

# Supplementary Material for: Towards high-throughput many-body perturbation theory: efficient algorithms and automated workflows

Miki Bonacci,<sup>1,2,\*</sup> Junfeng Qiao,<sup>3</sup> Nicola Spallanzani,<sup>2</sup> Antimo Marrazzo,<sup>4</sup> Giovanni Pizzi,<sup>3,5</sup> Elisa Molinari,<sup>1,2</sup> Daniele Varsano,<sup>2</sup> Andrea Ferretti,<sup>2</sup> and Deborah Prezzi<sup>2</sup>

<sup>1</sup>FIM Department, University of Modena and Reggio Emilia, Via Campi 213/a, Modena, Italy

<sup>2</sup>S3 Center, Istituto Nanoscienze, CNR, Via Campi 213/a, Modena, Italy

<sup>3</sup>Theory and Simulation of Materials (THEOS) and National Centre for Computational Design and Discovery of Novel Materials (MARVEL),

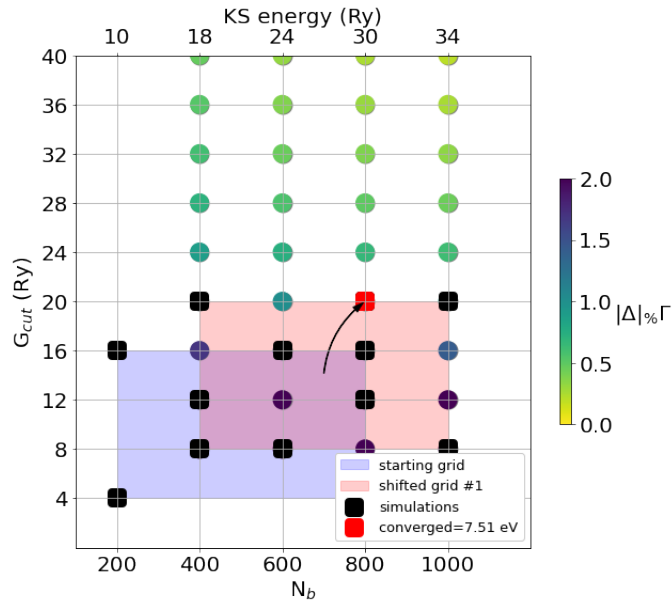
École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

<sup>4</sup>Dipartimento di Fisica, Università di Trieste, I-34151 Trieste, Italy

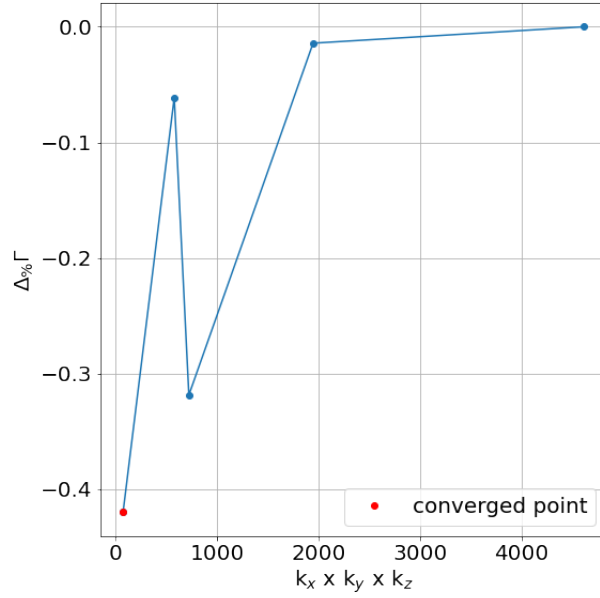
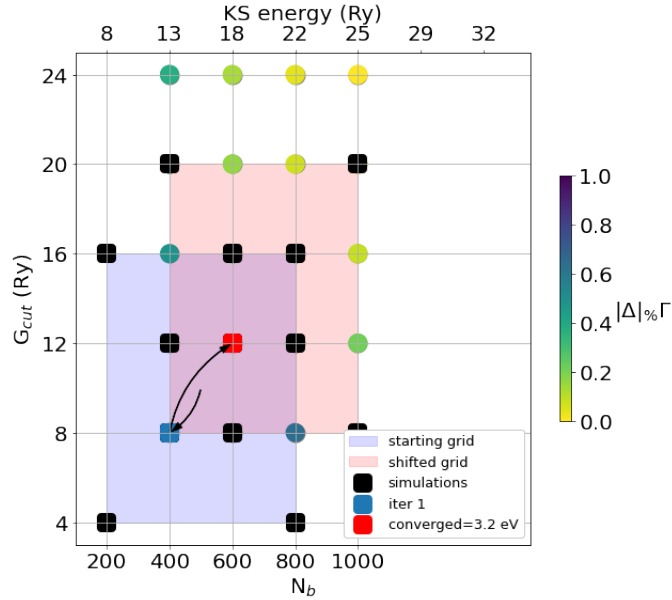
<sup>5</sup>Laboratory for Materials Simulations (LMS), Paul Scherrer Institut (PSI), CH-5232 Villigen PSI, Switzerland

