High-throughput discovery and development: software engines for fast autonomous screening and characterization of materials

PAOFLOW

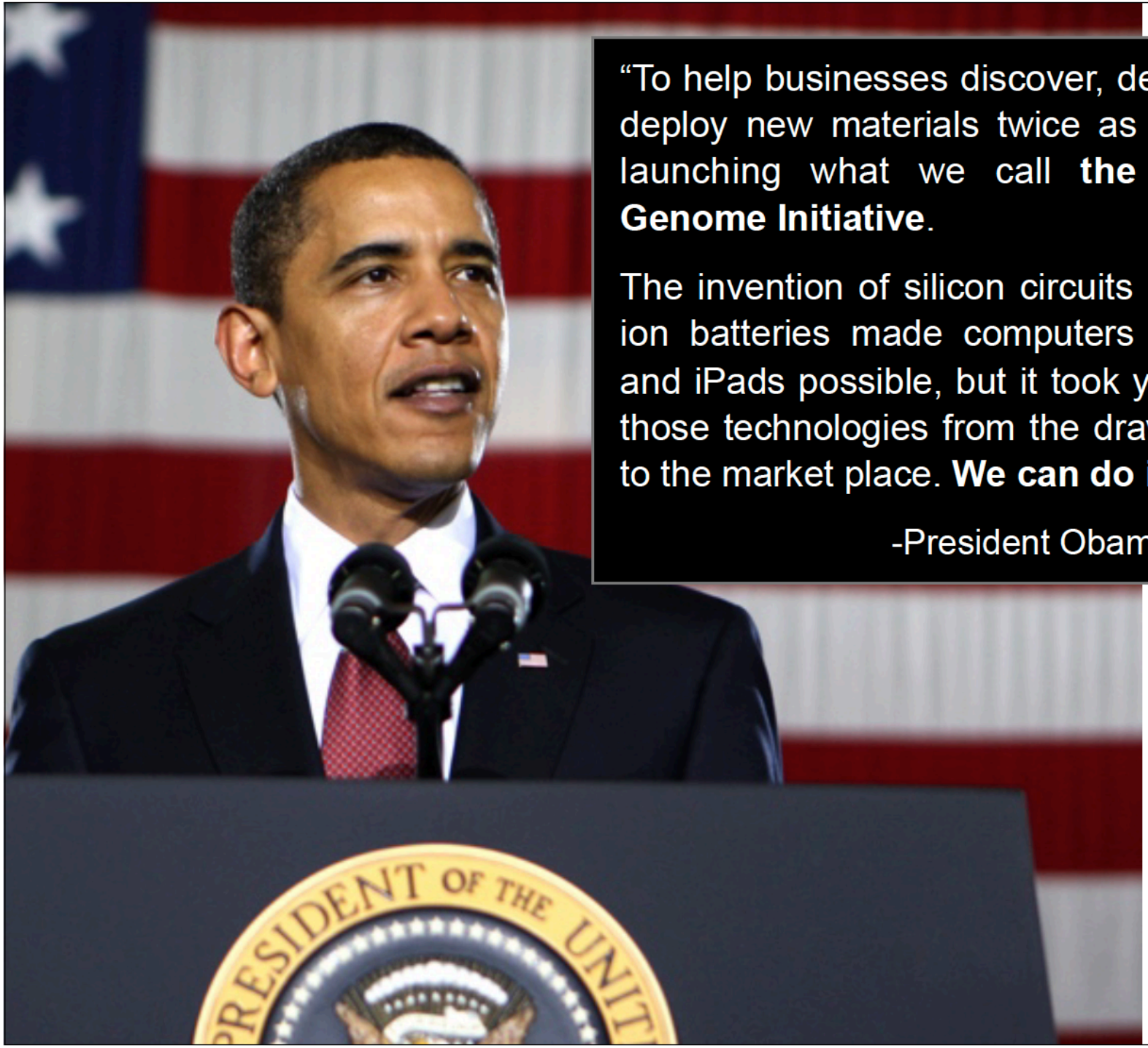
Marco Buongiorno Nardelli

University of North Texas

The AFLOW Consortium - Duke Center for Materials Genomics

QUANTUM ESPRESSO Foundation

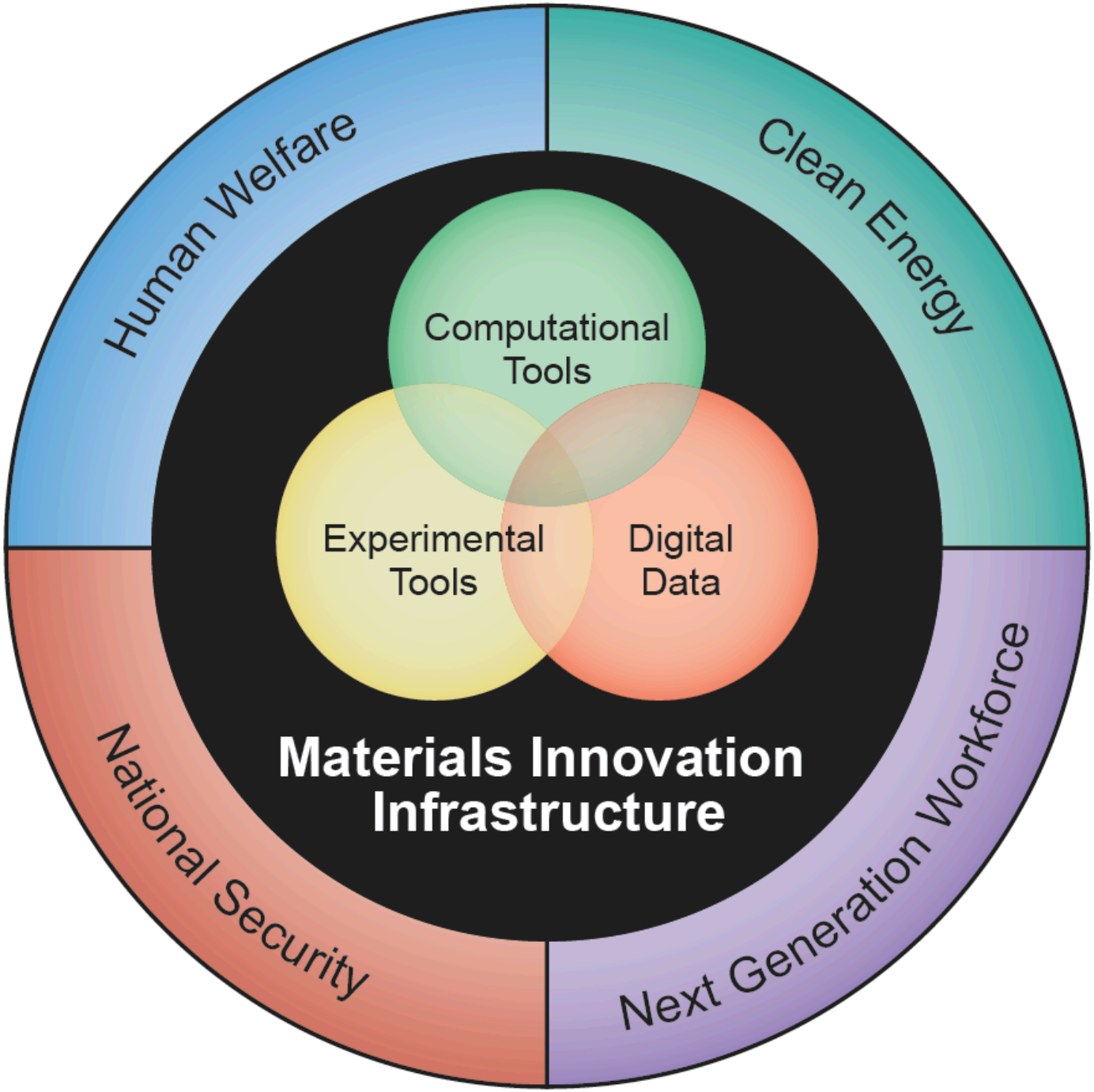
http://ermes.unt.edu



"To help businesses discover, develop, and deploy new materials twice as fast, we're launching what we call **the Materials Genome Initiative**.

The invention of silicon circuits and lithium ion batteries made computers and iPods and iPads possible, but it took years to get those technologies from the drawing board to the market place. **We can do it faster.**"

-President Obama (6/11)


Human Welfare
Clean Energy
Computational Tools
Experimental Tools
Digital Data
Materials Innovation Infrastructure
National Security
Next Generation Workforce

What is HT computational materials

AFLOW
Automatic-FLOW for Materials Discovery

procedures of synthesis. Is there another way? Indeed, this is the burgeoning area of computational materials science called 'high-throughput' (HT) computational materials design. It is based on the marriage between computational quantum-mechanical–thermodynamic approaches[1,2] and a multitude of techniques rooted in database construction and intelligent data mining[3]. The concept is simple yet powerful: create a large database containing the calculated thermodynamic and electronic properties of existing and hypothetical materials, and then intelligently interrogate the database in the search of materials with the desired properties. Clearly, the entire construct should be validated by reality, namely the existing materials must be predicted correctly and the hypothetical ones

nature materials

REVIEW ARTICLE

PUBLISHED ONLINE: 20 FEBRUARY 2013 | DOI: 10.1038/NMAT3568

The high-throughput highway to computational materials design

Curtarolo*, Hart, Buongiorno Nardelli, Mingo, Sanvito, Levy

DOI: 10.1038/NMAT3568 (March 2013)

COMPUTATIONAL high-throughput

AFLOW
Automatic-FLOW for Materials Discovery

The practical implementation of computational HT is highly non-trivial. The method is employed in three strictly connected steps: (i) virtual materials growth: thermodynamic and electronic structure calculations of materials[3,23]; (ii) rational materials storage: systematic storage of the information in database repositories[24,25]; (iii) materials characterization and selection: data analysis aimed at selecting novel materials or gaining new physical insights[15,19,26].

Volume, Variety, Velocity, Veracity

nature materials

REVIEW ARTICLE

PUBLISHED ONLINE: 20 FEBRUARY 2013 | DOI: 10.1038/NMAT3568

The high-throughput highway to computational materials design

Curtarolo*, Hart, Buongiorno Nardelli, Mingo, Sanvito, Levy

DOI: 10.1038/NMAT3568 (March 2013)

Consortium of QM calculations

Computational Materials Science 58 (2012) 227–235

Contents lists available at SciVerse ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

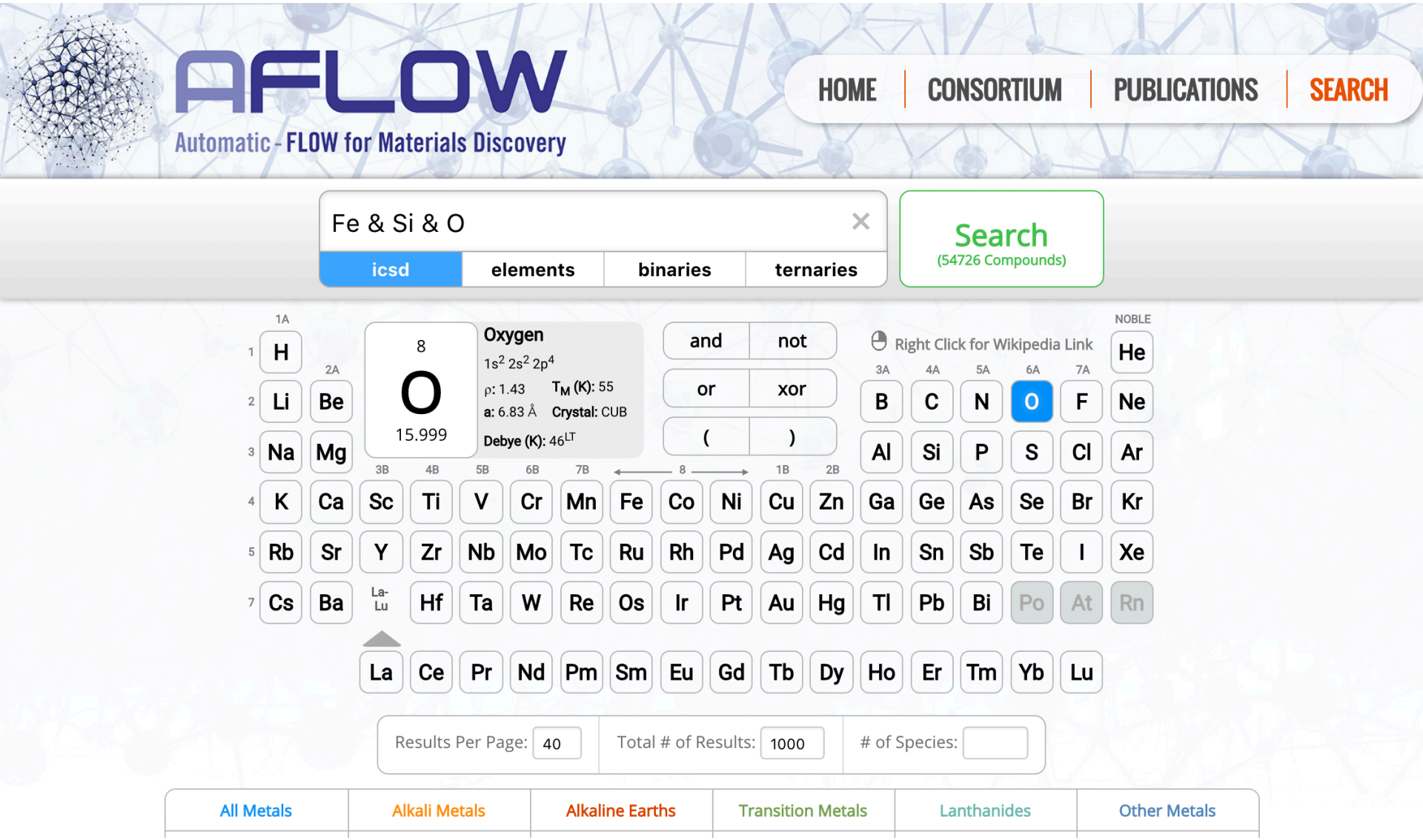
AFLOWLIB.ORG: A distributed materials properties repository from high-throughput *ab initio* calculations

Stefano Curtarolo [a,b,c,*], Wahyu Setyawan [a,1], Shidong Wang [a], Junkai Xue [a], Kesong Yang [a], Richard H. Taylor [a], Lance J. Nelson [d], Gus L.W. Hart [d], Stefano Sanvito [e], Marco Buongiorno-Nardelli [f,c], Natalio Mingo [g], Ohad Levy [h]

Advanced Search

> 1.64M entries
> 162M properties

Marco Buongiorno Nardelli - 2017



Marco Buongiorno Nardelli - 2017

AFLOW

Automatic - FLOW for Materials Discovery

Fe & Si & O

icsd | elements | binaries | ternaries

Search
(54726 Compounds)

Reset Search

Aflow is a globally available database of **1,281,724** material compounds - with over **115,355,160** calculated properties (and growing).

[First] [1-40] [41-80] [81-120] [121-136] [Last] *136 results found*

Index	▲chemical formula (compound)	AFLOWLIB Unique Identifier (auid) aflow:...	AFLOWLIB Uniform Resource Locator (aurl) aflowlib.duke.edu:/AFLOWDATA/...
1	$Al_4Ca_4Fe_2H_2O_{26}Si_6$	496e7120657dd276	ICSD_WEB/MCL/Al2Ca2Fe1H1O13Si3_ICSD_63661
2	$Al_4Ca_4Fe_2H_2O_{26}Si_6$	8d1e96ea7a827f93	ICSD_WEB/MCL/Al2Ca2Fe1H1O13Si3_ICSD_34209
3	$Al_8Fe_{12}O_{48}Si_{12}$	efbea507f1afa2a8	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_96736
4	$Al_8Fe_{12}O_{48}Si_{12}$	899f1cb308f34d3d	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_28031
5	$Al_8Fe_{12}O_{48}Si_{12}$	84ea3e5339926aac	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_80673
6	$Al_8Fe_{12}O_{48}Si_{12}$	4bf157c8f31063eb	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_28030
7	$Al_8Fe_{12}O_{48}Si_{12}$	8bbcf763a2a88c41	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_96737
8	$Al_8Fe_{12}O_{48}Si_{12}$	44c87fd8b9f299a8	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_89677
9	$Al_8Fe_{12}O_{48}Si_{12}$	5671101d5fcb5f55	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_71891
10	$Al_8Fe_{12}O_{48}Si_{12}$	cb9588884c145031	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_96735
11	$Al_8Fe_{12}O_{48}Si_{12}$	f7df4b337b406877	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_80672
12	$Al_8Fe_{12}O_{48}Si_{12}$	f8c5150c76cdcffd	ICSD_WEB/BCC/Al2Fe3O12Si3_ICSD_52395

AFLOW

Automatic - FLOW for Materials Discovery

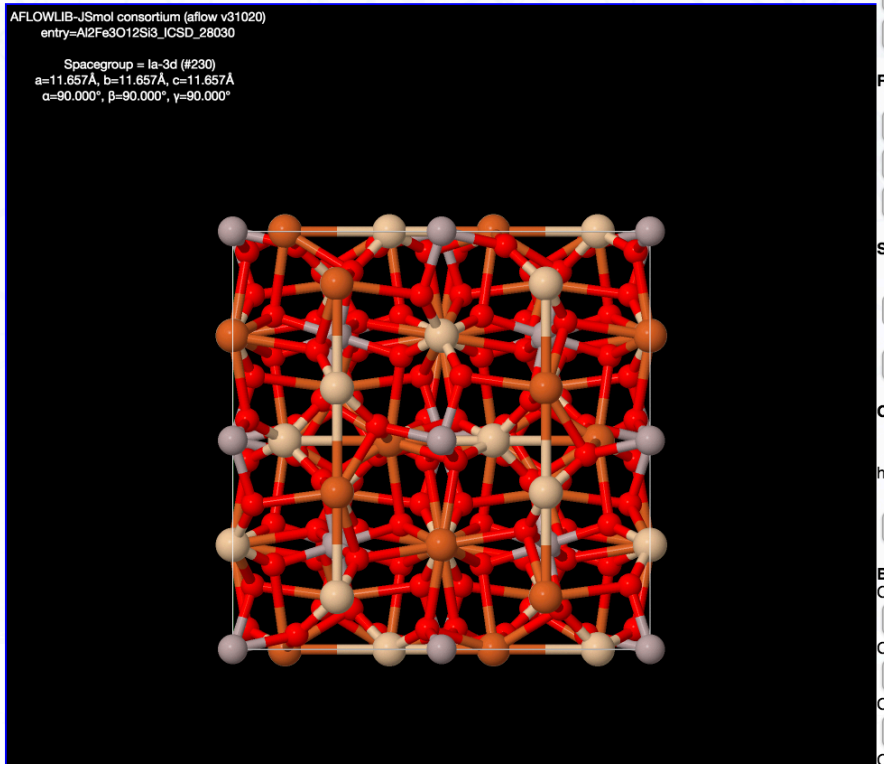
Al2Fe3O12Si3_ICSD_28030

Al8Fe12O48Si12

PERMANENT URL: http://aflowlib.org/material.php?id=Al2Fe3O12Si3_ICSD_28030

AFLOWLIB web entry generator V31020 [built=2015-11-24]

