



Cite this: *Phys. Chem. Chem. Phys.*,
2025, 27, 25636

Improving structure search with hyperspatial optimization and TETRIS seeding

Daviti Gochitashvili,^{id a} Maxwell Meyers,^a Cindy Wang^b and
Aleksey N. Kolmogorov^{id *a}

Advanced structure prediction methods developed over the past decades include an unorthodox strategy of allowing atoms to displace into extra dimensions. A recently implemented global optimization of structures from hyperspace (GOSH) has shown promise in accelerating the identification of global minima on potential energy surfaces defined by simple interatomic models. In this study, we extend the GOSH formalism to more accurate Behler–Parrinello neural network (NN) potentials, make it compatible with efficient local minimization algorithms, and test its performance on nanoparticles and crystalline solids. For clusters modeled with NN potentials, four-dimensional optimization offers fairly modest improvement in navigating geometric relaxation pathways and incurs increased computational cost largely offsetting the benefit, but it provides a significant advantage in facilitating atom swaps in nanoalloys. In comparison, the introduction of a moderate, controlled bias for generating more physically sensible starting configurations, achieved via TETRIS-inspired packing of atomic blocks, has a more direct impact on the efficiency of global structure searches. The benchmarked systems are Lennard–Jones clusters, Au or Cu–Pd–Ag nanoparticles and binary Sn alloys described by NN potentials, and compounds with covalent B or BC frameworks modeled with density functional theory.

Received 24th June 2025,
Accepted 21st October 2025

DOI: 10.1039/d5cp02412a

rsc.li/pccp

1. Introduction

Accurate and efficient prediction of atomic structures has evolved from an intractable challenge in the 1990s to an integral component of the materials discovery process.^{1–4} Advanced structure search methods are now routinely used to identify crystalline and nanoscale configurations synthesizable under specific experimental conditions.^{5–7} Prediction of thermodynamically stable compounds with unique morphologies beyond known prototypes^{8–11} has been particularly important in the exploration of matter subjected to extreme pressures. Notable predictions that guided materials synthesis or helped interpret experimental results include the ground states of elemental Na¹² and B¹³ exhibiting unusual properties under compression; LiB,^{14,15} MB₄ (M = Cr, Mn, or Fe),^{16–18} and CaB₆^{19,20} metal borides featuring unfamiliar frameworks; Na₂IrO₃,²¹ Na₃Ir₃O₈,²² and Cu₂IrO₃,²³ iridates showcasing rich strongly correlated physics; and SiH₄,^{24,25} H₃S,²⁶ and other compressed hydrides boasting record superconducting critical temperatures.

The complex optimization problem of locating global minima on multidimensional potential energy surfaces in

given chemical spaces has been addressed with a variety of strategies implemented in USPEX,⁴ XtalOpt,²⁷ CALYPSO,²⁸ AIRSS,²⁹ MAISE,³⁰ and other packages. The methodological diversity reflects the implications of the no-free-lunch theorem, which posits that no single optimization algorithm can outperform all others across every possible system.³¹ While the developed algorithms are often fine-tuned for specific materials classes, such as metal alloys,^{32,33} molecular crystals,^{34,35} or nanoparticles,^{36,37} their overall efficiency is largely determined by common factors: how structures are generated, evolved, and selected.

Creation of diverse yet physically meaningful starting configurations is critical for a large-scale efficient screening and ranges from purely random to highly constrained. It has been demonstrated that the introduction of atom exclusion zones,^{38,39} incorporation of lattice parameters from experiment,¹⁹ seeding searches with supercells of randomized known ground states,^{19,22} or generation of crystal structures with randomly chosen symmetries^{40,41} can speed up the determination of ground states by orders of magnitude.

Equally important is to control bias in structure optimization. Local relaxation of starting configurations with deterministic gradient-based algorithms introduces minimal bias and has been used extensively in *ab initio* random structure searches (AIRSS).²⁹ Variable degrees of implicit steering toward global minima can be achieved with the minima

^a Department of Physics, Applied Physics, and Astronomy, Binghamton University–SUNY, Binghamton, New York 13902, USA. E-mail: kolmogorov@binghamton.edu

^b Barrington High School, Barrington, Illinois 60010, USA



hopping method,⁴² particle swarm optimization,⁴³ evolutionary algorithm,^{4,44,45} and other strategies that allow structure populations to escape from local minima. Modified variable-composition^{46–48} and variable-size multitribes⁴⁹ evolutionary algorithms further facilitate the sharing of favorable motifs identified in large configuration spaces to iteratively construct improved configurations. The introduction of fingerprint-based energy penalty for low-symmetry structures has been shown to enhance convergence in various systems.⁵⁰

Screening candidate structures with classical interatomic potentials can be a promising means of accelerating *ab initio* structure searches. Inexpensive traditional interaction-specific models have been used extensively in studies of nanoclusters³⁷ but their generally limited transferability can lead to finding suboptimal solutions.^{49,51} Recently developed machine learning interatomic potentials offer significantly higher accuracy^{52,53} and have been successfully used to identify more favorable nanocluster configurations^{49,51} and thermodynamically stable compounds^{54,55} overlooked in previous *ab initio* searches.

An out-of-the box idea explored decades ago is based on performing global searches in higher dimensions. It was proposed that introducing extra spatial dimensions could help structures escape local minima by enabling exploration of directions unavailable in 3D space.^{56,57} In these studies, structures were first minimized in 3D, then displaced into a higher-dimensional configuration space where they performed constant-energy walks, before being returned to 3D for further relaxation. The added degrees of freedom provided access to relaxation pathways that more effectively traverse barriers on the potential energy surface, increasing the likelihood of reaching low-energy configurations. This strategy proved particularly useful for systems with complex landscapes, such as the double-funnel topology of the Lennard-Jones (LJ) 38-atom cluster. The unorthodox approach was investigated further in a recent study.⁵⁸ The global optimization of structures from hyperspace (GOSH) generates atomic configurations stochastically in higher special dimensions and gradually brings them back to the real 3D space during a local optimization. The algorithm was found to outperform the traditional 3D local relaxation for several non-periodic systems, from unary and binary clusters to linear chains and covalently bonded frameworks, modeled with simple pairwise potentials. The probability enhancement of finding the ground state reached two orders of magnitude in some cases and was especially evident for the connected systems.⁵⁸ One of the open questions was how effective the hyperdimensional optimization would be in the exploration of energy surfaces defined by more sophisticated potentials.

The present study focuses on implementing and examining two structure generation and optimization factors. First, we investigate the performance of GOSH by extending the formalism to neural network potentials (NNPs). We have modified our MAISE package³⁰ to perform GOSH for clusters and crystals described with more complex interatomic NNPs of the Behler-Parrinello type⁵⁹ that provide near- *ab initio* accuracy. Our implementation reproduces and enhances the previously

observed efficiency gains for LJ monoatomic clusters but shows little to no advantage over the deterministic 3D optimization in the case of Au clusters or Sn crystalline alloys, especially when the increased computational cost is taken into account. On the other hand, a notable improvement is observed for multicomponent clusters.

Second, we generalize our previously proposed TETRIS generation of atomic configurations⁴⁹ to periodic materials and benchmark its performance against more stochastic population initialization. The scheme was originally designed for creating compact starting nanoparticle configurations with physically meaningful interatomic distances by shooting atoms toward the cluster's center iteratively and keeping the closest positions. Our present tests reveal that this essentially parameter-free approach offers a better efficiency improvement than GOSH does for monoatomic LJ and multicomponent Cu–Pd–Ag clusters. The implementation of the TETRIS scheme for crystalline compounds enables a controlled injection of bias in the form of atomic blocks and targeted number of bonds. We demonstrate that the method has little effect on the probability of finding ground states in M–Sn alloys (M = Li, Pd, or Ag) with predominantly metallic bonding^{54,55} but offers significant improvement in locating more stable configurations in a layered metal borocarbide (ambient-pressure LiB₂C₂)⁶⁰ and a metal boride (CaB₆ at 50 GPa)¹⁹ featuring complex covalent networks. The customizable TETRIS seeding protocol, made available in the open-source MAISE package,³⁰ holds promise for accelerating the convergence of global optimization searches *via* a judicious introduction of beneficial motifs in structure populations.

II. Methods

A. GOSH original formulation

The global optimization method designed by Pickard⁵⁸ relies on generating atomic arrangements in higher dimensions, following a downhill trajectory on a hypersurface, and eventually forcing the system to materialize in the normal space. The creation of non-periodic structures involves random placement of hard hyperspheres of radius \tilde{a}_0 within a confining hypersphere of volume \tilde{V}/f , where \tilde{a}_0 is half the expected equilibrium bond length, \tilde{V} is the total volume of atom centered hyperspheres, and f is the packing fraction. The potential energy surface in the hyperspace of dimension $d = d_0 + d_+$ is a combination of the pairwise potential energy $V_{ij}(\tilde{r}_{ij})$ and a harmonic penalty term

$$\bar{E}(\{\tilde{\mathbf{x}}_i\}) = \tilde{E}(\{\tilde{\mathbf{x}}_i\}) + \frac{1}{2}\mu \sum_i l_i^2 = \sum_{ij} V_{ij}(\tilde{r}_{ij}) + \frac{1}{2}\mu \sum_i l_i^2,$$

where the Euclidian distances are defined as

$$\tilde{r}_{i,j} = \left[\sum_{n=1}^d (\tilde{x}_{j,n} - \tilde{x}_{i,n})^2 \right]^{1/2},$$

$$l_i = \left[\sum_{n=d_0+1}^d (\tilde{x}_{i,n})^2 \right]^{1/2}.$$



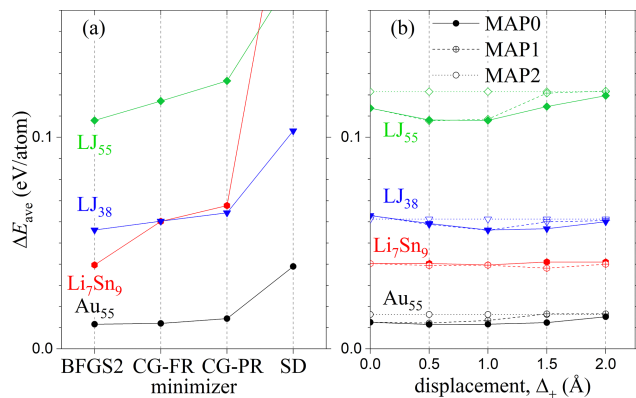


Fig. 1 Performance of different protocols in hyperdimensional optimization of LJ or Au clusters and Li–Sn crystals. ΔE_{ave} represents the average energy of the best half of 2000 structures relative to the corresponding ground state. (a) The local minimization algorithms were tested for MAP0 and $\Delta_+ = 1.0$ Å. (b) The mapping and displacement variants were tested with BFGS2. The MAP2 scheme is independent of Δ_+ .

