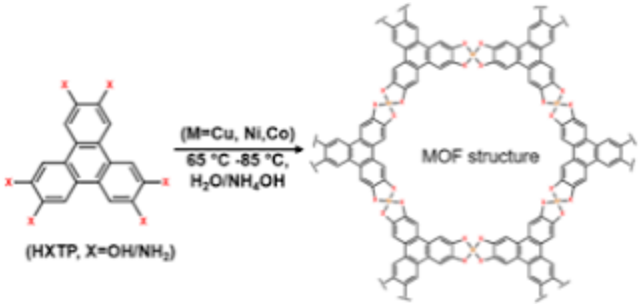
Materials informatics for DM

A. Balatsky



- Organic DM: Materials Database and search of novel organic MOFs
- Organic SC

omdb.mathub.io

Organic Materials Database
omdb.mathub.io

Joahn Helsvik

Roberto Díaz Pérez

A. Pertsova

S. Borysov

M. Geilhufe

B. Olsthoorn

J. Schaltegger

Organic Materials Database (OMDB)

Free to use!

https://omdb.mathub.io

S.S. Borysov, RM. Geilhufe, A.V. Balatsky,PloS one, 12:2, 2017

Electronic Structure

- Band structure, DOS

First dataset
- VASP, PBE
- ~ 41,000 materials

Second dataset
- VASP, SCAN+VdW (DFT-TS)
- In progress

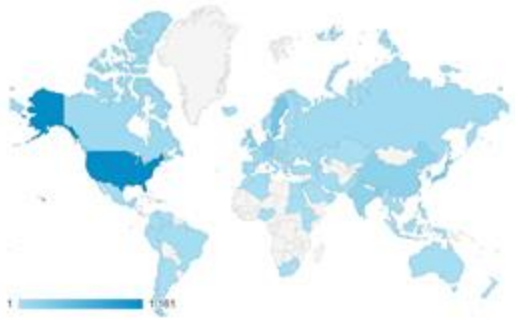
Magnetic Structure

- Magnon spectra
- Magnetic ground state

First dataset
- RSPT
- Linear spin-wave theory
- ~ 120 materials (in progress)

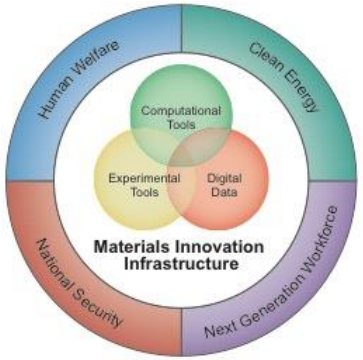
OMDB Community

- Provide / request services
  - Synthesis, theory, measurements
- Upload your own results
- “Follow” materials



General Context and Motivation

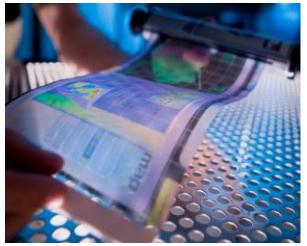
Accelerated Materials Discovery



https://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf

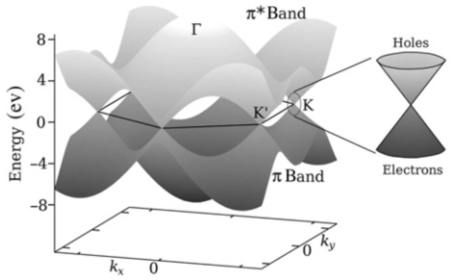
Organic Materials

- Flexible electronics
- Organic LED
- Organic solar cells



http://www.flickr.com/photos/rdecom/4146880795/

Dirac Materials



- Graphene
- Topological insulators
- Weyl semi-metals

Organic Materials Database
Online Database with Focus on Data Mining
omdb.mathub.io

The Materials Project

MatDat

Material Properties Database

Over 1000 detailed, fully referenced and verified **datasets** for steels, aluminium and titanium alloys, cast irons/steels, weld metals,...

Detailed mechanical properties (monotonic, cyclic, fatigue) for all materials in the database.

Sign up for **MatDat Free**

search materials in the database only

Sign up for **MatDat Edu**

access to all materials in the database for students and young researchers/scientists

Sign up for **MatDat Pro**

full access to all materials and advanced options

Curtarolo Materials Laboratory

AFLOW

Automatic - FLOW for Materials Discovery

MatWeb

MatNavi is NIMS Structural Datasheet Online. It is one of the largest databases in the world for materials.

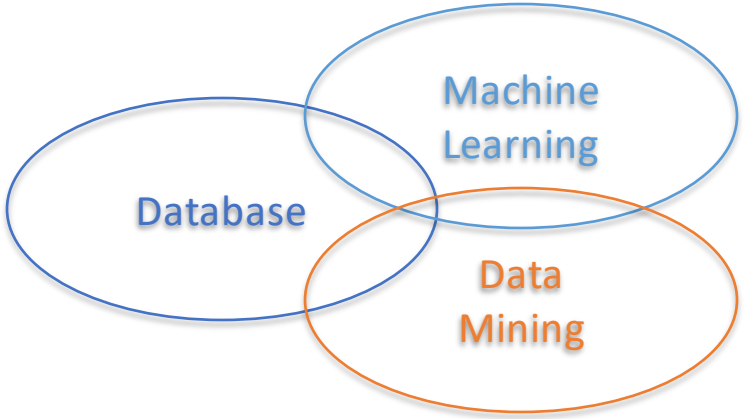
Overview

Data management and search systems for materials databases

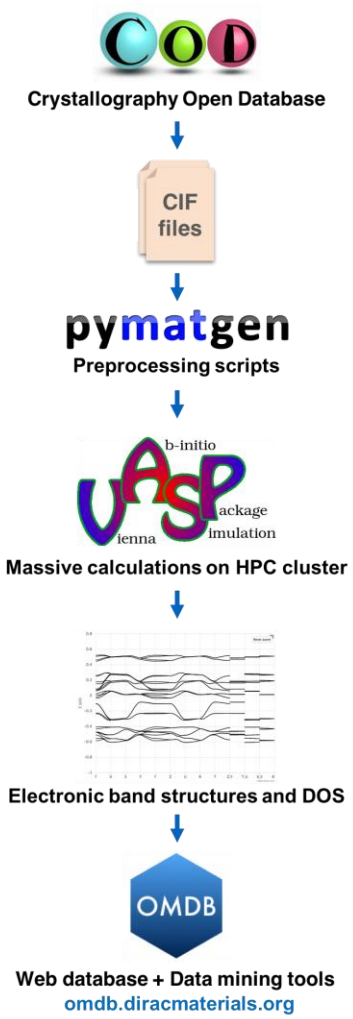
Organic materials

Building database

Development of advanced search tools and tools for predictive analysis for materials
incorporating machine learning and data mining techniques