Space Group: Ia-3d (#230)



Visualizer options:

Ball & Stick | Spacefill
Rotation On | Rotation Off
Label On | Label Off

Relaxed structure:

As calculated
Standard conventional [info]
Standard primitive [info]

Supercell:

2 X 2 X 2

BUILD | RESET

Crystallographic planes::

h: 2 k: 2 l: 2

Show plane | RESET

Bader Isosurfaces:

Cutoff = 0.20
Al | Fe | O | Si | All
Cutoff = 0.30
Al | Fe | O | Si | All
Cutoff = 0.40
Al | Fe | O | Si | All
Cutoff = 0.50

AFLOW

Automatic - FLOW for Materials Discovery

Point Group of the Crystal

Crystal Family:	cubic
Crystal System:	cubic
Crystal Class:	hexoctahedral
Point Group (Herm. Maug.):	m-3m
Point Group (Schoenflies):	O_h
Point Group Orbifold:	*432
Point Group Type:	centrosymmetric
Point Group Order:	48
Point Group Structure:	2 x symmetric

Bravais Lattice of the Lattice

Bravais Lattice Primitive	BCC
Bravais Lattice Variation:	BCC
Bravais Lattice System:	cubic

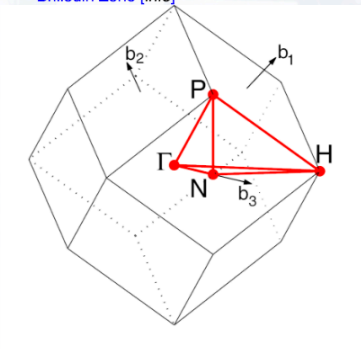
Superlattice

Superlattice Primitive unit cell:	BCC
Superlattice Variation:	BCC
Superlattice Lattice System :	cubic
Superlattice Pearson Symbol:	cI160

Reciprocal Space Lattice

Reciprocal Lattices:	a=0.762Å^{-1} b=0.762Å^{-1} c=0.762Å^{-1} α=60.000° β=60.000° γ=60.000°
Volume:	0.313 Å^{-3}
Lattice Primitive:	FCC
Lattice Variation:	FCC

Brillouin Zone [*info*]



BCC path: Γ-H-N-Γ-P-H|P-N

[Setyawan & Curtarolo, DOI: 10.1016/j.commatsci.2010.05.010]

Thermodynamics Properties

Formation Enthalpy_cell:	1e+09 (eV)
Formation Enthalpy_atom:	1e+09 (eV/atom)
Ab-initio energy_cell:	-597.5 (eV)
Ab-initio energy_atom:	-7.4688 (eV/atom)

Bader Atoms in Molecules Properties

Bader Output: [Al2Fe3O12Si3_ICSD_28030_abader.out]

Net Charges [*info*]: Al={2.4931, 2.4931, 2.4890, 2.4931, 2.4931, 2.4931, 2.4890, 2.4931}; Fe={1.4578, 1.4578, 1.4578, 1.4578, 1.4578, 1.4578, 1.4578, 1.4578, 1.4578, 1.4578, 1.4578, 1.4578}; O={-1.5574, -1.5558, -1.5533, -1.5656, -1.5558, -1.5574, -1.5656, -1.5533, -1.5656, -1.5533, -1.5558, -1.5582, -1.5574, -1.5558, -1.5541, -1.5656, -1.5558, -1.5582, -1.5656, -1.5541, -1.5656, -1.5541, -1.5558, -1.5582, -1.5582, -1.5558, -1.5541, -1.5656, -1.5558, -1.5582, -1.5656, -1.5541, -1.5656, -1.5541, -1.5558, -1.5574, -1.5582, -1.5558, -1.5533, -1.5656, -1.5558, -1.5574, -1.5656, -1.5533, -1.5656, -1.5533, -1.5558, -1.5574}; Si={3.1138, 3.1138, 3.1138, 3.1138, 3.1138, 3.1138, 3.1138, 3.1138, 3.1138, 3.1138, 3.1138, 3.1138} (electrons)

Atomic Volumes [*info*]: Al={3.4047, 3.4047, 3.4286, 3.4047, 3.4047, 3.4047, 3.4286, 3.4047}; Fe={10.0047, 10.0009, 10.0009, 10.0047, 10.0047, 10.0016, 10.0047, 10.0016, 10.0009, 10.0047, 10.0047, 10.0009}; O={12.5738, 12.5732, 12.5612, 12.6184, 12.5732, 12.5738, 12.6184, 12.5612, 12.6184, 12.5612, 12.5732, 12.5858, 12.5738, 12.5732, 12.5725, 12.6184, 12.5732, 12.5858, 12.6184, 12.5725, 12.6184, 12.5725, 12.5732, 12.5858, 12.5858, 12.5732, 12.5725, 12.6184, 12.5732, 12.5858, 12.6184, 12.5725, 12.6184, 12.5725, 12.5732, 12.5738, 12.5858, 12.5732, 12.5612, 12.6184, 12.5732, 12.5738, 12.6184, 12.5612, 12.6184, 12.5612, 12.5732, 12.5738}; Si={3.3890, 3.3890, 3.3890, 3.3890, 3.3890, 3.3890, 3.3890, 3.3890, 3.3890, 3.3890, 3.3890, 3.3890} (Angst^3)

AFLOW

Automatic - FLOW for Materials Discovery

Electronic Properties

Band Gap: 3.328 eV (insulator)

Fit Band Gap:

Magnetic Moment: 48.001 μ_B

Magnetic Moment/atom:

Electron Mass(FIX): XXX

Hole Mass(FIX):

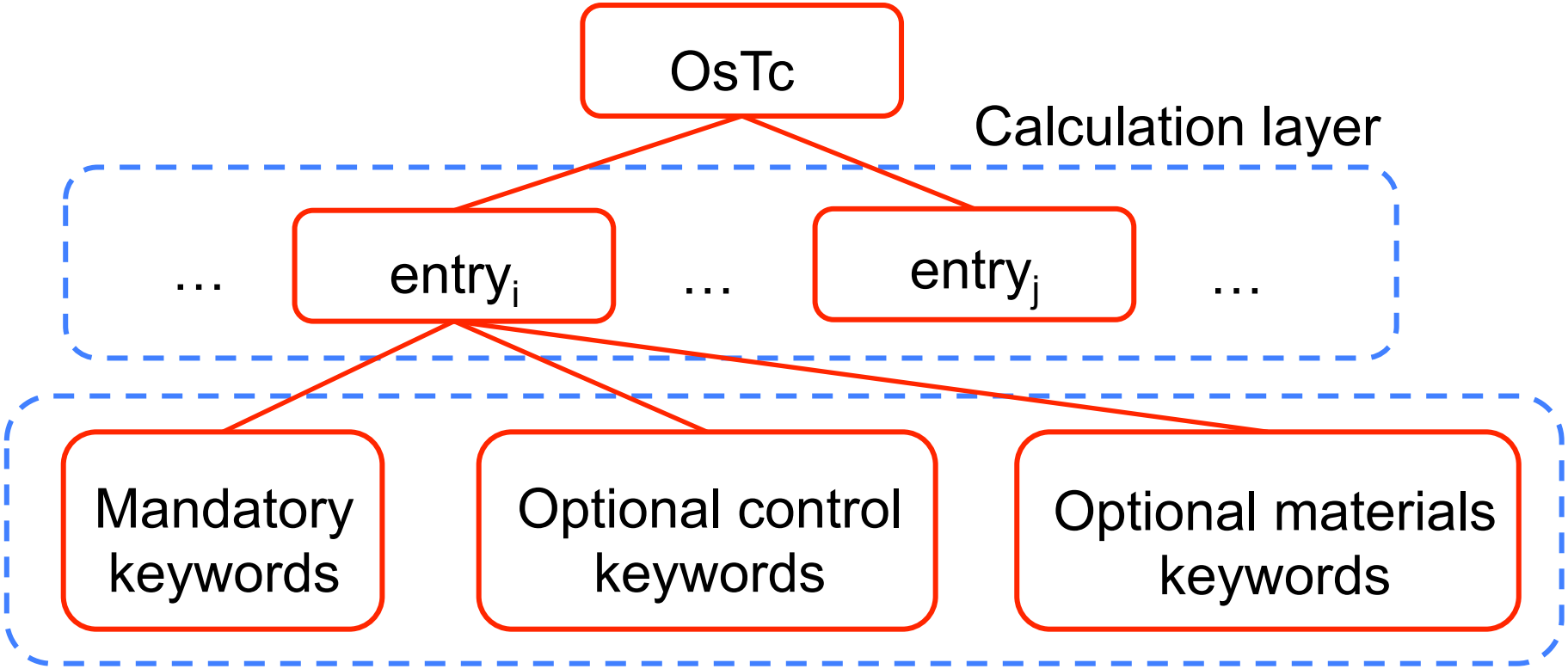
Band Gap Type: Direct

Spin Polarization (E_F): 0.000 μ_B

Spin Decomposition per atoms: {0.001,0.001,0.001,0.001,0.001,0.001,0.001,0.001,3.759,3.759,3.759,3.759,3.759,3.759,3.759,3.759,3.759,3.759,3.759,3.759,0.037,0.037,0.037

AFLOWLIB Rest-API

- For LIB2 binary alloys, each set contains entries with different stoichiometries and symmetries, denoted by numbers



- Each entry has set of keywords with material properties

An entry in detail: multi-layer system

```
aurl=aflowlib.duke.edu:AFLOWDATA/LIB3_RAW/AgCoMn_pv/T0002.A2BC
auid=aflow:AgCoMn_pv/T0002.A2BC:PAW_PBE
aapi=1.0
keywords=aurl,auid,aflow_api_version,code,compound,prototype,nspecies,...
aflowlib_entry_date=20140130_20:34:00_GMT-5
aflowlib_entry_version=30794
aflow_version=aflow30293
calculation_cores=1
calculation_memory=539
calculation_time=18347.2
corresponding=Stefano_Sanvito_sanvitos@tcd.ie
loop=thermodynamics,bands,magnetic
node_CPU_Cores=12
node_CPU_MHz=2661
node_CPU_Model=Intel(R)_Xeon(R)_CPU_X5650_@_2.67GHz
node_RAM_GB=24
code=vasp.4.6.35
composition=2,1,1
compound=Ag2Co1Mn1
density=8.94193
eentropy=0
eentropy_atom=0
Egap=0
energy=-20.4051
energy_atom=-5.10128
enthalpy=-20.4051
enthalpy_atom=-5.10128
enthalpy_formation=1.51248
```

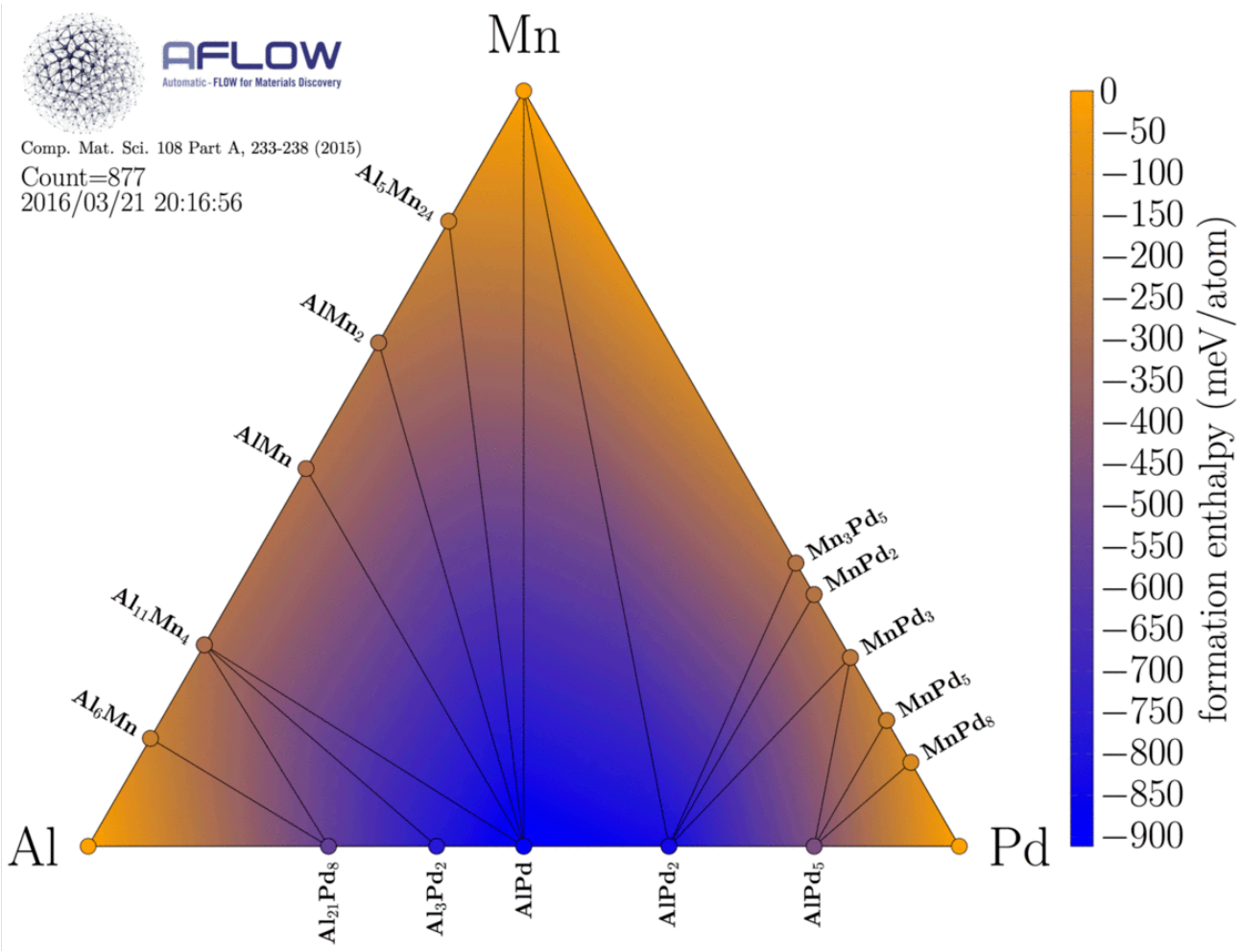
mandatory **keywords**

optional control **keywords**

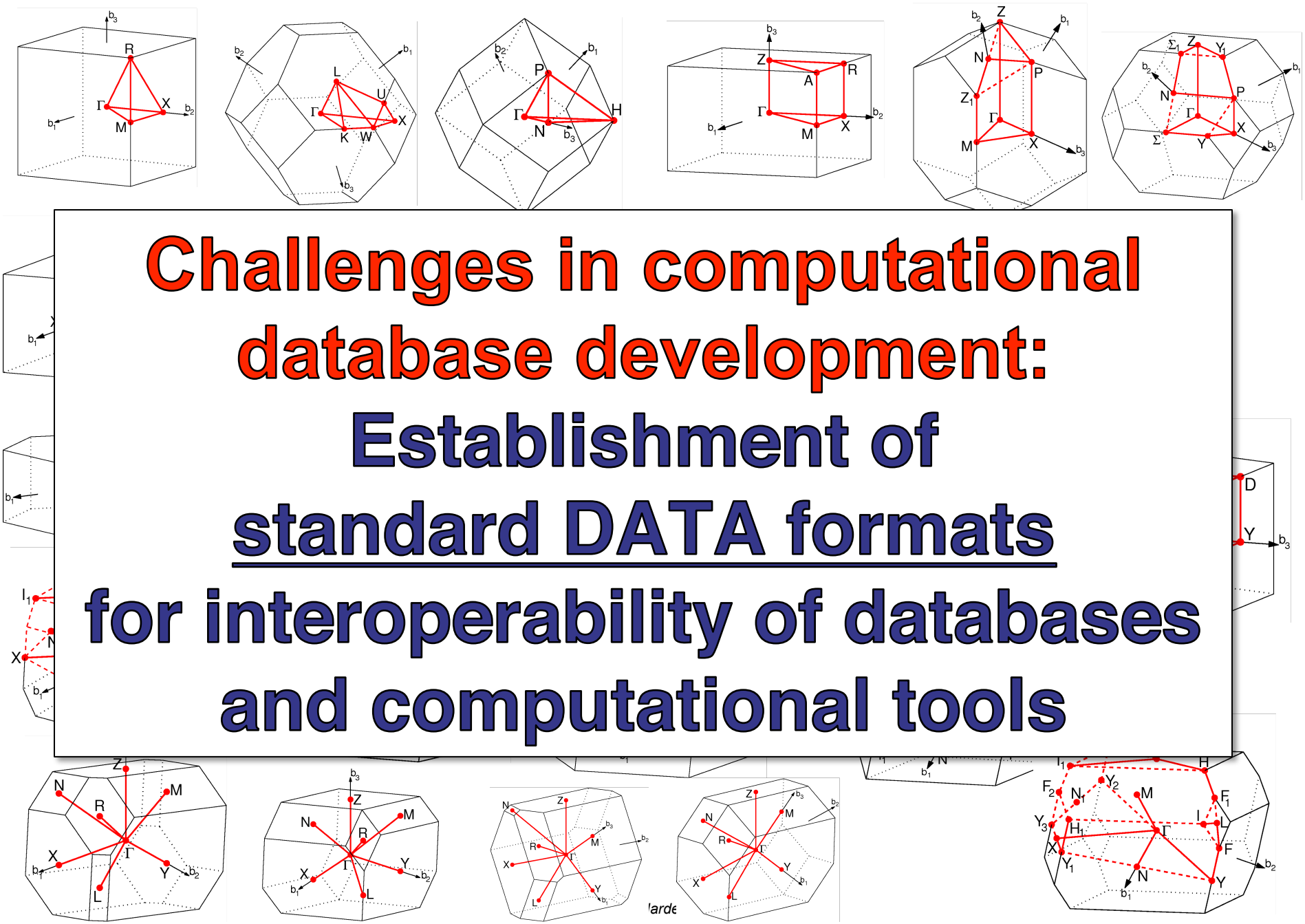
optional materials **keywords**

Comp. Mat. Sci. **93**, 178-192 (2014).

