I. INTRODUCTION

The evaluation of exact electronic exchange has long been noted as a major numerical bottleneck in extending first principles electronic structure calculations to systems of realistic size. This is due to the fact that in schemes based on nonorthogonal Gaussian bases the calculation of the two-electron exchange matrix elements usually scales as the fourth power of the number of basis functions. Efficient algorithms for the construction of the Fock matrix elements have been proposed recently.\(^1\) These methods, which use Gaussian basis sets, scale linearly for insulators and quadratically for metals. However, the calculation of Pulay forces makes schemes based on Gaussian bases inefficient for \textit{ab initio} molecular dynamics (AIMD) calculations\(^2\) for even moderate size clusters. Density functional (DF) schemes which instead seek to use an (yet unknown, but in principle universal) exchange-correlation (x-c) functional have thus become increasingly important both in the solid state physics and quantum chemistry communities.

Recently there has been renewed interest in the inclusion of exact exchange within DF schemes. Such “hybrid” x-c functionals have been shown to significantly improve the accuracy of molecular energies (in particular, binding energies\(^3\) and barrier heights\(^4\)) as well as spin couplings in magnetic systems.\(^5\) This is to be expected since, as noted by Becke,\(^3\) a consideration of the adiabatic connection formula\(^6\) shows that in the noninteracting limit the x-c functional limits to the exchange energy of a single determinant of Kohn–Sham orbitals. Indeed, the use of screened Hartree–Fock (HF) exchange also substantially improves calculated values of band gaps in semiconductors.\(^7,8\) In order that reliable electronic structure and AIMD calculations be done on systems of chemical interest, the efficient evaluation of exact exchange is thus of much importance.

In the present paper, a plane-wave (PW) based algorithm is developed which includes nonlocal exchange interactions within hybrid density functionals. The theory and its subsequent tests show that the method scales almost \textit{linearly} with the number of basis functions. While it is found that with a straightforward PW implementation the exact exchange calculations are slower than those for a generalized gradient approximation (GGA), they hold the promise of being made more efficient in the future. Even without further increases in efficiency, the improved accuracy is well worth the added overhead for many problems of interest.

II. THEORY: EXACT EXCHANGE IN PW BASIS

In this section the expressions for exchange within the Hartree–Fock approximation are obtained. The one-particle wave functions (henceforth labeled orbitals) are expanded in a PW basis. After simplification, it is shown that the exchange energy and its gradient with respect to the occupied orbitals (which is all that is needed in the AIMD approach) scale as \(N_{PW} \log N_{PW}\), where \(N_{PW}\) is the number of PWs.

To simplify the discussion, let us consider a closed shell system. The exchange contribution to the total energy is given by\(^9\)

\[
E_x^{\text{HF}} = -\frac{1}{2} \sum_{m,n} f_m f_n \int \frac{d \mathbf{r} d \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \rho_1(\mathbf{r}, \mathbf{r}') \phi_m^*(\mathbf{r}) \phi_m(\mathbf{r}) \phi_n^*(\mathbf{r}') \phi_n(\mathbf{r}')
\]

(1)

where \(\rho_1(\mathbf{r}, \mathbf{r}')\) is the one-electron density matrix, \(\phi_{m,n}(\mathbf{r})\) are occupied molecular orbitals, and \(f_m, f_n\) are the occupation numbers. By expanding the orbitals \(\phi_{m,n}(\mathbf{r})\) in a PW basis, viz.,

\[
\phi_{m}(\mathbf{r}) = \sum_{g} c_{m}(\mathbf{g}) e^{i \mathbf{g} \cdot \mathbf{r}},
\]

(2)

one obtains

\[
E_x^{\text{HF}} = -\frac{1}{4} \sum_{m,n} f_m f_n \sum_{\mathbf{g}, \mathbf{g}'} c_{m}(\mathbf{g}) c_{n}(\mathbf{g}') h_{m,n}(\mathbf{g}, \mathbf{g}'),
\]

(3)

where the exchange matrix elements \(h_{m,n}(\mathbf{g}, \mathbf{g}')\) are given by
\[ h_{mn}(\mathbf{g}, \mathbf{g}^\prime) = \langle e^{i \mathbf{g} \cdot \mathbf{r}} | h_{mn} | e^{i \mathbf{g}^\prime \cdot \mathbf{r}} \rangle \]
\[ = \int d\mathbf{r} e^{-i \mathbf{g} \cdot \mathbf{r}} \int d\mathbf{r}' \frac{\phi_m(\mathbf{r}') \phi_n^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} e^{i \mathbf{g}^\prime \cdot \mathbf{r}}. \tag{4} \]

