Scientific conference of FCS 2025

Talk: Application of machine-learning interatomic potentials in atomistic modeling

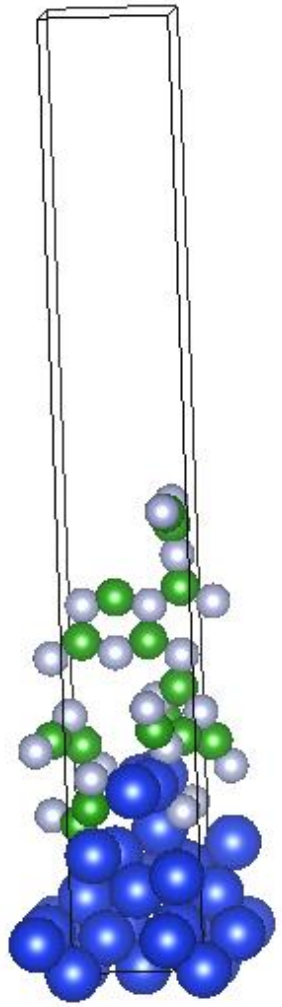
Speaker: Ivan Novikov, Associate Professor

About me

- Student, Moscow Institute of Physics and Technology, 2007-2013
- PhD student, Institute of Numerical Mathematics RAS, 2013-2016
- Candidate of Science, 2016 (Supervisor – Prof. Agoshkov V.I.)
- Research Intern/Junior Research Scientist/Research Scientist, Skolkovo Institute of Science and Technology (Skoltech), May 2016 – December 2019 (Supervisor – Prof. Alexander Shapeev)
- Research Scientist, University of Stuttgart, December 2019 – June 2020 (Research group of Prof. Blazej Grabowski)
- Research Scientist/Senior Research Scientist, Skoltech, August 2020 – March 2025 (Supervisor – Prof. Alexander Shapeev)
- Associate Professor, Faculty of Computer Science, HSE, March 2025 – present
- PI of the RScF grant “Development of machine-learning interatomic potentials with magnetic degrees of freedom”, 6 mln. rub./year, July 2022 – June 2025

Motivation to develop MLIPs

Quantum mechanical (QM) calculations demonstrated themselves as a reliable tool for accurate predicting materials properties. However, they are computationally expensive.



Example: time for calculation of Si-B-N atomic configuration (38 atoms)

MTP (MLIP code): 0.004 seconds, 1 core

Processor: Intel Core i5 7th Gen

QM (VASP code): 1140 seconds, 28 cores

Processor: Intel Xeon E5 2680 v4

QM calculations are not applicable to the systems including more than 1000 atoms!

It is also time-consuming to use QM calculate at finite temperatures even for small systems (of 100 of atoms).

Alternative: machine-learning interatomic potentials!

Moment Tensor Potential (local, non-magnetic)

Potential energy E^{MTP} of MTP is the sum of contributions $V(\mathfrak{n}_i)$ of individual atomic neighborhoods $\mathfrak{n}_i$, $i = \overline{1,n}$:

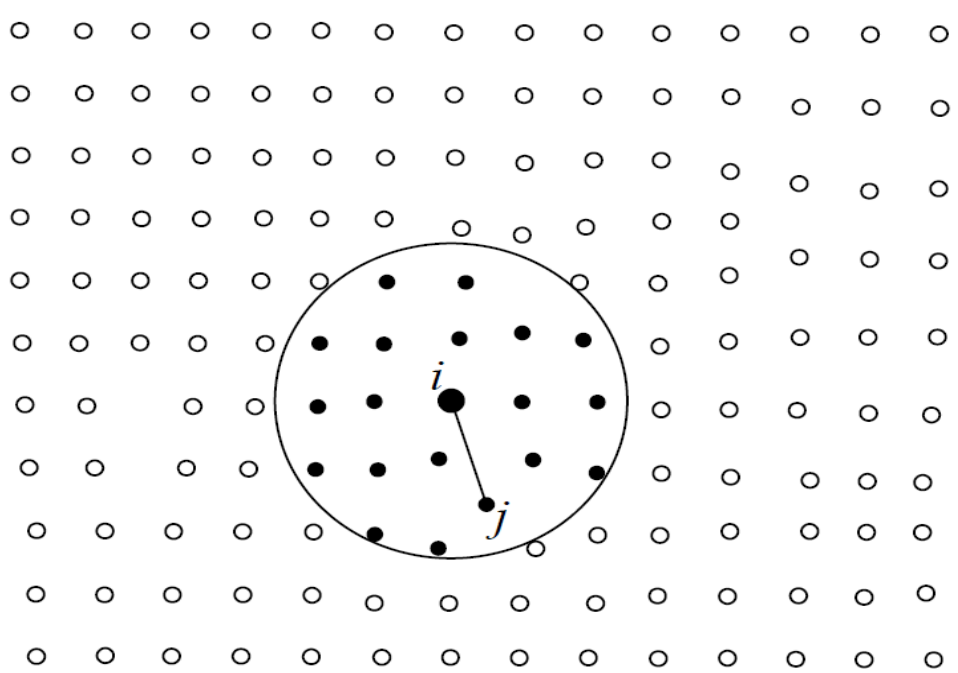
$$E^{\mathrm{MTP}} = \sum_{i=1}^{n} V(\mathfrak{n}_i) = \sum_{i=1}^{n} \sum_{\alpha} \xi_\alpha B_\alpha(\mathfrak{n}_i) \qquad (*)$$