The chosen two-step steepest-descent algorithm⁶¹ that does not require line minimization allows for incrementing the spring constant $\mu_{k+1} = \mu_k \beta$ after each atomic position optimization cycle k . The $\mu_0 = 10$ and $\beta = 0.001$ values provided an efficient convergence of the hyperspatial energy gradients and ensured the fall of extraspatial coordinates below an acceptable threshold for most systems, while a more conservative $\mu_0 = 1$ starting value was used for covalent materials.

B. GOSH implementation in MAISE

As discussed by Pickard,⁵⁸ GOSH can be naturally generalized for non-pairwise potentials and coupled with other minimization algorithms. In order to make the method compatible with the existing structure generation and optimization capabilities available in MAISE, we introduced a few modifications in the original scheme.

First, we enabled access to a wider range of local optimization algorithms with line minimization available in the GSL⁶² by switching from gradual stepwise to sharp plateau-wise change in the spring constant μ . We break down the maximum number of minimization steps M into N equal groups and set $\mu_n = \mu_0 \beta^n$. In the last set, we zero the extraspatial coordinates, which are almost always negligible at this stage, to ensure that the structure is fully optimized in the normal space. At each stage n , the relaxation halts if the energy change between successive steps falls below 10^{-8} eV (the same tight tolerance value is used in pure 3D minimizations). We observed little dependence of the optimization efficiency on these settings and used $M = 600$, $N = 6$, $\mu_0 = 2$, and $\beta = 2$ (for ease of comparison, we keep specifying unitless μ values but scale the parameter as eV Å^{-2} and set the pairwise LJ energy parameter to 1 eV in our simulations). Choosing a moderate starting spring constant value of $\mu_0 = 2$ proved important for converging structures with large initial extraspatial coordinates. Based on the benchmarking tests for Broyden–Fletcher–Goldfarb–Shanno (BFGS2), Fletcher–Reeves conjugate gradient (CG-FR), Polak–Ribiere conjugate

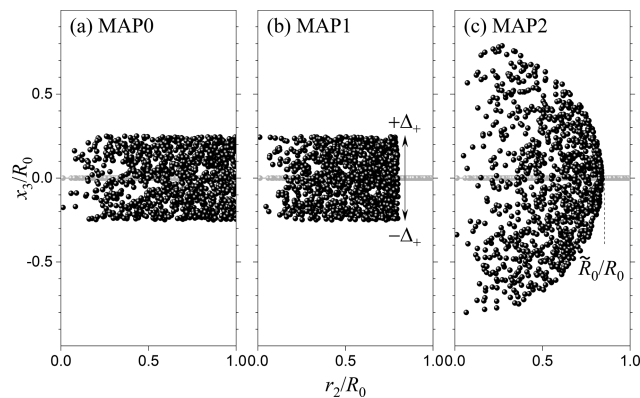


Fig. 2 Atom distributions in 20 50-atom 2D clusters (in gray) expanded into 3D (in black) using different mapping protocols. The hyperdimensional x_3 coordinate is plotted as a function of the 2D radial distance normalized to the largest radius R_0 . (a) x_3 is generated randomly in $[-\Delta_+, \Delta_+]$. (b) After the random x_3 generation, the 2D coordinates are contracted. (c) The 2D coordinates are remapped and x_3 is generated randomly in the $\pm\sqrt{\tilde{R}_0^2 - r_i^2}$ range as detailed in the main text.

gradient (CG-PR), and steepest descent (SD) algorithms^{63–66} in Fig. 1(a), we chose BFGS2 as the default minimizer.

Second, we implemented alternative procedures for extending structures beyond the normal space. For clusters, we aimed to retain the possibility of initializing configurations with the efficient TETRIS algorithm or generating offspring with various mutation or crossover operations in evolutionary searches.^{30,49,51} Therefore, instead of creating clusters with equally randomized atomic coordinates in all dimensions, we enabled random shifts into the extra dimensions for clusters with already defined 3D atomic positions. Protocol MAP0 illustrated in Fig. 2(a) involves uncorrelated random displacements of atoms into each extra dimension up to Δ_+ (for $d_0 = 2$ and $d_+ = 1$, this operation transforms a 2D disk into a fuzzy flat cylinder). Optionally, the normal coordinates are uniformly contracted to compensate for the increase of the interatomic distances in the hyperspace (MAP1 in Fig. 2(b)). Fig. 1(b) illustrates a good performance of the modified GOSH for $\Delta_+ \sim 1$ Å. The last tested protocol, MAP2 in Fig. 2(c), treats the normal and extra dimensions on a more equal footing and maintains a roughly uniform atomic density *via* a correlated adjustment of the normal and hyper coordinates (for $d_0 = 2$ and $d_+ = 1$, it converts a 2D disk into a near-spherical 3D shape). If the maximum distance to the origin in a cluster with N atoms is R_0 in the normal space and becomes \tilde{R}_0 upon expansion into one extra dimension, the normal space radial distance $r_{d_0}^{\text{old}}$ for each atom can be mapped approximately as

$$r_{d_0} = \tilde{R}_0 \left[1 - \left(1 - \left(\frac{r_{d_0}^{\text{old}}}{R_0} \right)^2 \right)^p \right]^{1/2},$$

with $p = 0.625$, $2/3$, and 0.694 for 1D, 2D, and 3D, respectively (see Note SI with Fig. S1 and Fig. S2 for further detail⁶⁷). The extra dimensional coordinate x_{d_0+1} should be generated randomly in the $\pm\sqrt{\tilde{R}_0^2 - r_{d_0}^2}$ range. The \tilde{R}_0/R_0 ratios, assessed



by rearranging N atoms on uniform grids in d_0 and $d_0 + 1$ dimensions, are $1.110N^{-1/6}$ for $d_0 = 2$ and $1.082N^{-1/12}$ for $d_0 = 3$. Since Fig. 1(b) illustrates a generally good performance of the MAP0 protocol with $\Delta_+ = 1 \text{ \AA}$, we have used it in all present simulations unless specified otherwise.

For crystals, the normal and extra dimensions are inherently different because the periodicity over hyperspatial coordinates, once introduced, cannot be naturally phased out. Therefore, we chose to extend crystalline structures with 3D periodicity by introducing random hyperspatial positions up to $\pm\Delta_+ = \pm 1 \text{ \AA}$ and uniformly shrinking the 3D volume by 0.8. It should be noted that the normal space contraction compensating for the average bond elongation due to the random shifts in the extra dimensions may bring some atoms too close to each other. To avoid that, we perform a series of hyperspatial coordinate adjustments with a simple repulsive potential for both clusters and crystals until all interatomic distances are larger than 0.7 of the expected equilibrium length, just as done for making random 3D structures.^{30,49}

C. TETRIS algorithm

MAISE features several options for generating initial structures.³⁰ Our default approach for constructing clusters, denoted BLOB in the code,³⁰ relies on randomly placing atoms around the origin and tweaking their positions to ensure realistic interatomic distances with a combination of a repulsive pairwise potential and attractive springs that keep atoms within a targeted cluster radius. An alternative TETRIS-inspired algorithm involves sending atoms from random directions toward the origin one by one and keeping the closest position out of $N^{3/2}$ tries to create spherical or ellipsoidal shapes.⁴⁹ Our previous studies illustrated benefits of initializing evolutionary searches with TETRIS-generated configurations. Fig. S3⁶⁷ confirms that TETRIS clusters are more compact than BLOB ones, with the volume definition adopted from ref. 68, even after the latter procedure is tightened with a compression factor of 1.3.

We now extend the method to crystals and introduce an option to seed structures with predefined motifs. While packing algorithms are popular for determining molecular crystal ground states,^{34,35} the use of atomic building blocks is uncommon in explorations of inorganic solid-state materials. Our TETRIS-based generation proceeds as follows: (i) create random lattice vectors; (ii) extend the c lattice vector to temporarily disable periodicity along this direction; (iii) randomly select atomic blocks, natural for the material's bonding type and producing the desired composition, from a user-defined set; (iv) optionally perform a random 3D rotation of each block using quaternions; (v) send blocks one by one from a random lateral position along the c axis, trying all possible orientations in the x - y plane at each downward step, until they either reach the bottom or come into close contact with another block; (vi) repeat the previous step for several starting lateral positions and keep the position with the lowest z ; and (vii) once all the blocks are dropped, turn the periodicity along c back on and shrink the axis until atom separations match the shortest allowed distance.

Due to the observed insignificant influence of implementation details on the method's overall performance, the degree of introduced bias can be controlled by customizing the set of atomic blocks. Additionally, one can specify the targeted number of nearest neighbors within a cutoff sphere for each species to discard configurations with unphysical local environments. Our tests illustrate that the designed algorithm works particularly well for generating layered covalent frameworks.

