Find descriptors of Zero-Point-Renormalization(ZPR) by SISSO

Jingkai Quan, coffee talk

25/07/2021
NOMAD Laboratory

Content

Introduction:

- Introduction of Zero-point renormalization
- Introduction of SISSO

Find descriptor of ZPR by SISSO:

- Focus on stable binary materials
- Including unstable structures
- Next Step

Content

Introduction:

- ☐ Introduction of Zero-point renormalization
- ☐ Introduction of SISSO

Find descriptor of ZPR by SISSO:

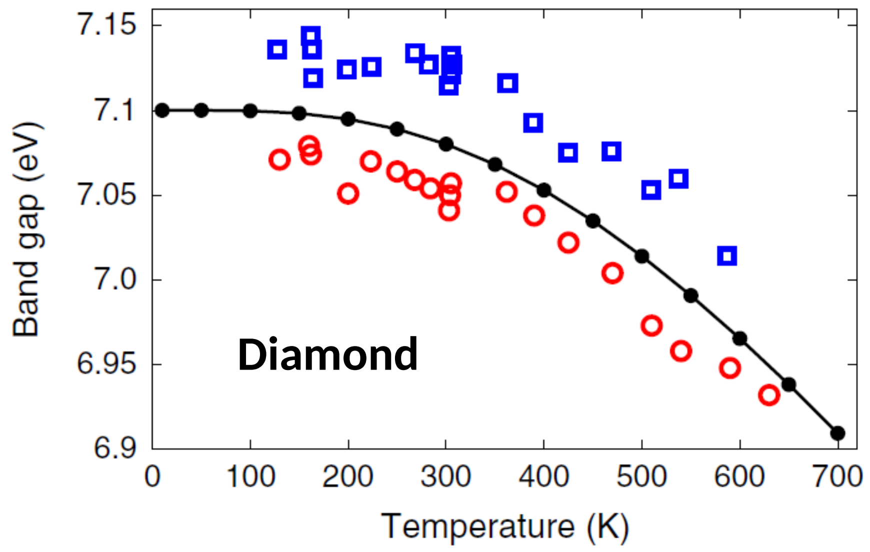
- ☐ Focus on stable binary materials
- ☐ Including unstable structures
- ☐ Next Step

Zero-point renormalization (ZPR)

- **Electronic structures can be affect by electron-phonon interaction (epi)：**

**PRL 105, 265501 (2010)*

e.g. temperature dependence of electronic band gap



→ Affect band structure even at 0K.

- **Harmonic approximation：**

Potential energy taylor expansion to the displacement ΔR：

$$E\left(\{\mathbf{R}^0+\Delta\mathbf{R}\}\right)\approx E\left(\{\mathbf{R}^0\}\right)+\sum_I \frac{\partial E}{\partial \mathbf{R}_I}\bigg|_{\mathbf{R}^0}\Delta\mathbf{R}_I+\frac{1}{2}\sum_{I,J}\frac{\partial^2 E}{\partial \mathbf{R}_I \partial \mathbf{R}_J}\bigg|_{\mathbf{R}^0}\Delta\mathbf{R}_I\Delta\mathbf{R}_J+\mathcal{O}(\Delta\mathbf{R}^3)$$

$$\Phi_{IJ}=\frac{\partial^2 E}{\partial \mathbf{R}_I \partial \mathbf{R}_J}\bigg|_{\mathbf{R}^0}$$

Hessian matrix (FORCE_CONSTANT) DFPT or Finite difference

Harmonic Approximation

NOMAD NOVEL MATERIALS DISCOVERY

$$\Phi_{IJ} = \left.\frac{\partial^2 E}{\partial \mathbf{R}_I \partial \mathbf{R}_J}\right|_{\mathbf{R}^0} \longrightarrow \mathbf{F}_I = -\sum_J \Phi_{IJ} \Delta \mathbf{R}_J$$

$$\begin{pmatrix} m_1 \, \Delta \ddot{\mathbf{R}}_1 \\ m_2 \, \Delta \ddot{\mathbf{R}}_2 \\ \vdots \\ m_N \, \Delta \ddot{\mathbf{R}}_N \end{pmatrix} = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \cdots & \Phi_{1N} \\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{N1} & \Phi_{N2} & \cdots & \Phi_{NN} \end{pmatrix} \begin{pmatrix} \Delta \mathbf{R}_1 \\ \Delta \mathbf{R}_2 \\ \vdots \\ \Delta \mathbf{R}_N \end{pmatrix}$$

tip1: differential property of fourier trans.

$$f^{(n)}(x) \overset{Fourier}{\leftrightarrow} (i\omega)^n F(\omega)$$

$$f^{(2)}(x) \overset{Fourier}{\leftrightarrow} -\omega^2 F(\omega)$$

↓ *Fourier*

Eigenvalue Problem:

$$D(\mathbf{q})\, e_s(\mathbf{q}) = \omega_s^2(\mathbf{q})\, e_s(\mathbf{q})$$

$e_s(\mathbf{q})$ **:eigenvector (displacement)**

$$D_{IJ}(\mathbf{q}) = \sum_{J'} \frac{\mathrm{e}^{\mathrm{i}(\mathbf{q}\cdot\mathbf{R}_{JJ'})}}{\sqrt{M_I M_J}} \Phi_{IJ'}$$

Dynamic matrix (*fourier trans. of* ***Φ***)

Zero-point renormalization (ZPR)

NOMAD
NOVEL MATERIALS DISCOVERY

Total zero-point renormalization:

$$\delta\varepsilon_i(\mathbf{k}) = \delta\varepsilon_i^{\,\mathrm{A}}(\mathbf{k}) + \delta\varepsilon_i^{\,\mathrm{P}}(\mathbf{k})$$

**J. Phys. Chem. C 2021, 125, 6479−6485*

**Phys. Rev. Lett. 112, 215501*

**Phys. Rev. Lett. 94, 036801*

$\delta\varepsilon_i^{\,\mathrm{A}}(\mathbf{k})$ **: Adiabatic term**

$\delta\varepsilon_i^{\,\mathrm{P}}(\mathbf{k})$ **: polar term**

$$D(\mathbf{q})\,\boldsymbol{e}_s(\mathbf{q}) = \omega_s^2(\mathbf{q})\,\boldsymbol{e}_s(\mathbf{q})$$

$\boldsymbol{e}_s(\mathbf{q})$ **:eigenvector (displacement)**

- **Adiabatic term:**

$$\delta\varepsilon_i^{\,\mathrm{A}}(\mathbf{k}) = \sum_\lambda \int \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} \frac{\hbar}{4\omega_\lambda(\mathbf{q})} \frac{\partial^2\varepsilon_i(\mathbf{k})}{\partial\delta_{\lambda q}^{\;2}}$$

Expand by eigen displacements

$$\delta\varepsilon_i^{\,\mathrm{A}}(k) = \sum_\lambda \int \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} \frac{\hbar}{4\omega_\lambda(\mathbf{q})} \frac{e_{I\alpha,\lambda}(\mathbf{q})}{\sqrt{M_I}} \frac{\partial^2\varepsilon_i(\mathbf{k})}{\partial R_{I\alpha}\partial R_{J\beta}} \frac{e_{J\beta,\lambda}(\mathbf{q})}{\sqrt{M_J}}$$

