

# A wavelet-based Projector Augmented-Wave (PAW) method: reaching frozen-core all-electron precision with a systematic, adaptive and localized wavelet basis set

T. Rangel<sup>a,b</sup>, D. Caliste<sup>c</sup>, L. Genovese<sup>c</sup>, M. Torrent<sup>a</sup>

<sup>a</sup>CEA, DAM, DIF, F-91297 Arpajon, France

<sup>b</sup>Current affiliation: Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

<sup>c</sup>Univ. Grenoble Alpes, CEA, INAC-MEM, L-Sim, F-38000 Grenoble, France

---

## Abstract

We present a Projector Augmented-Wave (PAW) method based on a wavelet basis set. We implemented our wavelet-PAW method as a PAW library in the ABINIT package [<http://www.abinit.org>] and into BigDFT [<http://www.bigdft.org>]. We test our implementation in prototypical systems to illustrate the potential usage of our code. By using the wavelet-PAW method, we can simulate charged and special boundary condition systems with frozen-core all-electron precision. Furthermore, our work paves the way to large-scale and potentially order- $N$  simulations within a PAW method.

*Keywords:* Wavelets, PAW, Density Functional Theory, Electronic Structure

*PACS:* 71.15.Ap

---

## 1. Introduction

Density Functional Theory (DFT) has recently gained popularity due to its inherent efficiency and simplicity. It is one of the most reliable first-principles methods to simulate material properties. Presently, DFT is widely used in the physics, chemistry and biology communities.

DFT codes have largely been developed to harness recent advances in modern computing. Parallel supercomputers consisting of thousands of processors have become the norm. Hence, highly-parallel architectures and efficient methods are desired in a DFT code. The BIGDFT library [1], which is integrated in the ABINIT software suite [2], has been conceived and implemented for massively-parallel environments in view of these large-scale calculations.

The quintessential characteristic of BIGDFT is the use of a Daubechies [3] wavelet (WVL) basis set to express the Kohn-Sham (KS) orbitals. WVLs form a localized, systematic, orthogonal and adaptive basis set, and hence combine most of the advantages of computational basis sets for DFT. A WVL basis set consists of a set of inter-related scaling functions and WVLs in a uniform grid of spacing  $h$ . Convergence is, thus, achieved by decreasing the value of  $h$  up to the required accuracy. Moreover, in BIGDFT a two-level grid is defined. There is a fine resolution grid close to the atoms to represent the chemical bonds and atomic orbitals, and secondly a coarse grid encompasses a larger volume further away from the atoms

until the wavefunction vanishes. This high-degree of adaptivity makes the BigDFT library optimal for calculations of polarized systems or systems with non Born-von Karman boundary conditions, such as material surfaces and isolated molecules.

Over the last few decades, the Projector Augmented-Wave (PAW) method [4] has proven successful in electronic structure calculations due to its frozen-core all-electron (AE) accuracy at low computational cost due to ultra-soft pseudopotentials (USPP). Within PAW, the AE wavefunction  $\Psi$  is obtained from an auxiliary smooth wavefunction  $\tilde{\Psi}$  by a linear transformation  $\mathcal{T}$ , as  $|\Psi\rangle = \mathcal{T}|\tilde{\Psi}\rangle$ . Close to the atoms in the so called “augmentation regions” wavefunctions tend to have rapid oscillations. This can be problematic since many basis elements are usually required to express the wavefunctions within norm-conserving pseudopotentials (NCP). However, within PAW a minimal basis set of atomic partial waves is sufficient to express the wavefunctions in these regions. Due to its accuracy and high computing performance, the PAW method has gained enormous popularity and it has been implemented in several DFT codes, which use either plane-wave [2, 5, 6, 7, 8, 9] or real-space [10, 11, 12, 13] approaches. To our knowledge the PAW method has never been implemented within WVLs.

In this work, we detail a WVL-based PAW implementation in the ABINIT package using the BIGDFT library. By making a PAW library the method has been ported into BIGDFT in stand-alone mode, and therefore it can be combined with different basis sets in other codes. Our approach presents the PAW AE accuracy, exploits the

---

*Email address:* [trange1@lbl.gov](mailto:trange1@lbl.gov) (T. Rangel)

wavelet adaptability and can potentially benefit from the BigDFT order- $N$ [14, 15] capabilities.

The paper is organized as follows. We give an overview of the theory of WVl-based DFT and PAW in Section 2. In Section 3, we explain the technical details of the implementation, detailing the approximations used to solve the KS equations. Input variables used in the code are introduced to help readers using the code. In Section 5, we show numerical tests illustrating the applicability of our code. Finally, we draw the conclusions and discuss future directions of this work in Section 7.

## 2. Background theory

### 2.1. Wavelet-based Density Functional Theory

In this Section, the theoretical background of the WVl formalism is briefly presented following the specific choices made by the BIGDFT library. We emphasize that we neither discuss the advantages of wavelets as a basis set nor the main operations related to the BigDFT code, as these have been previously detailed in Refs. 1, 16, 17, 18 and 19. In this work, we rather present our implementation of PAW associated with a wavelet basis set.

#### 2.1.1. Basis set

We use the so called maximally symmetric Daubechies WVl basis of order 16, [3] since they present virtually ideal properties of a basis set: they are orthogonal, systematic, can represent exactly up to 8-th order polynomials, and with translational invariance. [1, 3, 16, 17, 18, 19, 20] In general, a WVl basis set consists of two objects: WVls  $\psi(x)$  and scaling functions  $\phi(x)$ . An illustration of the WVls and scaling function is shown in Figure 1.

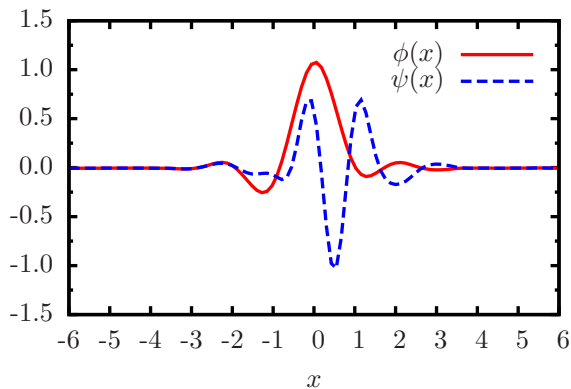


Figure 1: (Color online) Daubechies wavelet  $\psi(x)$  and scaling function  $\phi(x)$  of order 16.