AFLOWLIB Rest-API



STANDARD in Real Space and Reciprocal Space



Challenges in computational database development:

Establishment of

<u>standard DATA formats</u>

for interoperability of databases and computational tools

Computational Materials Science 108 (2015) 233–238

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Editor's Choice

The AFLOW standard for high-throughput materials science calculations

Camilo E. Calderon [a], Jose J. Plata [a], Cormac Toher [a], Corey Oses [a], Ohad Levy [a,1], Marco Fornari [b], Amir Natan [c], Michael J. Mehl [d], Gus Hart [e], Marco Buongiorno Nardelli [f], Stefano Curtarolo [g,*]

ARTICLE INFO

Article history:
Received 31 May 2015
Received in revised form 6 July 2015
Accepted 7 July 2015

Keywords:
High-throughput
Materials genomics
AFLOWLIB
VASP

ABSTRACT

The Automatic-Flow (AFLOW) standard for the high-throughput construction of materials science electronic structure databases is described. Electronic structure calculations of solid state materials depend on a large number of parameters which must be understood by researchers, and must be reported by originators to ensure reproducibility and enable collaborative database expansion. We therefore describe standard parameter values for *k*-point grid density, basis set plane wave kinetic energy cut-off, exchange–correlation functionals, pseudopotentials, DFT+U parameters, and convergence criteria used in AFLOW calculations.

© 2015 Elsevier B.V. All rights reserved.

URL for communication layer

http://aflowlib.duke.edu/AFLOWDATA/ICSD_WEB/BCT/Ag3Cu1S2_ICSD_163983/

Added logical operators:

- Captures some of the niceties of SQL but simple
- Fits neatly into the query portion of a URL
- Does not interfere with existing ?key=value# nomenclature
- Automatic in its property inspection

e.g. search for compounds containing potassium, chlorine and oxygen

http://aflowlib.duke.edu/search/API/?species(Ag,Cu,S)

Default format is an array of JSON objects, shown 40 at a time; which set is returned is controlled by “paging()” command

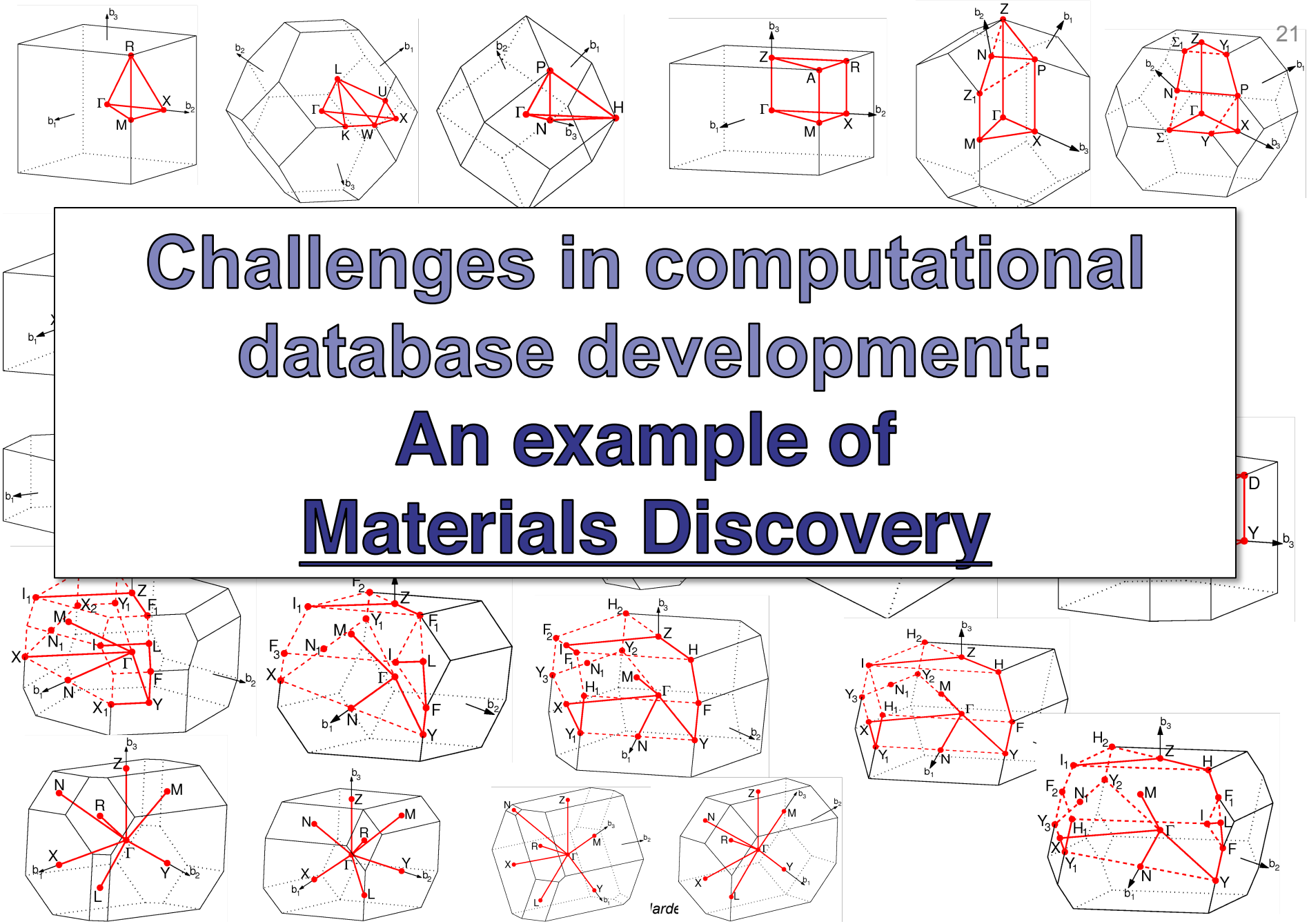
AFLOW
Automatic - FLOW for Materials Discovery

AFLUX

EXAMPLES

aflowlib.duke.edu/search/API/?species	Show default DB selection with the species property
aflowlib.duke.edu/search/API/?species(S)	Show entries that have sulfur
aflowlib.duke.edu/search/API/?species((Zn:Fe),S),$nspecies(2)	Show entries that have ZnS or FeS only
aflowlib.duke.edu/search/API/?species,catalog(icsd:lib2)	Show from both ICSD and Lib2 databases
aflowlib.duke.edu/search/API/?Egap(1.2*,*1.5),species(S)	Show sulfur containing at about Shockley-Queisser
aflowlib.duke.edu/search/API/?Egap(1.6*),energy_atom	Show entries with gap >= 1 along with energy per atom
aflowlib.duke.edu/search/API/?Egap(1*),paging(2)	Show entries with gap >=1, second set of 40 entries

STANDARD in Real Space and Reciprocal Space



Challenges in computational database development:

An example of Materials Discovery

Materials Genome: genes+descriptors

Table 1 | Examples of descriptors introduced in the literature. Nature Mater. **12**, 191 (2013)

Problem	Combination of materials properties (gene)	Descriptor
Structure stability: convex hull of an alloy system	Formation enthalpy (H_f) as a function of concentration (x) and the enthalpies (H) of A and B.	$H_f(x) = H(A_{1-x}B_x) - (1-x)H(A) - xH(B)$
Phase stability in off-lattice alloys PR	Spectral decomposition of alloy vector-energies (E_{np}, n-rows = species, p-columns = c alysis coefficients (	$E_{np} \simeq \alpha_1 E_{np} + \cdots + \alpha_{np} E_{np} + \epsilon(d)$
Nano PR	Ratio of the av) (ref. 15).	
Topo Nat	Variational ra rbit derivative stra s at k, a_0 lattice)[16].	
Powe l (spectroscopic limited maximum efficiency) PRL **108**, 068701 (2012)	Ratio of the m ident solar energy density (P_{in}) — a function (η) of the radiative electron–hole recombination current (f_r) and the photon absorptivity ($\alpha(E)$) — versus bandgap energy (E_g)[62].	$\eta(\alpha(E), f_r) = P_m/P_{in}; E_g$
Non-proportionality in scintillators IEEE Trans. Nucl. Sci. **56**, 2989 (2009)	Maximum mismatch between effective masses of electrons (m_e) and holes (m_h)[75].	$\hat{\chi}_{np} \equiv \max\left(\frac{m_e}{m_h}, \frac{m_h}{m_e}\right)$
Morphotropic phase boundary piezoelectrics PRB **84**, 014103 (2011)	Energy proximity between tetragonal, rhombohedra and rotational distortions (ΔE_p). Angular coordinate (α_{AB}) of the energy minimum in the A–B off-centerings energy map for ABO_3 systems[79].	$\Delta E_p \leq 0.5$ eV $\alpha_{AB} \approx 45°$

the "problem"

qualitative picture

quantitative picture

The search for functionality: Topological Insulators

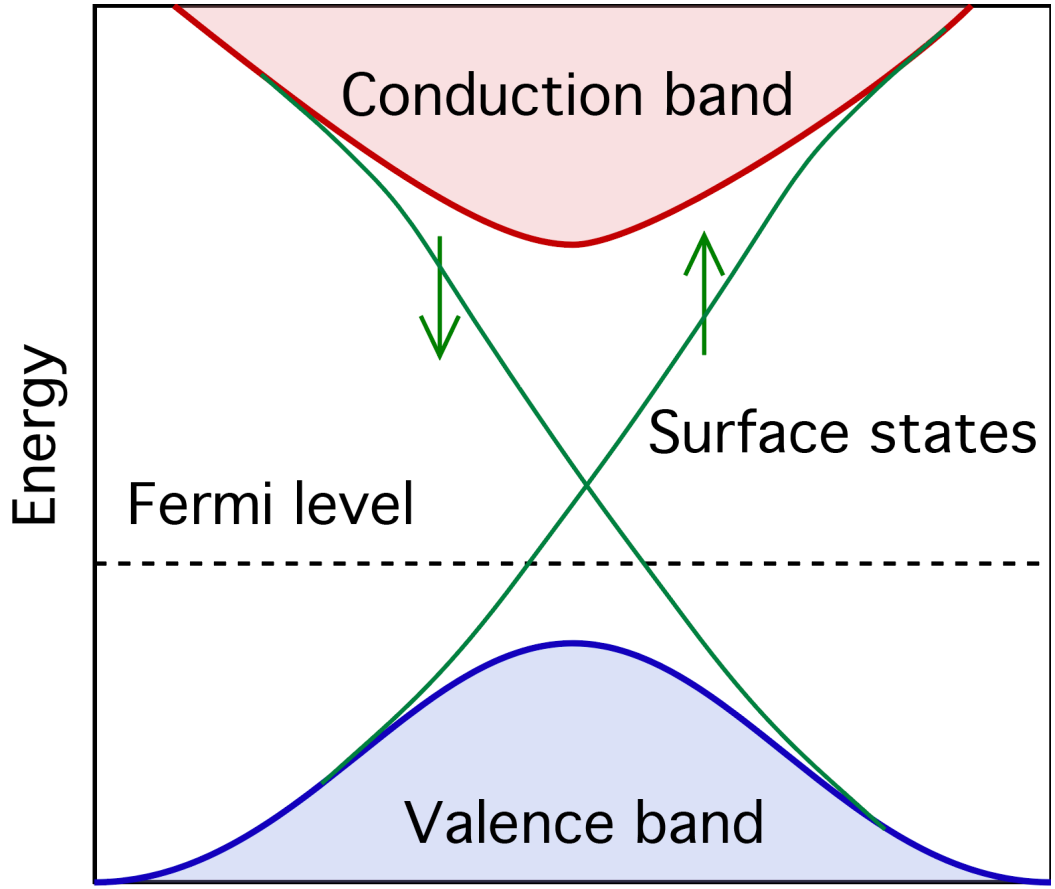
Table 1 | Examples of descriptors introduced in the literature. Nature Mater. 12, 191 (2013)

Problem	Combination of materials properties (gene)	Descriptor
Structure stability: convex hull of an alloy system	Formation enthalpy (H_f) as a function of concentration (x) and the enthalpies (H) of A and B.	$H_f(x) = H(A_{1-x}B_x) - (1-x)H(A) - xH(B)$
Phase stability in off-lattice alloys PRL 91, 135503 (2003)	Spectral decomposition of alloy vector-energies ($E_{n,p}$, n-rows = species, p-columns = configurations) with principal-component-analysis coefficients (α_i) and truncation error ($\epsilon(d)$) (ref. 3).	$E_{n,p} \simeq \alpha_1 E_{n,1} + \cdots + \alpha_{p-1} E_{n,p-1} + \epsilon(d)$
Nanosintered thermoelectrics PRX 1, 021012 (2011)	Ratio of the average power factor (<P>) to the grain size (L) (ref. 15).	$\hat{\chi}_{\text{thermo}} = \frac{\langle P \rangle}{L}$
Topological insulators (epitaxial growth) Nature Mater. 11, 614 (2012)	Variational ratio of spin–orbit distortion versus non-spin–orbit derivative strain (E_k^{SOC}, E_k^{noSOC}, spin/no spin–orbit bandgaps at k, a_0 lattice)[16].	$\hat{\chi}_{\text{TI}} = -\frac{E_k^{SOC}(a_0)/a_0}{\delta E_k^{noSOC}(a)_0/\delta a_0 \vert_{a_0}}$
Power conversion efficiency of a solar cell (spectroscopic limited maximum efficiency) PRL 108, 068701 (2012)	Ratio of the maximum output power density (P_m) to the incident solar energy density (P_{in}) — a function (η) of the radiative electron–hole recombination current (f_r) and the photon absorptivity ($\alpha(E)$) — versus bandgap energy (E_g)[62].	$\eta(\alpha(E), f_r) = P_m/P_{in}; E_g$
Non-proportionality in scintillators IEEE Trans. Nucl. Sci. 56, 2989 (2009)	Maximum mismatch between effective masses of electrons (m_e) and holes (m_h)[75].	$\hat{\chi}_{np} = \max\left(\frac{m_e}{m_h}, \frac{m_h}{m_e}\right)$
Morphotropic phase boundary piezoelectrics PRB 84, 014103 (2011)	Energy proximity between tetragonal, rhombohedra and rotational distortions (ΔE_p). Angular coordinate (α_{AB}) of the energy minimum in the A–B off-centerings energy map for ABO_3 systems[79].	$\Delta E_p \leq 0.5$ eV $\alpha_{AB} \approx 45°$

AFLOW
Automatic - FLOW for Materials Discovery

The search for functionality: Topological Insulators

material with non-trivial topological order that behaves as an insulator in its interior but whose surface contains conducting states that are symmetry protected by particle number conservation and time reversal symmetry

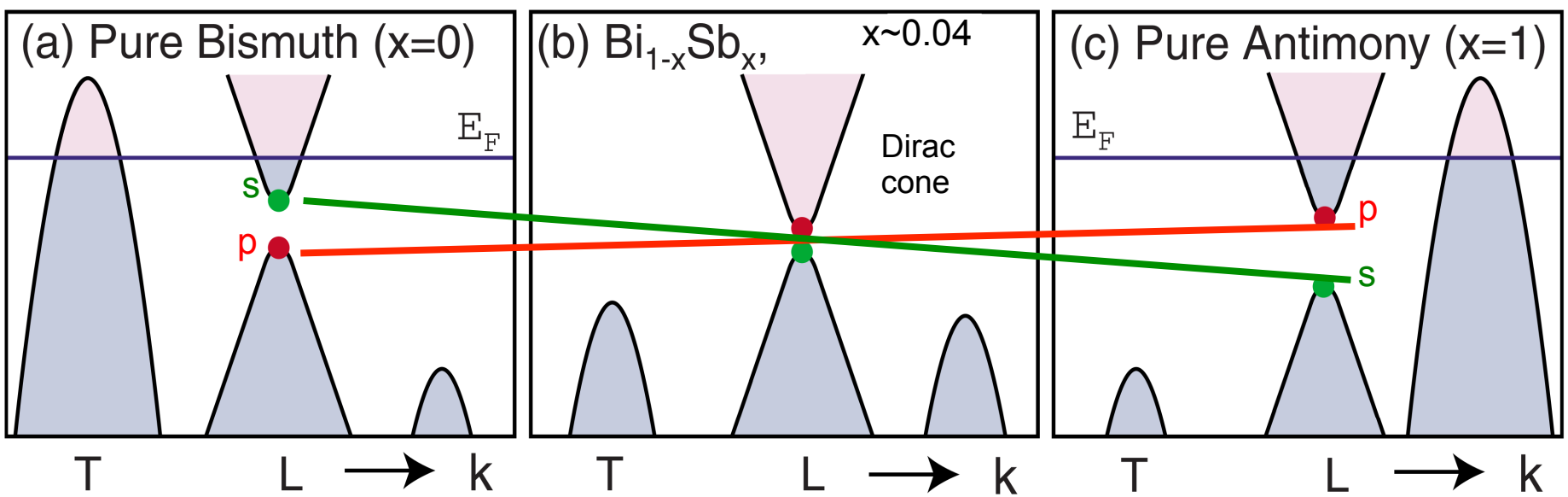


Example: a plan of action

- Scan the aflowlib.org library
- Need of a **DESCRIPTOR (need to grow… epixially).**
- search for combination of heavy metals (potential strong spin-orbit coupling)
- search for ideal band structures with appropriate gaps
- calculate band structure with LS (thousand of compounds)
- calculate the bands for surfaces to see localized conducting surface stares
- usually they contain Bi and/or Sb,Te, Pb.

