Design of high-hardness complex concentrated alloys from physics, machine learning, and experiments (EP)

Sharmila Karumuri,[1] (ID) Austin Hernandez,[2] (ID) Saswat Mishra,[3] (ID) Zachary McClure,[3] (ID) Victoria Tucker,[2] (ID)
Joseph C. Flanagan,[2] (ID) Sunghwan Hwang,[2] (ID) Kenneth H. Sandhage,[2] (ID) Ilias Bilionis,[1] (ID) Michael S. Titus,[2] (ID)
and Alejandro Strachan[3,a] (ID)

AFFILIATIONS

[1]School of Mechanical Engineering, Purdue University, West Lafayette, Indiana 47907, USA
[2]School of Materials Engineering, Purdue University, West Lafayette, Indiana 47907, USA
[3]School of Materials Engineering and Network for Computational Nanotechnology, Purdue University, West Lafayette, Indiana 47907, USA

[a]Author to whom correspondence should be addressed: strachan@purdue.edu

ABSTRACT

High-strength alloys are intimately connected to human development, from the bronze age to the current applications in aerospace and energy. State-of-the-art alloys are engineered to harness strengthening mechanisms across scales, from crystal-level processes to complex hierarchical microstructures that are designed to hinder the mobility of dislocations and other carriers of plasticity. In this context, complex concentrated alloys (CCAs) are attractive since they can exhibit very high strength at the single-phase level, which can be further enhanced via incorporation of the second phase and microstructural optimization. Unfortunately, the optimization of CCAs is notoriously difficult due to the high dimensionality of the design space. We demonstrate that a combination of physics-based modeling, machine learning, experimental fabrication, and multi-resolution characterization results in the discovery of the hardest Al-containing BCC-based alloy, surpassing the current state of the art by 31%. Importantly, this is accomplished with only 24 experiments within a design space consisting of 67 536 possible candidates. The approach can be generalized to other alloys, and the resulting materials are of interest in applications ranging from aerospace to nuclear power.

I. INTRODUCTION

Metallic alloys with desirable mechanical properties have been intimately related to human development, from bronze age materials to the advanced alloys used in energy and propulsion applications today. Our current understanding of the mechanisms underlying plastic deformation has enabled the design of alloys with impressive performance. For example, current Ni-based superalloys such as CMSX-10 used as high-pressure turbine blades in aerospace applications exhibit yield strengths around 1000 MPa[1] or 5.8× times greater strengths than that of Duralumin, the Al-based alloy used to fabricate the impellers in the first jet engine, Heinkel HE 3,[2,3] and 1.33× the "first" Ni-based superalloy N80A used in the Power Jets W.2B jet engine.[4,5] Furthermore, incredible gains in turbine entry temperatures have been made from 780 to 1507 °C (1100 °C metal capability) over 70 years,[6] thus helping to reduce specific fuel consumption by 47%.[7,8]

Improvements in strength have been achieved by a combination of composition and microstructure control, whereby the mobility of dislocations can be reduced by intrinsic properties, solute atoms surrounding individual dislocation cores and by extrinsic properties such as precipitates, interfaces, and grain boundaries. Composition control from the 1940s to the early 2000s

has largely been in the form of incremental changes to solute content, but with the advent of high-entropy alloys (HEAs), complex concentrated alloys (CCAs), and multi-principal element alloys (MPEAs), researchers can now explore billions of new compositions. Furthermore, there has been renewed interest in exploring body-centered cubic-based (BCC-based) refractory CCAs (RCCAs) for their potential to surpass the elevated temperature properties of Ni-based superalloys.[9]

Motivated, at least, in part by accelerating the optimization and deployment of CCAs, significant efforts have been devoted to model their thermo-mechanical properties using theory and physics-based simulations, see, for example, Refs. [10] and [11], as well as machine learning (ML).[12,13] In addition, active learning (AL) has been used to optimize the properties of CCAs. This technique uses machine learning models to evaluate and select possible experiments and performs a sequential optimization. AL has been used to drive physics-based simulations[14,15] as well as fabrication and testing experiments.[16,17] For example, Ref. [16] combined machine learning with physics-based simulations to find alloys with low thermal expansion coefficients.

Machine learning methods have been used to accelerate the discovery of alloys. For example, Chang et al.[18] used neural networks to predict candidate alloys with 600HV + hardness in AlCoCrFeMnNi alloys while Jain et al.[19] also used a neural network to predict the hardness albeit in AlCoCrFeMnNbNiV systems. Yang et al.[20] used a combination of multiple linear regression, neural networks, gradient boosting regression, random forests, and support vector regression to filter and predict the hardness of high melting point HEAs. Wen et al.[21] used the Gaussian Process (GP) model with active learning to improve their hardness by 10% in the same AlCoCrFeMnNi alloy design space showing that BGO has been successfully used in the literature to improve the hardness in CCAs. This paper explores the use of CALPHAD simulations to reduce the design space to alloys with the desired phase and a batch approach for experimental selection. Despite these reports, machine learning-guided discovery of ultra-high strength RCCAs remains challenging because of the multi-dimensional design space and the AL model limitations.

To tackle these challenges of difficult-to-model processes and microstructures as well as to ensure an efficient exploration of the desirable composition space is explored efficiently, we use a combination of physics-based simulations to reduce the design space, active learning to efficiently explore the remaining compositions, and surrogate experiments to accelerate data collection (as illustrated in Fig. 1). We use room temperature hardness as a surrogate for high-temperature strength and we restrict the search to Al-containing alloys for oxidation resistance. We note that neither of these guarantees high-temperature performance. We demonstrate, with only 24 experiments organized in six campaigns, the discovery of a BCC-based, intermetallic-free Al-containing alloy, $Al_{35}V_{30}Nb_{20}W_{15}$, which exhibits a hardness (775 $\pm$ 31 HV) 23% larger than any previously reported single-phase four-component RCCA($Cr_{25}Nb_{25}Ti_{25}W_{25}$ with a hardness of 630 HV).[22]

II. ACTIVE LEARNING APPROACH TO ALLOY DESIGN

This section describes our sequential optimization approach that combines machine learning, physics-based simulations,
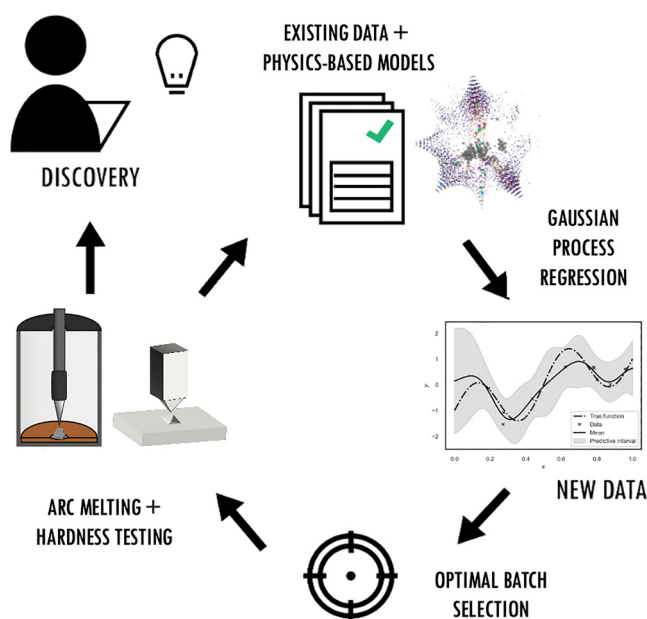

FIG. 1. Active learning approach for the design of high-hardness Al-containing RCCAs.

fabrication, and testing to discover the hardest Al-containing CCA in the fewest possible iterations.