In the present implementation we use a two-level adaptive grid. Away from the atoms, we use a “coarse” basis set, which consists of three-dimensional (3D) scaling functions on a grid with uniform mesh  $h$ :

$$\phi_{i,j,k}(\mathbf{r}) = \phi(x/h - i)\phi(y/h - j)\phi(z/h - k) \quad (1)$$

Close to the atoms we use a fine resolution degrees of freedom, with both scaling functions and wavelets, as high-resolution is needed to express the chemical bonds and the atomic orbitals. The basis set is here augmented by a set of seven 3D WVls:

$$\begin{aligned} \psi_{i,j,k}^1(\mathbf{r}) &= \psi(x/h - i)\phi(y/h - j)\phi(z/h - k), \\ \psi_{i,j,k}^2(\mathbf{r}) &= \phi(x/h - i)\psi(y/h - j)\phi(z/h - k), \\ \psi_{i,j,k}^3(\mathbf{r}) &= \psi(x/h - i)\psi(y/h - j)\phi(z/h - k), \\ \psi_{i,j,k}^4(\mathbf{r}) &= \phi(x/h - i)\phi(y/h - j)\psi(z/h - k), \\ \psi_{i,j,k}^5(\mathbf{r}) &= \psi(x/h - i)\phi(y/h - j)\psi(z/h - k), \\ \psi_{i,j,k}^6(\mathbf{r}) &= \phi(x/h - i)\psi(y/h - j)\psi(z/h - k), \text{ and} \\ \psi_{i,j,k}^7(\mathbf{r}) &= \psi(x/h - i)\psi(y/h - j)\psi(z/h - k). \end{aligned} \quad (2)$$

Due to the exponential localization of wavefunctions, far from the atoms, where the wavefunctions vanish, no basis elements are used. To illustrate, in Figure 2 the 2-level adaptive grid around a naphthalene molecule is shown.

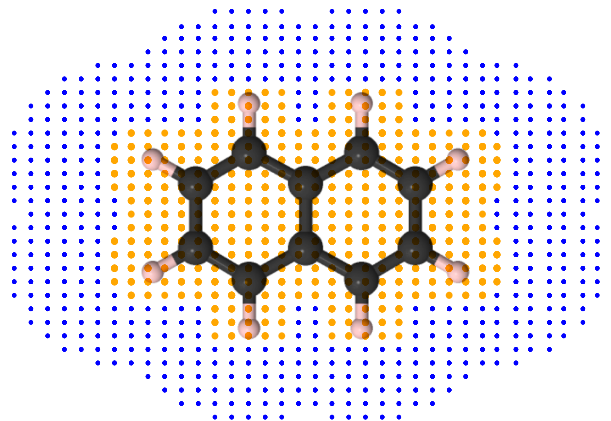


Figure 2: (Color online) Adaptive grid around a naphthalene molecule. The high- and low- resolution grids points are shown in orange and blue dots, respectively. Carbon and Hydrogen atoms are drawn with black and pink spheres, respectively.

The wavefunctions  $\Psi(\mathbf{r})$  are thus expanded as:

$$\Psi(\mathbf{r}) = \sum_{i_1, i_2, i_3} s_{i_1, i_2, i_3} \phi_{i_1, i_2, i_3}(\mathbf{r}) + \sum_{j_1, j_2, j_3} \sum_{\nu=1}^7 d_{j_1, j_2, j_3}^{\nu} \psi_{j_1, j_2, j_3}^{\nu}(\mathbf{r}), \quad (3)$$

where  $s_{i_1, i_2, i_3}$  and  $d_{j_1, j_2, j_3}^{\nu}$  are expansion coefficients. The sum over  $i_1$ ,  $i_2$  and  $i_3$  runs over all points on the coarse grid and the sum over  $j_1$ ,  $j_2$  and  $j_3$  runs over points on the fine grid.

In our implementation we exploit the *separability* property of the basis functions, in which the 3D scaling functions basis set, being a product decomposition of one-dimensional (1D) scaling functions/wavelets, is separable in the three Cartesian directions, see Eq. (1).

BIGDFT treats efficiently *Gaussian pseudopotentials* of the Goedecker-Teter-Hutter (GTH) [21] and Hartwigsen-Goedecker-Hutter (HGH) [22] kinds, since the intrinsic separability of both the basis set and Gaussian pseudopotentials allows for the simplification of several 3D operations into a sum of 1D products. Following the same spirit, for PAW it makes sense to use non-local projectors expressed as sum of Gaussians, as described in Section. 3.2.

## 2.2. The PAW method

In this Section, we briefly review the PAW formalism. We adopt the notation of Ref. 23.

### 2.2.1. The PAW transformation

In the PAW scheme, the true AE wavefunctions  $\Psi$  are obtained from the auxiliary wavefunctions  $\tilde{\Psi}$  (known as pseudo-wavefunctions) by applying a linear operator  $\mathcal{T}$ , expressed as a sum of atom-dependent contributions:

$$|\Psi\rangle = \mathcal{T}|\tilde{\Psi}\rangle = \left(1 + \sum_R \mathcal{T}_R\right)|\tilde{\Psi}\rangle, \quad (4)$$

where  $\mathcal{T}_R$  are local contributions acting only in an *augmentation region* around the atom. The PAW transformation takes the form:

$$|\Psi\rangle = |\tilde{\Psi}\rangle + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle\right) \langle\tilde{p}_i|\tilde{\Psi}\rangle, \quad (5)$$

where the  $i$  index runs on atom position  $\mathbf{R}$ , angular momentum ( $l, m$ ) and additional index  $m'$  for different partial waves with the same angular momentum and atom site. The AE  $\phi_i$  and pseudo  $\tilde{\phi}_i$  partial waves are identical outside the augmentation region. The partial waves and the projectors  $\tilde{p}_i$  are calculated in a spherical grid inside the augmentation region separating the radial and angular parts:

$$\phi_i(\mathbf{r}) = \frac{\phi_{n_i, l_i}(r)}{r} S_{l_i, m_i}; \quad \tilde{\phi}_i(\mathbf{r}) = \frac{\tilde{\phi}_{n_i, l_i}(r)}{r} S_{l_i, m_i}(\hat{r}) \quad (6)$$

$$\tilde{p}_i(\mathbf{r}) = \frac{\tilde{p}_{n_i, l_i}(r)}{r} S_{l_i, m_i}(\hat{r}), \quad (7)$$

where  $S_{l, m}(\hat{r})$  are the real spherical harmonics. A more extensive review of the method can be found in the original paper of Blöchl [4] and the implementation in ABINIT within plane-waves is detailed in Refs. [23, 24].

From Eq. (5), the AE valence charge density becomes:

$$n_v = \sum_{nk} f_{nk} |\Psi_{nk}|^2 = \tilde{n} + n^1 - \tilde{n}^1, \quad (8)$$

where  $f_{nk}$  is the occupation number of band  $n$  at  $k$ -point  $k$ . The pseudized density  $\tilde{n}$  is simply  $\tilde{n} = \sum f_{nk} |\tilde{\Psi}_{nk}|^2$ , akin the NCPP charge density.  $n^1$  and  $\tilde{n}^1$  are the AE and pseudized on-site densities, respectively. These are only defined in the augmentation regions. Note that the superscript "1" refers to atomic quantities.