$\delta\varepsilon_i^{\,\mathrm{A}}(\mathbf{k})$ $\quad \dfrac{1}{(M_I \cdot M_J)^{1/4}}$

Zero-point renormalization (ZPR)

- Polar term: Fröhlich model by Nery and Allen

*J. Phys. Chem. C 2021, 125, 6479−6485

*Phys. Rev. B 94, 115135

$$\delta\varepsilon_i^{\mathrm{P}}(\mathbf{k}) = \pm\frac{\alpha_i \hbar\omega_{\mathrm{LO}}}{\pi}\tan^{-1}\left(\frac{q_{\mathrm{F}}}{q_i}\right)$$

$$\alpha_i = \frac{e^2 q_i}{8\pi\tilde{\varepsilon}_0\hbar\omega_{\mathrm{LO}}}\left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0}\right) \qquad q_i = \sqrt{\frac{2m_i^*\omega_{\mathrm{LO}}}{\hbar}}$$

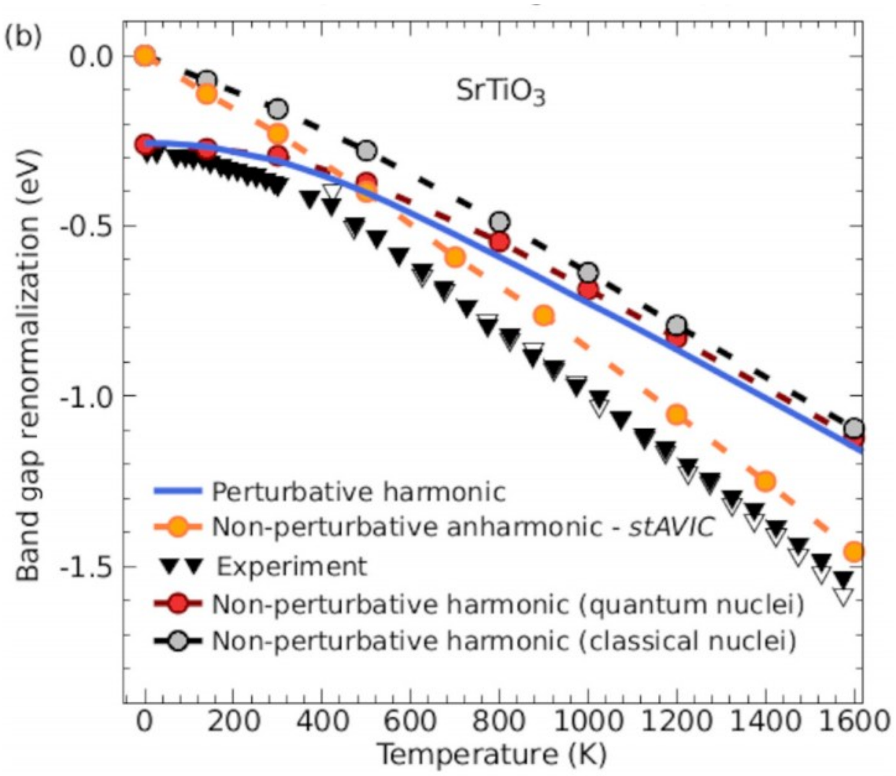
$\delta\varepsilon_i^{\mathrm{A}}(\mathbf{k})$ ω_{LO}

$\delta\varepsilon_i^{\mathrm{A}}(\mathbf{k})$ m^*

Beyong harmonic approximate

- There also beyond the harmonic/perturbative pictures exist

  but that they are not used here.



*Zacharias, M., Scheffler, M. & Carbogno, C. **Phys. Rev. B 102, 045126***

Content

Introduction:

- ☐ Introduction of Zero-point renormalization
- ☐ Introduction of SISSO

Find descriptor of ZPR by SISSO:

- ☐ Focus on stable binary materials
- ☐ Including unstable structures
- ☐ Next Step

SISSO: a compressed-sensing method

NOMAD NOVEL MATERIALS DISCOVERY

- **Sure Independence Screening (SIS) + Sparsifying Operator (SO)**

* *Phys. Rev. Materials* **2**, *083802*

STEP 1. Construct large feature space (hyperparameter: rung = n)

primary features

A, B, C, D, …

operators

+, -, ×, /, exp, log, …

rung-1 features

(A+B), (A-C), exp(B), …

operators

+, -, ×, /, exp, log, …

rung-2 features

(A+B)×C, (A-C)+D, exp(A+B), …

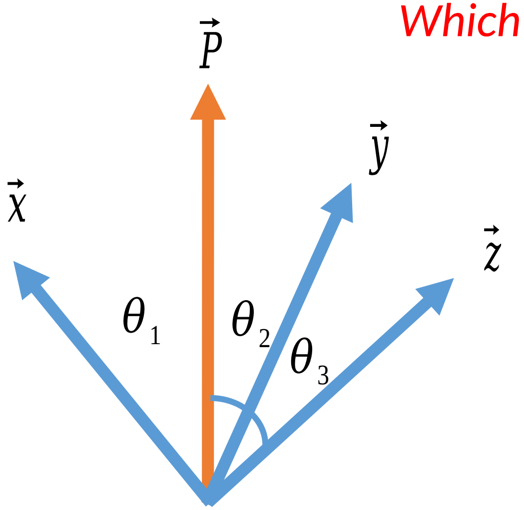
* Size of feature space grow rapidly with each iteration

SISSO: a compressed-sensing method

NOVEL MATERIALS DISCOVERY

STEP 2. SIS: reduce feature space size by pearson correlation

Which vector most correlated with target vector P?



- Pearson correlation: First **zero-centering** vectors, then calculate cosine

$$corr(x,P)=\frac{\sum_s (x_s-\overline{x})(P_s-\overline{P})}{\sqrt{\sum_s (x_s-\overline{x})^2}\sqrt{\sum_s (P_s-\overline{P})^2}}$$

$$\cos\theta=\frac{\sum_s (x_s P_s)}{\sqrt{\sum_s x_s^2}\sqrt{\sum_s P_s^2}}$$

$$\vec{x}=(x_1,x_2,\ldots\ldots,x_s)$$

$$\overline{x}=\frac{\sum_s x_s}{s}$$

We can select top ranked features to form a feature subspace (*hyperparameter*: *n_sis_select = n*)

e.g. choose 10000 top ranked features out of feature space.