A. Alloy design space using thermodynamic calculations

The initial design space consisted of four-component alloys containing three refractory elements selected from Ti, V, Cr, Zr, Nb, Mo, Hf, Ta, and W plus Al. Using composition step sizes of 5 at. % and restricting Al to a range of 5–40 at. % results in 67 536 distinct alloys across 84 four-component alloy systems. This initial design space is depicted in Fig. 2 after reducing the eight-dimensional space of refractory elements to 2D using multi-dimensional scaling[23] plots. The first iteration of the active learning optimization was performed over this space. However, starting in the second iteration, the possible alloys were restricted using thermodynamic calculations using CALPHAD to those that (i) are expected to be accessible using arc melting and (ii) predicted to form only BCC phases at 1000 °C. For the first condition, we restricted the design space to compositions with predicted liquidus temperatures below the boiling points of Al and Cr (2518 and 2669 °C), respectively.[24,25] The composition used in this step does not include Al or Cr since we first melt the high melting temperature element powders (Ti, V, Zr, Nb, Mo, Hf, Ta, and W) and then add Cr and Al. For the second condition, phase equilibria for all four-component compositions were calculated at 1000 °C, and compositions that exhibited any phase other than disordered BCC phases were excluded. We emphasize that two-phase (or greater) disordered BCC alloys were accepted, but they were not experimentally observed, except for $Al_{35}V_{30}Nb_{15}W_{20}$ in the final batch. In
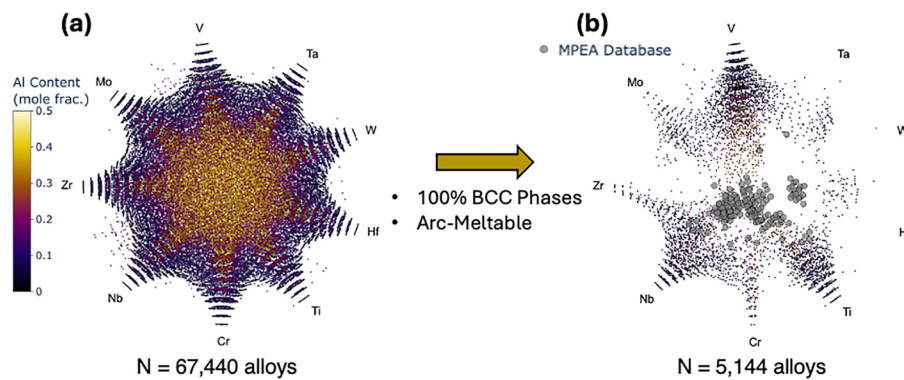
FIG. 2. (a) Full design space (mapped into 2D via MDS) and (b) design space restricted to alloys predicted to be single phase and processable via arc melting.

this case, we observed intermetallic formation, likely from non-equilibrium solidification conditions, as will be discussed in detail in Sec. IV. Beyond reducing the design space by over an order of magnitude to 5144 alloys, see Fig. 2, we observed from the preliminary experiments (not reported here) that this constraint significantly improved the matching between nominal and experimentally measured compositions and reduced the failure rate of arc-melted ingots by preventing significant evaporation of volatile elements and formation of high melting temperature intermetallic phases upon mixing.

In addition to reducing the design space, CALPHAD was used to generate physics-based descriptors of the alloys used as input to the predictive ML model described below. To inform the model with microstructural data, for each composition, we performed single-point equilibrium calculations to determine the equilibrium phases, their volume fraction, density, and their crystal structure at 1000 °C. Additionally, the liquidus, solidus, and highest BCC solvus (if present) temperatures were determined. All thermodynamic CALPHAD calculations were performed using Thermo-Calc (R) 2020a software with the TCHEA4 database implemented via the TC-Python API with the code available in Ref. 26.

B. Active learning

The initial model for the active learning loop was derived from experimental hardness measurements available in a database assembled by Borg *et al.*, downloaded from Citrination of Ref. 27. The initial model was trained with 1545 unique compositions and was later augmented via active learning with our in-house data during the sequential optimization.

Bayesian global optimization (BGO)[28] is an iterative approach for the optimization of an objective function (y) with few iterations, naturally capturing noisy data, and without the need for gradients. Mathematically, the optimization problem that we sought to solve was cast as

$$\mathbf{x}^* = \text{argmax}_{\mathbf{x}} \mathbb{E}(y|\mathbf{x}) \quad \text{where } y = f(\mathbf{x}) + \epsilon,$$

where $\mathbf{x}$, $f(\mathbf{x})$, and ϵ represented the alloy composition, the unknown average experimental hardness of alloy $\mathbf{x}$, and the experimental noise, respectively.

BGO performs an optimization by sequential information acquisition. The next experiment is found by maximizing an acquisition function, $a(\mathbf{x})$, that quantifies the expected value added if alloy $\mathbf{x}$ were tested. The key steps of the BGO process, depicted in Fig. S1 in the supplementary material, are as follows:

- Step 1: Construction of a statistical surrogate model from the data available.
- Step 2: Evaluation of the acquisition function for all alloys in our design space and selection of a set of alloys that maximized the acquisition function.
- Step 3: Experimental synthesis and characterization of the selected batch of alloys.
- Step 4: Addition of the new results to the original data set and repetition from step 1 until the desired property was achieved or the experimental budget was exhausted.

Predictive statistical model. We used Gaussian processes (GPs) to model the hardness of the alloys of interest; this choice was motivated by its rigorous treatment of uncertainties. Also important for the development of accurate models was the selection of features or descriptors to be used as model inputs. These need to be easy to evaluate for any composition within our design space and provide relevant information about the quantity of interest. We used a combination of fundamental atomic data information of the alloy, such as density, moduli, entropy, and so on calculated using the rule of mixtures, along with phase information, such as the list of phases and liquidus and solidus temperatures, calculated using the Thermo-Calc software. We used one-hot encoding (independent and mutually exclusive binary vectors to represent categorical data) delineating between BCC only, FCC only, and other/secondary phases. The complete list of descriptors employed is included in the supplementary material in Table S1. Finally, hardness was predicted using Gaussian Processes Regression (GPR).[29,30] We employed a homoscedastic GPR model for the initial hardness model for the active learning loop, which was based on the experimental hardness measurements available in Ref. 27. This approach was taken because each alloy had only a single hardness measurement reported. For subsequent models for the active learning loop, we transitioned to a heteroscedastic GPR model,[31,32] as we obtained multiple hardness measurements for each alloy in the suggested batches. The variability in the noise (i.e., standard deviation) of

30 August 2025 06:03:08

hardness measurements among different alloys in the batches necessitated the use of heteroscedastic GPR for these later models.

