MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH

MAX DRIVING THE EXASCALE TRANSITION

THEOS
THEORY AND SIMULATION OF MATERIALS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Automated high-throughput Wannierisation

Implementation of the entangled SCDM-k method in Quantum ESPRESSO and full automation with AiiDA workflows

Giovanni Pizzi,[1] Valerio Vitale,[2,3] Antimo Marrazzo,[1]
Nicola Marzari,[1] Jonathan R. Yates,[4] Arash A. Mostofi[3]

(1) Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), EPFL

(2) Cavendish Laboratory, Department of Physics, University of Cambridge, UK

(3) Departments of Materials and Physics, and the Thomas Young Centre for Theory and Simulation of Materials, Imperial College London, UK

(4) Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, UK

Hunting for projections

- Usually, code **needs user to specify initial projections** (guesses for the minimisation procedure)

- This needs a lot of chemical understanding and experience. Biggest challenges for new users, and very hard to automate

Some recent emails from the Wannier90 mailing list:

Dear Experts,
How can I define the correct projection of particular material? […]

Dear Sir,
I need to know the correct projection of Graphene for a converged wannier calculation. […]

Dear Wannier Community,
[...]
My question is how do I define three projections for the half-filled p states of the two As atoms?

Hunting for projections

- Usually, code **needs user to specify initial projections** (guesses for the minimisation procedure)

- This needs a lot of chemical understanding and experience. Biggest challenges for new users, and very hard to automate

- Recently: **SCDM** method (selected columns of the density matrix) [1,2] proposed, aiming at automatically finding Wannier functions

Some recent emails from the Wannier90 mailing list:

Dear Experts,
How can I define the correct projection of particular material? […]

Dear Sir,
I need to know the correct projection of Graphene for a converged wannier calculation. […]

Dear Wannier Community,
[...]
My question is how do I define three projections for the half-filled p states of the two As atoms?

[1] Damle, A., Lin, L. & Ying, L. Compressed representation of Kohn–Sham orbitals via selected columns of the density matrix. *Journal of Chemical Theory and Computation* 11, 1463–1469 (2015).
[2] Damle, A. & Lin, L. Disentanglement via entanglement: A unified method for Wannier localization. *Multiscale Modeling & Simulation* 16, 1392–1410 (2018).

Aim and outcome

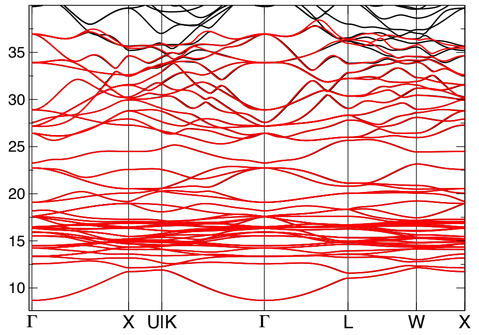
- **Implementation of the SCDM method** in Quantum ESPRESSO (pw2wannier90.x code)

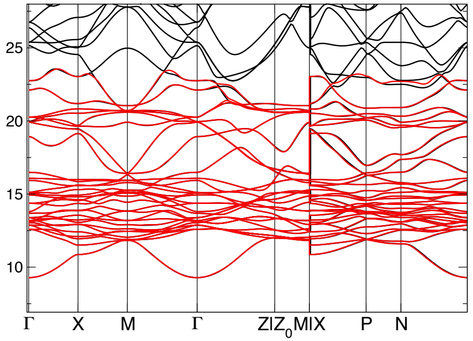
- **Determination of an algorithm** for the choice of "free" parameters in SCDM

$$N, \mu, \sigma$$

- **Implementation of automatic workflows** (within AiiDA) for fully automatic Wannierisation

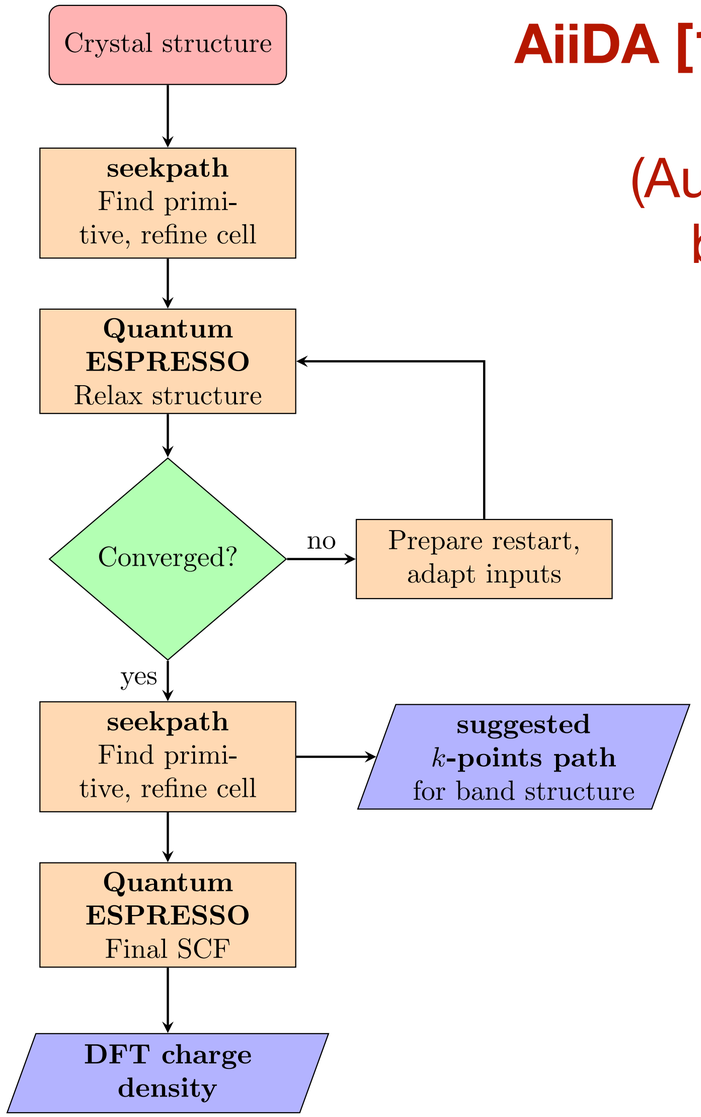
- **Validation** on a set of ~200 crystalline materials (~80 insulators) covering the chemical space





V. Vitale, GP et al., arXiv:1909.00433, accepted in npj Comp. Mat.

Automatic workflows: the implementation

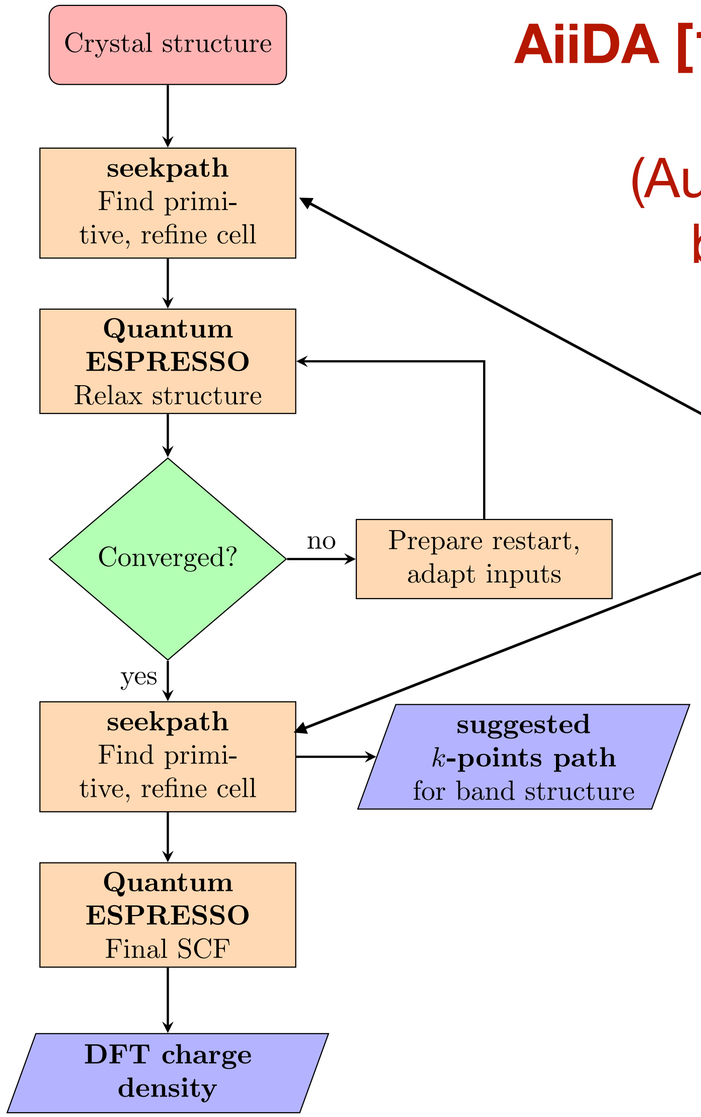


AiiDA [1] + AiiDA-Quantum ESPRESSO [2] plugin and workflows

(Automatic workflows implemented by Sebastiaan P. Huber, EPFL)

[1] http://www.aiida.net
[2] https://github.com/aiidateam/aiida-quantumespresso
[3] https://www.materialscloud.org/work/tools/seekpath

Automatic workflows: the implementation



AiiDA [1] + AiiDA-Quantum ESPRESSO [2] plugin and workflows

(Automatic workflows implemented by Sebastiaan P. Huber, EPFL)

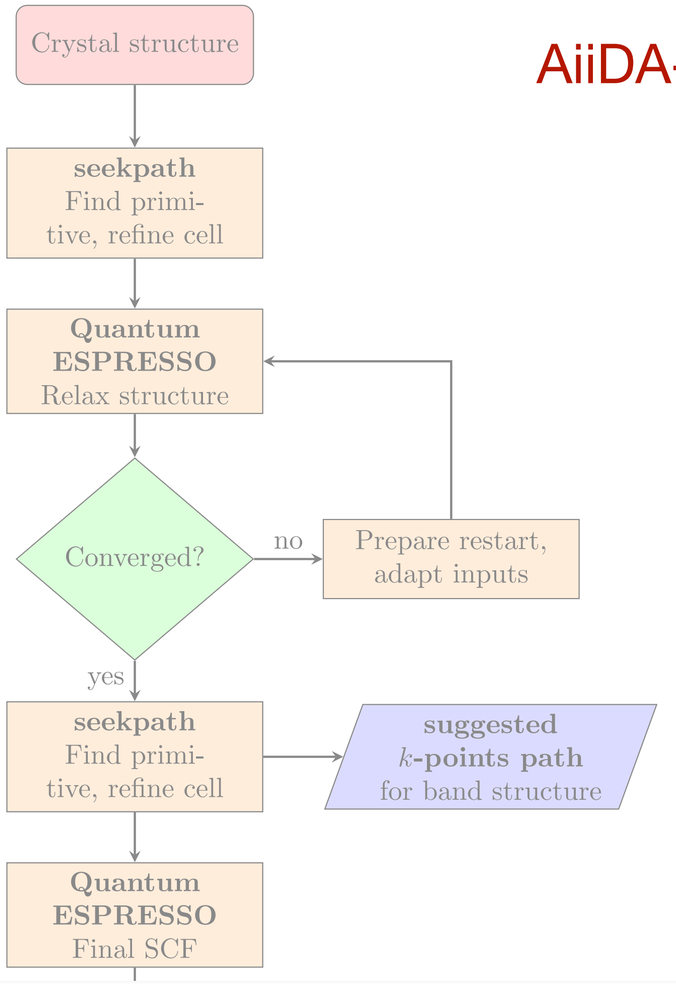
seekpath [3] (using internally spglib) to find primitive cell, refine it, standardize its orientation, find standard k-path for band structure

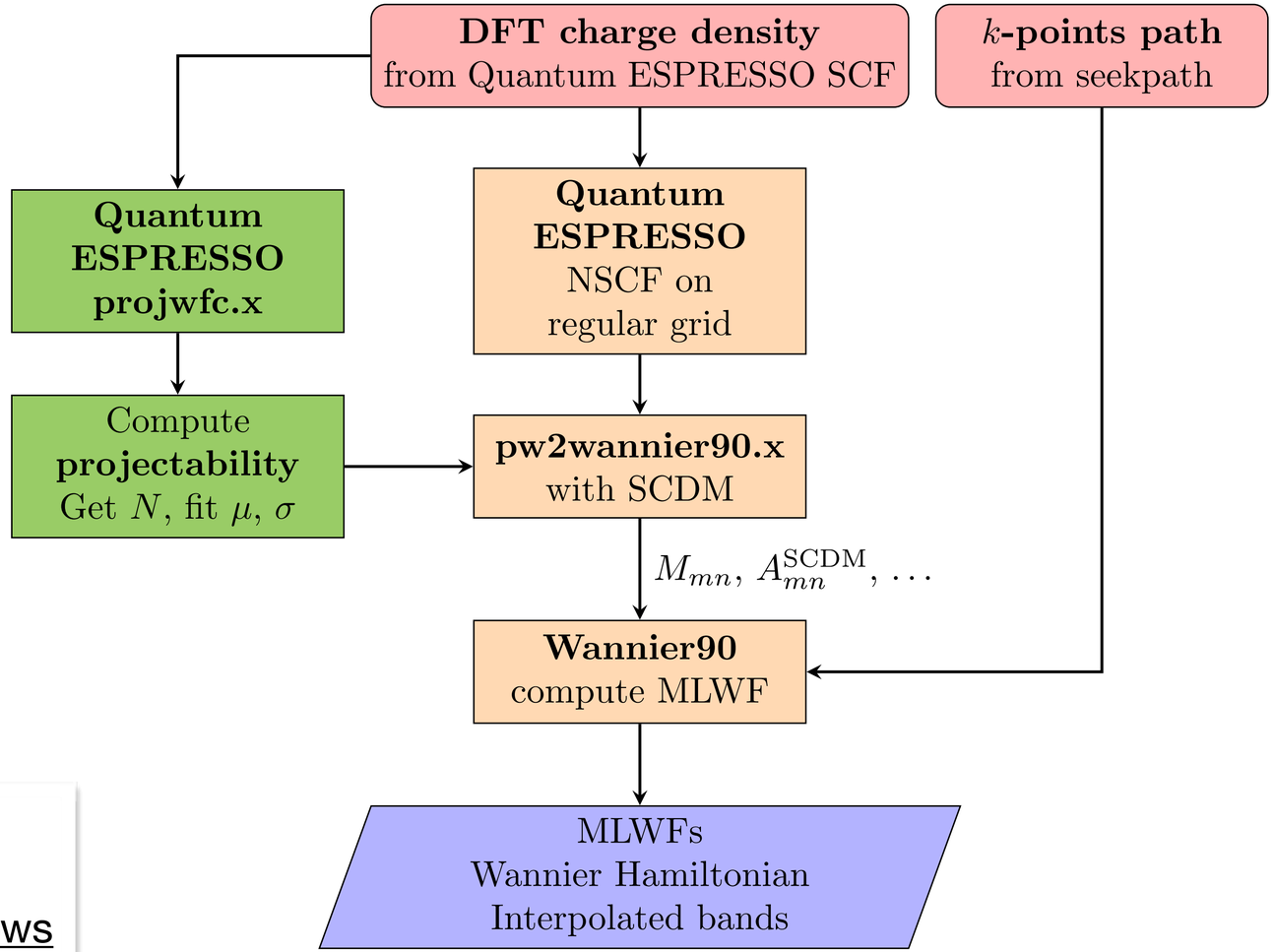
[1] http://www.aiida.net
[2] https://github.com/aiidateam/aiida-quantumespresso
[3] https://www.materialscloud.org/work/tools/seekpath

Automatic workflows: the implementation

AiiDA [1] + AiiDA-Wannier90 [2] plugin and workflows

AiiDA-Wannier90 plugin by *Dominik Gresch, Antimo Marrazzo, Daniel Marchand, Giovanni Pizzi, Junfeng Qiao, Norma Rivano*



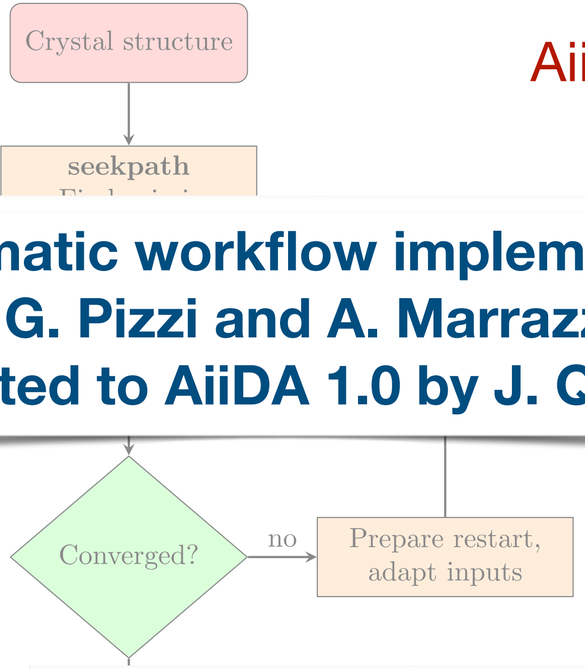


[1] http://www.aiida.net
[2] https://github.com/aiidateam/aiida-wannier90,
https://github.com/aiidateam/aiida-wannier90-workflows

Automatic workflows: the implementation

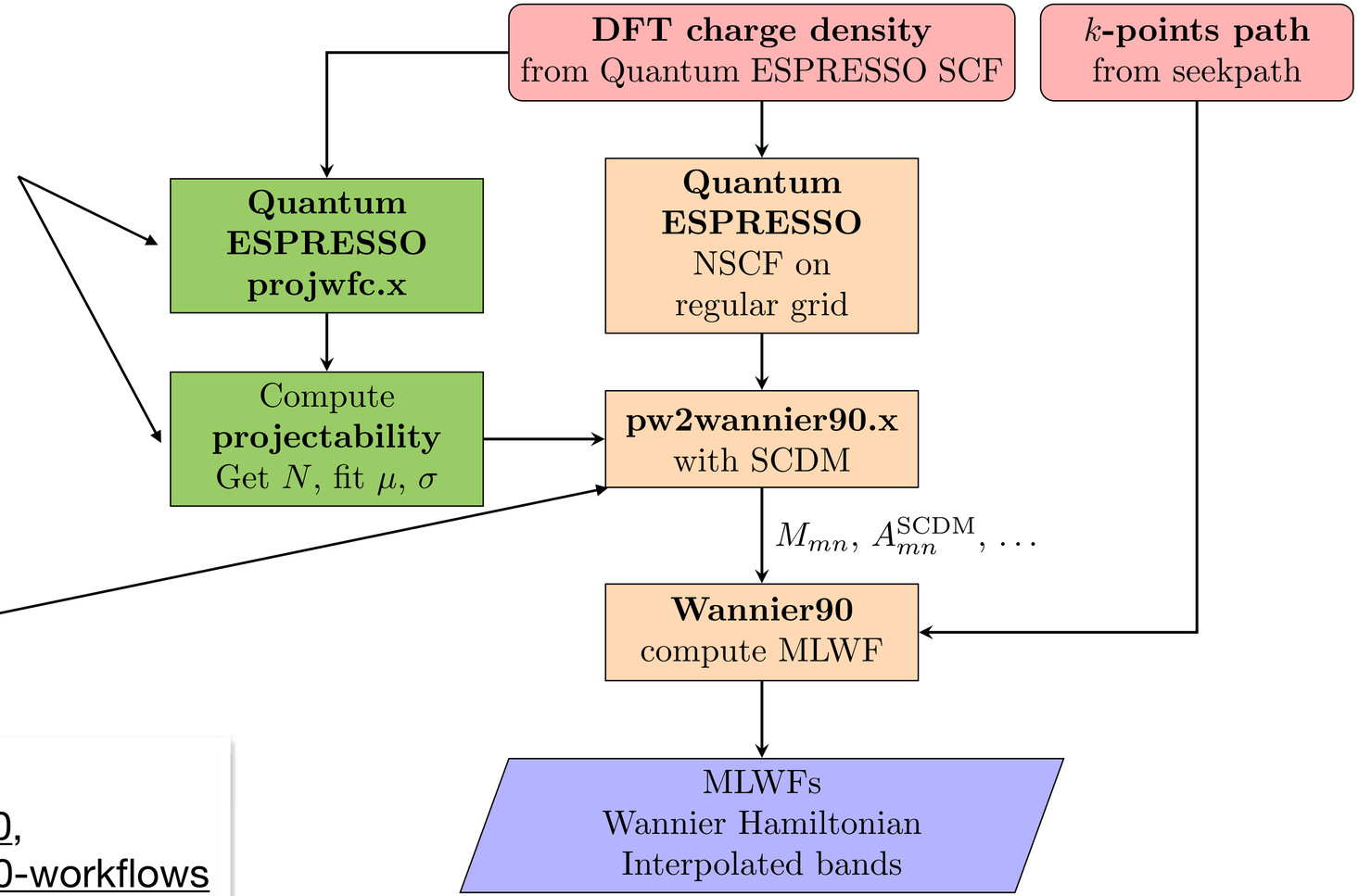
AiiDA [1] + AiiDA-Wannier90 [2] plugin and workflows

AiiDA-Wannier90 plugin by *Dominik Gresch, Antimo Marrazzo, Daniel Marchand, Giovanni Pizzi, Junfeng Qiao, Norma Rivano*



Automatic workflow implemented by G. Pizzi and A. Marrazzo, ported to AiiDA 1.0 by J. Qiao

SCDM implementation
by Valerio Vitale
in Quantum ESPRESSO [1]
(pw2wannier90.x code)

