Rensselaer

COLUMBIA UNIVERSITY

Emergent spin phenomena in the age of artificial intelligence

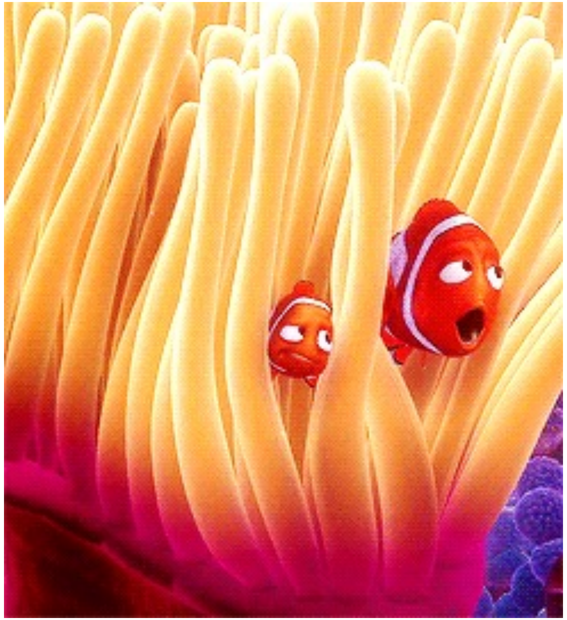
Trevor David Rhone

Department of Physics, Applied Physics and Astronomy,, Rensselaer Polytechnic Institute

Symposium to Honor the Life and Work of Aron Pinczuk

Aron’s Early Influence - Graduate school

Aron’s Early Influence - Graduate school




Lindt
SWISS PREMIUM
CHOCOLATE
CRISTA
Assorted chocolates
350 g
Chocolates surtidos


EMAIL

Opportunities and resources!

Biggest influence – Research career

- o Emergent behavior
  - Low dimensional electron systems
  - Spin degrees of freedom
  - Charge degrees of freedom
- o Data analytics and artificial intelligence
  - Molecular beam epitaxy (MBE) of graphene story
  - High-dimensional optimization problem?

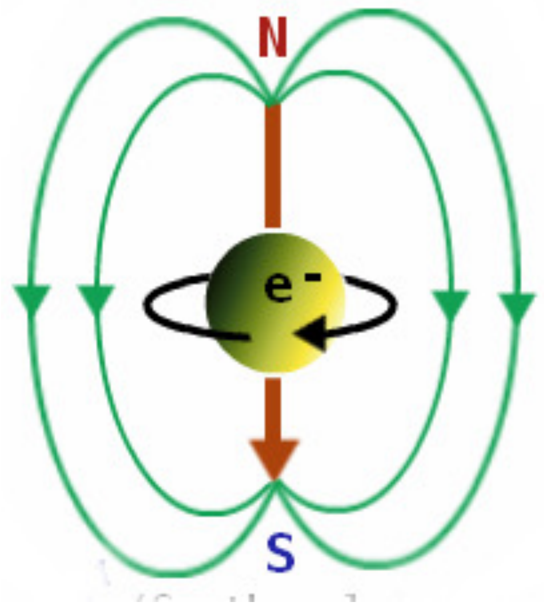
4 August 1972, Volume 177, Number 4047

SCIENCE

More Is Different

Broken symmetry and the nature of the hierarchical structure of science.

P. W. Anderson



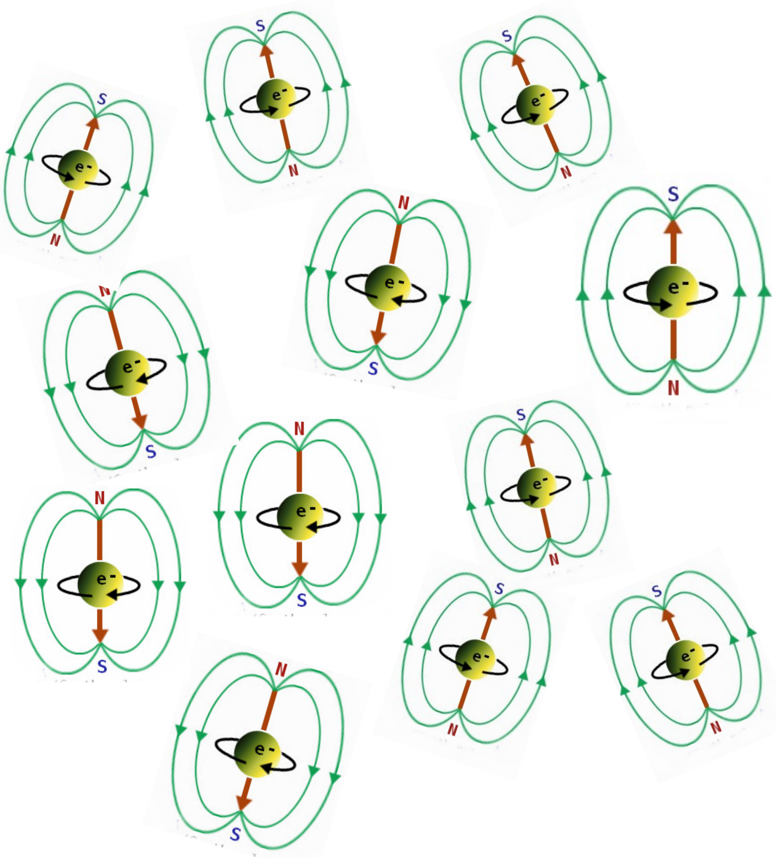
4 August 1972, Volume 177, Number 4047

SCIENCE

More Is Different

Broken symmetry and the nature of the hierarchical structure of science.

P. W. Anderson



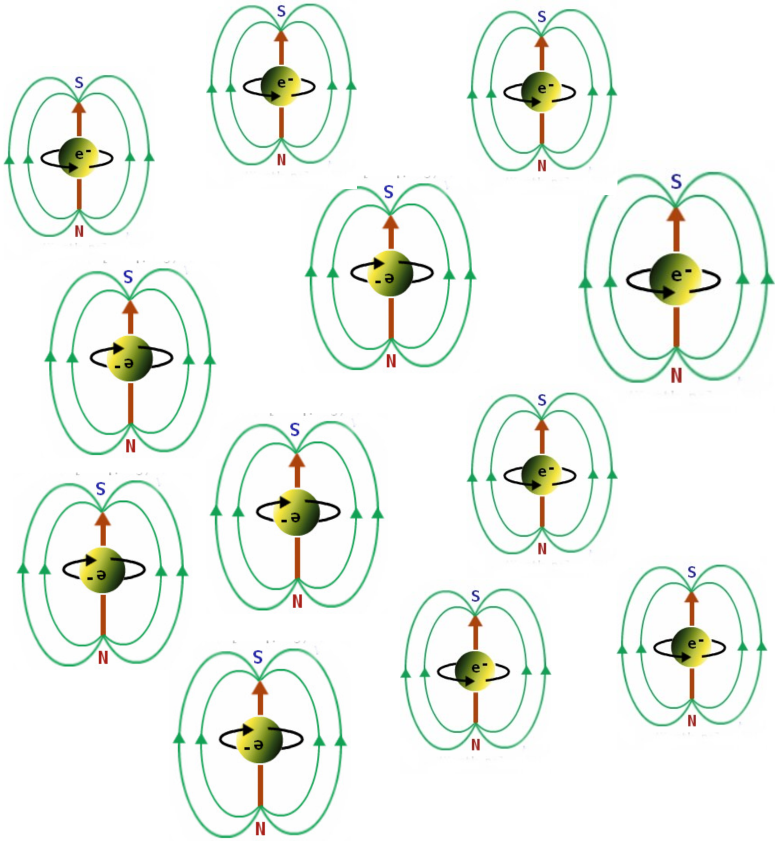
4 August 1972, Volume 177, Number 4047

SCIENCE

More Is Different

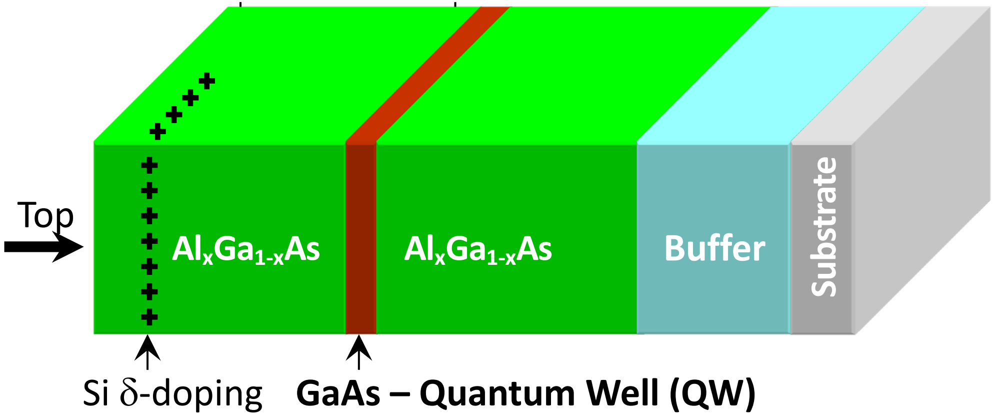
Broken symmetry and the nature of the hierarchical structure of science.

P. W. Anderson



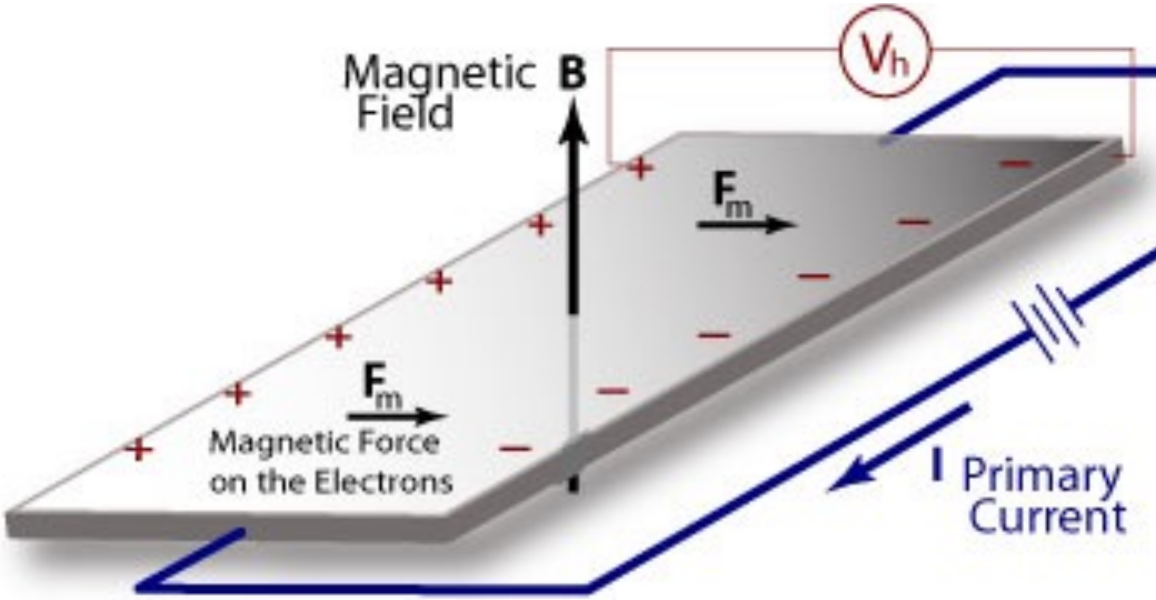
Spin properties in GaAs quantum wells

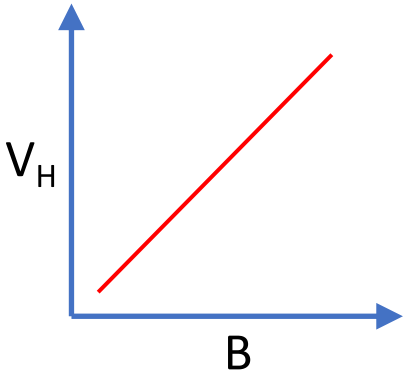
GaAs Quantum Well



- 270 Angstroms
- $n = [0.5, 4.2] \times 10^{11}$ cm^{-2}
- $\text{Mobility}_{max} = 1.1 \times 10^{7}$ cm^{2}V^{-1}s^{-1}
- V_{gate}

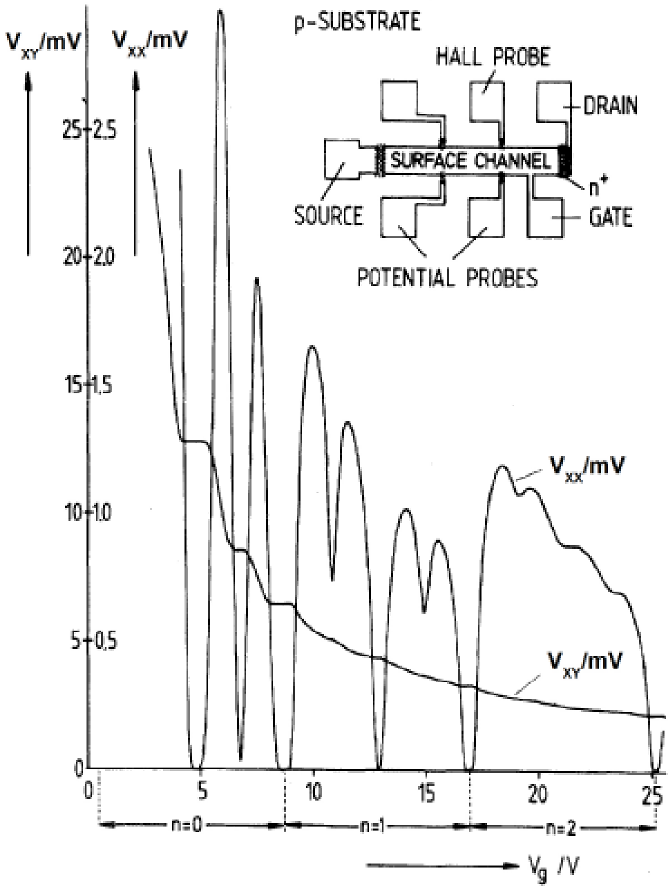
2DES in a magnetic field: Classical Hall Effect





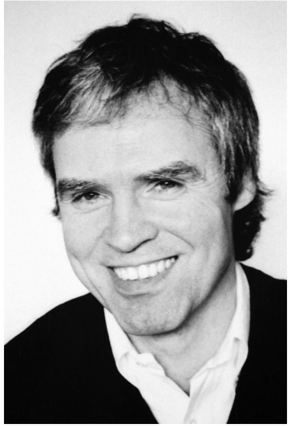
2DES in a magnetic field: Quantum Hall Effect

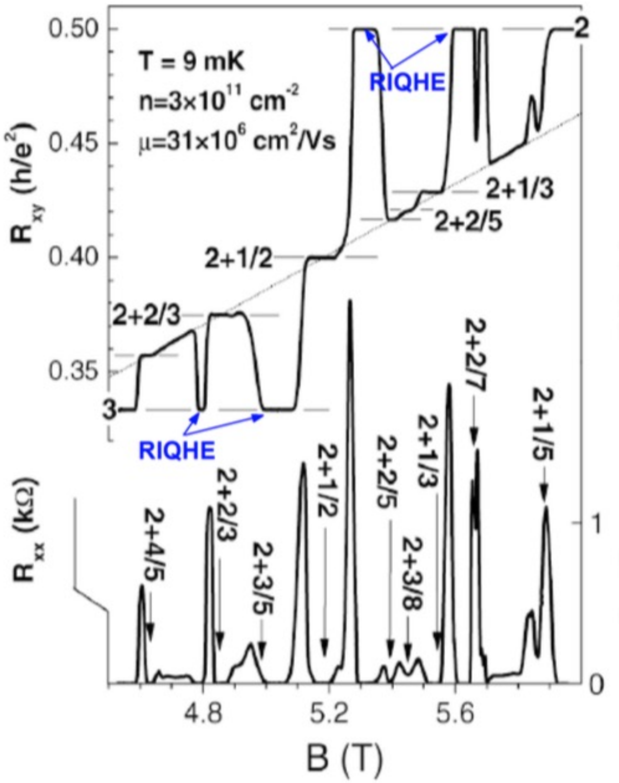
Integer Quantum Hall Effect



1. **K. von Klitzing**, G. Dorda, and M. Pepper.. Phys. Rev. Lett., 45(494), 1980.

2DES in a magnetic field: Fractional Quantum Hall Effect



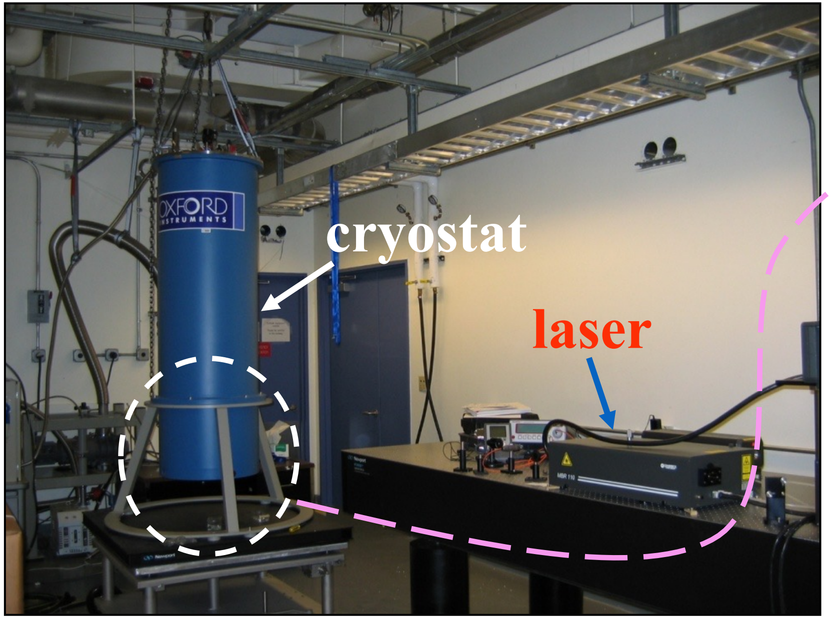


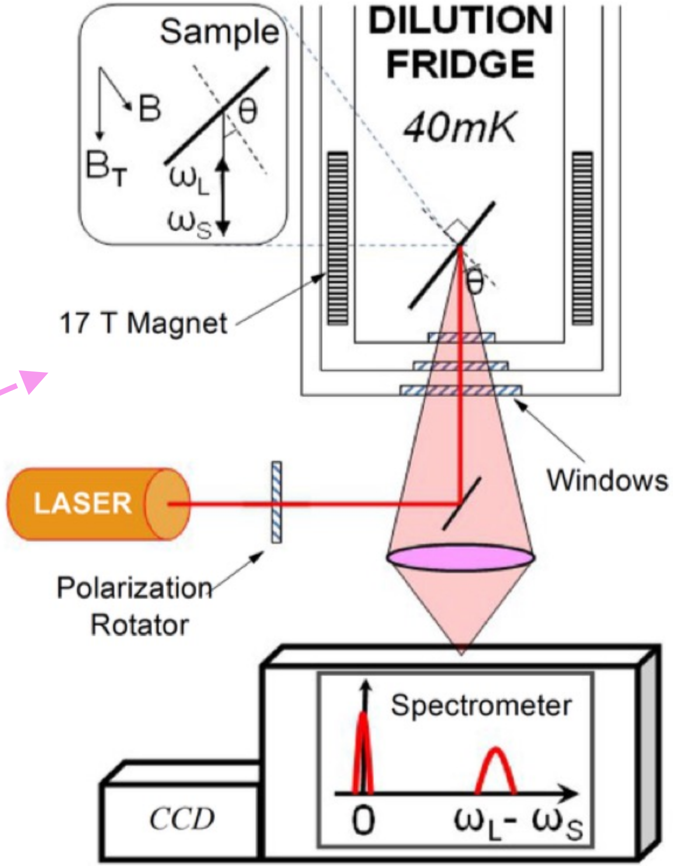
J. S. Xia, et. al. Phys. Rev. Lett., 93 176809 (2004)

- Exotic even-denominator state at $\nu = 5/2$
- Topologically protected?
- Implement topological quantum computation?

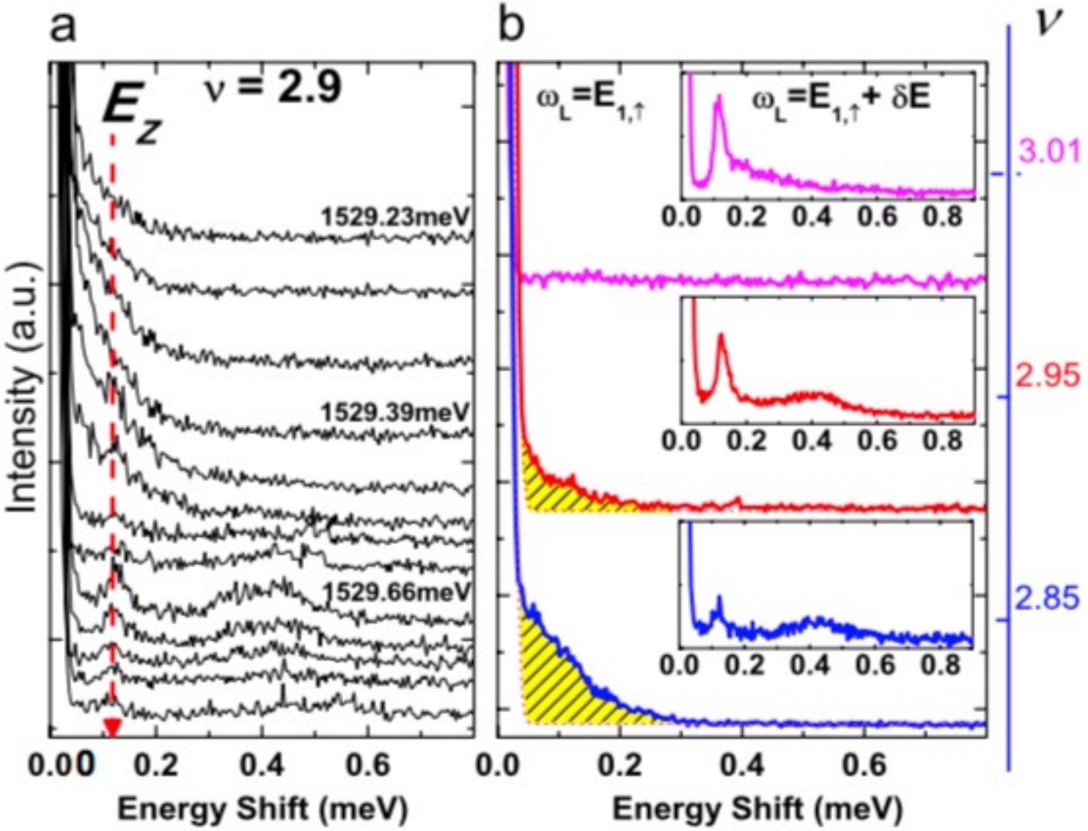
Milli-Kelvin Spectroscopy

- Dilution fridge: 40mK < T < 2K
- Tilt-angle, θ = 20 degrees
- Titanium:Sapphire laser with fine tunability



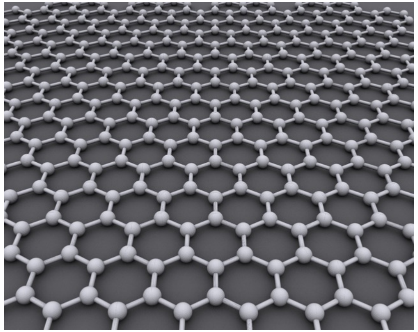


Spin properties of 2DES

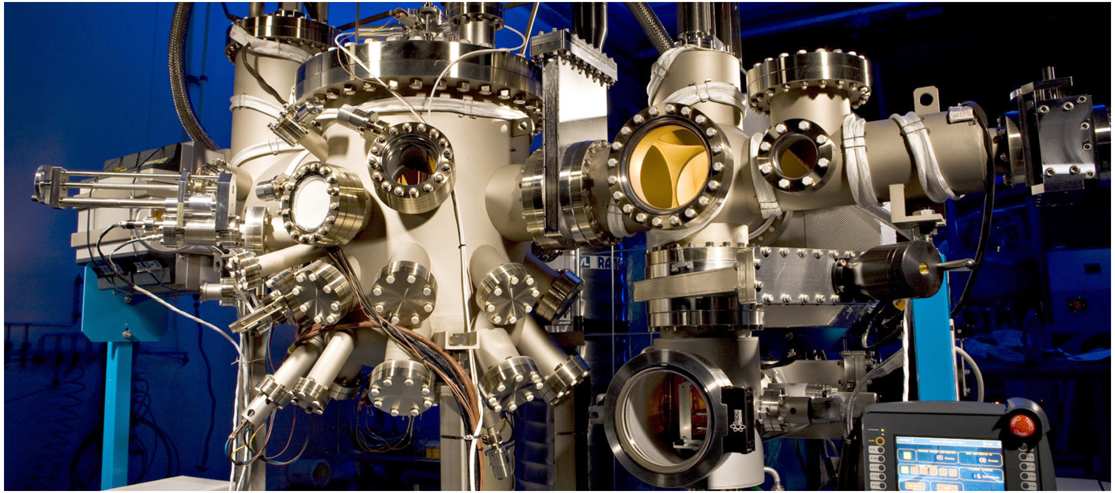


T.D. Rhone et al, PRL 106, 196805 (2011)