$$\mathfrak{n}_i = \{r_{ij}, z_i, z_j : j = 1, \dots, N_{nb}^i\}$$

B_α are some contractions of Moment Tensor Descriptors:

$$M_{\mu,\nu}(\mathfrak{n}_i) = \sum_j f_\mu(|r_{ij}|, z_i, z_j) \underbrace{\vec{r}_{ij} \otimes \cdots \otimes \vec{r}_{ij}}_{\nu \text{ times}}$$

$$f_\mu(|r_{ij}|, z_i, z_j) = \sum_\beta c_{\mu,z_i,z_j}^{(\beta)} Q_\beta(|r_{ij}|)(R_{cut} - |r_{ij}|)^2$$



We introduce lev $M_{\mu,\nu} = 2 + 4\mu + \nu$ and lev $\prod_{p=1}^{P} M_{\mu_p,\nu_p} = \sum_{p=1}^{P}(2 + 4\mu_p + \nu_p)$.

To define an MTP we choose some $\mathrm{lev}_{\max}$ and include in (*) any function B_α with $\mathrm{lev} B_\alpha \le \mathrm{lev}_{\max}$. We denote $\boldsymbol{\theta} = \{\boldsymbol{\xi}, \boldsymbol{c}\}$ and $E^{\mathrm{MTP}} = E^{\mathrm{MTP}}(\boldsymbol{\theta}; \boldsymbol{R}, \boldsymbol{Z}, \boldsymbol{L}) = E^{\mathrm{MTP}}(\boldsymbol{\theta})$.

Fitting of MTP

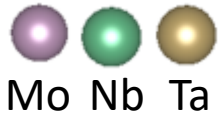
Let K be a number of configurations in a training set with QM energies, forces, and stresses. Denote a set of MTP parameters to be found by $\boldsymbol{\theta}$. The fitting consists of finding the parameters $\boldsymbol{\theta}$ that minimize the following loss function:

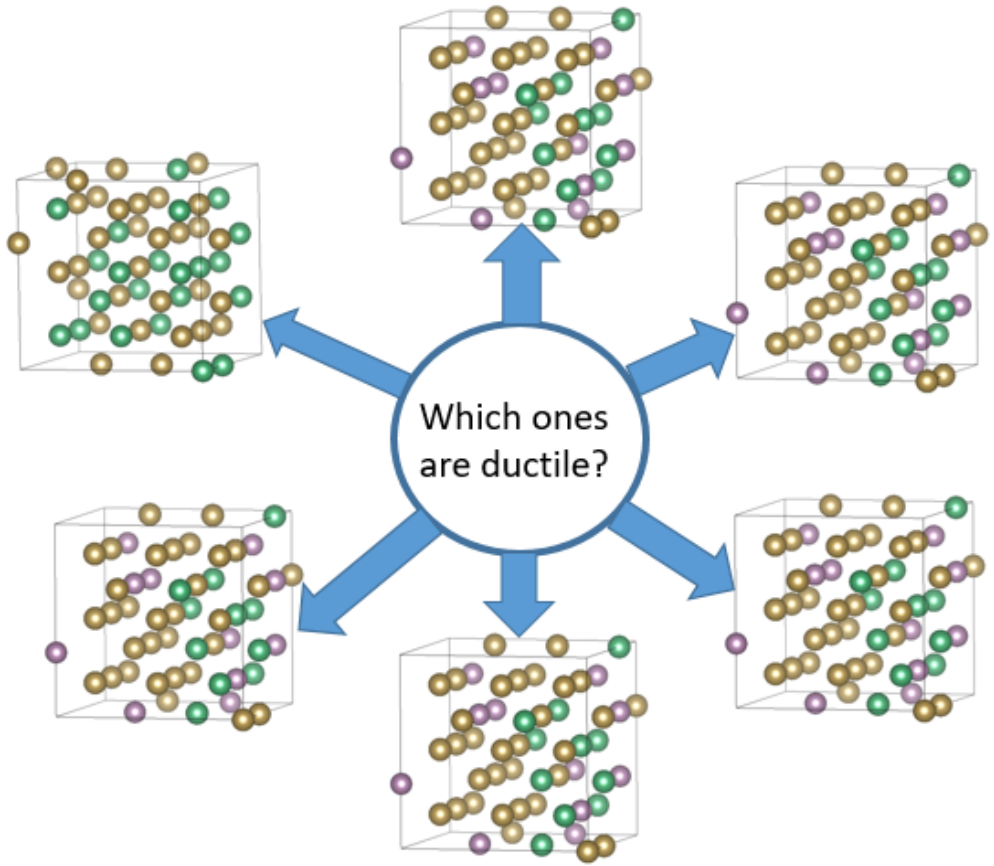
$$\sum_{k=1}^{K}\left[w_e\left(E_k^{\mathrm{QM}}-E_k^{\mathrm{MTP}}(\boldsymbol{\theta})\right)^2+w_f\sum_{i=1}^{n}\left|f_{i,k}^{\mathrm{QM}}-f_{i,k}^{\mathrm{MTP}}(\boldsymbol{\theta})\right|^2+w_s\sum_{i=1}^{6}(\sigma_{i,k}^{\mathrm{QM}}-\sigma_{i,k}^{\mathrm{MTP}}(\boldsymbol{\theta}))^2\right]\rightarrow \min,$$

