

ARTICLE OPEN



Bayesian optimization with active learning of design constraints using an entropy-based approach

Danial Khatamsaz¹, Brent Vela²✉, Prashant Singh^{2,3}, Duane D. Johnson^{3,4}, Douglas Allaire¹ and Raymundo Arróyave^{1,2,5}

The design of alloys for use in gas turbine engine blades is a complex task that involves balancing multiple objectives and constraints. Candidate alloys must be ductile at room temperature and retain their yield strength at high temperatures, as well as possess low density, high thermal conductivity, narrow solidification range, high solidus temperature, and a small linear thermal expansion coefficient. Traditional Integrated Computational Materials Engineering (ICME) methods are not sufficient for exploring combinatorially-vast alloy design spaces, optimizing for multiple objectives, nor ensuring that multiple constraints are met. In this work, we propose an approach for solving a constrained multi-objective materials design problem over a large composition space, specifically focusing on the Mo-Nb-Ti-V-W system as a representative Multi-Principal Element Alloy (MPEA) for potential use in next-generation gas turbine blades. Our approach is able to learn and adapt to unknown constraints in the design space, making decisions about the best course of action at each stage of the process. As a result, we identify 21 Pareto-optimal alloys that satisfy all constraints. Our proposed framework is significantly more efficient and faster than a brute force approach.

npj Computational Materials (2023)9:49; <https://doi.org/10.1038/s41524-023-01006-7>

INTRODUCTION

To improve their efficiency, gas turbine engines (GTEs) must be able to operate at higher temperatures. The development of materials capable of withstanding these demanding operating conditions has played a key role in the evolution of GTE technologies. Ni-based superalloys are currently the material of choice for GTE blades, and have been continually redesigned over the past 40 years to increase their ability to operate at higher temperatures. Starting with PWA1480 and culminating in TMS-238, six generations of single-crystal Ni-based superalloys have been developed¹. TMS-238 is the most advanced Ni-based superalloy to date, and is able to withstand 1000 hours of creep testing under 137 MPa tensile stress at 1100 °C¹. However, these alloys are approaching their operational limits as they are being designed to operate near their solidus temperatures. As a result, the discovery and development of ultrahigh-temperature materials are necessary to enable further increases in operating temperatures for GTE blades².

Refractory Multi-Principal-Element Alloys (MPEAs) have shown promise as structural materials for gas turbine engine blades³. These alloys consist of multiple alloying elements (typically 4 or more) at concentrations ranging from 5 to 35 at%. The diversity of MPEA compositions offers the potential to design alloys with desirable properties such as low density, high-temperature yield strength, creep resistance, and oxidation resistance. However, the MPEA design space has been largely unexplored to date⁴. The high dimensionality of this space and the combinatorial explosion of different constituent combinations makes it challenging to explore. For example, a 5-component alloy system sampled at 5 at % would result in over 10,000 candidate designs, not including the exploration of microstructure space. Due to the vast size of the MPEA space, it is impossible to explore it through traditional experimental (or even computational) approaches.

Moreover, candidate alloys for complex engineering applications such as GTE must meet multiple design objectives and constraints, all at once. For example, they must be ductile at room temperature for formability while retaining their yield strength at high temperatures. However, the ‘strength-ductility trade-off’⁵ makes it difficult to design such an alloy. In addition to these objectives, candidate alloys must also meet a number of performance constraints, including low density, high thermal conductivity, narrow solidification range, high solidus temperature, and a small linear thermal expansion coefficient. The design of structural materials for GTE blades is, therefore, a highly constrained problem, requiring the simultaneous satisfaction of multiple objectives and constraints. It is not possible to know beforehand whether a given alloy will meet all of these requirements, so each point in the design space must be individually evaluated. These multi-objective, multi-constraint problems are more complex and resource-intensive than conventional single-objective, loosely-constrained design problems.

The Integrated Computational Materials Engineering (ICME) paradigm⁶ offers a promising approach for designing alloys with tailored properties through computational means by inverting the process-structure-property-performance (PSPP) chain. However, constructing meaningful linkages along the PSPP chain is a resource-intensive process, both experimentally and computationally. Traditional ICME methods are not sufficient for efficiently exploring a vast, high-dimensional design space while simultaneously optimizing for multiple objectives and satisfying a range of constraints. This presents a major challenge in the field, as it is crucial to identify constraint-satisfying Pareto-optimal designs within limited resources. Without more efficient approaches for exploring and exploiting highly-constrained multi-objective design problems, it will be difficult to make significant progress in this area.

¹J. Mike Walker ‘66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843, USA. ²Department of Materials Science and Engineering, Texas A&M University, College Station, TX 77843, USA. ³Ames Laboratory, U.S. Department of Energy, Iowa State University, Ames, IA 50011, USA. ⁴Department of Materials Science & Engineering, Iowa State University, Ames, IA 50011, USA. ⁵Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX 77843, USA.

✉email: brentvela@tamu.edu

Multi-objective Bayesian optimization (MOBO) methods have been popular in materials design because they work with minimal data and employ a heuristic-based search to look for the possibly most informative observations to make and increase a system's state of knowledge in terms of optimal design. MOBO schemes have been successfully deployed in various contexts within the domain of materials science. For example, Arpan et al.⁷ leveraged MOBO to design interfacially controlled ferroelectric materials for superior energy storage and minimal energy loss. The authors performed 4-objective optimization of the following parameters: temperature, partial O₂ pressure, film thickness, and surface ion energy. Solomou et al.⁸ optimally explored the multi-objective Pareto front in precipitation-strengthened shape-memory alloys by maximizing the Expected Hyper Volume Improvement (EHVI) scalar metric⁹. In another work, Suzuki et al.¹⁰ proposed a MOBO scheme known as *Pareto-frontier entropy search* (PFES). The proposed acquisition function evaluates the information gain via the mutual information between the objective functions and the Pareto front and selects the design most likely to improve the system's knowledge of the Pareto front. The authors benchmarked the proposed optimization scheme against two datasets concerning the design of battery materials. Within the first dataset, the simultaneous maximization of ion conductivity and stability (minimization of formation energy) was performed within the $Bi_{1-x-y-z}Er_xNb_yW_zO_{48+y/3+2z}$ chemical space where a pool of 335 candidate designs were available. Likewise, simultaneous maximization of ion conductivity and stability was performed within the $La_{2/3-x}Li_{3x}TiO_3$ chemical space where a pool of 1119 candidate designs was available. The authors note that this entropy-based approach to MOBO converged faster than implementations such as ParEGO¹¹ in both design spaces.

An improvement to the Bayesian optimization paradigm is to employ multiple models representing the same quantity of interest. This is known as multi-fidelity BO and has been shown to effectively increase the robustness and efficiency of engineering design schemes^{12–16}. These models are built upon different assumptions and/or simplifications and vary in fidelity and cost of the evaluation. The models can then be considered to be *information sources* that provide useful knowledge about a given quantity of interest (QoI). In multi-fidelity BO, the assumption is that every source has some helpful information about the design space. By accurately fusing the information from all available sources, it is possible to construct a *fused model* that is a better approximation to the ground truth than any information source in isolation. In the earlier works of refs. 12–16, a multi-fidelity approach has been employed to optimize a single quantity of interest (single-objective optimization). Recently, this multi-fidelity setting has been expanded to multi-objective design problems as well¹⁷. However, none of these prior works have tackled problems for which constraints must be actively learned to identify the feasible design space.

Constrained design problems pose a significant challenge because it can be difficult to handle constraints and ensure the feasibility of proposed solutions. Without properly identifying the feasible region in the design space, there is a risk that optimal designs may be infeasible. Recently, Hickman and Aldeghi et al.¹⁸ proposed a method for using Bayesian optimization (BO) with constraints in their Python module, GRYFFIN. However, this method assumes that the constraints are already known and can be easily checked, which is often not the case. Additionally, checking if a design satisfies a constraint can often require expensive computational modeling or resource-intensive experiments. In such cases, machine learning approaches can be more effective at learning and modeling the constraints. The main focus when learning a constrained design space is the boundary of the feasible space, rather than the value of the constrained quantity of interest (QoI) at a particular location. Instead of a regression model, it may be more efficient to use a classifier to represent the

feasibility boundary that separates feasible and infeasible regions in the design space. Once this boundary has been correctly identified, optimization can be performed within the feasible design space, which increases the efficiency of the design process by limiting expensive queries against design objectives to only feasible design choices.

Of particular interest to this work is the Closed-loop autonomous materials exploration and optimization (CAMEO) framework, developed by ref. 19. The authors deployed CAMEO within the Ge-Sb-Te chemistry space in search of optimal phase-change memory materials for application in photonic switching devices. The authors first use GRENDEL (graph-based endmember extraction and labeling)²⁰ to determine where boundaries between phases lie in the chemistry space. Once the phase boundaries have been learned, the authors then use CAMEO to optimize within a phase of particular interest; priority is given to design near phase boundaries where significant changes in the optical contrast between amorphous and crystalline states (the target property) are expected. This framework first identifies the phase-boundaries in a particular design space. Once the phase-boundaries are identified, CAMEO will sample near the boundaries as this is likely where the local maxima are located. Despite this, CAMEO was limited to mapping phase boundaries. Furthermore, during constrained optimization in the context of alloy design, it is often the case that multiple constraints (not just phase boundaries) must be mapped in order to identify regions in the design space worth performing optimization in. In this work we propose a framework that actively learns the boundaries of multiple constraints and then searches within these boundaries for optimal materials.

In order to effectively use classifiers to represent constraint boundaries, it is necessary to learn the feasibility boundaries to ensure the accuracy of classifier predictions. In this work, we build upon our previous efforts²¹ in constraint-satisfaction multi-objective Bayesian optimization by introducing an entropy-based approach to the decision-making process. Our previous approach calculated entropy based on the difference between class membership probabilities predicted by Gaussian process classifiers, resulting in higher entropy for designs close to the predicted boundary. However, this approach did not take into account uncertainty in the probability predictions, and the entropy was heavily influenced by the location of the predicted boundary, which can change as the system learns more, potentially making previously queried data points less valuable. In this work, we propose calculating entropy based on uncertainty in class membership probability predictions so that designs with higher uncertainty about their class membership will have higher entropy regardless of their distance from the predicted boundary. This approach improves upon our previous efforts by considering uncertainty in probability predictions and reducing the reliance on the location of the predicted boundary.

Our proposed method for solving constrained design problems is not only faster than previous approaches but also allows for more informed decision-making at every stage of the process. By introducing an entropy-based approach to the Bayesian optimization (BO) framework, we are able to accurately learn the feasibility boundaries while also improving the system's knowledge of the optimal values of the quality of interests QoIs. This is exemplified in our application of the method to a tri-objective, multi-constrained design problem over the Mo-Nb-Ti-V-W system, a complex multi-physics problem space. The efficiency and effectiveness of our approach are further enhanced when it is implemented with a batch variant in the BO stage. With this method, we are able to make confident and strategic decisions that lead to successful design outcomes.