Synthesis of graphene by molecular beam epitaxy



- Grow graphene
- High-dimensional parameter space:
  - Substrate temperature
  - Partial pressure of Argon
  - Target temperature
  - Annealing temp.
  - Annealing time
- Measure quality of grown material



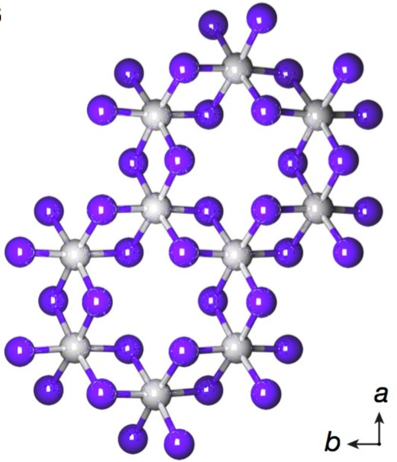
Aron the Influencer – Recently

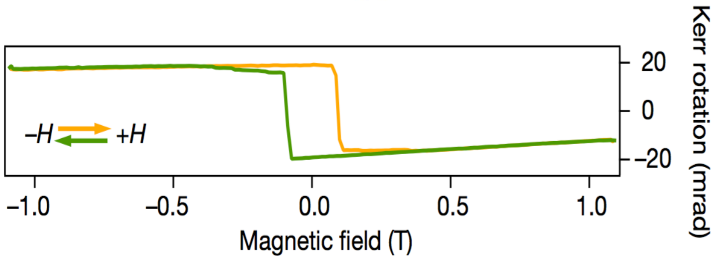
Harvard University

Rensselaer Polytechnic Institute

Two-dimensional Magnetic Materials

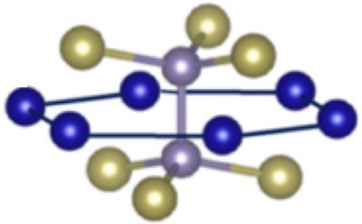
CrI_3

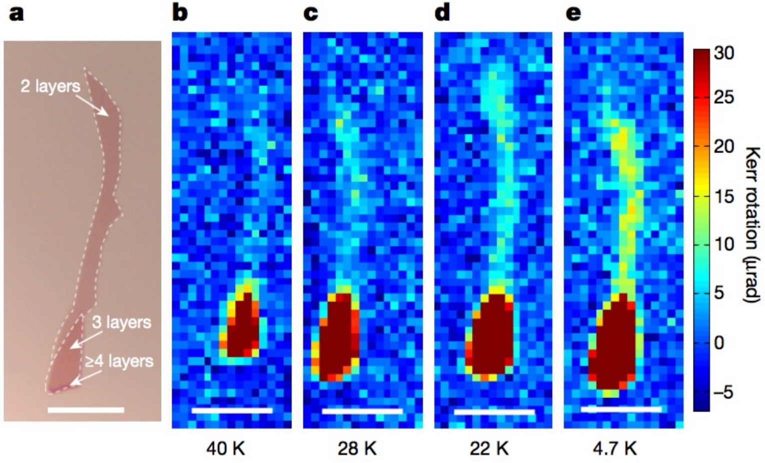




Huang et al., Nature **546**, 270 (2017)

$Cr_2Ge_2Te_6$

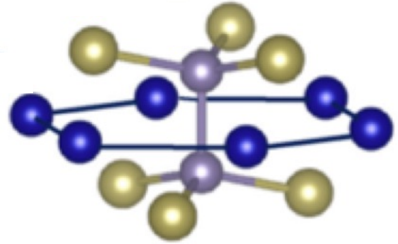




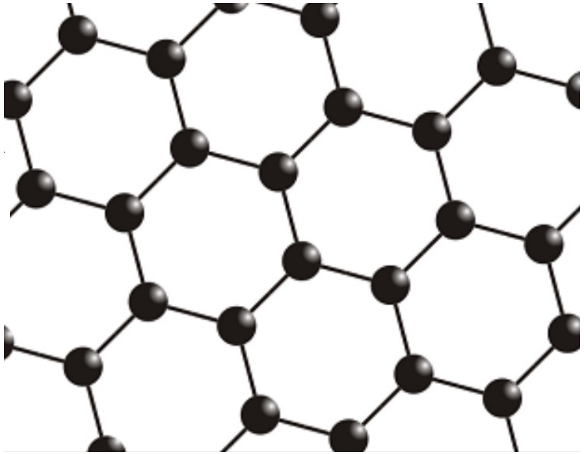
Gong et al., Nature **546**, 265 (2017)

Data-driven study of 2D materials?

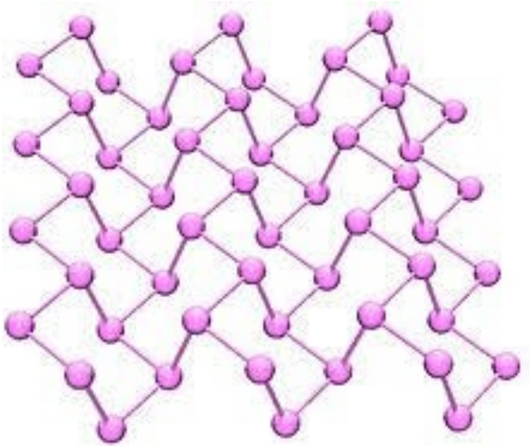
$CrGeTe_3$

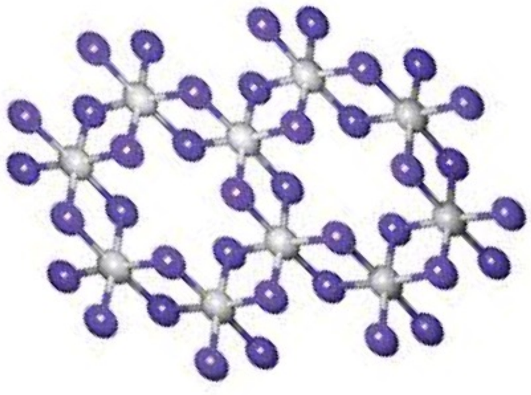


Graphene



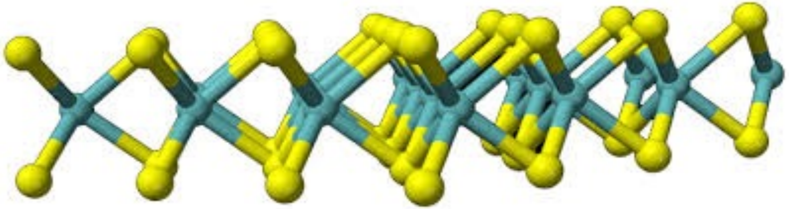
Black phosphorus





CrI_3

MoS_2



Data-driven study of 2D magnetic materials

Motivation:

1. Materials discovery
   - 2D magnetic materials
   - Chemical stability
2. Knowledge discovery
   - Magnetic properties of 2D materials

Discover novel magnetic 2D materials
using materials informatics

Data-driven study of 2D magnetic materials

Materials Informatics

Materials science + Artificial intelligence

$$\underbrace{y}_{\substack{\text{Magnetic}\\ \text{moment}}} = \underbrace{f(x_1, x_2, \ldots, x_N)}_{\substack{\text{Number of spin up electrons}\\ \text{Number of valence electrons}\\ \ldots\\ \text{Electronegativity}}}$$

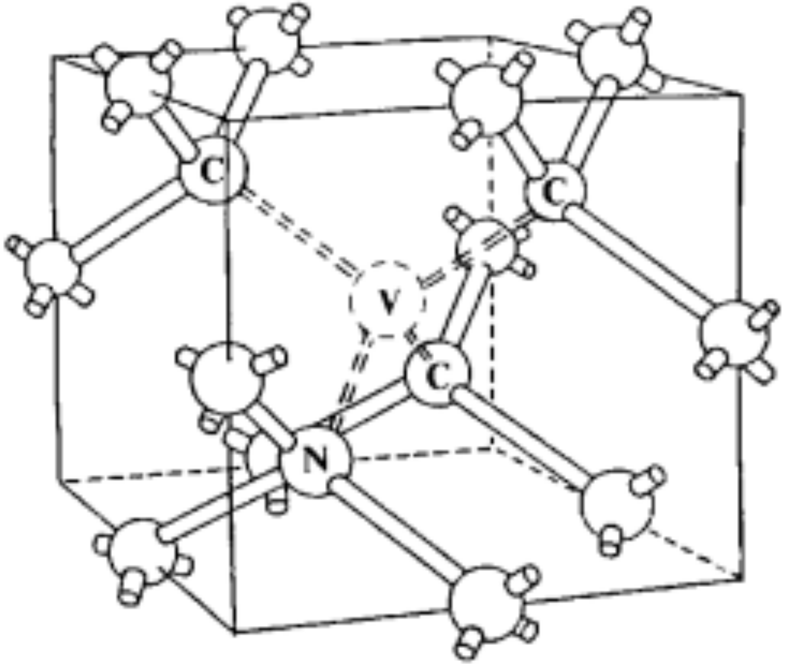
Why use artificial intelligence to study materials?

- Materials search space is huge
  - ICSD (inorganic crystal structure database)
    - 200,000 inorganic compounds
  - Chemical Abstracts Service
    - 49,037,297 organic/inorganic entries
  - Virtual Chemistry Space[1]
    - $\sim 10^{100}$ candidates for materials
- First principles calculations are expensive
- Experiments are slow and expensive

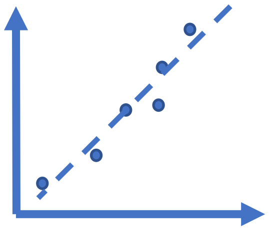
[1] W. P. Walters, M. T. Stahl, and M. A. Murcko, Drug Discov. Today 3, 160 (1998).

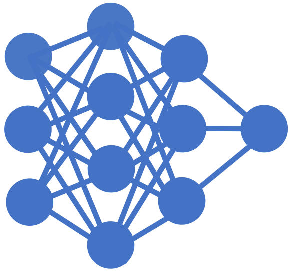
Crystal structure and materials properties

- Structure-property relationships:



Artificial Intelligence for Materials Studies





- The rise of the materials databases
  - Data are accessible
- Chemical space descriptors exist
  - Coulomb Kernel[1], Bag of bonds representation[2]
- Datascience tools exist
  - Scikit learn, Google's TensorFlow

[1] M. Rupp, et al., Phys Rev Lett. 108, 058301 (2012) [2] K. Hansen et al., J. Phys. Chem. Lett. 6, 2326–2331 (2015)

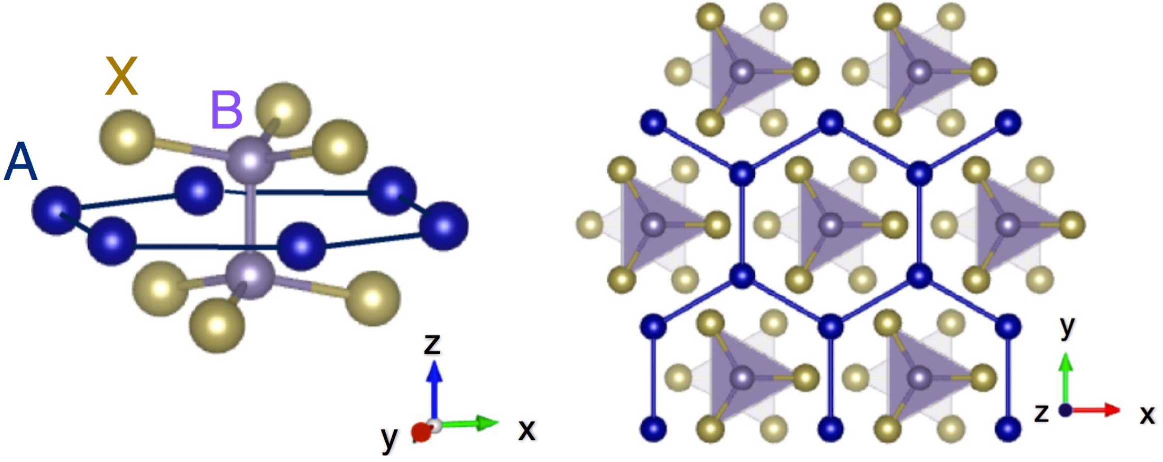
Artificial Intelligence for Materials Studies

Magnetic 2D crystals

Artificial Intelligence for Materials Studies

Magnetic 2D crystals

Transition metal trichalcogenides are magnetic 2D atomic crystals



$A_2B_2X_6$ crystal structure

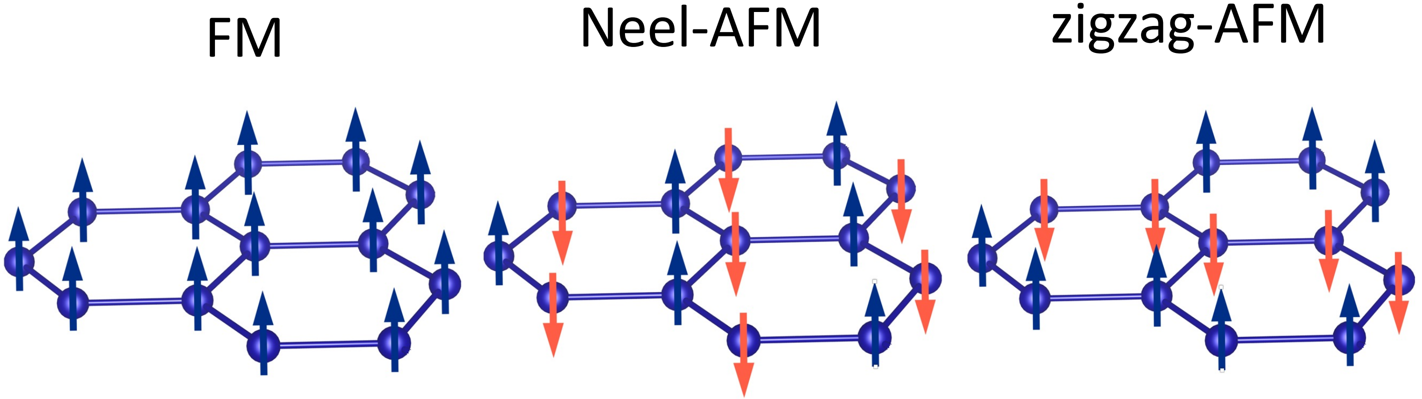
- $CrGeTe_3$ is a ferromagnet (FM)[1,2]
- Monolayer $CrSiTe_3$ is a zigzag antiferromagnet (zigzag-AFM)[1]

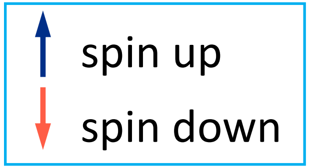
1. Sivadas et al., PRB 91 235425 (2015) 2. C. Gong et al., *Nature* **546**, 265 (2017)

Artificial Intelligence for Materials Studies

Magnetic 2D crystals

Transition metal trichalcogenides are magnetic 2D atomic crystals





Magnetic structures

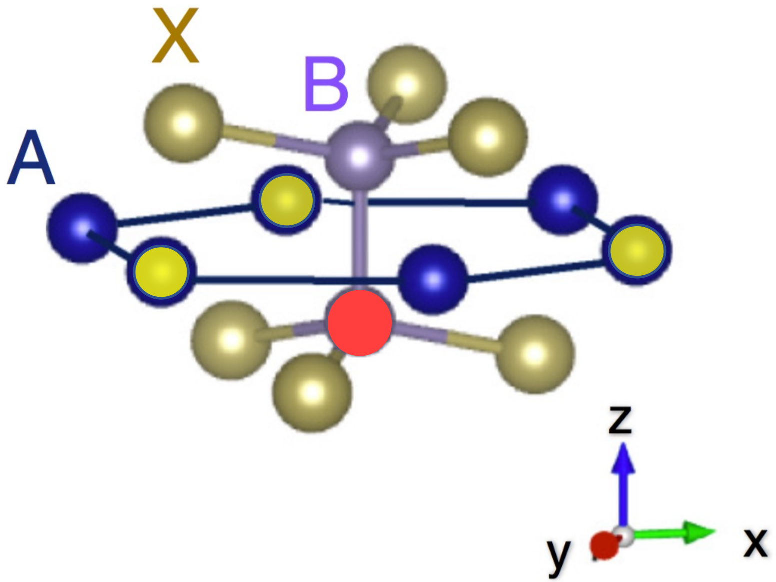
- $CrGeTe_3$ is a ferromagnet (FM)[1,2]
- Monolayer $CrSiTe_3$ is a zigzag antiferromagnet (zigzag-AFM)[1]

1. Sivadas et al., PRB 91 235425 (2015)   2. C. Gong et al., *Nature* **546**, 265 (2017)

Magnetic van der Waals materials

$A_2B_2X_6$ structures

- Create 198 ABX_3 structures using DFT
  - Total # of structures $\sim 10^4$
  - NM, FM and AFM spin configurations
- Extract data:
  - Formation energy
  - Magnetic order
  - Magnetic moment

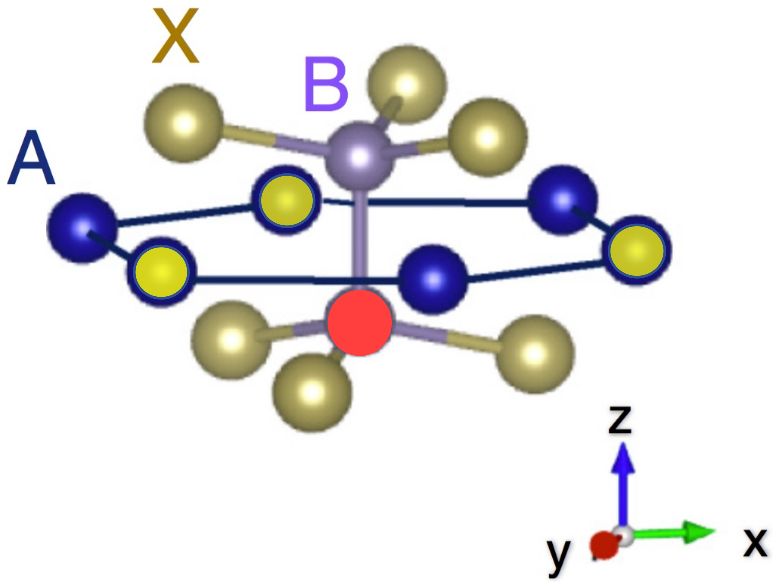


Magnetic van der Waals materials

$A_2B_2X_6$ structures

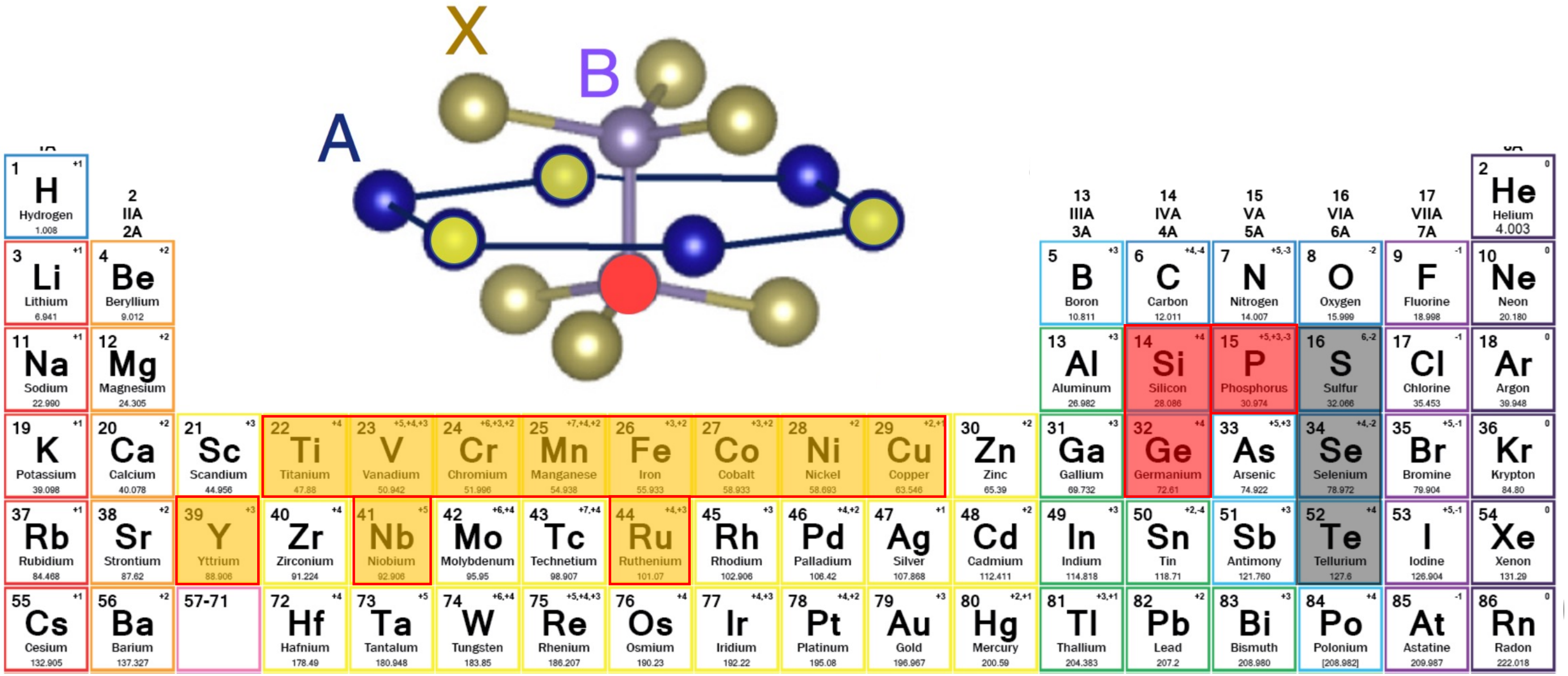
Substitutions:

- A site:
  - $Cr_{0.5}A_{0.5}$
  - A: Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Y, Nb, Ru
- B site:
  - Si, Ge, P combinations
  - B: Si, Ge, P, $Si_{0.5}Ge_{0.5}$, $Si_{0.5}P_{0.5}$, $Ge_{0.5}P_{0.5}$
- X site:
  - S, Se, Te

