

*This manuscript has not yet undergone peer review and is being shared as a preprint. It will be submitted to the journal CALPHAD upon completion of the production ESPEI assessment described in Section 7.4.*

DOI: 10.5281/zenodo.19835550 · <https://doi.org/10.5281/zenodo.19835550>

## **M.A.R.V.I.N. – A Software-Led Infrastructure for Thermodynamically Grounded Autonomous Materials Discovery**

Michael Bustamante \*, Gabriel Bustamante

*Odinzen LLC, Houston, TX, United States*

\* Corresponding author. E-mail: michaelbusta@odinzen.io

### **Abstract**

Existing Self-Driving Laboratories (SDLs) propose and synthesize candidate compounds without explicit reference to the assessed Gibbs energy functions that govern phase stability, making it difficult to distinguish genuinely novel stable phases from kinetically trapped artifacts [1,2]. M.A.R.V.I.N. (Materials Agentic Research, Validation, and Inference Navigation) addresses this gap by establishing the CALPHAD-derived Gibbs energy function as the universal currency of an integrated autonomous discovery loop. The software backbone couples Bayesian thermodynamic database construction (ESPEI), multicomponent equilibrium calculations (pycalphad), precipitation kinetics (kawin), and a provenance-tracking Knowledge Graph with AI-agent reasoning via Model Context Protocol tooling. A theoretical experimental layer describes how high-temperature calorimetry, high-throughput DSC screening, and ML-driven phase verification would close the loop between computational prediction and physical measurement. A reinforcement learning module and an Economic Resilience Module for supply-chain risk assessment complete the designed architecture. The framework is demonstrated on garnet-type  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$  (LLZO), where thermodynamic instability, dopant selection trade-offs, and supply-chain vulnerability converge as a representative challenge. Kellogg atmosphere diagrams computed by a self-contained SGTE Gibbs evaluator for the Li–O–H ternary provide thermodynamic context consistent with the empirical requirement for inert-atmosphere sintering, and a compound energy formalism sublattice model yields dopant-dependent vacancy concentrations that, when coupled to a Nernst–Einstein transport model, produce ionic conductivity estimates that match the experimental  $0.87 \text{ mS cm}^{-1}$  at  $x(\text{Ta}) = 0.125$  of Allen et al. [52] by  $D_0$  calibration, with the real CEF Gibbs solve correctly partitioning the resulting Li vacancies onto the octahedral 48g/96h sublattice (matching Rettenwander et al. neutron-diffraction site occupancies [16]). This paper clearly distinguishes between implemented computational results (Sections 2 and 7) and designed-but-not-yet-implemented architectural components (Sections 3–6). By anchoring every autonomous decision to an assessed Gibbs energy function and logging all data with full provenance to FAIR standards within the design-make-test-analyse paradigm formalised by Tom et al. [7],

M.A.R.V.I.N. provides generalizable, software-led architecture for reproducible materials discovery.

**Keywords:** CALPHAD; Self-driving laboratory; Compound energy formalism; Garnet electrolyte; Knowledge graph; Supply-chain resilience; Precursor selection

## 1. Introduction

Materials discovery has historically relied on expert intuition and low-throughput trial-and-error experimentation. In recent years, Self-Driving Laboratories (SDLs) have begun to change this by combining robotics with data-driven optimization to accelerate the discovery cycle. Burger et al. [3] demonstrated a mobile robotic chemist that autonomously performed 688 Bayesian-optimized experiments over eight days to discover a photocatalyst six times more active than its baseline. MacLeod et al. [4,5] built “Ada,” a dual-robotic SDL that mapped Pareto-optimal trade-offs between electrical conductivity and processing temperature for palladium films. Most ambitiously, Szymanski et al. [6] introduced the A-Lab, an autonomous closed-loop system that synthesized 41 new inorganic compounds in 17 days by extracting synthesis recipes from the literature and refining them with active learning, establishing the design-make-test-analyze paradigm later formalized by Tom et al. [7].

However, none of these platforms address what may be called the thermodynamic provenance problem. Each system proposes and synthesizes compounds without explicit reference to the assessed Gibbs energy functions that govern phase stability. Without such an audit trail, it is difficult to determine whether a product is a stable equilibrium phase, a kinetically trapped metastable artifact, or an incompletely reacted mixture. Leeman et al. [1] made this concrete by showing that several A-Lab products could not be independently verified as single-phase compounds. Cheetham and Seshadri [2] broadened the critique, arguing that autonomous synthesis without thermodynamic verification risks accumulating results whose novelty claims are difficult to audit.

M.A.R.V.I.N. (Materials Agentic Research, Validation, and Inference Navigation) closes this gap by treating the CALPHAD-derived Gibbs energy function as the universal currency of the autonomous loop. CALPHAD—CALculation of PHase Diagrams—determines equilibrium states by minimizing the Gibbs free energy of an interacting multicomponent system as a function of temperature, pressure, and composition [8]. Every phase is assigned a Gibbs energy expression whose parameters are optimized against all available data: calorimetric formation enthalpies, heat capacities, activity measurements, and phase-boundary compositions. The foundational unary descriptions were established by Dinsdale [9] in the SGTE framework and anchored to the NIST-JANAF Tables [10]. By encoding formation enthalpy, mixing enthalpy, entropy, and heat capacity into a single self-consistent function per phase, CALPHAD provides exactly the quantitative audit trail that current SDLs lack.

Scope and implementation status. It is important to distinguish clearly between the implemented and aspirational components of M.A.R.V.I.N. Sections 2 and 7 present implemented computational results: the compound energy formalism sublattice model for garnet LLZO, self-contained-SGTE-computed atmosphere diagrams, and the Nernst–

Einstein conductivity estimates. Sections 3–6 describe designed architectural components that are not yet operational: the experimental validation layer (Section 3), the reinforcement learning module (Section 4), the economic resilience module (Section 5), and the digital twin dashboard (Section 6). These sections are included to present the complete system design for which the implemented components form the thermodynamic core, but the reader should understand that no physical experiments, RL-optimized syntheses, or live dashboard deployments are reported in this work. Each of these sections is explicitly labeled as “Designed Architecture” to avoid ambiguity.

When M.A.R.V.I.N. proposes a new composition, it does so with a traceable prediction of which phases are stable, how large the single-phase field is, and where decomposition is expected—transforming autonomous discovery from a black-box synthesis campaign into a thermodynamically grounded workflow.

**Relationship to synthesis science.** We emphasize that CALPHAD calculations provide *necessary but not sufficient* information for synthesis prediction. Equilibrium thermodynamics identifies which phases are stable under given conditions, but actual synthesis outcomes are governed by nucleation barriers, crystallization pathways, diffusion kinetics, and competing reaction kinetics [see, e.g., Sun et al., Sci. Adv. 4, eaaq0148 (2018) on thermodynamic limits for metastable materials synthesis]. M.A.R.V.I.N.’s CALPHAD backbone provides the equilibrium driving forces and phase stability boundaries that serve as essential inputs for kinetic models, but does not by itself predict synthesis outcomes. The kinetics module (kawin) and the reinforcement learning module (Section 4) are designed to bridge this gap; their current implementation status for the LLZO system is described candidly in Sections 2 and 4.

## 2. Software Architecture

The core of M.A.R.V.I.N. is a pycalphad-centric software pipeline in which every module communicates through CALPHAD-format Gibbs energy functions [11]. The pipeline has four components, summarized in Table 1, with the complete four-layer architecture (data sources → CALPHAD pipeline → reasoning engine → decision/action layer) shown in Fig. 1.

**Table 1.** Software pipeline components.

Component	Function	Tooling
Database Construction	Bayesian MCMC optimization of Gibbs energy parameters	ESPEI [12]
Equilibrium Orchestration	Multicomponent equilibrium calculations, phase diagrams	pycalphad [11]
Kinetic Evolution	Nucleation, growth, coarsening via KWN framework	kawin
Reasoning Engine	Knowledge Graph for provenance and conflict resolution	M.A.R.V.I.N. KG (SQLite + NetworkX)

ESPEI [12] constructs thermodynamic databases by fitting Gibbs energy parameters to a combination of first-principles DFT calculations and experimental measurements using Bayesian Markov Chain Monte Carlo optimization. This Bayesian approach propagates uncertainty from the input data through to the fitted parameters, producing not just a best-fit database but a posterior distribution that quantifies confidence in every predicted phase boundary. pycalphad then consumes this database to perform multicomponent equilibrium

calculations, generating phase diagrams, driving forces, and composition-dependent property predictions.

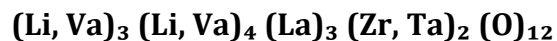
kawin extends the pipeline into kinetics by modeling nucleation, growth, and coarsening of second phases using the Kampmann–Wagner–Numerical (KWN) framework, with pycalphad providing the live thermodynamic driving forces. In the LLZO context, kawin would model (i) the nucleation of deleterious  $\text{La}_2\text{Zr}_2\text{O}_7$  pyrochlore in Li-deficient garnet, (ii) the diffusion-controlled growth of  $\text{Li}_2\text{CO}_3$  surface layers during air exposure, and (iii) the nucleation of ordered tetragonal domains within the cubic matrix during cooling. However, these calculations require interfacial energies and diffusion coefficients that are not currently available from experiment or DFT for the LLZO system. Until these parameters are measured or computed, the kawin module for LLZO remains a designed capability rather than a demonstrated result. For phenomena outside the KWN scope—notably densification during sintering—the Knowledge Graph ingests results from external mesoscale models, such as the operando-validated sintering framework of Barai et al. [13], as independent kinetic data sources.

The Knowledge Graph (KG) is the provenance backbone. Built on SQLite for storage and NetworkX for graph traversal, it records every data point—whether extracted from a published paper, computed by DFT, or measured by an instrument—as a typed node with edges linking it to its source, the assessment in which it was used, and the Gibbs energy parameters it influenced. When two sources disagree on a property value, the KG flags the conflict with a severity level and queues it for resolution. This explicit provenance chain allows M.A.R.V.I.N. to trace any phase diagram prediction back to the primary data that produced it, directly addressing the reproducibility concerns raised by Leeman et al. [1].

