Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

Published on 09 September 2026
Licensed under CC-BY-NC 4.0

Check for updates

Cite this: *Digital Discovery*, 2026, **5**, 4381

Data-driven, multi-objective QUBO-based optimization for multi-phase alloy design

Thomas Plehn, * Daniel Barragan-Yani, Eric Breitbarth, Guillermo Requena and David Melching

Received 17th March 2026
Accepted 17th July 2026

DOI: 10.1039/d6dd00126b

rsc.li/digitaldiscovery

Here, we present two complementary approaches that advance quadratic unconstrained binary optimization (QUBO) toward practical use in data-driven materials design and other real-valued black-box optimization tasks. First, we introduce a simple yet powerful preprocessing scheme that, when applied to a machine-learned QUBO model, entirely removes system-level equality constraints by construction. This makes cumbersome soft-penalty terms obsolete, simplifies QUBO formulation, and substantially accelerates solution search. Second, we develop a multi-objective optimization strategy inspired by Tchebycheff scalarization that is compatible with non-convex objective landscapes and outperforms existing QUBO-based Pareto front methods. We demonstrate the effectiveness of both approaches by applying them to the design of a multi-phase aluminum alloy using a simplified defect-free model. We then highlight the obtained design-gains in terms of efficiency and solution quality. Together, these methods broaden the applicability of QUBO-based optimization and provide practical tools for data-driven materials discovery and beyond.

1 Introduction

Solving many real-world problems critical for economic and scientific progress, such as discovering new innovative materials, finding optimal logistics strategies, planning the best travel routes, or designing the most efficient duty schedules, often involves tackling complex combinatorial optimization tasks. Quantum optimization is a promising route to deal with the exponential explosion of computational effort required to solve such problems. It does so by leveraging quantum mechanics (specifically quantum tunneling and superposition effects) to explore a vast search space more efficiently than classical methods.[1,2] Quantum optimization essentially means solving a quadratic unconstrained binary optimization (QUBO) problem.[3] Although conceptually alternative approaches do exist, most recent research, focused on finding practical applications for quantum optimization, is built on the QUBO-based methodologies: quantum approximate optimization algorithm (QAOA)[4,5] and quantum annealing (QA).[6] On the one hand, the D-Wave QA technology seems to be a very promising candidate for near to mid-future industrial scale application as it provides more than 5000 physical qubits.[7–9] On the other hand, circuit-based methods like QAOA provide extraordinary flexibility for algorithm engineering, which promises more rapid and continuous improvements. In addition to QAOA, the use of variable inverse time evolution has been repeatedly proposed as a promising means for solving combinatorial problems.[10,11]

In order to meet the mathematical QUBO formulation, considered optimization problems must first be binary encoded. Unfortunately, many realistic systems involve higher-order interactions among variables, which cannot directly be mapped on a quadratic model. Turning such problems into QUBO form usually comes at the cost of restricting binary interactions to second order, or involves the introduction of additional auxiliary variables.[12,13] Thankfully, data-driven, machine-learning-aided QUBO optimization ansatzes can treat realistic use-cases within the QUBO formalism efficiently.[14] Specifically, we highlight the potential of the data-driven FM + QO method for applied problems.[15–21] It was developed to combine the efficiency of either QA, QAOA or any QUBO-based quantum optimization algorithm, with the powerful quadratic regression of factorization machines (FM).[22] The core idea of FM + QO is the following: starting with an initial set of data points for the objective to be optimized, a FM surrogate regression model is trained to predict a target property. Then, the QUBO optimization (QO) method part optimizes the corresponding QUBO. The discovered optimum is checked against a validation source, *e.g.*, by simulation or experiment, and finally added to the data set. In this way, the FM model is iteratively refined in an overarching active learning process. Various studies have shown the great potential of FM + QO with QA as a black-box optimizer in the field of materials research,[15] for example, applied on optimization of optical multilayer systems and metamaterials,[17–19] core–shell nanoparticles[20] or the functional optimization of

Institute for Frontier Materials on Earth and in Space, German Aerospace Center (DLR), Linder Hoehe, Cologne, 51147, Germany. E-mail: Thomas.Plehn@dlr.de

© 2026 The Author(s). Published by the Royal Society of Chemistry

atomistic systems.[21] Most recently, the field-aware FM algorithm has been combined with QA in a high entropy alloy lattice optimization study.[16]

Despite all efforts, however, a practical advantage of quantum-computing-based QUBO solvers against state-of-the-art classical optimizers could not be demonstrated yet, particularly owing to thermal noise induced inaccuracies and other imperfections feeding decoherence within gate-based quantum hardware or QA devices.[2,23] While hardware developments will undoubtedly continue to advance and improve solution quality, it is the co-development of algorithms that must extract the full potential of today's quantum devices.

In this regard, especially the incorporation of problem-specific constraints describes an Achilles heel in practical QUBO formulations, which remain an active area of research.[24–29] Many real-world optimization tasks involve equality constraints, for instance, requiring the composition fractions, f_i, of a material system as a mixture of N_{comp} base components to properly satisfy $\sum_{i}^{N_{\mathrm{comp}}} f_i = 100.0\%$. Since such constraints cannot be naturally embedded into the QUBO formalism, instead they are commonly enforced through soft constraints in terms of quadratic penalty terms. While conceptually straightforward, soft constraints often make the resulting optimization problem significantly harder for QA and QAOA algorithms by narrowing the effective energy gap and increasing the complexity of the optimization landscape.[24,30–32] In addition, they exacerbate current hardware limitations by inflating the required minor embedding, needed to map the QUBO to the hardware architecture graph, and by consuming the limited dynamic range available for implementing the QUBO coefficients as physical spin couplings and local biases.[7,29,33] Consequently, penalty strengths must typically be adjusted through heuristic trial-and-error procedures. Several alternative approaches have recently been proposed, including grand-canonical linear corrections,[28,34] purely linear penalty formulations,[25–27] and constraint-preserving adiabatic state preparation schemes.[24,35] Although these methods can alleviate some of the limitations of quadratic penalties, they generally introduce new challenges, such as difficult parameter tuning susceptible to the optimization landscape[25] or non-trivial state initialization with complex mixer constructions that remain impractical for current quantum hardware.[36,37]

In this paper, we introduce two methods to advance data-driven quantum optimization. We emphasize that all calculations carried out for testing and benchmarking were carried out using simulated annealing (SA). Nevertheless, the developed approaches can be directly translated to a QUBO-based quantum solvers running on QA devices or gate-based quantum computers. The first method, which we call constraint-guided feature mapping (CGFM), is in the spirit of Hen *et al.*[24] It describes a novel pre-processing scheme for avoiding explicit treatment of equality constraints when solving real-valued, data-driven optimization problems with FM + QO. To do this, we exploit the feature of the data-driven FM + QO optimization method of not relying on a fixed known analytical QUBO model expression. Briefly, the method makes constraints unnecessary, because QUBO modeling and QO exclusively takes place within the sub space of feasible solutions. In addition, since the CGFM is part of our pre-processing, it remarkably retains its effectiveness even in the case of using quantum solvers on NISQ hardware. Although real-valued optimization problems pose a significant challenge for current QUBO-based quantum solvers due to their need to discretize continuous variables, we expect that addressing such problems becomes an increasingly important focus as quantum computing progresses beyond the NISQ era.[23]

Our second method, termed data-driven Tchebycheff scalarization (DDTS), addresses another blind spot of current QUBO optimization regarding real-world applications: the treatment of multi-objective Pareto problems. In such problems, conflicting objectives prevent the existence of a single optimum, which requires the determination of a Pareto front of non-dominated solutions.[38] Classical evolutionary algorithms represent the current state of the art,[38–40] but their performance deteriorates for large search spaces and computationally expensive objective functions. In contrast, quantum optimization offers a promising route for tackling the exponential growth of search spaces and can, in principle, find solutions on millisecond timescales.[33]

With the advent of mature quantum hardware, advantage through quantum optimization will not stop at multi-objective problems. Therefore, concepts for quantum Pareto optimization algorithms,[41–44] in particular multi-objective QUBO ansatzes,[41,42,44,45] are of recent interest. The "weighted sum" approach[46,47] provides a linear scalarization scheme that preserves QUBO structure of the optimization problem. It represents the current state of the art for multi-objective QUBO formulations. However, its applicability is fundamentally limited to convex optimization and it generally fails to recover non-convex regions of the Pareto front. Our DDTS ansatz is inspired by the Tchebycheff scalarization technique for multi-objective optimization[48] inheriting its ability to efficiently handle non-convex Pareto optimization. Similar to the CGFM framework, DDTS operates through a data preprocessing step that transforms the multi-objective dataset into a scalarized single-objective representation. Crucially, the scalarization is performed directly on the data level rather than on the analytical QUBO formulation itself. Consequently, there is no requirement to preserve the mathematical QUBO form character during scalarization. This circumvents the limitations of conventional QUBO scalarization schemes and enables the QUBO-based treatment of non-convex optimization problems. To the best of our knowledge, neither the FM + QO method specifically nor the Tchebycheff scalarization in general have previously been employed for data-driven QUBO-based multi-objective optimization. Our demonstration of a data-driven QUBO-based Pareto optimization is therefore the first of its kind.

To demonstrate the strength and potential of our CGFM and DDTS ansatzes, we introduce an optimization problem of a hypothetical Al-alloy design by varying the content of five microstructural phases that can be found in cast Al–Si alloys at

© 2026 The Author(s). Published by the Royal Society of Chemistry

the μm level: Al-matrix, eutectic Si, Mg_2Si, Al_3Ni and Al_2Cu. Such problem is technologically relevant as aluminum alloys exhibit low density, high mechanical strength, and excellent corrosion resistance, making them indispensable for a wide range of industrial applications.[49,50] Furthermore, such optimization task is a suitable benchmark to show that both CGFM and DDTS ansatzes are valuable building blocks for establishing quantum algorithms in the field of machine learning assisted materials discovery. We remark that our techniques could also be used to the fast growing field of high entropy alloy design,[51–55] where also quantum-assisted machine learning frameworks based on QUBO have been studied recently.[16,56]

For the demonstration studies presented here, the feature space is defined solely by the fractions of the precipitate phases. While this simplified representation is adequate for the present purpose, we acknowledge that machine-learning models for materials design can benefit from incorporating additional microstructure-related features.[51] As optimization targets we choose five thermal and mechanical properties: thermal conductivity (κ), Young's modulus (E), density (ρ), linear coefficient of thermal expansion (α) and the solidification interval (ΔT). To limit computational cost and reduce model complexity, we employ rule-of-mixtures estimates to obtain quantitative, rapid predictions of κ, E, ρ, and α. For ΔT, we introduce a coarse approximation derived from the binary Al–Si phase diagram. These five simplified property models were then used to generate synthetic data of randomly sampled alloy designs and their associated properties, providing initial input for our data-driven FM + QO framework. Naturally, the assumptions underlying our simplified models limit the predictive scope of the results. Nevertheless, the simplified models introduced here provide a suitable basis for benchmarking the proposed methods.

Our new methods stand out particularly in the optimization of material microstructures for three reasons. First, such optimizations have in common the requirement for a systematic constraint deciding about design feasibility. Second, they are typically multi-objective because they can show counteracting target properties. And third, in this field non-convex optimization problems are ubiquitous. In fact, thermal and mechanical material properties are ultimately based on energetic landscapes, which can be complex structured with local minima, energetic barriers and extended plateaus resulting in non-convex dependencies of system variables.

2 Results and discussion

2.1 Methodological developments

Our multi-phase alloy model is schematically illustrated in Fig. 1. In a nutshell, it is prepared by first introducing continuous variable fractions, f_k with $k = 1, \ldots, N_{\mathrm{ph}}$, for the $N_{\mathrm{ph}} = 4$ secondary phases: eutectic Si, Mg_2Si, Al_3Ni and Al_2Cu. In our model we define a fixed Al matrix content of 0.8, *i.e.*, $f_k \in [0, 0.2]$. For convenience, we continue with normalized phase fractions, which are finally one-hot encoded to binary variables. A discretization of the continuous variables is vital for this step. We elaborate on these processes in Sec. 3.2.

2.1.1 Constraint-guided feature mapping (CGFM).

Let $\bar{f}_k$ denote the normalized continuous variable phase fractions for the Al-alloy with $\bar{f}_k \in [0, 1]$ and $k = 1, \ldots, N_{\mathrm{ph}}$. Obviously, the fractions must sum to 100%, *i.e.*,

$$\sum_{k=1}^{N_{\mathrm{ph}}} \bar{f}_k = 1, \tag{1}$$

Any deviation from that would render an infeasible alloy composition. In order to incorporate this requirement in a QUBO optimization, one typically introduces soft equality constraint terms similar to

$$C(\vec{x}) = \left(\sum_{k=1}^{N_{\mathrm{ph}}} \sum_{i=1}^{N_{\mathrm{bits}}} \alpha_i x_i^{(k)} - 1 \right)^2, \tag{2}$$

already based on binary variables (or bits), $x_i^{(k)} \in \{0, 1\}$, used to encode the $\bar{f}_k$ in eqn (1) and the corresponding encoding-specific weights α_i (see eqn (17) in Sec. 3.2.2). The first sum runs through all N_{ph} phases and the second one sums over all N_{bits} binary variables used to encode one fraction variable.

