Acceleration schemes for *ab initio* molecular-dynamics simulations and electronic-structure calculations

F. Tassone
Institut Romand de Recherche Numerique en Physique des Materiaux (IRRMA), IN-Ecublens, CH-1015 Lausanne, Switzerland
and Scuola Normale Superiore, piazza dei Cavalieri 7, I-56126 Pisa, Italy

F. Mauri and R. Car
Institut Romand de Recherche Numerique en Physique des Materiaux (IRRMA), IN-Ecublens, CH-1015 Lausanne, Switzerland
and Department of Condensed Matter Physics, University of Geneva, CH-1211 Geneva, Switzerland
(Received 2 August 1993; revised manuscript received 26 May 1994)

We study the convergence and the stability of fictitious dynamical methods for electrons. First, we show that a particular damped second-order dynamics has a much faster rate of convergence to the ground state than first-order steepest-descent algorithms while retaining their numerical cost per time step. Our damped dynamics has efficiency comparable to that of conjugate gradient methods in typical electronic minimization problems. Then, we analyze the factors that limit the size of the integration time step in approaches based on plane-wave expansions. The maximum allowed time step is dictated by the highest frequency components of the fictitious electronic dynamics. These can result either from the large wave vector components of the kinetic energy or from the small wave vector components of the Coulomb potential giving rise to the so called charge sloshing problem. We show how to eliminate large wave vector instabilities by adopting a preconditioning scheme in the context of Car-Parrinello *ab initio* molecular-dynamics simulations of the ionic motion. We also show how to solve the charge sloshing problem when this is present. We substantiate our theoretical analysis with numerical tests on a number of different silicon and carbon systems having both insulating and metallic character.

I. INTRODUCTION

The introduction of a fictitious dynamics for the electrons[1,2] with driving forces obtained from the total energy within density functional theory (DFT) (Ref. 3) has provided a convenient approach to minimize the total energy of condensed matter systems and to perform *ab initio* molecular-dynamics simulations of the ionic motion. These techniques have been applied successfully to a variety of insulating, semiconducting, and metallic systems involving a large number of atoms in the context of structural optimization problems at zero temperature and of dynamical simulations of the atomic motion at finite temperature.[4]

It is a subject of current interest to study the factors that limit the efficiency of fictitious dynamical methods for electrons in order to improve their numerical efficiency. This depends on the choice made for the dynamics and on the size of the time step that can be used to integrate numerically the equations of motion.

When discussing how to choose a specific dynamics, it is convenient to consider total energy minimization separately from molecular dynamics. It has been shown by Car and Parrinello that to simulate the classical adiabatic motion of the atoms it is useful to adopt a second-order Newtonian dynamics also for the electronic degrees of freedom, since this exploits optimally the concept of continuous simultaneous evolution of electronic and atomic degrees of freedom.[1,2] Newtonian dynamics conserves energy. Different approaches should be used to minimize the electronic energy, as it is required to start a molecular-dynamics simulation or to solve an optimization problem at zero temperature. The simplest approach to minimization is provided by steepest-descent dynamics, which can be viewed as a dynamics of the first order in the time derivative.[2] Steepest-descent dynamics, which requires only knowledge of the gradients of the energy functional, is not very efficient, particularly in metallic situations. Better schemes require some knowledge also of the second derivatives of the energy functional either explicitly or implicitly. Conjugate gradient methods have been developed in this context[5–8] and have been shown to be superior to steepest-descent methods, particularly when the full energy functional was used in the line minimizations and full account was taken of the orthonormality constraints on the wave functions.[8]

In this paper we show that a minor modification of a steepest-descent algorithm, namely, replacing first-order dynamics with a specific damped second-order dynamics, improves substantially the rate of convergence of the wave functions to the ground state. The resulting scheme, which we call damped molecular dynamics, has efficiency comparable to that of the best conjugate gradient algorithms when used in typical electronic mini-

mization problems, with the additional advantage of having basically the same numerical complexity as simple steepest-descent algorithms.

We then investigate what determines the maximum allowed time step for numerical integration when using steepest-descent (SD), damped (D), or Newtonian molecular dynamics (MD). In all cases the time step is limited by the need to integrate the high frequency components of the fictitious dynamics. These arise either from the large wave vector components of the electronic kinetic energy or from the small wave vector components of the Hartree energy due to the divergence of the Coulomb potential at small wave vector. In the latter case the related numerical instability is usually referred to as the "charge sloshing" problem and it is expected to become serious when the size of the system becomes very large.

The large wave vector instability can be eliminated by preconditioning the equations of motion, since at large wave vectors the wave functions are dominated by the kinetic energy and are to a large extent free-particle-like. Indeed, it was already suggested earlier by several authors that this property could be used to speed up iterative schemes for electronic minimization. In particular, Ref. 9 proposed an analytical integration scheme for the large wave vector components of the wave functions within second-order dynamics. This scheme was subsequently extended in Ref. 10 to first-order steepest-descent equations. Since this approach can be less stable than standard steepest-descent algorithms[6] we will not discuss it any further. A successful preconditioning scheme in the context of conjugate gradient minimization of the electronic total energy has been proposed in Ref. 7. This has some similarities with the quasi-Newton step proposed in Ref. 11 and with the residual minimization/direct inversion in the iterative subspace described in Ref. 12.

In this paper, we propose a preconditioning scheme which is appropriate to all the dynamical methods referred to above, namely, SD, D, and MD dynamics. It consists in properly scaling the fictitious masses associated with the large wave vector components of the electronic wave functions, in order to compress the high frequency spectrum of the electronic dynamics and to use a larger integration time step. Our preconditioning method is similar in spirit to those of Ref. 7 but it is formulated as a modification of the differential equations leading to SD, D, and MD dynamics. In particular, we apply it here to the Car-Parrinello MD equations, which provide an efficient approach for *ab initio* molecular-dynamics simulations of the ionic motion. In this context our preconditioning scheme allows us to use a time step which is two to three times larger than in previous applications of this method, resulting in a considerable saving of computational time.

We now turn our attention to the charge sloshing problem. This has been discussed previously in the context of self-consistent diagonalization of the Kohn-Sham Hamiltonian.[13] The onset of this kind of instability depends on the algorithms used and is expected to occur at significantly large sizes in the context of fictitious dynamical methods since in these approaches the wave functions change little over a single time step. Indeed, recent MD simulations for metallic liquid silicon have shown no sign of a sloshing instability up to cubic cells containing 216 silicon atoms.[14] However, one expects that for sufficiently large cells the sloshing instability should appear, although a quantitative theoretical analysis of it in the context of fictitious dynamical methods for electrons has so far been missing. Sloshing instabilities have been found numerically within some iterative schemes for electronic minimization in the case of systems having a long linear dimension.[7] In this paper we present a theoretical analysis of the charge sloshing problem in the context of SD, D, and MD equations of motion. We find that the sloshing instability is absent for insulators, but it is present for metals. This is in accord with previous results of Ref. 13. A practical scheme to control the sloshing instability is discussed in the paper.

To summarize, we improve the numerical efficiency of fictitious dynamical methods for electrons in several ways. First, we replace steepest-descent dynamics by a more efficient damped second-order dynamics to minimize the total energy. Second, by preconditioning the fictitious electronic masses we increase the integration time step for total energy minimization and for simulation of the adiabatic ionic dynamics. Third, we show that in the context of fictitious dynamical methods the so-called charge sloshing problem, which is expected to arise for large systems, is less serious than expected. We support our theoretical analysis with detailed numerical tests on several systems involving Si and C atoms.

The paper is organized as follows. In Sec. II we discuss first-order SD dynamics and second-order conservative MD for the electronic degrees of freedom. In Sec. III we introduce a damped second-order dynamics which is substantially more efficient than SD and is competitive with the best conjugate gradient schemes for electron minimization. In Sec. IV we discuss large wave vector instabilities and the charge sloshing problem. In Sec. V we discuss the preconditioning of large wave vector components. In Sec. VI we present some details of the numerical implementation. Finally, in Sec. VII we present the results of realistic numerical tests on silicon and carbon systems. Section VIII is devoted to our conclusions.

II. FICTITIOUS DYNAMICS FOR THE ELECTRONS

Dynamical methods for minimizing the electronic total energy and for simulating the adiabatic motion of the atoms are based on a fictitious dynamics of the electronic degrees of freedom. Within these approaches the forces acting on the electronic degrees of freedom are derived from the total electronic energy $E[\{\psi\}]$ in the DFT local density approximation (LDA) form

$$E[\{\psi\}] = E_{\rm kin}[\{\psi\}] + E_{\rm ext}[\rho] + E_h[\rho] + E_{\rm XC}[\rho], \quad (1)$$

where $E_{\rm kin}$, $E_{\rm ext}$, E_h, and $E_{\rm XC}$ denote kinetic, external potential, Hartree, and exchange-correlation energy,

respectively.[3] The local density approximation is adopted for the last.[3] The electronic charge density ρ is given by

$$\rho(\mathbf{r}) = 2\langle\psi_i|\mathbf{r}\rangle\langle\mathbf{r}|\psi_i\rangle, \tag{2}$$

where the occupied orbitals $|\psi_i\rangle$ are orthonormal. The factor of 2 accounts for the occupation numbers, which here and in the following are supposed to be all equal to 2. Summation over repeated indices is understood.