To evaluate the on-site densities, the occupancy matrix  $\rho_{ij}$  is required [23],

$$\rho_{ij} = \sum_{nk} f_{nk} (c_{nk}^i)^* c_{nk}^j; \quad c_{nk}^i = \langle\tilde{p}_i|\tilde{\Psi}_{nk}\rangle. \quad (9)$$

The *compensation charge*  $\hat{n}$  [4, 7] is added to the pseudized densities  $\tilde{n}$  and  $\tilde{n}^1$ . The purpose of  $\hat{n}$ , as introduced in Ref. [25], is to add the correct amount of charge moments to the valence pseudo density  $\tilde{n}$  so that outside the augmentation region of all the atoms, the Coulomb (or Hartree) potential for the sum of the valence pseudo and compensation charge densities ( $\tilde{n}(\mathbf{r}) + \hat{n}(\mathbf{r})$ ) is the same as that for the fully nodal valence electron density  $n(\mathbf{r})$ . *i.e.*,  $\hat{n}$  is constructed ensuring that the Coulomb potential created from the on-site densities cancels out outside the augmentation regions, avoiding electrostatic interactions between PAW spheres. Further, a high resolution around the atoms is needed to ensure that the pseudo densities inside and outside the augmentation regions exactly cancel out. Therefore, a *double-grid* technique [26] is usually used to represent the density and potential terms.

### 2.2.2. The PAW Kohn-Sham Hamiltonian

The Kohn-Sham (KS) Hamiltonian  $\mathcal{H}$  is defined as [25, 27]:

$$\mathcal{H} = T + V_L + V_{NL} \quad (10)$$

where the effective potential consists of the kinetic  $T$ , the local (L) and the non-local (NL) potentials. The local potential  $V_L$  is given by

$$V_L = V_{xc} + V_H + \sum_i^{\text{ion}} V_{\text{loc}}^{(i)}, \quad (11)$$

where  $V_{xc}$ ,  $V_H$  and  $V_{\text{loc}}^{(i)}$  are the exchange-correlation, Hartree and local-ionic potentials, respectively. The non-local potential  $V_{NL}$  is

$$V_{NL} = \sum_{ij} |\tilde{p}_i\rangle D_{ij} \langle\tilde{p}_j|, \quad (12)$$

where, the non-local coefficients  $D_{ij}$  are calculated at each self-consistent field iteration.

Since the KS wavefunctions  $|\Psi_{nk}\rangle$  are non-orthogonal, the following generalized eigenvalue equation is solved,

$$(\mathcal{H} - \epsilon_{nk} S) |\Psi_{nk}\rangle = 0, \quad (13)$$

where  $\epsilon_{nk}$  is the eigenvalue of band  $n$  at  $k$ -point  $k$ , and the overlap matrix  $S$  is defined as:

$$S = 1 + \sum_{ij} |\tilde{p}_j\rangle s_{ij} \langle\tilde{p}_i|, \quad (14)$$

with  $s_{ij} = \langle\phi_i|\phi_j\rangle - \langle\tilde{\phi}_i|\tilde{\phi}_j\rangle$ .

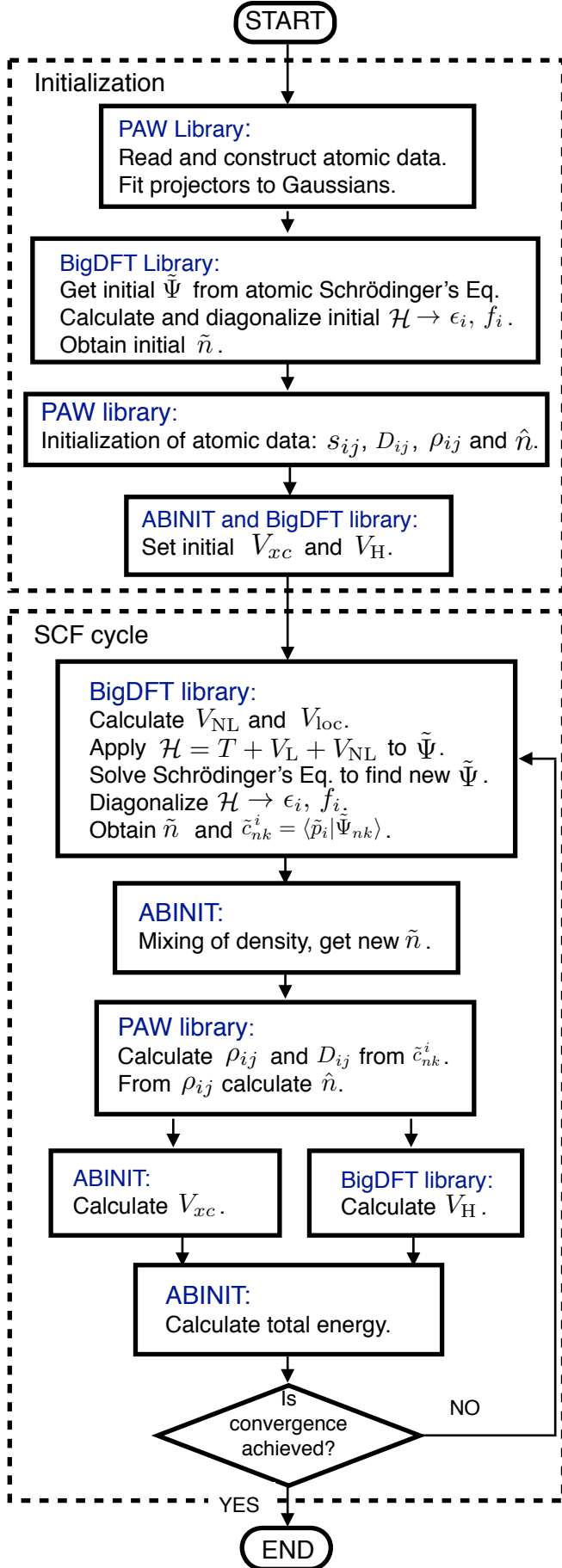


Figure 3: Flow-chart illustrating the the WVL-PAW method. The code/library used at each step in the Initialization or the Self-Consistent Field (SCF) cycle are indicated in blue fonts.

### 3. Methodology

In this Section, we describe the basic steps of our WVL-PAW method. The flowchart of this implementation is presented in Figure 3.

#### 3.1. Local ionic potential

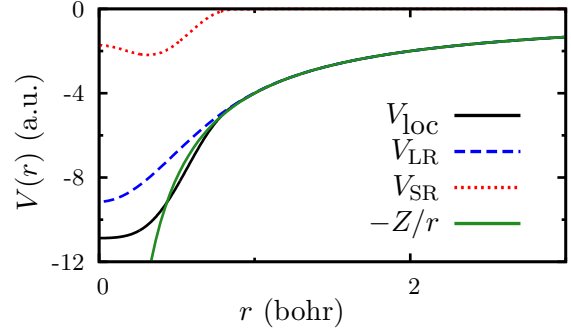


Figure 4: (Color online) Local ionic potential  $V_{\text{loc}}(r)$  (solid black lines) and its long-  $V_{\text{LR}}(r)$  (dashed blue lines) and short-range  $V_{\text{SR}}(r)$  (dotted red lines) components for Carbon with  $r_\ell = 0.35$  bohrs.

