Fast algorithm for periodic density fitting for Bloch waves[*]

Jianfeng Lu[†] and Lexing Ying[‡]

We propose an efficient algorithm for density fitting of Bloch waves for Hamiltonian operators with periodic potential. The algorithm is based on column selection and random Fourier projection of the orbital functions. The computational cost of the algorithm scales as $\mathcal{O}\big(N_{\text{grid}}N^2 + N_{\text{grid}}NK\log(NK)\big)$, where N_{grid} is number of spatial grid points, K is the number of sampling k-points in first Brillouin zone, and N is the number of bands under consideration. We validate the algorithm by numerical examples in both two and three dimensions.

AMS 2000 subject classifications: Primary 65Z05, 65F99.
Keywords and phrases: Density fitting, Bloch waves, column selection.

1. Introduction

Consider a Hamiltonian operator with periodic potential

$$(1) \qquad H = -\frac{1}{2}\Delta + V.$$

For simplicity of notation, we assume without loss of generality that V is periodic with respect to $\mathbb{Z}^d$, where d is the spatial dimension: *i.e.*, $V(x+e_i) = V(x)$ for $\{e_i,\ i = 1, \ldots, d\}$ the set of standard Cartesian basis vectors of $\mathbb{R}^d$. From the standard Bloch-Floquet theory (see *e.g.*, [15]), the spectrum of the operator H is given by the Bloch eigenvalue problem on the unit cell $\Gamma = [0, 1)^d$:

$$(2a) \qquad \Big(-\frac{1}{2}\Delta + V\Big)\psi_{n,k} = E_{n,k}\psi_{n,k}, \qquad \text{on } \Gamma;$$

(2b) $\qquad\qquad e^{-i(k\cdot x)}\psi_{n,k}(x)$ is periodic on Γ.

Here $k \in \Gamma^* = [-\pi, \pi)^d$ is known as the crystal momentum, where Γ^* is called the first Brillouin zone (FBZ), which is the unit cell of the reciprocal lattice. The eigenfunction $\psi_{n,k}$ for $n \in \mathbb{N}$ and $k \in \Gamma^*$ are called Bloch waves. We extend them from Γ to the whole $\mathbb{R}^3$ using the Bloch-periodicity (2b). It would be convenient to introduce the periodic part of the Bloch function, denoted as

(3) $\qquad\qquad u_{n,k}(x) = e^{-i(k\cdot x)}\psi_{n,k}(x), \qquad x \in \mathbb{R}^d.$

By definition, $u_{n,k}$ is periodic with respect to $\mathbb{Z}^d$.

In this work, we are interested in the pair density $\bar{u}_{n,k}(x)u_{m,l}(x)$ for $n, m \in \mathbb{N}, k, l \in \Gamma^*$, and $x \in \Gamma$. In particular, we aim at an approximation of the pair density of the type

(4) $\qquad\qquad \bar{u}_{n,k}(x)u_{m,l}(x) \approx \sum_\mu \bar{C}^\mu_{n,k}C^\mu_{m,l}P_\mu(x),$

where $\{P_\mu\}$ is a set of auxiliary basis functions we use to expand the pair density and $C^\mu_{n,k}$ gives the expansion coefficients. In terms of the Bloch waves, the approximation is then

(5) $\qquad\qquad \bar{\psi}_{n,k}(x)\psi_{m,l}(x) \approx \sum_\mu \bar{C}^\mu_{n,k}C^\mu_{m,l}P_\mu(x)e^{i(l-k)\cdot x}.$

The approximation of pair density in the form of

(6) $\qquad\qquad \bar{u}_{n,k}(x)u_{m,l}(x) \approx \sum_\mu T^\mu_{nk,ml}P_\mu(x)$

is known as periodic density fitting or density fitting for crystals in the literature [12, 10, 8]. Note that the approximation (4) is a special form, in which the expansion coefficient $T^\mu_{nk,ml} = \bar{C}^\mu_{n,k}C^\mu_{m,l}$ has a separable structure. The periodic density fitting is a generalization of the conventional density fitting for molecules (see *e.g.*, [16, 2, 17, 20, 21, 22]), which has wide applications in electronic structure calculations. In density fitting, the pair density for a set of orbital functions $\{\psi_n\}$ (without the k-dependence) is approximated as

(7) $\qquad\qquad \bar{\psi}_n(x)\psi_m(x) \approx \sum_\mu T^\mu_{nm}P_\mu(x).$

Here with slight abuse of notation, we use again T to denote the expansion coefficients.

Conventionally in density fitting, a set of auxiliary functions $P_\mu(x)$ are chosen, for example as Gaussian functions or atom centered orbital functions, the expansion coefficients T are obtained using least square fitting with respect to a suitable metric (L^2 or Coulomb metric, see Section 3 below). The computational cost to obtain these coefficients scales as $\mathcal{O}(N^4)$, where N is the total number of orbital functions. Moreover, the conventional procedure does not yield the separable structure, which is useful to reduce the computational cost for electronic structure calculations, see e.g., [6, 5, 13, 14, 19, 18].

In our previous work [11], we developed a $\mathcal{O}(N_{\mathrm{grid}}N^2 \log N)$ scaling algorithm for density fitting without the k point dependence, where N_{grid} is number of spatial grids and N is the number of orbital functions. The novelty is to recast the density fitting problem as a column selection procedure: Instead of fixing the auxiliary basis functions and compute the expansion coefficients, we look for a collection of grid points $\{x_\mu\}$ that well represent the pair densities, such that the auxiliary basis functions follow as a least square fitting.

$$
(8) \qquad \bar{\psi}_m(x)\psi_n(x) \approx \sum_\mu \bar{\psi}_m(x_\mu)\psi_n(x_\mu)P_\mu(x).
$$

This work extends our approach to periodic density fitting of Bloch waves. In order to further reduce the computational cost, we use a more efficient scheme for random projection based on the tensor product structure of the pair density ρ. Our resulting density fitting scheme scales as $\mathcal{O}\big(N_{\mathrm{grid}}N^2 + N_{\mathrm{grid}}NK \log(NK)\big)$, where N_{grid} is number of spatial grid points, K is the number of k-points in first Brillouin zone, and N is the number of bands under consideration. Recall that in the periodic case, the total number of Bloch waves is given by NK. Note that if we take $K = 1$ so that we go back to the situation considered in [11], our current algorithm costs as $\mathcal{O}(N_{\mathrm{grid}}N^2)$ and is hence more efficient than our previous algorithm. Moreover, the computational cost scales almost linearly with respect to the increase of K, with a "quadratic" coefficient $\mathcal{O}(N_{\mathrm{grid}}N)$; in fact, in practice, the actual running time is dominated by the $\mathcal{O}(N_{\mathrm{grid}}N^2)$ term, independent of K.