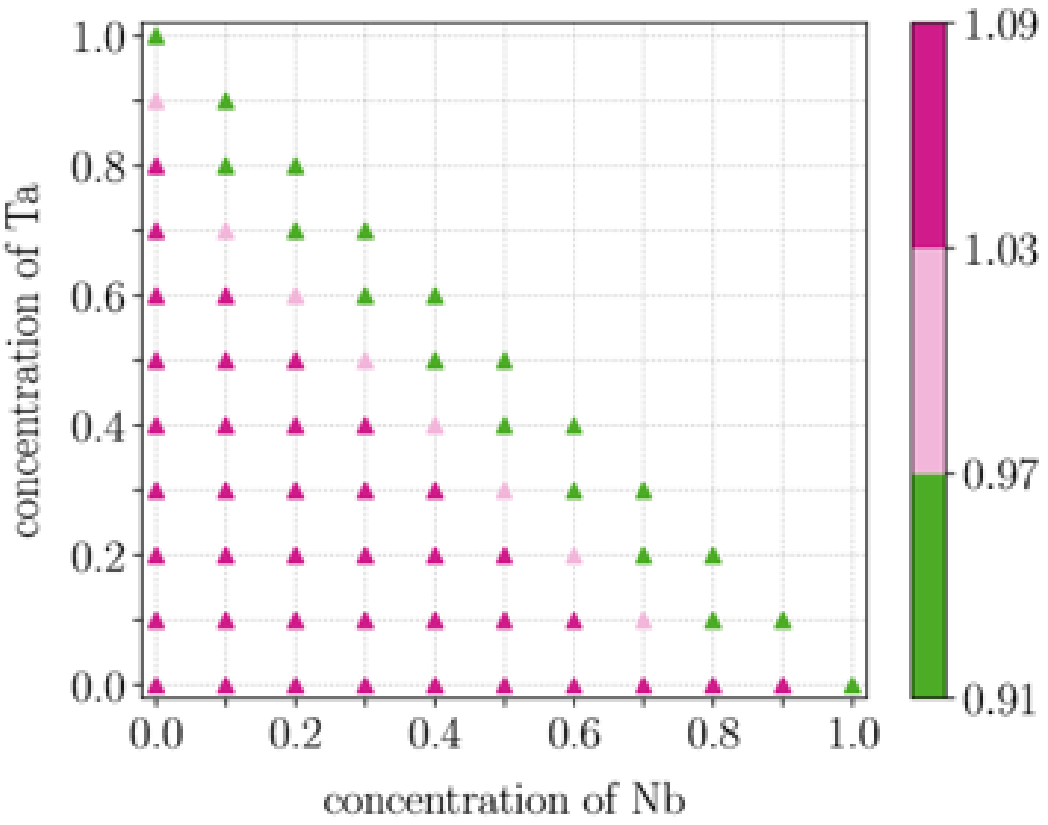
where w_e, w_f, and w_s are non-negative weights. We denote the loss function by $L=L(\boldsymbol{\theta})=L(\boldsymbol{\xi},\boldsymbol{c})$.

Training set can be constructed either manually (e.g., using MD simulations or random perturbations of atoms) or automatically with the MaxVol algorithm. The process of automatic construction of a training set for MTP fitting is called active learning [Gubaev, K., Podryabinkin, E. V., & Shapeev, A. V. (2018). Machine learning of molecular properties: Locality and active learning. *The Journal of chemical physics*, *148*(24)].