The local ionic potential  $V_{\text{loc}}^{(i)}$  for each atom indexed by  $i$  is calculated by the PAW atomic-data generator, and hence read by ABINIT in a radial grid. Aiming for linear-scaling, we circumvent the explicit evaluation of the local ionic potential which presents a slow decay, as shown in Figure 4. Instead, we calculate the potential locally by dividing it into long- (LR) and short-range (SR) components,

$$V_{\text{loc}}^{(i)} = V_{\text{SR}}^{(i)} + V_{\text{LR}}^{(i)}. \quad (15)$$

The LR part is chosen to be equal to the corresponding GTH-HGH pseudopotentials LR component,

$$V_{\text{LR}}^{(i)}(\mathbf{r}) = -\frac{Z_i}{r} \text{erf}\left(\frac{r}{\sqrt{2}r_\ell}\right), \quad (16)$$

being a smooth function with the correct decay of  $-Z_i/r$ , where  $Z$  is the atomic number (see Figure 4). The LR potential is associated to a given density  $n_{\text{LR}}$  via the Poisson equation  $\Delta V_{\text{LR}} = -4\pi n_{\text{LR}}$ , where

$$n_{\text{LR}}^{(i)}(\mathbf{r}) = -\frac{1}{(2\pi)^{3/2}} \frac{Z_i}{r_\ell^3} e^{-\frac{r^2}{2r_\ell^2}}, \quad (17)$$

and the radius  $r_\ell$  is chosen to be between 0.2 and 1.0 bohrs. Since the LR density is localized around the atoms, we can calculate locally the corresponding LR potential in a two step procedure; We first calculate  $n_{\text{LR}}$  at each atom, and afterwards, we find the corresponding potential in the entire the simulation box via the BigDFT Poisson-solver [19]. Moreover, the smooth shape of Eq. (17) is particularly convenient to achieve a fast convergence in the evaluation of  $V_{\text{loc}}$  with respect to the grid-spacing  $h$ .

From Eq. (15), the SR potential is given by

$$V_{\text{SR}}^{(i)} = V_{\text{loc}}^{(i)} - V_{\text{LR}}^{(i)}, \quad (18)$$

and can be easily evaluated locally at each atom. The LR part is analytic, see Eq. (16) and we calculate  $V_{\text{loc}}^{(i)}$  using splines. To avoid the divergence, the term  $V_{\text{LR}}^{(i)}(\mathbf{r} = 0)$  is evaluated using a quadratic interpolation of Eq. (16) from 3 points close to the origin. In summary, by dividing  $V_{\text{loc}}$  into LR and SR components, which can be evaluated locally at each atom, the calculation of  $V_{\text{loc}}$  scales linearly with the number of atoms  $N_{\text{at}}$ .

### 3.2. Non-local potential

We calculate the NL potential in Eq. (12),

$$V_{\text{NL}} = \sum_{ij} |\tilde{p}_i\rangle D_{ij} \langle \tilde{p}_j|,$$

by generalizing the method of BIGDFT to complex-Gaussian PAW pseudopotentials. The application of the Hamiltonian ( $\mathcal{H}|\Psi\rangle$ ) requires the calculation of wavefunction projections  $\tilde{c}_{nk}^i = \langle \tilde{p}_i | \Psi_{nk} \rangle$ . Both wavefunctions and projectors are expanded on the WVL basis in order to do this operation by simple scalar products, where the expansion coefficients for projectors are: [28]

$$\int \tilde{p}(\mathbf{r}) \phi_{i_1, i_2, i_3}(\mathbf{r}) d\mathbf{r} \quad \text{and} \quad \int \tilde{p}(\mathbf{r}) \psi_{i_1, i_2, i_3}^\nu(\mathbf{r}) d\mathbf{r}. \quad (19)$$

In BIGDFT, the integrals in Eq. (19) are simplified by using pseudopotentials with projectors of the form:

$$\tilde{p}(\mathbf{r}) = e^{-cr^2} x^{\ell_x} y^{\ell_y} z^{\ell_z}; \quad c \in \mathbb{R}. \quad (20)$$

Following the same spirit, PAW projectors are *fitted* to the general form:

$$\tilde{p}(\mathbf{r}) \approx x^{\ell_x} y^{\ell_y} z^{\ell_z} \sum_j^{N_g} a_j e^{b_j r^2}; \quad a_j, b_j \in \mathbb{C}, \quad (21)$$

where the number of complex Gaussian functions  $N_g$  is to be determined. This analytical form simplifies the 3D integrals in Eq. (19) into a sum of products of three 1D integrals.

$$\begin{aligned} \int \tilde{p}(\mathbf{r}) \phi_{i_1, i_2, i_3}(\mathbf{r}) d\mathbf{r} &= \sum_j^{N_g} W_{i_1}(a_j, b_j, \ell_x) \times \\ &W_{i_2}(a_j, b_j, \ell_y) \times W_{i_3}(a_j, b_j, \ell_z), \quad (22) \\ W_{i_k}(a, b, \ell) &= \int_{-\infty}^{+\infty} a e^{bt^2} t^\ell \phi(t/h - k) dt. \quad (23) \end{aligned}$$

The 1D integrals are calculated following the BIGDFT scheme, which is accurate to machine precision. Moreover, the fitting procedure introduced above is robust enough to obtain the desired accuracy in total energies, as explained more thoroughly in Section 4.

### 3.3. Direct minimization method

In a direct minimization approach, the gradient  $|g_i\rangle$  of the total energy with respect to the  $i$ th wavefunction  $|\Psi_i\rangle$  is defined in terms of the Hamiltonian  $\mathcal{H}$  and the overlap  $S$  operators:

$$|g_i\rangle = \mathcal{H}|\tilde{\Psi}_i\rangle - \sum_j \Lambda_{ij} S|\Psi_j\rangle, \quad (24)$$

$$\Lambda_{ij} = \langle \Psi_j | \mathcal{H} | \Psi_i \rangle \quad (25)$$

In BIGDFT, the preconditioned gradient  $|\tilde{g}\rangle$ , is found by solving the equation

$$\left(\frac{1}{2}\nabla^2 - \epsilon_i\right) \tilde{g}_i(\mathbf{r}) = g_i(\mathbf{r}). \quad (26)$$

The information coming from the wavefunction gradient is usually combined in the context of the *direct inversion in the iterative subspace (DIIS)*, method [25, 29] where the wavefunction  $i$  at iteration  $M + 1$  is found by suitable linear combination of the previous  $M$  trial functions  $|\Psi_i^j\rangle$  and  $|\tilde{g}^j\rangle$ :

$$|\Psi_i^{M+1}\rangle = |\bar{\Psi}^M\rangle + \lambda \bar{g}^M, \quad (27)$$

with

$$|\bar{\Psi}^M\rangle = \sum_{j=0}^M \alpha_j |\Psi_i^j\rangle, \quad |\bar{g}^M\rangle = \sum_{j=0}^M \alpha_j |\tilde{g}_i^j\rangle. \quad (28)$$

With the exceptions of the presence of the operator  $S$  in Eq. (24), all the wavefunction optimization flowchart is identical to the Norm-Conserving approach.