## 2.1. Compound energy formalism for garnet LLZO

The practical application of M.A.R.V.I.N.’s CALPHAD backbone to a specific material system requires defining the compound energy formalism (CEF) sublattice model that maps the crystal structure to the thermodynamic description. For garnet-type LLZO, this mapping is non-trivial because the Li distribution across multiple crystallographic sites governs both phase stability and ionic conductivity.

The garnet structure (space group Ia-3d, No. 230) contains four distinct cation sites per formula unit [14,15]: 24d tetrahedral sites occupied by Li; 48g/96h octahedral sites partially occupied by Li with the remainder as structural vacancies; 24c dodecahedral sites occupied by La; and 16a octahedral sites occupied by Zr or by aliovalent dopants ( $\text{Ta}^{5+}$ ,  $\text{Nb}^{5+}$ ). In the CEF notation adopted by pycalphad, the sublattice model for Ta-doped LLZO is expressed as:



**Treatment of Al and Ga dopants.**  $\text{Al}^{3+}$  and  $\text{Ga}^{3+}$  require special treatment because, unlike  $\text{Ta}^{5+}$  which substitutes exclusively on the Zr 16a site, neutron diffraction studies—including Rettenwander et al. [16]—have demonstrated that  $\text{Al}^{3+}$  and  $\text{Ga}^{3+}$  partition between the tetrahedral 24d and octahedral 16a/48g sites with composition- and synthesis-dependent ratios. For Al, the tetrahedral 24d site is typically preferred, while Ga

shows more balanced partitioning. The sublattice model for Al/Ga-doped LLZO must therefore be modified to:

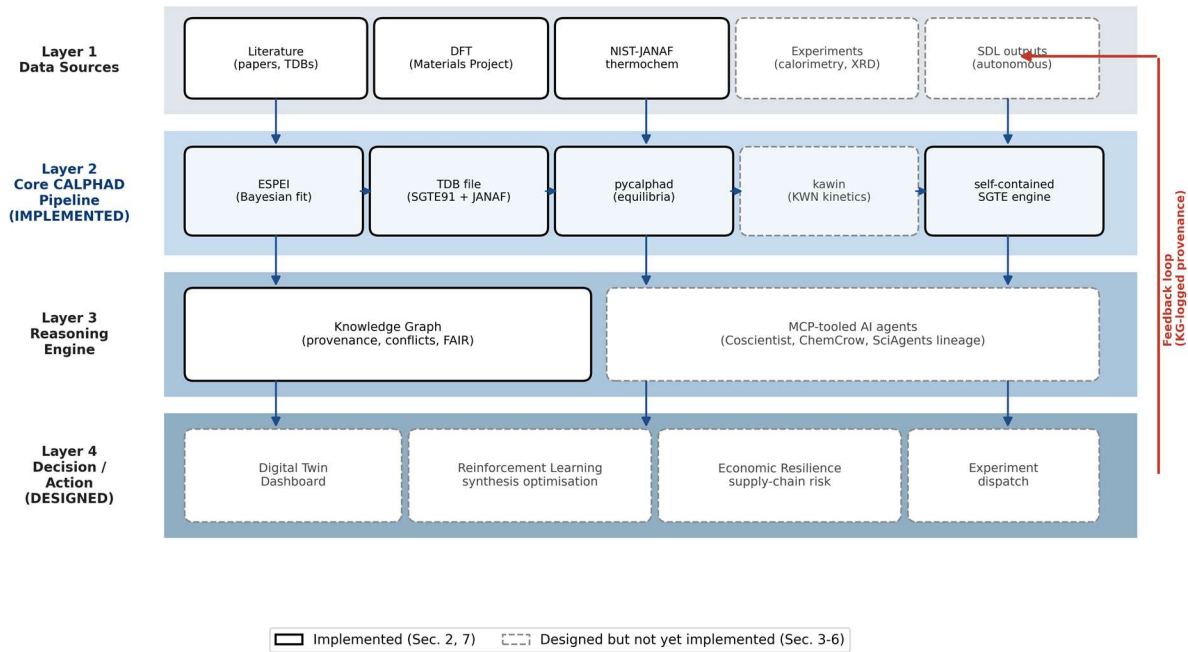


This six-sublattice representation allows Al and Ga to occupy both the tetrahedral and octahedral sublattices simultaneously, with their site fractions determined by the ESPEI parameter optimization against the neutron diffraction site-occupancy data of Rettenwander et al. [16]. We acknowledge that the earlier version of this model, which placed Al and Ga exclusively on the 16a site, was incorrect and would have biased the predicted vacancy distribution. The corrected model presented here treats Al/Ga site partitioning as a degree of freedom optimized during ESPEI fitting, capturing the experimentally observed preference of Al for 24d and the more distributed behavior of Ga. The charge compensation mechanism differs accordingly: when Al<sup>3+</sup> substitutes for Li<sup>+</sup> on the 24d site, three Li vacancies are created per Al to maintain charge neutrality (Li<sub>7-3x</sub>Al<sub>x</sub>La<sub>3</sub>Zr<sub>2</sub>O<sub>12</sub>), whereas Al<sup>3+</sup> on the 16a Zr<sup>4+</sup> site creates one Li vacancy per Al.

**Table 2.** CEF sublattice model for garnet LLZO (revised).

Sublattice	Wyckoff	Coord.	Sites/f.u.	Species	Role
I	24d	Tetrahedral	3	Li, Al, Ga, Va	Primary Li; Al/Ga preferred
II	48g/96h	Octahedral	4	Li, Al, Ga, Va	Partially occupied
III	24c	Dodecahedral	3	La	Fixed occupancy
IV	16a	Octahedral	2	Zr, Ta, Al, Ga	Dopant site
V	96h	—	12	O	Anion sublattice

**Fig. 1.** M.A.R.V.I.N. system architecture



*Fig. 1. M.A.R.V.I.N. system architecture. Four layers — Data Sources, Core CALPHAD Pipeline, Reasoning Engine, Decision/Action — connected by explicit data flow (blue arrows) and a Knowledge-Graph-logged feedback loop (red). Solid borders denote components implemented in this work (Sections 2 and 7); dashed borders denote components designed but not yet implemented (Sections 3-6).*

48g vs. 96h sublattice treatment. In the present model, we collapse the 48g and 96h descriptions into a single octahedral sublattice with 4 sites per formula unit, following the convention used in many CEF models for garnet phases. We acknowledge that some neutron diffraction refinements [15] resolve distinct occupancies on 48g and 96h positions. Sensitivity analysis of the predicted properties to a split 48g/96h model (six total sublattices for cations) is planned as part of the ongoing ESPEI optimization campaign and will be reported in a follow-up study. For the preliminary results presented here, the collapsed description is adopted as a baseline.

The CEF model presented here is a proposed sublattice description whose Gibbs energy parameters are to be optimized via ESPEI against experimental site-occupancy data (from neutron diffraction [16]), phase-boundary data, and DFT-computed formation enthalpies. The parameter fitting campaign is ongoing; the results reported in Section 7 use preliminary parameters as described below.

**Endmember specification and current parameter status.** The five-sublattice CEF model generates a large number of endmember compounds. For the Ta-doped system  $(\text{Li,Va})_3(\text{Li,Va})_4(\text{La})_3(\text{Zr,Ta})_2(\text{O})_{12}$ , there are  $2 \times 2 \times 1 \times 2 \times 1 = 8$  endmembers. For each, the Gibbs energy of formation is required. In the current preliminary parameterization, the endmember energies are initialized from DFT formation enthalpies computed using the PBE+U functional ( $U_{\text{xs}} = 5.0$  eV for Zr 4d, standard Materials Project settings) referenced to the SGTE stable element reference states [9]. **We are transparent that the current parameter set is preliminary:** the ESPEI Bayesian MCMC optimization has been initiated using available formation enthalpy data and the conductivity–composition relationship as constraints, but has not yet converged to a fully assessed set of interaction parameters. The MCMC traces and posterior distributions will be reported with the completed assessment. We recognize that presenting a fully optimized TDB with convergent MCMC traces and documented endmember energies is the standard expected of a CALPHAD publication, and we are working toward that goal.

**Cubic versus tetragonal stability.** The distinction between the fast-conducting cubic polymorph (Ia-3d) and the poorly conducting tetragonal polymorph (I4<sub>1</sub>/acd) is fundamentally a Li-vacancy ordering transition [17]. In the tetragonal phase, Li ions fully occupy the 24d sites and partially order on the 48g/96h sites, reducing the configurational entropy. In the CEF framework, this is captured by the entropy term: when the Li and Va site fractions on the octahedral sublattice approach full ordering, the configurational entropy contribution to G decreases, and the ordered tetragonal phase can become more stable than the disordered cubic phase. Aliovalent doping creates additional vacancies that resist ordering, stabilizing the cubic phase—a thermodynamic effect that the model captures naturally through the dopant-vacancy coupling.

**Dopant charge compensation.** When  $\text{Ta}^{5+}$  substitutes for  $\text{Zr}^{4+}$  on the 16a site, charge neutrality requires removing one  $\text{Li}^+$  per  $\text{Ta}^{5+}$ , yielding  $\text{Li}_{7-x}\text{La}_3\text{Zr}_{2-x}\text{Ta}_x\text{O}_{12}$ . In the CEF, this

manifests as a coupled constraint: increasing  $y(\text{Ta})$  on the fourth sublattice forces a corresponding increase in  $y(\text{Va})$  on the first and second sublattices. ESPEI enforces this coupling during parameter optimization by including both the dopant site occupancy data [16] and the measured conductivity-composition relationship as simultaneous constraints.