Note that these matrix elements have translational symmetry; thus \( h_{mn}(\mathbf{g}, \mathbf{g}^\prime) = h_{mn}(\mathbf{g} - \mathbf{g}^\prime) \). Now, using the property of the Fourier transform of the Coulomb potential, one gets
\[ \int d\mathbf{r}' \left[ \phi_m(\mathbf{r}') \phi_n^*(\mathbf{r}') \right] \frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{4\pi}{\Omega} \mathrm{FT}[\rho_m(\mathbf{g}^\prime)/g^{n^2}] \]
\[ = \mathrm{FT} \left[ \frac{4\pi}{\Omega g^{n^2}} \sum_{\mathbf{g}} c_m(\mathbf{g}^\prime + \mathbf{g})_n c_n^*(\mathbf{g}) \right], \tag{5} \]

where \( \Omega \) is the volume of the cell and \( \mathrm{FT} \) denotes a Fourier transform. One then obtains
\[ E_x^{\text{HF}} = -\frac{\pi}{\Omega} \sum_{m,n} \sum_{\mathbf{g}} c_m(\mathbf{g}) c_n(\mathbf{g}^\prime) \frac{C_{mn}(\mathbf{g} - \mathbf{g}^\prime)}{|\mathbf{g} - \mathbf{g}^\prime|^2}, \tag{6} \]

where
\[ C_{m,n}(\mathbf{g}) = \sum_{\mathbf{g}} c_n^*(\mathbf{g}) c_m(\mathbf{g} + \mathbf{g}). \tag{7} \]

i.e., \( C_{mn} \) is a correlation function. Substituting \( \mathbf{g} = \mathbf{g} - \mathbf{g}^\prime \) in Eq. (6), one has
\[ E_x^{\text{HF}} = -\frac{\pi}{\Omega} \sum_{m,n} f_n f_m \sum_{\mathbf{g}} \frac{1}{|\mathbf{g}|^2} |C_{mn}(\mathbf{g})|^2. \tag{8} \]

Given the orbitals, the exchange energy can be efficiently evaluated using a property of the correlation function, viz.,
\[ C_{m,n}(\mathbf{g}) = \mathrm{FT}[\phi_m(\mathbf{r}) \phi_n(\mathbf{r})], \tag{9} \]

making the calculation of the exchange energy an \( O(N_{\text{orb}}^2 N_{\text{PW}} \log N_{\text{PW}}) \) procedure (\( N_{\text{orb}} \) = number of occupied orbitals).

A self-consistent evaluation of exchange requires a variational optimization of the orbitals. Within an AIMP Car–Parrinello-type approach, \(^2\) this requires the evaluation of the gradient of the energy with respect to the orbital coefficients. From Eq. (1),
\[ \frac{\delta E_x^{\text{HF}}}{\delta \phi_n^*(\mathbf{r})} = -\frac{f_n}{2} \int d\mathbf{r}' \frac{\phi_n^*(\mathbf{r}',\mathbf{r}') \phi_n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{10} \]

Using the separable form for the density matrix, one has
\[ \frac{\delta E_x^{\text{HF}}}{\delta \phi_n^*(\mathbf{r})} = -\frac{f_n}{2} \sum_{m} f_m \phi_m(\mathbf{r}) \int d\mathbf{r}' \frac{\phi_n(\mathbf{r}') \phi_n^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{11} \]

As in Eq. (5),
\[ \int d\mathbf{r}' \frac{\phi_n(\mathbf{r}') \phi_n^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \frac{4\pi}{\Omega} \mathrm{FT} \left[ \frac{C_{m,n}(\mathbf{g})}{g^2} \right] = f_{m,n}(\mathbf{r}), \tag{12} \]

with \( C_{mn} \) as defined in Eq. (7). The required gradient is given by a Fourier transform of the gradient in real space, i.e.,
\[ \frac{\delta E_x^{\text{HF}}}{\delta \phi_n^*(\mathbf{r})} = \int d\mathbf{r} \frac{\delta E_x^{\text{HF}}}{\delta \phi_n^*(\mathbf{r})} e^{-i \mathbf{g} \cdot \mathbf{r}} \]
\[ = \mathrm{FT} \left[ \sum_{m} \phi_m(\mathbf{r}) f_{m,n}(\mathbf{r}) \right]. \tag{13} \]

Using the property of correlation functions, one obtains
\[ \frac{\delta E_x^{\text{HF}}}{\delta \phi_n^*(\mathbf{r})} = -\frac{2\pi}{\Omega} \sum_{m} f_m \sum_{\mathbf{g}} \frac{1}{|\mathbf{g} + \mathbf{g}|^2} C_{m,n}(\mathbf{g} + \mathbf{g}), \tag{14} \]

where \( G_{m,(mn)} \) are also correlation functions. The evaluation of the gradient of the exchange energy thus also scales as \( N_{\text{orb}}^2 N_{\text{PW}} \log N_{\text{PW}} \). It is important to note that such an algorithm could be devised for other orthogonal basis sets (e.g., compactly supported wavelets) which satisfy a similar convolution property. \(^9\)