D. DFT and NN calculations

For metal borides and borocarbides, DFT calculations were performed with VASP⁶⁹ using projector augmented wave potentials⁷⁰ and the optB88-*vdW* functional.⁷¹ A 500 eV plane-wave cutoff and dense Monkhorst-Pack k -meshes⁷² with $\Delta k \sim 2\pi \times 0.025 \text{ \AA}^{-1}$ ensured numerical convergence to typically within 1 meV per atom.

Detailed information about our previously developed NN interatomic potentials for Au, Cu-Pd-Sn, and Sn alloys, available on GitHub, can be found in ref. 30, 49, 51, 54 and 55. The local atomic environments were represented with 51 Behler-Parrinello symmetry functions per element, with the explicit expressions provided in Note SII⁶⁷ and the parameters detailed in ref. 49. The details of NN models of different systems studied in this paper are summarized in Table 1. NNs were fitted to PBE-level energies and forces and extensively tested in global structure searches. The iterative generation of reference structure sets and NN training were done with the automated MAISE-NET wrapper.³⁰

III. Results and discussion

A. Lennard-Jones clusters

Over a century, the LJ potential has served as a canonical model for systems with isotropic interactions and negligible many-body effects. Despite the simplicity of the functional form, the ground states of LJ clusters, established at least up to at $N = 150$ atoms,^{42,74,75} exhibit surprising structural complexity and have become standard benchmarks for developing and tuning global optimization algorithms.^{42,44,50,58,74-77}

We applied our implementations of the TETRIS and BLOB structure generation methods along with hyperdimensional

Table 1 Neural network (NN) parameterization details by chemical composition, including the total number of adjustable weights, the number of energy and force data points in the training sets, and the root mean square errors on the test sets for energies (meV per atom) and forces (meV \AA^{-1}). All models were trained and evaluated using PBE-level⁷³ reference data

| Chemical system | # of weights | # of E data | # of F data | Testing error E | Testing error F | Reference study |
|-----------------|--------------|---------------|---------------|-------------------|-------------------|-----------------|
| Au | 641 | 2912 | 21 621 | 6.5 | 37 | 51 |
| Cu-Pd | 3162 | 3725 | 32 223 | 6.6 | 63 | 49 |
| Cu-Ag | 3162 | 3724 | 32 034 | 3.5 | 38 | 49 |
| Pd-Ag | 3162 | 3705 | 32 166 | 4.8 | 58 | 49 |
| Cu-Pd-Ag | 8853 | 2191 | 29 163 | 5.2 | 53 | 49 |
| Li-Sn | 3162 | 6046 | 46 410 | 10.2 | 49 | 54 |
| Pd-Sn | 3162 | 5507 | 46 272 | 9.6 | 52 | 55 |
| Ag-Sn | 3162 | 5410 | 37 548 | 8.4 | 33 | 55 |



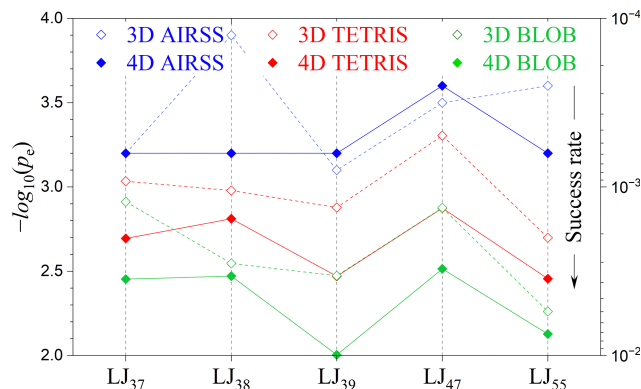


Fig. 3 Success rates of finding global minima for LJ clusters using different structure generation protocols (BLOB versus TETRIS) and optimization methods (3D versus 4D). The AIRSS results were taken from ref. 58. Each run contained between 10^4 and 2.5×10^5 structures that ensured the location of the ground state at least 100 times. Lower $-\log_{10} p_e$ values mean higher success rates.

optimization to five LJ clusters that feature ground states with different point group symmetries: 37 (C_1), 38 (O_h), 39 (C_{5v}), 47 (C_1), and 55 (I_h). These sizes were selected due to their non-trivial geometries, such as LJ_{38} known for its double-funnel energy landscape.⁷⁸ Fig. 3 shows the success rates of global minimum searches using various structure generation protocols and optimization strategies. To ensure statistical significance and direct comparison with the original GOSH study, we generated sufficient structure sets, up to 2.5×10^5 per size, to reach 100 ground state hits and picked the highest reported success rates obtained with AIRSS for packing fractions ranging between 0.1 and 0.3, as denser packing was determined to be not universally beneficial for locating ground states.⁵⁸ In fact, the consistently high compressions achieved with our TETRIS generation scheme (Fig. S3⁶⁷) were apparently detrimental for all considered sizes, as starting configurations constructed with the BLOB protocol converged to global minima ~ 2.3 times more frequently.

In contrast to the reported GOSH results, the addition of the extra coordinate through MAP0 improved the convergence across all considered sizes and seed generation strategies, effectively doubling the average probabilities of finding the LJ global minima. Unsurprisingly, the highest success rates, up to 1 in 100 tries, were observed for LJ_{39} and LJ_{55} with the most symmetric global minimum configurations. Histograms of local minima energies collected in Fig. S4⁶⁷ establish that the extra dimension helps LJ clusters transform into more favorable configurations not just near the bottom of the potential energy surface but across the full landscape. The presented findings for different sizes and structure initializations demonstrate that systematic gains in hyperspatial optimization of LJ clusters can be attained through relatively modest ± 1 Å maximum displacements of preconstructed 3D configurations into the extra dimension.

The net performance should be assessed by taking into account the increased computational cost of the introduced

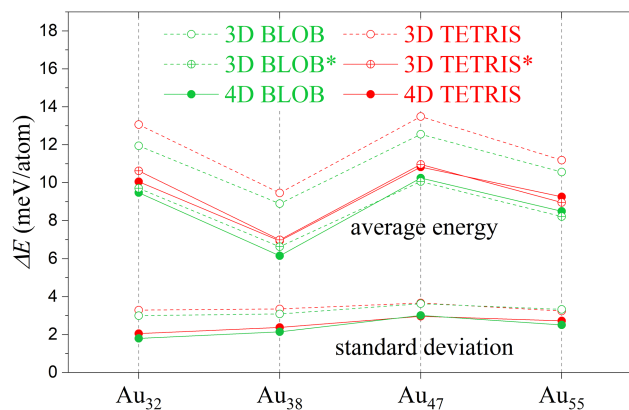


Fig. 4 Average energies (between 6 and 14 meV per atom) and standard deviations (between 1.8 and 3.6 meV per atom) of the best 2000 structures relative to the ground states for Au clusters modeled with NNPs. Each set generated with different protocols (random versus TETRIS) and optimized with different methods (3D versus 4D) consisted of 5000 structures. The full ΔE histograms for Au_{38} and Au_{55} are shown in Fig. S5.⁶⁷ The asterisk denotes CPU-time-adjusted average as explained in the text.

structure generation and optimization algorithms. According to our estimates, the BLOB and TETRIS operations take one to two orders of magnitude less time than typical local optimizations with the LJ potential. However, hyperspatial optimization can lead to significant slowdowns due to the overhead introduced by the extra spatial dimension and the increased number of energy and force evaluations required for convergence. For the LJ potential, the dimensional overhead added only about 8%, while the average number of relaxation steps increased by a factor of $413/212 \approx 1.9$, offsetting but not fully negating the benefit of 4D optimization.

B. Au clusters modeled with NNPs

As the first testbed for benchmarking hyperdimensional optimization on a chemically realistic many-body system, we chose small Au clusters described by a NNP. Nanosized Au has a complex potential energy surface due to relativistic effects and strong d-orbital hybridization, which lead to low-symmetry ground states including planar, hollow, and amorphous morphologies.^{79–82} Accurate structure prediction is particularly relevant for nanoparticles of this noble metal, as their rich catalytic properties are known to be highly sensitive to geometry.^{83,84} Extensive tests in our previous study established that the conventional Gupta potential or embedded atom model are inadequate for guiding *ab initio* structure searches, with Gupta-potential minima dispersing by nearly 30 meV per atom upon DFT relaxations.⁵¹ In contrast, our NNP trained to 6.5 meV per atom accuracy enabled refinement of ground-state candidates across the $30 \leq N \leq 80$ range and helped identify a more stable amorphous configuration for Au_{55} .

Since NNPs are roughly two orders of magnitude more expensive than the LJ potential and their ground states are subject to particular parametrization, we adopted a different benchmark strategy for Au clusters. Rather than measuring success rates based on the successful identification of global



minima, we generated 5000 structures for each cluster size and method and calculated the NN average energy of the lowest 40% (2000) minima relative to the best NN-relaxed structure obtained in this dataset. This metric reflects the sampling efficiency of each method while remaining robust to uncertainty in the true ground-state configurations. For example, the combined 2×10^4 tries for each size summarized in Fig. 4, reproduced the putative NN ground state for Au₄₇, converged to suboptimal minima for Au₃₂ and Au₅₅, and identified a more favorable configuration, by ~ 1.7 meV per atom, for Au₃₈ compared to the prior multi-tribe evolutionary searches with 2.5×10^4 local optimizations.⁵¹

Across all tested cluster sizes and initialization strategies, 4D optimization yielded average energy reductions of approximately 2.5 meV per atom compared to 3D relaxation. However, this gain is counterbalanced by an average 1.63-fold increase in the number of iterations leading to an average 1.78-fold increase in total wall time per relaxation. To account for this difference in computational cost, we reevaluated the effective sampling performance by proportionally reducing the size of the low-energy pool used for averaging. Namely, we adjusted the 3D data by computing the average over the lowest 2000/1.78 structures instead of 2000. When this correction is applied, the apparent advantage of 4D optimization largely vanishes, indicating that the observed improvements stem from extended local relaxation rather than more efficient exploration.

