CIF2WAN: A Tool to Generate Input Files for Electronic Structure Calculations with Wannier90

Christopher Sims

Department of Physics, University of Central Florida, Orlando, Florida 32816, USA

The generation of input files for density functional theory (DFT) programs must often be manually done by researchers. If one wishes to produce a maximally localized wannier functions (MLWFs) the calculation consists of several separate files that must be formatted correctly in order for the program to work properly. Many of the inputs are repeated throughout the files and can be easily automated. In this work, a program is presented to generate all of the input files needed to produce wannier functions with Wannier90 starting from open source DFT programs such as Quantum espresso, Abinit, and Siesta. In addition, the input files for WannierTools are also included for those that wish to produce surface green’s functions for the generation of surface state bands. The program presented allows for users new to DFT to use the programs with minimal understanding of parameters needed to produce good results, in addition, this program allows for researchers who are advanced DFT users to utilize this program for high throughput wannier calculations.

I. PROGRAM SUMMARY

Program title: CIF2WAN
Licensing provisions: GNU General Public Licence 3.0
Program obtainable from: Programming language: Python
Has the code been vectorised or parallelized?: no
Computer: Any computer that can run Python v3.6+
Operating system: Any operating system that can run Python v3.6+
external libraries: numpy, pymatgen (w/ registration), glob, shutil, csv
Running time: less than a minute (DFT runs are separate)

II. INTRODUCTION

Density functional theory (DFT) is a powerful tool that is widely used for calculations in solid-state physics such as electronic structure predictions, lattice relaxation calculations, magnetism, etc. Until the late 1990’s, DFT was considered to be too computationally expensive for such calculations\(^1\). However, with increasing technology and computational methods DFT is now considered to be accurate enough for quantum chemistry. There are a plethora of DFT packages for quantum chemistry such as VASP\(^5\), ABINIT\(^6\), Quantum ESPRESSO\(^7\), SIESTA\(^8\), and WIEN2K\(^9\) being the most popular. For real space electronic structure calculations, it is necessary to extract the real space maximally localized wannier functions (MLWFs)\(^10\) in order to calculate the surface band structure\(^11\), perform wannier charge center calculations, etc. There are two main programs that are capable of calculating MLWFs one being Wannier Transport (wanT)\(^12\) and the other, more popular being Wannier90\(^13\). With these programs one is able to calculate surface band structures utilizing the iterative surface green function matching technique such as the method implemented in Wannier90\(^14\) or chinook\(^15\). The main issue with DFT, especially if one wishes to conduct high throughput calculation is that one must meticulously format the input files for DFT programs so that there is no error in the calculation. Small errors cannot bypass the error detection built into DFT programs and learning the best methods for input will take some time for a beginner. Rather, having programs that assist in the pre-processing stage are more favorable for \textit{ab initio} DFT theorists. While tools such as CIF2CELL\(^16\), Aiida\(^17\), pymatgen\(^18\), jarvis\(^22-27\), etc. exist, none can generate all of the input files needed to generate the wannier tight binding (TB) Hamiltonian utilizing multiple different quantum chemistry programs.

Recently, there has been a rising interest in identifying materials by their surface states. This began with the discovery of a 3D topological insulator which was confirmed by angle-resolved photoemission spectroscopy and density functional theory\(^28-30\). Later, nodal-line and Weyl semimetal states have been discovered in real crystal systems such as ZrSiS\(^31-34\) and TaAs\(^35-37\), respectfully. All of these materials are confirmed to be topologically non-trivial via berry phase, wannier charge center, or Wilson loop analysis. All of these methods can be performed with wannier functions. However, in order to properly perform calculations of a large amount of materials that can be topologically non-trivial will require many input files to various DFT programs. The automation of these tasks will make these calculations much simpler to perform.

This work presents CIF2WAN, a python program that can generate the input files in order to perform wannier calculations from \textit{ab initio} calculations. In addition, slurm job handling files which are common in most supercomputer clusters are also generated for the user. The goal of this program is to make DFT calculations that utilize wannier TB Hamiltonians easier to understand for beginners and easier for more advance users of DFT programs.