### III. IMPLEMENTATION: PSEUDOPOTENTIAL FORMALISM

While the computational implementation of the above approach within an all-electron (AE) formalism is straightforward, the use of pseudopotentials (PPs), as used in most PW-based approaches to reduce the size of the basis, introduces at least two complications. The first, due to the neglect of core electrons, is less extreme since many properties of interest are insensitive to the core electrons. The second complication, though, arises due to the fact that, within the core region, the valence pseudo-orbitals could differ substantially from the AE orbitals. Thus in order to construct the density matrix the AE orbitals need to be constructed first. One such scheme has been recently devised. \(^11\) Working entirely from general principles, it has been formulated into an AE “projector augmented wave” (PAW) method, \(^12\) closely related to the Vanderbilt approach. \(^13\) In order to allow exact exchange calculations within a Vanderbilt PP scheme, we use an approach similar to that outlined in Ref. 12, and summarized below.

To put our method in perspective, consider first the DF pseudo-Hamiltonian \( H^{DF} \) within a Vanderbilt PP scheme, viz.,
\[ H^{DF} = -\frac{1}{2} \nabla^2 + V_{\text{eff}} + V_{NL}, \tag{15} \]

where the effective local potential \( V_{\text{eff}} \) includes the Hartree, local exchange-correlation and ionic terms, and \( V_{NL} \) is the nonlocal pseudopotential, given by \(^14\)
\[ V_{NL} = \sum_{ij} \left( D_{ij}^{(0)} + \int d\mathbf{r} V_{\text{eff}}(\mathbf{r}) Q_{ij}(\mathbf{r}) \right) |\beta_i(\beta_j)|. \tag{16} \]

Here \( i \) labels the atomic site, eigenvalue, and angular momenta quantum numbers for the functions \( |\beta_i\rangle \) in terms of
which the density $\rho(\mathbf{r})$ is augmented, and $Q_{ij}(\mathbf{r})$ are the corresponding atom-centered charge augmentation functions, viz.,

$$\rho(\mathbf{r}) = \rho_1(\mathbf{r}, \mathbf{r}) = \sum_m f_m \left| \phi_m^{PP}(\mathbf{r}) \right|^2 + \sum_{ij} Q_{ij}(\mathbf{r}) \langle \phi_m^{PP}(\mathbf{r}) | \beta_i \rangle \langle \beta_j | \phi_m^{PP}(\mathbf{r}) \rangle,$$

(17)

where $\phi_m^{PP}$ are molecular PP orbitals. The matrix elements $D_{ij}^{(0)}$ are given by the difference between the atomic contributions to the AE and PP energies.\(^{13}\) Now, $H^{GDF}$ can be generalized to include exact exchange by adding the HF exchange potential, $V_x^{HF} = \rho_1(\mathbf{r}, \mathbf{r}')/|\mathbf{r} - \mathbf{r}'|$, and generating the pseudopotential parameters $Q_{ij}(\mathbf{r})$ and $D_{ij}^{(0)}$ from corresponding atomic calculations. The resulting generalized pseudo-Hamiltonian $H^{GDF}$ is given by

$$H^{GDF} = \frac{1}{2} \nabla^2 + V_{eff} + V_{NL} + V_x^{HF}.$$  

(18)

To generate the density matrix-dependent exchange potential, one first constructs the AE orbitals,\(^{12}\) as summarized below. Given the PP for an atom, its valence PP states ($\chi_i^{PP}$) are related to the AE ones ($\chi_i^{AE}$) by a linear transformation. Now, within the “core region” of the PP, the PP orbitals of a molecule or solid can be expanded as a linear combination of the $\chi_i^{PP}$s, i.e.,

$$\phi_m^{PP}(\mathbf{r}) = \sum_i c_{mi} \chi_i^{PP}(\mathbf{r}).$$  

(19)

The AE orbitals can similarly be written as

$$\phi_m^{AE}(\mathbf{r}) = \sum_i c_{mi} \chi_i^{AE}(\mathbf{r}).$$  

(20)

Without loss of generality, one defines functions $p_i$ such that the coefficients $c_{mi}$ are given as $\langle \phi_m^{PP} | \phi_i^{AE} \rangle$. Since the PP orbitals differ from the AE ones only within the core region, the “projector” functions $p_i$ are localized within this region, and from Eq. (19) must satisfy an orthogonality constraint, $\langle p_i | \chi_j^{PP} \rangle = \delta_{ij}$.

