Automated high-throughput thermodynamics of disordered solid solutions with machine-learning interatomic potentials

Dominik Spath[1,†], Luka Wibmer[1], Kieran Bozier[2], Ralf Meyer[1], Simone Di Cataldo[3], Efisio Madeddu[3], Chris Pickard[2,4], Markus Aichhorn[1], Luiz T. F. Eleno[5], Christoph Heil[1], Pedro N. Ferreira[1,†]

[1]Institute of Theoretical and Computational Physics, Graz University of Technology, NAWI Graz, 8010 Graz, Austria
[2]Department of Materials Science and Metallurgy, University of Cambridge, 27 Charles Babbage Road, Cambridge CB3 0FS, United Kingdom
[3]Dipartimento di Fisica, Sapienza Università di Roma, Roma, Italy
[4]Advanced Institute for Materials Research, Tohoku University, Sendai 980-8577, Japan
[5]Computational Materials Science Group (ComputEEL/MatSci), Universidade de São Paulo, Escola de Engenharia de Lorena, DEMAR, Lorena, Brazil
† Corresponding authors: dominik.spath@tugraz.at, nunesferreira@tugraz.at

We present the EGQCA code package, an automated framework for modeling chemical disorder, composition-dependent properties, and phase stability in solid solutions. The package implements the extended generalized quasichemical approximation (EGQCA) [1], in which a disordered material is represented by a set of symmetrically non-equivalent atomic clusters whose equilibrium probabilities are obtained from free-energy minimization. From a compact set of cluster properties, the thermodynamics of the solid solution can then be reconstructed over the entire composition and temperature range, including configurational and, when available, vibrational contributions.

A central aim of the code is to automate the complete computational workflow preceding the thermodynamic assessment. Starting from a parent crystal structure, non-equivalent clusters can be generated and organized automatically, followed by structural relaxation and energy evaluation using either density-functional theory (DFT) or machine-learning interatomic potentials (MLIPs). The current workflow interfaces with standard first-principles codes and incorporates MLIPs as a computationally inexpensive alternative for screening large configurational spaces.

We integrate universal MLIPs directly into the workflow for structural relaxation, energy evaluation, and phonon calculations. This makes it possible to explore cluster spaces that would become prohibitively expensive using DFT alone. In addition, we have implemented the Ephemeral Data-Derived Potential (EDDP) framework, allowing dedicated interatomic potentials to be trained specifically for a target solid solution when universal models do not provide sufficient accuracy or transferability. The resulting workflow therefore spans configuration generation, atomistic relaxation and energetics, thermodynamic modeling, and phase-stability analysis within a common framework.

The MLIP-assisted workflow can also be used to screen the complete configurational space before expensive first-principles calculations are performed. Using the EGQCA cluster probabilities, the thermodynamically most relevant configurations can be identified and selected for subsequent calculations of more demanding physical properties. This provides a physically motivated reduction of the configurational space and concentrates computational effort on the clusters that dominate the behavior of the solid solution.

We will further present preliminary results from an ongoing high-throughput investigation of miscibility gaps across different solid solutions. By coupling automated MLIP calculations to the EGQCA package, composition- and temperature-dependent free energies can be generated efficiently and used to determine phase-separation regions, including binodal and spinodal boundaries. This ongoing study explores the extent to which universal and system-specific MLIPs can be used for rapid phase-stability screening, while identifying systems or composition ranges that require subsequent refinement with first-principles calculations.

Altogether, the EGQCA package provides a scalable route from atomistic calculations to finite-temperature thermodynamics of disordered materials. By combining automated workflows, universal and dedicated MLIPs, cluster selection, and high-throughput thermodynamic analysis, it enables rapid mapping of phase stability and composition-dependent properties while substantially reducing the number of expensive first-principles calculations required.

References

[1] P. N. Ferreira et al., "Ab initio modeling of superconducting alloys," *Materials Today Physics* **48**, 101547 (2024).