Time Reversal Invariant Momenta (TRIMs)

$$\Gamma_{i=(n_1 n_2 n_3)} = (n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2 + n_3 \mathbf{b}_3)/2, \quad \text{with} \quad n_j = 0, \ 1.$$



*Nature Materials,**11**(7), 614-619 (2012)*

Example: a plan of action

- Scan the aflowlib.org library
- Need of a **DESCRIPTOR (need to grow… epixially).**
- search for combination of heavy metals (potential strong spin-orbit coupling)
- search for ideal band structures with appropriate gaps
- calculate band structure with LS (thousand of compounds)
- calculate the bands for surfaces to see localized conducting surface stares
- usually they contain Bi and/or Sb,Te, Pb.

nature materials

ARTICLES

PUBLISHED ONLINE: 13 MAY 2012 | DOI: 10.1038/NMAT3332

A search model for topological insulators with high-throughput robustness descriptors

Kesong Yang[1], Wahyu Setyawan[2], Shidong Wang[1], Marco Buongiorno Nardelli[3,4] and Stefano Curtarolo[1,4,5]★

Nature Materials, ***11****(7), 614-619 (2012)*

New compounds: Cs{Sn,Pb,Ge}{Cl,Br,I}$_3$

Table 1 | Properties of bulk and simulated epitaxial structures.

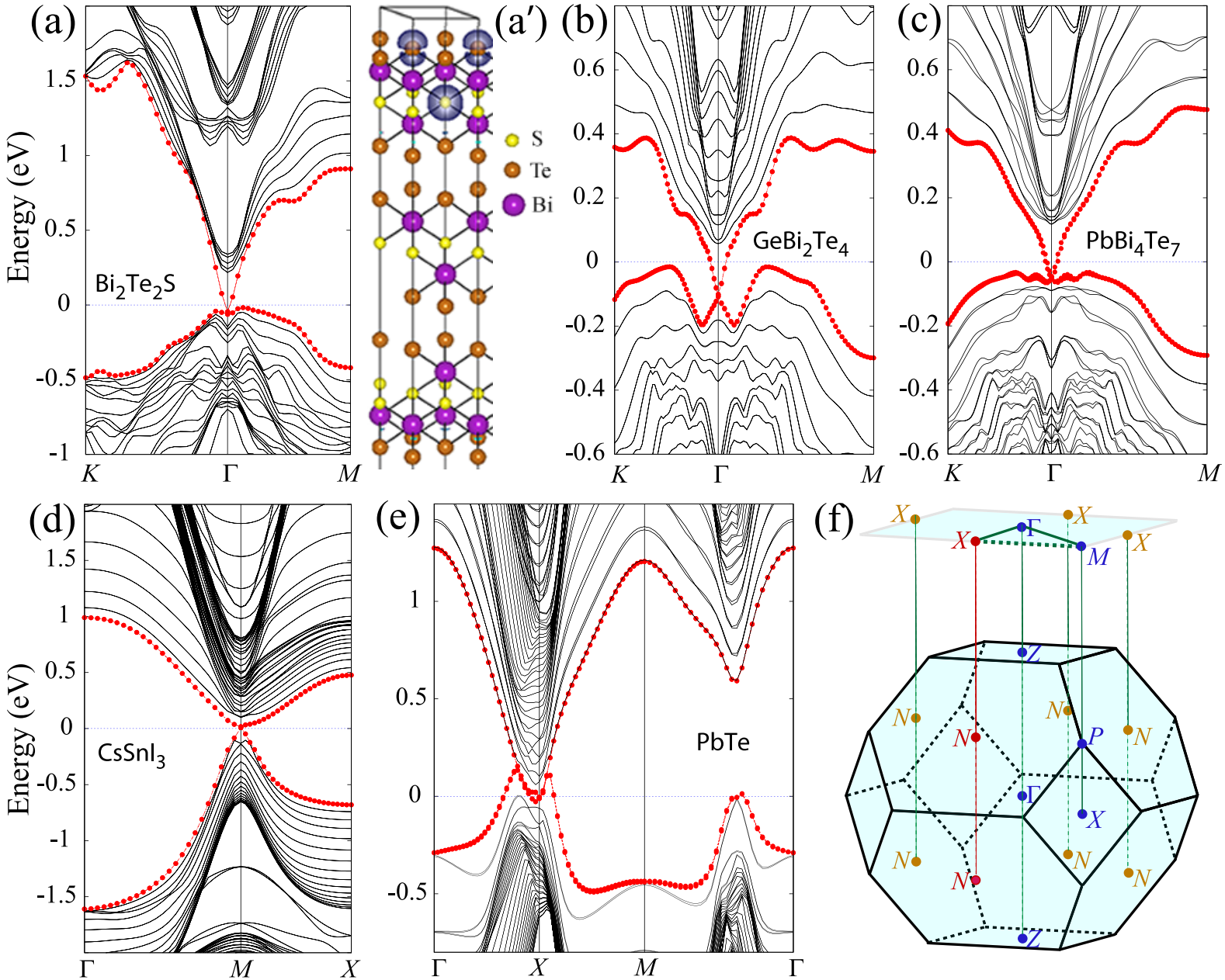
			Bulk				Simulated epitaxial growth (a optimized, c/a free)									
Compound	Space group	ICSD # ref. 31	Pearson symbol	Latt. ref. 39	Exp. a_0, c_0	DFT a_0, c_0	Pearson symbol	Latt. cleav.	a_{crit} (Å)	E_k^{SOC} 'ref'	Ref. lattice	E_g^{SOC} 'ref'	TRIM (mult.)	ΔE_k @ (a_0, c_0)	E_k^{SOC} @ (a_0, c_0)	$\hat{\chi}_{TI}$ (%)
Sb_2Te_2S	$R\bar{3}m$	–	hR5	rhl$_1$	–	4.192	hR5	rhl$_1$	$1.006a_0$	−0.106	$1.019a_0$	0.106	Γ (1)	0.21	0.043	0.6
						31.001		(0001)	$0.993c_0$		$0.975c_0$	(D)				PF
Bi_2Te_2S	$R\bar{3}m$	617050	hR5	rhl$_1$	4.33	4.297	hR5	rhl$_1$	$0.987a_0$	−0.089	a_0	0.089	Γ (1)	0.62	−0.089	−1.3
Fig. 2a					30.07	31.513		(0001)	$1.013c_0$		c_0	(D)				R
$SnSb_2Te_4$	$R\bar{3}m$	30392	hR7	rhl$_1$	4.312	4.389	hR7	rhl$_1$	$0.999a_0$	−0.065	$1.011a_0$	0.065	Z(1)	0.22	0.013	−0.1
					41.72	42.347		(0001)	$0.998c_0$		$0.984c_0$	(D)				PF
$PbSb_2Te_4$	$R\bar{3}m$	250250	hR7	rhl$_1$	4.35	4.413	hR7	rhl$_1$	$0.988a_0$	−0.017	a_0	0.017	Z(1)	0.35	−0.017	−1.2
					41.712	42.792		(0001)	$1.011c_0$		c_0	(D)				R
$PbBi_4Se_7$	$P\bar{3}m1$	ref. 49	hP12	hex	4.25	4.216	hP12	hex	$1.018a_0$	−0.016	$1.023a_0$	0.016	A(1)	0.41	0.128	2.3
					22.68	23.839		(0001)	$0.971c_0$		$0.966c_0$	(D)				PF
$CsSnCl_3$	$Pm\bar{3}m$	28082	cP5	cub	5.504	5.618	tP5	tet	$0.951a_0$	−0.281	$0.936a_0$	0.111	A(1)	0.34	0.646	−4.9
					5.504	5.618		(001)	$1.022c_0$		$1.209c_0$	(I)				HF
$CsPbCl_3$	$Pm\bar{3}m$	29072	cP5	cub	5.605	5.733	tP5	tet	$0.914a_0$	−0.450	$0.890a_0$	0.354	A(1)	1.11	1.073	−8.6
					5.605	5.733		(001)	$1.037c_0$		$1.050c_0$	(I)				HF
$CsGeBr_3$	$Pm\bar{3}m$	80320	cP5	cub	5.36	5.603	tP5	tet	$0.955a_0$	−0.055	$0.952a_0$	0.026	A(1)	0.16	0.591	−4.5
					5.36	5.603		(001)	$1.022c_0$		$1.023c_0$	(I)				HF
$CsSnBr_3$	$Pm\bar{3}m$	4071	cP5	cub	5.795	5.884	tP5	tet	$0.972a_0$	−0.099	$0.965a_0$	0.099	A(1)	0.34	0.288	−2.8
					5.795	5.884		(001)	$1.010c_0$		$1.013c_0$	(D)				PF
$CsPbBr_3$	$Pm\bar{3}m$	29073	cP5	cub	5.874	5.993	tP5	tet	$0.934a_0$	−0.120	$0.926a_0$	0.120	A(1)	1.11	0.641	−6.6
					5.874	5.993		(001)	$1.024c_0$		$1.027c_0$	(D)				HF
$CsSnI_3$	$Pm\bar{3}m$	69997	cP5	cub	6.219	6.272	tP5	tet	$0.993a_0$	−0.335	$0.960a_0$	0.169	A(1)	0.39	0.070	−0.7
Fig. 2d					6.219	6.272		(001)	$1.002c_0$		$1.013c_0$	(I)				PF
SnTe	$Fm\bar{3}m$	52489	cF8	fcc	4.471	4.528	tI4	bct$_2$	$1.027a_0$	−0.058	$1.010a_0$	0.058	N (4)	0.15	−0.107	2.7
					6.323	6.404		(001)	$0.998c_0$		$0.999c_0$	(D)				VR

Properties of bulk structure: compound (∗ indicates experimental validation), space group, ICSD number[31], Pearson symbol, Bravais lattice[39], experimental and DFT equilibrium lattices a, c in (Å). Properties under the simulated epitaxial growth condition: Pearson symbol, Bravais lattice with conventional cleavage Miller indices, critical value for band inversion (a_{crit}), SOC band energy difference ($E_k^{SOC}(ref.)$) at the TRIM with the reference lattice, reference lattice, SOC band-gap at the reference lattice (direct/indirect) ($E_g^{SOC}(ref.)$), TRIMs having band inversion with multiplicity[39], SOC energy-gap discrepancy (ΔE_k) at the *ab initio* equilibrium lattice, SOC band energy difference ($E_k^{SOC}(a_0)$) at the TRIM with the *ab initio* equilibrium lattice, HT-descriptor ($\hat{\chi}_{TI}$). The labels below $\hat{\chi}_{TI}$ indicate: F(fragile), R(robust), VR(very robust), PF(potentially feasible), and HF(hardly feasible) (structural and electronic data is available by following the links listed in the Supplementary Information Extended Table).

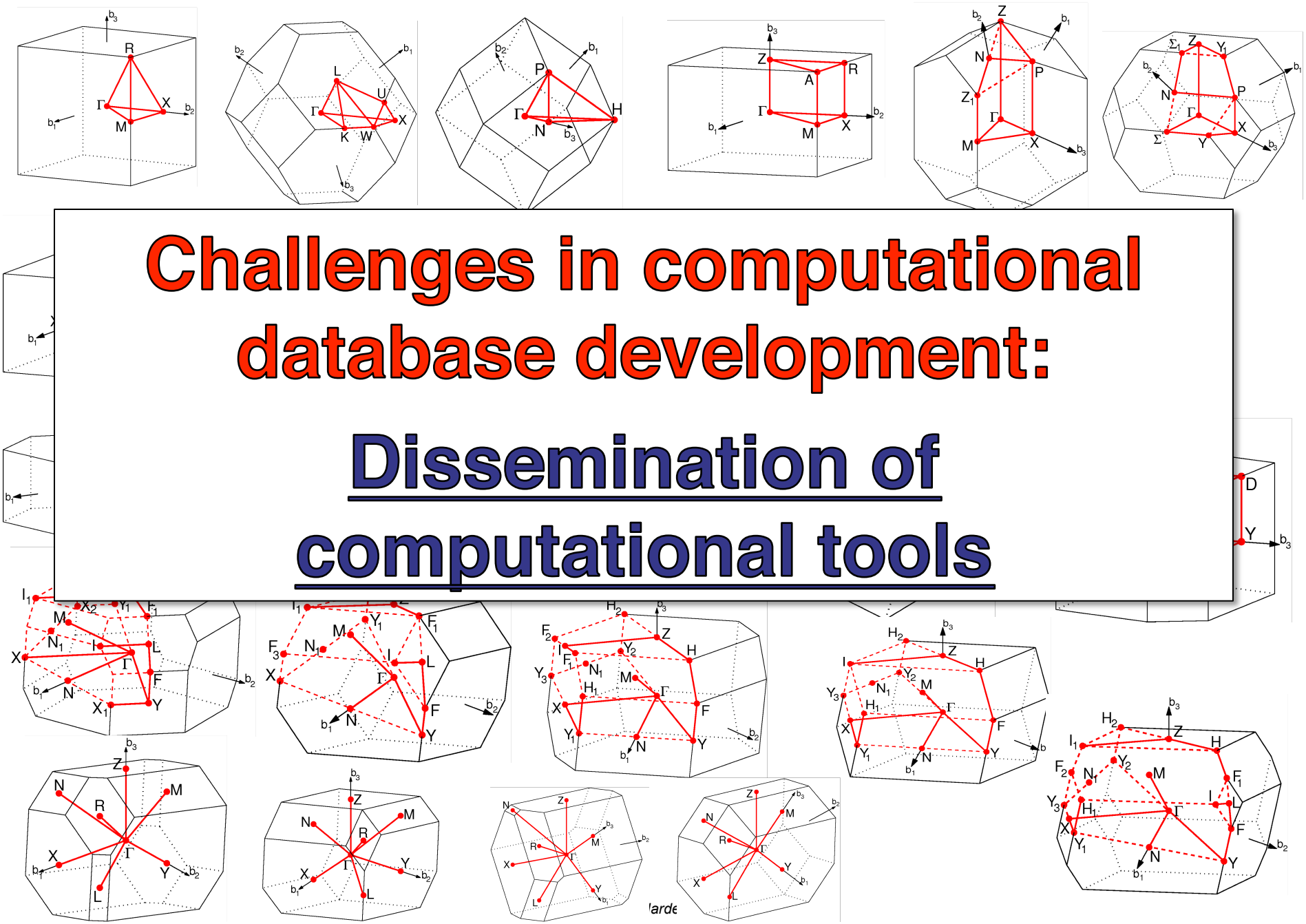
novel ternary halides

if $E_k^{SOC}(a_0) \leq 0$, **robustness**:
$|\hat{\chi}_{TI}| \leq 1\% \Rightarrow$ fragile,
$1\% < |\hat{\chi}_{TI}| \leq 2\% \Rightarrow$ robust,
$2\% < |\hat{\chi}_{TI}| \Rightarrow$ very-robust.

if $E_k^{SOC}(a_0) > 0$, **feasibility**:
$|\hat{\chi}_{TI}| \leq 3\% \Rightarrow$ potentially-feasible,
$3\% < |\hat{\chi}_{TI}| \Rightarrow$ hardly-feasible.


(a)
Energy (eV)
1.5
1
0.5
0
-0.5
-1
Bi_2Te_2S
K
Γ
M
(a′)
S
Te
Bi
(b)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
$GeBi_2Te_4$
K
Γ
M
(c)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
$PbBi_4Te_7$
K
Γ
M
(d)
Energy (eV)
1.5
1
0.5
0
-0.5
-1
-1.5
$CsSnI_3$
Γ
M
X
(e)
1
0.5
0
-0.5
PbTe
Γ
X
M
Γ
(f)
X
Γ
X
M
X
Z
N
N
N
P
N
Γ
X
N
N
N
N
Z

STANDARD in Real Space and Reciprocal Space



Challenges in computational database development:

Dissemination of computational tools

Density Functional Theory

- 1998 Nobel Prize in Chemistry to Walter Kohn and John Pople
- **Key point** - the ground state energy for the hard many-body problem can in principle be found by solving non-interacting electron equations in an effective potential determined only by electron density

Walter Kohn

$$\left(-\frac{1}{2}\nabla^2 + V_{eff}\right)\psi_i = \varepsilon_i \psi_i$$

- The effective potential contains all the information on the external potential and the many-body electron interactions - needs to be approximated! Accurate approximations for the functionals of the density exist: LDA, LSDA, GGA.
- Ab initio pseudopotentials to describe the ions (or all electron calculations possible)

QUANTUMESPRESSO

Home Downloads Docs Tutorials Examples Contact

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

AFLOWπ: A minimalist approach to high-throughput *ab initio* calculations including the generation of tight-binding hamiltonians

Andrew R. Supka[1,2], Troy E. Lyons[1], Laalitha Liyanage[3], Pino D'Amico[4,5], Rabih Al Rahal Al Orabi[1], Sharad Mahatara[1], Priya Gopal[1], Cormac Toher[7,8], Davide Ceresoli[6], Arrigo Calzolari[3,4,5,8], Stefano Curtarolo[3,8], Marco Buongiorno Nardelli[2,8], and Marco Fornari[1,2,8⋆]

ACBN0 calculations

The TB representation of the electronic structure can be exploited to efficiently compute two-electron integrals for the development of local exchange functionals. AFLOWπ implements the direct and self-consistent evaluation of the on-site Hubbard U correction that greatly improves the accuracy of standard DFT calculations [Agapito 2015]. Due to the accurate TB representation, the evaluation of the U parameters for atoms in dif- ferent chemical environments or close to topological defects becomes trivial and AFLOWπ facilitates the investigation of such systems. ACBN0 only demands computational resources comparable to regular LDA or PBE calculations. We have extensively investigated the improvements and the limits of ACBN0 calculation in Refs. [Gopal 2015, Shishkin 2016]. Overall, ACBN0 delivers better accuracy for lattice parameters and bulk moduli, improves the energy band gap on semiconductors and insulators as well as the relative position of occupied d-manifolds in transition metal oxides, and optimizes the phonon dispersions and associated thermal transport properties.