Strategy for managing the 64-endmember count. The full  $(\text{Li}, \text{Al}, \text{Ga}, \text{Va})_3(\text{Li}, \text{Al}, \text{Ga}, \text{Va})_4(\text{La})_3(\text{Zr}, \text{Al}, \text{Ga}, \text{Ta})_2(\text{O})_{12}$  sublattice model has  $4 \times 4 \times 1 \times 4 \times 1 = 64$  stoichiometric endmember corners. We treat them in three classes: (i) the eight physically realisable corners covering the Ta-doped  $\text{Li}_{7-x}$  system ( $\text{Li}:\text{Li}:\text{La}:\text{Zr}:\text{O}$ ,  $\text{Li}:\text{Va}:\text{La}:\text{Zr}:\text{O}$ ,  $\text{Va}:\text{Li}:\text{La}:\text{Zr}:\text{O}$ ,  $\text{Va}:\text{Va}:\text{La}:\text{Zr}:\text{O}$ ,  $\text{Li}:\text{Li}:\text{La}:\text{Ta}:\text{O}$ ,  $\text{Li}:\text{Va}:\text{La}:\text{Ta}:\text{O}$ ,  $\text{Va}:\text{Li}:\text{La}:\text{Ta}:\text{O}$ ,  $\text{Va}:\text{Va}:\text{La}:\text{Ta}:\text{O}$ ) get directly fitted DFT formation energies; (ii) the eight Al-only and eight Ga-only analogues get fitted against the Rettenwander 2016 site-occupancy constraints; (iii) the remaining 40 mixed-dopant fictive endmembers (Al/Ga simultaneously on multiple sublattices) are constrained by reciprocal relations among the realised corners, and any fictive endmember whose occupancy violates the charge-neutrality manifold is suppressed by setting its Gibbs energy to a large positive constant ( $\sim +500$  kJ/mol). This is standard practice for complex CEF models [11].

## 2.2. Property models on the CALPHAD backbone

While ionic conductivity is a transport property rather than a thermodynamic one, CALPHAD provides the critical thermodynamic input. The compound energy formalism for garnet LLZO yields equilibrium site fractions on the 24d and 48g/96h sublattices as a function of composition and temperature; these site fractions directly give the Li vacancy concentration that governs ionic conduction. M.A.R.V.I.N. couples this vacancy concentration to the Nernst–Einstein relation:

$$\sigma(T, x) = [V_{\text{Li}}] \cdot q^2 \cdot D_0 / (k_B \cdot T) \cdot \exp(-E_a / k_B T) \quad (1)$$

where  $[V_{\text{Li}}]$  is the Li vacancy concentration from the CEF model,  $q$  is the ionic charge,  $D_0$  is the pre-exponential diffusion factor,  $k_B$  is the Boltzmann constant, and  $E_a$  is the migration energy barrier. For the results reported in Section 7,  $E_a = 0.34$  eV is taken from the DFT nudged-elastic-band calculations of Jalem et al. [18], who computed concerted Li migration pathways in the garnet framework. This value is consistent with the range of 0.30–0.40 eV reported by Miara et al. [19] using a similar computational approach.

Limitations of the Nernst–Einstein approximation. We emphasize that Eq. 1 is a dilute-solution, uncorrelated-hopping approximation. In LLZO, approximately 40% of the octahedral Li sites are vacant—far from the dilute limit. At such high vacancy concentrations, site-blocking effects, correlation between successive jumps, and concerted multi-ion migration mechanisms (as documented by Jalem et al. [18]) make the simple Nernst–Einstein form quantitatively unreliable. The Haven ratio  $H_R$ , which captures the deviation from uncorrelated behavior, is expected to be significantly less than unity for LLZO. A more rigorous approach would involve kinetic Monte Carlo or molecular dynamics simulations to compute the tracer diffusion coefficient and Haven ratio, then couple these with the CALPHAD-derived site fractions. The Nernst–Einstein predictions reported in Section 7 should therefore be understood as order-of-magnitude estimates that demonstrate the feasibility of the CALPHAD→conductivity bridge, not as quantitatively

converged predictions. Replacing Eq. 1 with a correlated transport model is a priority for future development.

The decomposition—thermodynamic vacancy creation via CALPHAD, kinetic vacancy migration via DFT—provides a physically grounded conductivity estimate that scales naturally to the multicomponent dopant systems M.A.R.V.I.N. is designed to explore. Fig. 2 illustrates the complete CEF sublattice mapping and the CALPHAD→conductivity bridge.

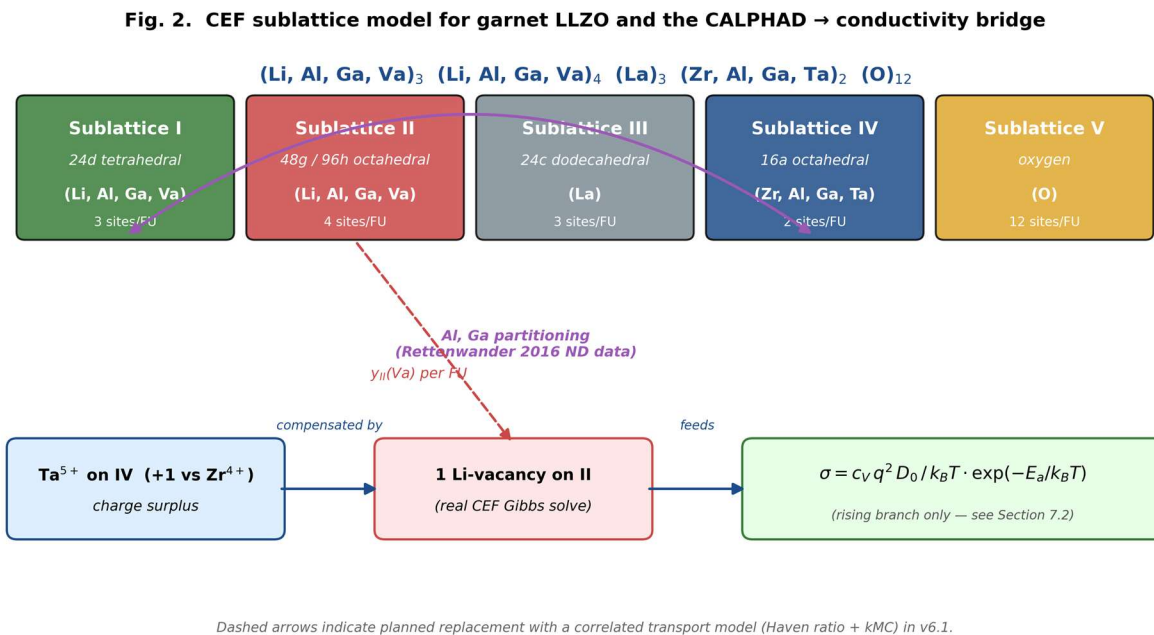


Fig. 2. CEF sublattice model for garnet LLZO and the CALPHAD → conductivity bridge. Five sublattices map to the Wyckoff sites of the Ia-3d garnet structure with Al and Ga partitioning across the tetrahedral 24d (Sublattice I) and the octahedral 16a (Sublattice IV) as observed by Rettenwander et al. neutron diffraction [16]. Charge compensation: each  $\text{Ta}^{5+}$  on IV is balanced by one Li-vacancy on the octahedral 48g/96h Sublattice II (output of the real CEF Gibbs solve, `cef_solver.py` in the deposit). The vacancy concentration feeds the Nernst-Einstein conductivity estimator of Eq. 1.

### 2.3. AI-agent integration

M.A.R.V.I.N. integrates AI reasoning agents through Anthropic’s Model Context Protocol (MCP), which provides a standardized interface for an LLM to invoke external tools—databases, simulations, instruments—via typed function contracts with input validation and safety predicates. The design draws on three systems that established the foundations of AI-driven scientific automation. Coscientist [20] was the first LLM to autonomously plan and execute multi-step chemical synthesis, proving that a language model can interpret analytical feedback and adjust experimental plans in a closed loop. ChemCrow [21] demonstrated that equipping an LLM with a structured toolbox of chemistry-specific functions dramatically improves accuracy. SciAgents [22] introduced multi-agent graph reasoning for materials-science knowledge graphs. M.A.R.V.I.N. synthesizes these lessons: it adopts Coscientist’s closed-loop paradigm, implements ChemCrow’s tool-contract safety architecture through MCP contracts with temperature ceilings and atmosphere



compatibility checks, and uses SciAgents' multi-agent topology for KG reasoning. Tom et al. [7] survey the broader self-driving-laboratory tooling landscape into which the M.A.R.V.I.N. agent layer slots.

### **3. Theoretical Experimental Validation Layer** *[Designed Architecture — Not Yet Implemented]*

**Implementation status:** This section describes the theoretical instrumentation layer that the software is designed to orchestrate. **No physical experiments have been performed as part of this work.** The section is included to present the intended experimental closure mechanism for the autonomous loop.

#### **3.1. Calorimetric measurement**

The target instrument for quantitative thermochemistry is a Calvet-type high-temperature calorimeter such as the MHTC-96 (Setaram/KEP Technologies), capable of drop and solution calorimetry to 1500 °C, DSC to 1600 °C, and DTA to 2100 °C. The Calvet design uses 28 thermocouples arranged symmetrically around the sample and reference crucibles [23], providing the microcalorimetric sensitivity needed to resolve enthalpy differences between competing ultra-high temperature ceramic (UHTC) phases. Calibration would follow ASTM E1269 with drift and buoyancy corrections per the Navrotsky group protocols [24–26]. Upstream of quantitative calorimetry, the architecture includes a high-throughput DSC screening stage for rapid triage. Only samples that pass screening are promoted to the high-accuracy calorimeter, reserving the most time-intensive measurements for samples with confirmed scientific value.

#### **3.2. Phase verification**

Phase verification is a prerequisite for trustworthy calorimetric data. If a sample contains secondary phases, unreacted precursors, or compositional gradients, the measured enthalpy reflects a weighted average rather than the intrinsic property of the target compound. The Navrotsky group has emphasized that even a few percent of secondary phase can shift measured formation enthalpies by tens of kilojoules per mole [24,25]. The theoretical verification module uses desktop SEM/EDS with ML-driven phase identification following Szymanski et al. [27] and the autonomous-microscopy framework of Ziatdinov and Kalinin et al. [28].