C. Cu–Pd–Ag clusters modeled with NNPs

Nanoalloys display a broader spectrum of tunable catalytic, optical, and magnetic properties due to their diverse composition-dependent core-shell, Janus, and mixed-species arrangements. To evaluate the performance of the hyperdimensional and TETRIS methods in identifying the most favorable configurations of multicomponent nanoparticles, we focused on 50-atom binary and ternary Cu–Pd–Ag clusters. Our previous investigation of the system with the evolutionary algorithm coupled with NNPs revealed a highly frustrated energy landscape across the composition range, making it a challenging benchmark for global optimization.⁴⁹ The reported most favorable Cu–Pd–Ag configurations have core-shell morphologies, with the larger metals preferentially residing on the surface (see Fig. 5).

Fig. 5 illustrates how the choice of local optimization strategy and initial population seeding influences search efficiency. All starting structures, 5000 for each data point, were generated using the TETRIS approach, but species assignment during cluster assembly followed three different schemes. In the unbiased case, the order in which atoms of different species were added was completely random (*e.g.*, BACABC for A₂B₂C₂). In the random core variant, atoms were grouped by species, but the sequence of species was randomized (*e.g.*, CCAABB). In the select core protocol, species were ordered by atomic size, from smallest to largest (*i.e.*, AABBC if $r_A < r_B < r_C$). Since the introduction of extraspatial coordinates introduces additional randomness to the atomic distributions, the resulting nine sets comprised 5000 different initial structures for each

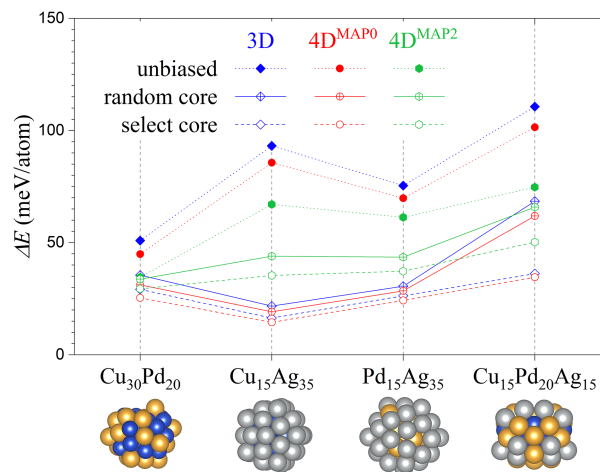


Fig. 5 Average NN energies of the best 2000 out of 5000 structures relative to the ground states for 50-atom clusters found in ref. 49 and displayed with blue (Cu), yellow (Pd), and silver (Ag) spheres. The solid, crossed, and open symbols correspond to clusters generated with TETRIS using different species sequences as explained in the text. Visual representations of atomic structures were created with VESTA.⁸⁵

composition. We chose to average the energies over the best 2000 minima, *i.e.*, 40% of each set, to clarify the performance of the random core generation for binary or ternary nanoalloys which have $1/(2!) = 50\%$ or $1/(3!) \approx 17\%$ chances of having the optimal core-shell sequences.

In unbiased searches, 4D optimizations consistently produced lower-energy structures than standard 3D relaxations across all tested nanoalloys. The average energies above the global minima decreased by $\sim 10\%$ with MAP0 and $\sim 30\%$ for MAP2, indicating that the added dimension, with larger average displacements in the latter case, allows the system to escape inefficient local geometries during relaxation. As in the Au case, the additional computational overhead, now measured at 2.24 for MAP0 and 3.25 for MAP2, diminishes these gains but the full factors behind this improvement become clearer when considered alongside the results with the other two structure initialization protocols. Growing clusters by grouping incoming atoms by species, either randomly or selectively, dramatically improved the search outcomes for both 3D and 4D optimizations. The moderate randomization of the extra dimensional coordinate up to ± 1 Å in MAP0 continued to provide a $\sim 10\%$ improvement, while the cluster ‘spherization’ in all dimensions in MAP2 likely promoted a certain degree of species mixing and led to less favorable minima than in the standard 3D case. The random and select core schemes yielded fairly similar results for binary clusters, which is consistent with the 50% probability of obtaining the correct species sequence when two species are randomly ordered. In contrast, the difference was more pronounced for ternary clusters, where the chance of realizing the correct sequence under random grouping drops to about 17%. Compared to our previous evolutionary searches with 5×10^4 local optimizations per nanoalloy starting with unbiased TETRIS configurations,⁴⁹ the present local relaxations totaling 4.5×10^4 tries yielded



comparable best minima for binary clusters, 0.1 meV per atom higher for $\text{Cu}_{30}\text{Pd}_{20}$, the same for $\text{Cu}_{15}\text{Ag}_{35}$, and 0.3 meV per atom lower for $\text{Pd}_{15}\text{Ag}_{35}$, but a significantly less optimal minimum, by 7.5 meV per atom, for $\text{Cu}_{15}\text{Pd}_{20}\text{Ag}_{15}$.

These trends suggest that the key limitation of standard 3D relaxation lies not in geometric constraints, but in its inability to reassign species to more favorable environments. Just as in the optimization of binary chains,⁵⁸ GOSH circumvents this by effectively enabling atom reordering during relaxation, but the same benefit in the more practical case of multicomponent nanoparticles can be attained more directly through incorporation of chemical knowledge in the initial population. These observations align with alternative efforts in the field to address compositional frustration in multicomponent clusters. Several methods have been developed to enable or approximate species swaps during optimization, including identity-exchange Monte Carlo steps in basin-hopping,⁷⁷ symmetry-constrained genetic algorithms,⁸⁶ and multi-agent search protocols with specialized roles such as the Flying-Landing-Hiking framework.⁸⁷

D. Li-Sn, Pd-Sn, and Ag-Sn modeled with NNPs

To test the performance of global optimization in periodic systems, we selected M-Sn intermetallics known to display a wide range of morphologies. As a group XIV element capable of covalent bonding, Sn readily forms directional extended frameworks and cage-like motifs, especially when alloyed with electropositive metals.^{54,55,88–90} For example, Sn-rich alloys can be represented with particular sequences of Kepler nets, while M-rich intermetallics often appear as simple decorations of common metallic lattices with isolated or only weakly connected Sn atoms. Our recent NNP-guided exploration of M-Sn binaries (M = Li, Na, Ca, Cu, Pd, and Ag) led to the identification of over 30 new phases stable under different (T, P) conditions.^{54,55} They include $\text{tI}36\text{-PdSn}_2$ with an $\text{A}^+\text{B}^0\text{A}^-\text{A}^0\text{B}^0\text{A}^+$ stacking of Pd atoms between Sn layers with rotated squares and $\text{hR}75\text{-Li}_{19}\text{Sn}_6$ with a^{34,54} stacking of layered units defining the bcc decoration.

We focused on Li_7Sn_9 , Li_9Sn_6 , $\text{Pd}_4\text{Sn}_{12}$, $\text{Pd}_{14}\text{Sn}_2$, and $\text{Ag}_{14}\text{Sn}_4$ compositions determined in our prior evolutionary searches to have complex ground states with sizes between 15 and 18 atoms per unit cell. To gauge the capability of the TETRIS method, we tried only broadly applicable structural fragments, such as M-M, M-Sn, and Sn-Sn dimers or bcc blocks, avoiding overly material-dependent structural motifs that could bias the searches toward previously known solutions. Our tests (not shown) demonstrated little advantage of block seeding over random structure generation, and we chose the latter in our investigation of the hyperspatial optimization.

Fig. 6 compares the average energies of the 2000 lowest structures out of 5000 relative to the ground state, evaluated by neural network potentials. Unlike in nanoalloys, 4D optimization did not yield systematic improvements in periodic crystals. Both 3D and 4D relaxations sampled similar energy distributions across the lowest 40% of minima and delivered comparable success in locating the lowest-energy states identified in earlier evolutionary runs. For Li_7Sn_9 and $\text{Pd}_{14}\text{Sn}_2$, both

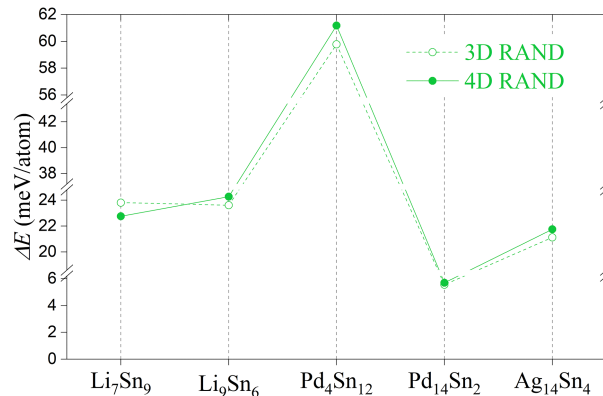


Fig. 6 Average energies of the best 2000 structures relative to the ground states for Li-Sn, Pd-Sn, and Ag-Sn crystals modeled with NNPs. Each set optimized with different methods (3D and 4D) consisted of 5000 structures.

methods repeatedly recovered the ground state. For $\text{Pd}_4\text{Sn}_{12}$, only 4D found the ground state, with 3D reaching a structure 2 meV per atom higher. For $\text{Ag}_{14}\text{Sn}_4$, the ground state was found only by 3D, with the closest 4D result 0.5 meV per atom less stable. The lack of significant advantage observed for 4D relaxation in these periodic systems may result from the indirect constraints imposed by lattice vectors, which evolve primarily in response to effective modulation of three-dimensional interatomic distances by the additional hyperspatial coordinates.

