# Building lattice models

**Jorge Íñiguez**

**Luxembourg Institute of Science and Technology (LIST)**

# Must-read references:

J. Phys.: Condens. Matter **25** (2013) 305401 (25pp) doi:10.1088/0953-8984/25/30/305401

## First-principles model potentials for lattice-dynamical studies: general methodology and example of application to ferroic perovskite oxides

Jacek C Wojdeł[1], Patrick Hermet[2,3], Mathias P Ljungberg[1], Philippe Ghosez[2] and Jorge Íñiguez[1]

PHYSICAL REVIEW B **95**, 094115 (2017)

## Efficient systematic scheme to construct second-principles lattice dynamical models

Carlos Escorihuela-Sayalero,[1] Jacek C. Wojdeł,[2] and Jorge Íñiguez[1,2]

# Very important note!

**SCALE-UP models may be analogous to pseudo-potentials in a DFT code…**

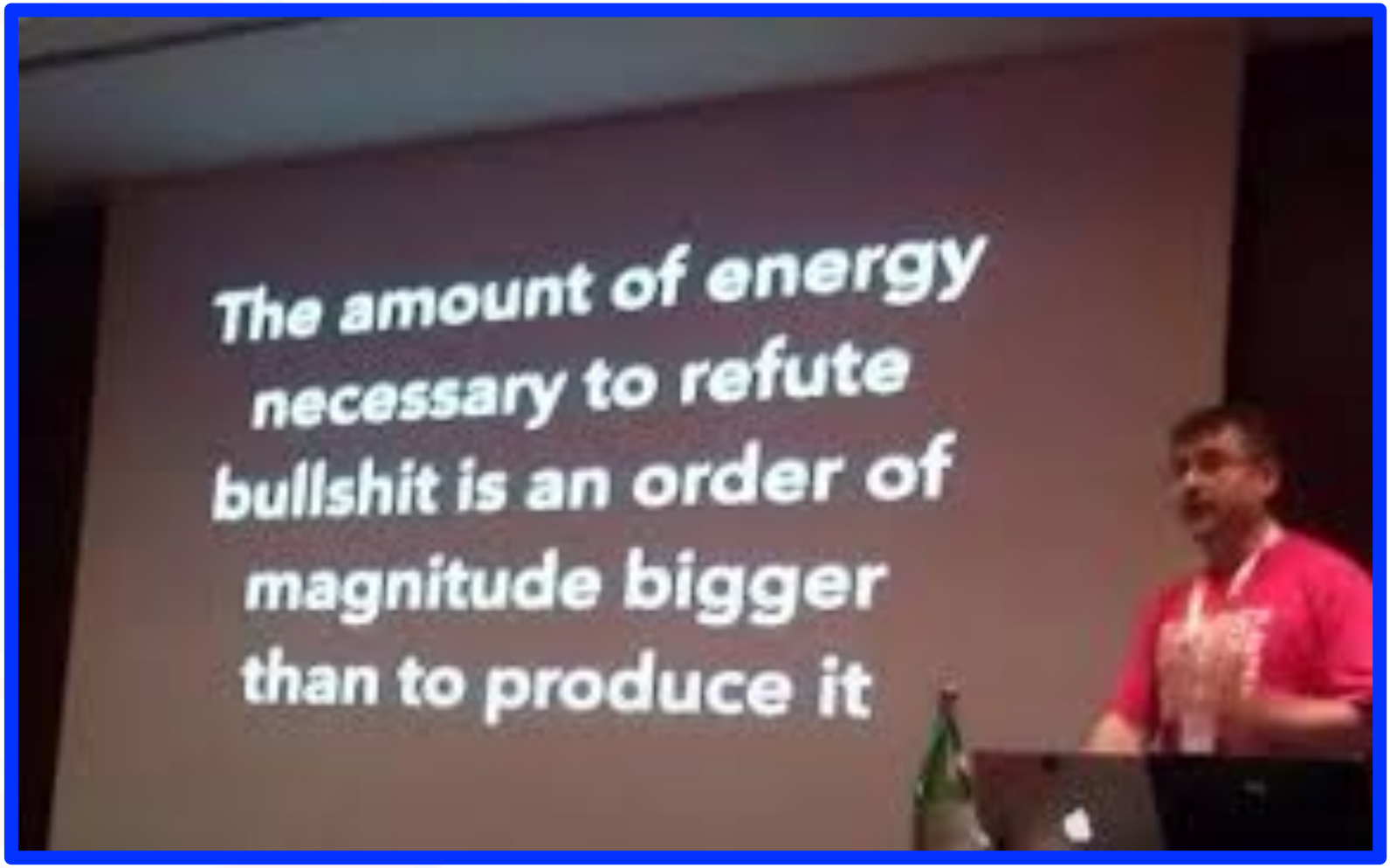
**BUT**

**They are not nearly as robust or simple!**

**SO**

**You have to know what you are doing!!**

# The bullshit axiom



# Our Goal

Compute lattice models:

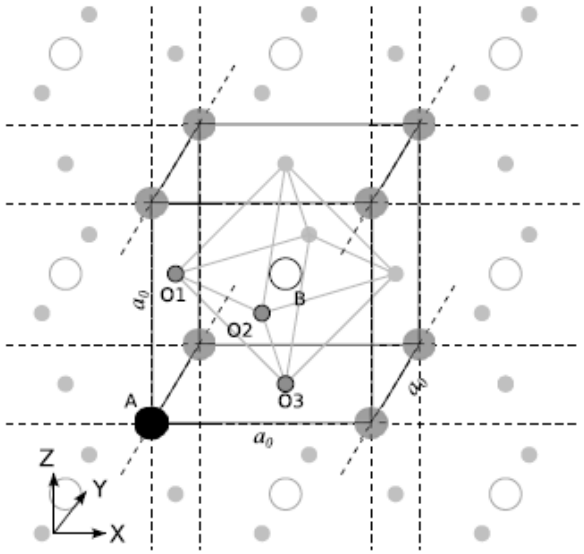
- In a way as automatic as possible
- Systematically improvable
- Without any prior knowledge
- Without any assumption
- From trivial to run FP simulations

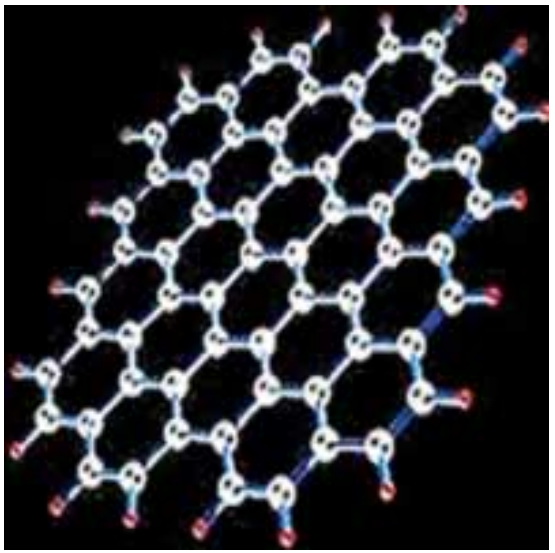
# The models

# SCALE-UP is a perturbative approach

- We assume an underlying reference structure

**RAG = Reference Atomic Geometry**



The RAG can be **anything**, from a 3D periodic lattice... to a molecule.

# The lattice energy $E^{(0)}$

$$E(\mathbf{u},\eta) = E_{\mathrm{RS}} + E_{\mathrm{p}}(\mathbf{u}) + E_{\mathrm{s}}(\eta) + E_{\mathrm{sp}}(\mathbf{u},\eta)$$

# The lattice energy $E^{(0)}$

$$E(\mathbf{u},\eta) = E_{\mathrm{RS}} + E_{\mathrm{p}}(\mathbf{u}) + E_{\mathrm{s}}(\eta) + E_{\mathrm{sp}}(\mathbf{u},\eta)$$

$$E_{\mathrm{s}}(\eta) = \frac{N}{2}\sum_{ab} C^{(2)}_{ab}\eta_a\eta_b + \frac{N}{6}\sum_{abc} C^{(3)}_{abc}\eta_a\eta_b\eta_c + \dots$$

# The lattice energy $E^{(0)}$

$$E(\mathbf{u},\eta) = E_{\mathrm{RS}} + E_{\mathrm{p}}(\mathbf{u}) + E_{\mathrm{s}}(\eta) + E_{\mathrm{sp}}(\mathbf{u},\eta)$$

$$E_{\mathrm{p}}(\mathbf{u}) = \frac{1}{2}\sum_{\substack{ijkh\\ \alpha\beta}} \widetilde{K}^{(2)}_{ij\alpha kh\beta}(u_{i\alpha}-u_{j\alpha})(u_{k\beta}-u_{h\beta})$$

$$+\frac{1}{6}\sum_{\substack{ijkhrt\\ \alpha\beta\gamma}} \widetilde{K}^{(3)}_{ij\alpha kh\beta rt\gamma}(u_{i\alpha}-u_{j\alpha})(u_{k\beta}-u_{h\beta})$$

$$\times (u_{r\gamma}-u_{t\gamma}) + \dots ,$$

$$K^{(n)}_{ijk\dots\alpha\beta\gamma\dots} = \frac{\partial^n E}{\partial u_{i\alpha}\partial u_{j\beta}\partial u_{k\gamma}\dots}$$

$$\boldsymbol{K}^{(n)} = \boldsymbol{K}^{(n),\mathrm{sr}} + \boldsymbol{K}^{(n),\mathrm{lr}}$$