The deployment of our framework within the Mo-Nb-Ti-V-W high entropy alloy system resulted in the identification of 21 constraint-satisfying Pareto optimal alloys. Importantly, the

framework converges on a Pareto front of alloys that is interpretable. With regard to *constraint satisfaction*, we find that alloys that meet constraints relevant to GTE blades are lean in W and Mo due to the dominance of the density constraint. On the other hand, Ti- and V-rich alloys failed the minimum solidus temperature constraint. When considering the multi-objective optimization problem, compositions along the tri-objective Pareto front were found to have more W when near the strength axis. At the same time, they were rich in Nb when the alloys were near the axes for both ductility indicators. We note that identifying these Pareto optimal alloys with a brute force approach would have required the querying of ~10,000 alloys for five constraints each, just to learn the feasible space. On the other hand, the proposed framework learns the feasible space and identifies the Pareto set in ~700 queries. Furthermore, we demonstrate that employing a batch querying policy after the feasible space has been identified can decrease the time required to identify the Pareto set by ~95%.

RESULTS

Definition of design problem

Alloys suitable for GTE blades must meet several objectives and satisfy numerous constraints. For the sake of simplification, in this work, we consider two opposing types of design objectives, summarized in Table 1. On the one hand, the alloy must have high strength at high temperatures in order to carry the necessary structural loads during operation. On the other hand, the alloy has to possess some degree of ductility at room temperature to minimize the risk of fracture.

In this work, we evaluate the HT (1300 °C) yield-strength objective using a physics-based model developed by Curtin and Maresca²². We consider this model to be the *truth model* for the HT yield strength objective, as detailed in the Methods section. This model relies on the assumption that a hypothetical homogeneous ‘average’ alloy has all the macroscopic properties of the true random alloy²². The model’s grounding assumption is that the intrinsic strength of compositionally complex BCC alloys originates from the increased ‘roughness’ of the landscape that dislocations must traverse to induce plastic deformation. The model is capable of incorporating temperature effects and has been found to be in moderately good agreement with available experimental data.

While models for the elongation at fracture (ϵ_f) of MPEAs are not available, the ductility of MPEAs can be *roughly* inferred from ground state properties of alloys, such as the Pugh ratio and the Cauchy pressure. These two indicators of ductility have been used extensively in the design of ductile MPEAs^{21,23–25}. In the context of metals, Pugh’s ratio is defined as the ratio of the bulk modulus to the shear modulus (B/G). Thus, B/G captures the extent of the plastic deformation (B) without fracture (G)²⁶. Pettifor²⁷ proposed Cauchy pressure as an indicator of intrinsic ductility/brittleness, which is the difference between two elastic constants C_{12} and C_{44} . A positive Cauchy pressure indicates non-directional metallic bonds resulting in intrinsic ductility of the crystal, whereas a negative Cauchy pressure corresponds to directional bonds and results in an intrinsically brittle crystal structure. Both indicators can be estimated with high-fidelity DFT frameworks at a great computational cost. However, as the MPEA composition space is combinatorially vast, sufficient exploration of the space is intractable using conventional brute-force approaches. In the case of this work, the *truth model* for both ductility objectives is the DFT-based Korringa–Kohn–Rostoker Green’s function (DFT-KKR-CPA) method, as detailed in the Methods section.

In addition to the objectives associated with strength and ductility, candidate alloys for next-generation GTE blades must satisfy several constraints. Feasible alloys must have a sufficiently high solidus temperature to operate in the hot zone of the

Table 1. The five constraints and the three objectives associated with the design problem addressed in this work.

Property	Constraint/Objective
Solidus Temperature	$T_s \geq 2000$ °C
Solidification Range	$\Delta T \leq 400$
Thermal Conductivity	$\kappa \geq 20$
Linear CTE	$\alpha_1 \leq 2\%$
Density	$\rho \leq 9\text{g/cc}$
Ductility	Max $C_{12} - C_{44}$
Ductility	Max B/G
HT Yield Strength	Max σ_{HT}

turbine. As such, we stipulate that the solidus temperature be greater than 2000 °C. Moreover, candidate alloys must also be lightweight, both to minimize centripetal forces caused by the rotation of the blades²⁸ and to reduce the total mass of the GTE system. For this reason, we stipulate feasible alloys must have a density of less than 9 g/cc. Alloys should also be designed with the thermal management system of the turbine blade in mind. As such, the material comprising the turbine blades must have high thermal conductivity to dissipate the large amounts of heat from the hot zone of the engine²⁹.

Additionally, the blade must be compatible with thermal barrier coatings. To ensure this, the linear thermal expansion from room temperature to 1300 °C must not exceed 2%/K. Furthermore, from the manufacturing standpoint, these alloys must be resistant to solidification tearing, a common concern during the synthesis/fabrication of metallic parts from melt precursors. While solidification tearing results from very complex physical processes, a narrow solidification range can protect against this failure mode. Here, we stipulate the solidification range not exceeding 400 °C. Finally, we want to note that the design constraints and objectives described above and summarized in Table 1 are derived directly from the challenge specifications by the Department of Energy’s ARPA-E ULTIMATE program². Thus, the present alloy design exercise has some practical relevance.

Deployment of framework

The proposed framework is structured by connecting Bayesian classification and Bayesian optimization loops. Starting with the Bayesian classification loop, the goal is to actively learn the boundaries separating the feasible and infeasible regions. Therefore, a binary classifier is a natural choice for such a condition. A Bayesian approach to learning the boundaries requires classifiers capable of providing uncertainty for class membership predictions. Thus, Gaussian process classifiers are employed to represent design constraint boundaries. A formal way to make uncertainty a comparable quantity is by representing it as entropy. Thus, active learning in the Bayesian classification framework is done by attempting to reduce the entropy associated with the classifiers via augmenting the prediction standard deviations provided by Gaussian process classifiers for a set of designs to the Shannon entropy formula. Once the reduction in entropy drops below a threshold, the predicted feasible regions are fed to the Bayesian optimization loop by generating feasible designs to be searched. The Bayesian optimization framework uses Gaussian process regressions (GPRs) to model objective functions and Expected HyperVolume Improvement (EHVI) as the acquisition function to suggest the most informative experiments to discover better approximations of the Pareto frontier. Note that the Bayesian classification loop runs in parallel to the Bayesian optimization loop in search of experiments that may significantly reduce

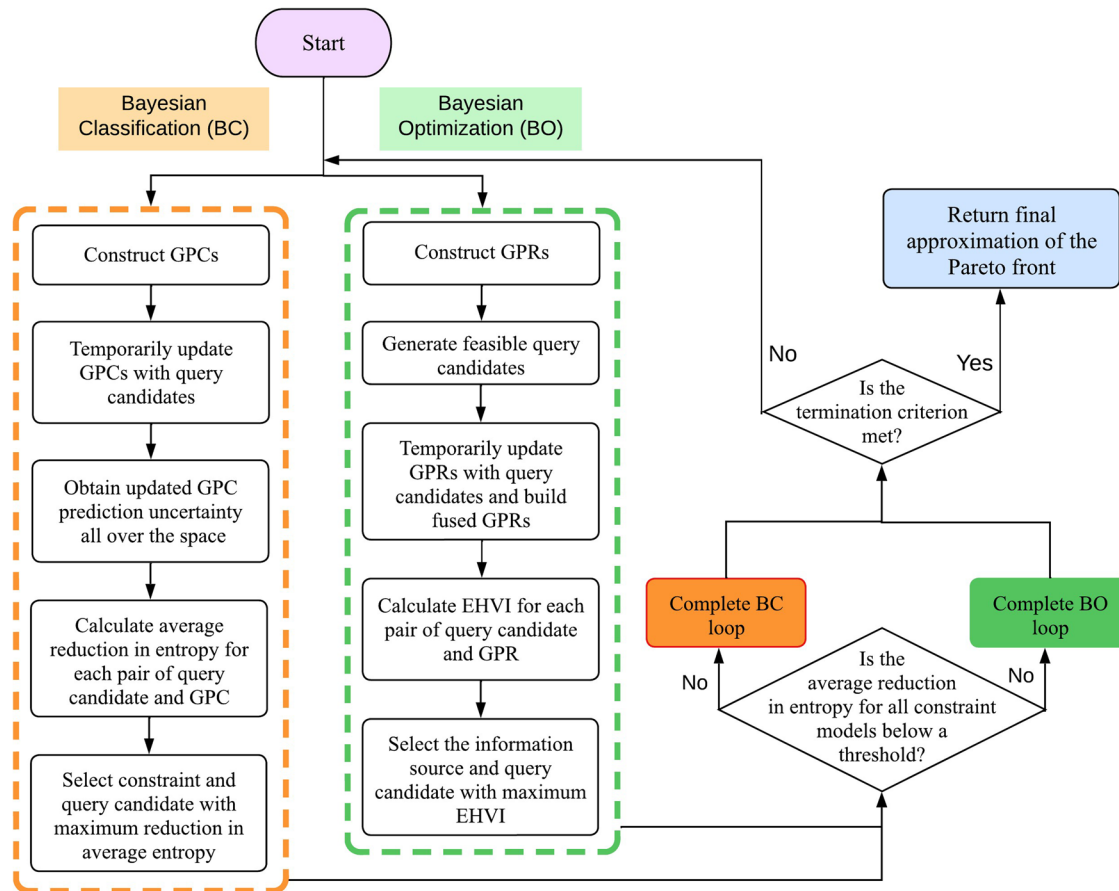


Fig. 1 Schematic of the Bayesian optimization framework with active learning of the design constraints. In every iteration of the framework, both Bayesian classification and Bayesian optimization loops run in parallel. The algorithm starts with Bayesian classification and switches to Bayesian optimization once the average reduction in entropy of all constraint models falls below a threshold. The framework switches back to Bayesian classification if a valuable experiment is suggested accordingly.

entropy. Thus, the framework is capable of *dynamically switching between both loops* depending on the expected information gain calculations. The schematic of the framework is illustrated in Fig. 1 (All codes will be publicly available upon the end of the project at the following Github repository: <https://github.com/Danialkh26/EBBC-MOBO>).

Figure 2 illustrates the overall results of implementing the proposed framework to solve the 3-objective, 5-constraint design problem in this study. A total of 700 iterations were completed, and the Bayesian optimization stage was initiated after iteration 420, when all the average entropy reduction plots flattened and dropped below 3%. At the beginning of the Bayesian optimization stage, classifiers were used to filter the design space, discarding the infeasible regions first. As more queries were made to the objective functions, better estimations of the Pareto front were obtained, as indicated by improvements in the hypervolume value. Initially, larger improvements were observed. However, the improvements gradually decreased, indicating convergence to the optimal Pareto front. It is important to note that all queries were made around one corner of the objective space corresponding to the maximum values of each quantity of interest, which confirms that the framework effectively recognized the optimal design region and is searching that area to discover better non-dominated solutions.

While the aforementioned results are obtained using a sequential approach during the Bayesian optimization stage, we also consider a batch Bayesian optimization approach. Since there is no change in the Bayesian classification stage and the related

results, the batch process begins after the optimization stage is triggered. Employing the batch Bayesian optimization scheme enables the execution of 48 experiments in parallel. This is equivalent to processing a batch of 48 samples at every single iteration at no or low additional costs. While economies of scale are likely to be more modest in the context of actual physical experiments, in this computational study, the batch of 48 simultaneous calculations was executed at no additional cost (per sample).

By employing the batch Bayesian optimization scheme, the same hypervolume improvement is obtained in only 13 iterations, as a comparison is shown in Fig. 3. In contrast, 280 iterations were needed while exploring the Pareto set using sequential MOBO. This corresponds to a reduction in the time necessary to discover the Pareto set of 95%. While the total cost (in terms of supercomputing time) associated with the calculations was roughly the same in both cases, there is a significant opportunity cost incurred during sequential BO by not learning the Pareto set early enough. Assuming each iteration lasts one day, it is much more valuable to learn the design capabilities of an alloy system in just 2 weeks rather than 9 months. In this context, batch-based strategies can significantly reduce opportunity costs related to long development times.

