Ground State Determination, Ground state preserving fit for Cluster Expansion and their integration for robust CE construction

by
Wenxuan Huang

Submitted to the Department of Materials Science and Engineering in Partial Fulfillment of the Requirements for the Degree of
Ph.D. in Materials Science and Engineering

at the
Massachusetts Institute of Technology
February 2018

Signature of Author.. Signature redacted
Wenxuan Huang
Department of Materials Science and Engineering
Sept 15, 2017

Certified by .... Signature redacted
Gerbrand Ceder
Visiting Professor of Materials Science and Engineering
Thesis Supervisor

Accepted by .... Signature redacted ....
Donald Sadoway
Professor of Materials Science and Engineering
Chair, Materials Science & Engineering Committee on Graduate Studies

Ground State Determination, Ground state preserving fit for Cluster Expansion and their integration for robust CE construction

by
Wenxuan Huang

Submitted to the Department of Materials Science and Engineering and in Partial
Fulfillment of the Requirement for the Degree of
Ph.D. in Materials Science and Engineering

Abstract:

In this thesis, we propose strategies to solve the general ground state problem for arbitrary effective cluster interactions and construct ground state preserving cluster expansions. A full mathematical definition of our problem has been formalized to illustrate its generality and clarify our discussion. We review previous methods in material science community: Monte Carlo based method, configurational polytope method, and basic ray method. Further, we investigate the connection of the ground state problem with deeper mathematical results about computational complexity and NP-hard combinatorial optimization (MAX-SAT). We have proposed a general scheme, upper bound and lower bound calculation to approach this problem. Firstly, based on the traditional configurational polytope method, we have proposed a method called cluster tree optimization method, which eliminates the necessity of introducing an exponential number of variables to counter frustration, and thus significantly improves tractability. Secondly, based on convex optimization and finite optimization without periodicity, we have introduced a beautiful MAX-MIN method to refine lower bound calculation. Finally, we

present a systematic and mathematically sound method to obtain cluster expansion models that are guaranteed to preserve the ground states of the reference data.

Acknowledgements

First and foremost, I would like to thank my advisor, Professor Gerbrand Ceder for his support and guidance throughout my PhD study. I am very grateful to his kindness and support in my personal and academic life events. Furthermore, I am really grateful for his effort in helping me to identify projects that suit my personal passion the most. It would be really hard for me to stretch my skills in both material science and mathematics this far without his support and guidance.

I would like to thank Professor Jeffrey Grossman and Professor Alfredo Alexander-Katz, for their time and effort being in the thesis committee. I get a lot of inspiration, hope and confidence every time from their supportive guidance and discussions.

I would like to thank my collaborators, Ziqin Rong, Daniil Kitchaev, Alexander Urban, Hena Das, Wenhao Sun, Penghao Xiao, Tina Chen, Junhe Chen, Huiwen Ji, Alexandra Toumar, Nong Artrith, Julia Yang, Olga Kononova, Zhiwei Ding, Shan Cao, Chuan Luo, William Richards and Stephen Dacek and Aziz Abdellahi. Our discussions and works together have boosted my research significantly.

I thank my parents for their full devoted support to my growth and education both financially and emotionally, which is indispensible to my achievements. Without their all in financial support to pay back all my tuitions during undergraduate study, I would not have a chance to come to MIT to finish my PhD degree.

Finally, I am very grateful to my dearest wife Jingling Liu for her willingness to join and devote to our odyssey of life, academics careers and achievements. Her emotional support and encouragement is the power that propels me to perform better and better everyday for greater and greater achievements. Life is very uncertain for starting PhD students. Yet, she is willing to commit herself and marry me in the very first year of my PhD study regardless of all of our difficulties. I really thank her and will bring her a brighter and brighter future.

Contents

Chapter 1 Introduction and Background

Section 1.1 Backgrounds

Lattice models have wide applicability in science [1-10], and have been used in a wide range of applications, such as magnetism [11], alloy thermodynamics [12], fluid dynamics [13], phase transitions in oxides [14], and thermal conductivity [15]. A lattice model, also referred to as generalized Ising model [16] or cluster expansion (CE) [12], is the discrete representation of materials properties, e.g., formation energies, in terms of lattice sites and site interactions. In first-principles thermodynamics, lattice models take on a particularly important role as they appear naturally through a coarse graining of the partition function [17] of systems with substitutional degrees of freedom. As such, they are invaluable tools for predicting the structure and phase diagrams of crystalline solids based on a limited set of *ab-initio* calculations [18-22]. In particular, the ground states of a lattice model determine the 0K phase diagram of the system. However, the procedure to find and prove the exact ground state of a lattice model, defined on an arbitrary lattice with any interaction range and number of species remains an unsolved problem, with only a limited number of special-case solutions known in the literature [23-29]. In general systems, an approximation of the ground state is typically obtained from Monte Carlo simulations, which by their stochastic nature can prove neither convergence nor optimality. Thus, in light of the wide applicability of the generalized Ising model, an efficient approach to finding and proving its true ground states would not only resolve long-standing uncertainties in the field and give significant insight into the behavior of lattice models, but would also facilitate their use in ab-initio thermodynamics.

In the first part of the thesis, we present an efficient algorithm that, in many cases, is able to find the global ground state of an arbitrary lattice model in any dimension and of any complexity, and to prove the optimality of the solution. We first introduce the formal structure of a general lattice model and the Hamiltonian used to the describe it. We proceed to derive a solution to this optimization problem by converging a periodicity-constrained upper bound and aperiodic lower bound on the total energy. For calculating the upper bound, we derive the equivalence between the optimization of the Hamiltonian under a fixed periodicity and MAX-SAT pseudo-Boolean optimization [30, 31] (PBO or MAXSAT), allowing us to leverage existing highly optimized and mathematically rigorous programing tools. To obtain the lower bound on the ground state energy, we demonstrate the two methods we develop throughout my PhD program. In the first method [32], we propose the cluster tree optimization algorithm. Compared to the previous state-of-the-art polytope method [33, 34], this algorithm eliminates the necessity of introducing an exponential number of variables to counter frustration, and thus significantly improves tractability. For the second method [35], we derive a computationally efficient approach based on a maximization of minimum-energy local configurations, which we called the MAXMIN method. We demonstrate the accuracy, robustness and efficiency of the MAXMIN approach using both an assortment of random Hamiltonians and an example of a realistic Hamiltonian of an existing material. Finally, while we are unable to guarantee that the global optimum can always be found and proven, we argue that our MAXMIN algorithm significantly improves on the state-of-the-art in computational efficiency, and does not sacrifice any guarantees of optimality with respect to established methodology.

The second part of the thesis focuses on construction of the ground state preserving CE: First-principles based cluster expansion models are the dominant approach in *ab initio* thermodynamics of crystalline mixtures enabling the prediction of phase diagrams and novel ground states [36]. However, despite recent advances, the construction of accurate models still requires a careful and time-consuming manual parameter tuning process for ground state preservation, since this property is not guaranteed by default. In this part of the thesis, we present a systematic and mathematically sound method to obtain cluster expansion models that are guaranteed to preserve the ground states of their reference data. The method builds on the recently introduced compressive sensing [37, 38] paradigm for cluster expansion and employs quadratic programming (QP) [39] to impose constraints on the model parameters. The robustness of our methodology is illustrated for two lithium transition-metal oxides with relevance for Li-ion battery cathodes, i.e., $Li_{2x}Fe_{2(1-x)}O_2$ and $Li_{2x}Ti_{2(1-x)}O_2$, for which the construction of cluster expansion models with compressive sensing alone has proven to be challenging. We demonstrate that our method not only guarantees ground state preservation on the set of reference structures used for the model construction, but also show that out-of-sample ground state preservation up to relatively large super cell size is achievable through a rapidly converging iterative refinement. This method provides a general tool for building robust, compressed and constrained physical models with predictive power.

Section 1.2 General statement of the ground state problem

A lattice model is a set of fixed sites on which objects (spins, atoms of different types, atoms and vacancies, etc.) are to be distributed. Its Hamiltonian consists of coupling terms between pairs, triplets, and other groups of sites, which we refer to as "clusters". A formal definition of effective cluster interactions can be found in [12]. Before discussing the algorithmic details of our method, it is essential to establish a precise mathematical definition of a general lattice model Hamiltonian and the task of determining its ground states. The ground state problem can formally be stated as follows: Given a set of effective cluster interactions (ECI's) $J \in \mathbf{R}^{\mathbf{C}}$, where $\mathbf{C}$ is the set of interacting clusters and $\mathbf{R}$ is the set of real numbers, what is the configuration $s:\mathbf{D}\rightarrow\{0,1\}$, where $\mathbf{D}$ is the domain of configuration space, such that the global Hamiltonian H is minimized:

$$H=\lim_{N\to\infty}\frac{1}{(2N+1)^3}\sum_{(i,j,k)\in\{-N,\dots,N\}^3}\sum_{\alpha\in\mathbf{C}}J_\alpha\prod_{(x,y,z,p,t)\in\alpha}s(i+x,j+y,k+z,p,t) \tag{1.1}$$

In the Hamiltonian of Eq. (1.1), each $\alpha\in\mathbf{C}$ is an individual interacting cluster of sites. In turn, each site within α is defined by a tuple (x,y,z,p,t), wherein (x,y,z) is the index of the primitive cell containing the interacting site, p denotes the index of the sub-site to distinguish between multiple sub-lattices in that cell, and t is the species occupying the site. To discretize the interactions, we introduce the "spin" variables $s_{x,y,z,p,t}$, where $s_{x,y,z,p,t}=1$ indicates that the p^{th} sub-site of the (x,y,z) primitive cell is occupied by species t, and otherwise $s_{x,y,z,p,t}=0$. The energy can be represented in terms of spin products, where each cluster α is associated with an ECI J_α denoting the energy associated with this particular cluster. To obtain the energy of the entire system, each cluster needs to be translated over all possible periodic images of the primitive cell, i.e., we have to consider all possible translations of the interacting cluster α, defined as a set of (x,y,z,p,t), by (i,j,k) lattice primitive cells translations, yielding the spin product $\prod_{(x,y,z,p,t)\in\alpha} s_{i+x,j+y,k+z,p,t}$. Finally, the prefactor $\frac{1}{(2N+1)^3}$ normalizes the energy to one lattice primitive cell, and the limit of N approaching infinity emphasizes our objective of minimizing the average energy over the entire infinitely large lattice. One remaining detail is that the Hamiltonian given in Eq. (1.1) is constrained such that that each site in the lattice must be occupied. For the sake of simplicity, lattice vacancies are included as explicit species in the Hamiltonian, so that all spin variables associated with the same site sum up to one:

$$\sum_{t\in\mathbf{c}(p)} s_{x,y,z,p,t} = 1\ \forall(x,y,z,p)\in\mathbf{F} \qquad (1.2)$$

In Eq. (1.2), $\mathbf{F}$ is the set of all sites in the form of (x,y,z,p), and $\mathbf{c}(p)$ denotes the set of species that can occupy sub-site p. The domain of configuration space $\mathbf{D}$ can be formally defined as the set of all (x,y,z,p,t), with $t\in\mathbf{c}(p)$.

To further illustrate the notation introduced above, **Figure 1** depicts an example of a two-dimensional lattice Hamiltonian for a square lattice with two sub-sites in each lattice primitive cell, i.e., $p\in\{0,1\}$. Each sub-site may be occupied by 3 types of species, so that $t\in\{0,1,2\}$, where $t=0$ shall be the reference (for example, vacancy) species. Hence, the energy of the system relative to the reference can be encoded into $t\in\{1,2\}$. Furthermore, the Hamiltonian shall be defined by only 2 different pairwise interaction types with the associated clusters $\alpha=\{(0,0,0,1,2),(1,2,0,0,1)\}$ and $\beta=\{(0,1,0,0,2),(0,0,0,1,2)\}$, and thus the set of all clusters is $\mathbf{C}=\{\alpha,\beta\}$. The first three of the five indices between "( )" brackets indicate the initial unit cell position, the forth index corresponds to the position in the unit cell (sub-site index), and the last index gives the species. The third component of the cell index (*x,y,z*) was retained for generality but set to 0 for this two-dimensional example. The example configuration shown in **Figure 1** depicts three specific interactions: The interaction represented on the bottom left in in the figure is of type α with $(i,j,k)=(0,0,0)$, corresponding to the spin product $J_\alpha s_{0,0,0,1,2} \cdot s_{1,2,0,0,1}$. The interaction in the center of the figure also belongs to type α but with $(i,j,k)=(1,1,0)$, corresponding to the spin product $J_\alpha s_{0+1,0+1,0,1,2} \cdot s_{1+1,2+1,0,0,1} = J_\alpha s_{1,1,0,1,2} \cdot s_{2,3,0,0,1}$. Lastly, the interaction on the right represents an interacting β cluster, with $(i,j,k)=(3,0,0)$, yielding a spin product of $J_\beta s_{0+3,1,0,0,2} s_{0+3,0,0,1,2} = J_\beta s_{3,1,0,0,2} s_{3,0,0,1,2}$.

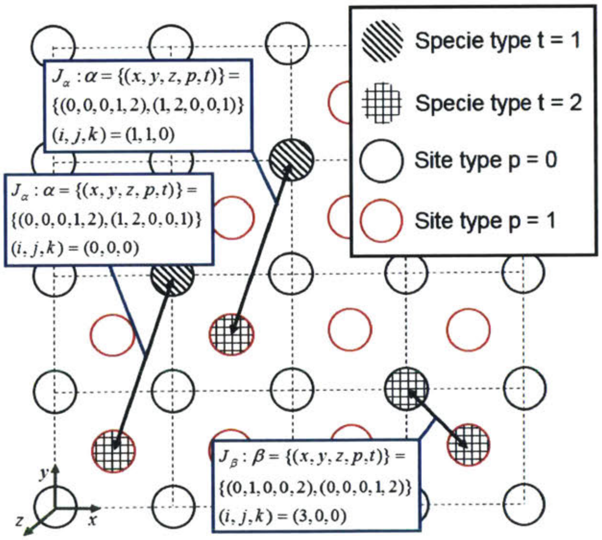

Figure 1: Illustration of a lattice Hamiltonian and examples of cluster interactions. The primitive unit of the lattice is indicated by a thin dashed line, and sites are represented by circles. Two different site types

are distinguished by black and red borders, respectively. The non-vacancy species that can occupy the sites are indicated by two different hatchings.

Currently, the most common approach to find the ground state of a generalized Ising model is simulated annealing [40] based on Metropolis Monte Carlo [41] in an *ad hoc* finite lattice cell. This approach has two major drawbacks. First, it is inherently an optimization over a finite set of sites, whereas the true objective function is defined over an infinite number of sites. Second, the result obtained from a Monte Carlo calculation is simply a particular low-energy configuration, a local minimum of energy, with no guarantee that it is the true ground state. This limitation becomes especially problematic when the size of the ground state structure increases, since the large number of degrees of freedom quickly renders it infeasible to sample the low-energy configurations in the cell. Hence, due to its stochastic nature and dependence on a particular lattice cell, simulated annealing can only identify possible ground state candidates, but it can hardly guarantee that the global ground state has been found.

An alternative approach that provides a provable ground state is the configurational polytope method [33, 34] combined with vertex enumeration [42]. This method provides a beautiful reformulation of the ground state problem as linear programming. Unfortunately, this approach has not been applied to finding the ground states of complex realistic Hamiltonians due to its computational inefficiency and the fact that the method yields a polytope with an large number of "inconstructible" vertices, i.e solutions that do not correspond to realizable lattice configurations, and there is no general, tractable algorithm to extract the true constructible polytope [43, 44]. Recently, the "basic ray" method has been proposed, and used to obtain the ground states of several small systems [23-25]. However, a universal algorithm based on this method is not known [24], and the number of systems solved by this approach is limited.

In the remainder of this chapter, we review the previous established methods in more detail and illustrate our solutions to ground state problems in depth in latter chapters.

Section 1.3 Monte Carlo based method

Being one of the most important class of Monte Carlo based method, simulated annealing is an optimization scheme based on a finite lattice system. However, this method alone cannot resolve our problems. One of the reasons is that our objective is optimization on the infinite lattice.

More importantly, simply based on simulated annealing, there is no way to guarantee that the state computed is the exact ground state even for a fixed periodicity. Although there is convergence proof that in infinite time the solution from simulated annealing converges to the exact ground state [45], the actual iterations needed to guarantee (at least in the proof) with high probability would be impractical [45]. More generally, if we consider finding ground states fixing periodicity as a combinatorial optimization problem, simulated annealing is one heuristic method [46], which is designed to obtain feasible low energy solutions without proof of optimality.

Besides simulated annealing, there are other heuristics methods available for minimization over other variant of ising model, say, genetic algorithm, particle swarm algorithm [47, 48] and quantum annealing [49]. But these researches are mostly oriented to solve the simpler model: standard 2D Ising model with only external applied field and nearest neighbor interaction with random distribution of bond strength. For this simpler model there are some research on its exact minimization mostly based on branch and bound [50] and graph cuts [51]. But these are neither very relevant nor applicable to our project.

To summarize, simulated annealing algorithm is not suitable for our project because of: its limitation of finiteness and more importantly its incapability of knowing when exact optimal is reached.

Section 1.4 Configurational polytope with vertex enumeration

Traditionally, the approach that could find and prove exact ground state is the configurational polytope method by Kaburagi and Kanamori [33, 34] combined with vertex enumeration [42]. However, the draw back for the traditional polytope method is that the constructed polytope has enormous amount of "unconstructible" vertices [43, 44] and there are no general tractable algorithm to obtain the true ground state polytope simply based on this method.

The general method without going into mathematical details looks like:

1. Set up inequalities corresponding to the cluster to examined
2. Loop over all corresponding vertexes

    a. For each vertex, check whether there exists some structure satisfying this correlation, i.e. check constructability of this vertex

3. Record all the constructible structures and vertexes

The polytope method does not go very far to solve every problem. Indeed, even for the 1D problem exact solution is only available up to 5th nearest neighbor interaction [26, 52]. This is very surprising actually. (As a side note from this background, we have arrived at algorithms capable of computing ground state up to 40th nearest neighbor interaction in reasonable time.)

For square lattice, as far as we know, exact solutions are available only for interactions up to 3rd nearest neighbor interactions [27]. For ternary system, exact solutions only up to 2nd nearest neighbor are published [53]. For faced-centered rectangular lattice, although there exist partial results of ground state results to 7th nearest neighbor interaction, the result is based on a heavy assumption that $J_k > 0$ [54]. The result with weaker assumption that $J_1 > 0$ is only available up to 4th nearest neighbor interaction [55].

There are indeed interesting progress on triangular lattice, a large part of the phase diagram for interactions up to 3rd nearest neighbor has been tackled [28]. We have to realize that the solution was not completed yet: the paper claims that some configurations are constructed based on physical consideration, which is not a proof in our context [28]. Only until recently, triangular lattice with pair interaction up to 2nd nearest neighbor and a triplet term is considered fully solved [24]. Some attempts has been made in trying to solve ground state with interaction up to the 7th nearest neighbor [56], but the result is just for some limited regions of J space.

For honeycomb lattice, it is interesting to note that interaction up to 3rd nearest neighbor is fully solved [57]. It is worthy to note that for some J, there are infinitely many corresponding ground states. The reason is that, in the configuration polytope, the vertex corresponding to the ground state has infinitely many ways to be constructed. From a pure mathematical point of view, this shall not be too surprising, because in mathematical language, it simply says that there are infinitely many ways to tile the plane using some specific set of tiles [58].

In 3D cases, results on cubic lattice with up to 3rd nearest neighbor with 3-body terms has been examined [59]. However, it also has the problem of unconstructible vertex which is intrinsic in the polytope method. For bcc structure, interactions up to 4th nearest neighbor is examined [60]. There are active researches devoted into fcc ground state calculations. Up to 5th nearest neighbor interaction is considered in one of the works [43, 61], but we shall still note its incompleteness. Later, ternary [44] fcc alloys with interactions up to 2nd nearest neighbor are examined. However, severe problem of configurational polytope combined with vertex enumeration is apparent: rapid explosion of number of vertexes as number configurations increase [43] and huge number of unconstructible vertexes.

Previous paragraphs summarize the result of the most fundamental method, configurational polytope combined with vertex enumeration. It is widely used to generate ground states in the last century. Due to rapid explosion of required computing time [43], researches on exact the ground state problem do not proceed much further.

Section 1.5 Recent basic ray method

Until recently [24, 62-65], another author has proposed a method called "basic ray" method. It has been used to solve some small systems [23-25]. The method boils down to:

1. Decompose global Hamiltonian into local Hamiltonian,
2. Introduce free parameters into local Hamiltonians,
3. Find specific free parameters such that there are as many local optimal configurations (which can tile the whole plane) as possible (such parameter sets are called basic rays),
4. Find the complete set of basic rays,
5. Argue that any Hamiltonian as positive combination of a set of basic rays such that the "overlapping" local configurations could tile the whole plane attains the ground state solution constructed by the "overlapping" local configurations.

The "basic ray" method, have been successfully applied in finding devil's step (infinitely many ground states) in triangular lattice with interaction up to 3rd nearest neighbor [62]. We mentioned that it provides a complete ground state solution to triangular lattice with interaction up to 2nd nearest neighbor and a triplet term [24]. This method is also applied to completely resolve the ground state problem of Shastry-Sutherland lattice involving edge and diagonal interactions [23, 66]. Finally, it has been successfully applied to solve completely the anisotropic triangular lattice with nearest neighbor interaction [64].

Section 1.6 Tiling and Undecidability

It is less commonly known in the cluster expansion community the deep relationship between the ground state problem and the undecidability of Wang tile problem [67-69]. To rephrase the Wang tile problem: given a set of squares with colored edge, the problem is to determine the possibility for this set of tiles to tile a plane, with the constraint that neighboring coloring must match. Unfortunately, because the existence of aperiodic set of tiles, this problem is shown to be undecidable [68, 69]. An aperiodic set of tiles is defined as a set of tiles where only aperiodic tiling (and no periodic tiling) is admissible. Later, small aperiodic set with only 13 and 14 elements are constructed [70, 71]. Although the initial problem of Wang tile is based on colored edges, Wang tile of colored corner type is considered and mutual conversion has been shown [72, 73]. Thus, the tiling problem for the corner type of Wang tile is also undecidable. For completeness, we mention that a small set of corner type of Wang tiles with 44 elements have been constructed [73].

For readers unfamiliar with the exact definition of undecidable, we would here review the definition: a problem is undecidable if and only if there does not exist a Turing machine (or equivalently a C program) such that for all instances of the problem, the program is guaranteed to produce an correct yes/no answer in finite time. For more details, refer to [74].

The reader might wonder why would tiling problem has anything to do with the ground state problem. We would like to explain the relationship for both configurational polytope method and basic ray method.

In (2-a) step of configurational polytope method, to check whether a vertex corresponds to a constructible structure is equivalent to the question of determining whether a set of tile could tile the whole plane, i.e. tileable [75].

In the "basic ray" method, each basic ray is determined in a way that maximize the number of optimal tiles and the optimal tile set need to be tileable [24]. Furthermore, to guarantee that the set of basic rays is complete, we need to guarantee that every "overlapping" tile set is tileable [24].

In this way, the ground state problem is intrinsically connected to the deep theoretical computer science undecidable problem of Wang tile. And this in a way explains the fundamental difficulty of the ground state problem, and the lack of complete solutions until now.

We show below that the general ground state problem is undecidable by using the aperiodic corner tile set in mathematics [73] and constructing J such that tiling from the aperiodic set of tiles results in the lowest energy. Another author has constructed aperiodic ground state using very long-range interactions [76].

