves the second derivative

$$\frac{1}{2}\left.\frac{\partial^2 E[\rho]}{\partial f_i \partial f_j}\right|_{f_i=\bar{f}_i, f_j=\bar{f}_j} = \left.\frac{\partial \bar{\varepsilon}_i}{\partial f_j}\right|_{f_j=\bar{f}_j} \quad (13)$$

which is known as the chemical hardness matrix in the context of the construction of transferable pseudopotentials [14]. This matrix is neither positive- nor negative-definite, so that the corrected energy is neither an upper

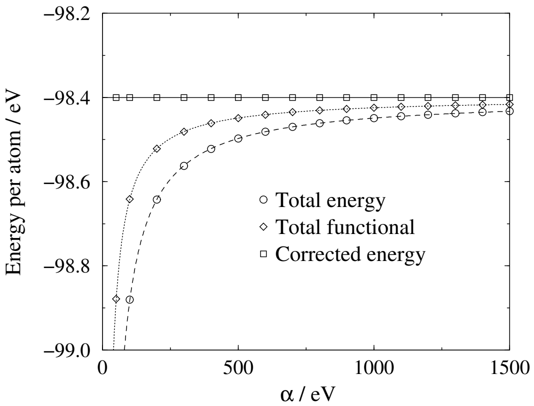

FIG. 2. Variation of the total energy, total functional and corrected energy with α. Dotted lines show the best fit to α^{-1} behaviour.

nor a lower bound to the true ground-state energy. However, this error is much smaller than those arising from the finite DM cut-off, so that in practice the corrected energy can be treated as a variational upper bound to the ground-state energy.

In practical $\mathcal{O}(N)$ calculations, we cannot work directly with the extended KS wave-functions and occupancies, and instead the trial DM is written in terms of a sparse matrix R^{ij} and a set of non-orthogonal functions $\{\phi_i\}$ localized within spherical regions of radius $r_{\rm reg}$:

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{ij} \phi_i(\mathbf{r}) R^{ij} \phi_j^*(\mathbf{r}'). \qquad (14)$$

The sparsity of R is controlled by the parameter r_R, forcing elements of R corresponding to localized functions with centres separated by more than r_R to vanish. $r_{\rm reg}$ and r_R together determine $r_{\rm cut}$. The localized functions are expanded in some localized basis, and then the total functional is minimized with respect to the expansion coefficients and the matrix elements R^{ij}.

The minimization is performed by two nested loops, the inner with respect to R and the outer with respect to the $\{\phi_i\}$. In terms of the Hamiltonian in the representation of the $\{\phi_i\}$, $H_{ij} = \langle\phi_i|\hat{H}|\phi_j\rangle$ and the overlap matrix $S_{ij} = \langle\phi_i|\phi_j\rangle$, the gradient with respect to R is

$$\frac{\partial Q[\rho]}{\partial R^{ij}} = 2H_{ji} + 2\alpha\left[SRS(1-RS)(1-2RS)\right]_{ji}. \qquad (15)$$

In contrast to other orbital-based methods, the condition number for this stage of the minimization is approximately unity (since the curvature of the functional with respect to variations of R is determined by the shape of P^2 alone) and hence the minimization is much more efficient [15]. By making the Löwdin transformation to a set of orthonormalized orbitals $\{\varphi_i = \phi_j S_{ji}^{-\frac{1}{2}}\}$ and defining the density-matrix and Hamiltonian in this representation as $\tilde{R} = S^{\frac{1}{2}} R S^{\frac{1}{2}}$ and $\tilde{H} = S^{-\frac{1}{2}} H S^{-\frac{1}{2}}$ respectively, then at the constrained minimum defined by (7);

$$2S^{\frac{1}{2}}\left[\tilde{H} + \alpha\tilde{R}(1-\tilde{R})(1-2\tilde{R}) - \lambda\right]S^{\frac{1}{2}} = 0 \qquad (16)$$

i.e. $\tilde{H} + \alpha\tilde{R}(1-\tilde{R})(1-2\tilde{R}) - \lambda = 0$ so that the density-matrix and the Hamiltonian commute and can therefore be diagonalized simultaneously. (In terms of this representation (14), $\lambda = 2\mathrm{Tr}(S^{-1}H)/N$, although its value is never required). The resulting eigenvalues obey (8).