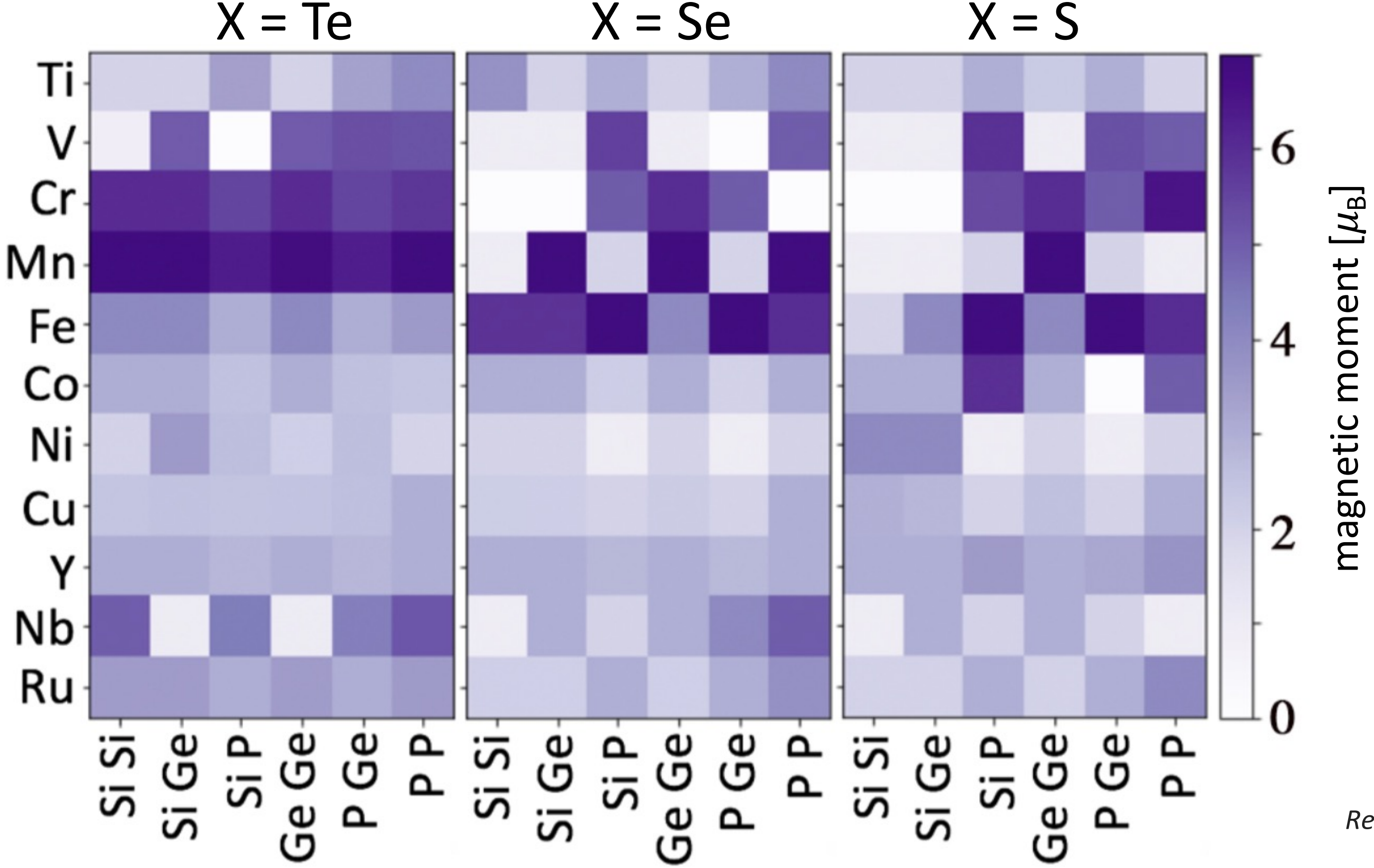


Magnetic van der Waals materials

$A_2B_2X_6$ *structures*



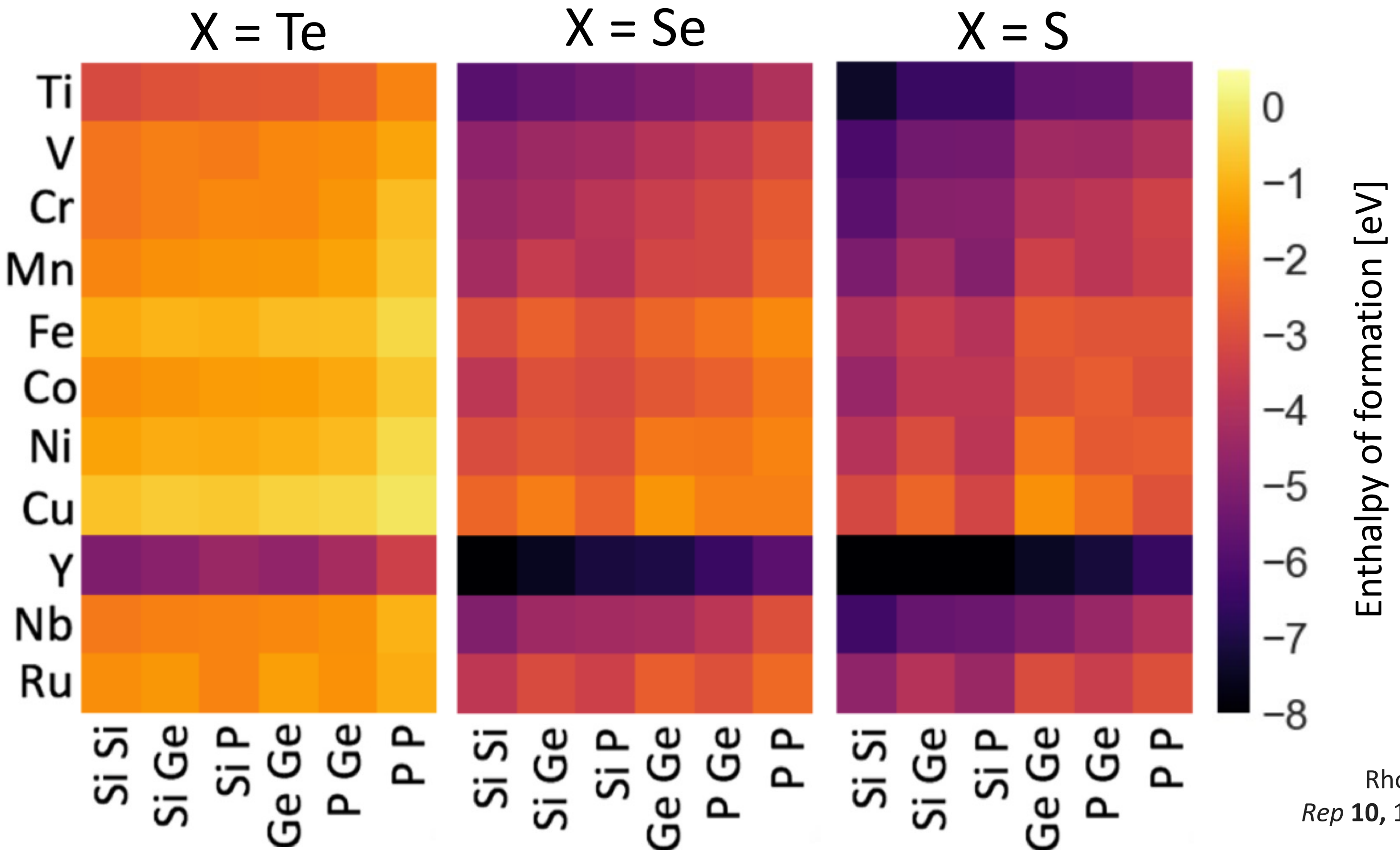
Magnetic moment of $A_2B_2X_6$



Rhone, *et al., Sci Rep* **10,** 15795 (2020).

Formation energy of $A_2B_2X_6$

Formation energy of $A_2B_2X_6$



Rhone, *et al.*, *Sci Rep* **10,** 15795 (2020).

Artificial intelligence in materials science

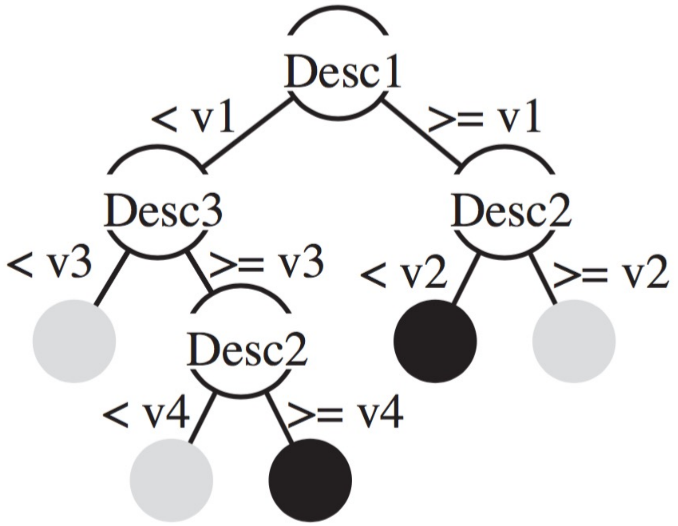
Materials descriptors

- Describe system using easily attainable components
  - Atomic properties, p
- Compound Property, P
  - P = mean(p(A), p(B), p(X))
  - P = std. deviation(p(A), p(B), p(X))
- Total # of descriptors:
  - 61

- Atomic property, p:
  - Number of spin up e's
  - atomic radius
  - etc.

Machine learning models

- o Random forest regression
  - Machine learning model
- o Inputs: X
  - Number of spin up e's
- o Output: Y
  - Magnetic moment
- o Training data
- o Test data



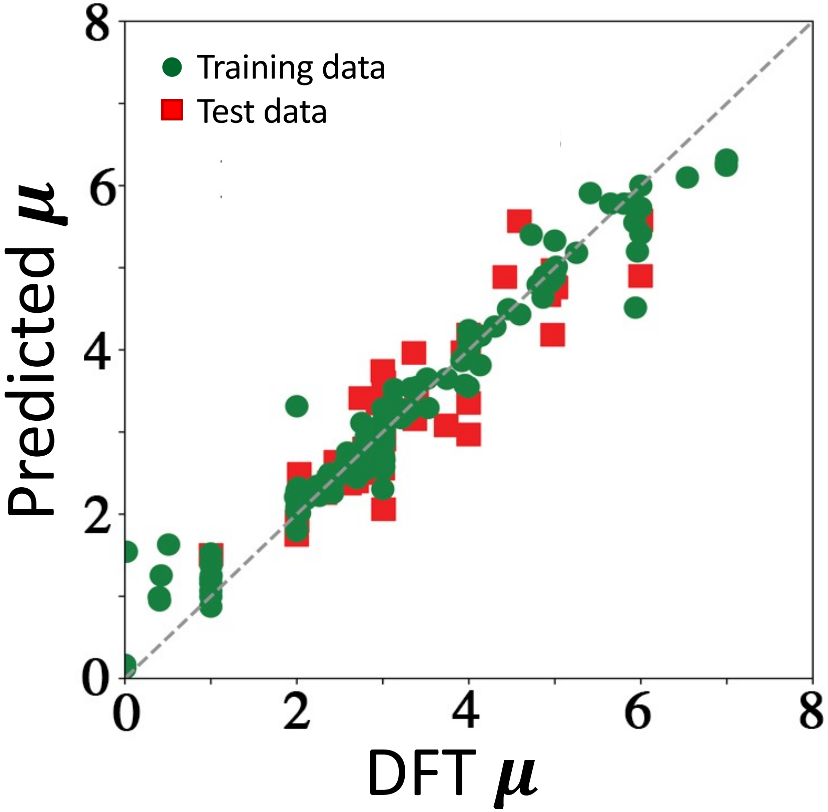
- Node m, region R_m, N_m observations
- Mean Absolute Error
  - minimize L1 error using median values at terminal nodes

$$L(X_m) = \frac{1}{N_m} \sum_{i \in N_m} |y_i - \bar{y_m}|$$

$$\bar{y_m} = \frac{1}{N_m} \sum_{i \in N_m} y_i$$

1. G.A. Landrum, H. Genin, Journal of Solid State Chemistry 176 (2003) 587–593

Machine learning predictions
Magnetic moment, X=Te



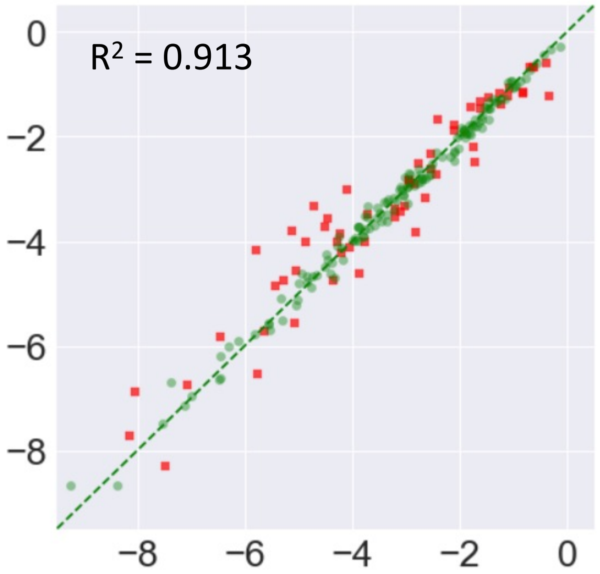
- N = 262
- Random forest $R^2 = 0.98$
- Mean absolute error (MAE) = 0.30 μ_B

DFT: first-principles quantum calculations
μ: magnetic moment ~ magnetization

Rhone, *et al., Sci Rep* **10,** 15795 (2020).

Machine learning predictions of DFT formation energy

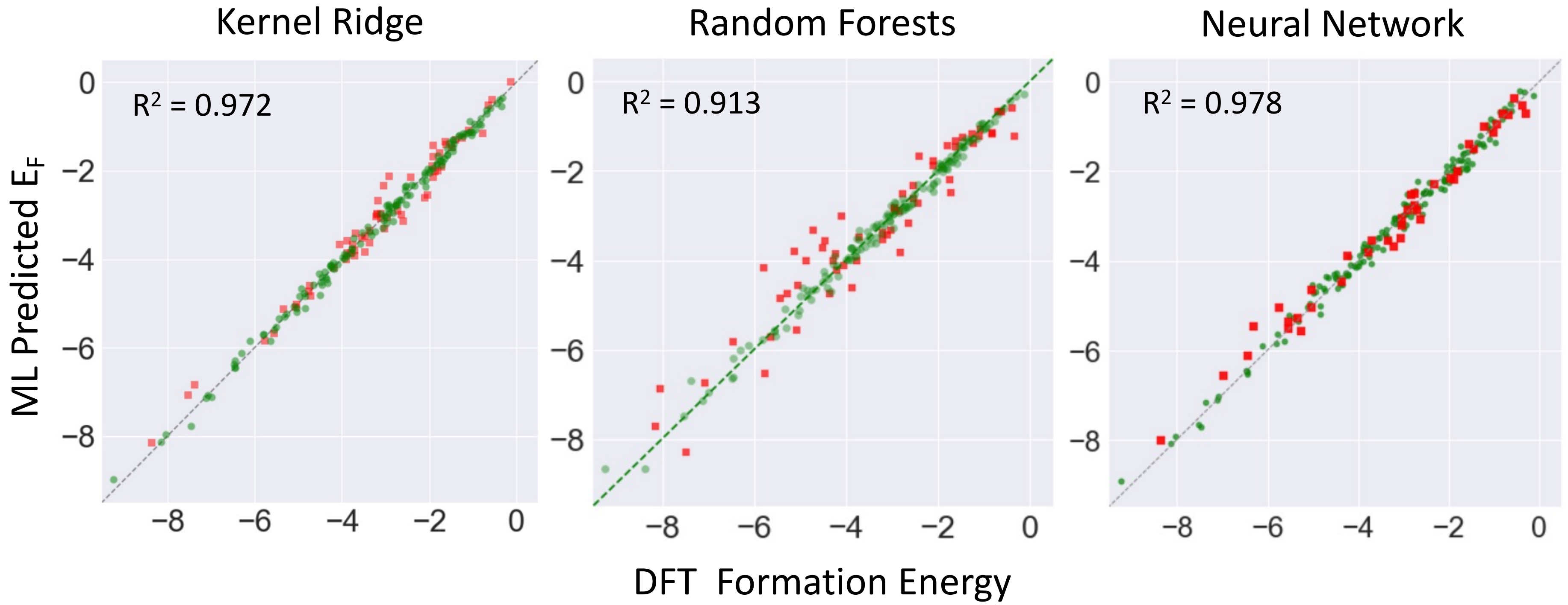
Random Forests



DFT Formation Energy

Rhone, *et al., Sci Rep* **10,** 15795 (2020).

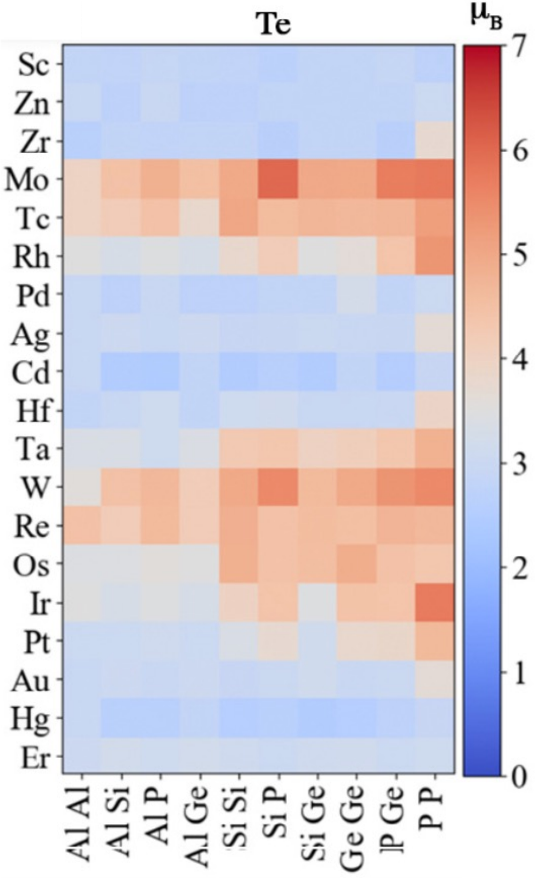
Machine learning predictions of DFT formation energy



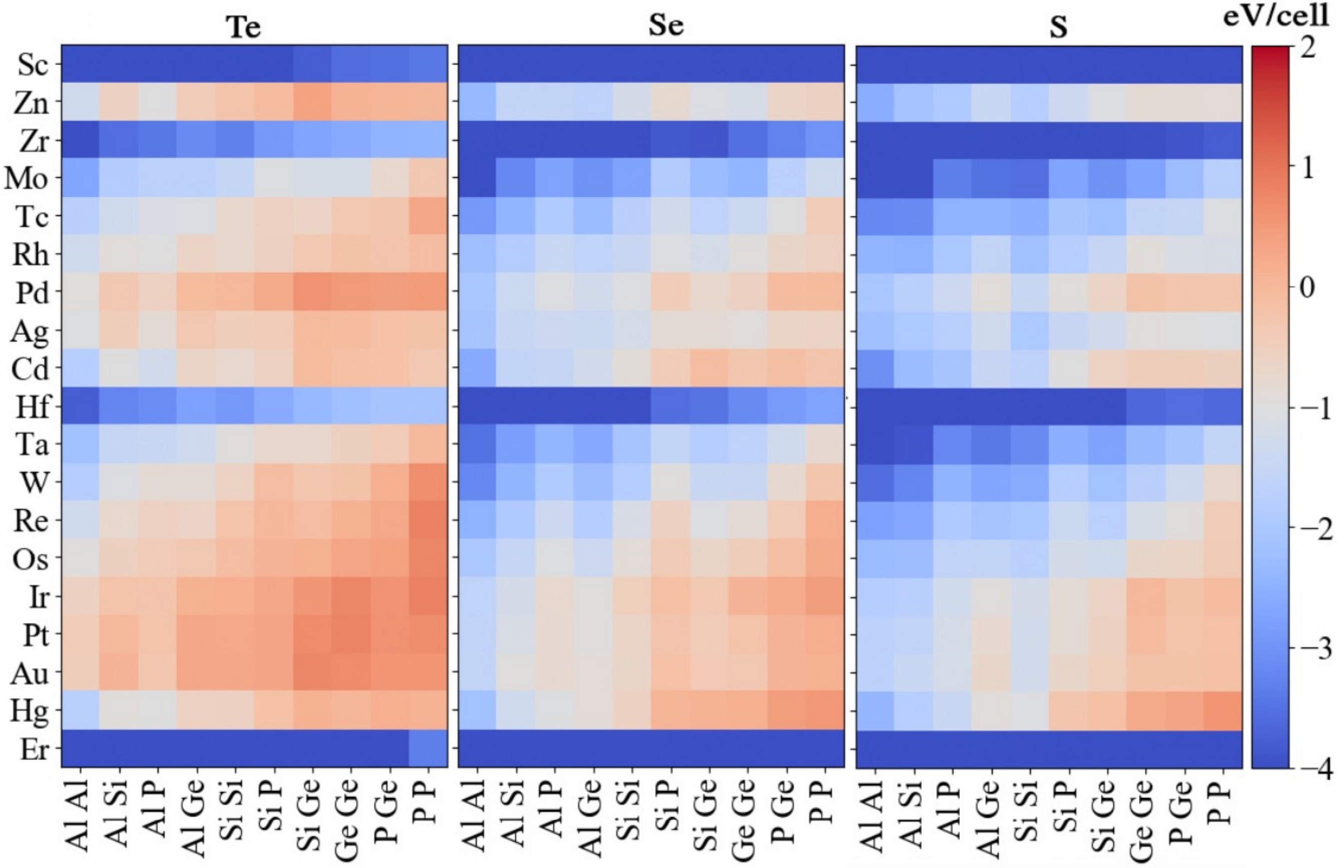
Rhone, *et al., Sci Rep* **10,** 15795 (2020).

Machine learning results

Magnetic moment

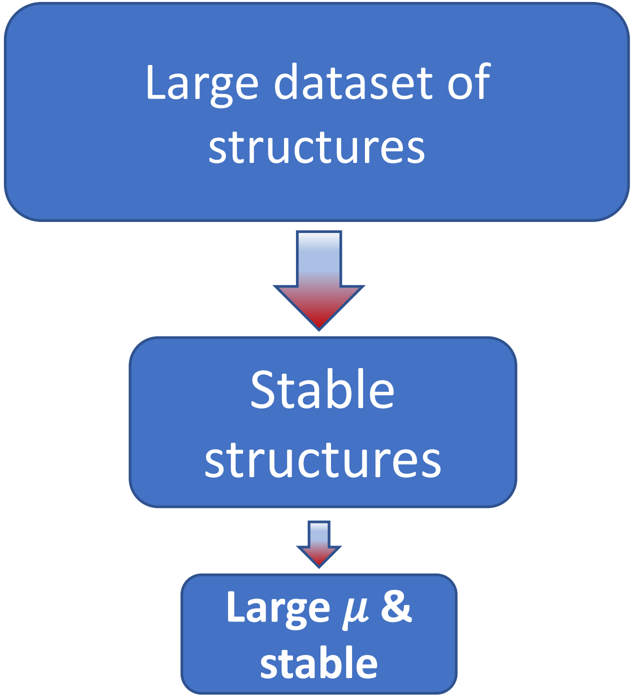


Formation Energy



Candidates for chemically stable 2D ferromagnets

- Initial DFT data (198 structures)
- ML predictions (> 4000 structures)



Screening Criteria:

- Magnetic moment > 5 μ_B & Formation energy < -1.1eV

DFT results:

- $(CrMn)Si_2Te_6$
- $(CrMn)Ge_2Se_6$
- $(CrFe)(SiP)S_6$

ML Results:

- $(CrMo)Si_2Te_6$
- $(CrW)Si_2Te_6$
- $(CrMo)(SiP)Te_6$

Rhone, *et al.*, *Sci Rep* **10,** 15795 (2020).

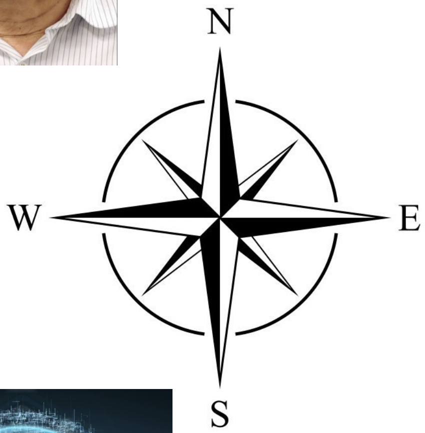
AI-guided materials discovery
Outcomes

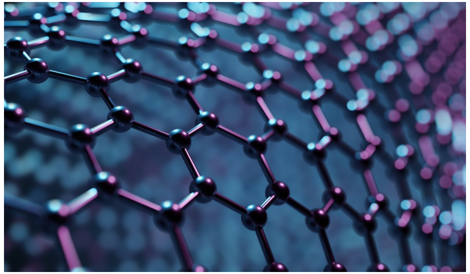
- Identified new chemically stable 2D ferromagnets
- AI provides fast estimates of 2D materials' properties
  - > 4000 $A_2B_2X_6$ composites
  - Magnetic properties
  - Chemical stability
- Non-traditional framework for studying emergent spin phenomena in 2D materials

Rhone, *et al., Sci Rep* **10,** 15795 (2020).

Summary

- Aron's invaluable guidance shaped the lives and careers of many
- A collection of people, under the right conditions, can give rise to 'emergent phenomena'
- AI-guided identification of emergent spin phenomena in materials





Acknowledgements

- This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562

- Some of the computations in this paper were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University

- This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357.