### 3.4. PAW atomic datasets

The first step in a PAW calculation is to read the atomic-data files (containing the description of atoms), or roughly speaking the "pseudopotentials", previously calculated. Numerous databases of atomic datasets are available and can be used within the present WVL-PAW implementation [30, 31, 32, 33].

For PAW, atomic-data files for WVLs and PWs are almost identical; In the PW case, projectors are given in a radial grid whereas in the WVL case projectors are expressed as a sum of complex Gaussian functions.

To support all atomic-data file formats, ABINIT contains now a converter from the conventional format (with projectors on a grid) to the Gaussian format. First projectors are fitted to a sum of Gaussian functions, as explained in 4. Later, the Gaussian coefficients are written into a new atomic-data file. This file can be used in further calculations to avoid running the fitting procedure at every time. From the user point of view, the above procedure is transparent.

#### 4. Fitting the PAW projectors

As previously discussed in Sec. 3.2, PAW non-local projectors are fitted to a sum of Gaussians with complex exponents in order to simplify scalar products with wavefunctions. Indeed, projectors expressed as linear combinations of Gaussian functions become analytical, and hence can be easily converted to WVL space. Moreover, since both projectors and Daubechies WVLS are separable in the three Cartesian directions, internal products  $\langle \tilde{p}_j | \phi_i \rangle$  are reduced to 1D operations.

In this work, we use the Levenberg-Marquardt formalism [34] to fit the projectors to a Gaussian form. A given set of points  $(x_i, \tilde{p}_i)$  are fitted to a given function  $f(x_i, \alpha)$ , where  $\alpha$  is a vector of independent and/or dependent parameters. As in other regression methods, the sum of the squares of deviations  $S(\alpha)$  is minimized, where

$$S(\alpha) = \sum_{i=1}^{N_i} [\tilde{p}_i - f(x_i, \alpha)]^2. \quad (29)$$

The radial part of the projectors  $\tilde{p}(r)$  is fitted to the analytical expression in Eq. (21):

$$\tilde{p}(r) \approx x^{\ell_x} y^{\ell_y} z^{\ell_z} \sum_j^{N_g} a_j e^{b_j r^2}; \quad a_j, b_j \in \mathbb{C}.$$

As other minimization techniques, the fitting algorithm is quite sensitive to the initial-guess. Moreover, a function which decays to zero (such as PAW projectors) is generally difficult to fit to periodic functions (sinus and cosinus functions). Therefore, we use a real Gaussian function as an envelope of a sum of sinus and cosinus functions, as follows,

$$\tilde{p}(r) \approx \sum_j^{N_g} a_{1,j} e^{-a_{2,j} x^2} (a_{3,j} \sin(k_j x^2) a_{4,j} \cos(k_j x^2)); \quad (30)$$

$$a_1, a_2, a_3, a_4 \in \mathbb{R}$$

To obtain a small  $S$ , the envelope function is constrained to decay almost to zero at the PAW radius. Further, the  $k^j$  are fixed to an exponential series (*i.e.*,  $k = 1.1^j$ ) to enhance the sinus & cosinus basis completeness, and hence avoiding duplicate basis set elements.

The fitting procedure is overall satisfactory. As expected, the accuracy can be improved by increasing the number of complex Gaussian functions  $N_g$ . For instance, in Figure 5 the 1s NL-projector for Hydrogen is fitted to a sum of complex Gaussians. For simplicity, spin-polarization is not taken into account. As expected, for larger  $N_g$  smaller  $S(\alpha)$  may be obtained, hence, the error due to fitting of  $\tilde{p}$  in total energy calculations decreases, reaching the same total energy, up to machine-precision, than the one obtained with a plane-waves-PAW calculation. In particular, the number of Gaussian functions  $N_g$  is a convergence parameter in the simulation, as exemplified in the next Section.

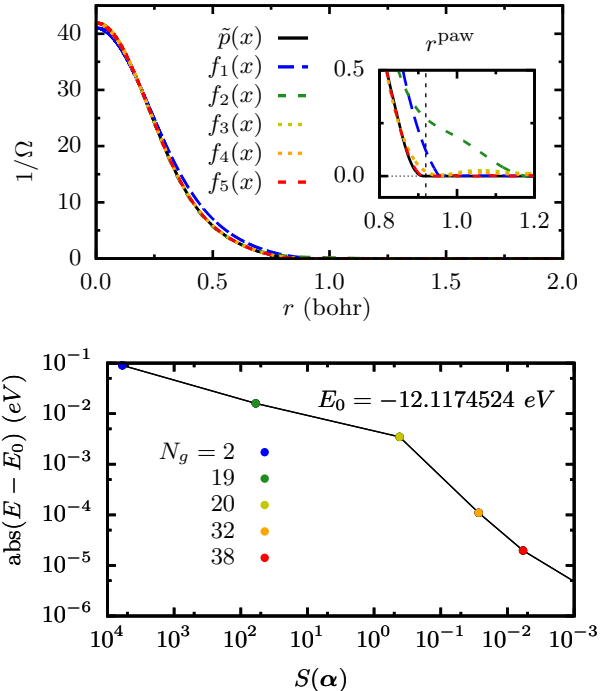


Figure 5: (Color online) The 1s Hydrogen NL PAW projector  $[\tilde{p}(x)]$  is fitted to a sum of complex Gaussians. The number of terms  $N_g$  in Eq. (30) is increased to achieve a better fit [a smaller  $S(\alpha)$ ]. **Top:** The original function  $[\tilde{p}(x)]$  is shown in solid-black lines on a radial grid. The fitted functions using  $N_g$  equal to 2 [ $f_1(x)$ ] [dashed-blue line], 19 [ $f_2(x)$ ] [dashed-green line], 20 [ $f_3(x)$ ] [dashed-yellow line], 32 [ $f_4(x)$ ] [dashed-orange line] and 38 [ $f_5(x)$ ] [dashed-red line] are also shown. In the inset, the area close to the PAW radius,  $r^{\text{paw}}$  is zoomed-in. **Bottom:** Error in total energies (in hartrees) with respect to a plane-waves-PAW calculation for Hydrogen in a box. The results for several  $N_g$  values corresponding to different  $S(\alpha)$  are shown. Increasing the number of Gaussian functions improves the accuracy in the calculated total energy, as expected.

#### 5. Numerical tests

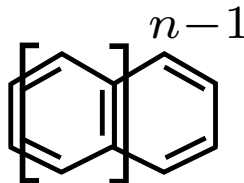


Figure 6: The acenes general formula.

In this Section, we show tests on simple systems to elucidate the performance of our implementation and the effect of the Gaussian expansion of the PAW non-local projectors. In particular, we present calculations of total-energy differences in finite systems, since wavelets are particularly adapted for charged species in non-periodic boundary conditions.