SISSO: a compressed-sensing method

STEP 3. SO: Find desired descriptor in smaller feature subspace (l0-norm)

Find the sparse solution c in feature subspace: minimization problem

$$\underset{c}{argmin}\left(\|P - Dc\|_2^2 + \lambda \|c\|_0\right)$$

$$\|c\|_0 = number\ of\ nonzero\ components\ of\ c$$

MT-SISSO (Multi Task SISSO):

Can find an overall descriptor for multiple different but related properties. (only coefficients are different)

* *Phys. Rev. Materials* **2**, *083802*

J. Phys. Mater. **2** *024002*

Content

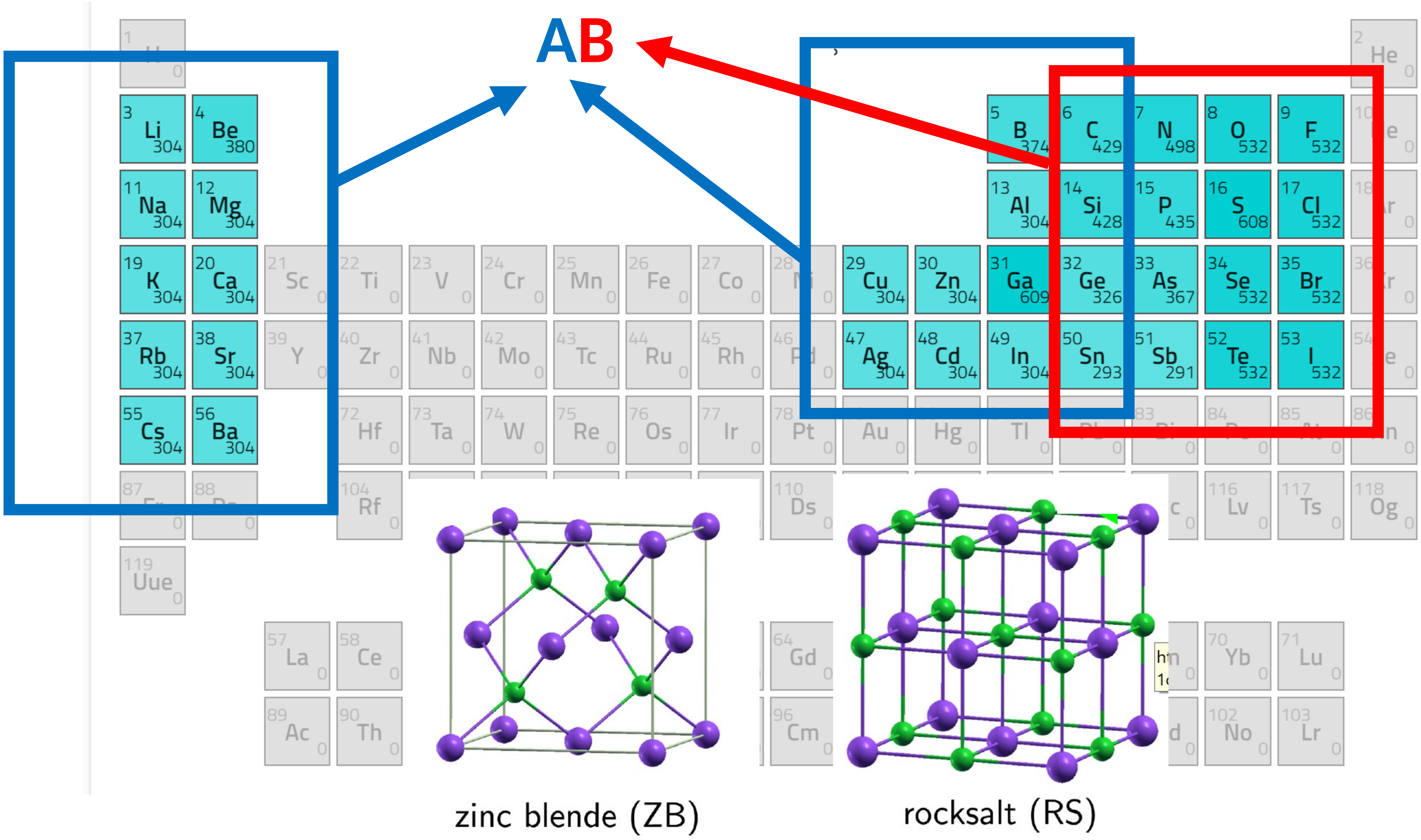
Introduction:

- ☐ Introduction of Zero-point renormalization
- ☐ Introduction of SISSO

Find descriptor of ZPR by SISSO:

- ☐ Focus on stable binary materials
- ☐ Including unstable structures
- ☐ Next Step

Overview: ZPR Dataset of binary materials



*J. Phys. Chem. C 2021, 125, 6479−6485

Content

Introduction:

- Introduction of Zero-point renormalization
- Introduction of SISSO

Find descriptor of ZPR by SISSO:

- Focus on stable binary materials
- Including unstable structures
- Next Step

Focus on stable structures

NOVEL MATERIALS DISCOVERY

1. ***82 rocksalt (RS) and 82 zincblende (ZB) structures in total.***

2. stable structures (total energy lowest) and leave metals out

- RS structure: 38 materials
- ZB structure: 37 materials

3. direct gap ZPR at Γ-point data is available with and w/o polar correlation term.

ZPR values calculated by Honghui, also uploaded on NOMAD.

**J. Phys. Chem. C 2021, 125, 6479−6485*

Primary Features

- **Property to learn: ZPR at Γ point**

30 primary features:

- **Effective mass at Γ point:**

- **Nuclear Mass related:**

$$*mass\ factor = \frac{1}{(mass_A \cdot mas}$$

- **Atomic features:**

  (IP=ionization potential, EA=electron affinity, r_s=radius of s orbital, r_val=radius of valance orbital)

- **Binary features:,**
  (dis=nearest atom distance, bond=bond length of AB, be=bond energy, hlg=Homo-Lumo gap)

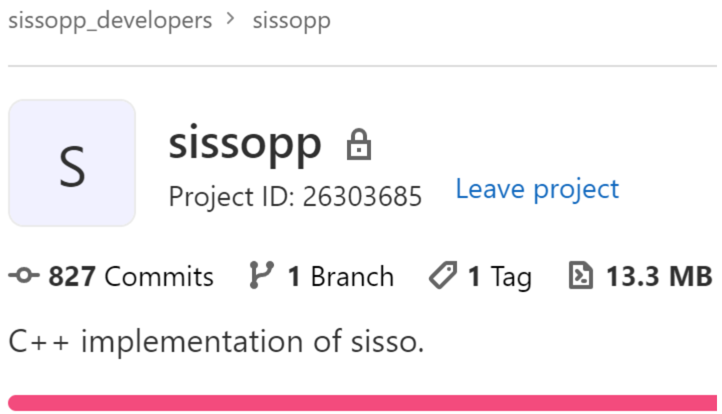
- **El-Ph-Related features:**

**Rev. Mod. Phys. 2005, 77, 1173−1224*

Softwares and codes

NOVEL MATERIALS DISCOVERY

1. C++ implementation of **SISSO** developed by **Tom**



**https://gitlab.com/sissopp_developers/sissopp*

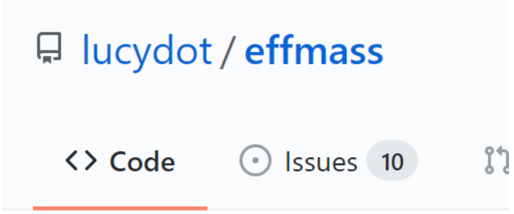
2. FHI-aims for structure relaxation and band structure calculation