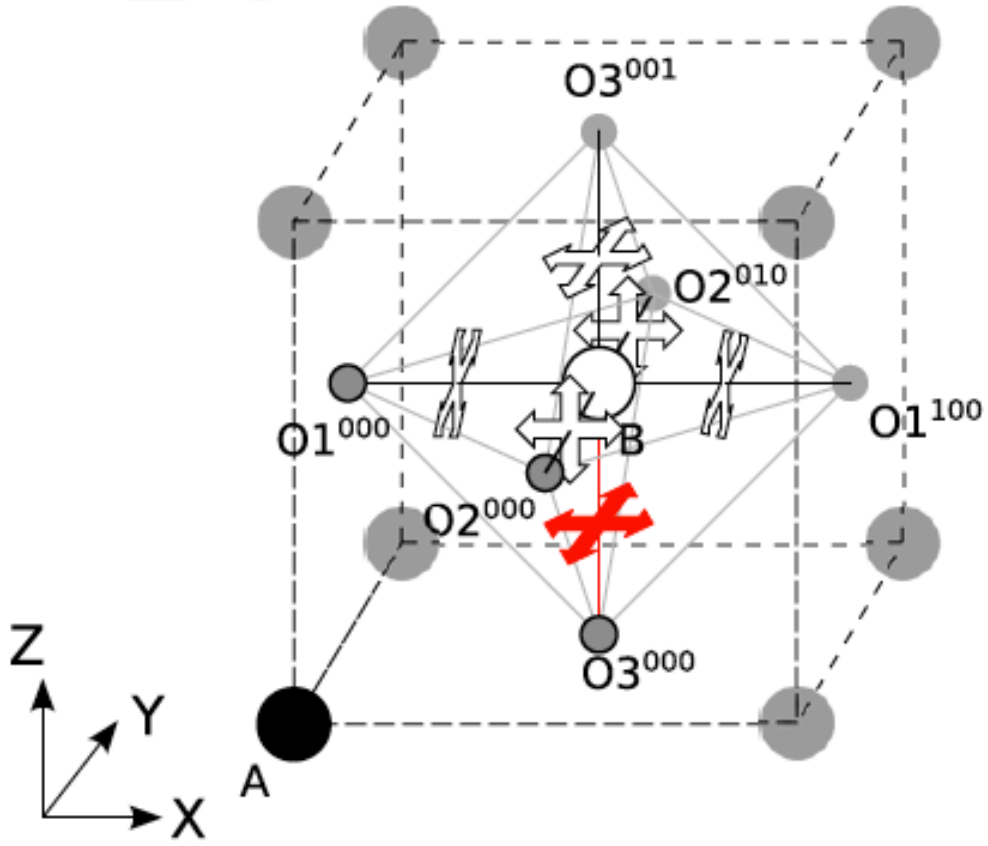
$$K^{(2),\mathrm{lr}}_{i\alpha j\beta} = \sum_{\gamma\delta} Z^*_{i\alpha\gamma} Z^*_{j\beta\delta}\left(\frac{(\epsilon_\infty^{-1})_{\gamma\delta}}{D^3} - \frac{3\Delta_\gamma\Delta_\delta}{D^5}\right)(\det\epsilon_\infty)^{-1/2},$$

# The lattice energy $E^{(0)}$

$$E(\mathbf{u},\eta) = E_{\mathrm{RS}} + E_{\mathrm{p}}(\mathbf{u}) + E_{\mathrm{s}}(\eta) + E_{\mathrm{sp}}(\mathbf{u},\eta)$$

$$E_{\mathrm{sp}}(\mathbf{u},\eta) = \frac{1}{2}\sum_{a}\sum_{ij\alpha}\tilde{\Lambda}^{(1,1)}_{aij\alpha}\eta_a(u_{i\alpha}-u_{j\alpha}) + \frac{1}{6}\sum_{a}\sum_{\substack{ijkh\\ \alpha\beta}}\tilde{\Lambda}^{(1,2)}_{aij\alpha kh\beta}\eta_a(u_{i\alpha}-u_{j\alpha})(u_{k\beta}-u_{h\beta}) + \dots, \quad (4)$$

# Symmetry-adapted terms (SATs)




$$(B_x - O3_x)^2 (B_y - O3_y)^2 \quad + \ldots$$

# Approach 1 (2013)

As much as possible, we wanted to compute the terms, not fit.

# Details (I): Harmonic terms of the Hamiltonian

- Phonon term (force constant matrix):

$$ {}^{2}K_{ij\alpha\beta}u_{i\alpha}u_{j\beta} \rightarrow {}^{2}K = {}^{2}K^{dd} + {}^{2}K^{sr} $$

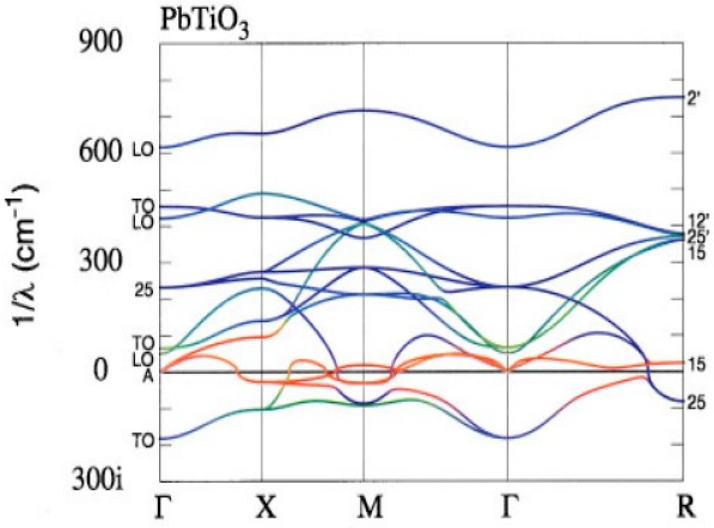
- Strain term (elastic tensor):

$$ {}^{2}C_{lm}\eta_{l}\eta_{m} $$

- Strain-phonon coupling:

$$ {}^{11}\Lambda_{li\alpha}\eta_{l}u_{i\alpha} $$

[ zero by symmetry in cubic perovskites ]




- All these terms are directly provided by DFT codes like ABINIT

# Details (II): Strain-phonon couplings

$$ {}^{12}\Lambda_{lij\alpha\beta}\eta_l u_{i\alpha}u_{i\beta} + {}^{22}\Lambda_{lmij\alpha\beta}\eta_l\eta_m u_{i\alpha}u_{i\beta} $$

Strategy:

- Consider reference structure under strain ($\eta$ = ± 2%)
- Compute force constant matrix $K'$

$$ K'_{ij\alpha\beta} = {}^{2}K_{ij\alpha\beta} + {}^{12}\Lambda_{lij\alpha\beta}\eta_l + {}^{22}\Lambda_{lmij\alpha\beta}\eta_l\eta_m $$

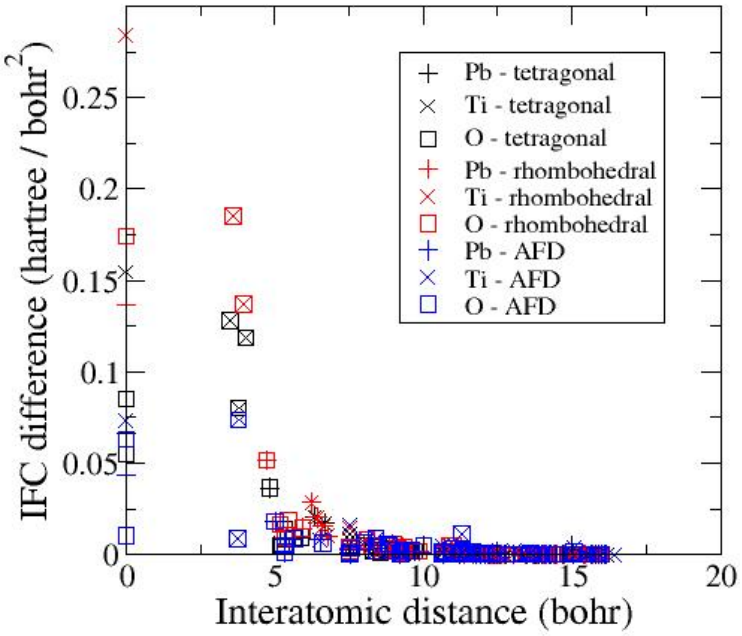
- Traditionally, $H_{eff}$'s have included only the ${}^{12}\Lambda$ term
- In $PbTiO_3$, relevant strains go up to 6% ⇨ higher order terms needed...

# Details (III): Anharmonic phonon-phonon

$$ {}^{3}K_{ijk\alpha\beta\gamma}u_{i\alpha}u_{j\beta}u_{k\gamma} + {}^{4}K_{ijkh\alpha\beta\gamma\delta}u_{i\alpha}u_{j\beta}u_{k\gamma}u_{h\delta} $$

Our choices for first models:

- Only pair-wise (two-body) terms
- Restricted to 1st nearest neighbors




We compute the anharmonic terms to get a good description of:

- Structure of low-symmetry phases
- Energy differences between phases
- Selected eigenmodes and eigenvalues of $K'$ for key low-symmetry phases

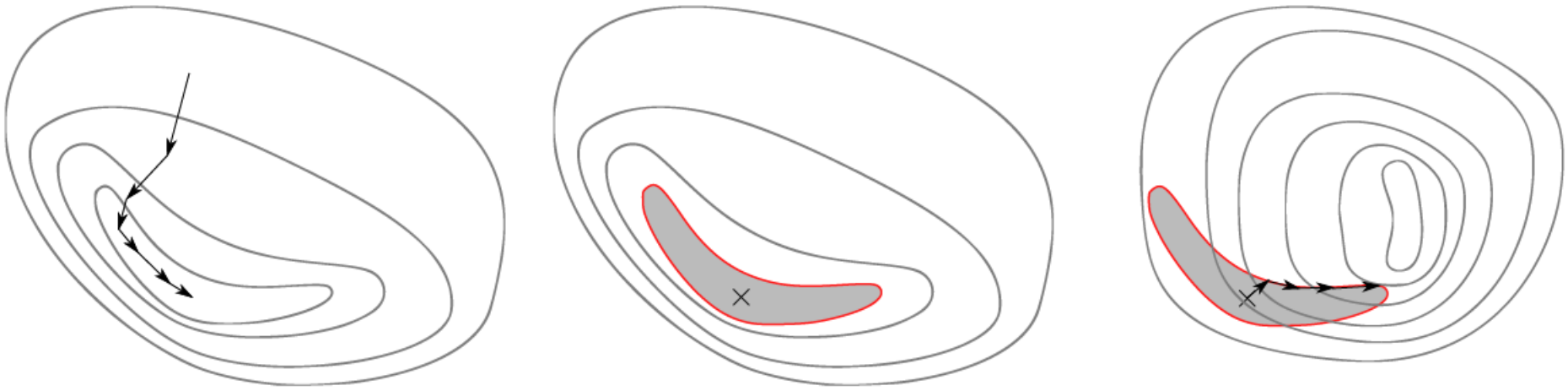
minimize $\left\| K'\xi_{qs} - \kappa_{qs}\xi_{qs} \right\|$ to get good energetics of FE and AFD dists.