To capture both nonlinear and linear trends in the data, we used a composite kernel comprising a Radial Basis Function (RBF) kernel and a Linear kernel, both with Automatic Relevance Determination (ARD),

$$k(\mathbf{x}, \mathbf{x}') = k_{\text{RBF}}(\mathbf{x}, \mathbf{x}') + k_{\text{Linear}}(\mathbf{x}, \mathbf{x}').$$

The RBF component enables the model to capture smooth, nonlinear variations, while the linear component accounts for global linear dependencies. The ARD mechanism allows the model to learn separate length scales for each input dimension, thereby identifying and down-weighting irrelevant features, which helps mitigate overfitting and improves model interpretability. Hyperparameters were optimized by maximizing the marginal log-likelihood using the **optimize_restarts()** routine in the GPy[33] library, with 10 random restarts to ensure robustness against local optima.

Overfitting was further addressed through several mechanisms: ARD helped suppress uninformative input dimensions; marginal likelihood optimization inherently applies Bayesian regularization, penalizing overly complex models, and for heteroscedastic models, noise variances were directly specified based on repeated hardness measurements, allowing the model to account for input-dependent noise without overfitting to high-variance regions.

Acquisition function. We used expected improvement (EI) Eq. (1) as our acquisition functions, $a(\mathbf{x})$. This choice combines exploration and exploitation of the design space using the predicted mean $m(\mathbf{x})$ and standard deviation $\sigma(\mathbf{x})$,

$$\begin{aligned}a(\mathbf{x}) &= \text{EI}(\mathbf{x}) \\ &= (m(\mathbf{x}) - m^*)\, \Phi\!\left(\frac{m(\mathbf{x}) - m^*}{\sigma(\mathbf{x})}\right) + \sigma(\mathbf{x})\, \phi\!\left(\frac{m(\mathbf{x}) - m^*}{\sigma(\mathbf{x}),}\right)\end{aligned} \quad (1)$$

where m^* is the maximum observed value at the current stage, Φ is the cumulative distribution function of the normal distribution, and ϕ represents the probability distribution function of the normal distribution.

As discussed above, our experimental output was maximized by modifying the BGO approach so as to select a batch of m compositions at a time with m being 10 in batches 1 and 4 for the subsequent batches. A naïve approach would be to select the top m candidates according to the AF, but this could result in similar alloys and assumes that the information gained is simply additive and independent. Instead, we designed an approach to overcome these issues, and the detailed flow chart of our active learning algorithm is shown in Fig. 3. The approach is similar to the general BGO algorithm, but with a change in the inner loop, highlighted in red, to suggest m experiments. The first experiment in the batch is chosen following the usual procedure, i.e., maximizing the AF. To suggest the second experiment location, we add the model prediction for the first alloy to a copy of the data set, re-build the surrogate, and the second experiment location is chosen using the AF with the updated model. The above procedure is repeated until prospective m compositions were obtained. All synthetic data added
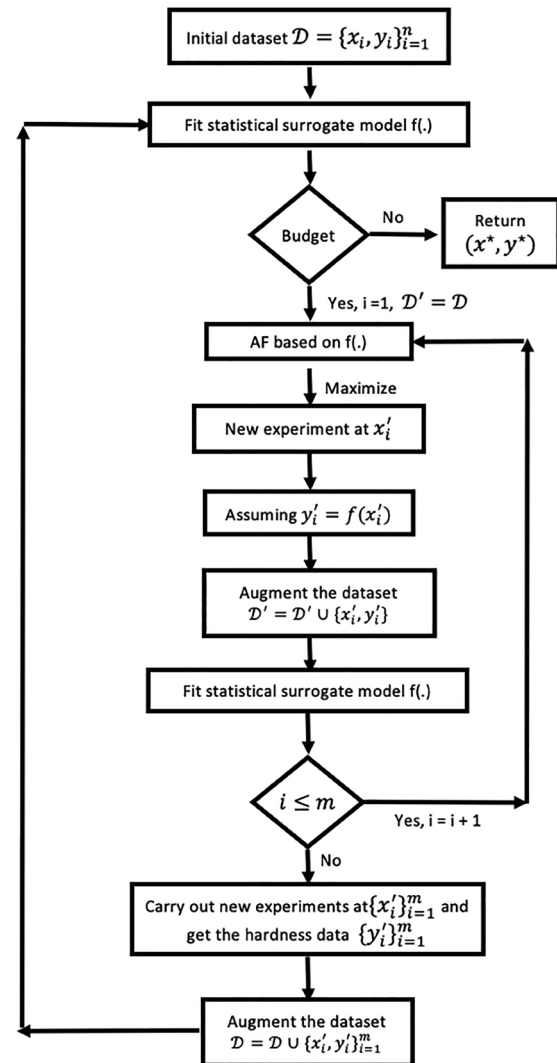

FIG. 3. Flow chart of our active learning procedure.

during the batch generation process are removed before the start of the next BGO cycle.

The complete procedures for building the hardness predictive models, including both homoscedastic and heteroscedastic GP models, are described in detail in the supplementary material. Additionally, the full implementation of our active learning strategy (b.atch—-BGO), is presented with step-by-step algorithmic descriptions to ensure clarity and reproducibility.

C. Fabrication, testing, and characterization

All alloys were fabricated via arc melting using elements in the form of granules of Al, Ti, Cr; a foil of V; wires of Mo and Hf; and W powder with purities of at least 99.7%. Arc melting was

conducted in a custom-made tri-arc chamber with a water-cooled Cu hearth. Elements with a total nominal mass of 5 g were weighed using an electronic microbalance (1 mg resolution). Melting was conducted after distributing the raw materials on the Cu hearth in the chamber of the tri-arc melter. The chamber was sealed and evacuated with a roughing pump for 20 min to achieve a pressure of 150 mTorr (20 Pa) and was subsequently backfilled with ultra-high-purity (UHP, 99.999% purity) Ar. Prior to entry into the chamber, UHP Ar was passed through an oxygen-gettering furnace (OG120 gas purification furnace, Oxy-Gon Industries) to further remove residual oxygen. This evacuation/backfilling process was repeated three times. After the final backfilling step, oxygen-gettered UHP Ar was passed through the chamber at a flow rate of 1.4 l/min. Prior to melting of the alloy, high-purity Zr was melted on the Cu hearth to act as a local oxygen getter.

Each alloy button was melted, allowed to solidify, and then flipped and remelted. Such melting/flipping was repeated five times to promote chemical homogeneity. All alloys were weighed before and after arc melting, and a maximum mass loss of 0.6 g was detected. For these high-throughput experiments, no additional alloy heat treatments to alleviate microsegregation were conducted, and we note that while, perhaps, a majority of compositions would exhibit softening from heat treatments, it has been reported that many alloys exhibit hardening post heat treatment as well.[34,35]

Characterization. All alloy ingots were sectioned, via wire electrical discharge machining, at a location 2 mm from the vertical centerline of each ingot and were subsequently mounted in electrically conductive bakelite. Standard grinding and polishing procedures were conducted with the last polishing steps utilizing a 1 μm diamond suspension in glycol on a woven polishing cloth followed by the use of 0.04 μm colloidal silica on a non-woven, low nap porous polyurethane pad.