In general, additional constraint terms in QUBO models lead to a more rugged structure of the potential energy surface of the system which, provided with many local minima, increases the probability that the optimization gets stuck in a local minimum instead of finding the global one.[23] Furthermore, soft constraints increase configuration effort, as they have to be added with a strength factor, λ, which must be precisely tuned against the original QUBO model, and apparently this is exacerbated when several constraints have to be taken into account. These two disadvantages generally occur with quantum QUBO solvers and classical heuristics, but equality constraints of type eqn (2) are especially critical and detrimental to the solution quality of quantum algorithms (*e.g.*, QA[28]).

In fact, the individual coupling terms in eqn (2) range over many orders of magnitude with realistic model scaling. When applying the one-hot type of encoding with $\alpha_i = i/N_{\mathrm{bits}}$, this can be observed straightly from

$$\max\{\alpha_i\}/\min\{\alpha_i\} = N_{\mathrm{bits}}. \tag{3}$$

In fact $N_{\mathrm{bits}} = 101$ are required with one-hot encoding in order to model our phase fractions, $\bar{f}_k$, with less than 0.01 resolution. According to the ratio, eqn (3), the coupling terms in eqn (2) would range over 4 orders of magnitude after squaring.

As described in Sec. 1, such a wide range of coupling terms is technically problematic regarding the analog nature of current QA devices. Because the strength factor of the constraint, λ, must be sufficiently large to guarantee feasibility, these terms often dominate the original QUBO contributions. When ranging over several orders of magnitude, the remaining QUBO coefficients become very small and increasingly prone to hardware noise. Moreover, the quadratic structure of eqn (2) couples all binary variables across all fractions. The resulting fully connected constraint graph is the worst case scenario for QA hardware, which makes ultimately long qubit chains for minor embedding of the QUBO unavoidable. These break more often

© 2026 The Author(s). Published by the Royal Society of Chemistry

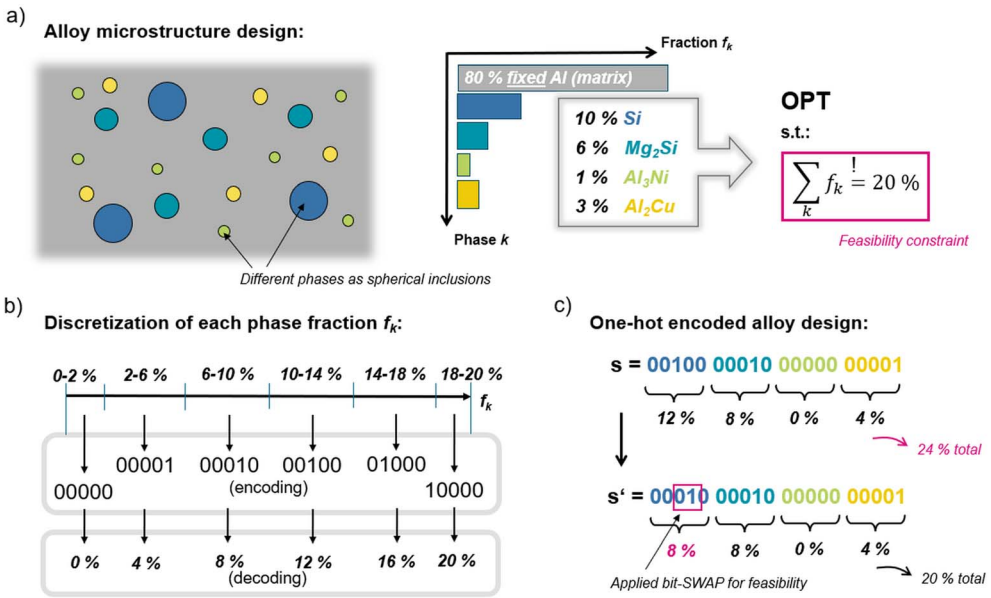


Fig. 1 Illustration of the multi-phase aluminum alloy model and its encoding. (a) Scheme of the simplified microstructure and phase morphology assumed for straight forward prediction of different thermal and mechanical material properties with the rule-of-mixtures: the Al-matrix dominate has the largest fraction with additional minor phases embedded in it. These phases are treated as isotropic and homogeneously distributed spherical particles (Si, Mg_2Si, Al_3Ni and Al_2Cu). The matrix fraction is fixed to 80%, while the four minor fractions, $f_k \in [0\%, 20\%]$, are optimization targets under the given constraint of feasible alloy designs. (b) Discretization of the continuous f_k to prepare the binary encoding (top). Each increment is encoded uniquely using the one-hot encoding (middle). Each one-hot code is uniquely decoded into continuous f_k again (bottom). Due to the finite encoding fidelity, the decoded fractions have limited accuracy. (c) Translation from fractions to the alloy design encoding with the microstructure configuration in figure a. As indicated, the limited accuracy results here in an infeasible solution once decoded back to continuous f_k. Those special cases are identified and treated by a curing bit-SWAP applied to a randomly chosen f_k. (Please note, this correction is only required on the randomly generated alloy designs in the initial data set. The optimization operates in the encoded search space).

and degrade solution quality despite post-processing. Gate-model platforms face a similar challenge. There, compilers insert SWAP operations to mediate interactions not supported by the hardware topology, increasing circuit depth and susceptibility to noise.

Our constraint-guided feature mapping (CGFM) ansatz opens the possibility to avoid the constraints in eqn (2) altogether when a QUBO model is learned from data, *e.g.*, by facilitating a factorization machine (FM). The ansatz is schematically illustrated in Fig. 2. The core idea is that differentiable constraints which satisfy regularity conditions define smooth submanifolds, which can be locally described using a reduced set of free parameters according to the implicit function theorem. The constraints can thus be explicitly incorporated into the data set through prior feature mapping. Feature mappings are a well-known pre-processing instrument from machine learning. Our approach is inspired by the kernel trick in support vector machines, where mapping the data into a higher-dimensional feature space enables more effective classification. This idea can be understood as a coordinate transformation of the system variables. However, we transform our original features, the phase fractions $\bar{f}_k$, in order to finally dismiss one dimension. The transformed coordinates are then binary encoded (analogously to the current procedure in Fig. 2a) and utilized for the QUBO training. By this dimensionality reduction, the surrogate QUBO model and the optimization are restricted to the subspace of feasible solutions making the system-related constraints obsolete. In general, this should simplify the search for global optima. Furthermore, due to the reduced dimensionality of the search space, its encoding requires less binary variables, which should further accelerate the optimization process.

For a concrete example, we focus on the constraint in eqn (1). First, we realize that eqn (1) defines a $(N_{\mathrm{elem}} - 1)$-dimensional submanifold in the N_{elem}-dimensional search space. Using the fact that the fractions $\bar{f}_k$ are non-negative, we substitute $\bar{f}_k = {a_k}^2$ to transform eqn (1) to the relation of an N_{elem}-dimensional sphere with radius 1, *i.e.*, $\sum_k {a_k}^2 = 1$. The CGFM ansatz now introduces a mapping which transforms the optimization problem from the original fraction variables $\bar{f}_k$ to angular coordinates ω_k parametrizing the sphere. The spherical coordinates related to this specific equation are well-known. Notably, the CGFM procedure generalizes straightly to satisfy linear equality constraints of type $\sum_k c_k {a_k}^2 = C$ with $c_k, C \in \mathbb{R}_{>0}$, where we refer to the SI S1.2 for further computational details. Finally, by expressing the problem in the new coordinates ω_k, the systemic constraint is satisfied by construction, ensuring that all downstream solutions remain feasible and eliminating the need to include additional constraint terms in the QUBO model. Consequently, the complexity of handling the

© 2026 The Author(s). Published by the Royal Society of Chemistry

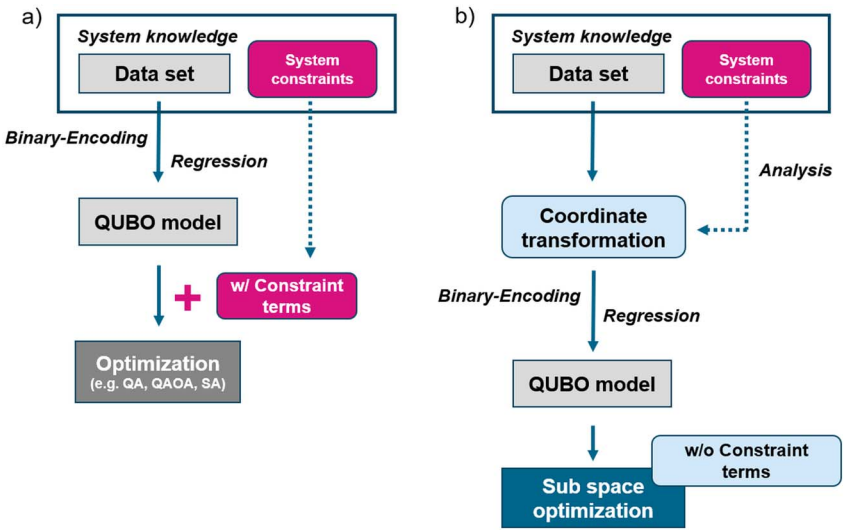


Fig. 2 Comparison of the CGFM ansatz against current data-driven QUBO modeling. (a) Current state-of-the-art: a binary quadratic regression model (*e.g.*, FM) is trained based on a data set and introducing a binary encoding for continuous system variables. Then, the regression model can be translated directly into a QUBO problem. The constraint terms are finally formulated according to the given feasibility conditions (*e.g.*, eqn (2)), configured with suitable scaling and added to the QUBO problem. (b) CGFM ansatz: by analysing the constraints a coordinate transformation is established which maps the original variable space (*i.e.*, data set features) to the sub space of feasible system states. Restricted to this sub space the QUBO modeling and subsequent optimization can be executed without the constraints.

constraints during the optimization is now shifted to a potentially more complicated process of training the FM surrogate model. We discuss this further in Sec. 2.2.1.

2.1.2 Data-driven Tchebycheff scalarization (DDTS). Concerning QO, the state-of-the-art multi-objective approach to search for Pareto optimal solutions is to implement the weighted-sum scalarization method (see eqn (20) in Sec. 3.3). The Pareto front, *i.e.*, the entire set of non-dominated solutions,[38,39] would then result from enough optimal solutions of the scalarized objective, $g_{\vec{w}}(\vec{x})$ in eqn (20), with different weight vectors, $\vec{w}$, expressing the preference of the considered objectives. For non-convex Pareto fronts, however, the weighted-sum method is unable to resolve the entire Pareto front.[46–48] The same challenge occurs when scalarizations of several objective-specific QUBO models produce non-convex Pareto fronts, thereby limiting the effectiveness of traditional weighted-sum approaches.

Although well-established alternative scalarization approaches exist for non-convex, continuous multi-objective optimization,[48] those usually lead to QUBO-incompatible expressions or numerical procedures. Recently, an alternative scalarization ansatz specialized for QUBO optimization has been proposed.[47] The method is based on the principles of the penalty-based boundary intersection method and works on non-convex Pareto fronts. However, the QUBO expression to be solved contains product terms of the involved objective functions. So, to finally stay in QUBO format, the original objective functions are restricted to linear dependence on the binary variables, which is an apparent drawback regarding many relevant use cases.

In our data-driven Tchebycheff scalarization (DDTS) framework, in the same spirit as for CGFM, we implement the scalarization step to the data set in terms of a pre-processing and, therefore, before the mathematical QUBO form of the system model restricts further algorithm engineering. Consequently, the DDTS ansatz is not restricted to linear descriptions of our system. Furthermore, since it originates from the Tchebycheff scalarization technique, the approach works well for non-convex Pareto fronts.