Example: prediction of ductility for Mo-Nb-Ta random alloy







We constructed MLIP (MTP) with the AL algorithm and predicted that all the compositions of Mo-Nb-Ta containing less than 20% of Mo, and both Nb and Ta, are ductile ($D/D_c < 1.03$, where $D_c = 1.26$ is the critical index below which the alloy is assumed to be ductile).

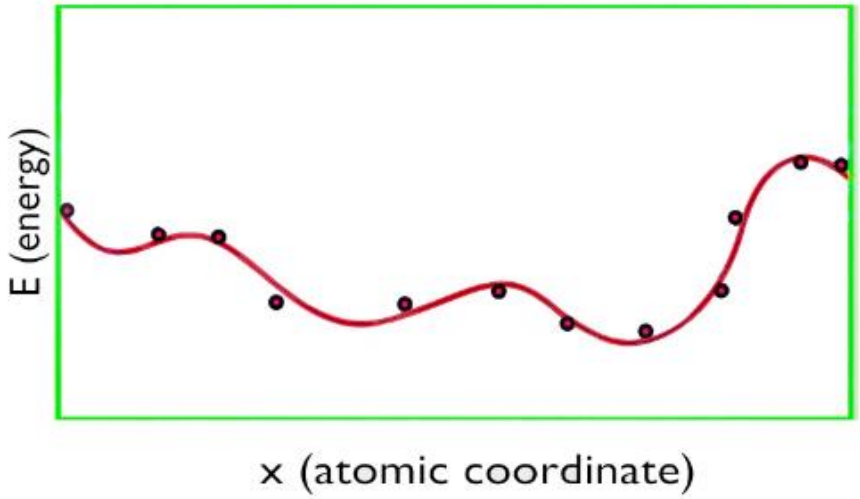
Training set included configurations with up to 100 atoms (240 configurations), but we calculated ductility for the configurations with 10000 atoms!

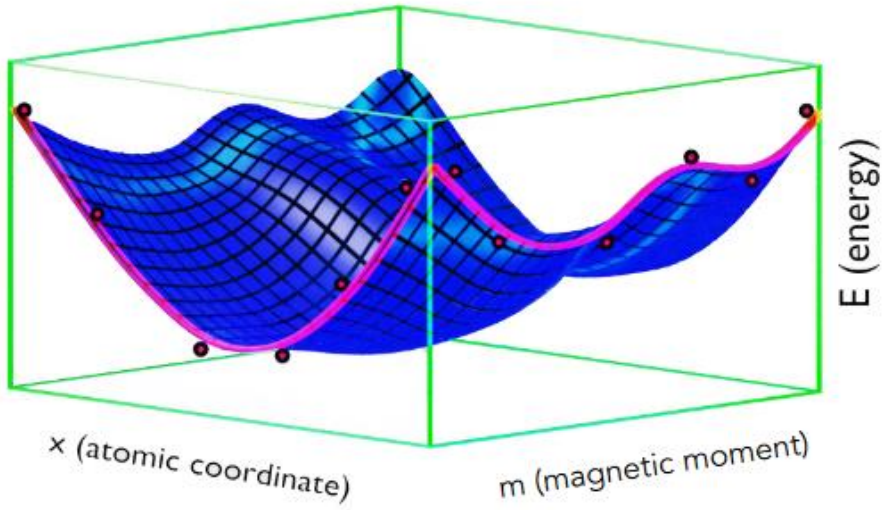
[**Novikov, I.**, Kovalyova, O., Shapeev, A., & Hodapp, M. (2022). AI-accelerated materials informatics method for the discovery of ductile alloys. *Journal of Materials Research*, *37*(21), 3491-3504]

Magnetic Moment Tensor Potential (local)

We explicitly include local magnetic moments (spins) of atoms in the functional form of MLIP, i.e. we consider them as an additional degree of freedom:

$$E^{\mathrm{MTP}}(\boldsymbol{\theta};\boldsymbol{r},\boldsymbol{Z},\boldsymbol{L}) \rightarrow E^{\mathrm{mMTP}}(\boldsymbol{\theta};\boldsymbol{r},\boldsymbol{Z},\boldsymbol{L},\boldsymbol{S})$$



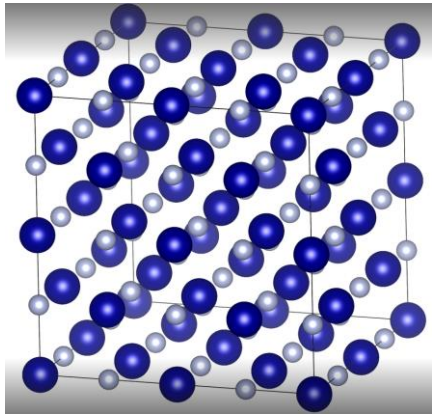


In Moment Tensor descriptors:

$$\sum_{\beta} c^{(\beta)}_{\mu,z_i,z_j} Q_\beta(|r_{ij}|)(R_{cut}-|r_{ij}|)^2 \rightarrow \sum_{\gamma}\sum_{\zeta}\sum_{\beta} c^{(\beta,\gamma,\zeta)}_{\mu,z_i,z_j} Q_\beta(|r_{ij}|)Q_\zeta(s_i)Q_\gamma(s_j)(R_{cut}-|r_{ij}|)^2$$

In this model we consider collinear spins!