- This material is based upon work supported by the National Science Foundation CAREER award under Grant No 2044842


TANGOBA
Buenos Aires Ciudad
TANGOBA
Buenos Aires Ciudad
TANGOBA

Aron’s Early Influence – Postdoctoral studies

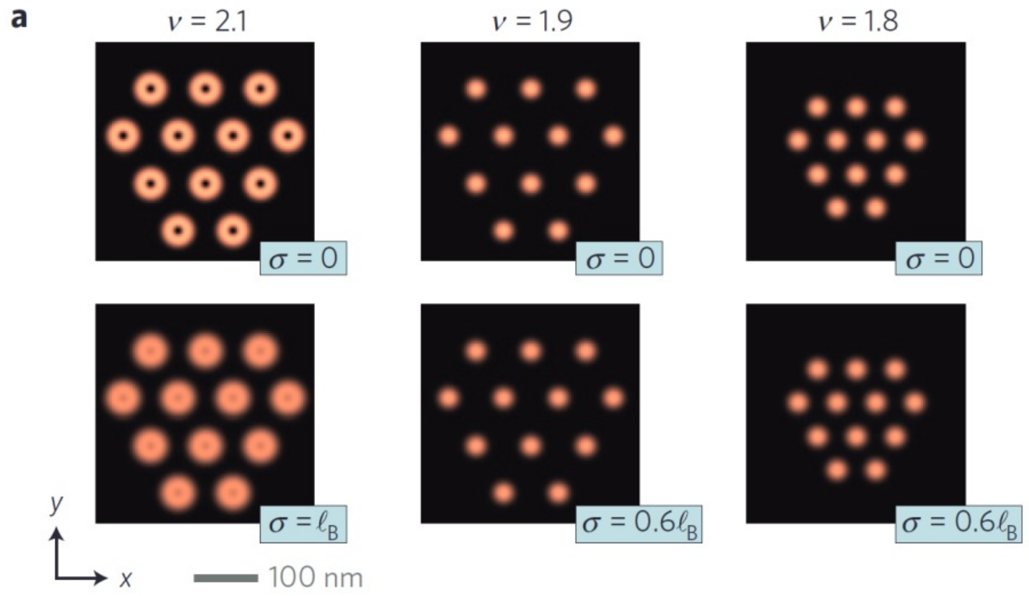
Postdoc at NTT, Japan

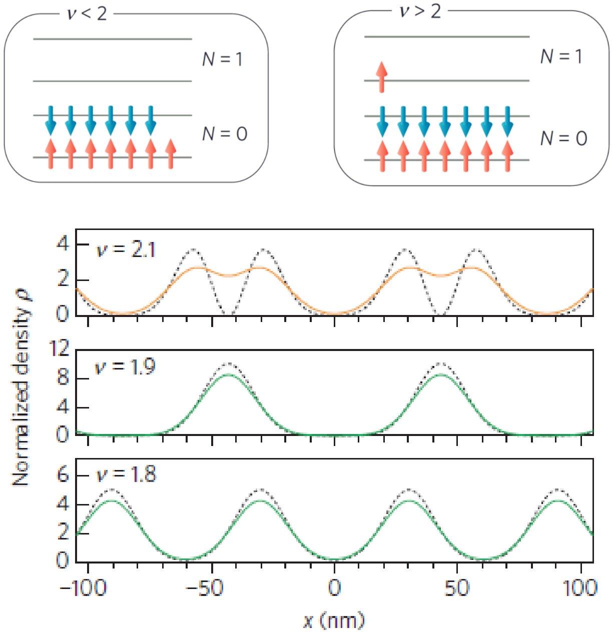
Charge properties of electrons in GaAs quantum wells

NMR probes Wigner Solids: *In-plane local density variations at $\nu \sim 2$*

nature physics
LETTERS
PUBLISHED ONLINE: 20 JULY 2014 | DOI: 10.1038/NPHYS3031

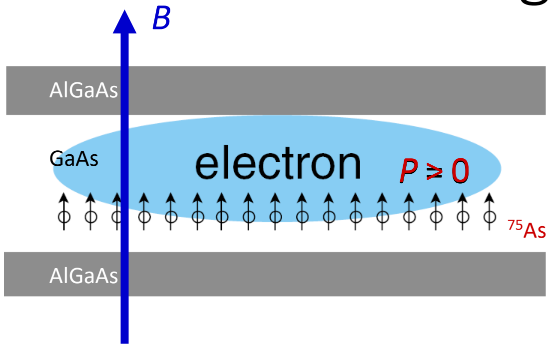
NMR profiling of quantum electron solids in high magnetic fields

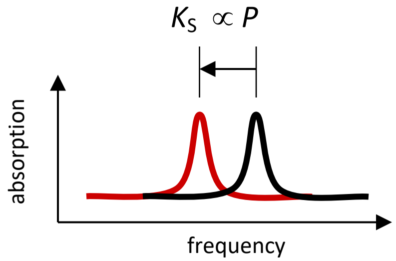


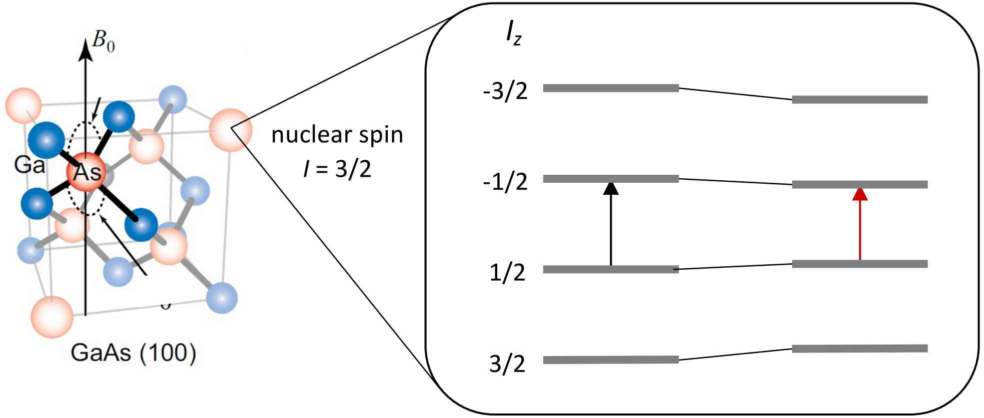


1) L. Tiemann(*), T.D. Rhone(*), N. Shibata, K. Muraki, Nature Physics, doi:10.1038/nphys3031
(*: equal contribution)

Nuclear Magnetic Resonance (NMR)
Knight Shift

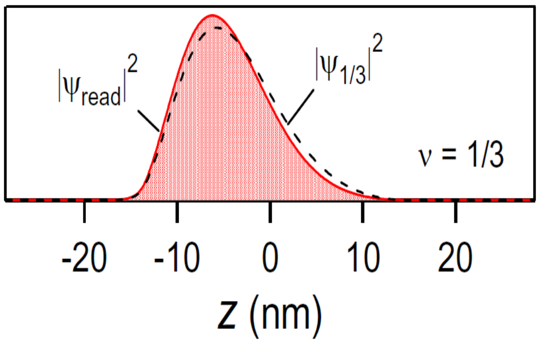




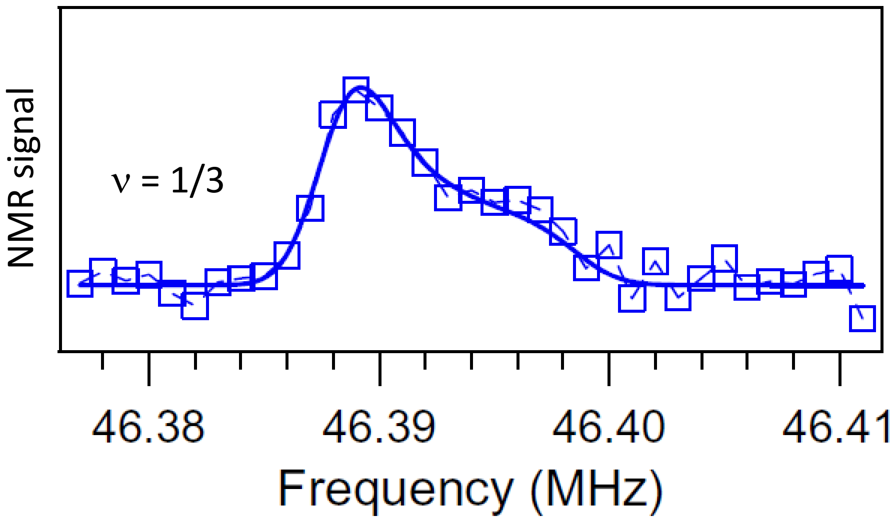


electron spin $P = 0$ $P > 0$

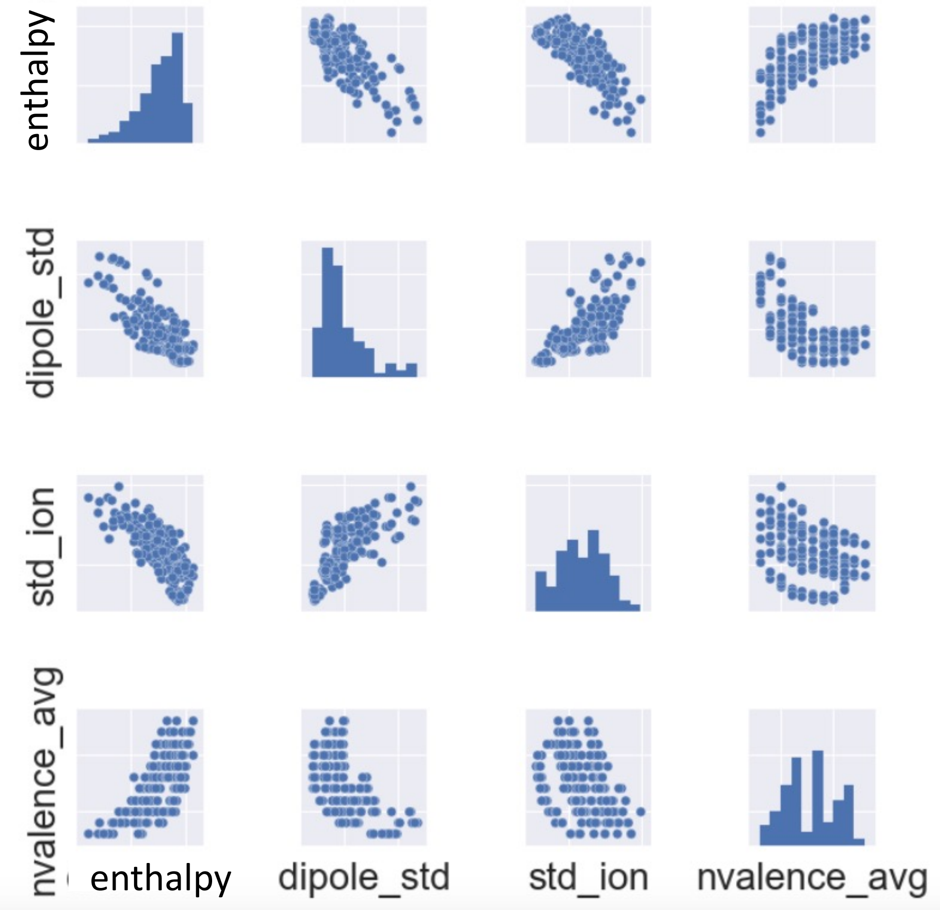
Simulating RD-NMR Spectra



$$I_\nu(f-f_0)=\int_{-w/2}^{w/2} g\big(f-f_0+\underbrace{\alpha_\nu|\psi_\nu(z)|^2}_{\text{(spin polarization) x (electron density)}}\big)|\psi_{read}(z)|^2\,dz$$

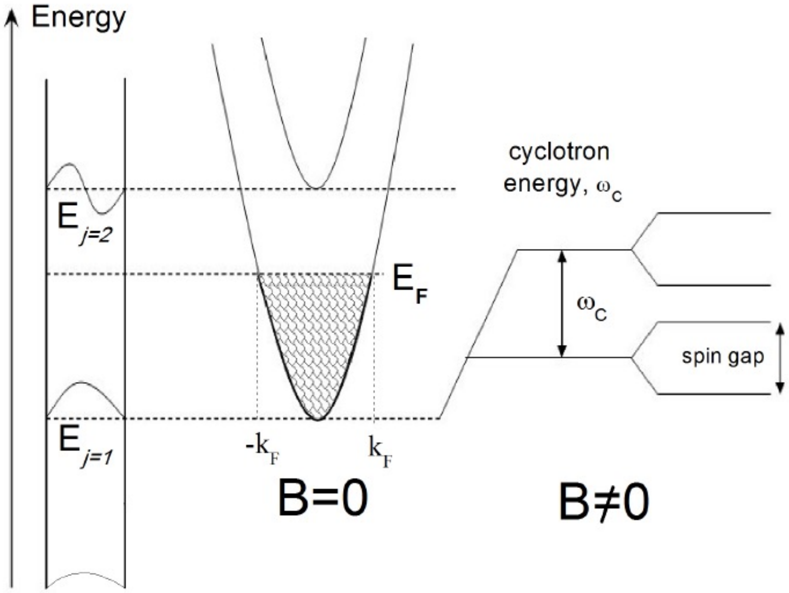


Materials descriptors: data visualization



- Compound Property, P
  - P = f(p(A), p(B), p(X)))
- p:
  - Enthalpy
  - Polarizability (dipole_std)
  - Ionization energy (std_ion)
  - # valence electrons (nvalence_avg)

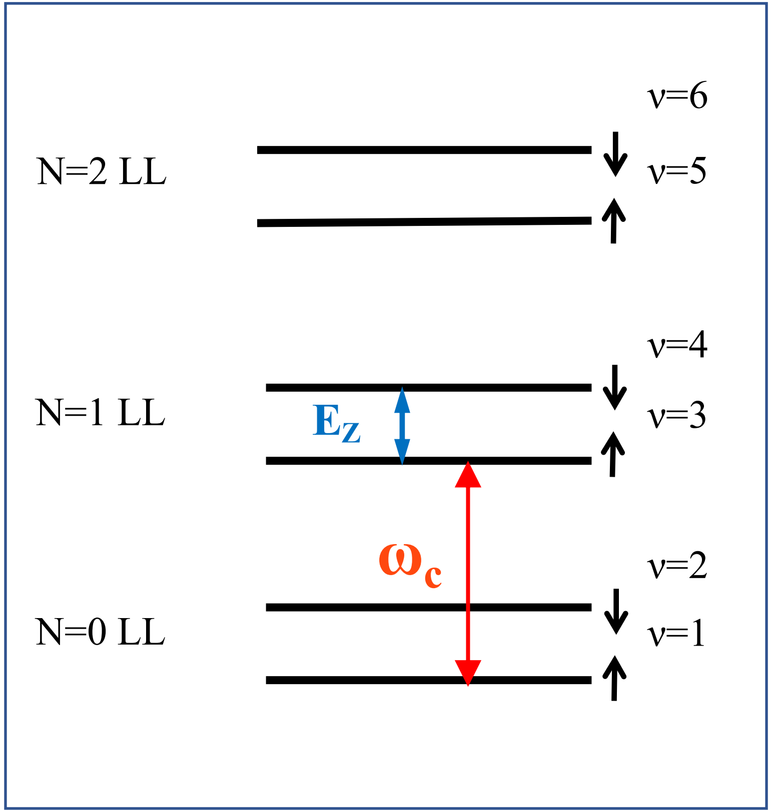
Quantum Well



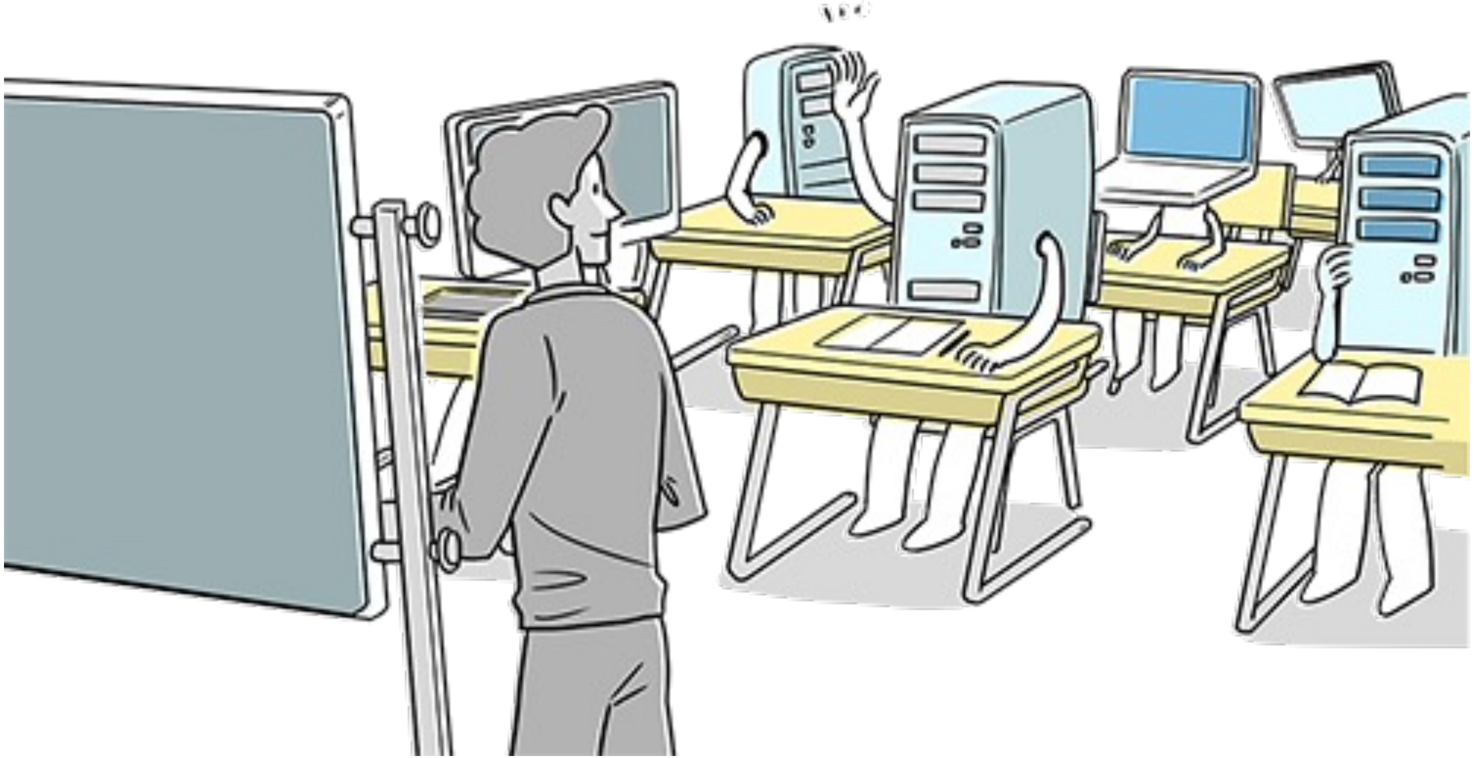
$\omega_c = eB/m^*c$

$\nu = n/(B/\phi_o)$

Spin gap is the Zeeman energy

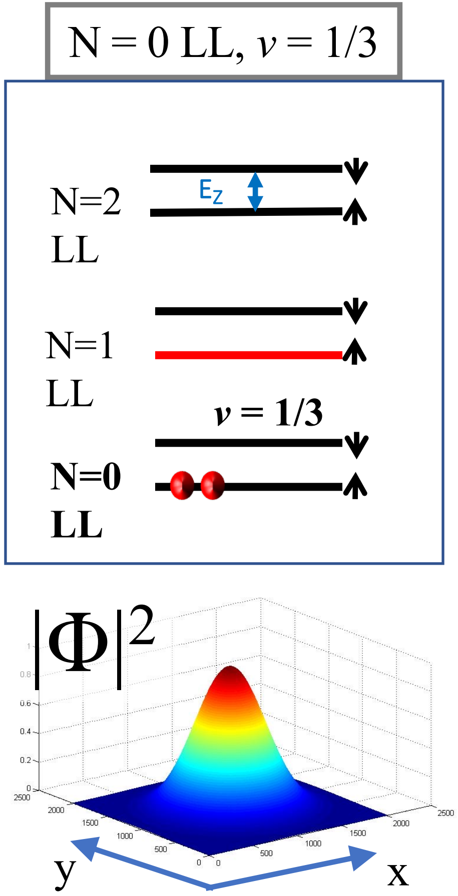


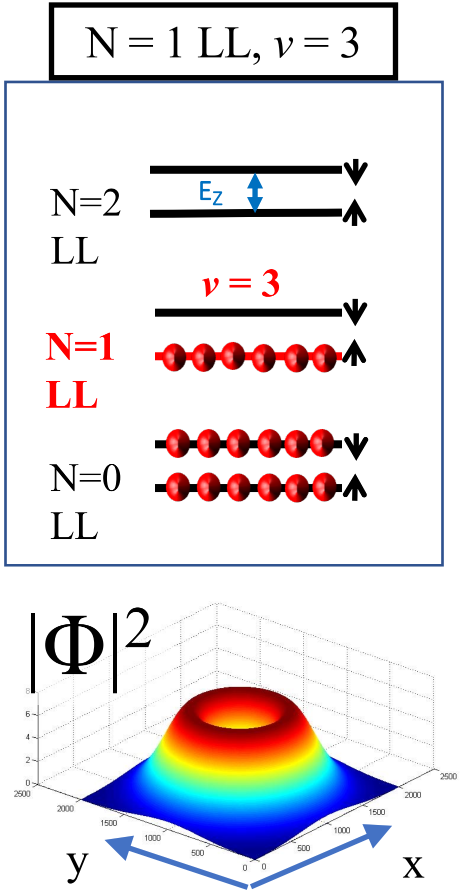
What is Materials Informatics?



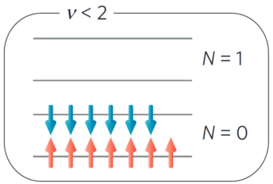
Teaching computers materials science using machine learning!

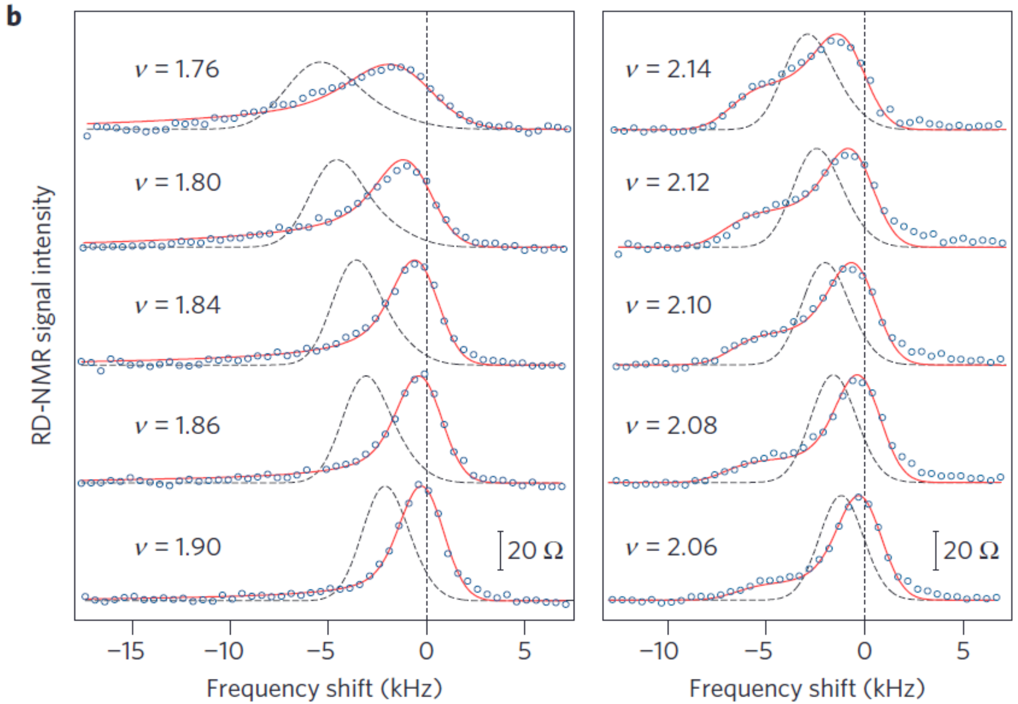
2DES in varying LL regimes

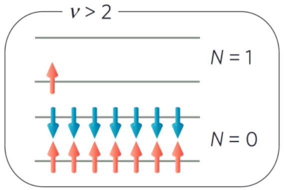




NMR probes Wigner Solids: *In-plane local density variations at ν ~ 2*



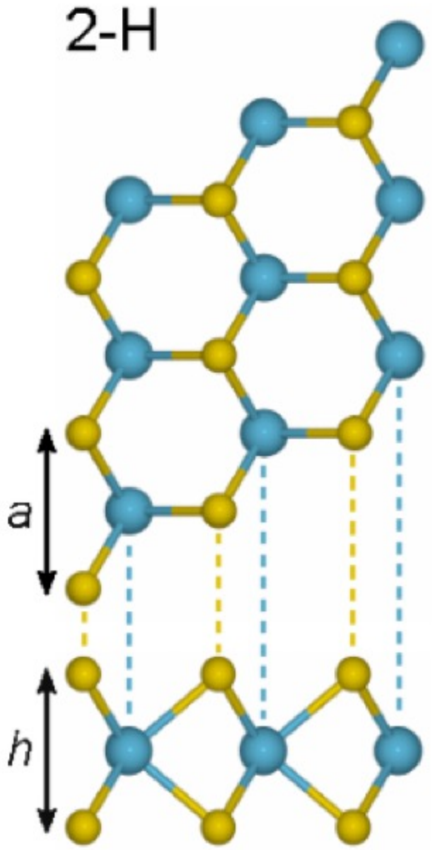


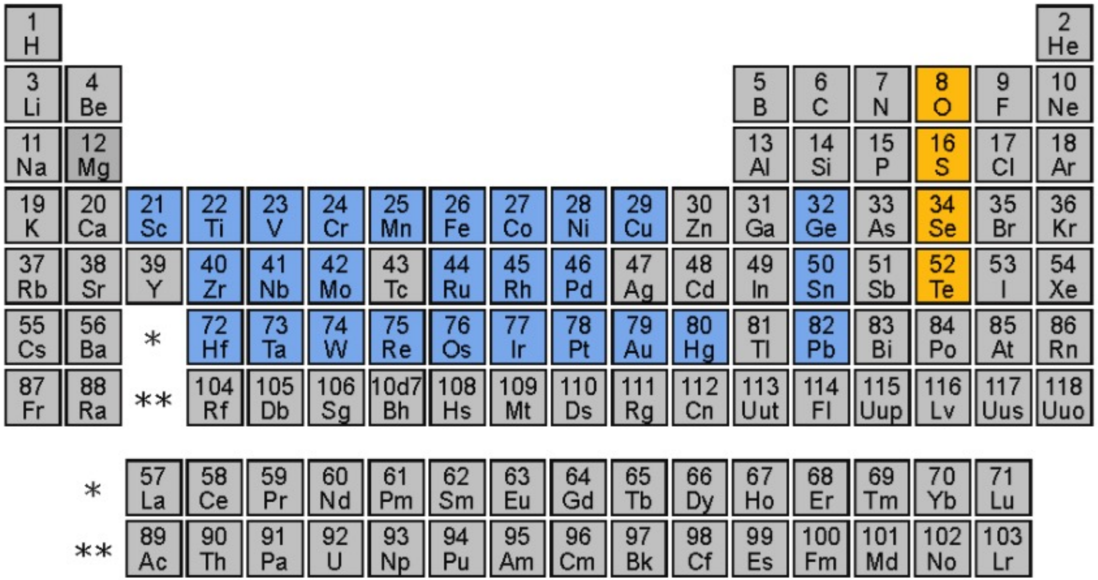


1) L. Tiemann(*), T.D. Rhone(*), N. Shibata, K. Muraki, Nature Physics, doi:10.1038/nphys3031
(*: equal contribution)