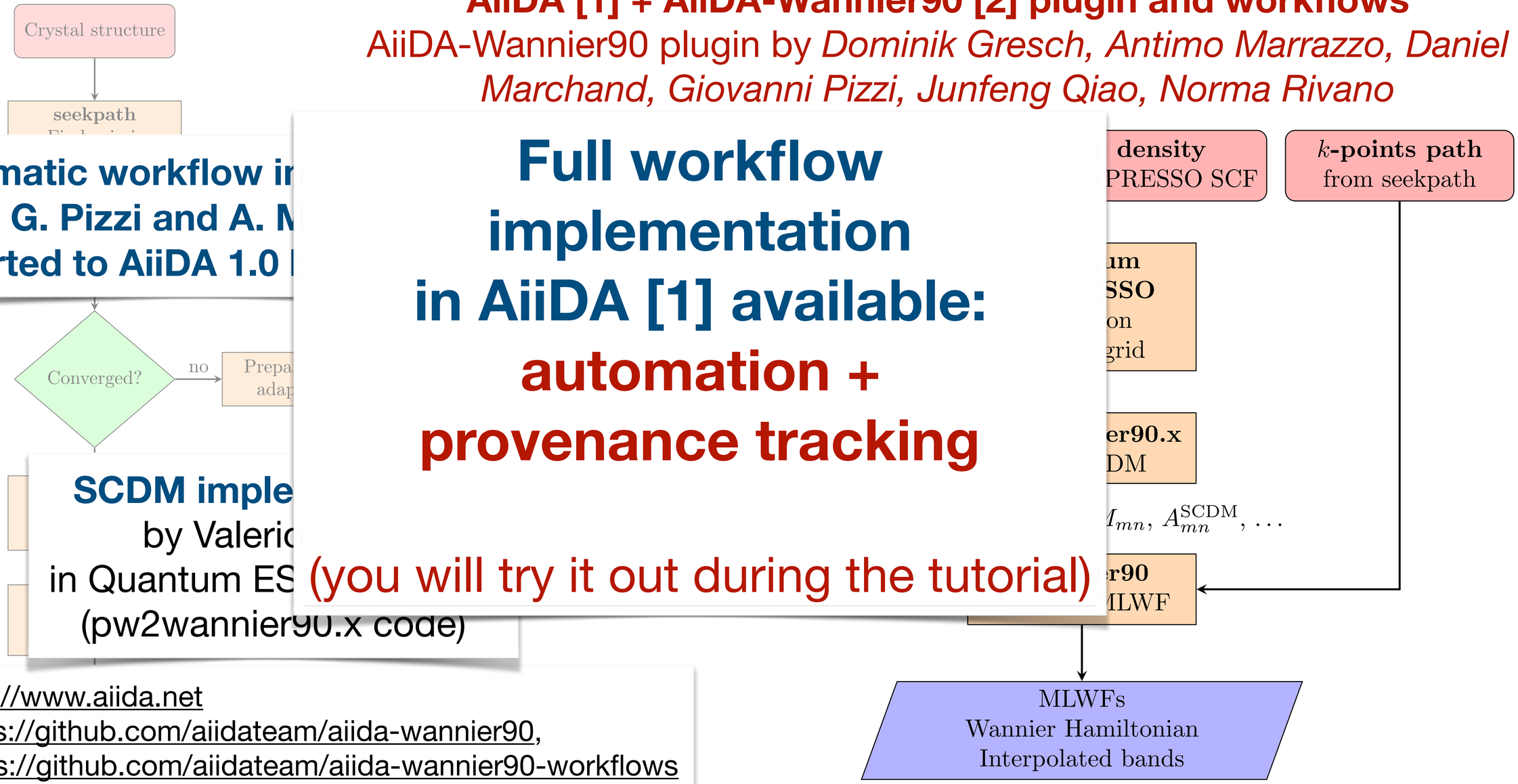


[1] http://www.aiida.net
[2] https://github.com/aiidateam/aiida-wannier90,
https://github.com/aiidateam/aiida-wannier90-workflows

Automatic workflows: the implementation

AiiDA [1] + AiiDA-Wannier90 [2] plugin and workflows

AiiDA-Wannier90 plugin by *Dominik Gresch, Antimo Marrazzo, Daniel Marchand, Giovanni Pizzi, Junfeng Qiao, Norma Rivano*



Automatic workflow ir
by G. Pizzi and A. M
ported to AiiDA 1.0 l

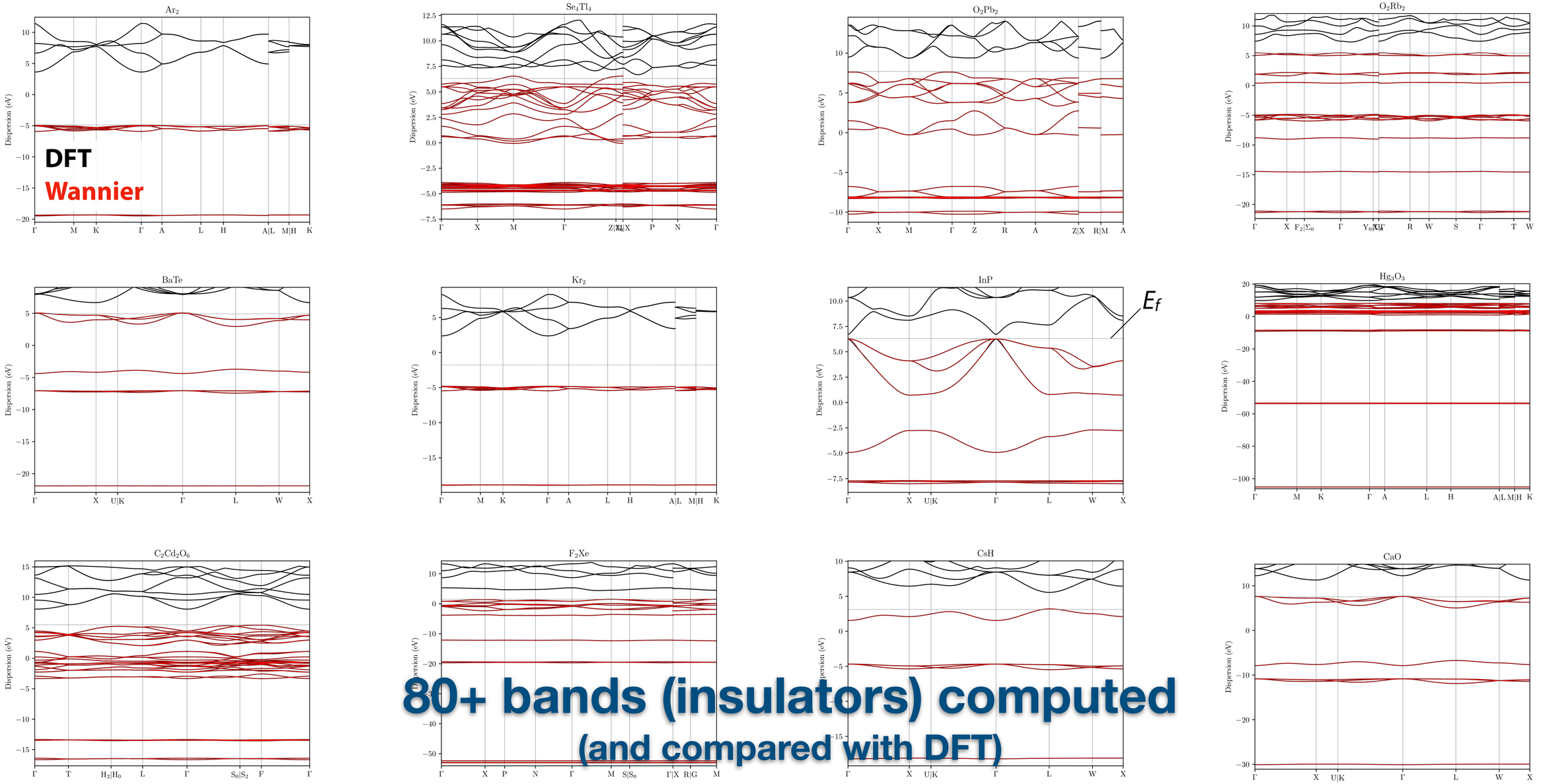
SCDM imple
by Valeric
in Quantum ES
(pw2wannier90.x code)

Full workflow implementation in AiiDA [1] available: automation + provenance tracking

(you will try it out during the tutorial)

[1] http://www.aiida.net
[2] https://github.com/aiidateam/aiida-wannier90,
https://github.com/aiidateam/aiida-wannier90-workflows

Insulators: (subset of the) results



80+ bands (insulators) computed

(and compared with DFT)

Bands distance

- To assess quality of Wannierisation and interpolation:
  we define a **bands distance**
  (between DFT bands and interpolated bands)

$$\eta = \sqrt{\sum_{n\mathbf{k}} \left(\varepsilon_{n\mathbf{k}}^{\mathrm{DFT}} - \varepsilon_{n\mathbf{k}}^{\mathrm{Wan}}\right)^2},$$

Average bands distance

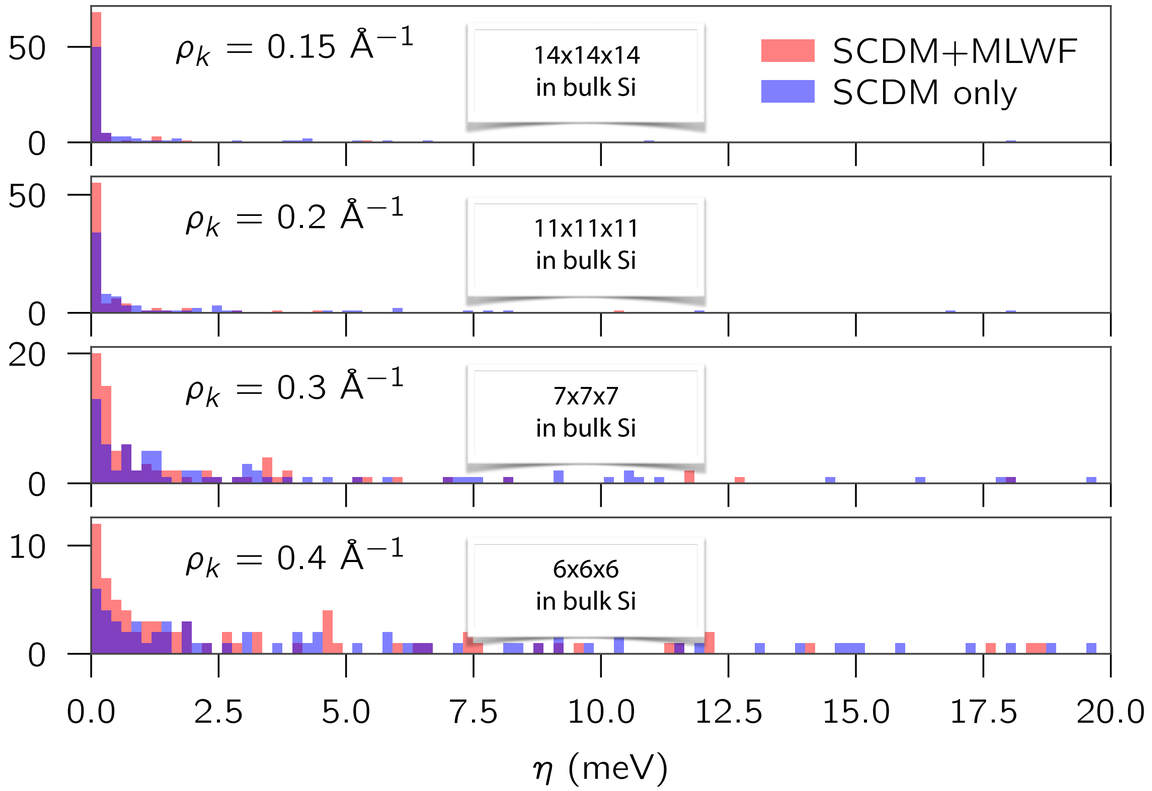
$$\eta^{\max} = \max_{n\mathbf{k}} \left(\left|\varepsilon_{n\mathbf{k}}^{\mathrm{DFT}} - \varepsilon_{n\mathbf{k}}^{\mathrm{Wan}}\right|\right)$$

Max bands distance

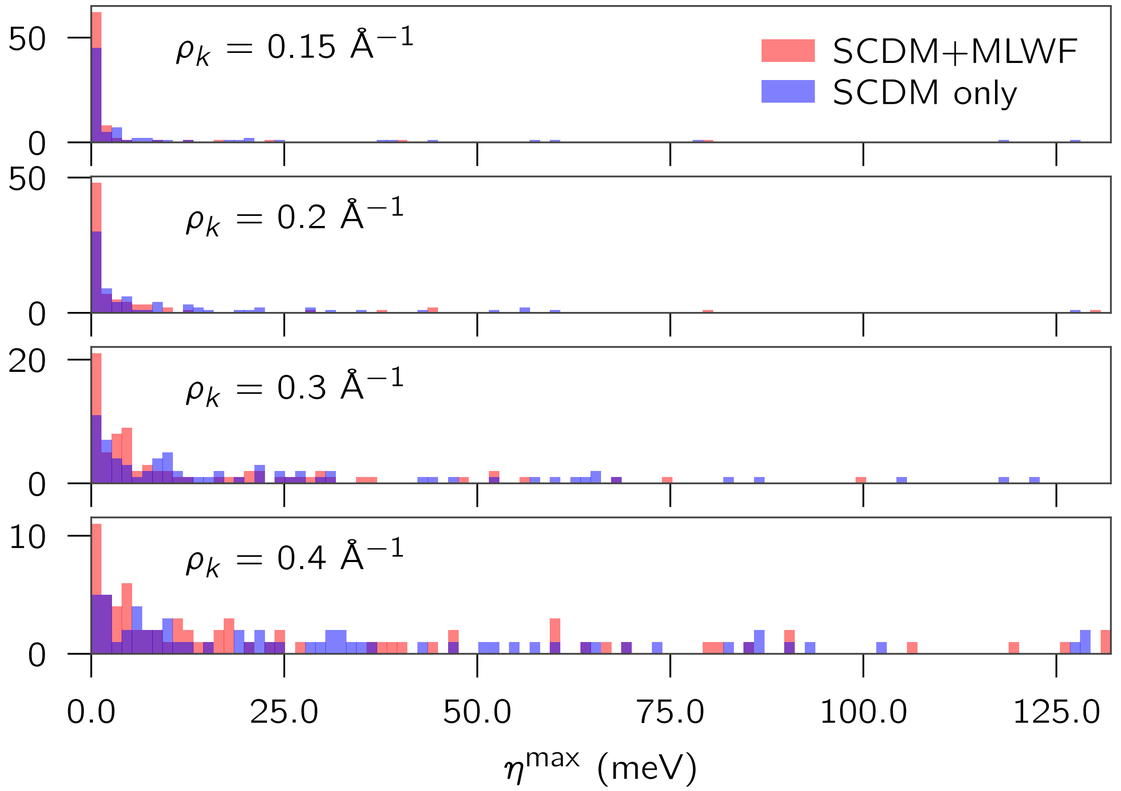
- Moreover, we want to assess the importance of the **density of *k*-points** in the NSCF/Wannierisation step
  We will use a linear density ρ_k in Å^{-1}

Comparison of SCDM + MLWF with SCDM only (80 insulators)

Max bands distance



Average bands distance

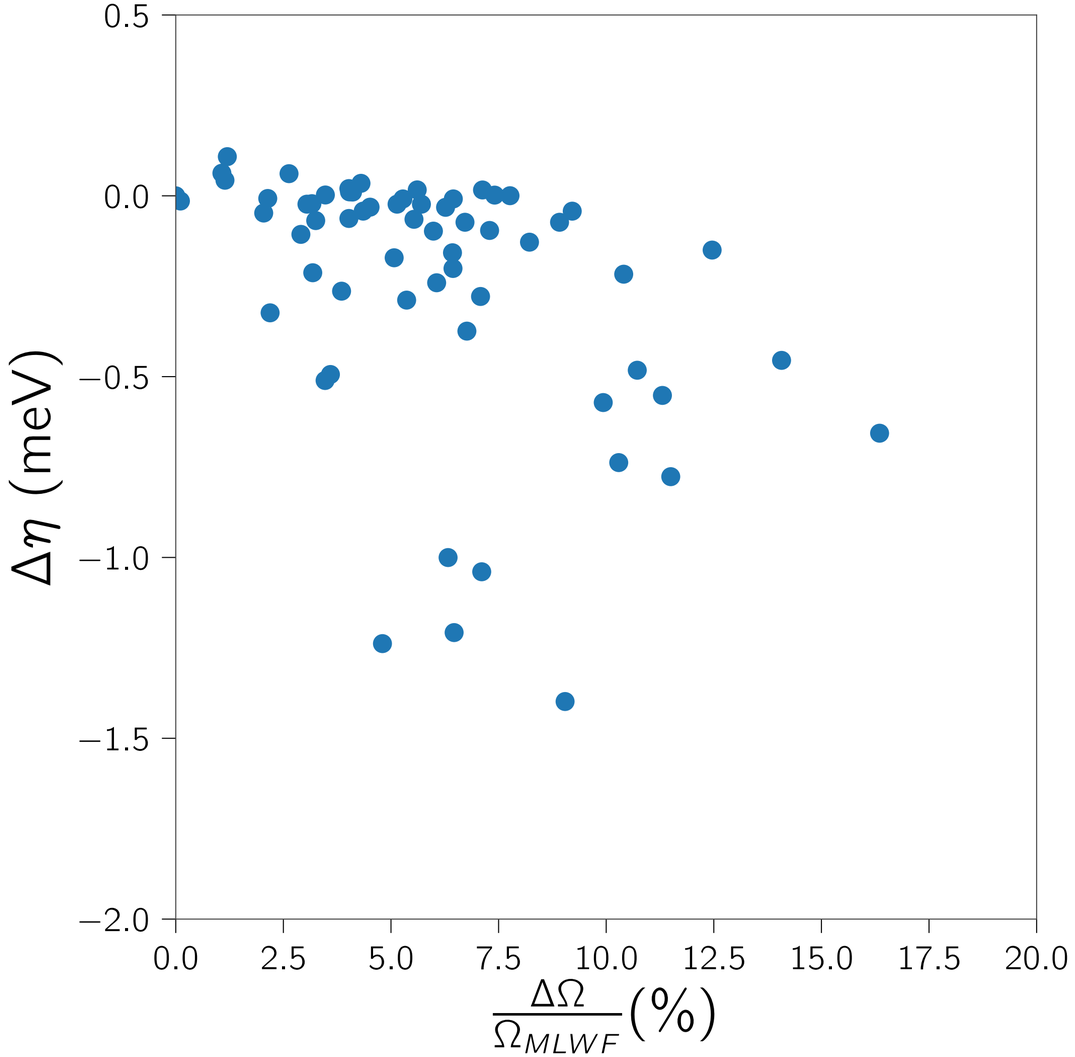


- Good results require a density of at least 0.2 Å^{-1} or more dense
- **For insulators**, SCDM-only already provides very good results; MLWF improves them
- In general, very small band distances (i.e. very good interpolation)

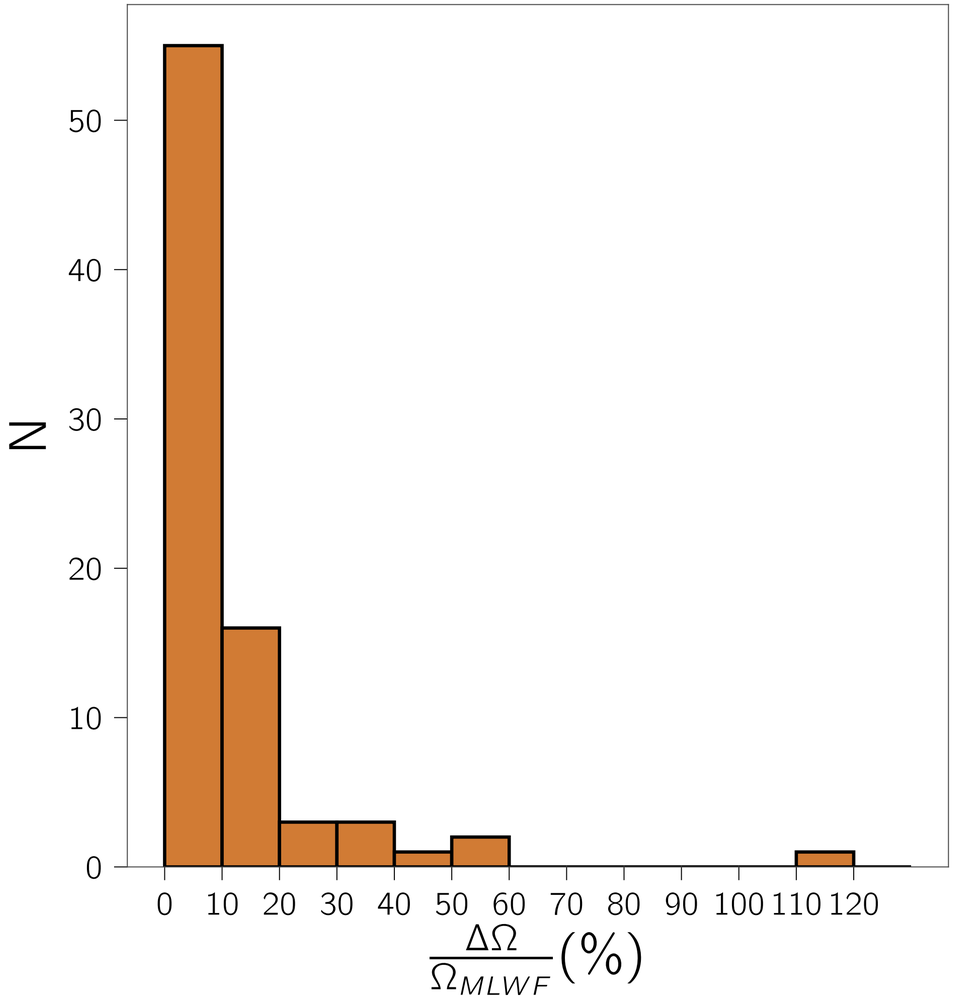
V. Vitale, GP et al., arXiv:1909.00433, accepted in npj Comp. Mat.