In order to ensure the orthonormalization of the electronic orbitals during a dynamical evolution it is convenient to add appropriate forces of constraints. These do not perform work on the electronic system, and can be conveniently calculated in terms of Lagrangian multipliers. The corresponding equations of motion for the first-order dynamics are

$$\mu|\dot{\psi}_i\rangle = -\frac{1}{2}\frac{\delta E[\{\psi\}]}{\delta\psi_i} + \Lambda_{ij}|\psi_j\rangle, \tag{3}$$

and those for the second-order dynamics are

$$\mu|\ddot{\psi}_i\rangle = -\frac{1}{2}\frac{\delta E[\{\psi\}]}{\delta\psi_i} + \Lambda_{ij}|\psi_j\rangle, \tag{4}$$

where we have assumed that the wave functions $|\psi\rangle$ are real. Here and in the following the indices i and j run over the occupied states only. The symmetric matrix Λ_{ij} of the Lagrange multipliers enforces the orthonormality condition, i.e., $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. The derivatives of $E[\{\psi\}]$ with respect to the $|\psi_i\rangle$ define the Kohn-Sham Hamiltonian:

$$\frac{1}{2}\frac{\delta E[\{\psi\}]}{\delta\psi_i} = 2\hat{H}_{\mathrm{KS}}|\psi_i\rangle. \tag{5}$$

The parameter μ is a fictitious electronic mass. It is used to tune the speed of the electronic dynamics and does not describe any other physical property. When the ions are held fixed, this mass can be included in the definition of the time step and it is irrelevant. However, when we allow the ions to move, the ratio between μ and the physical ionic masses is important since it defines the relative speed of the ionic and of the fictitious electronic motion.

Now let us suppose that the wave functions are close to the minimum of the energy $E[\{\psi^{(0)}\}]$,

$$|\psi_i\rangle = |\psi_i^{(0)}\rangle + |\delta\psi_i\rangle, \tag{6}$$

where $|\psi_i^{(0)}\rangle$ are the wave functions at the minimum and $|\delta\psi_i\rangle$ are the corresponding deviations. We notice that $|\delta\psi_i\rangle$ have to fulfill the orthonormality condition to linear order, i.e. $\langle\psi_i^{(0)}|\delta\psi_j\rangle = 0$. To linear order in $|\delta\psi_i\rangle$, the Lagrange multipliers Λ_{ij} are the same for first- and second-order dynamics, and are given by

$$\Lambda_{ij} = 2\langle\psi_i|H_{\mathrm{KS}}|\psi_j\rangle. \tag{7}$$

Thus, by retaining only the terms up to linear order in $|\delta\psi_i\rangle$ in the equations of motion (3) and (4) we obtain

$$|\delta\dot{\psi}_i\rangle = -\hat{K}_{ij}|\delta\psi_j\rangle \tag{8}$$

and

$$|\delta\ddot{\psi}_i\rangle = -\hat{K}_{ij}|\delta\psi_j\rangle, \tag{9}$$

respectively. Here $\hat{K}_{ij}$ is a linear operator, which acts on the single-particle Hilbert space and which has the same form for both first- and second-order dynamics. In the following we will use the notation $\hat{\mathbf{K}}$ to indicate a matrix of operators having for elements the $\hat{K}_{ij}$. Notice that $\hat{\mathbf{K}}$ is a positive definite linear operator since Eqs. (8) and (9) result from a quadratic expansion of $E[\{\psi\}]$ in $|\delta\psi_i\rangle$ about the minimum $E[\{\psi_0\}]$, i.e.,

$$E[\{\psi\}] - E[\{\psi_0\}] = \mu\langle\delta\psi_i|\hat{K}_{ij}|\delta\psi_j\rangle + O(\delta\psi^3). \tag{10}$$

The equations of motion (8) and (9) can be formally integrated, yielding

$$|\delta\psi_i(t)\rangle = [\exp(-\hat{\mathbf{K}}t)]_{ij}|\delta\psi_j(0)\rangle \tag{11}$$

and

$$|\delta\psi_i(t)\rangle = (\cos\sqrt{\hat{\mathbf{K}}}t)_{ij}|\delta\psi_j(0)\rangle + (\hat{\mathbf{K}}^{-1/2}\sin\sqrt{\hat{\mathbf{K}}}t)_{ij}|\delta\dot{\psi}_j(0)\rangle, \tag{12}$$

respectively. In the case of first-order evolution the wave functions decay exponentially towards the minimum $E[\psi_0]$, while in the second-order evolution they perform small oscillations around it. These motions take place with characteristic decay rates and frequencies which are equal to the eigenvalues K_α of the operator $\hat{\mathbf{K}}$ and to the square root $\sqrt{K_\alpha}$ of these eigenvalues for first- and second-order dynamics, respectively.

In numerical implementations the electronic states are expanded on a finite basis set, so that only a finite number of eigenfrequencies and eigenmodes occur. Let $K_{\min}$ and $K_{\max}$ be the minimum and maximum eigenvalues of $\hat{\mathbf{K}}$. The maximum allowed time step for numerical integration is proportional to the smallest period of the system, i.e., to $1/K_{\max}$ or to $1/\sqrt{K_{\max}}$ for first- and second-order dynamics, respectively.

In the case of first-order dynamics, the minimum eigenvalue dominates the long-time behavior of the decay to the ground state, so that a rough estimate of the convergence time is given by $1/K_{\min}$. Recalling that the size of the time step is proportional to $1/K_{\max}$, one finds that the number n_{O1} of integration steps needed to converge satisfies the condition

$$n_{O1} \propto K_{\max}/K_{\min}. \tag{13}$$

In the case of second-order dynamics we usually start a simulation from an electronic configuration close to the minimum of the electronic energy. Then if the ionic and the electronic frequencies are well decoupled,[1,2] the electrons remain adiabatically close to the instantaneous energy minimum during the ionic evolution. Let ω_{ion} be a typical ionic frequency. The adiabatic condition requires it to be much smaller than the minimum electronic frequency, i.e.,

$$\omega_{\mathrm{ion}} \ll \sqrt{K_{\min}}. \tag{14}$$

A meaningful measure of the simulation's workload is given by the number of time steps necessary to integrate a full ionic oscillation. Thus, recalling that the time step is inversely proportional to $\sqrt{K_{\max}}$, we find that the number n_{O2} of steps necessary to integrate a typical ionic oscillation satisfies the following condition:

$$n_{O2} \propto \sqrt{K_{\max}/K_{\min}}. \tag{15}$$

III. DAMPED SECOND-ORDER DYNAMICS FOR MINIMIZATION

In the preceding section we showed that the number of iterations necessary to minimize the electronic energy within steepest-descent dynamics is proportional to $K_{\max}/K_{\min}$. In this section we present an improved minimization dynamics in which the typical number of iterations is instead proportional to $\sqrt{K_{\max}/K_{\min}}$, i.e., a number significantly smaller than $K_{\max}/K_{\min}$. We attain this goal by inserting in Eq. (4) a damping term as follows:

$$\mu|\ddot{\psi}_i\rangle = -\frac{1}{2}\frac{\delta E[\{\psi\}]}{\delta\psi_i} - 2\gamma\mu|\dot{\psi}_i\rangle + \Lambda_{ij}|\psi_j\rangle. \tag{16}$$

This equation defines a damped second-order dynamics. As in the preceding section we study the resulting motion close to the energy minimum. We find that the deviations of the wave functions from the minimum are subject to damped oscillations given by

$$|\delta\psi_i(t)\rangle = \exp(i\hat{\mathbf{W}}_+ t)_{ij}|\psi_j^{(+)}\rangle + \exp(i\hat{\mathbf{W}}_- t)_{ij}|\psi_j^{(-)}\rangle. \tag{17}$$

Here $|\psi_j^{(+)}\rangle$ and $|\psi_j^{(-)}\rangle$ are determined by the initial conditions, and

$$\hat{\mathbf{W}}_\pm = i\gamma\mathbf{1} \pm \sqrt{\hat{\mathbf{K}} - \gamma^2\mathbf{1}}. \tag{18}$$

The real part of $\hat{\mathbf{W}}_\pm$ gives the frequencies of the oscillatory motion, while its imaginary part gives the decay rate to the minimum. In order to maximize the rate of convergence, we must use the maximum value of γ for which the argument of the square root remains positive. This optimal value of γ is given by

$$\gamma_{\text{opt}} = \sqrt{K_{\min}}, \tag{19}$$

since this value corresponds to critical damping of the smallest eigenvalue of $\hat{\mathbf{K}}$. In this case the imaginary part of all the eigenvalues of $\hat{\mathbf{W}}_\pm$ is equal to γ_{opt} and the time of convergence to the minimum is of the order of $1/\sqrt{K_{\min}}$. The integration time step is related to the maximum norm of the eigenvalues of $\hat{\mathbf{W}}_\pm$, which is equal to $\sqrt{K_{\max}}$. Thus, the number of integration steps necessary for minimization is given by

$$n_{DO2} \propto \sqrt{\frac{K_{\max}}{K_{\min}}} \propto \sqrt{n_{O1}}. \tag{20}$$

From this formula we see that a relevant gain of efficiency is obtained when using damped dynamics instead of steepest-descent dynamics to minimize the electronic energy. The gain is particularly important when a large number of iterations is needed to converge to the ground state, which is typically the case of metallic systems.