Computer Physics Communications **180, 2175-2196 (2009)*

3. FHI-vibes for phonon dispersion calculation

**Journal of Open Source Software, 5(56), 2671*

4. effmass for effective mass calculation



**https://github.com/lucydot/effmass*

Effective masses at Γ point

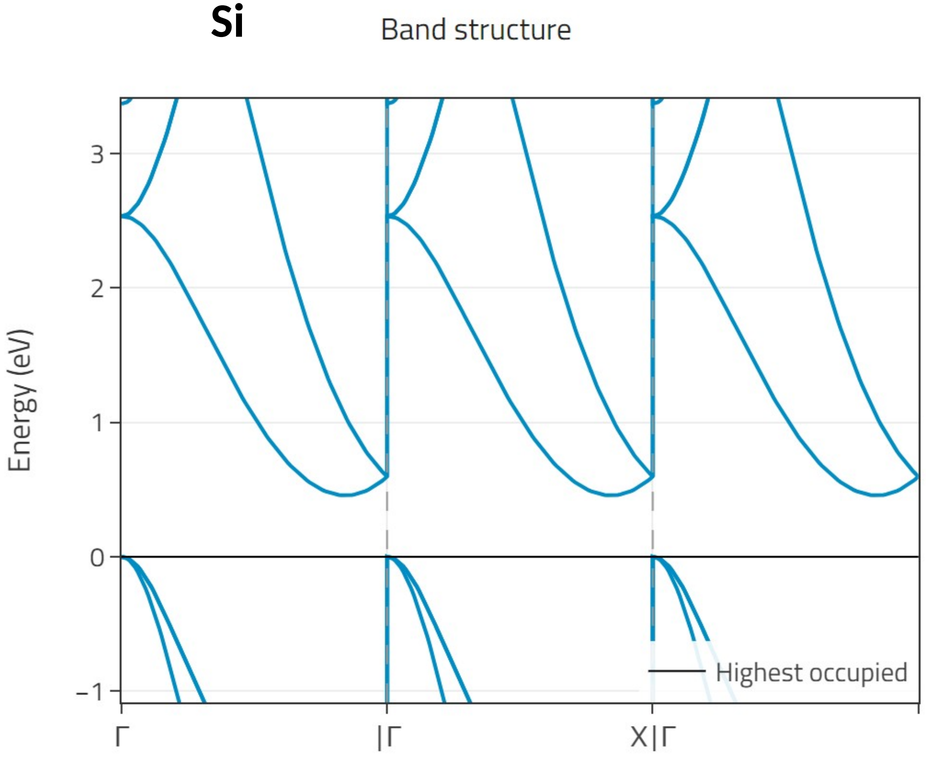
$$\frac{1}{m^*} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2}$$

Least square method： parabolic approximate

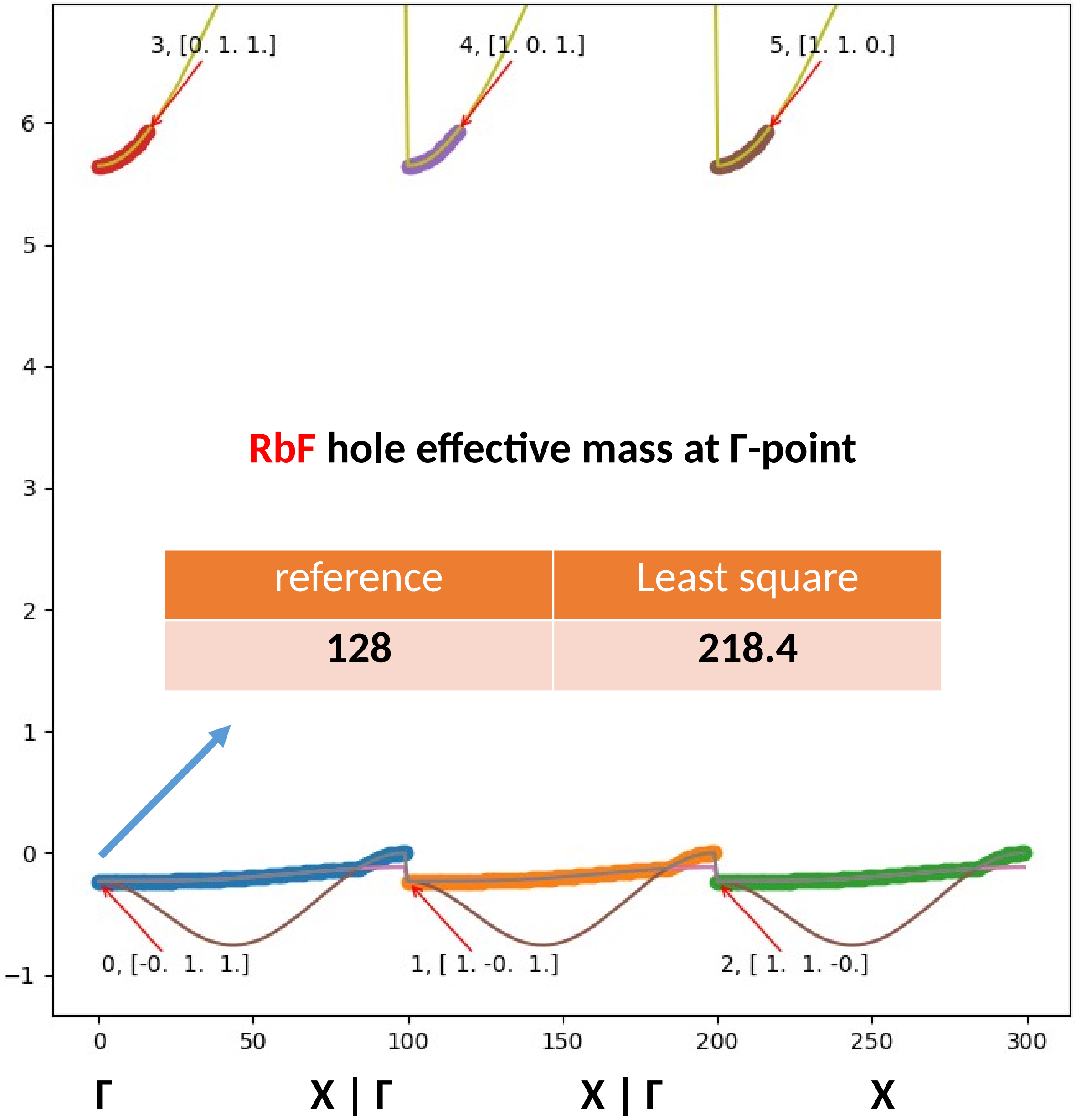
It's usually a good approximation, but:

Don't work for linear band dispersion!

Si



Effective masses at Γ point



RbF hole effective mass at Γ-point

reference	Least square
128	**218.4**

***Finite difference method：**

reference	Finite difference
128	**102.776**

- Thus use finite difference method to calculate effective masses by ***effmass*** code.

**https://github.com/lucydot/effmass*

**Phys. Rev. B 99(8), 085207*

SISSO on stable structures

NOVEL MATERIALS DISCOVERY

1. ***Use SISSO and Multitask-SISSO to find descriptors***

2. ***Performed 50 cross-validation (leave-10%-out) on less primary features***

#primary features	30	14
Time spent	~20h (Draco); ~8h (Raven)	~15min

Hyper-parameters:

- rung = 3
- n_sis_select = 10,000

(Tested n_sis_select = 5000; 10,000 and 50,000. They all find same descriptors.)