For each alloy, at least 16 Vickers hardness measurements were obtained using an automated indenter (AMH43, Leco Corporation) with a 500 g-force load applied for a 13 s dwell time. Indentations were arranged in a 4 × 4 square grid, with indentations spaced 300 μm apart. For a given ingot cross section, the grid was positioned so as to obtain hardness values at the middle, sides, and bottom regions of each ingot, thus sampling multiple as-solidified microstructures in the ingot. Vickers indents were approximately 40–60 μm in diagonal length, which samples randomly across multiple secondary dendrite arms that are spaced on order of 10–30 μm, depending on each individual alloy.

The phases present in the arc-melted alloys were characterized via x-ray diffraction analysis (D2 Phase diffractometer, Bruker) using Cu K-α radiation ($\lambda = 0.154$ nm) at a 0.23°/s scan rate, with a 99.8% purity Ni powder standard used to account for peak shifting due to minor specimen misalignment. Scanning electron microscopy (NanoNova 450, FEI) conducted at 15 keV and energy-dispersive x-ray (EDX) spectroscopy were employed for microstructural and microchemical analyses, respectively, and the measured EDX composition for each element for a given specimen was typically within 0–3 at. % of the targeted nominal composition, as shown in supplementary material Table 3. The detection of secondary, non-BCC phases that could influence hardness measurements was of particular interest. The nanostructures of four Al–V–Nb–W alloys, which included two alloys of the highest hardness, were

characterized using transmission electron microscopy. Lamellae of approximately 100 nm thickness were prepared by focused ion beam milling (Quanta 3D FEG, Thermo Fisher Scientific) conducted at 30 kV and 30 pA-15 nA. Specimen cross sections were transferred to a copper grid using an Omniprobe 200 lift-out system (Oxford Instruments, Abingdon, UK) and were subsequently analyzed using a Talos 200X TEM (Thermo Fisher Scientific, Waltham, Massachusetts, USA) operated at 200 keV. EDX analyses were obtained in a scanning transmission electron microscope (STEM) mode and images were captured with a high-angle annular dark-field (HAADF) detector at a camera length of 205 mm and a collection angle of 28–172 mrad.

III. RESULTS

Figure 4 shows the measured hardness values obtained from the arc-melted alloys as a function of the active learning cycle. Filled symbols represent the experimental measurements (error bars indicate standard deviation), and the empty symbols correspond to the predictions (mean and standard deviations) of the GPR model used to select the alloys. All alloys in the active learning iterations (batches 1–5), except for $Al_{30}V_{35}Nb_{20}W_{15}$ in batch 2, exhibit experimentally measured hardness values within one standard deviation of the mean prediction obtained from the GPR model. We also note that we were able to fabricate all alloys selected by the active learning process except for $Al_{15}V_{45}Ti_{10}W_{30}$, selected in batch 5. In batch 6, we selected the two alloys predicted by the GPR model, as will be discussed next.

The expected improvement acquisition function balances performance (i.e., high hardness) with uncertainty; this resulted in the selection of alloys with hardness no less than 387 HV (batch 4, $Al_5Mo_5Nb_{75}W_{15}$), and standard deviation predictions not less than 135 HV. The average standard deviation decreased from 164 HV (batch 1) to 135 HV (batch 5) during the sequential optimization. This reduction in model uncertainty and the reduction improvement is shown collectively by the EI vs batch in Fig. 5. Small dots represent the EI of all alloys within the design space, black circles represent the selected alloys, and the open symbol in batch 5 represents the alloy we were unable to fabricate. The EI decreased after the first experimental data acquisition and then increased in the third batch due to increased model uncertainty [$\sigma(\mathbf{x})$ in Eq. (1)] for the alloys in batch 3. This elevated model uncertainty is attributed to a high aleatoric uncertainty (experimental variability) in the hardness measurements of batch 2. The EI over the design space remained rather constant during cycles 4 and 5. Consequently, for the last experimental cycle, we fabricated and tested the two alloys with the highest hardness as predicted by the GPR model. These alloys, $Al_{35}V_{30}Nb_{15}W_{20}$ and $Al_{35}V_{30}Nb_{20}W_{15}$, were predicted to be the hardest of both the 5144 BCC-based alloy data set and the Al–V–Nb–W system. We confirmed these high values experimentally, with measured hardness values above 770 HV, as shown for the last iteration in Fig. 4. As is common when predicting extreme values, the GPR model underestimated the hardness of these alloys; this is more marked for the $Al_{35}V_{30}Nb_{15}W_{20}$ alloy, as will be discussed below.
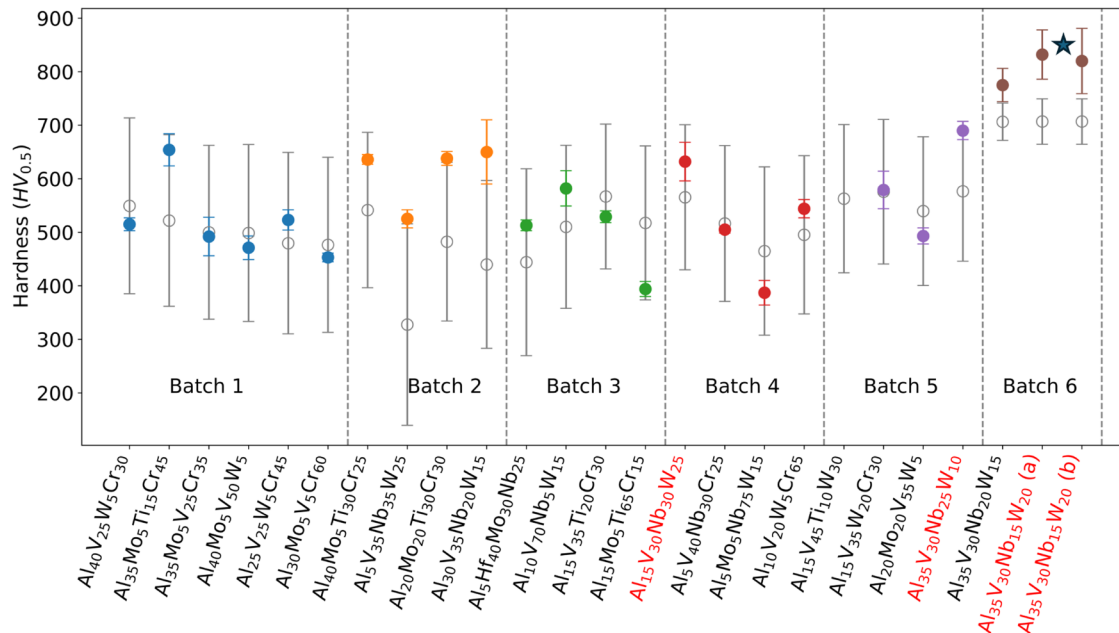
FIG. 4. Measured (full symbols) and predicted (open symbols) hardness of the alloys obtained by our active learning approach. Iterations 1–5 were obtained using expected improvement, and iteration 6 characterized the highest performer alloys as predicted by the GPR model. Highlighted in red are Al–V–Nb–W alloys with high hardness that were subjected to additional characterization. The asterisk marks the alloy with minor non-BCC phases.