Dataset: Transition Metal Dichalcogenides (TMDs)

MX_2

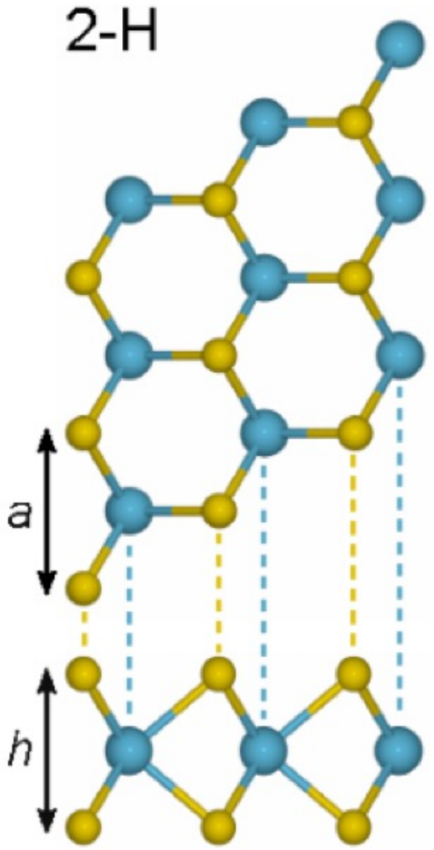


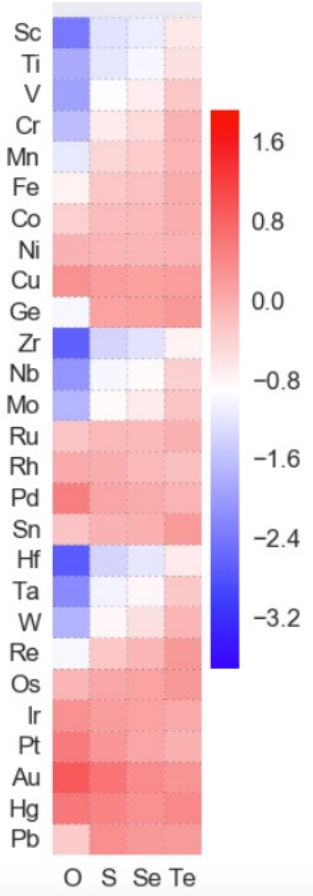


Rasmussen et. al., J. Phys. Chem. C 2015, 119, 13169–13183

Dataset: Transition Metal Dichalcogenides (TMDs)

MX_2

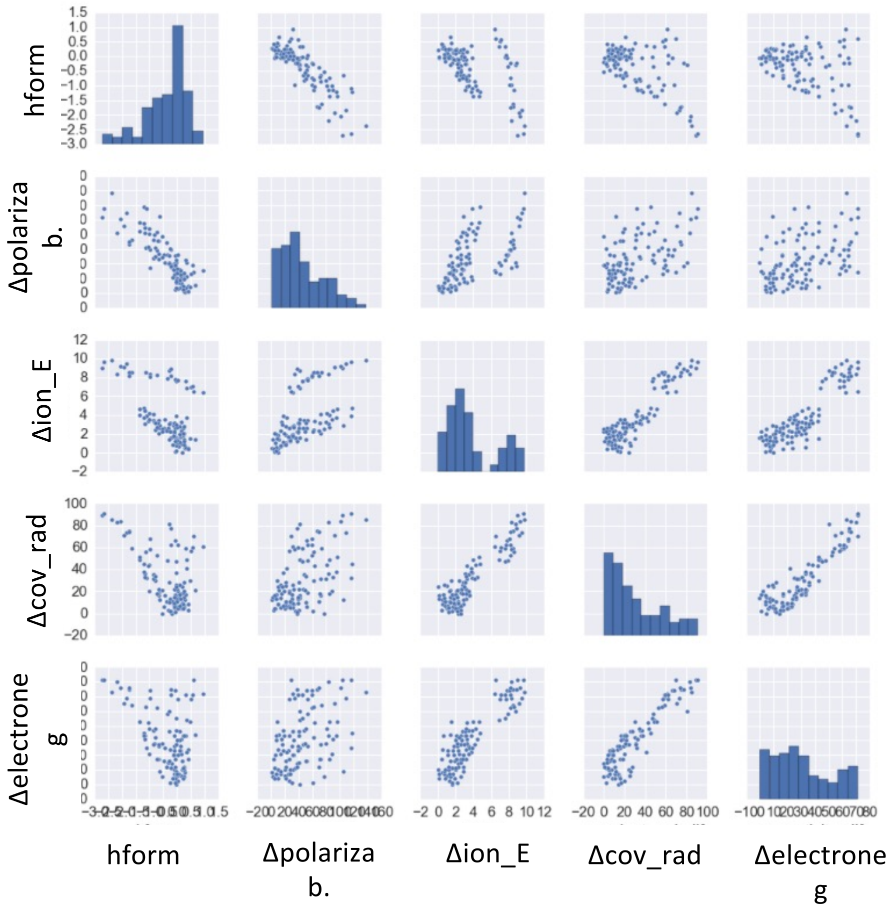




Heat of Formation of TMDs

Rasmussen et. al., J. Phys. Chem. C 2015, 119, 13169−13183

Data visualization: TMDs



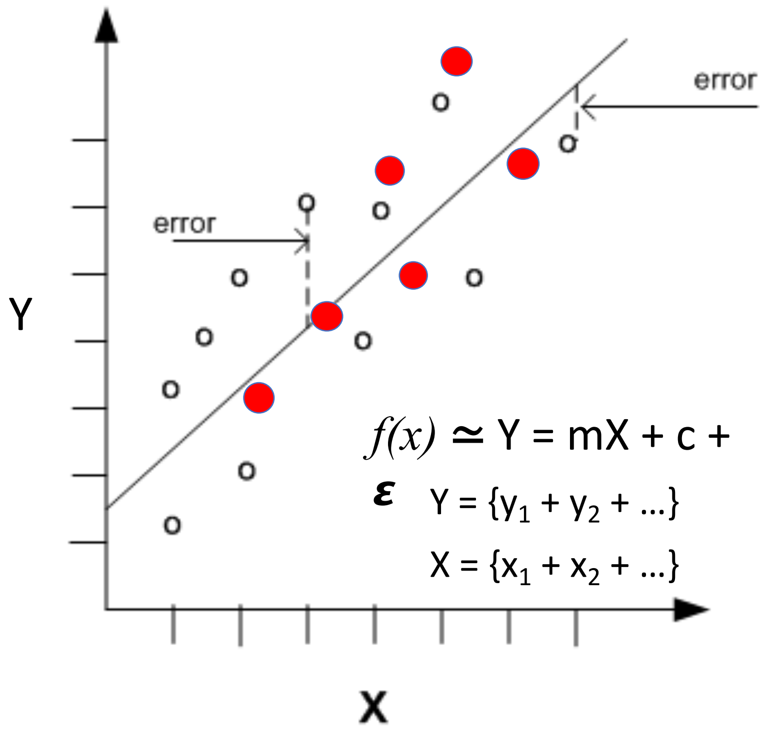
MX_2 materials ‘features’

- Atomic properties:
  - $\Delta x = x(M) - x(X)$
- x:
  - Polarizability
  - Ionization energy
  - Covalent radius
  - electronegativity

Machine Learning models

- Linear regression is a type of machine learning model
- Inputs: X
  - Dipole polarizability of A, B
- Output: Y
  - Heat of formation
- Training data
- Test data

- Algorithm estimates m & c
  - Reduce test error
  - $\min\{\sum_i (y_i - \hat{y_i})^2\}$



$f(x) \simeq Y = mX + c + \boldsymbol{\varepsilon}$

$Y = \{y_1 + y_2 + \ldots\}$

$X = \{x_1 + x_2 + \ldots\}$

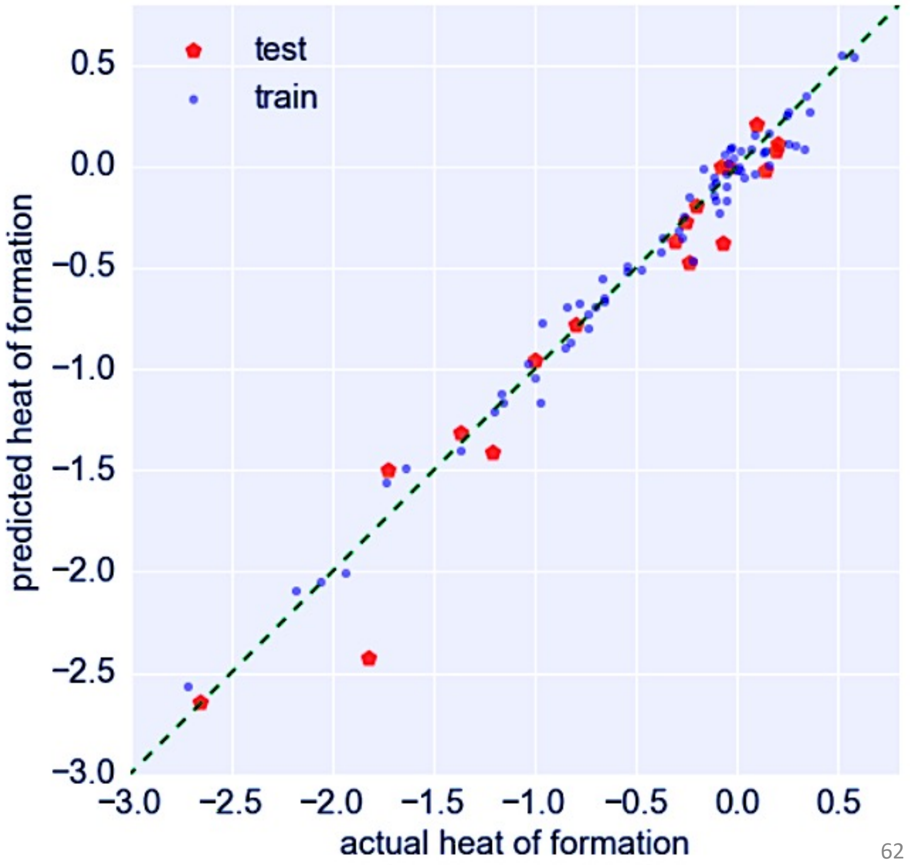
Machine learning TMD data to predict stability

Machine learning models

$$y = f(x_1,\ x_2,\ \ldots,\ x_N)$$

Heat of formation

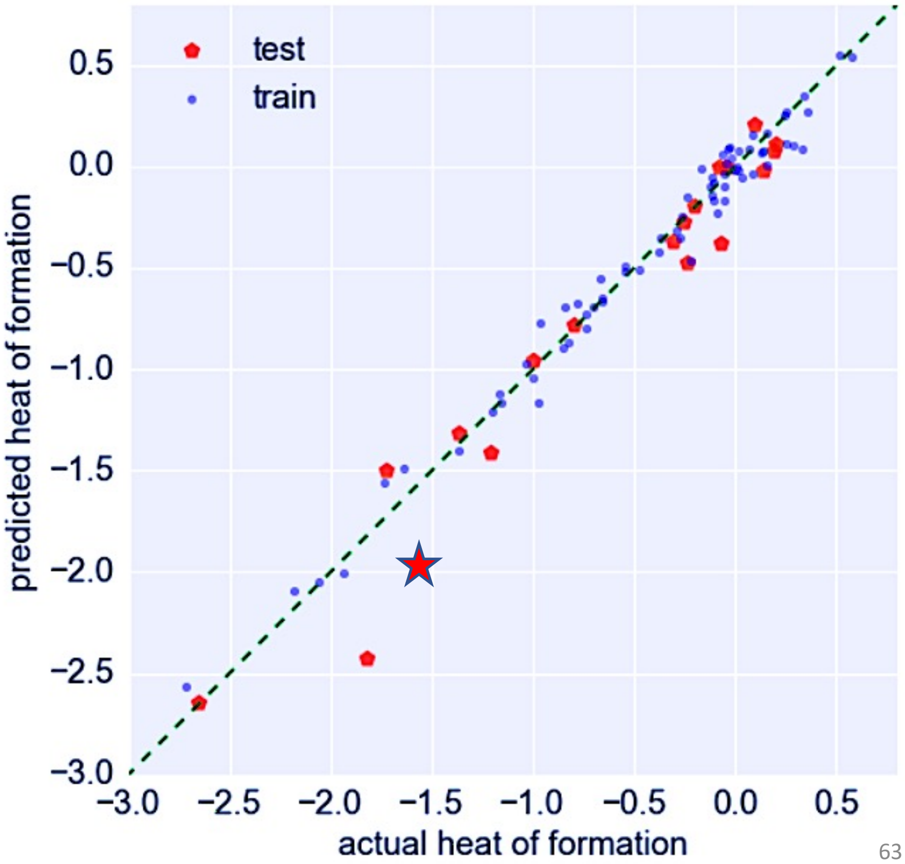
dipole polarizability difference
valence electrons difference
…
Electronegativity difference



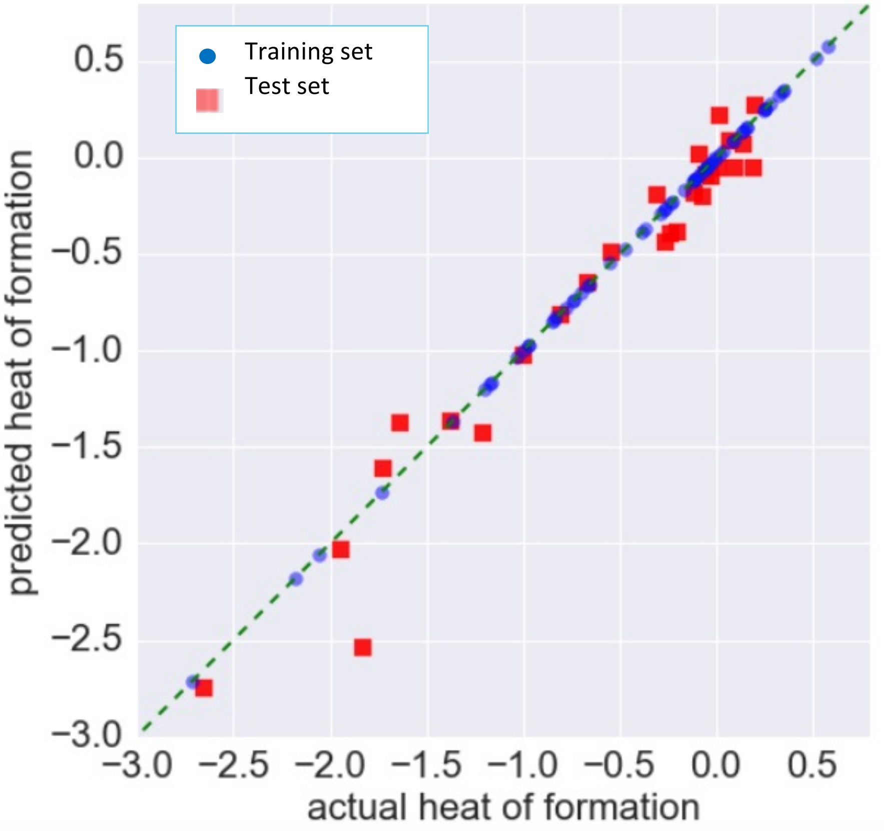
Machine learning TMD data to predict stability

Heat of formation of
Boron Nitride

- Predicted value: *-2eV*
- Expected value: *-1.56eV*



Machine Learning models: Kernel ridge regression

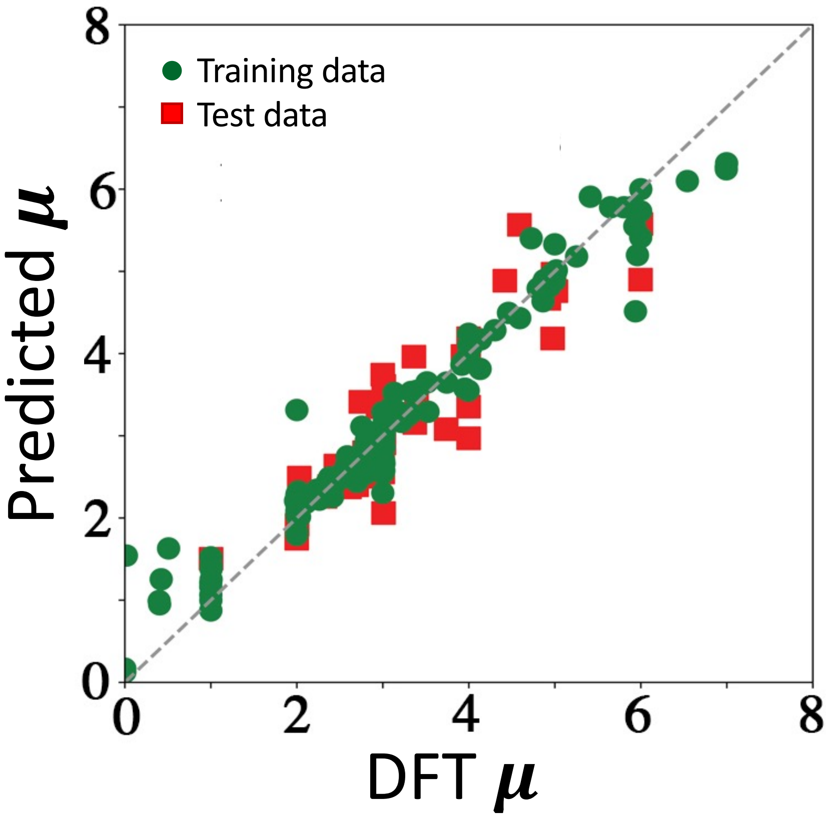


- # data points = 89
- # features = 30

$$C = \frac{1}{2}\sum_i (y_i - \mathbf{w}^T \mathbf{x}_i)^2 + \frac{1}{2}\lambda ||\mathbf{w}||^2$$

Test score , R^2 = 0.94
Train score, R^2 = 0.99

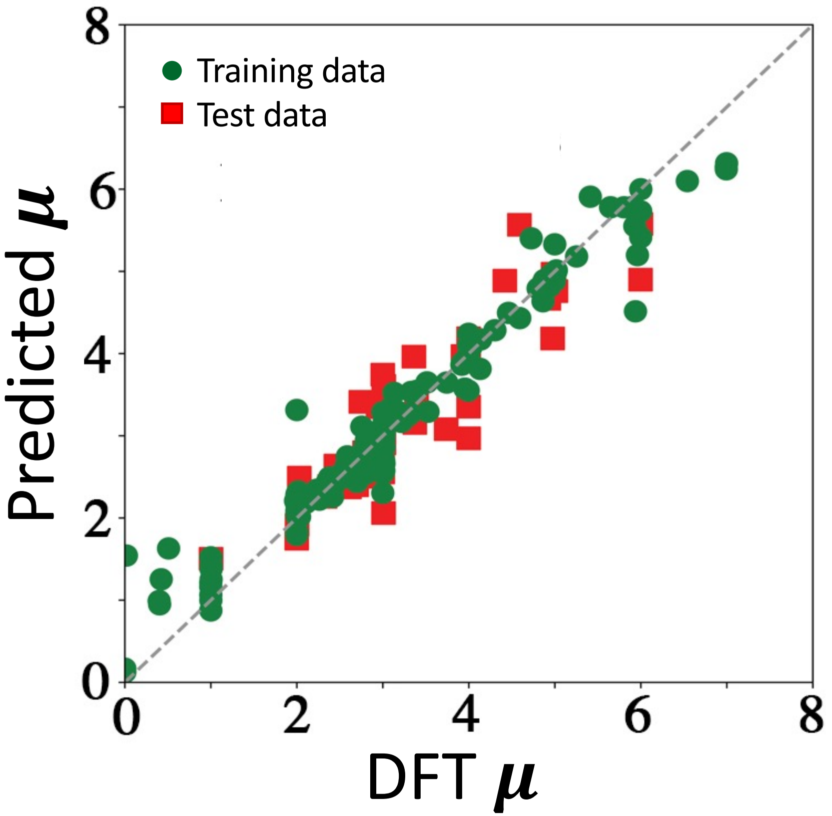
Machine learning predictions
Magnetic moment, X=Te

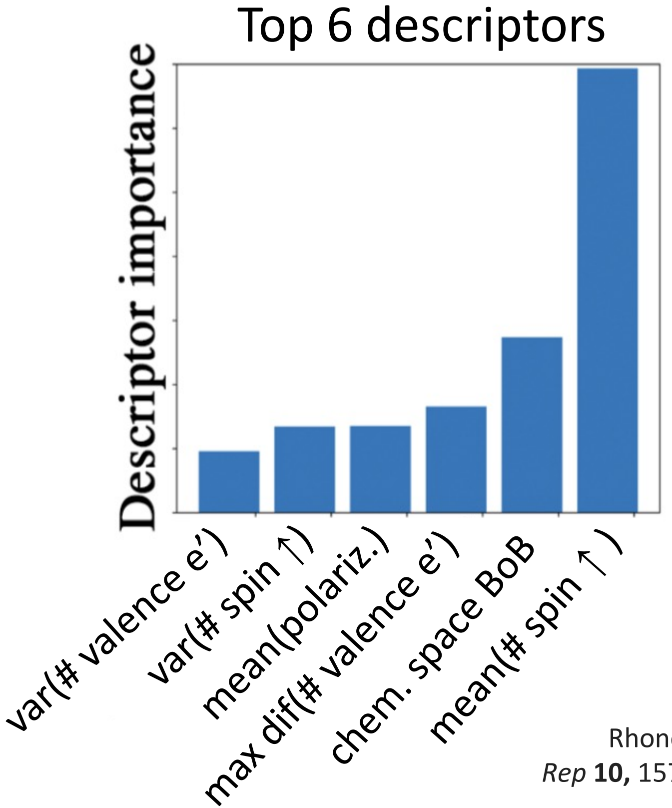


- N = 262
- Random forest R^2 = 0.98
- Mean absolute error (MAE) = 0.30 μ_B

DFT: first-principles quantum calculations
μ: magnetic moment ~ magnetization

Machine learning predictions
Magnetic moment, X=Te

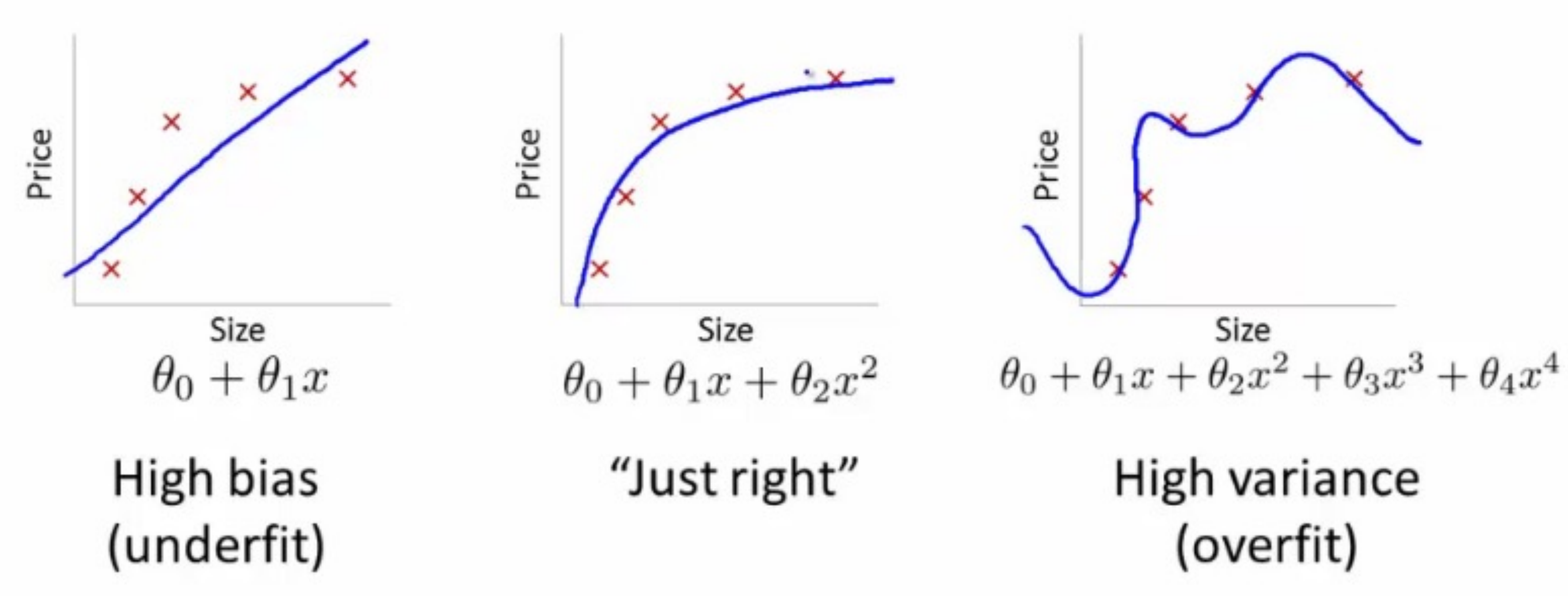




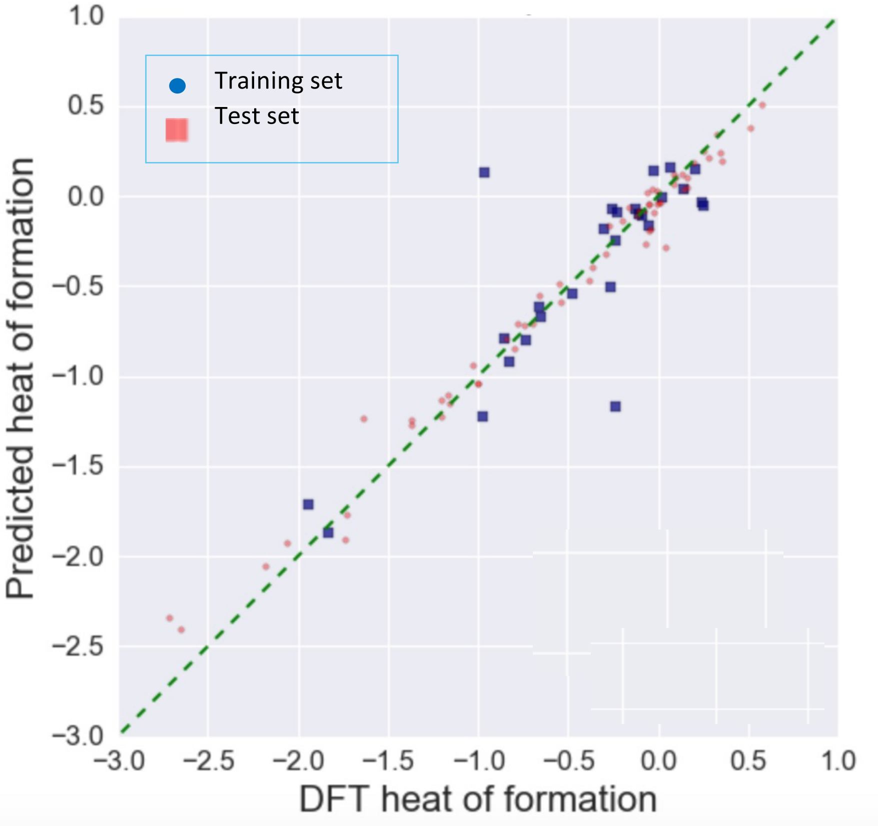
Rhone, *et al.*, *Sci Rep* **10,** 15795 (2020).

