Annealing for prediction of grand canonical crystal structures: Implementation of n-body atomic interactions

Yannick Couzinié [1,2,*] Yusuke Nishiya [1,2] Hirofumi Nishi [1,2] Taichi Kosugi [1,2]
Hidetoshi Nishimori [3,4,5] and Yu-ichiro Matsushita [1,2,6]

[1]*Laboratory for Materials and Structures, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama 226-8503, Japan*
[2]*Quemix Incorporated, Taiyo Life Nihombashi Building, 2-11-2, Nihombashi, Chuo-ku, Tokyo 103-0027, Japan*
[3]*International Research Frontiers Initiative, Tokyo Institute of Technology, Shibaura, Minato-ku, Tokyo 108-0023, Japan*
[4]*Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan*
[5]*RIKEN, Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS), Wako, Saitama 351-0198, Japan*
[6]*Quantum Material and Applications Research Center, National Institutes for Quantum Science and Technology (QST),
2-12-1, Ookayama, Meguro-ku, Tokyo 152-8552, Japan*

We propose an annealing scheme usable on modern Ising machines for crystal structures prediction (CSP) by taking into account the general n-body atomic interactions and in particular three-body interactions which are necessary to simulate covalent bonds. The crystal structure is represented by discretizing a unit cell and placing binary variables which express the existence or nonexistence of an atom on every grid point. The resulting quadratic unconstrained binary optimization (QUBO) or higher-order unconstrained binary optimization (HUBO) problems implement the CSP problem and is solved using simulated and quantum annealing. Using the example of Lennard-Jones clusters we show that it is not necessary to include the target atom number in the formulation allowing for simultaneous optimization of both the particle density and the configuration and argue that this is advantageous for use on annealing machines as it reduces the total amount of interactions. We further provide a scheme that allows for reduction of higher-order interaction terms that is inspired by the underlying physics. We show for a covalently bonded monolayer MoS_2 crystal that we can simultaneously optimize for the particle density as well as the crystal structure using simulated annealing. We also show that we reproduce ground states of the interatomic potential with high probability that are not represented on the initial discretization of the unit cell.

I. INTRODUCTION

Crystal structure prediction (CSP) from chemical composition alone is still one of the most difficult problems in materials science, even for the simplest structures [1]. The reason why this problem is still a challenge is that the variation of possible structures grows exponentially as the number of atoms increases, making an exhaustive search for the most stable structure, i.e., finding the global minimum on the Born-Oppenheimer surface, unfeasible even with today's supercomputers. For a small number of atoms, a brute force approach is possible, but reliably finding global optima for larger systems is out of reach of current computers.

Various approaches to develop searching algorithms that approximate solutions to the CSP have been developed [2], e.g., random search [3–6], simulated annealing (SA) [7–9],

minima hopping [10,11], evolutionary algorithm [12–15], and particle swarm optimization [16,17]. Various software suites such as USPEX [13–15], CALYPSO [16,17], and CRYSPY [18] that implement these algorithms continue to be developed and improve upon these algorithms. However, all of these approaches have one thing in common: As the system size increases, they become easily trapped by locally stable solutions, and to escape from these becomes a nontrivial problem. To address this, approaches incorporating experimental data such as x-ray diffraction patterns into the optimization process have been developed [19–21], e.g., the data assimilation technique which has been successfully applied to crystal structure and amorphous structure prediction [22–24].

In recent years, the use of quantum computers has attracted a great deal of attention as a means of searching for globally optimal solutions [25–30]. Quantum computers are characterized by their ability to escape from locally stable solutions and accelerate the search for globally optimal solutions by utilizing the quantum tunneling effect [27,31]. Quantum annealing (QA) machines [25,32–35] and gate-based quantum computers are the two main current architectures in development. Exhaustive structure search using gate-based quantum computers has been reported recently [36,37]. In the method described in Ref. [37], space is divided into meshes, and the

*couzinie.y.aa@m.titech.ac.jp

presence or absence of atoms on each mesh is represented as a $\{0,1\}$ digital number, allowing the crystal structure to be encoded onto qubits as a bit sequence. On the qubits, various atomic coordination structures can be prepared at once by using the quantum superposition states on the qubits. The idea is to perform exhaustive structural optimization by applying a probabilistic imaginary-time evolution technique reported in Ref. [30].

In this paper, we report a method to perform exhaustive structural optimization using QA. In particular, we discuss how to reformulate structural optimization as a quadratic unconstrained binary optimization problem (QUBO) or higher-order unconstrained binary optimization (HUBO). We provide a scheme for implementing an empirical three-body interatomic potential on QA hardware, and we provide a detailed analysis of preliminary SA and QA results. In particular we argue that providing more physical information in the form of penalty terms does not necessarily speed up the computation.

The remainder of the paper is structured as followed. In Sec. II we present the HUBO formulation for the CSP. In Sec. III we introduce the methods and general parameters used for optimization. In Sec. IV we outline the parameters for a Lennard-Jones cluster of krypton atoms for which we optimized both structure and density using SA and QA. In Sec. V we present a covalently bonded MoS_2 crystal modeled by a Stillinger-Weber potential for which we optimized again the structure and density using SA. We then close with the conclusions in Sec. VI.

II. HIGHER-ORDER UNCONSTRAINED BINARY OPTIMIZATION FORMULATION

In this section we discuss the construction of our HUBO. In Sec. II A we discuss the notation of our unit-cell discretization and the encoding into a HUBO of the CSP. In Sec. II B we discuss the penalty terms we use and finally in Sec. II C we discuss a physically motivated scheme to reduce the interaction terms of interaction terms of order higher than quadratic.

A. Crystal structures prediction problem encoding and Hamiltonian

Consider a unit cell that is spanned by a given basis $\{\vec{a}_i\}$ with periodic boundary conditions along a chosen set of basis vectors and a set of atom species $\mathcal{S}$. We look at a set of N lattice points X in this unit cell generated by partitioning each basis vectors into $g+1$ points and forming the corresponding lattice. The lattice points have the form $\sum_i \frac{k_i}{g}\vec{a}_i$ where $k_i \in \{0,\ldots,G_i\}$ with $G_i = g$ if we have no periodic boundary conditions along $\vec{a}_i$ and $G_i = g-1$ otherwise. Consider a set b_x^s of binary variables that we define such that if $b_x^s = 1$ there is an atom of species $s \in \mathcal{S}$ on $x \in X$. Assume that we have a set of potential functions $V_m^{s_1,\ldots,s_m}(x_1,\ldots,x_m)$ for a configuration of atoms of species s_i on x_i for $m \in \{1,\ldots,M\}$. As is usual for interatomic potential functions we assume that it does not depend on the order in which the argument, species pairs are supplied, i.e.,

$$V_m^{s_1,\ldots,s_m}(x_1,\ldots,x_m) \equiv V_m^{s_{\sigma(1)},\ldots,s_{\sigma(m)}}(x_{\sigma(1)},\ldots,x_{\sigma(m)}), \quad (1)$$

for any permutation σ. Assuming that we have no periodic boundary conditions, we define our Hamiltonian as

$$\begin{aligned}
H = &\sum_{\substack{x\in X\\ s\in\mathcal{S}}} V_1^s(x)b_x^s \\
&+ \frac{1}{2!}\sideset{}{'}\sum_{\substack{x_1,x_2\in X\\ s_1,s_2\in\mathcal{S}}} V_2^{s_i,s_j}(x_1,x_2)b_{x_1}^{s_1}b_{x_2}^{s_2} + \cdots \\
&+ \frac{1}{M!}\sideset{}{'}\sum_{\substack{x_1,\ldots,x_M\in X\\ s_1,\ldots,s_M\in\mathcal{S}}} V_M^{s_1,\ldots,s_M}(x_1,\ldots,x_M)b_{x_1}^{s_1}\cdots b_{x_M}^{s_M}, \quad (2)
\end{aligned}$$

where the prime indicates that the $x_i \in X$ should be chosen such that $x_i \neq x_j$ for any pair i, j (the species are chosen freely). Defined as such, finding the optimal nuclear structure on the lattice X corresponds to finding an optimal binary string that minimizes this Hamiltonian, as energy contributions only arise if all binary variables involved in an interaction are 1, i.e., all atoms involved in the interaction are present.

Generalizing this to the case with periodic boundary conditions requires a careful consideration of the self-interactions of atoms with their periodic images and a fitting definition of the Hamiltonian. This is done in detail in the Appendix A.

B. Penalty terms

Equation (2) allows us to calculate the cohesive energy of a given configuration (see Appendix B). Thus, for well-constructed interatomic potentials that accurately model a wide range of configurations of a material, Eq. (2) not only gives the optimal configuration, but by simultaneously finding the optimal amount of binary variables that should have the value 1 we optimize for the optimal density of atoms in the unit cell.