The (weighted) Tchebycheff scalarization[48] is well-known for continuous variable multi-objective problems. In addition to preference weights $\vec{w}\in\mathbb{R}^m$, it involves a reference point in the objective space, $\vec{u}\in\mathbb{R}^m$. This so-called utopian point, needs to fulfill $u_p < \min_{x\in\mathbb{A}} f^{(p)}(\vec{x})$, where $f^{(p)}(\vec{x})$ denotes the p-th objective function and $\mathbb{A}$ the (continuous variable) search space. The multi-objective Tchebycheff problem is then of the form

$$\min_{x\in\mathbb{A}}\left(\left[\max_{p=1,\ldots,m} w_p\left(f^{(p)}(\vec{x})-u_p\right)\right]+\gamma\sum_{p=1}^{m}\left(f^{(p)}(\vec{x})-u_p\right)\right), \quad (4)$$

attributed with different weights, $\vec{w}$ (with $0 < w_p < 1$ and $\sum_p w_p = 1$). The entire Pareto front can be found for convex and non-convex problems. The second term in eqn (4) quantifies the Manhattan distance of point $f^{(p)}(\vec{x})$ from the utopian in objective space. It regularizes the original Tchebycheff formulation to obtain only strong Pareto optimal solutions.[48,57]

Our multi-objective DDTS is motivated by the problem form of eqn (4). Fig. 3 illustrates the workflow. Since eqn (4) is clearly not a QUBO problem, this scalarization technique cannot be utilized on our problems. However, we can circumvent this issue thanks to the data-driven ansatz and introduce the scalarization directly on the data, *i.e.*, the multi-objective data set is translated into a data set with only a single artificial objective (see upper part in Fig. 3). After that, we train a single QUBO

© 2026 The Author(s). Published by the Royal Society of Chemistry

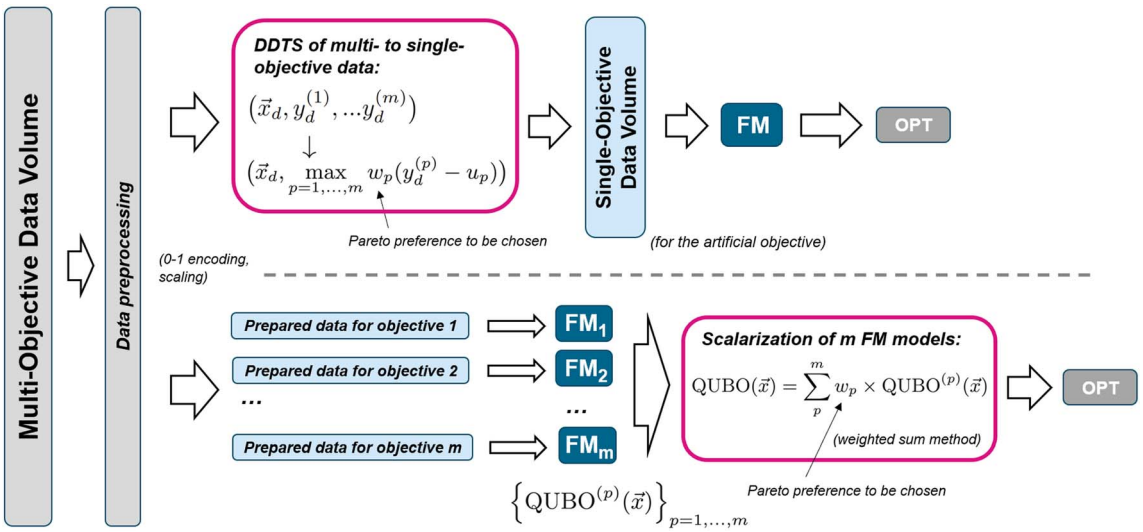


Fig. 3 Schematic processing of our data-driven Tchebycheff scalarization (DDTS) method (upper part). The sequences are part of the FM + QO method and repeated for each iteration of the active learning. In the lower part, the corresponding workflow is given based on the current state-of-the-art weighted sum method for QUBO scalarization.

model problem on the data set with the artificial objective, instead of applying the weighted-sum approach on individual QUBOs trained each for a different objective function (see lower part in Fig. 3).

Let $\vec{x}_d$ be the search space vector of the d-th data point, and $y_d^{(p)}$ the corresponding value of objective p, the DDTS ansatz for pre-processing of the data set follows as replacement according to

$$(\vec{x}_d, y_d^{(1)}, \ldots, y_d^{(m)}) \rightarrow (\vec{x}_d, \hat{y}_d). \tag{5}$$

By this procedure the data set becomes single-objective where $\hat{y}_d$ is defined as

$$\hat{y}_d = \max_{p=1,\ldots,m} w_p \left(y_d^{(p)} - u_p\right). \tag{6}$$

The DDTS is repeated within each active learning iteration. The numerical procedure works best after normalizing the range of objective values $y_d^{(p)}$. The utopian values are either known or can be chosen automatically based on knowledge from the data set. More details on the implementation are given in the Supplementary Information S1.3. Finally, the FM is trained on the data set $(\vec{x}_d, \hat{y}_d)$. In extreme cases of $w_p = 1$ and $w_{q\neq p} = 0$, the artificial objective value $\hat{y}_d$ resembles $y_d^{(p)}$ shifted by u_p, *i.e.*, the QUBO model is then trained to predict the objective values $y^{(p)}$ and solving the multi-objective QUBO resembles the single-objective optimization. In all other cases, the artificial objective of point d, $\hat{y}_d$, takes the value of the objective that is most distanced from its corresponding utopian point. This decision is made independently for each data point and the concrete weighting, $\vec{w}$, balances this preference.

Overall, the closer a data point $(\vec{x}_d, y_d^{(1)}, \ldots, y_d^{(m)})$ lies to the utopian value along objective p, the higher the weight w_p must be chosen in order to make that data point part of the FM training regarding the prediction of $f^{(p)}$. The number of points in the training data set which are attributed to objective $y^{(p)}$ decide how much the trained QUBO model is characterized by objective $f^{(p)}$. This effect allows DDTS to work as a multi-objective add-on for the FM + QO method. If one wishes to target a specific point on the Pareto front, the weights $\vec{w}$ can be adjusted, respectively. However, if the entire Pareto front is of interest, the preference weights can be chosen randomly for each iteration.

2.2 Simulation results

Results presented here are carried out using SA, but the developed approaches can be directly translated to a quantum algorithm.

2.2.1 Single-objective optimization with CGFM ansatz. We investigate the single-objective variant of the FM + QO active learning approach and search for multi-phase Al-alloys with

1. Maximum thermal conductivity, κ,
2. Maximum Young's modulus, E,
3. Minimal density, ρ,
4. Matching thermal expansion coefficient with a target value, $\delta\alpha$, and
5. Minimal solidification interval, ΔT.Regarding $\delta\alpha$ we defined the target value $\alpha_0 = 20$ ppm K^{-1}, which lies rather central in the range of possible values (see details in Sec. 3.1.2), and to remind, the normalized fractions, $\bar{f}_k \in [0, 1]$ with $k = 1, \ldots, N_{\mathrm{ph}}$, of the $N_{\mathrm{ph}} = 4$ secondary phases Si, Mg_2Si, Al_3Ni and Al_2Cu are our model system variables.

The impact of our ansatz, referred to as w/CGFM, is studied by comparing its overall optimization progress against the standard approach, referred to as w/o CGFM, which uses constraint terms as in eqn (2). The analytical single-objective optimal solutions of the property models (see Sec. 3.1.1) can be found in the SI S8.

© 2026 The Author(s). Published by the Royal Society of Chemistry

The five single-objective optimization cases are set up similarly. In the approach w/o CGFM ansatz, each of the four $\bar{f}_k$ is one-hot encoded using 50 binary variables leading to 200 binaries in total. Ultimately, this results in a 0.4 vol% resolution of the alloy composition. Similarly, the three constraint-guided feature variables, ω_k, are one-hot encoded with each 50 binary variables following the approach w/CGFM, which in sum consumes only 150 binary variables. However, the ultimate resolution of the alloy compositions can be very different. In fact, the mapping to transformed variables, ω_k, allows for diverse discretizations. Its analytical expression and reverse form can be read in the SI S1.2, eqn (S8) and (S9). Crucially, this is always in accordance with the constraint, eqn (2). Here, we conveniently stay with equal-sized increments of ω_k and observe that the straight back mapping to $\bar{f}_k$ introduces possible alloy fractions much smaller than 0.4 vol%. The Supplementary Information S2 elaborates on this intrinsic effect. Furthermore, the CGFM transformations, eqn (S8) and (S9), provide full flexibility of how the four $\bar{f}_k$ are assigned to the Cartesian coordinate axes. Here, we introduce a control parameter, $0 \leq p_{\mathrm{dc}} \leq 1$, governing the probability of selecting between two specific assignment strategies in each iteration. With probability p_{dc} the assignment of $\bar{f}_k$ is chosen depending on their correlation with the target property in descending order, particularly we employ the distance correlation, and with probability $1 - p_{\mathrm{dc}}$ the assignment is fully random. Given the large bit numbers, we substitute a classical SA algorithm for the optimizer part in the FM + QO, *i.e.*, here QO = SA. All initial data sets consist of 100 randomly selected Al-alloy designs. The algorithms w/and w/o CGFM are started from the same initial data sets. Full details on data generation and algorithm settings can be found in the SI S1.4.

The comparison of the optimization processes is shown in Fig. 4, where the accumulated best objective values, including all previous iterations, are pointed out. The statistics results from starting the FM + QO procedure w/and w/o the CGFM ansatz from 20 different initial data sets. The latter are the same for all optimization scenarios in Fig. 4. Table 1 shows the best designs of the 20 runs obtained w/and w/o CGFM (taken from Fig. 4), respectively. Concerning κ (Fig. 4a, to be maximized), E (Fig. 4b, to be maximized) and ρ (Fig. 4c, to be minimized), the optimization clearly performs better for w/CGFM, in particular, implementing random CGFM coordinate assignments, *i.e.*, $p_{\mathrm{dc}} > 0$, consistently supports the optimization progress, which is further discussed below. Contrarily, all runs w/o CGFM seem to converge very slowly and get stuck too early. Nevertheless, since even after 600 iterations the different statistical runs show rather wide spreading, we assume not a single local trap can be responsible for that. The different run repetitions w/CGFM and $p_{\mathrm{dc}} > 0$ show common convergence towards the known global optima ($\kappa_{\mathrm{opt}} = 154.5$ W mK; $E_{\mathrm{opt}} = 83.7$ GPa; $\rho_{\mathrm{opt}} = 2.56$ g cm^{-3}, see also SI S8).

Concerning the optimization of $\delta\alpha$, no pronounced performance differences are observed between w/and w/o CGFM

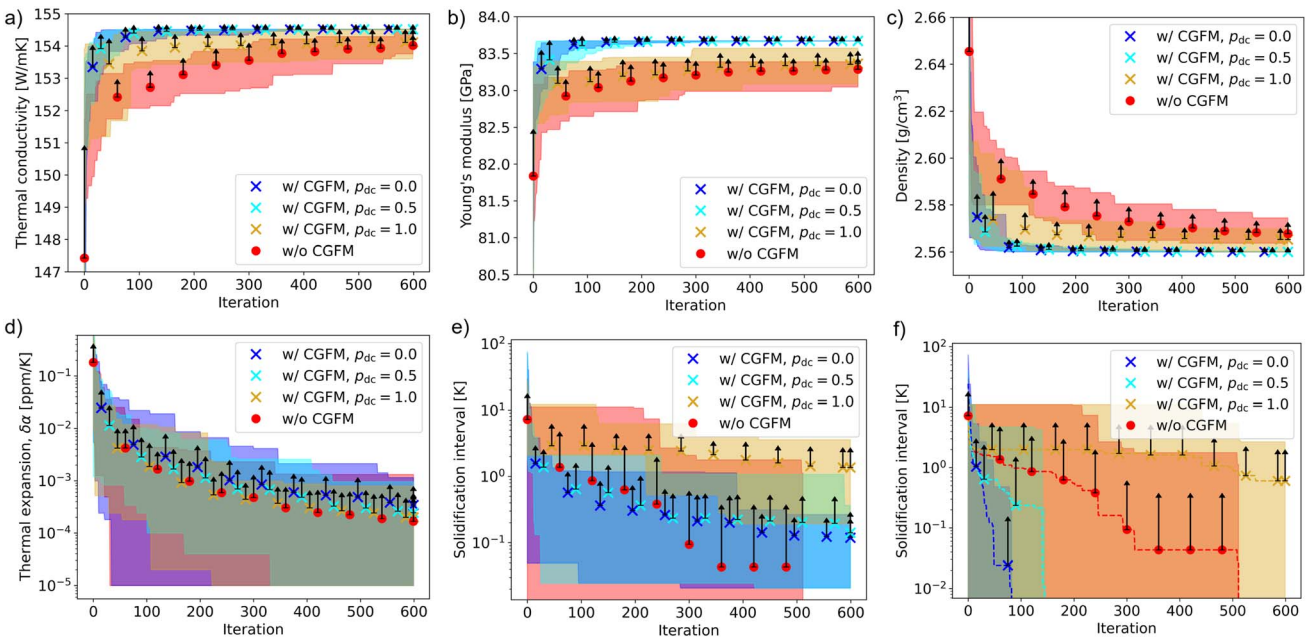


Fig. 4 Performance comparison of the FM + QO procedure w/CGFM (crosses) and w/o CGFM (red dots) over 600 active learning iterations implemented for single-objective optimization. Simulations w/CGFM are conducted applying three control parameter settings $p_{\mathrm{dc}} = 0.0$, 0.5, and 1.0, respectively. (a) Maximization of thermal conductivity; (b) maximization of Young's modulus, (c) minimization of density; (d) minimization of linear coefficient of thermal expansion (absolute difference to a defined target value), $|\alpha - \alpha_0|$ with $\alpha_0 = 20$ ppm K^{-1}); (e) minimization of solidification interval; (f) same as panel e, but a non-uniform variable discretization is applied to the ansatz w/CGFM restoring an equal-spaced grid in fraction variable space similar to w/o CGFM. In all plots the central markers give the mean values taken from 20 statistical replicates each started from randomized initial data. Arrows and shadings indicate the corresponding standard deviations and observed min-to-max ranges, respectively.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Table 1 Optimal single-objective Al-alloy designs discovered after 600 active learning iterations from the 20 statistical runs comparing FM + QO w/o and w/CGFM ansatz. Top to bottom: thermal conductivity (Fig. 4a), Young's modulus (Fig. 4b), density (Fig. 4c), linear coefficient of thermal expansion, *i.e.*, $\delta\alpha = |\alpha - \alpha_0|$ (Fig. 4d), and solidification interval (Fig. 4f). The alloy designs are described in terms of the phase fractions of Si, Mg_2Si, Al_3Ni and Al_2Cu, $\bar{f}_k \in [0, 1]$, where $\bar{f}_k = 1$ is here equivalent to the maximum phase contribution of 20% (reminding fixed 80% Al matrix). The corresponding objective values from the statistical ensembles are each given in the last column. Mean values and standard deviations are given in brackets. As a special case, the ΔT optimizations employ an adapted discretization that restores uniformly spaced increments in each $\bar{f}_k$