\(^{a)}\text{Christophersims@knights.ucf.edu}\)

arXiv:2006.12647v1 [physics.comp-ph] 22 Jun 2020
III. METHODS

Wannier functions can be seen as a derivation of the Bloch function with an associated phase phase factor $e^{-i \mathbf{k} \cdot \mathbf{R}_n}$ is the real space lattice vector. In addition, one can label each band with its associated real space lattice vector $\mathbf{R}_n$ where $n$ is the band index. With this we can construct the wannier functions for each band index $n$. \(^1\)

$$|\mathbf{R}_n\rangle = \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k} e^{-i \mathbf{k} \cdot \mathbf{R}_n} |\Psi_{nk}\rangle$$  \hspace{1cm} (1)

In order to compute wannier functions one must introduce a unitary mixing parameter $U^k_{mn}$ to make the Hamiltonian smooth in all of k-space

$$|\mathbf{R}_n\rangle = \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k} e^{-i \mathbf{k} \cdot \mathbf{R}_n} \sum_{m=1}^J U^k_{mn} |\Psi_{nk}\rangle$$  \hspace{1cm} (2)

A common way to compute the wannier functions is via the projection method. Starting from a set of trial projections ($J$), These trial functions are projected onto the Bloch manifold.

$$|\psi_{nk}\rangle = \sum_{m=1}^J |\Psi_{mk}\rangle \langle \Psi_{mk} | g_n\rangle$$  \hspace{1cm} (3)

In order to do this calculation the matrix of inner products must first be computed. ($A_k)_{mn} = \langle \Psi_{mk} | g_n\rangle$. By substituting this into equation 3, one can construct the trial wannier functions that are related to the real Bloch functions.

$$|\tilde{\psi}_{nk}\rangle = \sum_{m=1}^J |\Psi_{mk}\rangle \langle S_k^{-1/2} \rangle_{mn}$$  \hspace{1cm} (4)

Using these equations one can then minimize the localization function $\Omega$

$$\Omega = \sum_{n} \left[ (|0_n|^2 |0_m\rangle - \langle 0_n | r_0 | 0_n \rangle)^2 \right] = \sum_{n} [r_n^2 - \bar{r}_n^2]$$  \hspace{1cm} (5)

MLWFs can be derived by minimizing this function for each k point in the lattice with respect to $r_{mn}$ after the bands have been calculated during the self consistent step of a DFT calculation. This procedure is repeated until $\Delta \Omega$ is sufficient small. \(^13\)

IV. FEATURES OF CIF2WAN

CIF2WAN generates all of the input files needed to start from ab initio calculations in ABINIT, Quantum ESPRESSO, etc. to the generation and use of the wannier tight binding Hamiltonian (seedname.inr.dat) used in wannier tools. The program can load these files utilizing either a CIF file or by directly interfacing with materials project via pymatgen. The output is as follows for the following DFT programs.

V. INSTALLATION AND USAGE

In this section, we present how to install and use CIF2WAN.
A. Get CIF2WAN

CIF2WAN is an open source software package distributed under the GNU General Public license 3.0 (GPL). The code can be downloaded directly from the public code repository: https://github.com/ChristopherSims/CIF2WAN.

B. Installation and running the code

The code requires no installation, in the future this package may be developed into a python package. The version of python must be 3.6 or higher, in addition the user must install the following python packages: pymatgen, matplotlib, numpy, string, shutil, csv. These programs will likely automatically install in most modern python CDEs. In addition, one must obtain an APIKEY for usage of pymatgen (only if one uses the PYMATGEN flag), this key can be obtained by registering an account at materialsproject.org

C. File formats

File outputs are as described in section IV. The input file must be manually edited by the user in v1.0. However, a GUI may be introduced in later version.

D. Input file

```python
#INPUT FILE input.py
ncore = 50 # number of cores for calculation
USE = "PYMATGEN" # PYMATGEN OR CIF
DFT = "ESPRESSO" #ESPRESSO, SIESTA, ABINIT, VASP
APIKEY = "YOURAPIKEY" #APIKEY for mat
MATERIALID = "mp-22866" # input for pymatgen
SEEDNAME = "HB" #yours or default
SOC = True # TRUE or FALSE
KMESH = [8,8,8] # KMESH for nscf/wan
NUMBANDS = 120 # Number of bands
MAGNETISM = False
```

VI. WANNIER90 CONVERGENCE

Although all of the bands and energies are converged in the scf and nscf cycles there is the main issue of convergence with MLWFs that will cause most calculations to appear to not converge. A small error in the wannier minimization can lead to spurious bands that will lead to a non-real Fermi surface when conducting calculations with the results of that wannier90 run. There are many ways to correct this, such as iterating over many steps (>50,000), defining a good guess for the initial projection, changing the frozen window for disentanglement, etc. However, the main issue is two three fold, the initial projections (guess) must be manually defined or be made random, there should be an internal program that makes good projection guesses from the known valence states and the result of the nscf run. Secondly, the steepest decent algorithm implemented in wannier90 v3.1 is not the best algorithm in order to get the proper minimization of the wannier centers. In fact, this method often gets stuck in local minimum during the calculation. Several better alternatives to the algorithm have been proposed but have not been implemented in the main wannier90 code. Finally, many of the interface codes to wannier90 were not written by the original developers of the respective DFT functions, this means that they could have many bugs or not correctly calculate the .mmn and .amn files needed for wannier90. This creates of problem of there being errors in wannier90 runs that must be corrected manually by users or are so fatal that wannier90 cannot parse to input file correctly at all. detracting from the possibility of "high throughput" wannierization runs.