E. LiB_2C_2 and CaB_6 modeled with DFT

The usefulness of the TETRIS adaptation to periodic systems can be illustrated with *ab initio* searches for ground states in complex covalent materials, such as metal borides and borocarbides known to display a variety of 2D and 3D frameworks.^{91,92} To collect sufficient data for benchmarking different protocols at the DFT level, we generated 1000 or 2000 medium-sized configurations and assessed each variant's performance by calculating the average energy of the best 2% of locally optimized candidates. The dataset sizes were chosen to approximate fractions of random motifs injected into population during typical evolutionary runs totaling 4000–6000 local optimizations.^{19,30} Our recent screening of layered metal borocarbides for conventional high- T_c superconductors synthesizable under ambient pressure revealed that high levels of hole doping destabilize the honeycomb layer morphology in M_xBC (M = Li, Na, or Mg) essential for strong quasi-2D electron-phonon coupling.^{60,93} At the $\text{Li}_{1/2}\text{BC}$ composition, our evolutionary searches identified an $\text{mS}20$ (Cm) structure with a 3D framework ~ 70 meV per atom below the best $\text{oI}10$ ($Imm2$) layered counterpart.

We performed our present tests for $\text{Li}_3\text{B}_6\text{C}_6$ and referenced the resulting 8×1000 structures to the $\text{mS}15$ (Pm) polymorph with 3D morphology determined previously to have the lowest energy among 15-atom unit cells. Fig. 7 shows the effectiveness of generating low-energy motifs as a function of the introduced bias. For each protocol, we checked whether the additional



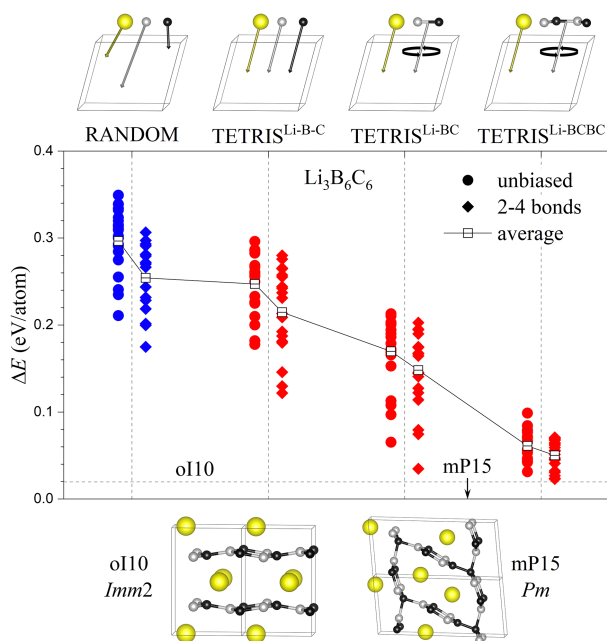


Fig. 7 Performance of random and TETRIS structure generation methods in the search for most favorable $\text{Li}_3\text{B}_6\text{C}_6$ configurations at the DFT level. The number and types of user-defined TETRIS blocks per set are indicated at the top. The sets denoted with diamonds were generated with an additional constraint for B and C atoms to have 2–4 neighbors within 1.9 Å. The energy is referenced to the most favorable mP15- $\text{Li}_3\text{B}_6\text{C}_6$ phase with 3 formula units found in our previous study,⁶⁰ with the dashed line corresponding the lowest-energy layered configuration.

condition on the number of nearest neighbors, 3–4 within 1.9 Å around B/C atoms expected to be in the sp^2 or sp^3 environments, leads to better starting configurations. The constraint indeed consistently lowered the average energy, with the largest improvement of ~ 40 meV per atom noted for the random scheme. The three TETRIS sets with different numbers of atomic blocks may help find a desired balance for bias in structure generation. A simple switch from a random placement of atoms within a unit cell to the TETRIS packing of individual atoms resulted in a noticeable drop of the average energy by ~ 50 meV per atom. The introduction of B and C atoms in pairs to promote a natural alternating order in heteronuclear frameworks further lowered the average energy by ~ 75 meV per atom. The larger four-atom preassembled units link into honeycomb layers far more efficiently while not prohibiting interlayer bridging, as we observed low-energy configurations with sp^3 -bonded C atoms.

Investigation of CaB_6 under pressure provides information on the methods' performance for non-layered configurations. Our previous studies revealed exceptional complexity of pressure-induced structural transformations in this compound, which starts as a simple cP7 at ambient pressure, becomes dynamically unstable above 23 GPa, and displays a variety of more stable exotic 3D frameworks above 13 GPa.^{19,20} The concurrent synthesis work indicated the formation of a new phase above 31 GPa but its structure could not be established from the collected XRD data and was solved with evolutionary

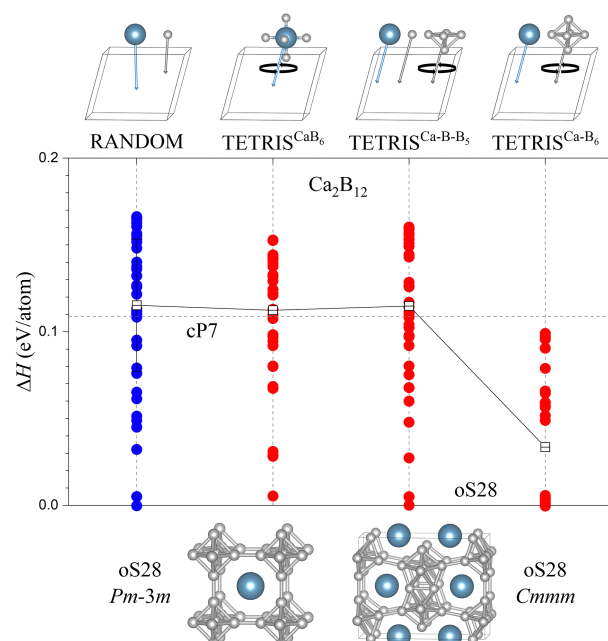


Fig. 8 Performance of random and TETRIS structure generation methods in the search for most favorable Ca_2B_{12} configurations at 50 GPa at the DFT level. The number and types of TETRIS blocks are indicated at the top. The enthalpy is referenced to the most favorable oS28- Ca_2B_{12} phase found in our previous study,¹⁹ with the dashed line corresponding to the cP7 ground state at ambient pressure.

searches without any structural input.¹⁹ Our benchmark tests demonstrated that the construction of the 28-atom tI56 ground state could not be accomplished with unbiased random searches in 10 000 tries and required either randomization of the cP7 supercells or use of the evolutionary algorithm.

Here, we considered smaller 14-atom Ca_2B_{12} unit cells at 50 GPa, determined to have oS28 ($Pm\bar{3}m$) as the lowest-enthalpy configuration with up to two formula units,¹⁹ and allowed an additional 3D random rotation of the TETRIS blocks with quaternions in their starting position. Fig. 8 displays the average enthalpy of the best 40 structures and shows that 2000 tries were sufficient to locate oS28 with each protocol. In fact, the TETRIS algorithm offered no improvement over the random method when we used blocks with one-three B atoms, a preassembled unit with a uniform distribution of six B atoms around Ca, or a five-atom fragment of the B octahedron (only the last two sets are shown). An appreciable enthalpy drop of nearly 80 meV per atom was observed only when full octahedron units were used.

We would like to stress that we view the TETRIS generation not as a stand-alone optimization method but rather as a protocol for seeding global structure searches with plausible configurations. The presented results of *ab initio* searches demonstrate that steering global exploration toward layered configurations can be accomplished naturally with minimal bias, such as heteronuclear dimers. Construction of complex 3D frameworks may require the use of more specific building blocks favorable for this material's class but the TETRIS



packing is not constrained to supercells of known prototypes to generate periodic configurations.

IV. Summary

We present a systematic extension and examination of two complementary strategies for improving global minima searches in nanoscale and crystalline systems: structure generation with physically biased seeding and hyperspatial optimization. To enable treatment of chemically realistic systems with a variety of structure initialization and minimization schemes, we expanded Behler–Parrinello-type NNPs to extra dimensions and reformulated several elements of the GOSH algorithm. Specifically, we introduced two coordinate mapping schemes, MAP0 and MAP2, that displace physically meaningful configurations constructed in the normal 3D space into extra dimensions using distinct extension strategies. MAP2 mimics the original GOSH approach by generating randomized displacements that effectively treat all spatial directions equally, producing near-hyperspherical distributions. In contrast, MAP0 introduces moderate, uncorrelated shifts of up to ± 1 Å into the extra dimensions. Additionally, we adjusted the protocol for changing the spring constant associated with extraspatial coordinates to plateau-wise ramp-up during local optimizations, enabling the use of advanced minimization algorithms, such as BFGS2.

Our benchmark tests showed that both the modified mapping scheme and enhanced relaxation protocol were essential for attaining consistent gains in complex systems. In multi-component Cu–Pd–Ag clusters initialized with unbiased species ordering, 4D optimization with the conservative MAP0 approach reduced the average energy above the global minimum by $\sim 10\%$, while the more aggressive MAP2 scheme yielded up to $\sim 30\%$ reductions, highlighting the benefit of using extra dimensions to facilitate atom swaps. However, a chemically informed TETRIS initialization, which assembles clusters with species grouped by atomic size to favor core–shell formation, led to significantly better convergence even with standard 3D optimization, and performed even more effectively when paired with MAP0. This balance between geometric flexibility and controlled species placement was especially important for ternary systems, where the probability of randomly assembling optimal core–shell configurations is inherently low. For bulk alloys, these factors had little impact, as neither hyperspatial optimization nor species ordering strategies produced measurable gains. In contrast, TETRIS-based seeding, by assembling unit cells from chemically sensible atomic blocks, substantially improved convergence in covalently bonded systems modeled with DFT.