Proof of the undecidability of the ground state problem

Suppose there exists an algorithm that, given arbitrary ECI, is guaranteed to produce the ground state configuration and ground state energy of a generalized Ising model/cluster expansion. Now consider arbitrary set of corner Wang tiles. We define the ECI such that all block energies corresponding to an element inside the set of Wang tiles to be -1 and all block energies corresponding to an element outside the set to be 0. Now, input this set of ECI into the presupposed algorithm. We could then get the ground state energy and ground state configuration. If the ground state energy is larger than -1, we can conclude that the set of tiles could not tile the plane. Otherwise, the ground state energy is -1, and we have the ground state spin configuration. Using arguments analogous to [77], we could show that there exists a tiling composed of only elements in the tile set and thus the tile set could tile the plane. Thus, the algorithm to calculate ground state corresponding to the given ECI can be modified to decide whether a given set of Wang tile can tile the plane, violating the undecidability of the Wang tile problem. Thus, the ground state problem must be undecidable. ■

Although we conclude that the general ground state problem is undecidable, this does not stop us from continuing to work on the ground state problem. This is because undecidability simply means that for any program we constructed, there might be some instances that are too hard to be solved exactly. But the instances used to prove undecidability involves very peculiar interaction terms, in general material science practice, we seldom come into those cases. Furthermore, a bounded approximation of ground state is always attainable. If we ever come into some J such that the ground state is aperiodic and our program cannot decide it, this is a new discovery anyway, and we might possibly apply advanced mathematical methods on tiling to decide it.

Chapter 2 General solution scheme and upper bound method

Section 2.1 General solution scheme

Our general scheme for finding an exact ground state of an Ising model is to calculate and converge upper and lower bounds on the energy. We note that the energy of any periodic configuration is an upper bound on the ground state energy [78]. Thus, by enumerating periodicities and finding the exact ground state for each, we can successively tighten the upper bound on the true ground state energy. If an exact periodic ground state structure exists, we are guaranteed to obtain the tightest upper bound possible once we reach the true periodicity by enumeration. However, there is no way of knowing when this condition has been reached, i.e., when the enumeration should stop. We therefore require an additional procedure to construct successfully tighter, rigorous lower bounds on the ground state energy, so that periodicity enumeration can be stopped when the upper and lower bounds match, indicating that the exact ground state has been found. Our approach for the construction of the lower bound is less intuitive and involves the optimization over a non-periodic domain. We discuss both upper and lower bound procedure separately.

Note that cluster expansions are often applied in both canonical and grand-canonical contexts, i.e., investigating the ground states that arise at both fixed composition and fixed chemical potential [79, 80]. Our derivation focuses on the grand-canonical case, where the chemical potential of each species is fixed by the "single-point" interaction terms. However, the grand-canonical solution obtained this way can be readily used to obtain the canonical ground states using a convex hull approach [81]. By construction it removes compositions that do not have a stable ground state. In thermodynamic language we minimize the Legendre transform of the energy (grand potential) with respect to composition, rather than the energy itself, as is the common procedure to find ground states as function of composition [79].

Section 2.2 Enumerating periodicity

We begin our optimization of the upper bound on the energy by enumerating all distinct periodicities up to a chosen maximal unit cell size, which can be iteratively increased until convergence of the upper bound and lower bound has been obtained. Note that all distinct periodic orderings on a lattice can be represented by an all-integer supercell matrix in Hermite normal form [82], whose determinant represents the size of the periodic super cell. Thus, we can systematically enumerate all periodicities by generating supercells of the lattice primitive cell from all integer Hermite normal form matrices with up to a given determinant. We then proceed to solve the fixed-periodicity ground state problem within each generated supercell, achieving successively tighter upper bounds on the infinite-lattice ground state energy.

Section 2.3 Obtaining the ground state at a fixed periodicity

The first element of our solution to the ground state problem is to efficiently find the ground state given a fixed periodicity of the solution. While this problem is typically solved by Metropolis Monte Carlo (MC) simulated annealing in a prescribed simulation cell, this approach cannot prove that the periodic solution found is in fact optimal, even for a given periodicity.

To arrive at a methodology that yields a provably optimal solution, we convert the problem of minimizing the Hamiltonian into a mathematical programming problem. The advantage of this approach is that mathematical programming algorithms not only yield good performance, but also require a rigorous proof of solution correctness, i.e. optimality, before termination. Classic examples of mathematical programming algorithms are the simplex method in linear programming [83] and the branch and bound method for integer programming [84], where the algorithm itself is also a schematic of the proof of optimality. As we will show in the following, the ground state problem for a fixed periodicity can be transformed into a maximum satisfiability problem [85], a well-researched class of optimization problems for which highly efficient solvers exist [30, 31].

Using the notation introduced in the previous section, we note that calculating the periodic ground state is equivalent to solving the finite optimization problem:

$$\min_{\widehat{s}} \sum_{\alpha \in \tilde{\mathbf{C}}} J_{\alpha} \prod_{(x,y,z,p,t)\in\alpha} \widehat{s}_{x,y,z,p,t} \tag{2.1}$$

subject to:

$$\sum_{t \in c(p)} \widehat{s}_{x,y,z,p,t} = 1 \ \forall (x,y,z,p) \in \mathbf{F}_{\text{finite}} \tag{2.2}$$

where $\widehat{s}_{x,y,z,p,t}$ is, as s in Eq. (1.2), the indicator variable of species t on site (x,y,z,p), with the difference that $\widehat{s}$ is now defined on a smaller domain determined by the periodicity, $\tilde{\mathbf{C}}$ is the set of all interacting clusters within the fixed periodic system and $\mathbf{F}_{\text{finite}}$ is the set of sites within the fixed periodic unit cell. Such an optimization over discrete $\{0,1\}$ variables can be equivalently posed as a logic problem by converting the minimization problem into the negative of a maximization problem and replacing the discrete variables by Boolean equivalents. Following this insight, the minimization of the finite Hamiltonian can thus be expressed in the form of a pseudo-boolean optimization (PBO) problem, allowing us to solve this optimization as a weighted partial maximum satisfiability (MAX-SAT) [30, 31] problem. The essence of MAX-SAT is to model the discrete optimization problem by maximizing the number of logical clauses that can be satisfied in a Boolean formula of conjunctive normal form, weighted by a set of arbitrary coefficients.

To illustrate this approach, we consider the example of a binary 1D system with a positive point term J_0 and a negative nearest-neighbor interaction J_{NN}, on a 2-site unit cell. For this system, the transformation is:

$$\begin{aligned}
E &= \min_{\widehat{s}_0,\widehat{s}_1}\left(J_0\widehat{s}_0 + J_0\widehat{s}_1 + J_{NN}\widehat{s}_0\widehat{s}_1\right) \\
&= -\max\left(J_0\left(1-\widehat{s}_0\right) - J_0 + J_0\left(1-\widehat{s}_1\right) - J_0 - J_{NN}\left(\widehat{s}_0\widehat{s}_1\right)\right) \\
&= -\max\left(J_0(\neg\widehat{s}_0) - J_0 + J_0\left(\neg\widehat{s}_1\right) - J_0 + \left(-J_{NN}\right)\left(\left(1-\neg\widehat{s}_0\right)\widehat{s}_1\right)\right) \\
&= -\max\left(J_0(\neg\widehat{s}_0) - J_0 + J_0\left(\neg\widehat{s}_1\right) - J_0 + \left(-J_{NN}\right)\widehat{s}_1 + \left(-J_{NN}\right)\left(1-\neg\widehat{s}_0\widehat{s}_1\right) - \left(-J_{NN}\right)\right) \\
&= \left(2J_0 - J_{NN}\right) - \mathrm{MAXSAT}\left(J_0(\neg\widehat{s}_0) \wedge J_0(\neg\widehat{s}_1) \wedge \left(-J_{NN}\right)(\widehat{s}_1) \wedge \left(-J_{NN}\right)(\widehat{s}_0 \vee \neg\widehat{s}_1)\right)
\end{aligned}$$

where the indicator variable $\widehat{s}_i$ is now also a Boolean variable in the MAX-SAT setting, and the $\wedge$, $\vee$ and $\neg$ operators correspond to logical "and", "or" and "not" respectively. Note that, although in a MAX-SAT problem the coefficient of each clause needs to be positive, it is still possible to transform an arbitrary set of cluster interactions J_i into a proper MAX-SAT input, as in the example above.

The advantage of formulating the ground-state problem in this form is that MAX-SAT is one of the most actively researched NP-hard problems [86], allowing us to leverage the extensive literature written on the topic [87-90]. Note that any complete MAX-SAT solver encodes a proof of optimality [88] and includes a published proof of algorithm correctness (that it is guaranteed to find the optimal solution) [30] and efficiency [89, 90]. Furthermore, the algorithms are run through an annual MAX-SAT competition [87], which tests their correctness, robustness and efficiency. Under such stringent criteria, by converting our problem into MAX-SAT, we can safely guarantee provability, as well as further investigate the advanced proof schemes and fast algorithms developed over the last twenty years of MAX-SAT research [30, 31, 88]. The particular MAX-SAT solver we choose, based on the results of the MAX-SAT 2014 competition benchmarking and our own testing, is CCLS_to_akmaxsat [31, 91].

Another notable advantage of MAX-SAT over MC for obtaining a solution of the ground-state problem is that state-of–the-art MAX-SAT solvers generally include sophisticated methods to escape from local minima [89, 90] to arrive at the global minimum faster and more robustly than MC.

To verify the efficiency, robustness, and accuracy of the MAX-SAT solver compared to conventional MC, we performed a series of tests of both algorithms. We constructed a set of random 1D and 2D pair-interaction Hamiltonians with interactions up to the 28th nearest neighbor in 1D and up to the 10th nearest neighbor in 2D. We then attempted to find the ground state of each system within unit cells containing up to 50 sites by MAX-SAT and Monte Carlo. Finally, we considered only those Hamiltonians which could be classified as "difficult", which we define as having a ground state unit cell with more than 4 sites in 1D, or more than 12 sites in 2D. Among these "difficult" Hamiltonians, while our MAX-SAT approach consistently provides a provable ground state under the imposed periodicity constraints, we find that MC is unable to find the ground state energy comparable to the MAX-SAT result in 10% of cases. Thus, the MAX-SAT approach by itself is an attractive method to obtain provably optimal, periodically-constrained ground states up to some maximum unit cell size, separate from the problem of proving the optimality of the solution over infinite space.

Chapter 3 Lower bound method 1: Cluster tree optimization algorithm

Section 3.1 General introduction of cluster tree optimization algorithm

In this chapter, we present a general approach to the ground state problem, which we refer to as the "cluster tree optimization algorithm." We demonstrate that this algorithm is guaranteed to construct and prove, within an arbitrarily small numerical factor, the exact ground state for an arbitrary multicomponent set of ECIs on an arbitrary lattice system, assuming that a periodic ground state exists. We derive the algorithm by systematically constructing higher order polytopes without introducing exponentially many variables. Finally, we show that even in the case that the true ground state is aperiodic, our approach yields a series of converging spin configurations within an arbitrarily small margin of the true optimum. Compared with the state-of-the-art configurational polytope method [33, 34], our method moves from correlation space to appearance-frequency space. This conversion allows us to incrementally establish higher order configurational constraints, which is not possible in the traditional method. This conversion, together with a detailed implementation of the cluster tree algorithm, provides a systematic approach to deriving at the exact ground state of any cluster expansion.

Finally, we note that cluster tree optimization represents a useful procedure to approximate the generally undecidable Wang tile problem. Our method offers an effective procedure to determine tilability by converting the original problem into a series of efficient linear programming steps, providing a measure of tilability in the form of energy and a general direction to how the tiling could be constructed.

Section 3.2 Methodology formalism

We begin by formally introducing the cluster tree optimization algorithm. We first show that for the purposes of cluster interactions, any lattice can be mapped to an orthorhombic multicomponent lattice without any symmetry in its interactions. We then prove that the total energy of this system can be written in terms of the energies of blocks of lattice sites. We proceed to define the basic polytope method for solving the ground state of such a system. Finally, we derive the "cluster tree optimization algorithm" and prove the correctness and generality of the method.

First, note that a binary generalized Ising Model on an arbitrary lattice with an arbitrary motif of n sites can always be represented by a generalized Ising model on an orthorhombic lattice with 2^n components without any symmetry.

Second, we introduce the notion of a **"block"**. A block is a local configuration – for example, (1,0,1,1) is a block in 1D binary system where a lattice site can be occupied by two species that we label "0" and "1". We define a **"minimal block"** as the smallest block that encapsulates all the interactions in the system – for example, (0,0,0), (0,0,1), and (0,1,0) would all be minimal blocks for a 1D system with interaction up to the next nearest neighbor.

Third, we introduce the term **"energy of a block"** in order to use this change of basis for rewriting the Hamiltonian in terms of the **energies of blocks** and **appearance frequency of blocks**. As an example, consider the Ising Hamiltonian of a 1D lattice with nearest neighbor, next nearest neighbor, and triplet interactions. This Hamiltonian can be transformed into a sum over spin-configurations, multiplying of block energies by their appearance frequencies:

$$\begin{aligned}H &= \mu\sum_{i\in\mathbb{Z}}\sigma_i + J_N\sum_{i\in\mathbb{Z}}\sigma_i\sigma_{i+1} + J_{NN}\sum_{i\in\mathbb{Z}}\sigma_i\sigma_{i+2} + J_{triplet}\sum_{i\in\mathbb{Z}}\sigma_i\sigma_{i+1}\sigma_{i+2}\\ &= \sum_{i\in\mathbb{Z}}\left(\mu\sigma_i + J_N\sigma_i\sigma_{i+1} + J_{NN}\sigma_i\sigma_{i+2} + J_{triplet}\sigma_i\sigma_{i+1}\sigma_{i+2}\right)\\ &= \sum_{i\in\mathbb{Z}}E(\sigma_i,\sigma_{i+1},\sigma_{i+2}) = \sum_{(\sigma_1,\sigma_2,\sigma_3)}E(\sigma_1,\sigma_2,\sigma_3)\rho(\sigma_1,\sigma_2,\sigma_3)\end{aligned}$$

where μ are the point energies, J_N are the nearest neighbor interactions, J_{NN} are the next nearest neighbor interactions, $J_{triplet}$ are the triplet interactions, σ are spins, $\rho(\sigma_1,\sigma_2,\sigma_3)$ are the appearance frequencies for blocks $(\sigma_1,\sigma_2,\sigma_3)$, and $E(\sigma_1,\sigma_2,\sigma_3)$ are the block energies. To further illustrate the definition of appearance frequency, consider the periodic 1D configuration "---001001001001001---". In this configuration, the appearance frequencies would be $\rho[0]=\frac{2}{3},\rho[1]=\frac{1}{3},\rho[00]=\frac{1}{3},\rho[01]=\frac{1}{3}\ \rho[10]=\frac{1}{3},\rho[11]=0$, and so on.

Similar arguments lead to results for 2D and 3D systems:

$$H=\sum_{(i,j)\in\mathbb{Z}^2} E\begin{pmatrix}\sigma_{i,j} & \cdots & \sigma_{i,j+M}\\ \vdots & \ddots & \vdots \\ \sigma_{i+N,j} & \cdots & \sigma_{i+N,j+M}\end{pmatrix} = \sum_{\{\sigma\}} E\begin{pmatrix}\sigma_{1,1} & \cdots & \sigma_{1,1+M}\\ \vdots & \ddots & \vdots \\ \sigma_{1+N,1} & \cdots & \sigma_{1+N,1+M}\end{pmatrix}\rho\begin{pmatrix}\sigma_{1,1} & \cdots & \sigma_{1,1+M}\\ \vdots & \ddots & \vdots \\ \sigma_{1+N,1} & \cdots & \sigma_{1+N,1+M}\end{pmatrix}$$

where the sum is over all possible configurations of $\{\sigma\}$. For the sake of brevity, we introduce a more compact notation:

$$\begin{pmatrix}\sigma_{i,j} & \cdots & \sigma_{i,j+M}\\ \vdots & \ddots & \vdots \\ \sigma_{i+N,j} & \cdots & \sigma_{i+N,j+M}\end{pmatrix} \equiv \sigma_{[i:i+N]\times[j:j+M]}$$

$$\begin{pmatrix}\sigma_{1,1} & \cdots & \sigma_{1,M}\\ \vdots & \ddots & \vdots \\ \sigma_{N,1} & \cdots & \sigma_{N,M}\end{pmatrix} \equiv \sigma_{[N]\times[M]}$$

which are adapted from mathematical convention that $[n]=\{1,2,\ldots,n\}$ and $[i:i+M]=\{i,i+1,\ldots,i+M\}$. Thus, the Hamiltonian can be rewritten as:

$$H=\sum_{(i,j)\in\mathbb{Z}^2} E\left(\sigma_{[i:i+N]\times[j:j+M]}\right) = \sum_{\sigma_{[1+N]\times[1+M]}} E\left(\sigma_{[1+N]\times[1+M]}\right)\rho\left(\sigma_{[1+N]\times[1+M]}\right)$$

Based on these definitions, we can write down the **basic polytope method** for finding the ground state of an Ising Hamiltonian with a given set of interaction parameters. Consider a 2D system written in terms of block energies as described above, where all interactions fall within a range m by n. By defining the appearance frequency ρ for each possible block, our objective is to:

$$\min_{\rho} \quad H[\{\sigma\}] = \sum_{\sigma_{[n]\times[m]}} \rho\left[\sigma_{[n]\times[m]}\right] E\left[\sigma_{[n]\times[m]}\right] \tag{3.1}$$

However, constraints are needed on the ρ variables in order for the solution to be physical. Thus, we introduce **compatibility equations** of order m by n as constraints for ρ. Formally, **compatibility equations** of order m by n are defined as the following equations – Eq. (3.2), Eq. (3.3), and Eq. (3.4):

$$\rho\left[\sigma_{[n]\times[2:m]}\right] = \sum_{\sigma_{[n]\times[1]}} \rho\left[\;\sigma_{[n]\times[2:m]} \quad \sigma_{[n]\times[1]}\;\right] = \sum_{\sigma_{[n]\times[1]}} \rho\left[\;\sigma_{[n]\times[1]} \quad \sigma_{[n]\times[2:m]}\;\right] \tag{3.2}$$

Eq. (3.2) is a valid equality constraint on ρ based on the simple observation that whenever $\left[\sigma\right]$ appears, its next neighbor must be either 0 or 1 (in the case of the binary system described earlier), corresponding to block $\left[\sigma 0\right]$ or block $\left[\sigma 1\right]$. Thus $\rho\left[\sigma\right] = \rho\left[\sigma 0\right] + \rho\left[\sigma 1\right]$, which is exactly the constraint given by Eq. (3.2).

Furthermore, Eq. (3.2) guarantees the **constructability** of ρ in the x direction, where ρ is deemed to be constructible if it corresponds to a physical lattice configuration, which will be proven later in this paper. A pictorial illustration of Eq. (3.2) is shown in Figure 2 a.

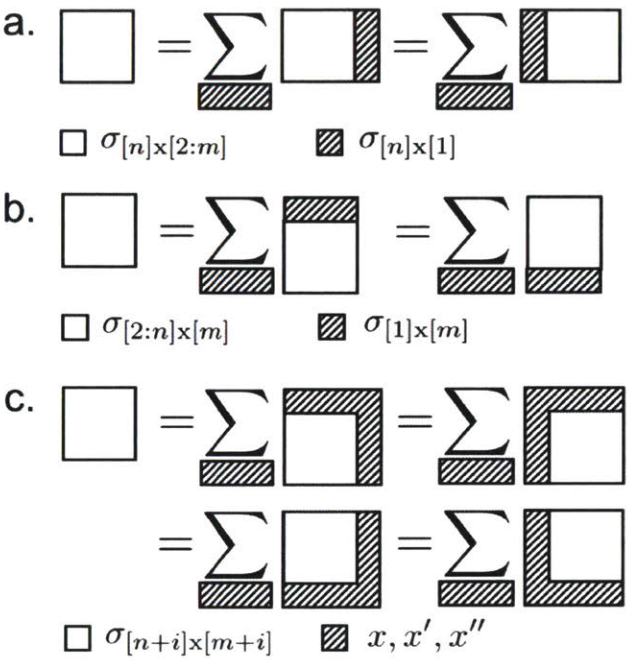
a.
□ $\sigma_{[n]\text{x}[2:m]}$
▨ $\sigma_{[n]\text{x}[1]}$
b.
□ $\sigma_{[2:n]\text{x}[m]}$
▨ $\sigma_{[1]\text{x}[m]}$
c.
□ $\sigma_{[n+i]\text{x}[m+i]}$
▨ x, x', x''

Figure 2. a. Pictorial illustration of the 1D block compatibility constraints defined in Eq. (3.2), where the white block corresponds $\sigma_{[n]\times[2:m]}$ and the hatched block corresponds to $\sigma_{[n]\times[1]}$. b. Pictorial illustration of the 2D compatibility constraint added in Eq. (3.3), where the white block corresponds $\sigma_{[2:n]\times[m]}$ and the hatched block corresponds to $\sigma_{[1]\times[m]}$. c. Pictorial illustration of the 2D perfect sum relationship, where the white block corresponds to a given [n+i] by [m+i] block, and the hatched block corresponds to all possible site configurations immediately adjacent to this block.

The next constraint is analogous, but given in the y-direction:

$$\rho\left[\sigma_{[2:n]\times[m]}\right] = \sum_{\sigma_{[1]\times[m]}} \rho\begin{bmatrix}\sigma_{[1]\times[m]}\\ \sigma_{[2:n]\times[m]}\end{bmatrix} = \sum_{\sigma_{[1]\times[m]}} \rho\begin{bmatrix}\sigma_{[2:n]\times[m]}\\ \sigma_{[1]\times[m]}\end{bmatrix} \qquad (3.3)$$

This constraint stems from similar reasoning as (3.2) and guarantees constructability of ρ in the y direction. A pictorial illustration is shown as in Figure 2 b.

The final constraint we must add is that the set of all ρ must correspond to a fully occupied lattice:

$$\sum_{\sigma_{[n]\times[m]}} \rho\left[\sigma_{[n]\times[m]}\right]=1 \tag{3.4}$$

The **basic polytope method** is formally defined as the linear programming minimization of Eq. (3.1) subject to Eq. (3.2), Eq. (3.3), and Eq. (3.4). Although we only show the formalism for 2D, the basic polytope method in 3D is exactly analogous.

Since every feasible solution must satisfy the compatibility equations, the linear system defined by these constraints provides a **lower bound** for the true ground state energy. Furthermore, one important result of this construction is that in any 1D problem, this lower bound is exact, meaning that any 1D problem can be fully solved by basic polytope method:

Proof:

Consider the interaction up to n^th^ nearest neighbor, after transforming the Hamiltonian in terms of blocks:

$$H=\sum_{\{\sigma\}} E(\sigma_1,\cdots,\sigma_{1+n})\rho(\sigma_1,\cdots,\sigma_{1+n})$$

We could then construct a directed graph with all vertexes being of the form $(\sigma_1,\cdots,\sigma_n)$. Then, if $(\sigma_2,\cdots,\sigma_n)=(\sigma'_1,\cdots,\sigma'_{n-1})$, meaning the two blocks are off-by-one translations of each other, we associate an edge connecting $(\sigma_1,\cdots,\sigma_n)$ to $(\sigma'_1,\cdots,\sigma'_n)$ with a flow of size $\rho(\sigma_1,\cdots,\sigma_n,\sigma'_n)$.