2. Algorithms

Let us consider a set of given orbitals $u_{n,k}(x)$ for $n = 1,\ldots,N$, $k \in \mathscr{K}$ a discretization of the first Brillouin zone with $|\mathscr{K}| = K$, and $x \in \mathscr{X}$ a grid

in the unit cell with $|\mathscr{X}| = N_{\mathrm{grid}}$. Similar to our previous work [11], the basic idea is to select a suitable subset of grid points to represent the pair densities $\rho_{nkml}(x) = \bar{u}_{n,k}(x)u_{m,l}(x)$.

$$(9) \qquad \rho_{nkml}(x) \approx \sum_\mu \rho_{nkml}(x_\mu)P_\mu(x) = \sum_\mu \bar{u}_{n,k}(x_\mu)u_{m,l}(x_\mu)P_\mu(x).$$

After the grid points $\{x_\mu\}$ are determined, the auxiliary basis functions P_μ follow from a least square fitting.

Note that if we view ρ as a matrix of dimension $N^2K^2 \times N_{\mathrm{grid}}$ with $(nkml)$ being the row index and x being the column index, the choice of x_μ amounts to select a sub-collection of the columns to represent all the matrix columns. One immediate approach for the column selection is to use a pivoted QR algorithm [3] on ρ. For completeness, we recall the column selection algorithm based on pivoted QR algorithm here.

Algorithm 1: Column selection based on pivoted QR

    Input   : $m \times n$ matrix M, error tolerance tol
    Output: An $m \times N_{\mathrm{col}}$ submatrix $\widetilde{M}$ of M and an $N_{\mathrm{col}} \times n$ matrix P,
            such that $M \approx \widetilde{M}P$

1 Compute the pivoted QR decomposition $[Q, R, E] = \mathrm{qr}(M)$, so that

$$QR = ME,$$

where E is an $n \times n$ permutation matrix, Q is an $m \times m$ unitary matrix and R is an $m \times n$ upper triangular matrix with diagonal entries in decreasing order;

2 Set N_{col} such that

$$R_{N_{\mathrm{col}},N_{\mathrm{col}}} \geq \mathsf{tol} \cdot R_{1,1} > R_{N_{\mathrm{col}}+1,N_{\mathrm{col}}+1};$$

3 Set $\widetilde{M} = (ME)_{:,1:N_{\mathrm{col}}}$, the first N_{col} columns of ME, where Matlab notations are used for submatrices. Note that ME amounts to a permutation of the columns of M.

4 Compute $P = R_{1:N_{\mathrm{col}},1:N_{\mathrm{col}}}^{-1} R_{1:N_{\mathrm{col}},:} E^{-1}$.

The computational cost of directly applying Algorithm 1 on the matrix ρ is however prohibitively expensive, since the computational complexity of the QR step scales as $\mathscr{O}(N_{\mathrm{grid}}N^4K^4)$ (recall that the matrix ρ has dimension $N^2K^2 \times N_{\mathrm{grid}}$).

In order to reduce the complexity, it is thus crucial to apply a random

projection on the matrix ρ, in the spirit of the pioneering works [9, 23]. The idea is to find an alternative representation of the column space of ρ by "compressing" the size of the matrix for the pivoted QR algorithm. In our previous work [11], we apply a random Fourier projection on ρ, which amounts to taking a sample of random linear combinations of rows of the ρ matrix:

$$(10) \qquad \varrho_\xi(x) = \sum_\gamma e^{-i2\pi\gamma\xi(N^2K^2)^{-1}} \eta_\gamma \rho_\gamma(x)$$

for random samples of $\xi \in \{1,\ldots,N^2K^2\}$ of total number proportional to N and η_γ being random unit complex number for each γ. Here, we have used $\gamma \in \{1,\ldots,N^2K^2\}$ to denote the row index $(nkml)$ for simplicity of notations. Fixing ξ, ϱ_ξ is a random linear combination of the rows ρ_γ. Thus, important columns of ρ remains important columns of ϱ with high probability.

The total size of the sample ξ to guarantee successful identification of important columns is $\mathscr{O}(N)$ and is much smaller than the original number of rows of ρ, and hence the cost of the pivoted QR algorithm is much reduced. The computational cost is dominated by using FFT to compute (10), and is hence $\mathscr{O}\big(N_{\mathrm{grid}}(NK)^2\log(NK)\big)$.

While the efficiency is much improved by using the algorithm in [11], it is still expensive in the current context due to the large number of rows of the matrix ρ. Our key observation to further reduce the computational cost is that ρ given by $\rho_{nkml}(x) = \bar{u}_{n,k}(x)u_{m,l}(x)$ has the tensor product structure, in the sense that the row index $(nkml)$ is naturally separated into two groups (nk) and (ml). We may explore this structure by taking the random Fourier projection on u following [1, 4] instead, which has a much smaller dimension than ρ, and hence makes the row projection more efficient. More concretely, we consider instead a projection of the type (cf. (10))

$$(11) \qquad M_{ij}(x) = \sum_\alpha e^{i2\pi\alpha\xi_i(NK)^{-1}} \bar{\eta}_\alpha\bar{u}_\alpha(x) \sum_\beta e^{-i2\pi\beta\xi_j(NK)^{-1}} \eta_\beta u_\beta(x)$$

for $\{\xi_i\} \subset \{1,\ldots,NK\}$ and $\alpha,\beta \in \{1,\ldots,NK\}$ being the row index of u, viewed as a $(NK) \times N_{\mathrm{grid}}$ matrix. Note that we can rewrite (11) as

$$(12) \qquad M_{ij}(x) = \sum_{\alpha,\beta} e^{i2\pi(\alpha\xi_i-\beta\xi_j)(NK)^{-1}} \bar{\eta}_\alpha\eta_\beta\bar{u}_\alpha(x)u_\beta(x)$$
$$= \sum_{\alpha,\beta} e^{i2\pi(\alpha\xi_i-\beta\xi_j)(NK)^{-1}} \bar{\eta}_\alpha\eta_\beta\rho_{\alpha\beta}(x).$$

and hence this still corresponds to a random linear combination of the rows of ρ, and hence could be still used to select important columns of ρ to represent the column space. The details of the algorithm are described in Algorithm 2.