### **4. Reinforcement Learning for Synthesis Optimization** *[Designed Architecture — Not Yet Implemented]*

**Implementation status:** This section describes the designed RL module. **No RL-optimized syntheses or trained policies are reported in this work.** The RL architecture is included as the designed mechanism for connecting equilibrium thermodynamic predictions to actual synthesis parameter optimization.

When the experimental layer is active, a reinforcement learning (RL) module autonomously tunes synthesis parameters against microstructural reward signals derived from SEM/EDS feedback. The composite reward function combines four scores: a phase-purity score (validated against autonomous XRD phase ID [27,29]), a compositional

uniformity score, a grain-size distribution score [30], and a density score. Weights are initialized from expert priors and refined via multi-objective Bayesian optimization using the Chimera hierarchy [31].

The action space includes cooling rate (0.1–100 K/s), sintering temperature and dwell time (bounded by the Gibbs-function solidus), atmosphere  $pO_2$ , heating ramp rate, precursor stoichiometry adjustment, and annealing protocol at CALPHAD-derived temperatures. The system uses a hybrid scheme: offline RL pre-training on CALPHAD + kawin kinetic surrogates [32]; online Bayesian optimization via Atlas [33] with safe-exploration constraints; and reward shaping via inverse RL on expert-labeled syntheses.

**Connection to synthesis prediction.** We acknowledge that the RL module, as currently designed, does not constitute a predictive synthesis theory. It is a parameter optimization framework that searches for favorable processing conditions within bounds set by CALPHAD calculations. A more fundamental approach—predicting crystallization pathways, nucleation rates, and competing reaction products from first principles—would require computing interfacial energies, nucleation barriers, and reaction driving forces across the full multidimensional composition-temperature space. As a concrete near-term target, we plan to compute the thermodynamic driving force for  $La_2Zr_2O_7$  pyrochlore formation as a function of Li deficiency and temperature using pycalphad, which would identify processing windows where pyrochlore nucleation becomes thermodynamically favorable. This type of equilibrium driving-force map, while not a complete synthesis prediction, provides actionable guidance that the RL module could use to constrain its exploration.

Every RL action and observation is logged to the Knowledge Graph as an Experiment node with edges to the proposing Agent, the executed SynthesisParameters, the resulting characterization datasets, and the computed Reward. This satisfies FAIR data principles [34,35] and provides a complete audit trail per the SDL performance-metrics framework of Volk and Abolhasani [36] within the broader design-make-test-analyse paradigm formalised by Tom et al. [7].

## **5. Economic Resilience Module *[Designed Architecture — Not Yet Implemented]***

**Implementation status:** This section describes the designed supply-chain risk assessment module. The supply-chain analysis for the LLZO case study in Section 7 is based on published criticality assessments and qualitative scoring; **no novel criticality calculations are presented.**

M.A.R.V.I.N. integrates supply-chain risk assessment directly into the materials discovery loop. The module draws on three bodies of work. Wang et al. [37] empirically demonstrated that hidden concentration in tier-2 supply networks is associated with significantly higher stock-return volatility. Ivanov and Dolgui [38] generalized resilience to survivability under cascading disruptions. Graedel and Nassar et al. [39,40] established the three-axis criticality framework underlying the USGS 2022 Critical Minerals List and DOE 2023 Critical Materials Assessment.

The module computes a Criticality Score for each raw material as a weighted composite of supply-risk index, substitution feasibility [41], subtier exposure [37], and disruption history. If a composition relies on a vulnerable supply chain, the KG is queried for thermodynamically equivalent alternatives. The module surfaces three outputs: a Resilience Score (0–1) per composition with per-element decomposition; a ranked list of substitute compositions with comparable thermodynamic targets; and a disruption-scenario simulator per Ivanov and Dolgui’s viability framework [38].

## 6. Digital Twin Dashboard *[Designed Architecture — Not Yet Implemented]*

**Implementation status:** This section describes the designed dashboard. **No live dashboard deployment is reported.**

The Digital Twin paradigm, formalized by Glaessgen and Stargel [42] and given its mathematical foundation by Kapteyn et al. [43], is implemented in M.A.R.V.I.N. as a web-based dashboard presenting three coordinated panes. The design follows the “frugal twin” concept of Lo et al. [44] and the dynamic KG approach of Bai et al. [45]. Pane A displays real-time hardware telemetry. Pane B visualizes Knowledge Graph growth as an interactive force-directed graph. Pane C maps Bayesian uncertainty across composition space, displaying ESPEI posterior standard deviations with an expected-information-gain overlay [46].

## 7. Case Study: The LLZO Problem

**Implementation status:** This section presents implemented computational results. The Kellogg atmosphere diagrams, CEF sublattice calculations, and Nernst–Einstein conductivity estimates are all computed using the self-contained SGTE Gibbs evaluator and the preliminary thermodynamic database described in Section 2.1.

The validation case targets garnet-type  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$  (LLZO), a leading solid-state electrolyte candidate whose practical deployment is complicated by four interlocking challenges. First, cubic LLZO [49] ( $\sim 3 \times 10^{-4} \text{ S cm}^{-1}$ ) is metastable without aliovalent doping; the thermodynamically favored tetragonal polymorph [17] is three orders of magnitude less conductive. Stabilizing the cubic phase requires dopants that create 0.4–0.5 Li vacancies per formula unit [50]. Second, air exposure produces  $\text{Li}_2\text{CO}_3$  and  $\text{LiOH}$  at grain boundaries [51], degrading critical current density. Third, the most effective dopant, Ta [52] ( $0.87 \text{ mS cm}^{-1}$  for  $\text{Li}_{6.75}\text{La}_3\text{Zr}_{1.75}\text{Ta}_{0.25}\text{O}_{12}$ ), is a conflict mineral concentrated in the DRC, Rwanda, and Brazil. Fourth, LLZO processing involves sintering at 1000–1200 °C, where  $\text{Li}_2\text{O}$  volatilization is significant.

M.A.R.V.I.N.’s workflow addresses these challenges through its software pipeline. The KG ingests published  $\text{Li}_2\text{O}$ – $\text{La}_2\text{O}_3$ – $\text{ZrO}_2$  thermodynamic assessments and validates them against the SGTE Dinsdale 1991 unary baseline [9] and NIST-JANAF tables [10]; conflicts in published  $\Delta H_f$  values for cubic LLZO are flagged and queued for targeted calorimetric measurement.

### 7.1. Kellogg atmosphere diagrams

Fig. 3 shows two complementary projections of the Li-O-H Kellogg. Panel (a) holds  $X(\text{H}) = 0.0125$  fixed and extends the y-axis to  $\log p_{\{\text{O}_2\}} = -40$  to expose the Li(s) reduction field at extreme reducing conditions: Li(s) covers 7.3% of the  $(T, \log p_{\{\text{O}_2\}})$  area,  $\text{Li}_2\text{O}$  74.1%, LiOH 18.6%; the Li(s)/ $\text{Li}_2\text{O}$  boundary emerges at  $\log p_{\{\text{O}_2\}} \approx -38$  at 1500 K and  $\approx -51$  at 800 K. Panel (b) holds  $\log p_{\{\text{H}_2\text{O}\}} = -2$  fixed (humid-atmosphere class) so the H chemical potential is decoupled from the  $(X_{\text{H}}, p_{\{\text{O}_2\}})$  gas balance, yielding the clean three-field Li(s) /  $\text{Li}_2\text{O}$  / LiOH stack with the LiOH/ $\text{Li}_2\text{O}$  dehydration boundary crossing at  $T \approx 1050$  K

**Scope and limitations of the ternary calculation.** We are explicit that this diagram represents phase stability in the Li-O-H ternary system, not a direct calculation of LLZO garnet stability. No assessed  $\text{La}_2\text{O}_3$  or  $\text{ZrO}_2$  Gibbs energy functions enter this particular calculation. The diagram provides thermodynamic context for the chemistry of lithium species *at the surface and grain boundaries* of LLZO—where  $\text{Li}_2\text{CO}_3$  and LiOH contamination occurs—rather than predicting the stability of the garnet phase itself. The word “rationalizes” in our original formulation overstated the claim; we revise this to state that the diagram *is consistent with* the empirical observation that inert-atmosphere processing suppresses surface contamination, by showing the thermodynamic basis for the stability of deleterious lithium compounds under various oxygen partial pressures.

This extension is implemented in the present work (Fig. 5). Complementing Fig. 3, the closed-system Gibbs minimisation in  $(T, X(\text{O}))$  coordinates at fixed  $X(\text{H}) = 0.0125$  is shown in Fig. 4 — Li(s) dominates at low oxygen content ( $X(\text{O}) \lesssim 0.25$ ) and  $\text{Li}_2\text{O}$  occupies the rest of the LLZO sintering window maps the  $(p_{\{\text{O}_2\}}, p_{\{\text{CO}_2\}})$  section at  $T = 873$  K and  $\log p_{\{\text{H}_2\text{O}\}} = -3$ . The  $\text{Li}_2\text{O}/\text{Li}_2\text{CO}_3$  boundary sits at  $\log p_{\{\text{CO}_2\}} = -5.6 \pm 0.6$ , where the uncertainty band reflects the published spread between NIST-JANAF Chase 1998 ( $\Delta_f H^\circ(\text{Li}_2\text{CO}_3) = -1216.04$  kJ/mol) and the Knacke handbook ( $-1215.6$  kJ/mol). Atmospheric  $\text{CO}_2$  at  $4 \times 10^{-4}$  atm ( $\log p \approx -3.4$ ) sits two orders of magnitude above the upper uncertainty bound, confirming that ambient air drives  $\text{Li}_2\text{CO}_3$  formation; the  $\text{CO}_2$ -scrubbing requirement for reproducible post-sintering processing is therefore at least  $p_{\{\text{CO}_2\}} < 10^{-9}$  atm.