How we do it



crystallography.net
More than 100,000 organic compounds to calculate (~40,000 are ready)

Standard DFT calculations to capture trends

- Vienna ab-initio simmulation package (VASP)
  G. Kresse & J. Furthmüller, *Phys. Rev. B*, 54, 11 169, 1996
  G. Kresse & D. Joubert, *Phys. Rev. B*, 59, 1758, 1999
- generalized gradient approximation (PBE)
  J. P. Perdew, K. Burke, & M. Ernzerhof, *Phys. Rev. Lett.*, 77, 3865, 1996
- spin polarized, no spin-orbit coupling

Borysov SS, Geilhufe RM, Balatsky AV,
Organic materials database: An open-access online database for data mining,
PLoS ONE 12(2), e0171501 (2017)

OMDB

Welcome to the

Organic Materials Database

The organic materials database is an open access electronic structure database for 3-dimensional organic crystals, developed and hosted at the Nordic Institute for Theoretical Physics – Nordita.

It provides tools for search queries based on data-mining and machine learning techniques. The universal features provided on our web interface facilitate the design of novel functional organic materials with a wide-range of applications.

Discover more

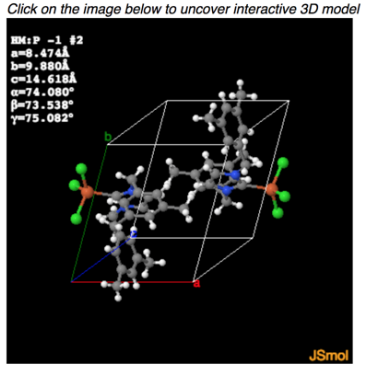
Organic Materials Database: Content

~40,000 Electronic Structures
(june, 2022)

Information about material with ID (COD) = 4082354

Basic information

Click on the image below to uncover interactive 3D model



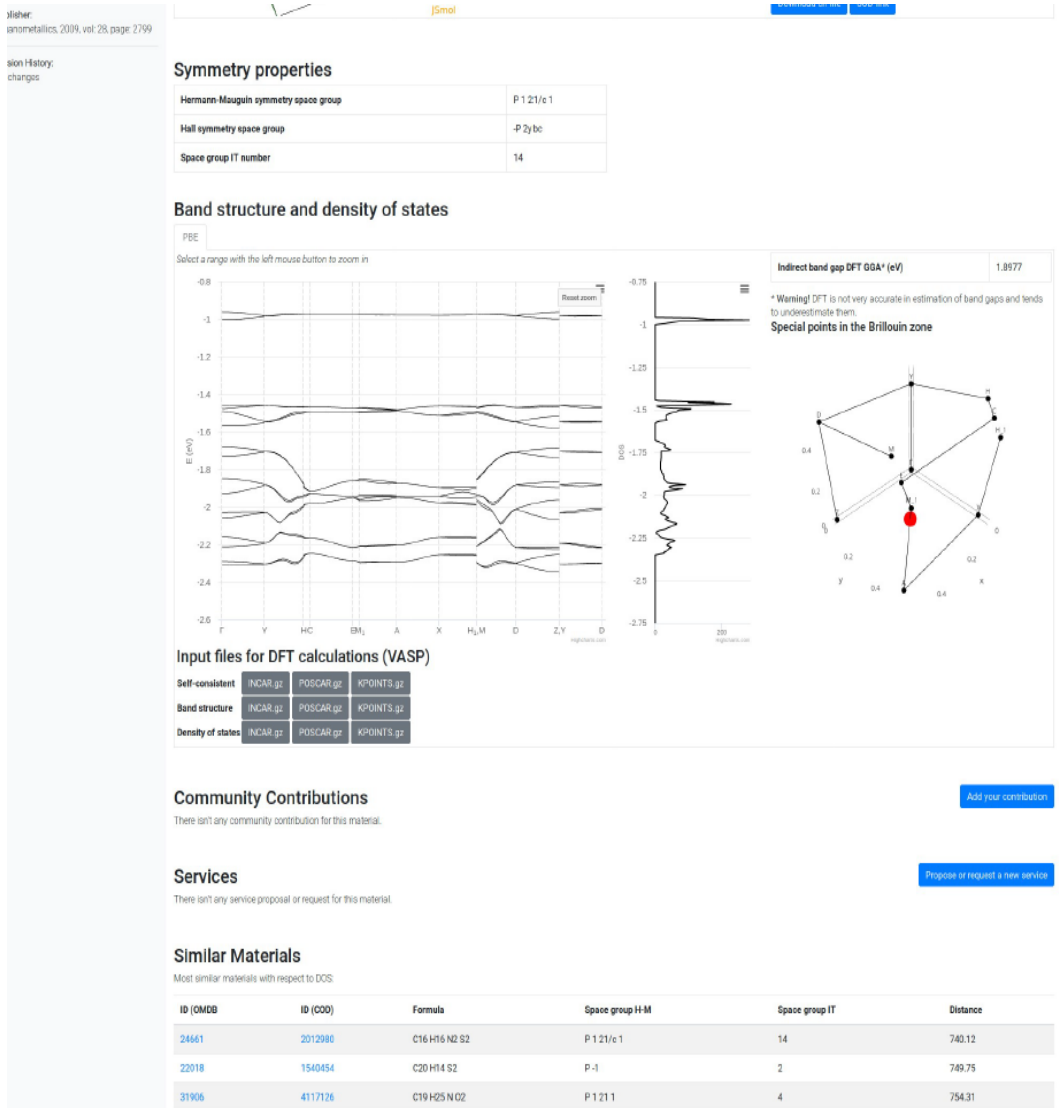
Explore
load unitcell | load 27 unitcells

