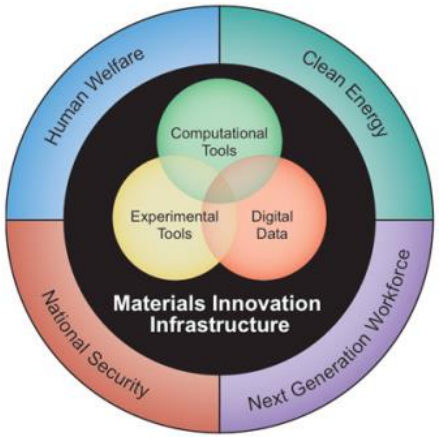
NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Accelerated Materials Discovery & Characterization with Quantum and Machine learning approaches

Kamal Choudhary, NIST
PARADIM, Johns Hopkins University
November 14, 2018

JARVIS

Web: https://jarvis.nist.gov



Acknowledgement and Collaboration

F. Tavazza
(NIST)

C. Campbell
(NIST)

J. Warren
(NIST)

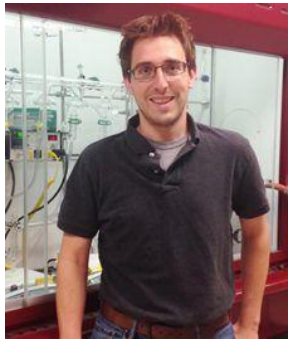
A. Reid
(NIST)

D. Wheeler
(NIST)

K. Garrity
(NIST)

B. DeCost
(NIST)

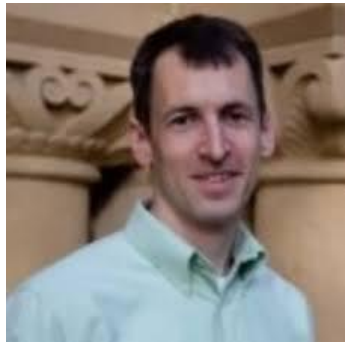
C. Becker
(NIST)



A. Biacchi
(NIST)

F. Congo
(NIST)

T. Liang
(Penn state)



E. Reed
(Stanford)

A. Agarwal
(Northwestern)

M. Bercx
(Antwerp)

Outline

- Motivation
- Basics: DFT
- Discovery & characterization:
  - low-dimensional materials (2D, 1D, 0D)
  - Topological materials
  - Efficient solar-cell materials
  - Thermoelectric materials
  - Flexible and negative Poisson ratio materials
- ML training
- Website/github demo

Navigate to:
https://jarvis.nist.gov

Motivation

Materials Genome Initiative

Combine different approaches (computation, experiments etc.)
Making materials data publicly available

https://obamawhitehouse.archives.gov/mgi

Density Functional Theory

Second derivative wrt X

Schrodinger Wave-function

$$\frac{\partial^2\psi}{\partial x^2}+\frac{8\pi^2 m}{h^2}(E-V)\psi=0$$

Position Energy Potential Energy

- Classical Newton's laws not applicable for electrons (**very fast, very tiny**);
- Schrödinger equation: mathematical equation that describes the evolution over time of a physical system in which quantum effects, such as **wave–particle duality**, are significant (such as electrons)
- Schrödinger equation of a **fictitious system** (the "Kohn–Sham system") of **non-interacting particles** (typically electrons) that generate the **same density** as any given system of interacting particles
- Uses **density** vs **wavefunction** quantity

Exchange-correlation

$$H\psi=E\psi \qquad \left[-\frac{\hbar^2}{2m}\nabla^2+V_{Eff}(r)\right]\psi_i(r)=E_i(r)\psi_i(r) \qquad V_{Eff}=T+V_{Ne}+V_{ee}+V_{XC}$$

Walter Kohn

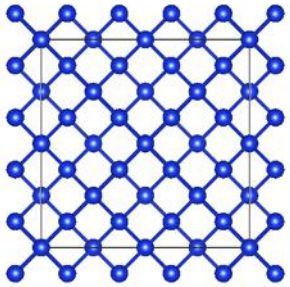
- **Different functionals: LDA, GGA (PBE, PW91), HSE06 etc.**
- **LDA**: developed for homogenous systems (Thomas-Fermi model), such as metals
- **GGA**: uses density gradient information on top of GGA
- **vdW-DF**: uses exchange from GGA, correlation from LDA, quantum-Monte-Carlo based non-local corrections, **optB88vdW** used here mainly

https://en.wikipedia.org/wiki/Classical_mechanics
http://www.psi-k.org/Update3.shtml
http://www.physlink.com/Education/AskExperts/ae329.cfm

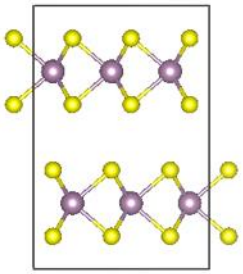
Hohenberg, Pierre; Walter Kohn (**1964**). "Inhomogeneous electron gas". *Physical Review*. **136** (3B): B864–B871

Discovery of 2D (low-D) materials

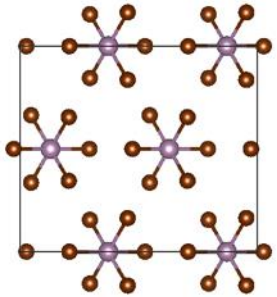
(a) 3D-bulk (Si)



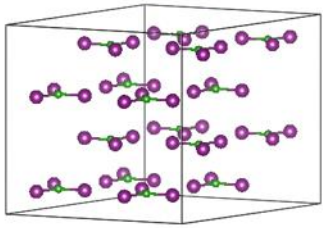
(b) 2D-bulk (MoS_2)



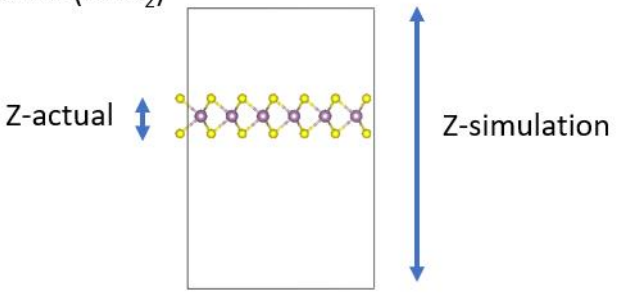
(c) 1D-bulk ($MoBr_3$)



(d) 0D-bulk (BiI_3)



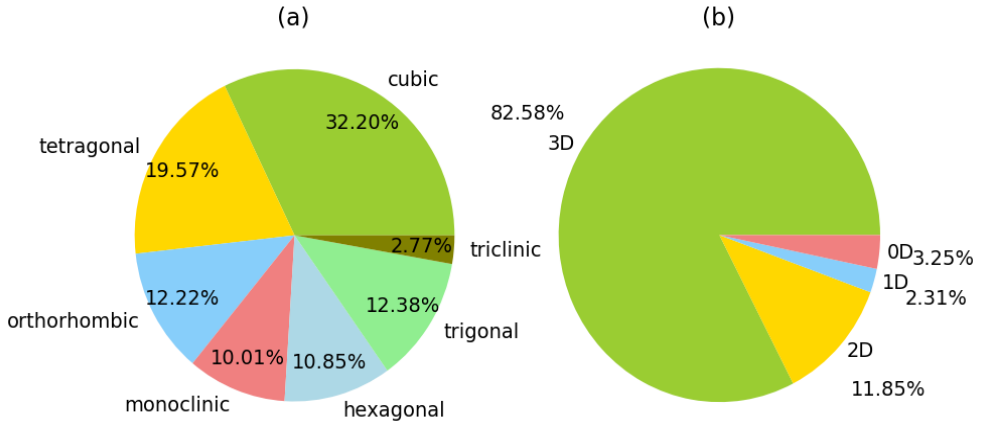
(e) 2D-1L (MoS_2)



$$\delta = \frac{\left| c_{PBE} - c_{ICSD} \right|}{c_{ICSD}}$$

Using lattice-constant error to discover low-D materials

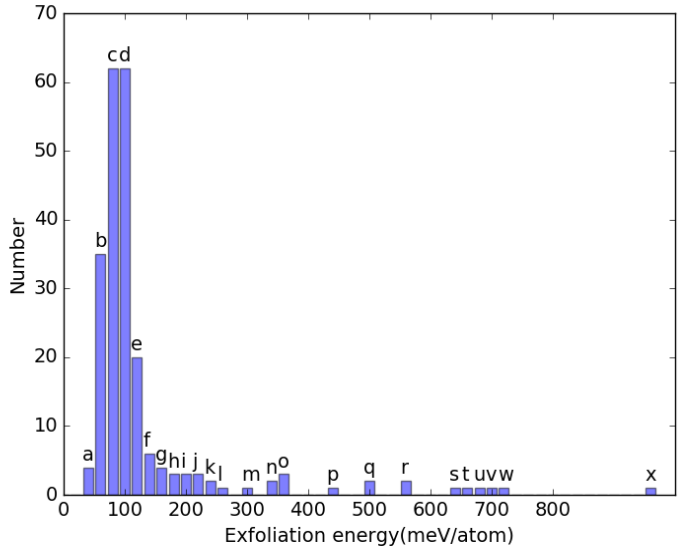
If the error is more than **5% > 1500 materials**