High-Throughput data generation

Electronic structure engine:

QE (LAPW, Gaussian, VASP…)

Testing and Data validation:

ACBN0

BNDs, DOS

Phonons

PAO-TB + e-transport

Development platform

New theories and models

Data analysis tools:

building SQL database

PCA, BiPlots,

Clustering + Classification

Integrated with **AFLOW**

Repository standards

REST-API

Minimalist!

```
import MTF.prep
import MTF.run
import MTF.plot

for pid in range(2,3):

†† init the frame
        aflowkeys = MTF.prep.init('BINARY', 'PROT'+str(pid))
        reffile = MTF.prep.protAflow(pid)
†† finds the variables (MTF variables in the reffile)
        allvars = MTF.prep.extractvars(reffile)
†† it's time to tell what you want to do: all tuples of strings
†† or tuple of numbers
        allvars.update(
        _MTF_PSEUDO_DIR_ = ('./',),
        _MTF_OUTDIR_ = ('./'),
        _MTF_KKK_ = ('4 4 4',),
        _MTF_SSS_ = ('1 1 1',),
        _MTF_ECUTWFC_ = (40.,),
        _MTF_ECUTRHO_ = (200.,),
        _MTF_A_ = ('Ti','Hf'),
        _MTF_B_ = ('C',),
        )
†† prepare the scf/relax/vc-relax calculations
        calcs = MTF.prep.scfs(allvars, reffile)
†† make the dir tree for the calculations, must find the GBRV pseudopo
udodir
        MTF.prep.maketree(aflowkeys, calcs, pseudodir='./PSEUDOs')
†† save the logs file for future calculations and error checking
        MTF.prep.updatelogs(aflowkeys, calcs)
†† run the tree with pw.x (default)
        MTF.run.pwx(calcs,execPrefix='mpiexec -np 12',execPostfix='')

        print '---------'
        print 'SCFs DONE'
        print '---------'

† prepare the dos calculatations
        calcsdos = MTF.prep.doss(calcs)
† save new logs
        MTF.prep.updatelogs(aflowkeys,calcsdos, True)
† run the dos calculations
        MTF.run.pwx(calcsdos,execPrefix='mpiexec -np 12')
† post processing with dos.x
        MTF.run.dos(calcsdos,execPrefix='mpiexec -np 12')
† plot all the DOS  and save the pdf files
        MTF.plot.dos(calcsdos, yLim=[-10,10])

        print '---------'
        print 'DOSs DONE'
        print '---------'
```

AFLOWπ

High-Throughput data generation

PHYSICAL REVIEW X **5,** 011006 (2015)

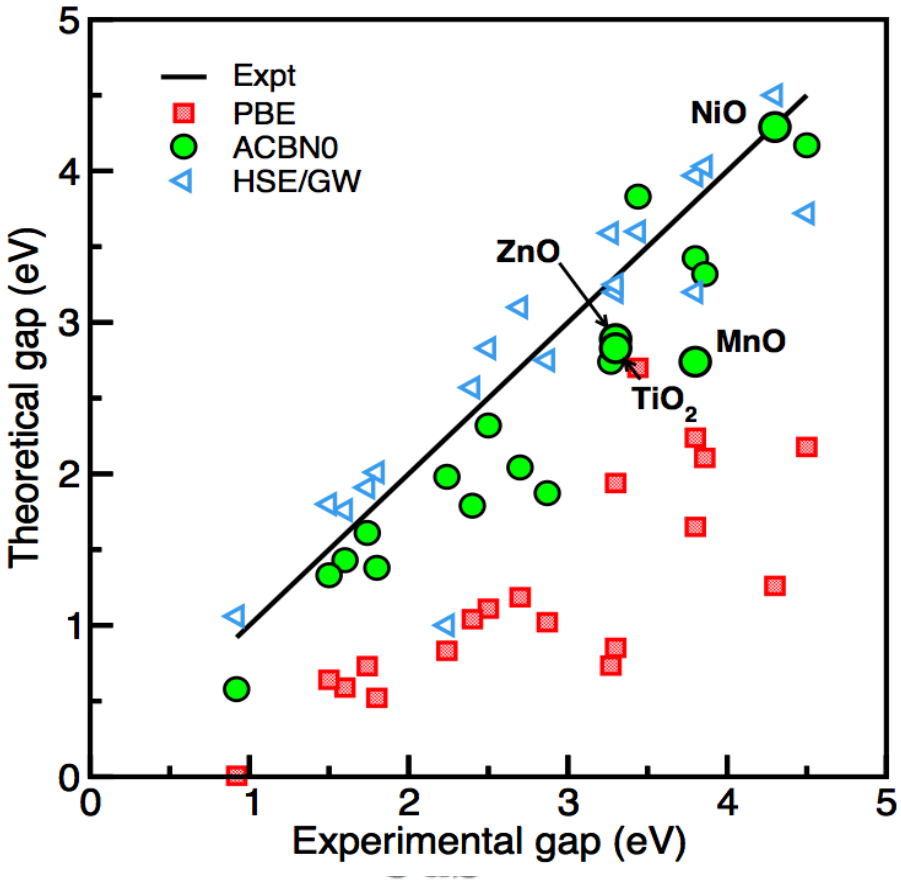
Reformulation of DFT + U as a Pseudohybrid Hubbard Density Functional for Accelerated Materials Discovery

Luis A. Agapito,[1,2] Stefano Curtarolo,[2,3] and Marco Buongiorno Nardelli[4,2,*]
[1]*Department of Physics, University of North Texas, Denton, Texas 76203, USA*
[2]*Center for Materials Genomics, Duke University, Durham, North Carolina 27708, USA*
[3]*Materials Science, Electrical Engineering, Physics and Chemistry, Duke University, Durham, North Carolina 27708, USA*
[4]*Department of Physics and Department of Chemistry, University of North Texas, Denton, Texas 76203, USA*
(Received 10 June 2014; published 28 January 2015)

High-Throughput data generation

Improve:
- ❑ Structure
- ❑ Energetics
- ❑ Energy bands
- ❑ Phonons
- ❑ Magnetism

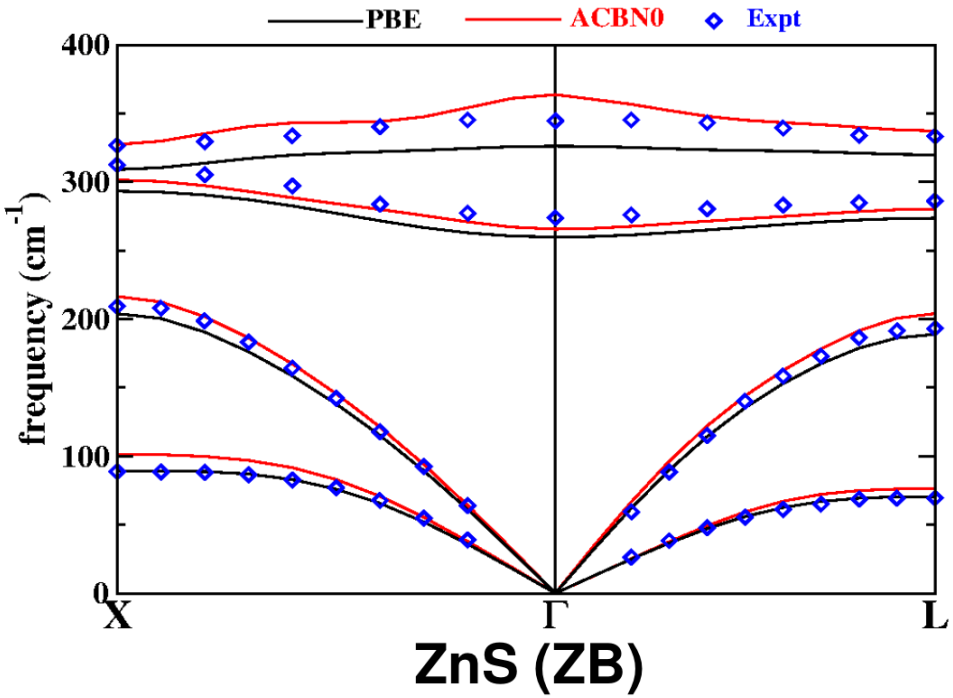
Crucial for:
- ❑ Mixed valence
- ❑ Optical



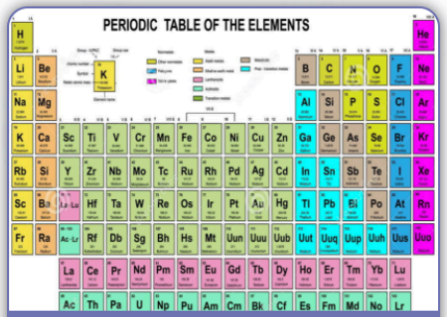
High-Throughput data generation

Improve:
- ❑ Structure
- ❑ Energetics
- ❑ Energy bands
- ❑ Phonons
- ❑ Magnetism

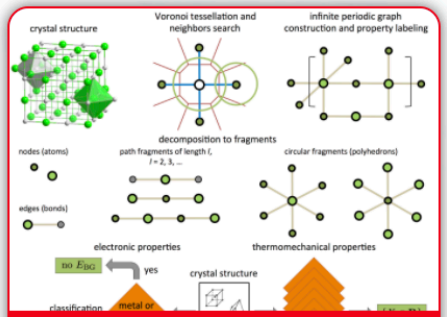
Crucial for:
- ❑ Mixed valence
- ❑ Optical



Apps and Docs

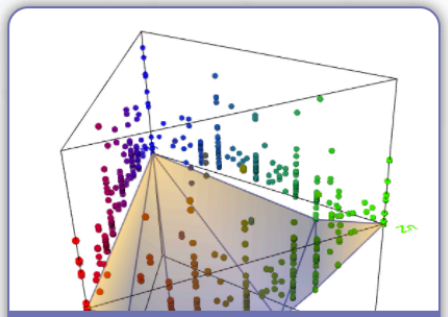


MendeLIB search

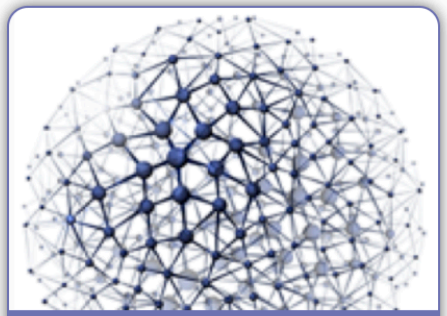


AFLOW-ML

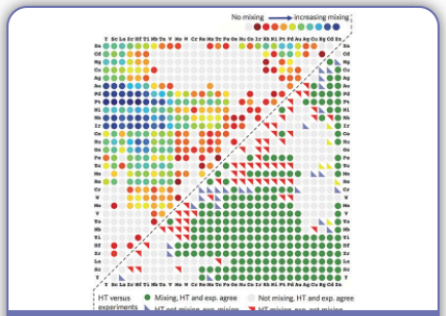
Crystal prototypes



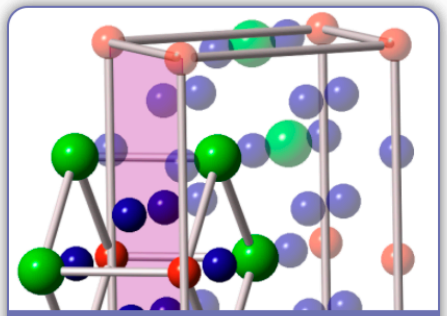
AFLOW hull beta



Aflow-online



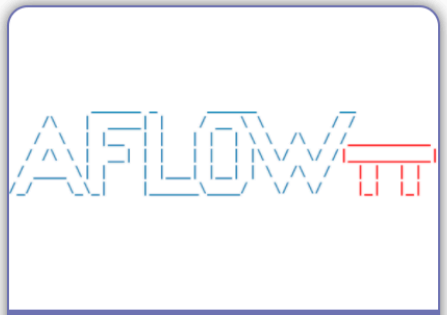
Binary alloy library



Superalloys search

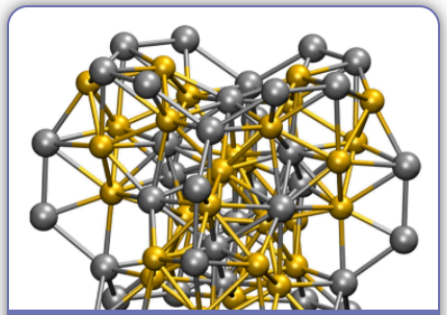
AFLOW
REST-API WIKI

Documentation



AFLOWπ

Geolocation data



Showcase material

PAOFLOW

PAOFLOW

PAOFLOW

Advanced characterization tools

PHYSICAL REVIEW B **93**, 035104 (2016)

Accurate tight-binding Hamiltonian matrices from *ab initio* calculations: Minimal basis sets

Luis A. Agapito,[1,2] Sohrab Ismail-Beigi,[3] Stefano Curtarolo,[4,5] Marco Fornari,[4,6] and Marco Buongiorno Nardelli[2,4,*]
[1]*Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708, USA*
[2]*Department of Physics, University of North Texas, Denton, Texas 76203, USA*
[3]*Department of Applied Physics and Center for Research on Interface Structures and Phenomena (CRISP), Yale University, New Haven, Connecticut 06511, USA*
[4]*Center for Materials Genomics, Duke University, Durham, North Carolina 27708, USA*
[5]*Materials Science, Electrical Engineering, Physics and Chemistry, Duke University, Durham, North Carolina 27708, USA*
[6]*Department of Physics, Central Michigan University, Mt. Pleasant, Michigan 48859, USA*

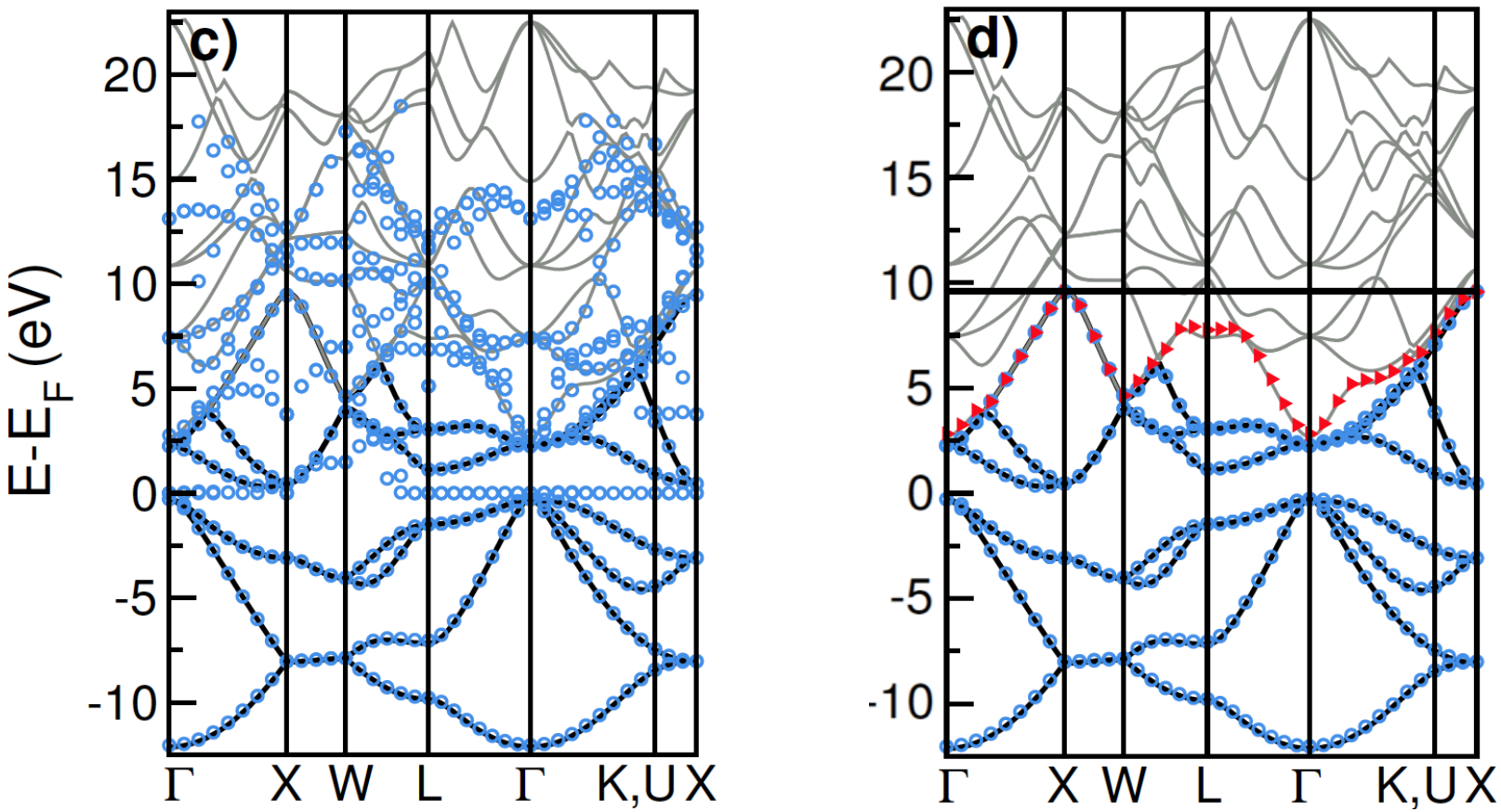
GOOD projectability

BAD projectability

PAOFLOW

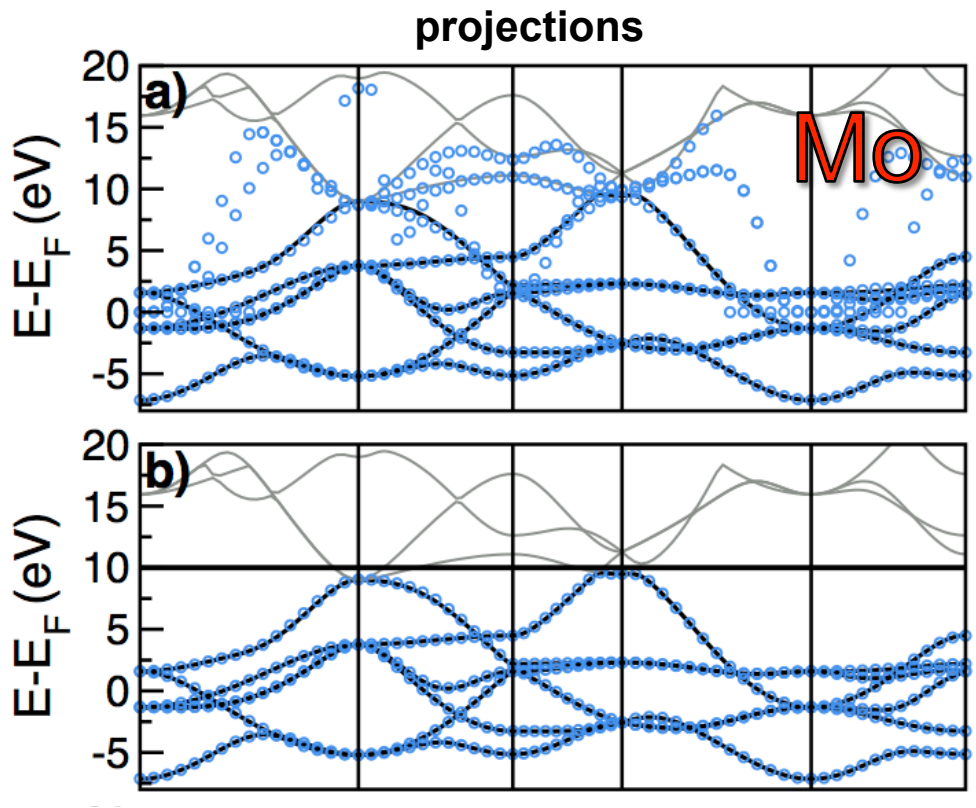
Advanced characterization tools

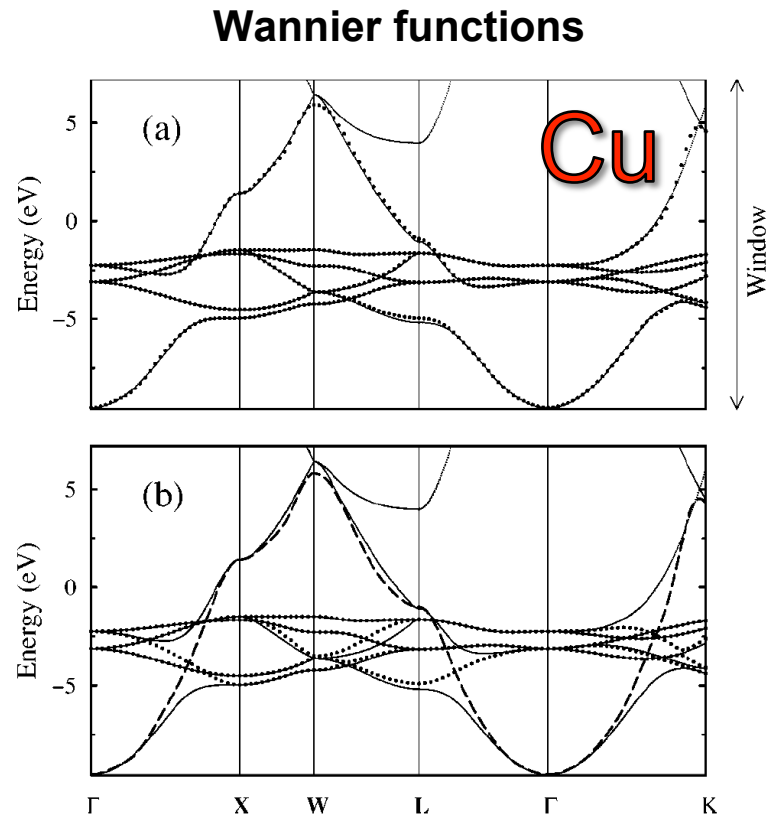
$$\bar{H} = AEA^{\dagger} \longrightarrow \bar{H}_{\kappa} = \bar{H} + \kappa Q_{\mathcal{N}}$$