Interact
rotate to best view | default zoom
hide symmetry
spin on

ID (OMDB)	1
ID (COD)	4082354
Publication details	Hashimoto, Takayoshi; Hoshino, Ryoko; Hatanaka, Tsubasa; et al. Dinuclear Iron(0) Complexes of N-Heterocyclic Carbenes
Publisher	Organometallics, 2014, vol: 33, page: 921
Chemical name	
Formula	C21 H24 Cl3 Fe N2
a	8.474(2)
b	9.880(3)
c	14.618(3)
α	74.080(13)
β	73.538(14)
γ	75.082(12)
Hermann-Mauguin symmetry space group	P -1
Hall symmetry space group	-P 1
Space group IT number	2
Indirect band gap DFT GGA* (eV)	0.9815
Magnetic moment DFT	9.9993

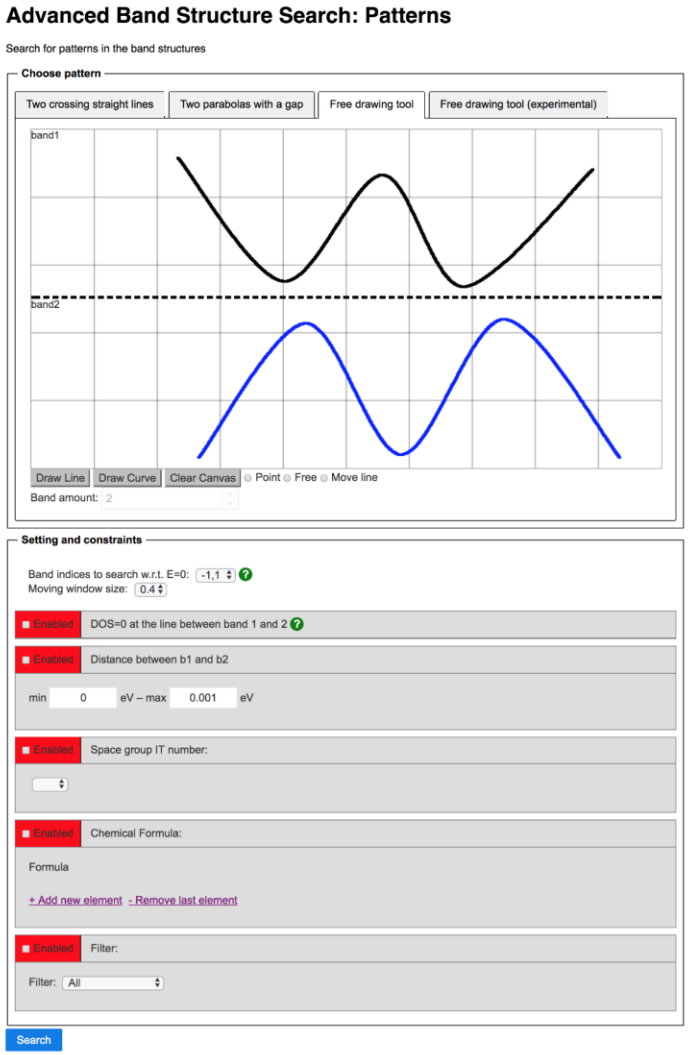
Download cif file | Find more information @ Crystallography Open Database (COD)

* **Warning!** DFT is not very accurate in estimation of band gaps and tends to underestimate them.