It is possible to *a priori* fix a target atom number in the unit cell by adding a *penalty term* such as

$$P\left(\sum_{x\in X} b_x^s - \mathcal{C}_s\right)^2 \quad (3)$$

to the Hamiltonian for an appropriately large positive P and all $s \in \mathcal{S}$, where $\mathcal{C}_s$ is the target particle number for species s atoms. We call this an *absolute penalty term*.

Equivalently, knowing the chemical formula [e.g., $Al_2(SO_4)_3$] but not the optimal density, a penalty term such as

$$P\left(\sum_{x\in X} b_x^{s_1} - c_{s_1,s_2}\sum_{x\in X} b_x^{s_2}\right)^2 \quad (4)$$

ensures that the ratios of atoms are respected, where c_{s_1,s_2} is the target ratio (in the above example $c_{S,O} = 1/4$). This penalty term allows for finding the optimal density in the range that the ratio is respected. We call this a *relative penalty term*.

C. Reduction of interaction terms

Interatomic potentials will usually include a cutoff distance. To reduce pairwise interaction terms, it is crucial to

choose the right penalty terms because an absolute penalty term will introduce interactions between any pair of binary variables for the same species, even if their pairwise distance is higher than the cutoff distance. Similarly, relative penalty terms introduce pairwise interactions between any pair of binary variables of the two involved species. Choosing the wrong penalty terms can make the difference between having a sparse or fully connected graph of pairwise interactions. Ideally, no penalty terms would be introduced, but this is dependent on the quality of the chosen potential.

The number of interaction terms for the cubic or higher-order terms in the HUBO will be orders of magnitudes higher than for the pair interactions. Often, alongside the total number of spins, the density of the interaction graph is the main bottleneck for modern annealing machines [35,38] and as such it is crucial to devise schemes that reduce the interaction number beyond just applying a cutoff. To this end we use the "deduc-reduc" method from Ref. [39]. In particular, we make the assumption that if the pairwise interaction between two binary variables is higher than a user-set threshold T, then any higher-order interaction containing this pair can safely be set to zero without influencing the ground state. At the same time we replace any pairwise interaction J_{ij} by $\min(J_{ij}, T)$. The intuition behind this is that for the interatomic potentials we use in this work, the pairwise interaction rapidly increases if the atoms are too close, and thus the ground state does not contain atoms on the two involved locations and we do not need to evaluate the higher-order terms. This is a simplification that does not lose any generality with respect to the ground state of the HUBO and which in particular also does not require any a priori knowledge like atomic radii of the involved species.

III. METHODS

We find optimal binary strings for the HUBO problems using SA and QA. In this section we outline the notation, parameters, and settings we used for the optimization.

A. Simulated annealing

Simulated annealing is a classic algorithm for optimizing cost functions with several local minima [40]. We assume some basic knowledge of the algorithm and will only discuss the specifics of our implementation. We use a geometric cooling schedule

$$T(x) = T_{\max}\left(\frac{T_{\min}}{T_{\max}}\right)^{x/N_{\mathrm{steps}}}, \quad x \in [0, N_{\mathrm{steps}}], \quad (5)$$

where $T_{\min}$ and $T_{\max}$ are the minimum and maximum temperature. The number of steps N_{steps} is the number of Monte Carlo steps per spin to perform.

Choosing the right neighborhood for a configuration in SA (i.e., defining legal transitions of the Markov chain) is crucial and generally one aims to have a smooth energy landscape with not too rugged local minima [41–43]. Traditionally, SA for HUBOs performs single bit flips. As this is equivalent to removing or adding an atom from the configuration, especially in the presence of penalty terms, this can be a costly operation. Thus, for each step in the schedule we loop over every binary variable and attempt to flip it and then we loop over every

opposite valued pair in the current configuration and attempt to exchange their values. This latter flip moves an existing atom to a random location and does not break penalty terms such as the absolute penalty (3) or relative penalty (4), thus ensuring a smoother energy landscape. So when we speak of Monte Carlo steps per spin we mean that we attempt $N|\mathcal{S}| + \binom{N|\mathcal{S}|}{2}$ spin flips where $N|\mathcal{S}|$ is the (unreduced) binary variable number.

B. Quantum annealing

We also assume familiarity with the basic concepts of quantum annealing [25,27]. We use the Advantage system available through the D-Wave leap cloud service [44]. Our HUBO and QUBO problems are very densely connected and if the cutoff of the potential function is large enough or the system small enough, the problem might even be fully connected. Embedding these onto the Pegasus architecture of the Advantage system [35] requires us to calculate a minor embedding [45–47]. Instead of manually calculating an embedding best fit for our problem, we use the standard implementation for clique embedding in the D-Wave Ocean SDK. This procedure can lead to results with broken chains which require a fitting unembedding. While there is evidence that designing a problem specific unembedding algorithm [48] can be advantageous we choose the simple majority vote which sets the binary value of a chain to the one that occurs most often on the chain.

C. Benchmarking

For benchmarking the various optimization schemes for the HUBO and QUBO formulation we use the time to solution (TTS) [49,50] given by

$$\mathrm{TTS}(\tau) = \tau \frac{\ln (1 - p_r)}{\ln [1 - \mathbb{P}_{\mathrm{GS}}(\tau)]} = \tau \frac{\ln (0.01)}{\ln [1 - \mathbb{P}_{\mathrm{GS}}(\tau)]}, \quad (6)$$

where τ is the running annealing time as measured on the local machine and $\mathbb{P}_{\mathrm{GS}}(\tau)$ is the probability of the corresponding algorithm to return the ground state with a running time of τ. The time-to-solution can be understood as the average time it takes to get the ground state with probability p_r which we set to 0.99.

IV. KRYPTON SYSTEM

In this section we introduce a Lennard-Jones cluster system consisting of krypton atoms in Sec. IV A and the related SA and QA results in Sec. IV B.

A. Setup

For the calculation of the potential functions we rely on the Open Knowledgebase of Interatomic Models (OpenKIM) [51]. In particular we look at a three-dimensional cubic unit cell of side length 5.653 Å with the Lennard-Jones potential parameters due to Bernades for krypton [51–55] and periodic boundary conditions along all three basis vectors. We look for the ground state configuration of krypton atoms in this unit cell discretized into an equipartitioned lattice of size g^3, which is equal to the face-centered cubic configuration
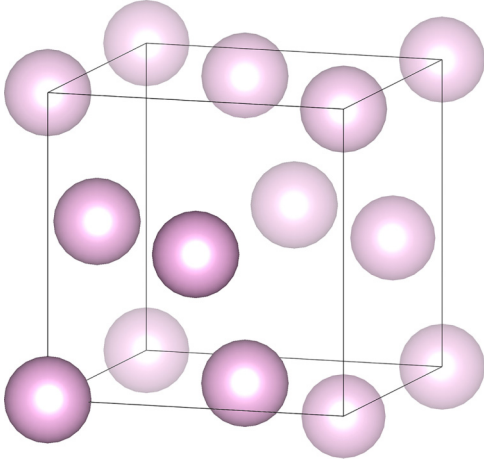
FIG. 1. The target fcc configuration of the krypton system with krypton atoms in pink (graphics due to Vesta). The solid atoms on the origin and the three incident face centers are the locations encoded in the HUBO while the remaining transparent ones are copies due to the periodic boundary conditions and not part of X.
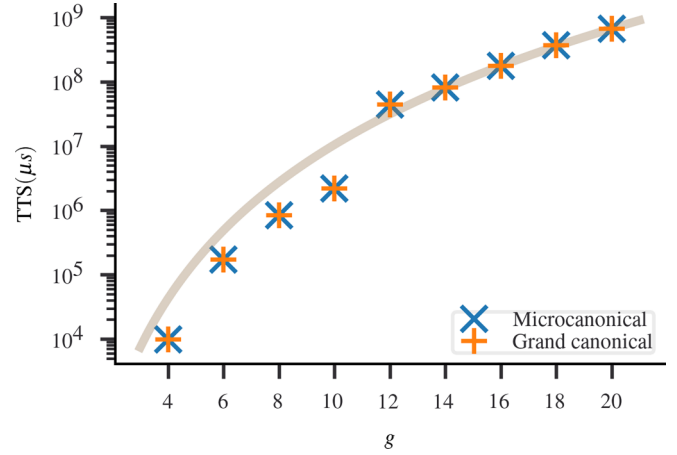

FIG. 2. The SA time to solution results for the krypton system with a penalty term in blue crosses and without in orange plus symbols plotted against various grid granularities g. The solid line corresponds to a fit of the measured points to $a[N + \binom{N}{2}]$, where $a = 21.015$ is the fitting parameter.

and can be seen in Fig. 1. The energy of the fcc configuration is -0.431 eV and for any interaction value J_{ij} we take $\min(J_{ij}, 1$ eV$)$. While for the SA calculations this is not strictly necessary, it helps for the QA calculations because the energy range is normalized to be between 0 and 1 on D-Wave machines, thus upper bounding the energy ensures that the physically interesting energy range takes up a larger portion of the renormalized energy range. We simply refer to this system as the krypton system. We perform SA calculations without any penalty terms and with an absolute number penalty term setting $\mathcal{C}_{\mathrm{Kr}} = 4$, we call the former grand canonical and the latter microcanonical. As the unit cell is smaller than the cutoff distance of the potential, even the grand canonical calculation QUBO is fully connected. We use a penalty strength of $P = 1$, and vary the temperature from 10^{-2} to 10^{-4}. The various probabilities correspond to the measured probability across 1000 annealing runs.