VII. EXAMPLES

In this section we present calculations utilizing the files that are generated by the programs. The only changes are made to the wannier90 input files in order to converge the MLWFs

A. 3D topological insulator Bi$_2$Se$_3$

Bi$_2$Se$_3$ has been discovered to be a 3D topological insulator with non-trivial surface states and an insulating bulk. Here we show a calculation utilizing quantum ESPRESSO with spin orbit coupling (SOC) on. Here we see the large Dirac cone at the Γ-point [Fig 1(A)]. In addition, we show the Fermi surface of the Dirac cone inside the band gap showing the topologically nontrivial surface states[Fig 1(B)].

![Bi$_2$Se$_3$](image)

**FIG. 1.** Bi$_2$Se$_3$: (A) Surface states of the (001) surface (B) Fermi surface of Bi$_2$Se$_3$

B. Nodal line semimetal HfP$_2$

HfP$_2$ is a nodal line semimetal where the non-trivial state lies above the Fermi level in the presence of neglig-
ble spin-orbit coupling, although the nodal-line is above the Fermi surface non-trivial surface states that originate from the nodal line still exist throughout the band structure. The nodal-line is located about 0.2 eV above the Fermi level along the Γ-X line [Fig 2(A)]. We also show the Fermi surface of HfP$_2$ [Fig 2(B)].

C. Weyl Semimetal TaAs

TaAs is the first material to be discovered to have a Weyl semimetal state. Weyl fermions materialize in condensed matter systems as chiral edge modes which are connected by Fermi arcs.

VIII. CONCLUSION

In conclusion we have develop a program that can generate the input files needed to conduct calculations all the way for generating the wannier TB Hamiltonian and utilizing the result in wanniertools. Our program is designed to be easy to use at the expense of complexity and being able to input several settings into DFT programs. Currently, CIF2WAN can interface with quantum ESPRESSO, ABINIT, VASP, and SIESTA, with the possibility of other quantum chemistry programs being added. In the future, we intend to add a graphical user interface once the program has been well tested.

IX. ACKNOWLEDGMENTS

The authors acknowledge the University of Central Florida Advanced Research Computing Center for providing computational resources and support that have contributed to results reported herein. URL: https://arcc.ist.ucf.edu.

Correspondence should be addressed to C.S (Email: Christophersims@knights.ucf.edu).

1W. Kohn and L. J. Sham, Physical Review 140, A1133 (1965).
2R. Jones, Reviews of Modern Physics 87, 897 (2015).
3G. Kresse and J. Hafner, Phys. Rev. B 48, 13115 (1993).
4J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996).
5X. Gonze, B. Amadon, P.-M. Anglade, J.-M. 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. Jamond, S. Leroux, M. Mancini, S. Mazevet, M. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G.-M. Rignanese, D. Sangalli, R. Shaltuf, M. Torrent, M. Verstraete, G. Zerah, and J. Zwanziger, Computer Physics Communications 180, 2582 (2009).
6X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, P. Ghosez, J.-Y. Raty, and D. Allan, Comptes Rendus Physique 25, 478 (2002).
7P. Giannozzi, O. Andreussi, T. Brumme, O. Bunka, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Kjolbien, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncè, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, Journal of Physics: Condensed Matter 29, 465901 (2017).
8P. Giannozzi, S. Baronii, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dal, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncè, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, Journal of Physics: Condensed Matter 29, 465901 (2017).
9J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, Journal of Physics: Condensed Matter 14, 2745 (2002).
10P. Blaha, K. Schwarz, F. Tran, R. Laskowski, G. K. H. Madsen, and L. D. Marks, The Journal of Chemical Physics 152, 074101 (2020).
11N. Marzari and D. Vanderbilt, Physical Review B 56, 12847 (1997).
12J. Souza, N. Marzari, and D. Vanderbilt, Physical Review B 65 (2001), 10.1103/physrevb.65.035109.
13N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Reviews of Modern Physics 84, 1419 (2012).
14M. P. L. Sancho, J. M. L. Sancho, J. M. L. Sancho, and J. Rubio, J. Phys. F: Met. Phys. 15, 851 (1985).
15A. Calzolari, N. Marzari, I. Souza, and M. B. Nardelli, Physical Review B 69 (2004), 10.1103/physrevb.69.035108.
16G. Pizzi, V. Vitale, R. Arita, S. Bigel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresh, C. Johnson, T. Koreszun, J. Ibañez-Azpiroz, H. Lee, J.-M. Lihm, D. Marchand, A. Marrazzo, Y. Mokrousov, J. I. Mustafa, Y. Nohara, Y. Nomura, L. Paulatto,
X. APPENDIX: Bi$_2$Se$_3$