This study provides the first evidence that hyperspatial optimization, previously applied to simple potential energy surfaces,^{56–58} can yield benefits in systems modeled with sophisticated many-body potentials. By independently varying the coordinate mapping and initialization strategies, we were able to disentangle the contributions of geometric relaxation flexibility and species ordering, a distinction particularly

important for nanoalloys. These insights offer a path toward improving hyperspatial optimization across a broader range of materials systems.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting the findings of this study are available within the article and its supplementary information (SI). Supplementary information is available. See DOI: <https://doi.org/10.1039/d5cp02412a>.

Additional data are available from the corresponding author upon request.

Acknowledgements

The authors acknowledge support from the National Science Foundation (NSF) (Award No. DMR-2320073). This work used the Expanse system at the San Diego Supercomputer Center *via* allocation TG-DMR180071. Expanse is supported by the Extreme Science and Engineering Discovery Environment (XSEDE) program⁹⁴ through NSF Award No. ACI-1548562.

References

- 1 S. M. Woodley and R. Catlow, Crystal structure prediction from first principles, *Nat. Mater.*, 2008, 7, 937–946, DOI: [10.1038/nmat2321](https://doi.org/10.1038/nmat2321).
- 2 D. J. Wales and H. A. Scheraga, Global optimization of clusters, crystals, and biomolecules, *Science*, 1999, 285, 1368–1372, DOI: [10.1126/science.285.5432.1368](https://doi.org/10.1126/science.285.5432.1368).
- 3 E. Zurek, Discovering new materials *via* A Priori crystal structure prediction, *Rev. Comput. Chem.*, 2016, 29, 274–326, DOI: [10.1002/9781119148739.ch5](https://doi.org/10.1002/9781119148739.ch5).
- 4 A. R. Oganov and C. W. Glass, Crystal structure prediction using ab initio evolutionary techniques: Principles and applications, *J. Chem. Phys.*, 2006, 124, 244704, DOI: [10.1063/1.2210932](https://doi.org/10.1063/1.2210932).
- 5 R. J. Needs and C. J. Pickard, Perspective: Role of structure prediction in materials discovery and design, *APL Mater.*, 2016, 4, 053210, DOI: [10.1063/1.4949361](https://doi.org/10.1063/1.4949361).
- 6 A. R. Oganov, C. J. Pickard, Q. Zhu and R. J. Needs, Structure prediction drives materials discovery, *Nat. Rev. Mater.*, 2019, 4, 331–348, DOI: [10.1038/s41578-019-0101-8](https://doi.org/10.1038/s41578-019-0101-8).
- 7 S. M. Woodley, G. M. Day and R. Catlow, Structure prediction of crystals, surfaces and nanoparticles, *Philos. Trans. R. Soc., A*, 2020, 378, 20190600, DOI: [10.1098/rsta.2019.0600](https://doi.org/10.1098/rsta.2019.0600).
- 8 A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards and S. Dacek, *et al.*, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater.*, 2013, 1, 011002, DOI: [10.1063/1.4812323](https://doi.org/10.1063/1.4812323).



- 9 S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito and O. Levy, The high-throughput highway to computational materials design, *Nat. Mater.*, 2013, **12**, 191–201, DOI: [10.1038/nmat3568](https://doi.org/10.1038/nmat3568).
- 10 J. E. Saal, S. Kirklin, M. Aykol, B. Meredig and C. Wolverton, Materials design and discovery with High-Throughput density functional theory: The open quantum materials database (OQMD), *JOM*, 2013, **65**, 1501–1509, DOI: [10.1007/s11837-013-0755-4](https://doi.org/10.1007/s11837-013-0755-4).
- 11 G. Bergerhoff and I. Brown, *Crystallographic databases*, ed. F. H. Allen, *et al.*, International Union of Crystallography, Chester, 1987, DOI: [10.1063/1.4812323](https://doi.org/10.1063/1.4812323).
- 12 X. Dong, A. R. Oganov, A. F. Goncharov, E. Stavrou, S. Lobanov and G. Saleh, *et al.*, A stable compound of helium and sodium at high pressure, *Nat. Chem.*, 2017, **9**, 440–445, DOI: [10.1038/nchem.2716](https://doi.org/10.1038/nchem.2716).
- 13 A. R. Oganov, J. Chen, C. Gatti, Y. Ma, Y. Ma and C. W. Glass, *et al.*, Ionic high-pressure form of elemental boron, *Nature*, 2009, **457**, 863–867, DOI: [10.1038/nature07736](https://doi.org/10.1038/nature07736).
- 14 A. N. Kolmogorov and S. Curtarolo, Prediction of different crystal structure phases in metal borides: A lithium monoboride analog to MgB₂, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2006, **73**, 180501, DOI: [10.1103/PhysRevB.73.180501](https://doi.org/10.1103/PhysRevB.73.180501).
- 15 A. N. Kolmogorov, S. Hajinazar, C. Angyal, V. L. Kuznetsov and A. P. Jephcoat, Synthesis of a predicted layered lib *via* cold compression, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2015, **92**, 144110, DOI: [10.1103/PhysRevB.92.144110](https://doi.org/10.1103/PhysRevB.92.144110).
- 16 A. N. Kolmogorov, S. Shah, E. R. Margine, A. F. Bialon, T. Hammerschmidt and R. Drautz, New superconducting and semiconducting fe-b compounds predicted with an ab initio evolutionary search, *Phys. Rev. Lett.*, 2010, **105**, 217003, DOI: [10.1103/PhysRevLett.105.217003](https://doi.org/10.1103/PhysRevLett.105.217003).
- 17 A. F. Bialon, T. Hammerschmidt, R. Drautz, S. Shah, E. R. Margine and A. N. Kolmogorov, Possible routes for synthesis of new boron-rich Fe–B and Fe_{1–x}Cr_xB₄ compounds, *Appl. Phys. Lett.*, 2011, **98**, 081901, DOI: [10.1063/1.3556564](https://doi.org/10.1063/1.3556564).
- 18 A. Van Der Geest and A. Kolmogorov, Stability of 41 metal–boron systems at 0 GPa and 30 GPa from first principles, *Calphad*, 2014, **46**, 184–204, DOI: [10.1016/j.calphad.2014.03.005](https://doi.org/10.1016/j.calphad.2014.03.005).
- 19 A. N. Kolmogorov, S. Shah, E. R. Margine, A. K. Kleppe and A. P. Jephcoat, Pressure-Driven Evolution of the Covalent Network in CaB₆, *Phys. Rev. Lett.*, 2012, **109**, 075501, DOI: [10.1103/physrevlett.109.075501](https://doi.org/10.1103/physrevlett.109.075501).
- 20 S. Shah and A. N. Kolmogorov, Stability and superconductivity of Ca–B phases at ambient and high pressure, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2013, **88**, 014107, DOI: [10.1103/physrevb.88.014107](https://doi.org/10.1103/physrevb.88.014107).
- 21 S. K. Choi, R. Coldea, A. N. Kolmogorov, T. Lancaster, I. I. Mazin and S. J. Blundell, *et al.*, Spin Waves and Revised Crystal Structure of Honeycomb Iridate Na₂IrO₃, *Phys. Rev. Lett.*, 2012, **108**, 127204, DOI: [10.1103/PhysRevLett.108.127204](https://doi.org/10.1103/PhysRevLett.108.127204).
- 22 F. Sun, H. Zheng, Y. Liu, E. D. Sandoval, C. Xu and J. Xu, *et al.*, Electronic and structural response to pressure in the hyperkagome-lattice Na₃Ir₃O₈, *Phys. Rev. B*, 2018, **98**, 085131, DOI: [10.1103/PhysRevB.98.085131](https://doi.org/10.1103/PhysRevB.98.085131).
- 23 G. Fabbri, A. Thorn, W. Bi, M. Abramchuk, F. Bahrami and J. H. Kim, *et al.*, Complex pressure-temperature structural phase diagram of the honeycomb iridate Cu₂IrO₃, *Phys. Rev. B*, 2021, **104**, 014102, DOI: [10.1103/PhysRevB.104.014102](https://doi.org/10.1103/PhysRevB.104.014102).
- 24 C. J. Pickard and R. J. Needs, Structures at high pressure from random searching, *Phys. Status Solidi B*, 2009, **246**, 536–540, DOI: [10.1002/pssb.200880546](https://doi.org/10.1002/pssb.200880546).
- 25 C. J. Pickard, Ephemeral data derived potentials for random structure search, *Phys. Rev. B*, 2022, **106**, 014102, DOI: [10.1103/PhysRevB.106.014102](https://doi.org/10.1103/PhysRevB.106.014102).
- 26 Y. Li, L. Wang, H. Liu, Y. Zhang, J. Hao and C. J. Pickard, *et al.*, Dissociation products and structures of solid h₂S at strong compression, *Phys. Rev. B*, 2016, **93**, 020103, DOI: [10.1103/PhysRevB.93.020103](https://doi.org/10.1103/PhysRevB.93.020103).
- 27 D. C. Lonie and E. Zurek, Xtalopt: An open-source evolutionary algorithm for crystal structure prediction, *Comput. Phys. Commun.*, 2011, **182**, 372–387, DOI: [10.1016/j.cpc.2010.07.048](https://doi.org/10.1016/j.cpc.2010.07.048).
- 28 Y. Wang, J. Lv, L. Zhu and Y. Ma, Calypso: A method for crystal structure prediction, *Comput. Phys. Commun.*, 2012, **183**, 2063–2070, DOI: [10.1016/j.cpc.2012.05.008](https://doi.org/10.1016/j.cpc.2012.05.008).
- 29 C. J. Pickard and R. J. Needs, Ab initio random structure searching, *J. Phys.:Condens. Matter*, 2011, **23**, 053201, DOI: [10.1088/0953-8984/23/5/053201](https://doi.org/10.1088/0953-8984/23/5/053201).
- 30 S. Hajinazar, A. Thorn, E. D. Sandoval, S. Kharabadze and A. N. Kolmogorov, MAISE: Construction of neural network interatomic models and evolutionary structure optimization, *Comput. Phys. Commun.*, 2021, **259**, 107679, DOI: [10.1016/j.cpc.2020.107679](https://doi.org/10.1016/j.cpc.2020.107679).
- 31 D. Wolpert and W. Macready, No free lunch theorems for optimization, *IEEE Trans. Evol. Comput.*, 1997, **1**, 67–82, DOI: [10.1109/4235.585893](https://doi.org/10.1109/4235.585893).
- 32 C. Wolverton and A. Zunger, Prediction of li intercalation and battery voltages in layered vs. cubic lix coo₂, *J. Electrochem. Soc.*, 1998, **145**, 2424, DOI: [10.1149/1.1838653](https://doi.org/10.1149/1.1838653).
- 33 T. Mueller and G. Ceder, Bayesian approach to cluster expansions, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2009, **80**, 024103, DOI: [10.1103/PhysRevB.80.024103](https://doi.org/10.1103/PhysRevB.80.024103).
- 34 G. R. Desiraju, Crystal Engineering: From Molecule to Crystal, *J. Am. Chem. Soc.*, 2013, **135**, 9952–9967, DOI: [10.1021/ja403264c](https://doi.org/10.1021/ja403264c).
- 35 L. M. Hunnisett, *et al.*, The seventh blind test of crystal structure prediction: structure generation methods, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2024, **80**, 517–547, DOI: [10.1107/S2052520624007492](https://doi.org/10.1107/S2052520624007492).
- 36 R. L. Johnston, Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries, *Dalton Trans.*, 2003, 4193–4207, DOI: [10.1039/B305686D](https://doi.org/10.1039/B305686D).
- 37 R. Ferrando, J. Jellinek and R. L. Johnston, Nanoalloys: from theory to applications of alloy clusters and nanoparticles, *Chem. Rev.*, 2008, **108**, 845–910, DOI: [10.1021/cr040090g](https://doi.org/10.1021/cr040090g).
- 38 S. M. Woodley, P. D. Battle, J. D. Gale and C. R. A. Catlow, Prediction of inorganic crystal framework structures part 1: Using a genetic algorithm and an indirect approach to exclusion zones, *Phys. Chem. Chem. Phys.*, 2004, **6**, 1815–1822, DOI: [10.1039/B315066F](https://doi.org/10.1039/B315066F).