# Hand-picked terms

| | Third-order A–O terms |
|---|---|
| 1 | $(A_z–O1_z)^3$ |
| 2 | $(A_z–O3_z)^2\,(A_x–O3_x)$ |
| 3 | $(A_x–O2_x)^2\,(A_z–O2_z)$ |
| | **Fourth-order A–O terms** |
| 4 | $(A_y–O2_y)^4$ |
| 5 | $(A_z–O2_z)^4$ |
| 6 | $(A_z–O3_z)^2\,(A_y–O3_y)^2$ |
| 7 | $(A_z–O2_z)^2\,(A_x–O2_x)^2$ |
| 8 | $(A_y–O1_y)^3\,(A_z–O1_z)$ |
| 9 | $(A_x–O1_x)^2\,(A_y–O1_y)\,(A_z–O1_z)$ |
| | **Third-order B–O terms** |
| 10 | $(B_x–O1_x)^3$ |
| 11 | $(B_z–O1_z)^2\,(B_x–O1_x)$ |
| | **Fourth-order B–O terms** |
| 12 | $(B_x–O1_x)^4$ |
| 13 | $(B_z–O1_z)^4$ |
| 14 | $(B_x–O3_x)^2\,(B_z–O3_z)^2$ |
| 15 | $(B_x–O3_x)^2\,(B_y–O3_y)^2$ |

| | A–O terms |
|---|---|
| 1 | $(A_y–O2_y)^2\eta_{yy}$ |
| 2 | $(A_z–O2_z)^2\eta_{zz}$ |
| 3 | $(A_z–O3_z)^2\eta_{yy}$ |
| 4 | $(A_x–O2_x)^2\eta_{xx}$ |
| 5 | $(A_z–O1_z)^2\eta_{yy}$ |
| 6 | $(A_x–O1_x)^2\eta_{zy}$ |
| 7 | $(A_y–O1_y)^2\eta_{zy}$ |
| 8 | $(A_x–O2_x)(A_z–O2_z)\eta_{yy}$ |
| 9 | $(A_y–O2_y)(A_x–O2_x)\eta_{yx}$ |
| 10 | $(A_z–O3_z)(A_y–O3_y)\eta_{xz}$ |
| 11 | $(A_x–O3_x)(A_y–O3_y)\eta_{yy}$ |
| 12 | $(A_y–O3_y)(A_x–O3_x)\eta_{yx}$ |
| | **B–O terms** |
| 13 | $(B_z–O3_z)^2\eta_{zz}$ |
| 14 | $(B_y–O3_y)^2\eta_{zz}$ |
| 15 | $(B_y–O2_y)^2\eta_{xx}$ |
| 16 | $(B_z–O2_z)^2\eta_{zz}$ |
| 17 | $(B_x–O2_x)^2\eta_{zz}$ |
| 18 | $(B_x–O3_x)(B_y–O3_y)\eta_{xy}$ |
| 19 | $(B_z–O1_z)(B_x–O1_x)\eta_{zx}$ |

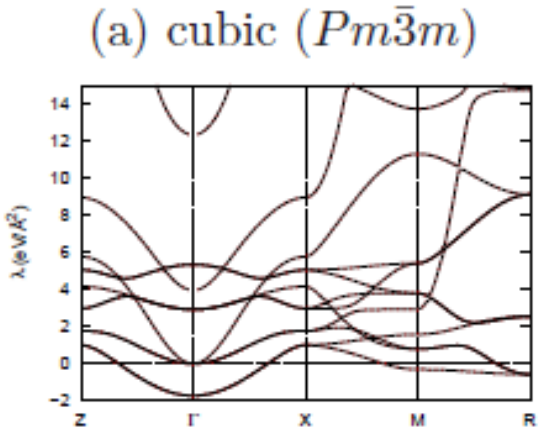
# Successive constrained optimization

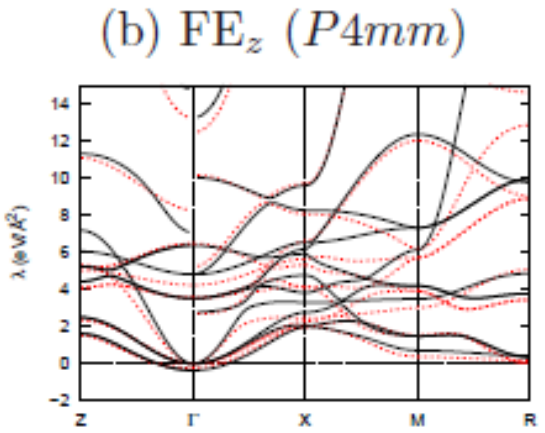


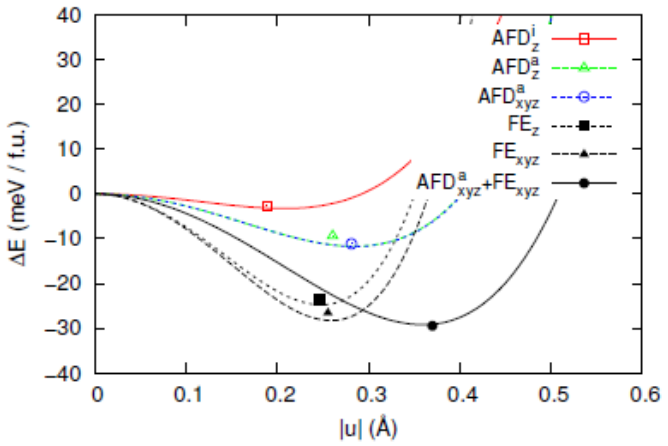
1 2 3

1. Fit energies
2. Energies OK
3. Fit structures while maintaining good accuracy for energies

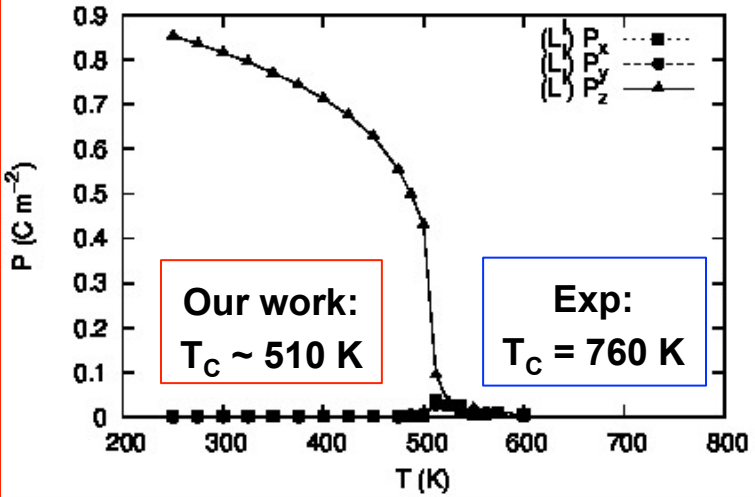
# Example of application to $PbTiO_3$




| Structure | Method | $u^{\Gamma}_{Pbz}$ | $u^{\Gamma}_{Tiz}$ | $u^{\Gamma}_{O1z}$ | $u^{\Gamma}_{O3z}$ | $O_6$ rot. | Energy |
|---|---|---|---|---|---|---|---|
| FE$_z$ ($P4mm$) | LDA | 0.179 | 0.072 | −0.104 | −0.043 | — | −23.7 |
| | Model | 0.180 | 0.073 | −0.105 | −0.043 | — | −24.8 |
| FE$_{xyz}$ ($R3m$) | LDA | 0.104 | 0.048 | −0.063 | −0.027 | — | −26.6 |
| | Model | 0.105 | 0.049 | −0.063 | −0.028 | — | −28.3 |
| AFD$_z^a$ ($I4/mcm$) | LDA | — | — | — | — | 5.4 | −9.4 |
| | Model | — | — | — | — | 5.9 | −11.7 |
| AFD$_{xyz}^a$ ($R\bar{3}c$) | LDA | — | — | — | — | 3.4 | −11.2 |
| | Model | — | — | — | — | 3.4 | −11.7 |
| AFD$_z^i$ ($P4/mbm$) | LDA | — | — | — | — | 3.9 | −2.7 |
| | Model | — | — | — | — | 4.3 | −3.3 |
| FE$_{xyz}$+AFD$_{xyz}^a$ ($R3c$) | LDA | 0.096 | 0.047 | −0.058 | −0.026 | 2.8 | −29.5 |
| | Model | 0.098 | 0.047 | −0.060 | −0.026 | 2.1 | −29.5 |

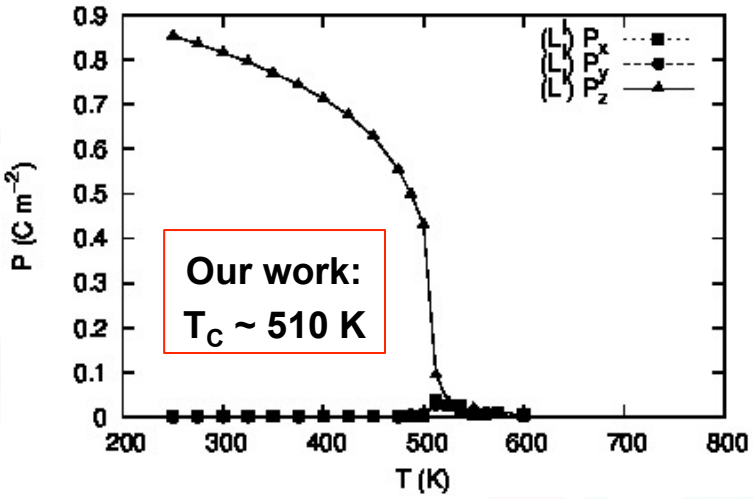


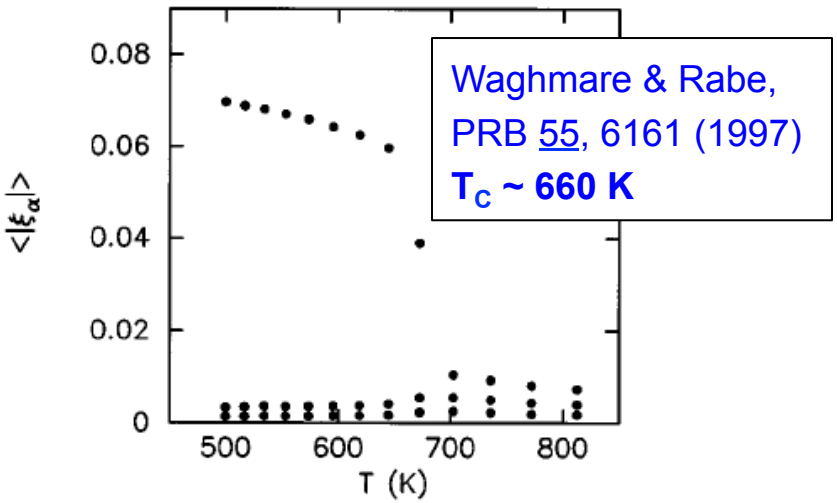

Displacements: Angstrom ; Rotations: degrees ; Energies: meV/f.u.