Note that in this system, each compatibility constraint is of the form:

$$\rho(\sigma_1,\cdots,\sigma_n)=\sum_s \rho(\sigma_1,\cdots,\sigma_n,s)=\sum_s \rho(s,\sigma_1,\cdots,\sigma_n)$$

meaning that in the directed graph, for each vertex, the sum of all the out-going flows from the vertex is equal to the sum of all in-coming flows into the vertex. By using the basic polytope method, one arrives at a flow solution ρ. Using analysis from linear programming and graph theory, specifically the network flow analysis [92], we know that this ρ corresponds to a cycle in the directed graph and thus ρ corresponds to a physical configuration. Thus, the ground state is given by such a configuration. ■

However in two dimensions and higher, the polytope method thus defined fails in that it can give solutions that do not correspond to a real lattice configuration – we call these solutions **unconstructible**. The primary reason for this failure is that up to now, the constraints on the system guaranteed constructability in the x and y directions independently, not accounting for the fact that the x- and y- constructible solutions must also be compatible with each other. For example, the block configuration:

$$\rho\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}=\rho\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}=\rho\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}=\rho\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}=0.25$$

satisfies the compatibility equations, but does not correspond to a real configuration on a lattice, making it an unconstructible solution. To be specific, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ connects $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ to the left; $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ connects $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ to the left, but $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ does not connect to any other block cluster with non-zero appearing frequency to the left.

To account for constructability, we need a higher order polytope with additional constraints. Traditional approaches to this problem have relied on the enumeration of lattice configurations, which requires an exponential number of variables and makes the solution intractable. Instead, we introduce the cluster-tree optimization algorithm, which iteratively adds variables as necessary to counter frustration, reducing the prefactor in computational complexity to a more tractable level in practical cases and allows us to solve for the true, constructible ground state efficiently.

Section 3.3 Definition of the cluster tree optimization algorithm

The first step in the algorithm is to obtain an initial solution from the basic polytope method: minimize (3.1) subject to the basic constraints given in (3.2), (3.3), and (3.4). This solution is a first, loose lower bound of ground state energy.

To refine this lower bound (if possible), we need to introduce variables for the appearance frequency of larger blocks. We generate these variables using a "spawning operation". A spawning operation on a variable, say $\rho\left[010\right]$, introduces a variable for the appearance frequency of a larger block, say $\rho[0100]$, such as to preserve the perfect sum relationship $\rho[010]=\rho[0100]+\rho[0101]$. To fully integrate the new $\rho[0100]$ variable, we add the necessary physical constraints $0\leq\rho[0100]\leq\rho[100]$ and $0\leq\rho[0101]\leq\rho[101]$ following the rules of the basic polytope method. Implicitly, this constraint $\rho[010]\leq\rho[100]+\rho[101]$ cuts out all unconstructible solutions where $\rho[010]\neq 0$, $\rho[100]=0$ and $\rho[101]=0$. Finally, after solving the new linear programming system, if we find that $\rho[0100]>0$ and correspondingly, $\rho[100]>0$ we introduce the perfect sum relationship constraint $\rho[0100]+\rho[1100]=\rho[100]$ into the linear programming system, which is a stronger condition than simply $\rho[0100]\leq\rho[100]$.

In 2D, the spawning operation follows the same concept but is more complex due to higher dimensionality and numerous possible shapes of the spawning block. However, as before, the spawning operation preserves the perfect sum relationship while introducing larger blocks into the linear programming system. The convergence and correctness of this approach will be proven in a later section.

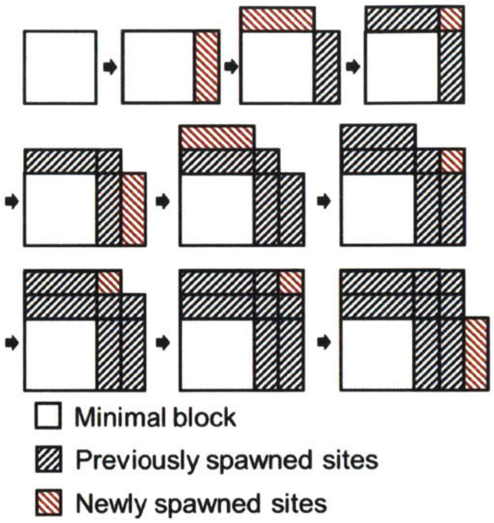

Figure 3 An illustration of the spawning procedure in 2D for the cluster tree optimization algorithm that generates blocks of increasing size while preserving the perfect-sum relationship and avoiding unnecessary variables. The hatched blocks indicate variables added to the original n by m block, where the red counter-hatched blocks specify the variables added in each specific spawning step.

Once again, consider a spin Hamiltonian in which all the interactions can be captured in a block of size m by n. As described earlier, we use a series of spawning operations to arrive at appearance frequencies of larger and larger blocks, giving us converging lower bounds on the total system energy. The general procedure for spawning is illustrated in Figure 3, where at each step, the red stars indicate the sites to be summed over. For example, the first iteration step illustrated in Figure 3 corresponds to the constraint:

$$\rho\begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,m} \\ \vdots & \ddots & \vdots \\ \sigma_{n,1} & \cdots & \sigma_{n,m} \end{bmatrix} = \sum_{*} \rho\begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,m} & * \\ \vdots & \ddots & \vdots & \vdots \\ \sigma_{n,1} & \cdots & \sigma_{n,m} & * \end{bmatrix}$$

Note that Figure 3 only demonstrates one direction of spawning, while in reality there are 3 other spawning directions as illustrated in Figure 2c. The other spawning directions can be derived by exact analogy to the procedure described above.

An essential detail to any spawning operation is that before spawning a variable of the form $\rho\left[\sigma_{[n+i]\times[m+i]}\right]$ with $i>0$, one needs to ensure that the perfect sum relationship holds for all block sizes below $[n+i]\times[m+i]$. Thus, for all $\rho\left[\sigma_{[n+i]\times[m+i]}\right]$, we need to ensure that $\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,1)\}}\right]$, $\rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,1)\}}\right]$, $\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,m+i)\}}\right]$, $\rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,m+i)\}}\right]$ have been generated in the calculation, so that we can impose the constraints:

$$\rho\left[\sigma_{[n+i]\times[m+i]}\right]\le\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,1)\}}\right]$$

$$\rho\left[\sigma_{[n+i]\times[m+i]}\right]\le\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,m+i)\}}\right]$$

$$\rho\left[\sigma_{[n+i]\times[m+i]}\right]\le\rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,1)\}}\right]$$

$$\rho\left[\sigma_{[n+i]\times[m+i]}\right]\le\rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,m+i)\}}\right] \qquad (3.5)$$

We refer this process as adding maximal constraints. Having introduced the maximal constraints and solved the linear optimization again, if$\rho\left[\sigma_{[n+i]\times[m+i]}\right]>0$, we can finally establish the constructability constraints:

$$\sum_{\sigma_{(1,1)}}\rho\left[\sigma_{[n+i]\times[m+i]}\right]=\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,1)\}}\right]$$

$$\sum_{\sigma_{(n+i,1)}} \rho\left[\sigma_{[n+i]\times[m+i]}\right] = \rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,1)\}}\right]$$

$$\sum_{\sigma_{(1,m+i)}} \rho\left[\sigma_{[n+i]\times[m+i]}\right] = \rho\left[\sigma_{[n+i]\times[m+i]-\{(1,m+i)\}}\right]$$

$$\sum_{\sigma_{(n+i,m+i)}} \rho\left[\sigma_{[n+i]\times[m+i]}\right] = \rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,m+i)\}}\right] \tag{3.6}$$

However, following the spawning procedure illustrated in Figure 3, it is possible that some of $\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,1)\}}\right]$, $\rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,1)\}}\right]$, $\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,m+i)\}}\right]$, $\rho\left[\sigma_{[n+i]\times[m+i]-\{(n+i,m+i)\}}\right]$ variables are not generated when $\rho\left[\sigma_{[n+i]\times[m+i]}\right]$ needs to be spawned. Without loss of generality, suppose the missing block is $\rho\left[\sigma_{[n+i]\times[m+i]-\{(1,m+i)\}}\right]$. In this case, we need to trace back in Figure 3 to find the closest block σ' that has already been generated, and impose the constraint:

$$\rho\left[\sigma_{[n+i]\times[m+i]}\right] \leq \rho\left[\sigma'\right]$$

We define this process as back tracing. So long as $\rho\left[\sigma_{[n+i]\times[m+i]}\right] > 0$ in subsequent computations, $\rho\left[\sigma'\right] > 0$ holds and $\rho\left[\sigma'\right]$ can be back traced to eventually yield all the missing blocks.

To summarize, if the algorithm is about to spawn $\rho\left[\sigma_{[n+i]\times[m+i]}\right]$, one needs to first either immediately add maximal constraints or back trace to ensure that $\rho\left[\sigma_{[n+i]\times[m+i]}\right]$ preserves the perfect sum relationship.

With basic polytope method, spawning, and adding maximal constraints defined, the pseudo code of the cluster tree optimization algorithm is as follows:

```
1. Use the basic polytope method to initiate a linear programming system to obtain the appearing frequency of minimal blocks
2. Collect the set of blocks with the smallest size and a positive appearing frequency, denote the set by $S$
3. If all elements of $S$ is in the form $\sigma_{[n+i]\times[m+i]}$ for some $i > 0$, then
    a. If for all $\sigma \in S$, the maximal constraints for $\rho[\sigma]$ have been added, spawning $\rho[\sigma]$ for all $\sigma \in S$ to generate a new set of larger blocks.
    b. Otherwise, try to add maximal constraints for $\rho[\sigma]$ for all $\sigma \in S$ either directly, or by back tracing.
4. Solve the linear programming system to obtain the refined lower bound and repeat from step 2.
```

The optimization loop terminates when either the computed lower bound matches the previously calculated upper bound, or when the spawning size *i* reaches some maximum defined threshold *N*.

When the cluster tree optimization algorithm terminates, if the lower bound and upper bound match, we can guarantee that the ground state solution has been found. Otherwise, based on the fact that the **perfect sum relationship** holds for all blocks with size below

$\left[n+N\right]\times\left[m+N\right]$, we arrive at a converging lower bound as N increases. In practical cases, we find that this convergence tends to be finite, meaning that the lower bound matches the upper bound after some finite number of iterations, as spawning directly corresponds to establishing larger and larger clusters in the traditional polytope method. However, this general finite convergence property cannot be proved.

In this method, we have introduced variables corresponding to interactions of a much higher order than those present in the original problem. Nonetheless, the performance of this approach is vastly superior to direct enumeration as required by traditional methods. The traditional polytope method in general requires 2^{n^2} variables in the binary case, or k^{n^2} variables in k-nary case, to account for clusters of size n by n, while for this method such exponentiation is not necessary. For example, we find that to solve a system with a maximum cluster size of 10 by 10, our method requires approximately 50,000 variables, compared to the completely intractable 2^{100} variables needed for direct enumeration.

As a final observation, there is one alternative termination condition for the optimization. If the algorithm reaches step 3.a with the i defined in that step, meaning that maximal constraints has been added for all $\sigma_{[n+i]\times[m+i]}$ such that $\rho\left[\sigma_{[n+i]\times[m+i]}\right]>0$, and if all such $\sigma_{[n+i]\times[m+i]}$ admit the same periodicity, then we immediately know that the current lower bound is the true ground state energy and ρ is constructible. The proof of this termination condition is given in the supplementary information [32].

Section 3.4 Examples and results

Having defined the cluster-tree optimization algorithm, we illustrate that our solver can reproduce and prove the correctness of ground states known in the literature [28]. In the following examples we look at a triangular lattice with interactions up to the third nearest neighbor. The first step is to define bijection between a triangular and square lattice by setting $(1,0) \rightarrow [1,0]$ and $\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right) \rightarrow \left[0,1\right]$, as shown in Figure 4.

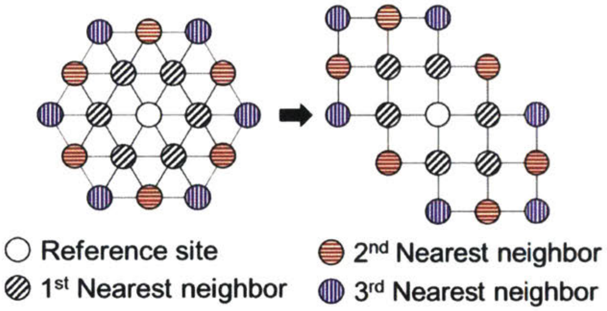

Figure 4 Mapping the interactions on a triangular lattice to an equivalent set on a square lattice, with loss of symmetry in the interactions.

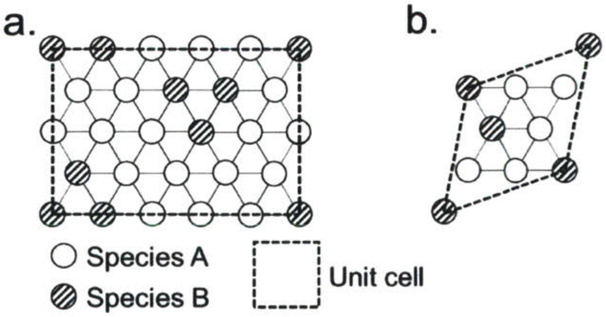

Figure 5 a. The known ground state structure of a pair-interaction Hamiltonian with $V_0 = -4, V_1 = 1, V_2 = 1, V_3 = 1$, where V_0, V_1, V_2, V_3 corresponds to the point term, nearest neighbor, next nearest neighbor and 3rd nearest neighbor interaction terms. b. Known ground state of the frustrated Hamiltonian with $V_1 = 2, V_2 = 1, V_3 = 1, V_0 = -6$, where V_0, V_1, V_2, V_3 are defined in the same way.

Example 1: It is known that the structure in Figure 5a corresponds to the ground state of with the interaction parameters $V_0 = -4 \quad V_1 = 1 \quad V_2 = 1 \quad V_3 = 1$, where V_0, V_1, V_2, V_3 correspond to the point term, nearest neighbor, next nearest neighbor and 3rd nearest neighbor interaction terms on a triangular lattice. Using only the basic polytope method and periodicity enumeration, we can already prove the ground state on an equivalent square lattice. Clearly, in the most basic cases, the polytope method can immediately yield a converged lower bound on the energy. The reason for this success is that this particular Hamiltonian is not frustrated. In the next example, we consider a frustrated system to see how the cluster-tree optimization algorithm efficiently counters frustration, giving a superior result to the basic polytope method.

Example 2: It is known that the ground state corresponding to interaction parameters $V_1 = 2, V_2 = 1, V_3 = 1, V_0 = -6$, where V are defined as before, is the given in structure in Figure 5b[37]. From periodicity enumeration, the ground state energy is suggested to be -1.143, yielding a structure symmetrically equivalent to the true ground state shown in Figure 4b. However, the basic polytope method produces a lower bound of -1.153, which does not match the energy obtained from site enumeration. The cluster tree algorithm in the other hand yields a lower bound energy of -1.143 after 4 iterations, consistent with that provided by this ground state structure.

First iteration: From the basic polytope method equation, we calculate blocks with non-zero appearing frequency to be:

$$\rho\begin{bmatrix}000\\000\\011\end{bmatrix}=\rho\begin{bmatrix}000\\001\\001\end{bmatrix}=\rho\begin{bmatrix}000\\011\\100\end{bmatrix}=\rho\begin{bmatrix}100\\000\\001\end{bmatrix}=\rho\begin{bmatrix}100\\100\\000\end{bmatrix}=\rho\begin{bmatrix}100\\001\\010\end{bmatrix}=\rho\begin{bmatrix}010\\100\\001\end{bmatrix}$$
$$=\rho\begin{bmatrix}010\\010\\100\end{bmatrix}=\rho\begin{bmatrix}110\\000\\000\end{bmatrix}=\rho\begin{bmatrix}001\\010\\010\end{bmatrix}=\rho\begin{bmatrix}001\\110\\000\end{bmatrix}=\rho\begin{bmatrix}001\\001\\110\end{bmatrix}=\rho\begin{bmatrix}011\\100\\100\end{bmatrix}=\frac{1}{13} \tag{3.7}$$

Second iteration: All the non-zero blocks are spawned in the horizontal direction after adding maximal constraints to the system. For example:

$$\rho\begin{bmatrix}000\\000\\011\end{bmatrix}=\rho\begin{bmatrix}0000\\0000\\0110\end{bmatrix}+\rho\begin{bmatrix}0001\\0000\\0110\end{bmatrix}+\rho\begin{bmatrix}0000\\0001\\0110\end{bmatrix}+\rho\begin{bmatrix}0001\\0001\\0110\end{bmatrix}+\rho\begin{bmatrix}0000\\0000\\0111\end{bmatrix}+\rho\begin{bmatrix}0001\\0000\\0111\end{bmatrix}+\rho\begin{bmatrix}0000\\0001\\0111\end{bmatrix}+\rho\begin{bmatrix}0001\\0001\\0111\end{bmatrix}$$

$$\rho\begin{bmatrix}000\\000\\110\end{bmatrix}\geq\rho\begin{bmatrix}0000\\0000\\0110\end{bmatrix}\quad\rho\begin{bmatrix}001\\000\\110\end{bmatrix}\geq\rho\begin{bmatrix}0001\\0000\\0110\end{bmatrix}\quad\rho\begin{bmatrix}000\\001\\110\end{bmatrix}\geq\rho\begin{bmatrix}0000\\0001\\0110\end{bmatrix}....$$

With these new constraints, Solving linear programming reproduces Eq. (3.7) and generates Eq. (3.8).

$$\rho\begin{bmatrix}0001\\0001\\0110\end{bmatrix}=\rho\begin{bmatrix}0001\\0010\\0010\end{bmatrix}=\rho\begin{bmatrix}0001\\0110\\1000\end{bmatrix}=\rho\begin{bmatrix}1000\\0000\\0011\end{bmatrix}=\rho\begin{bmatrix}1000\\1001\\0001\end{bmatrix}=\rho\begin{bmatrix}1000\\0011\\0100\end{bmatrix}=\rho\begin{bmatrix}0100\\1001\\0010\end{bmatrix}$$
$$=\rho\begin{bmatrix}0100\\0100\\1000\end{bmatrix}=\rho\begin{bmatrix}1100\\0000\\0001\end{bmatrix}=\rho\begin{bmatrix}0011\\0100\\0100\end{bmatrix}=\rho\begin{bmatrix}0010\\1100\\0001\end{bmatrix}=\rho\begin{bmatrix}0010\\0010\\1100\end{bmatrix}=\rho\begin{bmatrix}0110\\1000\\1000\end{bmatrix}=\frac{1}{13} \tag{3.8}$$

Third iteration: We then spawn all those nonzero blocks that have not been previously spawned, for example:

$$\rho\begin{bmatrix} 0011 \\ 0100 \\ 0100 \end{bmatrix} = \sum_{i,j,k} \rho\begin{bmatrix} ijkx \\ 0011 \\ 0100 \\ 0100 \end{bmatrix}$$

$$\rho\begin{bmatrix} 101x \\ 0011 \\ 0100 \\ 0100 \end{bmatrix} \le \rho\begin{bmatrix} 101 \\ 001 \\ 010 \end{bmatrix} \quad \cdots$$

where x is could be simply thought as empty space to make its representation clearer. In this step, linear programming results in Eq. (3.7), (3.8) and (3.9) where the x in equation 8 refers to an empty space. There are all together 52 such terms, where we only give a representative sample:

$$\rho\begin{bmatrix} 110x \\ 0001 \\ 0001 \\ 0110 \end{bmatrix} = \rho\begin{bmatrix} x000 \\ 0001 \\ 0001 \\ 0110 \end{bmatrix} = \rho\begin{bmatrix} 0001 \\ 0001 \\ 0110 \\ x000 \end{bmatrix} = \rho\begin{bmatrix} 0001 \\ 0001 \\ 0110 \\ 100x \end{bmatrix}$$
$$= \rho\begin{bmatrix} 100x \\ 1000 \\ 0000 \\ 0011 \end{bmatrix} = \rho\begin{bmatrix} x110 \\ 1000 \\ 0000 \\ 0011 \end{bmatrix} = \rho\begin{bmatrix} 1000 \\ 0000 \\ 0011 \\ x100 \end{bmatrix} = \rho\begin{bmatrix} 1000 \\ 0000 \\ 0011 \\ 001x \end{bmatrix}$$
$$= \rho\begin{bmatrix} 000x \\ 0110 \\ 1000 \\ 1000 \end{bmatrix} = \rho\begin{bmatrix} x001 \\ 0110 \\ 1000 \\ 1000 \end{bmatrix} = \rho\begin{bmatrix} 0110 \\ 1000 \\ 1000 \\ x011 \end{bmatrix} = \rho\begin{bmatrix} 0110 \\ 1000 \\ 1000 \\ 000x \end{bmatrix} = \cdots = \frac{1}{13} \tag{3.9}$$

Forth iteration: This step is crucial in countering the frustration effect. Again, every non-zero block is spawned. The most important of these for countering frustration in the system is:

$$\rho\begin{bmatrix}0110\\1000\\1000\\000x\end{bmatrix}=\rho\begin{bmatrix}0110\\1000\\1000\\0000\end{bmatrix}+\rho\begin{bmatrix}0110\\1000\\1000\\0001\end{bmatrix}$$

$$\rho\begin{bmatrix}0110\\1000\\1000\\0000\end{bmatrix}\leq\rho\begin{bmatrix}000\\000\\000\end{bmatrix}$$

$$\rho\begin{bmatrix}0110\\1000\\1000\\0001\end{bmatrix}\leq\rho\begin{bmatrix}000\\000\\001\end{bmatrix}$$

Note that neither $\rho\begin{bmatrix}000\\000\\000\end{bmatrix}$ nor $\rho\begin{bmatrix}000\\000\\001\end{bmatrix}$ is larger than 0 in the previous solution in Eq. (3.7), but the spawned term $\rho\begin{bmatrix}0110\\1000\\1000\\000x\end{bmatrix}$>0, meaning that that either one of $\rho\begin{bmatrix}000\\000\\000\end{bmatrix}$ or $\rho\begin{bmatrix}000\\000\\001\end{bmatrix}$ must be larger than 0.

Thus the next linear programming calculation forces $\rho\begin{bmatrix}0110\\1000\\1000\\000x\end{bmatrix}=0$ or $\rho\begin{bmatrix}000\\000\\000\end{bmatrix}>0$ or $\rho\begin{bmatrix}000\\000\\001\end{bmatrix}>0$.

Solving the linear system again, a brand new solution is obtained and the frustration effect has been countered:

$$\rho\begin{bmatrix}000\\010\\001\end{bmatrix}=\rho\begin{bmatrix}000\\101\\000\end{bmatrix}=\rho\begin{bmatrix}100\\010\\000\end{bmatrix}=\rho\begin{bmatrix}010\\000\\101\end{bmatrix}=\rho\begin{bmatrix}010\\001\\100\end{bmatrix}=\rho\begin{bmatrix}001\\100\\010\end{bmatrix}=\rho\begin{bmatrix}101\\000\\010\end{bmatrix}=\frac{1}{7}$$

As predicted, $\rho\begin{bmatrix}0110\\1000\\1000\\000x\end{bmatrix}=0$ and the lower bound is refined to be -1.143, which matches the periodic upper bound. Thus, we prove that structure given in Figure 5b is the true ground state.

