Fusion of quantum-mechanical and experimental data for phase diagram calculation

Timofei Miryashkin [1,2] **Alexander Shapeev** [1,3]

[1]*Skolkovo Institute of Science and Technology, Russia, 121205* [2]*Artificial Intelligence Research Institute, Russia, 105064* [3]*Digital Materials LLC, Russia, 143001.* Correspondence to: Timofei Miryashkin t.miryashkin@gmail.com, Alexander Shapeev alexander@shapeev.com.

Computational materials design aims to predict novel materials with desired properties, such as ductility and corrosion resistance, for industrial applications. Phase diagrams play a crucial role in materials design as they serve as maps that reveal which phase is stable under specific temperature, pressure, and concentration. Artificial intelligence algorithms accelerate quantum mechanical calculations, enabling the construction of phase diagrams with *ab initio* accuracy. However, computational errors can affect the predicted phase stability regions and should be properly accounted for. Moreover, systematic errors in quantum mechanical approximations—such as those arising from the decomposition of the wave function into a product of independent one-electron wave functions in density functional theory—can be corrected by learning from available experimental data. In this work, we develop a Bayesian framework for phase diagram construction that fuses density functional theory and experimental data. A schematic illustration of the framework is shown in Figure 1.

The framework starts with fitting of a Moment Tensor Potential [3] to the Density Functional Theory computations. Configurations for training the Moment Tensor Potential are actively selected using the maxvol algorithm [1], which has been demonstrated to sufficiently reduce the training dataset while maintaining the same accuracy as passively trained interatomic potentials.

The obtained interatomic potential is utilized in our previously developed Bayesian learning algorithm for phase diagram construction [2]. The algorithm is based on Gaussian process regression, which not only reconstructs the free energy from the computational data, but also propagates the statistical uncertainty in the data to the uncertainty of the free energy and, subsequently, to the phase boundaries. The algorithm accepts various data as input including melting points, ensemble-averaged potential energy and concentration, and results of phonon calculations. Our method is further equipped with an active learning algorithm that can suggest new points for molecular dynamics calculations to reduce the uncertainty in the phase diagram prediction in the most effective way.

We demonstrate the application of our Bayesian learning algorithm to the phase diagram of K-Na across its entire range in temperature-concentration coordinates in Figure 2. The resulting diagram is extrapolated for $N \to \infty$ in the system, and phase transition curves are presented with a 2-sigma confidence interval, which is on the order of $2K$. We also include the available experimental data in Figure 2. Most features of our phase diagram quantitatively agree with experimental studies, although the resulting phase transition curves are shifted by nearly $20K$. This observed temperature shift is known for Density Functional Theory computations with Perdew-Burke-Ernzerhof functionals, which typically underpredict melting points by 100-$300K$.

We extend the Bayesian algorithm to explicitly learn the dependence of free energy on a small number of selected interatomic potential parameters. This allows us to find the optimal interatomic potential parameters for which the predicted phase transition curves agree with experimental data. Thanks to the Bayesian nature of the algorithm, these parameters are optimized while accounting for uncertainty in both the predicted and experimental data.

We validate our algorithm on the unary Ti system, which features two transition points: the HCP-BCC and BCC-liquid phase transitions. We train the moment tensor potential on density functional theory data and reconstruct the free energies in the limit $N \to \infty$ with uncertainty for the HCP, BCC, and liquid phases, as shown by dashed lines in Figure 3. We obtain $T_{\text{HCP-BCC}} = 1329 \pm 5K$ and $T_{\text{melt}} = 1696 \pm 2K$ which significantly differ from the experimental values $T^{\text{exp}}_{\text{HCP-BCC}} = 1155K$ and $T^{\text{exp}}_{\text{melt}} = 1941K$, highlighting the need for experimental correction of the interatomic potential. To address this, we extend our Bayesian algorithm to learn the dependence of free energy on two parameters that scale the basis functions in the interatomic potential. By optimizing these parameters, we obtain the corrected phase transition temperatures $T_{\text{HCP-BCC}} = 1155 \pm 10K$ and $T_{\text{melt}} = 1941 \pm 4K$ which agree with experimental values (see Figure 3).

In summary, in this work we developed a framework to incorporate experimental data into a Bayesian algorithm for phase diagram construction and validated it on the unary Ti system which features HCP-BCC and BCC-liquid transition points. The framework is readily extendable to two-component phase diagrams and can be further modified to incorporate other thermodynamic data measured in experiments.