Discovered Al-alloy designs with maximal κ

Ansatz	Si	Mg_2Si	Al_3Ni	Al_2Cu	κ [W mK^{-1}]
w/CGFM	0.0	0.0	0.0	1.0	154.52 ($154.52 \pm 1.5 \times 10^{-3}$)
w/o CGFM	0.0	0.0	0.02	0.98	154.45 ($154.02 \pm 1.7 \times 10^{-1}$)

Discovered Al-alloy designs with maximal E

Ansatz	Si	Mg_2Si	Al_3Ni	Al_2Cu	E [GPa]
w/CGFM	1.0	0.0	0.0	0.0	83.67 ($83.67 \pm 1.1 \times 10^{-3}$)
w/o CGFM	0.9	0.0	0.1	0.0	83.39 ($83.28 \pm 9.0 \times 10^{-2}$)

Discovered Al-alloy designs with minimum ρ

Ansatz	Si	Mg_2Si	Al_3Ni	Al_2Cu	ρ [g cm^{-3}]
w/CGFM	0.0	1.0	0.0	0.0	2.560 ($2.560 \pm 3.0 \times 10^{-5}$)
w/o CGFM	0.06	0.94	0.0	0.0	2.564 ($2.568 \pm 2.9 \times 10^{-3}$)

Discovered Al-alloy designs with $\delta\alpha$ closest to target

Ansatz	Si	Mg_2Si	Al_3Ni	Al_2Cu	$\delta\alpha$ [ppm K^{-1}]
w/CGFM	0.376	0.378	0.049	0.197	1.0×10^{-5} ($2.0 \times 10^{-4} \pm 1.8 \times 10^{-4}$)
w/o CGFM	0.32	0.24	0.34	0.1	1.0×10^{-5} ($1.7 \times 10^{-4} \pm 2.7 \times 10^{-4}$)

Discovered Al-alloy designs with minimum ΔT

Ansatz	Si	Mg_2Si	Al_3Ni	Al_2Cu	ΔT [K]
w/CGFM	0.640	0.202	0.143	0.016	0.0 (0.0 ± 0.0)
w/o CGFM	0.64	0.04	0.02	0.3	0.0 (0.0 ± 0.0)

(Fig. 4d, to be minimized). Especially the optimizations w/ CGFM ($p_{dc} = 0$) and w/o CGFM converge very similarly. The underlying reason lies in the specific mechanism of FM training based on the encoded CGFM variables, ω_k, and in how this encoding shapes and simplifies the solution search.

To substantiate this interpretation, we first analyze the optimization scenarios and their global optima in more detail. Based on these findings, we subsequently draw conclusions regarding the strengths and limitations of the CGFM ansatz and finally discuss the ΔT optimization results. In general, the more complex the interplay among the four phase fractions becomes, the more approximate the training of the second-order FM surrogate model tends to be. As already indicated in Sec. 2.1.1, the CGFM ansatz simplifies the constrained optimization, but this simplification comes at the cost of presenting a more challenging learning task to the FM.

This trade-off can be understood by examining the feature mapping used in the CGFM formulation. A variation of an angle, ω_k, necessarily induces changes in at least one of the phase fractions $\bar{f}_k$. Because the mapping must remain within the feasible subspace defined by $\sum_k \bar{f}_k = 1$, alterations of ω_k typically modify multiple fractions simultaneously (see SI S1.2, eqn (S8) and (S9)). While this interdependence ensures feasibility by construction, it also amplifies the limitations of a second-order FM model. However, this effect is not critical when the optimum corresponds to a low-entropy solution, *i.e.*, an alloy design dominated by a few principal constituents, especially when the variable assignments within the mapping are reshuffled during each active-learning iteration or chosen according to their relevance (see below). In such cases, the benefits of unconstrained optimization can fully outweigh the approximation errors introduced by the mapping. In extreme scenarios, the CGFM ansatz may even be advantageous for the learning process itself.

This is exactly demonstrated by the κ, E and ρ optimizations in Fig. 4a–c, respectively. The optimal alloy designs are determined as Al + Al_2Cu, Al + Si, and Al + Mg_2Si, respectively, where we refer again to Table 1 and the SI S8 for details. These designs are extreme examples of minimum disorder, because in all three cases only a single secondary phase contributes with $\bar{f}_k = 1$ (*i.e.*, $f_k = 20\%$ plus 80% fixed Al matrix). But because alloy designs with $\bar{f}_k > 0.6$ are highly unlikely in our initial data sets, large individual phase fractions are unknown for the FM learning model and therefore it struggles to discover better κ, E or ρ designs right from the beginning of the active learning. Instead, the optimization progress can only successively move to such designs by means of iterative data set refinement. Thereby, we observe much weaker performance of the approach w/o CGFM ansatz. We identified two reasons for this behaviour. First, apparently the involved constraint, eqn (2), complicates the search for feasible low entropy designs. We demonstrate this by repeating the optimization runs w/o CGFM while spending substantially more search effort, *i.e.*, by increasing both the number of SA runs and sweeps (see Fig. S3 of the SI S3). We find the expected improvement of the optimization performance, but the gain saturates and cannot be scaled higher by applying more searching power to meet the performance of the approach w/CGFM ansatz. The second reason for the distinct performance of the CGFM ansatz is its specific training data representation. Namely, w/o CGFM ansatz all N_{ph} variables are relevant for learning any alloy design, even with only a single secondary phase. In contrast, for the same learning task, in the CGFM ansatz only a single variable might be relevant. Consequently, also the training in general can improve from the intrinsic constraint treatment of the CGFM ansatz, providing better exploration capabilities within less active learning iterations. In this context, assigning the four $\bar{f}_k$ variables to the CGFM coordinate axes according to their correlation with the target property, in descending order, substantially simplifies

© 2026 The Author(s). Published by the Royal Society of Chemistry

the learning of κ, E, and ρ. At the same time, however, the FM encounters difficulties similar to the ansatz w/o CGFM when exploring regions of the search space beyond the domain covered by the available dataset. Consequently, all optimizations w/CGFM exhibit rapid initial progress during the first 50 iterations, as shown in Fig. 4a–c. However, only the optimization runs w/CGFM that additionally include randomized assignments through $p_{\mathrm{dc}} > 0$ continue to converge efficiently towards the global optimum. These random CGFM coordinate assignments introduce an exploratory component that enhances the active-learning process and prevents premature convergence at the boundaries of the search space covered by the dataset. A comprehensive analysis of the observed dependence of the performance on p_{dc} for increasing system sizes is the subject of ongoing work.

For $\delta\alpha$, the optimal alloy design and surrounding candidate solutions found by FM + QO exhibit much greater diversity (see Table 1). We further analysed the solution density around the optimal alloy design by brute force search directly based on the underlying property models (with details given in the SI S5). In contrast to κ, E or ρ, solution density close to the ground state of $\delta\alpha$ is substantially higher. Furthermore, alloy designs close to the optimum for $\delta\alpha$ are characterized by relatively homogeneous compositions of all four phases, which primarily arises from the chosen target value $\alpha_0 = 20$ ppm K^{-1}, which can be approached by a wide range of alloy designs. The relevant ranges of the individual fractions are already well represented by the available datasets. In such scenarios, exploring previously unseen fraction values becomes less important than exploitative learning capabilities of the FM, explaining that FM + QO w/CGFM and $p_{\mathrm{cd}} = 1$ performs best in Fig. 4d. Assigning the fractions according to their correlation enables the FM to learn the most relevant interactions most efficiently, and interrelations of minor relevance are captured less efficiently. Overall, nevertheless, the FM training for $\delta\alpha$ is harder when encoding the feature variables of the CGFM ansatz, and its weakness against the direct encoding of the fraction hides its strengths.

The optimization of ΔT (Fig. 4e, to be minimized) represents a special case of the scenario type previously discussed for κ, E and ρ. The global optimum of ΔT is highly degenerated, since all designs with the eutectic Si phase fraction of 12.8%, *i.e.*, $\bar{f}_{\mathrm{Si}} = 0.64$, yield $\Delta T = 0$ K, and it is identified by all 20 conducted runs w/o CGFM within 500 iterations. Similar to the optimization of κ, E and ρ, only a single fraction is relevant, and its value to be found lies outside the range of the search space densely covered by the dataset. This constitutes a favourable scenario for the algorithm w/CGFM applying $p_{\mathrm{dc}} = 0.0$. In fact, the unexpected weak performance of all three runs w/CGFM observed in Fig. 4e originates from the different discretizations of the search space employed for w/CGFM and w/o CGFM. Specifically, the value of $\bar{f}_{\mathrm{Si}} = 0.64$ lies on a grid point of the encoding w/o CGFM, allowing the global optimum to be represented exactly. The discretization grid introduced by the CGFM differs to that and although it changes depending on the particular CGFM coordinate assignment, the optimal value $\bar{f}_{\mathrm{Si}} = 0.64$ cannot be represented accurately by any assignment with the uniform discretization scheme. We test this reasoning by attributing the CGFM ansatz with an adapted discretization scheme that restores equal-spaced grid sizes in fraction variable space, analogous to the ansatz without CGFM. This adaptation requires a non-uniform discretization of the CGFM variables, which can be implemented straightforwardly using the arccosine function. Importantly, the ability to freely modify the discretization scheme in this manner is an exclusive feature of the CGFM ansatz, since it avoids the constraint, eqn (2), coupling multiple variables. Fig. 4f shows the performance of the corresponding optimization w/and w/o CGFM. With the adapted discretization, the CGFM ansatz is then able to recover the global optimum similarly to the ansatz w/o CGFM. And in agreement with the above discussion on κ, E, and ρ, the ansatz w/CGFM and $p_{\mathrm{dc}} = 0.0$ yields the fastest convergence being 5× faster than w/o CGFM.

The bottom line is that FM + QO w/CGFM exhibits superior performance in identifying alloy designs characterized by only a few dominant secondary phases. Compared to FM + QO w/o CGFM, it explores regions of the search space that are not represented in the initial data set more efficiently. However, when targeting optimal designs with more evenly distributed component ratios, the benefits of the CGFM ansatz have to compensate for the more complicated FM training. Finally, it should be noted, that across all scenarios our CGFM ansatz consistently accelerates the active learning process when measured in real time, owing to QUBO model simplification, as demonstrated in detail in the SI Sec. S4.

2.2.2 Multi-objective optimization with DDTS ansatz. We now employ the multi-objective FM + QO active learning approach to systematically evaluate the performance of our DDTS ansatz. Therefore we maintain the Al-alloy model system already studied above having a fixed 80 vol.[%] Al-matrix and described in terms of normalized fraction variables, $\bar{f}_k \in [0, 1]$ with $k = 1, \ldots, N_{\mathrm{ph}}$, of the $N_{\mathrm{ph}} = 4$ phases: eutectic Si, Mg_2Si, Al_3Ni and Al_2Cu. Further, we keep the standard FM + QO without CGFM and apply one-hot encoding to translate each of the discretized $\bar{f}_k$ to a number of $N_{\mathrm{bits}} = 25$ binary variables. SI S1.3 holds further computational details.

Given the inherent conflict of optimizing κ, E and ρ simultaneously, these material properties are well-suited to demonstrate multi-objective Pareto optimization algorithms. Although we focus the following discussion on this scenario, other combinations of the five objectives show similar behaviour and all results can be found in SI S7.