A. Microstructural characterization

X-ray diffraction analyses of arc-melted alloys from each active-learning iteration (batch) are provided in Fig. 6(a). All 24 arc-melted alloys exhibit peaks corresponding to the BCC-like structures, with only the $Al_{35}V_{30}Nb_{15}W_{20}$ alloy (the composition of the highest hardness) containing minor non-BCC phase diffraction peaks, as indicated by the arrows. In the latter case, these minor diffraction peaks are consistent with an Al_3Nb-type (body centered tetragonal)
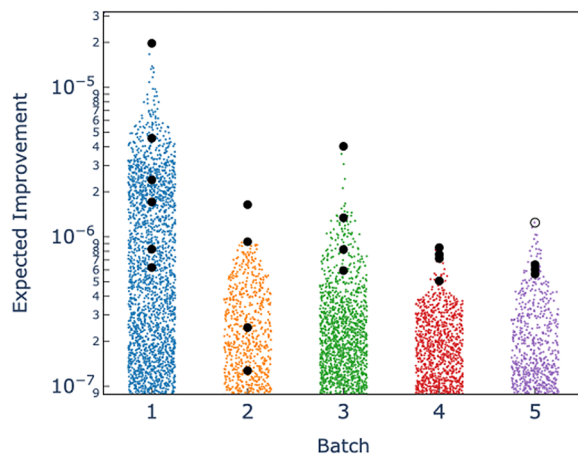

FIG. 5. Expected improvement of the selected alloys for iterations 1–5.

intermetallic phase.[36,37] To further characterize the alloys explored during the active learning interactions, backscattered electron images were obtained from polished cross sections for each active-learning iteration of the arc-melted alloys. Figure 6(b) shows the dendritic microstructure formed upon solidification, with, typically, higher average Z-number compositions comprising the dendrite cores.

Four alloys in the Al–V–Nb–W family, from active-learning iterations 4 ($Al_{15}V_{30}Nb_{30}W_{25}$), 5 ($Al_{35}V_{30}Nb_{25}W_{10}$), and 6 ($Al_{35}V_{30}Nb_{20}W_{15}$, $Al_{35}V_{30}Nb_{15}W_{20}$), were also examined by transmission electron microscopy. STEM images, EDX composition profiles, and selected area electron diffraction (SAED) patterns of these four alloys are provided in Fig. 7. The EDX profiles indicate that the dendrites present in these alloys are rich in W relative to the surrounding interdendritic regions. Despite these composition variations, SAED analyses obtained from the $Al_{15}V_{30}Nb_{30}W_{25}$, $Al_{35}V_{30}Nb_{25}W_{10}$, and $Al_{35}V_{30}Nb_{20}W_{15}$ alloys confirmed that the dendrites retain the BCC-like crystal structures. This is true in all cases except for $Al_{35}V_{30}Nb_{15}W_{20}$ which contains a needle-shaped phase dispersed within the interdendritic region. SAED analysis of this high-aspect-ratio phase was consistent with a body centered tetragonal structure, as has been reported for the Al_3Nb-type phase.[36,37] The measured spacings of the (002) and (−110) planes for this phase obtained from SAED analysis were 0.42 and 0.26 nm, respectively, which did not deviate significantly from the (002) and (−110) plane spacings for pure Al_3Nb (0.430 and 0.272 nm, respectively).[37] The presence of a second phase with a non-bcc crystal structure with its hardening effect explains the more significant underestimation of the hardness of this alloy by our GPR model.
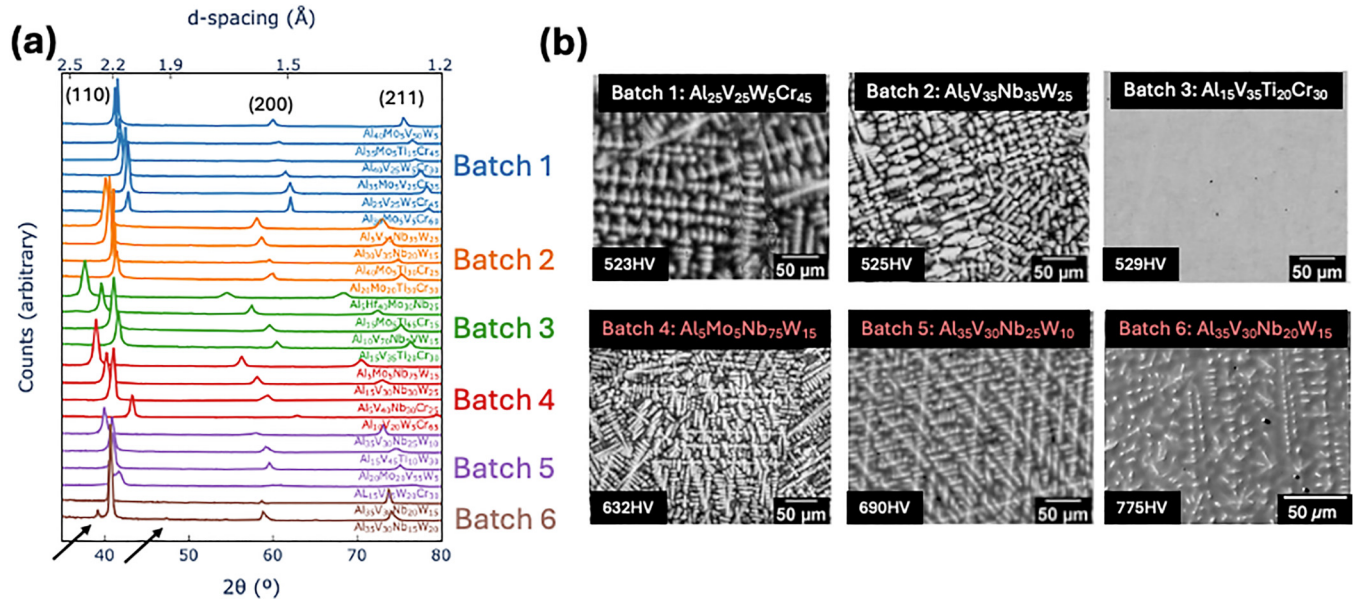
FIG. 6. (a) X-ray diffraction analyses and (b) backscattered electron images (SEM) obtained from polished cross sections of arc-melted alloys from active-learning iterations 1–6.