## I. SUPPLEMENTARY FIGURES

In the following, we show the convergence plot for the systems studied in this work, as done for in the main text for the monolayer hBN case. We converged the  $\Gamma - \Gamma$  band gap with respect to the two coupled parameters  $N_b$  and  $G_{\text{cut}}$  and the  $\mathbf{k}$ -point grid as well, except for  $\text{TiO}_2$  and  $\text{ZnO}$ . For diamond, we converged also the FFTGvecs parameter, governing the Fast-Fourier-Transform (FFT) size for the evaluation of the generalized dipole matrix elements.



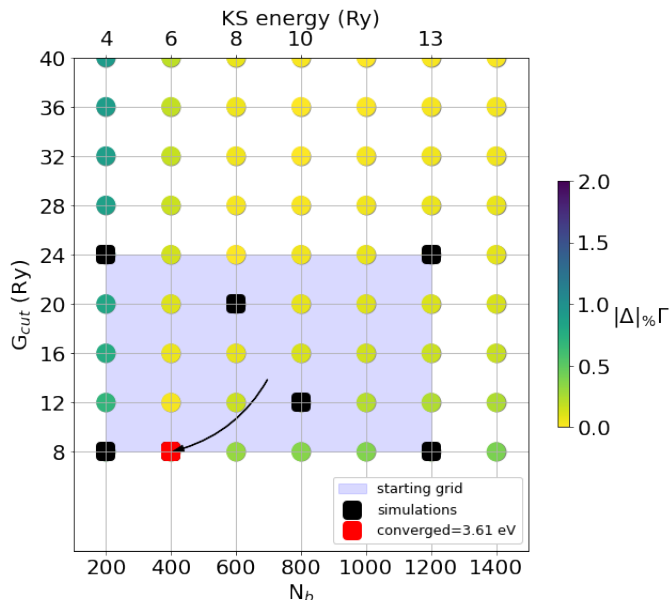
Supplementary Figure 1: Convergence of  $N_b$  and  $G_{\text{cut}}$  for bulk hBN.

Supplementary Figure 2:  $k$ -mesh convergence for bulk hBN.Supplementary Figure 3: Convergence of  $N_b$  and  $G_{cut}$  for Rutile  $\text{TiO}_2$ .