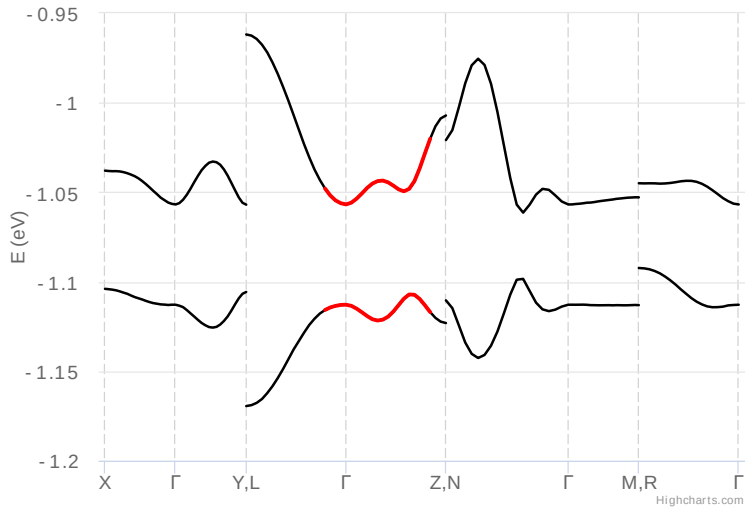


Organic Materials Database: Online Pattern Search

Arbitrary pattern



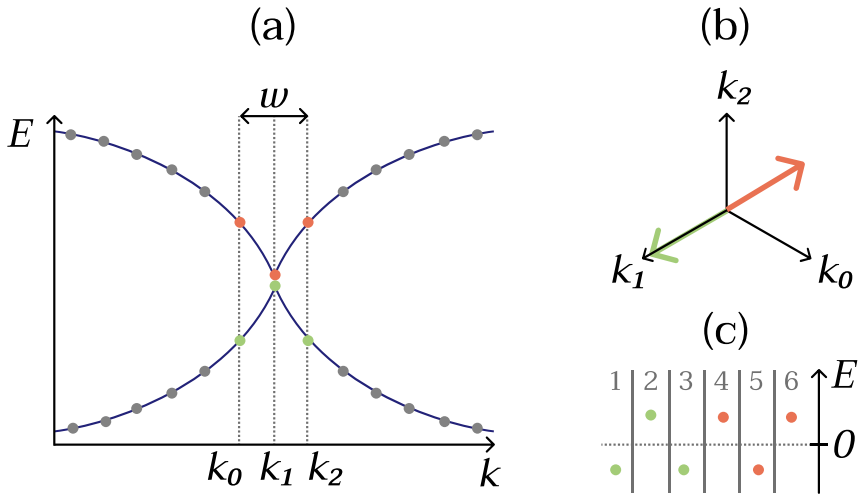
Search results



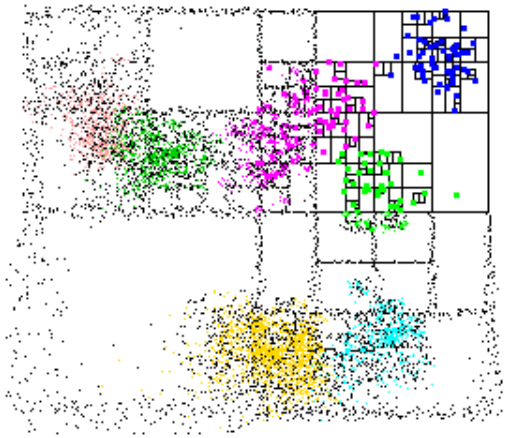
S.S. Borysov, B. Olsthoorn, M.B. Gedik, R.M. Geilhufe, A.V. Balatsky,
Nature Computatonal Materials Science (2018)

Organic Materials Database: Online Pattern Search

1. Represent data and query as vectors



2. Approximate nearest neighbor search



https://www.cs.umd.edu/~mount/ANN/

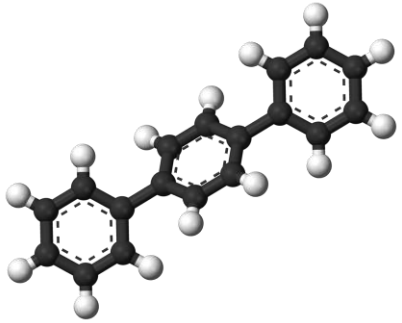
S.S. Borysov, B. Olsthoorn, M.B. Gedik, R.M. Geilhufe, A.V. Balatsky,
Nature Computational Materials Science (2018)

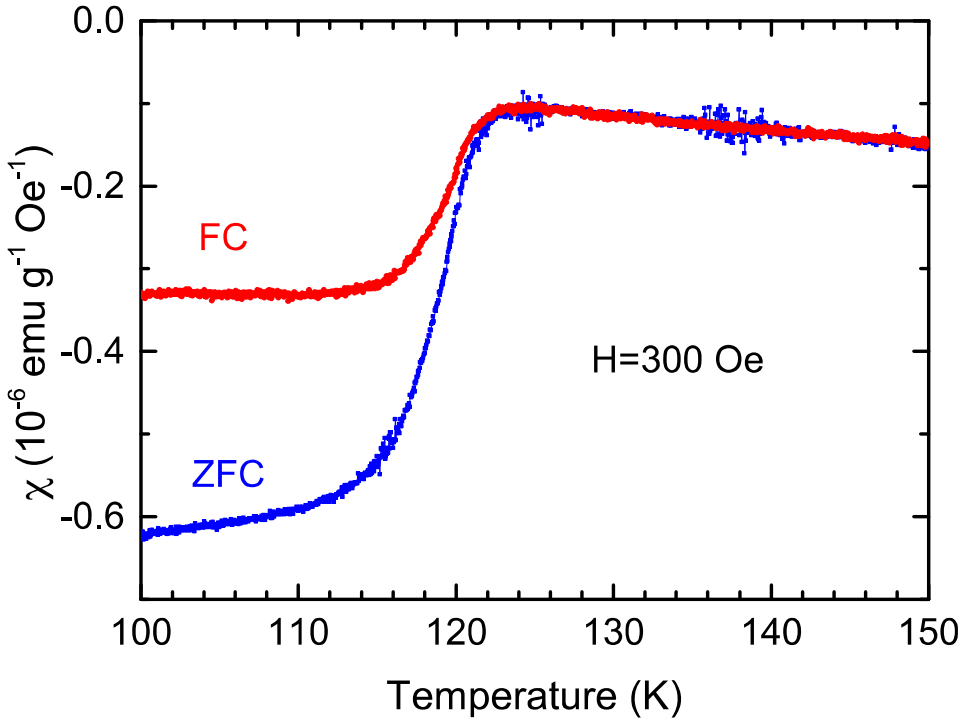
- Dirac Materials
- Organic Materials Database
- **Examples of a search of DM**
- ML and electronic structure inference
- Conclusion

omdb.mathub.io

Organic Materials Database: High-Tc SC Search

K doped P-Terphenyl (Tc ~ 120 K)

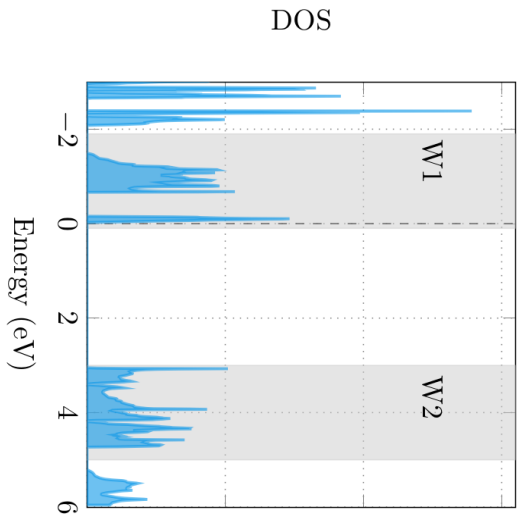




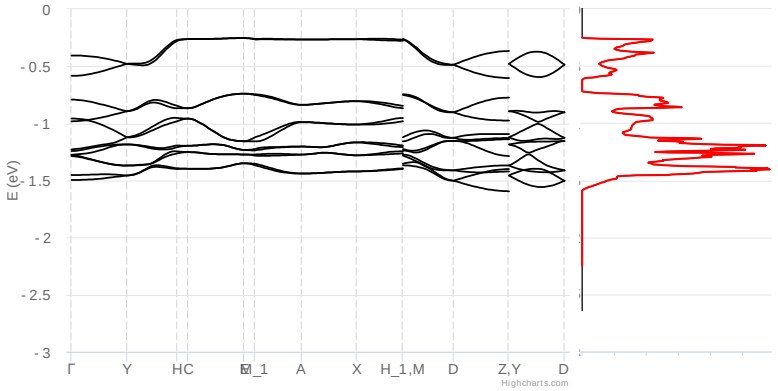
Ren-Shu Wang, Yun Gao, Zhong-Bing Huang, Xiao-Jia Chen,
Superconductivity above 120 kelvin in a chain link molecule,
arXiv preprint arXiv:1703.06641 (2017)