Related work

The most widely used practical approach is CALPHAD [8, 9], in which free energies are fitted with polynomial-like functions, primarily using experi-

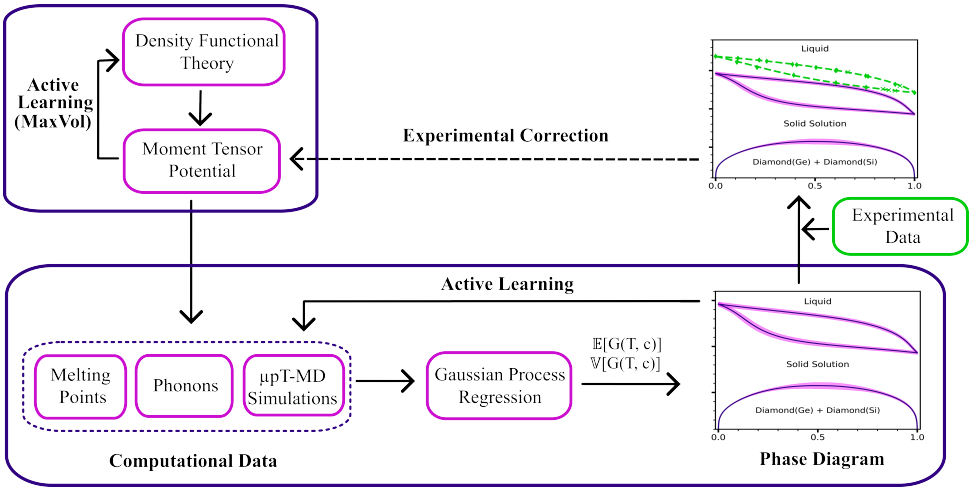

Fig. 1: Framework for phase diagram construction that incorporates experimental data into the learning process. First, a moment tensor potential is trained automatically via active learning [1]. Then, melting point calculations, phonon computations, and molecular dynamics simulations are performed using the trained potential and fed into a Gaussian process regression algorithm. The Gaussian process predicts free energies for each phase and constructs a final phase diagram with associated uncertainties which are further reduced through an active learning algorithm [2]. The predicted free energies depend on a selected set of interatomic potential parameters which are further optimized to reproduce experimental values.

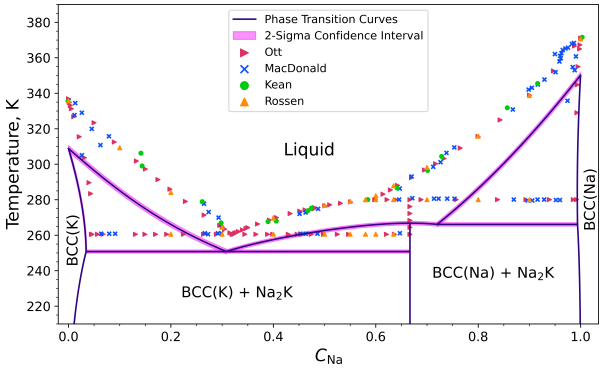

Fig. 2: K–Na phase diagram with a two-sigma confidence interval extrapolated for $N \to \infty$. Experimental data are taken from Ott [4], MacDonald [5], Kean [6], and Rossen [7]. Most features of the phase diagram quantitatively agree with experimental studies, although the resulting phase transition curves are shifted by nearly $20K$.

mental data. Further advancements in the CALPHAD method incorporate density functional theory data into the algorithm, integrate data from various sources, and include uncertainty quantification [10, 11, 12].

Correction of the embedded atom model interatomic potential to reproduce the melting point based on Gibbs-Duhem integration was proposed in [13] and later extended by Mendelev to account for the HCP-BCC transition temperature in unary titanium [14].

Several machine-learning interatomic potentials have been constructed for the titanium system [15, 16, 17], although, to our knowledge, none have been demonstrated to accurately reproduce the HCP-BCC and BCC-liquid transition temperatures.

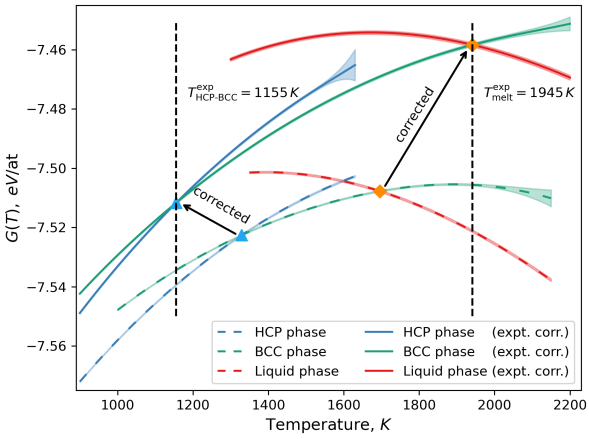

Fig. 3: Free energies of HCP, BCC and liquid phases of Titanium system. Free energies before experimental correction are shown as dashed lines, while those after correction are represented by solid lines. Phase transition temperatures for HCP-BCC and BCC-liqiud after correction coincide with experimental values $T^{\text{exp}}_{\text{HCP-BCC}} = 1155K$ and $T^{\text{exp}}_{\text{melt}} = 1941K$.