Algorithm 2: Density fitting via random Fourier projection and column selection

---

    Input : Orbitals $u_{nk}(x)$

    Output: Selected grids x_μ and auxiliary basis functions $P_\mu(x)$, such that $\bar{u}_{n,k}(x)u_{m,l}(x) \approx \sum_\mu \bar{u}_{n,k}(x_\mu)u_{m,l}(x_\mu)P_\mu(x)$

1 Reshape $u_{nk}(x)$ into an $(NK) \times N_{\text{grid}}$ matrix $U_\alpha(x)$ where $\alpha = 1, \ldots, NK$ is the row index;

2 Compute the discrete Fourier transform of U left multiplied by a random diagonal matrix:

$$\widehat{U}_\xi(x) = \sum_\alpha e^{-i2\pi\alpha\xi(NK)^{-1}}\eta_\alpha U_\alpha(x),$$

where η_α is a random unit complex number for each α;

3 Choose a submatrix $\mathscr{U}$ of $\widehat{U}$ by randomly choosing $r = c\sqrt{N}$ rows.

4 Construct a $r^2 \times N_{\text{grid}}$ matrix M

$$M_{ij}(x) = \overline{\mathscr{U}_i}(x)\mathscr{U}_j(x), \qquad i,j = 1, \ldots, r,$$

where we view (ij) as the row index of M.

5 Apply column selection Algorithm 1 on the $r^2 \times N_{\text{grid}}$ matrix M to find selected N_{col} columns $\{x_\mu\}$ and auxiliary basis functions $P_\mu(x)$.

The computationally costly steps in Algorithm 2 are:

1. Step 2 requires FFT of vectors of length (NK) for N_{grid} times, and hence has computational cost $\mathscr{O}\big(N_{\text{grid}}NK\log(NK)\big)$;
2. Step 4 scales as $\mathscr{O}(r^2 N_{\text{grid}}) = \mathscr{O}(NN_{\text{grid}})$ by the choice of $r = c\sqrt{N}$;
3. Step 5 requires pivoted QR algorithm applied on a $N \times N_{\text{grid}}$ matrix, and hence has cost $\mathscr{O}(N^2 N_{\text{grid}})$, which dominates the other steps in Algorithm 1.

Therefore, the total computational cost of the algorithm scales as $\mathscr{O}\big(N_{\text{grid}}N^2 + N_{\text{grid}}NK\log(NK)\big)$. In practice, the prefactor of the FFT is much smaller than the one of the pivoted QR factorization. As a result, the $\mathscr{O}\big(N_{\text{grid}}N^2\big)$ part dominates the actual running time.

3. Numerical examples

In this section, we consider a few model examples to demonstrate the effectiveness of the algorithms proposed above. In order to measure the error of the density fitting, two metrics can be used. The error in L^2 metric measures

$$\|\rho_{nkml} - \widetilde{\rho}_{nkml}\|_2, \tag{13}$$

for $\widetilde{\rho}_{nkml} = \sum_\mu \bar{C}^\mu_{nk} C^\mu_{ml} P_\mu$, while the error in Coulomb metric is given by

(14)

$$\|\rho_{nkml} - \widetilde{\rho}_{nkml}\|_C$$
$$= \left(\iint_{\Gamma \times \Gamma} \left(\rho_{nkml} - \widetilde{\rho}_{nkml}\right)(x) G(x-y) \left(\rho_{nkml} - \widetilde{\rho}_{nkml}\right)(y) \, \mathrm{d}x \, \mathrm{d}y \right)^{1/2},$$

where G is the periodic Coulomb kernel, solving

$$-\Delta G(\cdot - y) = 4\pi(\delta_y - 1) \tag{15}$$

with periodic boundary condition on $\Gamma = [0,1]^d$ and $\int_\Gamma G = 0$.

The Coulomb metric is a good measure of the error since density fitting is often used in approximating electron repulsion integral tensor, such as

(16)
$$E_{nkml} = \iint_{\Gamma \times \Gamma} \bar{\psi}_{n,k}(x)\psi_{m,l}(x)G(x-y)\bar{\psi}_{m,l}(y)\psi_{n,k}(y) \, \mathrm{d}x \, \mathrm{d}y$$
$$= \iint_{\Gamma \times \Gamma} \bar{u}_{n,k}(x)u_{m,l}(x)G(x-y)\bar{u}_{m,l}(y)u_{n,k}(y)e^{-i(k-l)\cdot(x-y)} \, \mathrm{d}x \, \mathrm{d}y.$$

Using the density fitting (4), we get the approximation

(17)
$$\widetilde{E}_{nkml} = \sum_{\mu\nu} \iint_{\Gamma \times \Gamma} \bar{C}^\mu_{n,k} C^\mu_{m,l} P_\mu(x) G(x-y) \bar{C}^\nu_{m,l} C^\nu_{n,k} P_\nu(y) e^{-i(k-l)\cdot(x-y)} \, \mathrm{d}x \, \mathrm{d}y$$
$$= \sum_{\mu\nu} \bar{C}^\mu_{n,k} C^\mu_{m,l} \bar{C}^\nu_{m,l} C^\nu_{n,k} \iint_{\Gamma \times \Gamma} P_\mu(x) G(x-y) P_\nu(y) e^{-i(k-l)\cdot(x-y)} \, \mathrm{d}x \, \mathrm{d}y.$$

A simple calculation gives

(18)

$$|E_{nkml} - \widetilde{E}_{nkml}|$$

$$\begin{aligned}
&= \left| \iint \Big(\rho_{nkml}(x)\bar{\rho}_{nkml}(y) - \widetilde{\rho}_{nkml}(x)\bar{\widetilde{\rho}}_{nkml}(y) \Big) \right. \\
&\qquad \left. \times\, G(x-y)e^{-i(k-l)\cdot(x-y)}\,\mathrm{d}x\,\mathrm{d}y \right| \\
&\leq \left| \iint \big(\rho_{nkml}(x) - \widetilde{\rho}_{nkml}(x)\big)\bar{\rho}_{nkml}(y)G(x-y)e^{-i(k-l)\cdot(x-y)}\,\mathrm{d}x\,\mathrm{d}y \right| \\
&\quad + \left| \iint \widetilde{\rho}_{nkml}(x)\big(\bar{\rho}_{nkml}(y) - \bar{\widetilde{\rho}}_{nkml}(y)\big)G(x-y)e^{-i(k-l)\cdot(x-y)}\,\mathrm{d}x\,\mathrm{d}y \right| \\
&\leq C\|\rho_{nkml} - \widetilde{\rho}_{nkml}\|_C \big(\|\rho_{nkml}\|_C + \|\widetilde{\rho}_{nkml}\|_C\big),
\end{aligned}$$

where the last step uses the Fourier representation and the Cauchy-Schwartz inequality.