**Excellent accuracy reproducing first-principles data**

**Wojdeł, Hermet, Ljungberg, Ghosez & Íñiguez, JPCM 25, 305401 (2013)**

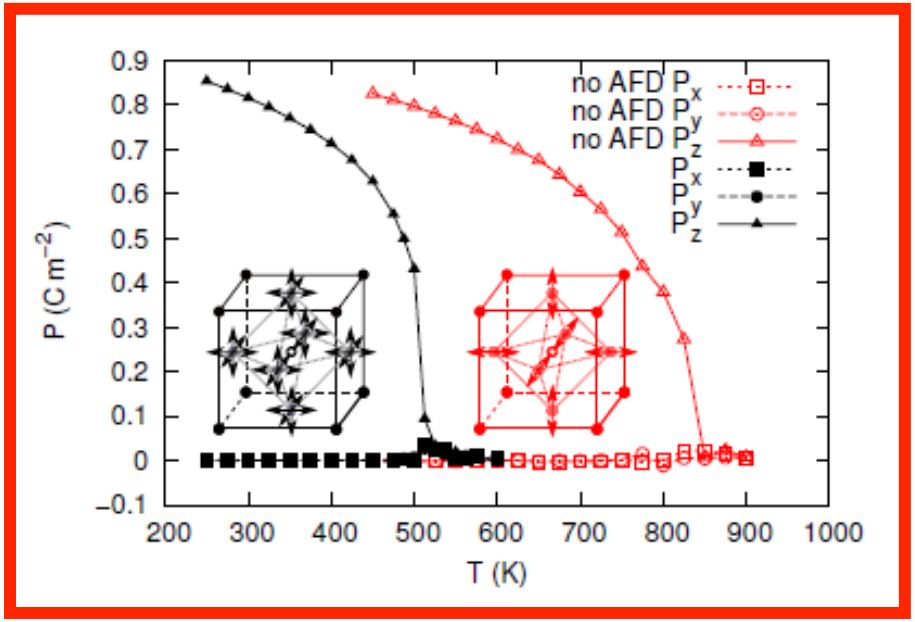
# Competing instabilities in $PbTiO_3$



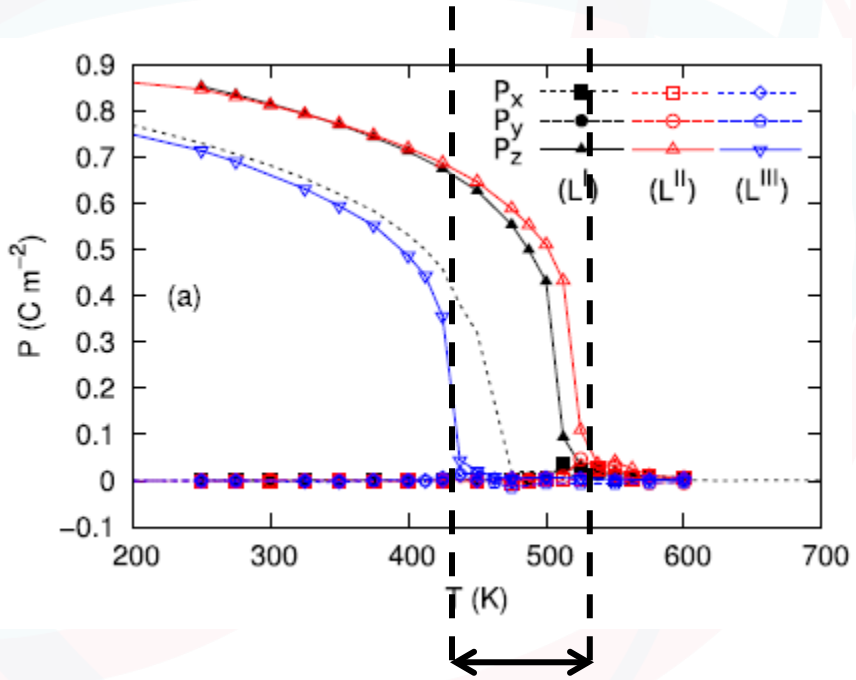

Experimental $T_C$ : 760 K

- Our energy surface for the polar distortions is very similar to the one of Waghmare and Rabe
- NOTE: they only included polar distortions, leaving out everything else!
- In particular, they left out the $O_6$-octahedral rotations that are known to compete with the polar distortions...
- We can remove the $O_6$-rots by hand...



# More on the $T_C$ issue

| Method | $a$ | $c$ | $c/a$ | $u_{Pb_z}$ | $u_{Ti_z}$ | $u_{O1_z}$ | $u_{O3_z}$ | Energy |
|---|---|---|---|---|---|---|---|---|
| LDA | 3.864 | 3.974 | 1.029 | 0.230 | 0.106 | −0.133 | −0.071 | −37.7 |
| model $L^0$ | 3.908 | 3.987 | 1.020 | 0.200 | 0.103 | −0.122 | −0.060 | −34.5 |
| model $L^{I}$ | 3.863 | 3.968 | 1.027 | 0.220 | 0.099 | −0.128 | −0.063 | −39.9 |
| model $L^{II}$ | 3.861 | 3.978 | 1.030 | 0.227 | 0.102 | −0.132 | −0.066 | −43.1 |
| model $L^{III}$ | 3.856 | 3.968 | 1.029 | 0.221 | 0.098 | −0.128 | −0.062 | −39.9 |



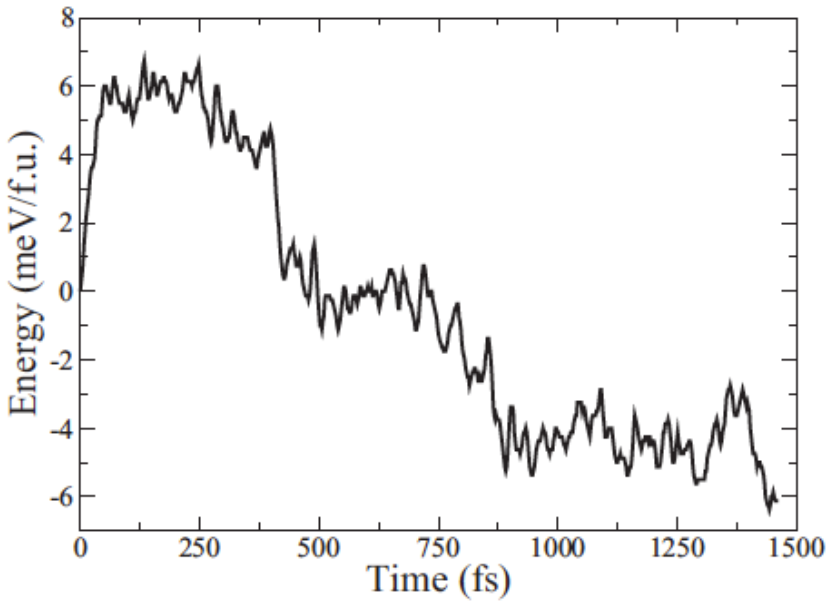

- Computed $T_C$ for various models of similar quality
- A priori negligible differences resulted in shifts of about 100 K in $T_C$ → Sensitivity is much stronger than expected.
- Extremelly challenging to get the experimental $T_C$: (1) need very complete models, (2) very accurate first-principles calculations.

# Approach 2 (2017)

**We fit everything.** Strongly recommended!!

# Trivial training set

- Molecular dynamics simulations at various temperatures
- Typically, we fit at 10 K and 300 K
- We use 200 K and 500 K as test sets




Details of the simulation ($SrTiO_3$):

- **40-atom cell**
- LDA potential
- Simulations start from the cubic phase
- NpT ensemble

# Best possible model

- 4th order
- Short-range interactions confined to 2x2x2 cell
- Up to 4-body terms

We then generate all the symmetry-allowed polynomial terms compatible with these cutoffs. For $E_p$, we obtain 45 harmonic terms, 79 third-order ones, and 275 fourth-order ones; for $E_s$, we obtain three terms, and 161 terms for $E_{sp}$. Note that these terms only depend on the structure and symmetry of the ideal cubic perovskite phase; hence they can be applied to the study of any such material, not only STO.

Total: ~ 500+ independent terms to be fitted

**How do we choose the most important ones??**

# New parameter-fitting method

$$GF[\boldsymbol{C}] = \frac{1}{M_1}\sum (F_i^{MODEL}[\boldsymbol{C}] - F_i^{TS})^2 + \frac{1}{M_2}\sum (\sigma_i^{MODEL}[\boldsymbol{C}] - \sigma_i^{TS})^2$$

- Goal function: *mean square error* in **forces and stresses.**
- It a is **convex, separable** and **strictly positive** function and, thus, it has exactly **one** minimum.

$$\frac{\partial GF[\boldsymbol{C}]}{\partial C_j} = 0 \rightarrow$$ **Linear** system of *N* equations →

- → Exact best solution
- → Faster than any minimization method
- → Scales as $N^{2.5}$

*Courtesy of Carlos Escorihuela-Sayalero*

# EMPHASIS is due here:

**The energy is linear in the fitting parameters**

**This is an outstanding advantage that our models have !!!**

We did not choose them for this reason, but this is probably the best reason to choose them.