Scientific Reports 7, 5179 (2017)
Physical Review B 98 (1), 014107 (2018)

Exfoliation energy calculations

$$E_f = \frac{E_{1L}}{N_{1L}} - \frac{E_{bulk}}{N_{bulk}}$$

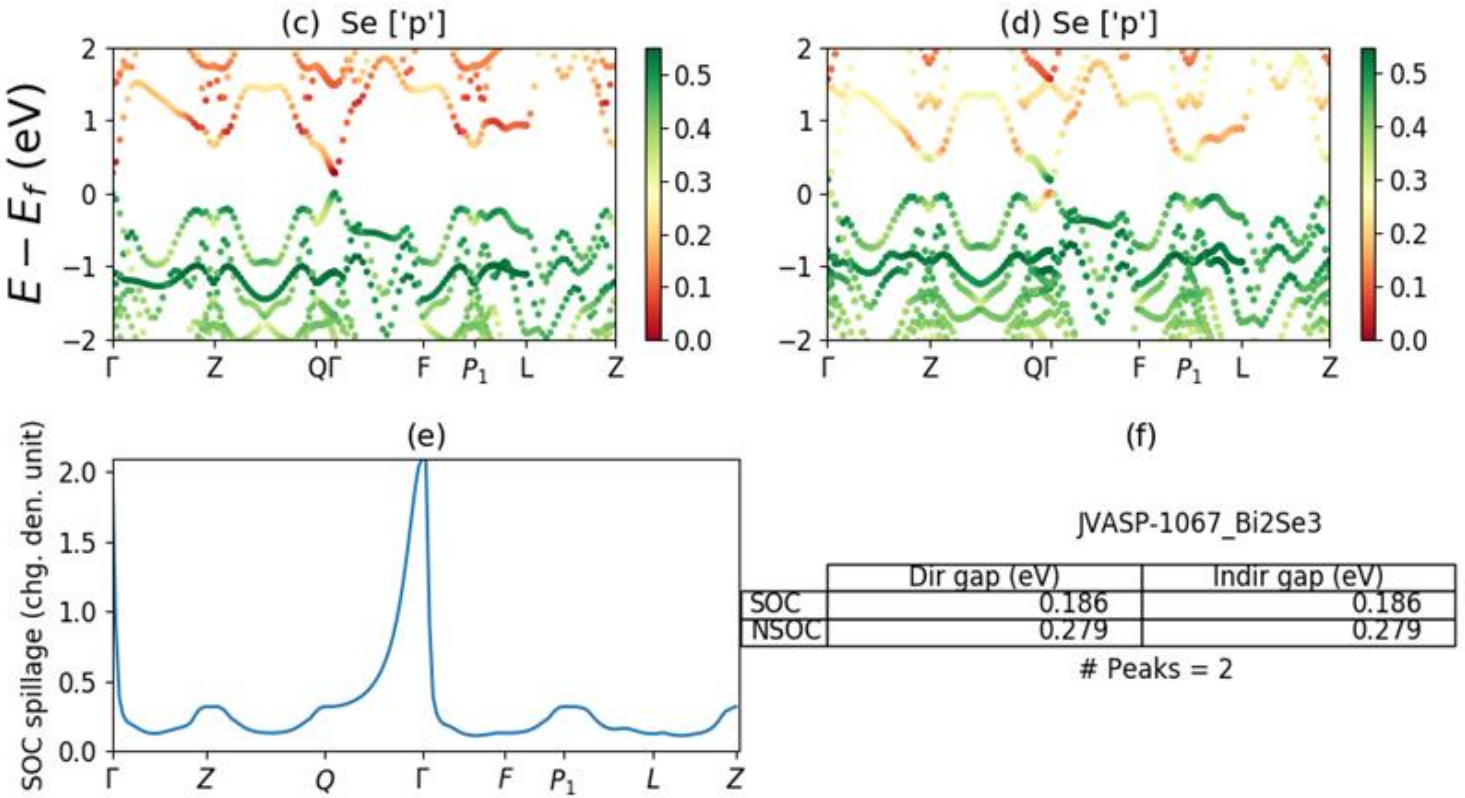


89% successful
Very simple criteria

Scientific Reports 7, 5179 (2017)

	Energy range (meV)	Materials
a)	0-40	$TiNCl$, SiH_4, $HfBrN$, $Mg(AlSe_2)_2$
b)	40-60	$GaSe$, CrS_2, ZrS_3, NiO_2, GaS, $ZrSe_3$, $NdTe_3$, US_3, TiS_3, $PrIO$, $DySI$, Sc_2CCl_2, $ThIN$, $TiBrN$, $InClO$, $LuSBr$, $SrHI$, $BiIO$, $BiBrO$, $KMnP$, $TiIN$, Sc_2NCl_2, $TlSbO_3$, $ZrCl$, $SmTe_3$, $PrTe_3$, As_2O_3, Nb_2CS_2, $RbMnAs$, SiH, Bi_2TeI, $ScCl$, $TbBr$, $Ge(BiTe2)_2$, GaS
c)	60-80	WSe_2, WS_2, MoS_2, C, SnO_2, PtO_2, $CdBr_2$, $ReSe_2$, $CrSe$, $MgCl_2$, $CoBr_2$, $ZrCl_2$, $MgBr_2$, TcS_2, $FeCl_2$, $MnCl_2$, $MnBr_2$, $InSe$, $CrBr_3$, VCl_3, USe_3, $IrCl_3$, $ScCl_3$, $RhCl_3$, TaI_2O, $DySBr$, $ErSeI$, $ErSCl$, $BiClO$, $OsCl_2O$, $CdCl_2$, BN, $Nb(SCl)_2$, Bi_2Te_2S, $ThBrN$, $HfCl_4$, Bi_2Te_2Se, $MgPSe_3$, $CdPS_3$, $ScPS_4$, $PPdS$, $TmAg(PSe_3)_2$, $ScAg(PSe_3)_2$, $ErAg(PSe_3)_2$, $ScAg(PS_3)_2$, Nb_3Cl_8, Nb_3TeCl_7, $InAg(PSe_3)_2$, Hf_3Te_2, $SNCl$, Sr_3Si_2, $TiCl_2$, $HfFeCl_6$, $GaTe$, CS_2, $Nb(SeCl)_2$, $CrCl_3$, BiI, $TiBr_2$, $GaAg(PSe_3)_2$, $CdPS_3$, TiS_2
d)	80-100	NbS_2, $MoSe_2$, $NbSe_2$, WTe_2, $MoTe_2$, VSe_2, ZrS_2, HfS_2, $HfSe_2$, MoS_2, PtO_2, PtS_2, SnS_2, $SnSe_2$, TiO_2, TiS_2, $TiSe_2$, $ZrSe_2$, TaS_2, $SiTe_2$, $TaSe_2$, VS_2, $TaSe_2$, MgI_2, SbI_3, PbI_2, GeI_2, SiS_2, MnI_2, CaI_2, $RhBr_3$, BiI_3, $MoBr_3$, $RuBr_3$, PCl_3, AuI, BPS_4, $IrBr_3$, $Re(AgCl_3)_2$, $AlPS_4$, $AlSiTe_3$, $PPdSe$, $CrSiTe_3$, Nb_3TeI_7, NdI_2, Al_2Te_3, S_5N_6, $AlTeI_7$, $AlSeBr_7$, CdI_2, PSe, Ta_3TeI_7, TmI_2, $SbBr_3$, P_4S_5
e)	100-120	$HfTe_2$, $PtSe_2$, $TiTe_2$, WO_2, SnO, BCl_3, Te_2Br, Te_2I, PBr_3, TiI_3, $BiTeCl$, $BiTeI$, $TlPt_2S_3$, $AlBr_3$, $BiSBr$, CaN, Mn_2Bi, HgI_2, $SrThBr_6$, P
f)	120-140	BBr_3, AlI_3, $TlTe_3Pt_2$, $SbSBr$, $TlPd_2Se_3$, P_2Se_5
g)	140-160	PdS_2, Te_2Pt, BI_3, $Ta(ICl)_2$
h)	160-180	$PdSe_2$, $NiTe_2$, NbI_5
i)	180-200	ZrS, PI_3, $BaBrCl$
j)	200-220	Te_2Pd, Te_2Ir, $BiSe_2$
k)	220-240	$SbSeI$, $BiSI$
l)	240-260	Bi_2Te_3
m)	260-300	VS_2
n)	300-340	$CaSn$, $KAuSe$
o)	340-360	$KAuS$, $RbAuS$, $RbAuSe$
p)	360-440	Sc_2C
q)	440-500	GaN, $TiSe_2$
r)	500-560	Sr_2H_3, Ca_2H_3
s)	560-640	AlN
t)	640-660	$TiTe_2$
u)	660-680	$CrSe_2$
v)	680-700	$ZrTe_2$
w)	700-720	$CoAs_2$
x)	720-960	Ti_2O