The gradient with respect to the localized functions is

$$\frac{\delta Q[\rho]}{\delta\phi_i^*(\mathbf{r})} = 2\Big\{R^{ij}\hat{H} + \alpha\left[RSR(1-SR)(1-2SR)\right]^{ij}\Big\}\phi_j(\mathbf{r}). \qquad (17)$$

When transformed to the common diagonal frame of the density-matrix and Hamiltonian by some unitary transformation $\psi_i = \varphi_j U_{ji}$, this gradient becomes

$$\frac{\delta Q}{\delta\psi_i^*(\mathbf{r})} = 2f_i\left[\hat{H} + \alpha f_i(1-f_i)(1-2f_i)\right]\psi_i(\mathbf{r}). \qquad (18)$$

At the minimum of Q with respect to both R and $\{\phi_i\}$ we thus have from (8):

$$\frac{\delta Q}{\delta\psi_i^*(\mathbf{r})} = 0 = 2\bar{f}_i\left[\hat{H} + \alpha\bar{f}_i(1-\bar{f}_i)(1-2\bar{f}_i)\right]\psi_i(\mathbf{r}) = 2\bar{f}_i\left[\hat{H} + (\bar{\varepsilon}_i - \lambda)\right]\psi_i(\mathbf{r}). \qquad (19)$$

These are simply the Kohn-Sham equations, with eigenvalues shifted by the Lagrange multiplier λ associated with the normalization constraint. Thus, in addition to kinetic-energy preconditioning [16], occupancy preconditioning [17] can be applied to improve the convergence.

This scheme has been implemented and applied to the test case of 216 atoms of crystalline silicon. The localized functions were expanded in a spherical-wave basis-set [18] using an energy cut-off of 200 eV and angular momentum components up to $\ell = 2$. In Fig. 3 the convergence of the corrected energy with respect to $r_{\rm reg}$ and r_R is plotted. As these spatial cut-offs are increased, the variational freedom of the DM is increased, so that the energy converges from above.

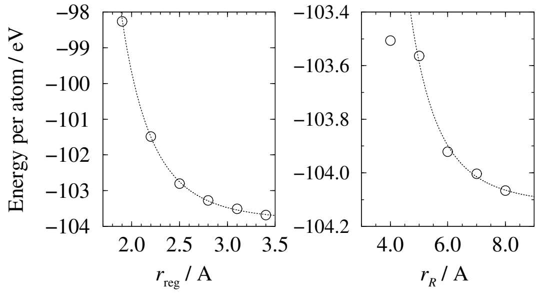

FIG. 3. Convergence of the corrected energy with respect to (left) the localization region radius $r_{\rm reg}$ and (right) the cut-off r_R applied to the matrix R.

Table I shows values for the equilibrium lattice parameter a and bulk modulus B obtained using values of

$r_{\rm reg} = 3.1$ Å, $r_R = 6.0$Å and $\alpha = 100$ eV. Experimental values and results of the CASTEP plane-wave code [19], using the same pseudopotential [20] and equivalent Brillouin zone sampling are also shown. The calculations of a agree to better than 1% and those of B (which is more sensitive to the data) to within 3%.

TABLE I. Comparison of calculated and experimental data for crystalline silicon.

Calculation	Linear-scaling	CASTEP	Experiment
a / Å	5.428	5.390	5.430
B / GPa	104.3	101.7	100.0

Finally, in Fig. 4 the computational effort for a single iteration of the outer loop for the same values of $r_{\rm reg}$, r_R and α is plotted against system-size to demonstrate the linear scaling of the method. Since fewer matrix multiplications are required, each iteration of this method is cheaper than the purifying method, also plotted (both methods require a similar number of iterations to converge the energy to a given accuracy). The number of iterations required to converge these linear-scaling calculations is currently an order of magnitude larger (about 200 for these calculations on silicon) than is the case for traditional cubic-scaling all-bands methods [21], but since the condition number of the functional depends only upon properties of the system being modelled, the number of iterations required does not increase with system-size. Since the method achieves linear scaling by partitioning real-space into overlapping regions, excellent scaling with respect to the number of processors is expected for implementations on massively-parallel computers [22].