# This allows a fast model optimization

$$\{C_1,\ldots,C_N\} \longrightarrow M_1$$

$$M_1 + \{C_1,\ldots,C_N\} \setminus \{C_i\} \longrightarrow M_2$$

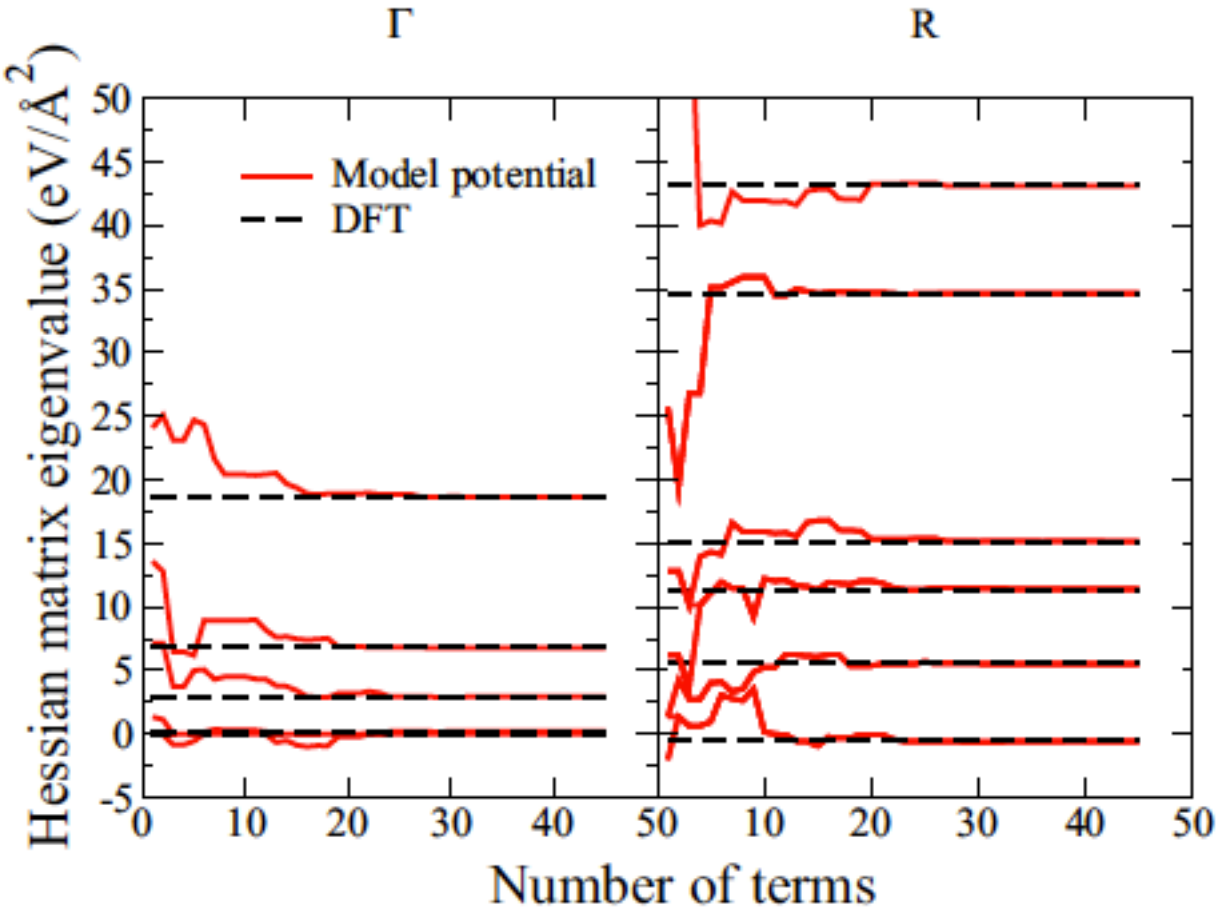
$$M_2 + \{C_1,\ldots,C_N\} \setminus \{C_i,C_j\} \longrightarrow M_3$$

$$\vdots$$

$$M_{n-1} + \{C_1,\ldots,C_N\} \setminus \underbrace{\{C_i,C_j,\ldots\}}_{n-1} \longrightarrow M_n$$

$M_n$ = best model of n parameters

# 1st test of the method: exact harmonic model



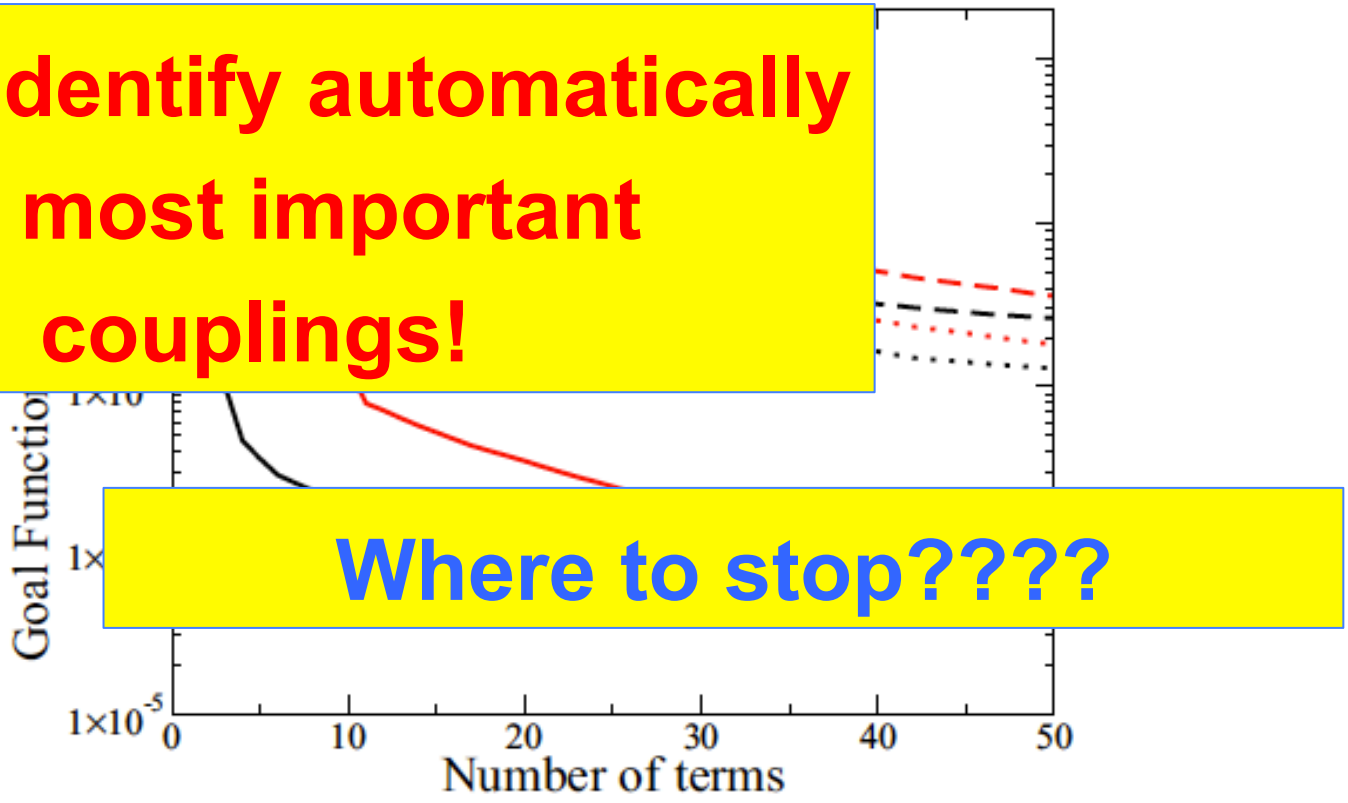
# Let's go for the real thing!

TABLE I. List of the STO models constructed in this work (see text). The number of parameters in the optimum cross-validated models is indicated in the last column. For the EHMs, we also indicate that we retain 45 harmonic terms in $E_p$, i.e., all the independent interactions within our $2 \times 2 \times 2$ supercell.

| Number | Fit type | TS | Parameters retained |
|---|---|---|---|
| 1 | FM | TS@10 | 33 |
| 2 | EHM | TS@10 | 45+10 |
| 3 | FM | TS@300 | 37 |
| 4 | EHM | TS@300 | 45+17 |
| 5 | FM | TS@10+300 | 44 |
| 6 | EHM | TS@10+300 | 45+14 |

# Families of models of increasing accuracy

**We can identify automatically the most important couplings!**

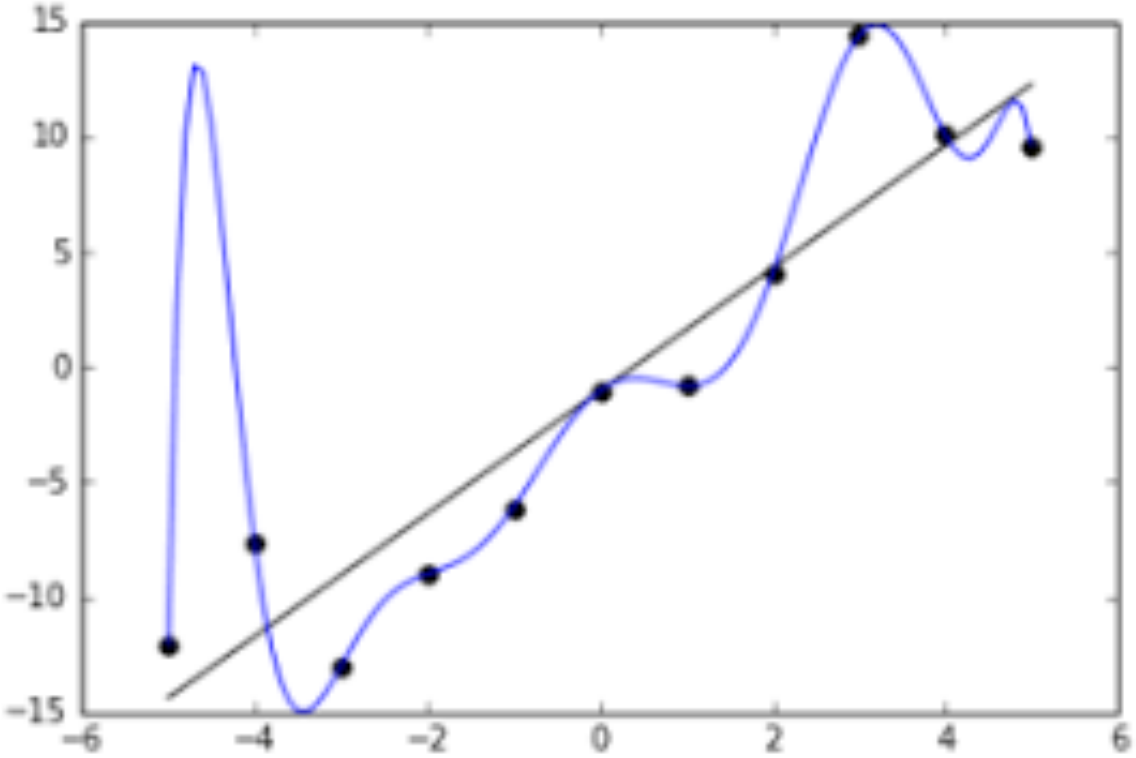



**Where to stop????**

FIG. 5. Optimized goal function, as a function of the number of terms, for the various models considered in this work. Red and black lines correspond to FM and EHM models, respectively. Solid, dashed, and dotted lines correspond to models fitted to the TS@10, TS@300, and TS@10+300, respectively. See text for details.