Discovery of topological materials

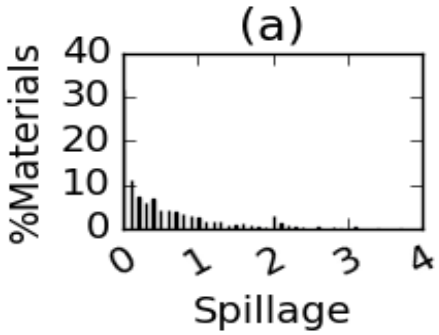


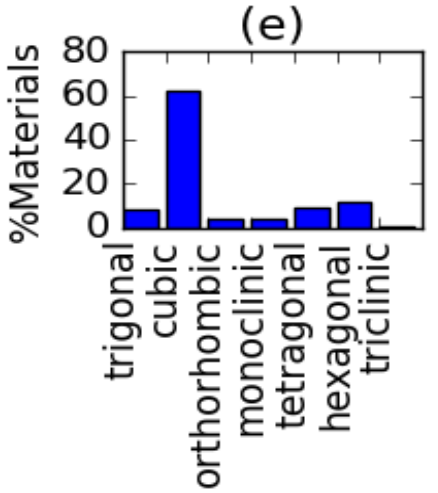
(f)

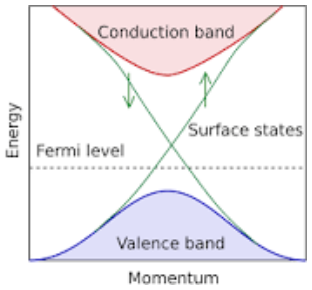
JVASP-1067_Bi2Se3

	Dir gap (eV)	Indir gap (eV)
SOC	0.186	0.186
NSOC	0.279	0.279

Peaks = 2





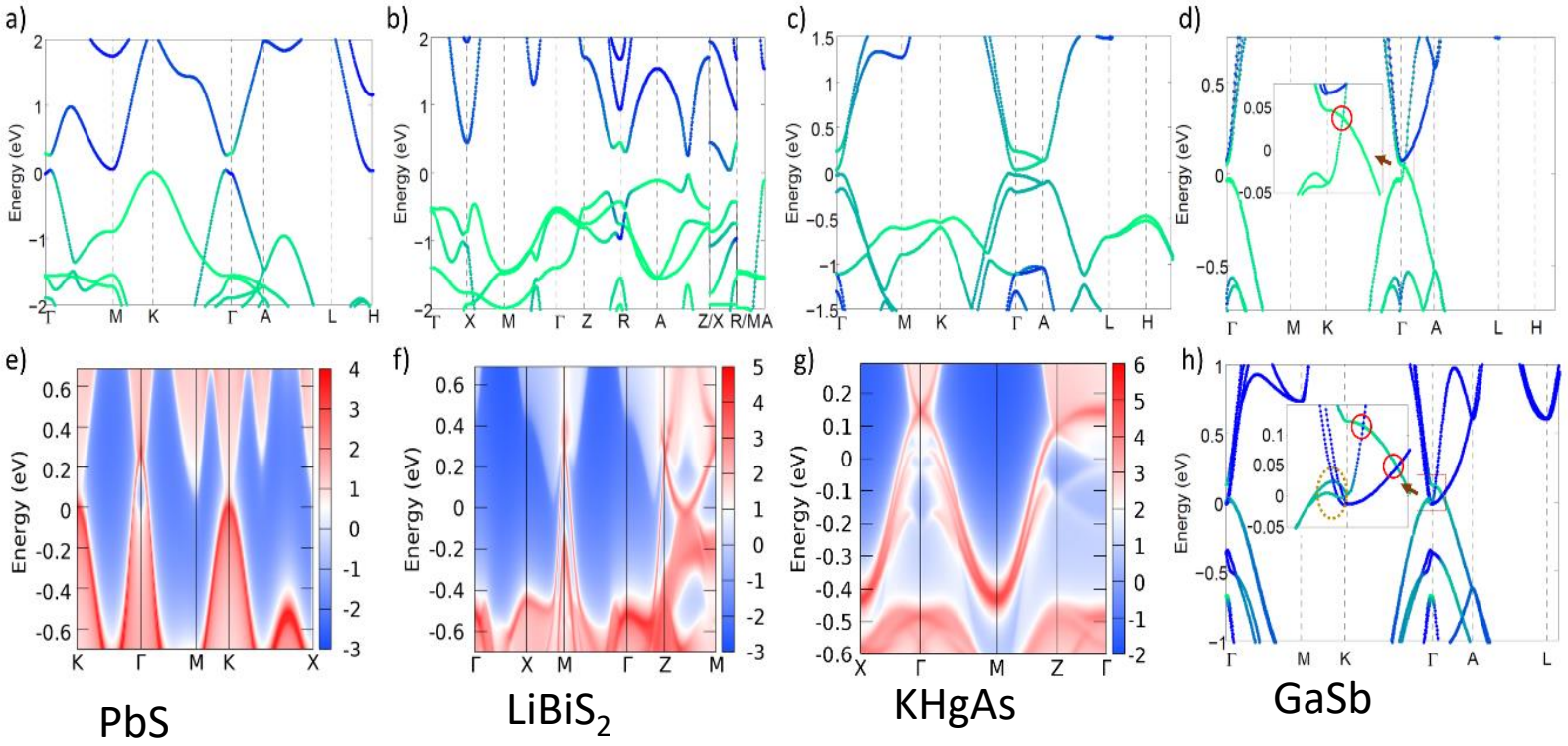


Surface states are topologically protected, aspect for Quantum computing

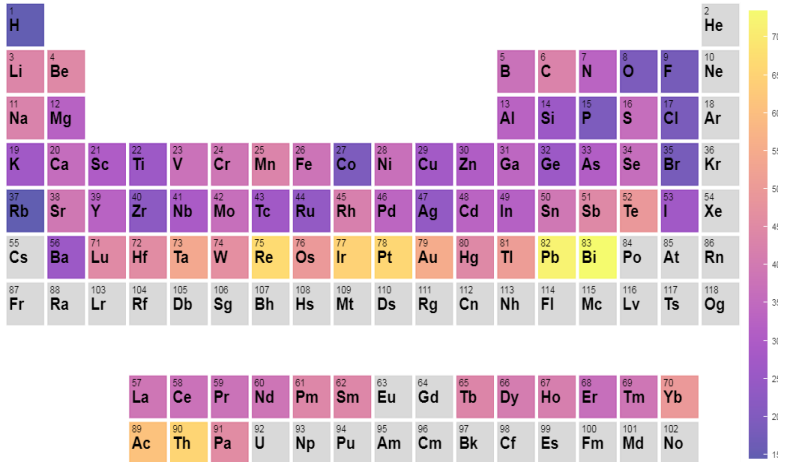
Using DFT wavefunctions w/o spin-orbit coupling to discover topological materials

https://arxiv.org/abs/1810.10640

Discovery of topological materials



PbS $LiBiS_2$ KHgAs GaSb



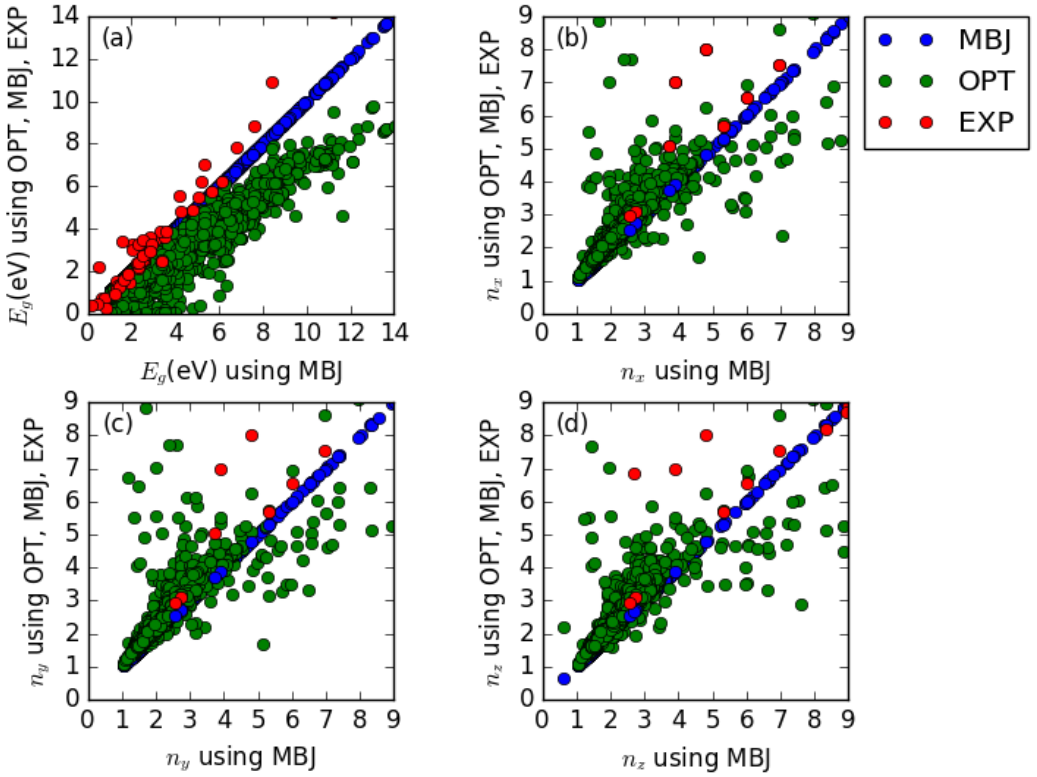
Wannier function, band-crossings, inversion, ARPES