IV. SECOND-ORDER EXPANSION OF THE LDA ENERGY FUNCTIONAL

In this section we compute explicitly the eigenvalues of the operator $\mathbf{K}$. For this purpose we consider the expansion of the energy functional around its minimum $\psi^{(0)}$ up to second order in $\delta\psi$. This is given by

$$E[\{\psi\}] - E[\{\psi^{(0)}\}] = 2\langle\delta\psi_i|H_{\text{KS}}|\delta\psi_i\rangle - 2\langle\delta\psi_i|\delta\psi_j\rangle\langle\psi_i^{(0)}|H_{\text{KS}}|\psi_j^{(0)}\rangle + \int d\mathbf{r}\int d\mathbf{r}'\delta\rho(\mathbf{r})\left[\frac{\delta^2 E_h}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} + \frac{\delta^2 E_{\text{XC}}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\right]\delta\rho(\mathbf{r}') + O(\{\delta\psi^3\}). \tag{21}$$

The second term on the right-hand side (rhs) of this equation comes from the Lagrange multipliers [see Eq. (7)], and

$$\delta\rho(\mathbf{r}) = 4\langle\psi_i^{(0)}|\mathbf{r}\rangle\langle\mathbf{r}|\delta\psi_i\rangle \tag{22}$$

gives the variation of the electronic density to first order in $\{\delta\psi\}$. We recall that $\{\psi\}$ and $\{\delta\psi\}$ are supposed to be real.

A. Non-self-consistent case

If we neglect the last two terms in Eq. (21), i.e., the terms corresponding to variations of the Hartree and exchange-correlation potentials, we recover the expansion of the total energy appropriate to a non-self-consistent Hamiltonian. Then we can expand the $\psi^{(0)}$ and $\delta\psi$ in terms of the the real eigenvectors $|\chi_l^0\rangle$ of H_{KS}, which have eigenvalues ε_l. Since the total energy is invariant under unitary transformations in the subspace of occupied states, we can suppose without loss of generality

$$|\psi_i^{(0)}\rangle = |\chi_i^0\rangle,$$

$$|\delta\psi_i\rangle = \sum_k c_k^i|\chi_k^0\rangle, \tag{23}$$

where c_k^i are real coefficients. Here and in the following the indices i and k refer to occupied and unoccupied

states, respectively. Hence, as shown in Ref. 15, we obtain for $\delta E_{\rm NSC}$, i.e., the second-order variation of the energy in which the self-consistency of the potential is not taken into account,

$$\delta E_{\rm NSC} = 2\sum_{ik}(c_k^i)^2(\varepsilon_k - \varepsilon_i), \tag{24}$$

where ε_i and ε_k are respectively the occupied and the unoccupied eigenvalues of $H_{\rm KS}$. By comparing Eq. (24) with Eq. (10) we see that the eigenvalues of $\hat{\mathbf{K}}$ are given by

$$K_{(i,k)} = 2\frac{\varepsilon_k - \varepsilon_i}{\mu}, \tag{25}$$

and the lowest eigenvalue of $\hat{\mathbf{K}}$ is given by $K_{\rm min} = 2E_{\rm gap}/\mu$, in terms of the energy gap $E_{\rm gap}$ separating the lowest unoccupied from the highest occupied electronic level.

In the case of an insulator, the energy gap has a finite positive value which, above a certain size, is independent of the simulation cell. In the case of a metal instead, the energy gap is still finite and positive for a finite sized system but it is no longer independent of the simulation cell. In fact the energy gap and $K_{\rm min}$ tend to zero for a cell size going to infinity. However, many properties of interest do not require an infinite energy resolution for the states around the Fermi energy. Typically a small but finite energy resolution $E_{\rm err}$ is sufficient. $E_{\rm err}$ does not depend on the size of the system, and $K_{\rm err} = 2E_{\rm err}/\mu$ replaces $K_{\rm min}$ in Eqs. (13), (20), and (19) to estimate the convergence rates n_{O1}, n_{DO2} and the optimal damping parameter $\gamma_{\rm opt}$. Since $E_{\rm err}$ is much smaller than a typical energy gap of an insulator, the number of iterations needed for ground state convergence is much larger in metals than in insulators. For the same reason, a perfectly adiabatic separation between ionic and electronic motions is not possible for metals. However, as shown in Ref. 16 a satisfactory solution to this problem, in the context of Car-Parrinello simulations, can be obtained by using two Nosé thermostats to control separately the respective temperatures of the ions and of the electrons.

When expanding the wave functions in terms of plane waves, we can define an effective cutoff energy $E_{\rm cut}$ given (in a.u.) by $q_{\rm max}^2/2$, where $q_{\rm max}$ is the largest wave vector in the basis set. The band of empty states is usually much larger than the band of occupied states. Thus when in Eq. (25) the index k refers to the highest unoccupied states, the eigenvalues $K_{(i,k)}$ of $\hat{\mathbf{K}}$ have a negligible dependence on the occupied state index i. Furthermore, since the highest unoccupied states are free-particle-like, the energy difference $\varepsilon_k - \varepsilon_i$ is dominated by the kinetic energy of the state k, i.e.,

$$\varepsilon_k - \varepsilon_i \sim q^2/2, \tag{26}$$

where q is the wave vector associated with the state k. The maximum eigenvalue of $\hat{\mathbf{K}}$ is therefore approximately given by $K_{\rm max} \simeq 2(q_{\rm max}^2/2)/\mu = 2E_{\rm cut}/\mu$. It is this eigenvalue that limits the maximum allowed time step for numerical integration in a non-self-consistent case: the numerical integration becomes unstable and the time step has to go to zero when $E_{\rm cut}$ goes to infinity.

B. Self-consistent case: charge sloshing

We now consider the terms of Eq. (21) that we neglected in the preceding subsection in order to see if they affect the maximum eigenvalue of $\hat{\mathbf{K}}$. In this case $\hat{\mathbf{K}}$ is not diagonal in the representation of the c_k^i and, for an arbitrary system, it is not possible to diagonalize it analytically. Thus we need some simplifying assumptions. Let us consider a crystal of given periodicity and use a supercell containing an arbitrary number of replicas of the crystal unit cell. In this case the $\{\psi_0\}$ are linear combinations of Bloch functions with the crystal periodicity whereas the fluctuations $\{\delta\psi\}$ may have all the wavelengths compatible with the supercell. In other words, we are restricting the periodicity of the unperturbed state but not the periodicity of the fluctuations. Based on the above simplifying assumption we find that charge sloshing affects metallic and nonmetallic systems differently. A numerical example presented in Sec. VII suggests that this result should hold also for nonperiodic systems.

In order to find out whether the maximum eigenvalue of $\hat{\mathbf{K}}$ diverges when the supercell size tends to infinity, we restrict our analysis to the Hartree term since the LDA exchange-correlation energy is well behaved and typically has a negligible effect compared to the kinetic energy on the maximum eigenvalue of $\hat{\mathbf{K}}$. The second-order variation of the Hartree energy is given by

$$\begin{aligned}\delta E_H &= \int d\mathbf{r}\int d\mathbf{r}'\delta\rho(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\delta\rho(\mathbf{r}') \\ &= \sum_{\mathbf{G}}\sum_{\mathbf{p}}\frac{4\pi}{\Omega|\mathbf{p}+\mathbf{G}|^2}|\delta\rho(\mathbf{p}+\mathbf{G})|^2.\end{aligned} \tag{27}$$

Here $\mathbf{p}$ is a vector belonging to the first Brillouin zone of the crystal, $\mathbf{G}$ is a vector of the reciprocal lattice of the crystal, and the sums extend over all the nonzero wave vectors $\mathbf{p}+\mathbf{G} = \mathbf{q}-\mathbf{q}'$ where $\mathbf{q}$ and $\mathbf{q}'$ are two generic plane waves of the basis set used to represent the electron wave functions in the supercell of volume Ω. $\delta\rho(\mathbf{p}+\mathbf{G})$ is the Fourier transform (FT) of $\delta\rho(\mathbf{r})$.

When a linear dimension L of the supercell becomes very large, $(\mathbf{p}+\mathbf{G})_{\rm min} = (\mathbf{p})_{\rm min}$, i.e., the smallest nonzero $\mathbf{q}-\mathbf{q}'$ vector, tends to zero like $1/L$. If, correspondingly, the maximum eigenvalue of $\hat{\mathbf{K}}$ diverges, we have the so called charge sloshing scenario. To study the effect of $\mathbf{p}_{\rm min}$ on the maximum eigenvalue of $\hat{\mathbf{K}}$ we consider only the terms with $\mathbf{G} = \mathbf{0}$ in Eq. (27). Then since $\delta\rho(\mathbf{r})$ is real, $|\delta\rho(\mathbf{p})| = |\delta\rho(-\mathbf{p})|$ and $\delta E_H(\mathbf{G}=\mathbf{0})$ can be written as

$$\delta E_H(\mathbf{G}=\mathbf{0}) = \sum_{\mathbf{p},p_z>0}\frac{8\pi}{\Omega p^2}|\delta\rho(\mathbf{p})|^2, \tag{28}$$

where

$$\begin{aligned}\delta\rho(\mathbf{p}) = \sum_{ik} 4[&\langle\chi_i^0|\cos(\mathbf{p}\cdot\mathbf{r})|\chi_k^0\rangle c_k^i \\ &+ i\langle\chi_i^0|\sin(\mathbf{p}\cdot\mathbf{r})|\chi_k^0\rangle c_k^i].\end{aligned} \tag{29}$$