3.1. 2D examples

We consider two examples in two dimensions. In both examples, the periodic model potential $V(x)$ is given by

$$V(x) = \sum_{n_1 \in \mathbb{Z}} \sum_{n_2 \in \mathbb{Z}} V_c(x - n_1 e_1 - n_2 e_2)$$

where V_c is a localized potential centered at the origin and e_1 and e_2 are the unit Cartesian basis vectors. This is a simple square lattice and the Brillouin zone is also a square. The potential $V_c(\cdot)$ is chosen to be centered at the origin and spherically symmetric. The unit cell $[0,1)^2$ is discretized with $N_{\mathrm{grid}} = 48$ points in each dimension. In each example, we consider up to $N = 41$ bands and the Brillouin zone $[-\pi, \pi)^2$ is sampled with up to $K = 16$ points per dimension. The band structure and the eigenfunctions are computed with spectral discretization for high order accuracy and the resulting discrete eigenvalue problem for each k point is solved using LOBPCG [7] with the simple inverse Laplacian preconditioning. The prescribed accuracy for the column selection algorithm is set to be of order 10^{-5}.

In the first example, the potential $V_c(x)$ is given by a Gaussian profile

(19)
$$V_c(x) = -144 \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right)$$

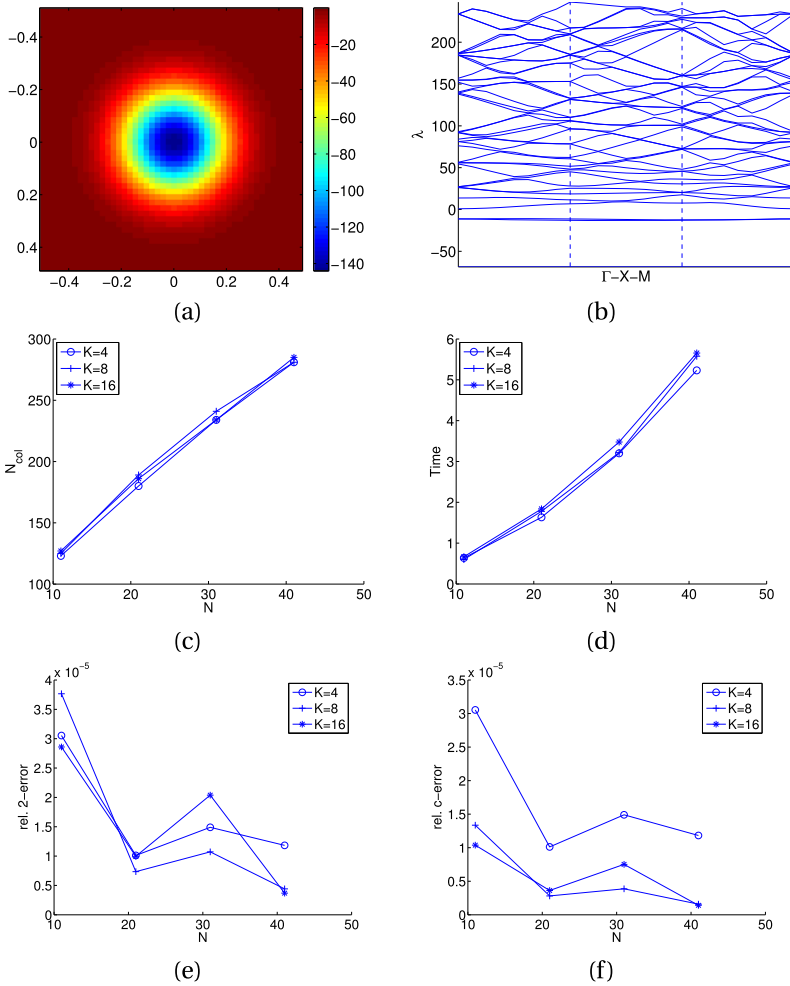with $\sigma = 0.1333$. Figure 1 summarizes the results for this first example.

Figure 1: Results of the first 2D example. (a) the potential V_c in unit cell. (b) the band structure plotted along the $\Gamma - X - M$ path. (c) N_{col} as a function N (the number of bands) for different values of K (the number of k points). (d) The time used by the column selection algorithm in seconds. (e) The relative error measured in the L^2 metric. (f) The relative error measured in the Coulomb metric.

- Figure 1(a) shows the potential profile in the unit cell $[-1/2, 1/2]^2$.
- Figure 1(b) gives the band structure along the standard $\Gamma - X - M$ path in the Brillouin zone for simple square lattice.
- Figure 1(c) plots the dependence of N_{col} (the number of auxiliary basis) as a function of N (the number of bands) for different choices of
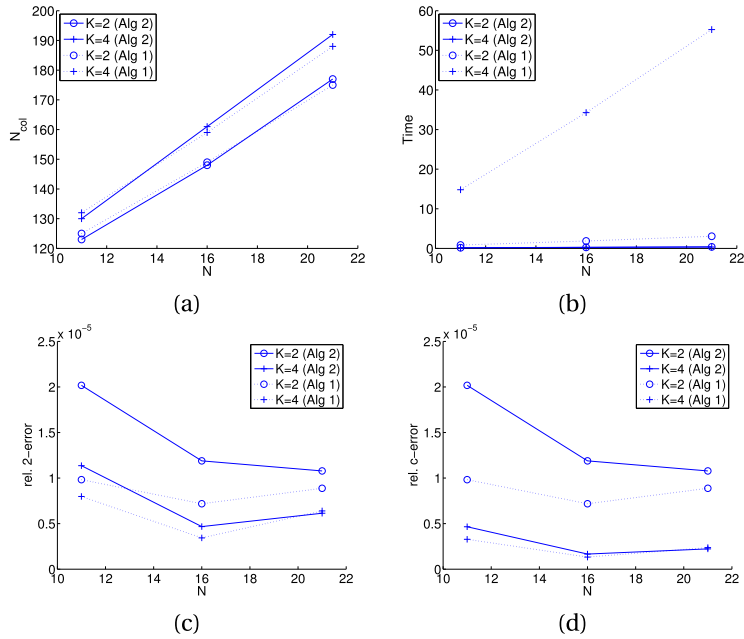
Figure 2: Comparison of Algorithm 2 (solid line) and Algorithm 1 (dotted line) for the first 2D example. (a) N_{col} as a function N (the number of bands) for different values of K (the number of k points). (b) The time used by the column selection algorithm in seconds. (c) The relative error measured in the L^2 metric. (d) The relative error measured in the Coulomb metric.