Comparison of SCDM + MLWF with SCDM only (80 insulators)

- Improvement of bands distance with MLWF on top of SCDM visible also from this plot



- Spread reduction typically of ~10-20%: SCDM for insulators already is quite localised

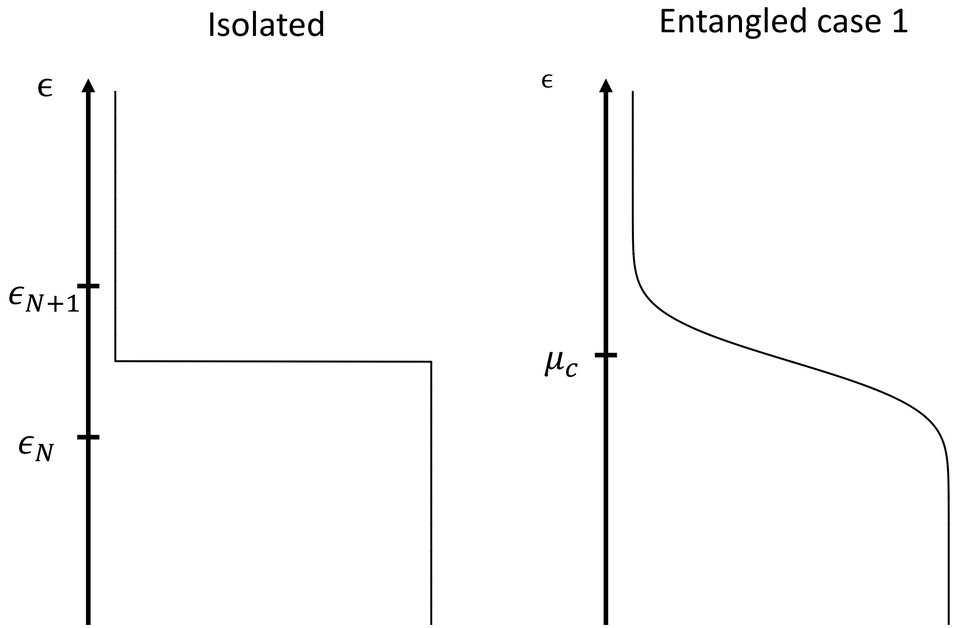


Entangled bands

- We consider (formally) all eigenstates, and give a weight in the *quasi-density-matrix P*

$$P = \sum_i |\psi_i\rangle f(\varepsilon_i) \langle\psi_i| = f(H)$$

- *f:* smooth function of energy, selecting relevant states. If *f* is smooth: *P*(**r**,**r'**) decays rapidly [2]
- We select the most N_w representative columns; procedure is analogous to isolated case



Damle, A. & Lin, L. Disentanglement via entanglement: A unified method for Wannier localization. *Multiscale Modeling & Simulation* 16, 1392–1410 (2018).

Entangled bands

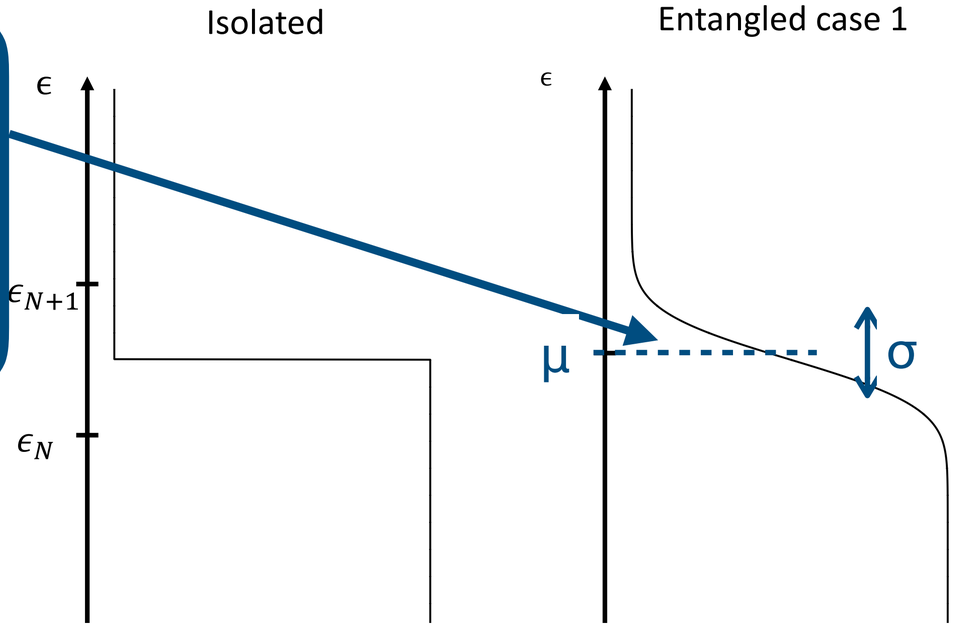
- We consider (formally) all eigenstates, and give a weight in the *quasi-density-matrix P*

$$P = \sum_i |\psi_i\rangle f(\varepsilon_i) \langle\psi_i| = f(H)$$

- *f:* smooth function of energy, selecting relevant states. If *f* is smooth: *P*(**r**,**r'**) decays rapidly [2]
- We select the most N_w representative columns; procedure is analogous to isolated case

$$f(\varepsilon) = \frac{1}{2}\mathrm{erfc}\left(\frac{\varepsilon - \mu}{\sigma}\right)$$

Arbitrary parameters to choose: μ and σ
(and *N*, the number of Wannier functions)

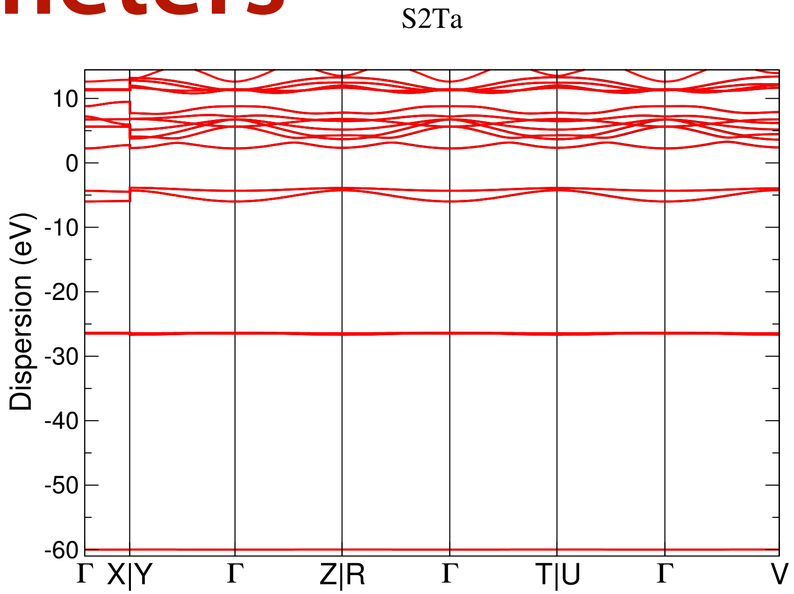


Damle, A. & Lin, L. Disentanglement via entanglement: A unified method for Wannier localization. *Multiscale Modeling & Simulation* 16, 1392–1410 (2018).

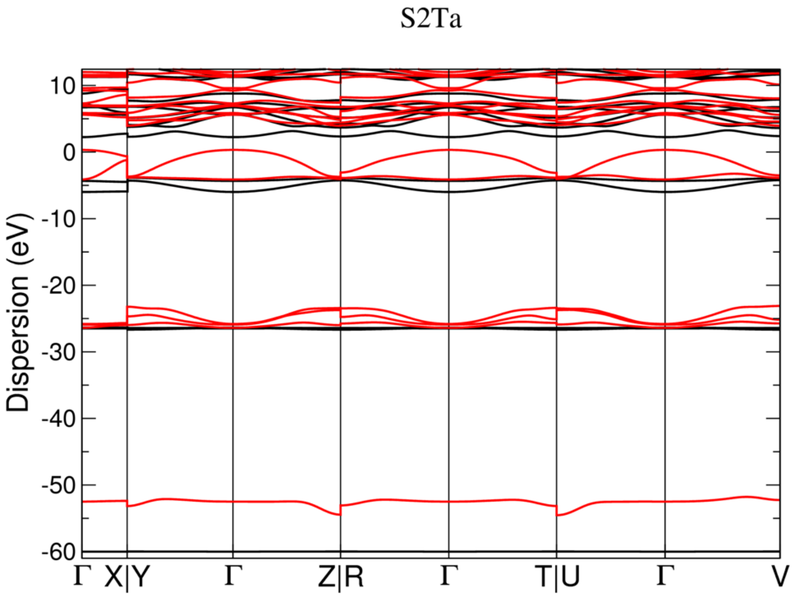
How to choose the parameters

- The SCDM method does not suggest how to choose the ***μ*** *and* ***σ*** *parameters* (and *neither the number* ***N*** *of Wannier functions*)
- The choice cannot be arbitrary: “bad” values generate bad interpolations
- **μ too small**: not enough information on high-energy bands: QRCP will pick top states randomly
- **μ too large**: high-energy states (that we are not interested into) might have a large weight and QRCP might prefer to select them: interpolation tends to have higher energy than the actual bands

How to choose these parameters (automatically)?



“Correct” μ: excellent interpolation



Too large μ: bad interpolation

Important ingredient: Projectability

$$p(|\psi_{n,\mathbf{k}}\rangle) = \sum_i |\langle o_i|\psi_{n,\mathbf{k}}\rangle|^2$$

atomic orbitals o_i described in the pseudopotential

- For each band (*n,k)*, it is the **projection** of that state on *all the pseudo-atomic orbitals described in the pseudopotential file*
- Easy to obtain from Quantum ESPRESSO's *projwfc.x*

Copper

Orbitals: s,p,d (no nodes) + s,p (1 node)

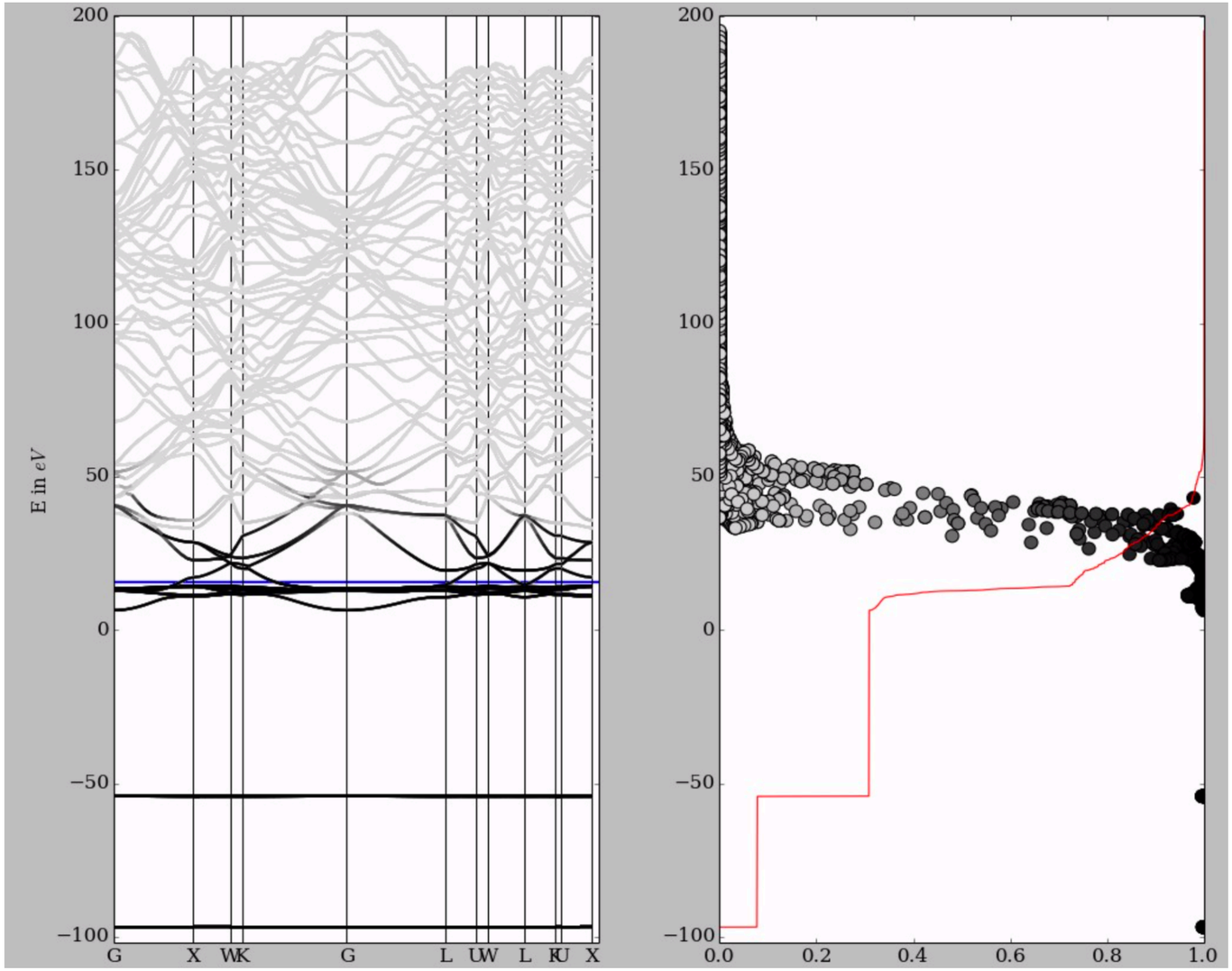


Image: courtesy of Daniel Marchand, EPFL

$$p(|\psi_{n,\mathbf{k}}\rangle) = \sum_i |\langle o_i|\psi_{n,\mathbf{k}}\rangle|^2$$

How to choose the parameters: Our recipe

- We aim at getting a good band interpolation for the low-lying bands
- 1: choose **N as the number of atomic orbitals for which we have information in the pseudopotential file** (see also *Agapito et al., PRB 88, 165127 (2013)*)

- 2: compute the "projectability" of each state as the projection of each state on the subspace of the atomic orbitals o_i described in the pseudopotential:

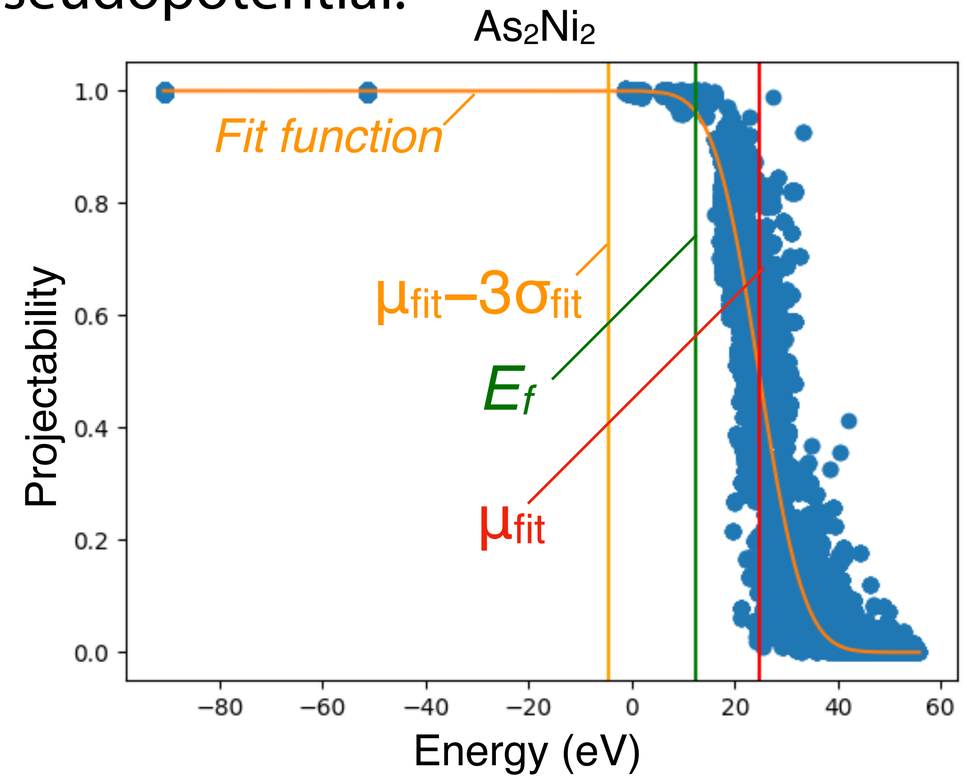
$$p(|\psi_{n,\mathbf{k}}\rangle) = \sum_i |\langle o_i|\psi_{n,\mathbf{k}}\rangle|^2$$

- 3: Fit the plot of the projectability vs. energy with

$$f(\varepsilon) = \frac{1}{2}\mathrm{erfc}\left(\frac{\varepsilon - \mu_{\mathrm{fit}}}{\sigma_{\mathrm{fit}}}\right)$$

- **4: choose the parameters μ and σ as follows**

$$\mu = \mu_{\mathrm{fit}} - 3\sigma_{\mathrm{fit}}; \qquad \sigma = \sigma_{\mathrm{fit}}$$



V. Vitale, GP et al., arXiv:1909.00433, accepted in npj Comp. Mat.

How to choose the parameters: Our recipe

- We aim at getting a good band interpolation for the low-lying bands
- 1: choose **N as the number of atomic orbitals for which we have information in the pseudopotential file** (see also *Agapito et al., PRB 88, 165127 (2013)*)

N: we select the **most relevant**, low-lying bands, and we **can compute the projectability**

- as the projection of each state on the in the pseudopotential:

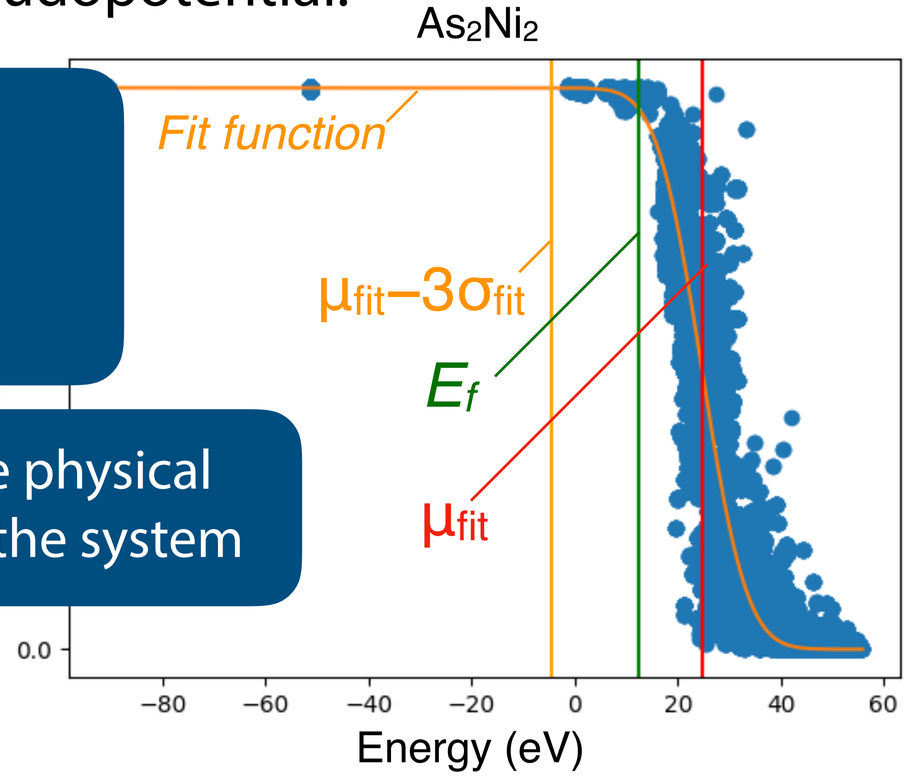
μ: when projectability≲0.9, weight≲10^{-3}

need to exponentially **suppress high-energy states** to affect SCDM choice

- $f(\epsilon) = \frac{1}{2}\mathrm{erfc}\left(\frac{\epsilon - \mu_{\mathrm{fit}}}{\sigma_{\mathrm{fit}}}\right)$
- **4: choose the parameters μ and σ**

σ: we select the physical "band width" of the system

$$\mu = \mu_{\mathrm{fit}} - 3\sigma_{\mathrm{fit}}; \qquad \sigma = \sigma_{\mathrm{fit}}$$



V. Vitale, GP et al., arXiv:1909.00433, accepted in npj Comp. Mat.

How to choose the parameters: Our recipe

- We aim at getting a good band interpolation for the low-lyir
- 1: choose **N as the number of atomic orbitals for which w pseudopotential file** (see also *Agapito et al., PRB 88, 165127 (2013)*)

N: we select the **most relevant**, low-lying bands, and we **can compute the projectability**

as the projectio in the pseudop

- **μ**: when projectability≲0.9, weight≲10^{-3}

need to exponentially **suppress high-energy states** to affect SCDM choice

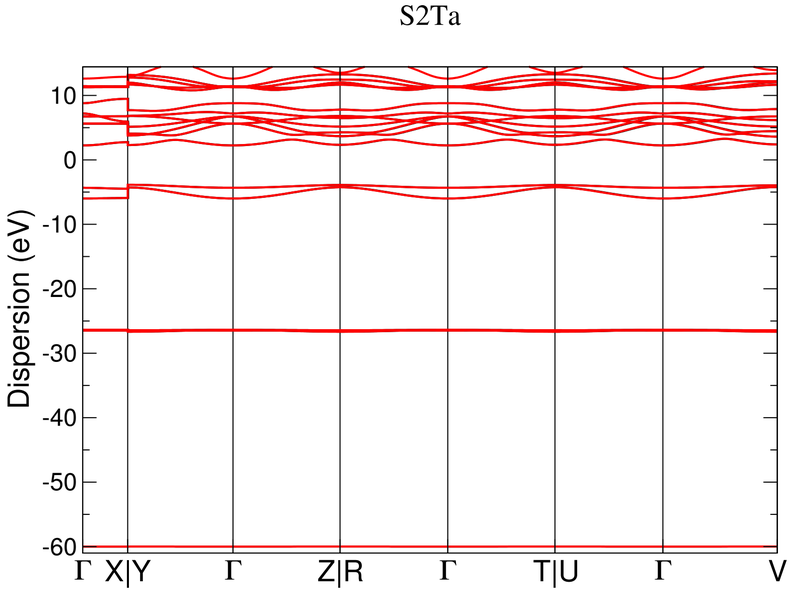
$$f(\epsilon) = \frac{1}{2}\mathrm{erfc}\left(\frac{\epsilon - \mu_{\mathrm{fit}}}{\sigma_{\mathrm{fit}}}\right)$$