PAOFLOW

Advanced characterization tools





Effective description of entangled bands in metals

PAOFLOW Advanced characterization tools

PHYSICAL REVIEW B **93**, 125137 (2016)

Accurate tight-binding Hamiltonians for two-dimensional and layered materials

Luis A. Agapito,[1,2] Marco Fornari,[3,4] Davide Ceresoli,[5] Andrea Ferretti,[6] Stefano Curtarolo,[4,7] and Marco Buongiorno Nardelli[2,4,*]

[1]*Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708, USA*
[2]*Department of Physics, University of North Texas, Denton, Texas 76203, USA*
[3]*Department of Physics, Central Michigan University, Mount Pleasant, Michigan 48859, USA*
[4]*Center for Materials Genomics, Duke University, Durham, North Carolina 27708, USA*
[5]*Consiglio Nazionale delle Ricerche, ISTM, Istituto di Scienze e Tecnologie Molecolari, I-20133 Milano, Italy*
[6]*Consiglio Nazionale delle Ricerche, NANO S3 Center, Istituto Nanoscienze, I-41125 Modena, Italy*
[7]*Materials Science, Electrical Engineering, Physics and Chemistry, Duke University, Durham, North Carolina 27708, USA*

(Received 11 January 2016; revised manuscript received 4 March 2016; published 22 March 2016)

projectability

BAD projectability

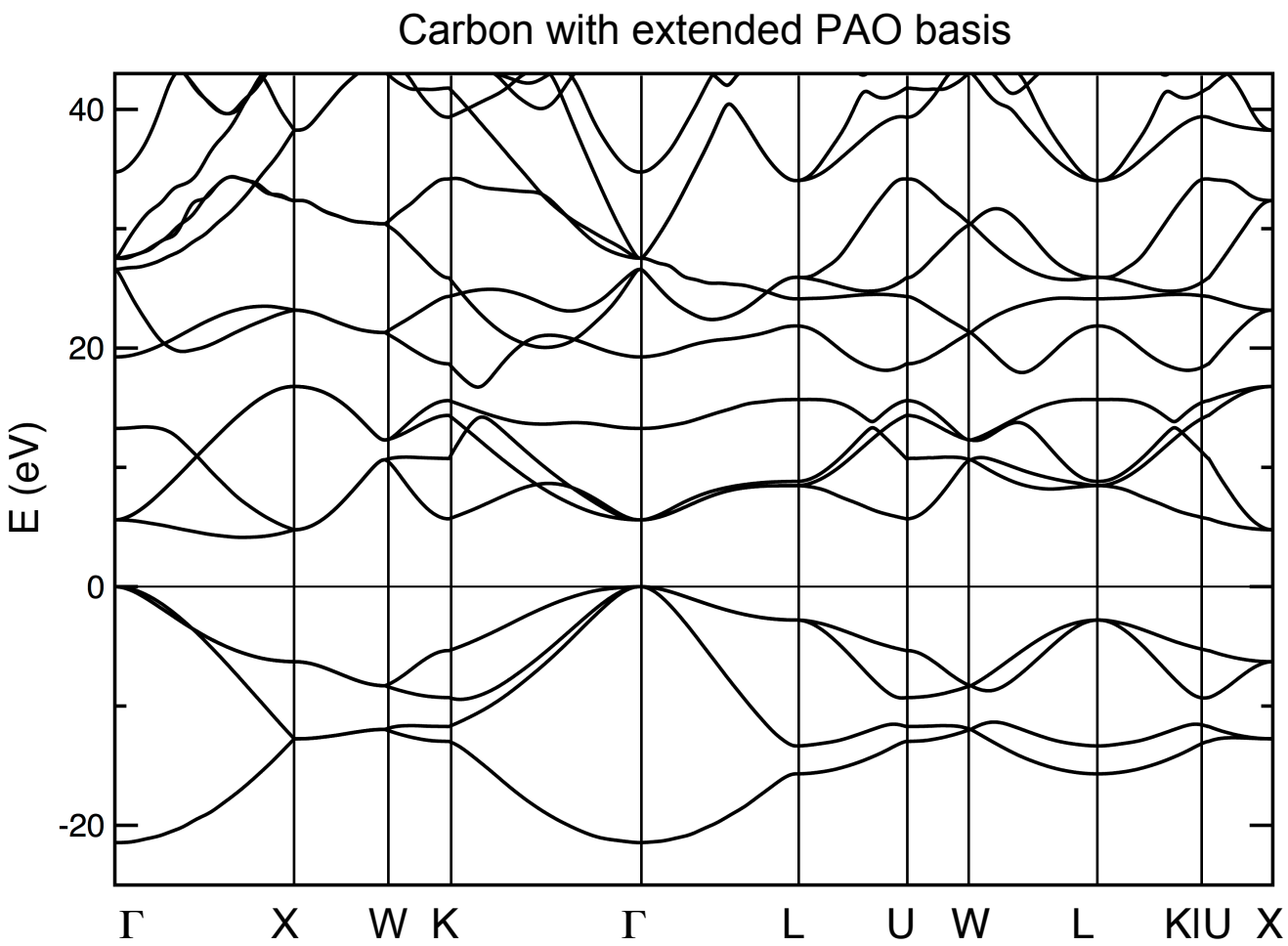
projectability

BAD projectability

From SZ to DZ to DZP to TZP…etc.

PAOFLOW

Advanced characterization tools



PAOFLOW

Advanced characterization tools

PAOFLOW: A utility to construct and operate on ab initio Hamiltonians from the Projections of electronic wavefunctions on Atomic Orbital bases (PAO), including characterization of topological materials

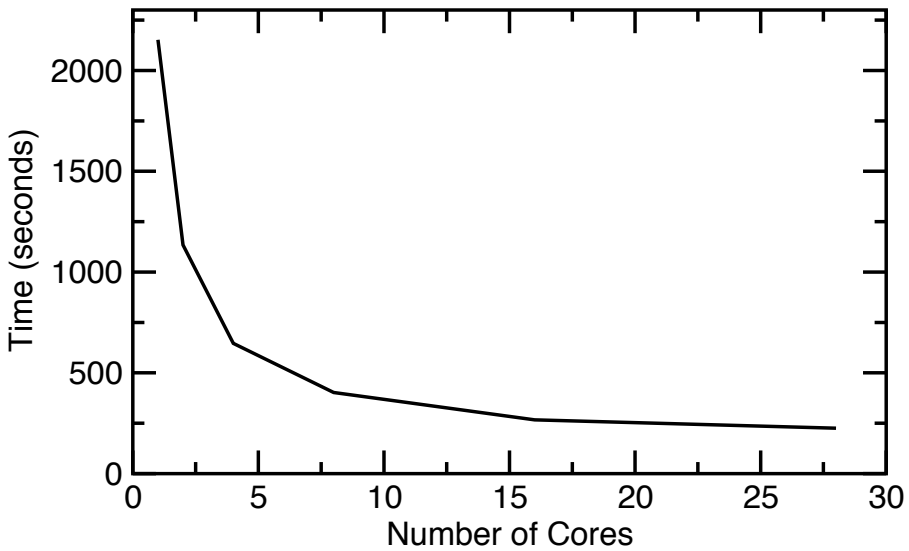
Marco Buongiorno Nardelli[1,5⋆], Frank T. Cerasoli[1], Marcio Costa[6], Stefano Curtarolo[2,5], Riccardo De Gennaro,[7], Marco Fornari[3,4,5], Laalitha Liyanage[1], Andrew R. Supka[3,4], Haihang Wang[1],

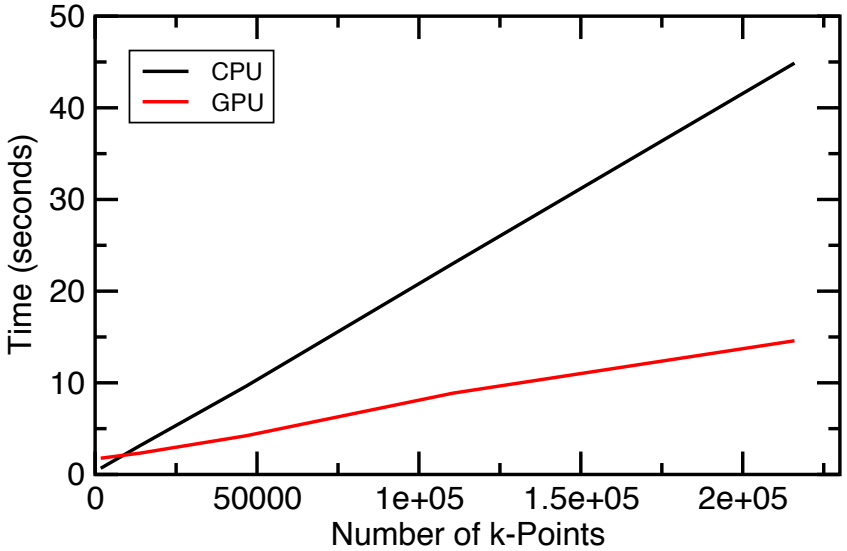
- Band structure and band topology
- DOS, Fermi surfaces, effective masses + warping
- spin-orbit correction on non-SO calculations
- Spin texture
- Boltzmann transport
- Dielectric functions and optical spectra
- Topological invariants ($\mathbb{Z}2$)
- Berry curvature and anomalous Hall conductivity
- Spin Hall conductivity
- Electron-phonon coupling
- …more to come…!

Submitted to Computational Materials Science

PAOFLOW software engineering (I)

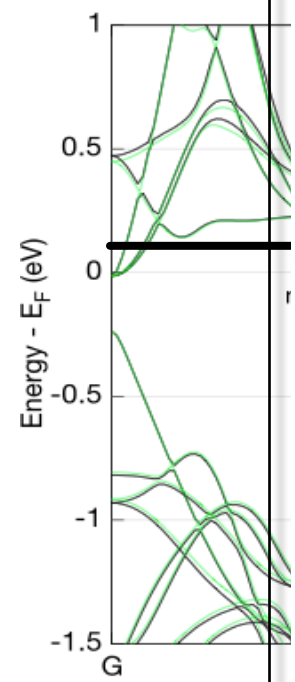
- 7200+ lines of python
- Highly parallelized with MPI and multithreading
- Added support for GPU hybrid computing (FFT)
- Full suite of examples and continuous integration development model
- Completely integrated in AFLOWπ

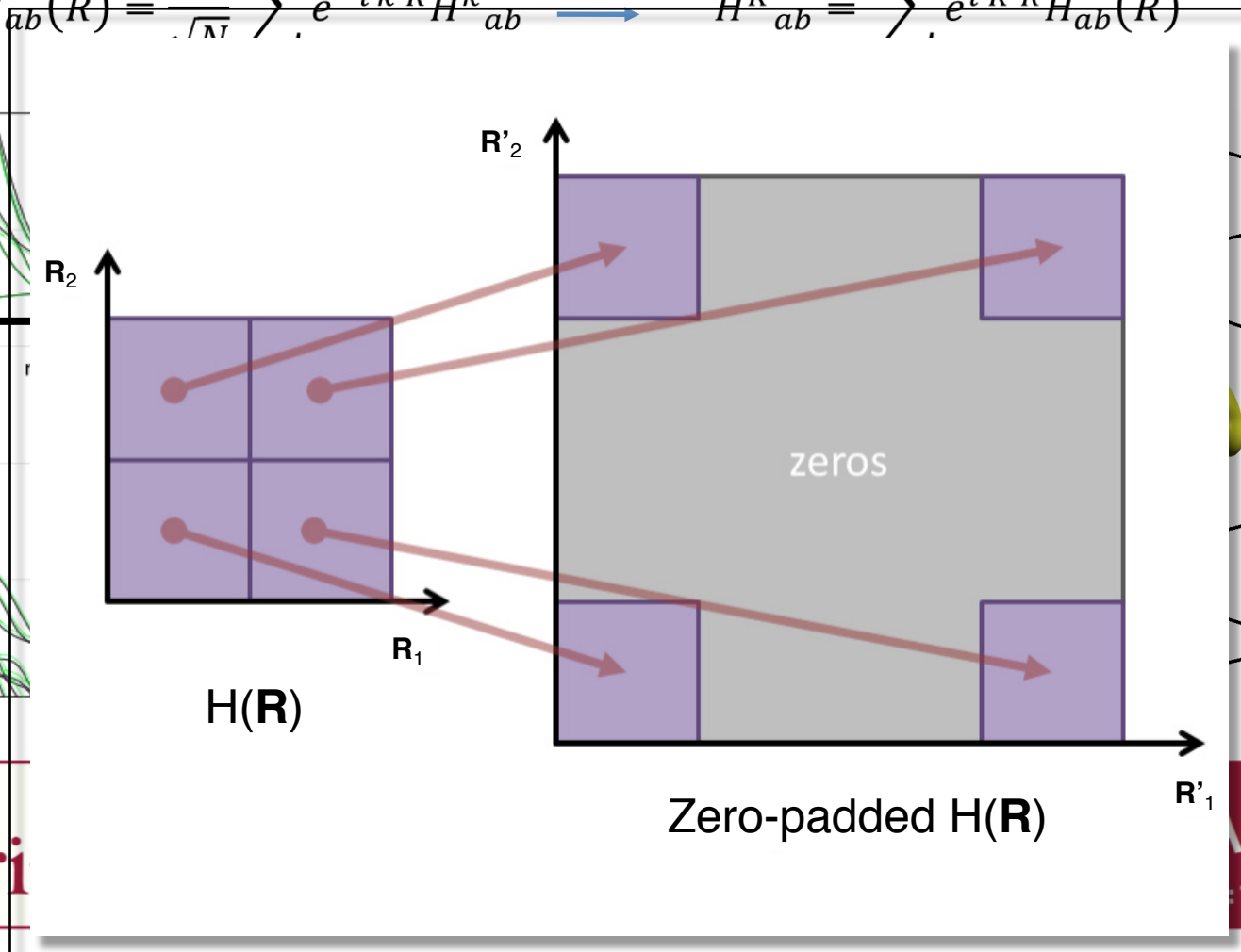


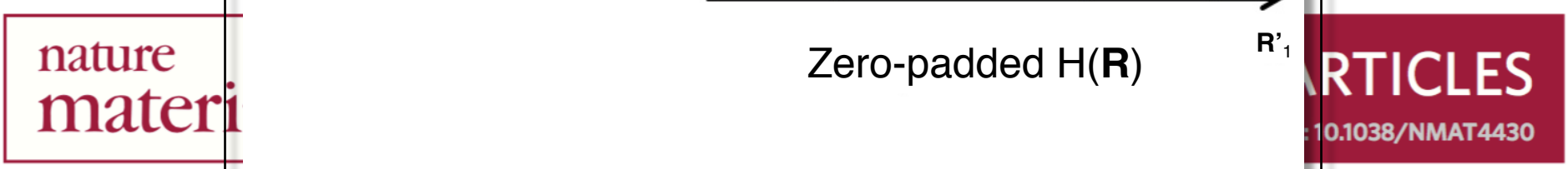


Ultra-fine mesh integration and encoding

$$H_{ab}(\vec{R}) = \frac{1}{\sqrt{N}} \sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{R}} H^{\vec{k}}_{ab} \longrightarrow H^{\vec{K}}_{ab} = \sum_{\vec{R}} e^{i\vec{K}\cdot\vec{R}} H_{ab}(\vec{R})$$







Convergence of multi-valley bands as the electronic origin of high thermoelectric performance in $CoSb_3$ skutterudites, *Y. Tang, Z.M. Gibbs, L.A. Agapito, G. Li, H. Kim, M. Buongiorno Nardelli, S. Curtarolo and G.J. Snyder*

PHYSICAL REVIEW B **94**, 165166 (2016)

Accurate *ab initio* tight-binding Hamiltonians: Effective tools for electronic transport and optical spectroscopy from first principles