Since the systems are fully connected, for the QA calculations, we simplify the QUBO by fixing the binary variable for the origin to be one and removing any binary variable that had an interaction with the origin of more than 1 eV. This corresponds in essence to removing the translational invariance of the problem. Furthermore, we use pausing [56,57]. We use a base length of the schedule of 20 μs and we pause for 3 μs. We consider the success probability, i.e., the ratio of obtained ground states over 40 000 annealing runs, plotted against the pause location $s_p \in (0, 1)$ so that the dimensionless time in the annealing schedule goes from 0 to s_p at $(17s_p)$ μs until $(17s_p)$ μs $+ 3$ μs and then goes to one linearly until 20 μs. We use a chain strength of 1.28. These parameters were heuristically found to provide reasonable results.

B. Results and discussions

In Fig. 2 we plot the TTS against various grid spacings g for SA calculations for the grand and microcanonical system. We performed SA until we found the ground state fcc configuration with a probability of more than 90% and take the

minimum TTS across the schedule steps as the data point for g. This takes at most 30 schedule steps for both systems and it is apparent that both systems have comparable performance. In particular note the fit to the function $N + \binom{N}{2}$, which is the scaling of the number of flips the SA algorithm attempts with the spin number N. There are two main mechanisms that increase the required TTS. The first is that, as we attempt more spin flips per schedule step with increasing g, SA requires more time per schedule step to perform the increasing amount of flips. The second is that with increasing g the atoms have more fine-grained displacement possibilities so that there are more local minima of the QUBO problem with energies closer to the actual ground state leading to an increased time to escape the local minima to find the ground state.

If the global minimum were harder to find due to increasing amounts of local minima, we would expect an increasing number of required schedule steps with increasing g. What we see is that the fit $a[N + \binom{N}{2}]$ with a constant $a = 21.015$ reconstructs the data well for $g \geqslant 12$ for both systems. Thus there is no significant scaling $\sim \mathrm{TTS}(\tau)/[N + \binom{N}{2}]$ of the required scheduled steps with g for the microcanonical and the grand canonical system. Furthermore, in Fig. 3 we show a representative energy histogram for the grand canonical calculations with three Monte Carlo steps per spin for $g \in \{12, 14, 16, 18, 20\}$. Despite not putting any particle number restrictions the annealing process, even for this low amount of schedule steps, only returns solutions with the correct atom density and in fact all returned energies are lower than the first-excited state energy corresponding to an fcc configuration with an atom taken out (see Appendix C), a state we call fcc-1. Using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [58–61] to converge to a local minimum off the grid X we confirmed that all states with four atoms converge to the ground state meaning that the TTS of the combination of annealing combined with BFGS is considerably lower than that of only annealing.

We also confirmed these tendencies on the D-Wave Advantage 4.1 system available on D-Wave Leap. We performed
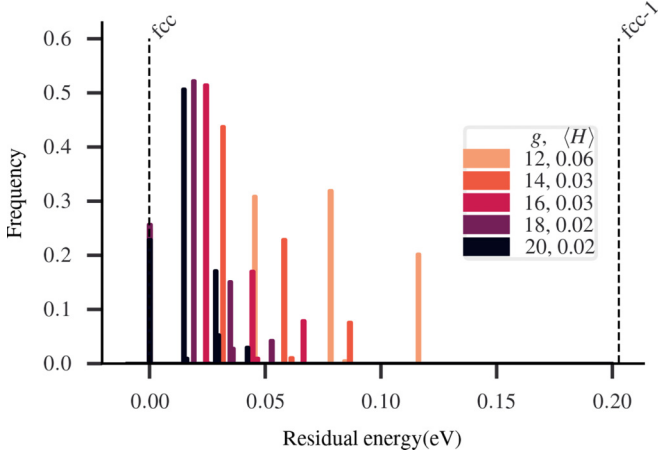
FIG. 3. The histogram for the residual energy of the krypton system after running SA for three Monte Carlo steps per spin $g \in \{12, 14, 16, 18, 20\}$ [colored from light red (gray) to dark purple (gray)] together with their average residual energy, i.e., energy above the ground state, $\langle H \rangle$. This is the full histogram, no results have been cut.

calculations only for the $g = 4$ system since the minor embedding for the $g = 6$ system had chain lengths of up to 20 spins which proved too hard to optimize. In Fig. 4 we plot the pause location s_p against the success probability for the grand and microcanonical system for just QA with pausing and without pausing and a schedule length of 18.9 μs and quantum annealing with pausing followed by BFGS. The penalty strength in the microcanonical calculations is 0.05 as it provided the best ground-state probability. First we see that pausing improves the performance as for both systems the probabilities without pausing are around 0.001 and with pausing the maximum probabilities for the grand canonical
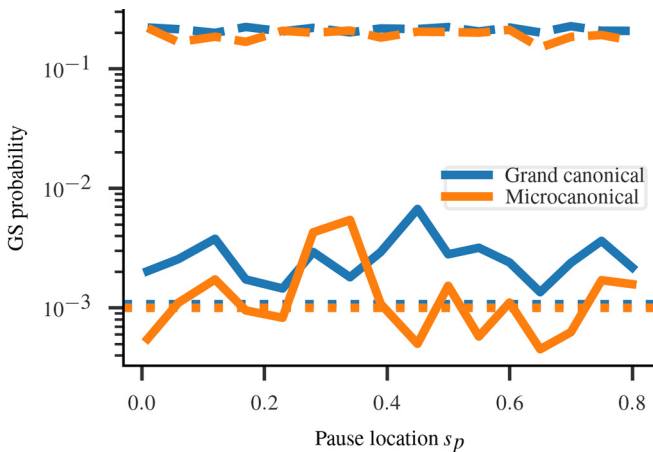

FIG. 4. Ground-state probabilities for the $g = 4$ krypton system using the D-Wave Advantage 4.1 system with various pause locations s_p ranging from 0.01 to 0.8. In blue (dark gray) the grand canonical calculation and in orange (light gray) the microcanonical with a penalty strength of 0.05. The dashed lines correspond to the ground-state probability after applying BFGS on the results from the solid lines and the dotted line to the probability of running annealing with no pauses and an annealing time of 18.9 μs.

system are 0.0067 at $s_p = 0.45$ and 0.005 425 at $s_p = 0.34$ for the microcanonical one. Since there are no same-density local minima, performing BFGS optimization on the results with pausing, is equivalent to looking at the results that have the correct density. We see that for QA + BFGS calculations both systems have success probabilities between 0.15 and 0.22 with the grand canonical consistently having a higher probability.

Without pausing QA has a TTS of around 0.9×10^5 μs comparable with the microcanonical system TTS for SA in the $g = 4$ case (see Fig. 2). With pausing we find a TTS of 13 700 and 16 931 μs, respectively, for the grand and microcanonical system providing comparable times to the grand canonical SA calculations albeit the QA calculations are a bit slower. Thus we find no indications of a quantum speedup. Possible reasons for this result may include the embedding of full connectivity on the sparse hardware graph and noise effects. We leave it for future research to analyze this problem with a wider set of parameters and using more intricate embedding techniques.

Note though, that while the SA + BFGS algorithm did not provide any other minima than the global one, QA + BFGS returns the fcc-1 configuration with probabilities between 0.3 and 0.33 across all pause locations s_p for the grand canonical system and 0.21 and 0.25 for the microcanonical system. Thus while we might not expect a quantum speedup there might be an advantage due to the higher breadth of results returned by QA compared with SA allowing a wider exploration of the potential-energy landscape.

Summarizing, we see that also for QA, at least in this very simple system, there are no performance costs in leaving out the penalty and in fact we can expect performance increases confirming the tendencies found in SA.

V. MoS_2 SYSTEM

In this section we introduce a MoS_2 system governed by the three-body Stillinger-Weber potential in Sec. V A and the related SA results in Sec. V B.