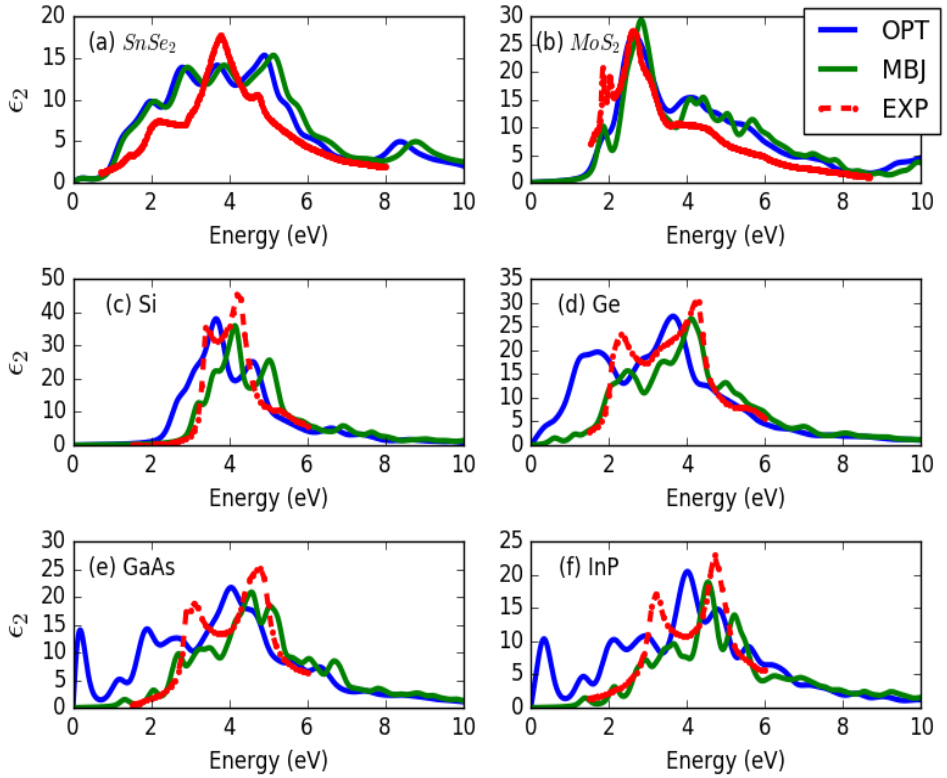
Periodic table trends

- **>1699 high-spillage materials identified**
- **Further screened with close to zero indirect bandgap [ΔEn] (to avoid trivial metallic cases)**
- **278 Wannier calculations**

https://arxiv.org/abs/1810.10640

Discovery of solar cell materials

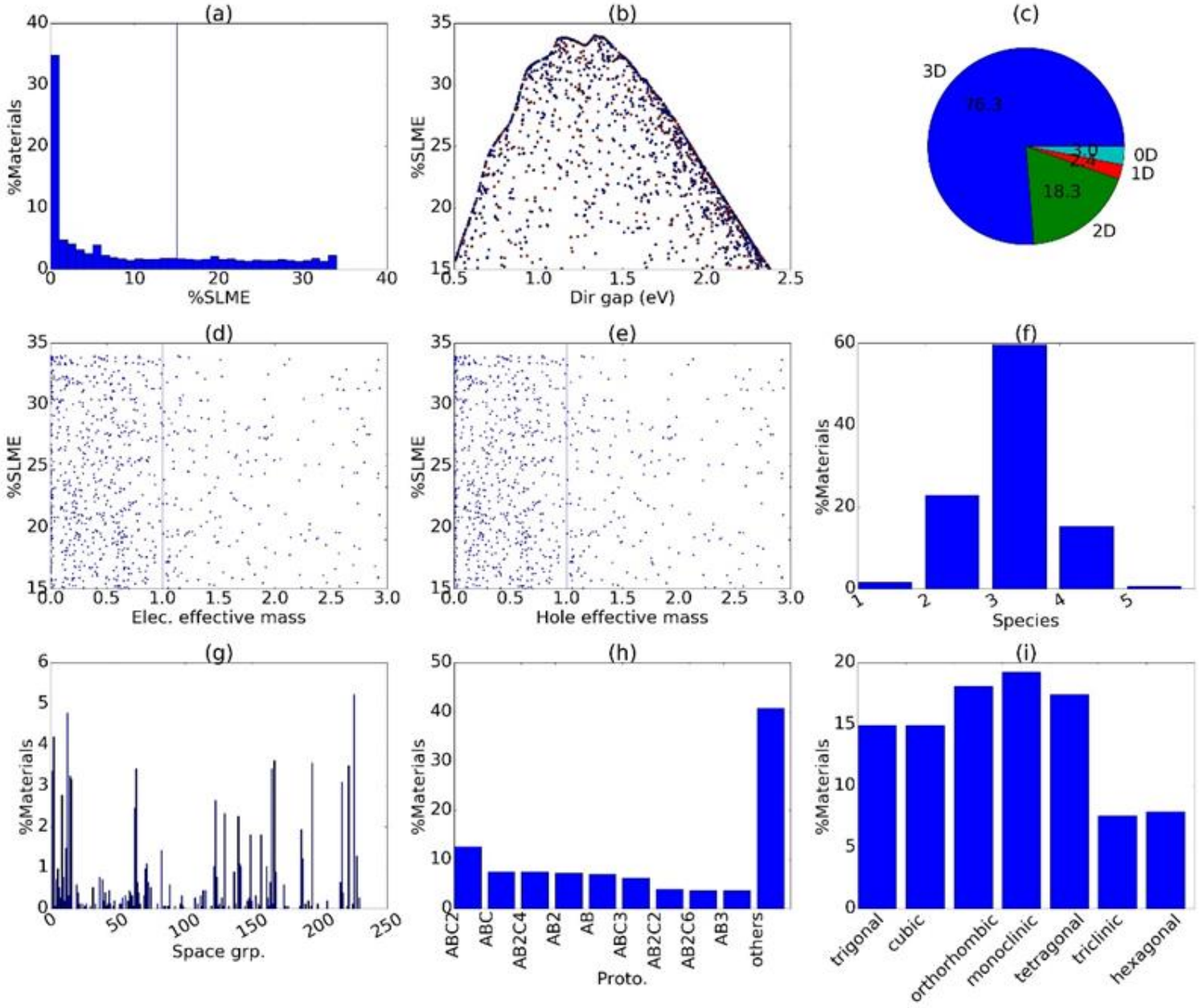


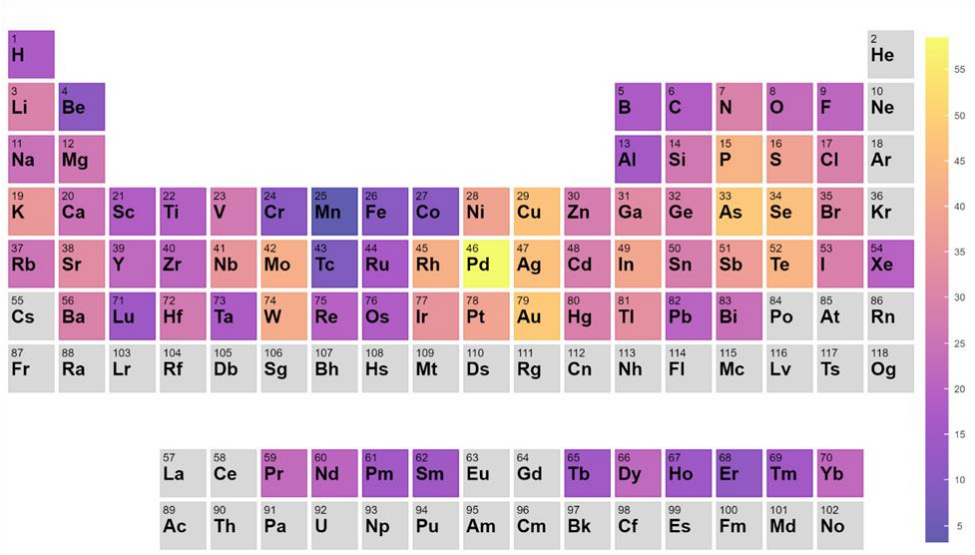


$$\varepsilon^2{}_{\alpha\beta}(\omega) = \frac{4\pi^2 e^2}{\Omega} \lim_{q\to 0} \frac{1}{q^2} \sum_{c,v,k} 2w_{\vec{k}}\,\delta(\zeta_{ck} - \zeta_{vk} - \omega) \left\langle \psi_{c\vec{k}+\vec{e}_\alpha q} \middle| \psi_{v\vec{k}} \right\rangle \left\langle \psi_{c\vec{k}+\vec{e}_\beta} \middle| \psi_{v\vec{k}} \right\rangle^*$$