The fact that the batch and sequential BO schemes show the same hypervolume improvement at convergence means that they achieve a predicted Pareto set of similar quality. However, the non-dominated designs found (i.e., the alloys comprising the Pareto sets) may not necessarily be the same due to the high

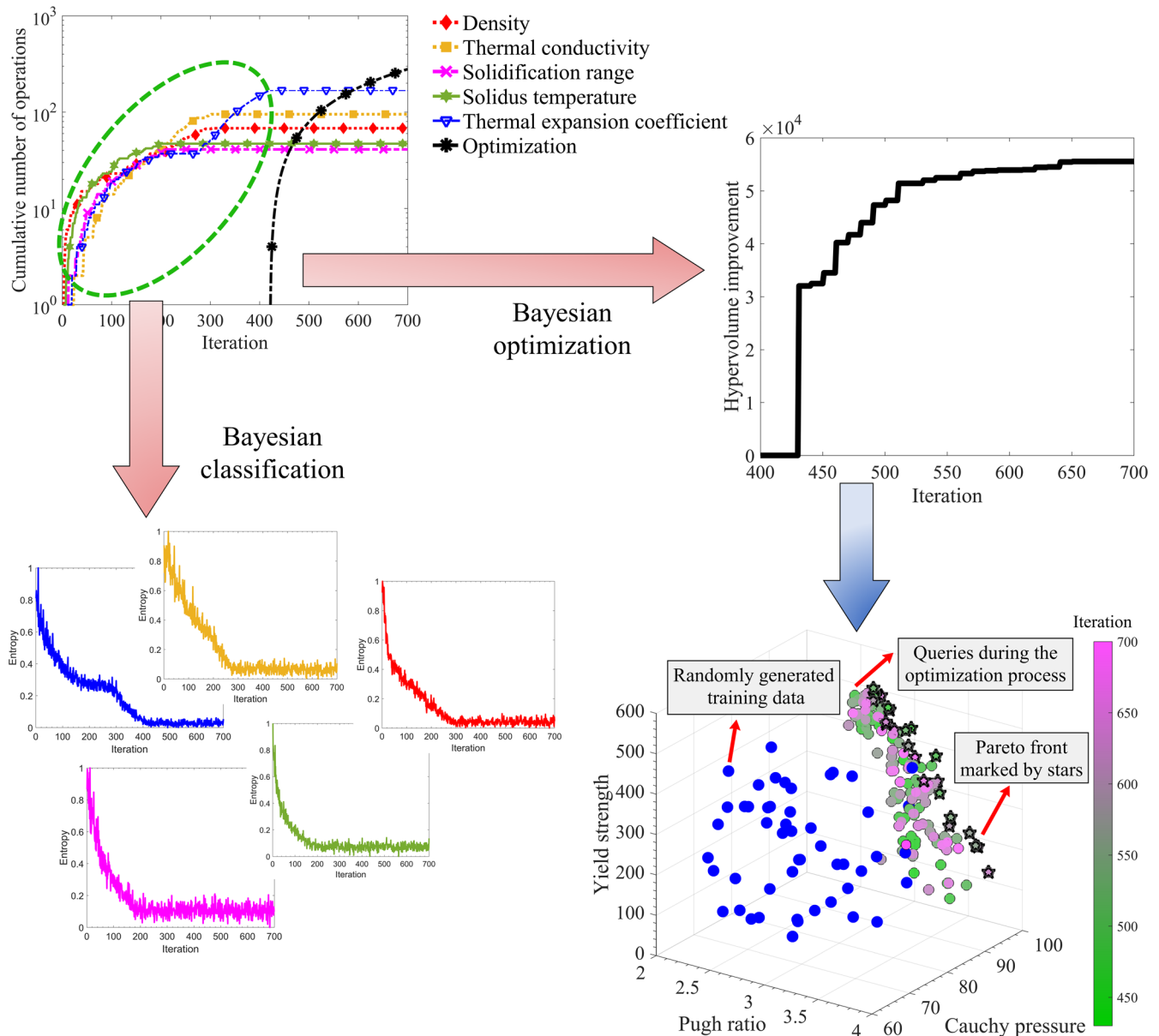


Fig. 2 Overall results of the 5-constraint 3-objective material design problem. The figure shows the application of the proposed framework to solve the problem. The process begins with learning the constraint boundaries by querying the constraints, effectively reducing the entropy associated with each classifier that represents a specific constraint. Once the entropy curves for all classifiers are flattened, Bayesian optimization begins to learn the non-dominated design region. As the estimations of the Pareto front improve, the hypervolume increases respectively. The figure also includes an illustration of the objective space, showing all the queries to the ground truth model and the final estimation of the Pareto front.

dimensionality of the input space and the stochastic nature of the BO process.

Regarding the discovery of constraint satisfying candidate alloys, UMAP (Uniform Manifold Approximation and Projection) in Fig. 4a–c shows that alloys rich in Ti, Mo, and particularly W fail one or more of the five constraints, depicted in gray. For a more quantitative view of this filtering process, in Fig. 4d, a Kernel Density Estimate (KDE) is fit over the frequency at which elements at various concentrations remain after filtering to visualize the *chemical signature* of the resultant feasible space. In these chemical signature plots, we see that the Ti and Mo signatures are slightly shifted back, indicating a slight depletion in these elements. On the other hand, the W signature is shifted back significantly, indicating W-rich alloys fail at least one of the design constraints.

The optimization portion of the framework converged on 21 Pareto-optimal alloys. The best-performing alloys with regard to the ductility indicators are rich in Nb. This can be seen in the UMAP, where the Pareto-optimal alloys, represented by stars, are located near the Nb-rich corner of the diagram. On the other hand, Pareto-optimal alloys that perform the best with regards to the HT yield strength metric have higher W-content. Again, this can be seen in the UMAP, where Pareto-optimal alloys approach the border of the diagram until reaching the edge of the feasible region. Likewise, the alloys that strike a trade-off between these three objectives have a wide range of potential Nb and W contents. This range of Nb and W contents can be seen in the chemical signature of the Pareto set, where the chemical signature of these two elements has broad peaks. These alloys and their associated objective and constraint values are summarized in

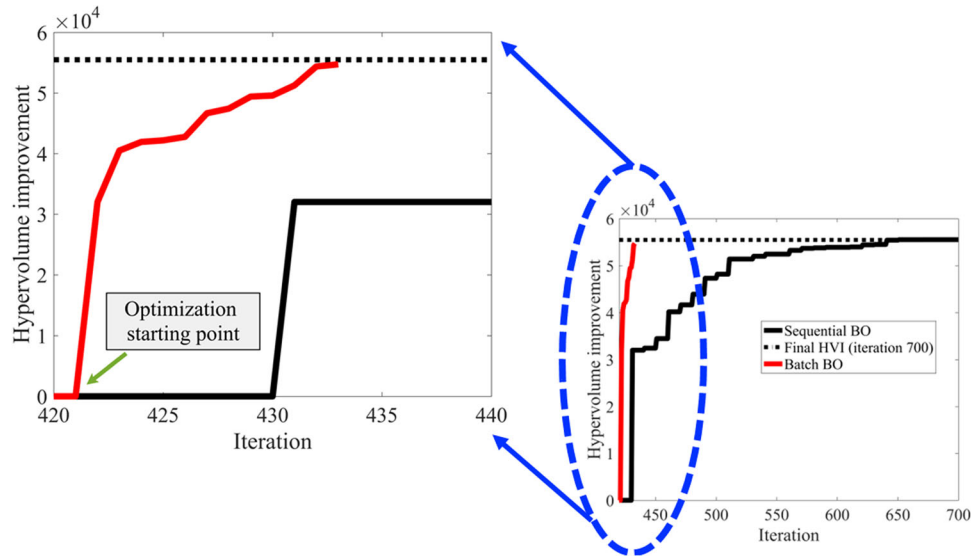


Fig. 3 Comparison of hyper-volume improvements in batch and sequential Bayesian optimization. Only 13 iteration is needed to reach the same Pareto front estimation quality in comparison to the sequential approach. 48 cores are accessible in our supercomputing system. Thus, it is possible to run 48 experiments in parallel without additional wall-time in batch Bayesian optimization case.

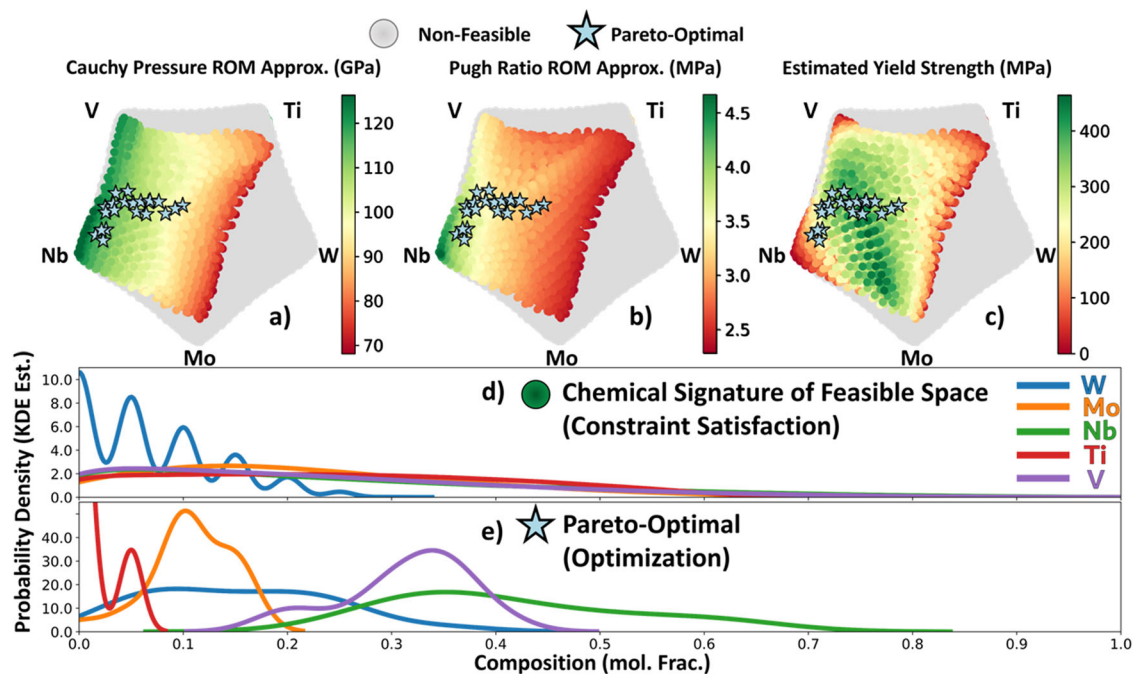


Fig. 4 Visualizations of constraint-passing and Pareto-optimal alloys. **a** ROM Cauchy Pressure (indicator of true objective) plotted over the design space. **b** ROM Pugh Ratio (indicator of true objective) objective plotted over the design space. **c** Estimated yield strength from the Curtin–Maresca model plotted over the design space. **d** Chemical signature of the feasible chemical space. **e** Chemical signature of Pareto-optimal set of alloys.

Table 2. We recommend further investigation of these 21 Pareto-optimal alloys to properly characterize their behavior in the context of GTE blade applications.