- **4: choose the parameters μ and σ**

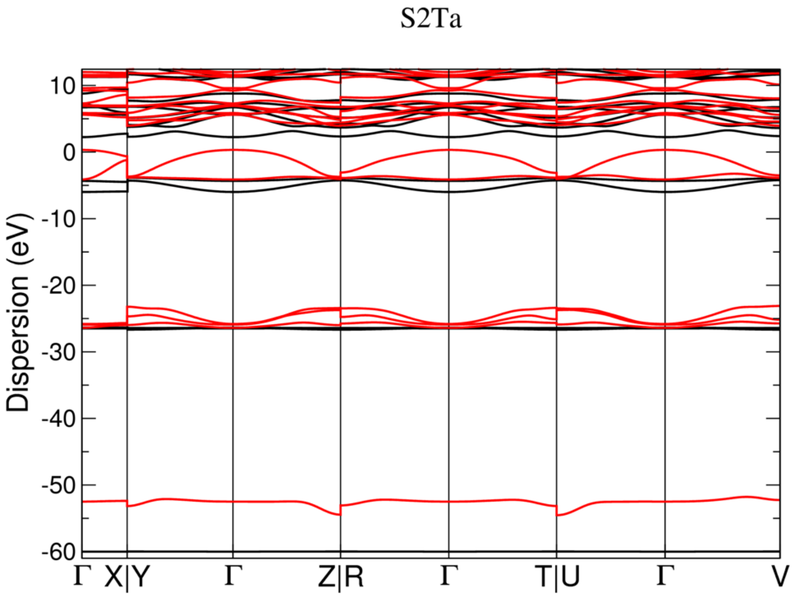
σ: we select the physi "band width" of the sys

$$\mu = \mu_{\mathrm{fit}} - 3\sigma_{\mathrm{fit}}; \qquad \sigma = \sigma_{\mathrm{fit}}$$

V. Vitale, GP et al., arXiv:1909.00433, accepted in npj Comp. Mat.



"Correct" μ: excellent interpolation



Too large μ: bad interpolation

Bands distance for metals

- For metals, we introduce a smearing function, to consider bands only up to a given energy:

$$\eta = \sqrt{\frac{\sum_{n\mathbf{k}} \left(\varepsilon^{\mathrm{DFT}}_{n\mathbf{k}} - \varepsilon^{\mathrm{Wan}}_{n\mathbf{k}}\right)^2 \tilde{f}_{n\mathbf{k}}}{\sum_{n\mathbf{k}} \tilde{f}_{n\mathbf{k}}}},$$

Average bands distance

$$\eta^{\max} = \max_{n\mathbf{k}} \left(\tilde{f}_{n\mathbf{k}} \left| \varepsilon^{\mathrm{DFT}}_{n\mathbf{k}} - \varepsilon^{\mathrm{Wan}}_{n\mathbf{k}} \right| \right)$$

Max bands distance

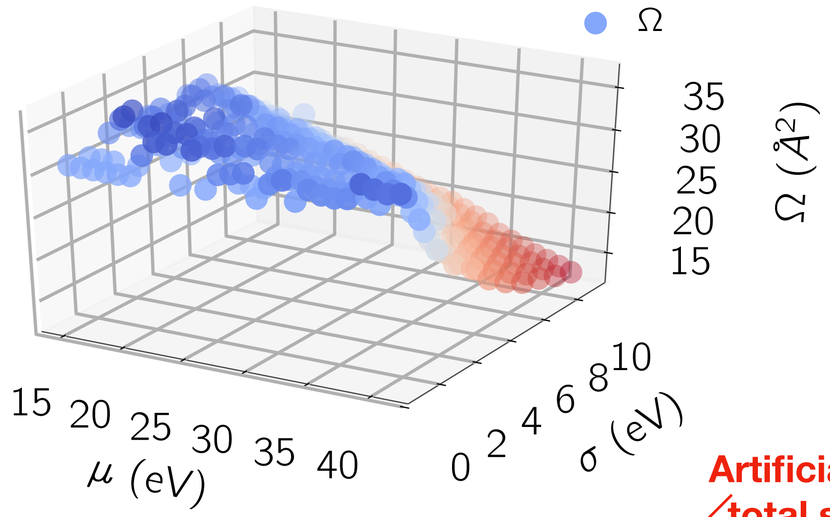
- Smearing function: $\tilde{f}_{n\mathbf{k}} = \sqrt{f^{\mathrm{DFT}}_{n\mathbf{k}}(\nu,\tau) f^{\mathrm{Wan}}_{n\mathbf{k}}(\nu,\tau)}$

- Typically, the "fake" chemical potential **ν** is set e.g. 1eV above the Fermi level (and we choose **τ** = 0.1eV)

Parameter choice validation: tungsten (W)

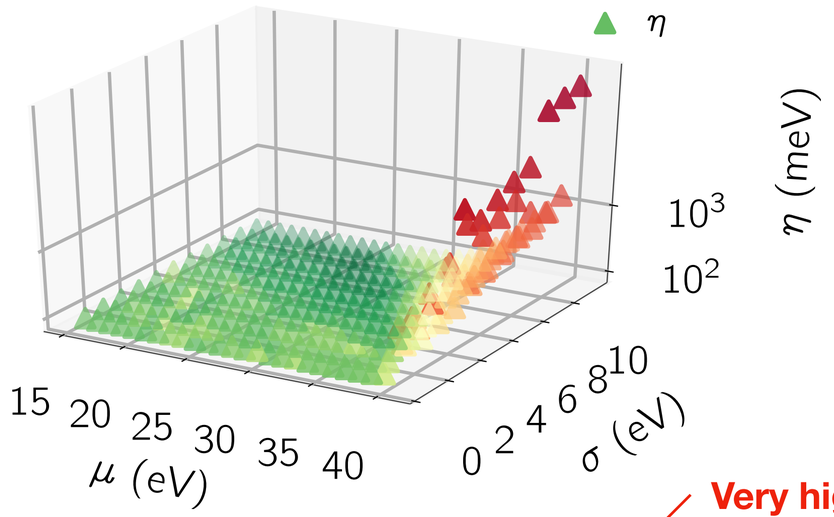
Spread

as a function of μ and σ



Bands distance

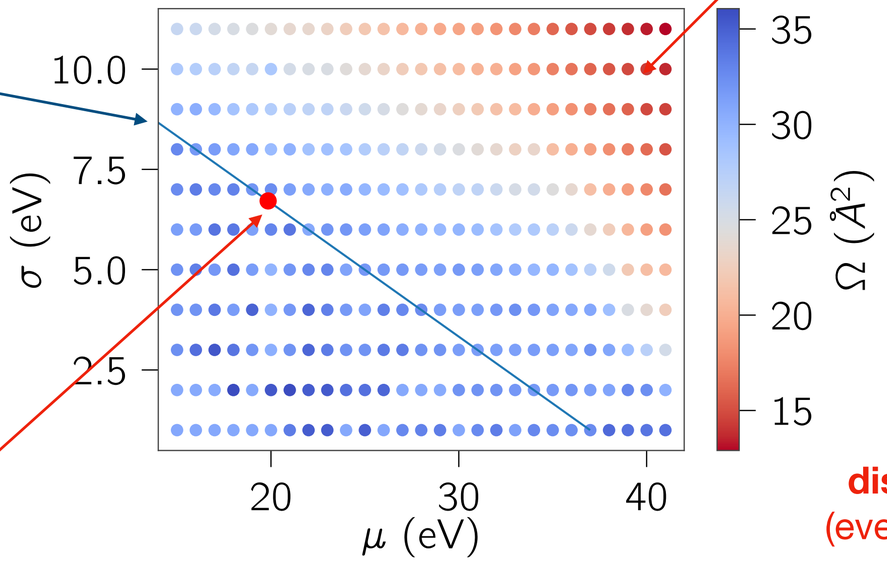
as a function of μ and σ



Artificially low total spread

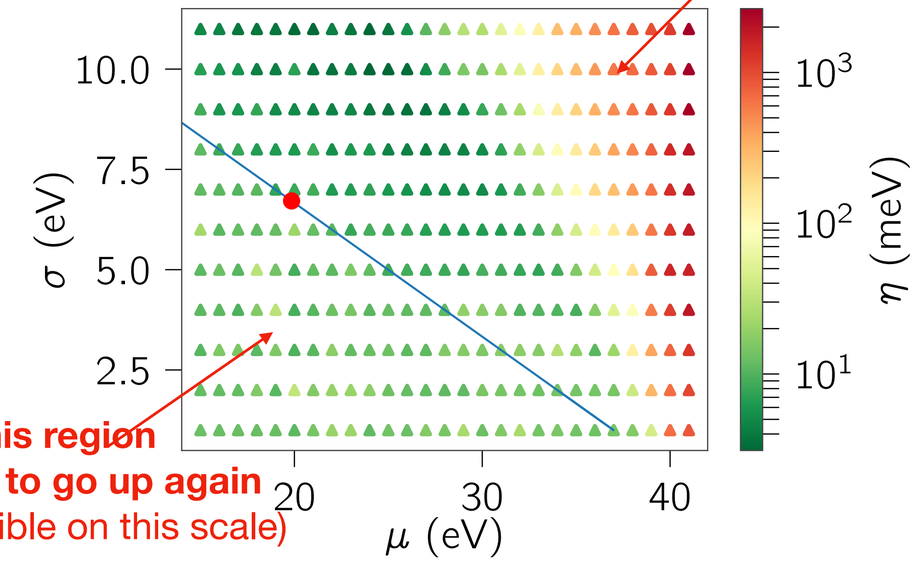
Very high bands distance

$\mu = \mu_{fit} - 3\sigma$



Our protocol:

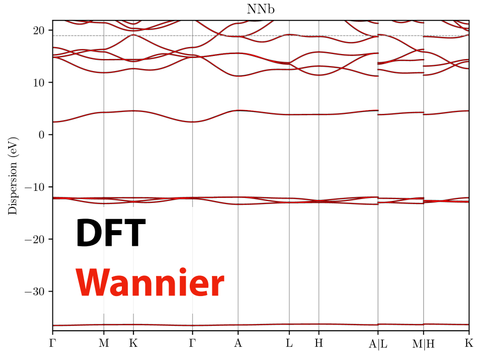
$\mu = \mu_{fit} - 3\sigma_{fit}$

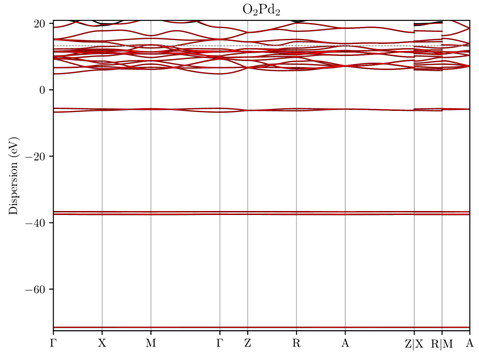


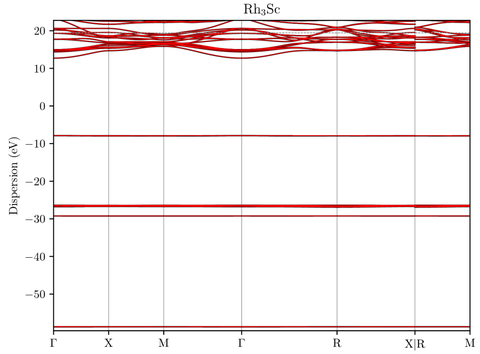
Also in this region distance starts to go up again

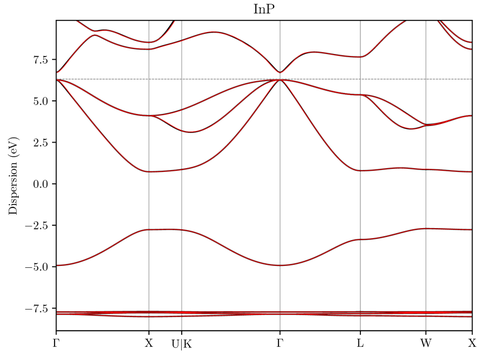
(even if not so visible on this scale)

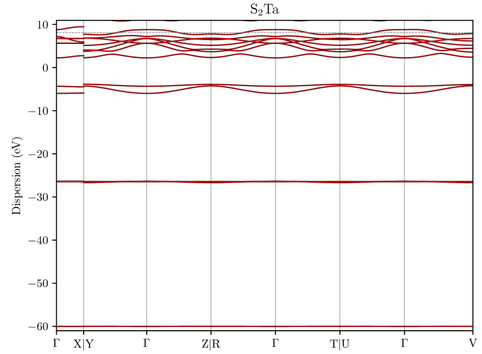
Entangled bands: (subset of the) results

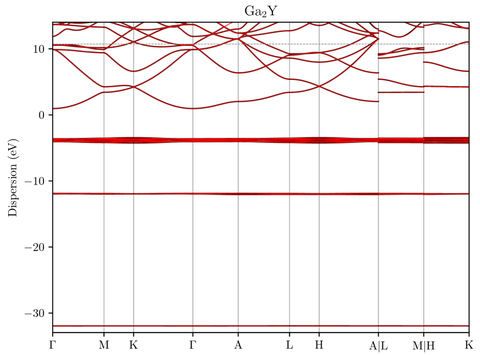


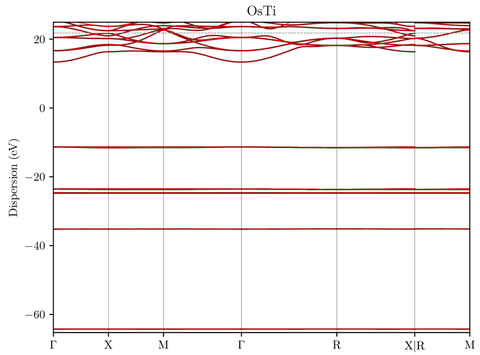


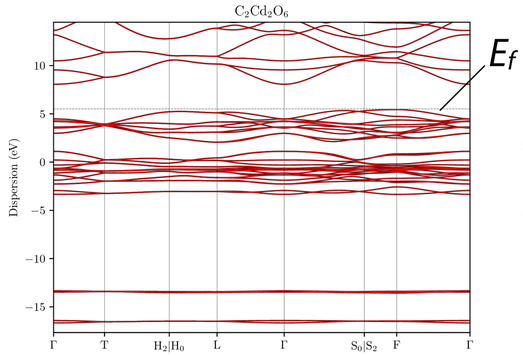


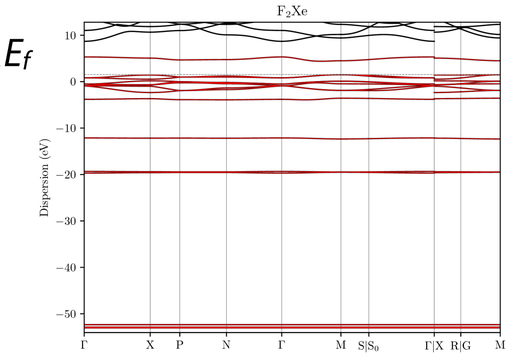


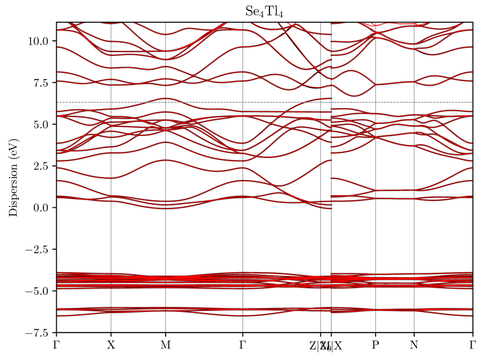


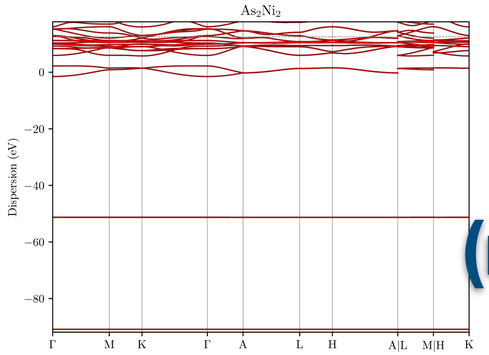


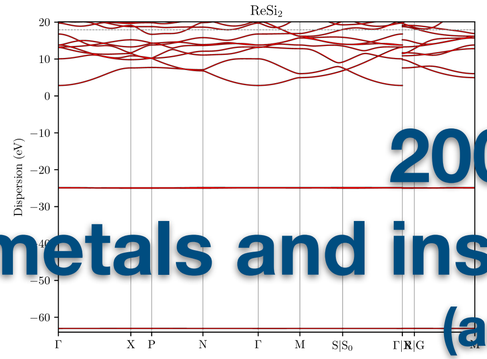


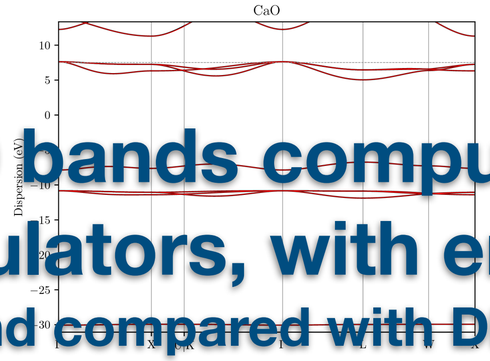


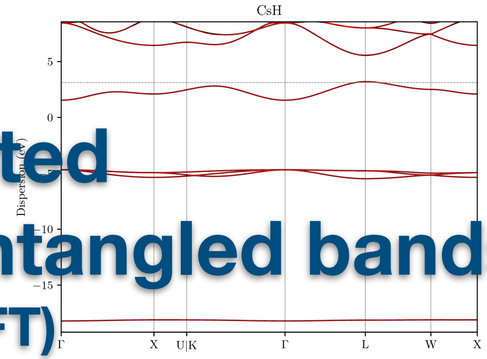


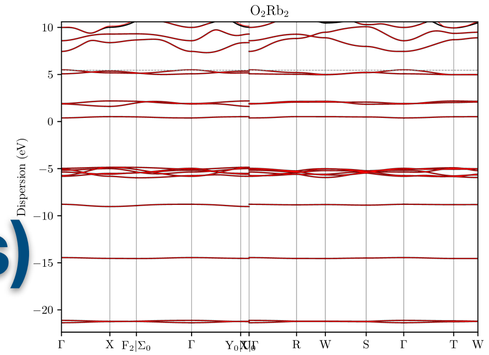






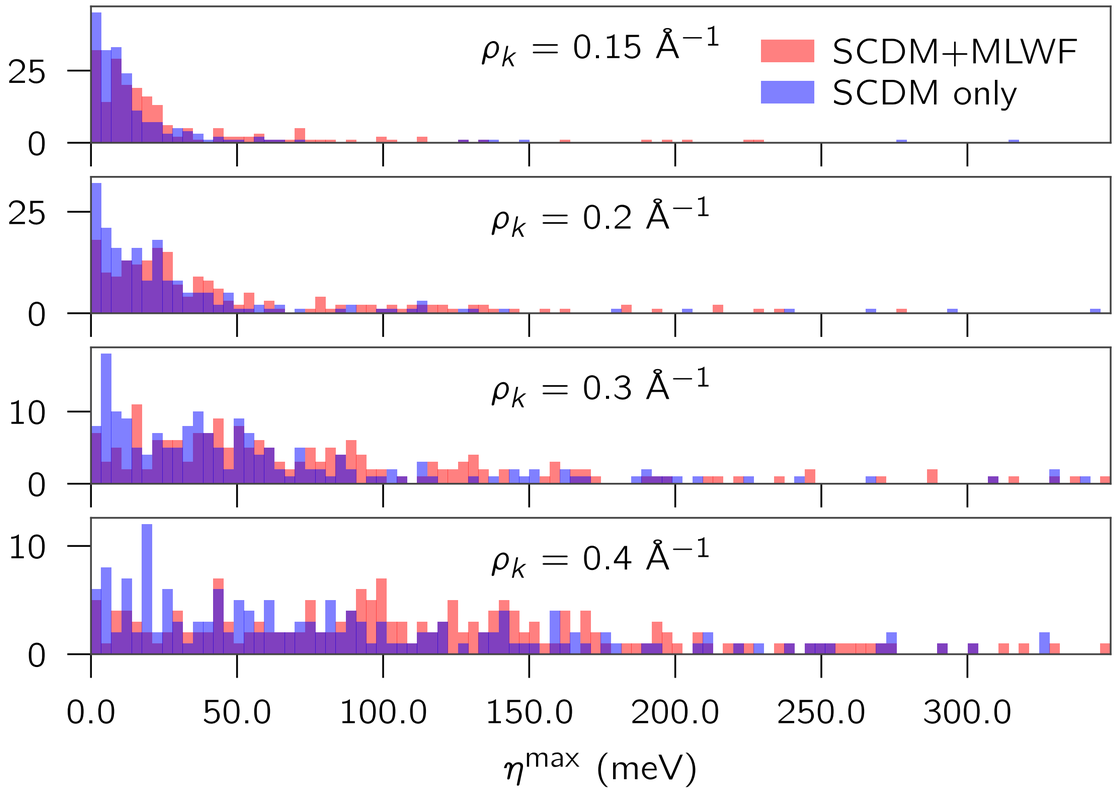


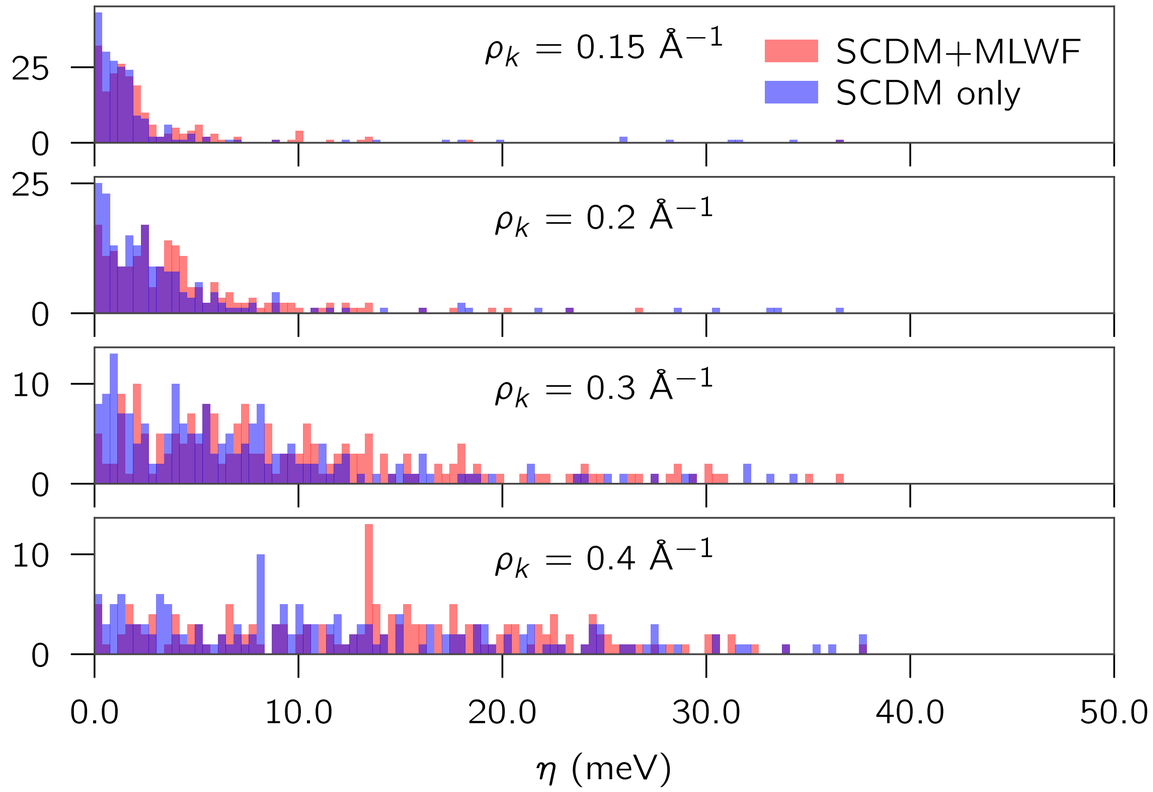




200 bands computed
(metals and insulators, with entangled bands)
(and compared with DFT)