Frequency dependent dielectric function with TBmBJ

Scientific Data 5, 180082 (2018)





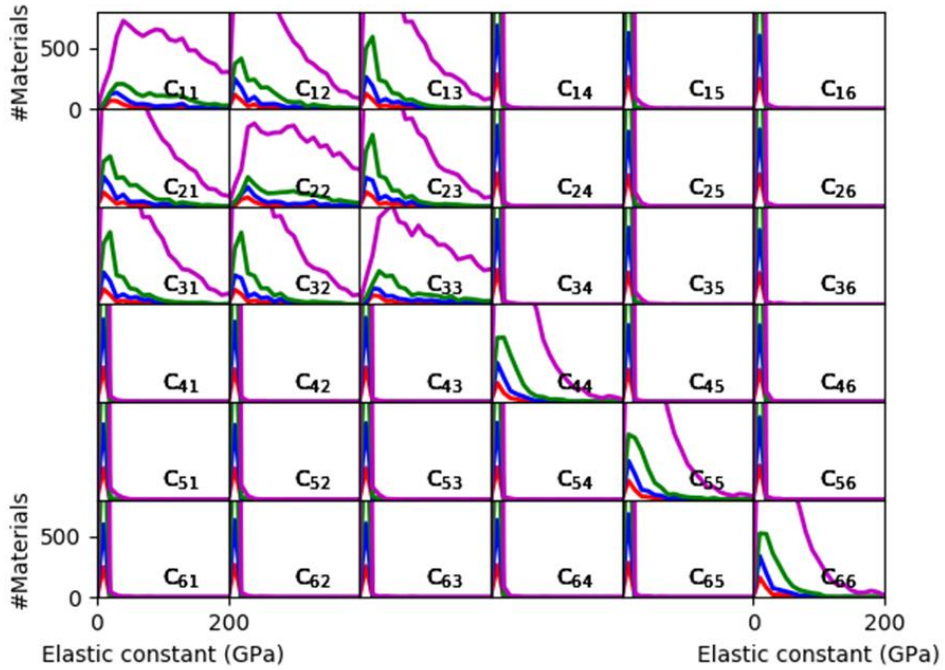
Spectroscopic Limited Maximum Efficiency (SLME)

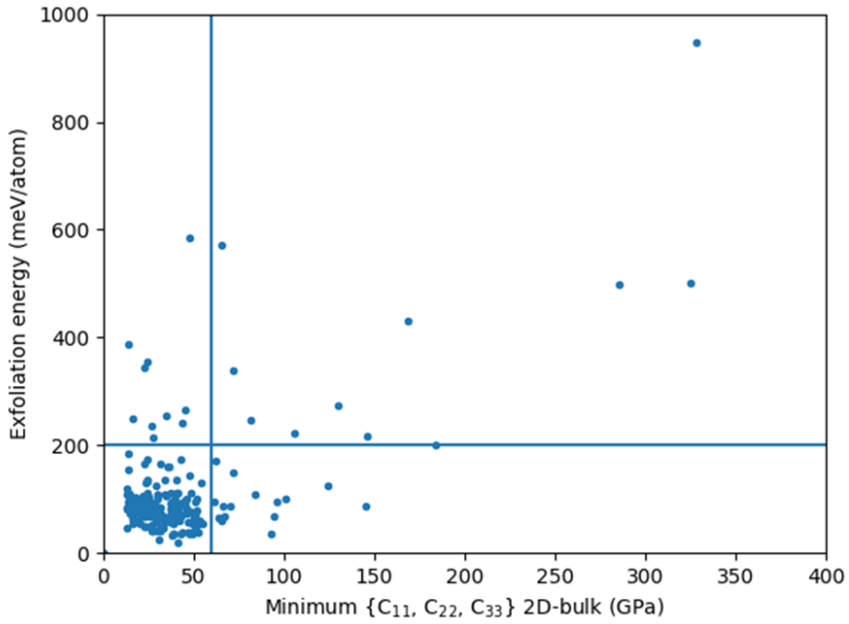
$$\eta = \frac{P_{mat}}{P_{inc}} = \frac{\left(J_{sc} - J_0\left(e^{\frac{eV}{KT}} - 1\right)\right)V}{1000}$$

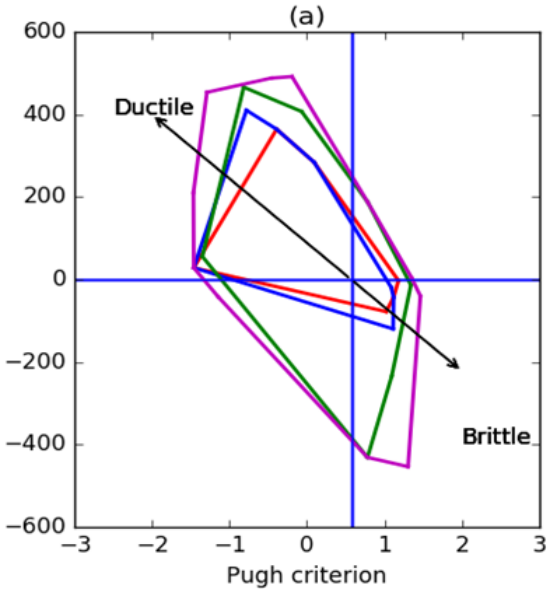
>1500 screened materials

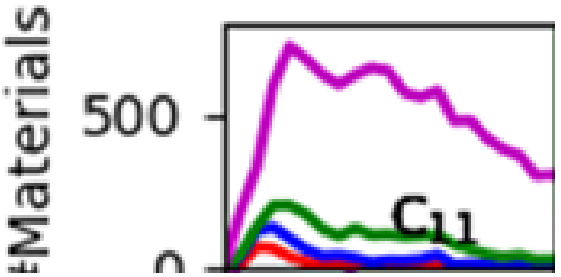
Paper in preparation

Discovery of negative Poisson and flexible materials









PbS (−0.5, Cmcm, JVASP-28369), Al (−6.2, Im−3m, JVASP-25408), CSi_2 (−0.13, P6/mmm, JVASP-16869), YbF3 (−0.06, Pnma, JVASP-14313), SiO_2 (−0.03, Pna21, JVASP-22571)

Physical Review B 98 (1), 014107 (2018)

Dimensionality dependent elastic properties

2D-Heterostructure design with elastic constant and work-function data

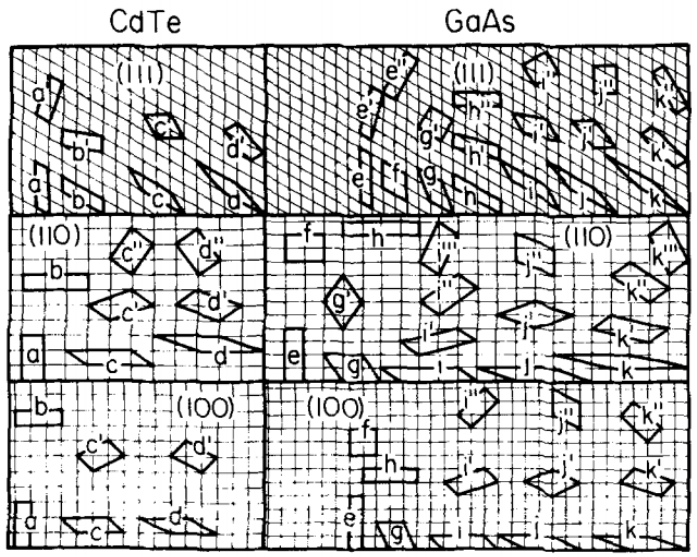
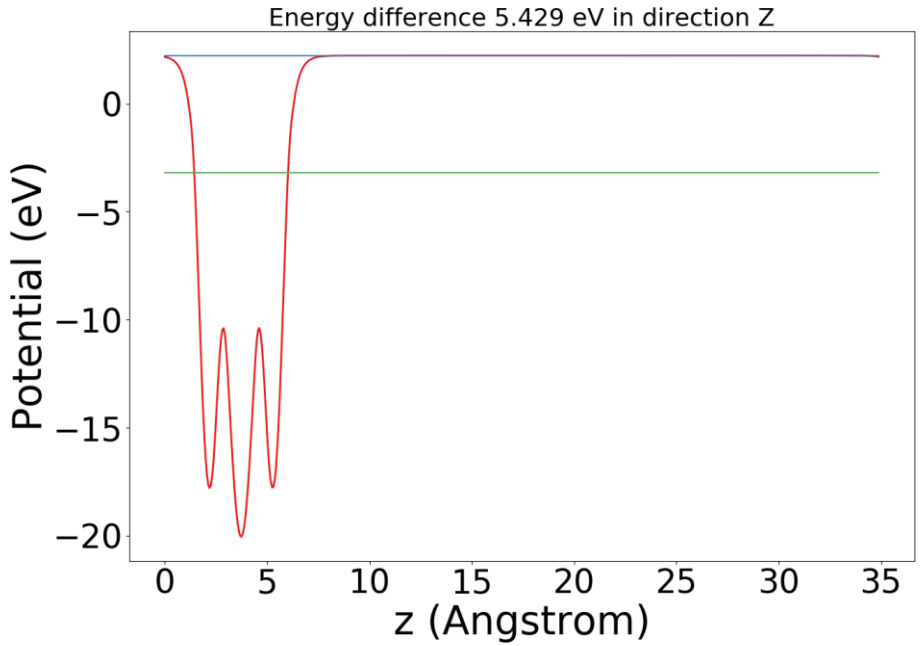


FIG. 3. Lattice translations parallel to the (100), (110), and (111) faces of CdTe and GaAs. Nonequivalent unit cells of order 3 for CdTe and 4 for GaAs are shown as parallelograms bounded by thick lines and denoted by unprimed letters. Various stages of the reduction procedure are denoted by primed letters.