A. Setup

For the second system we consider the Stillinger-Weber potential [62,63] which is a simple three-body potential that reflects covalent bond dynamics. We use the parametrization for hexagonal monolayer molybdenum-disulfide due to Wen *et al.* [64–67]. We do this on the supercell consisting of a 2×2 lattice of hexagonal lattice unit cells with a single unit cell having a lattice constant of 3.20 Å and thickness of 3.19 Å. Thus the lattice vectors for our system are $\vec{a}_1 = (3.2 \text{ Å}, -\sqrt{3} \times 3.2 \text{ Å}, 0)$, $\vec{a}_2 = (3.2 \text{ Å}, -\sqrt{3} \times 3.2 \text{ Å}, 0)$, $\vec{a}_3 = (0, 0, 3.19 \text{ Å})$ We build the lattice by partitioning both $\vec{a}_1$ and $\vec{a}_2$ into $g = 6$ equal parts each and applying periodic boundary conditions and partitioning $\vec{a}_3$ into three equal parts without periodic boundary conditions. Thus the amount of required bits scales like $6g^2$, where the additional two come from the amount of species. The target ground state is the 2 H configuration (see Fig. 5) and has an energy of -55.5283 eV. The first excited state that we expect to see is the 1 T configuration, with the same amount of atoms and an energy that is 1.4755 eV above the ground state (see Appendix C). We refer to this system as the MoS_2 system.
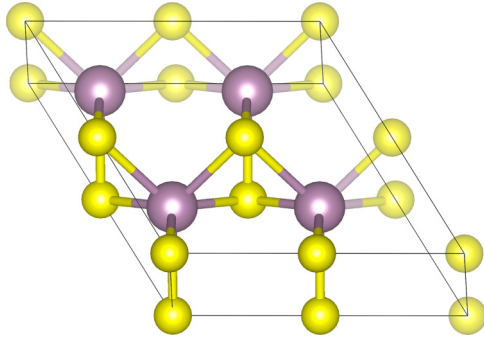
FIG. 5. The target 2 H ground-state configuration of the MoS_2 system with sulfur in yellow and molybdenum in violet (graphics due to Vesta). The bottom six sulfur atoms on the boundary, two at $(\vec{a}_1 + \vec{a}_2)/2$ and four molybdenum atoms with z coordinate given by $\vec{a}_3/2$ are the locations encoded in the HUBO (nontransparent atoms). The remaining ten sulfur atoms (transparent) are copies due to the periodic boundary conditions and not part of X.

We use our deduc-reduc with a threshold of 10 eV which in this particular case reduced the amount of nonzero three-body interaction terms by 18.8% (from 1 573 728 to 1 277 267) in the $g = 6$ system. Any lower threshold seemed to impact the ground-state configuration on our SA calculations. There is no general-use scheme known to the authors, that would allow us to quadratize this HUBO so as to make it runnable on any modern Ising machine [68–70] and so while our deduc-reduc step reduces the interactions it can only be a first step in conjunction with other approaches yet to be found and we perform no QA for this system.

We perform SA for the system with both absolute penalty terms ($\mathcal{C}_{Mo} = 4$, $\mathcal{C}_S = 8$) and relative penalty terms ($\mathcal{C}_{Mo,S} = 1/2$). For simplicity we call the former the absolute system and the latter the relative system. Grand canonical calculations as in the krypton system without penalty terms do not work for this potential, as it is more favorable to produce configurations with a single atom species rather than a MoS_2 mix, so we limit our analysis to the relative and absolute system and recall that the former retains the function of simultaneously optimizing for the atom density. The number of pairwise interaction terms without interactions increases by 1.2% using the absolute penalty (from 21 420 to 21 708) and by 8.4% using the relative penalty (23 220), underlining again the importance of finding potentials that can be used without penalties to reduce the number of pairwise interactions necessary. In fact, since this potential is parametrized for hexagonal MoS_2 we cannot expect it to yield accurate results for nonhexagonal configurations. This is a problem that does not pertain to the parametrization but the Stillinger-Weber potential in general. Since this one of the simplest three-body potentials we use it anyway for this proof-of-concept calculation.

We use a penalty strength of $P = 10$ and a temperature range of 10 to 0.1 for SA. The various probabilities correspond to the measured probability across 1000 annealing runs.

B. Results and discussions

The MoS_2 system proves harder to optimize than the krypton system. In Fig. 6 the ground-state probabilities for
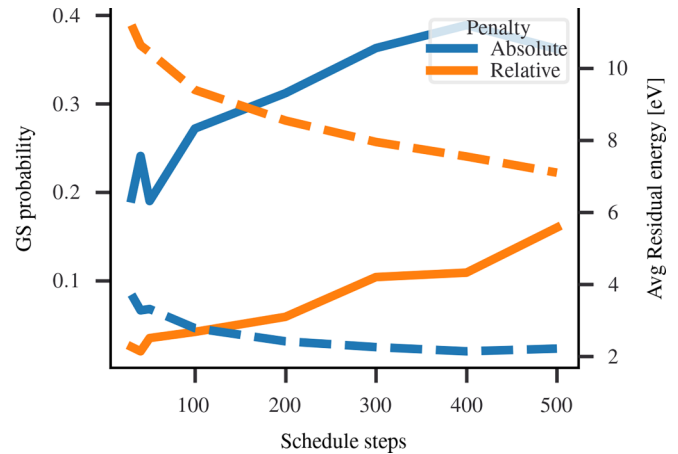

FIG. 6. Plot in solid lines of the ground-state probability for SA for the MoS_2 system with schedule steps going from 2 to 500 for both relative penalties and absolute penalties [blue (dark gray) and orange (light gray), respectively]. The scale for the probability is to the left. In dashed lines the average residual energy with the corresponding scale to the right.

schedule steps going from 2 to 500 are plotted. As opposed to the krypton system where even for $g = 20$ we need only 30 schedule steps to reach a ground-state probability of above 0.9 we see that it hovers around 0.4 for the absolute penalty and around 0.15 for the relative penalty at 500 schedule steps. In particular note that here the used penalty terms have an effect on the ground-state probability, and that supplying more information (in form of the absolute penalty) leads to higher ground-state probabilities. As expected the ground-state probability increases with increasing amount of schedule steps but the slope does not offset the increase in calculation length and so the TTS turns out to be minimized for a number of schedule steps in the single digits for both system. In Fig. 6 the average residual energies are plotted and we see that both systems seem to converge to an average residual energy that is well above the target 0 eV. To understand this, consider the energy histogram in Fig. 7 for the resulting states of only SA (top) and SA followed by BFGS with the same potential (bottom) after 500 schedule steps. First, note that despite not fixing an absolute number of atoms in the relative penalty, we find the correct density of Mo_4S_8 in 42.8% of the configurations (in dark purple in Fig. 7) and that the average residual energy for the states with the correct density is 2.3826 eV while it is 10.6117 eV for the states with the wrong density (in red) so that the relative penalty calculations allow for simultaneous optimization of the atom density and the optimal configuration. The probability to obtain either 2 H or 1 T configurations is 42% for the absolute penalty system and 18.9% for the relative penalty system. To understand the physical nature of the remaining local minima, which form the majority of found states, we performed BFGS on all the resulting states from SA. While the probability for 2 H and 1 T rose to 42.8% and 20.9% for the absolute penalty and relative penalty system, respectively, we see that most states converge to a local minimum that has an energy below that of 2 H. First, for the relative penalty system we see that 57.2% of
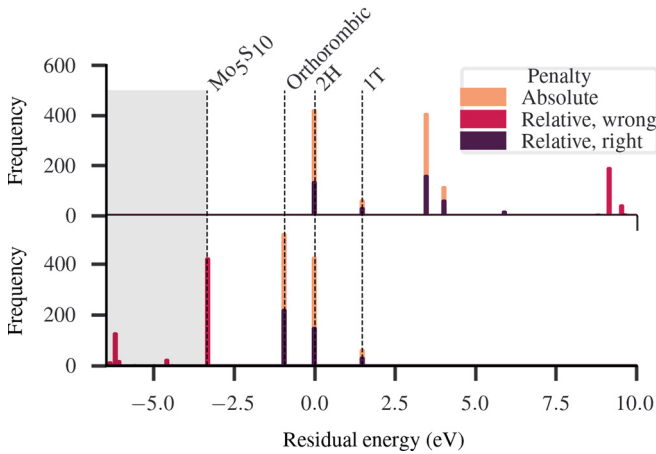
FIG. 7. Histogram of the residual energy for the MoS_2 system with SA with 500 schedule steps (top) and SA + BFGS (bottom) applied to the MoS_2 system with the absolute number penalty Eq. (3) in light red (light gray) and the relative number penalty Eq. (4) in dark purple (dark gray) for results with suboptimal density and red (gray) for the optimal density. Found local minima are marked by a dotted line and the shaded area to the left (see Appendix C for the configurations). This is not the full histogram, i.e., there are configurations with energies higher than 10 eV.

all observed configurations have 5 molybdenum atoms and 10 sulfur and form configurations that have an energy that is more than 2.5 eV lower than that of 2 H. In Fig. 7 we only shade the region as the BFGS algorithm does not converge well for these configurations so that we do not get well formed peaks but rather a distribution in the shaded area. The next lower state is a state we call orthorhombic (see Appendix C for an image of both the orthorhombic and an example Mo_5S_{10} configuration) and has an energy that is 0.9313 eV lower than that of 2 H. We find this configuration with a probability of 21.8% for the relative penalty system and 57.2% for the absolute penalty system.