Although we have only demonstrated this algorithm using small 2D binary systems, we have successfully applied cluster tree optimization to automatically solve systems with basic block sizes up to 4 by 5. We have also successfully applied it to a 3D binary system with a block size up to 2 by 3 by 3. Finally, we have successfully generalized this algorithm to multicomponent cases, although demonstrating the details of these solutions is exceedingly tedious. In terms of computational complexity, the bottleneck of this algorithm is the initial enumerations of elements in the minimal block, where the minimal block is the smallest block to capture all interactions. The complexity order is thus $O\left(k^{x\cdot y\cdot z}\right)$ where k is the numbers of components, and x, y, z is the minimal block size in the x, y, and z directions necessary to capture all interactions. As discussed earlier, while the complexity is exponential in the length scale of the interactions, the exponent is much smaller than that required for the traditional polytope method, making our algorithm much more tractable for solving realistic systems. While our current computational limit is $k^{x \cdot y \cdot z} < 2^{20}$, but this limit is not fundamental, and we intend to address methods to void the necessity to enumerate basic blocks in future work.

Section 3.5 Conclusion for cluster tree algorithm

We have presented a method for obtaining the ground state of a generalized Ising model by the novel cluster tree optimization algorithm. We have proven the correctness of this approach for finding periodic ground states, and shown that even when a periodic ground state solution cannot be found, this algorithm provides a sequence of states with energy converging to ground state energy.

Our approach voids the necessity of exponentially-difficult enumeration to counter frustration. Thus it enables us to probe the space of ground states by directly enumerating the vertices in the true polytope, automatically eliminating unconstructible vertices. Yet, we note that the initial exponential enumeration still limits the application of exact ground state solutions to realistic systems. Regardless, the cluster tree algorithm here is still very valuable as it establishes deep understanding in the ground state problem from the polytope standpoint. We introduce the MAXMIN algorithm in the next chapter, which applies successfully to complex and realistic systems.

Chapter 4 Lower bound method 2: the MAXMIN method: integrating convex optimization and MAXSAT

Section 4.1 Introduction to the MAXMIN method

In this chapter, we present an efficient algorithm that, in many cases, is able to find the global ground state of an arbitrary lattice model in any dimension and of any complexity, and to prove the optimality of the solution. For calculating the upper bound, as shown in Chapter 2 , we derive an equivalence between the optimization of the Hamiltonian under a fixed periodicity and MAX-SAT pseudo-Boolean optimization (PBO), allowing us to leverage existing highly optimized and mathematically rigorous programing tools. To obtain the lower bound on the ground state energy, we derive in this chapter a computationally efficient approach based on a maximization of minimum-energy local configurations. We demonstrate the accuracy, robustness and efficiency of our approach using both an assortment of random Hamiltonians and an example of a realistic Hamiltonian of an existing material. Finally, while we are unable to guarantee that the global optimum can always be found and proven, we argue that our algorithm significantly improves on the state-of-the-art in computational efficiency, and does not sacrifice any guarantees of optimality with respect to established methodology.

Section 4.2 Method formulation

Section 4.2.1 Lower bound calculation

The second element of our algorithm is the optimization of a lower bound to the ground state energy. The lower bound optimization provides both a proof of optimality of the ground state energy independent of periodicity, and a termination condition for the periodicity enumeration discussed in the previous section. We must reiterate that the periodicity-constrained upper-bound solutions described previously are already guaranteed to be optimal within the periodicity constraints provided, such that the lower bound calculation only serves as a proof that the solution obtained by periodicity enumeration is optimal over all other possible periodicities.

To start, we prove that minimization of the Hamiltonian on a finite group of sites, without any periodic constraints, provides a lower bound for the ground state energy. To see why this statement is true, consider the bounds on the Hamiltonian:

$$H\left(s\right)=\lim_{N\to\infty}\frac{1}{(2N+1)^3}\sum_{\substack{(i,j,k)\in\\ \{-N,\ldots,N\}^3}}\sum_{\alpha\in\mathbf{C}}J_\alpha\prod_{(x,y,z,p,t)\in\alpha}s_{i+x,j+y,k+z,p,t} \tag{4.1}$$

$$\equiv\lim_{N\to\infty}\frac{1}{(2N+1)^3}\sum_{\substack{(i,j,k)\in\\ \{-N,\ldots,N\}^3}}E_{i,j,k,s}\geq\min_{i,j,k}E_{i,j,k,s}\geq\min_{\mathbf{s}_0\in\{0,1\}^{\mathbf{B}}}E_{\mathbf{s}_0} \tag{4.2}$$

where $E_{i,j,k,s}$ is defined as $\sum_{\alpha\in\mathbf{C}}J_\alpha\prod_{(x,y,z,p,t)\in\alpha}s_{i+x,j+y,k+z,p,t}$ and represents the energy of a block configuration in the lattice at location (i,j,k) for a specific s. $\mathbf{B}$ is the block cluster containing the relevant (x,y,z,p,t); formally $\mathbf{B}=\bigcup_{\alpha\in\mathbf{C}}\alpha$. Then, $\mathbf{s}_0\in\{0,1\}^{\mathbf{B}}$ is naturally defined as a block configuration and $E_{\mathbf{s}_0}$ as the energy corresponding to block configuration $\mathbf{s}_0$.

The first part of equation (4.2) is a restatement of the total average energy as an average over (i,j,k) of block configuration energies $E_{i,j,k,s}$. As the LHS of equation (4.2) is an average over (i,j,k) of $E_{i,j,k,s}$ it must be greater than or equal to the minimum over (i,j,k) of $E_{i,j,k,s}$, the second part of equation (4.2). Hence, a minimization of $E_{i,j,k,s}$ over configuration space (the RHS), provides a lower bound:

$$H(s) \geq \min_{\mathbf{s}_0 \in \{0,1\}^{\mathbf{B}}} E_{\mathbf{s}_0}$$

As an example, consider a simple binary 1D lattice system with interactions up to the next nearest neighbor (NNN). The energy of this system is bounded from below by the energy of the lowest energy block configuration:

$$\begin{aligned} H &= \lim_{N\to\infty} \frac{1}{(2N+1)} \sum_{(i)\in\{-N,\ldots,N\}^3} \left(J_0 s_i + J_1 s_i s_{i+1} + J_2 s_i s_{i+2} \right) \\ &\geq \min_{s_0, s_1, s_2} \left(J_0 s_0 + J_1 s_0 s_1 + J_2 s_0 s_2 \right) \end{aligned} \tag{4.3}$$

Thus, minimization over the block configurations (s_0, s_1, s_2) produces a valid lower bound, $\min_{s_0, s_1, s_2} \left(J_0 s_0 + J_1 s_0 s_1 + J_2 s_0 s_2 \right)$, to the exact ground state energy.

Expressing the Hamiltonian in the above form assigns all weights of the point term interaction J_0 to the 0th site of the block cluster, corresponding to a $J_0 s_0$ term in the energy. We could have just as well redistributed the point term energy over all sites in the block cluster, transforming $J_0 s_0$ into $\frac{1}{3}(J_0 s_0 + J_0 s_1 + J_0 s_2)$, and similarly $J_1 s_0 s_2$ into $\frac{1}{2} J_1 (s_0 s_2 + s_1 s_3)$:

$$
\begin{aligned}
H &= \lim_{N\to\infty} \frac{1}{(2N+1)} \sum_{(i)\in\{-N,\dots,N\}^3} \begin{pmatrix} \frac{1}{3}\left(J_0 s_i + J_0 s_{i+1} + J_0 s_{i+2}\right) \\ + \frac{1}{2} J_1 \left(s_i s_{i+1} + s_{i+1} s_{i+2}\right) + J_2 s_i s_{i+2} \end{pmatrix} \\
&\geq \min_{s_0,s_1,s_2} \left(\frac{1}{3}\left(J_0 s_0 + J_0 s_1 + J_0 s_2\right) + \frac{1}{2} J_1 \left(s_0 s_2 + s_1 s_3\right) + J_2 s_0 s_2 \right)
\end{aligned}
\tag{4.4}
$$

In the case of the exact infinite system Hamiltonian, this transformation simply corresponds to interchanging the order of the summation and thus imparts no difference to the total energy. However, in the case of the lower bound, we have obtained a new bounding condition, which is a key insight we will use to systematically obtain the tightest possible lower bound on the ground state energy.

Section 4.2.2 Tightening the lower bound using translationally equivalent ECIs

Generally, the direct minimization of $E_{\mathbf{s}_0}$ as described in the previous section gives a very loose lower bound. In principle, a tighter lower bound could be generated systematically by enlarging the block size $|\mathbf{B}|$ used for finite minimization without periodicity constraints, thereby guaranteeing convergence to the exact ground state as we will show below. Furthermore by enlarging the periodicity used for periodic minimization, the upper bound energy is also guaranteed to converge to the exact ground state. To see why this statement is true, consider a minimization over larger and larger block clusters and the resulting block configuration $\mathbf{s}_0$. We could then translate and duplicate the configuration $\mathbf{s}_0$ to arrive at a periodic configuration $\mathbf{s}$ over the entire lattice. The energy of $\mathbf{s}$, $E_{\mathbf{s}}$, will differ from the configuration energy of $\mathbf{s}_0$, $E_{\mathbf{s}_0}$, only at the block boundaries, and the difference will diminish as blocks become larger and larger. This diminishing property results from the fact that the bulk energy scales as $r^{|D|}$ and the boundary energy scales as $r^{|D|-1}$, where D is the dimension of the physical system and r is the size of the cubic block cluster. $E_{\mathbf{s}_0} - E_{\mathbf{s}}$ therefore scales as $\frac{r^{|D|-1}}{r^{|D|}} = \frac{1}{r}$. Therefore the difference between $E_{\mathbf{s}}$ and $E_{\mathbf{s}_0}$ approaches 0, while $E_{\mathbf{s}}$ is an upper bound and $E_{\mathbf{s}_0}$ is a lower bound of the exact ground state energy, proving that the lower bound energy $E_{\mathbf{s}_0}$ converges to the exact ground state energy. We also note that when we perform periodic minimization with the same periodicity as $\mathbf{s}$, we arrive at an upper bound energy smaller than or equal to $E_{\mathbf{s}}$, while greater than the lower bound $E_{\mathbf{s}_0}$, which proves that the upper bound also converges to the exact ground state energy with increasing periodicity. Note that despite the looseness of this provable bound, we observe that in practice our algorithm yields convergence superior to $\frac{r^{|D|-1}}{r^{|D|}} = \frac{1}{r}$, as we elaborate in the discussion section.

Although in the limit of infinite block size, the lower bound converges to the exact ground state energy, this approach is not practical, since finite minimization is NP-hard with respect to the block size. In the following, we present a much more efficient algorithm preserving the convergence property.

Given the original set of cluster interactions $J \in \mathbf{R}^{\mathbf{C}}$, a tighter lower bound can be obtained by introducing the set of equivalent $J_\lambda \in \mathbf{R}^{\bar{\mathbf{C}}}$, which will be defined so as to leave the Hamiltonian of the infinite system (1.1) unchanged, but will modify the Hamiltonian on a finite block. The $J_\lambda \in \mathbf{R}^{\bar{\mathbf{C}}}$ are parameterized by $\lambda \in \mathbf{R}^n$, which we define as a shift parameter.

Note that although the J_λ will be defined to be equivalent to J in the sense that they leave the global Hamiltonian unchanged, finite minimization without periodicity constraints does not yield the same lower bound. Thus, we can maximize the lower bound energy over λ to obtain the tightest lower bound on the ground state energy:

$$\max_{\lambda} \min_{\mathbf{s} \in \{0,1\}^{\mathbf{B}}} E_{\lambda,\mathbf{s}} \tag{4.5}$$

One natural way to introduce equivalent J_λ is by redistributing an ECI over sites in the block cluster: given a fixed block cluster $\mathbf{B}$ to minimize over, for each cluster $\alpha \in \mathbf{C}$ such that $J_\alpha \neq 0$, we construct a set C_α such that all elements $\beta \in C_\alpha$ are equivalent to cluster α with respect to translations of the infinite lattice, and $\beta \subseteq \mathbf{B}$. For each element β in C_α, we assign weights λ_β such that $\sum_{\beta \in C_\alpha} \lambda_\beta = 1$, which relate the translationally equivalent ECIs J_λ to the original ECIs, so that for all $\alpha \in \mathbf{C}$ and $\beta \in C_\alpha$, $J_{\lambda,\beta} = \lambda_\beta J_\alpha$.

Returning to the 1D example of a NNN binary system given in Eq. (4.3), the conversion is:

$$
\begin{aligned}
H &= \lim_{N\to\infty} \frac{1}{(2N+1)} \sum_{(i)\in\{-N,\dots,N\}^3} \left(J_0 s_i + J_1 s_i s_{i+1} + J_2 s_i s_{i+2} \right) \\
&= \lim_{N\to\infty} \frac{1}{(2N+1)} \sum_{(i)\in\{-N,\dots,N\}^3} \begin{pmatrix} J_0 \left(\lambda_1 s_i + \lambda_2 s_{i+1} + \left(1-\lambda_1-\lambda_2\right) s_{i+2} \right) \\ +J_1 \left(\lambda_3 s_i s_{i+1} + \left(1-\lambda_3\right) s_{i+1} s_{i+2} \right) + J_2 s_i s_{i+2} \end{pmatrix} \\
&\geq \min_{s_0,s_1,s_2} \left(J_0 \left(\lambda_1 s_0 + \lambda_2 s_1 + \left(1-\lambda_1-\lambda_2\right) s_2 \right) + J_1 \left(\lambda_3 s_0 s_1 + \left(1-\lambda_3\right) s_1 s_2 \right) + J_2 s_0 s_2 \right)
\end{aligned}
\tag{4.6}
$$

where the last expression provides a lower bound on the ground-state energy, dependent on λ. The rationale behind the λ transform is analogous to that seen in Eq. (4.4): we exploit the fact that we can evenly distribute cluster interactions across sites, leaving the system unchanged, but obtaining a different lower bound on the ground state energy. Note that we are not limited to partitioning point terms equally over all sites, i.e., we could assign a contribution of the point term energy to site 0 with weight λ_1, to site 1 with λ_2 and to site 2 with $1-\lambda_1-\lambda_2$. In this way, we can generally convert $J_0 s_0$ to $J_0 \left(\lambda_1 s_0 + \lambda_2 s_1 + \left(1-\lambda_1-\lambda_2\right) s_2 \right)$, and $J_1 s_0 s_2$ into $J_1 \left(\lambda_3 s_0 s_1 + \left(1-\lambda_3\right) s_1 s_2 \right)$, arriving at the lower bound expression of Eq. (4.6).

From this algorithm, we arrive at $\min_{\mathbf{s}\in\{0,1\}^{\mathbf{B}}} E_{\mathbf{s},\lambda}$, which is a lower bound dependent on λ. Thus,

$$
\max_{\lambda} \min_{\mathbf{s}\in\{0,1\}^{\mathbf{B}}} E_{\lambda,\mathbf{s}} \tag{4.7}
$$

provides the maximal lower bound in the space defined by $\mathbf{B}$ and λ.

Finally, we note that Eq. (4.7) is a convex optimization problem. If $\mathbf{S}$ is fixed, $E_{\lambda,\mathbf{s}}$ is a linear function with respect to λ. Then $f(\lambda) = \min_{\mathbf{s}\in\{0,1\}^{\mathbf{B}}} E_{\lambda,\mathbf{s}}$ is the minimum of a set of linear functions evaluated at λ. Thus, $f(\lambda)$ is a concave function and $\max_{\lambda} f(\lambda)$ is a maximization over a concave function, which is equivalent to a minimization over a convex function, and thus is a convex optimization problem [93]. Due to its piece-wise linear characteristic, this problem belongs to the class of non-smooth convex optimization problems, where the objective function value is provided by MAX-SAT. In our implementation we use the level method [94] as a subclass of the bundle method [95] to efficiently solve this optimization.

Section 4.2.3 Demonstration using a 1D example

To further illustrate our methodology, we present a simple 1D example to demonstrate the key ideas of our algorithm. Consider the 1D system with nearest neighbor and point term interactions (chemical potentials) shown in Figure 2a. The central idea of our algorithm is that the Hamiltonian can be written as an average of block energies, shown in Figure 6b, under the constraint that the blocks whose energies we consider must sum up to the global Hamiltonian, as shown visually in Figure 6a, where the blocks must be able to tile to form the extended structure. Since the choice of how energy is partitioned into blocks is not unique, different ways can be used to describe the block energy, yielding the expression:

$$H = \left\langle J_0 s_i + J_1 s_i s_{i+1} \right\rangle = \left\langle \frac{1}{2} J_0 (s_i + s_{i+1}) + J_1 s_i s_{i+1} \right\rangle = \left\langle J_0 (l\, s_i + (1 - l\,) s_{i+1}) + J_1 s_i s_{i+1} \right\rangle \tag{14}$$

where brackets are used to represent averages and the terms inside the bracket are the so called “block energies”.

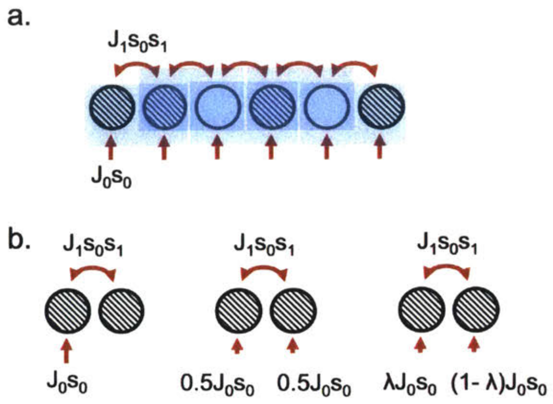

Figure 6. a. 1D schematic illustrating a. the energy of a global configuration, and b. several choices of block energies, with the interaction weighted by an arbitrary λ-shift. The red arrows represent the point term (J_0) and nearest neighbor (J_1) interactions, while the blue rectangles depict tiling compatibility between adjacent blocks, ensuring that the collection of blocks sums up to the global configuration in a.

Since the global energy is the average of block energies, it must be greater than or equal to the smallest possible block energy. Thus, given interaction parameters J, the smallest possible block energy can be found through minimization over spins in the chosen blocks, leading to a lower bound on the total energy:

$$H = \left\langle \frac{1}{2} J_0(s_i + s_{i+1}) + J_1 s_i s_{i+1} \right\rangle \geq \min_{s_i, s_{i+1}} \left(\frac{1}{2} J_0(s_i + s_{i+1}) + J_1 s_i s_{i+1} \right)$$

for the case of evenly distributed point term interactions, or more generally,

$$H = \left\langle J_0(l\, s_i + (1-l) s_{i+1}) + J_1 s_i s_{i+1} \right\rangle \geq \min_{s_i, s_{i+1}} \left(J_0(l\, s_i + (1-l) s_{i+1}) + J_1 s_i s_{i+1} \right) \tag{4.8}$$

for the case of a λ shift. Finally, since Eq. (4.8) is valid for all possible choices of λ, we can obtain a maximally tight lower bound by maximizing this expression over all possible choices of λ:

$$H \geq \max_{\lambda} \min_{s_i, s_{i+1}} \left(J_0(\lambda s_i + (1-\lambda)s_{i+1}) + J_1 s_i s_{i+1} \right) \tag{4.9}$$

In this simple example, it is clear that Eq. (4.9) presents a convex optimization problem. By enumerating all possible choices of $\left(s_i, s_{i+1}\right)$ for the 1D example of Figure 6: $(0,0),(0,1),(1,0),(1,1)$, Eq. (4.9) is

$$H \geq \max_{l} \min \left\{ 0, \left(1-l\right) J_0, l\, J_0, J_0 + J_1 \right\}$$

Each term in the curly bracket is a linear function of λ and therefore their minimum is concave. We illustrate this optimization in Figure 7 for the case where $J_0 = -1$ and $J_1 = 2$. At each λ, 4 linear functions corresponding to 4 different block configuration are considered; the minimum of them at each fixed λ is one valid lower bound. The maximization of the lower bound (maximization of a concave function = convex optimization) results in lower bound being -0.5 with $\lambda = 0.5$ and supporting hyper-planes being (0,1) and (1,0). Since (0,1) and (1,0) are the block configurations with minimum block energy and they can tile into a global configuration, we can guarantee that $\cdots 010101010 \cdots$ configuration is the global ground state even without knowing any upper bound. Furthermore, one easily verifies that $\cdots 010101010 \cdots$ configuration indeed has energy of -0.5. Thus, the upper and lower bounds match, proving the global optimality of this solution.

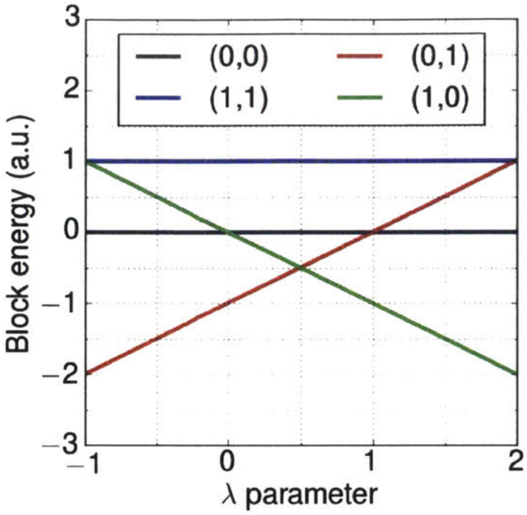

Figure 7. Illustration of block energy in terms of λ in the case of $J_0 = -1$ and $J_1 = 2$. Each line corresponds to one block configuration whose block energy is dependent on λ.

In general, for realistic systems, there are exponentially many hyper-planes with respect to the number of sites in block cluster. One distinct advantage of our method is that we use the MAX-SAT solver to search over such exponential complexity, which as a result provides the concave function and its sub-gradient, which can then be used by the convex optimization solver. Naturally, due to the intrinsic complexity of ground state problem, some NP-hard steps are unavoidable. However, with conversion of such computationally complex steps into MAX-SAT, we are handling such complexity with state of the art efficiency.

Section 4.3 Results and discussions

Section 4.3.1 Comparison to previous methods

The algorithm introduced above provides a range of important advantages over existing approaches towards ground state optimization in cluster expansions. The most common approaches to this problem in the literature are based on Monte Carlo and are not adequate for proving lattice-model ground states, as they only provide a loose upper bound on the energy with no way to determine convergence. Our method both improves the upper bound calculation (by significantly reducing the prefactor in this generally NP-hard problem) and introduces an approach to derive a (typically sufficient) lower bound. The basic-rays method [23-25], a previously reported approach to the ground-state problem, does not provide a lower bound on energy. Furthermore, its ground state solution to a particular set of ECI's (J) requires that basic rays are established at all vertices that define the configurational polytope facet containing J . No general approach to accomplish this has been demonstrated to work for cluster expansion systems of complexities relevant to physical systems. In contrast, our methodology is directly applicable to solving systems with a defined J -vector of relatively complex Hamiltonians, or at least provides tight bounds on the solution. To our knowledge, the only other method that can provide upper and lower bounds on lattice-model ground states is the configurational polytope method [33, 34] that establishes bounds on the solution by means of linear programming. However, as far as we know there is no reported algorithmic system to generate the linear programming constraints for complex cluster expansions. Additionally, even if these constraints were known, the linear programming requires exponentially many variables or constraints with respect to the size of the unit cell being considered, which renders it intractable for complex Hamiltonians. Therefore, we believe that our method is unique in its ability to tackle complex Hamiltonians in a mathematically rigorous way.

Section 4.3.2 Connection to the configurational polytope method

Even though our method has been derived independently of the configurational polytope method, there exists a strong connection between our lower bound calculation and one form of the configurational polytope method. Namely, in relation to the most rigorous form of the configurational polytope method, which involves an exponential number of variables as it relies on strict equality constraints rather than inequalities, our approach provides a route for efficient column generation (see ref. [96] for further details on column generation techniques in linear programming). This improvement expands the applicability of the method to complex, realistic Hamiltonians.