Pino D'Amico,[1,2] Luis Agapito,[3,4] Alessandra Catellani,[2] Alice Ruini,[1,2] Stefano Curtarolo,[3,5] Marco Fornari,[3,6] Marco Buongiorno Nardelli,[3,7,*] and Arrigo Calzolari[2,3,4,†]

[1]*Dipartimento di Fisica, Informatica e Matematica, Universitá di Modena and Reggio Emilia, Via Campi 213/a, 41125 Modena, Italy*

[2]*CNR-NANO Research Center S3, Via Campi 213/a, 41125 Modena, Italy*

[3]*Center for Materials Genomics, Duke University, Durham, North Carolina 27708, USA*

[4]*Department of Physics, University of North Texas, Denton, Texas 76203, USA*

[5]*Materials Science, Electrical Engineering, Physics and Chemistry, Duke University, Durham, North Carolina 27708, USA*

[6]*Department of Physics, Central Michigan University and Science of Advanced Materials Program, Mt. Pleasant, Michigan 48859, USA*

[7]*Department of Physics and Department of Chemistry, University of North Texas, Denton, Texas 76203, USA*

(Received 19 August 2016; published 26 October 2016)

Accurate evaluation of **<u>Boltzmann transport integrals</u>** and **<u>optical properties</u>** depends on the evaluation of the the expectation value of the momentum operator:

PHYSICAL REVIEW B VOLUME 51, NUMBER 8 15 FEBRUARY 1995-II

Electromagnetic fields and dielectric response in empirical tight-binding theory

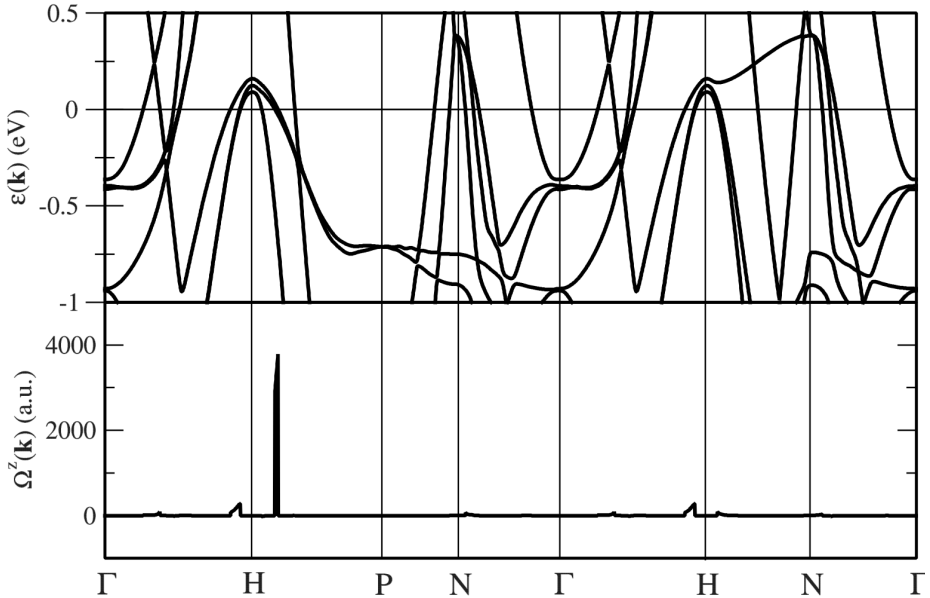
M. Graf and P. Vogl

Walter Schottky Institut and Physik Department, Technische Universität München, 85748 Garching, Germany

α

space

Berry curvature and Anomalous Hall conductivity: Fe



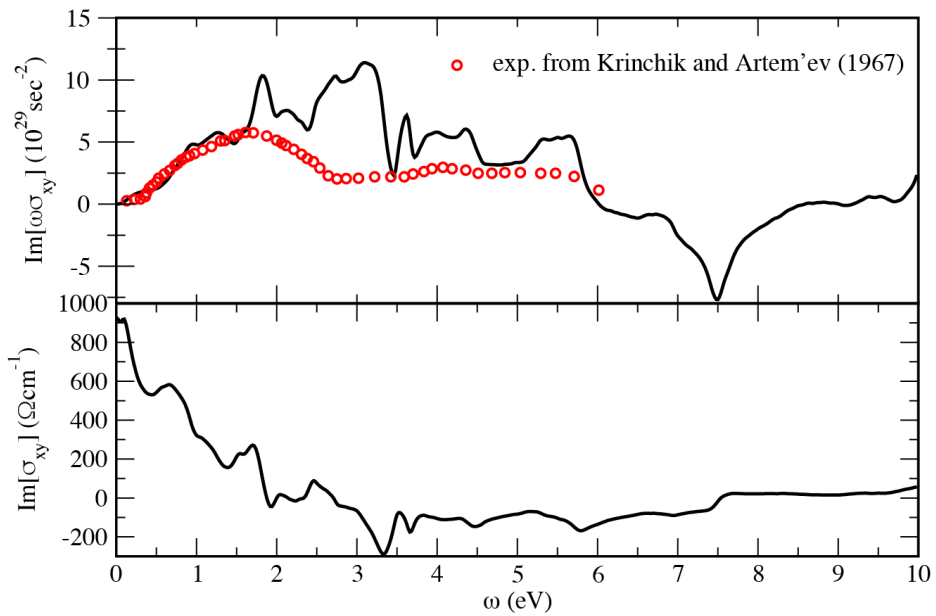
$$\Omega^z(\mathbf{k}) = \sum_n f_n \Omega_n^z(\mathbf{k})$$

Anomalous Hall conductivity:

$$\sigma_{xy} = -\frac{e^2}{\hbar}\int_{\mathrm{BZ}} \frac{d^3k}{(2\pi)^3}\Omega^z(\mathbf{k})$$

Berry curvature from Kubo formula:

$$\Omega_n^z(\mathbf{k}) = -\sum_{n'\neq n} \frac{2\mathrm{Im}\langle\psi_{n\mathbf{k}}|v_x|\psi_{n'\mathbf{k}}\rangle\langle\psi_{n'\mathbf{k}}|v_y|\psi_{n\mathbf{k}}\rangle}{(\omega_{n'}-\omega_n)^2}$$



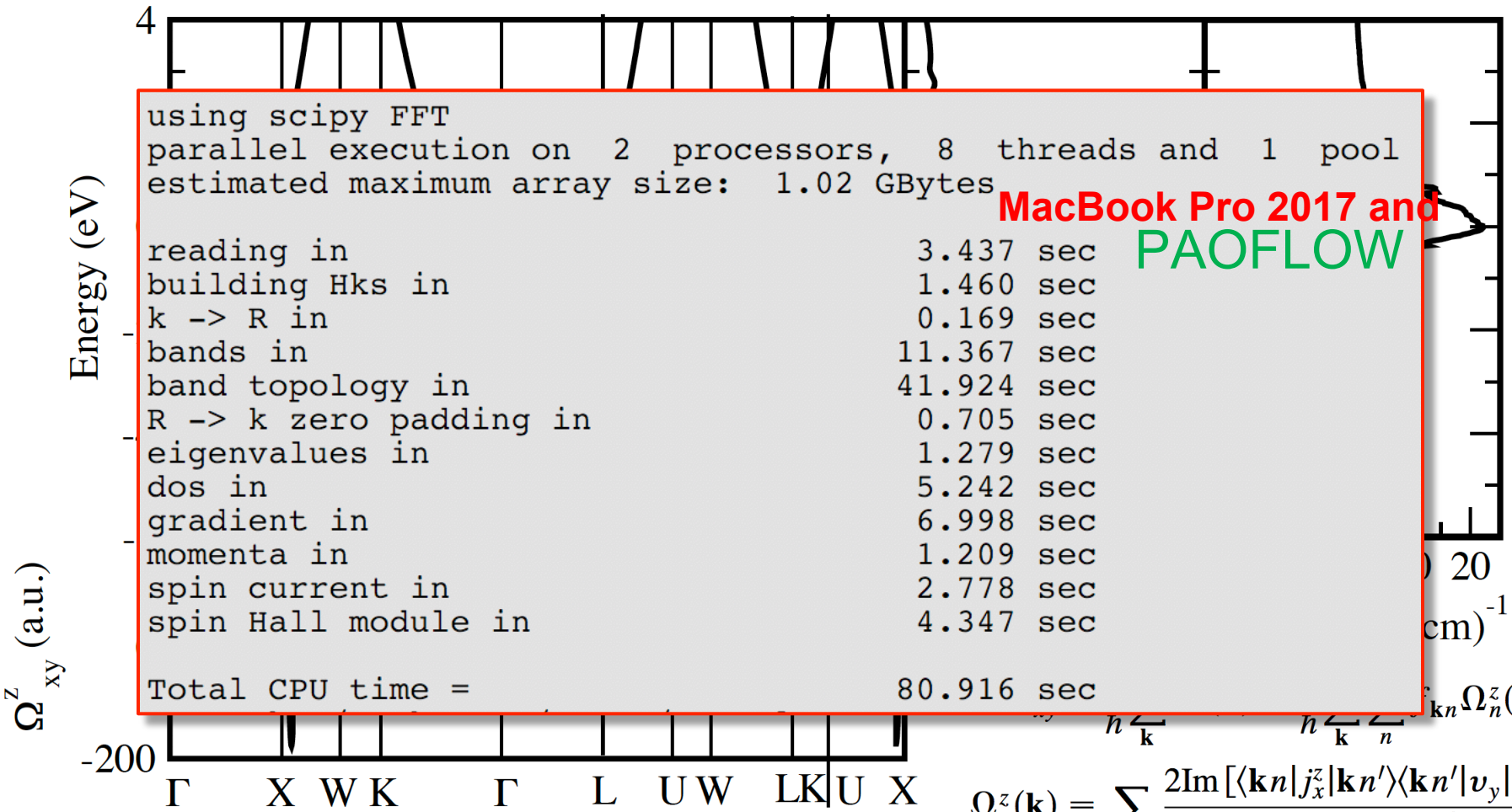
$$\sigma(\omega)_{xy} = \frac{e^2}{\hbar}\int_{V_G}\frac{d^3k}{(2\pi)^3}\sum_{n\neq n'}(f_{n,\mathbf{k}}-f_{n',\mathbf{k}}) \times \frac{\mathrm{Im}\langle\psi_{n\mathbf{k}}|v_x|\psi_{n'\mathbf{k}}\rangle\langle\psi_{n'\mathbf{k}}|v_y|\psi_{n\mathbf{k}}\rangle}{(\omega_{n'}-\omega_n)^2-(\omega+i\delta)^2}$$

VOLUME 92, NUMBER 3 PHYSICAL REVIEW LETTERS week ending 23 JANUARY 2004

First Principles Calculation of Anomalous Hall Conductivity in Ferromagnetic bcc Fe

Yugui Yao,[1,2,3] Leonard Kleinman,[1] A. H. MacDonald,[1] Jairo Sinova,[4,1] T. Jungwirth,[5,1] Ding-sheng Wang,[3] Enge Wang,[2,3] and Qian Niu[1]

Spin Hall conductivity: Pt




```
using scipy FFT
parallel execution on  2  processors,  8  threads and  1  pool
estimated maximum array size:  1.02 GBytes

reading in                          3.437 sec
building Hks in                     1.460 sec
k -> R in                           0.169 sec
bands in                           11.367 sec
band topology in                   41.924 sec
R -> k zero padding in              0.705 sec
eigenvalues in                      1.279 sec
dos in                              5.242 sec
gradient in                         6.998 sec
momenta in                          1.209 sec
spin current in                     2.778 sec
spin Hall module in                 4.347 sec

Total CPU time =                   80.916 sec
```

MacBook Pro 2017 and PAOFLOW

$$\Omega^z_n(\mathbf{k}) = \sum_{n' \neq n} \frac{2\mathrm{Im}\left[\langle \mathbf{k}n | j^z_x | \mathbf{k}n' \rangle \langle \mathbf{k}n' | v_y | \mathbf{k}n \rangle\right]}{(\epsilon_{\mathbf{k}n} - \epsilon_{\mathbf{k}n'})^2}$$

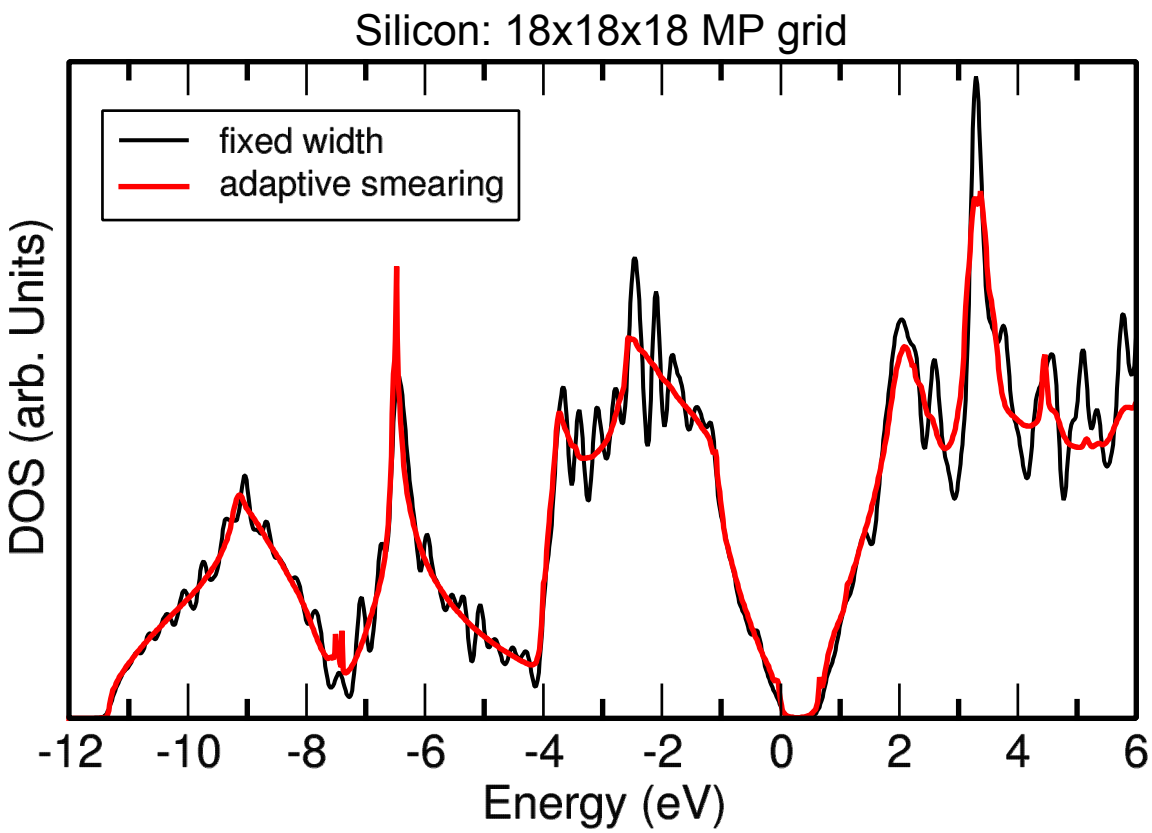
PRL **100,** 096401 (2008) PHYSICAL REVIEW LETTERS week ending 7 MARCH 2008

Intrinsic Spin Hall Effect in Platinum: First-Principles Calculations

G. Y. Guo,[1,*] S. Murakami,[2] T.-W. Chen,[1] and N. Nagaosa[3,4]

PAOFLOW software engineering (II)

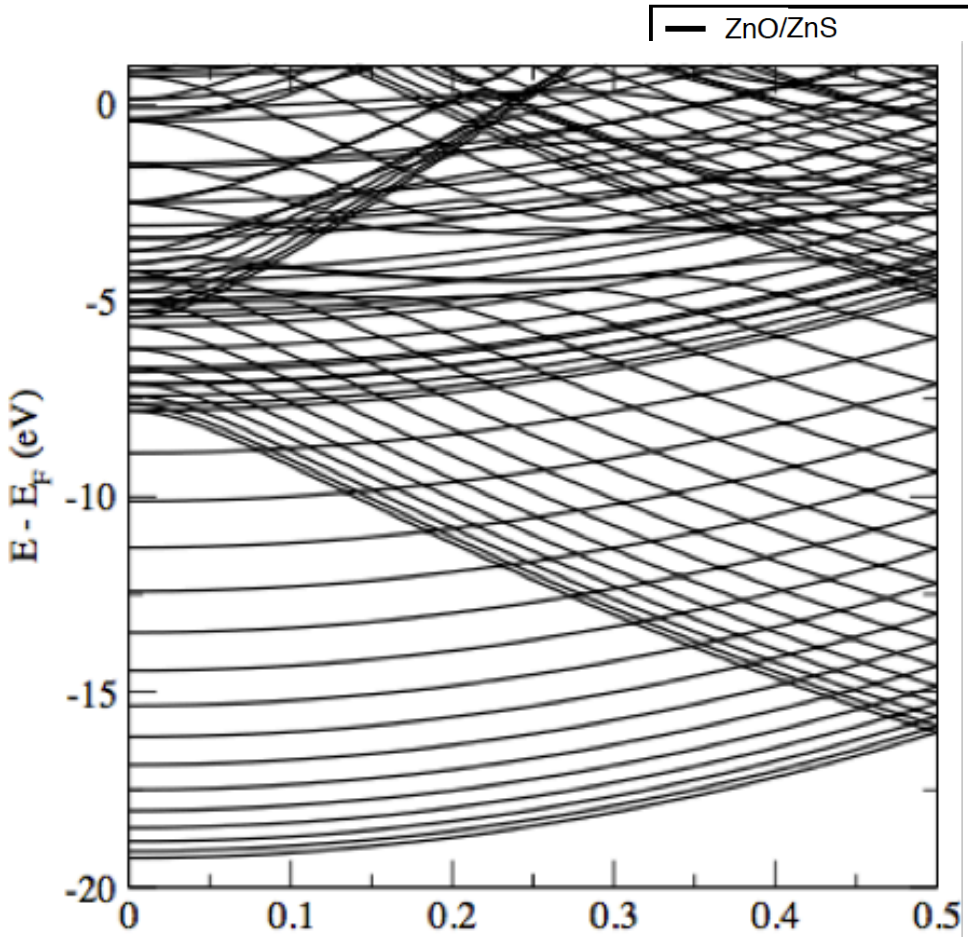
- Adaptive smearing for BZ integration: broadening widths using the band derivatives that are readily available from the knowledge of the momentum operator