Notice that Eq. (28) is a quadratic form in terms of the c_k^i. The real coefficients c_k^i can be considered as the components of a real vector $|\mathbf{c}\rangle$. Similarly we can introduce two real vectors $|\mathbf{A}(\mathbf{p})\rangle$ and $|\mathbf{B}(\mathbf{p})\rangle$ whose components, labeled by the composite index (ik), are given by $\langle\chi_i^0|\cos(\mathbf{p}\cdot\mathbf{r})|\chi_k^0\rangle$ and by $\langle\chi_i^0|\sin(\mathbf{p}\cdot\mathbf{r})|\chi_k^0\rangle$, respectively. In this notation $\delta\rho(\mathbf{p}) = 4[\langle\mathbf{c}|\mathbf{A}(\mathbf{p})\rangle + i\langle\mathbf{c}|\mathbf{B}(\mathbf{p})\rangle]$ and Eq. (28) becomes

$$\delta E_H(\mathbf{G}=\mathbf{0}) = \sum_{\mathbf{p},p_x>0} \frac{8\pi}{\Omega p^2}[\langle\mathbf{c}|\mathbf{A}(\mathbf{p})\rangle\langle\mathbf{A}(\mathbf{p})|\mathbf{c}\rangle + \langle\mathbf{c}|\mathbf{B}(\mathbf{p})\rangle\langle\mathbf{B}(\mathbf{p})|\mathbf{c}\rangle]. \tag{30}$$

Using the fact that the χ_i^0 are eigenstates of a periodic crystal, it is easy to show[17] that the vectors $|\mathbf{A}(\mathbf{p})\rangle$ and $|\mathbf{B}(\mathbf{p})\rangle$ constitute an orthogonal set:

$$\begin{aligned}\langle\mathbf{A}(\mathbf{p})|\mathbf{A}(\mathbf{p}')\rangle &= \delta_{\mathbf{p},\mathbf{p}'}\tfrac{\Omega}{2}S(\mathbf{p})\\ \langle\mathbf{B}(\mathbf{p})|\mathbf{B}(\mathbf{p}')\rangle &= \delta_{\mathbf{p},\mathbf{p}'}\tfrac{\Omega}{2}S(\mathbf{p})\\ \langle\mathbf{B}(\mathbf{p})|\mathbf{A}(\mathbf{p}')\rangle &= 0,\end{aligned} \tag{31}$$

where $p_x, p_x' > 0$ and the static structure factor $S(\mathbf{p})$ is defined by

$$S(\mathbf{p}) = \frac{1}{\Omega}\sum_{ik}\langle\chi_i^0|e^{-i\mathbf{p}\cdot\mathbf{r}}|\chi_k^0\rangle\langle\chi_k^0|e^{+i\mathbf{p}\cdot\mathbf{r}}|\chi_i^0\rangle. \tag{32}$$

Hence $|\mathbf{A}(\mathbf{p})\rangle$ and $|\mathbf{B}(\mathbf{p})\rangle$ are the vectors that diagonalize the quadratic form in Eq. (30). The corresponding eigenvalues of $\hat{\mathbf{K}}$ are given by

$$\begin{aligned}K_{\mathbf{A}(\mathbf{p})} = K_{\mathbf{B}(\mathbf{p})} &= \frac{128\pi}{\mu\Omega p^2}\langle\mathbf{A}(\mathbf{p})|\mathbf{A}(\mathbf{p})\rangle\\ &= \frac{128\pi}{\mu\Omega p^2}\langle\mathbf{B}(\mathbf{p})|\mathbf{B}(\mathbf{p})\rangle = \frac{64\pi}{\mu}\frac{S(\mathbf{p})}{p^2}.\end{aligned} \tag{33}$$

Therefore, when a linear dimension L of the supercell tends to infinity and, correspondingly, $\mathbf{p}_{\min}$ goes to zero, $K_{\mathbf{A}(\mathbf{p}_{\min})}$ and $K_{\mathbf{B}(\mathbf{p}_{\min})}$ do not diverge if $S(\mathbf{p})$ is of order $O(p^2)$.

We now consider a jellium model as a representative metallic system. In this case the χ_i^0 are plane waves and one finds[18] $S(\mathbf{p}) = p[1-(p/p_F)^2/12]p_F^2/8\pi^2$, where p_F is the Fermi momentum, and $p < 2p_F$. As a consequence, for L going to infinity, $K_{\mathbf{A}(\mathbf{p}_{\min})}$ and $K_{\mathbf{B}(\mathbf{p}_{\min})}$ diverge as L and the time step for numerical integration has to be reduced accordingly: this is a charge sloshing situation.

When the system is a periodic insulator one finds instead that $S(\mathbf{p})$ goes to zero as p^2 (see Appendix). As a consequence, for L going to infinity, $K_{\mathbf{A}(\mathbf{p}_{\min})}$ and $K_{\mathbf{B}(\mathbf{p}_{\min})}$ tend to a constant and the time step for numerical integration is independent of L: charge sloshing is absent here.

We stress that the above conclusions apply only if we consider small fluctuations around the ground state: this is the typical case of *ab initio* molecular-dynamics simulations of the ionic motion. However, in the initial steps of an electronic minimization procedure, the wave functions may be far from the ground state. In this case it is possible to observe sloshing effects also in periodic systems having an insulating ground state.

Since charge sloshing is a consequence of the singularity of the Coulomb potential at small $\mathbf{p}$, a simple way of eliminating charge sloshing instabilities consists in replacing the Coulomb potential $4\pi/p^2$ with a Yukawa potential $4\pi/(p^2+\alpha^2)$, where $2\pi/\alpha$ is a typical decay length of the order of the system size $L_{\min}$ that corresponds to the onset of the sloshing instability. In the case of an insulator we can use this technique to stabilize the numerical integration during the initial steps of an electronic minimization run. Then when we are sufficiently close to the ground state we can set $\alpha = 0$ and converge to the exact ground state. We will show with a numerical example in a subsequent section that this technique allows us to converge to the exact ground state of a disordered insulating system with a number of steps independent of the system size. In the case of a metal it is not possible to set α equal to zero, not even in the proximity of the ground state. However, we notice that $L_{\min}$ is usually much larger than the typical screening length of a metal. The results of a numerical simulation for a large but finite metallic system should not change appreciably if the Coulomb potential is replaced by a Yukawa potential that is equal to the Coulomb potential for distances smaller than $L_{\min}$.

V. PRECONDITIONING THE EQUATIONS OF MOTION

The numerical efficiency of all the fictitious dynamical methods previously introduced can be improved by preconditioning the dynamics in order to reduce the ratio $K_{\max}/K_{\min}$. This can be achieved by replacing the constant fictitious mass parameter μ in Eqs. (3), (4), and (16). with an arbitrary positive definite operator $\hat{\mu}$. The resulting increased arbitrariness in the choice of $\hat{\mu}$ can be exploited to compress the highest frequency components of the spectrum of the fictitious electron dynamics. Recalling that these are due basically to the high energy unoccupied states which are free-particle-like [see Eq. (26)], we choose an operator $\hat{\mu}$ which is diagonal in q space with eigenvalues $\mu(q)$ given by

$$\begin{aligned}\mu(q) &= \mu_0 \quad \text{if } \tfrac{1}{2}q^2 < E_p,\\ \mu(q) &= \mu_0\tfrac{q^2}{2E_p} \quad \text{if } \tfrac{1}{2}q^2 > E_p.\end{aligned} \tag{34}$$

Below a certain cutoff energy E_p, it is worth considering a constant mass μ_0, because the low energy eigenstates have a relevant potential energy contribution and are not free-particle like. The preconditioning cutoff E_p therefore represents the threshold above which the states are dominated by the kinetic energy.

It is easy to show that the solutions of the preconditioned equations of motions for small displacements are still given by Eqs. (11), (12), and (17) if the operator $\hat{\mathbf{K}}$ is replaced by the operator $\hat{\tilde{\mathbf{K}}}$ characteristic of the preconditioned dynamics. All the relations (13), (15), and (20) found for first- and second-order dynamics with and without damping hold therefore also in the case of the preconditioned dynamics but, in the latter case, $K_{\max}$ and $K_{\min}$ have to be replaced by the maximum and min-

imum eigenvalues of $\hat{\bar{\mathbf{K}}}$, i.e., by $\bar{K}_{\max}$ and $\bar{K}_{\min}$.

The preconditioning cutoff E_p that minimizes the ratio $\bar{K}_{\max}/\bar{K}_{\min}$ is called the optimal preconditioning cutoff. It depends strongly on the atomic species, i.e., on the pseudopotentials, and on the plane-wave cutoff that are used in the calculation. It depends only negligibly on the physical environment. Thus, for a given atomic species, it is possible to find the optimal preconditioning cutoff by performing calculations on a simple reference system. We present a typical example in Sec. VII.