Example: a case study of B1-CrN properties

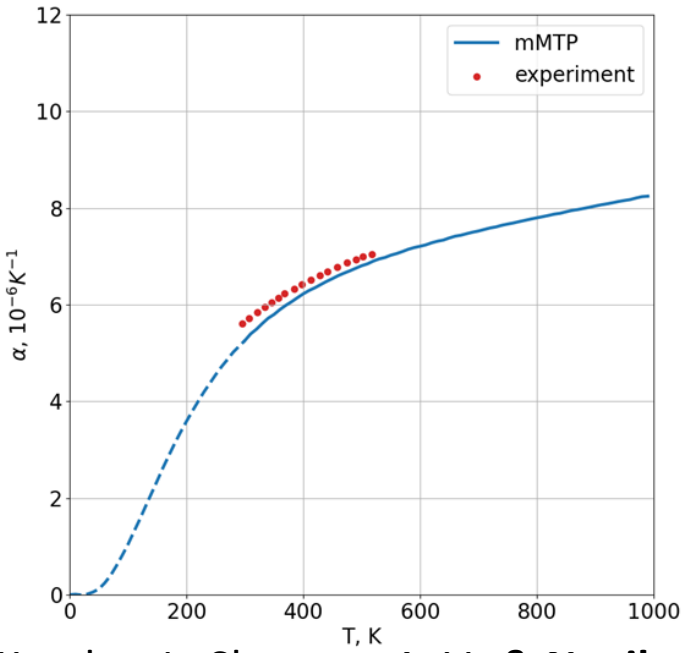


B1-CrN observed experimentally at the temperatures greater than 280 K in the paramagnetic state. We investigated some properties of this material with mMTP.

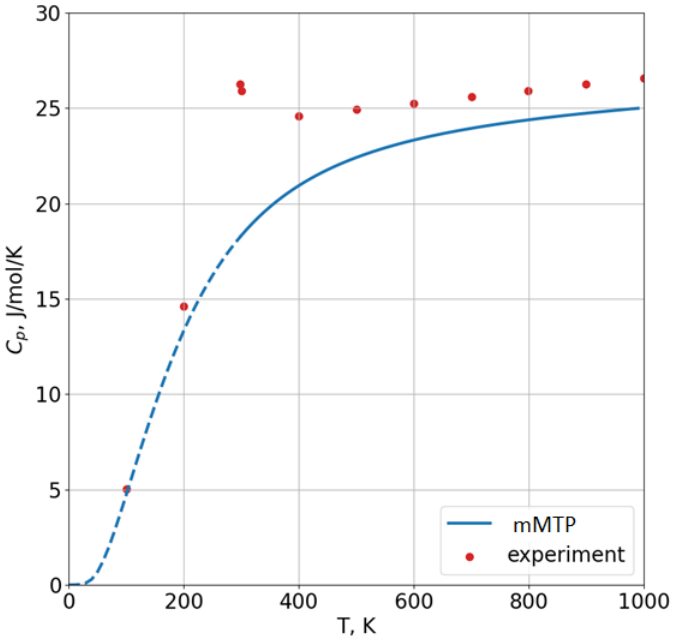
Elastic constants

	C_{11}, GPa	C_{12}, GPa	C_{44}, GPa
mMTP	645	98	160
DFT [this work]	635	93	149
DFT [Zhou et al.]	649	99	145

Lattice thermal expansion



Heat capacity



[Kotykhov, A. S., Hodapp, M., Tantardini, C., Kravtsov, K., Kruglov, I., Shapeev, A. V., & **Novikov, I. S.** (2025). Actively trained magnetic moment tensor potentials for mechanical, dynamical, and thermal properties of paramagnetic CrN. *Physical Review B*, *111*(9), 094438]

Compression of MTP

In non-magnetic MTP radial parameters are written in a tensor of the 4-th order $c^{(N_\beta)}_{N_\mu,N_t,N_t}$, where N_μ is the number of the radial functions f_μ, N_t is the number of atomic types, and N_β is the number of basis functions in radial basis functions. If, e.g., $N_\mu = 4$ and $N_\beta = 8$ then for the case of the F-Li-Na-K system we have $N_t = 4$ and 512 parameters. If $N_t = 7$ then we have 1568 parameters.