- 39 S. M. Woodley, Prediction of inorganic crystal framework structures part 2: Using a genetic algorithm and a direct approach to exclusion zones, *Phys. Chem. Chem. Phys.*, 2004, **6**, 1823–1829, DOI: [10.1039/B315216M](https://doi.org/10.1039/B315216M).
- 40 P. Avery and E. Zurek, Randspg: An open-source program for generating atomistic crystal structures with specific space-groups, *Comput. Phys. Commun.*, 2017, **213**, 208–216, DOI: [10.17632/v2nzgmpd37.1](https://doi.org/10.17632/v2nzgmpd37.1).
- 41 S. Fredericks, K. Parrish, D. Sayre and Q. Zhu, Pyxtal: A python library for crystal structure generation and symmetry analysis, *Comput. Phys. Commun.*, 2021, **261**, 107810, DOI: [10.1016/j.cpc.2020.107810](https://doi.org/10.1016/j.cpc.2020.107810).
- 42 D. J. Wales and J. P. K. Doye, Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms, *J. Phys. Chem. A*, 1997, **101**, 5111–5116, DOI: [10.1021/jp970984n](https://doi.org/10.1021/jp970984n).
- 43 Y. Wang, J. Lv, L. Zhu and Y. Ma, Crystal structure prediction via particle-swarm optimization, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2010, **82**, 094116, DOI: [10.1103/PhysRevB.82.094116](https://doi.org/10.1103/PhysRevB.82.094116).
- 44 D. Daven, N. Tit, J. Morris and K. Ho, Structural optimization of Lennard-Jones clusters by a genetic algorithm, *Chem. Phys. Lett.*, 1996, **256**, 195–200, DOI: [10.1016/0009-2614\(96\)00406-X](https://doi.org/10.1016/0009-2614(96)00406-X).
- 45 S. M. Woodley, A. A. Sokol and C. R. A. Catlow, Structure prediction of inorganic nanoparticles with predefined architecture using a genetic algorithm, *Z. Anorg. Allg. Chem.*, 2004, **630**, 2343–2353, DOI: [10.1002/zaac.200400338](https://doi.org/10.1002/zaac.200400338).
- 46 G. Trimarchi, A. J. Freeman and A. Zunger, Predicting stable stoichiometries of compounds via evolutionary global space-group optimization, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2009, **80**, 092101, DOI: [10.1103/PhysRevB.80.092101](https://doi.org/10.1103/PhysRevB.80.092101).
- 47 Q. Zhu, V. Sharma, A. R. Oganov and R. Ramprasad, Predicting polymeric crystal structures by evolutionary algorithms, *J. Chem. Phys.*, 2014, **141**, 154102, DOI: [10.1063/1.4897337](https://doi.org/10.1063/1.4897337).
- 48 S. V. Lepeshkin, V. S. Baturin, Y. A. Uspenskii and A. R. Oganov, Method for Simultaneous Prediction of Atomic Structure and Stability of Nanoclusters in a Wide Area of Compositions, *J. Phys. Chem. Lett.*, 2019, **10**, 102–106, DOI: [10.1021/acs.jpcclett.8b03510](https://doi.org/10.1021/acs.jpcclett.8b03510).
- 49 S. Hajinazar, E. D. Sandoval, A. J. Cullo and A. N. Kolmogorov, Multitribe evolutionary search for stable Cu–Pd–Ag nanoparticles using neural network models, *Phys. Chem. Chem. Phys.*, 2019, **21**, 8729–8742, DOI: [10.1039/c9cp00837c](https://doi.org/10.1039/c9cp00837c).
- 50 S. Tao, X. Shao and L. Zhu, Accelerating Structural Optimization through Fingerprinting Space Integration on the Potential Energy Surface, *J. Phys. Chem. Lett.*, 2024, **15**, 3185–3190, DOI: [10.1021/acs.jpcclett.4c00275](https://doi.org/10.1021/acs.jpcclett.4c00275).
- 51 A. Thorn, J. Rojas-Nunez, S. Hajinazar, S. E. Baltazar and A. N. Kolmogorov, Toward ab Initio Ground States of Gold Clusters via Neural Network Modeling, *J. Phys. Chem. C*, 2019, **123**, 30088–30098, DOI: [10.1021/acs.jpcc.9b08517](https://doi.org/10.1021/acs.jpcc.9b08517).
- 52 J. Behler, Perspective: Machine learning potentials for atomistic simulations, *J. Chem. Phys.*, 2016, **145**, 170901, DOI: [10.1063/1.4966192](https://doi.org/10.1063/1.4966192).
- 53 A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode and G. Csányi, *et al.*, Machine learning unifies the modeling of materials and molecules, *Sci. Adv.*, 2017, **3**, e1701816, DOI: [10.1126/sciadv.1701816](https://doi.org/10.1126/sciadv.1701816).
- 54 S. Kharabadze, A. Thorn, E. A. Koulakova and A. N. Kolmogorov, Prediction of stable Li–Sn compounds: boosting ab initio searches with neural network potentials, *npj Comput. Mater.*, 2022, **8**, 136, DOI: [10.1038/s41524-022-00825-4](https://doi.org/10.1038/s41524-022-00825-4).
- 55 A. Thorn, D. Gochitashvili, S. Kharabadze and A. N. Kolmogorov, Machine learning search for stable binary Sn alloys with Na, Ca, Cu, Pd, and Ag, *Phys. Chem. Chem. Phys.*, 2023, **25**, 22415–22436, DOI: [10.1039/D3CP02817H](https://doi.org/10.1039/D3CP02817H).
- 56 D. B. Faken, A. F. Voter, D. L. Freeman and J. D. Doll, Dimensional strategies and the minimization problem: Barrier-avoiding algorithms, *J. Phys. Chem. A*, 1999, **103**, 9521–9526, DOI: [10.1021/jp9920949](https://doi.org/10.1021/jp9920949).
- 57 A. S. Bayden and K. D. Jordan, Use of extended dimensions in global optimization, *Chem. Phys. Lett.*, 2004, **385**, 101–104, DOI: [10.1016/j.cplett.2003.12.059](https://doi.org/10.1016/j.cplett.2003.12.059).
- 58 C. J. Pickard, Hyperspatial optimization of structures, *Phys. Rev. B*, 2019, **99**, 054102, DOI: [10.1103/PhysRevB.99.054102](https://doi.org/10.1103/PhysRevB.99.054102).
- 59 J. Behler and M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, *Phys. Rev. Lett.*, 2007, **98**, 146401, DOI: [10.1103/PhysRevLett.98.146401](https://doi.org/10.1103/PhysRevLett.98.146401).
- 60 C. R. Tomassetti, G. P. Kafle, E. T. Marcial, E. R. Margine and A. N. Kolmogorov, Prospect of high-temperature superconductivity in layered metal borocarbides, *J. Mater. Chem. C*, 2024, **12**, 4870–4884, DOI: [10.1039/D4TC00210E](https://doi.org/10.1039/D4TC00210E).
- 61 J. Barzilai and J. M. Borwein, Two-point step size gradient methods, *IMA J. Nume. Anal.*, 1988, **8**, 141–148, DOI: [10.1093/imanum/8.1.141](https://doi.org/10.1093/imanum/8.1.141).
- 62 M. Galassi, *et al.*, *Gnu scientific library reference manual*, 2018, <https://www.gnu.org/software/gsl/>.
- 63 R. Fletcher and C. M. Reeves, Function minimization by conjugate gradients, *Comput. J.*, 1964, **7**, 149–154, DOI: [10.1093/comjnl/7.2.149](https://doi.org/10.1093/comjnl/7.2.149).
- 64 R. Fletcher, A new approach to variable metric algorithms, *Comput. J.*, 1970, **13**, 317–322, DOI: [10.1093/comjnl/13.3.317](https://doi.org/10.1093/comjnl/13.3.317).
- 65 R. Klessig and E. Polak, Efficient implementations of the polak–ribière conjugate gradient algorithm, *SIAM J. Control*, 1972, **10**, 524–549, DOI: [10.1137/0310040](https://doi.org/10.1137/0310040).
- 66 J. Fliege and B. F. Svaiter, Steepest descent methods for multicriteria optimization, *Math. Methods Oper. Res.*, 2000, **51**, 479–494, DOI: [10.1007/s001860000043](https://doi.org/10.1007/s001860000043).
- 67 See SI for Notes I and II, Fig. S1–S5. A description for each figure is given in the body of the paper, wherever ref. 67 is cited.
- 68 S. Chiriki, S. Jindal and S. S. Bulusu, Neural network potentials for dynamics and thermodynamics of gold nanoparticles, *J. Chem. Phys.*, 2017, **146**, 084314, DOI: [10.1063/1.4977050](https://doi.org/10.1063/1.4977050).
- 69 G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis



- set, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1996, **54**, 11169, DOI: [10.1103/PhysRevB.54.11169](https://doi.org/10.1103/PhysRevB.54.11169).
- 70 P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1994, **50**, 17953, DOI: [10.1103/PhysRevB.50.17953](https://doi.org/10.1103/PhysRevB.50.17953).
- 71 J. Klimeš, D. R. Bowler and A. Michaelides, Chemical accuracy for the van der Waals density functional, *J. Phys.:Condens. Matter*, 2009, **22**, 022201, DOI: [10.1088/0953-8984/22/2/022201](https://doi.org/10.1088/0953-8984/22/2/022201).
- 72 H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, *Phys. Rev. B*, 1976, **13**, 5188, DOI: [10.1103/PhysRevB.13.5188](https://doi.org/10.1103/PhysRevB.13.5188).
- 73 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, 1996, **77**, 3865, DOI: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- 74 R. H. Leary and J. P. K. Doye, Tetrahedral global minimum for the 98-atom Lennard-Jones cluster, *Phys. Rev. E:Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1999, **60**, R6320–R6322, DOI: [10.1103/PhysRevE.60.R6320](https://doi.org/10.1103/PhysRevE.60.R6320).
- 75 R. H. Leary, Global Optimization on Funneling Landscapes, *J. Global Optim.*, 2000, **18**, 367–383, DOI: [10.1023/A:1026500301312](https://doi.org/10.1023/A:1026500301312).
- 76 R. Brüning, D. A. St-Onge, S. Patterson and W. Kob, Glass transitions in one-, two-, three-, and four-dimensional binary Lennard-Jones systems, *J. Phys.:Condens. Matter*, 2008, **21**, 035117, DOI: [10.1088/0953-8984/21/3/035117](https://doi.org/10.1088/0953-8984/21/3/035117).
- 77 G. G. Rondina and J. L. F. Da Silva, Revised Basin-Hopping Monte Carlo Algorithm for Structure Optimization of Clusters and Nanoparticles, *J. Chem. Inf. Model.*, 2013, **53**, 2282–2298, DOI: [10.1021/ci400224z](https://doi.org/10.1021/ci400224z).
- 78 J. P. K. Doye, M. A. Miller and D. J. Wales, Evolution of the potential energy surface with size for Lennard-Jones clusters, *J. Chem. Phys.*, 1999, **111**, 8417–8428, DOI: [10.1063/1.480217](https://doi.org/10.1063/1.480217).
- 79 I. L. Garzón and A. Posada-Amarillas, Structural and vibrational analysis of amorphous Au₅₅ clusters, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1996, **54**, 11796–11802, DOI: [10.1103/PhysRevB.54.11796](https://doi.org/10.1103/PhysRevB.54.11796).
- 80 M. P. Johansson, D. Sundholm and J. Vaara, Au₃₂: A 24-carat golden fullerene, *Angew. Chem., Int. Ed.*, 2004, **43**, 2678–2681, DOI: [10.1002/anie.200453986](https://doi.org/10.1002/anie.200453986).
- 81 J. Wang, J. Jellinek, J. Zhao, Z. Chen, R. B. King and P. von Ragué Schleyer, Hollow Cages versus Space-Filling Structures for Medium-Sized Gold Clusters: The Spherical Aromaticity of the Au₅₀ Cage, *J. Phys. Chem. A*, 2005, **109**, 9265–9269, DOI: [10.1021/jp052414q](https://doi.org/10.1021/jp052414q).
- 82 N. Tarrat, M. Rapacioli, J. Cuny, J. Morillo, J.-L. Heully and F. Spiegelman, Global optimization of neutral and charged 20- and 55-atom silver and gold clusters at the dftb level, *Comput. Theor. Chem.*, 2017, **1107**, 102–114, DOI: [10.1016/j.comptc.2017.01.022](https://doi.org/10.1016/j.comptc.2017.01.022). Structure prediction of nanoclusters from global optimization techniques: computational strategies.
- 83 N. Lopez, J. Nørskov, T. Janssens, A. Carlsson, A. Puig-Molina and B. S. Clausen, *et al.*, The adhesion and shape of nanosized Au particles in a Au/TiO₂ catalyst, *J. Catal.*, 2004, **225**, 86–94, DOI: [10.1016/j.jcat.2004.03.036](https://doi.org/10.1016/j.jcat.2004.03.036).
- 84 P. Suchomel, L. Kvitek, R. Prucek, A. Panacek, A. Halder and S. Vajda, *et al.*, Simple size-controlled synthesis of Au nanoparticles and their size-dependent catalytic activity, *Sci. Rep.*, 2018, **8**, 4589, DOI: [10.1038/s41598-018-22976-5](https://doi.org/10.1038/s41598-018-22976-5).
- 85 K. Momma and F. Izumi, VESTA3 for three-dimensional visualization of crystal, volumetric and morphology data, *J. Appl. Crystallogr.*, 2011, **44**, 1272–1276, DOI: [10.1107/S0021889811038970](https://doi.org/10.1107/S0021889811038970).
- 86 S. Han, G. Barcaro, A. Fortunelli, S. Lysgaard, T. Vegge and H. A. Hansen, Unfolding the structural stability of nanoalloys *via* symmetry-constrained genetic algorithm and neural network potential, *npj Comput. Mater.*, 2022, **8**, 121, DOI: [10.1038/s41524-022-00807-6](https://doi.org/10.1038/s41524-022-00807-6).
- 87 D. Rapetti, C. Roncaglia and R. Ferrando, Optimizing the Shape and Chemical Ordering of Nanoalloys with Specialized Walkers, *Adv. Theory Simul.*, 2023, **6**, 2300268, DOI: [10.1002/adts.202300268](https://doi.org/10.1002/adts.202300268).
- 88 J. Nylén, F. J. García García, B. D. Mosel, R. Pöttgen and U. Häussermann, Structural relationships, phase stability and bonding of compounds PdSn_n (*n* = 2, 3, 4), *Solid State Sci.*, 2004, **6**, 147–155, DOI: [10.1016/j.solidstatesciences.2003.09.011](https://doi.org/10.1016/j.solidstatesciences.2003.09.011).
- 89 J. Engelkemier, V. M. Berns and D. C. Fredrickson, First-principles elucidation of atomic size effects using dft-chemical pressure analysis: Origins of Ca₃₆Sn₂₃'s long-period superstructure, *J. Chem. Theory Comput.*, 2013, **9**, 3170–3180, DOI: [10.1021/ct400274f](https://doi.org/10.1021/ct400274f).
- 90 J. Shao, C. Beaufils and A. N. Kolmogorov, Ab initio engineering of materials with stacked hexagonal tin frameworks, *Sci. Rep.*, 2016, **6**, 28369.
- 91 J. Kouvetakis, R. B. Kaner, M. L. Sattler and N. Bartlett, A novel graphite-like material of composition BC₃, and nitrogen-carbon graphites, *J. Chem. Soc., Chem. Commun.*, 1986, 1758–1759, DOI: [10.1039/C39860001758](https://doi.org/10.1039/C39860001758).
- 92 P. F. Rogl, J. Vreštal, T. Tanaka and S. Takenouchi, The B-rich side of the B–C phase diagram, *Calphad*, 2014, **44**, 3–9, DOI: [10.1016/j.calphad.2013.07.016](https://doi.org/10.1016/j.calphad.2013.07.016). Special Issue: TOFA 2012 – Discussion Meeting on Thermodynamics of Alloys.
- 93 C. R. Tomassetti, D. Gochitashvili, C. Renskers, E. R. Margine and A. N. Kolmogorov, First-principles design of ambient-pressure Mg_xB₂C₂ and Na_xBC superconductors, *Phys. Rev. Mater.*, 2024, **8**, 114801, DOI: [10.1103/PhysRevMaterials.8.114801](https://doi.org/10.1103/PhysRevMaterials.8.114801).
- 94 J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither and A. Grimshaw, *et al.*, XSEDE: Accelerating scientific discovery, *Comput. Sci. Eng.*, 2014, **16**, 62–74, DOI: [10.1109/MCSE.2014.80](https://doi.org/10.1109/MCSE.2014.80).