We first show that our method provides results in agreement with previous and well-established DFT codes, for the sake of validation. With this aim, we compare total-

Molecule	Ionization potential (eV)	
	PAW	HGH
Benzene	13.05	12.99
Naphthalene	10.87	10.81
Anthracene	9.55	9.57
Tetracene	8.76	8.77

Table 1: Ionization potential (IP) of the oligoacenes calculated with energy differences of the neutral and charged species within the LDA. Several methods are considered: our WVl-PAW implementation and HGH PPs with WVl basis, for comparison.

energy differences calculated with our WVl-PAW implementation and with HGH PPs with WVl basis. Here, we study the acene family of molecules: from benzene to hexacene, as a prototype finite system. The general formula is shown in Figure 6. In Table 1, we show the ionization potential (IP) of the acenes calculated as

$$\text{IP} = \frac{1}{2} (E(N+2) - E(N)), \quad (31)$$

and neglecting spin as a first-order approximation. The IPs calculated with our WVl-PAW implementation are within 0.05 eV of those calculated with HGH PPs. As a further check, the resulting IPs were verified with the former PW-PAW implementation in ABINIT.

Having verified the accuracy of our method, we now simulate a relatively more complex and inhomogeneous case consisting of a C60 molecule doped with a Ti atom at its center. In Figure 7, we show the convergence of the calculated ionization potential energy (IP) with respect to the number of complex Gaussian projectors. As before, the IP is calculated from total energy differences with Eq. 31. Interestingly, a relatively small number of  $N_g = 10$  terms is needed to converge the IP energy to 0.1 eV. This proves that the Gaussian fit of the PAW projectors might provide an interesting strategy to express in analytic form a  $\mathcal{T}_R$  operator. Work is in progress to identify the better strategy to reduce the computational overhead in the case when the fit needs a large (e.g. bigger than 40) number of Gaussians.

In summary, we showed that our WVl-PAW implementation is accurate and efficient and is appropriate to simulating complex in-homogeneous systems in non-periodic boundary conditions.

## 6. The PAW library

While implementing the WVl-PAW code, the core PAW routines of ABINIT were assembled, modularized, and disentangled from the PWs part of the code. The outcome of this work is a basis-independent, modular and stand-alone *PAW library* written in Fortran 2003. It is worth noticing that the XC potential is calculated by using ABINIT or the LIBXC package [35]. The library is now stable and ready to be used in other codes.

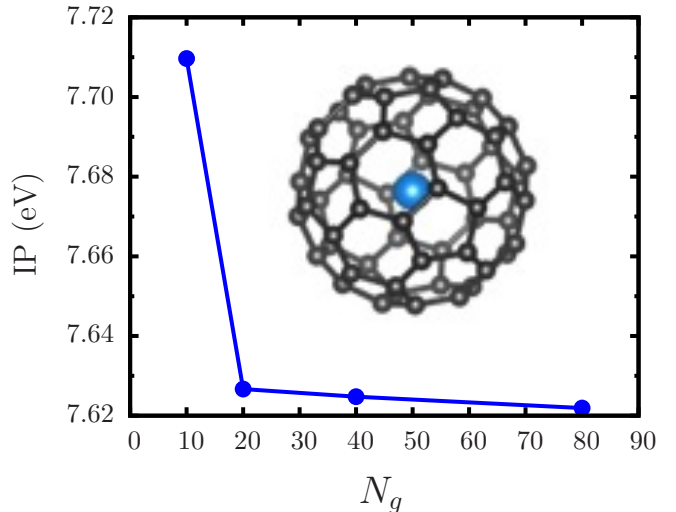


Figure 7: Endohedral doped fullerene with a Ti atom enclosed at its center; carbon and Ti atoms are shown in gray and blue spheres, respectively. We show the calculated ionization potential (IP) energy with respect to the number of complex Gaussian projectors ( $N_g$  terms).

We emphasize that the PAW library is basis independent, and hence it can be easily ported to another basis, as proven by our implementation of PAW in BigDFT. The only part that remains dependent on the given basis set are the projections of the pseudo-wavefunctions onto the non-local projectors  $\langle \tilde{\Psi} | \tilde{p}_i \rangle$ . From the knowledge of these projections, all quantities involved in the PAW formalism can be computed: occupancy matrix, on-site densities, self-consistent contribution to the non-local operator, etc. And this is exactly the purpose of the PAW library. For more details, please consult the ABINIT documentation, or the code sources, all modules are self-explanatory.

In the next, the basic modules of the PAW library are briefly described.

Modules related to basic datatypes:

- **m\_pawrad**: contains all functions related to the PAW radial meshes, such as datatypes and derivation/integration routines for the different kinds of meshes (linear or logarithmic).
- **m\_pawtab**: used to define tabulated PAW data which is read from atomic data files, such as core- and local-potentials and PAW projectors, etc.
- **m\_pawang**: contains definitions and functions related to the angular mesh discretization of the PAW spheres.
- **m\_paw\_ij**: contains objects expressed in the on-site partial-wave basis, e.g., the calculated  $D_{ij}$  terms.
- **m\_paw\_an**: on-site potentials and densities are defined here. The potentials/densities are stored in terms of angular mesh or angular moments for each atom.

Modules related to high-level objects:

- `m_pawpsp`: module to read PAW atomic data files. Used to read and calculate data inside the augmentation spheres, such as potentials and its derivatives.
- `m_pawcprj`: calculates, stores and manipulates  $\tilde{c}_{nk}^i$ . Here we include routines to compute the on-site contribution to the overlap between two states and MPI communication routines dealing with  $\tilde{c}_{nk}^i$  objects.
- `m_pawrhoij`: computes and symmetrizes occupancy matrix  $\rho_{ij} = \sum_{nk} f_{nk}(\tilde{c}_{nk}^i)^* \tilde{c}_{nk}^j$ . This module also contains MPI routines to send/distribute/gather the  $\rho_{ij}$  occupations.
- `m_pawdij`: calculates all contributions to the  $D_{ij}$  non-local pseudopotential terms. This module also computes local Hubbard-U and local exact-exchange contributions to the PAW onsite-potentials.

Low-level modules:

- `m_paw_gaussfit`: contains routines used to fit the numeric NL projectors to sums of complex Gaussians, see 4.
- `m_pawxc`: computes the exchange-correlation potential/energy in the augmentation regions using developments over spherical harmonics, see Ref. [23].
- `m_paw_finegrid`: contains all routines performing operations (*e.g.*, integrations) on the fine-grid around each atom.

## 7. Conclusions

In this work we presented a PAW method in a WVL basis set. In order to take advantage of WVL properties, we modified the PAW non-local operator to a sum of Gaussians. Hence, the application of the non-local part of the Hamiltonian is performed analytically, reducing computational costs. Our WVL-PAW method was implemented in ABINIT using the BIGDFT library and by creating a PAW library the method was ported into BIGDFT in stand-alone mode. In addition, the PAW library is stable and can be used in other codes. In summary, our new method presents the PAW frozen-core AE accuracy and the WVL adaptability, locality and systematic convergence. This opens up the possibility to treat large heterogeneous systems and large molecules within PAW and of potentially having an order- $N$  code in PAW.