The DDTS workflow is depicted in the upper part of Fig. 3. Now let us compare the generated Pareto fronts based on the DDTS to the results obtained from the FM + QO approach with standard weighted-sum scalarization (shown in the lower part of Fig. 3). For convenience, we refer to both ansatzes as FM + QO w/DDTS and w/o DDTS in the following. Pareto results from the multi-objective FM + QO algorithm are presented in the 3D objective space spanned by κ, E and ρ in Fig. 5. Fig. 5a and b refer to optimization runs w/and w/o DDTS, respectively. The active learning process is stopped after completing 1000 iterations. The preference weights, w_i, for setting the different scalarizations in both ansatzes are randomly sampled for each

© 2026 The Author(s). Published by the Royal Society of Chemistry

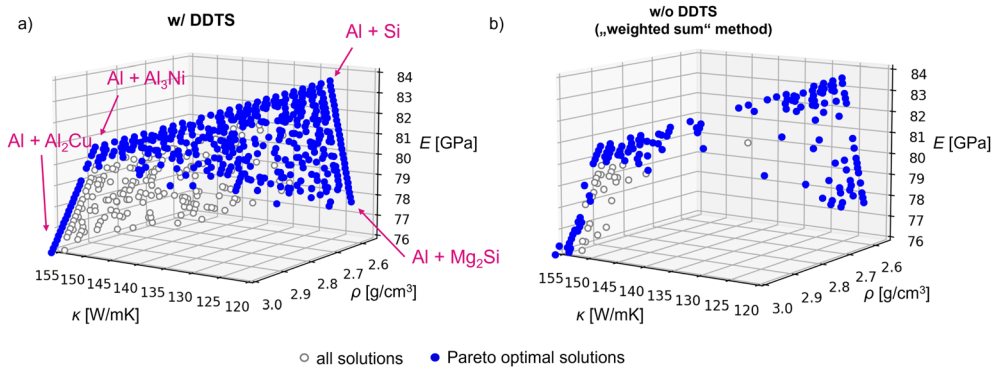


Fig. 5 Multi-objective optimization results with the FM + QO method based on either the DDTS ansatz (w/DDTS) or the standard weighted-sum scalarization (w/o DDTS). Thermal conductivity (κ), Young's modulus (E) and mass density (ρ) span the objective space. (a and b) Solutions found within 1000 iterations w/and w/o DDTS, respectively. Gray points are all QUBO solutions (best alloy designs) found. (These can appear multiple times, the substituted random designs are not illustrated.) The blue points illustrate the 3D Pareto front (non-dominating solutions). DDTS ansatz delivers a homogeneous distribution of Pareto optimal alloy designs across the entire front, which is in stark contrast to the ansatz w/o the DDTS. According to our property models, the alloy designs in the Pareto "corners" are given by the four possible binary Al alloys: Al + Si, Al + Mg_2Si, Al + Al_3Ni and Al + Al_2Cu.

active learning iteration in order to generate the entire Pareto front. The comparison of the Pareto fronts in Fig. 5a and b demonstrates the advantage of our DDTS ansatz against the standard weighted-sum scalarization. The blue dots give each the obtained set of Pareto optimal solutions called Pareto front. Also the remaining gray dots are alloy designs proposed by the QO, however, those are not in the final Pareto front. The DDTS ansatz delivers a homogeneous distribution of Pareto optimal alloy designs across the entire Pareto front. To prove this, the SI S6 provides the exact Pareto front based on brute force search. Apparently, the FM + QO w/DDTS is capable to generate Pareto optimal designs with equal probability. In this regard, the ansatz w/o DDTS performs poorly. In Fig. 5b, we observe the found Pareto solutions concentrate on the peripheral regions. In fact, the planar parts of the Pareto front are inaccessible when solution search is based on the weighted-sum scalarization. Only solutions around the corresponding single-objective optima (extreme preference weights): Al + Al_2Cu, Al + Si and Al + Mg_2Si can be returned, as well as in the forth "corner" of the front (around Al + Al_3Ni). Noteworthy, the latter binary composition is not a single-objective optimum in our model but it is responsible for a local convexity which can be straightly accessed by the convex approach. The time-resolved generation of QUBO solutions, *i.e.*, new optimal design candidates found by the QUBO solver, is visualized in Fig. 6a. The temporal evolution demonstrates the ability of the ansatz w/DDTS to homogeneously improve the front over time. Moreover, it shows that throughout the entire active learning process the ansatz w/o DDTS samples only solutions from peripheral regions. The evolution of the Pareto front hypervolume and the inverted generational distance, both quantitatively measuring the temporal optimization progress, are compared in Fig. 6b and c, respectively. The comparison pronounces the advantage of the DDTS algorithm.

We conclude this section by noting that the FM + QO with DDTS ansatz performs exceptionally well for data-driven multi-objective optimization problems. It clearly outperforms standard weighted-sum scalarization and enables the treatment of inherently non-convex QUBO problem settings. In contrast to the weighted-sum method, our DDTS ansatz enables the identification of Pareto-optimal Al-alloy designs even in non-convex regions of the Pareto front, ranging here from binary to quaternary compositions.

Future work will involve systematic studies comparing FM + QO w/DDTS to established non-convex multi-objective optimizers, such as evolutionary algorithms or particle swarm optimization.

2.3 Implications

We see significant potential in QUBO-based quantum optimization for real-world applications. However, significant research challenges remain – both in advancing quantum hardware and in developing algorithmic approaches capable of achieving quantum advantage.

Contributing to the development of tailored algorithms, we present two straightforward and easily implementable methods customized to data-driven, real-valued black-box optimization problems. We test and benchmark them *via* SA, but they are seamlessly transferable to quantum devices.

Constraint-guided feature mapping (CGFM): this approach enables the reformulation of QUBO problems with equality constraints by transforming the problem into specially oriented coordinates. This coordinate transformation reduces the solution space to the lower-dimensional subspace of feasible solutions, effectively eliminating equality constraints. As a result, the QUBO formulation becomes more compact, parameterization is simplified, and subsequent optimization, whether performed *via* quantum annealing (QA), quantum approximate optimization algorithms (QAOA), or classical heuristics, is significantly accelerated.

Data-driven Tchebycheff scalarization (DDTS): this preprocessing method supports efficient QUBO-based multi-

© 2026 The Author(s). Published by the Royal Society of Chemistry

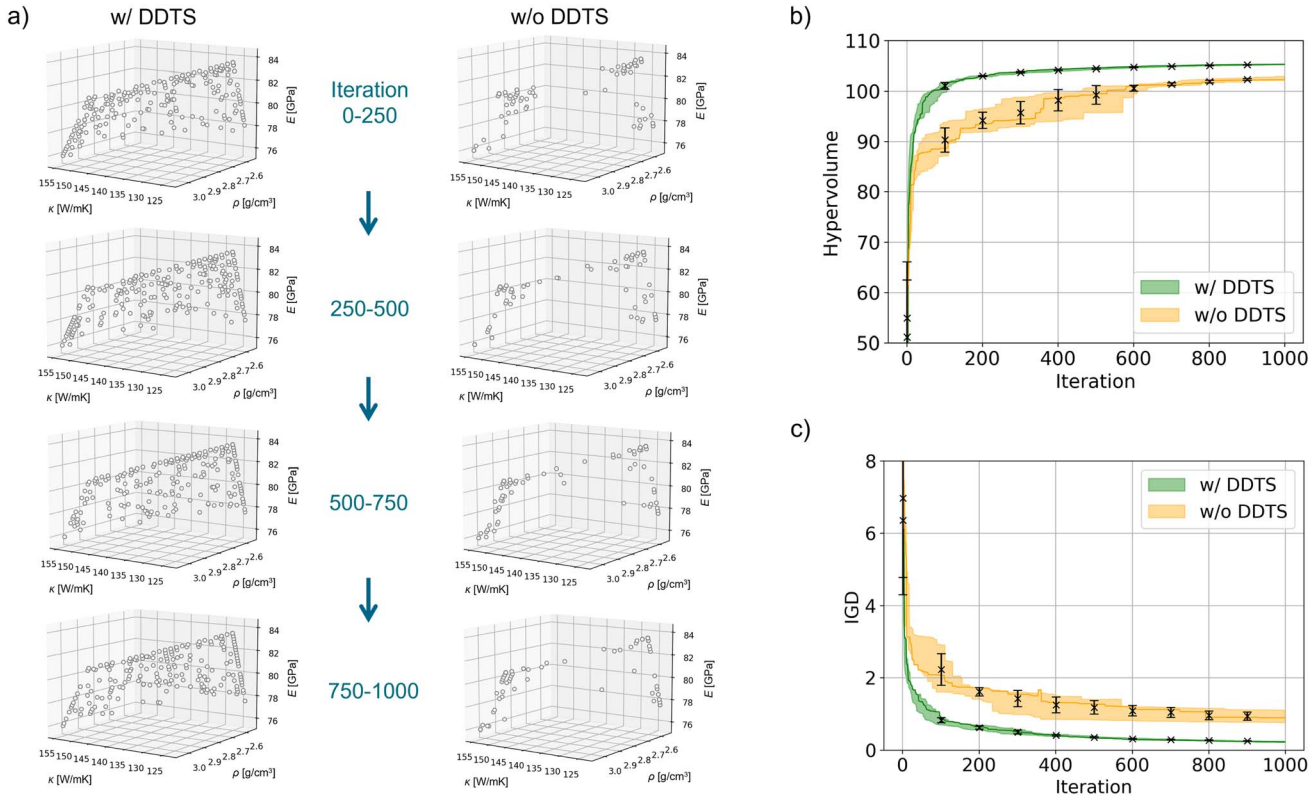


Fig. 6 Iterative Pareto front comparison of FM + QO w/and w/o DDTS ansatz. (a) Temporal progress of objective space discovery of the single FM + QO run w/DDTS and w/o DDTS highlighted in Fig. 5a and b, respectively. Top to bottom: solutions sampled within iterations 0–250, 250–500, 500–750 and 750–1000, respectively. (b) Pareto front hypervolume evolution. (c) Comparison of inverted generational distance (IGD). Shadings indicate min–max spread with respect to each 5 statistical repetitions, single darker green and yellow lines represent the specific FM + QO run corresponding to panel a left and right hand side column, respectively.

objective optimization. Inspired by the Tchebycheff scalarization method, which is well-established for non-convex, multi-objective problems outside the QUBO domain, we adapt the technique to the data-driven QUBO setting. Our formulation guarantees the generation of optimal solutions across the entire Pareto front including non-convex regions. This is a notable improvement over current scalarization strategies in QUBO, which often fail to explore large regions of the front.

We apply both methods to a complex, multi-dimensional, multi-objective optimization task from materials research, namely the design of a multi-phase alloy considering a hypothetical microstructure at μ m-scale. In this paper we use simplified models that limit the reach of our predictions, but are accurate enough to benchmark the methods. More realistic microstructures considering actual descriptors for engineering materials will be considered in the next developments. Looking ahead, we plan to explore the CGFM approach in conjunction with quantum hardware, particularly quantum annealers. These platforms may still offer sufficient qubit capacity to model complex composite ratios effectively. Additionally, a systematic evaluation of CGFM under different QUBO encoding schemes remains an open question. For the latter issue we will expand our studies by investigating the alternative options of regular binary and domain-wall encoding. We consider such extensions as promising avenues for future research.

3 Materials and methods

3.1 Multi-phase alloy model

Our optimization problem describes a hypothetical Al-alloy with four components (phases), which can be found in the microstructure of many cast Al-alloys at μ m-level. Besides the Al-matrix these are: eutectic Si, Mg_2Si, Al_3Ni and Al_2Cu. The task is to vary the fractions [%] of the $N_{ph} = 4$ phases, f_k with $k \in N_{ph}$, in order to maximize κ and E as well as to minimize ρ, $\delta\alpha$ and ΔT. Note further, that with

$$\delta\alpha = |\alpha - \alpha_0| \tag{7}$$

we minimize the absolute value difference of the thermal expansion, α, with respect to a given target value α_0. The content of the aluminum matrix is fixed to 80% vol. fraction, which a realistic value for near eutectic Al–Si alloy.

3.1.1 Thermal and mechanical objectives. We are aware of the generally complex 3D-architecture that phases can form in cast Al–Si alloys,[58–60] however, we keep a very simplified description of the alloy, as shown in Fig. 1a (left). Thereby, we assume that all four phases can be considered as spherical isotropic particles which are homogeneously distributed in the aluminum matrix and do not interact with each other.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Following these assumptions, we employ the Mori-Tanaka model[61] to predict the thermal conductivity of the alloy according to

$$\kappa = \kappa_{\mathrm{Al}} \frac{1+2Q}{1-Q} \tag{8}$$

with

$$Q = \sum_{k=1}^{N_{\mathrm{ph}}} \frac{\kappa_k - \kappa_{\mathrm{Al}}}{2\kappa_{\mathrm{Al}} + \kappa_k} f_k. \tag{9}$$

Conveniently, the k-th phase fraction is introduced in relative units here, *i.e.* $f_k \in [0, 0.2]$, which we use hereafter. The κ_k and κ_{Al} indicate the thermal conductivity of phase k and the Al-matrix (see parameterization in Sec. 3.1.2).

The Young's modulus is estimated in an iterative manner by the Hill model.[62] Hill's model suites well to our morphology assumptions and the resulting E values lie within the upper and lower boundaries given by Voigt[63] and Reuss[64] for fibre-like and plate-like morphology, respectively. The Hill model is initialized by determining

$$\begin{aligned} E_{\mathrm{Hill}} &= \frac{1}{2}\left(E_{\mathrm{Reuss}} + E_{\mathrm{Voigt}}\right) \\ S_{\mathrm{Hill}} &= \frac{1}{2}\left({E_{\mathrm{Reuss}}}^{-1} + {E_{\mathrm{Voigt}}}^{-1}\right) \end{aligned} \tag{10}$$

with the Young's moduli, E_{Reuss} and E_{Voigt}, according to the Reuss and Voigt models,

$$\begin{aligned} E_{\mathrm{Reuss}} &= \left[\frac{f_{\mathrm{Al}}}{E_{\mathrm{Al}}} + \sum_{k=1}^{N_{\mathrm{ph}}} \frac{f_k}{E_k}\right]^{-1} \\ E_{\mathrm{Voigt}} &= f_{\mathrm{Al}} E_{\mathrm{Al}} + \sum_{k=1}^{N_{\mathrm{ph}}} f_k E_k. \end{aligned} \tag{11}$$

The E_{Al} and E_k in eqn (11) indicate the Young's modulus of the matrix and the k-th phases, respectively (see parameterization in Sec. 3.1.2). Subsequently, the initial values, eqn (10), are utilized to solve the equations

$$\begin{aligned} E_{\mathrm{Hill}} &= \frac{1}{2}\left(E_{\mathrm{Hill}} + {S_{\mathrm{Hill}}}^{-1}\right) \\ S_{\mathrm{Hill}} &= \frac{1}{2}\left({E_{\mathrm{Hill}}}^{-1} + S_{\mathrm{Hill}}\right) \end{aligned} \tag{12}$$

self-consistently until convergence is found in terms of $E_{\mathrm{Hill}} = {S_{\mathrm{Hill}}}^{-1}$.