Cross-Validation on less primary features

1D descriptor:

RS: $$p=c_0+a_0\cdot\frac{\sqrt[3]{mass_{electron}}\cdot beBB\cdot homo_B\cdot IP_B}{dis_{AB}\cdot ¿\,IP_A\vee ¿¿}$$

ZB: $$p=c_0+a_0\cdot\frac{mass_{hole}\cdot be_{BB}\cdot be_{AB}}{mass_B\cdot(IP\,¿¿\,A\cdot mass_B+IP_B\cdot mass_A)¿}$$

MT: $$p=c_0+a_0\cdot\frac{\Delta_{a-a}{}^3+\frac{homo_B}{homo_A}}{\sqrt[3]{average_{mass}}\cdot rval_B\cdot homo_B}$$

choose all features appearing above (and the respective A/B counterparts):

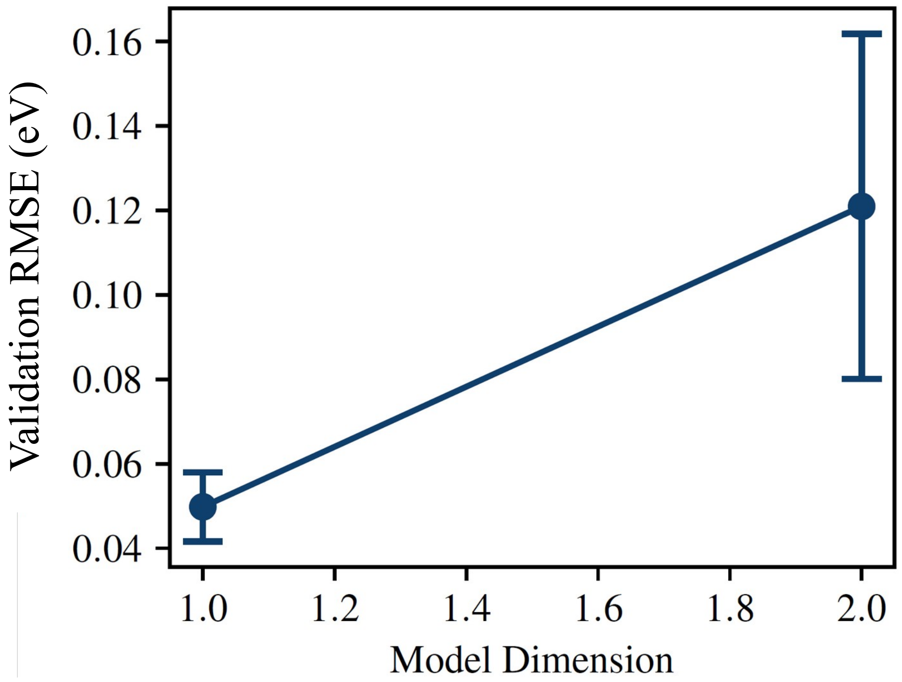
, , , , , , , , , , , , ,

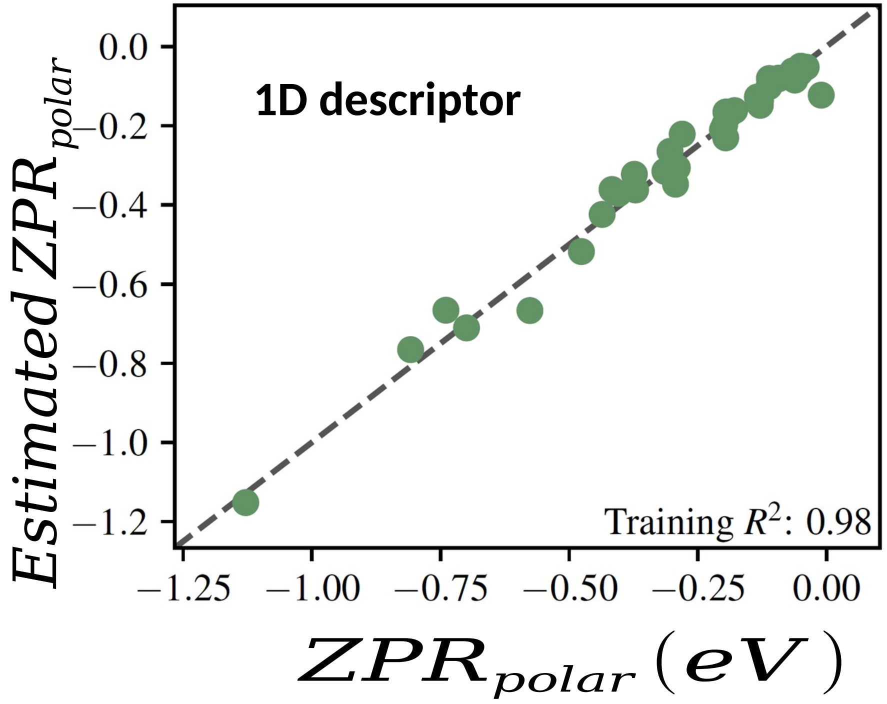
Do cross-validation on **total 14 features**

Cross-Validation for RS structure

mean RMSE on 50 leave-10%-out cross validations
- rung = ***3***
- n_sis_select = ***10,000***

Same parameters as our regression.





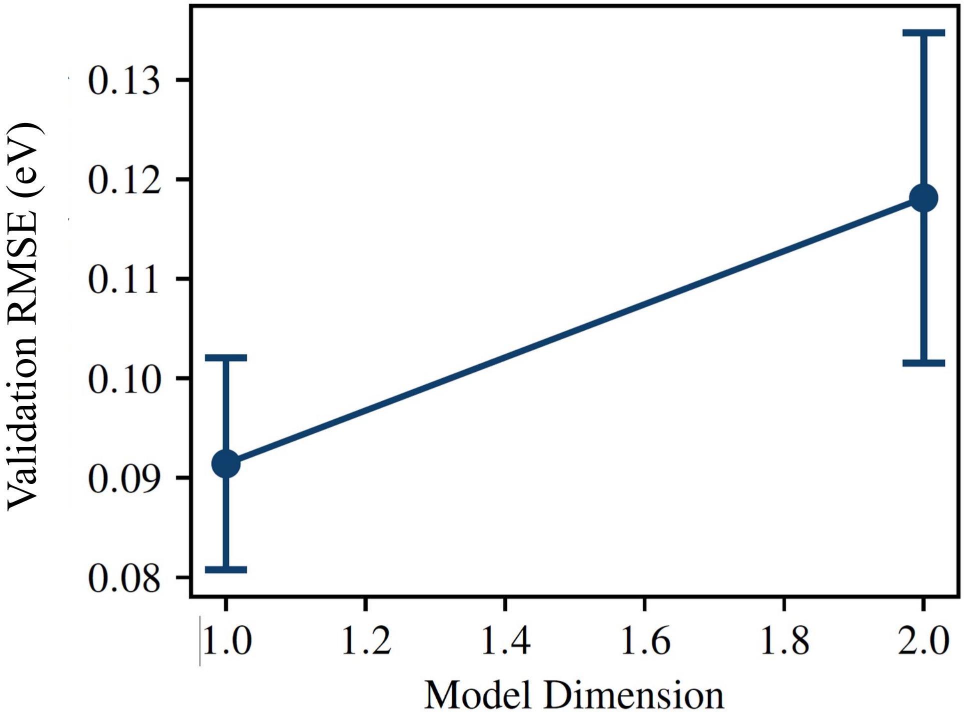
1D descriptor: $p = c_0 + a_0 \cdot \frac{\sqrt[3]{mass_{electron}} \cdot beBB \cdot homo_B \cdot IP_B}{dis_{AB} \cdot ¿ IP_A \vee ¿¿}$