IV. RESULTS AND DISCUSSION

Figure 8 compares the hardness of Al-bearing alloys obtained in this work with comparable alloys, with and without Al, in the UCSB/Citrine MPEA database[27] used to train the initial GPR model. While the original database includes alloys with hardness exceeding those measured in this work, those materials comprised multiple phases, including intermetallics. Therefore, the $Al_{35}V_{30}Nb_{20}W_{15}$ alloy presented here surpasses all BCC-based (i.e., non-intermetallic-containing) alloys in the database. More specifically, the measured hardness of the $Al_{35}V_{30}Nb_{20}W_{15}$ alloy
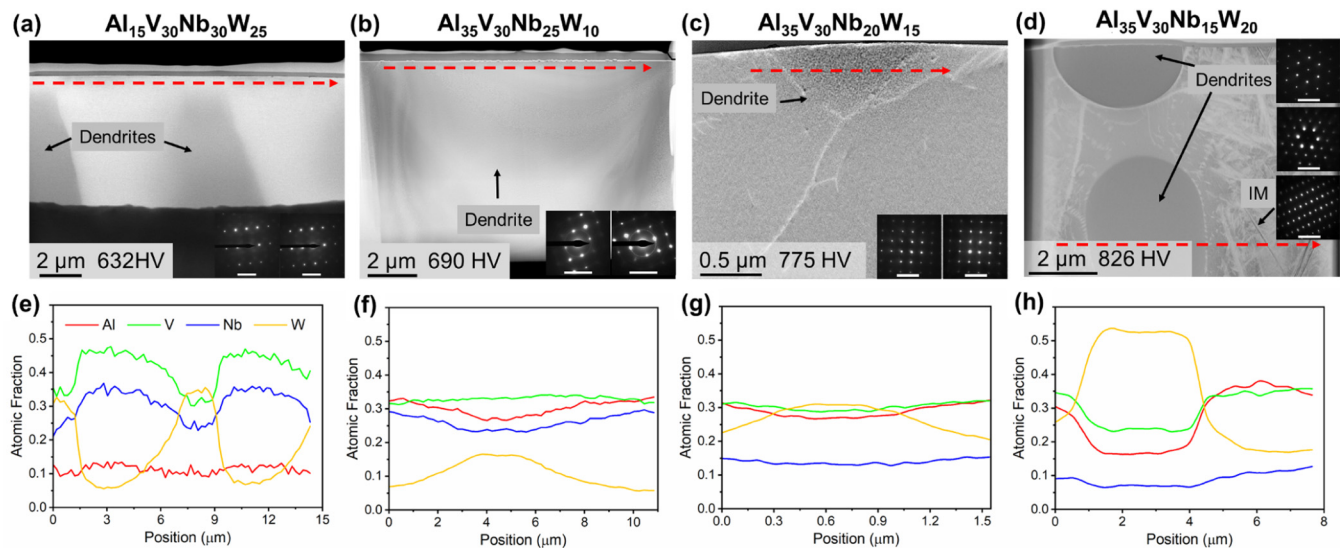
FIG. 7. STEM and SAED analyses of an ion-milled cross section of (a) $Al_{15}V_{30}Nb_{30}W_{25}$, (b) $Al_{35}V_{30}Nb_{25}W_{10}$, (c) $Al_{35}V_{30}Nb_{20}W_{15}$, (d) $Al_{35}V_{30}Nb_{15}W_{20}$, and their corresponding EDX (e)–(h).
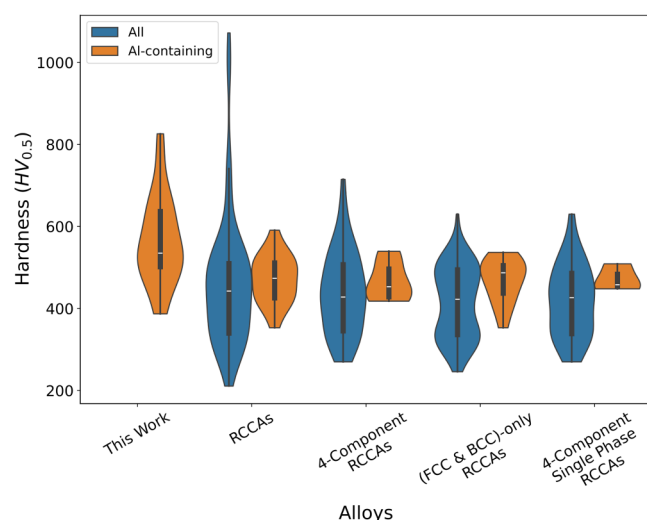
FIG. 8. Distribution of the hardness values obtained in this work compared to various subsets of the existing literature data. Dark colors are aluminum-containing, while lighter colors include all alloys.

FIG. 9. Design space restricted to alloys predicted to be single phase and processable via arc melting, with alloys from the literature and the synthesized alloys in large size colored by their hardness.

(775 ± 31 HV) from this work is 31% higher than the highest hardness previously reported for an Al-containing RCCA ($Al_{20}Mo_{10}Nb_{20}Ta_{10}Ti_{20}Zr_{20}$ with a hardness of 591 HV)[38] and 23% harder than the hardest single-phase four-component RCCA previously reported ($Cr_{25}Nb_{25}Ti_{25}W_{25}$ with a hardness of 630 HV).[22] There are two Al-containing alloys in the Citrine database with hardness comparable to the values achieved here but our comparison is restricted to our design space. However, the original paper[34] reporting those values does not specify the structure of AlFeNiTiVZr (800 HV) or AlCoFeNiTiVZr (790 HV) explicitly (despite the Borg database indicating it is BCC). Thermo-Calc TCHEA4 predicts both alloys to be multi-phase below the solidus temperature. In addition, the paper is entitled "Nanostructure high-entropy alloys...," and thus most of the alloys investigated likely contain small volume fractions of secondary (strengthening) phases. Overall, Al-free RCCAs from the UCSB/Citrine database exhibit higher hardness values compared to Al-containing ones, and we attribute this trend to researchers prioritizing high melting temperatures and phase stability, which Al does not typically provide for refractory-based compositions. Therefore, our work significantly expands the potential Al-bearing RCCA composition space and identifies promising candidate materials.

While aluminum was included in our design space to promote high-temperature oxidation resistance, this alone does not ensure the desired outcome. It is worth noting that the hard Al-rich$Al_{35}V_{30}Nb_{20}W_{15}$ alloy identified in this work also includes a high vanadium content. The low melting point and relatively high vapor pressure of V_2O_5 have contributed to the poor high-temperature oxidation resistance of vanadium, vanadium–aluminum alloys, and some vanadium-bearing refractory multiple-principal-element alloys (W–Mo–Ta–Nb–V).[39–41] Evaluation and optimization of the high-temperature oxidation resistance of
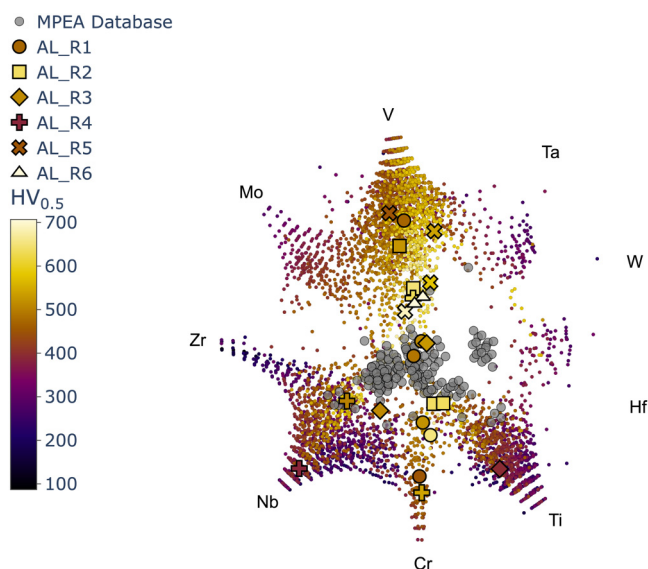
Al-bearing refractory complex concentrated alloys were beyond the scope of the present paper and are important research topics for future works.