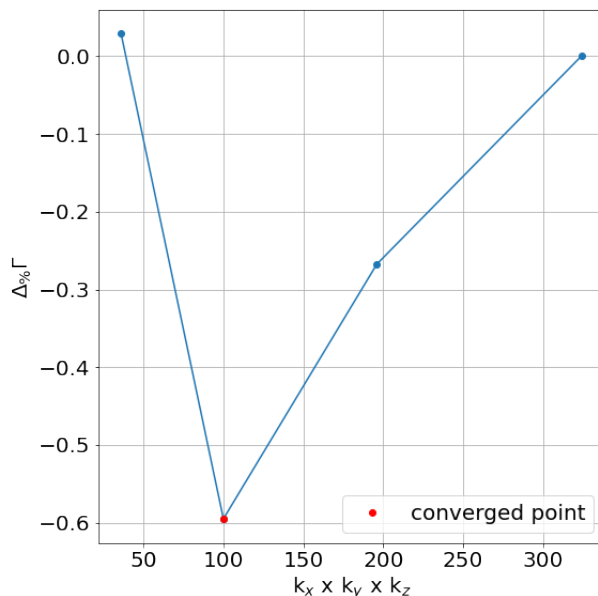
## II. SUPPLEMENTARY NOTES

### A. Wannierization at the GW level and restart from DFT Wannierization

Supplementary Figure 13 shows a detailed flowchart of the Wannierization workflow for GW band structure. The right column shows further details of Wannierization with respect to the ones provided in the main text. Firstly, a `Wannier90Calculation` is launched to generate the `mnkp` file for the subsequent `YppRestart`, which performs the `ypp` calculation for writing the QP correction in the `eig` file format for `Wannier90`. Then, a `Wannier90BandsWorkChain` is launched for Wannier interpolated bands at DFT level. Next, a `Gw2wannier90Calculation` sorts the new eigenenergies in ascending order, and rewrites relevant files. Finally, a `Wannier90BaseWorkChain` is launched to obtain the Wannier interpolated bands at  $G_0W_0$  level.

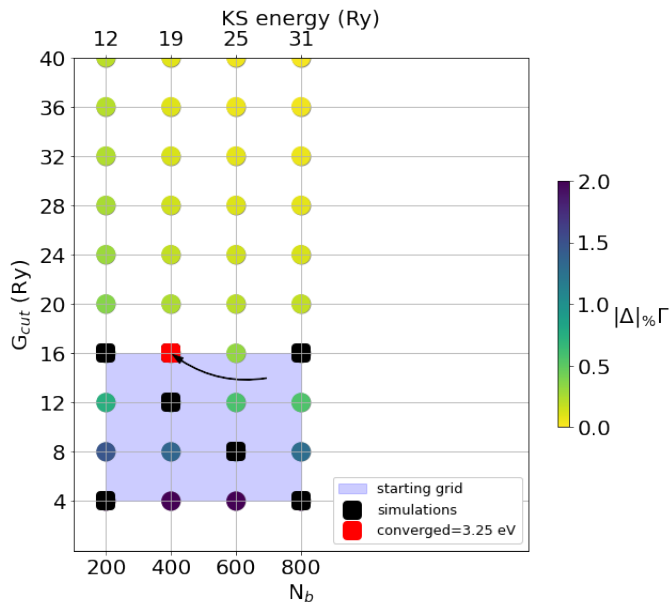


Supplementary Figure 4: Convergence of  $N_b$  and  $G_{\text{cut}}$  for monolayer  $\text{MoS}_2$ .

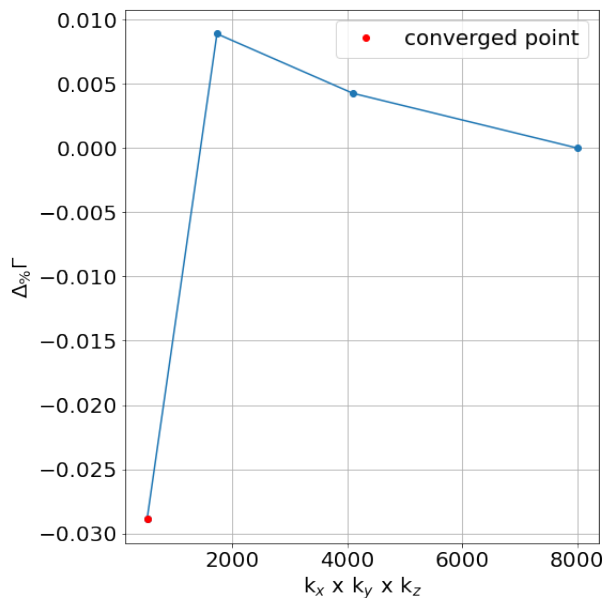


Supplementary Figure 5:  $\mathbf{k}$ -mesh convergence for monolayer  $\text{MoS}_2$ .