Machine Learning models: Kernel ridge regression

Overfitting



Machine Learning models: Random forest regression



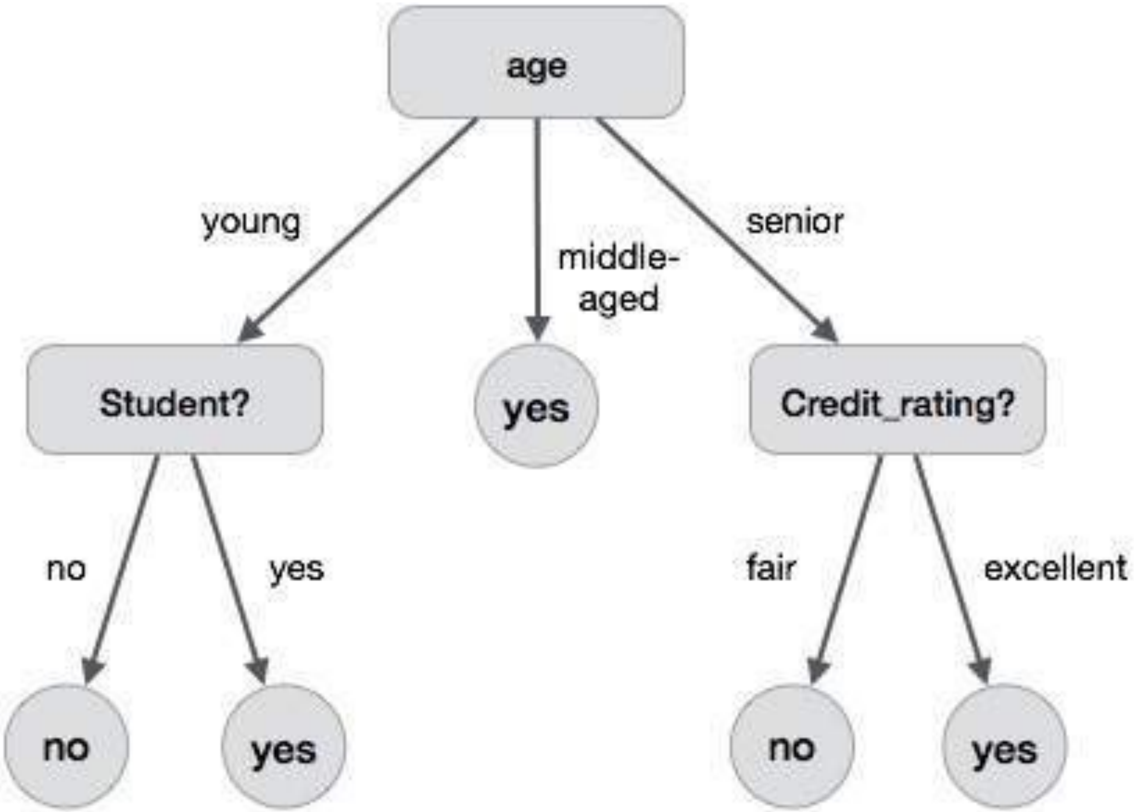
Random forest regression is based on decision trees

- # data points = 89
- # features = 39

Test score, MSE = 0.095

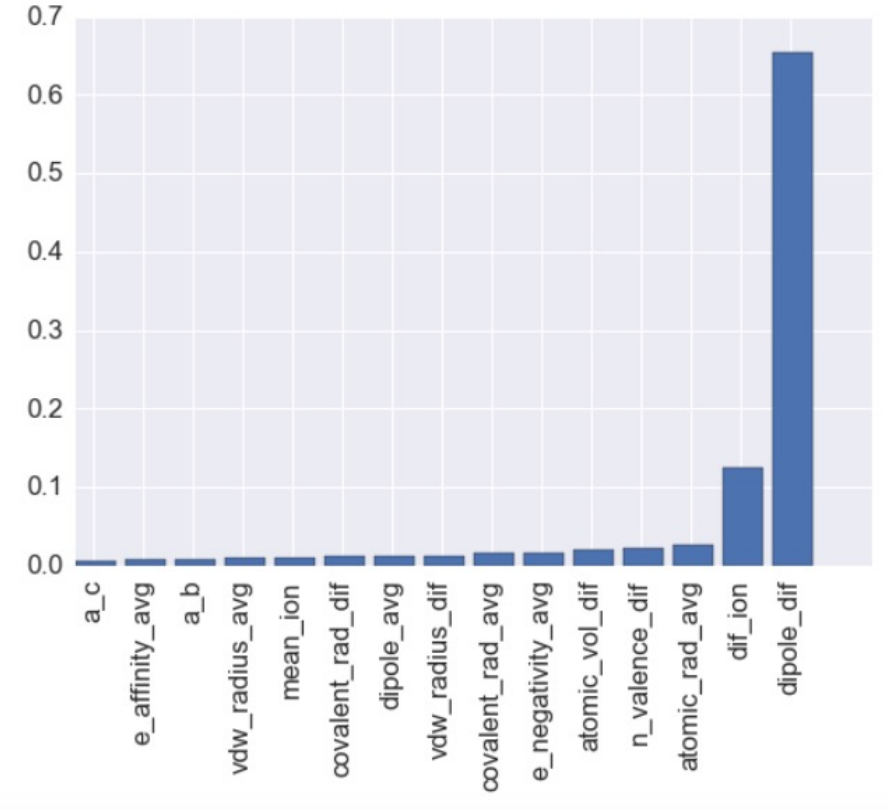
Random Forest model and Decision Trees

Is a customer at a company likely to buy a computer or not?



Machine Learning models: Random forest regression

Feature Importances



Top ten descriptors

1	Dipole dif
2	Dif in ionization E
3	atomic radius avg
.	num valence dif
.	atomic vol dif
.	e_negativity avg
	covalent rad avg
	vdw radius dif
	dipole avg
10	covalent rad dif

Machine Learning models

Challenges

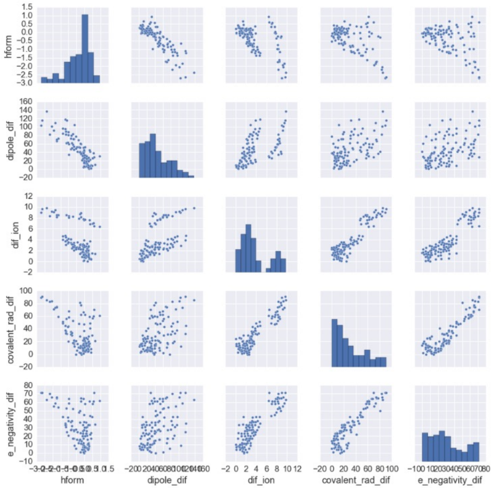
1. Can our ML models generalize to other 2D materials?
2. Are our descriptors sufficient?
3. Develop descriptors to describe crystal structure?

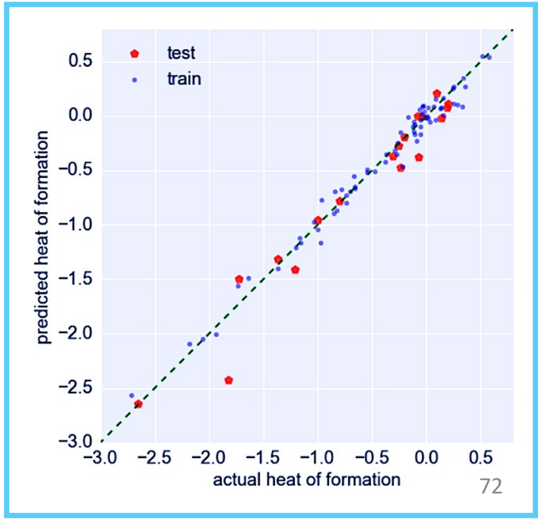
Top ten descriptors

1	Dipole dif
2	Dif in ionization E
3	atomic radius avg
.	num valence dif
.	atomic vol dif
.	e_negativity avg
	covalent rad avg
	vdw radius dif
	dipole avg
10	covalent rad dif

Summary

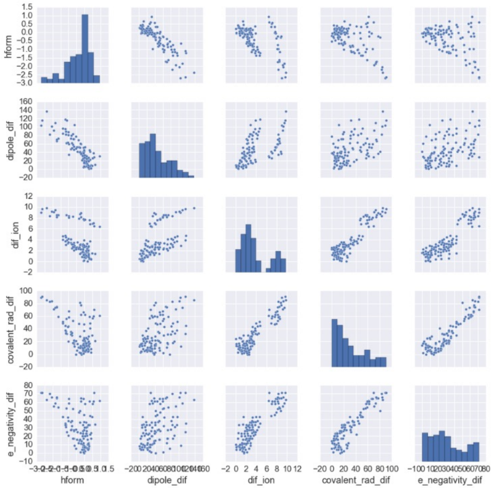
- Materials informatics
  - Identified and exploited trends in materials' properties data for vdW materials
  - Predicted heat of formation of TMDs
- Materials databases + access to machine learning tools makes materials informatics possible

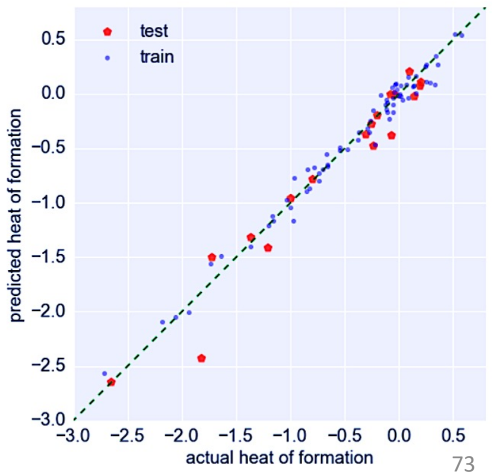




Outlook

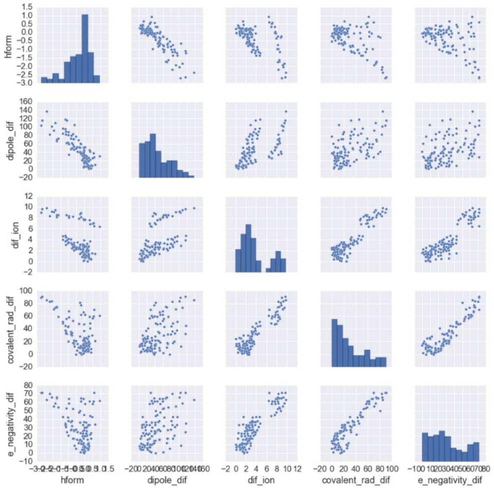
- Discovery of new 2D materials?
- Tailor 2D materials to have desired properties?
- Use statistical inference to guide the discovery of new physics?

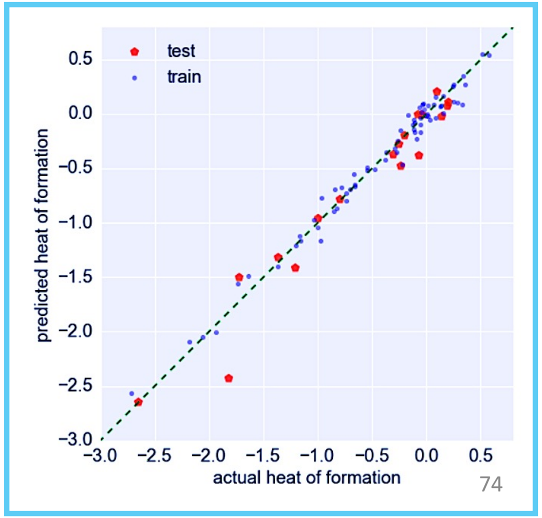




Outlook

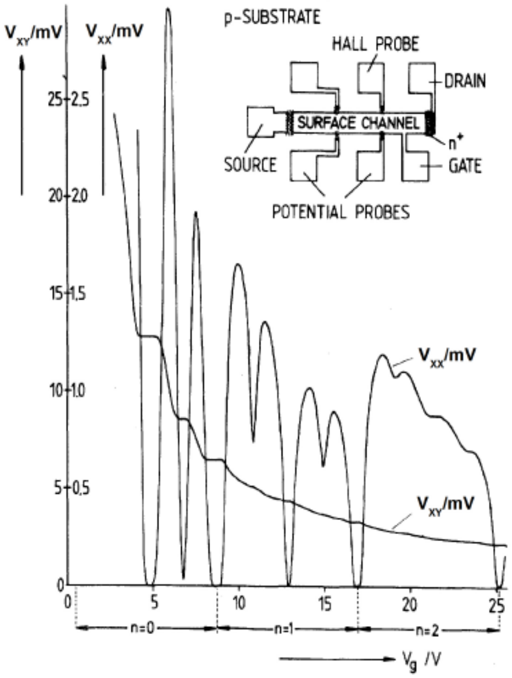
- Academic research
  - Universities
  - Materials Project (Materials Genome Initiative)
  - NIMS
- Industry
  - Citrine
  - Google
  - Applications:
    - ❖Photovoltaics
    - ❖Thermoelectrics
    - ❖Organic LEDs
    - ❖Magnetocalorics





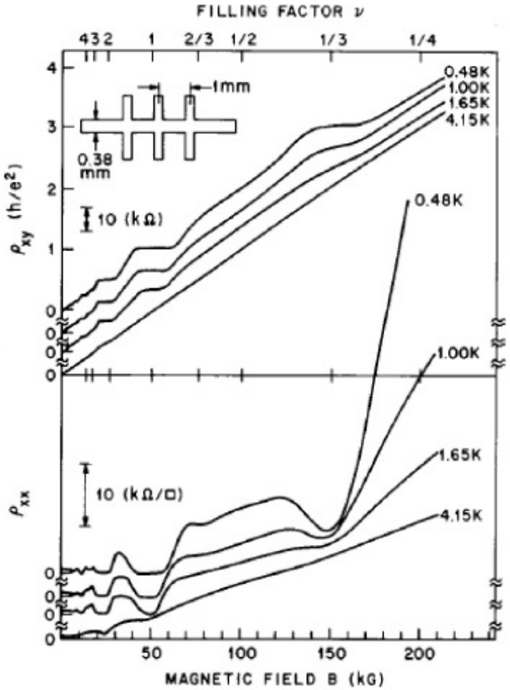
2DES in a magnetic field: Integer Quantum Hall Effect

Integer Quantum Hall Effect



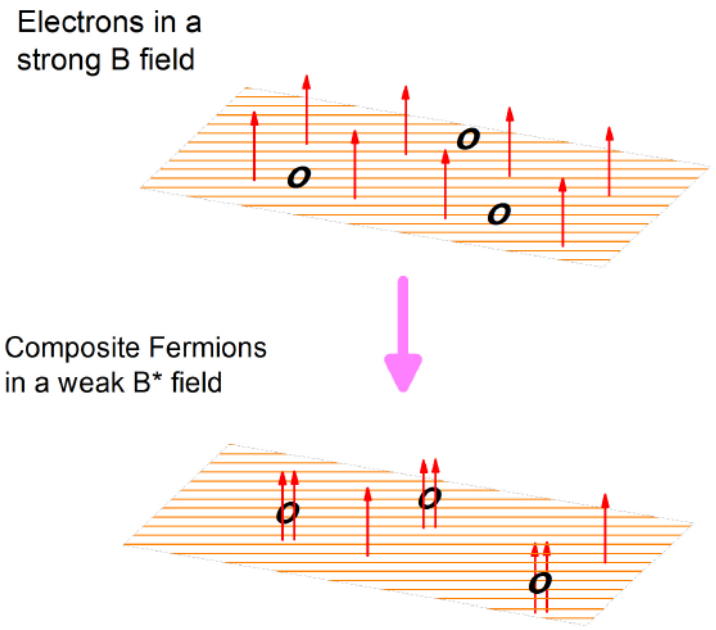
1. K. von Klitzing, G. Dorda, and M. Pepper. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. Phys. Rev. Lett., 45(494), 1980.

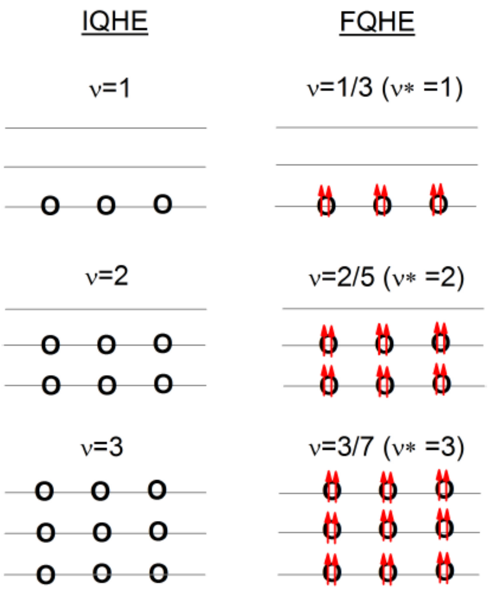
Fractional Quantum Hall Effect



2. D.C. Tsui, H.L. Stormer, and A.C. Gossard. Two-dimensional magnetotransport in the extreme quantum limit. Phys. Rev. Lett., 48 (1559), 1982.

Emergent Quasiparticles: Composite Fermions





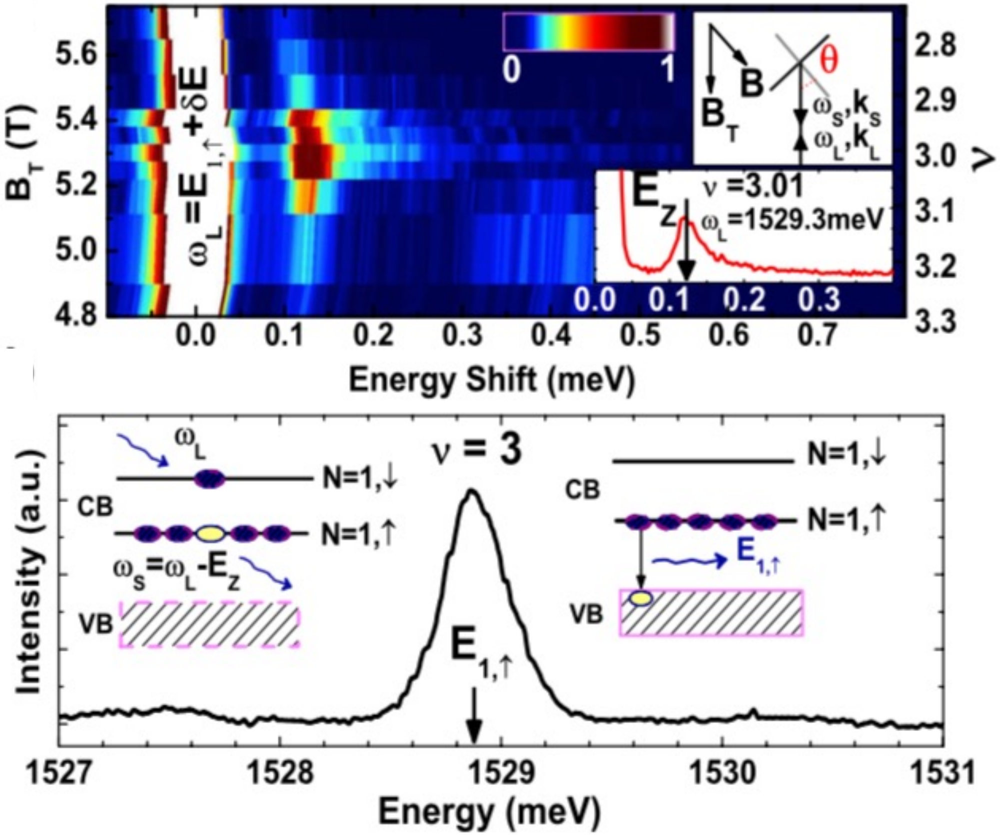
IQHE $\Psi_{\nu=1} = \prod_{j<k}(z_j - z_k)exp[\frac{-1}{4l_o^2}\sum_l |z_l|^2]$

FQHE $\Psi_{p/(\phi p+1)} = \hat{P}_{LLL}\prod_{j<k}(z_j - z_k)^{\phi}\Psi_{\nu=p}$

$$\nu^* = n\phi_o/|B^*| \quad \nu = \frac{\nu^*}{\phi\nu^* \pm 1}$$

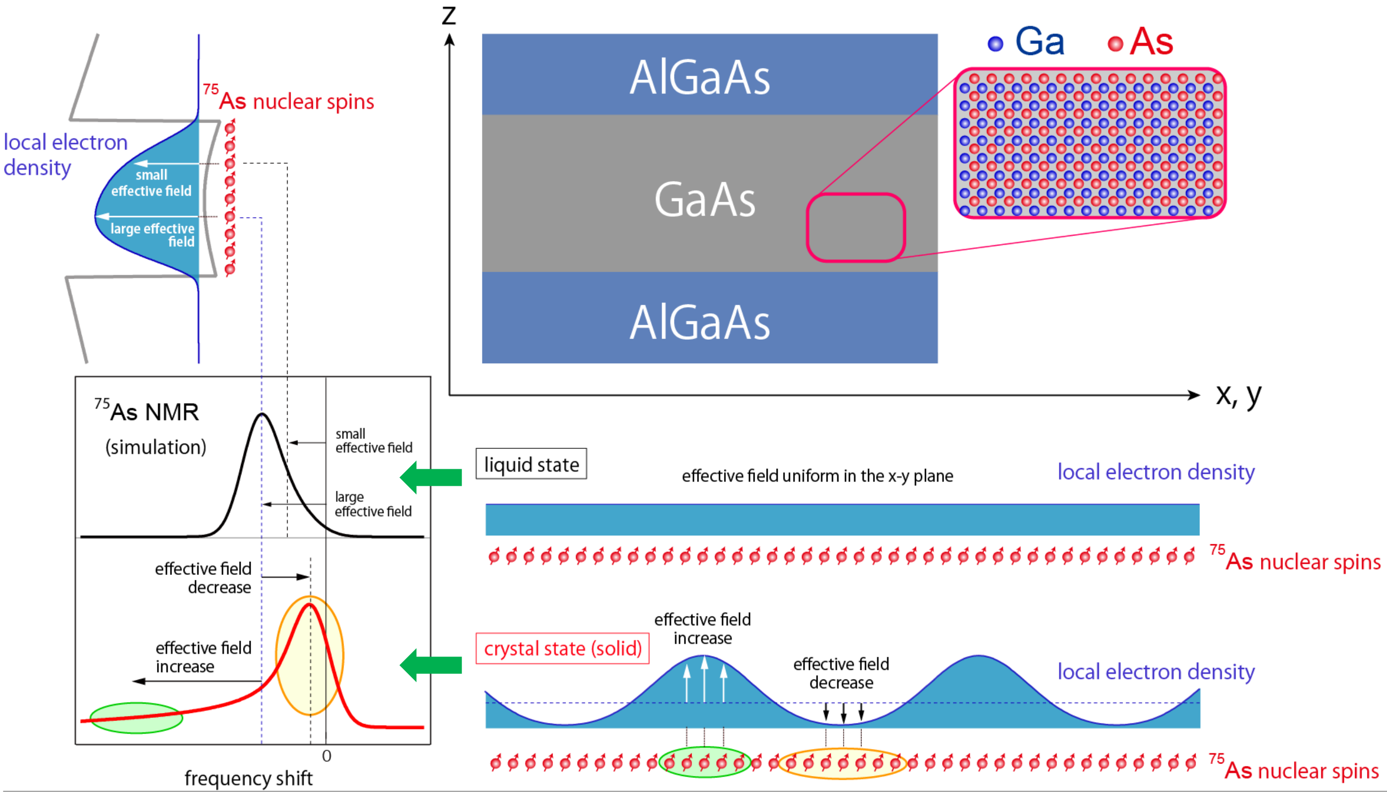
J.K. Jain. Composite Fermions. Cambridge University Press, 1st edition, 2007.