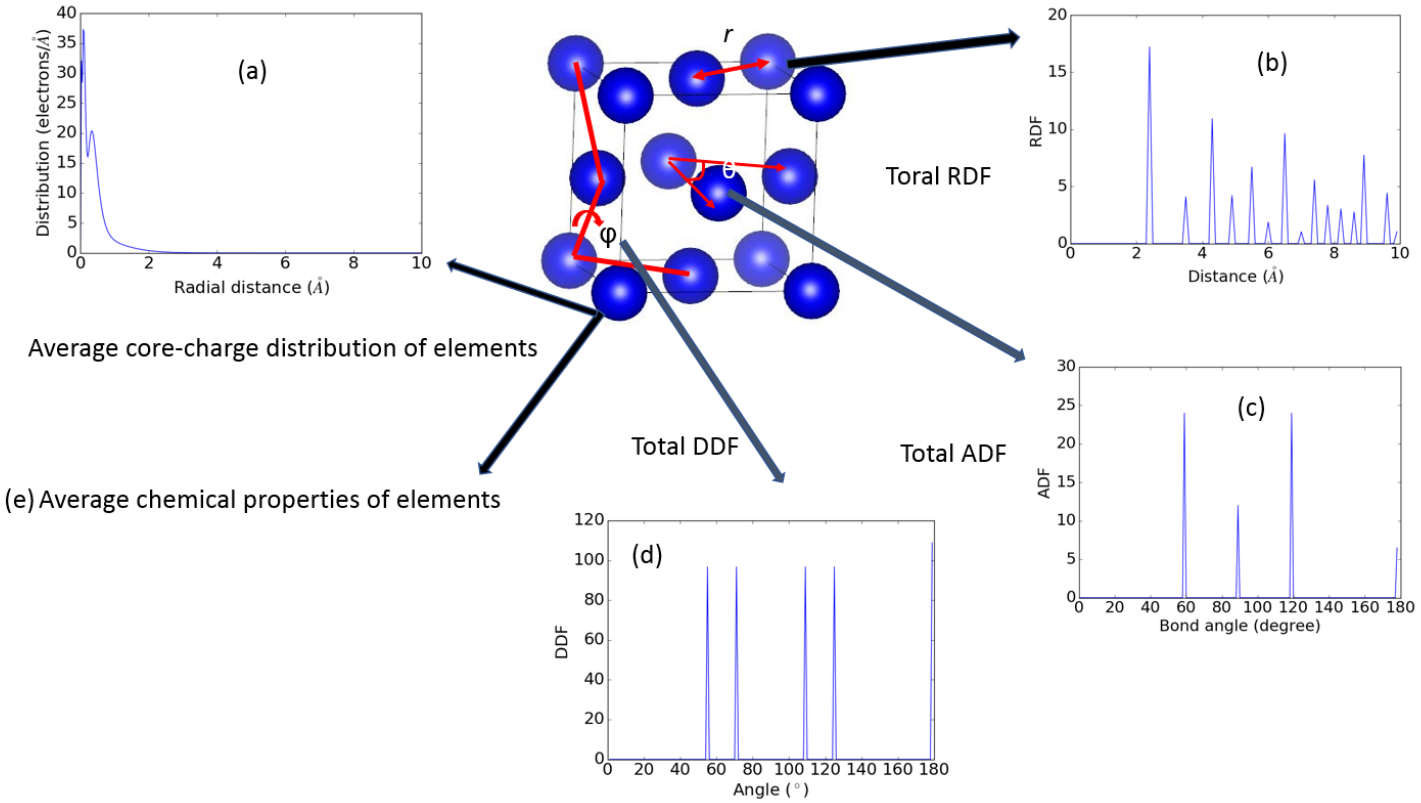


Zur, Journal of Applied Physics **55**, 378 (1984)

- **Lattice mismatch and elastic tensor to predict elastic energy of the interface**
- **Work-function/band-alignment match**

In preparation

Finding the right features/descriptors: ML

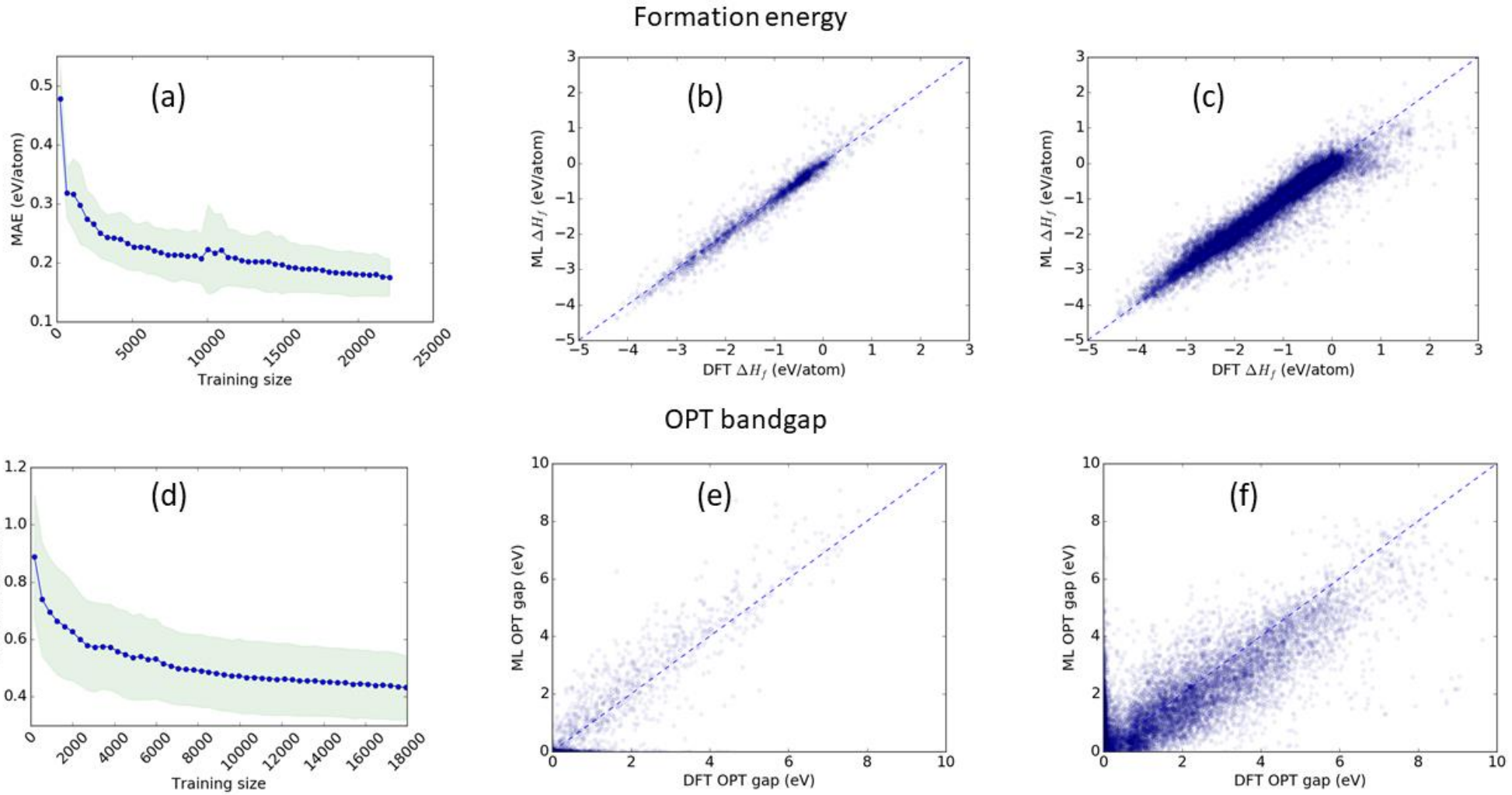


- Arithmetic operations (mean, sum, std. deviation…) of **electronegativity, atomic radii, heat of fusion**,….
of atoms at each site
(example: Electronegativity of Mo+Mo+S+S+S+S)/6 = 0.15
- Atomic bond distance based descriptors
- Angle based descriptors