Comparison of SCDM + MLWF with SCDM only (200 materials with entangled bands)



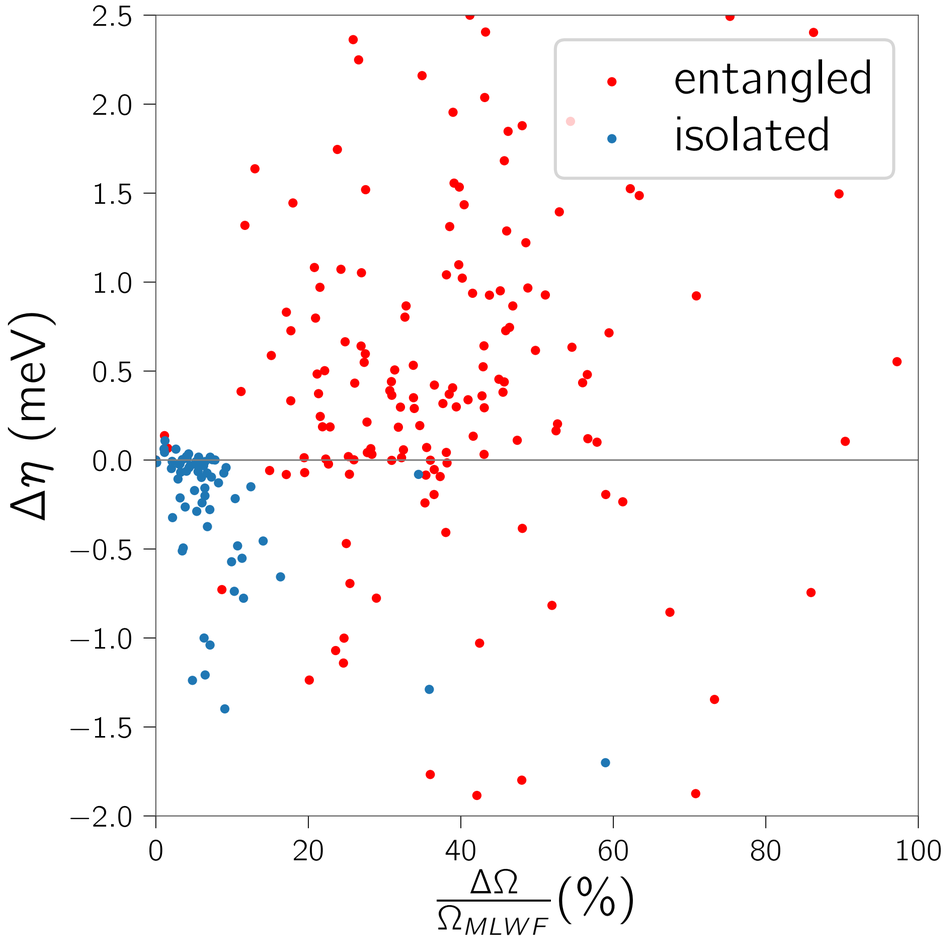


- Good results require a density of at least 0.2 Å^{-1} or more dense
- **For insulators**, SCDM-only already provides very good results; MLWF improves them
- In general, very small band distances (i.e. very good interpolation)

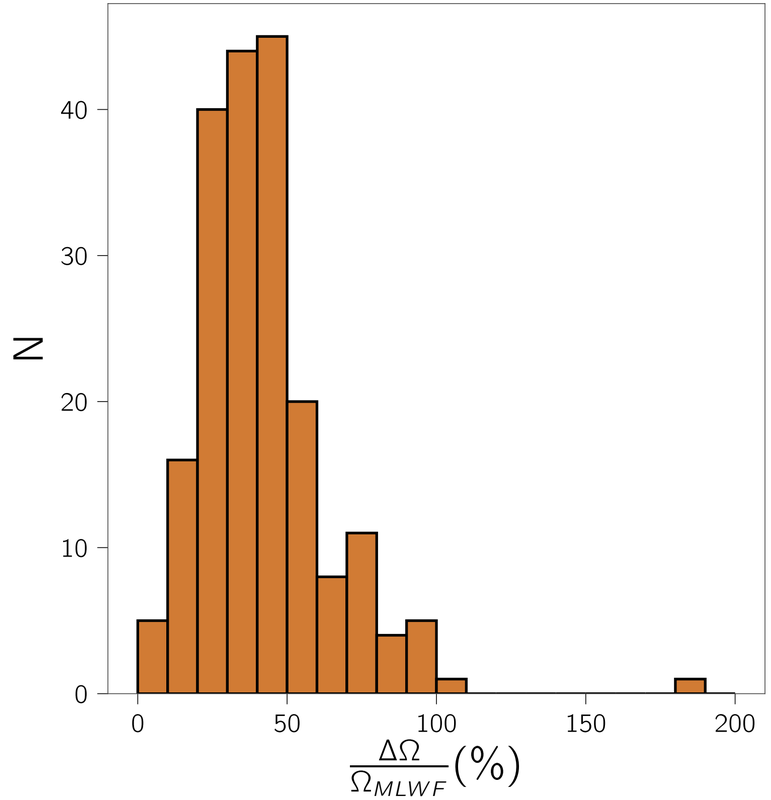
V. Vitale, GP et al., arXiv:1909.00433, accepted in npj Comp. Mat.

Comparison of SCDM + MLWF with SCDM only (200 materials with entangled bands)

- In many cases, running a MLWF in the entangled case worsens the bands distance (even if only by a little)



- Spread reduction is relatively large, often 50% or more
- SCDM with entangled bands is not *maximally* localised



Conclusions (part 1)

V. Vitale et al., arXiv:1909.00433,
accepted in npj Comp. Mat.

- **SCDM method** now implemented in Quantum ESPRESSO (since QE 6.3)
- We have **defined an algorithm** for the choice of "free" parameters in SCDM (*N*, μ, σ)
- We have **implemented automatic workflows** within AiiDA for fully automatic Wannierisation
- **We have validated** the SCDM method and our algorithm on a set of 200 crystalline materials (~80 insulators) covering the chemical space

Acknowledgements

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

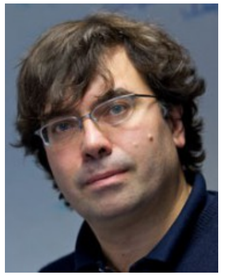
MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH

THEOS
THEORY AND SIMULATION OF MATERIALS

MAX DRIVING THE EXASCALE TRANSITION


Valerio Vitale


Antimo Marrazzo


Nicola Marzari


Jonathan R. Yates


Arash A. Mostofi

Open Science Platform for Materials Science: AiiDA and the Materials Cloud

Giovanni Pizzi (EPFL)

AiiDA

http://www.aiida.net

G. Pizzi et al., Comp. Mat. Sci. 111, 218-230 (2016)

MATERIALSCLOUD

http://www.materialscloud.org

FNSNF

MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH

MAX DRIVING THE EXASCALE TRANSITION

swissuniversities

EPFL

Our two core infrastructures

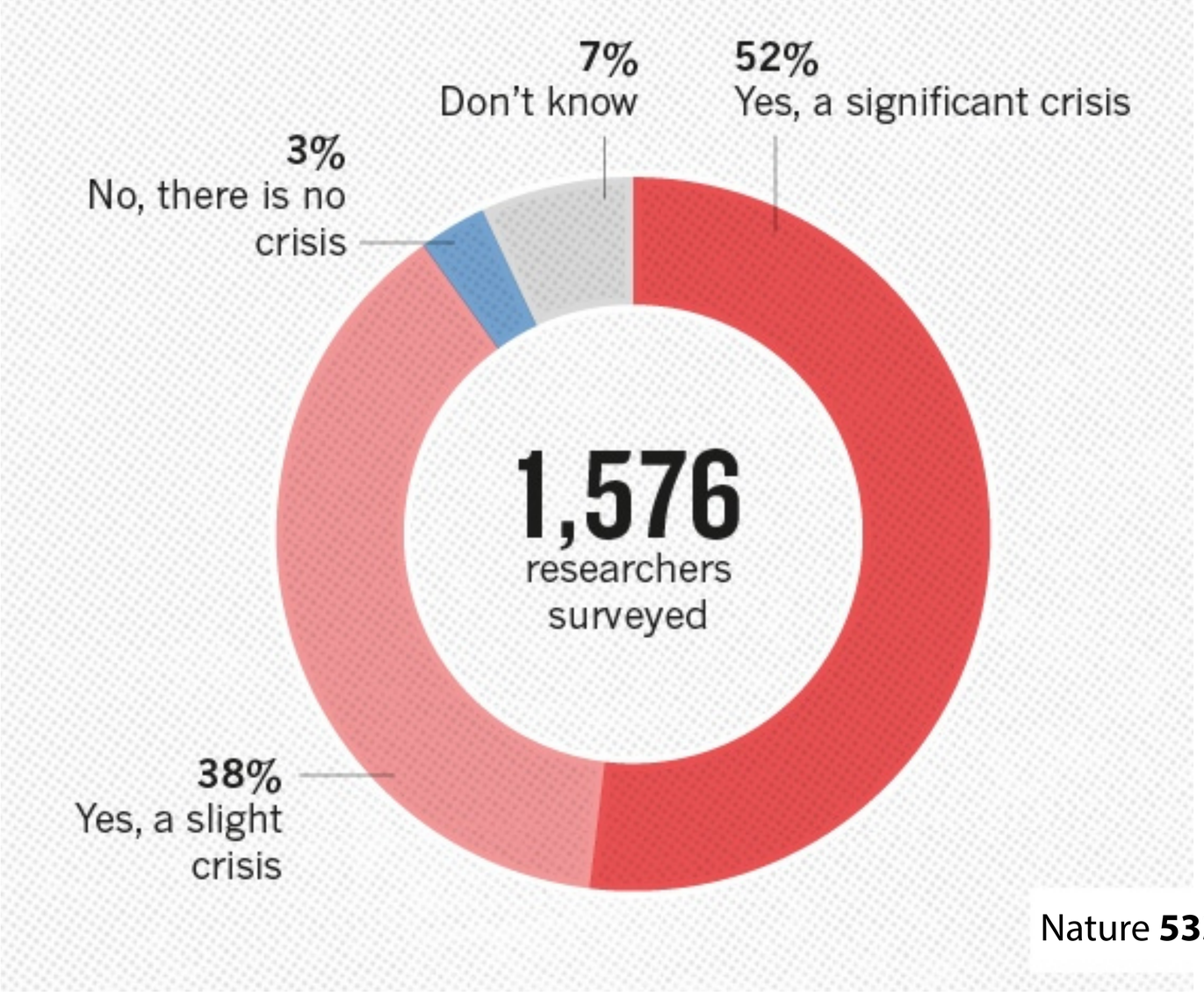
AiiDA as the "operating system" to manage, automate and store simulations and their results

and

Materials Cloud as the open-science dissemination portal and cloud simulation platform

Reproducibility: a cornerstone of the scientific method

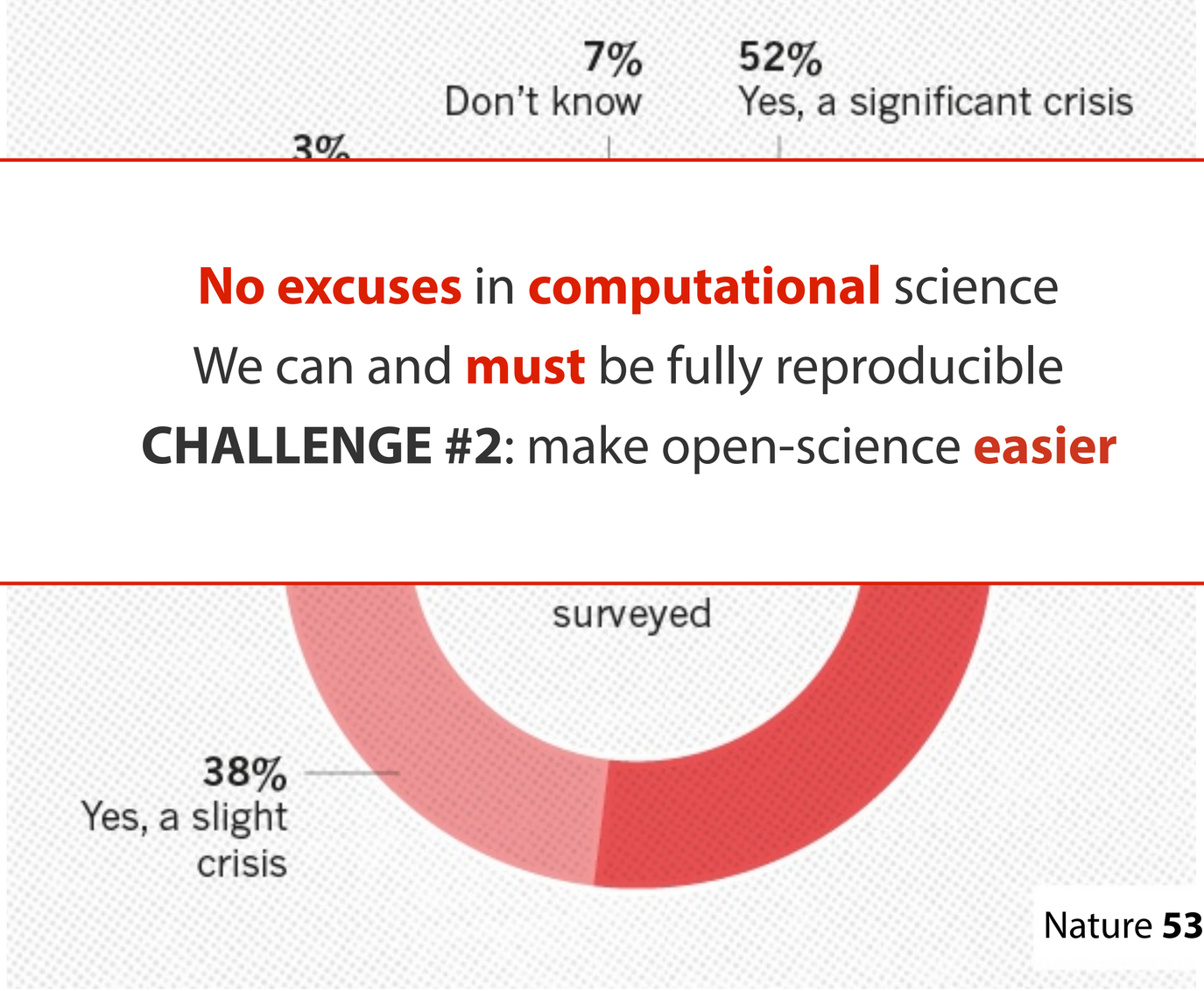
IS THERE A REPRODUCIBILITY CRISIS?



Nature **533**, 452–454 (2016)

Reproducibility: a cornerstone of the scientific method

IS THERE A REPRODUCIBILITY CRISIS?



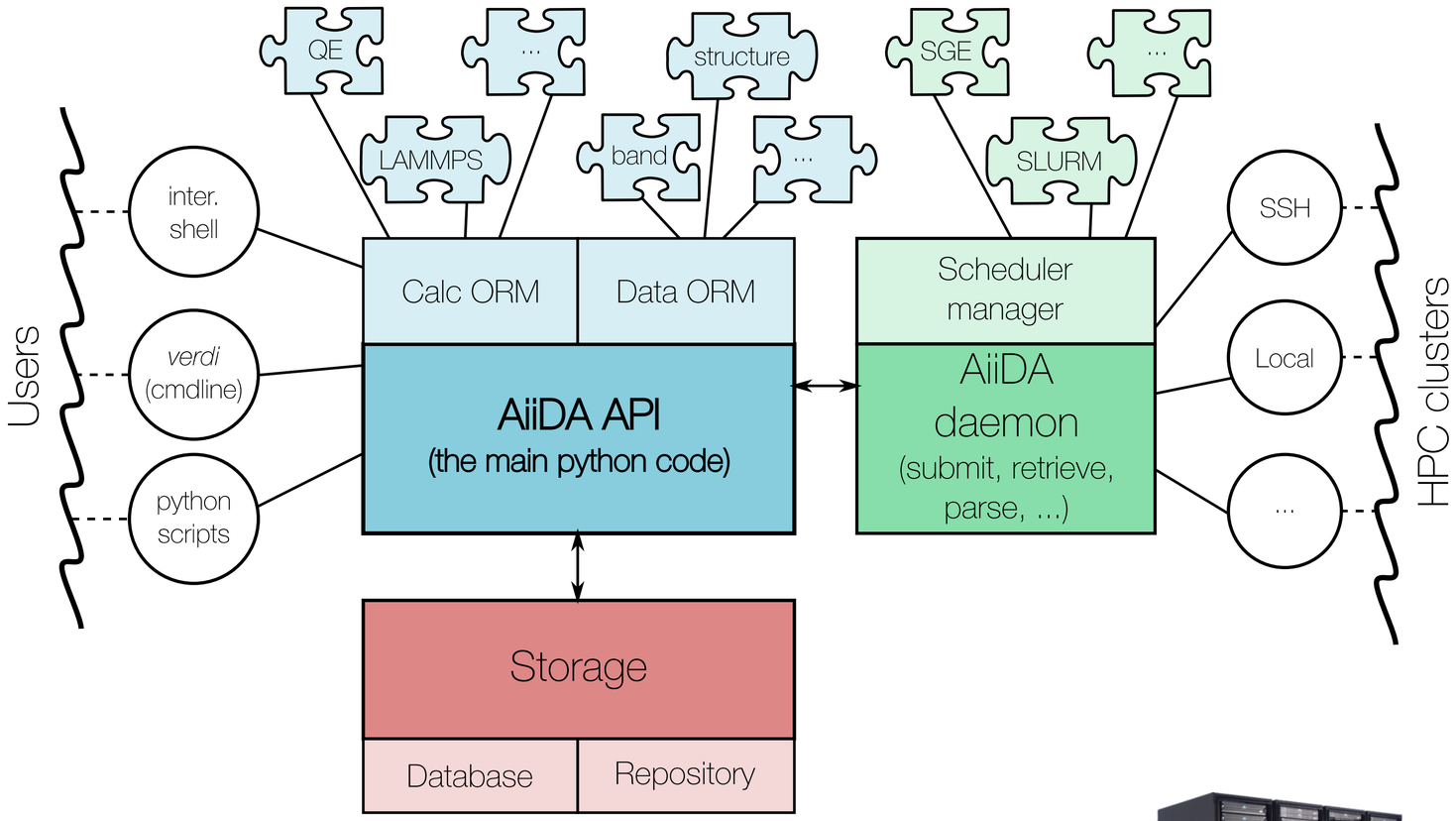
No excuses in **computational** science

We can and **must** be fully reproducible

CHALLENGE #2: make open-science **easier**

Nature **533**, 452–454 (2016)