VI. NUMERICAL IMPLEMENTATION

In our numerical implementation we adopt the standard procedures described in Refs. 2 and 4 to integrate the equations of motion for first- and second-order dynamics. In the case of damped second-order dynamics we follow the procedure introduced in Ref. 16 to integrate Car-Parrinello dynamics in the presence of a friction term. We obtain for first- and second-order and damped dynamics, respectively,

$$|\psi_i(t+\Delta)\rangle = |\psi_i(t)\rangle - 2\hat{\mu}^{-1}\hat{H}_{\rm KS}|\psi_i\rangle\Delta + X_{ij}\hat{\mu}^{-1}|\psi_j(t)\rangle, \tag{35}$$

$$|\psi_i(t+\Delta)\rangle = -|\psi_i(t-\Delta)\rangle + 2|\psi_i(t)\rangle - 2\hat{\mu}^{-1}\hat{H}_{\rm KS}|\psi_i\rangle\Delta^2 + X_{ij}\hat{\mu}^{-1}|\psi_j(t)\rangle, \tag{36}$$

$$|\psi_i(t+\Delta)\rangle = |\psi_i(t-\Delta)\rangle + \left(|\psi_i(t)\rangle - |\psi_i(t-\Delta)\rangle - \hat{\mu}^{-1}\hat{H}_{\rm KS}|\psi_i\rangle\frac{\Delta^2}{2}\right)\frac{2}{1+\gamma\Delta} + X_{ij}\hat{\mu}^{-1}|\psi_j(t)\rangle, \tag{37}$$

where Δ is the integration time step and X_{ij} is a symmetric matrix equal to $\Lambda_{ij}\Delta$ for first-order dynamics, equal to $\Lambda_{ij}\Delta^2$ for conservative second-order dynamics, and equal to $\Lambda_{ij}\Delta^2/(1+\gamma\Delta)$ for damped second-order dynamics. The matrix $\mathbf{X}$ is found by imposing the orthonormality of the wave functions at time $t+\Delta$:

$$\langle\psi_i(t+\Delta)|\psi_j(t+\Delta)\rangle = \delta_{ij}. \tag{38}$$

We notice that the inversion of the mass operator $\hat{\mu}$ is straightforward in q space where it is diagonal. For the calculation of X_{ij} we define the wave functions $|\bar{\psi}_i(t+\Delta)\rangle$ as the rhs of Eqs. (35)–(37) without the orthonormalization terms $X_{ij}\hat{\mu}^{-1}|\psi_j(t)\rangle$. Then Eq. (38) becomes

$$\mathbf{X}\mathbf{M}\mathbf{X}^{\dagger} + \mathbf{B}\mathbf{X}^{\dagger} + \mathbf{X}\mathbf{B}^{\dagger} = \mathbf{1} - \mathbf{A}, \tag{39}$$

where the matrices $\mathbf{M}, \mathbf{B}, \mathbf{A}$ are given, respectively, by

$$M_{ij} = \langle\psi_i(t)|\hat{\mu}^{-2}|\psi_j(t)\rangle, \tag{40}$$

$$B_{ij} = \langle\bar{\psi}_i(t+\Delta)|\hat{\mu}^{-1}|\psi_j(t)\rangle, \tag{41}$$

$$A_{ij} = \langle\bar{\psi}_i(t+\Delta)|\bar{\psi}_j(t+\Delta)\rangle. \tag{42}$$

The scalar products are easily evaluated in q space where the mass operator $\hat{\mu}$ is diagonal.

Equation (39) is formally identical to the matrix equation that expresses the orthonormality condition for Car-Parrinello dynamics when using Vanderbilt's ultrasoft pseudopotentials.[19] It can be solved as described in Ref. 19. The matrix $\mathbf{B}$ can be conveniently split into a symmetric part $\mathbf{B}_s$ and an antisymmetric part $\mathbf{B}_a$. The antisymmetric part $\mathbf{B}_a$ is first order in Δ, while $\mathbf{X}$ and $\mathbf{1}-\mathbf{A}$ are first (second) order in Δ for first- (second-) order dynamics. Using these properties, we can solve Eq. (39) iteratively in terms of increasing powers of Δ (Δ^2):

$$\mathbf{B}_s X^{(n+1)} + \mathbf{X}^{(n+1)}\mathbf{B}_s = \mathbf{1} - \mathbf{A} - \mathbf{X}^{(n)}\mathbf{M}\mathbf{X}^{(n)} - \mathbf{B}_a X^{(n)} + \mathbf{X}^{(n)}\mathbf{B}_a. \tag{43}$$

Here $\mathbf{X}^{(0)}$ is the solution of the equation

$$\mathbf{B}_s X^{(0)} + \mathbf{X}^{(0)}\mathbf{B}_s = \mathbf{1} - \mathbf{A} \tag{44}$$

and the left-hand sides of Eqs. (43) and (44) are inverted after transforming to the basis where $\mathbf{B}_s$ is diagonal.

An alternative approach based on an unconstrained total energy functional which avoids explicit orthonormalization has been recently proposed in Refs. 20 and 21. The electronic mass preconditioning scheme discussed in the present paper can be easily applied to the unconstrained energy functional method without any overload.

In some applications of fictitious dynamical methods for electrons the orthonormalization of the electronic wave functions can be achieved via a Gram-Schmidt procedure. We stress that this approach is not justified in connection with the mass preconditioning scheme described above. Indeed, if the Gram-Schmidt orthogonalization procedure is used within preconditioned steepest-descent dynamics, one is not guaranteed that the energy will decrease at any integration step for a sufficiently small time step. The origin of this instability is related to the nonholonomic character of the constraints imposed via a Gram-Schmidt procedure. We found that this instability is rather severe in practical numerical applications, where it spoils all the efficiency gains of the mass preconditioning scheme.

VII. NUMERICAL RESULTS

We tested the different dynamical schemes described above on various physical systems within a DFT LDA formulation. In particular we considered Si and C

systems. We used pseudopotentials of the Bachelet-Hamann-Schlüter type,[22] with s and p nonlocality in the Kleinmann-Bylander form.[23] The cutoff for the plane-wave expansion of the electronic orbitals was 12 Ry for silicon, and 35 Ry for carbon. We carried out all the calculations at the Γ point of the Brillouin zone only. Moreover, in order to compare the dynamical schemes with conjugate gradient minimization, we used a tight-binding energy functional for carbon.[24]

A. Preconditioning

We start by presenting the results obtained with preconditioning. In order to determine the optimal preconditioning cutoff E_p we had to minimize the ratio $\bar{K}_{\max}/\bar{K}_{\min}$. We measured $\bar{K}_{\max}$ and $\bar{K}_{\min}$ within first-order and second-order dynamics, by giving a small displacement to the system from its energy minimum. In the case of first-order dynamics, the numerical integration of Eq. (35) becomes unstable and results in an exponential increase of the energy, when $\Delta > 2/\bar{K}_{\max}$. Therefore the maximum allowed integration time step provides an accurate way of estimating $\bar{K}_{\max}$. $\bar{K}_{\min}$, i.e., the lowest eigenvalue of $\hat{\bar{\mathbf{K}}}$, gives instead the slowest rate of decay of the energy. This rate is conveniently sampled at large times, i.e., when only the slowest exponential is left in the decay.

We report in Fig. 1 the ratio $\bar{K}_{\max}/\bar{K}_{\min}$ as a function of E_p for a Si_3 molecule. We notice that for the highest E_p values the ratio decreases linearly with decreasing E_p. The behavior of the ratio $\bar{K}_{\max}/\bar{K}_{\min}$ in this range is explained by the following considerations. First, the minimum frequency is unchanged, since it is related to the lowest excited state which has small components at high q. Second, all the excited modes at energies higher than E_p are compressed to the same maximum frequency $2E_p/\mu_0$, as long as they are kinetically dominated. Instead, in the range of low preconditioning cutoffs E_p, the minimum frequency $\bar{K}_{\min}$ decreases and the highest excited modes become less efficiently compressed. Thus a minimum value of the ratio $\bar{K}_{\max}/\bar{K}_{\min}$ is found, as we can see in Fig. 1. This minimum occurs at E_p=1 Ry. The corresponding reduction of the ratio $\bar{K}_{\max}/\bar{K}_{\min}$ is by a factor of 5 compared to the nonpreconditioned case. We obtained very similar results for a sample of crystalline silicon in the diamond structure, where the optimal E_p was also close to 1 Ry.

In the case of second-order dynamics $\sqrt{\bar{K}_{\max}}$ and $\sqrt{\bar{K}_{\min}}$ can be found as the maximum and minimum frequencies of the power spectrum of the fictitious electronic dynamics. This is easily evaluated by computing the velocity autocorrelation function corresponding to the wave function dynamics. The power spectrum of the velocity autocorrelation function corresponding to the electronic fictitious dynamics is given in Fig. 2 for the case of the Si_3 molecule. In particular, we show results obtained with optimal preconditioning (E_p = 1 Ry) and unpreconditioned dynamics. In calculating the spectra we chose the value of μ_0 in such a way that the lowest frequency $\sqrt{\bar{K}_{\min}}$ of the preconditioned dynamics coincided with $\sqrt{K_{\min}}$, i.e., the lowest frequency of the dynamics without preconditioning. This is achieved by setting μ_0=260 a.u. when the mass associated with the dynamics without preconditioning is μ=300 a.u. The significant compression of the high frequency modes resulting from preconditioning is clearly evident in Fig. 2.