More precisely, if we start with Eq. (4.5) and reformulate its linear programming (LP) problem and its dual, Eq. (4.5) is equivalent to

$$\max_{\lambda} z$$

$$s.t.: \quad z \le E_{\lambda,\mathbf{s}} = \sum_{\alpha\in\mathbf{C}} \sum_{\beta\in C_\alpha} \lambda_\beta J_\alpha \prod_{i\in\beta} s_i \quad \forall \mathbf{s} \in \{0,1\}^{\mathbf{B}}$$

$$\sum_{\beta\in C_\alpha} \lambda_\beta J_\alpha = J_\alpha \quad \forall \alpha \in \mathbf{C}$$

By re-organizing the terms, we have:

$$\max_{\lambda} z$$

$$s.t.: \quad z-\sum_{\alpha\in\mathbf{C}}\sum_{\beta\in C_\alpha}\lambda_\beta J_\alpha\prod_{i\in\beta}s_i\le 0 \;\; \forall\mathbf{s}\in\{0,1\}^{\mathbf{B}} \qquad (4.10)$$

$$\sum_{\beta\in C_\alpha}\lambda_\beta J_\alpha = J_\alpha \;\; \forall\alpha\in\mathbf{C}$$

We could then apply linear programming to obtain its dual (grouping $\lambda_\beta J_\alpha$ as one variable to allow for $J_\alpha = 0$). For each constraint, we construct corresponding dual variables, $\rho_{\mathbf{s}}$ and ρ_α (which physically represent the "appearance frequency" of spin configuration $\mathbf{s}$ and cluster α, as derived later), yielding the dual problem:

$$\min_{\rho}\sum_{\alpha\in\mathbf{C}}\rho_\alpha J_\alpha \qquad (4.11)$$

$$s.t.: \quad \sum_{\mathbf{s}\in\{0,1\}^{\mathbf{B}}}\rho_{\mathbf{s}} = 1 \qquad (4.12)$$

$$-\sum_{\mathbf{s}\in\{0,1\}^{\mathbf{B}}}\rho_{\mathbf{s}}\prod_{i\in\beta}s_i+\rho_\alpha = 0 \;\; \forall\alpha\in\mathbf{C},\beta\in\mathbf{C}_\alpha \qquad (4.13)$$

$$\rho_{\mathbf{s}}\ge 0 \qquad \forall\mathbf{s}\in\{0,1\}^{\mathbf{B}} \qquad (4.14)$$

We arrive at one formulation of the configurational polytope method in its general form, with an exponential number of variables and constraints. We interpret $\rho_{\mathbf{s}}$ as the probability of finding spin configuration $\mathbf{s}$ when we look at some block $\mathbf{B}$, i.e., the "appearance frequency" in the configurational polytope method formulation. Therefore, the natural interpretation of Eq. (4.12) is that some block configuration must appear within a given block cluster, and the probability of observing all possible configurations sums to one. Similarly, Eq. (4.14) can be interpreted to mean that the appearance probability of each block configuration is greater than 0. Finally, Eq. (4.13) is the central equality that constrains the configurational polytope. It is equivalent to $\rho_\alpha = \sum_{\mathbf{s}\in\{0,1\}^{\mathbf{B}}} \rho_{\mathbf{s}} \prod_{i\in\beta} s_i \ \forall\alpha\in\mathbf{C}, \beta\in\mathbf{C}_\alpha$, where $\prod_{i\in\beta} s_i = 1$ if the block configuration exactly matches β, the translated interacting cluster, and $\prod_{i\in\beta} s_i = 0$ otherwise. Therefore ρ_α is the appearance probability of an interacting cluster, constrained by Eq. (4.13) to be translationally invariant. As a result, Eq. (4.11) can naturally be interpreted as a minimization of the configurational energy, with the appearance frequency of each interacting cluster ρ_α multiplied by its interaction constant J_α.

Eq. (4.13) is the central equation that derives constraints in the configurational polytope method. For a 1D system with nearest neighbor and next nearest neighbor interaction, where we take $\mathbf{B}$ to be a 3 site block cluster, these constraints are:

$$\rho[11] = \rho[11*] = \rho[111] + \rho[110] = \rho[011] + \rho[111] = \rho[*11]$$

$$\rho[1*1] = \rho[111] + \rho[101]$$

$$\begin{aligned} \rho[1] &= \rho[1**] = \rho[100] + \rho[101] + \rho[110] + \rho[111] \\ &= \rho[*1*] = \rho[010] + \rho[011] + \rho[110] + \rho[111] \\ &= \rho[**1] = \rho[001] + \rho[011] + \rho[101] + \rho[111] \end{aligned}$$

These relations establish links between the appearance frequencies of small clusters $\rho[1]$, $\rho[11]$, $\rho[1*1]$ and that of larger clusters $r[000] \sim r[111]$ [33]. In order to avoid an exponential growth in the number of variables to be generated, which would render LP optimization infeasible, the configurational polytope method introduces different types of constraints [33, 75], replacing equalities with inequalities. These alternate constraints weaken the formulation relative to the underlying relationship given in Eq. (4.13), but are necessary for all but the simplest systems [75]. Our approach to the problem of exponential growth in the number of variables is equivalent to a column generation scheme [96], where we use the MAX-SAT solver as our column generation oracle [96], enabling us to solve this LP system without departing from its most rigorous constraints. Therefore, our approach holds an advantage over the configurational polytope method in terms of accuracy. Naturally, our lower-bound calculation may still yield inconstructible solutions, which means that we cannot guarantee convergence between our periodically-optimal upper bound solution and lower bound energy in the most general case. Nonetheless, the net result is a solution at least as rigorous as that offered by the traditional polytope method, with numerous advantages in computational efficiency, rigidity of constraints, and feasibility in handling exponential complexity. Thus, we have demonstrated that our formulation is at least as strong as the configurational polytope method if all intermediate clusters used in deriving constraints in the configurational polytope method are introduced into our MAX-MIN formulation.

An important detail regarding the computational complexity of the problem is that the interacting cluster set $\mathbf{C}$ in the primal problem [Eq. (4.10)] and dual problem [Eq. (4.13)] must be the same for the equivalence relationship to hold. As a result, if we include only the nonzero interacting clusters $\mathbf{C}$ in the MAX-MIN formulation, it is equivalent to the configurational polytope method in its general form incorporated with only nonzero interacting clusters in Eq. (4.13). Even in this case, both the MAX-MIN formulation and the configurational polytope method have exponential complexity, since the primal problem has an exponential number of constraints, each of which is associated with one configuration $\mathbf{s} \in \{0,1\}^{\mathbf{B}}$. Similarly, the dual problem has exponentially many nonzero variables $\rho_{\mathbf{s}}$, for each $\mathbf{s} \in \{0,1\}^{\mathbf{B}}$. Therefore, exponential complexity persists in both primal and dual problems. The advantage that the MAX-MIN formulation offers is that we can tackle such exponential complexity with a state-of-the-art procedure, implemented as the MAX-SAT solver. Nonetheless, we must emphasize that the duality transformation itself does not change the computational complexity of the problem.

Section 4.3.3 Empirically observed finite convergence property

An important assertion in our derivation of the method is that, even though we can only mathematically prove that the convergence rate of the lower bound (without a λ-shift) is $\frac{1}{r}$, where r is the length of cubic block cluster, the empirical performance for most realistic or hypothetical systems is that they exhibit a finite convergence property when the λ-shift is introduced, meaning that the upper and lower bounds can be matched exactly at some computationally feasible block size. The formal statement of finite convergence is, given an arbitrary cluster expansion, there exists some particular block size N such that the ground state algorithm would terminate (i.e., the lower bound equals the upper bound) when we consider the block size up to N. Unfortunately, in the most general sense, this statement cannot be proven, as presence of aperiodic ground states means that the ground state problem is generally undecidable, meaning that there is no algorithm that could always guarantee an exact solution [32]. Correspondingly, there exists no algorithm with the finite convergence property for the ground state problem, as finite convergence implies decidability.

Although finite convergence cannot be mathematically proven for our or any other algorithm, it is possible to heuristically reason when we could expect good performance as we empirically observe for our approach. A key feature of our method that enables such performance is the λ-shift and the corresponding tilability of block clusters. Without the λ-shift procedure (meaning that λ is fixed to some constant), the convergence rate of the lower bound with block size is indeed $\frac{1}{r}$. Optimization in λ space yields an improvement as it requires $N \cong \mathrm{Dim}(\lambda)$ supporting hyperplanes at the optimal vertex in λ-space. Each supporting hyperplane corresponds to a block configuration, meaning that these $N \cong \mathrm{Dim}(\lambda)$ block configurations have the same block configurational energy at the optimal λ. The lower bound obtained does not equal to the exact ground state energy only if these $N \cong \mathrm{Dim}(\lambda)$ block configurations do not tile infinite space. Generally, larger $N \cong \mathrm{Dim}(\lambda)$ makes it more likely that block configurations tile the space. Therefore we expect that there is an additional $H(\mathrm{Dim}(\lambda))$ improvement on the convergence rate, where $H(\bullet)$ is some non-decreasing function, yielding an overall convergence of rate of approximately $\frac{1}{r \cdot H(\dim(\lambda))}$. While we cannot determine the mathematical form of $H(\bullet)$, we speculate that $H(\bullet)$ is the main advantage of our method that provides finite convergence in most cases, as it is based on the strong mathematical and physical intuition of tilability. Nonetheless, we must emphasize that this argument remains a speculation on the convergence behavior, and we are unable to prove convergence beyond $\frac{1}{r}$. Thus, most generally, we refer to our earlier proof that our method converges at least as quickly and is at least as rigorous as the configurational polytope method, while offering definite computational advantages by means of the MAX-SAT column-generation oracle.

Section 4.3.4 Computational performance

To test the performance of this approach on practically relevant systems, we measure the runtime of our algorithm on binary 1D, 2D square, and 3D cubic lattices over random sets of ECIs across a spectrum of interaction ranges. First, we restrict ourselves to only pair interactions, calculating runtimes for up to 28 pair interactions on unit cells with up to 50 sites, where the energy of each interaction takes on a random value. In the 1D, 2D and 3D cases, this limit corresponds to all interactions up to and including the 28th, 10th and 5th nearest neighbors respectively. The results of these calculations on a single Intel E5-1650 3.20 GHz core are given in Figure 8. It is important to note that the code performance could be significantly improved by parallelization - the upper bound implementation is perfectly parallelizable up to at least hundreds of compute cores, and the lower-bound calculation parallelizes favorably based on the method chosen for the non-smooth convex optimization.

The results reveal that the primary source of runtime complexity is the range and number of interactions included in the Hamiltonian, with a secondary dependence on the dimensionality of the problem. As could be expected, increasing the range of interactions results in an exponential increase in runtime due to the exponential increase in the size of the spin configuration space. Fortunately, the increase in runtime with the number of interactions at a given range is polynomial. The effect of dimensionality is more subtle: dimensionality determines the number of distinct interactions at a given interaction range, and the number of possible unit cells containing no more than a set number of sites. We find that the former condition is important to the lower bound calculation runtime, while the latter condition determines the variation in the upper bound runtime.

In all cases, our implementation gives a very promising single-core runtime on the order of hours for realistic Hamiltonians, which typically include fewer than 100 interactions. The runtime scales more favorably when all the interactions included in the Hamiltonian are kept below some maximum range – for example, a Hamiltonian with 100 interactions limited to the 8th nearest neighbor-range in 3D can be solved in 3 hours on a single core. This performance is consistent with the trends presented in Figure 8. In a 3D cubic system, there are 61 pair interactions at or below the 8th nearest neighbor range, which, based on the trend in Figure 4 would indicate a runtime of approximately 10^4 seconds, or 2.7 hours. Thus, if we include three and four body terms in the Hamiltonian, the runtime is comparable to that of a pair-interaction Hamiltonian with the same interaction range.

A curious detail of the runtime data presented in Figure 8 is that the computation time required for the 1D problem is similar to that of the 3D problem, considering that the number of periodicities is on the order of $O(N)$ in 1D and $O(N^3)$ in 3D, where N is the maximum unit cell size under consideration. One qualitative explanation of this behavior relies on the fact that in our implementation, we first compute the lower bound on energy, and then attempt to converge the upper bound. Given a fixed number of pair interactions M, the convex optimization in the lower bound calculations has $Dim(\lambda)=O\left(M\sqrt[D]{M}\right)$ for a D dimensional system. Therefore, we may expect that the lower bound calculation is more expensive in 1D than in 3D at a fixed number of interactions M. Furthermore, we find that in 3D, we are often able to terminate the upper bound calculation quickly as it converges to the lower bound at a relatively small periodicity, foregoing the general $O\left(N^3\right)$ MAX-SAT calculations. One possible explanation of this behavior relates to the fact that we measure problem complexity by the number of interactions rather than their range. Consequently, in 3D the M interactions are more mutually exclusive in relation to 1D as they are confined to a much shorter range, reducing the effective number of interactions relevant to the solution.

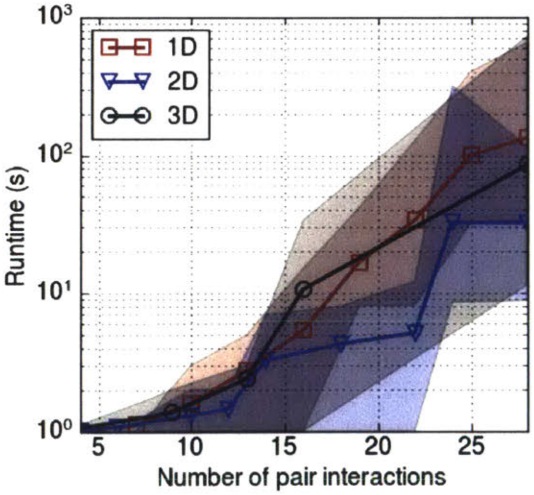

Figure 8. Single-core computation time needed to find and prove the ground state of a 1D, 2D, and 3D pair-interaction Hamiltonian for unit cells up to 50 sites in size across an increasing range of pair-interactions. In all cases, the solver finds the ground state for all unit cells up to 50 atoms in size, and

calculates a tight lower bound on the true ground state energy without enlarging $|\mathbf{B}|$. Each point corresponds to the geometric average runtime of 100 such calculations with random interaction coefficients, while the shading gives the spread between the 20th and 80th percentiles

Section 4.3.5 Application to a realistic Hamiltonian

Finally, we apply our method to obtain the exact ground state of a cluster expansion Hamiltonian used to model sodium-vacancy orderings in the layered Na_xNiO_2 compound as a function of composition. The J interactions for this system are determined from density-functional theory (DFT) calculations of 400 structures through standard approaches [12, 38]. In this cluster expansion, there are 72 interacting clusters, including pair, triplet and quadruplet terms. We emphasize that no previous method exists that could in practice prove the exact ground states for a system with such interactions. In the configurational polytope method [33, 34], a linear programming system with about 2^{32} variables and 2^{32} constraints would be required to capture the frustration effect necessary to provide a tight lower bound. Such a linear programming system cannot be solved in a practically relevant amount of time. In contrast, our method not only finds the exact ground states, but also proves their optimality on a time scale of minutes to hours.

As can be observed in Figure 9, our algorithm finds ground states at x=2/5, 1/2, and 3/5 that were not within the set of DFT input structures initially used to derive the cluster expansion. As we are able to prove that the solutions are optimal, we can guarantee that there are no other configurations that are lower in energy. The inset shows the unusual ground state predicted at x=1/2 which is unlikely to be proposed from intuition.

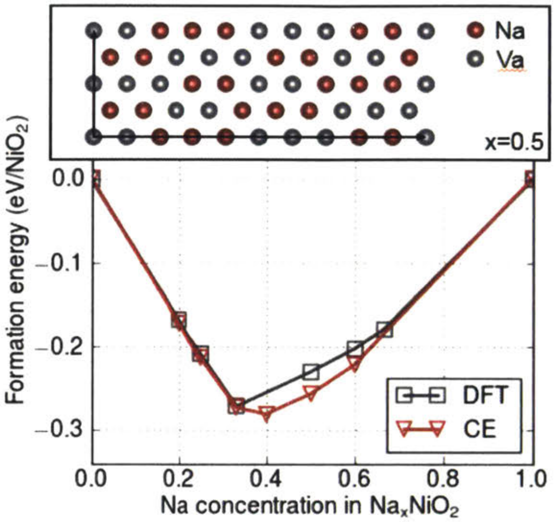

Figure 9. Ground states found for a cluster expansion Hamiltonian of sodium-vacancy orderings in layered Na$_x$NiO$_2$. The red triangles indicate the mathematically proven ground states of the lattice model, whereas the gray squares are the originally proposed ground states from DFT calculations of 400 possible Na-vacancy arrangements. The ground state configuration for x=1/2 is shown in the inset.

Section 4.4 Chapter conclusion

We have introduced a computationally efficient and mathematically rigorous MAX-MIN procedure which is, in many cases, able to obtain the exact ground state of a generalized Ising model and prove its optimality both within a constrained periodicity and with respect to all possible periodicities. To the best of our knowledge, our approach is the only known method of approaching the problem of proving exact ground states of generalized Ising Hamiltonians with interactions of practically relevant complexity. In developing our procedure, we have derived an efficient approach to find an upper bound on the energy by transforming the finite optimization into a Boolean problem in the form of MAX-SAT, which provides a provable periodically-constrained optimum for the Hamiltonian. We also derived a lower bound on the energy from convex optimization over translationally-equivalent clusters. We then converged the upper and lower bounds on the energy to attempt to prove the global optimality of the periodic ground state.

We find that our method is formally related to the most rigorous form of the traditional configurational polytope method, but provides a practical approach for handing the generally exponential number of variables and constraints. Thus, while we are unable to guarantee a provable global ground state solution in all cases, we are able to **(1)** find and prove a periodically-constrained ground state up to a pre-determined unit cell size, **(2)** guarantee global optimality in certain cases, and **(3)** guarantee accuracy at least at the level of the most rigorous configurational polytope method, while offering numerous advantages in terms of computational efficiency. We demonstrate that in practice, our procedure performs very well and has made it possible to determine the exact ground states of many formerly intractable systems, e.g. the cluster expansions of battery systems.

In the next chapter, we will look into another ground state related topics: the construction of cluster expansions that preserves ground state.

Chapter 5 Construction of ground-state preserving cluster expansion models

Section 5.1 Introduction

First-principles density functional theory (DFT) calculations have established themselves as a routine and reliable tool in computational materials science research [97-100] and have enabled important advancements in materials discovery [97, 98, 101]. Although implementations with increasing numerical efficiency and growing computational power have made it possible to simulate ever larger structures with DFT, the method's intrinsic scaling with the number of electrons prevents applications that require large structures (thousands of atoms) and intensive sampling (millions of configurations). Approximate energy models fitted to DFT reference data, such as cluster expansion (CE) lattice models [12, 102-104] or machine learning regression [105, 106], can overcome these limitations by constructing computationally more efficient models with accuracies that are close to DFT for a chosen structural and chemical space. One prototypical application for approximate energy models is the prediction of ordered ground states based on an underlying lattice topology [107, 108].

The key challenge in constructing CE models is to determine the expansion coefficients, the *effective cluster interactions* (ECIs), in a robust fashion through a fit to reference configurations [38, 97, 109]. Conventional ECI fitting procedures [38, 97, 109] focus on minimizing the overall difference between the CE fit and the input configurations with respect to the expanded quantity, such as the energy. In many cases, that input quantity may be determined by an accurate ab-intio method such as Density Functional Theory, One essential requirement that each CE fit must meet for practical applications is ground state preservation, i.e., a physically accurate CE model must reproduce the *ground states* of the input if only the input configurations are considered. This requirement is important as the ground states usually govern the material properties at relevant temperature [110], such as finite temperature voltage profiles [110] and phase diagrams [110]. In this chapter, we revisit the ECI fitting problem with a focus on constraints that guarantee ground-state preservation. We propose a robust and efficient scheme to construct ground state preserving CE models based on compressive sensing [37, 38] and quadratic programming [111].

Section 5.2 Results and discussion

Section 5.2.1 Compressed sensing and cluster expansion

For a rigorous mathematical introduction to cluster expansions and their formal relationship to the partition function of crystalline solids, we refer the readers to references [12, 17, 104]. Here, we only illustrate the key features of cluster expansions that are of relevance to the present work [38].

The general expression of a cluster expansion Hamiltonian is

$$\mathbf{E}_{\mathrm{CE}}(\sigma) = J_0 + \sum_{i}^{\text{sites}} J_i \sigma_i + \sum_{i,j}^{\text{pairs}} J_{i,j} \sigma_i \sigma_j + \sum_{i,j,k}^{\text{triplets}} J_{i,j,k} \sigma_i \sigma_j \sigma_k + \ldots = \sum_{c \in \mathbf{C}} J_c \sigma_c \tag{5.1}$$

where σ is the spin representation of an atomic configuration in which each component σ_i (a *spin variable*) denotes the occupancy of site *i*. Following the Ising-model convention, σ_i takes on values of ±1 in a binary system, encoding the atomic species on site *i*. Each product of spin variables, $\sigma_i \sigma_j \ldots$, (*spin product*) corresponds to a *cluster* of lattice sites, and the cluster-expansion energy E_{CE} is a polynomial of the spin variables weighted by the expansion coefficients $\mathbf{J}$, the *effective cluster interactions* (ECIs). For brevity, we denote the set of interacting clusters as $\mathbf{C}$. For any cluster $c \in \mathbf{C}$, J_c is the corresponding ECI and σ_c is the corresponding spin product. Note that typically multiple clusters of the same *type* exist (e.g., the point term for each equivalent site or the cluster corresponding to the nearest-neighbor pair interaction), and symmetry requires the coefficients of equivalent clusters have to be identical [112]. The summation in Eq. (5.1) is therefore actually over cluster types, and the individual spin products can be replaced by their average over all equivalent clusters, the *cluster correlations.*

From Eq. (5.1) it is obvious that the CE energy is linearly dependent on the ECIs, $\mathbf{J}$, when the configuration σ is fixed. We can thus write

$$\mathbf{E}_{CE}(\sigma) = \mathbf{\Pi}(\sigma)\mathbf{J} \tag{5.2}$$

where $\Pi(\sigma)$ is the row vector of cluster correlations (with multiplicity incorporated) corresponding to configuration σ. Given a set of input atomic configurations S and their DFT energies $\mathbf{E}_{\mathrm{DFT},S}$, the problem of determining the ECIs can then be naïvely expressed as minimization of the L_2 norm

$$\min_{\mathbf{J}} \left\| \mathbf{E}_{\mathrm{DFT},S} - \mathbf{\Pi}_S \mathbf{J} \right\|_2 \tag{5.3}$$

where the rows of the feature matrix Π_S are the cluster correlations of the configurations in *S*. Note that the L_2 norm is the conventional Euclidean norm, and the general L_p norm $\|\mathbf{u}\|_p$ is defined as:

$$\|\mathbf{u}\|_p = \left(\sum_i |u_i|^p \right)^{1/p} \tag{5.4}$$

Simply minimizing the L_2 norm in Eq. (5.3) essentially means that the ECIs are fitted such that the average squared difference between the DFT energies and the CE-predicted energies of all structures is minimized. However, such a direct minimization of the error function leads to *overfitting* when the number of ECIs (the model parameters) exceeds or becomes close to the number of reference configurations (the fitting parameters), i.e., when the system of linear equations Eq. (5.3) is underdetermined. Overfitting means that the ECIs accurately reproduce the energies of the reference structures (*in-sample data*) but deliver poor generalization, i.e., the CE model does not reliably predict the energy of other unseen structures (*out-of-sample data*). A standard method to avoid overfitting is regularization [113], i.e., the simultaneous minimization of the sum of the error function and the magnitude of the model parameters. Compressive sensing [37, 38] implements L_1

norm regularization, which has been shown to be a nearly optimal and robust way to reconstruct signals from a small number of data points [114]. The compressive sensing formulation of the cluster expansion problem is:

$$\min_{\mathbf{J}} \left\| \mathbf{E}_{\mathrm{DFT,S}} - \mathbf{\Pi}_S \mathbf{J} \right\|_2^2 + \mu \left\| \mathbf{J} \right\|_1 \tag{5.5}$$

where μ is a parameter controlling the sparseness of the fit. A higher value of μ shifts the weight towards minimizing the L_1 norm, when μ is small the minimization of the L_2 error dominates. The L_1 norm of a vector is a measure of the vector's sparseness [114], thus larger μ values result in fewer ECIs not equal to zero and thereby reduce overfitting. An optimal μ value can be determined through minimizing the error of the CE model on unseen data [38].