DISCUSSION

To benchmark, the performance of the constraint-satisfaction aspect of the proposed framework, a factorial exploration of the space was performed. The information sources for the 5 constraints were queried at increments of 5 at% considering binaries to quaternaries resulting in 10,626 queries of each model

(53,130 queries in total). Using the proposed batch active learning of constraints, only 420 queries were required to learn the extent of the feasible design space, demonstrating the improved efficiency of the proposed framework over a brute-force approach, with a total reduction in the effort of ~96%. Here we note that while in this work, the constraints were evaluated computationally at relatively modest cost, in a real physical setting such a reduction in effort would have a dramatic impact on the feasibility of experimental campaigns.

We note that the classification of the feasible space has arrived at interpretable results. Regarding the solidus temperature,

Table 2. The set of alloys that lie on the tri-objective strength-ductility Pareto-front identified by the proposed framework.

Mo	Nb	Ti	V	W	B/G	$C_{12} - C_{44}$ (GPa)	σ_y (MPa)
0.073	0.299	0.002	0.37	0.256	3.06	91.3	529
0.093	0.407	0.002	0.27	0.228	3.15	94.7	466
0.154	0.289	0.008	0.322	0.227	2.99	91.5	524
0.051	0.399	0.009	0.374	0.167	3.35	95.6	408
0.093	0.463	0.028	0.301	0.115	3.45	93.6	333
0.117	0.355	0.008	0.332	0.188	3.16	93.0	471
0.151	0.277	0.013	0.301	0.258	2.91	90.5	535
0.084	0.489	0.03	0.319	0.078	3.56	93.4	282
0.129	0.584	0.025	0.233	0.029	3.71	96.0	209
0.082	0.519	0.009	0.348	0.042	3.69	94.7	247
0.111	0.37	0.011	0.341	0.167	3.23	91.8	444
0.009	0.313	0.029	0.372	0.277	3.08	90.3	494
0.18	0.29	0.002	0.333	0.195	3.04	92.3	516
0.125	0.553	0.019	0.238	0.065	3.54	93.8	257
0.077	0.614	0.026	0.198	0.085	3.71	96.6	203
0.078	0.375	0.011	0.429	0.107	3.43	93.7	361
0.096	0.655	0.024	0.195	0.030	3.83	95.8	154
0.047	0.337	0.015	0.342	0.259	3.07	90.1	500
0.143	0.343	0.006	0.328	0.180	3.12	91.2	479
0.116	0.421	0.015	0.345	0.103	3.37	92.0	366
0.075	0.37	0.011	0.39	0.154	3.27	92.5	414

85.96% percent of alloys pass the $T_s \geq 2000$ °C constraint. Alloys that fail this constraint are rich in Ti and V. This is to be expected, as Ti and V are the least refractory elements comprising this design space. Most alloys in the space (99.46%) pass the thermal conductivity constraint $\kappa \geq 20$ W/m/K. The few alloys that fail this constraint are, again, rich in the two elements with the lowest thermal conductivities, Ti and V. Again, this is likely due to the fact that Ti and V are the least refractory elements in this design space. In addition to Ti-rich and V-rich alloys, compositionally complex alloys are also more likely to fail this constraint due enhanced phonon and electron scattering leading to a decreased thermal conductivity, putting a slight penalty on more high-entropy alloys.

All alloys in the Mo-Nb-Ti-V-W space pass the thermal expansion coefficient constraint $CTE < 2\%$ /K. Regarding solidification range, 97.12% of candidate alloys pass the $\Delta T \leq 400$ K constraint. Alloys that fail this constraint are rich in W and Ti. This is to be expected as these W and Ti have the biggest difference in their melting temperatures i.e., 3422 °C and 1668 °C, respectively. Furthermore, increased alloy complexity increases the solidification range, again putting a penalty on high entropy alloys. Regarding density, 42.55% alloys pass the $\rho \leq 9$ g/cc constraint. The three most refractory elements, W, Mo, and to a lesser extent Nb, fail this constraint. Figure 5 depicts a summary of this filtering.

Likewise, the *optimization* aspect of the framework has converged on results that can be understood using metallurgical intuition. The fact that Nb-rich alloys perform well concerning the ductility objectives agrees with other works where Nb is to enhance the ductility of otherwise brittle RHEAs³⁰. The Pugh ratios and Cauchy pressures of these 21 Pareto-optimal alloys are on the order of 3.32 ± 0.266 and 93.1 ± 1.92 GPa, respectively. These values are comparable to the ductile refractory MPEAs TiHfV/NbTa ($B/G = 3.817$, $C_{12} - C_{44} = 75$ GPa, $\epsilon_{frac} = 12.6\%$)³¹ and NbMoTaWTi ($B/G = 2.74$, $C_{12} - C_{44} = 73$, $\epsilon_{frac} = 13\%$)³². Regarding yield strength, increasing the W content within MPEAs has been shown to increase the yield strength of alloys³³.

To further benchmark the performance of the optimization aspect of the proposed framework, we carried out a DFT analysis of the Pareto-front. For example, in Fig. 6, we analyzed the correlation of at.% Nb and V, (as both are from same group in the periodic table) on key DFT quantities such as formation energy (E_{form}), intrinsic-strength, and Pugh's ratio³⁴.

In Fig. 6a, b, we plot E_{form} with respect to (Mo + Nb) and V concentration, respectively, where an increase in at.% (Nb with Mo) increases the alloy stability while increasing at.% V destabilizes the BCC phase. We found that there is an optimal V (<50 at%) or Mo + Nb (>50 at%) concentration that stabilizes the alloy. On the temperature scale, the 25 meV is equivalent to 300 K (RT; 27 °C), i.e., all predicted HEAs (except one) show RT stability.

The intrinsic strength (bulk moduli, B) and Pugh's ratio (i.e., ductility indicator) in Fig. 6c, d shows a strong correlation with V+Nb composition for predicted HEAs in Table 2. As seen in Fig. 6c, the intrinsic strength decreases sharply with increasing V + Nb concentration, while Pugh's ratio (shown in Fig. 6d) increases. Alloys with Pugh's ratio (G/B) < 0.57 are considered ductile based on Pugh's criteria²⁶. Furthermore, a good correlation is observed between framework-predicted properties in Fig. 4 and DFT calculations in Fig. 6 for increasing Nb composition. This correlation suggests the utility of such frameworks for reliable exploration and understanding of the strength-ductility trade-offs in HEAs.

In light of recent initiatives for ICME-enabled closed-loop design platforms and autonomous materials discovery, it is important to note that the methodology used in this work, while conducted in silico, can also be used to guide experimental exploration of design spaces. One possible approach would be to use computational models to initially reduce the design space by applying relaxed constraints to eliminate candidates that are likely to fail one or more constraints, such as predicted thermal conductivity greater than $10W/m - K$. This initial filtering could then be followed by experimental campaigns to more accurately determine the true boundaries of the constraint-satisfying regions in the design space using stricter constraints, such as thermal conductivity greater than $20W/m - K$. A possible design of experiments could include using a dilatometer to measure the thermal expansion coefficient³⁵, a densimeter to measure the density³⁶, a laser flash apparatus to measure thermal conductivity³⁷, and a high-temperature tensile testing rig to measure the yield strength and elongation at yield³⁸. For constraints related to the solidus and solidification range, the design space could be reduced by relying on CALPHAD-based predictions, as it is currently not feasible to experimentally determine the melting temperature for such refractory alloys in an HTP manner. After reducing the design space, an experimental campaign could be undertaken to optimize simultaneously for strength and ductility. The proposed framework can be useful for autonomous and closed-loop material design campaigns, as depicted in Fig. 7.

METHODS

In this study, we proposed and implemented an approach to solving constrained multi-objective design problems by deploying a Bayesian classification and optimization-based active learning strategy. The framework is capable of handling an arbitrary number of objectives and constraints. Moreover, the Bayesian classification scheme uses an entropy-based measure to select an optimal sequence of informative experiments. As a result, this approach can identify the feasible boundaries on the design space in a more efficient manner compared to previous approaches²¹ by incorporating the uncertainty provided by Gaussian process classifiers regarding the class membership predictions. The superiority of our MOBO framework is that it employs a Bayesian classification approach that can handle any number of constraints and recognizes the feasible regions regarding each constraint

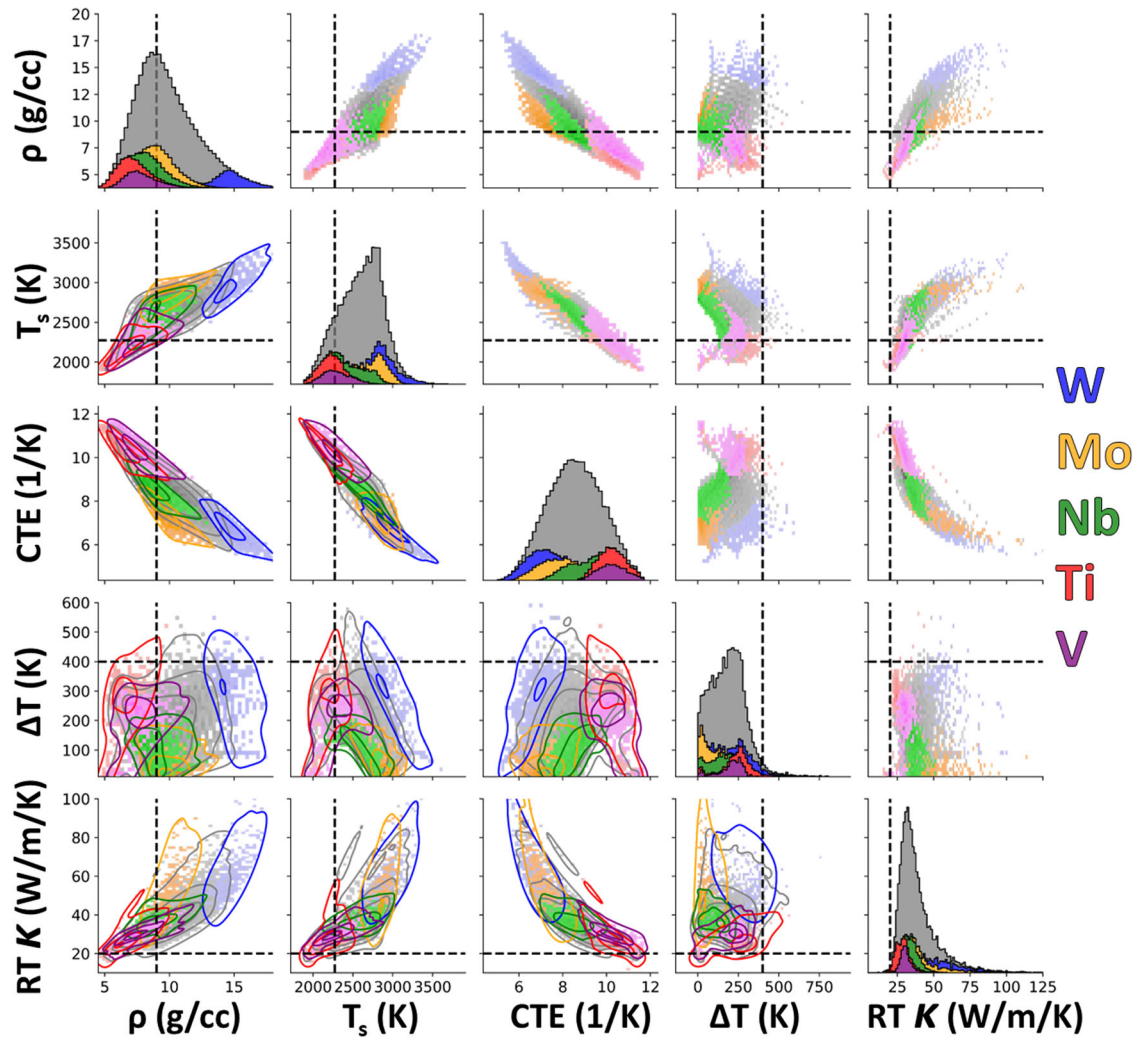


Fig. 5 Pairwise plot demonstrating correlations and trade-offs between the 5 constraints applied to the design space. Alloys that are comprised of more than 50% of a particular element are colored accordingly. Alloys that do not have a majority element are colored in gray. Diagonal rows depict property distributions for each class of alloy. The lower-left triangle depicts Kernel Density Estimate (KDE) estimates over joint property distributions to better visualize the structure of the data.

without spending a substantial computational budget to obtain training data required for accurately distinguishing the feasible and infeasible regions. Since the models representing the constraints are not computationally cheap to evaluate, it is vital to manage the available resources to make observations on alloys with the greatest values in them.