The mass density of the alloy, ρ, is the weighted arithmetic mean considering the density of the aluminum matrix, ρ_{Al}, and the other phases, ρ_k,

$$\rho = f_{\mathrm{Al}}\rho_{\mathrm{Al}} + \sum_{k=1}^{N_{\mathrm{ph}}} f_k \rho_k. \tag{13}$$

In particular, aerospace and automotive industries are interested in lightweight Al-alloys to reduce fuel consumption. The coefficient of thermal expansion indicates the volumetric change of a material as a function of temperature. In particular, the linear coefficient of thermal expansion, α, is defined by the ratio of the relative change in length and the applied temperature difference. Here, we utilize the simple and frequently used Turner model,[65] which follows as

$$\alpha = \frac{f_{\mathrm{Al}}\alpha_{\mathrm{Al}}K_{\mathrm{Al}} + \sum_{k=1}^{N_{\mathrm{ph}}} f_k \alpha_k K_k}{f_{\mathrm{Al}}K_{\mathrm{Al}} + \sum_{k=1}^{N_{\mathrm{ph}}} f_k K_k} \tag{14}$$

and is based on the bulk modulus of the aluminum matrix, K_{Al}, and the k-th phase, K_k (see parameterization in Sec. 3.1.2). Eqn (14) assumes isostrain conditions, *i.e.*, it is suitable for predicting α in alloys with continuous reinforcements, which, though beyond the assumptions introduced above, describes our hypothetical microstructure reliably. Instead of optimizing α directly, we intend to minimize the absolute difference, $\delta\alpha$ (eqn (7)) to a predefined target value. Controlling the thermal expansion of Al alloys can be important in multi-component systems that operate at different temperatures.

Finally, we introduce the solidification interval, ΔT, that gives an indication of the castability of an alloy. As a rule of thumb, it can be considered that the castability decreases with increasing solidification intervals. For a very first estimate, we simplify our picture of the multi-phase alloy and derive ΔT from the Al–Si binary phase diagram, in particular, the temperature difference between L-phase and α + Si-phase. To do that analytically, we apply a linear fitting to the phase boundaries in the range of 0–20% Si. The resulting linear expression approximately models ΔT only depending on the silicon fraction, f_{Si},

$$\Delta T = \begin{cases} -M_1 f_{\mathrm{Si}} + N_1 & \text{if } f_{\mathrm{Si}} < 0.128 \\ 0 & \text{if } f_{\mathrm{Si}} = 0.128 \\ M_2 f_{\mathrm{Si}} - N_2 & \text{if } f_{\mathrm{Si}} > 0.128 \end{cases} \tag{15}$$

with slopes M_1 and M_2 and shifts N_1 and N_2 (see Sec. 3.1.2 for details). Silicon content of 12.8% approximately marks the eutectic point, *i.e.* $\Delta T = 0$ *K*, in the Al–Si binary alloy.

3.1.2 Material parameters. The eqn (7)–(15) are utilized to determine values for our five target properties: thermal conductivity (κ), Young's modulus (E), density (ρ), linear coefficient of thermal expansion (α) and the solidification interval (ΔT) within our Al-alloy design model (see Sec. 3.1). For the phases Si, Mg_2Si, Al_3Ni and Al_2Cu, we utilize the materials parameters given in Table 2. Note that we directly maximize κ (see eqn (8)) and E (see eqn (12)) as well as directly minimize ρ (see eqn (13)) and ΔT (see eqn (15)). The thermal expansion coefficient, however, is optimized to match best a defined target, *i.e.*, our third optimization goal regarding α is to minimize the absolute difference $|\alpha - \alpha_0|$ with α determined by eqn (14) and the target value defined as $\alpha_0 = 20$ ppm K. α_0 is chosen without particular purpose but to be in the mid of range of possible α values resulting from the variation of the four phase fractions. This range stretches from $\alpha = 18.2$ ppm K^{-1} for (80% Al + 20% Si) to $\alpha = 21.9$ ppm K^{-1} for (80% Al + 20% Al_2Cu).

© 2026 The Author(s). Published by the Royal Society of Chemistry

Table 2 Materials properties of the five phases considered in the Al alloy, as introduced in Sec. 3.1 (κ: thermal conductivity, E: Young's modulus, ρ: density, α: linear coefficient of thermal expansion and K: bulk modulus). The parameters are required for computing the optimization objectives of the Al-alloy along the prediction models described in eqn (7)–(15). References from literature related to the parameters are stated

Precipitate phase types and material properties					
Phase	κ [W m^{-1} K]	E [GPa]	ρ [g cm^{-3}]	α [ppm K^{-1}]	K [GPa]
Al (matrix)	164[a]	70 (ref. 66)	2.7	23.6	70 (ref. 67)
Eutectic Si	15 (ref. 68)	163[b]	2.3	2.6 (ref. 69 and 70)	97[b]
Mg_2Si	11 (ref. 71 and 72)	116 (ref. 72 and 73)	2.0	7.5 (ref. 73)	59 (ref. 74)
Al_3Ni	104 (ref. 67)	140 (ref. 75–77)	4.0	13.0 (ref. 67)	110 (ref. 67 and 77)
Al_2Cu	120 (ref. 78 and 79)	105 (ref. 80 and 81)	4.4	17.0 (ref. 82 and 83)	97 (ref. 84)

[a] Please note we assume thermal conductivity of the aluminum matrix, $\kappa_{Al} = 164$ W mK^{-1}. The value lies below the value of 237 W mK^{-1} of pure aluminum due to inherent minor impurities. Density and elastic modulus are less effected by those. [b] The bulk modulus of silicon, K_{Si}, is approximately determined based on $K_{Si} = E_{Si}/(3 - 6\nu_{Si})$ for isotropic materials with the Young's modulus, $E_{Si} = 163$ GPa, and the Poisson ratio, $\nu_{Si} = 0.22$, of silicon. The value E_{Si} is chosen as a reasonable estimate for eutectic silicon.

Furthermore, we implement a simple approximation based on the binary Al–Si system to determine the solidification interval, ΔT. The analytic expression in eqn (15) describes a linear fitting model of the temperature difference between the boundaries of the liquid and solid phase. For this purpose, we approximate the multi-phase alloy with the solidification interval of the Al–Si binary composition. This assumption crucially allows for straightly determining the fitting parameters M_1, M_2, N_1 and N_2 in eqn (15) based on the phase boundaries in the Al–Si phase diagram. A schematic representation of the phase diagram and the fitted boundaries are given in Fig. 7. We are aware of the limitation of our simplification but it helps us to demonstrate the methodology used for the optimization of a multi-component system adding a further variable, namely the solidification interval. A more realistic approach to actually discover a new alloy composition could, *e.g.*, consider determining the solidification interval using the CALPHAD method or by experimental studies of the real alloy compositions. The three points indicated in the phase diagram are used to fit $\Delta T[K] = -M_1 f_{Si} + N_1$ with $M_1 = 669.6$ and $N_1 = 85.7$ for $f_{Si} \in [0, 0.128]$, as well as $\Delta T[K] = M_2 f_{Si} - N_2$ with $M_2 = 1388.9$ and $N_2 = 177.8$ for $f_{Si} \in [0.128, 0.2]$ (see the magenta lines in Fig. 7). For simplicity, we exclude the α-Al phase from our consideration (in the range $f_{Si} \in [0, 0.016]$). Point (2) in Fig. 7 is the eutectic point of $\Delta T = 0$ K. The ΔT is to be minimized in our optimization scenarios, therefore we expect FM + QO to return ideal alloys with $f_{Si} = 0.128$ (*i.e.*, $\bar{f}_{Si} = 0.64$), as far as the introduced search space discretization allows.

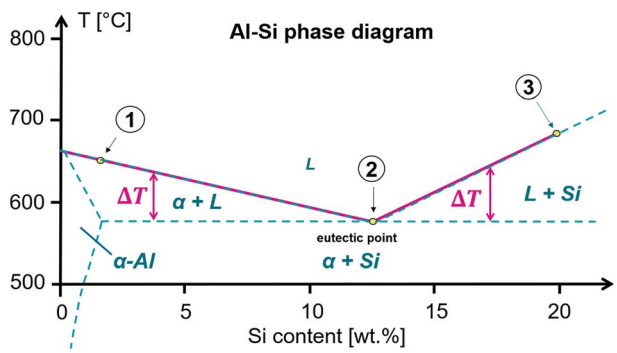


Fig. 7 Phase diagram of binary Al–Si alloy. The solidification interval, ΔT, describes the temperature difference between the liquid and solid phase (L and α + Si phase, respectively) and depends on Si content. Here, the range from 0–20% Si is of interest. According to our assumption, we require the indicated 3 points for linearly fitting ΔT. 1: (1.6, T_S+75), 2: (12.8, T_S), 3: (20, T_S+100) with the solidification temperature $T_S = 577$.

3.2 Model encoding

3.2.1 Discretization of phase fractions. The microstructure of our hypothetical Al-alloy is illustrated in Fig. 1a. The design is defined by the fractions of the phases, f_k, with $k = 1, \ldots, N_{ph}$, which serve our model as variables in the range [0, 0.2], reminding the fixed Al matrix content of 0.8. For the FM + QO method, this continuous variable space must be translated into binary variable space. In the most straightforward manner, an equally-spaced discretization of the [0, 0.2] range can be introduced for this purpose.[2,23] The mesh is applied to all variables and determines the resolution of the design model. For instance, using a discretization with 9 mesh points would allow describing fractions in multiples of 0.02 (*i.e.*, 2%). The discretized variable axes are finally converted into N_{bits} binary variables, x_i, with $i = 1, \ldots, N_{bits}$ (see Sec. 3.2.2). Though more types of 0–1 encoding are well known from literature, we focus on one-hot encoding types.[2,85] The total number of required variables is determined by multiplying the number of bits, N_{bits}, with the number of phases, N_{ph}.

3.2.2 Encoding of phase fractions. To utilize FM + QO the discretized variable ranges of the phase fractions must be further encoded in binary variables. In preparation, we first conveniently introduce *normalized* phase fractions, $\bar{f}_k \in [0, 1]$, according to

$$f_k = L + (U - L)\bar{f}_k \tag{16}$$

with L and U indicating the lower and upper boundaries of the f_k variable range, respectively. Please note, in our model we

© 2026 The Author(s). Published by the Royal Society of Chemistry

identify $L = 0$ and $U = 0.2$. Given a certain number of bits, N_{bits}, available for the encoding of a single normalized variable, $\bar{f}_k$, we write

$$\bar{f}_k = \sum_{i=1}^{N_{\text{bits}}} \alpha_i x_i^{(k)} \tag{17}$$

Here, $x_i^{(k)} \in \{0, 1\}$ is the i-th binary variable (or bit) used to encode the fraction $\bar{f}_k$. The factors α_i depend on the used type of encoding. In general, α_i could be k-dependent as well, however, here we process each $\bar{f}_k$ similarly.

For the one-hot encoding, we have

$$\alpha_i = i \cdot \Delta, \tag{18}$$

with $i = 1, \ldots, N_{\text{bits}}$, where $\Delta = 1/N_{\text{bits}}$ is the increment size resulting from the discretization of normalized fractions, $\bar{f}_k$. Note that to obtain the absolute increment size we would simply have to multiply this value by $U - L$ (*i.e.*, here with a factor of 0.2). The name one-hot encoding refers to the fact that, for this type of encoding, at most one bit can be 1 while the rest has to be 0. In contrast to the classical one-hot encoding, we encode $\bar{f}_k = 0$ by the zero bit-string $x_i = 0$ for all i with the advantage of saving one bit. In this way, intuitively, each bit corresponds to a specific contribution to $\bar{f}_k$. If bit i is set to 1, a part of i/Δ contributes to the total sum in eqn (17). Consequently, the set of representable values spans uniformly the interval [0, 1].

3.3 FM + QO active learning

In general, the FM + QO methodology works through an iterative process of step-wise improvements repeating:[15,17]

1. Data-based FM modeling (FM),
2. QUBO optimization (QO),
3. Enlargement of the data set with the validated solutions

until some form of convergence is reached (see Fig. 2). According to the state-of-the-art weigted-sum scalarization, this straightly generalizes to multi-objective optimization tasks with m objectives as follows: let the data set consist of vectors, $\vec{x}$, describing the system in terms of binary variables, $x_i \in \{0, 1\}$, and the corresponding m optimization target properties, $y^{(1)}, \ldots, y^{(m)}$, *i.e.*, the d-th data point follows as $(\vec{x}_d, y_d{}^{(1)}, \ldots, y_d{}^{(m)})$. Then, one FM is trained independently for each of the m objectives. Details on the FM modeling analytics can be found in the SI S1.1.