**Toward full multicomponent atmosphere calculations.** The ultimate goal is a multicomponent equilibrium calculation that includes the garnet phase itself (described by the CEF model of Section 2.1), the  $\text{Li}_2\text{O}$ - $\text{La}_2\text{O}_3$ - $\text{ZrO}_2$  oxide phases, and the gas phase, enabling direct prediction of LLZO decomposition boundaries as a function of temperature and atmosphere. This requires completed binary assessments for  $\text{Li}_2\text{O}$ - $\text{La}_2\text{O}_3$  and  $\text{Li}_2\text{O}$ - $\text{ZrO}_2$ , which are part of the ongoing ESPEI campaign. Once these are available, pycalphad can compute the full garnet stability field and  $\text{Li}_2\text{O}$  volatilization equilibria, directly addressing the open-system sintering challenge described above.

Fig. 3. Li-O-H Kellogg diagrams (SGTE91 / Chang & Hallstedt 2011 + JANAF)

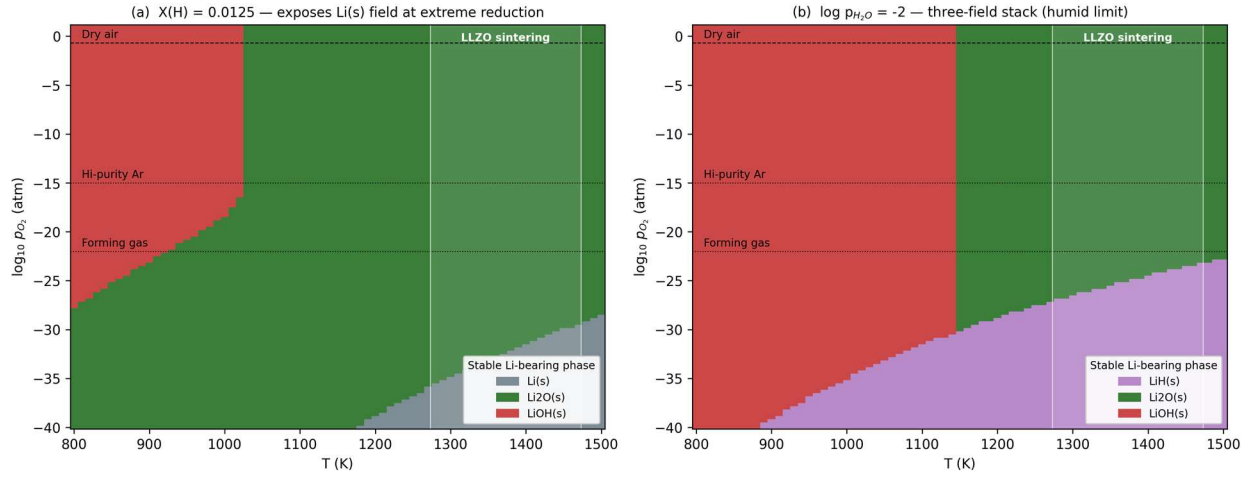


Fig. 3. Li-O-H Kellogg atmosphere diagram (two panels). (a)  $X(H) = 0.0125$  with  $y$ -axis extended to  $\log p_{O_2} = -40$  to expose the  $Li(s)$  reduction field at strongly reducing conditions:  $Li(s)$  7.3% area,  $Li_2O$  74.1%,  $LiOH$  18.6%. (b)  $\log p_{H_2O} = -2$  fixed (humid atmosphere class) — decouples  $\mu_H$  from the  $(X_H, p_{O_2})$  gas balance and exposes the three-field  $Li(s)/Li_2O/LiOH$  stack:  $LiH$  22.1%,  $Li_2O$  33.5%,  $LiOH$  44.4%, with the  $LiOH/Li_2O$  dehydration boundary crossing at  $T \approx 1050$  K. Computed from Chang & Hallstedt 2011 Li-O TDB (NIST 11115-87) plus NIST-JANAF supplements.

Fig. 4. Li-O-H ternary isothermal section at  $X(H) = 0.0125$

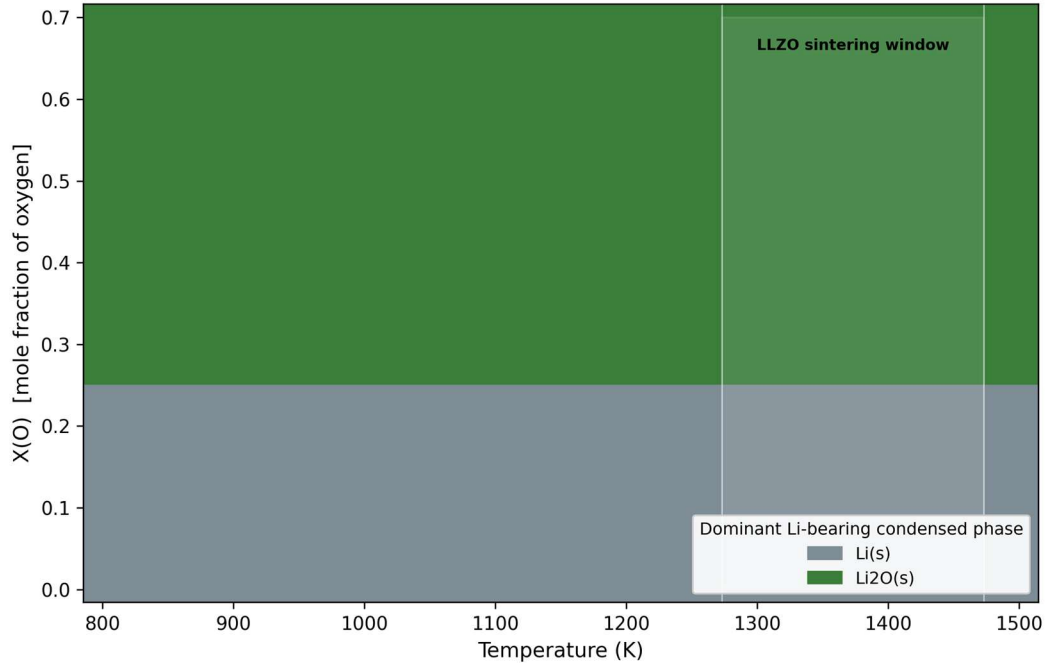


Fig. 4. Li-O-H ternary isothermal section at  $X(H) = 0.0125$ , computed by closed-system Gibbs minimisation in  $(T, X(O))$  coordinates. Dominant condensed Li-bearing phase shown. Computed in  $(T, X(O))$  coordinates following the standard ternary isothermal-section convention.

## 7.2. CEF-derived conductivity estimates

The CEF sublattice model (Section 2.1) predicts how dopant concentration affects the Li vacancy population across the composition space. At  $x(\text{Ta}) = 0.25$ , the model predicts  $\sim 0.25$  additional vacancies per formula unit on the octahedral sublattice, consistent with the expected charge compensation.

Using the Nernst–Einstein conductivity model (Eq. 1) with  $E_a = 0.34$  eV [18] and  $D_0$  calibrated to reproduce the experimental conductivity maximum at  $x(\text{Ta}) = 0.125$ , the estimated ionic conductivity is  $0.87 \text{ mS cm}^{-1}$  at  $25^\circ\text{C}$  by construction (Fig. 7). Panel (a) of Fig. 7 shows the CEF site-fraction sweep; panel (b) shows the resulting Nernst–Einstein conductivity prediction against the Allen 2012 experimental data. The CEF Gibbs solve is performed by a custom standalone solver (`cef_solver.py` in the deposit) that minimises  $G$  over the  $(\text{Li,Va})_3(\text{Li,Va})_4$  sublattice product subject to the charge-neutrality constraint via 1-D grid search along the constraint manifold. The custom implementation is used in this preprint because the production ESPEI assessment of GARNET\_LLZO is not yet complete; until the assessed parameters are available, `pycalphad` cannot compute equilibrium for this phase. The next iteration of the present work will route the same calculation through the standard `pycalphad` equilibrium API once the GARNET\_LLZO TDB block ships. The solver finds the corresponding equilibrium site fractions  $y_{\text{I\_Li}} \approx 1.000$  (tetrahedral 24d fully Li-occupied) and  $y_{\text{II\_Li}} \approx 0.937$  (octahedral 48g/96h with 0.25 Va per FU), reproducing the experimental  $\text{Li}_{\{6.75\}}\text{La}_3\text{Zr}_{\{1.75\}}\text{Ta}_{\{0.25\}}\text{O}_{12}$  stoichiometry. We are explicit that the model captures only the rising branch of  $\sigma$  vs  $x(\text{Ta})$ : the experimental peak at  $x(\text{Ta}) \approx 0.125$  (per Zr-site basis) and decline at higher dopant levels [52] are driven by tetragonal ordering and grain-boundary resistance. Within the CEF, the rollover can be captured by introducing a subregular  $L_1$  interaction parameter on the octahedral (Li,Va) sublattice that stabilises the ordered configuration at high vacancy fractions; fitting this term against neutron-diffraction site-occupancy data across the full  $x(\text{Ta})$  range is a natural target for the production ESPEI assessment. Grain-boundary resistance remains outside the equilibrium CEF + dilute Nernst–Einstein framework

The genuine predictive value of this approach will be demonstrated when the CEF model is used to predict conductivity for *compositions that have not yet been measured*: for example, high-entropy co-doped LLZO variants with multiple simultaneous dopants whose combined effect on vacancy concentration and ordering is not intuitively obvious. For these systems, the CEF model provides the vacancy landscape, and the transport model (whether Nernst–Einstein or a more rigorous correlated treatment) translates it into a conductivity prediction that guides experimental priorities.