For a sample of crystalline carbon in the diamond structure we found an optimal value of E_p equal to 2.7 Ry. This reduced by a factor of 9 the ratio $\bar{K}_{\max}/\bar{K}_{\min}$ compared to the unpreconditioned case. In this case in order to make the lowest frequencies $\sqrt{\bar{K}_{\min}}$ and $\sqrt{K_{\min}}$

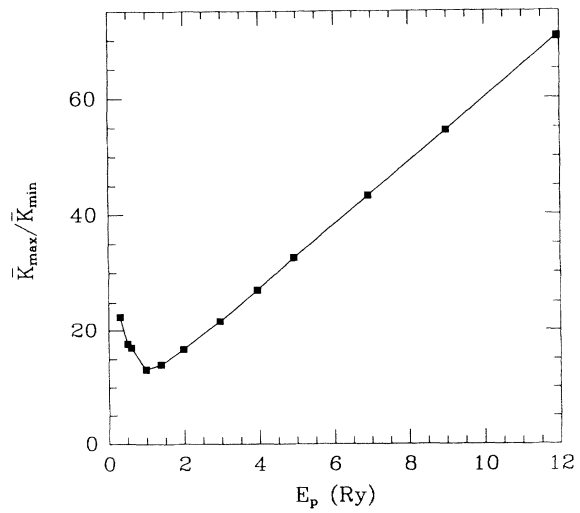

FIG. 1. $\bar{K}_{\max}/\bar{K}_{\min}$ as a function of the preconditioning cutoff E_p for the Si_3 molecule. A periodically repeated cubic cell of 20 a.u. is used in all the calculations for the Si_3 molecule.

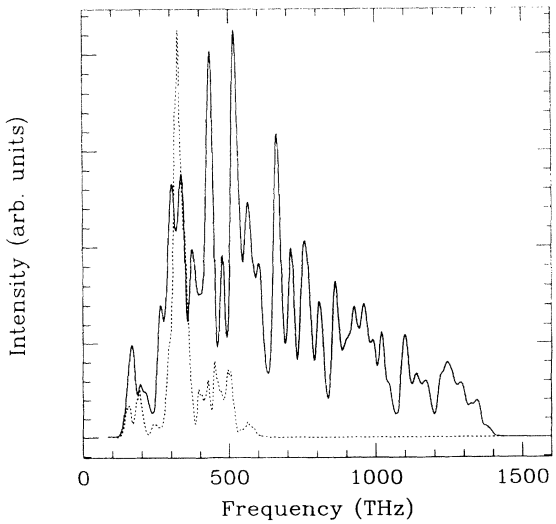

FIG. 2. Spectra of the electronic frequencies for the Si_3 molecule. The solid line refers to second-order dynamics without preconditioning. The dashed line refers to second-order dynamics with preconditioning (E_p=1 Ry).

coincide, the mass μ of the unpreconditioned dynamics had to be rescaled by a factor of 0.93 in order to obtain the mass μ_0 of the preconditioned dynamics.

We notice that in a different context the authors of Ref. 7 proposed using a preconditioning cutoff E_p equal to the expectation value of the kinetic energy divided by the number of electrons. In the cases discussed above this corresponds to a value of E_p = 0.8 Ry and E_p =2 Ry for Si and C, respectively. These values are close to the optimal values of E_p.

B. Ionic molecular dynamics

In order to test the effect of preconditioning on *ab initio* molecular-dynamics simulations of the ionic motion, we considered the coupled set of equations given by Eq. (4) for the electronic degrees of freedom and by

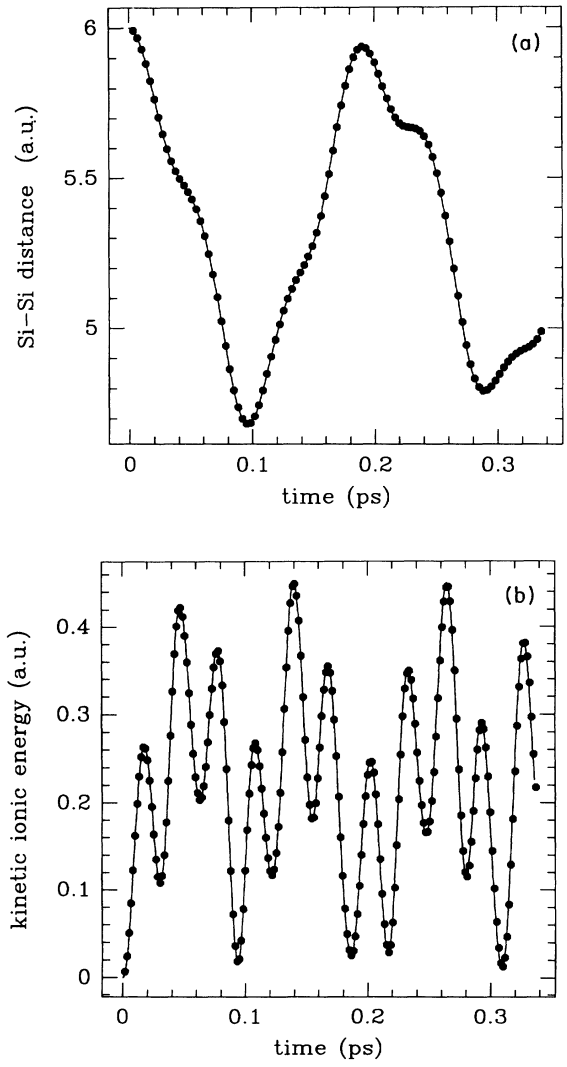

FIG. 3. Ionic dynamics of the Si_3 molecule without (solid line) and with (dots) preconditioning. In (a) we report the oscillations of the long side of the Si_3 molecule, and in (b) the oscillations of the ionic kinetic energy as a function of time.

$$M_i\ddot{\mathbf{R}}_i = -\frac{\delta E[\{\psi\},\mathbf{R}]}{\delta \mathbf{R}_i} \tag{45}$$

for the ionic coordinates $\mathbf{R}_i$. Here M_i are the physical ionic masses and the mass μ in Eq. (4) has to be replaced by the mass operator $\hat{\mu}$ in the preconditioned case. Equations (4) and (45) reproduce the adiabatic dynamics of the ions when the appropriate decoupling condition, Eq. (14) discussed in Sec. II, is satisfied.[1,2] We considered the vibrational motion of a Si_3 molecule during a time span of about 0.3 ps. In the unpreconditioned case we used a time step Δ=7 a.u. to integrate the equations of motion. This is close to the maximum allowed time step for a fictitious electronic mass μ=300 a.u. Preconditioning allowed to increase this time step to Δ =15 a.u. for a mass μ_0=260 a.u. and a preconditioning cut-off E_p=1 Ry. In spite of the significantly larger time step the preconditioned dynamics proceeded adiabatically in the same way as the one without preconditioning. In particular, any systematic energy transfer from the ionic system to the electronic one was absent. We plot in Fig. 3 the temporal evolution of the ionic kinetic energy and of the longest side of the Si_3 molecule as a function of time in both the preconditioned and the unpreconditioned cases. Differences between the two dynamics are not noticeable.

C. Damped dynamics in insulators

In order to assess the efficiency of the various minimization dynamics discussed in this work, we considered a 64 atom amorphous Si sample generated by *ab initio* molecular dynamics.[25] We notice that this system has a finite gap, and therefore a nonzero $K_{\min}$. In all our total energy minimizations we used the same set of starting trial wave functions. These were obtained by minimizing the total energy with a very small energy cutoff $E_{\rm cut}$ of 2 Ry. We then minimized the total energy with a cutoff of 12 Ry using four types of dynamics, namely, steepest-descent and second-order damped dynamics both without and with optimal preconditioning. We report the results in Fig. 4. In particular, we found that, when using the optimal value $\gamma_{\rm opt}$ of Eq. (19), the rate of convergence of second-order damped dynamics is faster than that of steepest-descent dynamics by the amount expected from the theoretical analysis in Sec. III. Preconditioning accelerated further the rate of convergence, so that finally the rate of convergence of preconditioned second-order damped dynamics was 14 times faster than the one of unpreconditioned steepest descent.

We determined the value $\gamma_{\rm opt}$ by a rough estimate of $\bar{K}_{\min}$ based on steepest-descent dynamics. In particular, a three point fit of the exponential decay of the total energy gives

$$\gamma_{\rm opt}\Delta \sim \sqrt{\frac{1}{2}\ln\left(\frac{E_1-E_2}{E_2-E_3}\right)}, \tag{46}$$

where E_1, E_2, E_3 are the energies at three successive steps of steepest descent. We waited until only the slowest exponential was left in the decay. If faster exponentials are still present, Eq. (46) overestimates $\gamma_{\rm opt}\Delta$. In

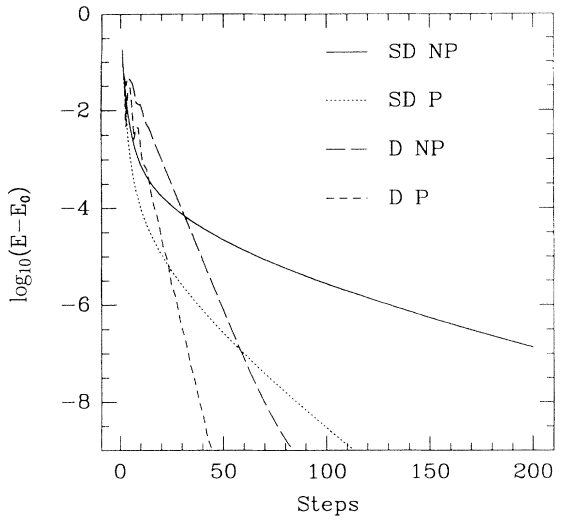

FIG. 4. Total energy minimization for a 64 atom amorphous silicon sample, using nonpreconditioned steepest-descent (SD NP), preconditioned steepest-descent (SD P), nonpreconditioned damped dynamics (D NP), and preconditioned damped dynamics (D P). We plot the logarithm of the difference between the energy per atom (E) and the ground state energy per atom (E_0) in Hartree units vs the number of integration steps.

