Data-Driven Materials Design and Synthesis

Kristin A. Persson[1]
[1]Department of Materials Science and Engineering, UC Berkeley, Berkeley CA 94720, USA

kristinpresson@berkeley.edu

Fueled by increased availability of materials data, machine learning is poised to revolutionize materials science by enabling accelerated discovery, design, and optimization of materials. As one of the most popular materials data providers, the Materials Project (www.materialsproject.org) [1] uses supercomputing and an industry-standard software infrastructure together with state-of-the-art quantum mechanical theory to compute the properties of all known inorganic materials and beyond. The data, currently covering over 160,000 materials and millions of properties, is offered for free to the community together with online analysis and design algorithms. This wealth of data is inspiring data-driven materials design, including the development of machine learning algorithms aimed at predicting material properties and characteristics. However, we note that truly accelerating materials innovation also requires rapid synthesis, testing and feedback, seamlessly coupled to existing data-driven predictions and computations. In this respect, the ability to devise data-driven methodologies to guide synthesis efforts is needed as well as rapid interrogation and recording of results – including 'non-successful' ones.

Current state-of-the-art atomistic modeling of solid state synthesis describe reaction behavior in terms of bulk thermodynamic properties from the Materials Project. Prominent examples include reaction networks which produce thermodynamically favorable reaction pathways linking products and reactants and active learning algorithms [2,3] which propose recipes based on thermodynamics and then updates the recipes according to experimental results. While thermodynamics defines the possible products, predictions based solely on reaction energetics can be inaccurate — especially for systems with competing phases that have similar formation driving forces. [4] In such cases, limited transport of essential constituents may prevent the formation of globally stable products, hindering attainment of thermodynamic equilibrium. from the degree of conversion of the reactants.

Here, we demonstrate a computational inorganic synthesis framework by incorporating machine learning-derived transport properties through "liquid-like" product layers into a thermodynamic cellular reaction model. In the Ba-Ti-O system, known for its competitive polymorphism, we obtain accurate predictions of phase formation with varying BaO:TiO2 ratios as a function of time and temperature. Specifically, we find that diffusion-thermodynamic interplay governs phase compositions, with cross-ion transport coefficients critical for predicting diffusion-limited selectivity.

References

[1] Anubhav Jain et al. "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation". *APL Materials* 1.1 (July 2013), p. 011002. ISSN: 2166-532X. DOI: 10.1063/1.4812323.

[2] Matthew J. McDermott, Shyam S. Dwaraknath, and Kristin A. Persson. "A graph-based network for predicting chemical reaction pathways in solid-state materials synthesis". *Nature Communications* 12.1 (May 2021), p. 3097. ISSN: 2041-1723. DOI: 10.1038/s41467-021-23339-x.

[3] Nathan J. Szymanski et al. "Autonomous and dynamic precursor selection for solid-state materials synthesis". *Nature Communications* 14.1 (Oct. 2023). Publisher: Nature Publishing Group, p. 6956. ISSN: 2041-1723. DOI: 10.1038/s41467-023-42329-9.

[4] Nathan J. Szymanski et al. "Quantifying the regime of thermodynamic control for solid-state reactions during ternary metal oxide synthesis". In: *Science Advances* 10.27 (2024), eadp3309. DOI: 10.1126/sciadv.adp3309.