# OMG, my perfect model is not predictive!!

## The curse of overfitting



# Classic & non-classic cross-validation

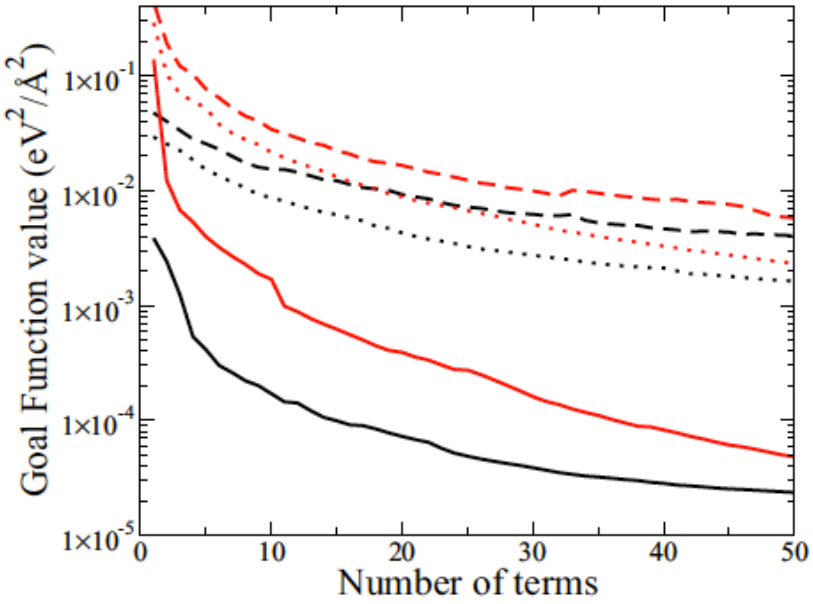



FIG. 6. Behavior of the goal function evaluated using fitted models and test-set configurations, as described in the text. The colors and line types are as in Fig. 5.

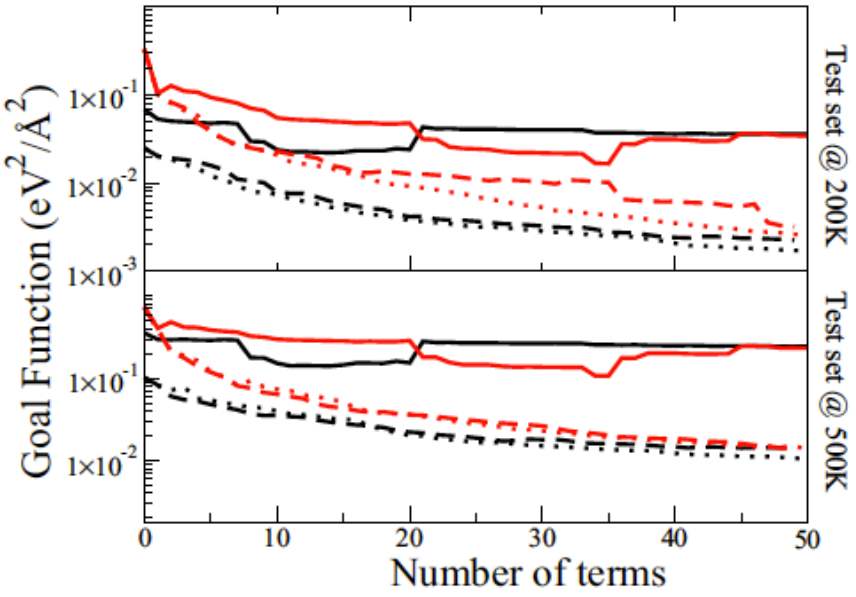



FIG. 7. Behavior of the goal function evaluated using fitted models and test-set configurations, as described in the text. The colors and line types are as in Fig. 5.

# Our practical cross-validation criterion

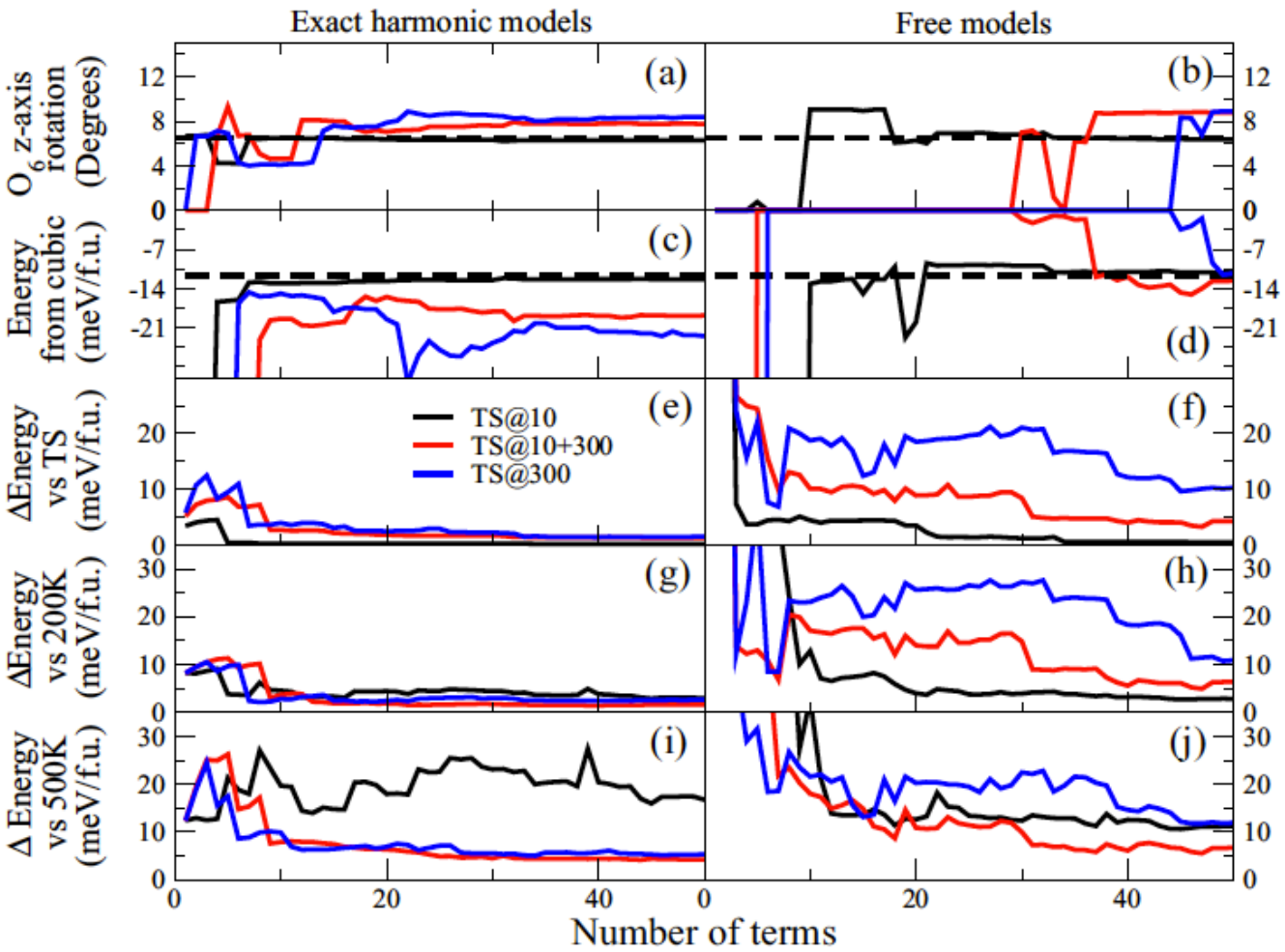
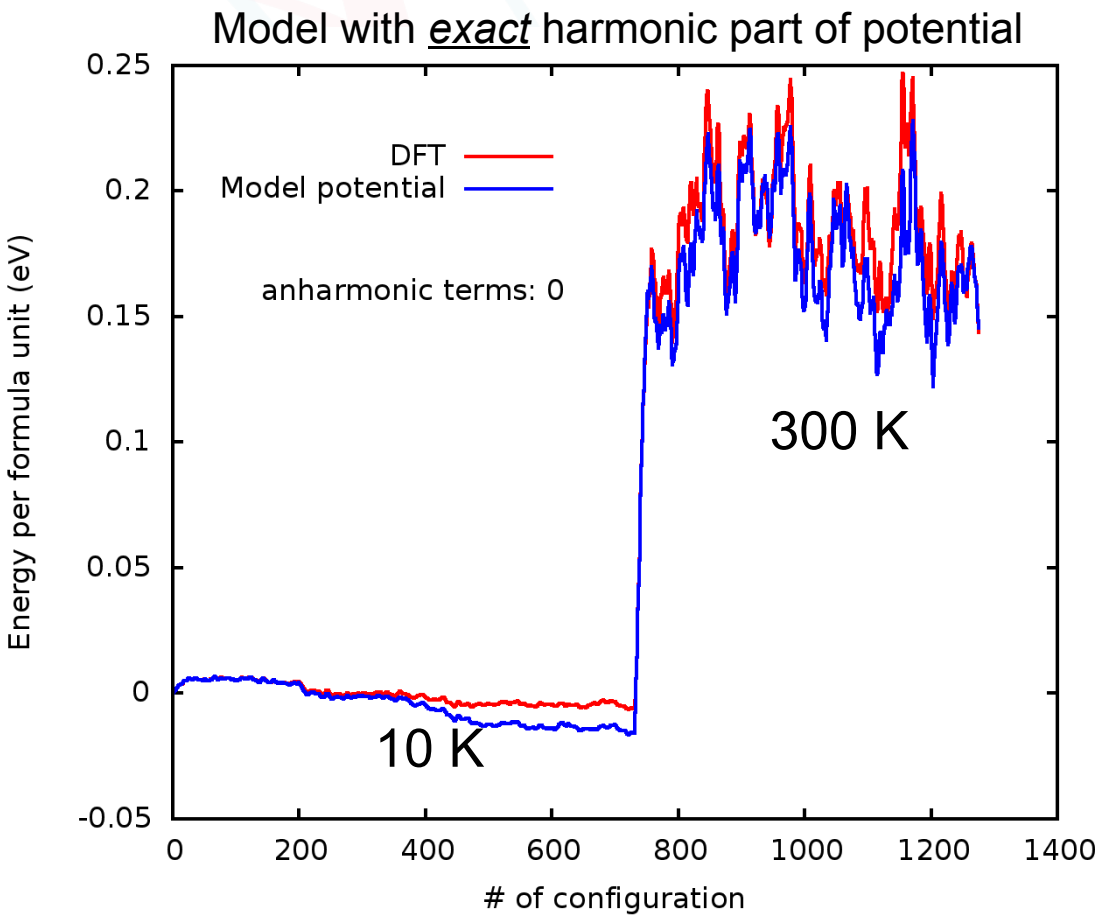