AiiDA



AiiDA

Main features

- Python (2.7 & 3.6) infrastructure
- SQL database backend, access via a Python ORM
- Local connection to clusters, or via ssh using a python API
- Interface to various job schedulers (SGE, Torque, LSF, PBS Pro, SLURM, ...)
- Event-based daemon with remote management and workflow execution manager
- REST APIs using Flask to expose one own's data
- Plugin management system and extended code support
- Easy sharing of the results with other users in the community

Storage and provenance

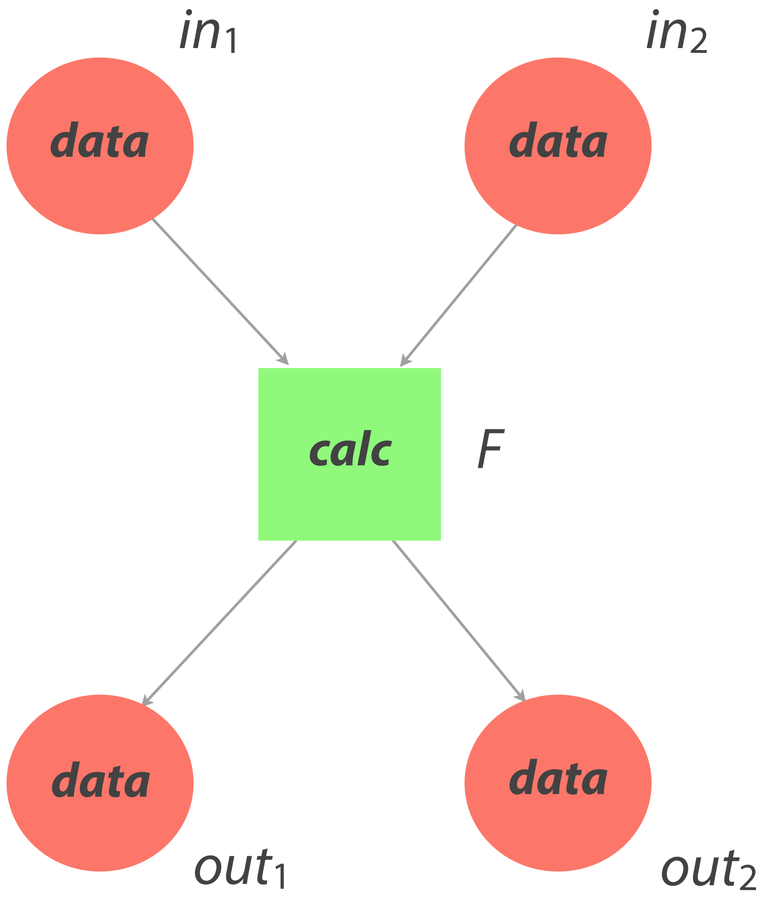
Open Provenance Model

- *Calculated properties*: result of complex, connected calculations
- How do we store simulations **preserving the connected structure?**

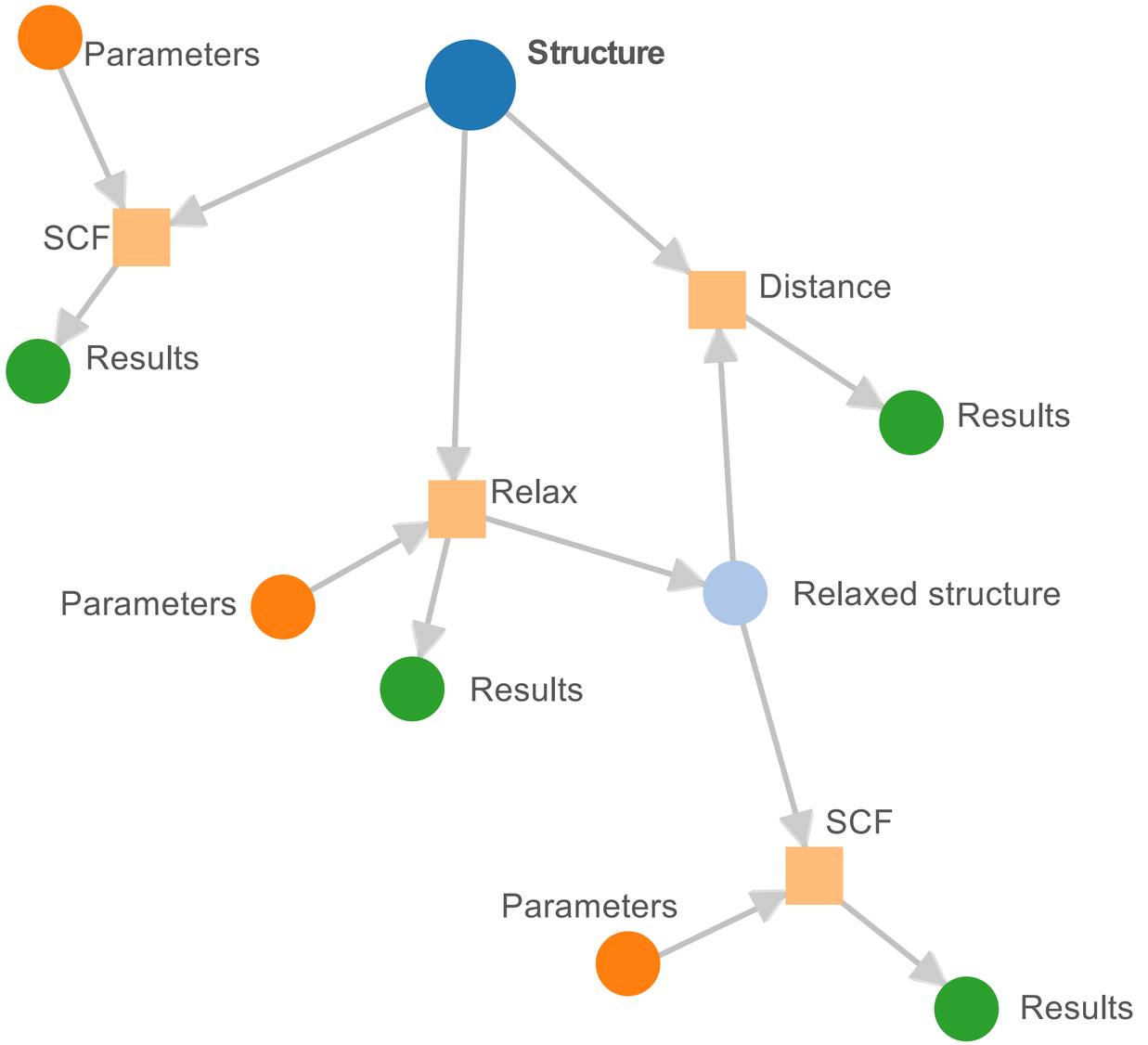
- Inspiration from the *open provenance model*
- **Any calculation: a function,**
  converting inputs to outputs:

  $$out_1, out_2 = F(in_1, in_2)$$

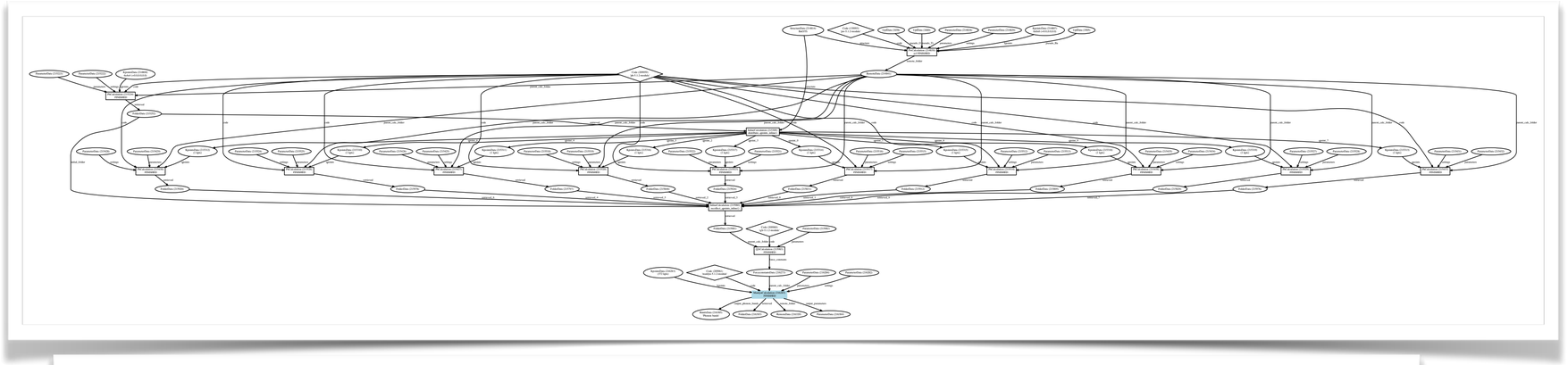
- **Each object is a node in a graph**,
  connected by directional labeled links
- Output nodes can be used as inputs



Data provenance: directed acyclic graphs

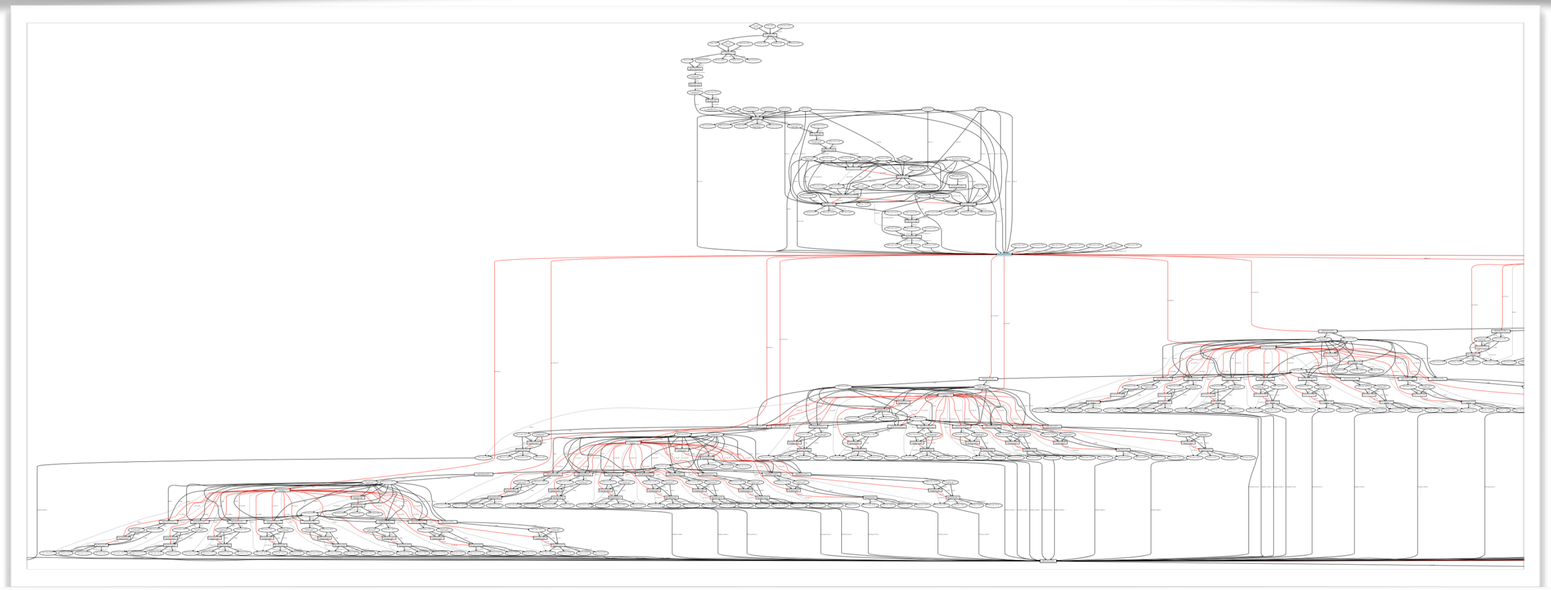


“Simple” graphs of workflows for a single material



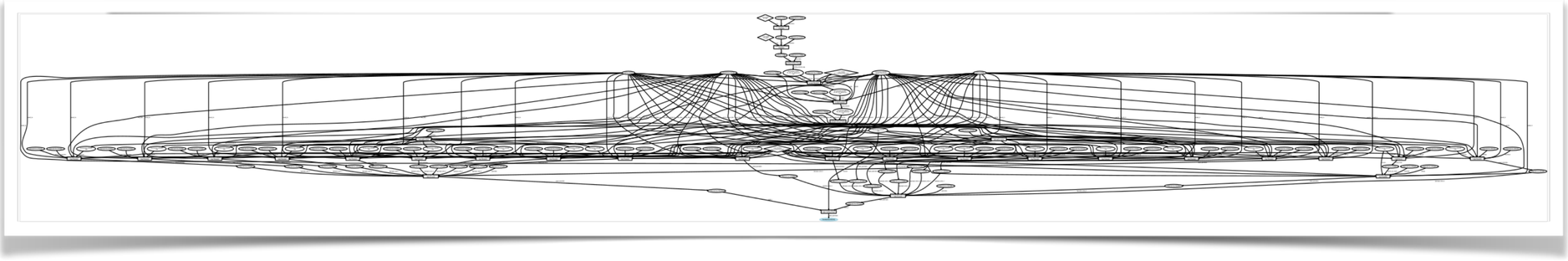
Phonon dispersion

Thermal transport, electronic mobility, …



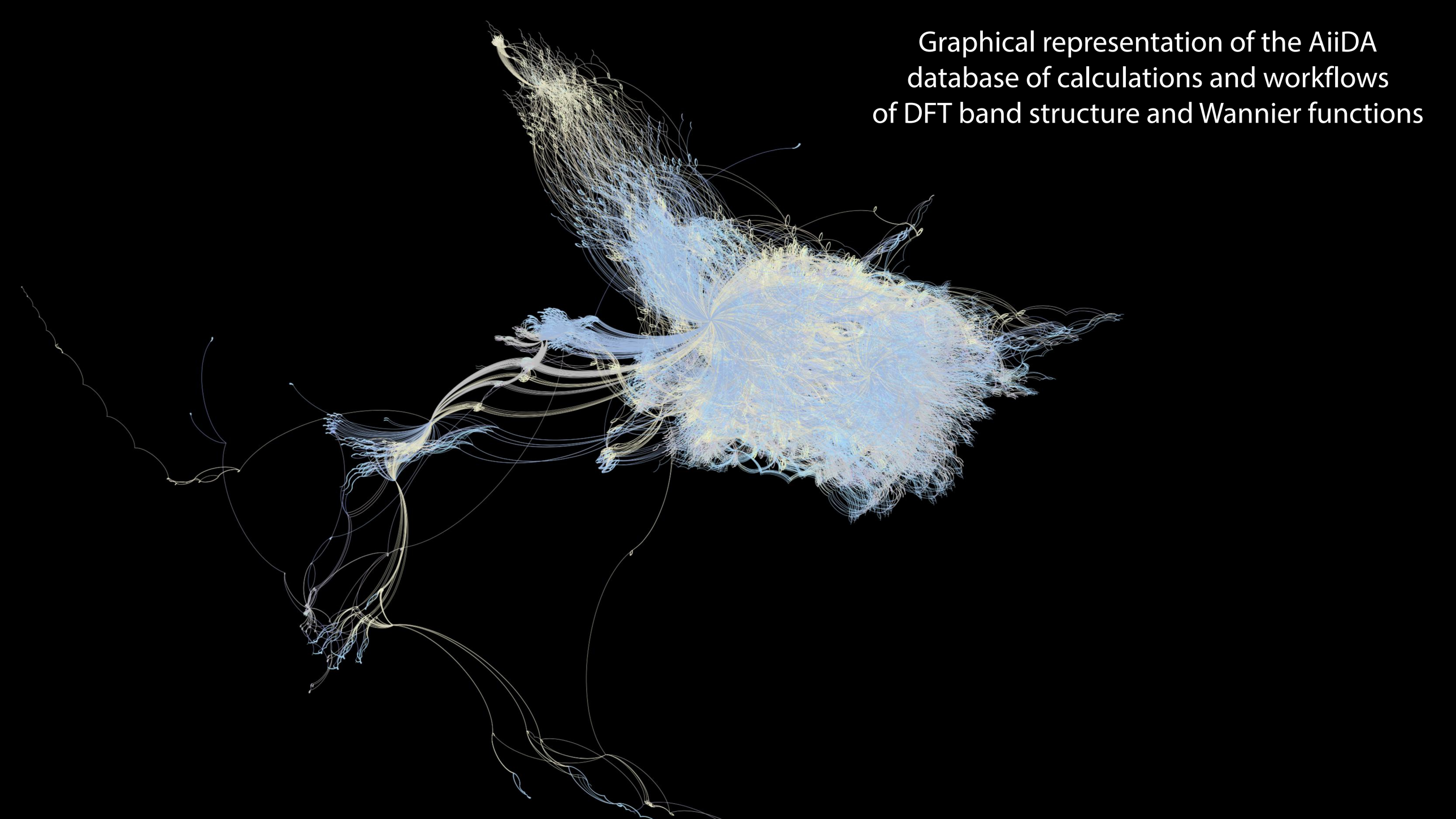
Molecular dynamics of Li ions in a solid-state electrolyte

Novel, safe and efficient Li-batteries

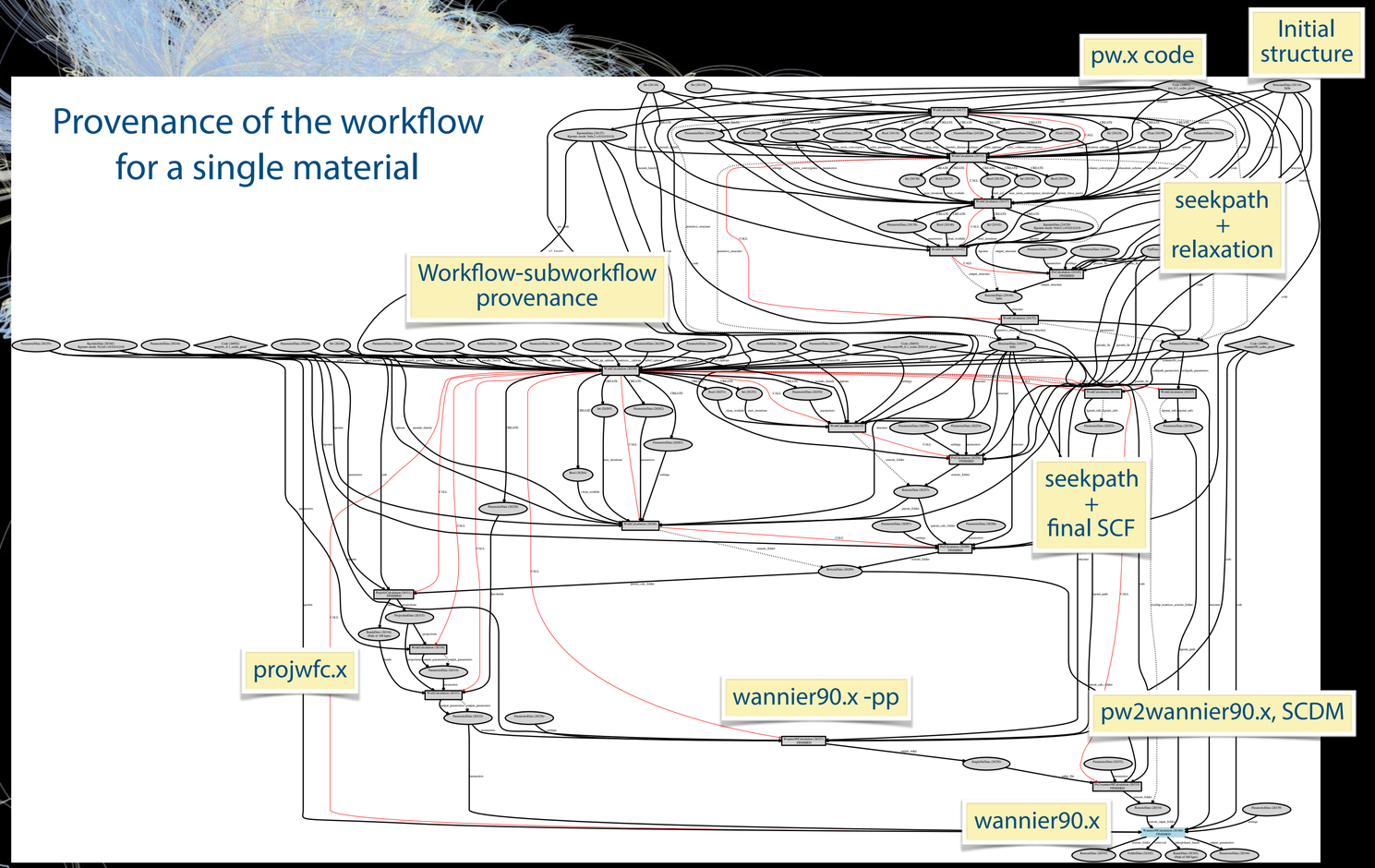


Elastic constants

Response of materials to stresses and deformations

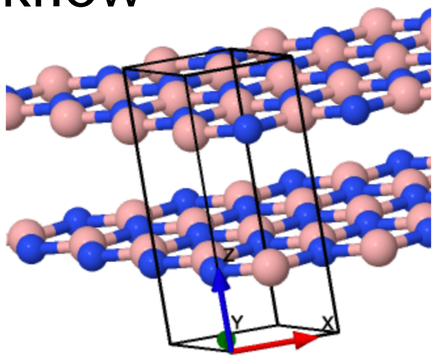

Graphical representation of the AiiDA
database of calculations and workflows
of DFT band structure and Wannier functions