A. BS.scf.in

```plaintext
ACONTROL
calculation = 'scf'
prefix = 'BS'
outdir = 'bin/'
pseudo_dir = 'PP/'
verbosity = 'high'
/

&system
ibrav = 0
nat = 5
ntyp = 2
ecutwfc = 60 ! Ryberg
ecutrho = 500
occupations = 'smearing'
smearing = 'gaussian'
degauss = 0.01
!
ISOC
noncolin = .TRUE.
lepinorb = .TRUE.
starting_magnetization(1) = 0
starting_magnetization(2) = 0
/

&ELECTRONS
conv_thr = 1.0d-7
mixing_beta = 0.495
mixing_mode = 'TF'
diagonalization = 'david'
adaptive_thr = .true.
/

ATOMIC_SPECIES
Bi 208.9804 Bi.rel-pbesol-dn-kjpaw_psl.1.0.0.UPF
Se 78.96 Se.rel-pbesol-dn-kjpaw_psl.1.0.0.UPF
/

ATOMIC_POSITIONS (crystal)
Bi 0.000 0.000 0.000
Se 0.794 0.794 0.882
/

CELL_PARAMETERS (angstrom)
-2.069 -3.583614 0.000000 ! crystal lattice
information
2.069 -3.583614 0.000000
0.000 2.389075 9.546667
0.000 0.000 0.5000
0.000 0.2060 0.1180
0.000 0.000 0.3030
0.794 0.794 0.8820
/

K_POINTS (automatic)
8 8 8 1 1 1
```
B. BS.nscf.in

%CONTROL

calculation = 'nscf'
prefix = 'BS'
outdir = 'bin/
pseudo_dir = 'PP/
verbosity = 'high'
/

&system
ibrav = 0
nat = 5
ntyp = 2
ecutwfc = 60 !Ryberg
ecutrho = 500
occupations = 'smearing'
smearing = 'gaussian'
degauss = 0.01

nosym = .true.
nbnd = 200 !SOC
noncolin = .TRUE.
lspinorb = .TRUE.
starting_magnetization(1) = 0
starting_magnetization(2) = 0
/

&ELECTRONS
conv_thr = 1.0d-7
mixing_beta = 0.495
mixing_mode = 'TF'
diagonalization = 'david'
adaptive_thr = .true.
/

ATOMIC_SPECIES
Bi 208.9804 Bi.rel-pbesol-dn-kjpaw_psl.1.0.0.UPF
Se 78.96 Se.rel-pbesol-dn-kjpaw_psl.1.0.0.UPF
/

ATOMIC_POSITIONS (crystal)
Bi 0.3990 0.3990 0.6970
Bi 0.6010 0.6010 0.3030
Se 0.0000 0.0000 0.5000
Se 0.2060 0.2060 0.1180
Se 0.7940 0.7940 0.8820
/

CELL_PARAMETERS (angstrom)
-2.069 -3.583614 0.000000 ! crystal lattice
information
2.069 -3.583614 0.000000
0.000 2.389075 9.546667
/

K_POINTS (crystal)
64
0.0000000 0.0000000 0.0000000 1
0.0000000 0.2500000 0.0000000 1
0.0000000 0.5000000 0.0000000 1
0.0000000 0.2500000 0.7500000 1
0.0000000 0.0000000 0.7500000 1
0.0000000 0.2500000 0.0000000 1

C. BS.win

write_hr = .TRUE.
write_xyz = .TRUE.
!wannier_plot = .TRUE.
spinors = .TRUE.
num_wann = 50
dis_num_iter=1000
trial_step=50
num_iter=2000
exclude_bands = 1-50, 101-200

begin unit_cell_cart
-2.069 -3.583614 0.000000 ! crystal lattice information
 2.069 -3.583614 0.000000
 0.000 2.389075 9.546667
end unit_cell_cart

begin atoms_frac
Bi 0.3990 0.3990 0.6970
Bi 0.6010 0.6010 0.3030
Se 0.0000 0.0000 0.5000
Se 0.2060 0.2060 0.1180
Se 0.7940 0.7940 0.8820
end atoms_frac