In magnetic MTP, we have a tensor of the 6-th order $c^{(N_\beta,N_\zeta,N_\gamma)}_{N_\mu,N_t,N_t}$ and even for the 2-component Fe-Al system with $N_\mu = 3$, $N_\beta = 8$, and $N_\zeta = N_\gamma = 4$ we have 1536 radial parameters of 1567 total parameters.

A way to solve this problem: low-rank approximations!

Low-rank approximations

Skeleton decomposition:

$C = UV^T$ where C is $n \times m$ matrix, U is $n \times r$ matrix, and V is $m \times r$ matrix.

$$c(i_1, i_2) = \sum_{\alpha=1}^{r} u(i_1, \alpha) v(\alpha, i_2)$$

We have $(n + m) \times r$ parameters instead of $n \times m$ parameters. We refer to MTP compressed in this manner as MF MTP.

To apply the skeleton decomposition to the tensor of radial parameters we, first, reshape it: $\text{reshape}(c(N_\mu, N_\beta, N_t, N_t)) = c(N_\mu N_\beta, N_t^2)$ and then apply skeleton decomposition to the reshaped tensor $c(N_\mu, N_\beta, N_t, N_t)$.

Tensor train decomposition:

$$c(i_1, i_2, i_3, i_4) = \sum_{\alpha_1=1, \alpha_2=1, \alpha_3=1}^{r} u_1(i_1, \alpha_1) u_2(\alpha_1, i_2, \alpha_2) u_3(\alpha_2, i_3, \alpha_3) u_4(\alpha_3, i_4)$$

We have $2nr + 2nr^2$ parameters instead of n^4 parameters. We refer to MTP compressed in this manner as TF MTP.

We also used Riemannian optimization (R-MF MTP) and matrix factorization with rank augmentation (MFRA).

Algorithm of matrix factorization with rank augmentation

- Matrix $\hat{C}$ of size $N_\mu N_\beta \times N_t^2$, initial ranks $r_{\min}, r_{\max}$, augmentation interval s (number of iterations of BFGS), rank increment Δr, loss function $L = L\left(\Xi, \hat{C}\right)$.
- Input: $r \coloneqq r_{\min}$, randomly initialized U of size $N_\mu N_\beta \times r$ and V of size $N_t^2 \times r$.

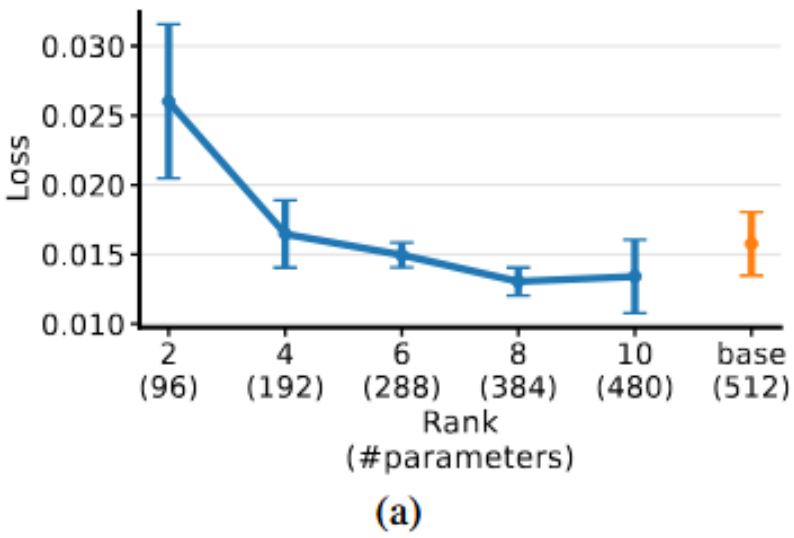
for $k = 1,2, \ldots$ **do**

- Conduct one BFGS iteration and update U and V.
- **If** $k \bmod s = 0$ **and** $r < r_{\max}$ **then**