Acknowledgments

This work was supported by Russian Science Foundation (grant number 23-13-00332, https://rscf.ru/project/23-13-00332/).

References

[1] Ivan S Novikov, Konstantin Gubaev, Evgeny V Podryabinkin, and Alexander V Shapeev. The mlip package: moment tensor potentials with mpi and active learning. *Machine Learning: Science and Technology*, 2(2):025002, dec 2020.

[2] Timofei Miryashkin, Olga Klimanova, Vladimir Ladygin, and Alexander Shapeev. Bayesian inference of composition-dependent phase diagrams. *Phys. Rev. B*, 108:174103, Nov 2023.

[3] Alexander V. Shapeev. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Modeling & Simulation*, 14(3):1153–1173, 2016.

[4] J. B. Ott, J. R. Goates, D. R. Anderson, and H. T. Hall. Solid-liquid phase equilibria in the sodium + potassium system. *Trans. Faraday Soc.*, 65:2870–2878, 1969.

[5] D. K. C. MacDonald, W. B. Pearson, and Lois T. Towle. An investigation of the sodium-potassium equilibrium diagram. *Canadian Journal of Physics*, 34(4):389–394, 1956.

[6] C. H. Kean. The pressure-temperature phase diagram of na-k alloys and the effect of pressure on the resistance of the liquid phase. *Phys. Rev.*, 55:750–754, Apr 1939.

[7] G. L. C. M. van Rossen and Hoogendijk van Bleiswijk. Über das zustandsdiagramm der kalium-natriumlegierungen. *Zeitschrift für anorganische Chemie*, 74(1):152–156, 1912.

[8] Nigel Saunders and Peter Miodownik. *CALPHAD (Calculation of Phase Diagrams): A Comprehensive Guide*. Elsevier, New York, 1998.

[9] P.J. Spencer. A brief history of calphad. *Calphad*, 32(1):1–8, 2008.

[10] Brandon Bocklund, Richard Otis, Aleksei Egorov, Abdulmonem Obaied, Irina Roslyakova, and Zi-Kui Liu. Espei for efficient thermodynamic database development, modification, and uncertainty quantification: application to cu–mg. *MRS Communications*, 9(2):618–627, 2019.

[11] Noah H. Paulson, Brandon J. Bocklund, Richard A. Otis, Zi-Kui Liu, and Marius Stan. Quantified uncertainty in thermodynamic modeling for materials design. *Acta Materialia*, 174:9–15, 2019.

[12] Richard A. Otis and Zi-Kui Liu. High-throughput thermodynamic modeling and uncertainty quantification for icme. *The Journal of The Minerals, Metals and Materials Society*, 69:886–892, 2017.

[13] Jess B. Sturgeon and Brian B. Laird. Adjusting the melting point of a model system via gibbs-duhem integration: Application to a model of aluminum. *Phys. Rev. B*, 62:14720–14727, Dec 2000.

[14] M. I. Mendelev, T. L. Underwood, and G. J. Ackland. Development of an interatomic potential for the simulation of defects, plasticity, and phase transformations in titanium. *The Journal of Chemical Physics*, 145(15):154102, 10 2016.

[15] Tongqi Wen, Rui Wang, Lingyu Zhu, Linfeng Zhang, Han Wang, David J. Srolovitz, and Zhaoxuan Wu. Specialising neural network potentials for accurate properties and application to the mechanical response of titanium. *npj Computational Materials*, 7(1):206, 2021.

[16] Mashroor S. Nitol, Doyl E. Dickel, and Christopher D. Barrett. Machine learning models for predictive materials science from fundamental physics: An application to titanium and zirconium. *Acta Materialia*, 224:117347, 2022.

[17] A.I. Kartamyshev, A.G. Lipnitskii, V.N. Saveliev, V.N. Maksimenko, I.V. Nelasov, and D.O. Poletaev. Development of an interatomic potential for titanium with high predictive accuracy of thermal properties up to melting point. *Computational Materials Science*, 160:30–41, 2019.