Section 5.2.2 Constrained cluster expansion models

For practical applications it is often desirable that a CE model preserves some invariants on the input data. For example, predicting the qualitative features of a phase diagram may require that the energetic order of all structures is exactly preserved while quantitative errors in the structural energies might be tolerable [2]. This is because the set of ground states and the ranking of excited states close in energy determines the topology of a phase diagram more than the actual energies themselves [36]. As the energy difference between competing structures is typically small, minimization of the average error in reproducing the DFT energy does not by itself enforce the structural energy order one wants to preserve. As a result even very small energy errors in the CE can qualitatively change a phase diagram when it leads to new ground states [36]. We have found practically that trying to preserve the structural ordering and ground states by increasing the relative weights of these input data rapidly leads to overfitting in the CE. In the following, we will develop a methodology that allows including constraints in the ECI optimization problem in a systematic, unbiased fashion and without overfitting.

In recent years, mathematical programming has been a rapidly growing field that enables the highly efficient, systematic and rigorous solution of problems in different standard forms [115]. One rapidly growing area is quadratic programming (QP) [39], for which robust solvers exist [111], and a variety of different approaches have been researched and implemented, such as the interior point method, the active set method and the augmented Lagrangian method [39]. In essence, quadratic programming is a mathematical optimization technique for problems of the following specific form:

$$\begin{aligned} &\min_{\mathbf{x}} \ \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x}+\mathbf{c}^T\mathbf{x} \\ &s.t. \ \ \mathbf{A}\mathbf{x}\leq\mathbf{b} \\ &\qquad \mathbf{C}\mathbf{x}=\mathbf{d} \end{aligned} \tag{5.6}$$

where $\mathbf{Q}$ is a positive semidefinite matrix, $\mathbf{A}$ and $\mathbf{C}$ are real matrices, and $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ are real vectors. Note that a matrix is positive semidefinite if and only if for all real vectors $\mathbf{x}$, $\mathbf{x}^T\mathbf{Q}\mathbf{x}\geq 0$. The semidefinite property is essential so that the optimization problem is convex. Also note that when $\mathbf{Q}=\mathbf{0}$ Eq. (5.6) reduces to a standard linear programming problem which was introduced to CE optimization in reference [2].

Our key strategy for CE fitting is to cast the compressive sensing problem Eq. (5.5) into a quadratic programming problem Eq. (5.6) [116] and to add constraints that guarantee ground-state preservation. Explicitly, Eq. (5.5) can be rewritten as:

$$
\begin{aligned}
& \min_{\mathbf{J}} \left\| \mathbf{E}_{\mathrm{DFT,S}} - \mathbf{\Pi}_{\mathrm{S}}\mathbf{J} \right\|_2^2 + \mu \left\| \mathbf{J} \right\|_1 \\
\Leftrightarrow \quad & \min_{\mathbf{J}} \left\| \mathbf{E}_{\mathrm{DFT,S}} - \mathbf{\Pi}_{\mathrm{S}}\mathbf{J} \right\|_2^2 + \mu \sum_{c\in\mathbf{C}} \left| J_c \right| \\
\Leftrightarrow \quad & \min_{\mathbf{J},\mathbf{z}} \left\| \mathbf{E}_{\mathrm{DFT,S}} - \mathbf{\Pi}_{\mathrm{S}}\mathbf{J} \right\|_2^2 + \mu \sum_{c\in\mathbf{C}} z_c \\
& s.t. \quad z_c \geq J_c \quad \forall c \in \mathbf{C} \\
& \qquad\;\; z_c \geq -J_c \quad \forall c \in \mathbf{C}
\end{aligned}
\tag{5.7}
$$

$$
\begin{aligned}
\Leftrightarrow \quad & \min_{\mathbf{J},\mathbf{z}} \mathbf{J}^T \mathbf{\Pi}_{\mathrm{S}}{}^T \mathbf{\Pi}_{\mathrm{S}} \mathbf{J} - 2\mathbf{E}_{\mathrm{DFT,S}}{}^T \mathbf{\Pi}_{\mathrm{S}} \mathbf{J} + \mu \sum_{c\in\mathbf{C}} z_c + \mathbf{E}_{\mathrm{DFT,S}}{}^T \mathbf{E}_{\mathrm{DFT,S}} \\
& s.t. \quad z_c \geq J_c \quad \forall c \in \mathbf{C} \\
& \qquad\;\; z_c \geq -J_c \quad \forall c \in \mathbf{C}
\end{aligned}
\tag{5.8}
$$

In the conversion step in Eq. (5.7) auxiliary variables z_c, corresponding to constraints on $\mathbf{J}$, have been introduced to remove the L_1 norm of Eq. (5.5). The equivalence in Eq. (5.7) holds because every z_c can be independently minimized while it is constrained to be larger than $\pm J_c$. Note that the QP formulation in Eq. (5.6) does not allow absolute value operations, so that two separate linear constraints are required in Eq. (5.7), $z_c \geq J_c$ and $z_c \geq -J_c$, even though they are in combination essentially expressing the absolute value constraint $z_c \geq |J_c|$. The conversion step in Eq. (5.8) is a direct expansion of the L_2 norm into vector multiplication. Note that $\mathbf{\Pi}_{\mathrm{S}}{}^T \mathbf{\Pi}_{\mathrm{S}}$ is always positive semidefinite for every vector

$\mathbf{x}$, since $\mathbf{x}^T \mathbf{\Pi}_S{}^T \mathbf{\Pi}_S \mathbf{x} = (\mathbf{\Pi}_S \mathbf{x})^T \mathbf{\Pi}_S \mathbf{x} \geq 0$. Hence, we have arrived at a formulation of the compressive sensing ECI problem Eq. (5.5) in terms of a QP problem.

The second key step of our methodology is to include suitable constraints for ground state preservation in the QP formulation. Ground states, i.e., those configurations that are thermodynamically stable at zero temperature (0 K), can be identified by constructing the lower convex hull of the formation energies [117]. When the energy of a configuration is above the ground-state hull it is thermodynamically unstable with respect to decomposition into neighboring ground states.

Note that there are 2 different scenarios that lead to inconsistent ground states from an ECI fit: The first type of ground state inconsistency occurs when the energy of some non-ground-state configuration is underestimated so much that it erroneously becomes a ground state of the CE model. This problem is illustrated in **Figure 10** (labeled with P_1), where the energy of configuration s_1 is predicted to be below its decomposition line in the input data, i.e., below the convex combination of configurations h_2 and h_3 (shown as the line connecting the points). To constrain the QP system such that no inconsistency of type 1 occurs, we add the first constraint:

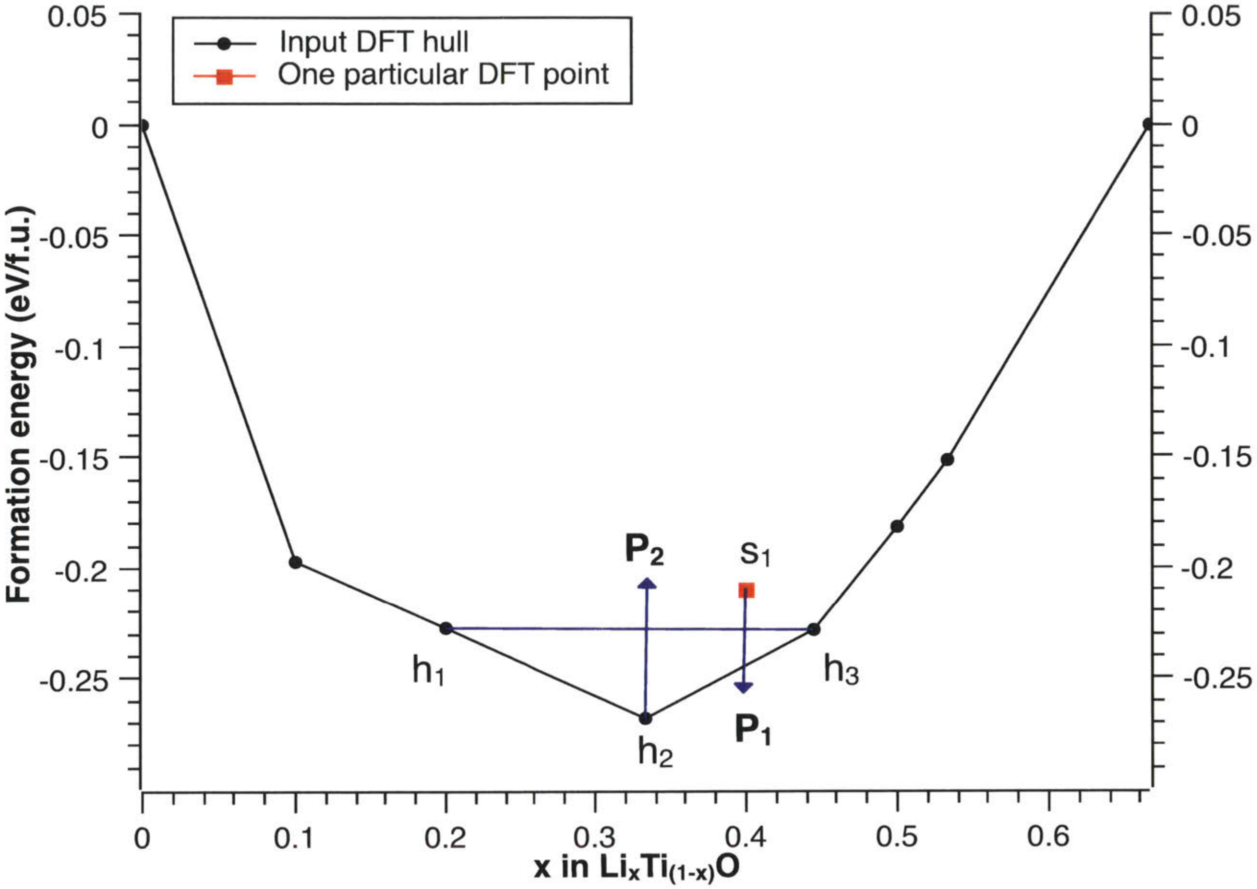

Figure 10 Schematic of the two types of ground-state inconsistencies that may arise during the fit of a cluster expansion (CE) model to DFT reference data. (P1) illustrates the situation in which one particular configuration, s1, that is unstable based on the DFT input data becomes a stable ground state of the CE model, as its CE energy is below the convex combination of its decomposition line defined by the ground state configurations h2 and h3. (P2) illustrates the converse situation in which the CE energy of one ground state configuration, h2, is greater than the convex combination of the neighboring ground states, h1 and h3, which causes h2 to be unstable in the CE model.

(C1) for each configuration that is not on the ground state hull (i.e., a configuration that would thermodynamically decompose into ground states), we require that its CE configuration energy is greater than its CE decomposition line. To express this condition formally, we denote the *i*-th ground state configuration (i.e., the *i*-th configuration on the lower convex hull) by $\sigma_{h,i}$ with $i \in H$. With this notation, the decomposition of an unstable configuration σ_s into the stable ground states can be expressed as $Dec_H(\sigma_s) = \left\{(x_i(\sigma_s), \sigma_{h,i})_{i\in H}\right\}$ where $x_i(\sigma_s)$ is the fraction of $\sigma_{h,i}$ in the decomposition products. The constraint to remove ground state inconsistencies of type 1 becomes

$$\mathbf{\Pi}(\sigma_s)\mathbf{J} \geq \sum_{i\in H} x_i(\sigma_s)\mathbf{\Pi}(\sigma_{h,i})\mathbf{J} + \varepsilon \quad \text{for all unstable configurations } \sigma_s \tag{5.9}$$

where ε is some small number used as numerical tolerance.

Introducing constraint (C1) in Eq. (5.9) to the QP problem in Eq. (5.8) guarantees that all ground states of the CE model are also ground states of the DFT input data. However, the converse is not necessarily true, i.e., a DFT ground state configuration might not be a ground state of the CE model. This scenario is shown in **Figure 10** (P_2), where configuration h_2 has a greater CE energy than its convex hull decomposition line defined by h_1 and h_3. To remove this second type of ground state inconsistency, we introduce a second constraint:

(C2) for each ground state configuration σ_h (i.e., for each configuration σ_h on the lower convex hull), we require that its energy is smaller than the energy of a modified hull that results from removing σ_h from the set of input ground states. Formally, given a ground state configuration σ_h on the input hull, we consider its decomposition into a modified ground state hull as $Dec_{H\backslash h}(\sigma_h)=\left\{\left(x_{i,H\backslash h}(\sigma_h),\sigma_{h,i}\right)_{i\in H\backslash h}\right\}$ where $H\backslash h$ is the index set of all input hull configurations not including σ_h, and $x_{i,H\backslash h}(\sigma_h)$ is the fraction of decomposition product $\sigma_{h,i}$. The constraint to remove ground-state inconsistencies of type 2 thus becomes

$$\mathbf{\Pi}(\sigma_h)\mathbf{J} \le \sum_{i\in H\backslash h} x_{i,H\backslash h}(\sigma_h)\mathbf{\Pi}(\sigma_{h,i})\mathbf{J}-\varepsilon \quad \text{for all ground state configurations } \sigma_h \qquad (5.10)$$

Constraint (C2) in Eq. (5.10) guarantees that all ground state configurations in the (DFT) input data are also ground states of the CE model. Consequently, by combining (C1) and (C2), a configuration is a ground state of the resulting CE model if and only if it is a ground state of the input data. The full quadratic programming formulation for ground-state preserving CE fitting is

$$\begin{aligned}
&\min_{\mathbf{J},\mathbf{z}} \mathbf{J}^T\mathbf{\Pi}_{\mathrm{S}}{}^T\mathbf{\Pi}_{\mathrm{S}}\mathbf{J}-2\mathbf{E}_{\mathrm{DFT,S}}{}^T\mathbf{\Pi}_{\mathrm{S}}\mathbf{J}+\mu\sum_{c\in\mathbf{C}} z_c+\mathbf{E}_{\mathrm{DFT,S}}{}^T\mathbf{E}_{\mathrm{DFT,S}}\\
&s.t.\quad z_c \ge J_c \quad \forall c\in\mathbf{C}\\
&\qquad\; z_c \ge -J_c \quad \forall c\in\mathbf{C} \qquad\qquad (5.11)\\
&\qquad\; \mathbf{\Pi}(\sigma_s)\mathbf{J} \ge \sum_{i\in H} x_i(\sigma_{h,i})\mathbf{\Pi}(\sigma_{h,i})\mathbf{J}+\varepsilon \quad \text{for all unstable configurations } \sigma_s\\
&\qquad\; \mathbf{\Pi}(\sigma_h)\mathbf{J} \le \sum_{i\in H\backslash h} x_{i,H\backslash h}(\sigma_{h,i})\mathbf{\Pi}(\sigma_{h,i})\mathbf{J}-\varepsilon \quad \text{for all ground state configurations } \sigma_h
\end{aligned}$$

Section 5.2.3 Cation ordering in the rocksalt-type lithium transition metal oxide systems $Li_xFe_{(1-x)}O$ and $Li_xTi_{(1-x)}O$

We demonstrate the effectiveness of the QP approach to cation ordering in two oxide systems. Rocksalt-type lithium transition metal oxides, $LiMO_2$ (M = one or more transition metal species), are the most important class of cathode materials for lithium-ion batteries in consumer electronics [118]. During the last decade, materials with lithium excess compositions, $Li_{(1+x)}M_{(1-x)}O_2$, have attracted much interest owing to their high lithium storage capacities [119, 120]. One criterion for the suitability of $Li_{(1+x)}M_{(1-x)}O_2$ as cathode material is whether the material is a sufficiently good conductor for Li ions, which critically depends on the cation (Li, M) ordering in the structure [121, 122]. While conventional oxide-based cathode materials form in ordered crystal structures (such as layered $LiCoO_2$ [123]), several cation-disordered lithium-excess materials with high practical capacities have recently been discovered [122, 124]. Some of these new compositions contain Ti [125] and Fe [124] which makes them attractive for technological applications because of the metals' high abundance and nontoxicity. However, $LiTiO_2$ [126, 127] and $LiFeO_2$ [128, 129] are the only $LiMO_2$ with single transition metal species that form in cation-disordered structures in solid-state synthesis, and consequently their configurational phase diagrams are challenging to investigate experimentally.

In the following we employ the ground-state preserving QP methodology developed above to investigate the phase diagrams of $Li_xFe_{(1-x)}O$ and $Li_xTi_{(1-x)}O$ [126-129] to obtain a better understanding of the relevant atomic configurations. The input consists of 863 and 602 reference configurations for $Li_xFe_{(1-x)}O$ and $Li_xTi_{(1-x)}O$ respectively. DFT calculations for $Li_xFe_{(1-x)}O$ configurations were performed within the Hubbard-U corrected Generalized Gradient Approximation (GGA+U), using the PBE exchange-correlation functional [130, 131]. The U values are taken from the work of Jain et al [132]. DFT calculations for $Li_xTi_{(1-x)}O$ configurations did not employ a Hubbard-U correction. For both systems, an initial set of configurations at $x = 0.5$ with supercell sizes up to 8 sites was generated using the enumerating algorithms by Hart at al. [133], and the reference sets were subsequently refined by including ground state configurations of preliminary cluster expansions determined using a recently published ground-state search algorithm for lattice models [35]. The corresponding ground state input hulls are shown in Figure 11 as black dots and lines. We note that both systems, $Li_xFe_{(1-x)}O$ and $Li_xTi_{(1-x)}O$, cannot easily be fitted using the conventional (unconstrained) compressive sensing technique, as the approach gives rise to a number of spurious ground states as shown in Figure 11 (with optimal μ parameter as will be discussed below). Specifically, some $Li_xFe_{(1-x)}O$ configurations with x=1/3, 5/8 and 9/16 and $Li_xTi_{(1-x)}O$ configurations with x=1/8, 1/6, 1/4, 5/9, 3/5, 5/8 are erroneously predicted to be ground states (i.e., inconsistencies of type 1 as defined above). These over-stabilized configurations are marked with arrows in Figure 11. In addition, the actual $Li_xTi_{(1-x)}O$ ground state configurations with x=1/10, 1/5, 8/15 become unstable in the compressive sensing CE model (inconsistencies of type 2). These examples demonstrate that ground state preservation is not an automatic feature inherent to the compressive sensing approach, and the problem needs to be addressed before predictive simulations of materials systems are possible.

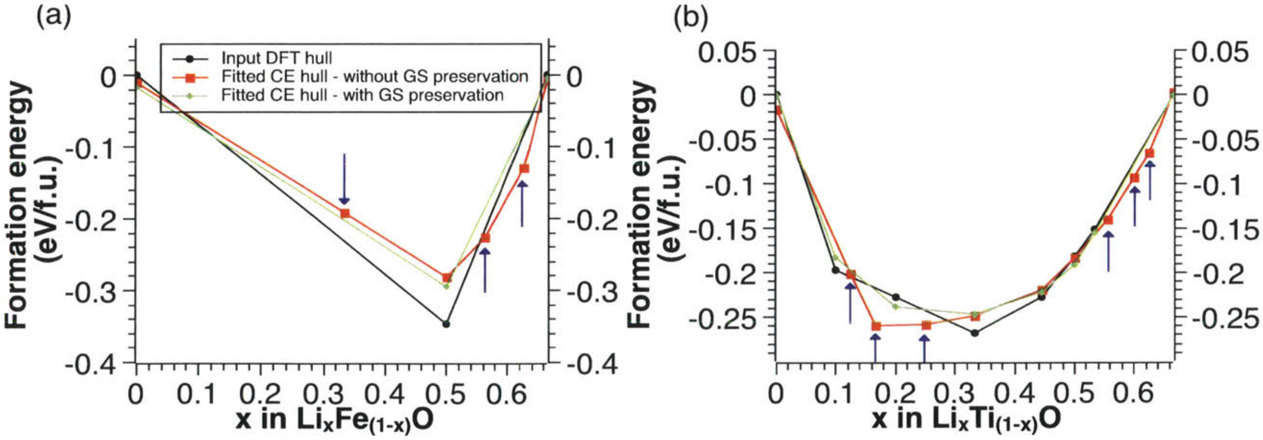

Figure 11 Input hull as obtained from density-functional theory (black lines and circles), cluster-expansion hull fitted using the conventional (unconstrained) compressive sensing method (red lines and squares), and ground-state preserving cluster-expansion hull fitted with the quadratic programming method of the present work (green lines and diamonds) for (a) LixFe(1-x)O and (b) LixTi(1-x)O. The input ground state configurations for (a) LixFe(1-x)O correspond to FeO, LiFeO2 and Li2FeO3 and for (b) LixTi(1-x)O correspond to TiO, Li1Ti9O10, Li2Ti8O10, Li3Ti6O9, Li4Ti5O9,Li2TiO3, LiTiO2, Li8Ti7O15 and Li2TiO3. Arrows are used to mark the over-stabilized unstable configurations. The sparseness parameter, μ, was chosen to be 0.144 for both systems.

As seen in **Figure 11**, the QP fitting scheme achieves ground-state preservation for both materials and yields CE hulls that are spanned by the same configurations as the input DFT hulls. However, it is worth noting that the sparseness parameter, μ, of Eq. (5.11) needs to be carefully selected to arrive at this result. In the following section we will show that μ should be chosen such that the *cross validation error* is minimized. The discussion of the cross validation error is essential in that it provides a standard measure of predictive power of our fitting scheme which sets the method apart from other approaches for ground state preservation, such as the adjustment of configuration weights, which will also be shown below.

Section 5.2.4 Cross validation of the choice of the sparseness parameter

Cross validation is the standard way to decide the optimal sparseness of a numerical model, which is generally referred to as *bias-variance trade-off* in statistical inference [134]. To determine the sparseness parameter μ by means of cross validation we randomly split the

DFT data, D, into $N = 10$ equal parts. For each part D_i, we define its complement $\overline{D}_i$ as all the DFT data points except those belonging to D_i (formally, $\overline{D}_i \equiv D - D_i$). Next, the QP scheme of Eq. (5.11) is applied to the complement set $\overline{D}_i$ to obtain a CE fit without the information in part D_i, so that an out-of-sample validation can be performed by calculating the root mean square error (RMSE) of the unseen data D_i. We denote the resulting out-of-sample RMSE as $e_{i,\mu}$. The cross validation (cv) score cv_μ given a sparseness parameter μ is then defined as the root mean square of the out-of-sample RMSE over all N data parts, i.e., formally $cv_\mu = \sqrt{\sum_{i=1\ldots N} e_{i,\mu}^2 / N}$. Using this definition, the optimal μ resulting in the model with greatest predictive power can be determined by plotting cv_μ against μ and selecting the value of μ that minimizes the cv score.