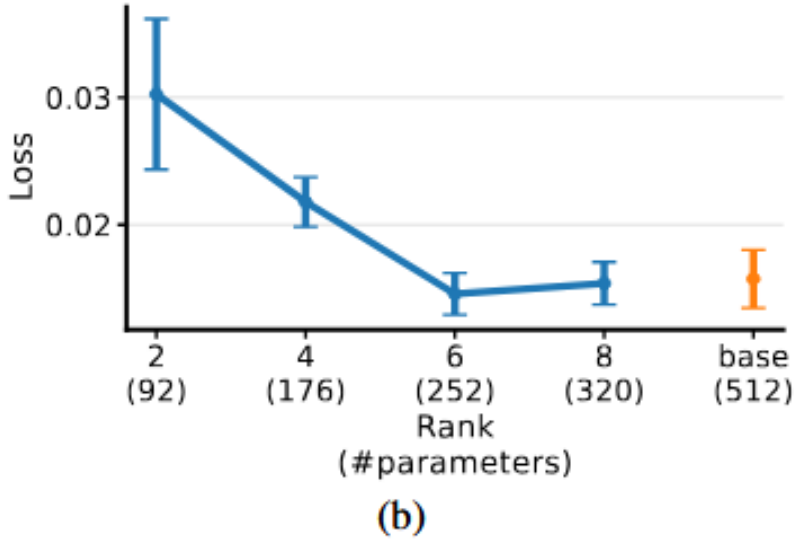
1. Form approximation $\hat{C} = UV^T$.
2. Compute gradient $G = \partial L / \partial \hat{C}$ and compute its SVD: $G \approx \tilde{U}\Sigma\tilde{V}^T$.
3. Extract top Δr singular vectors U_Δ, V_Δ and singular values Σ_Δ.
4. Find $h = \arg\min_{\tilde{h}} L\left(\Xi, \hat{C} + \tilde{h} U_\Delta \Sigma_\Delta V_\Delta^T\right)$.
5. Increase rank: $U \coloneqq [U | h U_\Delta \Sigma_\Delta^{0.5}]$, $V \coloneqq [V | V_\Delta \Sigma_\Delta^{0.5}]$, $r \coloneqq r + \Delta r$.

Example: optimal ranks of compressed MTPs (MoNbTaW)

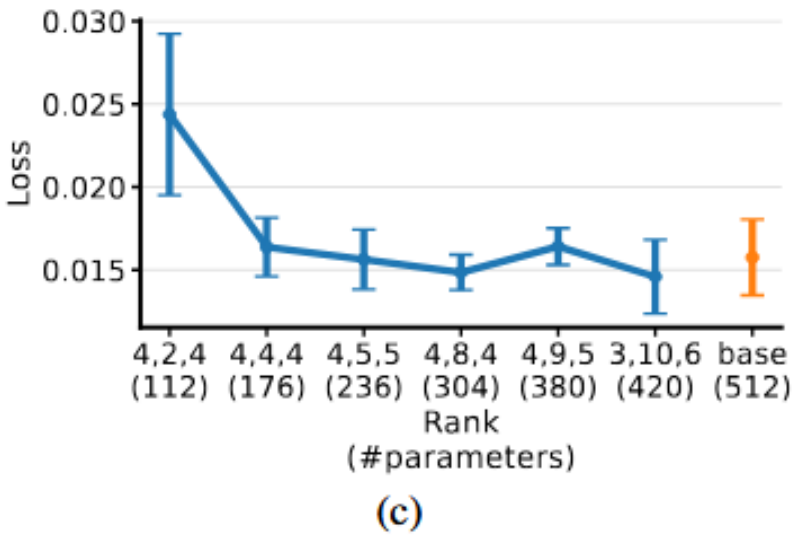
We took a training set for MoNaTaW of 4983 configurations and a validation set of 546 configurations from [Hodapp, Max, and Alexander Shapeev. "Equivariant tensor network potentials." *Machine Learning: Science and Technology* 5.3 (2024): 035075]. We fitted both original (base) and different compressed MTPs on this training set and compared accuracy.



MF



R-MF

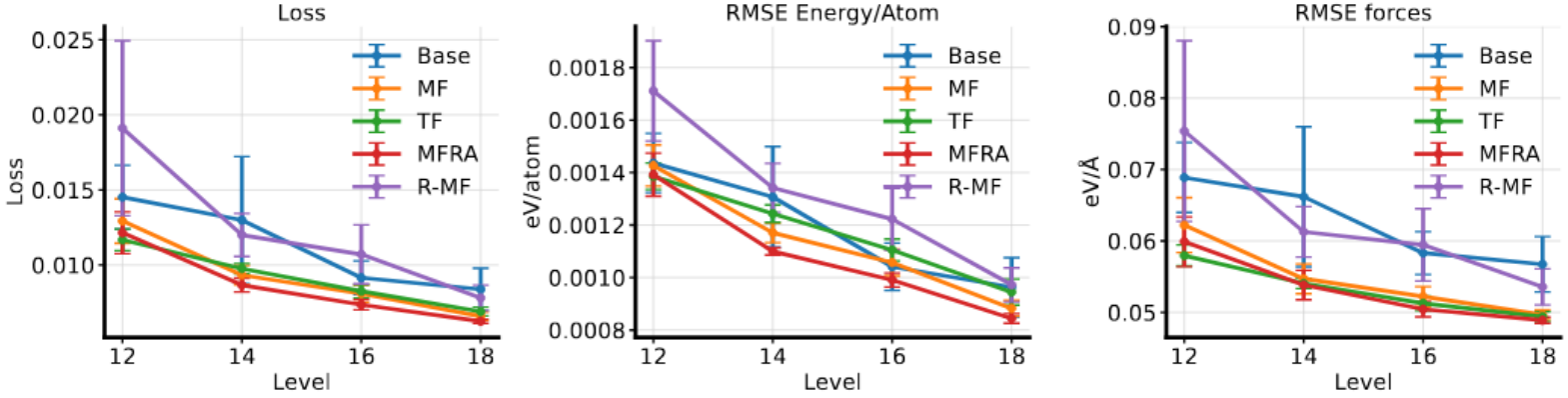


TF

We did not lose in accuracy and even improved it after reduction of the number of MTP radial parameters!