> K (the number of k-point samples in each dimension) at the targeted accuracy level 10^{-5}. From this plot, one can clearly see that N_{col} grows roughly linearly with respect to N and that it depends only mildly on the number of k point samples.

- The time for performing the column selection in seconds is given in Figure 1(d) as a function of N for different values of K. This plot shows that the complexity depends quadratically on N but only grows mildly with K.

- Finally, Figures 1(e) and (f) give the relative errors in the L^2 metric and the Coulomb metric, respectively. These two plots show that the estimated relative errors are bounded by a small constant times the prescribed accuracy level.

In order to demonstrate the effectiveness of Algorithm 2, a comparison study for Algorithm 1 and Algorithm 2 is performed using this example and the results are summarized in Figure 2. Since Algorithm 1 is quite time-

consuming, a lower resolution study is adopted here with $N_{\text{grid}} = 24$ and smaller values for both N and K. The plots in Figure 2 clearly show that Algorithm 2 is much more efficient while at the same time maintains the same level of accuracy.

In the second example, the potential $V_c(x)$ is chosen to be

$$(20) \qquad V_c(x) = -144 \exp\left(-\frac{\max(\|x\| - 1/4, 0)^2}{2\sigma^2}\right)$$

with $\sigma = 0.0667$. The results for this example are summarized in Figure 3. The results show similar characteristics as the previous example.

3.2. 3D examples

For the 3D case, we consider two similar examples. In each example, the periodic model potential $V(x)$ is given by

$$V(x) = \sum_{n_1 \in \mathbb{Z}} \sum_{n_2 \in \mathbb{Z}} \sum_{n_3 \in \mathbb{Z}} V_c(x - n_1 e_1 - n_2 e_2 - n_3 e_3)$$

where V_c is a localized potential centered at the origin and e_1, e_2, and e_3 are the unit Cartesian basis vectors. This is a simple cubic lattice and the Brillouin zone is also a cube. The potential $V_c(\cdot)$ is again chosen to be centered at the origin and spherically symmetric. The unit cell $[0, 1)^3$ is discretized with $N_{\text{grid}} = 24$ points in each dimension. In each example, we consider up to $N = 41$ bands and the Brillouin zone $[-\pi, \pi)^3$ is sampled with up to $K = 12$ points per dimension. Similar to the 2D examples, the band structure and the eigenfunctions are computed with pseudo-spectral discretization for high order accuracy and the resulting discrete eigenvalue problem for each k point is solved using LOBPCG with simple inverse Laplacian preconditioning. The prescribed accuracy for the column selection algorithm is set to be of order 10^{-5}.

The potential $V_c(x)$ in the first example is given by (19) but now in 3D and with $\sigma = 0.1667$. Figure 4 summarizes the results for this examples. The meaning of each plot is similar to the ones of the 2D examples, except that