begin projections
random
end projections

mp_grid = 4 4 4
begin kpoints
 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.25000000
 0.00000000 0.00000000 0.50000000
 0.00000000 0.00000000 0.75000000
 0.00000000 0.25000000 0.00000000
 0.00000000 0.25000000 0.25000000
 0.00000000 0.25000000 0.50000000
 0.00000000 0.25000000 0.75000000
 0.00000000 0.50000000 0.00000000
 0.00000000 0.50000000 0.25000000
 0.00000000 0.50000000 0.50000000
 0.00000000 0.50000000 0.75000000
 0.00000000 0.75000000 0.00000000
 0.00000000 0.75000000 0.25000000
 0.00000000 0.75000000 0.50000000
 0.00000000 0.75000000 0.75000000
 0.25000000 0.00000000 0.00000000
 0.25000000 0.00000000 0.25000000
 0.25000000 0.00000000 0.50000000
 0.25000000 0.00000000 0.75000000
 0.25000000 0.25000000 0.00000000
 0.25000000 0.25000000 0.25000000
 0.25000000 0.25000000 0.50000000
 0.25000000 0.25000000 0.75000000
 0.25000000 0.50000000 0.00000000
 0.25000000 0.50000000 0.25000000
 0.25000000 0.50000000 0.50000000
 0.25000000 0.50000000 0.75000000
 0.25000000 0.75000000 0.00000000
 0.25000000 0.75000000 0.25000000
 0.25000000 0.75000000 0.50000000
 0.25000000 0.75000000 0.75000000
 0.50000000 0.00000000 0.00000000
 0.50000000 0.00000000 0.25000000
 0.50000000 0.00000000 0.50000000
 0.50000000 0.00000000 0.75000000
 0.50000000 0.25000000 0.00000000
 0.50000000 0.25000000 0.25000000
 0.50000000 0.25000000 0.50000000
 0.50000000 0.25000000 0.75000000
 0.50000000 0.50000000 0.00000000
 0.50000000 0.50000000 0.25000000
 0.50000000 0.50000000 0.50000000
 0.50000000 0.50000000 0.75000000
 0.50000000 0.75000000 0.00000000
 0.50000000 0.75000000 0.25000000
 0.50000000 0.75000000 0.50000000
 0.50000000 0.75000000 0.75000000
 0.75000000 0.00000000 0.00000000
 0.75000000 0.00000000 0.25000000
 0.75000000 0.00000000 0.50000000
 0.75000000 0.00000000 0.75000000
 0.75000000 0.25000000 0.00000000
 0.75000000 0.25000000 0.25000000
 0.75000000 0.25000000 0.50000000
 0.75000000 0.25000000 0.75000000
 0.75000000 0.50000000 0.00000000
 0.75000000 0.50000000 0.25000000
 0.75000000 0.50000000 0.50000000
 0.75000000 0.50000000 0.75000000
 0.75000000 0.75000000 0.00000000
 0.75000000 0.75000000 0.25000000
 0.75000000 0.75000000 0.50000000
 0.75000000 0.75000000 0.75000000
end kpoints

XI. APPENDIX: HfP₂

A. HfP₂.scf.in

&CONTROL
calculation = 'scf'
prefix = 'HfP₂'
outdir = './bin'
pseudo_dir = './PP/'
verbosity='high'
/

&system
ibrav = 0
nat = 12
ntyp = 2
ecutwfc = 50 !Ryberg
occupations = 'tetrahedra_opt'
occupations = 'smearing'
smearing = 'gaussian'
degauss = 0.001

&SOC
noncolin = .FALSE.
lspinorb = .FALSE.
starting_magnetization(1) = 0
starting_magnetization(2) = 0
/

&ELECTRONS
mixing_beta = 0.495
conv_thr = 1.0d-6
mixing_mode = 'TF'
diagonalization = 'david'

ATOMIC_SPECIES
P 30.974 P.upf
Hf 178.49 Hf.upf

ATOMIC_POSITIONS (crystal)
Hf 0.222920 0.750000 0.335788
Hf 0.777080 0.250000 0.664212
Hf 0.277080 0.250000 0.835788
Hf 0.722920 0.750000 0.164212
P 0.092791 0.750000 0.642458
P 0.907209 0.250000 0.357542
P 0.407209 0.250000 0.142458
P 0.592791 0.750000 0.857542
P 0.110700 0.750000 0.039379
P 0.889300 0.250000 0.960621
P 0.389300 0.250000 0.539379
P 0.610700 0.750000 0.460621

CELL_PARAMETERS (angstrom)
6.460131 0.000000 0.000000
0.000000 3.509322 0.000000
0.000000 0.000000 8.687298