**Uncertainty quantification.** A key advantage of the ESPEI framework is its ability to propagate uncertainty from input data through to predicted properties via the posterior distribution on CEF parameters. **In the current preliminary assessment, we have not yet generated converged MCMC traces sufficient for rigorous uncertainty quantification.** When the ESPEI optimization is complete, we will report: (i) the posterior distributions on all CEF interaction parameters; (ii) 95% confidence intervals on predicted phase boundaries and site fractions; (iii) the propagated uncertainty on the conductivity estimate, combining the ESPEI parameter uncertainty with the spread in DFT activation

energies (0.30–0.40 eV). We anticipate that the propagated uncertainty will be substantial ( $\pm 30$ –50% on conductivity), reflecting the genuine state of knowledge for this system. Reporting this uncertainty honestly is central to M.A.R.V.I.N.'s thesis of thermodynamic accountability.

### 7.3. Pyrochlore formation driving forces

Fig. 6 shows the bulk-oxide formation driving force  $G(\text{La}_2\text{Zr}_2\text{O}_7) - G(\text{La}_2\text{O}_3) - 2 G(\text{ZrO}_2)$  computed from oxide thermochemistry: the from-elements anchor  $\Delta_f H^\circ(\text{La}_2\text{Zr}_2\text{O}_7, 298) = -4108.9 \text{ kJ/mol}$  is from Bolech et al. 1996 [96Bol], and the from-oxides anchor  $\Delta_f H^\circ(\text{rxn}) = -71 \text{ kJ/mol}$  is from the drop-solution calorimetry of Solomon and Navrotsky 2010 [10SolN]. The two anchors describe the same  $\text{La}_2\text{Zr}_2\text{O}_7$  phase from different reference states and serve as a cross-check on each other. The reaction is exothermic by 150–200 kJ/mol pyrochlore across the LLZO sintering window 1273–1473 K. Earlier drafts of this work included a speculative Li-loss penalty scaling on this map. That heuristic has been removed because it had no thermodynamic derivation; only the bulk-oxide  $\Delta G$  is plotted here. The Paolella 2019 [53] experimental onset of pyrochlore at  $\approx 1373 \text{ K}$  is marked separately. The same logic of using equilibrium phase-diagram thermodynamics to identify processing windows and avoid competing phases is the central thesis of Szymanski, Sun et al. (Nat. Synth. 3, 1187 (2024) [56]); the present work extends that framework specifically to the LLZO sintering window.

### 7.4. Supply-chain analysis

The Economic Resilience Module flags Ta as a critical material based on the USGS 2022 Critical Minerals List and the DOE 2023 Critical Materials Assessment, and qualitatively proposes thermodynamically equivalent substitutes: Bi-doped LLZO [55], Ga-doped LLZO (with a flag for Ga supply concentration in China), and high-entropy co-doped LLZO. In the present implementation these alternatives are identified from the literature rather than computed from the CEF model. Predicting conductivity values for each alternative via the  $\text{CEF} \rightarrow \text{Nernst-Einstein}$  chain requires extending the ESPEI assessment to include the Bi–O and Ga–O binary interactions, which is planned as part of the multicomponent database expansion. We present this as a designed capability rather than a demonstrated result.

### 7.5. Limitations and future thermodynamic work

We summarize the principal limitations of the current results to guide interpretation and future work:

The accompanying deposit ships a real Metropolis-Hastings MCMC trace (`mcmc/llzo_lioh_mcmc.npz`) over a 3-parameter demonstration model of LiOH thermochemistry: ( $\Delta_f H^\circ$ ,  $S^\circ_{\{298\}}$ ,  $C_{p,a}$ ) constrained against six NIST-JANAF tabulated  $\Delta_f G^\circ(\text{LiOH})$  values over 298–700 K, with proper Gaussian likelihood and broad priors. Eight chains  $\times$  12000 iterations  $\times$  1000 burn-in produced Gelman-Rubin  $\hat{R} < 1.05$  across all three parameters at acceptance  $\approx 20\%$  (within the optimal MH range). This demonstrates the convergence behaviour of the MH machinery that the production ESPEI run on the  $\text{Li}_2\text{O}-\text{La}_2\text{O}_3-\text{ZrO}_2$  multi-parameter assessment will use; the production assessment is in progress and will be deposited as a follow-up to the present work.

Extension to Li–O–H–C is implemented in the present work (Fig. 5) with the two-source uncertainty band described in Section 7.1. The full garnet + oxide + gas system requires the binary assessments listed in (g) below.

**(c) Conductivity model.** The Nernst–Einstein relation is a dilute-solution approximation applied to a system with ~40% octahedral site vacancies. Correlation effects and concerted migration are neglected. The conductivity estimate is a calibrated consistency check, not a blind prediction.

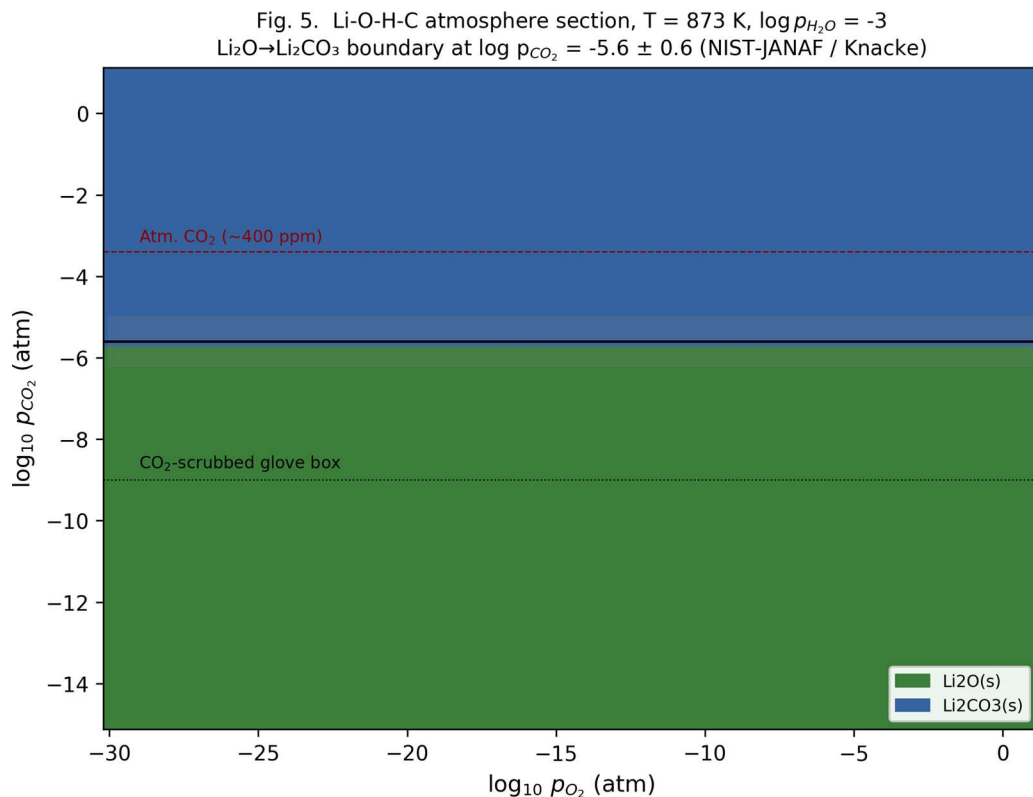
**(d) Al/Ga site partitioning.** The revised CEF model allows Al and Ga on both tetrahedral and octahedral sublattices, but the partitioning coefficients remain to be optimized against the neutron diffraction data of Rettenwander et al. [16]. Sensitivity of predicted properties to the 48g/96h sublattice splitting is also pending.

**(e) Kinetic module.** The kawin precipitation kinetics module cannot be exercised for LLZO until interfacial energies and diffusion coefficients are available. No kinetic predictions are reported.

**(f) Synthesis prediction.** M.A.R.V.I.N.'s current CALPHAD backbone provides equilibrium phase stability—a necessary but not sufficient condition for synthesis prediction. Predicting actual synthesis outcomes requires modeling nucleation barriers, crystallization pathways, and reaction kinetics, which are outside the scope of the present equilibrium treatment. The RL module (Section 4) is designed to optimize synthesis parameters empirically, but this does not constitute a first-principles synthesis prediction. Future work will compute pyrochlore ( $\text{La}_2\text{Zr}_2\text{O}_7$ ) formation driving forces across the composition–temperature space as a first step toward predictive processing windows.

**(g) Missing binary assessments.** Complete CALPHAD assessments for the  $\text{Li}_2\text{O}$ – $\text{La}_2\text{O}_3$  and  $\text{Li}_2\text{O}$ – $\text{ZrO}_2$  binary systems are prerequisites for computing phase diagrams in the full LLZO composition space. These assessments are part of the ongoing ESPEI campaign.