Organic Materials Database: DOS Similarity Search

Reference material

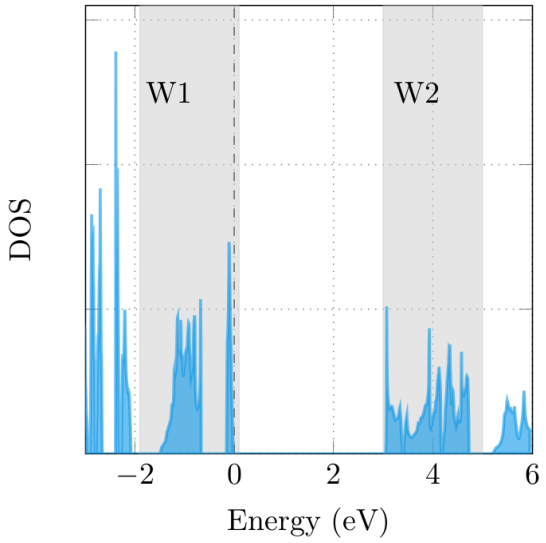


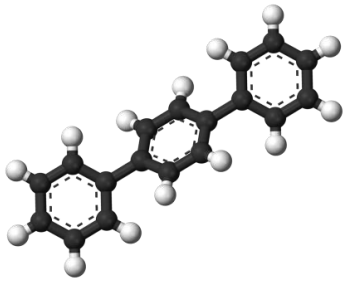
Search results



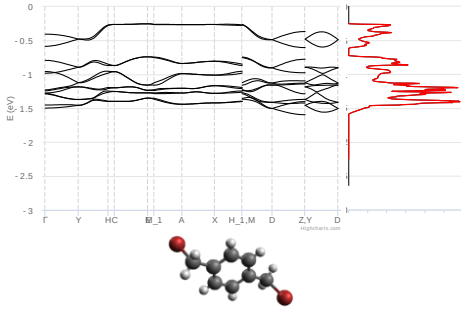
Organic Materials Database: DOS Similarity Search

Reference material (p-terphenyl)

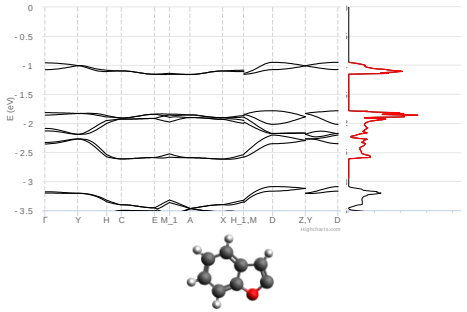




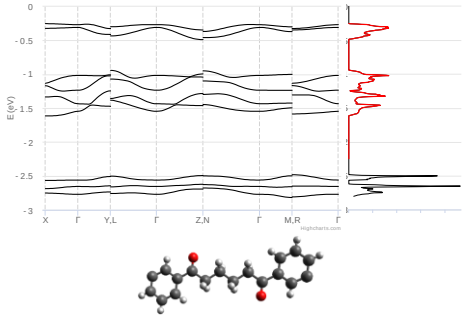
Valence Bands



(a) OMDB-ID 12336, cod-id 2212018, $C_8H_8Br_2$

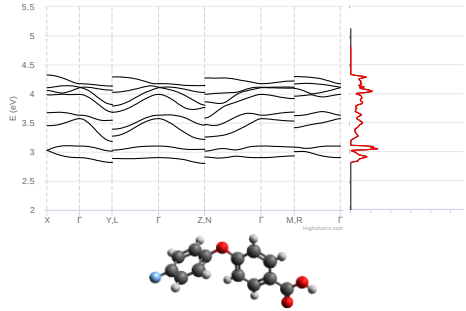


(b) OMDB-ID 261, cod-id 4062949, C_8H_5O

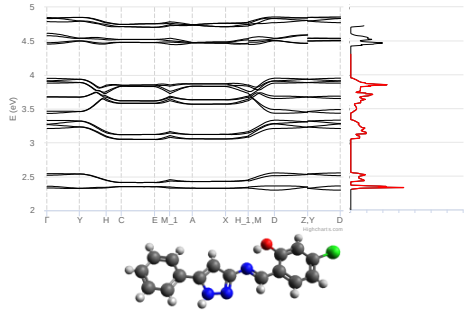


(c) OMDB-ID 12231, cod-id 2203730, $C_{18}H_{18}O_2$

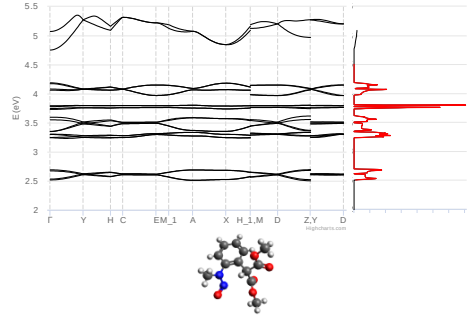
Conduction Bands



(a) OMDB-ID 20042, cod-id 2222905, $C_{13}H_9FO_3$



(b) OMDB-ID 21466, cod-id 7153064, $C_{16}H_{12}ClN_3O$



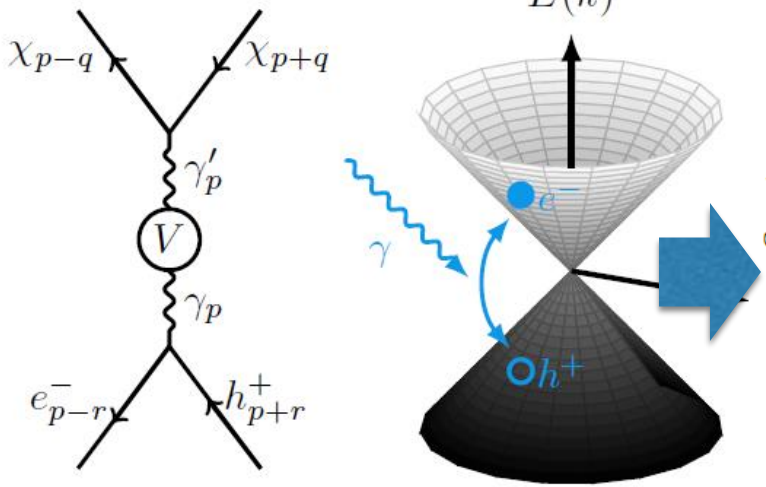
(c) OMDB-ID 19325, cod-id 1542706, $C_{12}H_{14}N_2O_5$