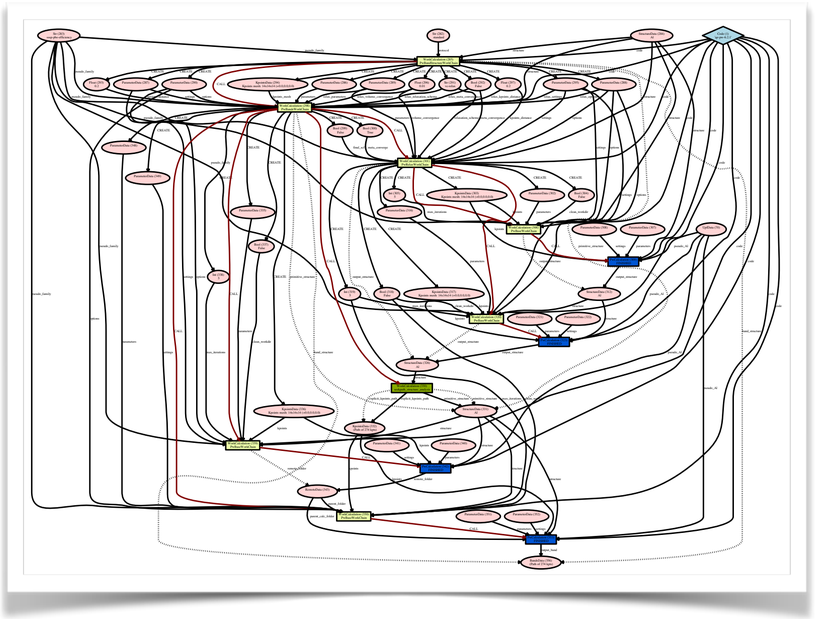
Graphical representation of the AiiDA database of calculations and workflows of DFT band structure and Wannier functions

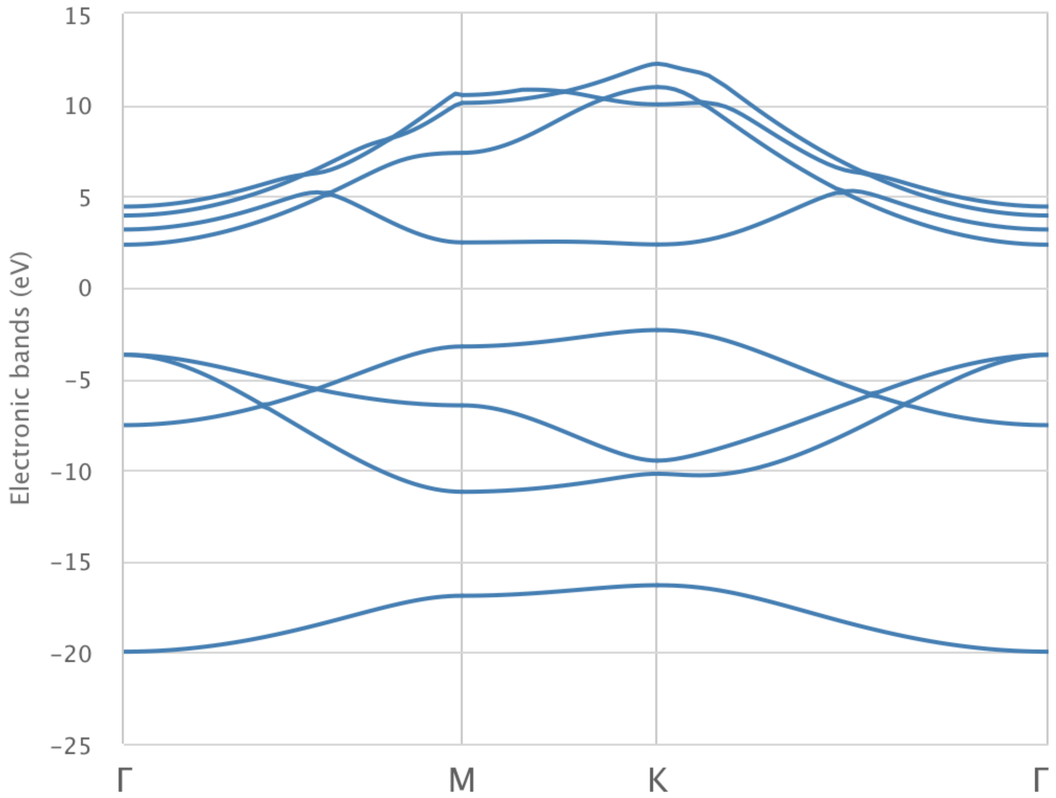


Turn-key workflows in AiiDA

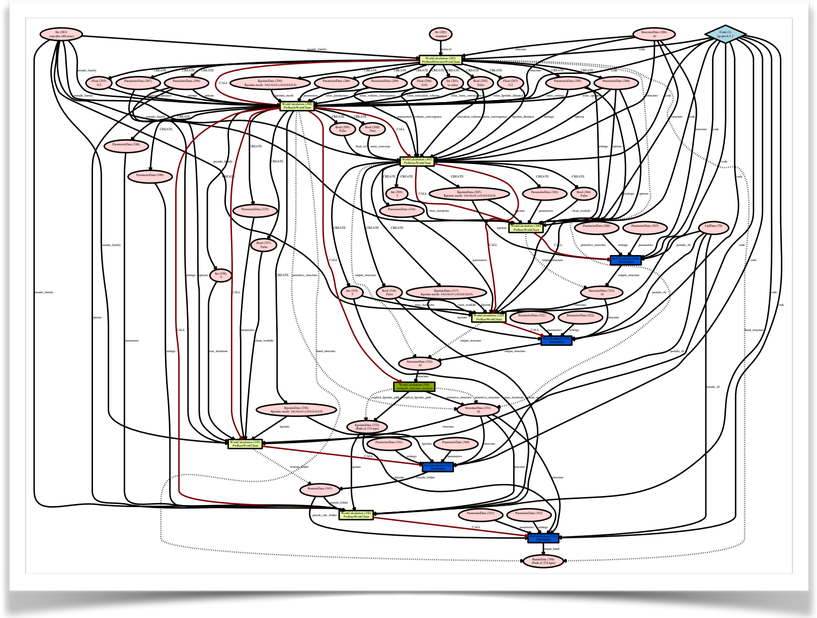
- Given a material, we often need to compute advanced quantities
- These are often non-trivial and result from a complex workflow

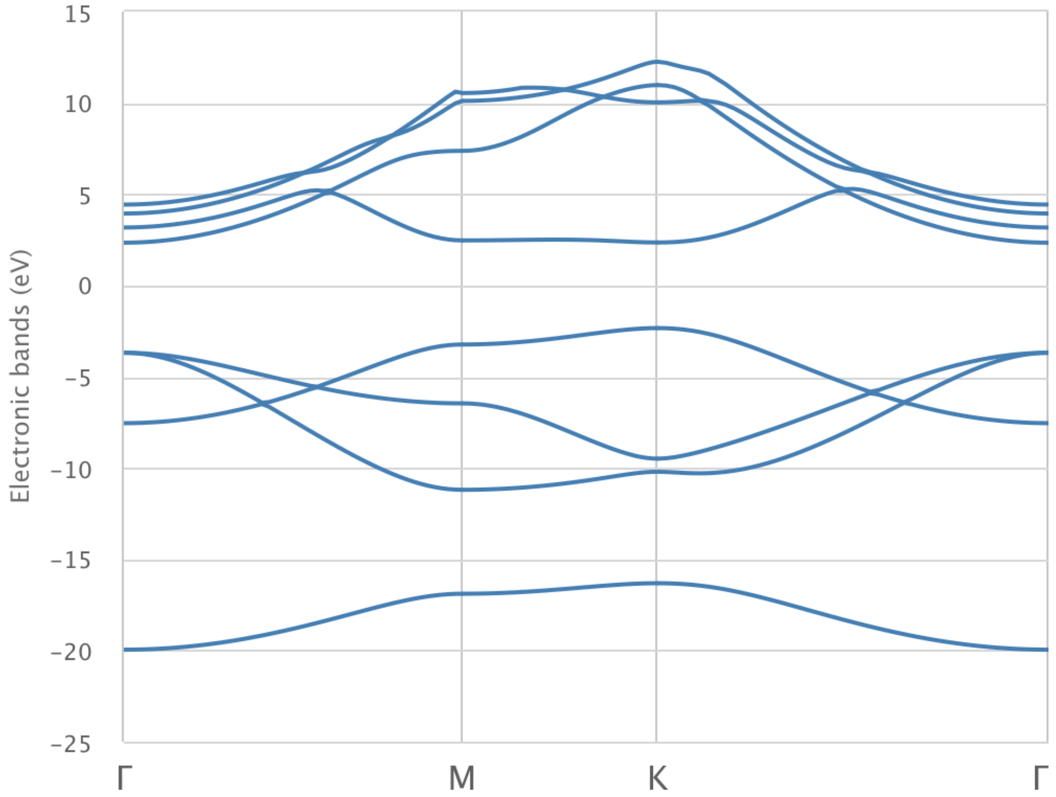






Turn-key workflows in AiiDA





- The AiiDA provenance graph allows to know how the structure was computed and to **reproduce that single specific calculation:** *log of "what happened in the past"*
- We need also an **easy way to re-run the same calculation again** with different parameters or for a different material: **turn-key workflows**

Turn-key workflows in AiiDA

```
class PwBandsWorkChain(WorkChain):
    @classmethod
    def define(cls, spec):

        spec.input('code',
                   valid_type=Code)
        spec.input('structure',
                   valid_type=StructureData)
        spec.input('pseudo_family',
                   valid_type=Str)

        spec.outline(
            cls.setup,
            cls.validate_inputs,
            if_(cls.should_do_relax)(
                cls.run_relax,
            ),
            cls.run_seekpath,
            cls.run_scf,
            cls.run_bands,
            cls.results,
        )
```

- **"Operating system" for all calculations**
- Automatic provenance tracking in the DB
- Control provenance granularity
  store level of detail relevant to the workflows
- Progress checkpointing
  restart from arbitrary step, retry on failure, allows to shut down daemon and continue later
- Easy debugging, self-documenting

Turn-key workflows in AiiDA

```
class PwBandsWorkChain(WorkChain):
    @classmethod
    def define(cls, spec):

        spec.input('code',
                   valid_type=Code)
        spec.input('structure',
                   valid_type=StructureData)
        spec.input('pseudo_family',
                   valid_type=Str)

        spec.outline(
            cls.setup,
            cls.validate_inputs,
            if_(cls.should_do_relax)(
                cls.ru
            ),
            cls.run_s
            cls.run_s
            cls.run_b
            cls.resul
        )
```

- **"Operating system" for all calculations**
- Automatic provenance tracking in the DB
- Control provenance granularity
  store level of detail relevant to the workflows
- Progress checkpointing
  om arbitrary step, retry on
  ows to shut down daemon and
  ater
  ugging, self-documenting

- Turn-key solution:

```
PwBandStructureWorkChain.run(
    code=Code.get_from_string(
        'qe-pw-6.2.1@localhost'),
    structure=StructureData(
        ase=ase.build.bulk('Al')),
     pseudo_family=Str('sssp-pbe-efficiency'))
```

Simple provenance tracking also for pre-/post-processing

```
def add(a, b):
    return a+b


def multiply(a, b):
    return a*b


def add_multiply(a, b, c):
    return multiply(add(a, b), c)


final_value = add_multiply(
    a=3,
    b=4,
    c=5)
```

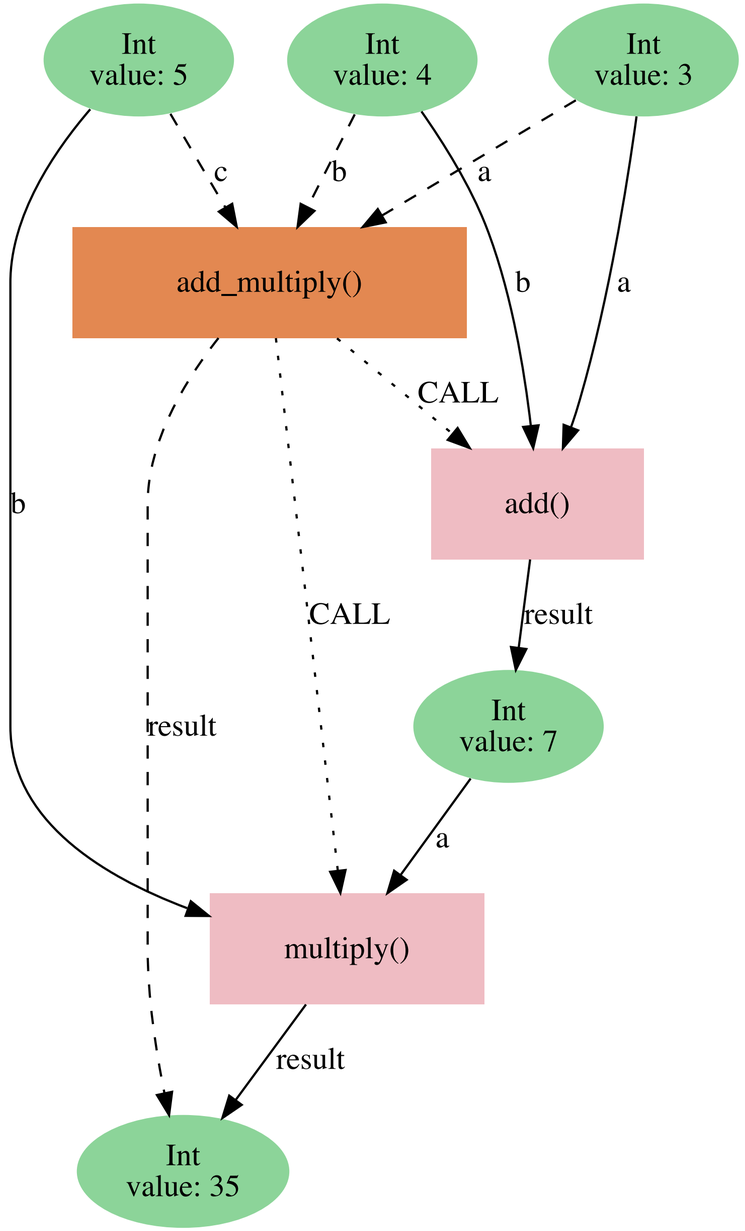
final_value: (3+4)*5= 35

Simple provenance tracking also for pre-/post-processing

```python
from aiida.engine import calcfunction, workfunction

@calcfunction
def add(a, b):
    return a+b


@calcfunction
def multiply(a, b):
    return a*b


@workfunction
def add_multiply(a, b, c):
    return multiply(add(a, b), c)


final_value = add_multiply(
    a=Int(3),
    b=Int(4),
    c=Int(5))
```

final_value:

```
<Int: uuid: e25b98c2-db06-4451-86dc-4e1b179d00cc (pk: 1547) value: 35>
```



Simple provenance tracking also for pre-/post-processing

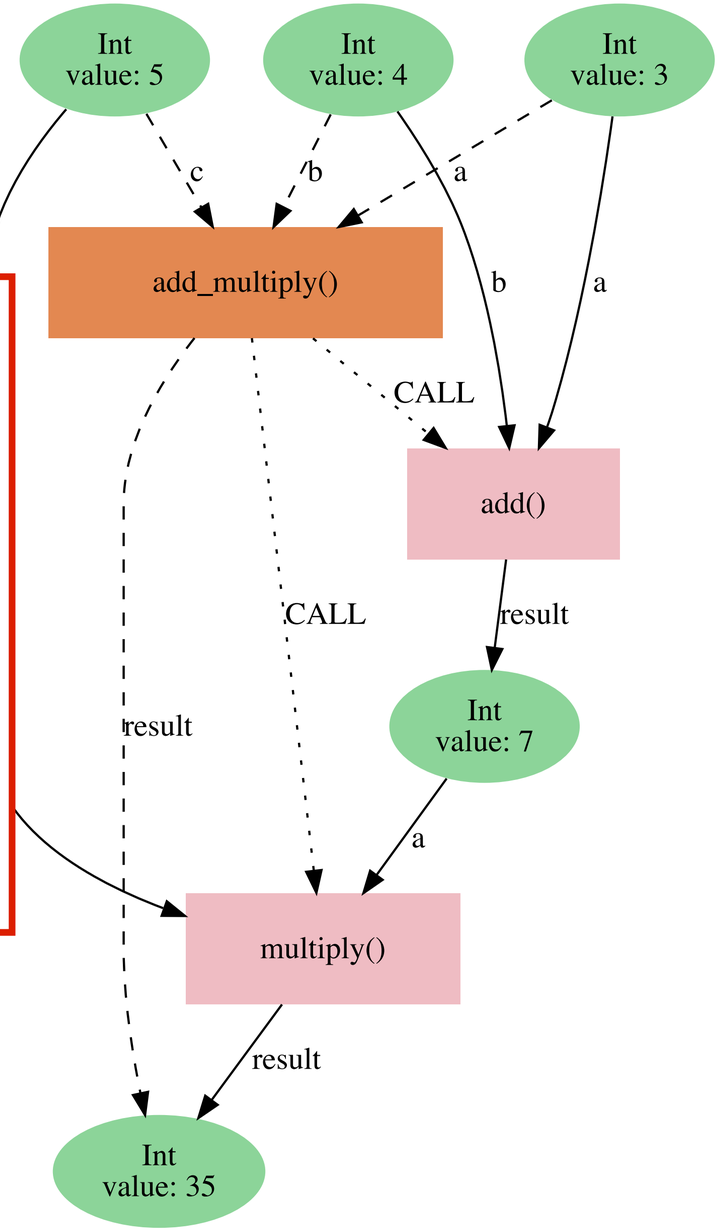
```python
from aiida.engine import calcfunction, workfunction

@calcfunction
def add(a, b):
    return a+b


@calcfunction
def multiply(a, b):
    return a*b


@workfunction
def add_multiply(a, b, c):
    return multiply(add(a, b), c


final_value = add_multiply(
    a=Int(3),
    b=Int(4),
    c=Int(5))
```

final_value:

```
<Int: uuid: e25b98c2-db06-4451-86dc-4e1b179d00cc (pk: 1547) value: 35>
```

AUTOMATIC PROVENANCE TRACKING

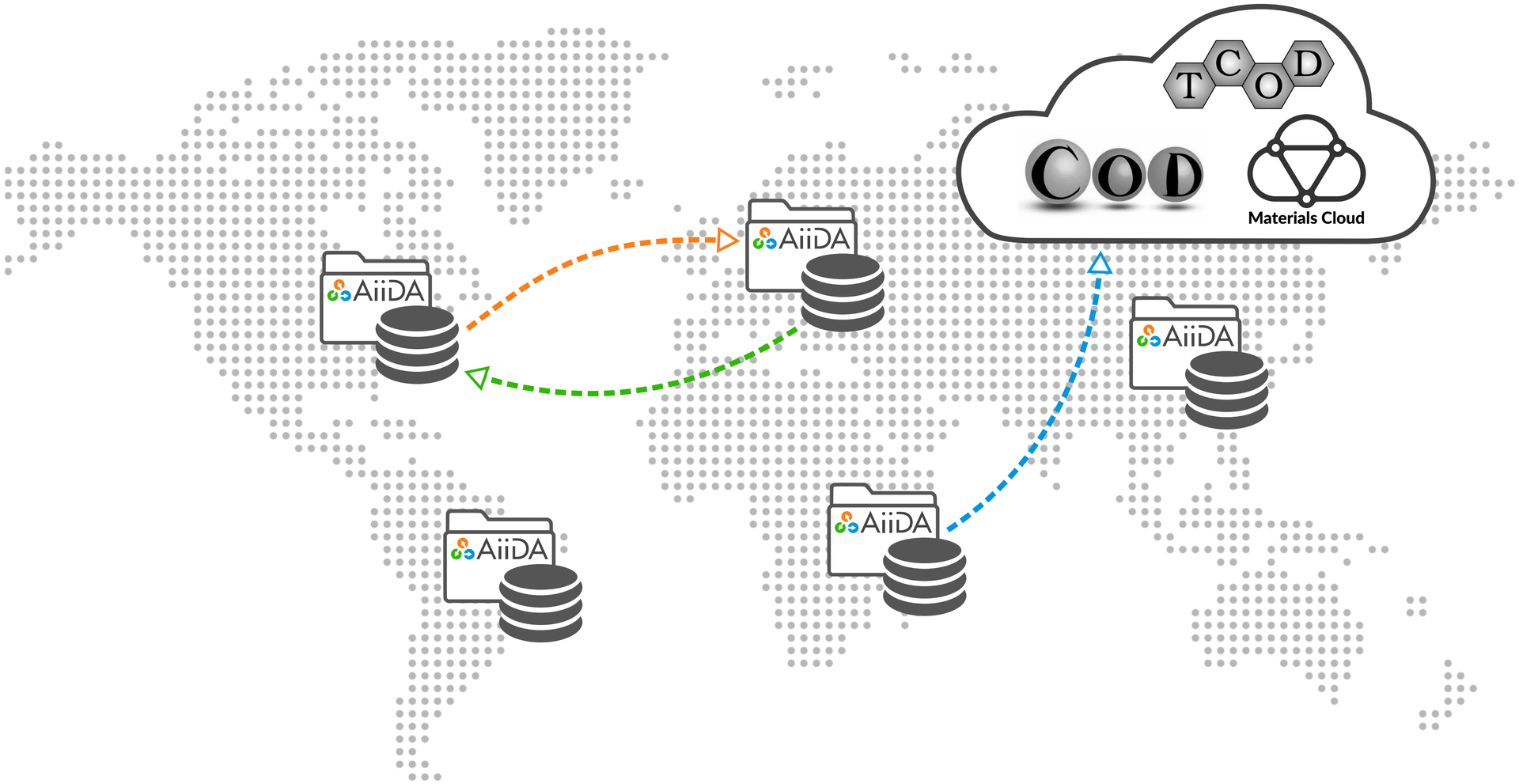
By very minimal modifications, provenance is automatically tracked also for simple, local data processing python functions