Using the Vienna *ab initio* simulation package [71–73] with the projector augmented-wave method [74,75] we find that the energy of the 2 H configuration is in fact the lowest of the four found local minima, followed by the 1 T, the orthorhombic and finally the Mo_5S_{10} configurations. The fact that this order is not represented is due to the fact that the Stillinger-Weber potential is parametrized to model hexagonally ordered MoS_2 configurations and thus does not correctly model other configurations. The potential is not fit to provide new physical insights in our application and these results should be taken merely as a proof of concept.

Noteworthy about these results is that, despite the orthorhombic and locally optimal Mo_5S_{10} states not being representable on the discretization of the unit cell, the combination of SA and BFGS managed to find these states in a majority of attempts. This is a strong indication that if we are able to provide a fitting potential or directly a fitting HUBO we can find a wide array of globally and locally optimal configurations even if they are not part of the initial discretization. Thus, in particular it might suffice to have rougher discretizations with spin numbers that fit onto current

quantum hardware instead of trying to be fine grained enough to represent all possible local minima.

VI. CONCLUSIONS

In this paper we have presented an annealing scheme for crystal structure prediction based on n-body atomic interactions. We discretized a given unit cell with a lattice and placed binary variables on the lattice points to express the existence or nonexistence of an atom at every grid point. In particular this is done for three-body atomic interactions which is the minimum order necessary for covalent crystals. We solved the resulting HUBOs using SA and QA giving insights into the crystal structure. We have shown that a grand canonical calculation without penalty terms allows for the simultaneous optimization of both the nuclear structure as well as the particle density inside the unit cell. Furthermore, we have also shown evidence that the difficulty of solving the nuclear structure problem does not necessarily scale with the mesh size. These results show that it might not always be advantageous to put all the available information into the QUBO to speed up calculations in particular as this also increases the amount of total interaction terms the reduction of which is crucial for embedding problems into modern hardware with limited graphs.

We also considered a molybdenum-disulfide monolayer system modeled by a three-body Stillinger–Weber-type potential. Using our interaction number reduction scheme we reduced the amount of cubic interactions by 18.8% while maintaining physical accuracy to the extent of the used potential. We have shown that the potential contained unphysical ground states that are due to the limited transferability of the potential outside the context of hexagonal monolayer MoS_2. While these results do not provide physical insights, we show that our algorithm reproduces the ground state of the system even if they are not representable on the chosen discretization of the unit cell in the annealing step of the algorithm. Thus, while we could only optimize the roughest discretization for the krypton system on the D-Wave quantum annealer, this could be a hint that rougher discretizations, that are easier to embed onto quantum annealers, are enough for the local optimization algorithm to find a wide array of ground-state and locally optimal configurations.

An immediate future research question is to choose a more fitting potential to construct a HUBO that accurately models a wide array of covalent crystal configurations to test the performance with rough unit-cell meshes on larger unit cells.

Another research direction is to investigate the nature of returned local minima by QA and to confirm the tendency we found where QA provided a more varied insight into the energy than SA which tended to favor only ground states.

Note added. Recently, we have become aware of a similar proposal for the construction of the QUBO [76] for ionic crystals. That paper does not address higher-order optimization problems and thus does not address covalent bonds and did not consider the grand canonical case, their focus is on classical computation and providing guarantees that ground truths to the crystal structure prediction problem are found using their algorithm. They have similar findings with respect

to the reproducibility of the ground state even if it is not contained in the initial discretization.

APPENDIX A: PERIODIC BOUNDARY CONDITION IMPLEMENTATIONS

Recall that we work with charge neutral atoms and short-range (i.e., integrable) interatomic potentials with cutoffs. Usually in such cases to calculate interaction terms with periodic boundary conditions, the *minimum image convention* is employed, in which the simulation cell is chosen such that for any set of interacting atoms only one image of the involved atoms should be within the cutoff distance of each other, so there is a unique choice of which atoms interact [78]. This requires the unit cell to be at least twice the size of the cutoff distance. As we cannot choose the cutoff distance and the size of the required qubit number scales exponentially with the unit cell size we cannot use the minimum image convention.

In this section we derive the direct sum formula for an m-body potential with periodic boundary conditions and then show how to calculate the coefficients in the HUBO.

The energy of an infinite system due to an m-body potential V_m with atoms located on $x_1, x_2, \ldots \in \mathbb{R}^3$ is given as

$$\frac{1}{m!} \sum_{i_1 \in \mathbb{N}} \sum_{\substack{i_2 \in \mathbb{N} \\ i_2 \neq i_1}} \cdots \sum_{\substack{i_m \in \mathbb{N} \\ i_m \neq i_1, \ldots, i_{m-1}}} V_m(x_{i_1}, \ldots, x_{i_m}). \quad \text{(A1)}$$

Note that this includes the case where the atoms are of different species, for which the actual parametrized form of V_m would change depending on the input and the case where we have periodic boundary conditions only on a subset of basis vectors. We use the word atom on a location to mean an atom of a specific species on a given location to simplify the notation from Eq. (2) from the main text.

Assume now that the infinite system is generated by atoms on a unit cell on locations $x_1, x_2, \ldots, x_N$ replicated following a set of lattice vectors $\mathcal{L}$ so that Eq. (A1) becomes

$$\frac{1}{m!} \sum_{i \in [N]^m} \sideset{}{'}\sum_{\vec{n}_1, \ldots, \vec{n}_m \in \mathcal{L}} V_m(x_{i_1} + \vec{n}_1, \ldots, x_{i_m} + \vec{n}_m), \quad \text{(A2)}$$

where we write $[N] := \{1, \ldots, N\}$ and the prime on the sum indicates that if $i = j$ then $\vec{n}_i \neq \vec{n}_j$, i.e., we exclude

interactions with two or more atoms on the same location. This sum can be interpreted as the interaction terms of the unit cell given on $x_1 + \vec{n}_1$ with the surrounding supercell generated by the other lattice vectors. We thus define the energy of a single unit cell by setting $\vec{n}_1 = 0$ as

$$\frac{1}{m!} \sum_{i \in [N]^m} \sideset{}{'}\sum_{\vec{n}_2, \ldots, \vec{n}_m \in \mathcal{L}} V_m(x_{i_1}, x_{i_2} + \vec{n}_2, \ldots, x_{i_m} + \vec{n}_m), \quad \text{(A3)}$$

where the prime condition on the sum is the same as before with $\vec{n}_1$ replaced by zero. For example for the two-body potential given by $q_i q_j / |r_i - r_j|$, where q_i and q_j are the charges of the atoms on x_i and x_j, we recover the well-known formula [77]

$$\frac{1}{2} \sum_{i \in [N]} \sum_{j \in [N]} \sideset{}{'}\sum_{\vec{n} \in \mathcal{L}} \frac{q_i q_j}{|x_i - x_j - \vec{n}|}, \quad \text{(A4)}$$

to calculate Coulomb interactions with periodic boundary conditions. For the case with potentials of various order governing the system, e.g., Stillinger-Weber with a two- and three-body part, we take the sum over m to obtain the total energy of a unit cell with periodic boundary conditions given as

$$E(\{x_1, x_2, \ldots, x_N\})$$

$$:= \sum_{m \in [M]} \frac{1}{m!} \sum_{i \in [N]^m} \sideset{}{'}\sum_{\vec{n}_2, \ldots, \vec{n}_m \in \mathcal{L}} V_m(x_{i_1}, x_{i_2} + \vec{n}_2, \ldots, x_{i_m} + \vec{n}_m), \quad \text{(A5)}$$

where M is the highest-order potential involved.

Let us now come to the calculation of the HUBO coefficients so that the sum over binary variables in Eq. (2) from the main text reproduces Eq. (A5). Consider a set of lattice points $\{x_1, \ldots, x_m\} \subset X$ and associate to each point a species so that we consider an atom of species s_1 on x_1, where $\{s_1, s_2, \ldots, s_m\}$ is such that $s_i \in \mathcal{S}, i \in [m]$. We define the HUBO coefficients $H_{x_1, \ldots, x_m}^{s_1, \ldots, s_m}$ as

$$\sum_{\substack{\ell \in [M] \\ \ell \geqslant m}} \frac{1}{\ell!} \sum_{\substack{i \in [m]^\ell \\ [m] \subset i}} \sideset{}{'}\sum_{\vec{n}_2, \ldots, \vec{n}_\ell \in \mathcal{L}} V_\ell(x_{i_1}, x_{i_2} + \vec{n}_2, \ldots, x_{i_\ell} + \vec{n}_\ell), \quad \text{(A6)}$$

where for simplicity we leave out the explicit writing of the species and the condition $[m] \subset i$ on the second summation ensures that every index is contained in i. This condition is needed to ensure that we only consider potential contributions that require all the atoms and not only a subset which would be part of a different HUBO coefficient.