Example: comparison of the fitted compressed and non-compressed MTPs (F-Li-Na-K)

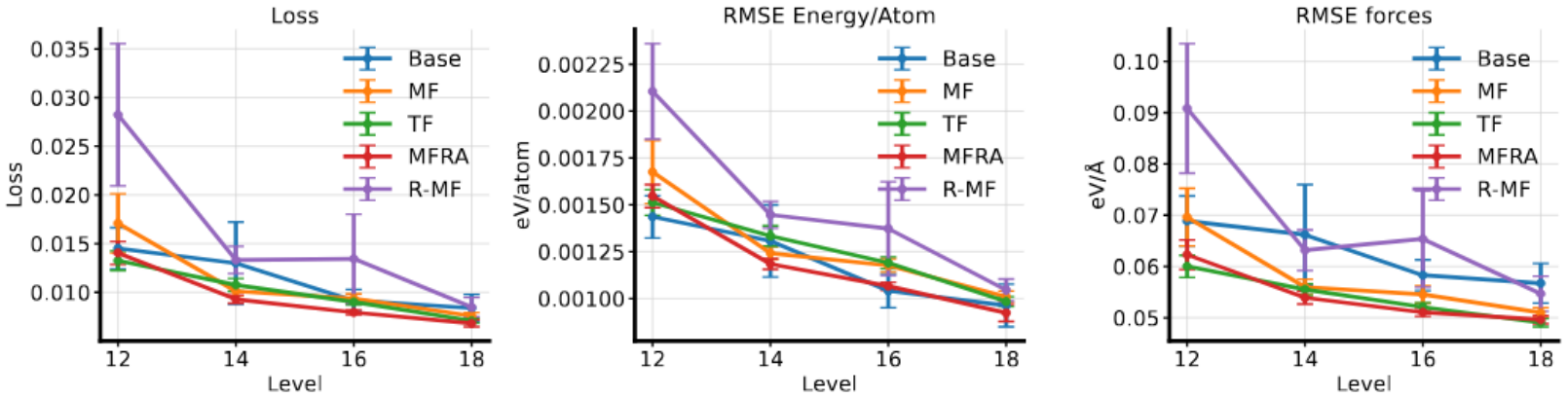
We fitted compressed and non-compressed (base) MTPs on F-Li-Na-K (754 configurations) and compared fitting errors. We also compared density of F-Li-Na-K at different temperatures.



density

Method	# Radial params.	800 K	1000 K	1200 K
Base	512	2.018 ± 0.002	1.904 ± 0.005	1.803 ± 0.003
MF opt	384	2.016 ± 0.003	1.905 ± 0.002	1.802 ± 0.002
MF 50%	240	2.015 ± 0.006	1.907 ± 0.006	1.807 ± 0.003
TF opt	380	2.016 ± 0.003	1.906 ± 0.002	1.801 ± 0.002
TF 50%	276	2.015 ± 0.002	1.904 ± 0.002	1.798 ± 0.004
MFRA opt	384	2.013 ± 0.002	1.904 ± 0.002	1.803 ± 0.003
MFRA 50 %	192	2.015 ± 0.004	1.905 ± 0.004	1.798 ± 0.004
R-MF opt	361	2.015 ± 0.003	1.906 ± 0.003	1.800 ± 0.008
R-MF 50 %	287	2.018 ± 0.002	1.908 ± 0.004	1.798 ± 0.002
Experiment	-	2.080	1.955	1.830

Optimal ranks



Compression of 50% of radial parameters

[Vorotnikov, I., Romashov, F., Rybin, N., Rakhuba, M., & **Novikov, I. S.** (2025). Low-rank matrix and tensor approximations: advancing efficiency of machine-learning interatomic potentials. *arXiv preprint arXiv:2509.04440*]

This work was supported by the Basic Research Program at the HSE University, Russian Federation.

Future plans

- To develop rank adaptation of TF MTP (together with Maxim Rakhuba)
- To actively train the compressed MTP and to compare a size of the obtained training set with the size of the training set needed for the base MTP
- To develop compressed magnetic MTP

Thank you for your attention!