To determine the overall uncertainty of a classifier, the class memberships of a set of randomly generated samples are checked. However, the labels are not informative here, but the uncertainty of the predictions in the form of standard deviation is used to calculate the entropy. As the classifiers get more information in terms of boundaries, the standard deviations get smaller, and so does the entropy. Here, a criterion is defined by the user to make the transition to the Bayesian optimization stage once all classifiers are confident enough in terms of label predictions. Since the entropy data is noisy because, at every iteration, a different set of samples are generated in the composition space to make sure it does not overlook any part of the space, a window of 50 iterations is considered to calculate the average reduction in entropy (in distances of 25 iterations). In this case, we stop considering a constraint among the possible experiments for the next step if this average drops below 3 percent.

Once all constraints meet the defined criteria, the Bayesian optimization stage begins; however, the framework still keeps track of entropy values for all constraints at every iteration of the process so that if it finds an experiment of great value (when the average entropy reduction jumps greater than 3 percent), it may switch to the classification stage and perform that experiment. This dynamic decision-making approach makes the framework capable of switching between classification and optimization stages when necessary. Below, all the ingredients of this framework are introduced.

Gaussian process regression

Surrogate models are essential for a Bayesian optimization framework to model black-box functions, given prior observations made from these functions. Moreover, surrogate models make it possible to search the space at low computational costs, looking for the best next experiment that adds the most information about the optimum design to the system. This work uses GPRs as surrogates to model the objective functions³⁹. Gaussian process models are powerful tools for probabilistic modeling due to the ease with which models can be updated with newly acquired information. Moreover, they provide probabilistic predictions that model the uncertainty associated with unobserved regions in a given design space. Finally, GPs are

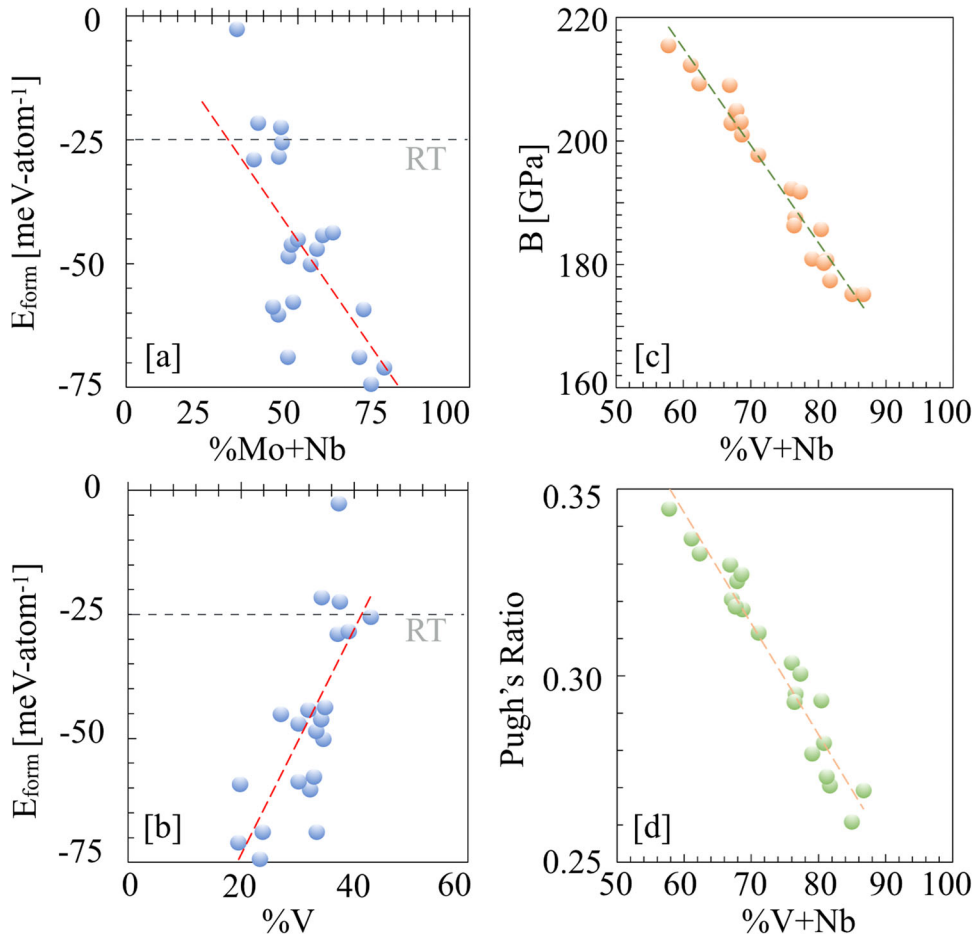


Fig. 6 Phase stability, intrinsic-strength, and ductility. **a, b** Formation energy plotted with respect to Mo+Nb and V concentration for 21 MPEAs in Table 2. **c** Intrinsic strength (bulk-moduli), and **d** Pugh's ratio with respect to V + Nb concentration.

constructed with an intrinsic notion of distance (or correlation) between points in a design space. This correlation is exploited when predicting the model uncertainty.

Since more than one model may represent the same quantity of interest, each model needs its own GPR. These models are considered as different sources that the system has access to gain required information about a quantity of interest—such frameworks are known as multi-information source approaches. Following refs. ^{15,16}, we formulate the surrogates (GPRs) by assuming we have available multiple information sources, $f_i(\mathbf{x})$, where $i \in \{1, 2, \dots, S\}$, to estimate a quantity of interest, $f(\mathbf{x})$, at design point \mathbf{x} . These surrogates are indicated by $f_{GP,i}(\mathbf{x})$. Assuming there are N_i evaluations of information source i denoted by $\{\mathbf{X}_{N_i}, \mathbf{y}_{N_i}\}$, where $\mathbf{X}_{N_i} = (\mathbf{x}_{1,i}, \dots, \mathbf{x}_{N_i,i})$ represents the N_i input samples to information source i and $\mathbf{y}_{N_i} = (f_i(\mathbf{x}_{1,i}), \dots, f_i(\mathbf{x}_{N_i,i}))$ represents the corresponding outputs from information source i , then the posterior distribution of information source i at design point \mathbf{x} is given as

$$f_{GP,i}(\mathbf{x}) | \mathbf{X}_{N_i}, \mathbf{y}_{N_i} \sim \mathcal{N}(\mu_i(\mathbf{x}), \sigma_{GP,i}^2(\mathbf{x})) \quad (1)$$

where

$$\begin{aligned} \mu_i(\mathbf{x}) &= K_i(\mathbf{X}_{N_i}, \mathbf{x})^T [K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i}) + \sigma_{n,i}^2 \mathbf{I}]^{-1} \mathbf{y}_{N_i} \\ \sigma_{GP,i}^2(\mathbf{x}) &= k_i(\mathbf{x}, \mathbf{x}) - K_i(\mathbf{X}_{N_i}, \mathbf{x})^T \\ &\quad [K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i}) + \sigma_{n,i}^2 \mathbf{I}]^{-1} K_i(\mathbf{X}_{N_i}, \mathbf{x}) \end{aligned} \quad (2)$$

where k_i is a real-valued kernel function, $K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i})$ is the $N_i \times N_i$ matrix whose m, n entry is $k_i(\mathbf{x}_{m,i}, \mathbf{x}_{n,i})$, and $K_i(\mathbf{X}_{N_i}, \mathbf{x})$ is the $N_i \times 1$ vector whose m^{th} entry is $k_i(\mathbf{x}_{m,i}, \mathbf{x})$ for information

source i . We have also included the term $\sigma_{n,i}^2$, which is used to model observation error for information sources based on experiments or expert's opinion. Note that the term signal variance is to cover two sources of uncertainty: the variance associated to the GPR estimation of the objective function and the variance associated to the information source with respect to the highest fidelity model, also known as the ground truth.

Gaussian process classification

In Bayesian classification frameworks, similar to Bayesian optimization technique, Bayes' theorem can be employed but to calculate the joint probability $p(y, \mathbf{x})$, where y is the class label:

$$p(y | \mathbf{x}) = \frac{p(y)p(\mathbf{x}|y)}{\sum_{c=1}^C p(C_c)p(\mathbf{x}|C_c)} \quad (3)$$

Gaussian process classifications (GPCs) are probabilistic models that predict the probability of belonging to a specific class by putting a Gaussian process prior over a latent function $f(\mathbf{X})$ and computing the posterior distribution at a desired location \mathbf{x} ^{39,40}. GPCs are formulated similar to GPRs but with labeled data, instead of a continuous objective value, as follows:

$$\begin{aligned} \mu_i(\mathbf{x}) &= K_i(\mathbf{X}_{N_i}, \mathbf{x})^T [K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i})]^{-1} \mathbf{f}(\mathbf{X}) \\ \Sigma_i(\mathbf{x}) &= k_i(\mathbf{x}, \mathbf{x}) - K_i(\mathbf{X}_{N_i}, \mathbf{x})^T \\ &\quad [K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i})]^{-1} K_i(\mathbf{X}_{N_i}, \mathbf{x}) \end{aligned} \quad (4)$$

The class label predictions are obtained by performing Monte Carlo sampling from the calculated posterior distribution and then

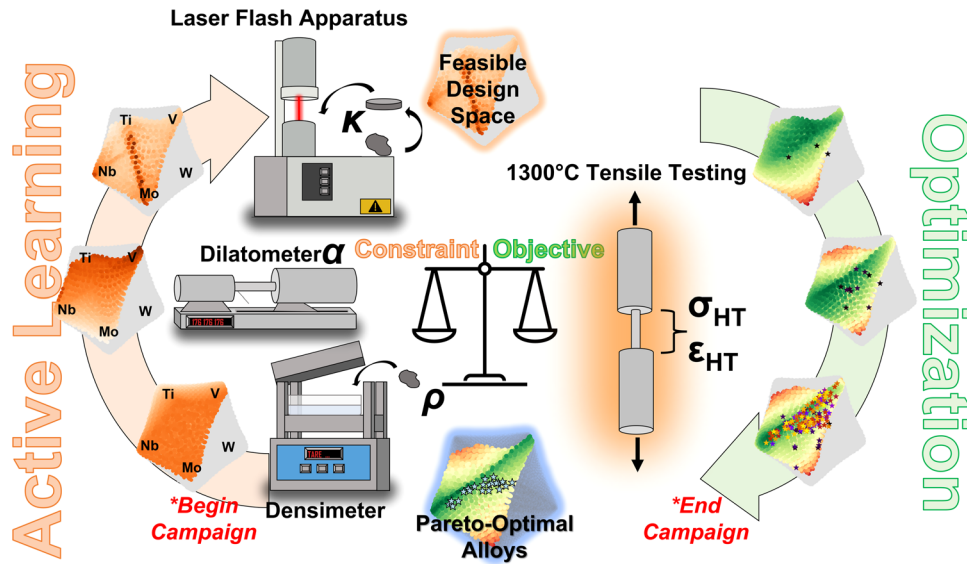


Fig. 7 Schematic representation of an experimental campaign utilizing the proposed framework. The process begins with a closed-loop exploration of the design space to identify the range of compositions that meet all requirements. This forms the feasible region of the design space. Optimization is then carried out within this region to identify a set of Pareto-optimal alloys as the final outcome of the campaign.

passing samples through a sigmoid function σ to ensure the output is bounded to $[0,1]$. Then the mean and variance of the obtained distribution define the class membership probability and associated uncertainty to the predicted label.