1557 descriptors/features for one material
GBDT algorithm

Physical Review Materials 2, 083801 (2018)

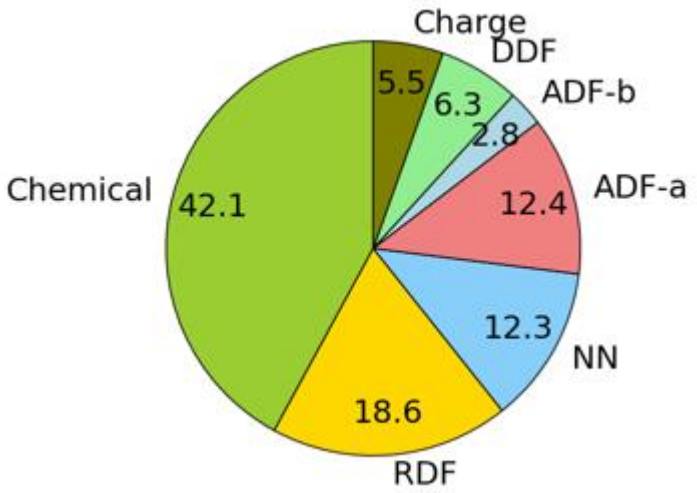
Regression models: formation energy and bandgap model

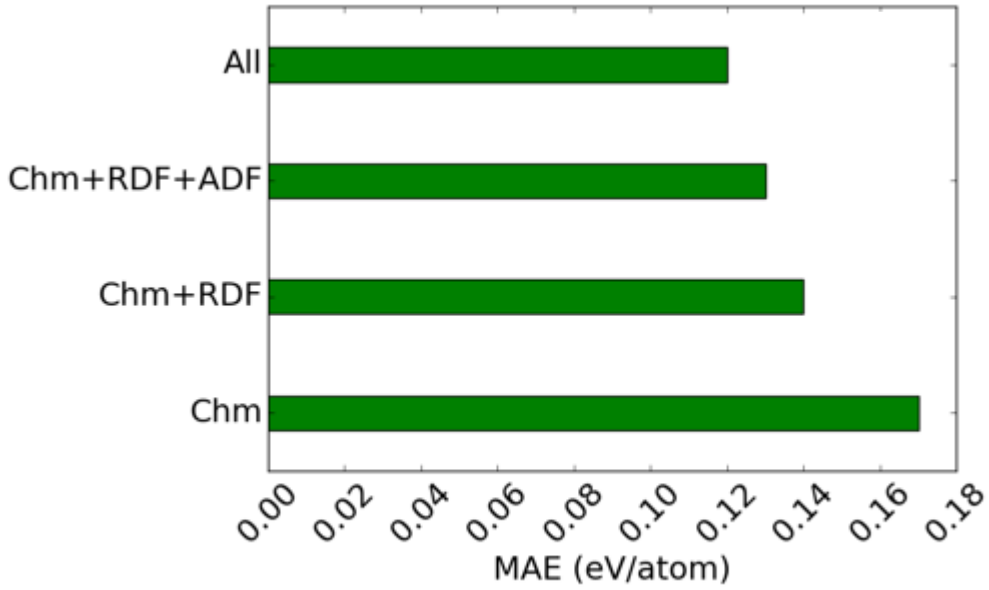


Learning curve, 1 and 5-fold CV plots

Physical Review Materials 2, 083801 (2018)

Explainability: feature importance

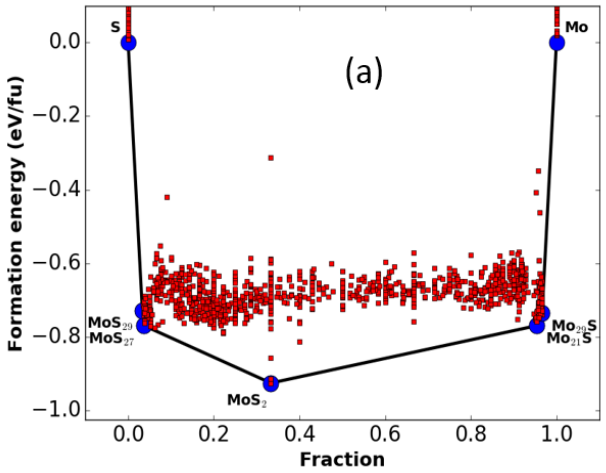


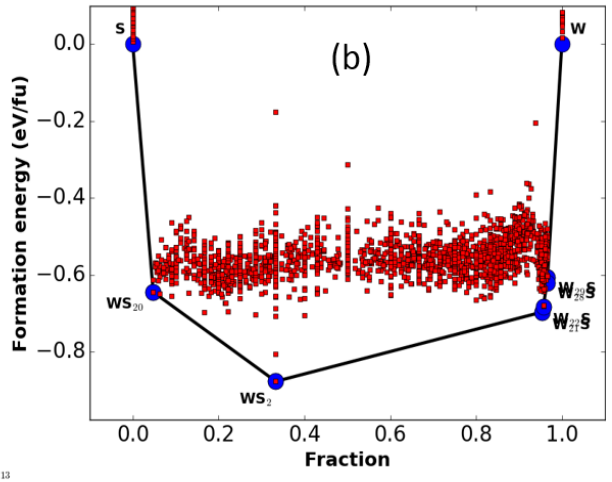


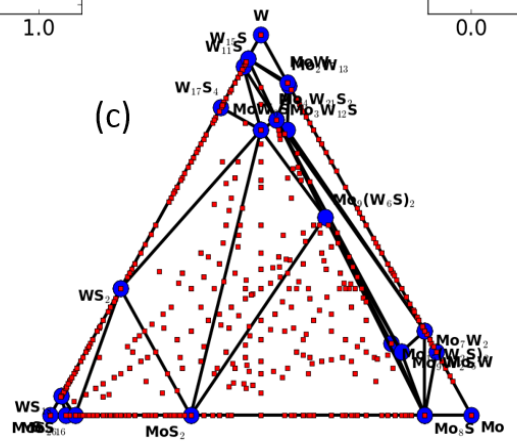
- **Chemical features most important followed by RDF and NN**
- **Incrementally adding structural features decreases MAE**

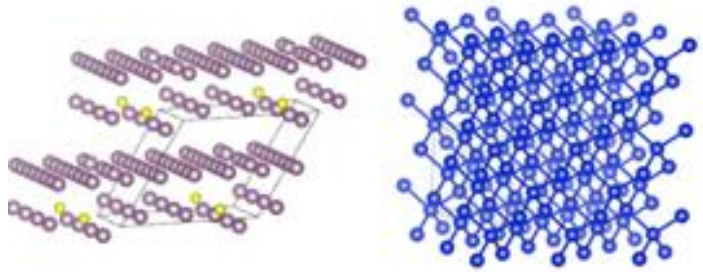
Physical Review Materials 2, 083801 (2018)