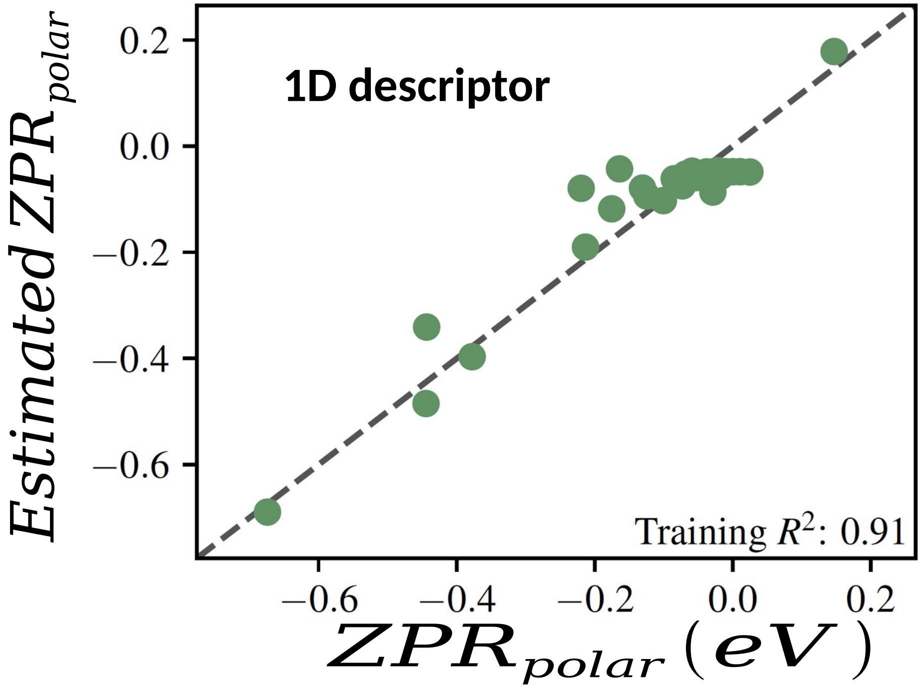
Cross-Validation for ZB structure

mean RMSE on 50 leave-10%-out cross validations

- rung = ***3***
- n_sis_select = ***10,000***

Same parameters as our regression.





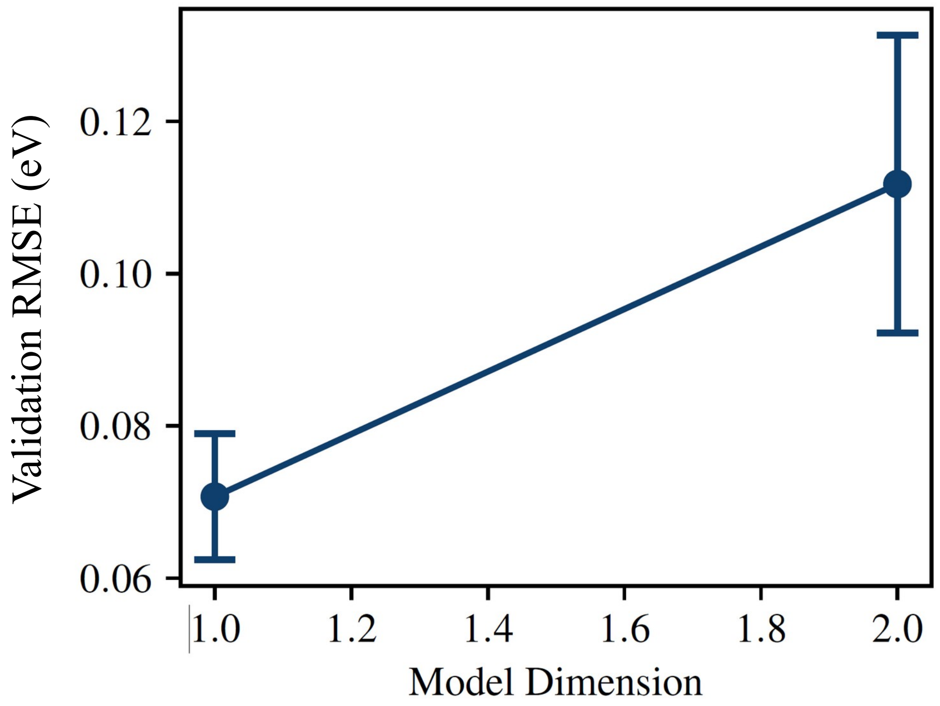
1D descriptor: $p = c_0 + a_0 \cdot \frac{mass_{hole} \cdot be_{BB} \cdot be_{AB}}{mass_B \cdot (IP \text{¿¿} A \cdot mass_B + IP_B \cdot mass_A) \text{¿}}$