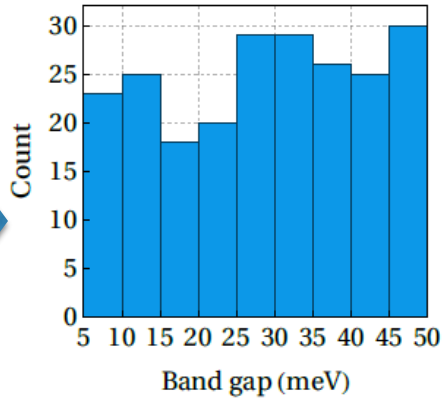
Geilhufe RM, Borysov SS, Kalpakchi D, Balatsky AV,
Novel Organic High-Tc Superconductors: Data Mining using Density of States Similarity Search, arXiv preprint arXiv:1709.03151 (2017)

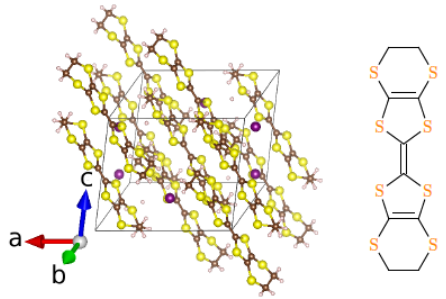
Light dark matter – small gap sensors Small gap natural for Dirac Materials

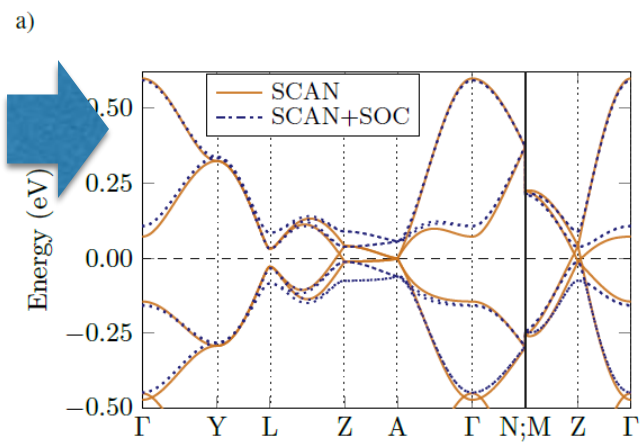
Y. Hochberg, et al, Physics Letters B 772, 239 – 246 (2017). Y. Hochberg et al, Phys. Rev. D 97, 015004 (2018)

DM for DM detection









(BEDT-TTF)Br.

Database search: M. Geilhufe et al, arxiv18

Mass fluctuations and absorption rates in Dirac materials sensorsB Olsthoorn, AV Balatsky
arXiv preprint arXiv:1909.10394

Goals: close the loop

Crystallography Open Database

COD Home

Home
What's new?

Accessing COD Data

Browse
Search
Search by structural formula

Add Your Data

Deposit your data
Manage depositions
Manage/release prepublications

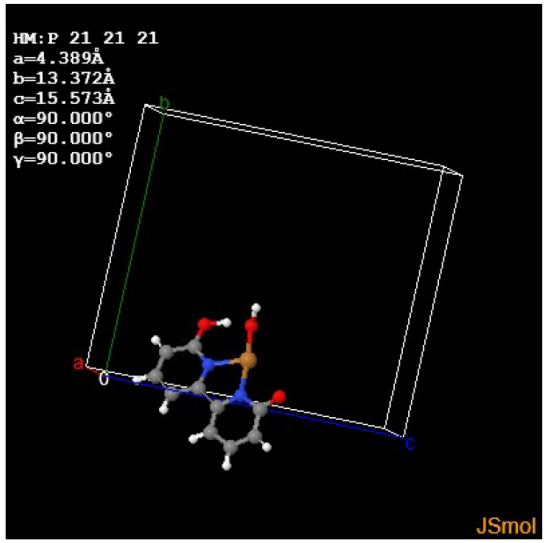
Documentation

COD Wiki
Obtaining COD
Querying COD
Citing COD
COD Mirrors
Advices to donators
Useful links