Search for new materials: Genetic algorithm with ML





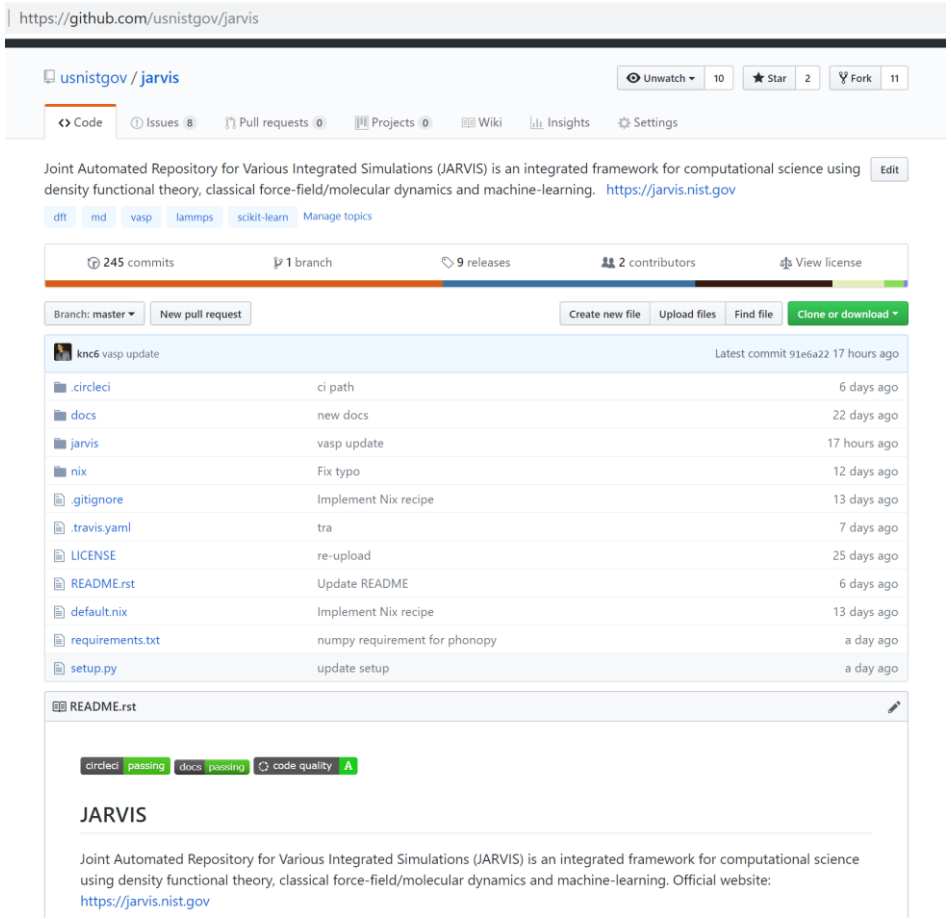


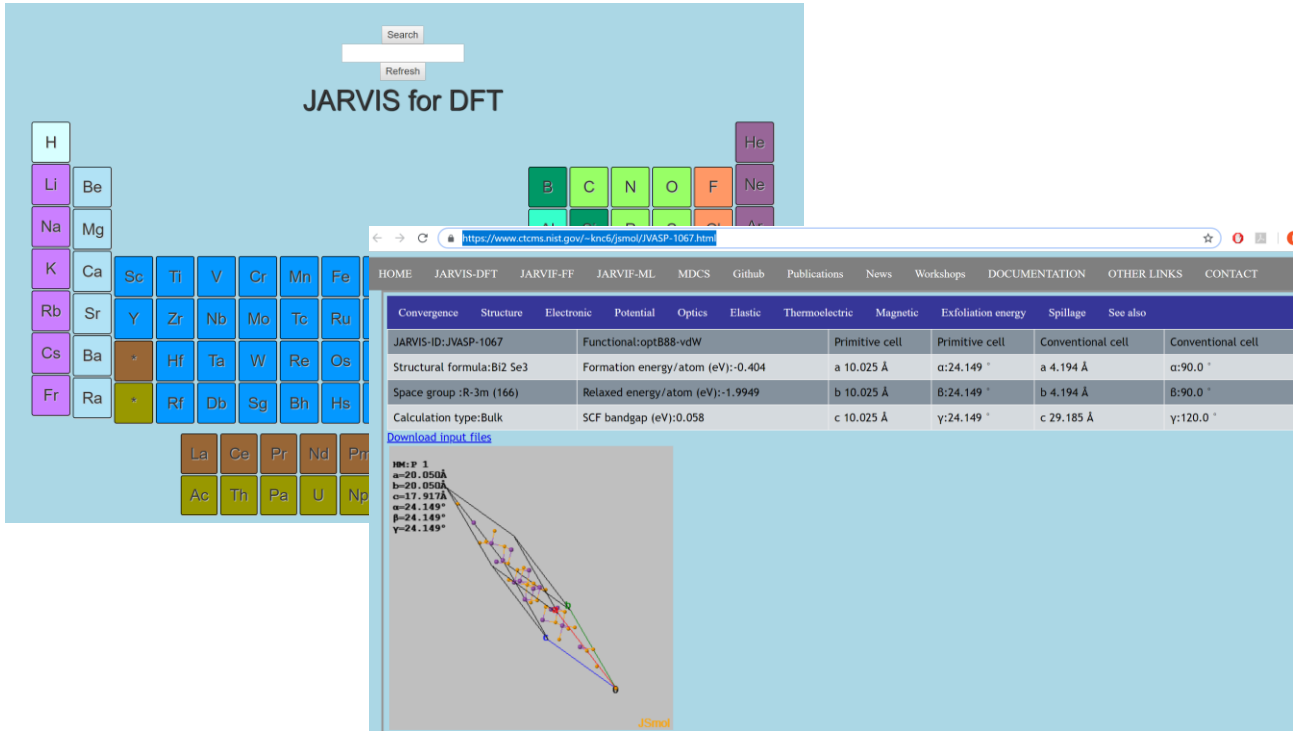


- **New way of validating ML model for materials**
- MoS_2, WS_2 indeed stable as in DFT and experiments
- Need further verification foor low-lying energy structures with DFT

Physical Review Materials 2, 083801 (2018)

Exploring the github and webpages





[Example: https://www.ctcms.nist.gov/~knc6/jsmol/JVASP-1067.html](https://www.ctcms.nist.gov/~knc6/jsmol/JVASP-1067.html)

Github:
https://github.com/usnistgov/jarvis

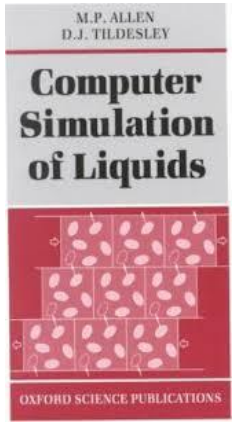
Continuous integration, various modules, notebooks etc.

Exploring material properties

JARVIS-FF

Force-field (classical)

$$F = ma = -\nabla V(r)$$



- **Solve Newton's equation for atomic positions**
- **Approximations for V (force-fields)**:

EAM, EIM, MEAM, AIREBO, REAXFF, COMB, COMB3, TERSOFF, SW *etc.*

- **Contains:**

Automated LAMMPS based force-field calculations on DFT geometries. Some of the properties included in JARVIS-FF are energetics, elastic constants, surface energies, defect formations energies and phonon frequencies of materials

- **Time:** Takes years to fit FFs, relatively quick calculations
- **Website**: https://www.ctcms.nist.gov/~knc6/periodic.html
- **Publications**:

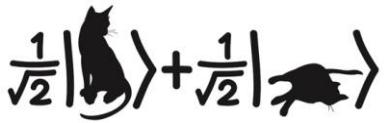
➢ Nature:Scientific Data 4, 160125 (2017)
➢ arXiv:1804.01024 (2018)

Email: knc6@nist.gov or kamal.choudhary@nist.gov

JARVIS-DFT

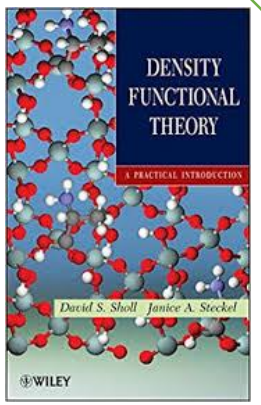
Density-functional theory (quantum)

Schrödinger's cat



$$H\psi = E\psi$$

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{Eff}(r)\right]\psi_i(r) = E_i(r)\psi_i(r)$$



$$V_{Eff} = T + V_{Ne} + V_{ee} + V_{XC}$$

- **Solve Schrödinger equation for electrons**
- >30,000 materials data (3D, 2D, 1D, 0D)
- **Contains:**

Formation energy, exfoliation energy, diffraction pattern, radial distribution function, band-structure (SOC/Non-SOC), density of states, carrier effective mass, temperature and carrier concentration dependent thermoelectric properties, elastic constants and gamma-point phonons

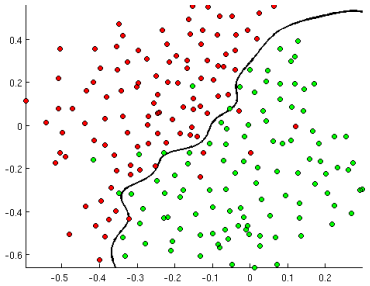
- **Time:** 5000 cores for last 4 years
- **Website**: https://www.ctcms.nist.gov/~knc6/JVASP.html
- **Publications**:

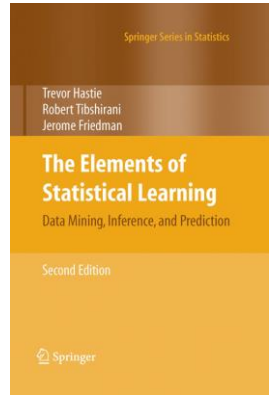
➢ Nature:Scientific Reports 7, 5179 (2017)
➢ Nature:Scientific Data 5, 180082 (2018)
➢ Phys. Rev. B 98, 014107 (2018)

Web: https://jarvis.nist.gov

JARVIS-ML

Machine learning (data-driven)





- **Drawing the line, dimensionality reduction, curve-fitting?**
- Neural nets, decision trees, fuzzy-logic etc.
- Uses gradient boosting decision tree
- **Contains:**

Machine learning prediction tools, trained on JARVIS-DFT data.

Some of the ML-predictions focus on energetics, heat of formation, GGA/METAGGA bandgaps, bulk and shear modulus, exfoliation energy, refractive index, magnetic moment, carrier effective masses

- **Time:** Much easier and faster to train
- **Website**: https://www.ctcms.nist.gov/jarvisml/
- **Publication**:

➢ Phys. Rev. Materials 2, 08380 (2018)

Github: https://github.com/usnistgov/jarvis