FIG. 8. Performance of the six models discussed in this work as regards the prediction of various ground state [(a)–(d)] and MD [(e)–(j)] related quantities. To obtain the data in (e)–(j), we use our model potentials to compute the energy of the configurations visited in the MD trajectories computed from DFT at various temperatures; we report the average difference between the DFT energy and that obtained from the model.

# The cross-validated models are great !!!

$SrTiO_3$ simulated in a 2x2x2 (40 atom) box




$\Delta E$

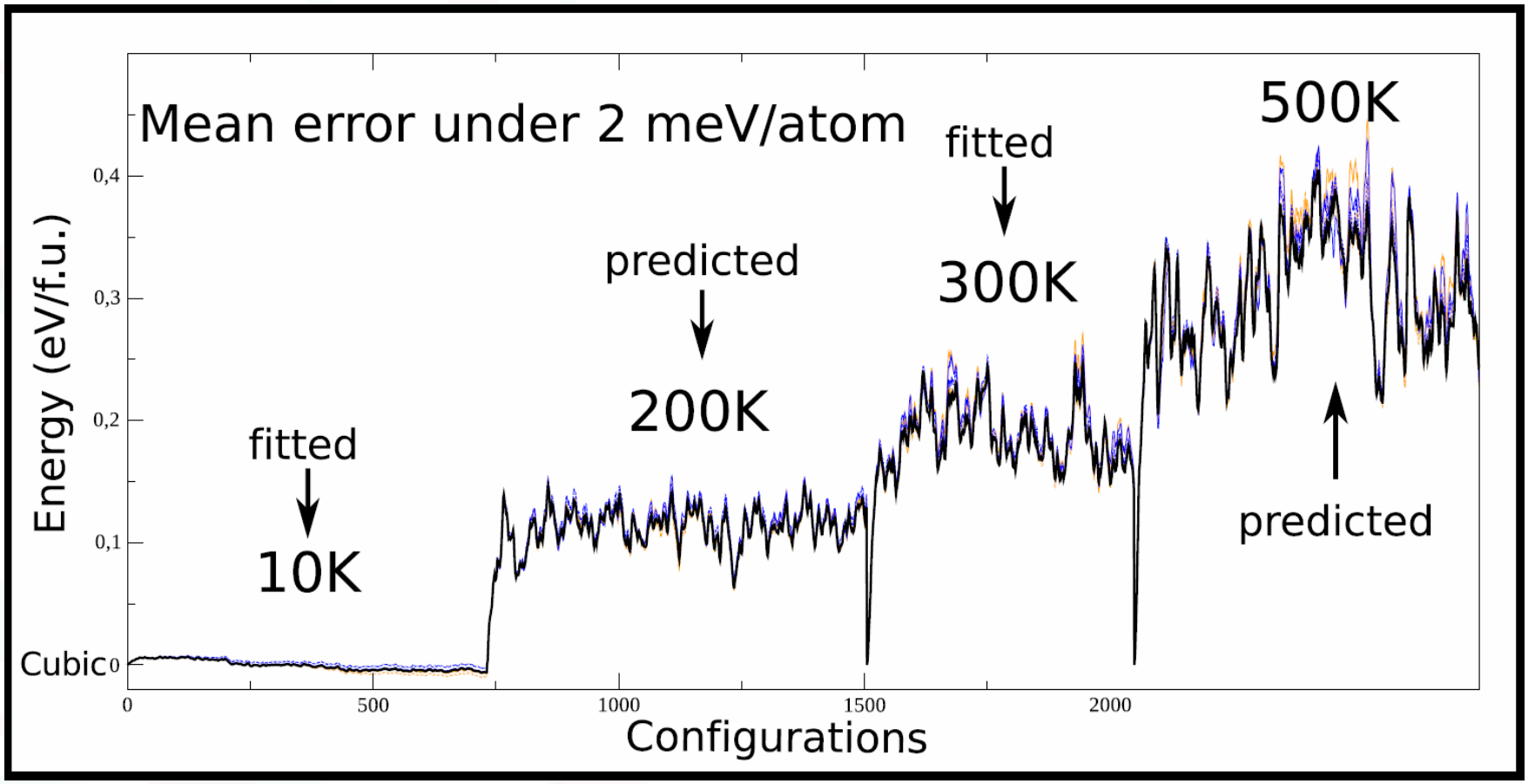
@ 10 K

0.6 meV/fu

@ 300 K

1.5 meV/fu

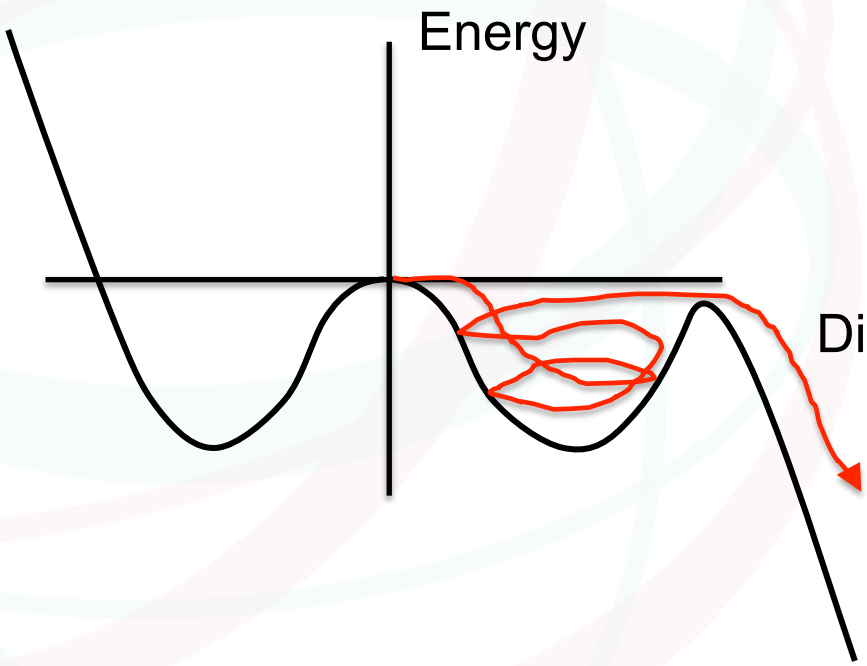
# The cross-validated models are great !!!




**but we cannot declare victory yet…**

# The one weakness of this method

Hey! Let's run a Monte Carlo temperature annealing to find the ground state! ☺




**Unbounded** direction of the PES

**Ups!**

# Dealing with unboundedness

## II. FORMALISM

### D. Energy boundedness

## III. APPLICATION TO $SrTiO_3$

### G. Energy boundedness

#### *1. Quasi-automatic approaches*

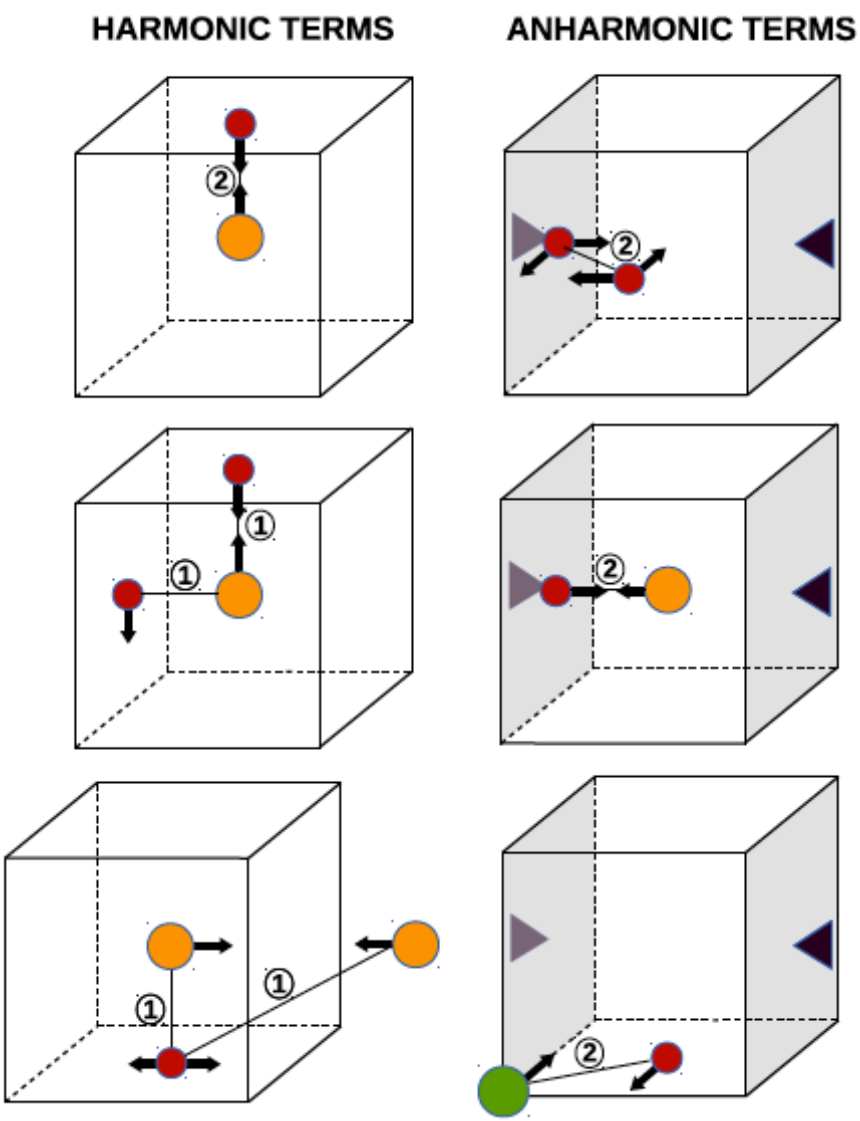
#### *2. Practical approach*

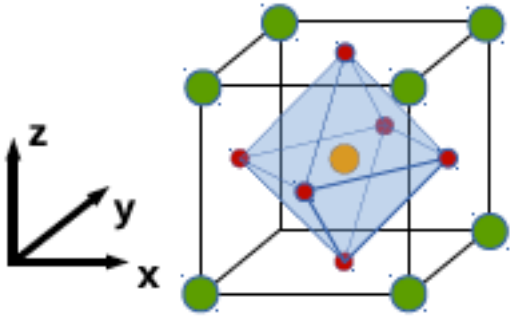
**We have a way to correct the problem that works very well!**