Information card for entry 4120693

4120692 << **4120693** >> 4120694

Preview

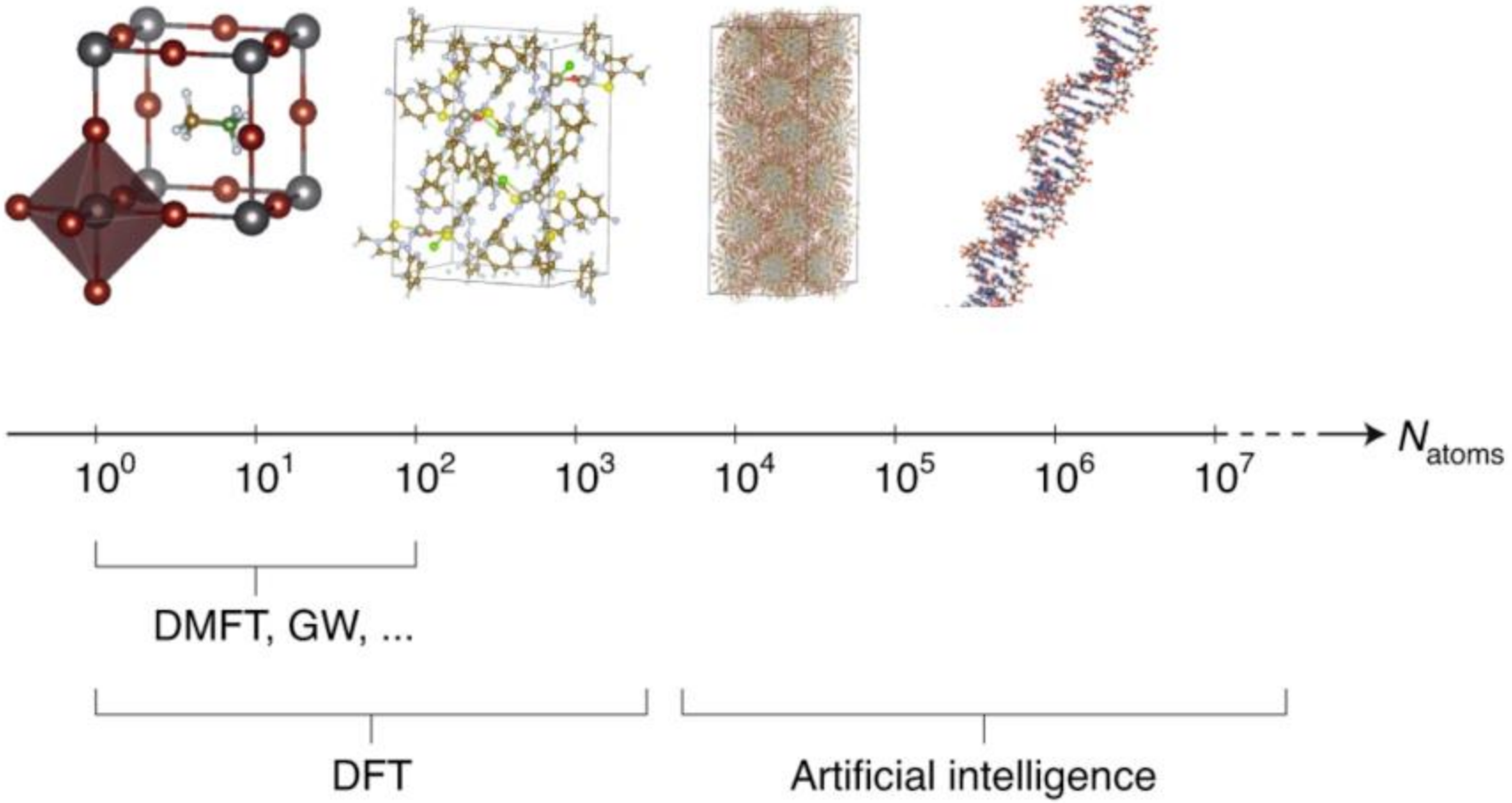


Display in Jmol

Mined- predicted

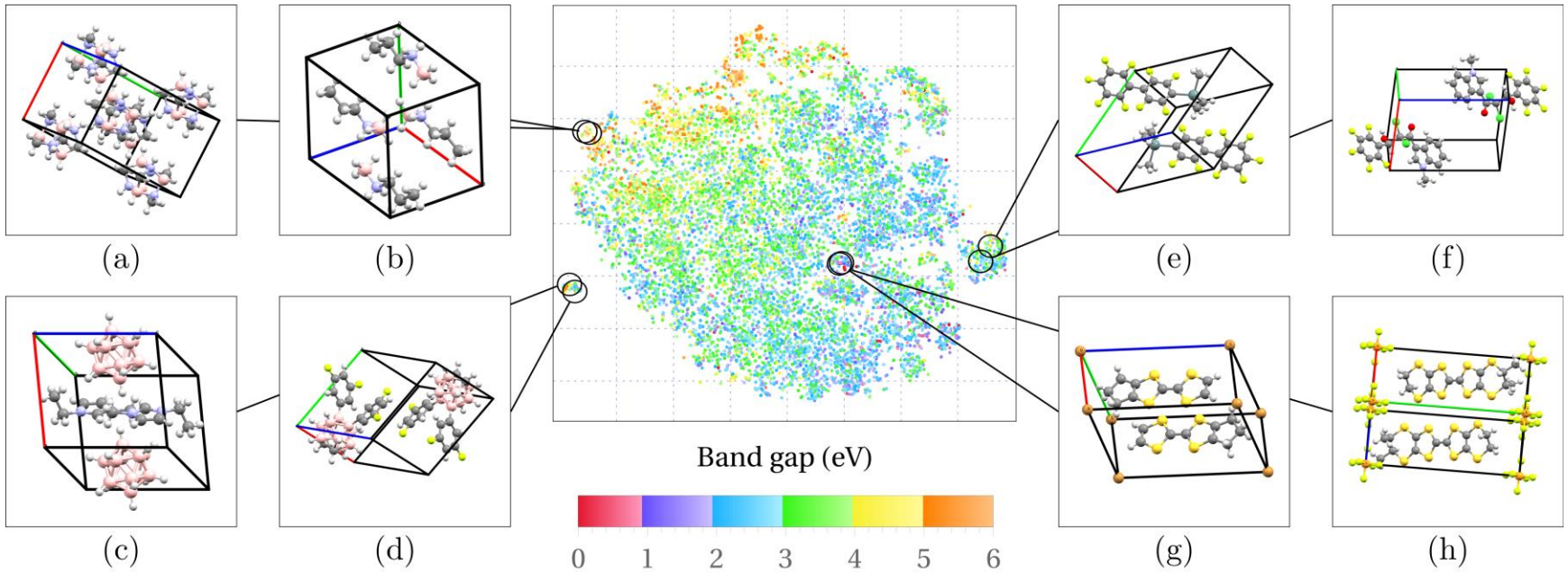
Formula C10 H8 Cu N2 O3
A biomimetic copper water oxidation catalyst with low overpotential
Zhang et al, JACS, v 136, 273(2014)

Ab initio saturation- certainty



Structural similarity (dimensionality reduction) and band gaps

- Chemical hydrogen storage
- Similar molecular crystals



- Boron-containing clusters
- Organic metals/superconductors
- Interactive version available: https://omdb.diracmaterials.org/dataset/OMDB-GAP1/interactive