## 8. Acknowledgments

T. Rangel thanks Jean-Michel Beuken for his technical support on ABINIT and Alessandro Mirone for his help and discussions on the Gaussian fitting procedure. This work was supported by the ANR NEWCASTLE project, grant ANR- 2010-COSI-005-01 of the French National Research Agency.

## Appendix A. The BigDFT Poisson solver for planewave based-PAW calculations

Finite, isolated or low-dimensional systems are often simulated using supercells in periodic boundary conditions (BC), such as in plane-wave approaches. In order to eliminate spurious electrostatic interactions due to neighboring cells, numerous methods are well established and have been documented in Refs. [36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52]. Among these approaches, the BIGDFT Poisson solver [45, 49] has proven successful in treating isolated- or surface-BC by solving the Hartree Potential in the correct BCs.

In this work, we also add the capability of treating reduced BCs to the plane-wave PAW implementation of ABINIT, by generalizing to PAW the BIGDFT Poisson solver. Here, all potentials and densities are calculated in real-space, as done within WVLs. In particular, the local-ionic potential is calculated as explained in Section 3.1 and the Hartree potential is obtained by using the BIGDFT Poisson solver. This approach can be accessed by setting the user-variable `icoulomb` to 1 or 2 for free- or surface-BCs, respectively.

To illustrate advantages of this implementation, we show in Figure A.8 the convergence with respect to the lateral unit-cell size ( $a$ ) for Na in a squared-box. The resulting total energies are  $-1304.75$  and  $-1299.59$  eV for the neutral and charged systems, respectively. For the reciprocal space approach (`ic=0`), these values are obtained after fitting the total energies to an infinite cell-size ( $a \rightarrow \infty$ ), using the Makov-Payne method [37]. As expected, convergence is achieved much faster using the WVL Poisson solver (`ic=1`), since spurious interactions between neighboring cells are eliminated.

## References

- [1] L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman, R. Schneider, Daubechies wavelets as a basis set for density functional pseudopotential calculations, *J. Chem. Phys.* 129 (2008) 014109.
- [2] X. Gonze, B. Amadon, P. Anglade, J. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, T. Deutsch, L. Genovese, P. Ghosez, M. Giantomassi, S. Goedecker, D. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M. Verstraete, G. Zerah, J. Zwanziger, ABINIT: first-principles approach to material and nanosystem properties, *Comput. Phys. Commun.* 180 (12) (2009) 2582–2615.
- [3] I. Daubechies, 10 Lectures on Wavelets, Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1992.
- [4] P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* 50 (24) (1994) 17953–17979.
- [5] The CP-PAW code, <https://www2.pt.tu-clausthal.de/paw>.
- [6] Plane-wave self-consistent field (PWSCF) in quantum espresso package, <http://www.quantum-espresso.org> (2007).
- [7] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* 59 (3) (1999) 1758.



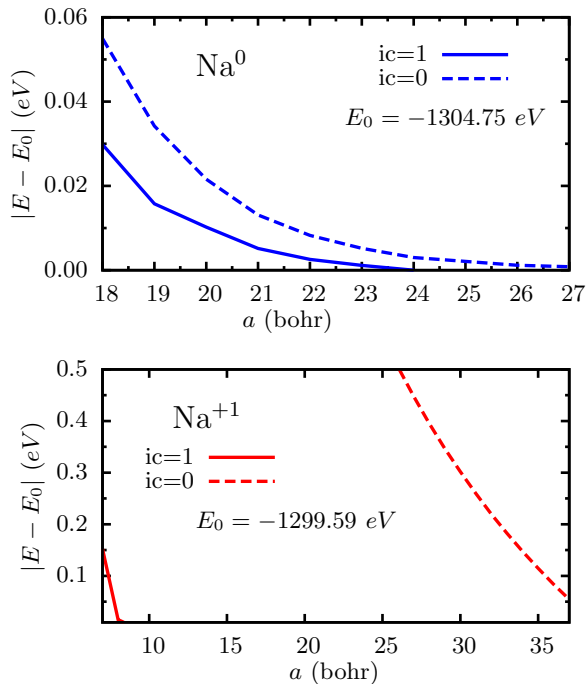


Figure A.8: (Color online) Convergence on the lateral cell-size ( $a$ ) for Na in a box. The neutral (top) and charged (bottom) systems are considered. The results with- ( $ic=1$ ) and without- ( $ic=0$ ) the WVL Poisson solver are shown. It is evident that using the WVL Poisson Solver ( $ic=0$ ) drastically improves the convergence of charged species with respect to the cell-size.