Sharing in AiiDA: data and graphs

- *Private* AiiDA instances
- UUIDs to uniquely identify nodes

- Data can be shared to other AiiDA repositories **or to online repositories**



Sharing in AiiDA: *not only data*
Codes, plugins and workflows

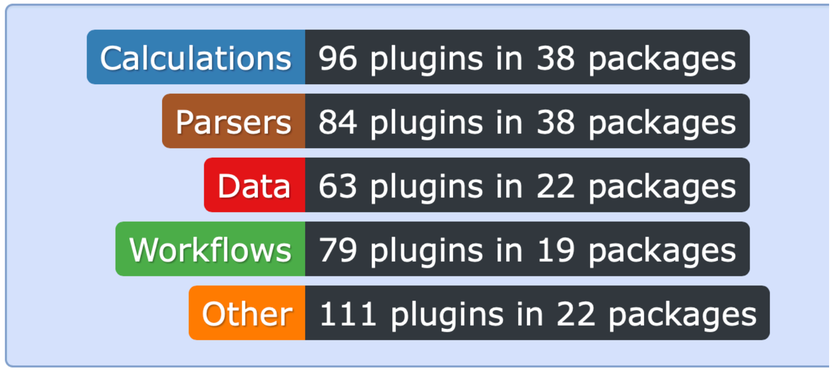





Calculation | **Data** | **Parsers** | **Transport and scheduler** | **Workflows** | **Importers & exporters**

AiiDA plugin registry

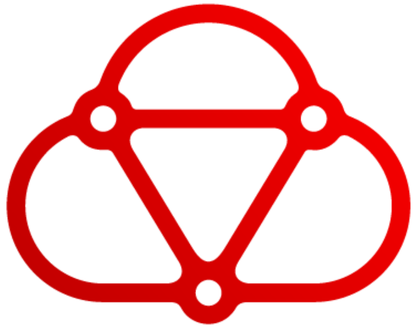
Registered plugin packages: 48



- Plugins are collected in the AiiDA plugin registry
- Over **90 different codes currently supported,** with almost 80 workflows
- Many are **community-contributed**

https://aiidateam.github.io/aiida-registry/

OPEN SCIENCE PLATFORM:

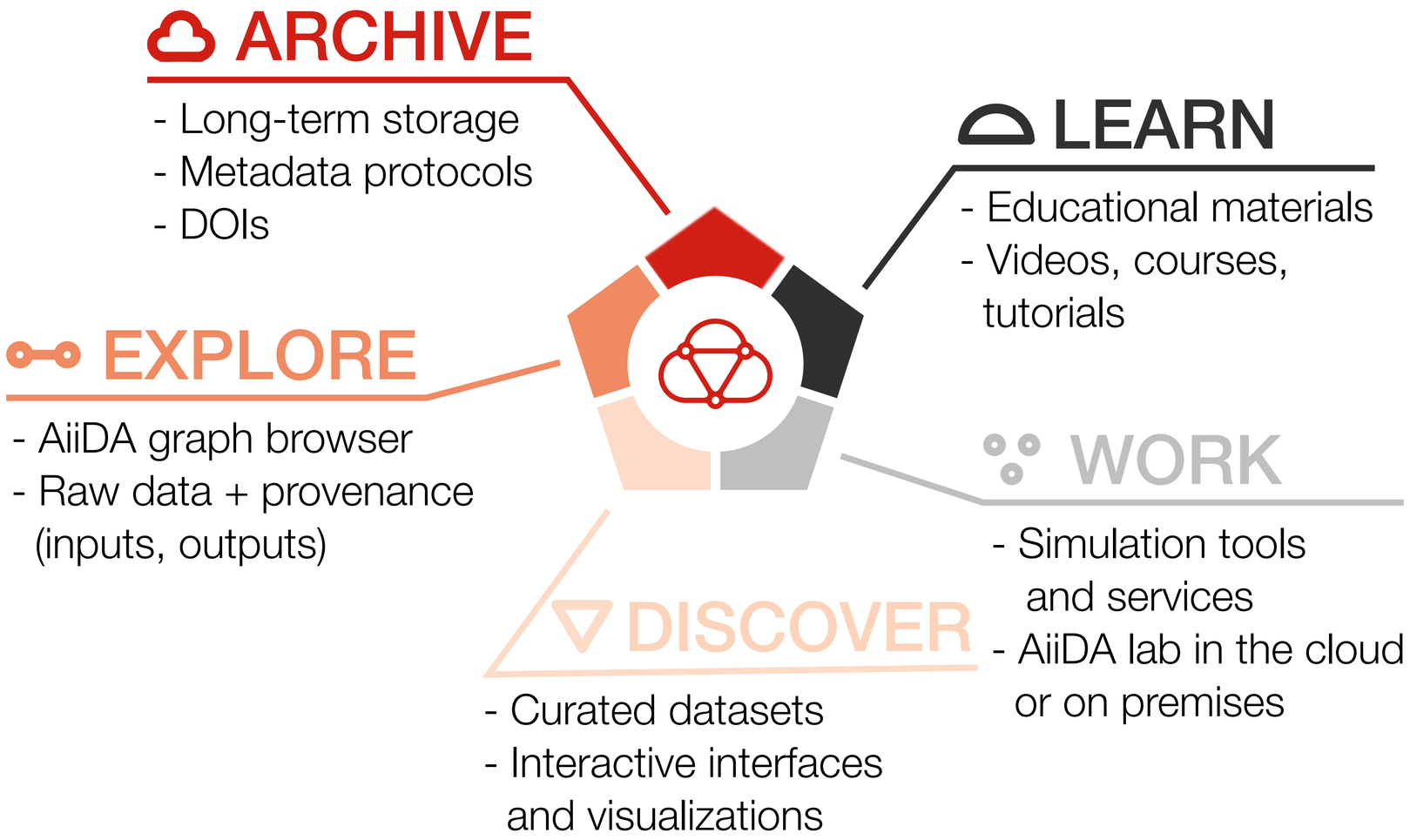


MATERIALSCLOUD

https://www.materialscloud.org

Materials Cloud

- **AiiDA** is the 'engine', like **Git -** used in production *since 2015*
- **Materials Cloud** is the dissemination platform (like **GitHub**) **and more** (cloud computing and data generation platform) - online since *Dec 2017*



Open data sharing: **Archive, Discover, Explore**

DOIs assigned

materialscloud:2017.0008

SCIENTIFIC DATA re3data.org http://doi.org/10.17616/R3ZJ5W Materials Cloud FAIRsharing.org standards, databases, policies

Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds

Authors: Nicolas Mounet[1*], Marco Gibertini[1], Philippe Schwaller[1], Davide Campi[1], Andrius Merkys[1,2], Antimo Marrazzo[1], Thibault Sohier[1], Ivano E. Castelli[1], Andrea Cepellotti[1], Giovanni Pizzi[1], Nicola Marzari[1*]

1 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

2 Vilnius University Institute of Biotechnology, Sauletekio al. 7, LT-10257 Vilnius, Lithuania

* Corresponding authors emails: nicolas.mounet@epfl.ch, nicola.marzari@epfl.ch

DOI 10.24435/materialscloud:2017.0008/v2 (version v2, submitted on 21 March 2018)

How to cite this entry

Nicolas Mounet, Marco Gibertini, Philippe Schwaller, Davide Campi, Andrius Merkys, Antimo Marrazzo, Thibault Sohier, Ivano E. Castelli, Andrea Cepellotti, Giovanni Pizzi, Nicola Marzari, *Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds*, Materials Cloud Archive (2018), doi: 10.24435/materialscloud:2017.0008/v2.

Description

Two-dimensional (2D) materials have emerged as promising candidates for next-generation electronic and optoelectronic applications. Yet, only a few dozens of 2D materials have been successfully synthesized or exfoliated. Here, we search for novel 2D materials that can be easily exfoliated from their parent compounds. Starting from 108423 unique, experimentally known three-dimensional compounds we identify a subset of 5619 that appear layered according to robust geometric and bonding criteria. High-throughput calculations using van-der-Waals density-functional theory, validated against experimental structural data and calculated random-phase-approximation binding energies, allow to identify 1825 compounds that are either easily or potentially exfoliable. In particular, the subset of 1036 easily exfoliable cases provides novel structural prototypes and simple ternary compounds as well as a large portfolio of materials to search from for optimal properties. For a subset of 258 compounds we explore vibrational, electronic, magnetic, and topological properties, identifying 56 ferromagnetic and antiferromagnetic systems, including half-metals and half-semiconductors. This archive entry contains the database of 2D materials (structural parameters, band structures, binding energies, etc.) together with the provenance of all data and calculations as stored by AiiDA.

Direct links to Discover & Explore

Materials Cloud sections using this data

Select 2d materials via interactive periodic table and view their properties (with links to provenance)

Explore interface providing access to the full database

FAIRsharing.org
re3data.org

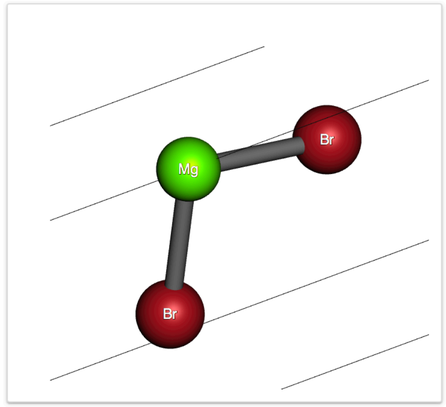
+

Recommended data repository by Nature's journal *Scientific Data*

Open data sharing: **Archive, Discover, Explore**

DISCOVER

Compound: $MgBr_2$



Info and properties

See definitions...

Formula: $MgBr_2$
Spacegroup: P-3m1
Pointgroup: -3m
Prototype: CdI2
Band gap [eV]: 4.8

Magnetic properties:

Magnetic State: non-magnetic
Tot. Magnetization [μB/cell]: -
Abs. Magnetization [μB/cell]: -

Binding Energies:

DF2-C09 Binding energy [meV/Å^2]:
(From parent COD 9009107)
rVV10 Binding energy [meV/Å^2]: 15
(From parent COD 9009107)

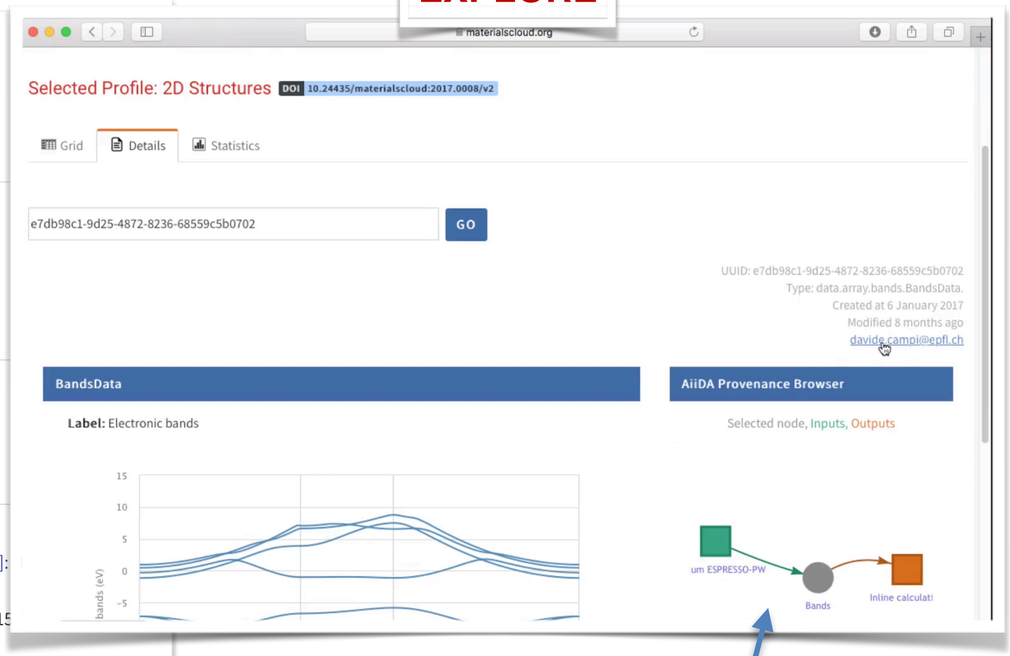
Delta in interlayer distance (vdW vs revPBE):

Δ_{DF2} [%]: 17.1 (From parent COD 9009107)
Δ_{rVV10} [%]: 18.3 (From parent COD 9009107)

Band structure

UUID links to jump to the provenance graph in the EXPLORE section

EXPLORE

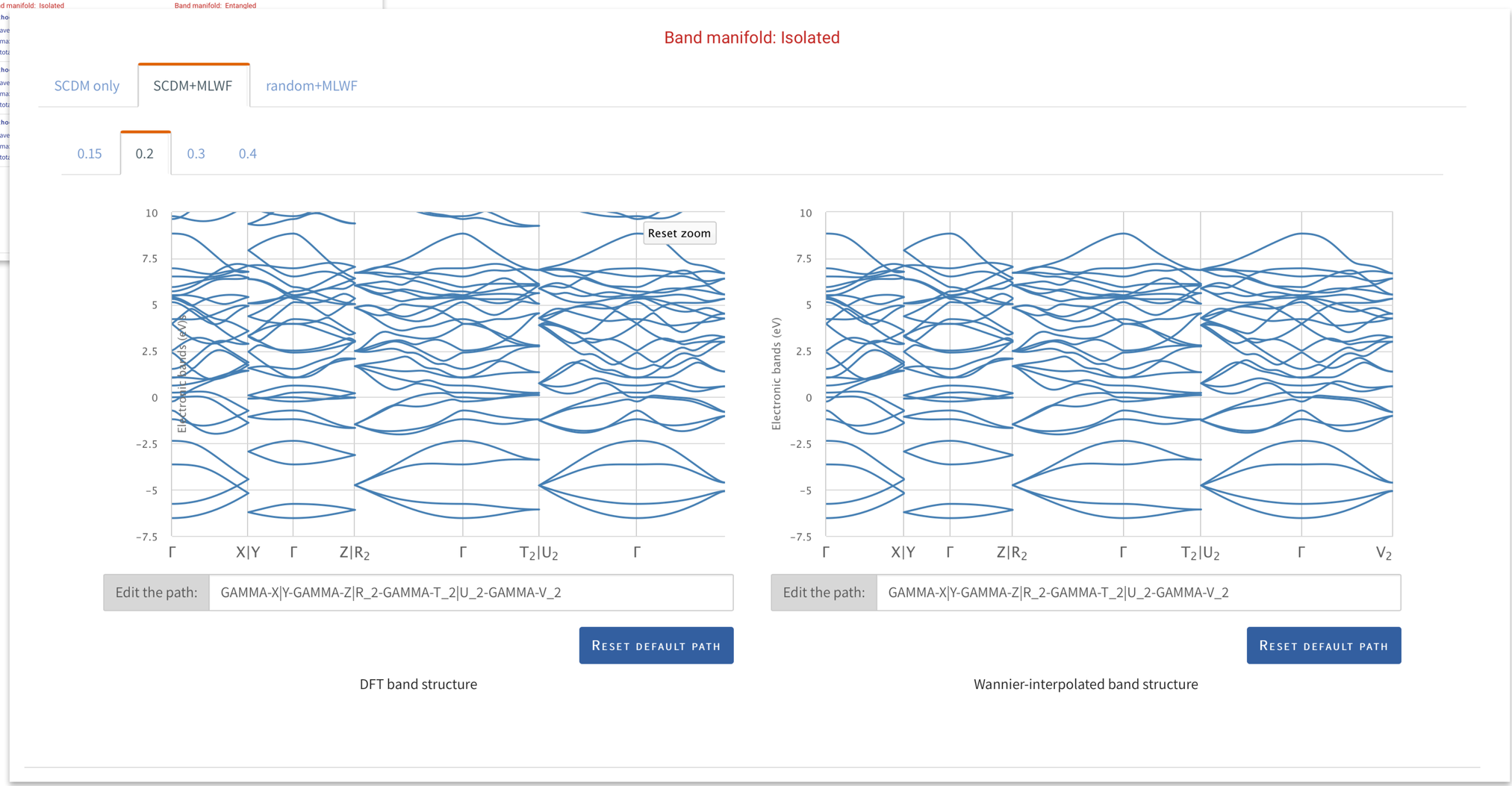


Browse the full AiiDA provenance graph (inputs, outputs, …) at any level

Explore+Discover section for Automatic Wannier functions will be online soon

Paper: V. Vitale, GP et al., arXiv:1909.00433, accepted in npj Comp. Mat.

Data: V. Vitale, GP et al., Materials Cloud Archive (2019),
doi: 10.24435/materialscloud:2019.0044/v2)



Data Management Plans (DMPs) and FAIR principles

- Combination of **AiiDA + Materials Cloud** (Discover, Explore, Archive): **FAIR-compliant sharing**

- **Findable**: DOIs with standardized metadata
- **Accessible**: web interface to browse data, calculations and provenance, curated data in Discover section
- **Interoperable**: data linked via the AiiDA directed graph; data structures reusable between different codes
- **Reusable**: downloadable data, encourage open (CC) licences, reproduce in the AiiDA Lab thanks to *full provenance*

- We provide DMP templates for researchers using Materials Cloud (and we are coordinating with EMMC for a EU H2020 template)

Funding Body	DMP template (using AiiDA)	DMP template (no AiiDA)
SNF	.docx .odt .pdf	.docx .odt .pdf
H2020	.docx .odt .pdf	.docx .odt .pdf

Materials Cloud WORK section

Work with your data in the cloud

You can access a number of services related to AiiDA and computational materials science that require almost no setup. They are either cloud online services or downloadable directly to your machine.

Tools

Online web tools to work with your data

AiiDA lab

Run your own simulations using AiiDA on the cloud

Quantum Mobile

A virtual machine with quantum codes ready to be used via AiiDA

AiiDA registry

The official registry of AiiDA plugins

Materials Cloud WORK section

Work with your data in the cloud

You can access a number of services related to AiiDA and computational materials science that require almost no setup. They are either cloud online services or downloadable directly to your machine.

Tools

Online web tools to work with your data

AiiDA lab

Run your own simulations using AiiDA on the cloud

Quantum Mobile

A virtual machine with quantum codes ready to be used via AiiDA

AiiDA registry

The official registry of AiiDA plugins

For instance: cloud/web version of **seekpath**

Materials Cloud WORK section

Work with your data in the cloud

You can access a number of services related to AiiDA and computational materials science that require almost no setup. They are either cloud online services or downloadable directly to your machine.

Tools

Online web tools to work with your data

AiiDA lab

Run your own simulations using AiiDA on the cloud

Quantum Mobile

A virtual machine with quantum codes ready to be used via AiiDA

AiiDA registry

The official registry of AiiDA plugins

You are using it
for this tutorial
Feel free to use it also
for other events and tutorials

Materials Cloud WORK section

Work with your data in the cloud

You can access a number of services related to AiiDA and computational materials science that require almost no setup. They are either cloud online services or downloadable directly to your machine.

Tools

Online web tools to work with your data

AiiDA lab

Run your own simulations using AiiDA on the cloud

Quantum Mobile

A virtual machine with quantum codes ready to be used via AiiDA

AiiDA registry

The official registry of AiiDA plugins

AiiDA on the cloud
using a jupyter interface

Acknowledgements and funding

AiiDA core and Materials Cloud authors

Carl S. Adorf, Casper W. Andersen, Marco Borelli, Andrea Cepellotti, Fernando Gargiulo, Valeria Granata, Dominik Gresch, Sebastiaan P. Huber, Rico Häuselmann, Conrad Johnston, Leonid Kahle, Boris Kozinsky, Snehal Kumbhar, Nicola Marzari, Andrius Merkis, Nicolas Mounet, Tiziano Müller, Elsa Passaro, Daniele Passerone, Carlo A. Pignedoli, Giovanni Pizzi, Francisco F. Ramirez, Thomas C. Schulthess, Ole Schütt, Berend Smit, Leopold Talirz, Martin Uhrin, Joost VandeVondele, Aliaksandr V. Yakutovich, Spyros Zoupanos

All AiiDA core contributors
All contributors for the 45+ plugins

The CSCS support teams

MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH
FNSNF

SNSF NCCR "MARVEL"
Discovery of new materials via simulations and dissemination of curated data

MAX DRIVING THE EXASCALE TRANSITION

H2020 Centre of Excellence "MaX"
Scaling towards exascale machines and high-throughput efficiency

swissuniversities

Swissuniversities P-5 "Materials Cloud"
Scaling the web platform, extending to more disciplines

Moreover:

H2020 Marketplace (providing data and simulation services in a EU Markeplace platform also for industry)

H2020 Intersect (develop AiiDA workflows to compute transport properties of materials)

Conclusions (part 2 - AiiDA and Materials Cloud)

- **AiiDA**: reproducibility and automation engine
  - turn-key workflows and automation with full provenance tracking as a graph
  - Share data, plugins and workflows
- **Materials Cloud**: FAIR dissemination portal
  - *Archive*: Findable long-term storage + *Discover* (curated data) + *Explore* (raw AiiDA data)
  - Quantum Mobile as a tool for tutorials and schools

Conclusions

- Automated Wannier functions: **all data and provenance on Materials Cloud Archive**
  (Valerio Vitale, GP *et al*, Materials Cloud Archive (2019), doi: 10.24435/materialscloud:2019.0044/v2)
- Archive entry contains a **reproducible Virtual Machine** to run the same (or new) simulations
- Workflows now updated to most recent version of AiiDA, you will try in the tutorial
- *Discover+Explore* sections for the automatic Wannier functions project: soon online

Contacts

Website: http://www.aiida.net

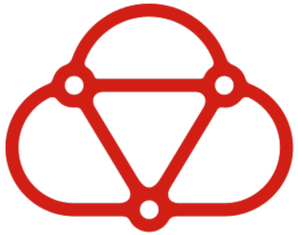
Docs: http://aiida-core.readthedocs.io

Git repo: https://github.com/aiidateam/aiida_core/

Plugin registry: http://aiidateam.github.io/aiida-registry

https://www.facebook.com/aiidateam

@aiidateam



Materials Cloud: http://www.materialscloud.org

- **AiiDA Lab**: http://aiidalab.materialscloud.org
- **Archive**: http://archive.materialscloud.org

Quantum Mobile: http://www.materialscloud.org/work/quantum-mobile