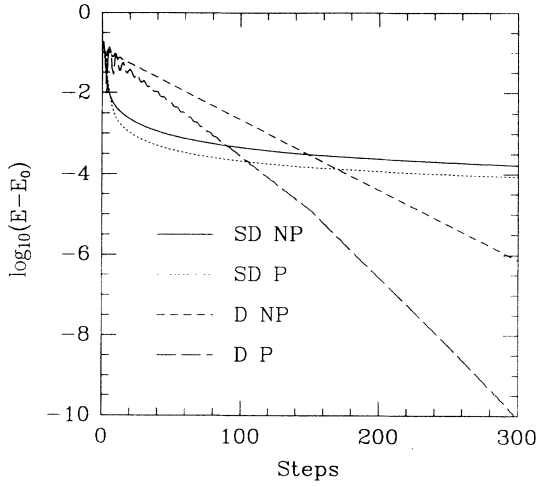

FIG. 5. Total energy minimization for a 64 atom liquid silicon sample, using nonpreconditioned steepest descent (SD NP), preconditioned steepest-descent (SD P), nonpreconditioned damped dynamics (D NP), and preconditioned damped dynamics (D P). We plot the logarithm of the difference between the energy per atom (E) and the ground state energy per atom (E_0) in Hartree units vs the number of integration steps.

a practical calculation, we therefore suggest starting the minimization with a few steps of steepest descent and using Eq. (46) to obtain an upper bound for the optimal $\gamma\Delta$. Then, we suggest proceeding with the damped second-order minimization, readjusting $\gamma\Delta$ in order to achieve the optimal limit of critical damping. As we can see from Fig. 4, it is indeed convenient to use steepest descent in the first steps of minimization when the highest frequency components dominate the deviation of the energy from the minimum. Subsequently, when only the slowest frequencies are left, damped dynamics becomes much more convenient, especially in those cases of extremely slow convergence rate.

D. Damped dynamics in metallic systems

In order to test the efficiency of our damped dynamical scheme for minimization in the case of metallic systems we applied it to liquid silicon which is a metal. We use a 64 atom sample generated by *ab initio* molecular dynamics.[14] As explained in Sec. IV, the damping constant γ can be fixed on the basis of the required energy resolution $E_{\rm err}$, for which we chose here a value of about 20 meV. We minimized the total energy with a cutoff of 12 Ry using four types of dynamics, similarly to what we did in the insulating case. Again the starting trial wave functions were obtained by a minimization using a small cutoff of 2 Ry. We report the results in Fig. 5. Notice that, in the present metallic case, steepest-descent dynamics is particularly inefficient, while damped dynamics is very effective, since it improves by many orders of magnitude the convergence rate of steepest descent. A further gain results from preconditioning.

E. Comparison with conjugate gradient minimization

In this subsection we compare our damped dynamical method with a conjugate gradient minimization scheme. The standard conjugate gradient procedure, described, e.g., in Ref. 5, cannot be directly applied to a constrained functional, unless some additional simplifying assumptions are invoked which can reduce the minimization efficiency.[6,7] To fully exploit the power of the conjugate gradient procedure the authors of Ref. 8 proposed use of an unconstrained energy functional. We adopt the same procedure of Ref. 8 but we use a different form for the unconstrained energy functional. We use the form suggested in Refs. 20 and 21 in the context of electronic-structure calculations with linear size scaling but without imposing any localization constraints on the electronic orbitals.[26] For reasons of numerical simplicity we adopt here a total energy functional based on a non-self-consistent tight-binding Hamiltonian. This choice simplifies considerably the line minimization in the conjugate gradient scheme, which can be performed exactly.[21]

We used a tight-binding Hamiltonian for carbon,[24] and we considered an ionic liquid configuration of 64 atoms at a temperature of 5000 K. For this configuration the system is metallic. Our results are reported in Fig. 6 where we plot the logarithmic error in the total energy per atom versus the number of iterations for various minimization schemes, namely, damped dynamics, conjugate gradient, and steepest-descent minimization. In the case of conjugate gradient minimization the number of iterations was multiplied by a factor of 2 in order to take into account the increase in computational cost arising from line minimization. From Fig. 6 it is evident that the nu-

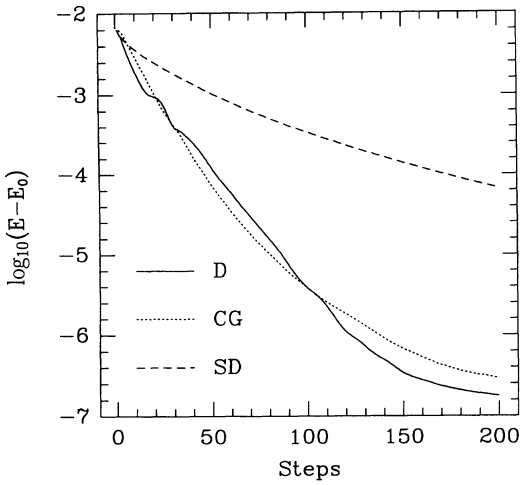

FIG. 6. Total energy minimization for a 64 atom liquid carbon sample, using steepest-descent (SD), conjugate gradient (CG), and damped dynamics (D). The total energy corresponds to a parametrized tight-binding Hamiltonian (see text). We plot the logarithm of the difference between the energy per atom (E) and the ground state energy per atom (E_0) in Hartree units vs the number of integration steps. The number of integration steps of the conjugate gradient calculation has been multiplied by 2 to take into account the increase in computational cost compared to the other methods.

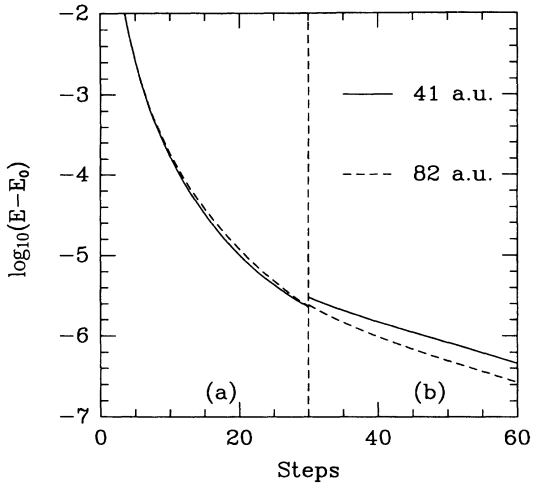

FIG. 7. Total energy minimization for two randomized crystalline silicon samples using preconditioned steepest-descent. The continuous line refers to a 32 atom cell with a long side of 41 a.u. The dashed line refers to a 64 atom cell with a long side of 82 a.u. We plot the logarithm of the difference between the energy per atom (E) and the ground state energy per atom (E_0) in Hartree units vs the number of integration steps. In (a) we use a Yukawa potential to compute E and E_0, while in (b) the Yukawa potential is replaced by a Coulomb potential (see text).

merical efficiency of both conjugate gradient and damped molecular dynamics is considerably superior to that of steepest-descent minimization. In the present example the numerical efficiency of conjugate gradient and that of damped molecular-dynamics minimization are practically the same.

We expect that the results that we have found here should remain valid also in the case of a self-consistent LDA Hamiltonian.

F. Charge sloshing on very long cells

In order to study charge sloshing effects we used a tetragonal supercell having a long side. In particular, we considered crystalline silicon in the diamond structure and we constructed two supercells by repeating four or eight elementary cubic cells along the crystallographic (100) direction. The resulting supercells contain 32 and 64 atoms, respectively. Then we broke the translational invariance of the diamond lattice by giving the silicon atoms a random displacement of about 5% of the bond length. This did not modify the insulating character of the system.

In the present example we have considered only preconditioned steepest-descent minimization. As in the previous subsections we prepared the initial trial state by minimizing the total energy with a small cutoff of 2 Ry starting from a set of random wave functions. Severe charge sloshing instabilities immediately showed up during this initial minimization in which the starting random wave functions were very far from the converged insulating ground state. In particular, already for the 32 atom cell the time step for numerical integration had to be reduced by an order of magnitude compared to the time step that we could use in an equivalent situation with a smaller cell. Such instability was completely eliminated by replacing the Coulomb by a Yukawa potential as described in Sec. IV. We adopted here a parameter $2\pi/\alpha = 20.5$ a.u. for the Yukawa potential. Once we obtained the initial trial state, we performed a total energy minimization on the 32 and on the 64 atom cell with a cutoff of 12 Ry. The results are shown in Fig. 7 which reports the deviation from the converged ground state energy as a function of the number of numerical time steps. During the initial 30 steps we used the Yukawa potential. This allowed us to use for both 32 and 64 atom cells an integration time step equal to the one usually adopted for the same system when using cells sufficiently small that no charge sloshing effects are present. Then we switched from Yukawa to Coulomb potential. After 30 minimization steps with the Yukawa potential the system was insulating and already very close to its exact ground state. In this case, as shown analytically in Sec. IV, charge sloshing instabilities are not expected to occur in a periodic system. Indeed, not even in our disordered sample did they occur. Therefore we could use in the final 30 minimization steps the same time step used with the Yukawa potential. The overall convergence rate, as can be clearly seen in the figure, is independent of the cell size.