Some special care need to be taken at the step of incorporating the GW correction to the **Wannier90** inputs, i.e., the initial projection matrices  $A_{m\mathbf{n}\mathbf{k}}$ , the overlap matrices  $M_{m\mathbf{n}\mathbf{k}\mathbf{b}}$ , the eigenvalues  $\epsilon_{n\mathbf{k}}$ , and the Bloch wavefunctions  $|u_{n\mathbf{k}}\rangle$ , where the  $\mathbf{b}$  is the  $\mathbf{b}$  vectors connecting neighboring  $\mathbf{k}$ points.<sup>2</sup> Since the GW correction might change the order of eigenvalues, and the Wannierization disentanglement process requires the eigenvalues to be sorted in ascending order, we need to sort the GW corrected eigenenergies and rewrite the  $A_{m\mathbf{n}\mathbf{k}}$ ,  $M_{m\mathbf{n}\mathbf{k}\mathbf{b}}$ , and  $|u_{n\mathbf{k}}\rangle$  files accordingly. Then the Wannierization and interpolation proceed as usual. Another option is reusing the unitary transformation matrices  $U_{m\mathbf{n}\mathbf{k}}$  from the Wannierization at DFT level. Since the  $G_0W_0$  only corrects energies, while the wavefunctions are left unchanged, it is expected that the orbital characters of the wavefunctions are unchanged, thus the disentanglement process should still choose the same set of orbitals, i.e., the previous  $U_{m\mathbf{n}\mathbf{k}}$  should work as well for the  $G_0W_0$  corrections. Therefore, we can also reuse the  $U_{m\mathbf{n}\mathbf{k}}$  for the Wannier interpolation of  $G_0W_0$  eigenenergies thus skipping the second disentanglement step. The Wannier interpolation is the inverse Fourier transform of the Hamiltonian on the  $\mathbf{k}$ point



Supplementary Figure 6: Convergence of  $N_b$  and  $G_{\text{cut}}$  for silicon.

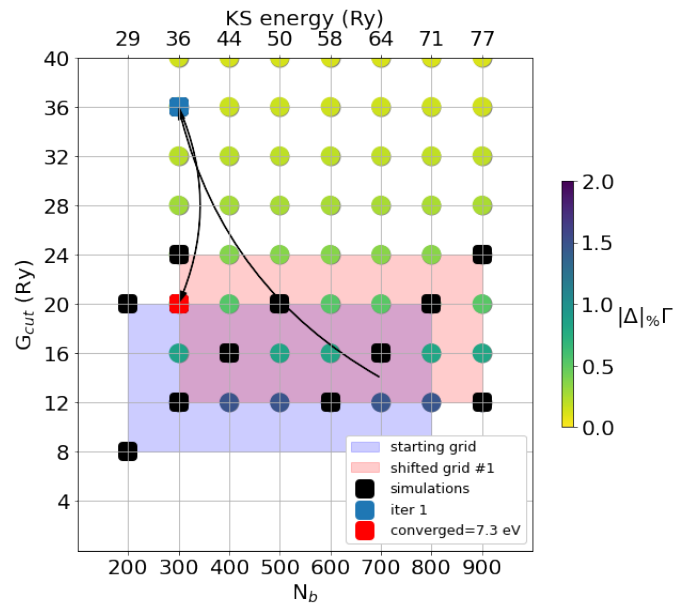


Supplementary Figure 7:  $\mathbf{k}$ -mesh convergence for silicon.