K_POINTS (crystal)
8 8 8 0 0 0

B.    HfP2.nscf.in

&CONTROL
calculation = 'nscf'
prefix = 'HfP2'
outdir = './bin'
pseudo_dir = './PP/'
verbosity = 'high'
/

&system
ibrav = 0
nat = 12
ntyp = 2
ecutwfc = 50 !Ryberg
occupations = 'smearing'
smearing = 'gaussian'
degauss = 0.01
nosym = .TRUE.
nbnd = 120
!SOC
noncolin = .FALSE.
lspinorb = .FALSE.
starting_magnetization(1) = 0
starting_magnetization(2) = 0
/

& ELECTRONS
mixing_beta = 0.495
conv_thr = 1.0d-6
mixing_mode = 'TF'
diagonalization = 'david'
/
begin kpoints
mp_grid = 8 8 8
end projections

random
begin projections
end atoms_frac

Hf 0.222920 0.750000 0.335788
Hf 0.777080 0.250000 0.664212
Hf 0.277080 0.250000 0.835788
Hf 0.722920 0.750000 0.164212
P 0.092791 0.750000 0.642458
P 0.907209 0.250000 0.357542
P 0.407209 0.250000 0.142458
P 0.592791 0.750000 0.857542
P 0.110700 0.750000 0.039379
P 0.889300 0.250000 0.960621
P 0.389300 0.250000 0.539379
P 0.610700 0.750000 0.460621
end unit_cell_cart

begin projections
random

!dis_froz_max = 14
!inner
!dis_win_max = 16.0
!dis_froz_min = 12
!dis_froz_max = 14

begin unit_cell_cart
6.460131 0.000000 0.000000
0.000000 3.509322 0.000000
0.000000 0.000000 8.687298
end unit_cell_cart

begin kpoints
0.8750000 0.7500000 0.6250000 1
0.8750000 0.7500000 0.7500000 1
0.8750000 0.7500000 0.8750000 1
0.8750000 0.8750000 0.0000000 1
0.8750000 0.8750000 0.1250000 1
0.8750000 0.8750000 0.2500000 1
0.8750000 0.8750000 0.3750000 1
0.8750000 0.8750000 0.5000000 1
0.8750000 0.8750000 0.6250000 1
0.8750000 0.8750000 0.7500000 1
0.8750000 0.8750000 0.8750000 1
0.8750000 0.7500000 0.6250000 1
0.8750000 0.7500000 0.7500000 1
0.8750000 0.7500000 0.8750000 1
0.8750000 0.8750000 0.0000000 1
0.8750000 0.8750000 0.1250000 1
0.8750000 0.8750000 0.2500000 1
0.8750000 0.8750000 0.3750000 1
0.8750000 0.8750000 0.5000000 1
0.8750000 0.8750000 0.6250000 1
0.8750000 0.8750000 0.7500000 1
0.8750000 0.8750000 0.8750000 1
XII. APPENDIX: HfP$_2$

A. TA.scf.in

```plaintext
&CONTROL
  calculation = 'scf'
  prefix = 'TA'
  outdir = './bin'
  pseudo_dir = './PP/'
  verbosity = 'high'
/

&system
  ibrav = 0
  nat = 4
  ntyp = 2
  ecutwfc = 55 !Ryberg
  ecutrho = 550
  occupations = 'smearing'
  smearing = 'mv'
  degauss = 0.002
end system
```

```plaintext
0.75000000 0.25000000 0.62500000
0.75000000 0.25000000 0.75000000
0.75000000 0.25000000 0.87500000
0.75000000 0.37500000 0.00000000
0.75000000 0.37500000 0.12500000
0.75000000 0.37500000 0.25000000
0.75000000 0.37500000 0.37500000
0.75000000 0.37500000 0.50000000
0.75000000 0.37500000 0.62500000
0.75000000 0.37500000 0.75000000
0.75000000 0.37500000 0.87500000
0.75000000 0.50000000 0.00000000
0.75000000 0.50000000 0.12500000
0.75000000 0.50000000 0.25000000
0.75000000 0.50000000 0.37500000
0.75000000 0.50000000 0.50000000
0.75000000 0.50000000 0.62500000
0.75000000 0.50000000 0.75000000
0.75000000 0.50000000 0.87500000
0.75000000 0.62500000 0.00000000
0.75000000 0.62500000 0.12500000
0.75000000 0.62500000 0.25000000
0.75000000 0.62500000 0.37500000
0.75000000 0.62500000 0.50000000
0.75000000 0.62500000 0.62500000
0.75000000 0.62500000 0.75000000
0.75000000 0.62500000 0.87500000
0.75000000 0.75000000 0.00000000
0.75000000 0.75000000 0.12500000
0.75000000 0.75000000 0.25000000
0.75000000 0.75000000 0.37500000
0.75000000 0.75000000 0.50000000
0.75000000 0.75000000 0.62500000
0.75000000 0.75000000 0.75000000
0.75000000 0.75000000 0.87500000
0.75000000 0.87500000 0.00000000
0.75000000 0.87500000 0.12500000
0.75000000 0.87500000 0.25000000
0.75000000 0.87500000 0.37500000
0.75000000 0.87500000 0.50000000
0.75000000 0.87500000 0.62500000
0.75000000 0.87500000 0.75000000
0.75000000 0.87500000 0.87500000
0.87500000 0.00000000 0.00000000
0.87500000 0.00000000 0.12500000
0.87500000 0.00000000 0.25000000
0.87500000 0.00000000 0.37500000
0.87500000 0.00000000 0.50000000
0.87500000 0.00000000 0.62500000
0.87500000 0.00000000 0.75000000
0.87500000 0.12500000 0.00000000
0.87500000 0.12500000 0.12500000
0.87500000 0.12500000 0.25000000
0.87500000 0.12500000 0.37500000
0.87500000 0.12500000 0.50000000
0.87500000 0.12500000 0.62500000
0.87500000 0.12500000 0.75000000
0.87500000 0.12500000 0.87500000
```