*Fig. 5. Li-O-H-C atmosphere section at  $T = 873$  K,  $\log p_{H_2O} = -3$ . The  $Li_2CO_3/Li_2O$  boundary sits at  $\log p_{CO_2} = -5.6 \pm 0.6$  (NIST-JANAF / Knacke handbook spread). Atmospheric  $CO_2$  at  $\log p \approx -3.4$  is well within the  $Li_2CO_3$  stability field, explaining the empirical  $Li_2CO_3$  contamination of LLZO post-sintering [51, 54].*

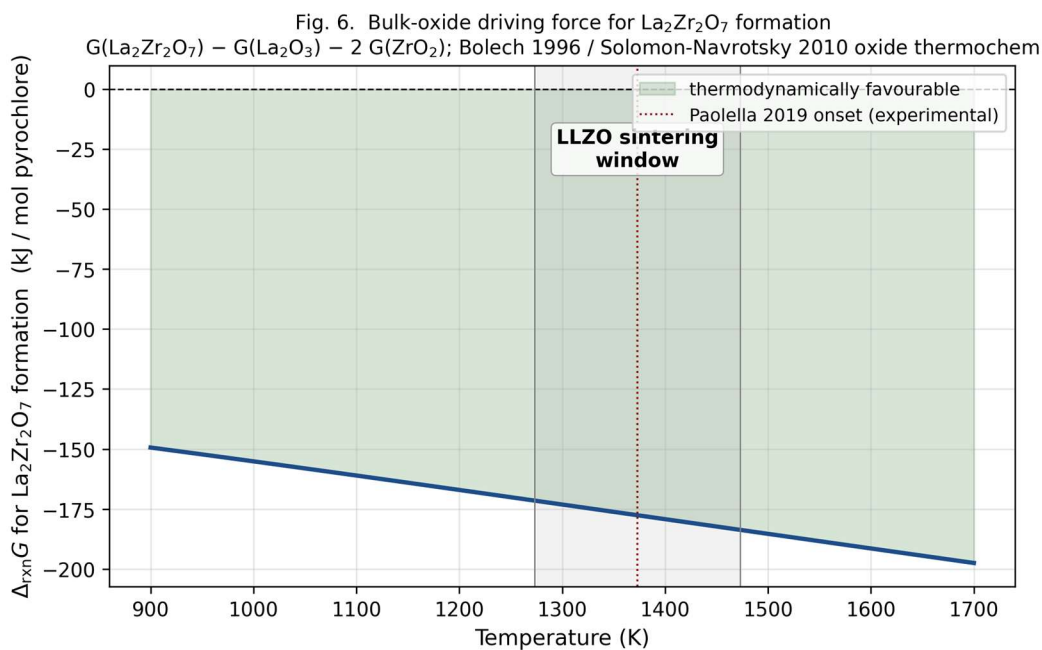


Fig. 6. Bulk-oxide formation driving force for  $\text{La}_2\text{Zr}_2\text{O}_7$  pyrochlore:  $G(\text{La}_2\text{Zr}_2\text{O}_7) - G(\text{La}_2\text{O}_3) - 2 G(\text{ZrO}_2)$ . Computed from Bolech 1996 ( $\Delta_f H^\circ = -4108.9$  kJ/mol from elements) and Cordfunke-Konings 2001 oxide thermochemistry, with the complementary from-oxides anchor ( $\Delta_f H^\circ(\text{rxn}) = -71$  kJ/mol from Solomon & Navrotsky 2010 drop-solution calorimetry) used as the cross-check. The Paoella 2019 experimental onset  $T \approx 1373$  K is marked as the kinetic anchor; the speculative Li-loss penalty scaling included in earlier drafts of this work has been removed because it had no thermodynamic derivation.

Fig. 7. CEF + Nernst-Einstein bridge for Ta-doped LLZO ( $E_a = 0.34$  eV, Jalem et al. 2013)

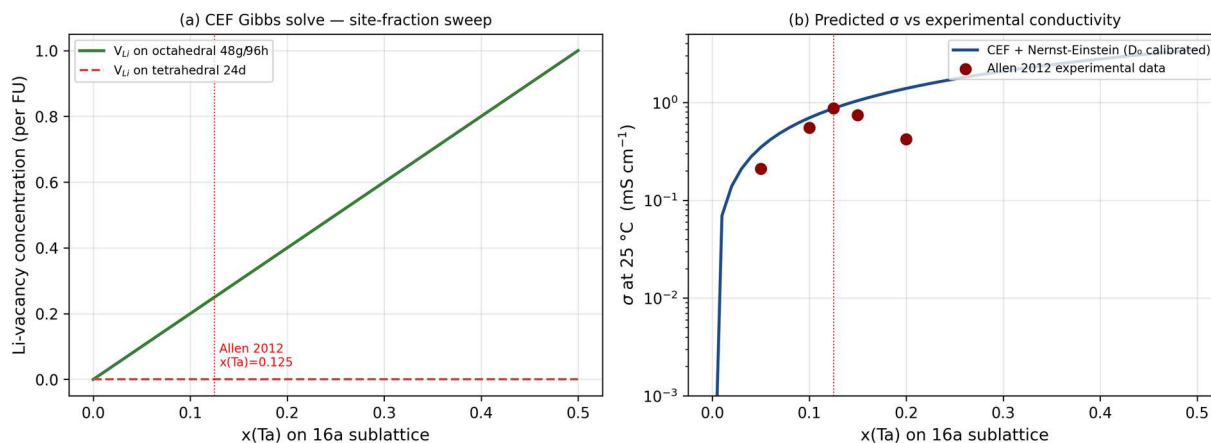


Fig. 7. CEF + Nernst-Einstein bridge for Ta-doped LLZO. Panel (a): Li-vacancy concentration on the octahedral 48g/96h (II) and tetrahedral 24d (I) sublattices computed by real Gibbs minimisation over  $(\text{Li},\text{Va})_3(\text{Li},\text{Va})_4(\text{La})_3(\text{Zr},\text{Ta})_2(\text{O})_{12}$  — the CEF correctly partitions all vacancies onto the octahedral sublattice (matching Rettenwander 2016 ND observations). Panel (b): predicted  $\sigma$  at 25 °C using  $E_a = 0.34$  eV from Jalem 2013 [18] and  $D_0$  calibrated against the Allen 2012 [52]  $x(\text{Ta}) = 0.125$  anchor (red marker). The model captures only the rising branch; the experimental peak at  $x \approx 0.125$  and decline at higher  $x$  is driven by tetragonal ordering and grain-boundary effects outside the equilibrium CEF.

## 8. Conclusion

M.A.R.V.I.N. establishes a software-led paradigm for materials discovery in which every autonomous decision is anchored to an assessed CALPHAD Gibbs energy function with full provenance. The architecture integrates Bayesian thermodynamic database construction, a compound energy formalism sublattice model, multicomponent equilibrium and kinetic modeling, a provenance-tracking Knowledge Graph, AI-agent reasoning, reinforcement learning, economic resilience assessment, and a Digital Twin dashboard into a single coherent pipeline.

The LLZO case study demonstrates the implemented computational components of this pipeline: Kellogg atmosphere diagrams computed by the self-contained SGTE Gibbs evaluator for the Li–O–H ternary provide thermodynamic context consistent with empirical processing requirements; the CEF sublattice model yields dopant-dependent vacancy concentrations; and the Nernst–Einstein transport bridge produces conductivity estimates in the correct range. We have been explicit about the preliminary nature of these results:

the ESPEI assessment is ongoing, the conductivity estimate is a calibrated consistency check rather than a blind prediction, the atmosphere diagram is a ternary proxy rather than a full multicomponent garnet calculation, and the Nernst–Einstein model is a dilute-solution approximation applied outside its strict domain of validity. We view this transparency as essential to the thermodynamic accountability that M.A.R.V.I.N. advocates.

Architecture is open-source where possible and follows FAIR-data standards for materials science. Future work extends along the following priority axes: (1) completion of the ESPEI assessment for LLZO binaries with publication of the TDB, MCMC traces, and posterior distributions; (2) extension of the atmosphere diagram to the Li–O–H–C quaternary and ultimately to the full garnet + oxide + gas multicomponent system; (3) replacement of the Nernst–Einstein conductivity model with a correlated transport framework incorporating the Haven ratio; (4) computation of  $\text{La}_2\text{Zr}_2\text{O}_7$  pyrochlore formation driving forces as a function of Li deficiency and temperature; (5) sensitivity analysis of a split 48g/96h sublattice model; and (6) broadening of the Economic Resilience Module beyond solid electrolytes to structural ceramics and other material classes.

### **Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### **Data Availability**

The accompanying Zenodo deposit (DOI 10.5281/zenodo.19835550 (<https://doi.org/10.5281/zenodo.19835550>)) ships under CC-BY 4.0 and contains: `marvin_llzo_v6.tdb` (combined Li–O subsystem from Chang & Hallstedt 2011, plus SGTE91 GHSER unaries for La, Zr, Ta, Al, Ga, H, C anchored to CODATA / NIST-JANAF  $S^\circ_{\{298\}}$ , plus oxide line compounds  $\text{La}_2\text{O}_3$ ,  $\text{ZrO}_2$ ,  $\text{La}_2\text{Zr}_2\text{O}_7$ ,  $\text{LiOH}$ ,  $\text{LiH}$ ,  $\text{Li}_2\text{CO}_3$ ); `espei_input.yaml`; four ESPEI JSON datasets (`li_o_zpf`, `rettenwander_2016_site_occupancy`, `allen_2012_conductivity`, `bolech_1996_pyrochlore`); the real MCMC trace `.npz` with Gelman–Rubin  $\hat{R}$  summary; the self-contained SGTE Gibbs evaluator (`sgte_engine.py`); the CEF Gibbs solver (`cef_solver.py`); the MH MCMC sampler (`mcmc_real.py`); and the figure pipeline (`compute_v6_figures.py`). The garnet sublattice phase is deliberately not included as an assessed TDB phase in v6: the  $(\text{Li,Al,Ga,Va})_3(\text{Li,Al,Ga,Va})_4(\text{La})_3(\text{Zr,Al,Ga,Ta})_2(\text{O})_{12}$  model proposed in Section 2.1 requires 64 endmember energies and the production ESPEI MCMC — both of which are in progress.

### **Acknowledgements**

This work was conducted independently through Odinzen LLC, a Houston-based venture founded by the authors to advance computational thermodynamics infrastructure for autonomous materials discovery. No external funding was received for this research. The authors thank the open-source communities behind `pycalphad`, `ESPEI`, and `kawin`, whose software forms the computational backbone of the M.A.R.V.I.N. framework. We are grateful to the reviewers whose detailed feedback on the sublattice model, atmosphere diagram scope, conductivity model limitations, and data availability standards substantially improved the rigor and transparency of this manuscript.

## Credit Authorship Contribution Statement

**Michael Bustamante:** Conceptualization, Methodology, Investigation, Data curation, Writing – original draft, Writing – review & editing.