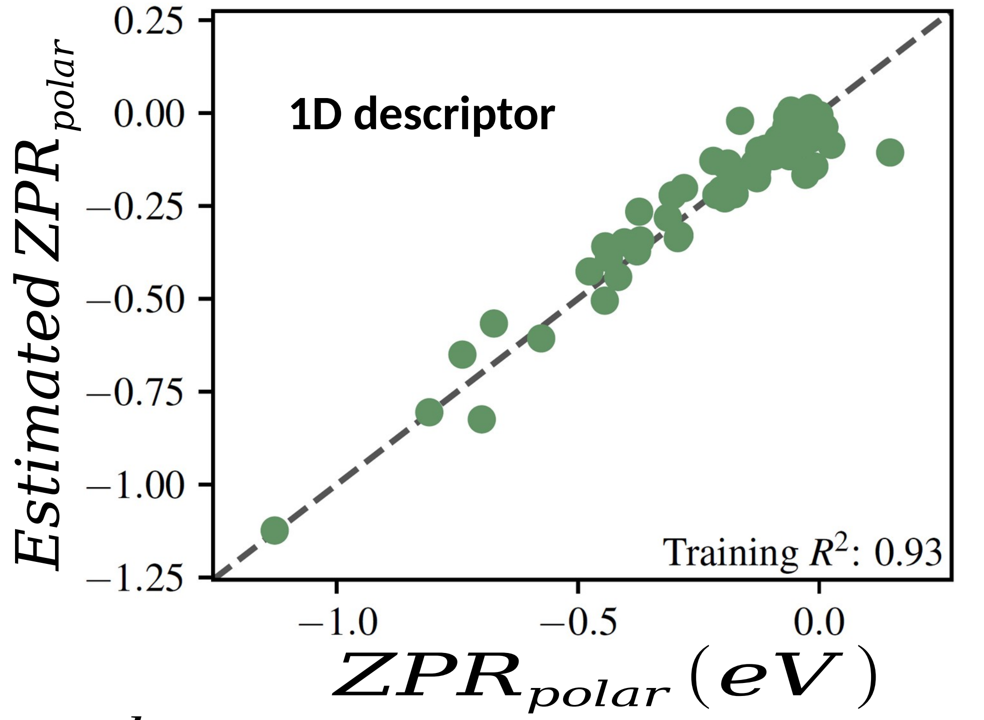
Cross-Validation for Multi-Task SISSO

mean RMSE on 50 leave-10%-out cross validations

- rung = ***3***
- n_sis_select = ***10,000***

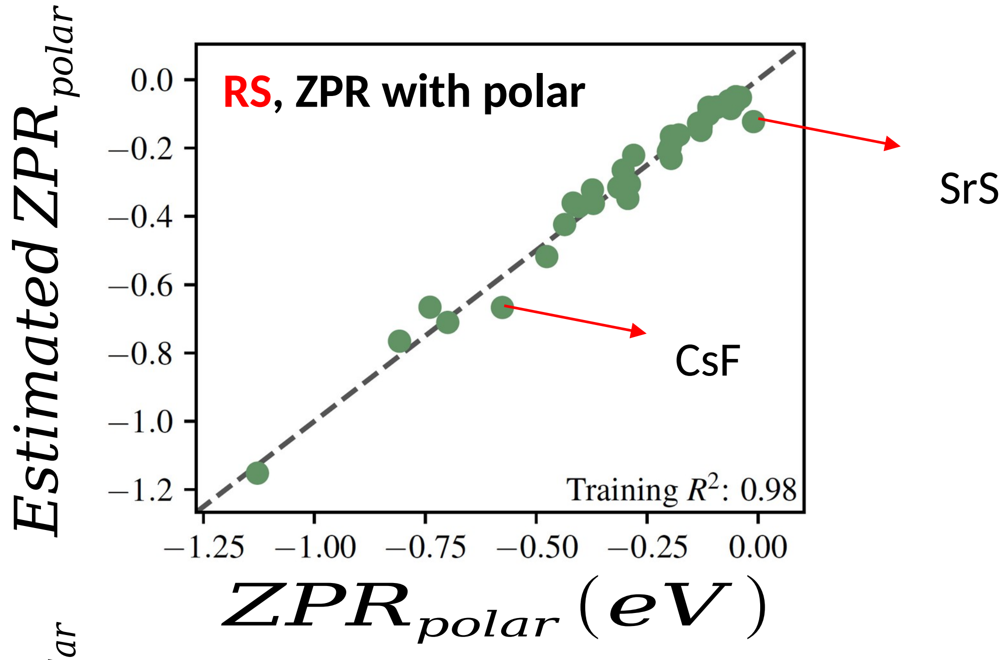
Same parameters as our regression.