By utilizing a Bayesian methodology, the inclusion of uncertainty in the predictions is a crucial aspect in determining the expected utility value. Importantly, this feature differentiates Gaussian Process Classification (GPC) as a probabilistic model from other classification methods. As a result, GPC is particularly well-suited for applications that involve probabilistic frameworks and machine learning tasks. A more detailed discussion is presented in ref. ³⁹.

Active learning in Bayesian classification

As mentioned earlier, GPCs are probabilistic models well-suited for Bayesian classification frameworks because they provide uncertainty associated with the predicted class memberships. The class membership predicted by a GPC of information source i is a random variable defined via a normal distribution $\mathbb{Y} \sim \mathcal{N}(\mu_i(\mathbf{x}), \sigma_i^2(\mathbf{x}))$. A Bayesian classification framework aims to reduce the overall classifier's uncertainty associated with class membership predictions. To further quantify the uncertainty of a classifier, a measure is needed to compare how newly added information to the system may help to achieve more accurate classifiers. Entropy is a natural choice here to determine the uncertainty of different models.

Herein, we propose to use the uncertainty in form of standard deviation assigned to class membership predictions of a GPC. Then, we employ the discrete entropy formula to determine the entropy:

$$H = - \sum_{j=1}^k \sigma_j \log(\sigma_j) \quad (5)$$

where we have predicted the labels of k samples randomly generated, and σ_j is the standard deviation of the predicted class membership provided by the GPC. The more accurate a classifier is about the boundary, the less uncertain it will be about the assigned labels. Such a decrease in uncertainty is manifested as a lower model/classifier entropy. Employing the entropy measure as the utility function in a Bayesian classification framework, we can

recognize the best next experiment to make and update the system that results in the most significant reduction in a classifier's entropy.

Information fusion of multiple sources

Several approaches exist for fusing multiple sources of information, such as Bayesian modeling averaging^{41–46}, the use of adjustment factors^{47–50}, covariance intersection methods⁵¹, and fusion under known correlation^{52–54}.

In engineering design, there often exist multiple models that represent the same system of interest. Each model provides valuable information about the quantity of interest. By combining all of this knowledge through a process known as model fusion, more accurate and less biased models can be produced. As more sources are incorporated into the fusion process, it is commonly expected to see a reduction in the variance of the quantity of interest estimates. However, this is not always the case with other fusion techniques such as Bayesian model averaging^{41–46,55}, the use of adjustment factors^{47–50}, covariance intersection methods⁵¹, with the exception of fusion under known correlation^{52–54}.

Unlike some multi-fidelity methods, our approach does not rely on any assumptions about the relative importance of the information sources. As a result, it is crucial to establish the correlation between information sources prior to the fusion process. We use a technique called the reification process to estimate the correlation coefficients between the different information sources^{56,57}. In accordance with the methodology outlined in previous studies such as refs. ^{15,56–58}, once the correlation coefficients are determined, the fused mean and variance at a specific design point \mathbf{x} can be defined using the method proposed in ref. ⁵⁴.

$$\mathbb{E}[\hat{f}(\mathbf{x})] = \frac{\mathbf{e}^T \tilde{\Sigma}(\mathbf{x})^{-1} \boldsymbol{\mu}(\mathbf{x})}{\mathbf{e}^T \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \quad (6)$$

$$\text{Var}(\hat{f}(\mathbf{x})) = \frac{1}{\mathbf{e}^T \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \quad (7)$$

where $\mathbf{e} = [1, \dots, 1]^T$, $\boldsymbol{\mu}(\mathbf{x}) = [\mu_1(\mathbf{x}), \dots, \mu_S(\mathbf{x})]^T$ given S models, and $\tilde{\Sigma}(\mathbf{x})^{-1}$ is the inverse of the covariance matrix between the information sources. A more detailed discussion and examples can be found in Refs. ^{12,14,16,56,59–62}.

Multi-objective optimization

A multi-objective optimization problem is defined as

$$\text{minimize } \{f_1(\mathbf{x}), \dots, f_n(\mathbf{x})\}, \mathbf{x} \in \mathcal{X} \quad (8)$$

where $f_1(\mathbf{x}), \dots, f_n(\mathbf{x})$ are the objectives and \mathcal{X} is the feasible design space. In multi-objective optimization, it is typical that there is no single solution that simultaneously optimizes all objectives. Rather, the optimal solutions are represented by a set of non-dominated designs, which form the Pareto front in the objective space. In this context, the optimal solutions, \mathbf{y} , to a multi-objective optimization problem with n objectives, are denoted as $\mathbf{y} \prec \mathbf{y}'$ and can be expressed as

$$\{\mathbf{y} : \mathbf{y} = (y_1, y_2, \dots, y_n), y_i \leq y'_i \quad \forall i \in \{1, 2, \dots, n\}, \exists j \in \{1, 2, \dots, n\} : y_j < y'_j\} \quad (9)$$

where $\mathbf{y}' = (y'_1, y'_2, \dots, y'_n)$ denotes any possible objective output. The set of $\mathbf{y} \in \mathcal{Y}$, where \mathcal{Y} is the objective space, is known as the Pareto front.

There are various techniques for estimating the Pareto front in multi-objective optimization problems, such as the weighted sum approach⁶³, the adaptive weighted sum approach⁶⁴, normal boundary intersection methods⁶⁵ and hypervolume indicator methods^{66–72}, among others. In Bayesian optimization frameworks, hypervolume indicator approaches are well-suited to handle the probabilistic nature of these frameworks and to approximate the Pareto front of solutions efficiently. We adopt the methodology presented in refs. 17,73 to conduct Bayesian optimization of multi-objective functions in multi-fidelity settings. For a detailed explanation of the calculation of the EHV, we refer readers to ref. 9.

Models for constraints

In this work, the truth models for all five constraints were derived from high-fidelity CALCULATION of PHase Diagrams (CALPHAD)-based simulations. The high entropy alloy database TCHEA5 and Thermo-Calc's equilibrium simulation were used to calculate the density, solidus, solidification range, CTE, and thermal conductivity. Specifically, the solidus and solidification range (difference between solidus and liquidus temperatures) were extracted from phase diagrams generated from CALPHAD models⁷⁴. The coefficient of thermal expansion (CTE) was calculated by using Thermo-Calc to estimate the equation of state of the system at a given reference temperature T_0 and a target temperature T_f and then determining the amount of expansion at those two temperatures⁷⁴. The density was calculated in the same manner. Finally, thermal conductivity was determined by querying fitted polynomials within the CALPHAD database. In situations where separate data for electronic and lattice thermal conductivities is not available, the thermal conductivity was estimated using the Slack model^{74–76} for lattice thermal conductivity and the Wiedemann–Franz Law^{74,77} for electronic thermal conductivity, then summing these two estimates to obtain the overall thermal conductivity⁷⁴. The Thermo-Calc's API, TC-Python, was used to integrate these models with the proposed framework.

Models for objectives

The yield strength objective in this study is modeled using the analytical framework proposed by Curtin and Maresca in ref. 22. This model considers the behavior of an edge dislocation within a random solute field present in a body-centered cubic high-entropy alloy. In order to minimize the energy associated with the dislocation in this random alloy, the dislocation adopts a wavy configuration. This allows the dislocation to avoid high-energy areas in the medium due to dislocation-solute interactions while being attracted to and pinned to areas with lower energy from such interactions. This wavy configuration results in an increased line tension, which represents the energy cost of this

configuration. However, the characteristic waviness also minimizes the overall energy of the dislocation by simultaneously reducing the energy associated with the interaction between the edge dislocation and the solute field and the energy associated with the line tension of the edge dislocation. A statistical analysis of the energy barrier required for thermally activated edge glide was carried out, leading to the following equations:

$$\tau_{y0} = 0.040a^{-1/3}\bar{\mu}\left(\frac{1+\bar{\nu}}{1-\bar{\nu}}\right)^{4/3}\left[\frac{\sum_n c_n \Delta V_n^2}{b^{-6}}\right]^{2/3} \quad (10)$$

$$\Delta E_b = 2.00a^{1/3}b^{-3}\bar{\mu}\left(\frac{1+\bar{\nu}}{1-\bar{\nu}}\right)^{2/3}\left[\frac{\sum_n c_n \Delta V_n^2}{b^{-6}}\right]^{1/3} \quad (11)$$

$$\tau_y(T, \dot{\epsilon}) = \tau_{y0} \left[-\frac{1}{0.55} \left(\frac{k_B T}{\Delta E_b} \ln \frac{\dot{\epsilon}_0}{\dot{\epsilon}} \right)^{0.91} \right] \quad (12)$$

$$\sigma_y(T, \dot{\epsilon}) = M\tau_y \quad (13)$$

The variables in the above equation are as follows: a is the line tension parameter and is set to 1/12 for edge dislocations; $\bar{\mu}$ is the average shear modulus of the alloy; $\bar{\nu}$ is the average Poisson ratio of the alloy; b is the Burger vector associate with the BCC edge dislocation within the random alloy; ΔV is the misfit volume of the n th solute, which can be accurately estimated as $\Delta V_n = V_n - \sum_{n=1}^n c_n V_n$ according to Vegard's law; τ_{y0} is the zero-temperature yield stress; ΔE_b is the energy barrier for the thermal-activated flow; $\dot{\epsilon}_0$ is the reference strain rate which is typically set to 10^4 s^{-1} ; $\dot{\epsilon}$ is the applied strain rate which is typically set to 10^3 s^{-1} and is indeed set to this value in the current work; M is the Taylor factor for edge glide in a random BCC polycrystal; k_B is the Boltzmann constant; $\sigma_y(T, \epsilon)$ is the yield strength estimated at a finite temperature and strain rate, T and $\dot{\epsilon}$.

In this study, we employed the DFT-based KKR (Korringa-Kohn-Rostoker Green's function) method as the reference model for calculating key properties such as intrinsic strength and phase stability for arbitrary compositions. The method uses a coherent-potential approximation (CPA) to account for direct configurational averages over chemical disorder^{78,79}. The gradient-corrected exchange-correlation functional provided by Perdew, Burke, and Ernzerhof (PBE)⁸⁰ was used in the DFT-KKR-CPA calculations, which were used to obtain bulk moduli and derived quantities such as shear moduli and Pugh's ratio. A $24 \times 24 \times 24$ Monkhorst-Pack mesh was used for Brillouin zone integrations and core-electrons were treated relativistically, while valence-electrons were handled scalar-relativistically, with no spin-orbit coupling⁷⁸. The Fermi energy was determined by integrating the complex Green's function on a semicircular energy contour of 25 points on a Gauss-Chebyshev mesh⁷⁸.