Significant efforts have been made toward the discovery of RCCAs using computational methods. Elder *et al.*[42] used thermodynamics and dislocation-based models to explore RCCAs with high yield strength and BCC solid solution stability and found a few medium entropy alloys with better specific yield strength than alloys with more elements. In our work, we limited our search space to Al -containing 4-element alloys in line with this work. Rao *et al.*[43] also used successive down selection by filtering desirable composition, followed by mechanical properties and thermodynamic stability to obtain a set of promising candidates. They validated the methodology with the fabrication of a quinary HfMoNbTaTi alloy. In our work, we start from an existing experimental data set instead and down select to form the core of our training set. In contrast to their methodology, we use this training set for active learning to predict potential candidates iteratively. In each iteration, we fabricate a set of potential candidates. If the fabricated alloy does not have the desired properties, it is still useful as it goes into the active learning model for the next iteration, which makes this methodology different from the ones discussed above.

It is instructive to explore how the AL approach navigated the compositional space available. Figure 9 shows the 5144 alloys in our design space, now colored according to the predicted hardness from the GPR model trained with data up to batch 5. As before, we use multi-dimensional scaling to represent the alloy composition in 2D. Due to the exceptional performance of V-containing alloys, many of the higher hardness alloys are observed to be located nearer to the V-rich spike of the plot, but other high-hardness

30 August 2025 06:03:08

alloys are observed toward the center of the plot (indicating richer in Al content and near equi-molar compositions) as well as toward Cr- and Nb-containing compositions.

The efficacy of the exploration component of the acquisition function is indicated by comparing the location of the UCSB/Citrine MPEA database compositions (gray points) to the locations of our fabricated samples (colored, enlarged points). While the batched BGO process selected compositions near the existing compositions, the majority of the compositions selected were far from the initial training data set. Notably, the AlVNbW system is located near the center, slightly above the MPEA database points, and we note that the batch BGO frequently returned to this alloy system to explore this space, returning in batches 2, 3, 4, 5, and, finally, 6. Analyzing the highest hardness alloys predicted by the GPR model for each batch, the Al–V–W–Nb system did not breach top 10 until Batch 3 where its compositions were ranked 7–10. In Batch 4, its presence remained the same, but it increased in Batch 5 to ranks 4–10 and, finally, dominated the top 10 at the start of Batch 6.

V. CONCLUSIONS

In this work, we used a combination of physics-based modeling and machine learning to accelerate the discovery of high-strength complex concentrated alloys. CALPHAD calculations restricting the search to compositions with the desired BCC structure reduced the number of compositions by approximately an order of magnitude. Using BGO sequential optimization, we discovered an Al-containing four-element BCC RCCA with a 31% improvement over the highest hardness previously reported. This was achieved in only 24 experiments (out of 67 000 possible candidates) organized in six campaigns. Our approach integrates CALPHAD simulations with GP regression and sequential optimization based on EI using a batch approach to enable for parallel experimental processing. The approach is easily generalizable to other alloy discovery tasks.

We also observed that the single-crystal strength models and microstructural features showed no strong correlation with the measured experimental hardness. Predictions from the Maresca and Curtin model,[10,11] implemented in Ref. 44, showed a weak Pearson correlation of 0.32 with the measured values (supplementary material Figs. S2 and S3), suggesting that dislocation–solute interactions are not the primary cause of high strength. Correlations with elemental fractions (supplementary material Fig. S4) and physical descriptors (supplementary material Fig. S5) were also modest, with W content and atomic radius mismatch showing the strongest negative trends (supplementary material Figs. S6 and S7), though none exceeded $|p| > 0.7$. The exceptional hardness in Al–V–Nb–W likely stems from mechanisms beyond the resolution of the current theoretical models, warranting further experimental and theoretical works.

SUPPLEMENTARY MATERIAL

The supplementary material includes the Bayesian global optimization details used for active learning, the descriptors used for the HEA hardness prediction model, different correlations in the model/data, and a comparison of our model hardness prediction with the strength predictions from literature.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

S.K., A.H., and S.M. have contributed equally in this work.

Sharmila Karumuri: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Austin Hernandez:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Visualization (equal). **Saswat Mishra:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Project administration (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Zachary McClure:** Data curation (equal); Formal analysis (equal); Methodology (equal); Validation (equal); Visualization (equal). **Victoria Tucker:** Data curation (supporting); Writing – review & editing (supporting). **Joseph C. Flanagan:** Data curation (supporting); Investigation (supporting). **Sunghwan Hwang:** Data curation (supporting); Formal analysis (supporting); Methodology (supporting); Visualization (supporting); Writing – original draft (equal). **Kenneth H. Sandhage:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **Ilias Bilionis:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal). **Michael S. Titus:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Visualization (supporting); Writing – original draft (equal); Writing – review & editing (equal). **Alejandro Strachan:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]G. L. Erickson, "The development and application of CMSX-10," Superalloys 1996, 35–44 (1996), available at https://www.tms.org/Superalloys/10.7449/1996/Superalloys_1996_35_44.pdf.

[2]C. Rodgers, "A performance diagnosis of the 1939 Heinkel He S3B turbojet," in *Turbo Expo: Power for Land, Sea, and Air* (ASME, 2004), Vol. 41677, pp. 1–11.

[3]R. W. Daniels, "Duralumin," SAE Trans. **14**, 751–763 (1922).

[4]Y. Xu, C. Yang, X. Xiao, X. Cao, G. Jia, and Z. Shen, "Evolution of microstructure and mechanical properties of Ti modified superalloy Nimonic 80A," Mater. Sci. Eng.: A **530**, 315–326 (2011).

[5]C. B. Meher-Homji *et al.*, "The historical evolution of turbomachinery," in *Proceedings of the 29th Turbomachinery Symposium* (Texas A&M University, Turbomachinery Laboratories, 2000).

[6]R. C. Reed, *The Superalloys: Fundamentals and Applications* (Cambridge University Press, 2008).

[7]F. T. Jane, *Jane's Fighting Aircraft of World War II* (Studio Editions Ltd, London, 1998). ISBN: 0-517-67964-7.

[8]E. Roux, *Turbofan and Turbojet Engines: Database Handbook* (Elodie Roux, 2007).

[9]F. G. Coury, M. Kaufman, and A. J. Clarke, "Solid-solution strengthening in refractory high entropy alloys," Acta Mater. **175**, 66–81 (2019).

[10]F. Maresca and W. A. Curtin, "Mechanistic origin of high strength in refractory BCC high entropy alloys up to 1900 K," Acta Mater. **182**, 235–249 (2020).

[11]F. Maresca and W. A. Curtin, "Theory of screw dislocation strengthening in random BCC alloys from dilute to 'high-entropy' alloys," Acta Mater. **182**, 144–162 (2020).