The cross validation score cv_μ of the $Li_xFe_{(1-x)}O$ system as function of μ is shown in **Figure 12**(a) for various different numbers of input clusters. Note that the number of input clusters to draw from is determined by a maximum interaction order (e.g., *triplets*) and a radial cutoff. Across all 5 curves, the cv score initially decreases and then increases with increasing μ. We consider the concept of bias variance tradeoff [134] to understand this behavior: The input DFT energies may conceptually be understood as the sum of an ideal cluster expansion and a certain degree of noise ε, i.e., $E_{\mathrm{DFT}} = \mathbf{\Pi}(\sigma)J + \varepsilon$. Here, the noise could originate from numerical errors in the DFT energies. For small values of μ, the CE fit uses all available degrees of freedom (i.e., all ECIs) to incorporate the noise ε into the CE model, resulting in severe overfitting. As the value of μ increases, the number of non-zero

ECIs decreases and the effect of noise, i.e., the variance in fitting, becomes less severe. However, when μ becomes too large, the bias that ECIs should tend to 0 becomes dominating over the data itself, resulting in severe underfitting and thus increasing cv scores. As a consequence, the cv score has a pronounced minimum allowing to determine the optimal μ corresponding to the best tradeoff between the variance and bias during fitting.

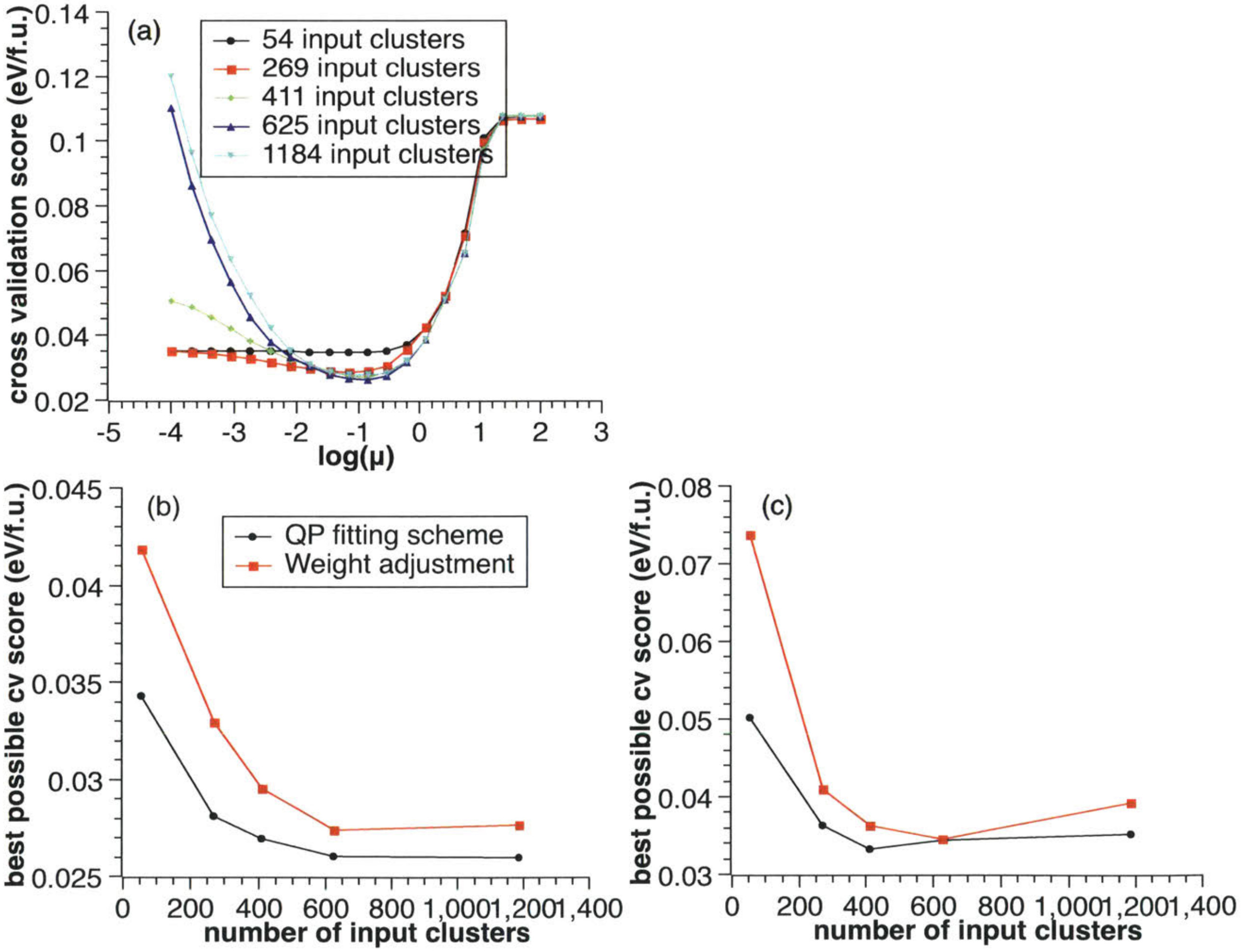

Figure 12 (a) Cross validation (cv) score cv_m plotted as function of the sparseness parameter μ for the example of the LixFe(1-x)O system for various numbers of input clusters. The optimal μ minimizes the cv score. (b)(c) Optimal cv score for different numbers of input clusters as determined using the QP methodology and the conventional weight adjustment scheme for (b) LixFe(1-x)O and (c) LixTi(1-x)O.

As seen in **Figure 12**(a), for small values of μ where $\log(\mu) < -2$ the cv score increases dramatically with the number of input clusters indicating overfitting as the result of insufficient regularization. As the sparseness parameter increases above $\log(\mu) > -2$, the cv score becomes less sensitive with respect to the number of input clusters, indicating that the regularization is effective and that most non-essential ECIs are fitted to zero regardless of the number of input clusters. The optimal cv scores are found for $-1 \le \log(\mu) \le 0$ and are plotted in Figure 12(b) for different numbers of input clusters (labeled "QP methodology"). As seen in the figure, the optimal cv score decreases from 0.0345 eV/f.u. (formula unit) for 54 input clusters to 0.0261 eV/f.u. for 625 input clusters. The cv score stabilizes at 625 input clusters and barely changes for 1184 input clusters (0.0260 eV/f.u.). Hence, we conclude that 625 input clusters and a sparseness parameter $\mu = 0.144$ result in a CE model with optimal predictive power for the $Li_xFe_{(1-x)}O$ system. The corresponding analysis for the $Li_xTi_{(1-x)}O$ system is shown in Figure 12(c), and the optimal parameters, $\mu = 0.144$ and 411 input clusters, yield a cv score of 0.0331 eV/f.u.

Section 5.2.5 In-sample ground state preservation and comparison with conventional weight adjustment

Per construction, the QP form Eq. (5.11) guarantees that the CE fit preserves the ground states of the reference data set. Conventionally, such *in-sample* gound state preservation is often achieved by assigning weights to the reference configurations to manually bias the fit. In the following, we compare the performance of the QP methodology with the conventional weight adjustment technique to further assess the utility of our approach. Before we detail the weight adjustment method, we briefly consider how configuration weights can be included in the QP approach in practice. For this purpose, we define a diagonal weight matrix $\mathbf{W}$ whose diagonal entries $w_{i,i}$ correspond to the *weight* of the i$^{\text{th}}$ input configuration. With this definition, $\mathbf{W}$ can be incorporated into Eq. (5.5) to achieve multiplying weights to the in-sample fitting error:

$$\min_{\mathbf{J}} \left\| \mathbf{W}\left(\mathbf{E}_{\mathrm{DFT,S}} - \mathbf{\Pi}_S \mathbf{J}\right) \right\|_2^2 + \mu \left\| \mathbf{J} \right\|_1 . \qquad (5.12)$$

Note that large $w_{i,i}$ result in a strong bias of the fitting error for the i$^{\text{th}}$ input configuration to be 0. The concrete weight adjustment procedure that we employed in this work is as follows:

(1) Initialize all weights to be 1.

(2) Perform QP to construct a CE model.

(3) Check if the CE model preserves in-sample ground states. If it does, the ground state preserving fit is completed. If it does not, we define the set $\mathbf{T}$ of all DFT and CE ground state configurations

$$\mathbf{T} \equiv \left\{ i \in \mathrm{S} \text{: } \mathrm{S}_i \text{ is DFT Ground state or CE Ground State} \right\} \qquad (5.13)$$

Further, we define the maximum CE hull as $err_{\text{hull}} \equiv \max_{j \in \mathbf{T}} \left| \mathbf{E}_{\text{CE},j} - \mathbf{E}_{\text{DFT},j} \right|$ and introduce a *weight-increment* set $\mathbf{T}' \equiv \left\{ i \in \mathbf{T} : \left| \text{E}_{\text{CE},i} - \text{E}_{\text{DFT},i} \right| > \frac{1}{2} err_{\text{hull}} \right\}$. For each configuration $i \in \mathbf{T}'$, $w_{i,i}$ is increased $\sqrt[4]{2} \approx 1.19$ times. The procedure is continued with step (2).

This weight adjustment scheme guarantees that in-sample ground states are preserved, since it iteratively converges the CE hull to the DFT hull and corrects spurious ground state configurations.

A comparison of the optimal cv scores obtained for different numbers of input clusters using both methods is shown in Figure 12(b,c). The cv score is, once again, used as a standard measure for the predictive power of the fits. In case of the $Li_xFe_{(1-x)}O$ system, the predictive power of the QP fit is consistently better than the fit obtained using the weight adjustment technique, and the improvement of the cv score is generally found to be around 2 meV/f.u. or 10%. For the $Li_xTi_{(1-x)}O$ system, the cv score of the QP CE fit also improves about 3 meV/f.u. or 10% of the CE fit from weight adjustment, except for 625 input clusters for which both methods give equivalent results. However, considering all numbers of input clusters, the overall best cv score for the QP method is 1.5 meV/f.u. or 5% better. While absolute energy errors on the order of a few meV/f.u. are close to the inherent error of density-functional theory, similar errors in the relative energies of different configurations may add up and thereby give rise to qualitatively different phase diagrams.

Generally, we observed that the weight adjustment method biases some configurations by more than a factor of one thousand ($w_{i,i} > 10^3$), resulting in overfitting of those particular configurations, whereas the QP scheme shows no evidence of such a partial over-fitting.

In summary, we conclude that the QP methodology of this work has significant advantages over conventional weight adjustment for the preservation of in-sample ground states. However, as will be demonstrated in the following section, the superiority of the QP approach becomes truly evident when out-of-sample configurations are considered.

Section 5.2.6 Out-of-sample ground state preservation

Suppose that we are certain that the set of reference configurations available for the construction of the CE model comprises all physical ground states of the system. This situation could occur after an extensive exploration of the configurational space or when the DFT data agrees exceptionally well with experiment. With such confidence in the reference data set, we would like to *guarantee* that the fitted CE model not only reproduces the ground states of the reference data, but also does not possess any additional ground states that are not already present in the reference data. We call this property *out-of-sample* ground state preservation. In the following, we describe an iterative procedure for constructing CE models that guarantee out-of-sample ground state preservation up to a given number of periodic sites. We will further show that this procedure is generally a useful strategy to construct CE models even when, initially, it is not known whether ground states outside of the reference set exist.

The QP formulation established in Eq. (5.11) provides ground state preservation within the set of input data. However, out of sample ground state preservation is not guaranteed. In principle, if the true configuration polytope [33], P, is known for a set of possible ECIs, i.e., $\sigma \in P$ can be added and solved within a QP, one could add the following constraint:

$$\mathbf{\Pi}(\sigma)\mathbf{J} \geq \sum_{i \in H} x_i(\sigma)\mathbf{\Pi}(\sigma)\mathbf{J} \quad \forall \sigma \in P \quad (5.14)$$

to Eq. (5.11) and the corresponding optimization problem will result in a globally ground state preserved CE fit. In practice, however, solving the configurational polytope for an arbitrary CE is an undecidable problem [32]. Although this does not necessarily mean that finding a ground state preserving fit is globally undecidable as well, this fact hints at the intrinsic difficulty of the out-of-sample ground state preservation problem.

Instead of determining *a priori* constraints that guarantee out-of-sample preservation, we first examine a CE fit with in-sample ground state preservation obtained from the QP methodology with optimal parameters (sparseness, number of input clusters) and determine all ground states of the CE model up to a defined system size using the methodology of reference [35]. The ground state hull defined by the input configurations is denoted as the *in-sample hull* whereas we refer to the hull that is based on all identified ground states as the *out-of-sample hull.* A comparison of the in-sample and out-of-sample hulls for $Li_xFe_{(1-x)}O$ and $Li_xTi_{(1-x)}O$ for supercell sizes with up to 16 sites is shown in **Figure**

13. For $Li_xFe_{(1-x)}O$, one extra ground state at x=5/8 is identified that is predicted to be 6 meV below the in-sample hull. Even though the distance between the in-sample and out-of-sample hulls is small (6 meV), this CE would produce a qualitatively wrong phase diagram due to the spurious ground state at x=5/8. For $Li_xTi_{(1-x)}O$, the discrepancy between the in-sample CE hull and the out-of-sample CE hull is even more severe, as shown in **Figure 13**(b).

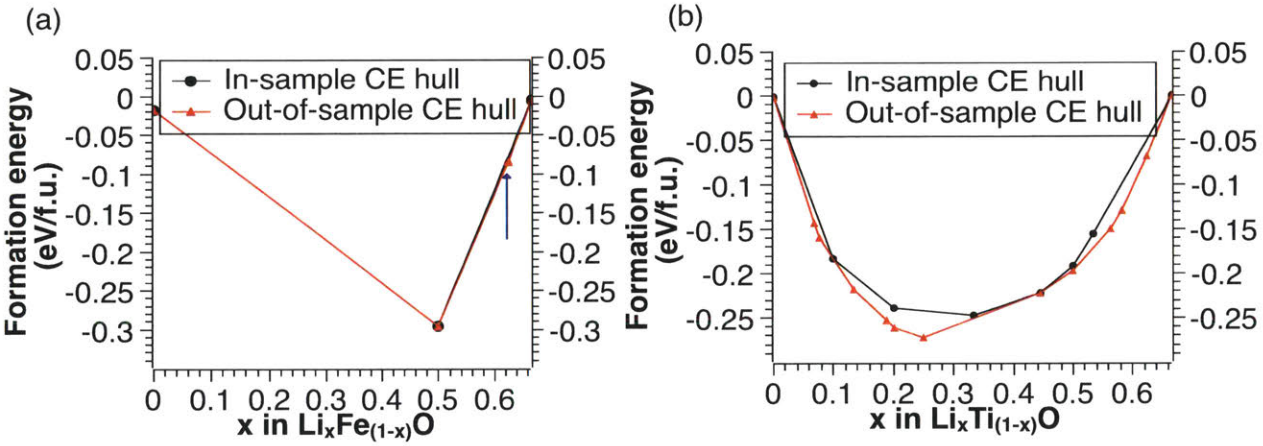

Figure 13 Comparison of the convex hulls based only on the reference input data (in-sample CE hull) and based on all ground state configurations of the cluster expansion model (out-of-sample CE hull) for (a) LixFe(1-x)O and (b) LixTi(1-x)O. The LixTi(1-x)O out-of-sample hull is below the in sample hull with a maximal distance of 6 meV. An arrow is added to emphasize the out of sample ground state at x=0.625 for LixFe(1-x)O.

In the following we would like to arrive at a scheme to construct a CE that does not lead to additional ground states, i.e., out-of-sample ground state preservation. Such scheme is useful in efficient determination of new ground state configurations and self-consistent CE. Instead of determining the true configurational polytope of Eq. (5.14), we arrive at a CE model with out-of-sample ground state preservation iteratively by determining the global ground states of preliminary CE models (as above) to identify those configuration $\sigma \in P$ for which Eq. (5.14) is not satisfied. Afterwards, without additional DFT calculations, the constraint corresponding to these configurations σ are added to the QP form as in Eq. (5.14). By iteratively calculating the ground state hulls and adding further constraints, global ground state preservation up to a large super cell size can be achieved. The procedure is illustrated in Figure 14(a).

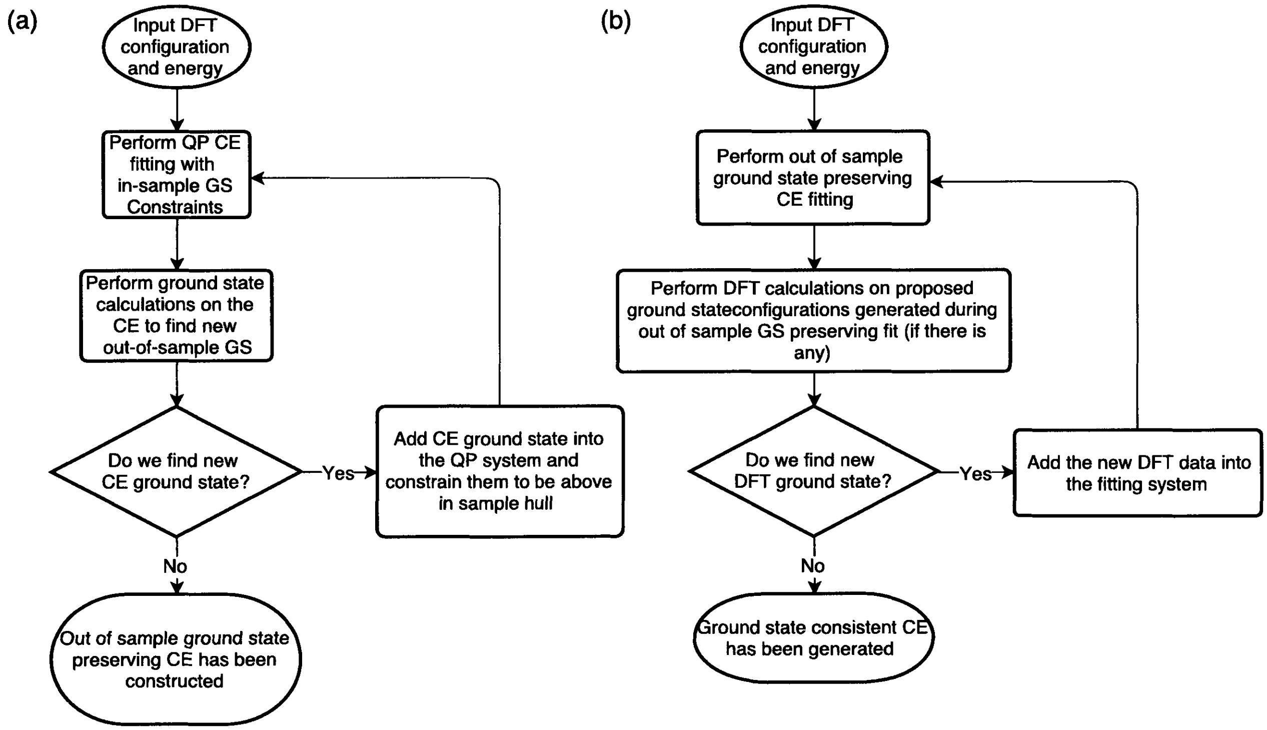
(a)
Input DFT configuration and energy
Perform QP CE fitting with in-sample GS Constraints
Perform ground state calculations on the CE to find new out-of-sample GS
Do we find new CE ground state?
Yes
Add CE ground state into the QP system and constrain them to be above in sample hull
No
Out of sample ground state preserving CE has been constructed
(b)
Input DFT configuration and energy
Perform out of sample ground state preserving CE fitting
Perform DFT calculations on proposed ground stateconfigurations generated during out of sample GS preserving fit (if there is any)
Do we find new DFT ground state?
Yes
Add the new DFT data into the fitting system
No
Ground state consistent CE has been generated

Figure 14 (a) Flow chart of the iterative procedure for constructing out-of-sample ground state preserving cluster expansion models. (b) Flow chart of a combined DFT-CE configurational sampling resulting in the construction of ground state consistent cluster expansion models.

To demonstrate the convergence of this iterative refinement, we applied the procedure to the two model systems for super-cell size of up to 16 sites. The weight adjustment procedure described above is used for comparison. Small initial weights, 10^{-4}, and energies of about 1 meV above the hull are assigned to the predicted new ground states. The results are shown in Figure 15(a) for $Li_xFe_{(1-x)}O$ and Figure 15(b) for $Li_xTi_{(1-x)}O$. For reference, the figure also shows the results of an iterative refinement using the weight adjustment method. The maximum distance between the in-sample hull and the out-of-sample hull is plotted in the upper panel as a measure of the difference between the two hulls as the iteration progresses. The corresponding cv score is plotted in the lower panel as a measure of the predictive power of the CE fit.

As seen in Figure 15, for both systems, $Li_xFe_{(1-x)}O$ and $Li_xTi_{(1-x)}O$, the maximum distance (defined as the difference of energy under the same x) between the out-of-sample and in-sample hulls decreases monotonously to 0 with the QP methodology. The iterative weight adjustment also converges for $Li_xFe_{(1-x)}O$, though the distance between the hulls fluctuates and does not decay monotonously. For $Li_xTi_{(1-x)}O$ the weight adjustment method does not converge. More importantly, the cv scores of the QP fits are nearly constant throughout the iterations, whereas the cv score continuously increases for the weight adjustment algorithms. This means that, using the QP methodology, out-of-sample ground state preservation can be achieved without sacrificing the predictive power of the CE fit. On the other hand, the weight adjustment technique that is often used for CE construction is not guaranteed to converge and rends to achieve gound-state preservation at the cost of predictive power (increasing cv score). We therefore conclude that the QP methodology developed in the present work allows for the systematic construction of CE models with in-sample and out-of-sample ground state preservation.

The results above are based on an exact ground state search for system sizes of up to 16 sites, however, for the purpose of phase diagram calculations via Monte Carlo simulations much larger supercell sizes may be required. To construct CE models that are in practice ground state preserving even for sufficiently large system sizes, the exact ground state search may be replaced by simulated annealing simulations, which allow to determine plausible ground states for larger super cell size (but cannot provide prove that all ground states have been identified, see reference [35] for a more detailed discussion). We repeated the iterative procedure of Figure 14(a) for LixTi(1-x)O for supercell sizes with up to 512 sites using simulated annealing, and the results are depicted in Figure 15(c).

