## **Automated high-throughput wannierisation**

This section contains the results from the validation of a procedure for automatically generating maximally-localized Wannier functions in a high-throughput framework.

Maximally-localised Wannier functions (MLWFs) are routinely used to compute from firstprinciples advanced materials properties that require very dense Brillouin zone integration and to build accurate tight-binding models for scale-bridging simulations. At the same time, highthroughput (HT) computational materials design is an emergent field that promises to accelerate the reliable and cost-effective design and optimisation of new materials with target properties. The use of MLWFs in HT workflows has been hampered by the fact that generating MLWFs automatically and robustly without any user intervention and for arbitrary materials is, in general, very challenging. We address this problem directly by proposing a procedure for automatically generating MLWFs for HT frameworks. Our approach is based on the selected columns of the density matrix method and we present the details of its implementation in an AiiDA workflow. We apply our approach to a dataset of 200 bulk crystalline materials that span a wide structural and chemical space. We assess the quality of our MLWFs in terms of the accuracy of the bandstructure interpolation that they provide as compared to the band-structure obtained via full firstprinciples calculations. Finally, we provide a downloadable virtual machine that allows to reproduce the results of our paper, including all first-principles and atomistic simulations as well as the computational workflows.

In this section, for all the materials we studied we provide:

- optimised crystal structure (variable-cell relaxation)
- electronic band structure obtained with Quantum ESPRESSO at the DFT-PBE level
- interpolated band structure obtained with Wannier90
- bands distance  $(\eta, \eta_{max})$  and total localization spread  $(\Omega)$
- comparison between different linear k-point spacings (0.15,0.2,0.3,0.4 Å<sup>-1</sup>) and techniques (SCDM with/without MLWF, random projections with MLWF)

More details can be found in: Valerio Vitale, Giovanni Pizzi, Antimo Marrazzo, Jonathan R. Yates, Nicola Marzari, Arash A. Mostofi, npj Computational Materials 6, 66 (2020) doi: 10.1038/s41524-020-0312-y