The mechanical properties (bulk and shear moduli) of down-selected compositions were calculated and assessed by designing Super-Cell Random Approximates (SCRAPs)³⁴ and employing a computationally intensive stress-strain method as implemented within DFT based Vienna Ab-initio Simulation Package (VASP)^{81–84}. The PBE generalized-gradient approximation (GGA) functional⁸⁰ was employed for geometrical relaxations with total-energy and force convergence criteria of 10^{-6} eV and 0.01 eV/Å, respectively. The Brillouin-zone integration during ionic relaxation was performed on $1 \times 1 \times 1$ while mechanical properties calculations were calculated on $3 \times 3 \times 3$ k-mesh grid using Monkhorst-Pack method⁸⁵ with a plane-wave cutoff energy of 520 eV. The effect of the core electrons and interaction between the nuclei and the valence was treated by the projector-augmented wave (PAW)⁸⁶.

LIMITATIONS OF MODELS

Regarding the accuracy of the yield strength truth model, in their original publication using a similar refractory MPEA system and

its subsystems (Mo-Nb-Ta-V-W) Maresca et al.²² compared their model to experimental yield strength measurements captured over a range of temperatures. They determined their model has acceptable agreement with experiment (MAE = 126 MPa, RMSE = 138 MPa from 800 K to 1800 K) to be used for HTP alloy design. Furthermore, the Curtin–Maresca model has been successfully used in alloy design^{87,88}. Specifically, Rao et al.⁸⁷ showed that the Curtin–Maresca models accurately (MAE = 167 MPa, RSME = 209 MPa) predicts temperature-dependent yield strength of 4 refractory MPEAs at 5 temperatures ranging from 25 °C to 1200 °C.

DFT is the truth model for the two ductility indicators in this work, the Pugh ratio and Cauchy pressure. A large potential source of error in these calculations is the exchange–correlation functional. However, the exchange–correlation functional used in this work has been extensively tested in MPEA composition spaces^{21,89}. Specifically, DFT-calculated elastic constants were consistently within 10% to experimental values⁸⁹, which represents a smaller scatter compared to one calculated from elemental average.

Regarding the accuracy of these models, Thermo-Calc's equilibrium simulations have been able to successfully predicted phase stability⁹⁰, solidus temperatures⁹¹, thermal-conductivity^{92,93}, and thermal expansion coefficient^{93,94}. Specifically, Abu-Odeh et al.⁹⁰ benchmarked the TCHEA1 database against experiments and found that phase predictions from Thermo-Calc's equilibrium simulation were in 70.8% agreement with experimental data; the authors note that discrepancies between Thermo-Calc may lie in experimental procedures such as not providing enough time for the alloys to reach thermodynamic equilibrium. In this work, the 5th iteration of this database is used, likely increasing the accuracy of the model. Regarding the solidus constraint, Kirk et al. demonstrated that Thermo-Calc equipped with the TCHEA4 database accurately predicts the melting temperature of MPEAs (MAE = 10.5 K), outperforming ROM predictions of solidus temperatures. Regarding thermal conductivity, preliminary works conducted throughout the BIRDSHOT collaboration² (a project conducted for ARPA-e's ULTIMATE program) indicate that the Thermo-Calc's property module equipped with the TCHEA5 database is able to accurately predict thermal conductivity of BCC alloys (MAE = 14.9 W/m/K)⁹⁵. As previously stated, when data for a particular system is sparse, Thermo-Calc's thermal conductivity is informed by the Slack model⁹⁶. As such, the model used for thermal conductivity is *at least* as accurate as the commonly used⁹⁶ Slack model.

DATA AVAILABILITY

Data generated from the Bayesian Optimization/Classification framework are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

Codes associated with this work will be publicly available upon the end of the ULTIMATE program at the following Github repository: <https://github.com/Danielkh26/EBBC-MOBO>.

Received: 3 November 2022; Accepted: 16 March 2023;

Published online: 01 April 2023

REFERENCES

- Long, H., Mao, S., Liu, Y., Zhang, Z. & Han, X. Microstructural and compositional design of Ni-based single crystalline superalloys - a review. *J. Alloy. Compd.* **743**, 203–220 (2018).
- Ultrahigh Temperature Impervious Materials Advancing Turbine Efficiency (ULTIMATE). Advanced Research Projects Agency-Energy. <https://arpa.energy.gov/technologies/programs/ultimate> (2020).
- Yeh, J.-W. & Lin, S.-J. Breakthrough applications of high-entropy materials. *J. Mater. Res.* **33**, 3129–3137 (2018).
- Liu, X., Zhang, J. & Pei, Z. Machine learning for high-entropy alloys: Progress, challenges and opportunities. *Prog. Mater. Sci.* **131**, 101018 (2023).
- Jung, Y. et al. Investigation of phase-transformation path in TiZrHf(VNbTa), refractory high-entropy alloys and its effect on mechanical property. *J. Alloy. Compd.* **886**, 161187 (2021).
- Allison, J. Integrated computational materials engineering: A perspective on progress and future steps. *JOM* **63**, 15–18 (2011).
- Biswas, A., Morozovska, A. N., Ziatdinov, M., Eliseev, E. A. & Kalinin, S. V. Multi-objective bayesian optimization of ferroelectric materials with interfacial control for memory and energy storage applications. *J. Appl. Phys.* **130**, 204102 (2021).
- Solomou, A. et al. Multi-objective bayesian materials discovery: application on the discovery of precipitation strengthened NiTi shape memory alloys through micromechanical modeling. *Mater. Des.* **160**, 810–827 (2018).
- Zhao, G., Arróyave, R., Qian, X. Fast exact computation of expected hypervolume improvement. <https://arxiv.org/abs/1812.07692> (2018).
- Suzuki, S., Takeno, S., Tamura, T., Shitara, K., Karasuyama, M. Multi-objective bayesian optimization using pareto-frontier entropy. In *International Conference on Machine Learning*, pp. 9279–9288 (2020).
- Knowles, J. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Trans. Evol. Comput.* **10**, 50–66 (2006).
- Khatamsaz, D. et al. Efficiently exploiting process-structure-property relationships in material design by multi-information source fusion. *Acta Mater.* **206**, 116619 (2021).
- Khatamsaz, D. et al. Adaptive active subspace-based efficient multifidelity materials design. *Mater. Des.* **209**, 110001 (2021).
- Ghoreishi, S.F., Allaire, D.L. A fusion-based multi-information source optimization approach using knowledge gradient policies. In *AIAA/ASCE/AHS/ASC Struct. Dyn. Mater. Conf.*, p. 1159 (2018).
- Ghoreishi, S. F., Molkeri, A., Srivastava, A., Arroyave, R. & Allaire, D. Multi-information source fusion and optimization to realize ICME: Application to dual-phase materials. *J. Mech. Des.* **140**, 111409 (2018).
- Ghoreishi, S. F. & Allaire, D. Multi-information source constrained bayesian optimization. *Struct. Multidiscip. Optim.* **59**, 977–991 (2019).
- Khatamsaz, D., Peddareddygar, L., Friedman, S., Allaire, D. Bayesian optimization of multiobjective functions using multi information sources. *AIAA J.* 1–11 <https://doi.org/10.2514/1.J059803> (2021).
- Hickman, R. J., Aldeghi, M., Häse, F. & Aspuru-Guzik, A. Bayesian optimization with known experimental and design constraints for chemistry applications. *Digit. Discov.* **1**, 732–744 (2022).
- Kusne, A. G. et al. On-the-fly closed-loop materials discovery via bayesian active learning. *Nat. Commun.* **11**, 5966 (2020).
- Kusne, A. G., Keller, D., Anderson, A., Zaban, A. & Takeuchi, I. High-throughput determination of structural phase diagram and constituent phases using GRENDEL. *Nanotechnology* **26**, 444002 (2015).
- Khatamsaz, D. et al. Multi-objective materials bayesian optimization with active learning of design constraints: Design of ductile refractory multi-principal-element alloys. *Acta Mater.* **236**, 118133 (2022).
- Maresca, F. & Curtin, W. A. Mechanistic origin of high strength in refractory bcc high entropy alloys up to 1900K. *Acta Mater.* **182**, 235–249 (2020).
- Shaikh, S. M., Hariharan, V. S., Yadav, S. K. & Murty, B. S. CALPHAD and rule-of-mixtures: A comparative study for refractory high entropy alloys. *Intermetallics* **127**, 106926 (2020).
- Chen, L., Zhang, X., Wang, Y., Hao, X. & Liu, H. Microstructure and elastic constants of AlTiVMoNb refractory high-entropy alloy coating on Ti6Al4V by laser cladding. *Mater. Res. Express* **6**, 116571 (2019).
- Ye, Y. X. et al. Evaluating elastic properties of a body-centered cubic NbHfZrTi high-entropy alloy - A direct comparison between experiments and ab initio calculations. *Intermetallics* **109**, 167–173 (2019).
- Pugh, S. F. Relations between the elastic moduli and the plastic properties of polycrystalline pure metals. *Philos. Mag.* **45**, 823–843 (1954).
- Pettifor, D. G. Theoretical predictions of structure and related properties of intermetallics. *Mater. Sci. Technol.* **8**, 345–349 (1992).
- Kobayashi, T. Advances in turbine materials design and manufacturing. In *Proc. 4th Int. Charles Parsons Turbine Conference*, vol. 4, p. 766 (1997).
- Wee, S. et al. Review on mechanical thermal properties of superalloys and thermal barrier coating used in gas turbines. *Appl. Sci.* **10**, 5476 (2020).
- Sheikh, S. et al. Alloy design for intrinsically ductile refractory high-entropy alloys. *J. Appl. Phys.* **120**, 164902 (2016).
- Li, W. et al. An ambient ductile TiHfVNBa refractory high-entropy alloy: Cold rolling, mechanical properties, lattice distortion, and first-principles prediction. *Mater. Sci. Eng. A* **856**, 144046 (2022).