To see that Eq. (A6) is the correct way to define the HUBO coefficients, we need to show that the sum in Eq. (2) from the main text reproduces Eq. (A5). Let us consider a subset $\{y_1, \ldots, y_N\} = Y \subset X$ and a set $\{s_1, \ldots, s_N\}$ of species such that $b_{y_i}^{s_i} = 1$ for $i \in [N]$ and $b_x^s = 0$ otherwise. The sum in

Eq. (2) from the main text then resolves to

$$\sum_{m \in [M]} \sum_{\vec{i} \in [N]^m} H^{s_{i_1},\ldots,s_{i_m}}_{x_{i_1},\ldots,x_{i_m}} = \sum_{m \in [M]} \sum_{\vec{i} \in [N]^m} \sum_{\substack{\ell \in [M] \\ \ell \geqslant m}} \frac{1}{\ell!} \sum_{\substack{j \in \vec{i}^\ell \\ \vec{i} \subset j}} \sum_{\vec{n}_2,\ldots,\vec{n}_\ell \in \mathcal{L}}' V_\ell(x_{j_1}, x_{j_2} + \vec{n}_2, \ldots, x_{j_\ell} + \vec{n}_\ell) \tag{A7}$$

$$= \sum_{\ell \in [M]} \frac{1}{\ell!} \sum_{\substack{m \in [M] \\ m \leqslant \ell}} \sum_{\vec{i} \in [N]^m} \sum_{\substack{j \in \vec{i}^\ell \\ \vec{i} \subset j}} \sum_{\vec{n}_2,\ldots,\vec{n}_\ell \in \mathcal{L}}' V_\ell(x_{j_1}, x_{j_2} + \vec{n}_2, \ldots, x_{j_\ell} + \vec{n}_\ell), \tag{A8}$$

where the prime on the sum is in reference to the j index, i.e., if $j_k = j_{k'}$ then $\vec{n}_k \neq \vec{n}_{k'}$. Now use that the sums $\sum_{\vec{i} \in [N]^m} \sum_{j \in \vec{i}^\ell, \vec{i} \subset j}$ can be written as the sum over all ℓ-element multisets with elements from $[N]$ that have exactly m distinct elements, i.e., in an abuse of notation we can write

$$\sum_{i \in [N]^m} \sum_{\substack{j \in [i]^\ell \\ [i] \subset j}} = \sum_{j \in [N]^\ell} \mathbf{1}_{j \text{ has } m \text{ distinct elements}}, \tag{A9}$$

where $\mathbf{1}$ is the indicator function. Finally, since j has ℓ elements we have

$$\sum_{\substack{m \in [M] \\ m \leqslant \ell}} \mathbf{1}_{j \text{ has } m \text{ distinct elements}} = 1 \tag{A10}$$

and thus Eq. (A8) can be written as

$$\sum_{\ell \in [M]} \frac{1}{\ell!} \sum_{j \in [N]^\ell} \sum_{\vec{n}_2,\ldots,\vec{n}_\ell \in \mathcal{L}}' V_\ell(x_{j_1}, x_{j_2} + \vec{n}_2, \ldots, x_{j_\ell} + \vec{n}_\ell), \tag{A11}$$

and we recovered Eq. (A5).

There is an efficient way to calculate Eq. (A6) when you have access to an oracle that calculates the total energy as is, for example, the case in the OpenKIM API. This oracle for atoms on some locations $Y = y_1, \ldots, y_N \in \mathbb{R}$ returns

$$F_\ell(Y) := \sum_{\substack{i \in [N]^\ell \\ i_1 < i_2 < \cdots < i_\ell}} V_\ell(y_{i_1}, y_{i_2}, \ldots, y_{i_\ell}) \tag{A12}$$

$$= \frac{1}{\ell!} \sum_{i \in [N]^\ell}'' V_\ell(y_{i_1}, y_{i_2}, \ldots, y_{i_\ell}), \tag{A13}$$

where again we leave out the explicit mention of the species on the potential, use that the potential is constant under permutation of arguments and the double prime indicates that no two indices i_k, $i_{k'}$ should be the same in the summation [this is to simplify the notation from Eq. (A1)]. Recall that the potentials that we use have a hard cutoff. To calculate $H^{s_1,\ldots,s_m}_{x_1,\ldots,x_m}$ construct a supercell by adding copies of the configuration in the unit cell around the unit cell in the directions in which we have periodic boundary conditions up until the atoms in the unit cell have no nonzero interaction with the newly copied unit cells. As an example, for the MoS_2 system this means that we create a 5×5 cell of unit cells with the copied configurations. Call this set SC and their elements $y_1, y_2, \ldots, y_{|SC|}$ and note that the set $\mathcal{L}$ of lattice vectors is given by the basis vectors of

the unit cell. Now,

$$F_\ell(SC) - F_\ell(SC \setminus \{x_j\})$$

$$= \frac{1}{\ell!} \sum_{\substack{z_1,\ldots,z_\ell \in SC \\ \exists k \in [\ell]: z_k = x_j}}'' V_\ell(z_1, z_2, \ldots, z_\ell) \tag{A14}$$

$$= \frac{1}{(\ell-1)!} \sum_{z_2,\ldots,z_\ell \in SC}'' V_\ell(x_j, z_2, \ldots, z_\ell), \tag{A15}$$

so that

$$\sum_{j \in [m]} F_\ell(SC) - F_\ell(SC \setminus \{x_j\})$$

$$= \frac{1}{(\ell-1)!} \sum_{i \in [m]^\ell} \sum_{\vec{n}_2,\ldots,\vec{n}_\ell \in \mathcal{L}}' V_\ell(x_{i_1}, x_{i_2} + \vec{n}_2, \ldots, x_{i_\ell} + \vec{n}_\ell), \tag{A16}$$

where we used again that the potential is constant under permutation of arguments. The configuration energy with periodic boundary conditions Eq. (A5) is thus obtained by

$$E(UC) = \sum_{\ell \in [M]} \frac{1}{\ell} \sum_{j \in [m]} [F_\ell(SC) - F_\ell(SC \setminus \{x_j\})]. \tag{A17}$$

We can now calculate the linear HUBO coefficients in Eq. (A6) as

$$H^t_x = \sum_{\ell \in [M]} \frac{1}{\ell!} \sum_{\vec{n}_2,\ldots,\vec{n}_\ell \in \mathcal{L}}' V_\ell(x, x + \vec{n}_2, \ldots, x + \vec{n}_\ell) \tag{A18}$$

$$= E(\{x\}). \tag{A19}$$

Now, for quadratic terms we find

$$H^{s_1,s_2}_{x_1,x_2} = E(\{x_1, x_2\}) - E(\{x_1\}) - E(\{x_2\}), \tag{A20}$$

which is easily seen by looking at the second sum in Eq. (A6) which considers any multiset of indices that contains the entirety of the original set, i.e., here $\{1, 2\}$ and by subtracting the single atom energies on the right-hand side, we subtract those contributions that arise from the summands in which only a single index, either 1 or 2 is present. It is now clear how to generalize this:

$$H^{s_1,\ldots,s_m}_{x_1,\ldots,x_m} = E(\{x_1, \ldots, x_m\}) - \sum_{Y \subsetneq \{x_1,\ldots,x_m\}} H^{s_Y}_Y, \tag{A21}$$

where on the right-hand side we write $H^{s_Y}_Y$ for the coefficient with atoms on positions given by Y and the appropriate species set s_Y.
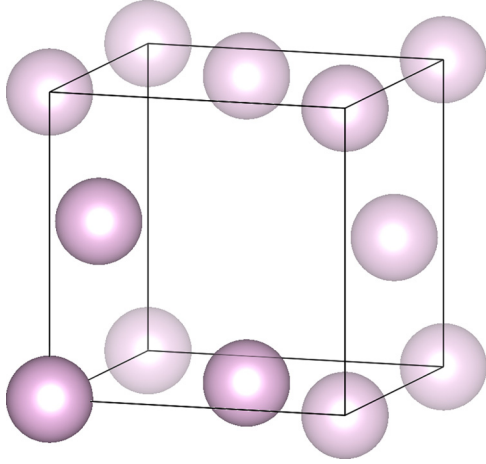
FIG. 8. Kr_3 configuration that corresponds to an fcc configuration with a single atom taken out and which has a residual energy of 0.2029 eV.