VIII. CONCLUSIONS

We have presented a detailed analysis of the stability and of the convergence rate of fictitious dynamical

methods for electrons. We have succeeded in improving considerably the efficiency of currently used algorithms for total energy minimization and for *ab initio* molecular dynamics.

In the case of *ab initio* molecular-dynamics simulations of the ionic motion we have introduced a preconditioning scheme which gives rise to an overall saving of CPU time of the order of 2–3 in typical applications. In the case of total energy minimization we have introduced an optimal damped preconditioned dynamics which has a convergence rate substantially faster than steepest-descent algorithms and comparable to that of the best conjugate gradient schemes for electronic-structure calculations. This is especially important in metallic situations.

Although in this paper we confine our analysis to electronic minimization, we stress that the damped dynamics algorithm can also be applied to ionic minimization. In this case the optimal ionic damping parameter is related to the phonon frequencies of the system under study.

In addition, we have presented a detailed analysis of the charge sloshing instability and we have indicated a practical way to control it. We have shown with a numerical example that, in the case of insulators, our approach allows us to converge to the ground state with a number of iterations that is independent of the system size.

Note added in proof: In Sec. VII E we have shown numerically on a specific example that conjugate gradient minimization and our damped dynamics scheme have similar convergence rates. Recently, we became aware that this can also be shown analytically. In fact, the convergence rate of the conjugate gradient approach is proportional to $\sqrt{K_{\min}/K_{\max}}$, as is that of our damped dynamical method. An analytical estimate of the convergence rate of conjugate gradient minimization can be found in Jennings and McKeown, *Matrix Computation*, 2nd ed. (Wiley, Chichester, 1992) and in Stoer and Bulirsch, *Introduction to Numerical Analysis* (Springer-Verlag, New York, 1992). This has been extended to the minimization of the DFT-LDA functional in Annett (unpublished).

ACKNOWLEDGMENTS

We gratefully acknowledge many useful discussions with Giulia Galli and Osamu Sugino. We acknowledge support by the Swiss National Science Foundation under Grant No. 20-39528.93. We thank J. Annett for sending us a copy of his work prior to its publication.

APPENDIX

In this Appendix we show that for a periodic insulator the function $S(\mathbf{p})$ given in Eq. (32) goes to zero like p^2 for p going to zero.

$$S(\mathbf{p}) = \frac{1}{\Omega}\sum_{ik}\langle\chi_i^0|e^{-i\mathbf{p}\cdot\mathbf{r}}|\chi_k^0\rangle\langle\chi_k^0|e^{+i\mathbf{p}\cdot\mathbf{r}}|\chi_i^0\rangle$$
$$= \frac{1}{\Omega}\sum_{i}\left(1-\sum_{j}\langle\chi_i^0|e^{-i\mathbf{p}\cdot\mathbf{r}}|\chi_j^0\rangle\langle\chi_j^0|e^{+i\mathbf{p}\cdot\mathbf{r}}|\chi_i^0\rangle\right), \tag{A1}$$

where the indices i and j refer to occupied states and the index k refers to empty states. In Eq. (A1) we used the completeness relation $\sum_k |\chi_k^0\rangle\langle\chi_k^0| = \mathbf{1}-\sum_j |\chi_j^0\rangle\langle\chi_j^0|$. Let us suppose for simplicity that we have a single occupied band. Since the expression in Eq. (A1) is invariant under unitary transformations on the occupied subspace, we can write it in terms of Wannier functions, i.e.,

$$S(\mathbf{p}) = \frac{1}{\Omega_{\min}}\left(1-\sum_{\mathbf{R}}|\langle W_{\mathbf{0}}|e^{-i\mathbf{p}\cdot\mathbf{r}}|W_{\mathbf{R}}\rangle|^2\right), \tag{A2}$$

where $W_{\mathbf{R}}$ is the Wannier function centered on site $\mathbf{R}$, and $\Omega_{\min}$ is the volume of the elementary cell. The Wannier functions are exponentially localized in the case of an insulator: this allows us to expand the exponentials in Eq. (A2) in a Taylor series for $\mathbf{p}$ going to zero, In particular, if we consider the term with $\mathbf{R} = \mathbf{0}$ in Eq. (A2), and expand the exponentials in $\mathbf{p}\cdot\mathbf{r}$ around $\mathbf{p}\cdot\langle\mathbf{r}\rangle = \mathbf{p}\cdot\langle W_{\mathbf{0}}|\mathbf{r}|W_{\mathbf{0}}\rangle$, we get

$$1-|\langle W_{\mathbf{0}}|e^{-i\mathbf{p}\cdot\mathbf{r}}|W_{\mathbf{0}}\rangle|^2 = 1-|\langle W_{\mathbf{0}}|\mathbf{1}-i\mathbf{p}\cdot(\mathbf{r}-\langle\mathbf{r}\rangle)$$
$$-[\mathbf{p}\cdot(\mathbf{r}-\langle\mathbf{r}\rangle)]^2/2|W_{\mathbf{0}}\rangle|^2$$
$$+o(p^2)$$
$$=+\langle W_{\mathbf{0}}|[\mathbf{p}\cdot(\mathbf{r}-\langle\mathbf{r}\rangle)]^2|W_{\mathbf{0}}\rangle+o(p^2). \tag{A3}$$

This term tends to zero as p^2. In a similar way one can show that the terms with $\mathbf{R}$ different from zero in Eq. (A2) also go to zero as p^2.

[1] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[2] R. Car and P. Parrinello, in *Simple Molecular Systems and Very High Densities*, edited by A. Polian, P. Loubeyre, and N. Boccara (Plenum Press, New York, 1988), p. 455.
[3] *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N.H. March (Plenum Press, New York, 1983).
[4] For a review, see, e.g., G. Galli and A. Pasquarello, *New Perspectives on Computer Simulations in Chemical Physics* (Kluwer, Dordrecht, in press).
[5] W.H. Press, A.A. Teukoisky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes* (Cambridge University Press, New York, 1992).
[6] I. Stich, R. Car, M. Parrinello, and S. Baroni, Phys. Rev. B, **39**, 4997 (1989).
[7] M.P. Teter, M.C. Payne and D.C. Allan, Phys. Rev. B **40**,

12 225 (1989).
[8] T.A. Arias, T.A. Payne, and J.D. Joannopoulos, Phys. Rev. Lett. **69**, 1077 (1992).
[9] M.C. Payne, D. Joannopoulos, D.C. Allan, M.P. Teter, and D.H. Vanderbilt, Phys. Rev. Lett. **56**, 2656 (1986).
[10] A. Williams and J. Soler, Bull. Am. Phys. Soc. **32**, 562 (1987).
[11] R.K. Nesbet, J. Chem. Phys. **43**, 311 (1965).
[12] P. Bendt and A. Zunger, Bull. Am. Soc. **27**, 248 (1982); D.M. Wood and A. Zunger, J. Phys. A **18**, 1343 (1985).
[13] R.M. Martin and K. Kunc, in *Ab Initio Calculation of Phonon Spectra*, edited by J.T. Devreese (Plenum Press, New York, 1983), p. 49.
[14] O. Sugino and R. Car, Bull. Am. Phys. Soc. **39**, 278 (1994); O. Sugino and R. Car (unpublished).
[15] G. Pastore, E. Smargiassi, and F. Buda, Phys. Rev. A **44**, 6334 (1991).
[16] P.E. Blöchl and M. Parrinello, Phys. Rev. B **45**, 9413 (1993).
[17] To this purpose it is useful to apply the relation $\sum_{ik}\langle\chi_i^0|e^{-i\mathbf{p}\cdot\mathbf{r}}|\chi_k^0\rangle\langle\chi_k^0|e^{+i\mathbf{p}'\cdot\mathbf{r}}|\chi_i^0\rangle = \Omega S(\mathbf{p})\delta_{\mathbf{p},\mathbf{p}'}$, where $S(\mathbf{p})$ is given in Eq. (32).
[18] See, e.g., V.D. Gorobchenko, V.N. Kohon, and E.G. Maksimov, in *Modern Problems in Condensed Matter Sciences*, edited by V.M. Agranovich and A.A. Maradudin (North-Holland, Amsterdam, 1989), Vol. 24, p. 113.
[19] A. Pasquarello, K. Laasonen, R. Car, C. Lee, and D. Vanderbilt, Phys. Rev. Lett. **69**, 1982 (1992); K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, Phys. Rev. B **47**, 10 142 (1993).
[20] F. Mauri, G. Galli, and R. Car, Phys. Rev. B **47**, 9973 (1993).
[21] F. Mauri and G. Galli, Phys. Rev. B **50**, 4316 (1994).
[22] G.B. Bachelet, D.R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
[23] L. Kleinmann and D.M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[24] C. Xu, C. Wang, C. Chan, and K. Ho, J. Phys. Condens. Matter **4**, 6047 (1992).
[25] I. Stich, R. Car, and M. Parrinello, Phys. Rev. B **44**, 4262 (1991).
[26] Since we do not impose localization constraints on the electronic orbitals, the cost of the calculation still grows as the cube of the system size, as in standard electronic-structure calculations. As shown in Ref. 20 the minimum of the unconstrained energy functional coincides with that of the constrained functional. Moreover, since we adopt a truncation parameter $\mathcal{N} = 1$ (see Ref. 20) for the inverse of the overlap matrix, the minimization rate is the same for the unconstrained and the constrained functional (Ref. 21).