The resulting QUBO models, $f^{(p)}(\vec{x})$, with objective index $p = 1, \ldots, m$, read

$$f^{(p)}(\vec{x}) = \vec{x}^{\mathrm{T}} \hat{Q}^{(\mathrm{p})} \vec{x}, \tag{19}$$

based on the learned matrices $\hat{Q}^{(p)}$, the multi-objective model finally follows from merging all $f^{(p)}(\vec{x})$ into a single QUBO,

$$g(\vec{x}) = \sum_{p=1}^{m} w_p \vec{x}^T \hat{Q}^{(p)} \vec{x} + \lambda C(\vec{x}). \tag{20}$$

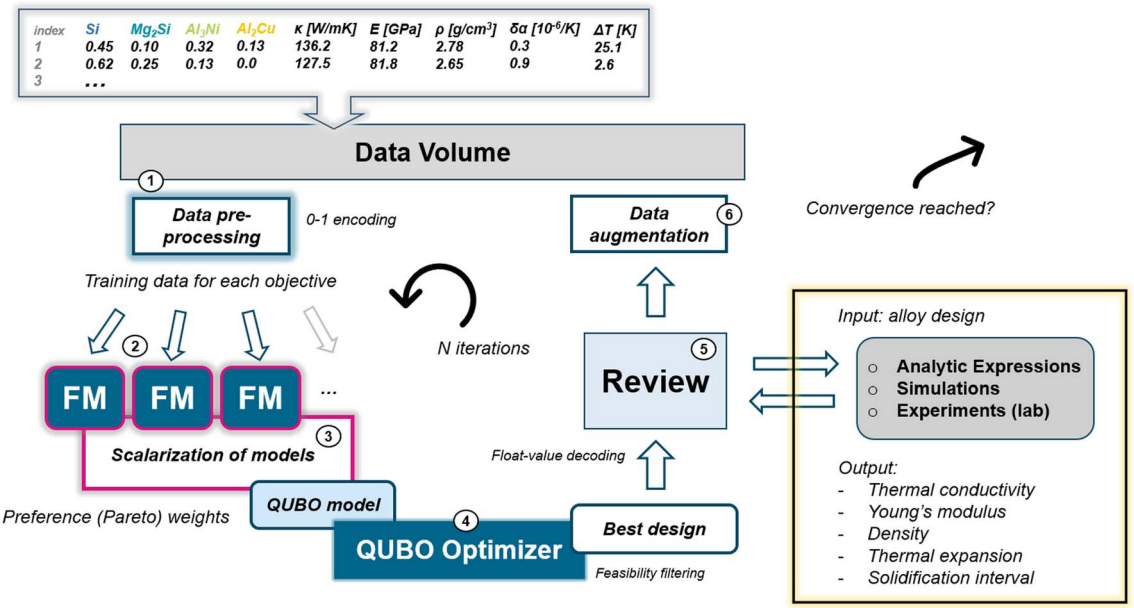


Fig. 8 The multi-objective implementation of the active learning FM + QO method based on recent state-of-the-art QUBO scalarization, here applied for optimizing the microstructure of a hypothetical Al-alloy in terms of normalized phase fractions, $\bar{f}_k \in [0, 1]$ with $k = 1, \ldots, N_{\text{ph}}$ (see eqn (16)). Each iteration starts from encoding and pre-processing the existing data set (1). A factorization machine (FM) is trained for regression of each of the given thermal and mechanical objectives (2). The FMs are scalarized to merge the surrogate models into a single one (3). The resulting QUBO model is solved for searching a new optimal design (4). Finally, the objective's values of the found design are reviewed by means of a validation source, *e.g.*, analytic expressions, simulations or experiments (5). Here, we make use of eqn (7)–(15). This gives a new data point which is added to the data set for improving the modeling in the next iteration (6). Note that the FM + QO pipeline based on our new DDTS ansatz differs in steps (1) – (3), see Fig. 3.

© 2026 The Author(s). Published by the Royal Society of Chemistry

In principle, choosing specific weights, w_p, where usually $\sum_p w_p = 1$, controls the merger preferences regarding the objectives. Finally, system required constraints $C(\vec{x})$ must be added to the QUBO with penalty strength factor λ, to support feasible solutions. Please note that here we conveniently suppressed constraints which additionally arise for specific encodings like the one-hot type and instead refer the interested reader to eqn (S4) in SI S1.2. After solving $g(\vec{x})$ with QO, the learned model is validated and improved. Practically this is realised by determining the "true" objective values for the QO solution and by adding this new, validated data point to the data set. This review process might be facilitated by simulations, experiments, or available analytic formulae. By refining the data set in this manner, the FM model improves iteratively.

In our specific use case, each initial data point represents a random alloy design. Its search space vector, $\vec{x}_d$, describes a concatenation of N_{ph} vectors, $\vec{x}^{(k)} = (x_1{}^{(k)}, \ldots, x_{N_{\text{bits}}}{}^{(k)})$, with $x_i{}^{(k)} \in \{0, 1\}$ and $k = 1, \ldots, N_{\text{ph}}$, obtained by one-hot encoding the normalized phase fraction variables, $\bar{f}_k$ (see Sec. 3.2). Moreover, each data point holds the thermal and mechanical property values, $y_d{}^{(p)}$ with $p \in \{\kappa, E, \rho, \delta\alpha, \Delta T\}$, introduced in Sec. 3.1 and determined by querying the prediction models in eqn (7)–(15). Importantly, these expressions are also our validation sources to review the predicted properties of the optimal Al-alloy design in each iteration (step 5 in Fig. 8).

Author contributions

T. P. led the project, performed all coding and data analysis, and wrote the main draft of the manuscript. D. B and D. M. contributed to the conceptual development of the work and provided guidance and feedback throughout the research process. E. B. and G. R. provided guidance and assessment on the use case selection. All authors contributed to the writing and revision of the manuscript.

Conflicts of interest

T. P., D. M., D. B. and E. B. are named inventors on the pending patent application DE102025118818.1, and all authors are named inventors in the pending patent application DE102025147173.8, both filed by the German Aerospace Center (DLR), related to the results described in this work.

Abbreviations

CGFM	Constraint-guided feature mapping
DDTS	Data-driven Tchebycheff scalarization
QUBO	Quadratic unconstraint binary optimization
FM	Factorization machine
QO	QUBO optimization
QA	Quantum annealing
QAOA	Quantum approximate optimization algorithm
SA	Simulated annealing
κ	Thermal conductivity
E	Young's modulus
ρ	Density
α	Linear coefficient of thermal expansion
ΔT	Solidification interval

Data availability

The source code of the FM + QO algorithm and all data sets utilized for initializing the active learning processes of the presented studies are available on GitHub: https://github.com/dlr-wf/fmqo-alloy-discovery. The source code includes the CGFM and DDTS algorithms, as well as the standard single and multi-objective variant. Configuration files are prepared to make all studies conducted in this work reproducible.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d6dd00126b.

Acknowledgements

We acknowledge the financial support of the DLR Quantum Computing Initiative through the project QuantiCoM (https://qci.dlr.de/quanticom/).

References

1 J. Preskill, Quantum Computing in the NISQ era and beyond, *Quantum*, 2018, **2**, 79, DOI: 10.22331/q-2018-08-06-79.
2 S. Yarkoni, E. Raponi, T. Bäck and S. Schmitt, Quantum annealing for industry applications: introduction and review, *Rep. Prog. Phys.*, 2022, **85**, 104001, DOI: 10.1088/1361-6633/ac8c54.
3 A. Lucas, Ising formulations of many np problems, *Front. Phys.*, 2014, **2**, DOI: 10.3389/fphy.2014.00005.
4 E. Farhi, J. Goldstone & S. Gutmann A quantum approximate optimization algorithm, *arXiv*, 2014, preprint, arXiv:1411.4028, DOI: 10.48550/arXiv.1411.4028.
5 L. Zhou, S.-T. Wang, S. Choi, H. Pichler and M. D. Lukin, Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices, *Phys. Rev. X*, 2020, **10**, 021067.
6 M. W. Johnson, *et al.*, Quantum annealing with manufactured spins, *Nature*, 2011, **473**, 194–198, DOI: 10.1038/nature10012.
7 D-Wave Systems Inc. Advantage processor overview. Tech. Rep., D-Wave Systems Inc, 2022, White Paper.
8 K. Boothby, P. Bunyk, J. Raymond & A. Roy. Next-generation topology of d-wave quantum processors, *arXiv*, 2020, preprint, arXiv:2003.00133, DOI: 10.48550/arXiv.2003.00133.
9 A. D. King, *et al.*, Beyond-classical computation in quantum simulation, *Science*, 2025, **388**, 199–204.
10 S. McArdle, *et al.*, Variational ansatz-based quantum simulation of imaginary time evolution, *npj Quantum Inf.*, 2019, **5**, DOI: 10.1038/s41534-019-0187-2.
11 T. D. Morris, A. Kaushik, M. Roetteler & P. C. Lotshaw Performant near-term quantum combinatorial optimization, *arXiv*, 2024, preprint, arXiv:2404.16135, DOI: 10.48550/arXiv.2404.16135.

© 2026 The Author(s). Published by the Royal Society of Chemistry

12 M. Anthony, E. Boros, Y. Crama and A. Gruber, Quadratic reformulations of nonlinear binary optimization problems, *Math. Program.*, 2017, **162**, 115–144, DOI: 10.1007/s10107-016-1032-4.
13 A. Mandal, A. Roy, S. Upadhyay & H. Ushijima-Mwesigwa Compressed quadratization of higher order binary optimization problems, *arXiv*, 2020, preprint, arXiv:2001.00658, DOI: 10.48550/arXiv.2001.00658.
14 B. A. Wilson, *et al.*, Machine learning framework for quantum sampling of highly constrained, continuous optimization problems, *Applied Physics Reviews*, 2021, **8**, 041418, DOI: 10.1063/5.0060481.
15 R. Tamura, *et al.*, Black-box optimization using factorization and ising machines, *Applied Physics Reviews*, 2026, **13**, 021307, DOI: 10.1063/5.0275450.
16 Z. Xu, W. Shang, S. Kim, E. Lee and T. Luo, Quantum annealing-assisted lattice optimization, *npj Comput. Mater.*, 2025, **11**, 4, DOI: 10.1038/s41524-024-01505-1.
17 K. Kitai, *et al.*, Designing metamaterials with quantum annealing and factorization machines, *Phys. Rev. Res.*, 2020, **2**, 013319, DOI: 10.1103/PhysRevResearch.2.013319.
18 S. Kim, *et al.*, High-performance transparent radiative cooler designed by quantum computing, *ACS Energy Lett.*, 2022, **7**, 4134–4141, DOI: 10.1021/acsenergylett.2c01969.
19 S. Kim, S. Wu, R. Jian, G. Xiong and T. Luo, Design of a high-performance titanium nitride metastructure-based solar absorber using quantum computing-assisted optimization, *ACS Appl. Mater. Interfaces*, 2023, **15**, 40606–40613, DOI: 10.1021/acsami.3c08214.
20 M. Urushihara, M. Karube, K. Yamaguchi & R. Tamura, *Optimization of core–shell nanoparticles using a combination of machine learning and ising machine*, Advanced Photonics Research, 4, 2023, DOI: 10.1002/adpr.202300226.
21 K. Nawa, T. Suzuki, K. Masuda, S. Tanaka and Y. Miura, Quantum annealing optimization method for the design of barrier materials in magnetic tunnel junctions, *Phys. Rev. Appl.*, 2023, **20**, 024044, DOI: 10.1103/PhysRevApplied.20.024044.
22 S. Rendle Factorization machines, In *2010 IEEE International Conference on Data Mining*, 2010, DOI: 10.1109/ICDM.2010.127.
23 S. Arai, H. Oshiyama and H. Nishimori, Effectiveness of quantum annealing for continuous-variable optimization, *Phys. Rev. A*, 2023, **108**, 042403, DOI: 10.1103/PhysRevA.108.042403.
24 I. Hen and M. S. Sarandy, Driver hamiltonians for constrained optimization in quantum annealing, *Phys. Rev. A*, 2016, **93**, 062314, DOI: 10.1103/PhysRevA.93.062312.
25 M. Ohzeki, Breaking limitation of quantum annealer in solving optimization problems under constraints, *Sci. Rep.*, 2020, **10**, 3126, DOI: 10.1038/s41598-020-60022-5.
26 S. Yu and T. Nabil, Applying the hubbard-stratonovich transformation to solve scheduling problems under inequality constraints with quantum annealing, *Front. Phys.*, 2021, **9**, DOI: 10.3389/fphy.2021.730685.
27 P. Mirkarimi, D. C. Hoyle, R. Williams and N. Chancellor, Experimental demonstration of improved quantum optimization with linear ising penalties, *New J. Phys.*, 2024, **26**, 103005, DOI: 10.1088/1367-2630/ad7e4a.
28 T. Binninger, Y.-Y. Ting, P. M. Kowalski and M. H. Eikerling, Optimization of ionic configurations in battery materials by quantum annealing, *Phys. Rev. B*, 2024, **110**, DOI: 10.1103/PhysRevB.110.L180202.
29 J. A. Montañez-Barrera, D. Willsch, A. Maldonado-Romo and K. Michielsen, Unbalanced penalization: a new approach to encode inequality constraints of combinatorial problems for quantum optimization algorithms, *Quantum Sci. Technol.*, 2024, **9**, 025022, DOI: 10.1088/2058-9565/ad35e4.
30 A. Miessen, D. J. Egger, I. Tavernelli and G. Mazzola, Benchmarking digital quantum simulations above hundreds of qubits using quantum critical dynamics, *PRX Quantum*, 2024, **5**, 040320, DOI: 10.1103/PRXQuantum.5.040320.
31 W. H. Zurek, U. Dorner and P. Zoller, Dynamics of a quantum phase transition, *Phys. Rev. Lett.*, 2005, **95**, 105701, DOI: 10.1103/PhysRevLett.95.105701.
32 C. Roch, D. Ratke, J. Nüßlein, T. Gabor and S. Feld, The effect of penalty factors of constrained hamiltonians on the eigenspectrum in quantum annealing, *ACM Trans. Quantum Comput.*, 2023, **4**, 1–18, DOI: 10.1145/3577202.
33 D-Wave Systems Inc, *Ocean Sofware Documentation*, D-Wave Systems Inc, https://docs.ocean.dwavesys.com/.
34 T. Binninger *et al.*, Simulating charging characteristics of lithium iron phosphate by electro-ionic optimization on a quantum annealer, *arXiv*, 2025, preprint, arXiv:2503.10581, DOI: 10.48550/arXiv.2503.10581.
35 S. Hadfield, *et al.*, From the quantum approximate optimization algorithm to a quantum alternating operator ansatz, *Algorithms*, 2019, **12**, 34, DOI: 10.3390/a12020034.
36 Z. Wang, N. C. Rubin, J. M. Dominy and E. G. Rieffel, xy mixers: Analytical and numerical results for the quantum alternating operator ansatz, *Phys. Rev. A*, 2020, **101**, 012320.
37 P. Niroula, *et al.*, Constrained quantum optimization for extractive summarization on a trapped-ion quantum computer, *Sci. Rep.*, 2022, **12**, 17171, DOI: 10.1038/s41598-022-20853-w.
38 M. T. M. Emmerich and A. H. Deutz, A tutorial on multiobjective optimization: fundamentals and evolutionary methods, *Nat. Comput.*, 2018, **17**, 585–609, DOI: 10.1007/s11047-018-9685-y.
39 K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, A fast and elitist multiobjective genetic algorithm: Nsga-ii, *IEEE Trans. Evol. Comput.*, 2002, **6**, 182–197, DOI: 10.1109/4235.996017.
40 C. Olvera, O. Montiel and Y. Rubio, Quantum-inspired evolutionary algorithms on continuous space multiobjective problems, *Soft Comput.*, 2022, **27**, 13143–13164, DOI: 10.1007/s00500-022-06916-0.
41 P. Schworm, X. Wu, M. Klar, M. Glatt and J. C. Aurich, Multi-objective quantum annealing approach for solving flexible job shop scheduling in manufacturing, *J. Manuf. Syst.*, 2024, **72**, 142–153, DOI: 10.1016/j.jmsy.2023.11.015.
42 E. Aguilera, J. de Jong, F. Phillipson, S. Taamallah and M. Vos, Multi-objective portfolio optimization using