- Figure 4 (a) is the cross-section of the potential $V_c(x)$ taken at the plane $x_3 = 0$;
- Figure 4 (b) gives the band structure along the standard $\Gamma - X - M - \Gamma - R - X | M - R$ path in the Brillouin zone for the simple cubic lattice.
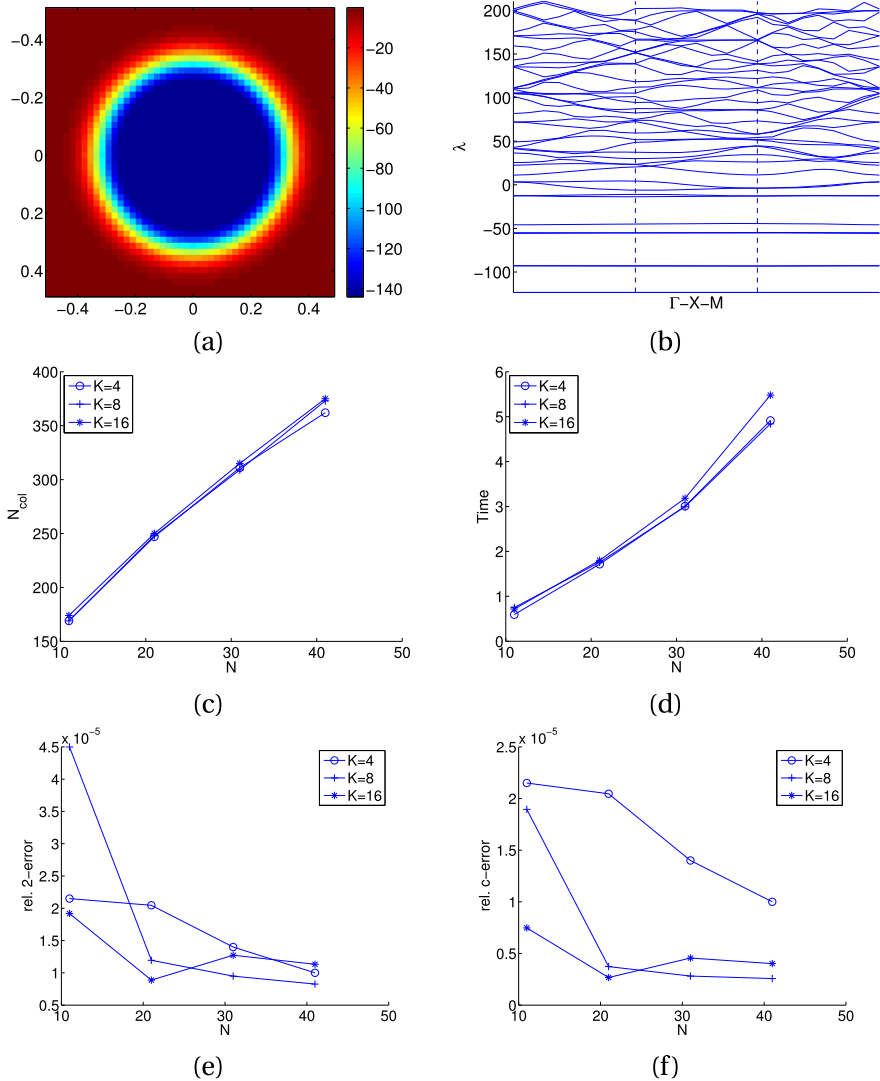
Figure 3: Results of the second 2D example. (a) the potential V_c in unit cell. (b) the band structure plotted along the $\Gamma - X - M$ path. (c) N_{col} as a function N (the number of bands) for different values of K (the number of k points). (d) The time used by the column selection algorithm in seconds. (e) The relative error measured in the L^2 metric. (f) The relative error measured in the Coulomb metric.
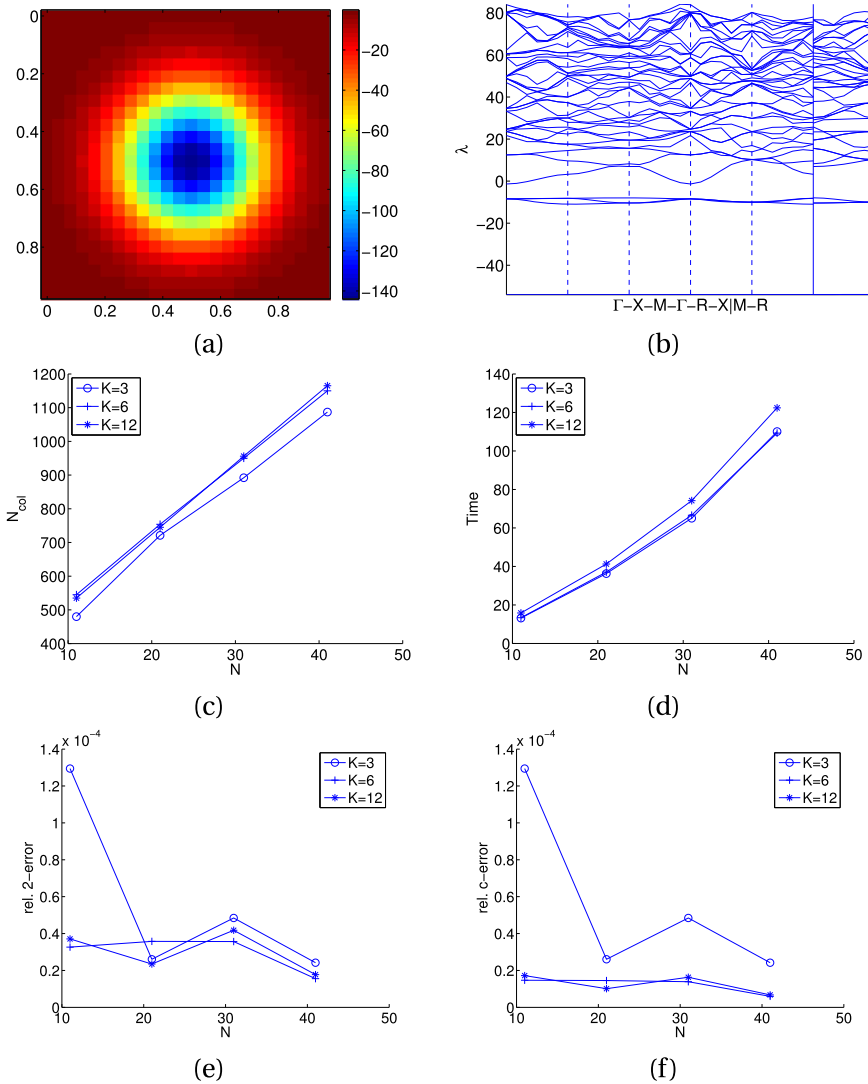
Figure 4: Results of the first 3D example. (a) the cross-section of the potential V_c in unit cell taken at $x_3 = 0$. (b) the band structure plotted along the $\Gamma - X - M - \Gamma - R - X | M - R$ path. (c) N_{col} as a function N (the number of bands) for different values of K (the number of k points). (d) The time used by the column selection algorithm in seconds. (e) The relative error measured in the L^2 metric. (f) The relative error measured in the Coulomb metric.
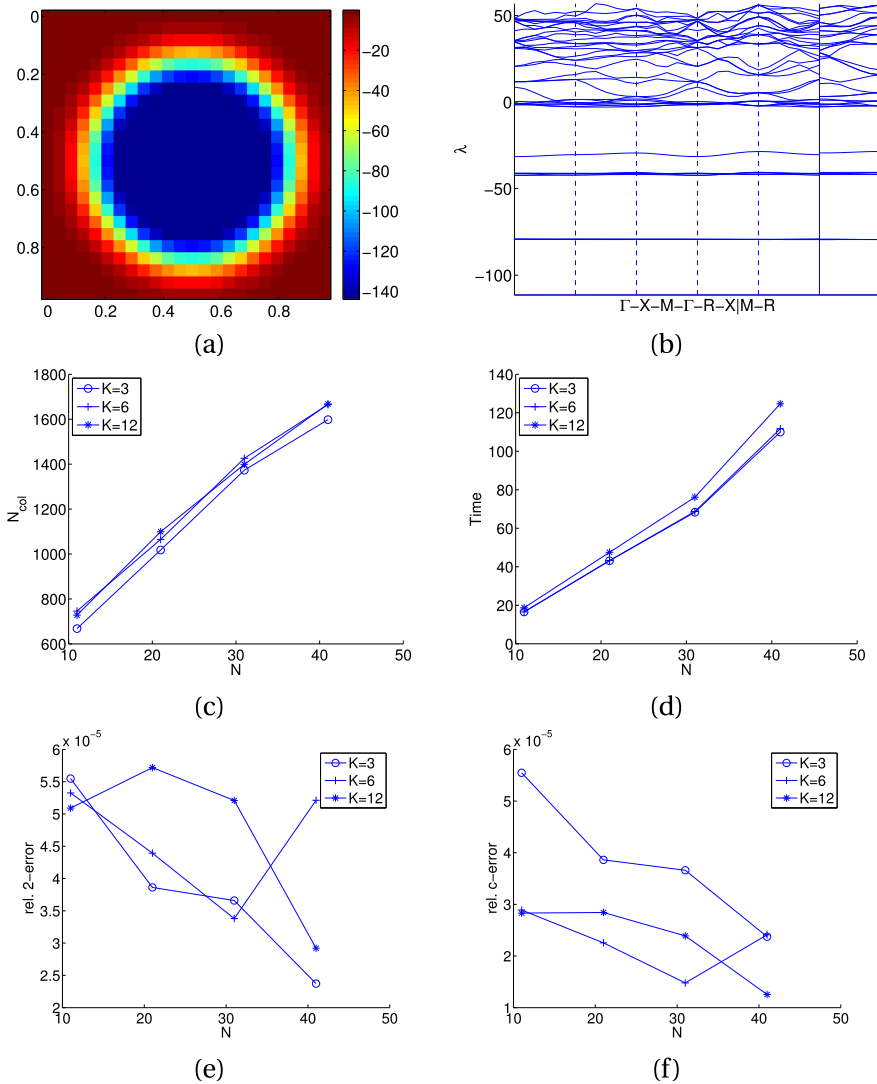
(a)