end kpoints
&SOC
noncolin = .TRUE.
lspinorb = .TRUE.
starting_magnetization(1) = 0
starting_magnetization(2) = 0
/

&SYSTEM
ibrav = 0
nat = 4
ntyp = 2
ecutwfc = 55 !Ryberg
ecutrho = 550
!occupations = 'fixed'
occupations = 'smearing'
smearing = 'mv'
degauss = 0.002
nosym = .true.
nhnd = 120
!SOC
noncolin = .TRUE.
lspinorb = .TRUE.
starting_magnetization(1) = 0
starting_magnetization(2) = 0
/

&ELECTRONS
conv_thr = 1.0d-7
mixing_beta = 0.495
diagonalization = 'david'
adaptive_thr = .true.
/

ATOMIC_SPECIES
Ta 180.94788 Ta_ONCV_PBE_fr.upf
As 74.9216 As_ONCV_PBE_fr.upf

ATOMIC_POSITIONS (crystal)
Ta 0.25000 0.75000 0.50000
Ta 0.00000 0.00000 0.00000
As 0.66700 0.16700 0.33400
As 0.41700 0.41700 0.83400

CELL_PARAMETERS (angstrom)
3.437000 -0.000000 -0.000000
-0.000000 3.437000 0.000000
-1.718500 -1.718500 5.828000

K_POINTS (crystal)
64
0.0000000 0.0000000 0.0000000 1
0.0000000 0.0000000 0.2500000 1
0.0000000 0.0000000 0.5000000 1
0.0000000 0.0000000 0.7500000 1
0.0000000 0.2500000 0.0000000 1
0.0000000 0.2500000 0.2500000 1
0.0000000 0.2500000 0.5000000 1
0.0000000 0.2500000 0.7500000 1
0.0000000 0.5000000 0.0000000 1
0.0000000 0.5000000 0.2500000 1
0.0000000 0.5000000 0.5000000 1
0.0000000 0.5000000 0.7500000 1
0.0000000 0.7500000 0.0000000 1
0.0000000 0.7500000 0.2500000 1
0.0000000 0.7500000 0.5000000 1
0.0000000 0.7500000 0.7500000 1
0.2500000 0.0000000 0.0000000 1
0.2500000 0.0000000 0.2500000 1
0.2500000 0.0000000 0.5000000 1
0.2500000 0.0000000 0.7500000 1
0.2500000 0.2500000 0.0000000 1
0.2500000 0.2500000 0.2500000 1
0.2500000 0.2500000 0.5000000 1
0.2500000 0.2500000 0.7500000 1
0.2500000 0.5000000 0.0000000 1
0.2500000 0.5000000 0.2500000 1
0.2500000 0.5000000 0.5000000 1
0.2500000 0.5000000 0.7500000 1
0.2500000 0.7500000 0.0000000 1
0.2500000 0.7500000 0.2500000 1
0.2500000 0.7500000 0.5000000 1
0.2500000 0.7500000 0.7500000 1
0.5000000 0.0000000 0.0000000 1

B. TA.nscf.in

&CONTROL
calculation = 'nscf'
prefix = 'TA'
outdir = './bin'
pseudo_dir = './PP/'
verbosity = 'high'
/