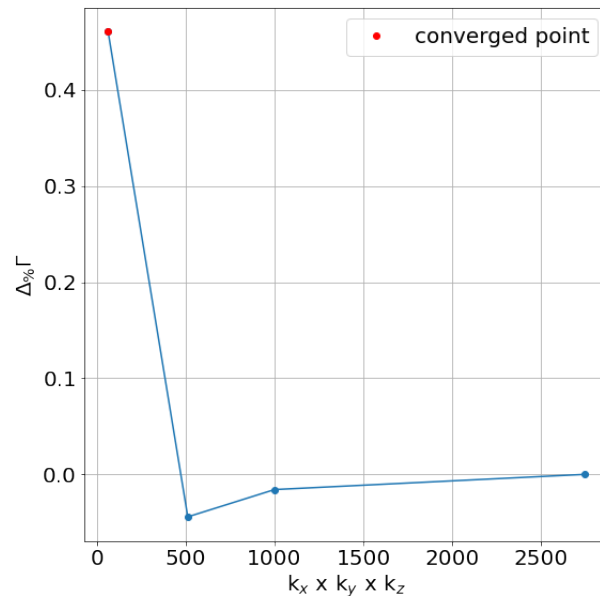
path for the band structure, the order of eigenenergies is irrelevant, thus the  $A_{m\mathbf{n}\mathbf{k}}$ ,  $M_{m\mathbf{n}\mathbf{k}\mathbf{b}}$ , and  $|u_{n\mathbf{k}}\rangle$  can be left intact. This can speed up large-scale calculations where the file size of these matrices are large, since usually the disk read/write speed is the slowest during the calculations. Moreover, reusing the  $U_{m\mathbf{n}\mathbf{k}}$  means it will reach the same minimum as the Wannierization at DFT level, thus the Wannier interpolation accuracy of GW eigenenergies is expected to be similar to that at DFT level.

### B. Effective band interpolation around given energy windows to reduce computational costs

The calculation of a large number of quasiparticle corrections can become easily computationally unfeasible for large systems. For example, in the case of  $\text{PdCoO}_2$ ,<sup>3</sup> a bulk metallic system of interest for Fermi surface measurements, the

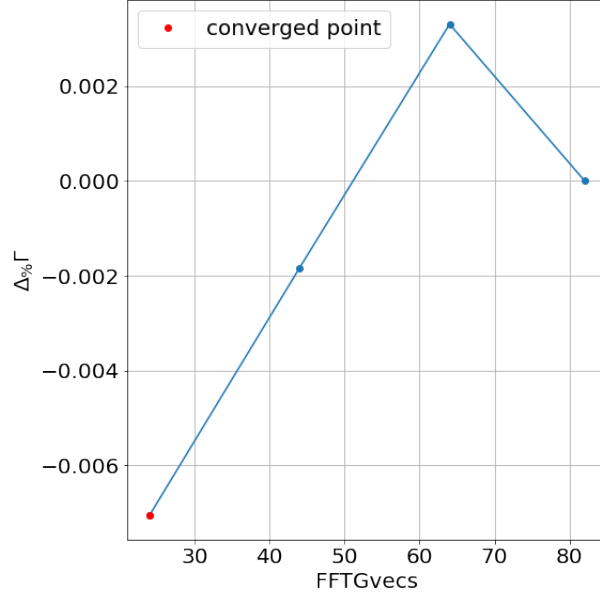


Supplementary Figure 8: Convergence of  $N_b$  and  $G_{\text{cut}}$  for diamond.

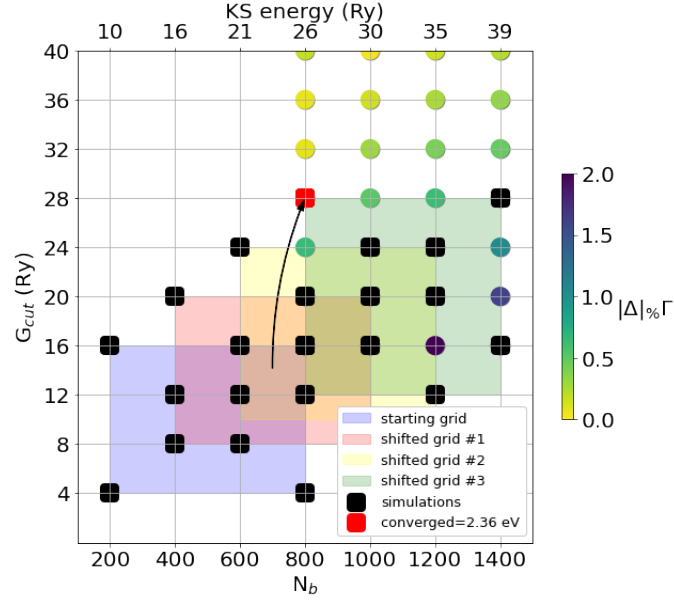


Supplementary Figure 9:  $\mathbf{k}$ -mesh convergence for diamond.