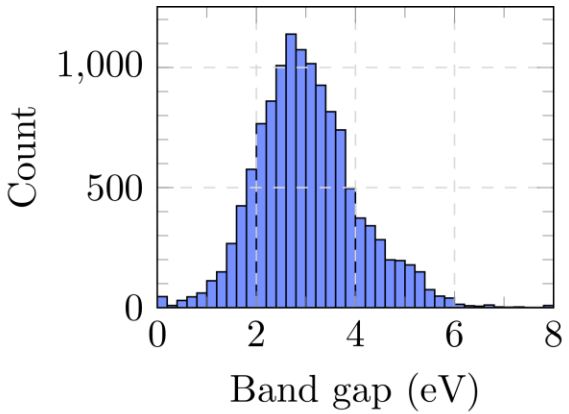
Band gap prediction for large organic crystal structures with machine learning

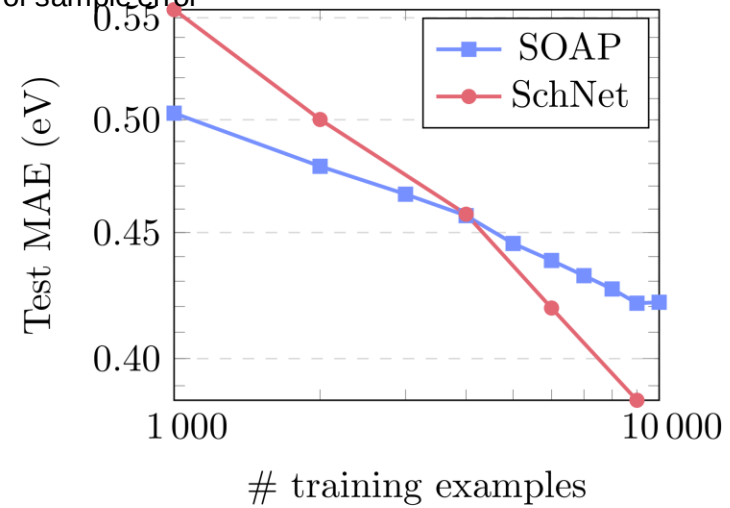
Bart Olsthoorn, **arXiv:1810.12814v4**

Band gap prediction for large organic crystal structures with machine learning

ArXiv:1810.12814 (2018, under review)
Also part of: Machine Learning for Molecules and Materials **NIPS 2018 Workshop**

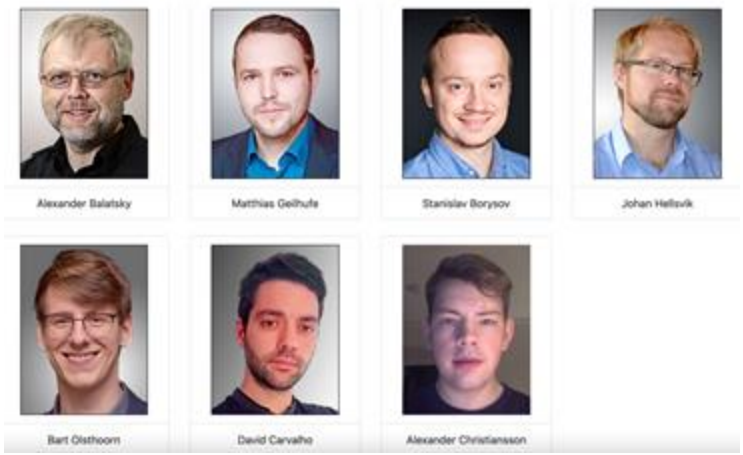
10,000 training examples, 2500 to test out of sample error





SOAP: a structural similarity kernel for kernel ridge regression (10.1103/PhysRevB.87.184115)
SchNet: a deep neural network (arXiv:1809.01072)
OMDB-GAP1: A new dataset for machine learning – `https://omdb.diracmaterials.org/dataset/`

Band gap prediction for large organic crystal structures with machine learning B Olsthoorn, RM Geilhufe, SS Borysov, AV Balatsky
Advanced Quantum Technologies 2 (7-8), 1900023



Papers ~200+ citations

1. R. Matthias Geilhufe, Bart Olsthoorn, Alexander V. Balatsky, Nature Physics (2021).
2. R. Matthias Geilhufe and Bart Olsthoorn, Phys. Rev. B 102, 205134 (2020).
3. Johan Hellsvik, Roberto Díaz Pérez, R. Matthias Geilhufe, Martin Månsson, and Alexander V. Balatsky, Phys. Rev. Materials 4, 024409 (2020).
4. Bart Olsthoorn, R. Matthias Geilhufe, Stanislav S. Borysov, Alexander V. Balatsky, Advanced Quantum Technologies (2019).
5. R. Matthias Geilhufe, Bart Olsthoorn, Alfredo Ferella, Timo Koski, Felix Kahlhoefer, Jan Conrad, Alexander V. Balatsky, Phys. Status Solidi RRL, 12: 1800293 (2018).
6. R. M. Geilhufe, B. Commeau, and G. W. Fernando, physica status solidi (RRL)–Rapid Research Letters (2018)
7. B. Commeau, R. M. Geilhufe, G. W. Fernando, A. V. Balatsky, Physical Review B 96.12 (2017): 125135.
8. R. M. Geilhufe, A. Bouhon, S. S. Borysov, A. V. Balatsky, Physical Review B 95:4 (2017): 041103.
9. R. M. Geilhufe, S. S. Borysov, A. Bouhon, A. V. Balatsky, Scientific Reports 7:1 (2017): 7298.
10. R. M. Geilhufe, S. S. Borysov, D. Kalpakchi, A.V. Balatsky, Physical Review Materials 2:2, (2018): 024802
11. S. S. Borysov, B. Olsthoorn, M. B. Gedik, R. M. Geilhufe, A. V. Balatsky, npj Computational Materials 4, 46 (2018).
12. S. S. Borysov, R. M. Geilhufe, A. V. Balatsky, PLOS ONE 12(2): e0171501 (2017).