Spin properties of 2DES



T.D. Rhone et al, PRL 106, 196805 (2011)

Resistively detected NMR (RDNMR) as a Local Density Probe



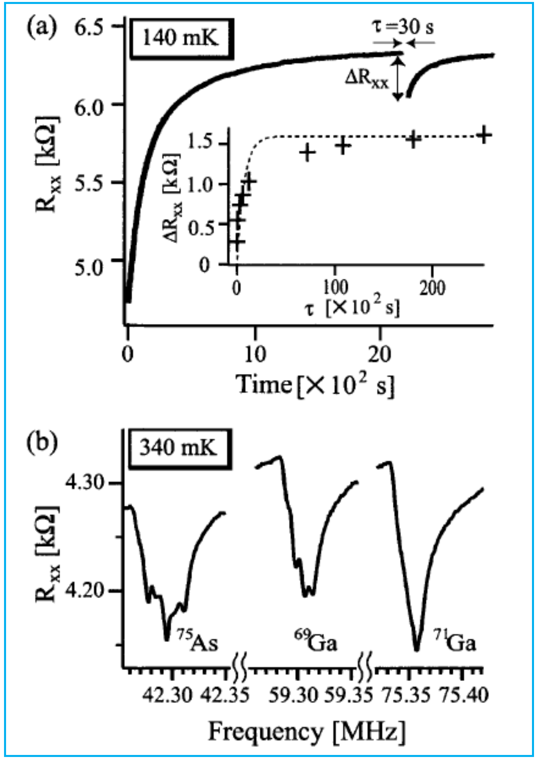
(effective field exerted on nuclear spins) ∝ (electron spin polarization) × (local electron density)

Resistively Detected NMR (RD-NMR)

VOLUME 88, NUMBER 17 PHYSICAL REVIEW LETTERS 29 APRIL 2002

Electrically Controlled Nuclear Spin Polarization and Relaxation by Quantum-Hall States

K. Hashimoto,[1,2,*] K. Muraki,[1] T. Saku,[1] and Y. Hirayama[1,2]



$E_z = g\mu_B(B_0 + b_n \langle I_z \rangle)$

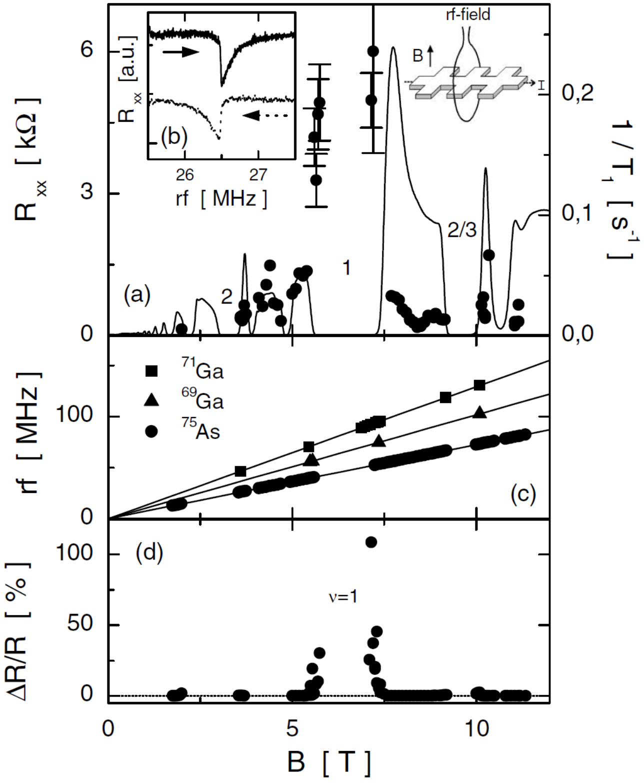
$\Delta E_z \propto \Delta \langle I_z \rangle$

if $\Delta R_{xx} \propto \Delta E_z$, then $\Delta R_{xx} \propto \Delta \langle I_z \rangle$

only applicable to those states with:

- sensitivity to ΔE_z
- finite R_{xx}

Resistively Detected NMR (RD-NMR)



W. Desrat *et al.*, PRL **88**, 256807 (2002).

$E_z = g\mu_B(B_0 + b_n \langle I_z\rangle)$

$\Delta E_z \propto \Delta\langle I_z\rangle$

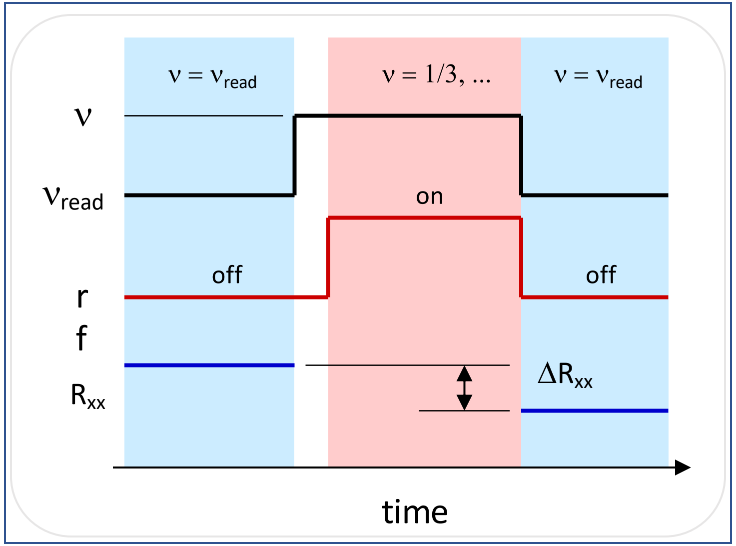
if $\Delta R_{xx} \propto \Delta E_z$, then $\Delta R_{xx} \propto \Delta\langle I_z\rangle$

only applicable to those states with:

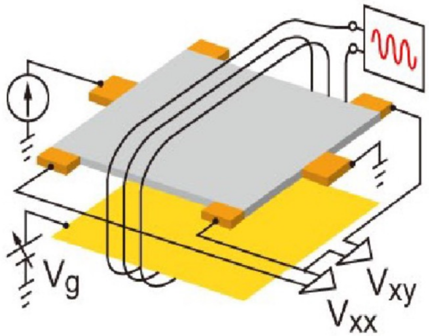
- sensitivity to ΔE_z
- finite R_{xx}

Our Method: Modified RD-NMR

Key idea:
Use gate voltage to go to a different filling factor ν_{read} for the readout



Detect resonant rf absorption at ν
as a change in R_{xx} at ν_{read}

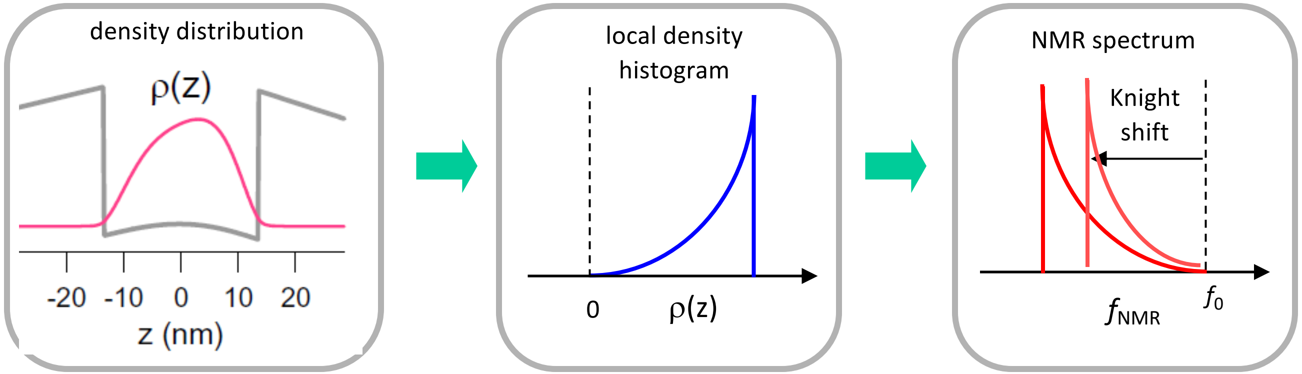


- Nuclear spins interact with electron spins at ν_{read}=2/3
- R_{xx} at ν_{read} can be used to measure spin degrees of freedom

NMR Line Shape

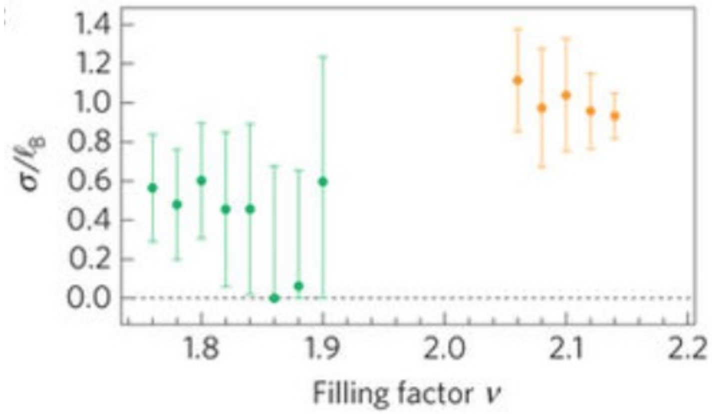
$K_s(\mathbf{r})$ (Knight shift of nucleus at position $\mathbf{r}$) $= \alpha_\nu$(spin polarization * n) x $\rho(\mathbf{r})$ (local probability density)

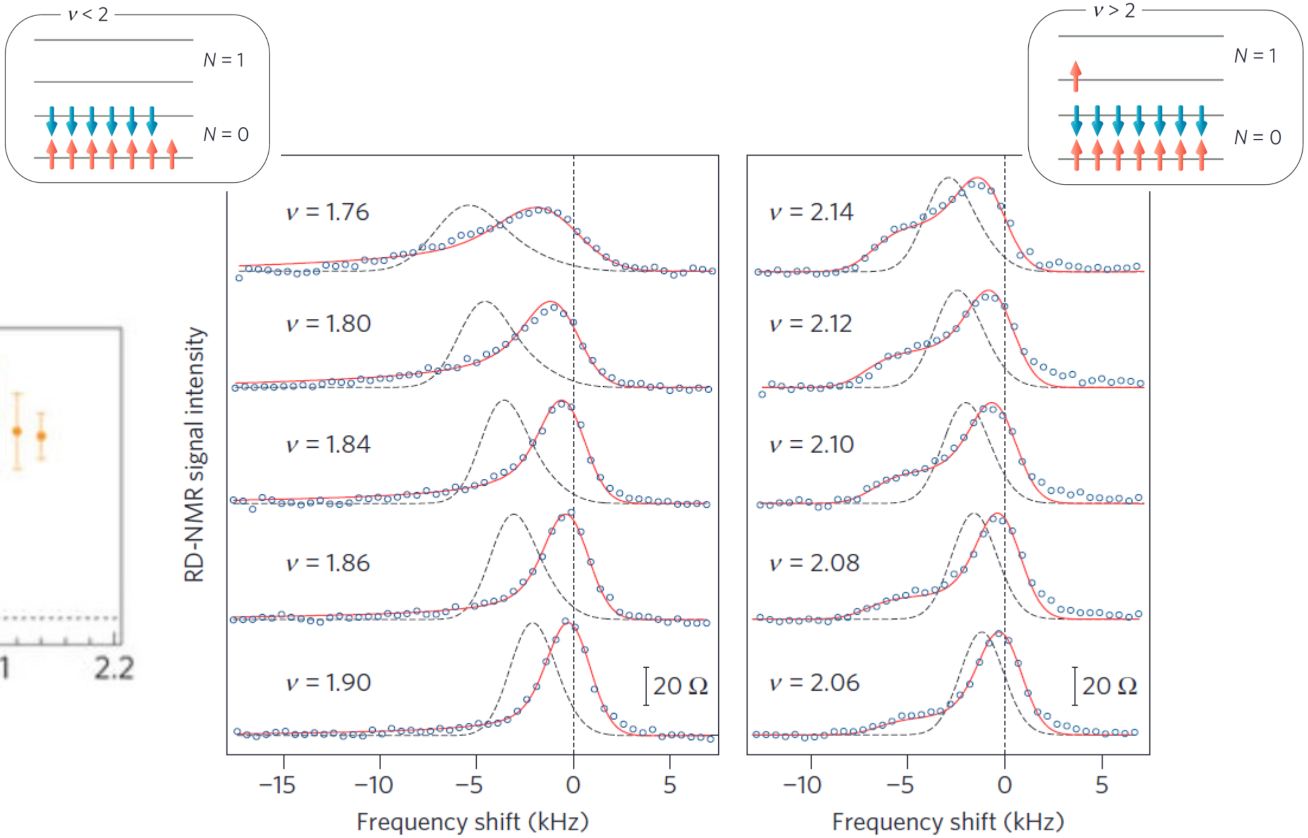
$$P = \alpha_\nu / \alpha_{\max,\nu}$$



- For $P > 0$, NMR line shape ≅ mirror image of local density histogram
- spectral mapping of subband wavefunction

NMR probes Wigner Solids: *In-plane local density variations at ν ~ 2*





1) L. Tiemann(*), T.D. Rhone(*), N. Shibata, K. Muraki, Nature Physics, doi:10.1038/nphys3031
(*: equal contribution)

"I Keep Six Honest Serving Men ..."

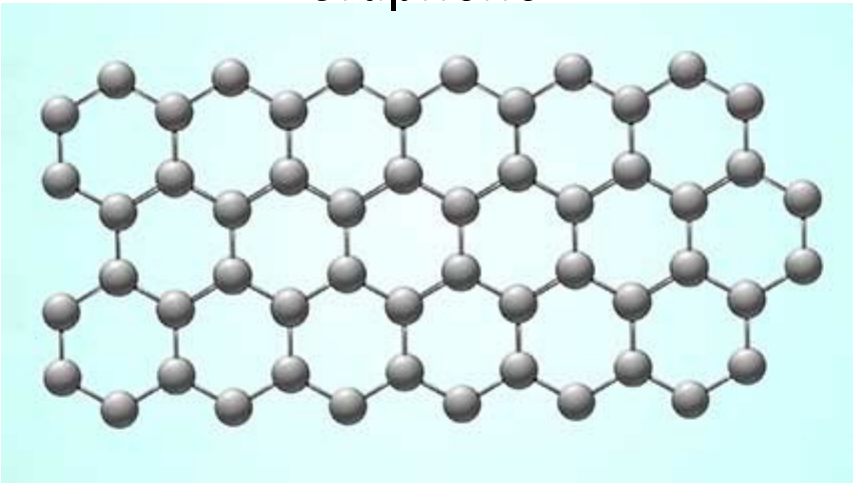
I keep six honest serving-men
(They taught me all I knew);
Their names are What and Why and When
And How and Where and Who…

by Rudyard Kipling

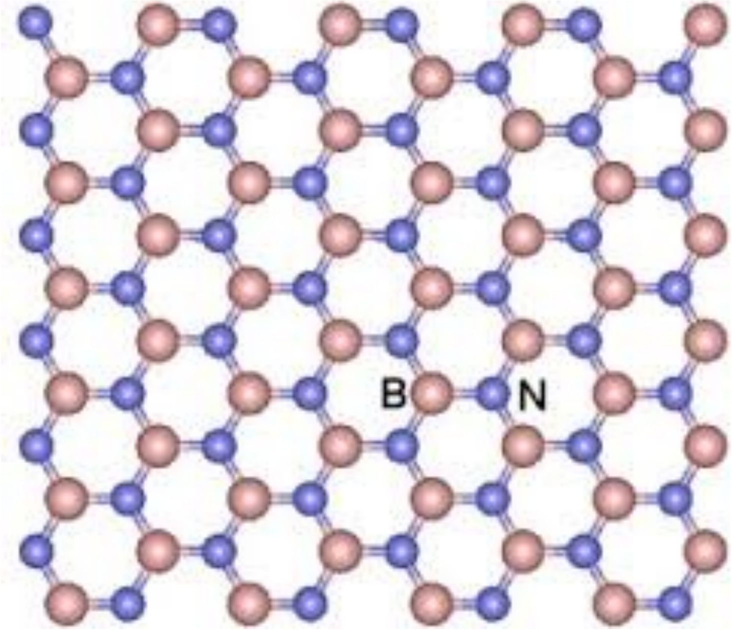
Crystal structure and materials properties

- Structure-property relationships:

Graphene



Boron Nitride



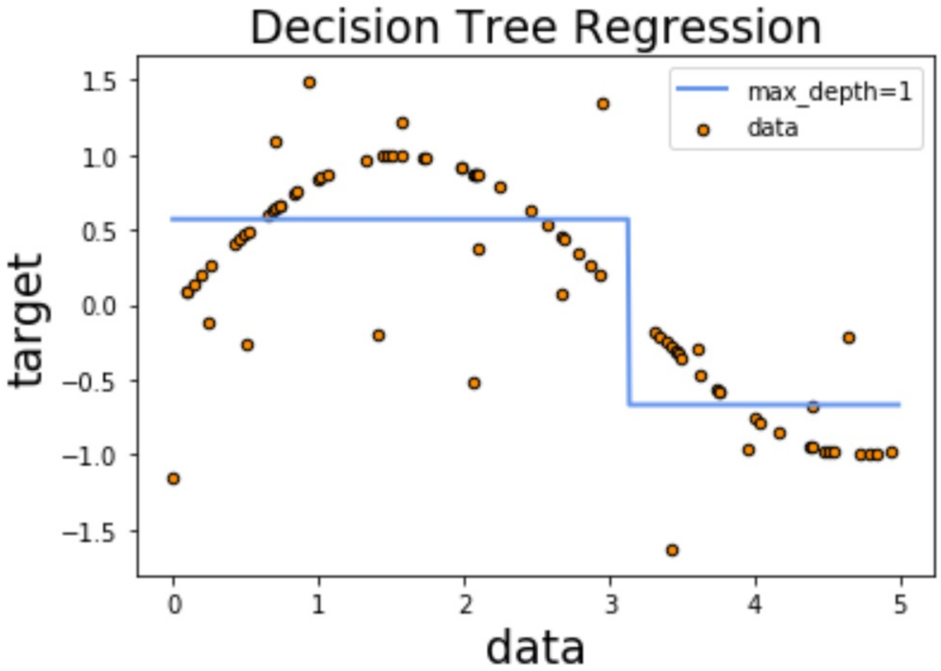
Total set of descriptors

1 num_p
2 num_d
3 num_f
4 atomic_rad
5 atomic_vol
6 covalent_rad
7 dipole
8 eaffinity
9 num_electrons
10 atomic_rad_avg
11 atomic_rad_max_dif
12 atomic_rad_sum_dif
13 atomic_rad_std_dif
14 atomic_rad_std
15 atomic_vol_avg
16 atomic_vol_max_dif
17 atomic_vol_sum_dif
18 atomic_vol_std_dif
19 atomic_vol_std
20 covalentrad_avg
21 covalentrad_max_dif
22 covalentrad_sum_dif
23 covalentrad_std_dif
24 covalentrad_std
25 dipole_avg
26 dipole_max_dif
27 dipole_sum_dif
28 dipole_std_dif
29 dipole_std
30 numelectron_avg
31 numelectron_max_dif
32 numelectron_sum_dif
33 numelectron_std_dif
34 numelectron_std
35 vdwradius_avg
36 vdwradius_max_dif
37 vdwradius_sum_dif
38 vdwradius_std_dif
39 vdwradius_std
40 e_negativity_avg
41 e_negativity_max_dif
42 e_negativity_sum_dif
43 e_negativity_std_dif
44 e_negativity_std
45 nvalence_avg
46 nvalence_max_dif
47 nvalence_sum_dif
48 nvalence_std_dif
49 nvalence_std
50 lastsubshell_avg
51 cmpd_skew_p
52 cmpd_skew_d
53 cmpd_skew_f
54 cmpd_sigma_p
55 atomE_AB
56 frac_f
57 std_ion
58 cmpd_sigma_d
59 cmpd_sigma_f
60 sum_ion
61 mean_ion

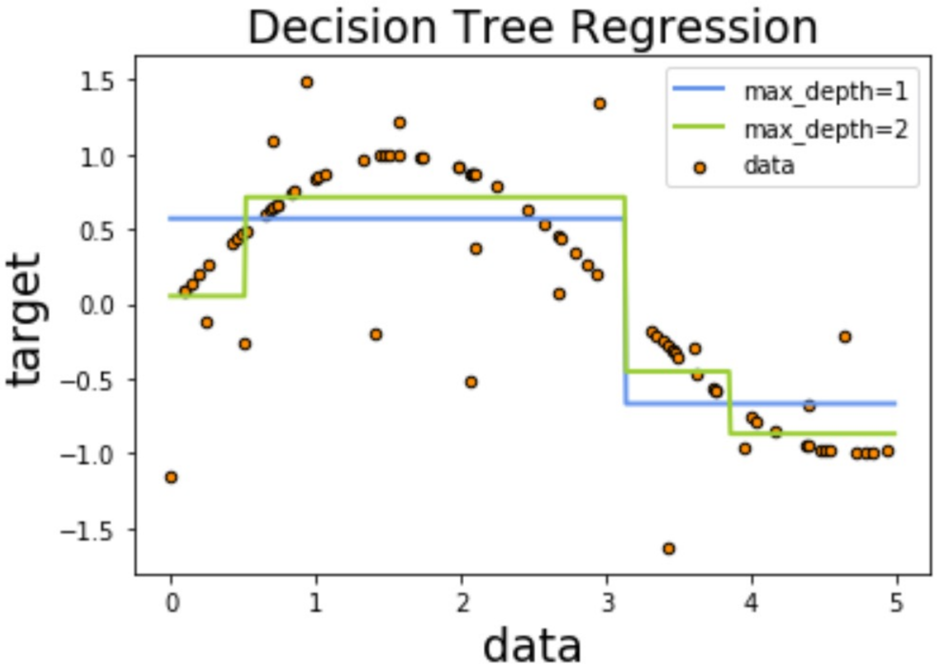
Reduced set of descriptors

1 covalentrad_max_dif
2 hardness_var
3 hardness_mean
4 covalentrad_avg
5 sum_ion
6 Nup_mean
7 Nup_var
8 atomic_rad_std_dif
9 atomic_vol_avg
10 atomic_vol_max_dif
11 nvalence_std_dif
12 dipole_std
13 covalentrad_std_dif
14 dipole_max_dif
15 nvalence_sum_dif
16 atomic_vol_std_dif
17 nvalence_max_dif
18 atomic_vol_sum_dif
19 nvalence_std
20 nvalence_avg