We close this Appendix with a remark on nonparametrized potentials in which you do not have access to the *n*-body potential part separately so that the oracle (A22) looks like

$$F(Y) := \sum_{\ell \in [M]} \frac{1}{\ell!} \sum_{i \in [N]^\ell}'' V_\ell(y_{i_1}, y_{i_2}, \ldots, y_{i_\ell}). \quad \text{(A22)}$$

In this case we have

$$F(SC) - F(SC \setminus \{x_j\})$$

$$= \sum_{\ell \in [M]} \frac{1}{(\ell - 1)!} \sum_{i \in [m]^{\ell-1}} \sum_{\vec{n}_2, \ldots, \vec{n}_\ell \in \mathcal{L}}' V_\ell$$

$$\times (x_j, x_{i_2} + \vec{n}_2, \ldots, x_{i_\ell} + \vec{n}_\ell), \quad \text{(A23)}$$

and thus it is not clear whether there exists an efficient algorithm to calculate $E(UC)$ with such an oracle.

APPENDIX B: COHESIVE ENERGY

When doing grand canonical calculations we need to ensure that the energies with different numbers of atoms are comparable. We use the notion of cohesive energy for this, which is usually defined as the difference in energy between
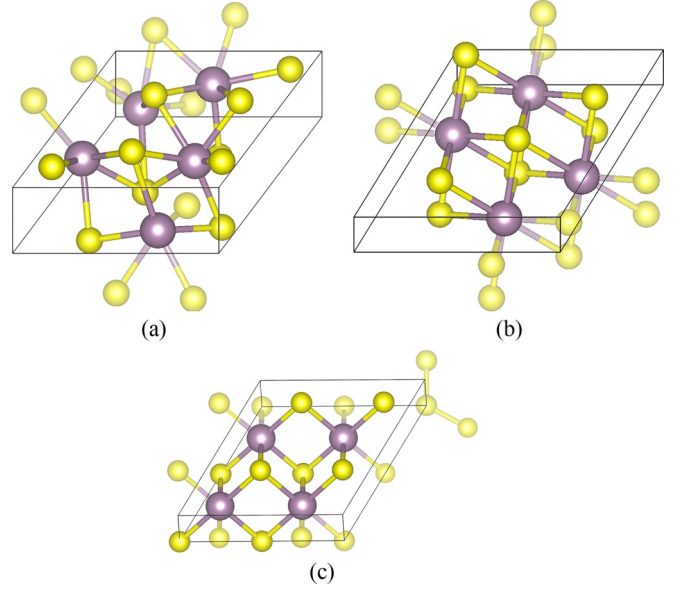

FIG. 9. Local minima of the MoS_2 system marked with a dotted line in Fig. 7 from the main text. From left to right, (a) an example Mo_5S_{10} configuration with a residual energy of -6.2161 eV, (b) the orthorhombic state with a residual energy of -0.9313 eV, and (c) the 1 T configuration of MoS_2, 1.4755 eV.

the atoms in a specific configuration and the energy of all the involved atoms at an infinite pairwise distance. In our case this means that we compare the energy of a configuration on the lattice with the regular lattice constant *a* and the energy with $a \to \infty$. For these energy calculations we use interatomic potentials with a hard cutoff and thus the energy of the atoms with an infinite pairwise distance is zero while it is nonzero for the regular lattice constant. Thus the cohesive energy in our case is calculated by Eq. (2) from the main text as claimed in the main text.

APPENDIX C: LOCAL MINIMA

We give an overview of the local minima indicated by dotted lines in the histograms Figs. 3 and 7 in the main text. The local minima for the krypton system are given in Fig. 8 and for the MoS_2 system in Fig. 9.

---

[1] J. Maddox, Crystals from first principles, Nature (London) **335**, 201 (1988).

[2] A. R. Oganov, C. J. Pickard, Q. Zhu, and R. J. Needs, Structure prediction drives materials discovery, Nat. Rev. Mater. **4**, 331 (2019).

[3] C. J. Pickard and R. J. Needs, High-pressure phases of silane, Phys. Rev. Lett. **97**, 045504 (2006).

[4] C. J. Pickard and R. J. Needs, Structure of phase III of solid hydrogen, Nat. Phys. **3**, 473 (2007).

[5] C. J. Pickard and R. J. Needs, *Ab initio* random structure searching, J. Phys.: Condens. Matter **23**, 053201 (2011).

[6] R. J. Needs and C. J. Pickard, Perspective: Role of structure prediction in materials discovery and design, APL Mater. **4**, 053210 (2016).

[7] L. Wille, Minimum-energy configurations of atomic clusters: New results obtained by simulated annealing, Chem. Phys. Lett. **133**, 405 (1987).

[8] L. Wille, Simulated annealing and the topology of the potential energy surface of Lennard-Jones clusters, Comput. Mater. Sci. **17**, 551 (2000).

[9] X. Yin and C. E. Gounaris, Search methods for inorganic materials crystal structure prediction, Curr. Opin. Chem. Eng. **35**, 100726 (2022).

[10] S. Goedecker, Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems, J. Chem. Phys. **120**, 9911 (2004).

[11] M. Amsler and S. Goedecker, Crystal structure prediction using the minima hopping method, J. Chem. Phys. **133**, 224104 (2010).

[12] T. S. Bush, C. R. A. Catlow, and P. D. Battle, Evolutionary programming techniques for predicting inorganic crystal structures, J. Mater. Chem. **5**, 1269 (1995).

[13] A. R. Oganov and C. W. Glass, Crystal structure prediction using *ab initio* evolutionary techniques: Principles and applications, J. Chem. Phys. **124**, 244704 (2006).

[14] A. R. Oganov, A. O. Lyakhov, and M. Valle, How evolutionary crystal structure prediction works—and why, Acc. Chem. Res. **44**, 227 (2011).

[15] A. O. Lyakhov, A. R. Oganov, H. T. Stokes, and Q. Zhu, New developments in evolutionary structure prediction algorithm USPEX, Comput. Phys. Commun. **184**, 1172 (2013).

[16] Y. Wang, J. Lv, L. Zhu, and Y. Ma, Crystal structure prediction via particle-swarm optimization, Phys. Rev. B **82**, 094116 (2010).

[17] Y. Zhang, H. Wang, Y. Wang, L. Zhang, and Y. Ma, Computer-assisted inverse design of inorganic electrides, Phys. Rev. X **7**, 011017 (2017).

[18] T. Yamashita, S. Kanehira, N. Sato, H. Kino, K. Terayama, H. Sawahata, T. Sato, F. Utsuno, K. Tsuda, T. Miyake, and T. Oguchi, Cryspy: A crystal structure prediction tool accelerated by machine learning, Sci. Technol. Adv. Mater.: Methods **1**, 87 (2021).

[19] B. Meredig and C. Wolverton, A hybrid computational–experimental approach for automated crystal structure solution, Nat. Mater. **12**, 123 (2013).

[20] P. Gao, Q. Tong, J. Lv, Y. Wang, and Y. Ma, X-ray diffraction data-assisted structure searches, Comput. Phys. Commun. **213**, 40 (2017).

[21] N. Tsujimoto, D. Adachi, R. Akashi, S. Todo, and S. Tsuneyuki, Crystal structure prediction supported by incomplete experimental data, Phys. Rev. Mater. **2**, 053801 (2018).

[22] D. Adachi, N. Tsujimoto, R. Akashi, S. Todo, and S. Tsuneyuki, Search for common minima in joint optimization of multiple cost functions, Comput. Phys. Commun. **241**, 92 (2019).

[23] S. Yoshikawa, R. Sato, R. Akashi, S. Todo, and S. Tsuneyuki, A noise-robust data assimilation method for crystal structure determination using powder diffraction intensity, J. Chem. Phys. **157**, 224112 (2022).

[24] Y. Zhao, R. Sato, and S. Tsuneyuki, Accelerating simulated annealing of glassy materials with data assimilation, J. Non-Cryst. Solids **600**, 122028 (2023).

[25] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, Phys. Rev. E **58**, 5355 (1998).

[26] C. Durr and P. Hoyer, A quantum algorithm for finding the minimum, arXiv:quant-ph/9607014.

[27] T. Albash and D. A. Lidar, Adiabatic quantum computation, Rev. Mod. Phys. **90**, 015002 (2018).

[28] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, Variational quantum algorithms for discovering Hamiltonian spectra, Phys. Rev. A **99**, 062304 (2019).

[29] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational ansatz-based quantum simulation of imaginary time evolution, npj Quantum Inf. **5**, 75 (2019).

[30] T. Kosugi, Y. Nishiya, H. Nishi, and Y.-I. Matsushita, Imaginary-time evolution using forward and backward real-time evolution with a single ancilla: First-quantized eigensolver algorithm for quantum chemistry, Phys. Rev. Res. **4**, 033121 (2022).