- [8] The Socorro code, <http://dft.sandia.gov/socorro>.
- [9] A. R. Tackett, N. A. W. Holzwarth, G. E. Matthews, A projector augmented wave (PAW) code for electronic structure calculations, Part II: pwpaw for periodic solids in a plane wave basis, *Comput. Phys. Commun.* 135 (2001) 348.
- [10] W. Kromen, The projector augmented wave method: A fast all-electron procedure for the ab-initio molecular dynamics, Ph.D. thesis, Forschungszentrum Juelich GmbH. Inst. fuer Festkoerperforschung; Technische Hochschule Aachen (2001).
- [11] J. J. Mortensen, L. B. Hansen, K. W. Jacobsen, Real-space grid implementation of the projector augmented wave method, *Phys. Rev. B* 71 (3) (2005) 035109.
- [12] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, M. C. Payne, Introducing ONETEP: linear-scaling density functional simulations on parallel computers, *J. Chem. Phys.* 122 (8) (2005) 084119.
- [13] Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method, *J. Phys.: Condens. Matter* 22 (25) (2010) 253202. doi:10.1088/0953-8984/22/25/253202.
- [14] S. Mohr, L. E. Ratcliff, P. Boulanger, L. Genovese, D. Caliste, T. Deutsch, S. t. Goedecker, Daubechies wavelets for linear scaling density functional theory, *J. Chem. Phys.* 140 (20) (2014) 204110. doi:10.1063/1.4871876.
- [15] S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker, T. i. Deutsch, Accurate and efficient linear scaling DFT calculations with universal applicability, *Phys. Chem. Chem. Phys.* 17 (47) (2015) 31360–31370. doi:10.1039/C5CP00437C.
- [16] G. Beylkin, On the representation of operators in bases of compactly supported wavelets, *SIAM Journal on Numerical Analysis* 29 (6) (1992) 1716.
- [17] A. Neelov, S. Goedecker, An efficient numerical quadrature for the calculation of the potential energy of wavefunctions expressed in the daubechies wavelet basis, *J. Comput. Phys.* 217 (2) (2006) 312–339.
- [18] T. Deutsch, L. Genovese, Wavelets for electronic structure calculations, *Collection SFN 12* (2011) 33. doi:10.1051/sfn/201112004.
- [19] A. Cerioni, L. Genovese, A. Mirone, V. A. Sole, Efficient and accurate solver of the three-dimensional screened and unscreened poisson's equation with generic boundary conditions, *J. Chem. Phys.* 137 (13) (2012) 134108. doi:http://dx.doi.org/10.1063/1.4755349.
- [20] L. Genovese, T. Deutsch, Multipole-preserving quadratures for the discretization of functions in real-space electronic structure calculations, *Phys. Chem. Chem. Phys.* 17 (47) (2015) 31582. doi:10.1039/C5CP01236H.
- [21] S. Goedecker, M. Teter, J. Hutter, Separable dual-space gaussian pseudopotentials, *Phys. Rev. B* 54 (3) (1996) 1703–1710.
- [22] C. Hartwigsen, S. Goedecker, J. Hutter, Relativistic separable dual-space Gaussian pseudopotentials from H to Rn, *Phys. Rev. B* 58 (7) (1998) 3641.
- [23] M. Torrent, F. Jollet, F. Bottin, G. Zérah, X. Gonze, Implementation of the projector augmented-wave method in the ABINIT code: Application to the study of iron under pressure, *Comput. Mater. Sci.* 42 (2) (2008) 337–351.
- [24] M. Torrent, N. Holzwarth, F. Jollet, D. Harris, N. Lepley, X. Xu, Electronic structure packages: Two implementations of the projector augmented wave (PAW) formalism, *Comput. Phys. Commun.* 181 (11) (2010) 1862 – 1867.
- [25] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* 54 (16) (1996) 11169–11186.
- [26] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* 59 (1999) 1758–1775.
- [27] D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism, *Phys. Rev. B* 41 (11) (1990) 7892–7895.
- [28] S. Goedecker, Wavelets and their application for the solution of partial differential equations, *Presses Polytechniques Universitaires et Romandes, Lausanne, Switzerland*, 1998.
- [29] P. Péter, Convergence acceleration of iterative sequences. the case of scf iteration, *Chem. Phys. Lett.* 73 (2) (1980) 393–398.
- [30] F. Jollet, M. Torrent, N. Holzwarth, Generation of Projector Augmented-Wave atomic data: A 71 element validated table in the XML format, *Comput. Phys. Commun.* 185 (4) (2014) 1246–1254.
- [31] XML GPAW format for atomic datasets. URL <https://wiki.fysik.dtu.dk>
- [32] K. F. Garrity, J. W. Bennett, K. M. Rabe, D. Vanderbilt, Pseudopotentials for high-throughput DFT calculations, *Comput. Mater. Sci.* 81 (2014) 446–452.
- [33] PWPaw atomic dataset format. URL <http://users.wfu.edu/natalie/papers/pwpaw>
- [34] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, *Quart. Appl. Math.* (2) (1944) 164.
- [35] M. A. L. Marques, M. J. T. Oliveira, T. Burnus, Libxc: A library of exchange and correlation functionals for density functional theory, *Comput. Phys. Commun.* 183 (10) (2012) 2272–2281.
- [36] M. Leslie, N. J. Gillan, The energy and elastic dipole tensor of defects in ionic crystals calculated by the supercell method, *J. Phys. C: Solid State Phys.* 18 (5) (1985) 973. doi:10.1088/0022-3719/18/5/005.
- [37] G. Makov, M. C. Payne, Periodic boundary conditions in ab initio calculations, *Phys. Rev. B* 51 (7) (1995) 4014.
- [38] M. R. Jarvis, I. D. White, R. W. Godby, M. C. Payne, Supercell technique for total-energy calculations of finite charged and polar systems, *Phys. Rev. B* 56 (23) (1997) 14972–14978. doi:10.1103/PhysRevB.56.14972.
- [39] L. N. Kantorovich, Elimination of the long-range dipole interaction in calculations with periodic boundary conditions, *Phys. Rev. B* 60 (23) (1999) 15476–15479. doi:10.1103/PhysRevB.60.15476.

- [40] L. Bengtsson, Dipole correction for surface supercell calculations, *Phys. Rev. B* 59 (19) (1999) 12301–12304. doi:10.1103/PhysRevB.59.12301.
- [41] G. J. Martyna, M. E. Tuckerman, A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters, *J. Chem. Phys.* 110 (6) (1999) 2810–2821. doi:10.1063/1.477923.
- [42] P. A. Schultz, Charged Local Defects in Extended Systems, *Phys. Rev. Lett.* 84 (9) (2000) 1942–1945. doi:10.1103/PhysRevLett.84.1942.
- [43] H. Nozaki, S. Itoh, Energy correction for isolated impurities under periodic boundary conditions, *Phys. Rev. E* 62 (1) (2000) 1390–1396. doi:10.1103/PhysRevE.62.1390.
- [44] A. Castro, A. Rubio, M. J. Stott, Solution of Poisson’s equation for finite systems using plane-wave methods, *Can. J. Phys.* 81 (10) (2003) 1151–1164. doi:10.1139/p03-078.
- [45] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, G. Beylkin, Efficient solution of Poisson’s equation with free boundary conditions, *J. Chem. Phys.* 125 (7) (2006) 074105.
- [46] A. F. Wright, N. A. Modine, Comparison of two methods for circumventing the Coulomb divergence in supercell calculations for charged point defects, *Phys. Rev. B* 74 (23) (2006) 235209. doi:10.1103/PhysRevB.74.235209.
- [47] S. Ismail-Beigi, Truncation of periodic image interactions for confined systems, *Phys. Rev. B* 73 (23) (2006) 233103. doi:10.1103/PhysRevB.73.233103.
- [48] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, A. Rubio, Exact Coulomb cutoff technique for supercell calculations, *Phys. Rev. B* 73 (20) (2006) 205119. doi:10.1103/PhysRevB.73.205119.
- [49] L. Genovese, T. Deutsch, S. Goedecker, Efficient and accurate three-dimensional Poisson solver for surface problems, *J. Chem. Phys.* 127 (5) (2007) 054704.
- [50] L. Yu, V. Ranjan, W. Lu, J. Bernholc, M. B. Nardelli, Equivalence of dipole correction and Coulomb cutoff techniques in supercell calculations, *Phys. Rev. B* 77 (24) (2008) 245102. doi:10.1103/PhysRevB.77.245102.
- [51] I. Dabo, B. Kozinsky, N. E. Singh-Miller, N. Marzari, Electrostatics in periodic boundary conditions and real-space corrections, *Phys. Rev. B* 77 (11) (2008) 115139. doi:10.1103/PhysRevB.77.115139.
- [52] N. D. M. Hine, J. Dziedzic, P. D. Haynes, C.-K. Skylaris, Electrostatic interactions in finite systems treated with periodic boundary conditions: Application to linear-scaling density functional theory, *J. Chem. Phys.* 135 (20) (2011) 204103. doi:10.1063/1.3662863.