TABLE II. Properties characterizing the behavior of the optimum models that we consider for MC simulations. We give the values corresponding to the models modified so that the energy is bounded from below, as well as (in parenthesis) those obtained from the original models. In the last row, $\Delta E$ is the average energy difference, for the configurations included in the corresponding TS, between the LDA results and the energies evaluated with our potentials.

| | FM TS@10 | EHM TS@10 | FM TS@10+300 | EHM TS@10+300 | LDA |
|---|---|---|---|---|---|
| $O_6$ rotation (degrees) | 6.5 (6.5) | 6.5 (6.5) | 9.9 (8.7) | 8.3 (8.1) | 6.5 |
| Ground state energy (meV/f.u.) | −11.1 (−11.1) | −13.2 (−13.1) | −13.2 (−11.9) | −22.9 (−20.5) | −11.7 |
| Goal function ($eV^2$ $Å^{-2}$ $\times 10^{-3}$) | 0.14 (0.14) | 0.18 (0.19) | 2.79 (2.81) | 4.94 (5.04) | |
| $\Delta E$ (meV/f.u.) | 0.58 (1.45) | 0.30 (0.27) | 4.74 (4.78) | 2.31 (2.57) | |

# The physical meaning of our models




Strontium

Titanium

Oxygen

Surely, the obtained dominant terms mean something!

# The physical meaning of our models

Surely, the obtained dominant terms mean something! …

**… but let's not get too attached to them!**

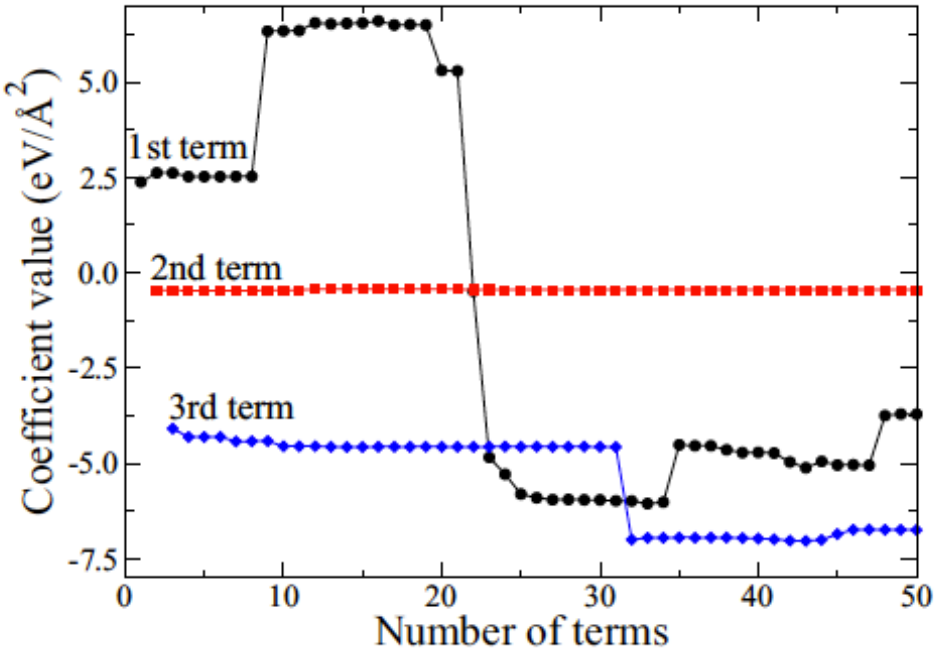
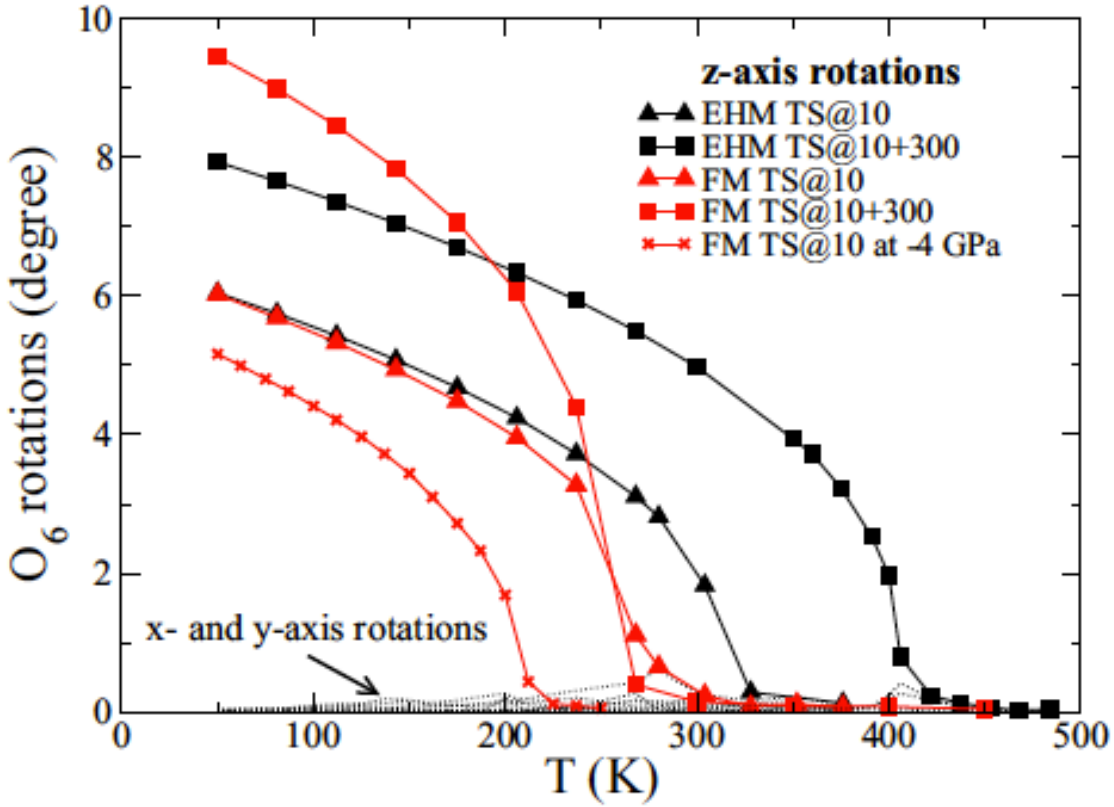



TABLE III. Each row corresponds to one of the three parameters identified to be the most relevant ones for the EHM model fitted to TS@10, indicated by Nos. 1, 2, and 3, respectively. We indicate the corresponding interaction (in the notation of Appendix C) and, in parenthesis, the value of the GF corresponding to the best 1-model, 2-model, and 3-model, respectively. Further, we indicate the terms that are related with these most important ones and whose inclusion in the model causes the discontinuities in their values shown in Fig. 10. Thus, for parameter No. 1, we also include parameters Nos. 9 and 22. In such cases, we give in parenthesis the value of the GF that corresponds to considering a 1-model composed of parameter No. 9 or 22; note that these values are very close to the minimum GF obtained for parameter No. 1. All GF values are given in eV$^2$ Å$^{-2}$.

| | | |
|---|---|---|
| No. 1: $\eta_2(O1_x - O2_x)(O1_y - O2_y)$ (0.00320) | No. 9: $\eta_1(Ti_x - O3_x)^2$ (0.00328) | No. 22: $\eta_3(O2_y - O3_y)^2$ (0.00324) |
| No. 2: $\eta_1(O2_z - O3_z)^2$ (0.00185) | | |
| No. 3: $\eta_4(Sr_z - O1_z)(Sr_y - O1_y)$ (0.00101) | No. 32: $\eta_4(O1_y - O3_y)(O1_x - O3_x)$ (0.00102) | |

# Our model(s) for $SrTiO_3$




| | FM TS@10 | EHM TS@10 | FM TS@10+300 | EHM TS@10+300 | LDA |
|---|---|---|---|---|---|
| $O_6$ rotation (degrees) | 6.5 (6.5) | 6.5 (6.5) | 9.9 (8.7) | 8.3 (8.1) | 6.5 |
| Ground state energy (meV/f.u.) | −11.1 (−11.1) | −13.2 (−13.1) | −13.2 (−11.9) | −22.9 (−20.5) | −11.7 |
| Goal function ($eV^2$ $Å^{-2}$ $\times 10^{-3}$) | 0.14 (0.14) | 0.18 (0.19) | 2.79 (2.81) | 4.94 (5.04) | |
| $\Delta E$ (meV/f.u.) | 0.58 (1.45) | 0.30 (0.27) | 4.74 (4.78) | 2.31 (2.57) | |

# In summary

• We have an automatic scheme to produce light and accurate models that match DFT results beyond your wildest dreams, accuracy wise.

• We make sure our models are bound, taking some drastic measures that, nevertheless, have a very small impact on accuracy.

• They are great to predict novel behaviors (DFT checked!). At the same time, they are not so great when it comes to reproducing some experimental quantities (most notably, $T_C$ in ferroic perovskite oxides)

• We encourage you to use them and have fun…

… but **always know what you are doing!**