[12]M. S. Farnell, Z. D. McClure, S. Tripathi, and A. Strachan, "Modeling environment-dependent atomic-level properties in complex-concentrated alloys," J. Chem. Phys. **156**(11), 114102 (2022).

[13]K. Kaufmann and K. S. Vecchio, "Searching for high entropy alloys: A machine learning approach," Acta Mater. **198**, 178–222 (2020).

[14]D. Khatamsaz, B. Vela, P. Singh, D. D. Johnson, D. Allaire, and R. Arróyave, "Multi-objective materials bayesian optimization with active learning of design constraints: Design of ductile refractory multi-principal-element alloys," Acta Mater. **236**, 118133 (2022).

[15]D. E. Farache, J. C. Verduzco, Z. D. McClure, S. Desai, and A. Strachan, "Active learning and molecular dynamics simulations to find high melting temperature alloys," Comput. Mater. Sci. **209**, 111386 (2022).

[16]Z. Rao, P.-Y. Tung, R. Xie, Y. Wei, H. Zhang, A. Ferrari, T. P. C. Klaver, F. Körmann, P. T. Sukumar, A. K. da Silva *et al.*, "Machine learning–enabled high-entropy alloy discovery," Science **378**(6615), 78–85 (2022).

[17]H. Li, R. Yuan, H. Liang, W. Y. Wang, J. Li, and J. Wang, "Towards high entropy alloy with enhanced strength and ductility using domain knowledge constrained active learning," Mater. Des. **223**, 111186 (2022).

[18]Y.-J. Chang, C.-Y. Jui, W.-J. Lee, and A.-C. Yeh, "Prediction of the composition and hardness of high-entropy alloys by machine learning," JOM **71**(10), 3433–3442 (2019).

[19]R. Jain, U. Lee, S. Samal, and N. Park, "Machine-learning-guided phase identification and hardness prediction of Al-Co-Cr-Fe-Mn-Nb-Ni-V containing high entropy alloys," J. Alloys Compd. **956**, 170193 (2023).

[20]C. Yang, C. Ren, Y. Jia, G. Wang, M. Li, and W. Lu, "A machine learning-based alloy design system to facilitate the rational design of high entropy alloys with enhanced hardness," Acta Mater. **222**, 117431 (2022).

[21]C. Wen, Y. Zhang, C. Wang, D. Xue, Y. Bai, S. Antonov, L. Dai, T. Lookman, and Y. Su, "Machine learning assisted design of high entropy alloys with desired property," Acta Mater. **170**, 109–117 (2019).

[22]F. G. Coury, T. Butler, K. Chaput, A. Saville, J. Copley, J. Foltz, P. Mason, K. Clarke, M. Kaufman, and A. Clarke, "Phase equilibria, mechanical properties and design of quaternary refractory high entropy alloys," Mater. Des. **155**, 244–256 (2018).

[23]F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," J. Mach. Learn. Res. **12**, 2825–2830 (2011), available at https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?source=post_page.

[24]O. Kubaschewski and G. Heymer, "The thermodynamics of the chromium-iron system," Acta Metall. **8**(7), 416–423 (1960).

[25]A. Peterlongo, A. Miotello, and R. Kelly, "Laser-pulse sputtering of aluminum: Vaporization, boiling, superheating, and gas-dynamic effects," Phys. Rev. E **50**(6), 4716 (1994).

[26]M. S. Titus, A. Hernandez, V. Tucker, and A. Bejjipurapu, see https://github.itap.purdue.edu/michaeltitusgroup/pyTCPlotter for "pyTCPlotter" (2024) (last accessed August 8, 2024).

[27]C. K. H. Borg, C. Frey, J. Moh, T. M. Pollock, S. Gorsse, D. B. Miracle, O. N. Senkov, B. Meredig, and J. E. Saal, "Expanded dataset of mechanical properties and observed phases of multi-principal element alloys," Sci. Data **7**(1), 430 (2020).

[28]D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," J. Glob. Optim. **13**, 455–492 (1998).

[29]C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*, Adaptive Computation and Machine Learning Series (MIT Press, 2005). ISBN: 9780262182539, see https://books.google.com/books?id=Tr34DwAAQBAJ.

[30]E. Schulz, M. Speekenbrink, and A. Krause, "A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions," J. Math. Psychol. **85**, 1–16 (2018).

[31]C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning* (Springer, 2006), Vol. 4.

[32]P. Goldberg, C. Williams, and C. Bishop, "Regression with input-dependent noise: A Gaussian process treatment," in *Advances in Neural Information Processing Systems, 10* (NeurIPS, 1997), available at https://proceedings.neurips.cc/paper/1997/hash/afe434653a898da20044041262b3ac74-Abstract.html.

[33]See https://github.com/SheffieldML/GPy for "Gpy: A Gaussian process framework in python" (2002).

[34]J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, and S.-Y. Chang, "Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes," Adv. Eng. Mater. **6**(5), 299–303 (2004).

[35]C. Acemi, B. Vela, E. Norris, W. Trehern, K. C. Atli, C. Cleek, R. Arróyave, and I. Karaman, "Multi-objective, multi-constraint high-throughput design, synthesis, and characterization of tungsten-containing refractory multi-principal element alloys," Acta Mater. **281**, 120379 (2024).

[36]The International Center for Diffraction Data. File 04-065-2666 for Al$_3$Nb, 2020.

[37]C.-P. Reip and G. Sauthoff, "Deformation behaviour of the intermetallic phase Al$_3$Nb with DO22 structure and of Al$_3$Nb-base alloys: Part I. Physical properties and short-term behaviour," Intermetallics **1**(3), 159–169 (1993).

[38]O. N. Senkov, C. Woodward, and D. B. Miracle, "Microstructure and properties of aluminum-containing refractory high-entropy alloys," JOM **66**, 2030–2042 (2014).

[39]W. R. Price and J. Stringer, "The oxidation of vanadium at high temperatures," J. Less Common Met. **8**(3), 165–185 (1965).

[40]J. G. Keller and D. L. Douglass, "The high-temperature oxidation behavior of vanadium-aluminum alloys," Oxid. Met. **36**, 439–464 (1991).

[41]R. Su, H. Zhang, G. Ouyang, L. Liu, D. D. Johnson, and J. H. Perepezko, "Oxidation mechanism in a refractory multiple-principal-element alloy at high temperature," Acta Mater. **246**, 118719 (2023).

[42]K. L. M. Elder, J. Berry, B. Bocklund, S. K. McCall, A. Perron, and J. T. McKeown, "Computational discovery of ultra-strong, stable, and lightweight refractory multi-principal element alloys. Part I: Design principles and rapid down-selection," npj Comput. Mater. **9**(1), 84 (2023).

[43]Y. Rao, C. Baruffi, A. De Luca, C. Leinenbach, and W. A. Curtin, "Theory-guided design of high-strength, high-melting point, ductile, low-density, single-phase BCC high entropy alloys," Acta Mater. **237**, 118132 (2022).

[44]D. Wen and M. S. Titus, "pySSpredict: A python-based solid-solution strength prediction toolkit for complex concentrated alloys," Comput. Mater. Sci. **220**, 111977 (2023).