minimum number of GW eigenvalues needed to obtain the interpolated band structure is  $N_{QP} \sim 3 \cdot 10^4$ . We propose here a method to effectively reduce the number of GW evaluation needed, if only a given energy region of interest requires an accurate band interpolation. Within this approximation, the quasiparticle correction to KS eigenvalues is introduced in such a way to be the exact one in a given energy range, and replaced by a scissor and stretching evaluation outside. The exact GW correction goes continuously into the scissored and stretched one, and viceversa,



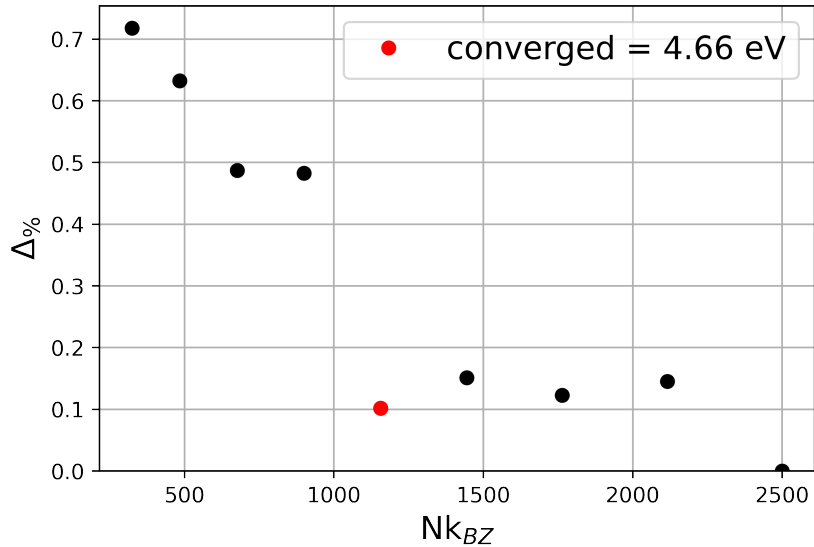
Supplementary Figure 10: FFT convergence for diamond.

Supplementary Figure 11: Convergence of  $N_b$  and  $G_{\text{cut}}$  for ZnO.

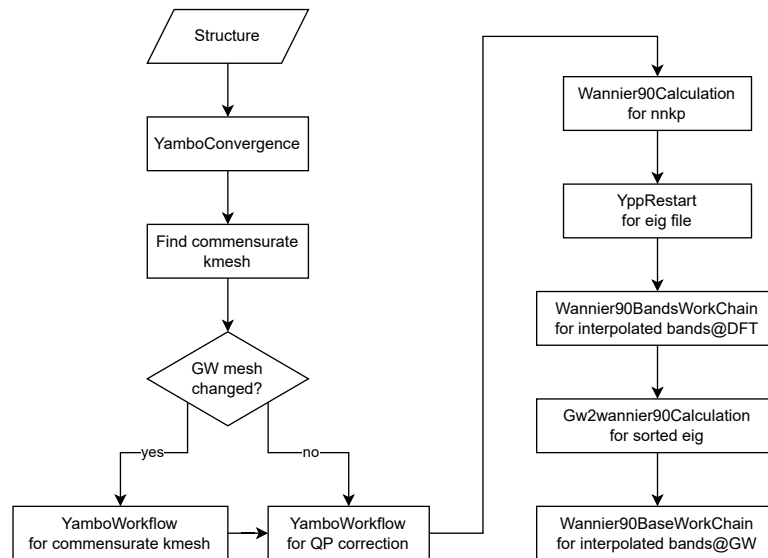
through a Fermi-Dirac smearing, considering an energy window centered at  $E_{\text{center}}$  and of width  $\mu$ :

$$E_{\text{GW}}^{\text{approx}} = \frac{E_{\text{GW}}^{\text{true}}}{1 + e^{\frac{|E_{\text{KS}} - E_{\text{center}}| - \mu}{T}}} + \left(1 - \frac{1}{1 + e^{\frac{|E_{\text{KS}} - E_{\text{center}}| - \mu}{T}}}\right) \times (A * E_{\text{KS}} + b) \quad (1)$$