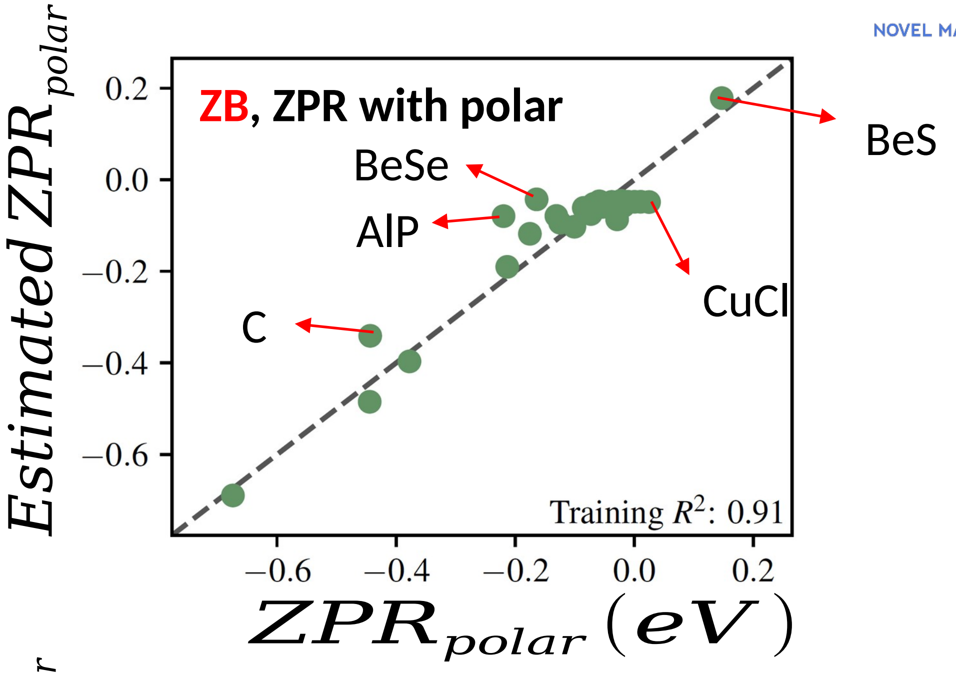


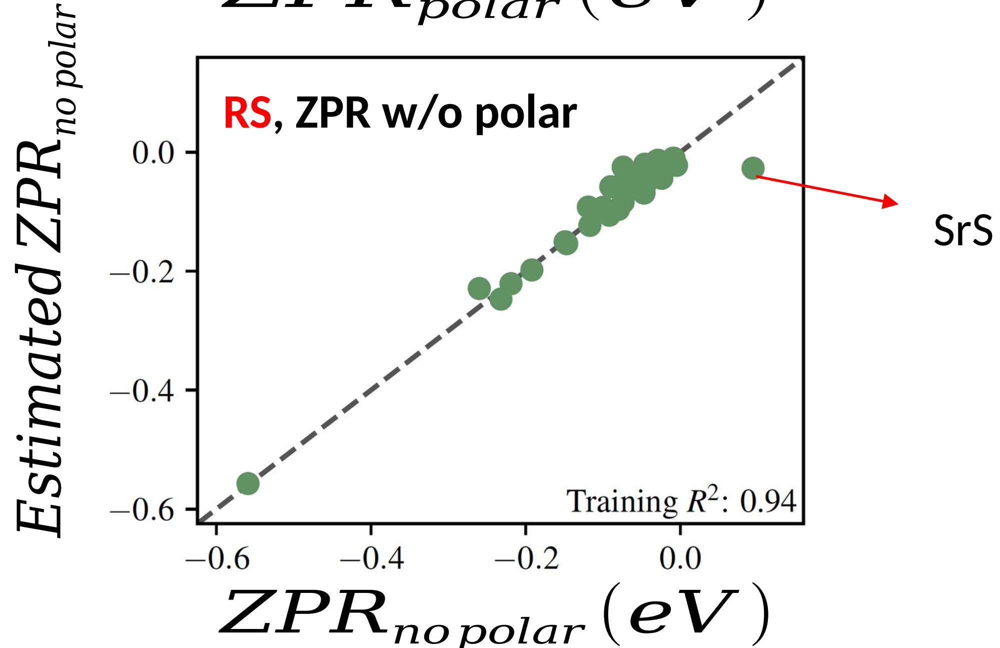


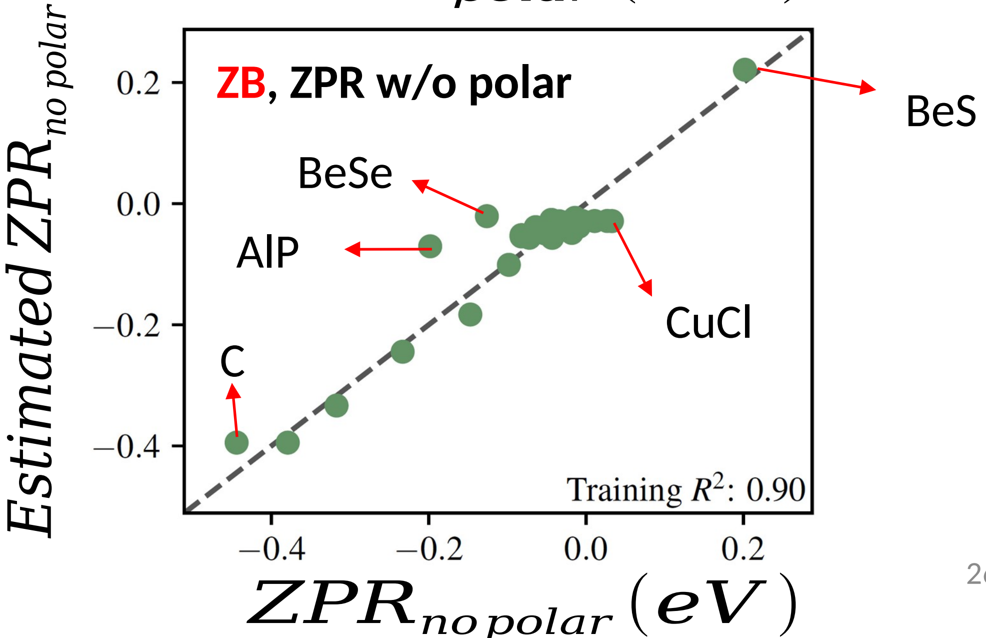
1D descriptor: $p = c_0 + a_0 \cdot \frac{{\Delta_{a-a}}^3 + \frac{homo_B}{homo_A}}{\sqrt[3]{average_{mass} \cdot rval_B \cdot homo_B}}$

Including metastable structures









Content

Introduction:

- ☐ Introduction of Zero-point renormalization
- ☐ Introduction of SISSO

Find descriptor of ZPR by SISSO:

- ☐ Focus on stable binary materials
- ☐ Including metastable structures
- ☐ Next Step

Including metastable structures

- More data points can help find better descriptor

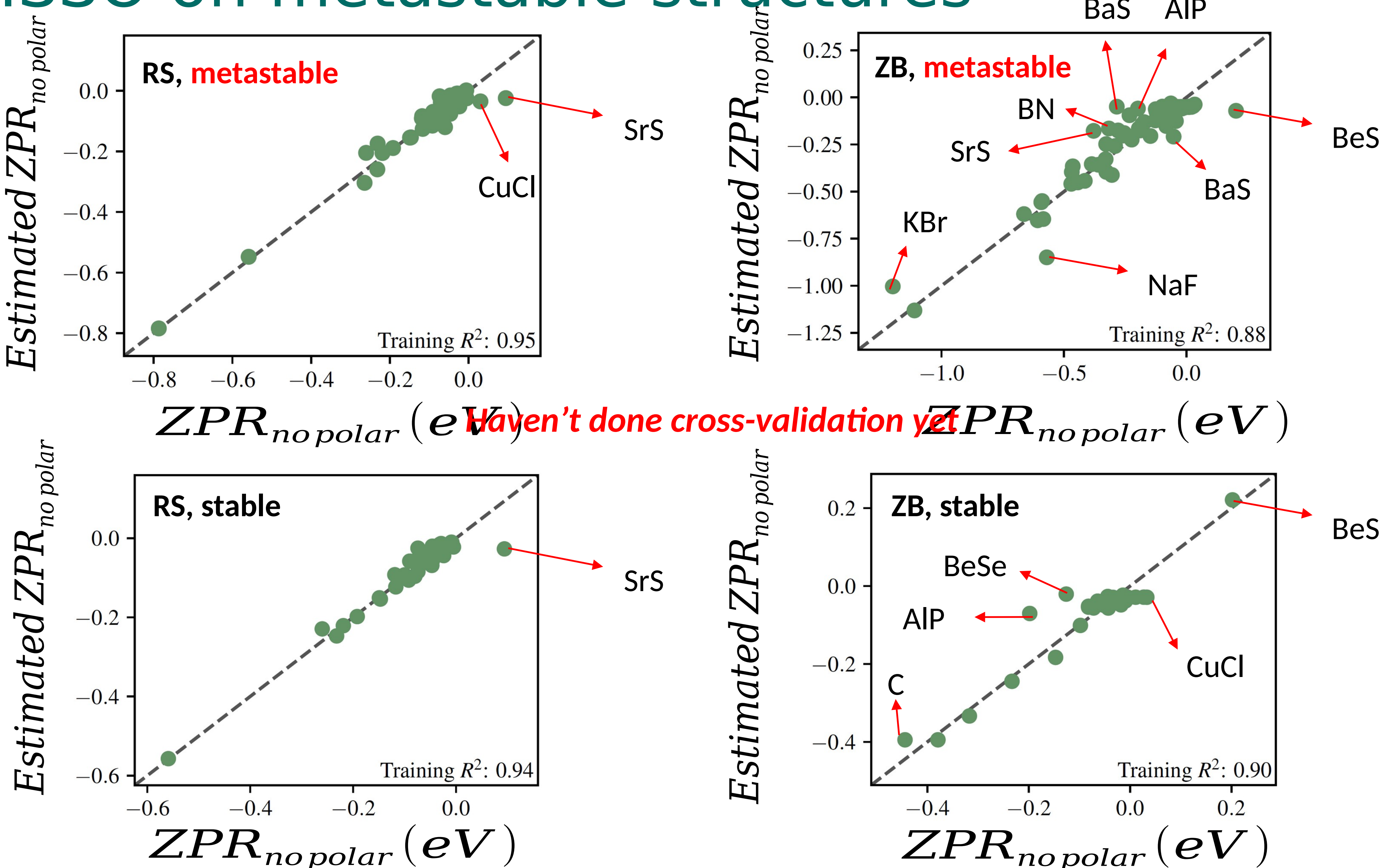
**Metastable means: 1. total energy a little higher than most stable structures.*
2. Have no imaginary phonon frequencies.

82 materilas in total → ***leave all metals out*** → ***leave unreasonable structures out (imaginary phonon frequencies)***

↓

44 RS and 72 ZB structures

SISSO on metastable structures



Next Steps

NOVEL MATERIALS DISCOVERY

- *Further Cross-validation of descriptors for metastable structures*
- *Investigate [correctness of] outliers and physical effects leading to a different behaviour.*
- *Compare with experimental and theoretical literature results [only few data available, unfortunately!] and add them to the learning set.*

Thank you !