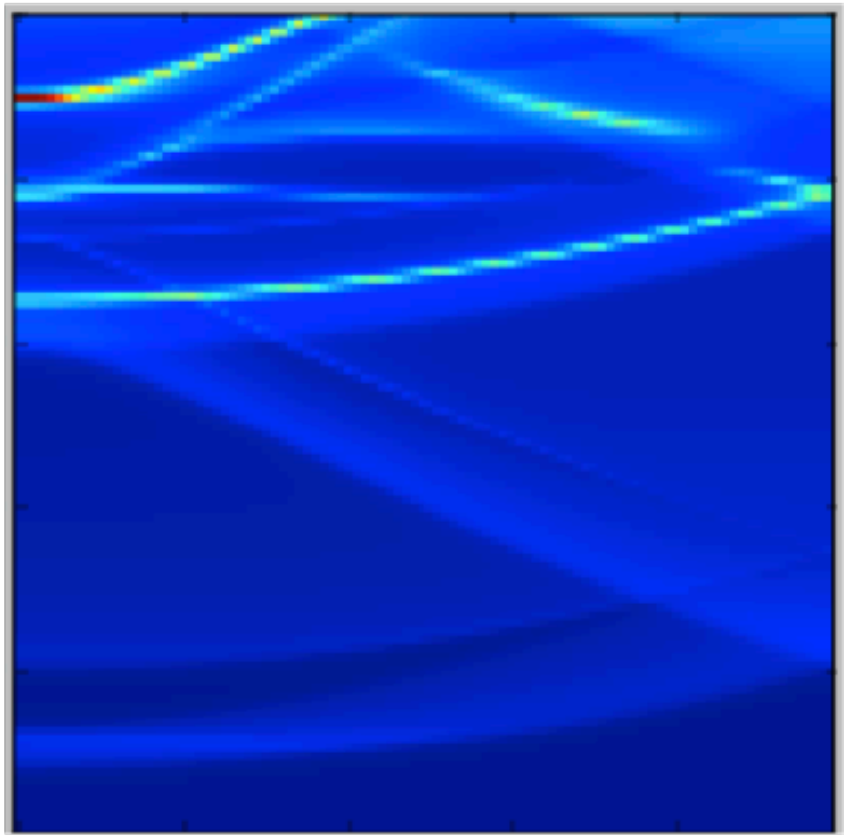
$$W_{n,\mathbf{k}} = a \left| \frac{\partial \epsilon_{n,\mathbf{k}}}{\partial k} \right|$$

Yates et al. PRB (2007)

Externals: Landauer transport and surface bands



32 atoms slab

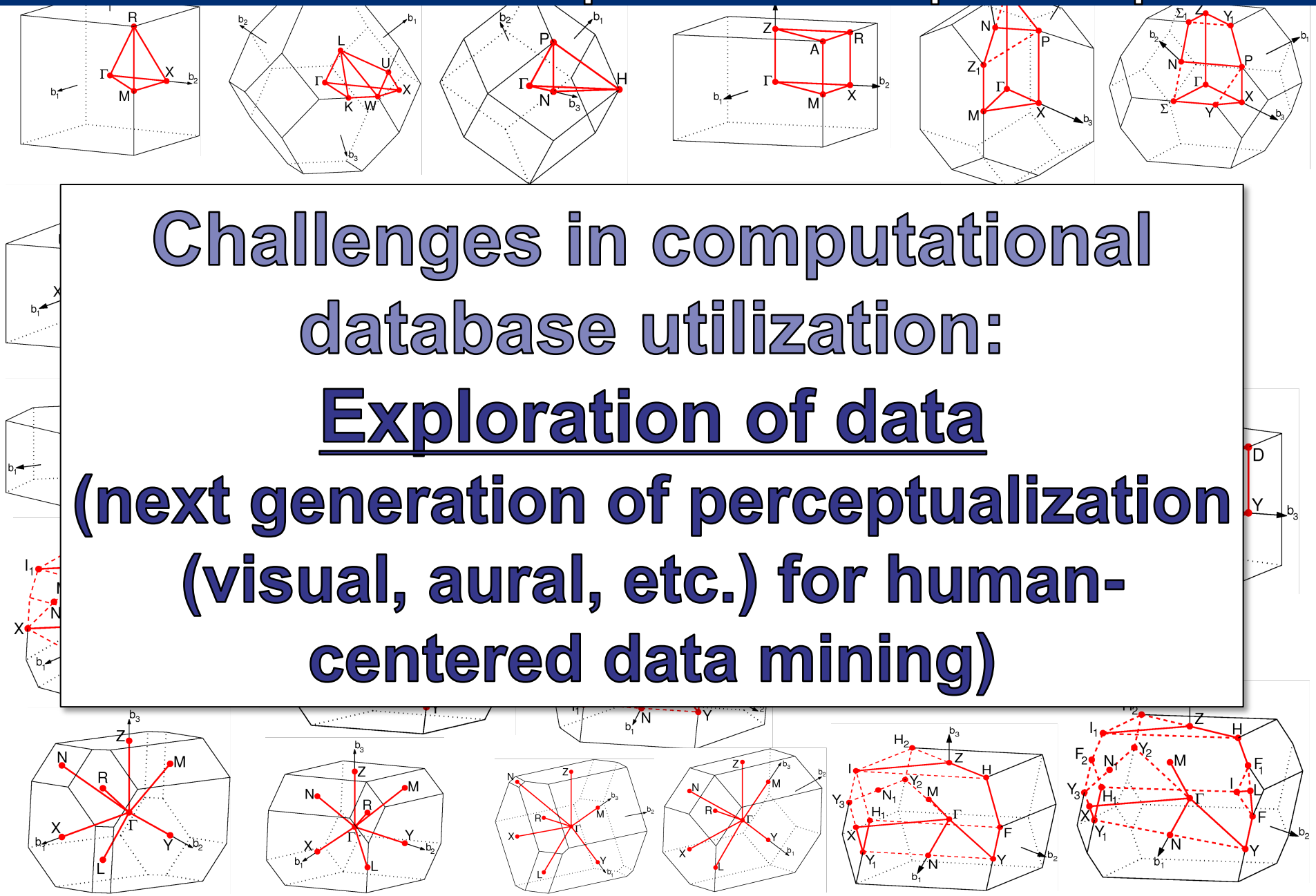


Semi-infinite slab with PAOs

-4 -2 0 2 4

energy (eV)

STANDARD in Real Space and Reciprocal Space

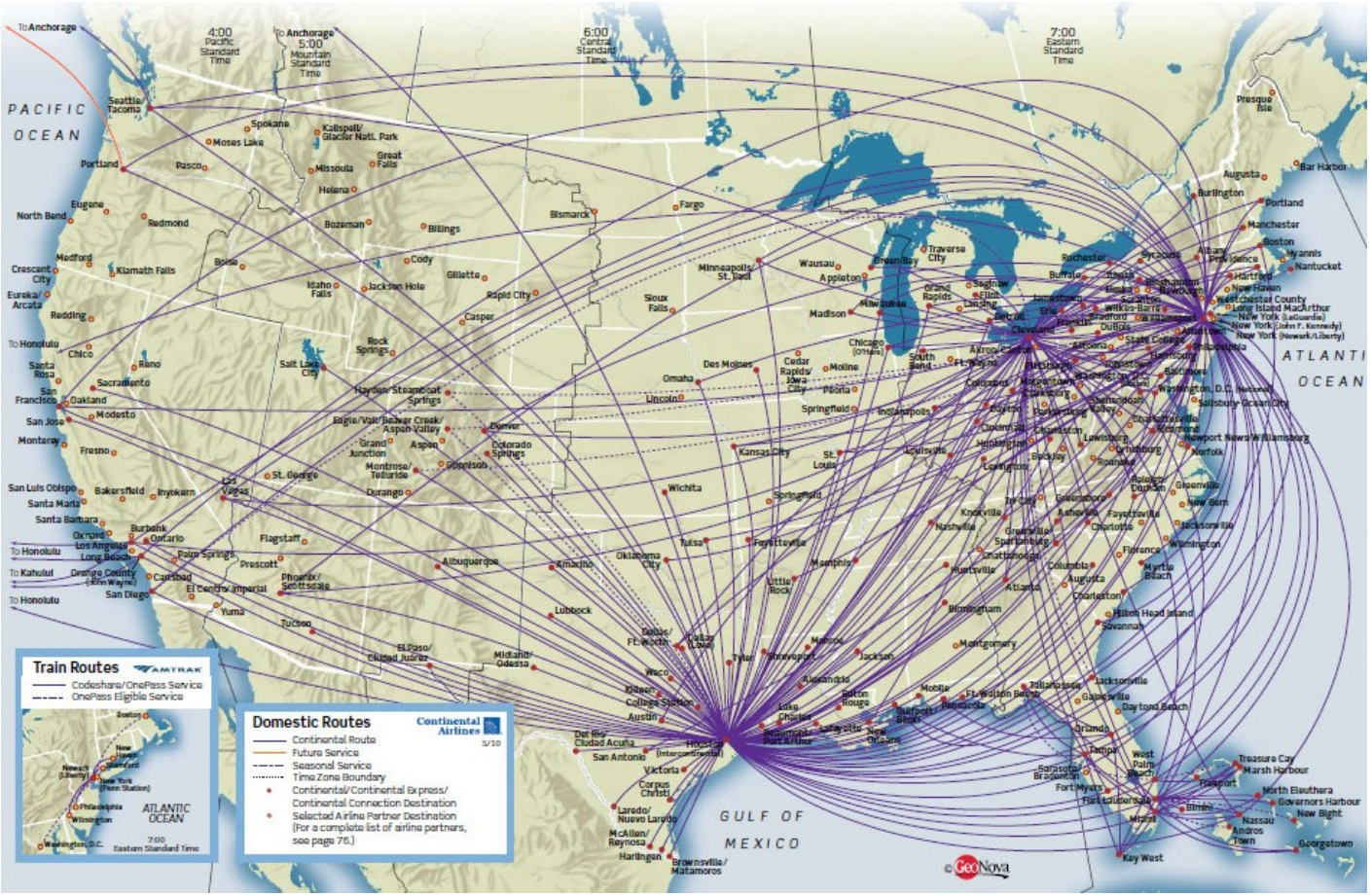


Challenges in computational database utilization:

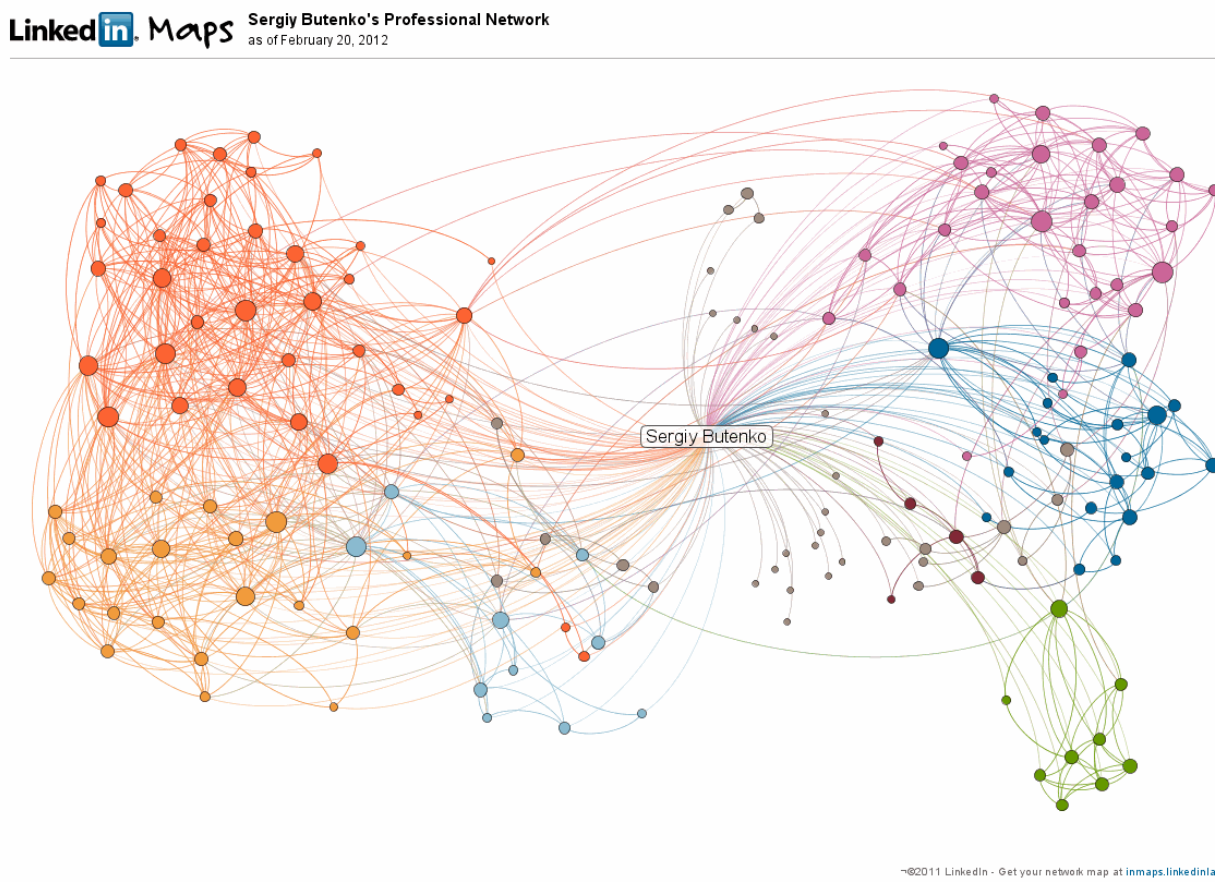
<u>Exploration of data</u>

(next generation of perceptualization (visual, aural, etc.) for human-centered data mining)

Network-based analysis of materials data



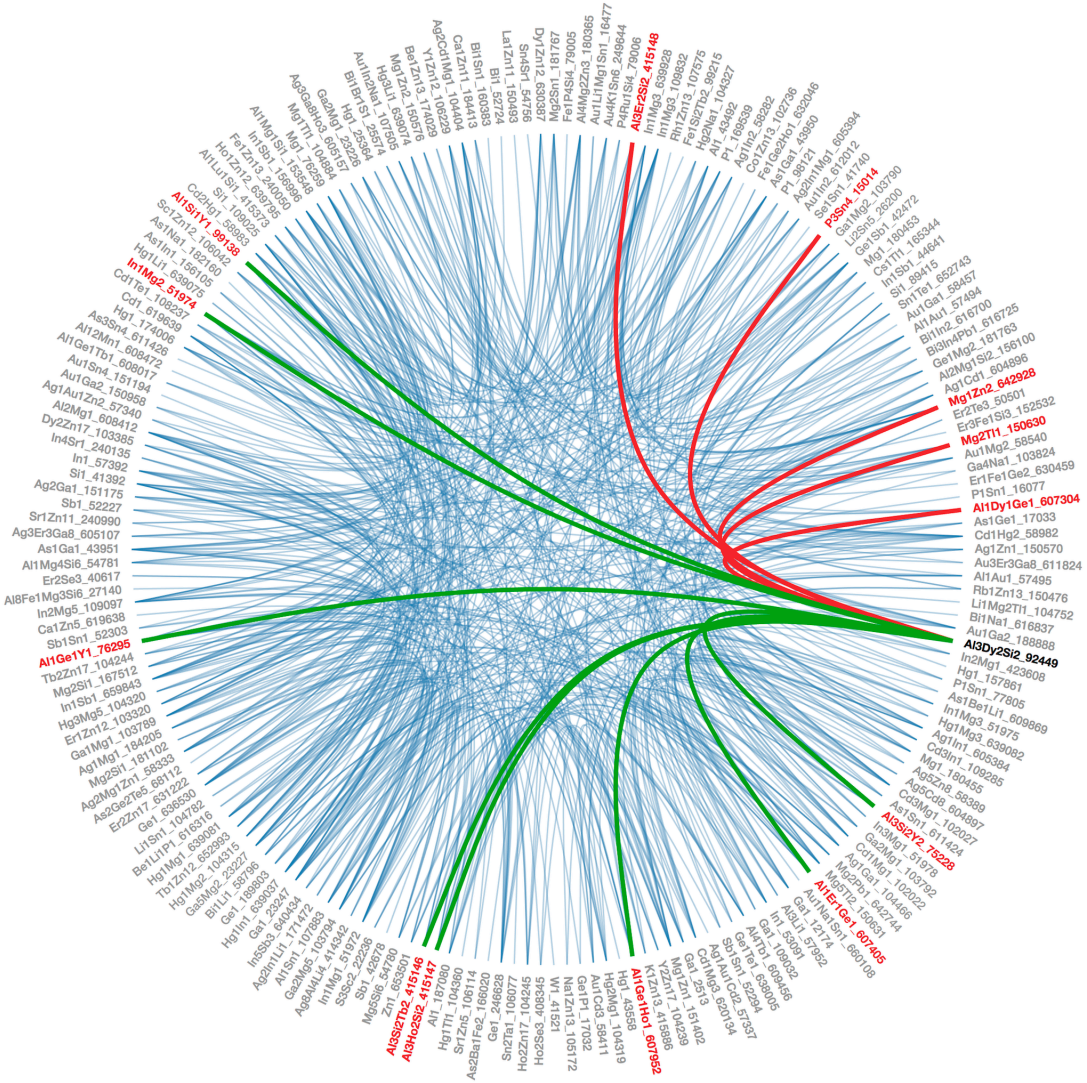
Network-based analysis of materials data



Network-based analysis of materials data

- Identifying subsets of materials with properties similar to the ones of interest (i.e., narrowing down the "list of candidates" for further investigation)
- Identifying materials with high degree of similarity to many other materials
- Cluster materials space into groups of materials that share similar properties
- Provide an intuitive "navigation" tool in the materials space (given a certain material one can visually see its "neighborhood" for substitution or compatibility problems)

Network-based analysis of materials data



QUANTUMESPRESSO

AFLOW
Automatic-FLOW for Materials Discovery

http://ermes.unt.edu

http://AFLOW.org

http://www.quantum-espresso.org

www.materialssoundmusic.com

M. Fornari, S. Curtarolo, P. Gopal, L. Liyanage, L.A. Agapito, H. Wang, C. Toher, M. Costa, M. dos Santos Guzmao, A. Supka, P. D'Amico, A. Calzolari

DEPARTMENT OF THE NAVY
USN
ONR
Science & Technology