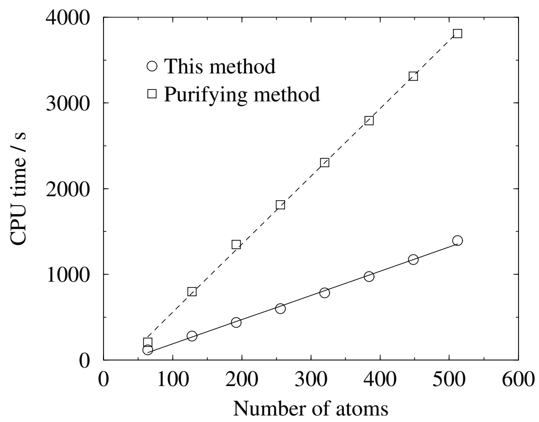

FIG. 4. Scaling of computational effort with system-size for this method compared with a method based on the purifying transformation (DEC 500au workstation).

In conclusion, we have presented a new DM-based linear-scaling method for DFT, in which the approximate imposition of idempotency, enforced by means of a penalty-functional, permits the use of efficient minimization methods, and from which accurate estimates of the energy derived from the true idempotent ground-state DM can be obtained by using a correction derived from the form of the penalty-functional.

We acknowledge useful discussions with I. D. White, and PDH acknowledges the support of an EPSRC studentship. This work is covered by British Patent Application No. 9814931.3.

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964).
[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133 (1965).
[3] X.-P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993); M. S. Daw, Phys. Rev. B **47**, 10895 (1993).
[4] S. Ismail-Beigi and T. Arias (unpublished).
[5] L.-W. Wang and M. P. Teter, Phys. Rev. B **46**, 12798 (1992); G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992); F. Mauri, G. Galli, and R. Car, Phys. Rev. B **47**, 9973 (1993); F. Mauri and G. Galli, *ibid.* **50**, 4316 (1994); P. Ordejón, D. A. Drabold, R. M. Martin, and M. P. Grumbach, *ibid.* **51**, 1456 (1995); J. Kim, F. Mauri, and G. Galli, *ibid.* **52**, 1640 (1995).
[6] E. Hernández and M. J. Gillan, Phys. Rev. B **51**, 10157 (1995); E. Hernández, M. J. Gillan, and C. M. Goringe, *ibid.* **53**, 7147 (1996).
[7] S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994); S. Goedecker and M. Teter, Phys. Rev. B **51**, 9455 (1995); U. Stephan and D. A. Drabold, *ibid.* **57**, 6391 (1998).
[8] G. Galli, Current Opinion in Solid State & Materials Science **1**, 864 (1996).
[9] R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[10] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).
[11] The penalty-functional is minimal for all idempotent DMs (which can describe both ground- and excited states of the fictitious KS system) but in a fully self-consistent calculation, the total functional is only minimal for ground-states, only one of which obeys (3).
[12] P. D. Haynes and M. C. Payne (unpublished).
[13] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).
[14] M. Teter, Phys. Rev. B **48**, 5031 (1993).
[15] S. Goedecker (unpublished).
[16] D. R. Bowler and M. J. Gillan (unpublished).
[17] N. Marzari, D. Vanderbilt, and M. C. Payne, Phys. Rev. Lett. **79**, 1337 (1997).
[18] P. D. Haynes and M. C. Payne, Comp. Phys. Comm. **102**, 17 (1997).
[19] M. C. Payne *et al.*, Rev. Mod. Phys. **64**, 1045 (1992). CASTEP is available from Molecular Simulations.
[20] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[21] M. J. Gillan, J. Phys.: Condens. Matter **1**, 689 (1989).
[22] C. M. Goringe, E. Hernández, M. J. Gillan, and I. J. Bush, Comp. Phys. Comm. **102**, 1 (1997).