[31] V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. Smelyanskiy, J. Martinis, and H. Neven, What is the computational value of finite-range tunneling? Phys. Rev. X **6**, 031015 (2016).

[32] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Perspectives of quantum annealing: Methods and implementations, Rep. Prog. Phys. **83**, 054401 (2020).

[33] A. D. King, S. Suzuki, J. Raymond, A. Zucca, T. Lanting, F. Altomare, A. J. Berkley, S. Ejtemaee, E. Hoskinson, S. Huang, E. Ladizinsky, A. J. R. MacDonald, G. Marsden, T. Oh, G. Poulin-Lamarre, M. Reis, C. Rich, Y. Sato, J. D. Whittaker, J. Yao *et al.*, Coherent quantum annealing in a programmable 2,000 qubit Ising chain, Nat. Phys. **18**, 1324 (2022).

[34] C. C. McGeoch, R. Harris, S. P. Reinhardt, and P. I. Bunyk, Practical annealing-based quantum computing, Computer **52**, 38 (2019).

[35] K. Boothby, P. Bunyk, J. Raymond, and A. Roy, Next-generation topology of D-Wave quantum processors, arXiv:2003.00133.

[36] H. Hirai, T. Horiba, S. Shirai, K. Kanno, K. Omiya, Y. O. Nakagawa, and S. Koh, Molecular structure optimization based on electrons–nuclei quantum dynamics computation, ACS Omega **7**, 19784 (2022).

[37] T. Kosugi, H. Nishi, and Y.-I. Matsushita, Exhaustive search for optimal molecular geometries using imaginary-time evolution on a quantum computer, npj Quantum Inf. **9**, 112 (2023).

[38] T. Kanao and H. Goto, Simulated bifurcation for higher-order cost functions, Appl. Phys. Express **16**, 014501 (2023).

[39] R. Tanburn, E. Okada, and N. Dattani, Reducing multi-qubit interactions in adiabatic quantum computation without adding auxiliary qubits. Part 1: The "deduc-reduc" method and its application to quantum factorization of numbers, arXiv:1508.04816.

[40] D. Bertsimas and J. Tsitsiklis, Simulated annealing, Stat. Sci. **8**, 10 (1993).

[41] S. A. Solla, G. B. Sorkin, and S. R. White, Configuration space analysis for optimization problems, in *Disordered Systems and Biological Organization*, edited by E. Bienenstock, F. F. Soulié, and G. Weisbuch (Springer, Berlin, Heidelberg, 1986), pp. 283–293.

[42] R. Eglese, Simulated annealing: A tool for operational research, Eur. J. Oper. Res. **46**, 271 (1990).

[43] D. Henderson, S. H. Jacobson, and A. W. Johnson, The theory and practice of simulated annealing, in *Handbook of Metaheuristics*, edited by F. Glover and G. A. Kochenberger (Springer US, Boston, 2003), pp. 287–319.

[44] C. McGeoch and P. Farré, *Advantage Processor Overview, D-Wave Technical Report Series 14-1058A-A* (D-Wave Systems Inc., 3033 Beta Ave, Burnaby, BC V5G 4M9 Canada, 2022).

[45] V. Choi, Minor-embedding in adiabatic quantum computation: I. The parameter setting problem, Quantum Inf. Process. **7**, 193 (2008).

[46] V. Choi, Minor-embedding in adiabatic quantum computation: II. minor-universal graph design, Quantum Inf. Process. **10**, 343 (2011).

[47] S. Zbinden, A. Bärtschi, H. Djidjev, and S. Eidenbenz, Embedding algorithms for quantum annealers with chimera and pegasus connection topologies, in *High Performance Computing*, edited by P. Sadayappan, B. L. Chamberlain, G. Juckeland, and H. Ltaief (Springer International Publishing, Cham, 2020), pp. 187–206.

[48] E. Pelofske, G. Hahn, and H. Djidjev, Advanced unembedding techniques for quantum annealers, in *2020 International Conference on Rebooting Computing (ICRC)* (Atlanta, GA, USA, 2020), pp. 34–41.

[49] L. Prielinger, A. Hartmann, Y. Yamashiro, K. Nishimura, W. Lechner, and H. Nishimori, Two-parameter counter-diabatic driving in quantum annealing, Phys. Rev. Res. **3**, 013227 (2021).

[50] T. Kadowaki and H. Nishimori, Greedy parameter optimization for diabatic quantum annealing, Philos. Trans. R. Soc., A **381**, 20210416 (2023).

[51] E. B. Tadmor, R. S. Elliott, J. P. Sethna, R. E. Miller, and C. A. Becker, The potential of atomistic simulations and the knowledgebase of interatomic models, JOM **63**, 17 (2011).

[52] N. Bernardes, Theory of solid Ne, A, Kr, and Xe at 0 K, Phys. Rev. **112**, 1534 (1958).

[53] R. S. Elliott and E. B. Tadmor, Knowledgebase of interatomic models (KIM) application programming interface (API), https://openkim.org/kim-api (2011).

[54] E. Tadmor, Driver for the Lennard-Jones model uniformly shifted to have zero energy at the cutoff radius v004, OpenKIM, https://doi.org/10.25950/bdffd6a6 (2020).

[55] E. Tadmor, Lennard-Jones model (shifted) for Kr with parameters from Bernardes (1958) (low precision cutoff) v004, OpenKIM, https://doi.org/10.25950/29c4bf9b (2020).

[56] H. Chen and D. A. Lidar, Why and when pausing is beneficial in quantum annealing, Phys. Rev. Appl. **14**, 014100 (2020).

[57] Z. Gonzalez Izquierdo, S. Grabbe, H. Idris, Z. Wang, J. Marshall, and E. Rieffel, Advantage of pausing: Parameter setting for quantum annealers, Phys. Rev. Appl. **18**, 054056 (2022).

[58] C. G. Broyden, The convergence of a class of double-rank minimization algorithms 1. General considerations, IMA J. Appl. Math. **6**, 76 (1970).

[59] R. Fletcher, A new approach to variable metric algorithms, Comput. J. **13**, 317 (1970).

[60] D. Goldfarb, A family of variable-metric methods derived by variational means, Math. Comput. **24**, 23 (1970).

[61] D. F. Shanno, Conditioning of quasi-Newton methods for function minimization, Math. Comput. **24**, 647 (1970).

[62] F. H. Stillinger and T. A. Weber, Computer simulation of local order in condensed phases of silicon, Phys. Rev. B **31**, 5262 (1985).

[63] F. H. Stillinger and T. A. Weber, Erratum: Computer simulation of local order in condensed phases of silicon, Phys. Rev. B **33**, 1451(E) (1986).

[64] M. Wen, S. N. Shirodkar, P. Plecháč, E. Kaxiras, R. S. Elliott, and E. B. Tadmor, A force-matching Stillinger-Weber potential for MoS_2: Parameterization and Fisher information theory based sensitivity analysis, J. Appl. Phys. **122**, 244301 (2017).

[65] M. Wen, Stillinger-Weber model driver for monolayer MX2 systems v001, OpenKIM, https://doi.org/10.25950/7d664757 (2018).

[66] Y. Kurniawan, C. Petrie, K. Williams, M. K. Transtrum, R. S. Elliott, E. B. Tadmor, D. S. Karls, and M. Wen, Modified Stillinger-Weber potential (MX2) for monolayer MoS_2 by Kurniawan *et al.* (2022) v000, OpenKIM, https://doi.org/10.25950/328bfabb (2022).

[67] Y. Kurniawan, C. L. Petrie, K. J. Williams Jr., M. K. Transtrum, E. B. Tadmor, R. S. Elliott, D. S. Karls, and M. Wen, Bayesian, frequentist, and information geometric approaches to parametric uncertainty quantification of classical empirical interatomic potentials, J. Chem. Phys. **156**, 214103 (2022).

[68] E. Boros and A. Gruber, On quadratization of pseudo-Boolean functions, arXiv:1404.6538.

[69] M. Anthony, E. Boros, Y. Crama, and A. Gruber, Quadratic reformulations of nonlinear binary optimization problems, Math. Program. **162**, 115 (2017).

[70] N. Dattani, Quadratization in discrete optimization and quantum mechanics, arXiv:1901.04405.

[71] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, Phys. Rev. B **47**, 558 (1993).

[72] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[73] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. **6**, 15 (1996).

[74] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[75] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59**, 1758 (1999).

[76] V. V. Gusev, D. Adamson, A. Deligkas, D. Antypov, C. M. Collins, P. Krysta, I. Potapov, G. R. Darling, M. S. Dyer, P. Spirakis, and M. J. Rosseinsky, Optimality guarantees for crystal structure prediction, Nature (London) **619**, 68 (2023).

[77] P. P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale, Ann. Phys. **369**, 253 (1921).