32. Bai, L. et al. Titanium alloying enhancement of mechanical properties of NbTa-MoW refractory high-entropy alloy: First-principles and experiments perspective. *J. Alloy. Compd.* **857**, 157542 (2021).
33. Jiang, H. et al. Effects of tungsten on microstructure and mechanical properties of CrFeNiV_{0.5}W_x and CrFeNi₂V_{0.5}W_x high-entropy alloys. *J. Mater. Eng. Perform.* **24**, 4594–4600 (2015).
34. Singh, R., Sharma, A., Singh, P., Balasubramanian, G. & Johnson, D. D. Accelerating computational modeling and design of high-entropy alloys. *Nat. Comput. Sci.* **1**, 54–61 (2021).
35. Behera, M., Panigrahi, A., Bönisch, M., Shankar, G. & Mishra, P. K. Structural stability and thermal expansion of TiTaNbMoZr refractory high entropy alloy. *J. Alloy. Compd.* **892**, 162154 (2022).
36. Lin, D. et al. Effects of annealing on the structure and mechanical properties of fecocmri high-entropy alloy fabricated via selective laser melting. *Addit. Manuf.* **32**, 101058 (2020).
37. Riva, S. et al. A novel high-entropy alloy-based composite material. *J. Alloy. Compd.* **730**, 544–551 (2018).
38. Daoud, H., Manzoni, A., Wanderka, N. & Glatzel, U. High-temperature tensile strength of Al₁₀Co₂₅Cr₈Fe₁₅Ni₆Ti₆ compositionally complex alloy (high-entropy alloy). *JOM* **67**, 2271–2277 (2015).
39. Rasmussen, C.E., Williams, C.K.I. *Gaussian processes for machine learning (Adaptive Computation and Machine Learning)*, pp. 8–29. (The MIT Press, Cambridge, MA, USA, 2005).
40. Costabal, F. S., Perdikaris, P., Kuhl, E. & Hurtado, D. E. Multi-fidelity classification using gaussian processes: accelerating the prediction of large-scale computational models. *Comput. Methods Appl. Mech. Eng.* **357**, 112602 (2019).
41. Clyde, M.A. *Model Averaging* 2nd edn, pp. 320–335 Ch. 13 (Wiley-Interscience, Hoboken, NJ, USA, 2003). <https://doi.org/10.1002/9780470317105.ch13>
42. Clyde, M. & George, E. Model uncertainty. *Stat. Sci.* **19**, 81–94 (2004).
43. Draper, D. Assessment and propagation of model uncertainty. *J. R. Stat. Soc. Ser. B* **57**, 45–97 (1995).
44. Hoeting, J., Madigan, D., Raftery, A. & Volinsky, C. Bayesian model averaging: a tutorial. *Stat. Sci.* **14**, 382–417 (1999).
45. Leamer, E. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. John Wiley & Sons, New York, NY. <https://doi.org/10.2307/1057568> (1978).
46. Madigan, D. & Raftery, A. Model selection and accounting for model uncertainty in graphical models using Occam's window. *J. Am. Stat. Assoc.* **89**, 1535–1546 (1994).
47. Mosleh, A. & Apostolakis, G. The assessment of probability distributions from expert opinions with an application to seismic fragility curves. *Risk Anal.* **6**, 447–461 (1986).
48. Reinert, J. & Apostolakis, G. Including model uncertainty in risk-informed decision making. *Ann. Nucl. Energy* **33**, 354–369 (2006).
49. Riley, M. & Grandhi, R. Quantification of modeling uncertainty in aeroelastic analyses. *J. Aircr.* **48**, 866–873 (2011).
50. Zio, E. & Apostolakis, G. Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories. *Reliab. Eng. Syst. Saf.* **54**, 225–241 (1996).
51. Julier, S., Uhlmann, J. A non-divergent estimation algorithm in the presence of unknown correlations. In *Proc. Am. Control Conf.* pp. 2369–2373. <https://doi.org/10.1109/ACC.1997.609105> (1997).
52. Geisser, S. A bayes approach for combining correlated estimates. *J. Am. Stat. Assoc.* **60**, 602–607 (1965).
53. Morris, P. Combining expert judgments: a bayesian approach. *J. Manag. Sci.* **23**, 679–693 (1977).
54. Winkler, R. Combining probability distributions from dependent information sources. *J. Manag. Sci.* **27**, 479–488 (1981).
55. Talapatra, A. et al. Autonomous efficient experiment design for materials discovery with bayesian model averaging. *Phys. Rev. Mater.* **2**, 113803 (2018).
56. Allaire, D., Willcox, K. Fusing information from multifidelity computer models of physical systems. In *15th Int. Conf. Inf. Fusion* pp. 2458–2465 (2012)
57. Thomison, W.D., Allaire, D.L. A model reification approach to fusing information from multifidelity information sources. In *19th AIAA Non-Deterministic Approaches Conf.* p. 1949. <https://doi.org/10.2514/6.2017-1949> (2017).
58. Ghoreishi, S. F., Thomison, W. D. & Allaire, D. Sequential information-theoretic and reification-based approach for querying multi-information sources. *J. Aerosp. Inf. Syst.* **16**, 575–587 (2019).
59. Winkler, R. L. Combining probability distributions from dependent information sources. *J. Manag. Sci.* **27**, 479–488 (1981).
60. Khatamsaz, D., Allaire, D.L. A comparison of reification and cokriging for sequential multi-information source fusion. In *AIAA Scitech 2021 Forum* p. 1477 (2021).
61. Ghoreishi, S. F., Molkeri, A., Arróyave, R., Allaire, D. & Srivastava, A. Efficient use of multiple information sources in material design. *Acta Mater.* **180**, 260–271 (2019).
62. Ghoreishi, S. F., Friedman, S. & Allaire, D. L. Adaptive dimensionality reduction for fast sequential optimization with gaussian processes. *J. Mech. Des.* **141**, 071404 (2019).
63. Marler, R. T. & Arora, J. S. The weighted sum method for multi-objective optimization: New insights. *Struct. Multidiscip. Optim.* **41**, 853–862 (2010).
64. Kim, I. Y. & de Weck, O. L. Adaptive weighted-sum method for bi-objective optimization: Pareto front generation. *Struct. Multidiscip. Optim.* **29**, 149–158 (2005).
65. Das, I. & Dennis, J. E. Normal-boundary intersection: a new method for generating the pareto surface in nonlinear multicriteria optimization problems. *SIAM J. Control Optim.* **8**, 631–657 (1998).
66. Beume, N. S-metric calculation by considering dominated hypervolume as klee's measure problem. *Evol. Comput.* **17**, 477–492 (2009).
67. Bradstreet, L., While, L., Barone, L. A fast many-objective hypervolume algorithm using iterated incremental calculations. In *IEEE Congr. Evol. Comput.* pp. 1–8. <https://doi.org/10.1109/CEC.2010.5586344> (2010).
68. Emmerich, M.T., Deutz, A.H., Klinkenberg, J.W. Hypervolume-based expected improvement: Monotonicity properties and exact computation. In *2011 IEEE Congress of Evol. Comput. (CEC)* pp. 2147–2154. <https://doi.org/10.2514/6.2015-0143> (2011).
69. Fonseca, C.M., Paquete, L., López-Ibáñez, M. An improved dimension-sweep algorithm for the hypervolume indicator. In *2006 IEEE Int. Conf. Evol. Comput.* pp. 1157–1163. <https://doi.org/10.1109/CEC.2006.1688440> (2006).
70. Russo, L. M. & Francisco, A. P. Quick hypervolume. *IEEE Trans. Evol. Comput.* **18**, 481–502 (2013).
71. Yang, Q., Ding, S. Novel algorithm to calculate hypervolume indicator of pareto approximation set. In *Int. Conf. Intell. Comput.*, pp. 235–244. https://doi.org/10.1007/978-3-540-74282-1_2 (2007).
72. Zitzler, E. & Thiele, L. Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. *IEEE Trans. Evol. Comput.* **3**, 257–271 (1999).
73. Khatamsaz, D., Peddareddygar, L., Friedman, S., Allaire, D.L. Efficient multi-information source multiobjective bayesian optimization. In *AIAA Scitech 2020 Forum* p. 2127. <https://doi.org/10.2514/6.2020-2127> (2020).
74. Thermo-Calc Documentation Set. Thermo-Calc Software. <https://thermocalc.com/support/documentation/>(2022).
75. Slack, G. A. The thermal conductivity of nonmetallic crystals. *J. Solid State Phys.* **34**, 1–71 (1979).
76. Morelli, D., Heremans, J. & Slack, G. Estimation of the isotope effect on the lattice thermal conductivity of group iv and group iii-v semiconductors. *Phys. Rev. B* **66**, 195304 (2002).
77. Jones, W., March, N.H. *Theoretical Solid State Physics*, Volume 1: Perfect Lattices in Equilibrium. John Wiley & Sons Ltd, London, UK (1973)
78. Johnson, D. D., Nicholson, D. M., Pinski, F. J., Gyorffy, B. L. & Stocks, G. M. Density-functional theory for random alloys: Total energy within the coherent-potential approximation. *Phys. Rev. Lett.* **56**, 2088–2091 (1986).
79. Singh, P., Smirnov, A. V. & Johnson, D. D. Atomic short-range order and incipient long-range order in high-entropy alloys. *Phys. Rev. B* **91**, 224204 (2015).
80. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
81. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
82. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
83. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
84. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
85. Monkhorst, H. J. & Pack, J. D. Special points for brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
86. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
87. Rao, Y., Baruffi, C., De Luca, A., Leinenbach, C. & Curtin, W. A. Theory-guided design of high-strength, high-melting point, ductile, low-density, single-phase bcc high entropy alloys. *Acta Mater.* **237**, 118132 (2022).
88. Ferrari, A., Lysovorskiy, Y. & Drautz, R. Design of refractory compositionally complex alloys with optimal mechanical properties. *Phys. Rev. Mater.* **5**, 063606 (2021).
89. Vazquez, G. et al. Efficient machine-learning model for fast assessment of elastic properties of high-entropy alloys. *Acta Mater.* **232**, 117924 (2022).
90. Abu-Odeh, A. et al. Efficient exploration of the high entropy alloy composition-phase space. *Acta Mater.* **152**, 41–57 (2018).
91. Kirk, T., Vela, B., Mehalic, S., Youssef, K. & Arróyave, R. Entropy-driven melting point depression in fcc heas. *Scr. Mater.* **208**, 114336 (2022).
92. Vela, B. et al. Evaluating the intrinsic resistance to balling of alloys: A high-throughput physics-informed and data-enabled approach. *Addit. Manuf. Lett.* **3**, 100085 (2022).

93. Rai, A. K., Trpathy, H., Hajra, R. N., Raju, S. & Saroja, S. Thermophysical properties of Ni based super alloy 617. *J. Alloy. Compd.* **698**, 442–450 (2017).
94. Hellström, K., Diaconu, V.-L. & Diószegi, A. Density and thermal expansion coefficients of liquid and austenite phase in lamellar cast iron. *China Foundry* **17**, 127–136 (2020).
95. Singh, P. et al. A systematic first principles study of transport behavior of high-entropy alloys with experimental validation (in preparation).
96. Slack, G. A. Nonmetallic crystals with high thermal conductivity. *J. Phys. Chem. Solids* **34**, 321–335 (1973).

ACKNOWLEDGEMENTS

The authors acknowledge the support from the U.S. Department of Energy (DOE) ARPA-E ULTIMATE Program through Project DE-AR0001427 and DEVCOM-ARL under Contract No. W911NF2220106 (HTMDEC). B.V. acknowledges the support of NSF through Grant No. DGE-1545403. D.K. acknowledges the support of NSF through Grant No. CDSE-2001333. R.A. acknowledges the support from Grants No. NSF-CISE-1835690 and NSF-DMREF-2119103. High-throughput CALPHAD and DFT calculations were carried out partly at the Texas A&M High-Performance Research Computing (HPRC) Facility. ARPA-E supported the applications of theory in this work. In contrast, the theory development (KKR-CPA and SCRAPs by DDJ/PS) at Ames National Laboratory were supported by the U.S. DOE, Office of Science, Basic Energy Sciences, Materials Science and Engineering Department. Ames Laboratory is operated by Iowa State University for the U.S. DOE under contract DE-AC02-07CH11358.

AUTHOR CONTRIBUTIONS

D.A., R.A., and D.J. designed the problem. D.K. implemented the Bayesian Optimization/Classification framework in collaboration with B.V. P.S. performed DFT calculation and analysis. D.K. and B.V. wrote the first version of the manuscript. Finally, all the authors edited and reviewed the final version of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Brent Vela.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023