(b)

(c)

(d)

(e)

(f)

Figure 5: Results of the second 3D example. (a) the cross-section of the potential V_c in unit cell taken at $x_3 = 0$. (b) the band structure plotted along the $\Gamma - X - M - \Gamma - R - X | M - R$ path. (c) N_{col} as a function N (the number of bands) for different values of K (the number of k points). (d) The time used by the column selection algorithm in seconds. (e) The relative error measured in the L^2 metric. (f) The relative error measured in the Coulomb metric.
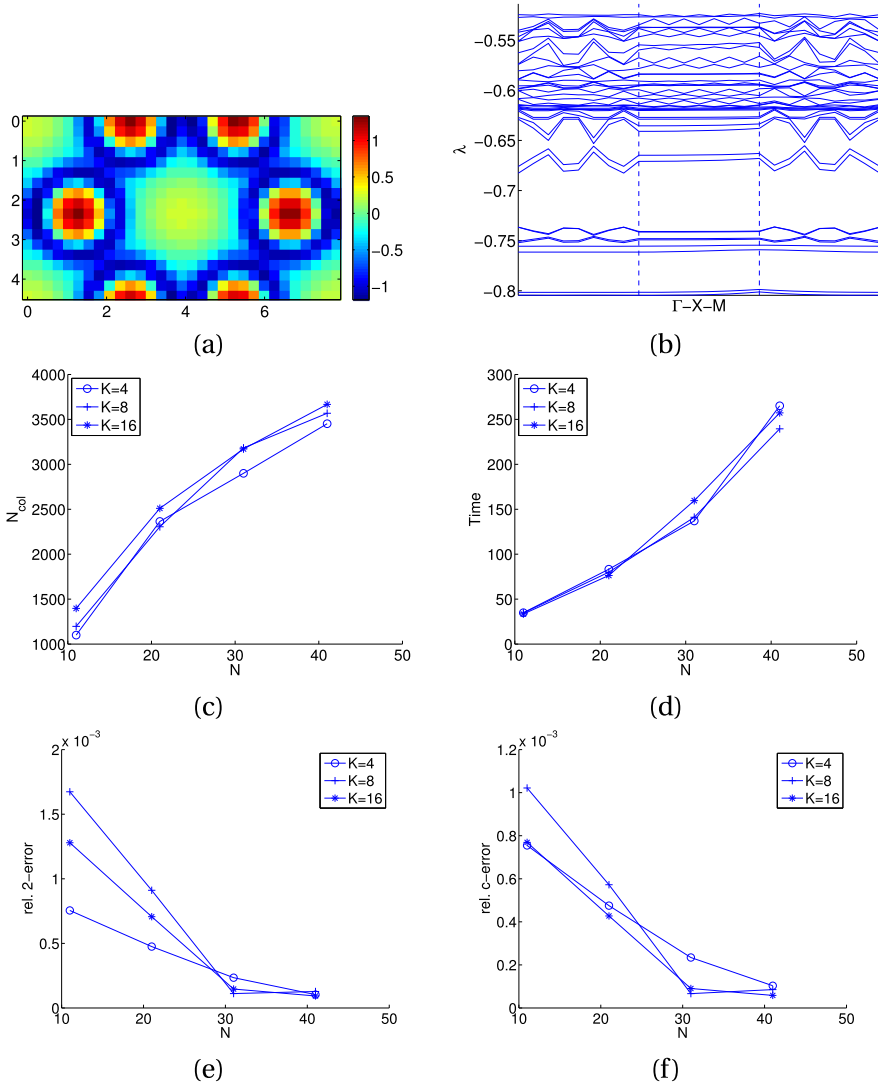
(a)

(b)

(c)

(d)

(e)

(f)

Figure 6: Results of the graphene sheet example. (a) the cross-section of the potential V_c through the atom centers. (b) the band structure plotted along the $\Gamma - X - M$ path. (c) N_{col} as a function N (the number of bands) for different values of K (the number of k points). (d) The time used by the column selection algorithm in seconds. (e) The relative error measured in the L^2 metric. (f) The relative error measured in the Coulomb metric.

Similar to the results from the 2D examples, we can draw the following conclusions:

- N_{col} depends roughly linearly on N but is almost independent of K;
- the time of the column selection algorithm grows quadratically in N but is almost independent of K;
- the relative errors of the algorithm are bounded by a small factor times the accuracy level used in the column selection algorithm both in the L^2 metric and the Coulomb metric.

In the second example, the potential $V_c(x)$ given by (20) but again interpreted in 3D and with $\sigma = 0.0833$. The results are summarized in Figure 5.