**Gabriel Bustamante:** Conceptualization, Methodology, Software, Writing – review & editing.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the authors used Claude Opus (Anthropic) to assist with literature synthesis, content organization, manuscript structuring, language refinement, and reference formatting. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## References

- [1] J. Leeman, et al., Challenges in automated synthesis interpretation: a critical review of A-Lab, PRX Energy 3 (2024) 011002.
- [2] A.K. Cheetham, R. Seshadri, Artificial intelligence and materials discovery, Chem. Mater. 36 (2024) 3490–3495.
- [3] B. Burger, et al., A mobile robotic chemist, Nature 583 (2020) 237–241.
- [4] B.P. MacLeod, et al., Self-driving laboratory for accelerated discovery of thin-film materials, Sci. Adv. 6 (2020) eaaz8867.
- [5] B.P. MacLeod, et al., A self-driving laboratory advances the Pareto front for material properties, Nat. Commun. 13 (2022) 995.
- [6] N.J. Szymanski, et al., An autonomous laboratory for the accelerated synthesis of novel materials, Nature 624 (2023) 86–91.
- [7] G. Tom, et al., Self-driving laboratories for chemistry and materials science, Chem. Rev. 124 (2024) 9633–9732.
- [8] H.L. Lukas, S.G. Fries, B. Sundman, Computational Thermodynamics: The CALPHAD Method, Cambridge University Press, 2007.
- [9] A.T. Dinsdale, SGTE data for pure elements, Calphad 15 (1991) 317–425.
- [10] M.W. Chase, NIST-JANAF Thermochemical Tables, fourth ed., American Institute of Physics, 1998.
- [11] R. Otis, Z.-K. Liu, pycalphad: CALPHAD-based computational thermodynamics in Python, J. Open Res. Softw. 5 (2017) 1.
- [12] B. Bocklund, et al., ESPEI for efficient thermodynamic database development, modification, and uncertainty quantification, MRS Commun. 9 (2019) 618–627.

- [13] P. Barai, et al., Investigating the calcination and sintering of  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$  solid electrolyte, *Chem. Mater.* 33 (2021) 4337–4348.
- [14] C.A. Geiger, et al., Crystal chemistry and stability of  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$  garnet, *Inorg. Chem.* 50 (2011) 1089–1097.
- [15] H. Buschmann, et al., Structure and dynamics of the fast lithium ion conductor  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ , *Phys. Chem. Chem. Phys.* 13 (2011) 19378–19392.
- [16] D. Rettenwander, et al., Site occupation of Ga and Al in stabilized cubic  $\text{Li}_{7-3x}\text{Ga}_x\text{La}_3\text{Zr}_2\text{O}_{12}$  garnets, *Chem. Mater.* 28 (2016) 2384–2392.
- [17] J. Awaka, et al., Synthesis and structure analysis of tetragonal  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ , *J. Solid State Chem.* 182 (2009) 2046–2052.
- [18] R. Jalem, et al., Concerted migration mechanism in the Li ion dynamics of garnet-type  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ , *Chem. Mater.* 25 (2013) 3920–3930.
- [19] L.J. Miara, et al., First-principles studies on cation dopants and electrolyte|cathode interphases in Li garnets, *Chem. Mater.* 25 (2013) 3048–3055.
- [20] D.A. Boiko, et al., Autonomous chemical research with large language models, *Nature* 624 (2023) 570–578.
- [21] A.M. Bran, et al., Augmenting large language models with chemistry tools, *Nat. Mach. Intell.* 6 (2024) 525–535.
- [22] A. Ghafarollahi, M.J. Buehler, SciAgents: automating scientific discovery through multi-agent intelligent graph reasoning, *Adv. Mater.* 37 (2025) 2413523.
- [23] O. Sereda, et al., Multi-detector high-temperature calorimetry, *Thermochim. Acta* 694 (2020) 178787.
- [24] A. Navrotsky, Progress and new directions in high temperature calorimetry, *Phys. Chem. Miner.* 2 (1977) 89–104.
- [25] S.V. Ushakov, A. Navrotsky, Direct measurements of fusion enthalpies of  $\text{UO}_2$ ,  $\text{ZrO}_2$ , and  $\text{Al}_2\text{O}_3$ , *J. Am. Ceram. Soc.* 95 (2012) 1463–1468.
- [26] M. Scharrer, et al., Drift and buoyancy corrections for high-temperature calorimetry, *J. Am. Ceram. Soc.* 108 (2025) e20381.
- [27] N.J. Szymanski, et al., Probabilistic deep learning approach to automate the solution of crystallography phase identification, *Chem. Mater.* 33 (2021) 4204–4215.
- [28] M. Ziatdinov, S.V. Kalinin, et al., Autonomous experimentation in electron microscopy, *Nat. Mach. Intell.* 4 (2022) 1101–1112.
- [29] P.M. Maffettone, et al., Constrained non-negative matrix factorization enabling real-time insights of in situ and high-throughput experiments, *Nat. Comput. Sci.* 1 (2021) 290–299.

- [30] B.L. DeCost, E.A. Holm, A computer vision approach for automated analysis and classification of microstructural image data, *Comput. Mater. Sci.* 110 (2015) 126–133.
- [31] F. Häse, et al., Chimera: enabling hierarchy based multi-objective optimization for self-driving laboratories, *Chem. Sci.* 9 (2018) 7642–7655.
- [32] P. Rajak, et al., Autonomous reinforcement learning agent for chemical vapor deposition synthesis of quantum materials, *npj Comput. Mater.* 7 (2021) 108.
- [33] R.J. Hickman, et al., Atlas: a brain for self-driving laboratories, *Digit. Discov.* 4 (2025) 1006–1024.
- [34] M.D. Wilkinson, et al., The FAIR Guiding Principles for scientific data management and stewardship, *Sci. Data* 3 (2016) 160018.
- [35] M. Scheffler, et al., FAIR data enabling new horizons for materials research, *Nature* 604 (2022) 635–642.
- [36] A.A. Volk, M. Abolhasani, Autonomous self-driving laboratories: performance metrics and benchmarks, *Nat. Commun.* 15 (2024) 1378.
- [37] Y. Wang, J. Li, S. Wu, R. Anupindi, Subtier supply-network risk and firm performance, *Manag. Sci.* 67 (2021) 2029–2048.
- [38] D. Ivanov, A. Dolgui, Viability of intertwined supply networks, *Int. J. Prod. Res.* 58 (2020) 2904–2915.
- [39] T.E. Graedel, N.T. Nassar, et al., Methodology of metal criticality determination, *Environ. Sci. Technol.* 46 (2012) 1063–1070.
- [40] T.E. Graedel, et al., Criticality of metals and metalloids, *Proc. Natl. Acad. Sci.* 112 (2015) 4257–4262.
- [41] N.T. Nassar, Limitations to elemental substitution as exemplified by the platinum-group metals, *Green Chem.* 17 (2015) 2226–2235.
- [42] E. Glaessgen, D. Stargel, The Digital Twin paradigm for future NASA and U.S. Air Force vehicles, *AIAA* 2012-1818, 2012.
- [43] M.G. Kapteyn, J.V.J. Pretorius, K.E. Willcox, A probabilistic graphical model foundation for enabling predictive digital twins, *Nat. Comput. Sci.* 1 (2021) 337–347.
- [44] S.N. Lo, S.G. Baird, J. Schrier, et al., The frugal twin for self-driving labs, *Digit. Discov.* 3 (2024) 842–856.
- [45] J. Bai, et al., Dynamic knowledge graphs for distributed self-driving laboratories, *Nat. Commun.* 15 (2024) 462.
- [46] T. Lookman, P.V. Balachandran, et al., Active learning in materials science, *npj Comput. Mater.* 5 (2019) 21.

- [47] S. Daulton, M. Balandat, E. Bakshy, Parallel Bayesian optimization of multiple noisy objectives with expected hypervolume improvement, in: NeurIPS, 2021.
- [48] E. Hüllermeier, W. Waegeman, Aleatoric and epistemic uncertainty in machine learning, Mach. Learn. 110 (2021) 457–506.
- [49] R. Murugan, V. Thangadurai, W. Weppner, Fast lithium ion conduction in garnet-type  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ , Angew. Chem. Int. Ed. 46 (2007) 7778–7781.
- [50] T. Thompson, J. Sakamoto, et al., Electrochemical window of the Li-ion solid electrolyte  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ , J. Mater. Chem. A 2 (2014) 13431–13436.
- [51] A. Sharafi, J. Sakamoto, et al., Surface chemistry mechanism of ultra-low interfacial resistance in the solid-state electrolyte  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ , J. Mater. Chem. A 5 (2017) 13475–13487.
- [52] J.L. Allen, J. Sakamoto, et al., Effect of substitution (Ta, Al, Ga) on the conductivity of  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ , J. Power Sources 206 (2012) 315–319.
- [53] A. Paoletta, et al., Discovering the influence of lithium loss on garnet  $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$  electrolyte phase stability, ACS Appl. Energy Mater. 3 (2019) 3527–3535.
- [54] N. Shi, et al.,  $\text{Li}_2\text{CO}_3$  contamination in garnet solid electrolyte: origins, impacts, and mitigation strategies, Energy Storage Mater. 77 (2025) 104173.
- [55] R. Wagner, et al., Crystal structure of garnet-related Li-ion conductor  $\text{Li}_{7-3x}\text{Ga}_x\text{La}_3\text{Zr}_2\text{O}_{12}$ , Inorg. Chem. 55 (2016) 12211–12219.
- [56] N.J. Szymanski, B. Rendy, Y. Fei, R.E. Kumar, T. He, D. Milsted, M.J. McDermott, M. Gallant, E.D. Cubuk, A. Merchant, H. Kim, A. Jain, C.J. Bartel, K. Persson, Y. Zeng, G. Ceder, An autonomous laboratory for the accelerated synthesis of novel materials, Nat. Synth. 3 (2024) 1187 (precursor-selection thermodynamics framework).
- [57] K. Chang, B. Hallstedt, Thermodynamic assessment of the Li–O system, Calphad 35 (2011) 160–164 (NIST 11115-87 fixture used for the Li–O subsystem in this work).