The smearing parameter  $T$  is tuned in such a way to reduce the bump between the two different types of corrections. The  $E_{\text{center}}$ , in terms of KS energies, is user defined. In Supplementary Figure 14 are shown the results of the interpolation for different values of  $\mu$  and  $T$ , considering  $E_{\text{center}} = E_{\text{Fermi}}^{\text{KS}}$ . We observe that results are in very good agreement ( $\leq 10^{-2}$  eV) for  $T \leq 10^{-2}$  eV. In particular for the metals case, we are often interested only in the

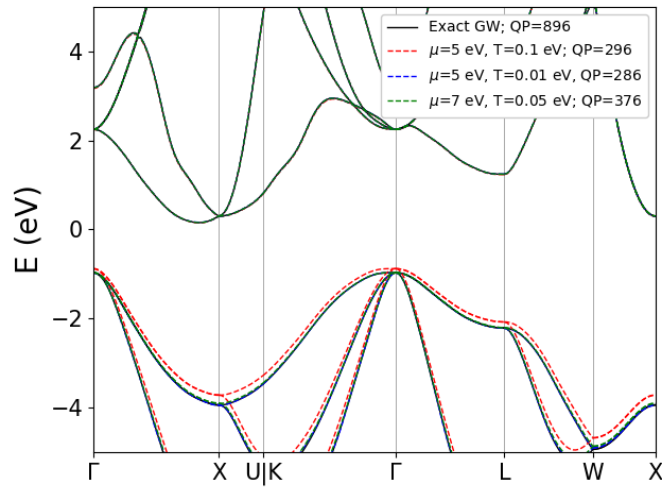


Supplementary Figure 12: **k**-mesh convergence of the lowest excitonic eigenvalue (BSE) for the double-sided fully hydrogenated graphene, as contained in Materials Cloud<sup>1</sup> at <https://archive.materialscloud.org/record/2023.40>. The  $\Delta\%$  represents the relative difference with respect to the last computed point, where a  $\Delta\%$  of  $\sim 0.1$  corresponds to an absolute difference of  $\sim 5$  meV.



Supplementary Figure 13: Detailed flowchart of the `YamboWannier90WorkChain` for automated GW convergence and Wannier interpolated GW band structure. The workflow performs the `Yambo` convergence, searching of commensurate kpoint mesh between `Yambo` and `Wannier90`, and running the `Yambo` quasiparticle calculation. The QP correction is provided and the final steps of the flow comprise the Wannierizations and the band interpolations at DFT level and GW level.

region around the Fermi level, for example to compute the Fermi surface. This means that we require the exact GW correction to DFT only for a small region belonging to an energy window supposed to include the Fermi level, approximating the corrections for the other eigenvalues in such a way to still have a good Wannier interpolation but reducing the number of explicit GW eigenvalues. Anyway, we stress that since  $G_0W_0$  is not particle conserving, a large region centered at the DFT Fermi energy is needed in order to possibly include the new and unknown GW Fermi level, that can be calculated a posteriori.



Supplementary Figure 14: **FD and scissor corrections for Silicon.** We observe a rather good agreement, around  $10^{-2}$  eV, between the exact solution and the approximated ones for  $T \leq 10^{-2}$  eV.  $E_{\text{center}}$  here is chosen to be the KS-DFT Fermi level.