3.3. Graphene sheet

Finally, we consider the graphene sheet, a three dimensional system with two dimensional periodic structure, with the results summarized in Figure 6. We calculate the Bloch functions for an effective potential obtained from a Γ-point calculation using KSSOLV [24], a Matlab toolbox for the Kohn-Sham equations. This example is significantly more challenging that the previous example since the potential profile in the unit cell has more structure. As a result, the number of auxiliary basis N_{col} is significantly increased, but it is still much smaller compared to the number of grid points. The results for the running times and relative errors are qualitatively similar to the ones of the previous examples.

4. Conclusion

We present and demonstrate an efficient algorithm for periodic density fitting for Bloch waves. The proposed algorithm is based on randomized pivoted QR algorithm with a choice of random projection adapted to the tensor product structure of the matrix. The resulting algorithm is validated in several numerical examples in two and three dimensions.

References

[1] Ailon, N. and Chazelle, B. (2009). The fast Johnson-Lindenstrauss transform and approximate nearest neighbors. *SIAM J. Comput.* **39** 302–322. MR2506527

[2] Dunlap, B. I., Connolly, J. W. D. and Sabin, J. R. (1979). On first-row diatomic molecules and local density models. *J. Chem. Phys.* **71** 4993–4999.

[3] Golub, G. H. and Van Loan, C. F. (2013). *Matrix computations*, fourth ed. *Johns Hopkins Studies in the Mathematical Sciences*. Johns Hopkins University Press, Baltimore, MD. MR3024913

[4] Halko, N., Martinsson, P. G. and Tropp, J. A. (2011). Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.* **53** 217–288. MR2806637

[5] Hohenstein, E. G., Parrish, R. M. and Martinez, T. J. (2012). Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **137** 044103.

[6] Hohenstein, E. G., Kokkila, S. I. L., Parrish, R. M. and Martinez, T. J. (2013). Quartic scaling second-order approximate coupled cluster singles and doubles via tensor hypercontraction: THC-CC2. *J. Chem. Phys.* **138** 124111.

[7] Knyazev, A. V. (2001). Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method. *SIAM J. Sci. Comput.* **23** 517–541. MR1861263

[8] Levchenko, S. V., Ren, X., Wieferink, J., Johanni, R., Rinke, P., Blum, V. and Scheffler, M. (2015). Hybrid functionals for large periodic systems in an all-electron, numeric atom-centered basis framework. *Comp. Phys. Commun.* **192** 60–69. MR3336695

[9] Liberty, E., Woolfe, F., Martinsson, P. G., Rokhlin, V. and Tygert, M. (2007). Randomized algorithms for the low-rank approximation of matrices. *Proc. Natl. Acad. Sci. USA* **104** 20167–20172. MR2366406

[10] Lorenz, M., Usvyat, D. and Schütz, M. (2011). Local ab initio methods for calculating optical band gaps in periodic systems. I. Periodic density fitted local configuration interactino singles method for polymers. *J. Chem. Phys.* **134** 094101.

[11] Lu, J. and Ying, L. (2015). Compression of the electron repulsion integral tensor in tensor hypercontraction format with cubic scaling cost. *J. Comput. Phys.* **302** 329–335. MR3404522

[12] Maschio, L., Usvyat, D., Manby, F. R., Casassa, S., Pisani, C. and Schütz, M. (2007). Fast local-MP2 method with density-fitting for crystals. I. Theory and algorithms. *Phys. Rev. B* **76** 075101.

[13] Parrish, R. M., Hohenstein, E. G., Martinez, T. J. and David Sherrill, C. (2012). Tensor hypercontraction. II. Least-squares renormalization. *J. Chem. Phys.* **137** 224106.

[14] Parrish, R. M., David Sherrill, C., Hohenstein, E. G., Kokkila, S. I. L. and Martinez, T. J. (2014). Communication: Acceleration of coupled cluster singles and doubles via orbital-weighted least-squares tensor hypercontraction. *J. Chem. Phys.* **140** 181102.

[15] Reed, M. and Simon, B. (1980). *Methods of Modern Mathematical Physics, Vol IV*. Academic Press, New York. MR0751959

[16] Ren, X., Rinke, P., Blum, V., Wieferink, J., Tkatchenko, A., Sanfilippo, A., Reuter, K. and Scheffler, M. (2012). Resolution-of-identity approach to Hartree-Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions. *New J. Phys.* **14** 053020.

[17] Schütz, M., Usvyat, D., Lorenz, M., Pisani, C., Maschio, L., Casassa, S. and Halo, M. (2010). Density fitting for correlated calculations in periodic systems. In *Accurate Condensed-Phase Quantum Chemistry*, (F. Manby, ed.). *Computation in Chemistry* 27. CRC Press.

[18] Shenvi, N., van Aggelen, H., Yang, Y. and Yang, W. (2014). Tensor hypercontracted ppRPA: Reducing the cost of the particle-particle random phase approximation from $O(r^6)$ to $O(r^4)$. *J. Chem. Phys.* **141** 024119.

[19] Shenvi, N., van Aggelen, H., Yang, Y., Yang, W., Schwerdtfeger, C. and Mazziotti, D. (2013). The tensor hypercontracted parametric reduced density matrix algorithm: Coupled-cluster accuracy with $O(r^4)$ scaling. *J. Chem. Phys.* **139** 054110.

[20] Sodt, A., Subotnik, J. E. and Head-Gordon, M. (2006). Linear scaling density fitting. *J. Chem. Phys.* **125** 194109.

[21] Vahtras, O., Almlöf, J. and Feyereisen, M. W. (1993). Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **213** 514–518.

[22] Weigend, F., Häser, M., Patzelt, H. and Ahlrichs, R. (1998). RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **294** 143–152.

[23] Woolfe, F., Liberty, E., Rokhlin, V. and Tygert, M. (2008). A fast randomized algorithm for the approximation of matrices. *Appl. Comput. Harmon. Anal.* **25** 335–366. MR2455599

[24] Yang, C., Meza, J. C., Lee, B. and Wang, L. W. (2009). KSSOLV– a MATLAB toolbox for solving the Kohn-Sham equations. *ACM T. Math. Software* **36** 10. MR2738191

Jianfeng Lu
Department of Mathematics
Department of Physics, and Department of Chemistry
Duke University, Box 90320
Durham NC 27708
USA
E-mail address: jianfeng@math.duke.edu

Lexing Ying
Department of Mathematics and
Institute of Computational and Mathematical Engineering
Stanford University
450 Serra Mall, Bldg 380
Stanford CA 94305
USA
E-mail address: lexing@stanford.edu