© 2026 The Author(s). Published by the Royal Society of Chemistry

a quantum annealer, *Mathematics*, 2024, **12**, 1291, DOI: 10.3390/math12091291.
43 L. Ekstrom, H. Wang and S. Schmitt, Variational quantum multiobjective optimization, *Phys. Rev. Res.*, 2025, **7**, 023141, DOI: 10.1103/PhysRevResearch.7.023141.
44 A. Kotil, *et al.*, Quantum approximate multi-objective optimization, *Nat. Comput. Sci.*, 2025, 1168–1177, DOI: 10.1038/s43588-025-00873-y.
45 Y. Imanaka, T. Anazawa, F. Kumasaka and H. Jippo, Optimization of the composition in a composite material for microelectronics application using the ising model, *Sci. Rep.*, 2021, **11**, 3057, DOI: 10.1038/s41598-021-81243-2.
46 J. L. J. Pereira, G. A. Oliver, M. B. Francisco, S. S. Cunha and G. F. Gomes, A review of multi-objective optimization: Methods and algorithms in mechanical engineering problems, *Arch. Comput. Methods Eng.*, 2021, **29**, 2285–2308, DOI: 10.1007/s11831-021-09663-x.
47 H. Ikeda and T. Yamazaki, Multi-objective optimization technique based on qubo and an ising machine, *IEEE Access*, 2024, **12**, 8957–8969, DOI: 10.1109/access.2024.3353222.
48 P. M. Pardalos, A. Žilinskas & J. Žilinskas, *Non-Convex Multi-Objective Optimization*, Springer International Publishing, 2017, DOI: 10.1007/978-3-319-61007-8.
49 J. Hirsch and T. Al-Samman, Superior light metals by texture engineering: Optimized aluminum and magnesium alloys for automotive applications, *Acta Mater.*, 2013, **61**, 818–843, DOI: 10.1016/j.actamat.2012.10.044.
50 I. Polmear, D. St. John, J. F. Nie & M. Qian, *Light Alloys: Metallurgy of the Light Metals*, Butterworth-Heinemann, United States of America, 2017.
51 C. Wen, *et al.*, Machine learning assisted design of high entropy alloys with desired property, *Acta Mater.*, 2019, **170**, 109–117), https://www.sciencedirect.com/science/article/pii/S1359645419301430.
52 X. Liu, J. Zhang and Z. Pei, Machine learning for high-entropy alloys: Progress, challenges and opportunities, *Prog. Mater. Sci.*, 2023, **131**, 101018), https://www.sciencedirect.com/science/article/pii/S0079642522000998.
53 L. Zhichao, M. Dong, L. Xiongjun and Z. Lu, High-throughput and data-driven machine learning techniques for discovering high-entropy alloys, *Commun. Mater.*, 2024, **5**, 76, DOI: 10.1038/s43246-024-00487-3.
54 J. He, *et al.*, Machine learning-assisted design of high-entropy alloys with superior mechanical properties, *J. Mater. Res. Technol.*, 2024, **33**, 260–286), https://www.sciencedirect.com/science/article/pii/S2238785424020192.
55 Q. Wang and Y. Yao, Harnessing machine learning for high-entropy alloy catalysis: a focus on adsorption energy prediction, *npj Comput. Mater.*, 2025, **11**, 91, DOI: 10.1038/s41524-025-01579-5.
56 D. Ibarra-Hoyos, *et al.*, Quantum-annealed machine learning discovers ductile, high strength and corrosion-resistant high-entropy alloy, *npj Comput. Mater.*, 2026, **12**, 159, DOI: 10.1038/s41524-026-02032-x.
57 R. E. Steuer and E.-U. Choo, An interactive weighted tchebycheff procedure for multiple objective programming, *Math. Program.*, 1983, **26**, 326–344, DOI: 10.1007/BF02591870.
58 Z. Asghar, G. Requena, H. Degischer and P. Cloetens, Three-dimensional study of ni aluminides in an alsi12 alloy by means of light optical and synchrotron microtomography, *Acta Mater.*, 2009, **57**, 4125–4132), https://www.sciencedirect.com/science/article/pii/S1359645409003073.
59 Z. Asghar, G. Requena and F. Kubel, The role of ni and fe aluminides on the elevated temperature strength of an alsi12 alloy, *Mater. Sci. Eng. A*, 2010, **527**, 5691–5698, DOI: 10.1016/j.msea.2010.05.033.
60 D. Tolnai, G. Requena, P. Cloetens, J. Lendvai and H. Degischer, Effect of solution heat treatment on the internal architecture and compressive strength of an almg4.7si8 alloy, *Mater. Sci. Eng. A*, 2013, **585**, 480–487, DOI: 10.1016/j.msea.2013.06.033.
61 Y. Benveniste, On the effective thermal conductivity of multiphase composites, *Z. Angew. Math. Phys.*, 1986, **37**, 696–713, DOI: 10.1007/BF00947917.
62 H. Bunge, R. Kiewel, T. Reinert and L. Fritsche, Elastic properties of polycrystals—influence of texture and stereology, *J. Mech. Phys. Solids*, 2000, **48**, 29–66, DOI: 10.1016/S0022-5096(99)00020-4.
63 W. Voigt, *Lehrbuch der Kristallphysik*, BG Teubner, 1910.
64 A. Reuss, Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle, *ZAMM Z. für Angew. Math. Mech.*, 1929, **9**, 49–58, DOI: 10.1002/zamm.19290090104.
65 P. S. Turner, Thermal-expansion stresses in reinforced plastics, *J. Res. Natl. Bur. Stand.*, 1946, **37**, 239–250.
66 A. Barnoush, Correlation between dislocation density and nanomechanical response during nanoindentation, *Acta Mater.*, 2012, **60**, 1268–1277, DOI: 10.1016/j.actamat.2011.11.034.
67 L. Mo, *et al.*, Microstructure and thermal-physical properties of hypereutectic al-ni alloys, *J. Mater. Res. Technol.*, 2023, **24**, 6227–6237, DOI: 10.1016/j.jmrt.2023.04.168.
68 A. Zhang and Y. Li, Thermal conductivity of aluminum alloys—a review, *Materials*, 2023, **16**, 2972, DOI: 10.3390/ma16082972.
69 H. Watanabe, N. Yamada and M. Okaji, Linear thermal expansion coefficient of silicon from 293 to 1000 k, *Int. J. Thermophys.*, 2004, **25**, 221–236, DOI: 10.1023/B:IJOT.0000022336.83719.43.
70 T. Middelmann, A. Walkov, G. Bartl and R. Schödel, Thermal expansion coefficient of single-crystal silicon from 7 k to 293 k, *Phys. Rev. B*, 2015, **92**, 174113, DOI: 10.1103/PhysRevB.92.174113.
71 J. Martin, Thermal conductivity of mg2si, mg2ge and mg2sn, *J. Phys. Chem. Solids*, 1972, **33**, 1139–1148, DOI: 10.1016/S0022-3697(72)80273-7.
72 J.-i. Tani and H. Kido, Lattice dynamics of mg2si and mg2ge compounds from first-principles calculations, *Comput.*

© 2026 The Author(s). Published by the Royal Society of Chemistry

Mater. Sci., 2008, **42**, 531–536, DOI: 10.1016/j.commatsci.2007.08.018.

73 X. Lan, *et al.*, Preparation of millimeter scale second phase particles in aluminum alloys and determination of their mechanical properties, *J. Alloys Compd.*, 2019, **784**, 68–75, DOI: 10.1016/j.jallcom.2018.12.395.

74 B. Yu, D. Chen, Q. Tang, C. Wang and D. Shi, Structural, electronic, elastic and thermal properties of mg2si, *J. Phys. Chem. Solids*, 2010, **71**, 758–763, DOI: 10.1016/j.jpcs.2010.01.017.

75 Y. Fukui, K. Takashima and C. B. Ponton, Measurement of young's modulus and internal friction of an in situ al-al3ni functionally gradient material, *J. Mater. Sci.*, 1994, **29**, 2281–2288, DOI: 10.1007/BF00363415.

76 R. W. Hertzberg, F. D. Lemkey & J. A. Ford Mechanical behavior of lamellar (al-cual2) and whisker type (al-al3ni) unidirectionally-solidified eutectic alloys. In *Transactions of the Metallurgical Society of AIME*, The American Institute of Mining, Metallurgical, and Petroleum Engineers, 1965.

77 C. Xiao, *et al.*, First-principles-based structural and mechanical properties of al3ni under high pressure, *Crystals*, 2024, **15**, 3, DOI: 10.3390/cryst15010003.

78 C. Zhang, Y. Du, S. Liu, Y. Liu and B. Sundman, Thermal conductivity of al–cu–mg–si alloys: Experimental measurement and calphad modeling, *Thermochim. Acta*, 2016, **635**, 8–16, DOI: 10.1016/j.tca.2016.04.019.

79 G. Wei, *et al.*, Thermophysical property measurements and thermal energy storage capacity analysis of aluminum alloys, *Sol. Energy*, 2016, **137**, 66–72, DOI: 10.1016/j.solener.2016.07.054.

80 J. Wang, B. Li and C.-H. Zhang, First-principles study of the effect of v, co, ca, sr, ga, as doping on the mechanical properties of al2cu, *Appl. Phys. A:Mater. Sci. Process.*, 2021, **127**, 637, DOI: 10.1007/s00339-021-04791-4.

81 J. Tian, Y. Zhao, H. Hou and P. Han, First-principles investigation of the structural, mechanical and thermodynamic properties of al 2 cu phase under various pressure and temperature conditions, *Solid State Commun.*, 2017, **268**, 44–50, DOI: 10.1016/j.ssc.2017.09.016.

82 C.-L. Chen and R. Thomson, Study on thermal expansion of intermetallics in multicomponent al–si alloys by high temperature x-ray diffraction, *Intermetallics*, 2010, **18**, 1750–1757, DOI: 10.1016/j.intermet.2010.05.015.

83 J. P. Lokker, *et al.*, Localized stress near and the thermal expansion of al2cu precipitates in an al thin film matrix, *J. Appl. Phys.*, 2000, **87**, 682–688, DOI: 10.1063/1.371926.

84 Y. Jin, *et al.*, First-principles investigation of the properties of copper–aluminum intermetallic compounds, *Phys. Met. Metallogr.*, 2024, **125**, S121–S132, DOI: 10.1134/S0031918X24601446.

85 F. Dominguez, *et al.*, Encoding-independent optimization problem formulation for quantum computing, *Front. Quantum Sci. Technol.*, 2023, **2**, DOI: 10.3389/frqst.2023.1229471.

© 2026 The Author(s). Published by the Royal Society of Chemistry