&SYSTEM
ibrav = 0
nat = 4
ntyp = 2
ecutwfc = 55 !Ryberg
ecutrho = 550
!occupations = 'fixed'
occupations = 'smearing'
smearing = 'mv'
degauss = 0.002
nosym = .true.
nhnd = 120
!SOC
noncolin = .TRUE.
lspinorb = .TRUE.
starting_magnetization(1) = 0
starting_magnetization(2) = 0
/

&ELECTRONS
conv_thr = 1.0d-7
mixing_beta = 0.495
diagonalization = 'david'
adaptive_thr = .true.
/

ATOMIC_SPECIES
Ta 180.94788 Ta_ONCV_PBE_fr.upf
As 74.9216 As_ONCV_PBE_fr.upf

ATOMIC_POSITIONS (crystal)
Ta 0.25000 0.75000 0.50000
Ta 0.00000 0.00000 0.00000
As 0.66700 0.16700 0.33400
As 0.41700 0.41700 0.83400

CELL_PARAMETERS (angstrom)
3.437000 -0.000000 -0.000000
-0.000000 3.437000 0.000000
-1.718500 -1.718500 5.828000

K_POINTS (crystal)
8 8 8 0 0 0
C. TA.win

write_hr = .TRUE.
write_xyz = .TRUE.
!wannier_plot = .TRUE.
spinors = .TRUE.
num_wann = 120
dis_num_iter=1000
!trial_step=50
num_iter = 200
!guiding_centres = .TRUE.
exclude_bands: 41-120

begin unit_cell_cart
3.437000 -0.000000 -0.000000
-0.000000 3.437000 0.000000
-1.718500 -1.718500 5.828000
end unit_cell_cart

begin atoms_frac
Ta 0.25000 0.75000 0.50000
Ta 0.00000 0.00000 0.00000
As 0.66700 0.16700 0.33400
As 0.41700 0.41700 0.83400
end atoms_frac

begin projections
random

mp_grid = 4 4 4
begin kpoints
0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.25000000
0.00000000 0.00000000 0.50000000
0.00000000 0.00000000 0.75000000
0.00000000 0.25000000 0.00000000
0.00000000 0.25000000 0.25000000
0.00000000 0.25000000 0.50000000
0.00000000 0.25000000 0.75000000
0.00000000 0.50000000 0.00000000
0.00000000 0.50000000 0.25000000
0.00000000 0.50000000 0.50000000
0.00000000 0.50000000 0.75000000
0.00000000 0.75000000 0.00000000
0.00000000 0.75000000 0.25000000
0.00000000 0.75000000 0.50000000
0.00000000 0.75000000 0.75000000
0.25000000 0.00000000 0.00000000
0.25000000 0.00000000 0.25000000
0.25000000 0.00000000 0.50000000
0.25000000 0.00000000 0.75000000
0.25000000 0.25000000 0.00000000
0.25000000 0.25000000 0.25000000
0.25000000 0.25000000 0.50000000
0.25000000 0.25000000 0.75000000
0.25000000 0.50000000 0.00000000
0.25000000 0.50000000 0.25000000
0.25000000 0.50000000 0.50000000
0.25000000 0.50000000 0.75000000
0.25000000 0.75000000 0.00000000
0.25000000 0.75000000 0.25000000
0.25000000 0.75000000 0.50000000
0.25000000 0.75000000 0.75000000
0.50000000 0.00000000 0.00000000
0.50000000 0.00000000 0.25000000
0.50000000 0.00000000 0.50000000
0.50000000 0.00000000 0.75000000
0.50000000 0.25000000 0.00000000
0.50000000 0.25000000 0.25000000
0.50000000 0.25000000 0.50000000
0.50000000 0.25000000 0.75000000
0.50000000 0.50000000 0.00000000
0.50000000 0.50000000 0.25000000
0.50000000 0.50000000 0.50000000
0.50000000 0.50000000 0.75000000
0.50000000 0.75000000 0.00000000
0.50000000 0.75000000 0.25000000
0.75000000 0.00000000 0.00000000
0.75000000 0.00000000 0.25000000
0.75000000 0.00000000 0.50000000
0.75000000 0.00000000 0.75000000
0.75000000 0.25000000 0.00000000
0.75000000 0.25000000 0.25000000
0.75000000 0.25000000 0.50000000
0.75000000 0.25000000 0.75000000
0.75000000 0.50000000 0.00000000
0.75000000 0.50000000 0.25000000
0.75000000 0.50000000 0.50000000
0.75000000 0.50000000 0.75000000
0.75000000 0.75000000 0.00000000
0.75000000 0.75000000 0.25000000

0.75000000  0.75000000  0.50000000  
0.75000000  0.75000000  0.75000000  
end kpoints