A question then arises as to how one chooses the functions $p_i$. Blochl\(^{12}\) has described one such procedure. However, it is easy to see that within Vanderbilt’s\(^{13}\) generalized PP approach such $p_i$’s satisfying all the above constraints arise naturally—they are the functions $\beta_i$ in Eq. (16). From Eqs. (19) and (20) one then has

$$\phi_n^{AE}(\mathbf{r}) = \phi_n^{PP}(\mathbf{r}) + \sum_i \langle \beta_i | \phi_n^{PP} \rangle \left( \chi_i^{AE}(\mathbf{r}) - \chi_i^{PP}(\mathbf{r}) \right).$$

(21)

This equation is valid both within and outside the core regions, since the PP orbitals limit to AE orbitals outside the core regions. Using the generalized orthogonality relation for the pseudo-orbitals, it can be easily seen that these AE functions are orthogonal in the usual sense, i.e., $\langle \phi_m^{AE} | \phi_n^{AE} \rangle = \delta_{mn}$, and that the charge density $\Sigma_m \phi_m^{AE}(\mathbf{r})^2$ is identical to the generalized density defined in the Vanderbilt scheme.\(^{12,13}\) The density matrix, as given by

$$\rho_1(\mathbf{r}, \mathbf{r}') = \sum_n f_n \phi_n^{AE}(\mathbf{r}) \phi_n^{AE}(\mathbf{r}')$$

(22)

thus allows a natural extension of the Vanderbilt PP approach to a density matrix method. Indeed, using the method outlined in Ref. 12 it can be shown that the exchange potential contribution to the pseudo-Hamiltonian is the functional $V_x^{HF} [\rho_1]$, with $\rho_1$ as given above.

**IV. RESULTS AND DISCUSSION**

In order to gauge the scaling of our HF method, as well as timings of various generalized DF approximations, we have performed calculations within the Vanderbilt scheme using a plane-wave basis set (kinetic energy cutoff=25 Ryd). Supercell calculations were done with orbitals expanded around the $k=0$ (Gamma) point.

First, the scaling behavior was obtained as a function of the number of plane waves for a fixed number of time steps in an AIMD simulation of the water monomer. To this end, cutoff energies between 100 and 400 Ryd were used to expand the density matrix in the PW basis. In order to evaluate the AE orbitals efficiently, the core contribution has been evaluated on the fine grid within a small cell surrounding each atomic site, in manner similar to that used earlier for the augmentation functions.\(^{14}\) Figure 1 shows the relative CPU times for these calculations as a function of the number of plane waves on this fine grid (which, for a given ratio of energy cutoffs, is proportional to the number of PWs in the wave function grid). As can be seen from the figure, the timings scale almost linearly with the basis set size, even for large basis sets.

![FIG. 1. CPU timings (in msec) for a fixed number of Hartree–Fock AIMD time steps for the water monomer as a function of the number of plane waves used to expand the density matrix. A linear fit is shown for comparison. (The minor scatter in the points is related to the efficiency of the nonradix-2 FFT routine.)](image)
It is also important to consider the CPU time it actually takes to do generalized DF calculations with and without exact exchange. Results were thus obtained within the local density approximation (LDA) as well as within two popular DF approximations—one which includes Becke’s exchange gradient \( E^X_B \) and the Lee-Yang-Parr (LYP) correlation \( E^C_{LYP} \)—known as the BLYP approximation, and another which includes, in addition, the HF exchange, known as the B3LYP functional. While a density matrix cutoff of 100 Ryd was found to be adequate for LDA, acceptable energy conserving dynamics for BLYP and B3LYP calculations required a cutoff of 250 Ryd. The calculations were done for a water monomer and tetramer, keeping the same cell size and PW cutoffs. It was found that BLYP calculations take about three times longer than do LDA, and HF or B3LYP calculations take between 3 and 12 times longer than LDA.

The variation in relative times for exact exchange calculations follows from its \( N_{orb}^2 \) scaling. However, methods outlined in Refs. 18 and 19 might be used to localize the orbitals so that exchange interactions can be evaluated locally. These procedures scale linearly with system size, and could reduce the complexity of our algorithm to \( O(N_{orb}N_{PW} \log N_{PW}) \), making its timings more in line with gradient approximations. This approach will be explored in the future.

V. CONCLUDING REMARKS

In summary, an algorithm has been developed within the pseudopotential+plane-wave approach for the efficient evaluation of \textit{ab initio} exchange within a Hartree–Fock approximation. As a prelude to one future application, the inclusion of exact exchange has, e.g., been found to be important in our previous studies of proton motion in the \( \text{H}_2\text{O}_2^+ \) complex, and thus it could be important for calculating the proton transfer barrier in aqueous solution. The application of the present exact exchange methodology within AIMD simulations of this and other systems will be the subject of future research.

ACKNOWLEDGMENTS

We thank Dan Berard for key discussions and David Vanderbilt for a suggestion regarding the construction of the density matrix and for drawing our attention to Ref. 11. This work was supported by the Air Force Office of Scientific Research.