### Supplementary References

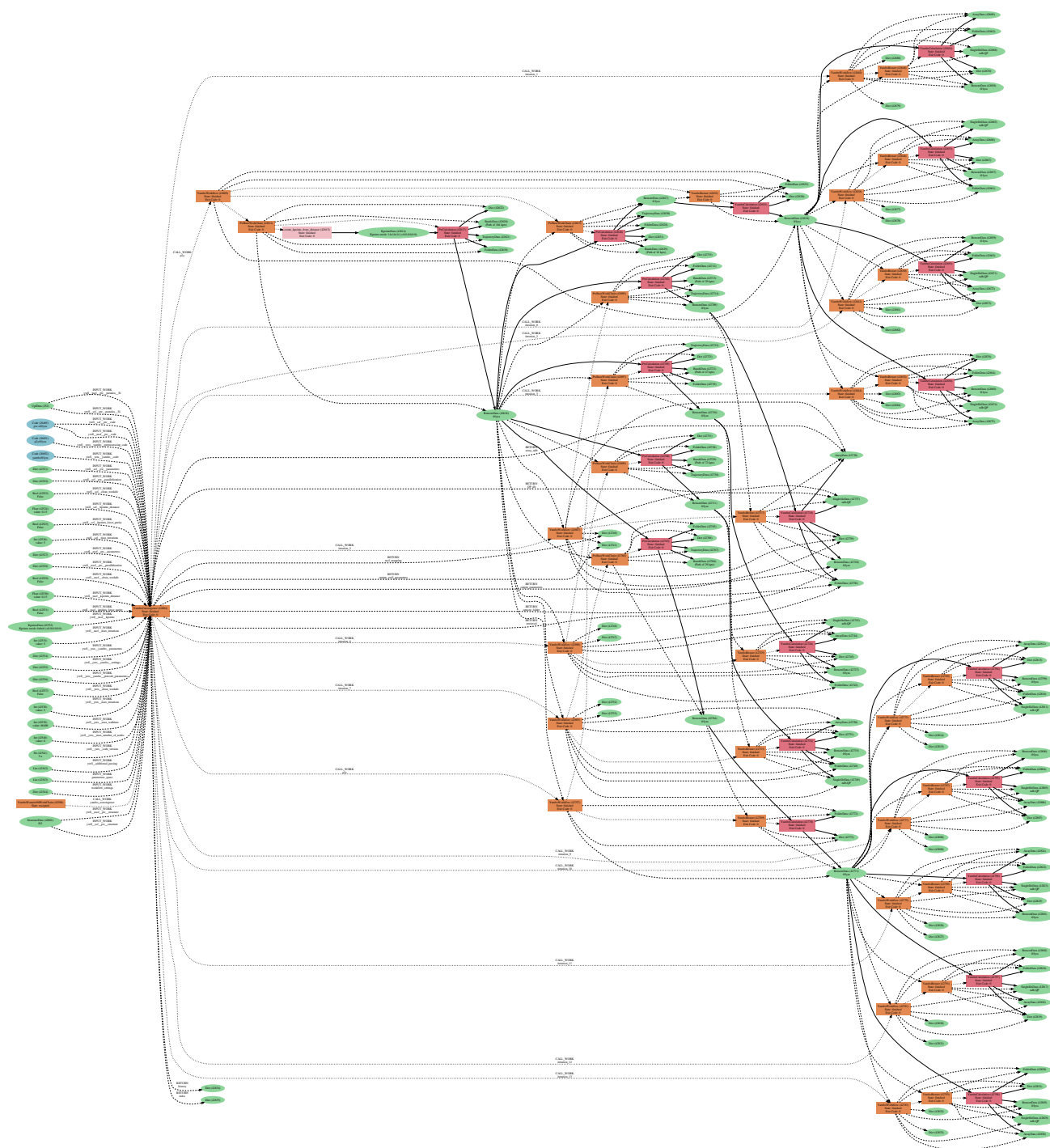
\* corresponding author: [miki.bonacci@nano.cnr.it](mailto:miki.bonacci@nano.cnr.it)

<sup>1</sup> L. Talirz, et al., Materials Cloud, a platform for open computational science, *Scientific Data* **7**, 299 (2020).

<sup>2</sup> G. Pizzi, et al., Wannier90 as a community code: new features and applications, *Journal of Physics Condensed Matter* **32**, 165902 (2020).

<sup>3</sup> M. Gruner, U. Eckern, and R. Pentcheva, Impact of strain-induced electronic topological transition on the thermoelectric properties of  $\text{ptcoo}_2$  and  $\text{pdcoo}_2$ , *Physical Review B* **92** (2015).





Supplementary Figure 15: **Example of provenance graph for YamboConvergence.** Provided the inputs (the first column of boxes from the left), the YamboConvergence calls a series of YamboWorkflows to perform the needed runs of PwCalculation and YamboCalculation, represented by the red boxes. Outputs are then generated and analysed by the workflow.