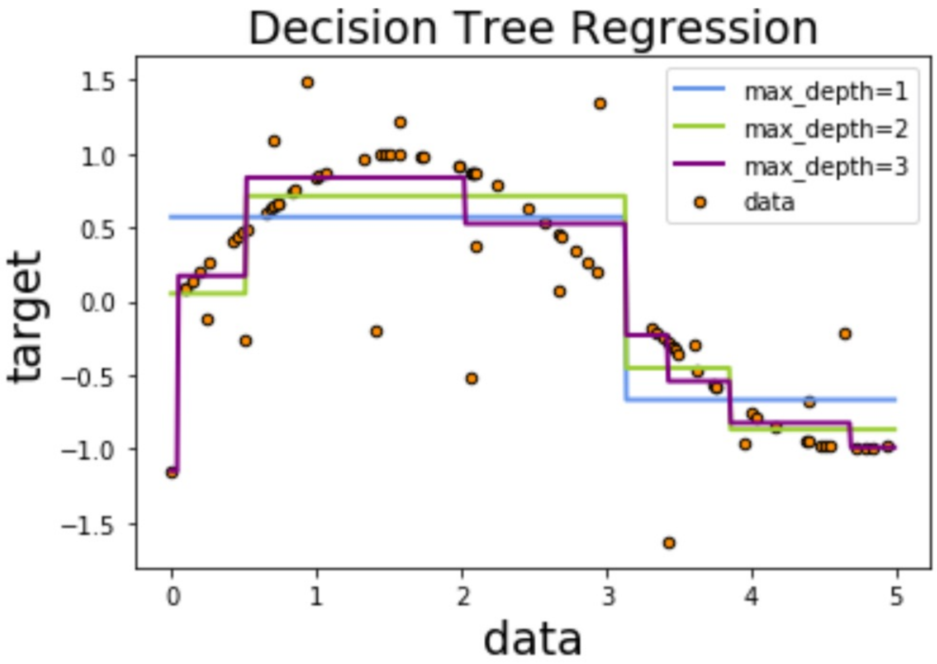
Random Forests Regression



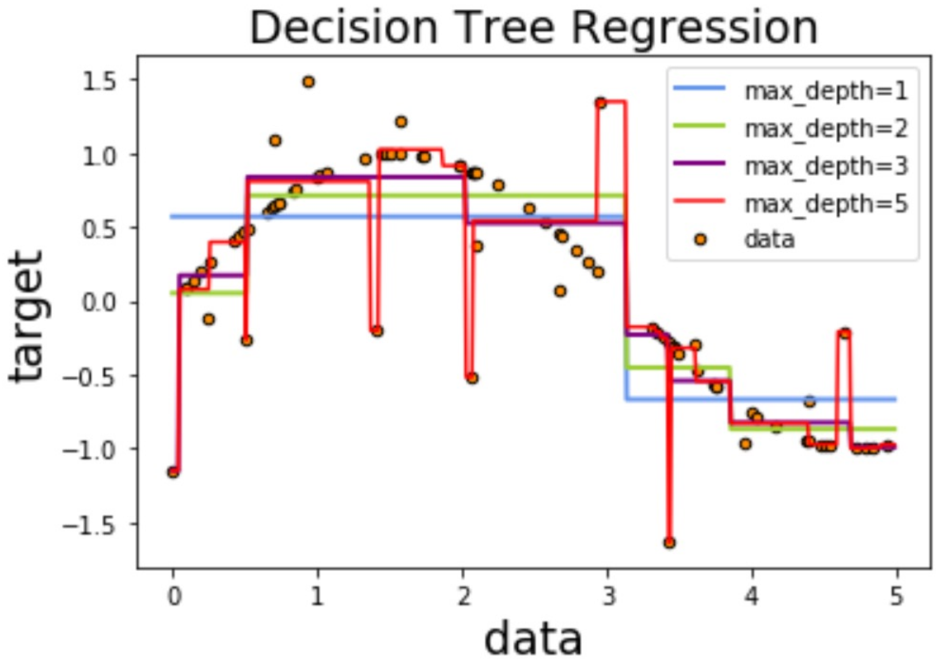
Random Forests Regression



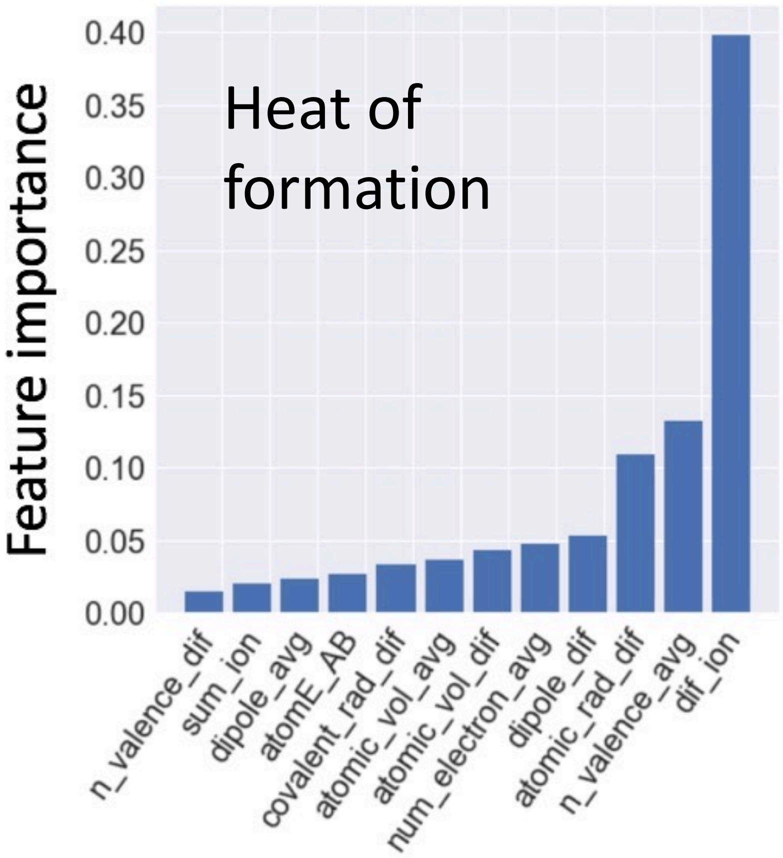
Random Forests Regression



Random Forests Regression



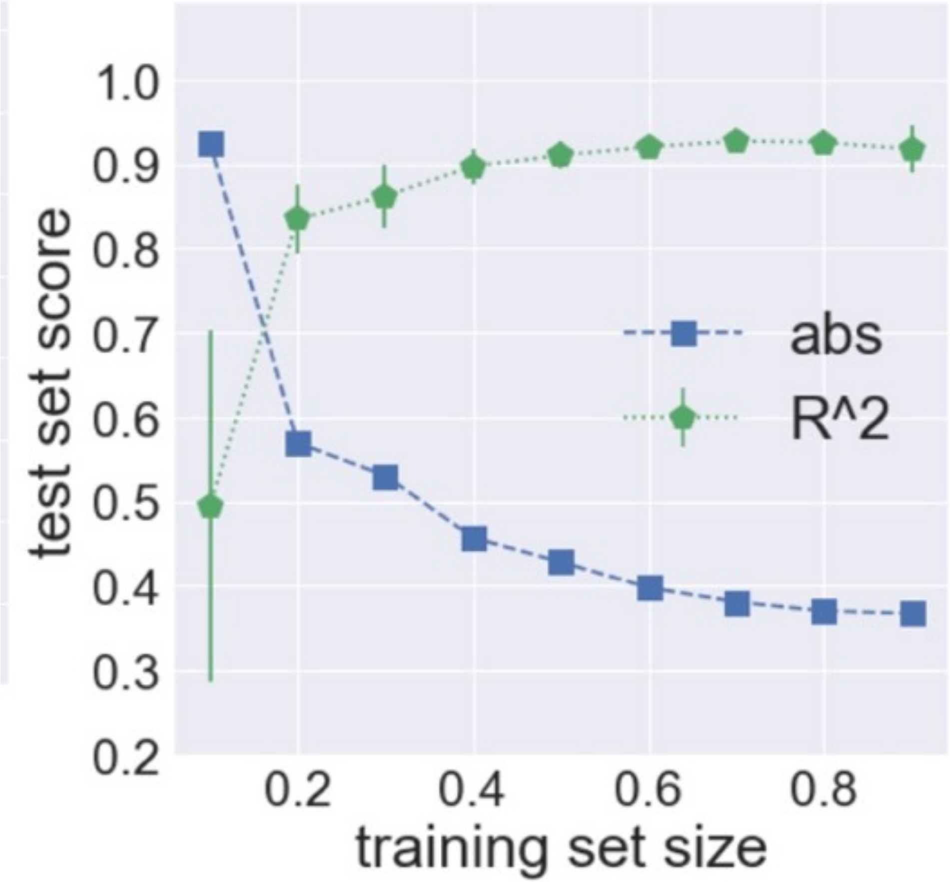
Extra forest regression: Feature extraction



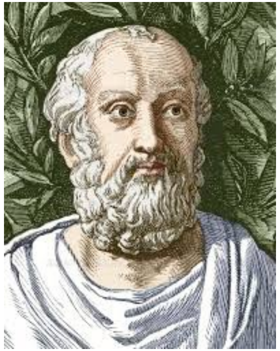
Feature Importance Ranking:

1. Ionization energy dif.
2. # valence electrons avg.
3. Atomic radius dif.
4. Dipole polarizability dif
5. Num of electrons avg.
6. Atomic volume avg
7. Covalent radius dif
8. Atom E_AB
9. Dipole poarizability avg
10. Ionization energy sum
11. # of valence electrons dif

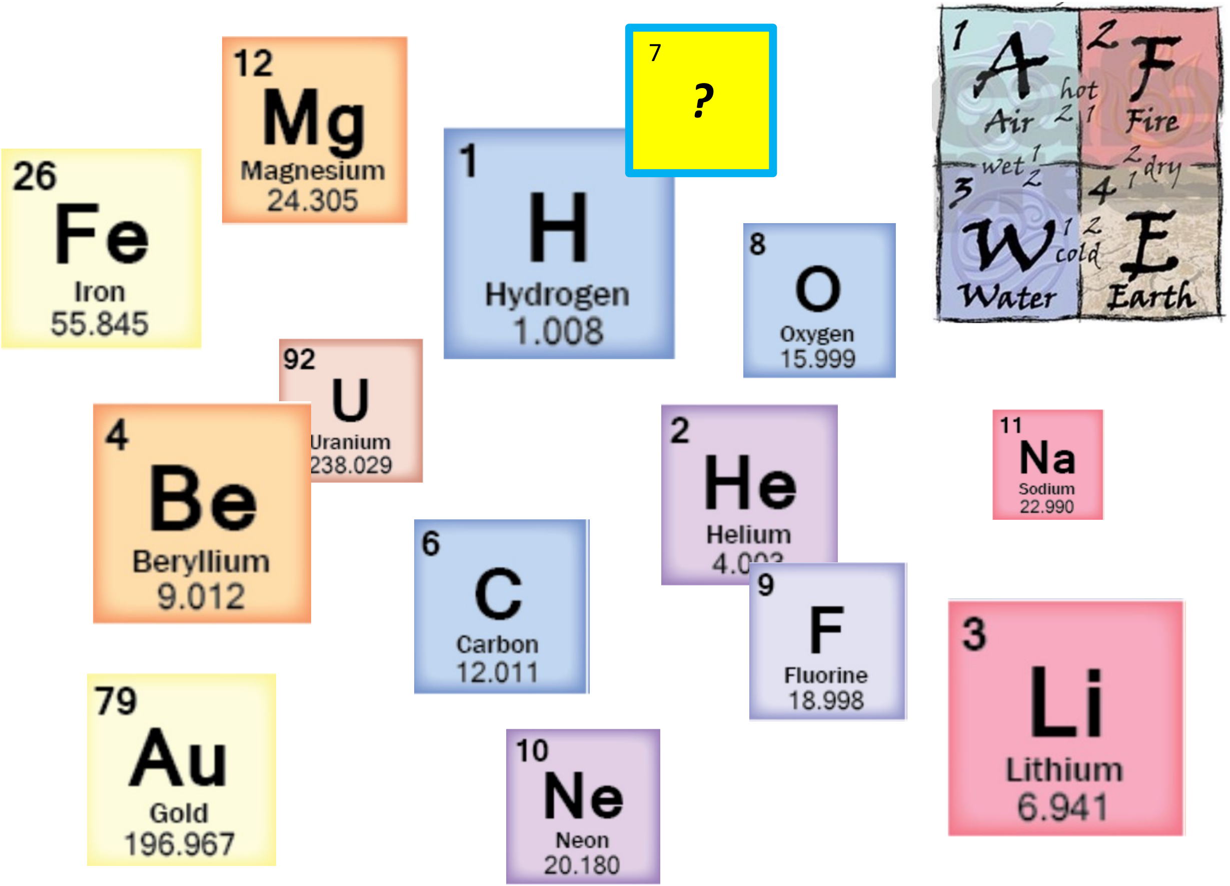
Test score vs training set size



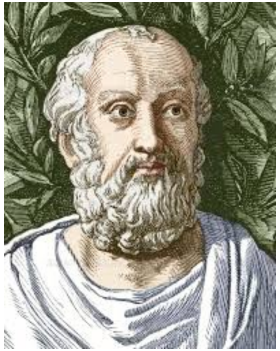
Where and who of Materials Informatics

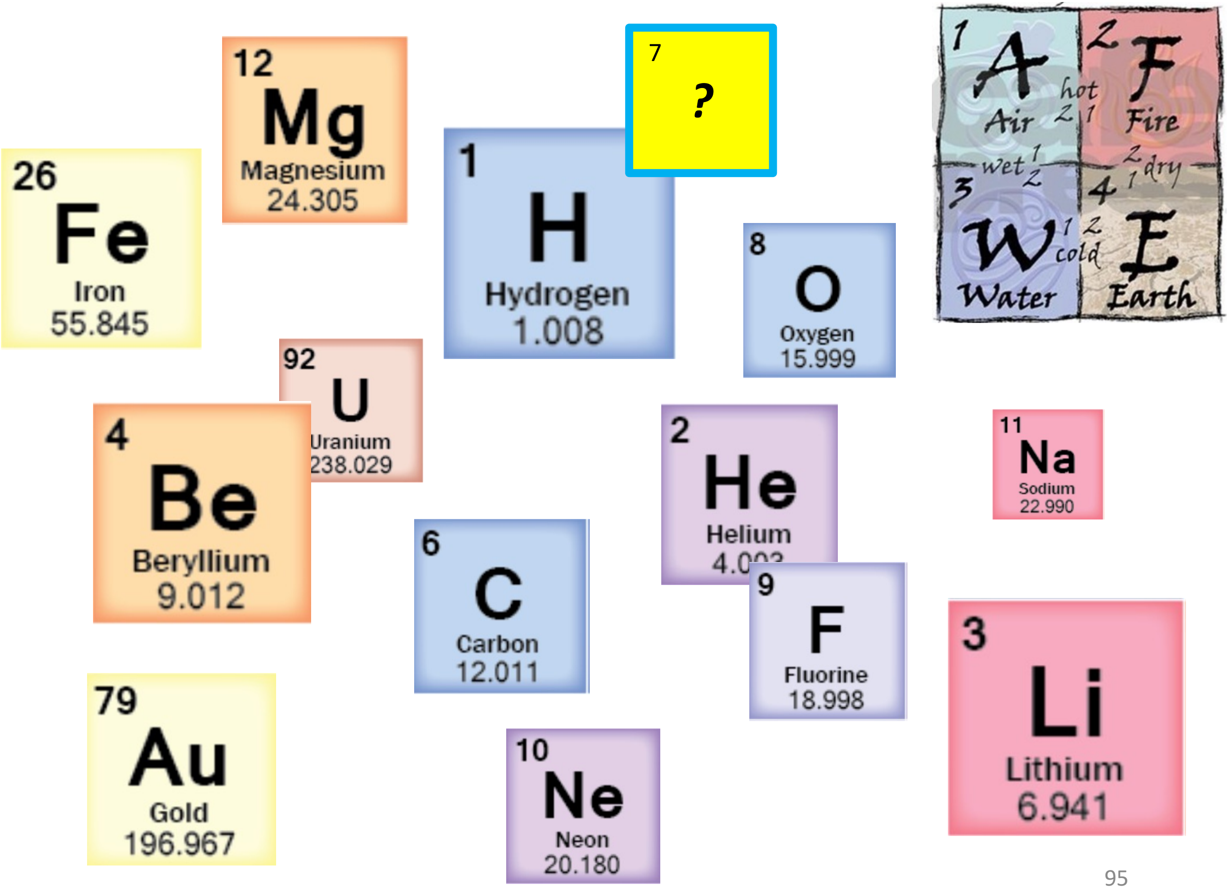


- You are an alchemist living in ancient Greece
- You have a small random set of elements



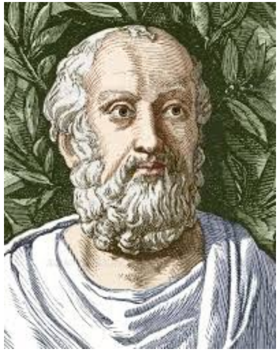
Where and who of Materials Informatics

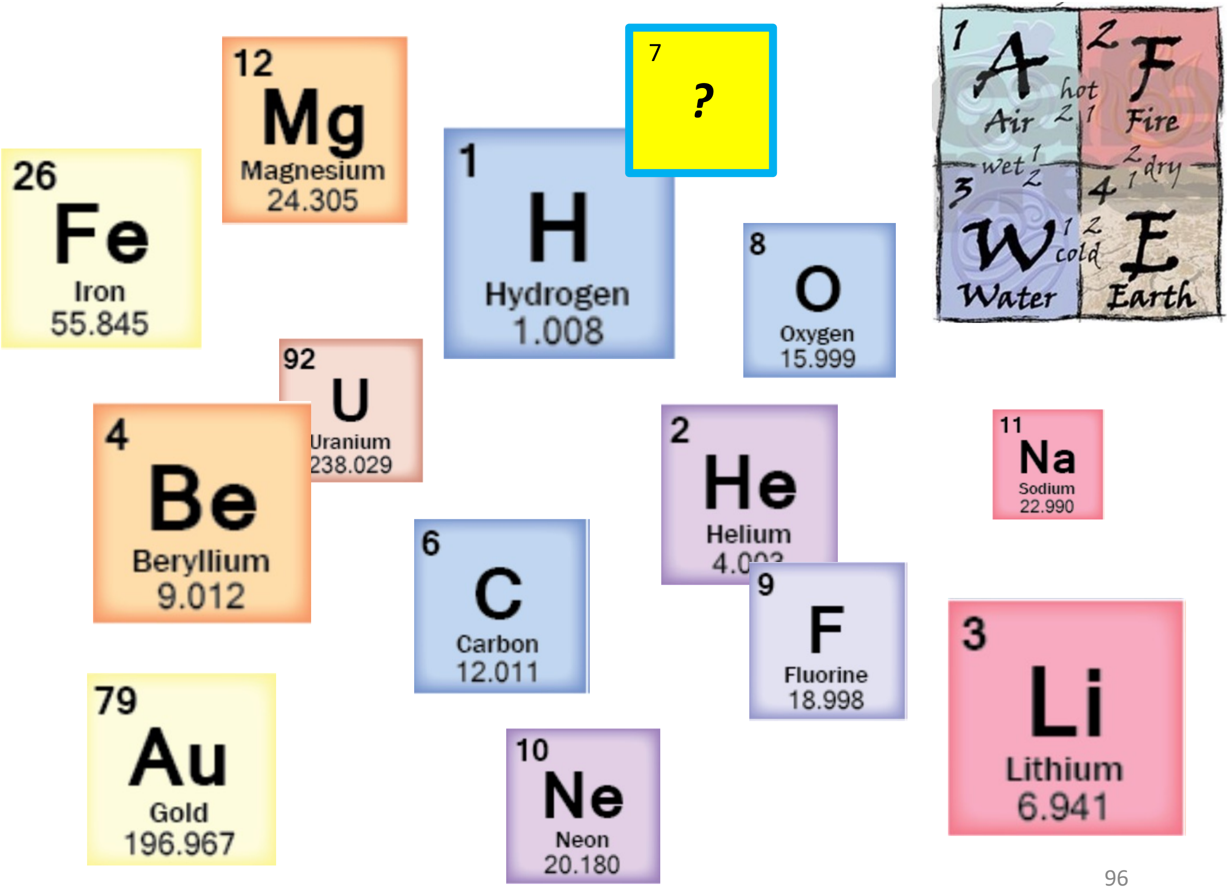




- What is a good descriptor?
- What are trends in the data?

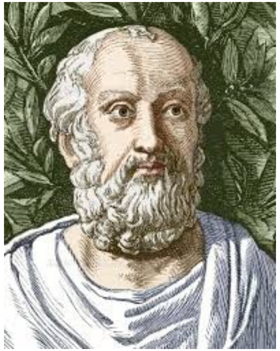
Where and who of Materials Informatics



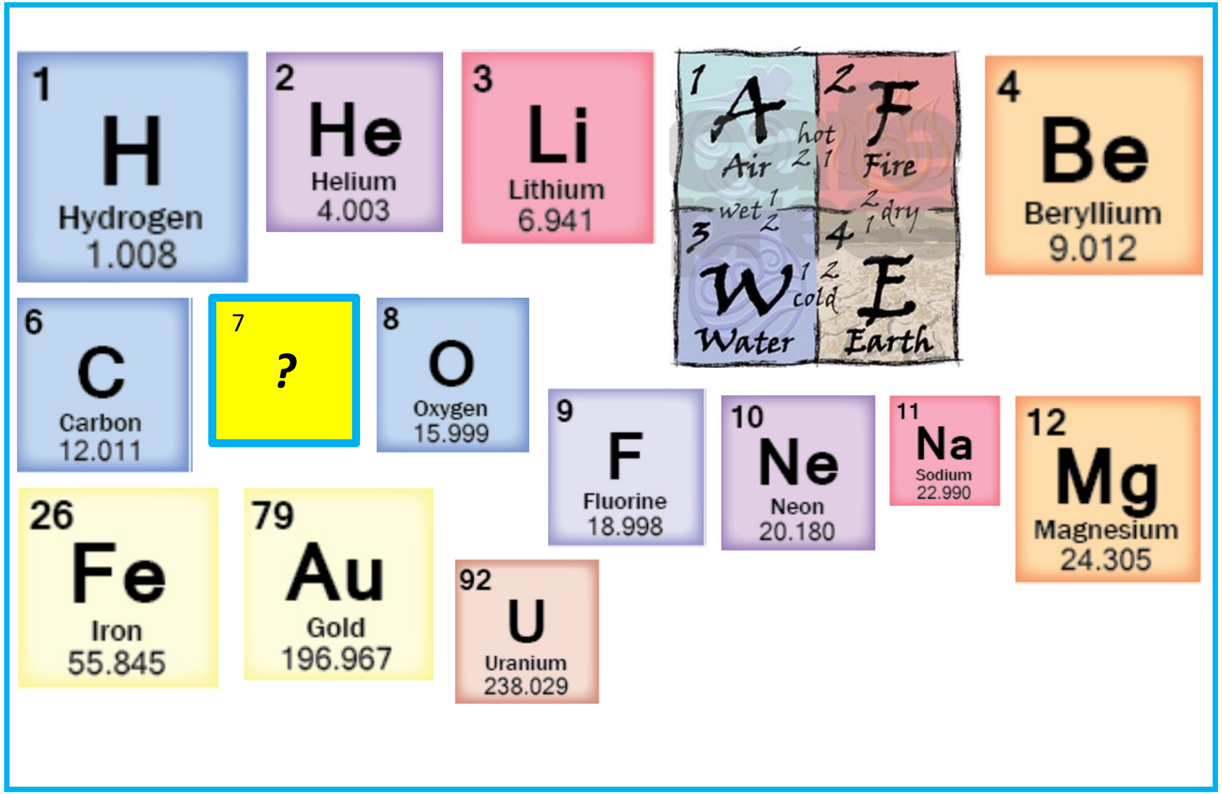


- Descriptor:
  - Atomic number
- Apply a model:
  - Sort atomic numbers

Where and who of Materials Informatics

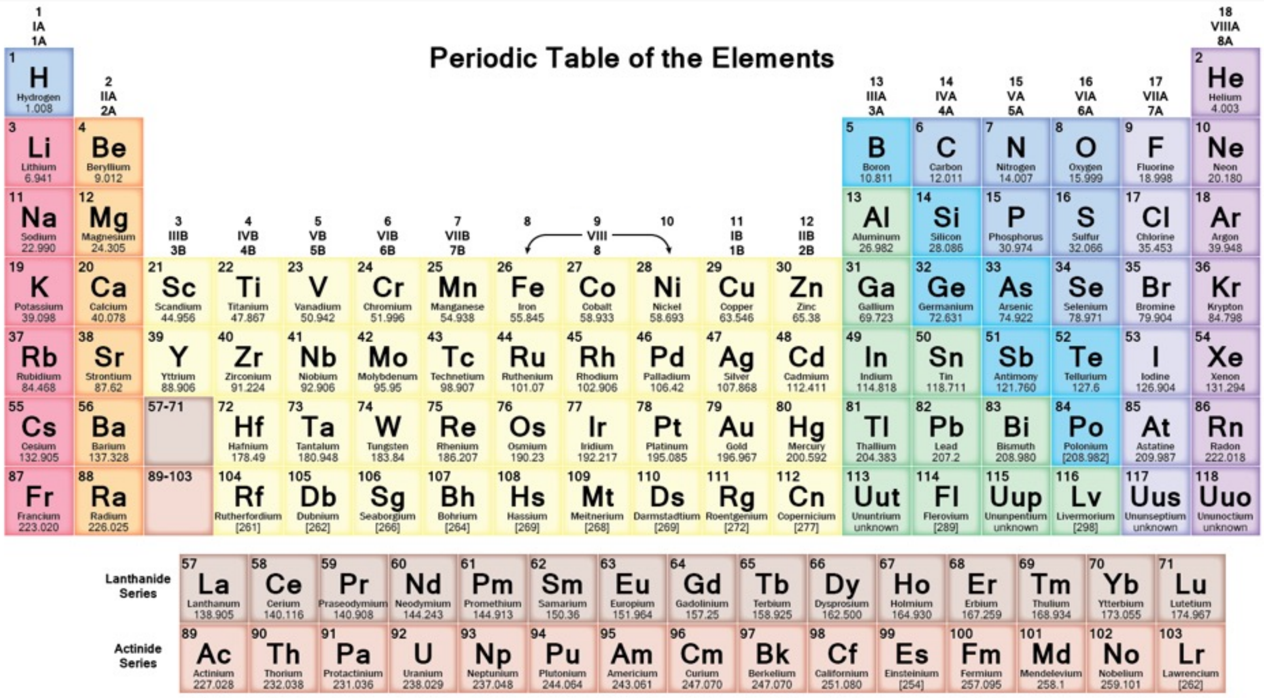


- Descriptor:
  - Atomic number
- Apply a model:
  - Sort atomic numbers



Where and who of Materials Informatics

- Descriptor:
  - Atomic number
- Apply a model:
  - Sort atomic numbers



Machine Learning for Materials studies

Data

Materials Project

AFLOW
Automatic - FLOW for Materials Discovery

MatNavi
NIMS Materials Database

MATDAT.COM
Material Properties Database and Estimation Tools

Inorganic Crystal Structure Database

Machine Learning for Materials studies

Descriptors

- Kitchen sink method:
  - Use statistical analysis to learn relevant descriptors
- Domain knowledge:
  - Construct features that are important for describing a system
- Example:
  - Ionic compounds have atoms in different columns of periodic table
  - Descriptor: column # of the periodic table

Machine Learning for Materials studies

Datascience tools

Statistical models