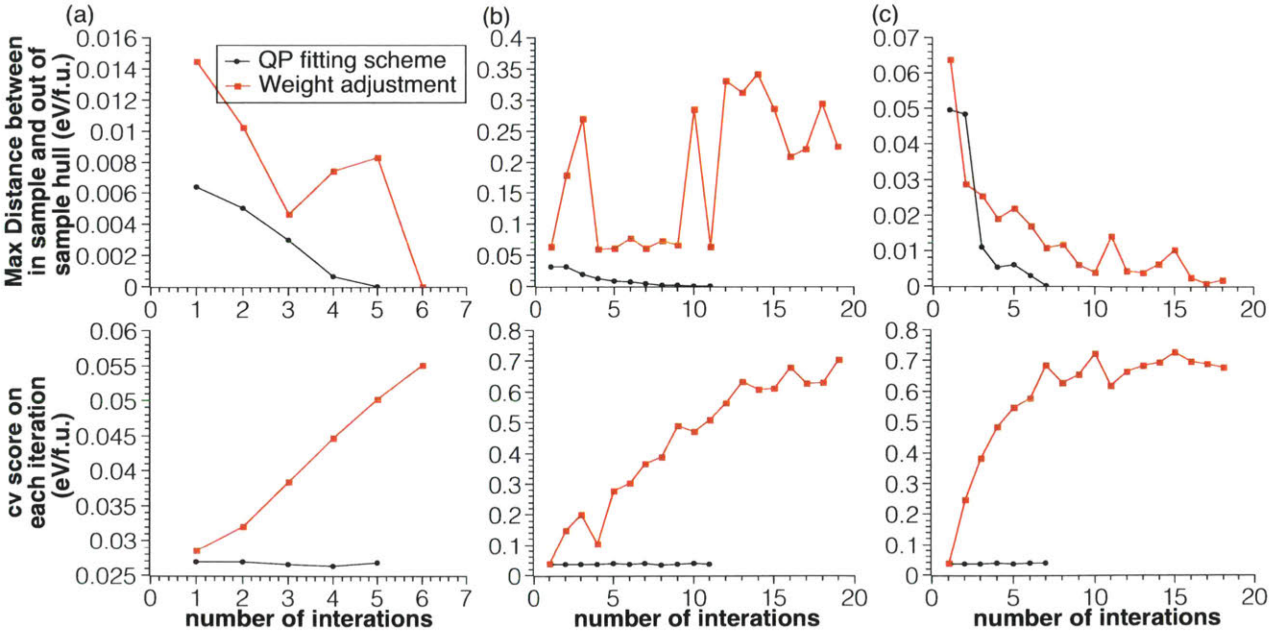

Figure 15 Results of the iterative construction of cluster expansion models with out-of-sample ground state preservation for (a)$Li_xFe_{(1-x)}O$, (b)$Li_xTi_{(1-x)}O$ based on the exact ground states up to a super cell size of 16 sites and (c) $Li_xTi_{(1-x)}O$ based on the simulated annealing with a super cell size of 8×8×8 (512 sites). The upper panels show the maximal distance between the in-sample and out-of-sample hulls, and the lower panels show the corresponding cross-validation scores. For comparison, the corresponding results using iterative adjustment of configuration weights is shown as red lines and squares.

As shown in Figure 15(a) for smaller cell sizes, using the QP methodology, the distance between the in-sample and out-of-sample CE hulls decreases monotonously to 0 within 7 iterations, and the cv score remains nearly constant. As before, for the iterative weight adjustment algorithm does not achieve complete convergence even after 18 iterations and gives rise to a dramatic increase of the cv score. This final example demonstrates again that the QP methodology is a robust scheme to obtain ground state preserving CE fits even for large system sizes that are suitable for realistic Monte Carlo simulations.

Finally, we point out that the iterative procedure for out-of-sample ground state preservation is not only useful, when the ground states of the system are known *a priori*. Instead, the procedure may also serve as a means for the sampling of the configurational space to generate additional reference data. For this purpose, the configurations that were identified as "spurious" ground states may be evaluated with the reference method (i.e., DFT) to confirm whether any unknown ground state has been discovered. By construction, this approach also provides a good stopping criterion for the cluster expansion fit when no additional ground states have been identified. This procedure is illustrated in Figure 14(b).

If DFT calculations for all prospective new ground states are carried out and none of them turns out to be an actual ground state, the out-of-sample ground state preserving fit has the correct assumption and the resulting CE fit is a valid fit with consistently low cv errors. No further iteration is necessary, and the CE fit is finalized. On the other hand, if additional DFT ground states are found within the proposed set, then the out of sample ground state preserving fit would have to be re-started.

Section 5.3 Chapter conclusion

To summarize, in this chapter, we presented a robust and efficient procedure to obtain ground state preserving cluster expansion models. The method is formulated in terms of quadratic programming and compressive sensing and is mathematically rigorous. We demonstrated the robustness of the approach by application to the phase diagrams of $Li_xFe_{(1-x)}O$ and $Li_xTi_{(1-x)}O$ that are challenging to describe with conventional cluster expansion techniques. We further showed that out-of-sample ground state preservation can be achieved up to large supercell sizes. These properties make the presented quadratic programming approach an attractive tool for the fit of general constraint lattice models and point the way towards the fully automated construction of cluster expansion models for materials simulations.

Chapter 6 Concluding remark

Lattice models, also known as generalized Ising models or cluster expansions, are widely used in many areas of science and are routinely applied to alloy thermodynamics, solid-solid phase transitions, magnetic and thermal properties of solids, and fluid mechanics, among others. However, the problem of finding the true global ground state of a lattice model, which is essential for all of the aforementioned applications, has remained unresolved, with only a limited number of results for highly simplified systems known.

In the first part of the thesis, we present, an approach to find the exact ground states of complex lattice models and to prove their global optimality. We believe this work to be a breakthrough in condensed matter theory, as exact solutions of practical lattice models are exceedingly rare. By approaching the ground state problem with modern mathematics and computer science techniques, namely maximum satisfiability (MAX-SAT) and convex optimization, we arrive at an universal algorithm to determine the exact global ground state of a lattice model, defined on an arbitrary lattice with an arbitrary number of components and interactions, that performs very efficiently on systems of physically relevant complexity.

Our algorithm is, to our knowledge, the first general and scalable method for finding provable global energy minima of lattice Hamiltonians. Furthermore, we demonstrate that our approach is practically useful for finding the ground states of realistic Hamiltonians, such as those used for representing lattice orderings in battery systems. Considering that currently such Hamiltonians are solved using simulated annealing and genetic algorithms that are often unable to find the true global energy minimum, and never able to prove the optimality of their result, our work opens the door to resolving long-standing uncertainties in lattice models of physical phenomena.

Cluster expansions enable the study of metallic alloys and other solid solutions with many configurational degrees of freedom, but despite recent developments, the construction of reliable lattice models remains a manual and time-consuming process. A major limitation of existing methods for the construction of lattice models is that they do not guarantee the preservation of the certain important properties, for example the ground states, giving rise to large errors in computed phase diagrams. Thus, manual adjustment of the model parameters, for example through ad-hoc weighting of select configurations, is usually required.

In the second part of the thesis, we introduce an entirely new paradigm for the construction of lattice models that could automatically preserve such important properties. To accomplish this, we mathematically reformulate the compressive-sensing cluster-expansion (CE) method in terms of a quadratic programming (QP) problem. This allows us to introduce constraints on the lattice model, such as ground-state preservation, in a straightforward fashion. We demonstrate for the example of lithium transition-metal oxides that the QP-CE approach converges rapidly to ground-state-preserving CE models and consistently improves the model accuracy compared to conventional manual weight adjustment. This work constitutes significant progress towards the fully-automated construction of lattice models with predictive accuracy.

Reference

1. Li, X., et al., *Direct visualization of the Jahn-Teller effect coupled to Na ordering in Na5/8MnO2.* Nat Mater, 2014. **13**(6): p. 586-92.
2. Garbulsky, G.D. and G. Ceder, *Linear-programming method for obtaining effective cluster interactions in alloys from total-energy calculations: Application to the fcc Pd-V system.* Phys Rev B Condens Matter, 1995. **51**(1): p. 67-72.
3. Struck, J., et al., *Engineering Ising-XY spin-models in a triangular lattice using tunable artificial gauge fields.* Nature Physics, 2013. **9**(11): p. 738-743.
4. Aidun, C.K. and J.R. Clausen, *Lattice-Boltzmann method for complex flows.* Annual review of fluid mechanics, 2010. **42**: p. 439-472.
5. Mueller, T. and G. Ceder, *Effective interactions between the N-H bond orientations in lithium imide and a proposed ground-state structure.* Physical Review B, 2006. **74**(13): p. 134104.
6. Kremer, K. and K. Binder, *Monte Carlo simulation of lattice models for macromolecules.* Computer Physics Reports, 1988. **7**(6): p. 259-310.
7. Seko, A., et al., *Prediction of ground-state structures and order-disorder phase transitions in II-III spinel oxides: A combined cluster-expansion method and first-principles study.* Physical Review B, 2006. **73**(18): p. 184117.
8. Rothman, D.H. and S. Zaleski, *Lattice-gas cellular automata: simple models of complex hydrodynamics*. Vol. 5. 2004: Cambridge University Press.
9. van de Walle, A., *A complete representation of structure-property relationships in crystals.* Nat Mater, 2008. **7**(6): p. 455-8.
10. Van der Ven, A. and G. Ceder, *Vacancies in ordered and disordered binary alloys treated with the cluster expansion.* Physical Review B, 2005. **71**(5): p. 054102.
11. Casola, F., et al., *Direct observation of impurity-induced magnetism in a spin-(1/2) antiferromagnetic Heisenberg two-leg spin ladder.* Phys Rev Lett, 2010. **105**(6): p. 067203.
12. Sanchez, J.M., F. Ducastelle, and D. Gratias, *Generalized cluster description of multicomponent systems.* Physica A: Statistical Mechanics and its Applications, 1984. **128**(1): p. 334-350.
13. Frisch, U., B. Hasslacher, and Y. Pomeau, *Lattice-gas automata for the Navier-Stokes equation.* Phys Rev Lett, 1986. **56**(14): p. 1505-1508.
14. Li, W., J.N. Reimers, and J.R. Dahn, *Crystal structure of Li x Ni 2-x O 2 and a lattice-gas model for the order-disorder transition.* Physical Review B, 1992. **46**(6): p. 3236.
15. Chan, M.K.Y., et al., *Cluster expansion and optimization of thermal conductivity in SiGe nanowires.* Physical Review B, 2010. **81**(17): p. 174303.
16. Ising, E., *Beitrag zur Theorie des Ferromagnetismus.* Zeitschrift für Physik, 1925. **31**(1): p. 253-258.
17. Ceder, G., *A derivation of the Ising model for the computation of phase diagrams.* Computational Materials Science, 1993. **1**(2): p. 144-150.
18. Hinuma, Y., Y.S. Meng, and G. Ceder, *Temperature-concentration phase diagram of P 2-Na x CoO 2 from first-principles calculations.* Physical Review B, 2008. **77**(22): p. 224111.
19. Ozoliņš, V., C. Wolverton, and A. Zunger, *Cu-Au, Ag-Au, Cu-Ag, and Ni-Au intermetallics: First-principles study of temperature-composition phase diagrams and structures.* Physical Review B, 1998. **57**(11): p. 6427.
20. Asta, M. and V. Ozoliņš, *Structural, vibrational, and thermodynamic properties of Al-Sc alloys and intermetallic compounds.* Physical Review B, 2001. **64**(9): p. 094104.
21. Burton, B.P. and A. van de Walle, *First principles phase diagram calculations for the octahedral-interstitial system.* Calphad, 2012. **37**(0): p. 151-157.

22. Zhou, F., T. Maxisch, and G. Ceder, *Configurational electronic entropy and the phase diagram of mixed-valence oxides: The case of Li x FePO 4.* Physical review letters, 2006. **97**(15): p. 155704.
23. Dublenych, Y.I., *Ground states of the Ising model on the Shastry-Sutherland lattice and the origin of the fractional magnetization plateaus in rare-earth-metal tetraborides.* Phys Rev Lett, 2012. **109**(16): p. 167202.
24. Dublenych, Y.I., *Ground states of the lattice-gas model on the triangular lattice with nearest- and next-nearest-neighbor pairwise interactions and with three-particle interaction: full-dimensional ground states.* Phys Rev E Stat Nonlin Soft Matter Phys, 2011. **84**(1 Pt 1): p. 011106.
25. Dublenych, Y.I., *Ground states of the lattice-gas model on the triangular lattice with nearest- and next-nearest-neighbor pairwise interactions and with three-particle interaction: ground states at boundaries of full-dimensional regions.* Phys Rev E Stat Nonlin Soft Matter Phys, 2011. **84**(6 Pt 1): p. 061102.
26. Teubner, M., *Ground states of classical one-dimensional lattice models.* Physica A: Statistical Mechanics and its Applications, 1990. **169**(3): p. 407-420.
27. Kanamori, J. and M. Kaburagi, *Exact Ground States of the Lattice Gas and the Ising Model on the Square Lattice.* Journal of the Physical Society of Japan, 1983. **52**(12): p. 4184-4191.
28. Kaburagi, M. and J. Kanamori, *Ground State Structure of Triangular Lattice Gas Model with up to 3rd Neighbor Interactions.* Journal of the Physical Society of Japan, 1978. **44**(3): p. 718-727.
29. Finel, A. and F. Ducastelle, *On the phase diagram of the FCC Ising model with antiferromagnetic first-neighbour interactions.* EPL (Europhysics Letters), 1986. **1**(3): p. 135.
30. Ansótegui, C., M.L. Bonet, and J. Levy, *Solving (weighted) partial MaxSAT through satisfiability testing*, in *Theory and Applications of Satisfiability Testing-SAT 2009*. 2009, Springer. p. 427-440.
31. Luo, C., et al., *CCLS: An efficient local search algorithm for weighted maximum satisfiability.* IEEE Transactions on Computers, 2014.
32. Huang, W., et al., *Constructing and proving the ground state of a generalized Ising model by the cluster tree optimization algorithm.* arXiv preprint arXiv:1606.07429, 2016.
33. Ducastelle, F., *Order and Phase Stability in Alloys*. 1991: North-Holland.
34. Kaburagi, M. and J. Kanamori, *A method of determining the ground state of the extended-range classical lattice gas model.* Progress of Theoretical Physics, 1975. **54**(1): p. 30-44.
35. Huang, W., et al., *Finding and proving the exact ground state of a generalized Ising model by convex optimization and MAX-SAT.* Physical Review B, 2016. **94**(13): p. 134424.
36. Kohan, A., et al., *Computation of alloy phase diagrams at low temperatures.* Computational materials science, 1998. **9**(3-4): p. 389-396.
37. Candès, E.J. and M.B. Wakin, *An introduction to compressive sampling.* IEEE signal processing magazine, 2008. **25**(2): p. 21-30.
38. Nelson, L.J., et al., *Compressive sensing as a paradigm for building physics models.* Physical Review B, 2013. **87**(3): p. 035125.
39. Gill, P.E. and E. Wong, *Methods for convex and general quadratic programming.* Mathematical Programming Computation, 2014. **7**(1): p. 71-112.
40. Aarts, E. and J. Korst, *Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing.* 1988.
41. Metropolis, N., et al., *Equation of State Calculations by Fast Computing Machines.* The Journal of Chemical Physics, 1953. **21**(6): p. 1087-1092.
42. Fontaine, D.D., *Cluster approach to order-disorder transformations in alloys.* Solid state physics, 1994. **47**: p. 33-176.
43. Garbulsky, G.D., P.D. Tepesch, and G. Ceder. *Ground state analysis on the fcc lattice with four pair interactions*. in *MRS Proceedings*. 1992. Cambridge Univ Press.

44. Ceder, G., et al., *Ground states of a ternary fcc lattice model with nearest-and next-nearest-neighbor interactions.* Physical Review B, 1994. **49**(1): p. 1.

45. Granville, V., M. Krivánek, and J.P. Rasson, *Simulated annealing: A proof of convergence.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1994. **16**(6): p. 652-656.

46. Kirkpatrick, S., C.D. Gelatt, and M.P. Vecchi, *Optimization by simmulated annealing.* science, 1983. **220**(4598): p. 671-680.

47. Bautu, A. and E. Bautu, *Searching ground states of Ising spin glasses with particle swarms.* Romanian Journal of Physics, 2007. **52**(3/4): p. 337.

48. Băutu, A. and E. Băutu, *Searching Ground States of Ising Spin Glasses with Genetic Algorithms and Binary Particle Swarm Optimization*, in *Nature Inspired Cooperative Strategies for Optimization (NICSO 2007)*. 2008, Springer. p. 85-94.

49. McGeoch, C.C. and C. Wang. *Experimental evaluation of an adiabiatic quantum system for combinatorial optimization.* in *Proceedings of the ACM International Conference on Computing Frontiers*. 2013. ACM.

50. Hartmann, A.K., *Ground States of Two-Dimensional Ising Spin Glasses: Fast Algorithms, Recent Developments and a Ferromagnet-Spin Glass Mixture.* Journal of Statistical Physics, 2011. **144**(3): p. 519-540.

51. Rieger, H., *Frustrated systems: Ground state properties via combinatorial optimization*, in *Advances in Computer Simulation*. 1998, Springer. p. 122-158.

52. Bundaru, M., N. Angelescu, and G. Nenciu, *On the ground state of Ising chains with finite range interactions.* Physics Letters A, 1973. **43**(1): p. 5-6.

53. Takasaki, K., T. Tonegawa, and M. Kaburagi, *Exact Ground States of the Lattice Gas Model with Two Kinds of Particles and the Finite-Field Three-State Potts Model on the Square Lattice.* Journal of the Physical Society of Japan, 1988. **57**(2): p. 570-579.

54. Kaburagi, M. and J. Kanamori, *Theoretical Analysis of Adsorbate Structures on Metal Surface–Sodium Atoms on Tungsten (110) Plane–.* Journal of the Physical Society of Japan, 1977. **43**(5): p. 1686-1693.

55. Kaburagi, M., *Ground State Structures of the Face Centered Rectangular Lattice Gas Model with up to 4th Neighbor Interactions.* Journal of the Physical Society of Japan, 1978. **44**(2): p. 394-401.

56. Kanamori, J., *Lattice Gas Model Analysis of the (111) Surface Structures of Si, Ge and Related Systems. II. On the DAS Model.* Journal of the Physical Society of Japan, 1986. **55**(8): p. 2723-2734.

57. Kanamori, J., *Infinite Series of Ground States of the Ising Model on the Honeycomb Lattice.* Journal of the Physical Society of Japan, 1984. **53**(1): p. 250-260.

58. Grünbaum, B. and G.C. Shephard, *Tilings and Patterns*. 2013: Dover Publications, Incorporated.

59. Lipkin, M.D., *Ground states of a frustrated Ising model with one-, two- and three-body interactions.* Physica A: Statistical Mechanics and its Applications, 1988. **150**(1): p. 18-39.

60. Kanamori, J. and Y. Kakehashi, *Conditions for the existence of ordered structure in binary alloy systems.* Le Journal de Physique Colloques, 1977. **38**(C7): p. C7-274-C7-279.

61. Cowley, J.M., *Modulated Structures 1979: Kailus Kona, Hawaii.* 1979: American Institute of Physics.

62. Dublenych, Y.I., *Ground states of lattice-gas models on the triangular and honeycomb lattices: devil's step and quasicrystals.* Phys Rev E Stat Nonlin Soft Matter Phys, 2009. **80**(1 Pt 1): p. 011123.

63. Dublenych, Y.I., *Ground states of an Ising model on an extended Shastry-Sutherland lattice and the 1/2-magnetization plateau in some rare-earth-metal tetraborides.* Phys Rev E Stat Nonlin Soft Matter Phys, 2013. **88**(2): p. 022111.

64. Dublenych, Y.I., *Ground states of the Ising model on an anisotropic triangular lattice: stripes and zigzags.* J Phys Condens Matter, 2013. **25**(40): p. 406003.

65. Dublenych, Y.I., *Structures on lattices: some useful relations.* Phys Rev E Stat Nonlin Soft Matter Phys, 2011. **83**(2 Pt 1): p. 022101.

66. Dublenych, Y.I., *Ground-state structures in Ising magnets on the Shastry-Sutherland lattice with long-range interactions and fractional magnetization plateaus in TmB4.* arXiv preprint arXiv:1403.0874, 2014.

67. Wang, H., *Proving theorems by pattern recognition—II.* Bell system technical journal, 1961. **40**(1): p. 1-41.

68. Robinson, R.M., *Undecidability and nonperiodicity for tilings of the plane.* Inventiones mathematicae, 1971. **12**(3): p. 177-209.

69. Berger, R., *The undecidability of the domino problem*. 1966: American Mathematical Soc.

70. Culik, II, *An aperiodic set of 13 Wang tiles.* Discrete Mathematics, 1996. **160**(1): p. 245-251.

71. Kari, J., *A small aperiodic set of Wang tiles.* Discrete Mathematics, 1996. **160**(1): p. 259-264.

72. Lagae, A. and P. Dutré, *An alternative for Wang tiles: colored edges versus colored corners.* ACM Transactions on Graphics (TOG), 2006. **25**(4): p. 1442-1459.

73. Lagae, A., J. Kari, and P. Dutré, *Aperiodic sets of square tiles with colored corners.* Report CW, 2006. **460**.

74. Sipser, M., *Introduction to the Theory of Computation*. 2005: Cengage Learning.

75. Walle, A.v.d., *The effect of lattice vibrations on substitutional alloy thermodynamics*. 2000.

76. Dublenych, Y.I., *Continuum of ground states and aperiodic structures in a lattice gas on the triangular lattice with finite-range interactions.* Physical Review B, 2012. **86**(1): p. 014201.

77. Wang, H., *Notes on a class of tiling problems.* Fundamenta Mathematicae, 1975. **82**(4): p. 295-305.

78. Graf, P.A., et al., *Direct enumeration of alloy configurations for electronic structural properties.* Applied Physics Letters, 2005. **87**(24): p. 243111.

79. Kohan, A.F., et al., *Computation of alloy phase diagrams at low temperatures.* Computational materials science, 1998. **9**(3): p. 389-396.

80. Zhang, Y., et al., *Prediction of new stable compounds and promising thermoelectrics in the Cu–Sb–Se system.* Chemistry of Materials, 2014. **26**(11): p. 3427-3435.

81. Van der Ven, A., et al., *First-principles investigation of phase stability in Li x CoO 2.* Physical Review B, 1998. **58**(6): p. 2975.

82. Hart, G.L.W. and R.W. Forcade, *Algorithm for generating derivative structures.* Physical Review B, 2008. **77**(22): p. 224115.

83. Luenberger, D.G., *Introduction to linear and nonlinear programming*. Vol. 28. 1973: Addison-Wesley Reading, MA.

84. Garfinkel, R.S. and G.L. Nemhauser, *Integer programming*. Vol. 4. 1972: Wiley New York.

85. Boros, E. and P.L. Hammer, *Pseudo-boolean optimization.* Discrete applied mathematics, 2002. **123**(1): p. 155-225.

86. Gomes, C.P., et al., *Satisfiability solvers.* Handbook of Knowledge Representation, 2008. **3**: p. 89-134.

87. Argelich, J., et al., *Max-SAT Evaluation 2010.*

88. AnsóTegui, C., M.L. Bonet, and J. Levy, *Sat-based maxsat algorithms.* Artificial Intelligence, 2013. **196**: p. 77-105.

89. Hoos, H.H. and T. Stützle, *Stochastic local search: Foundations & applications*. 2004: Elsevier.

90. Stützle, T., H. Hoos, and A. Roli, *A review of the literature on local search algorithms for MAX-SAT.* Rapport technique AIDA-01-02, Intellectics Group, Darmstadt University of Technology, Germany, 2001.

91. Kuegel, A. *Improved Exact Solver for the Weighted MAX-SAT Problem*. in *Pos@ sat*. 2010.

92. Ahuja, R.K., T.L. Magnanti, and J.B. Orlin, *Network flows*. 1988, DTIC Document.

93. Boyd, S.P. and L. Vandenberghe, *Convex Optimization*. 2004: Cambridge University Press.

94. Kiwiel, K.C., *The efficiency of subgradient projection methods for convex optimization, part I: General level methods.* SIAM Journal on Control and Optimization, 1996. **34**(2): p. 660-676.

95. Kiwiel, K.C., *Proximity control in bundle methods for convex nondifferentiable minimization.* Mathematical Programming, 1990. **46**(1-3): p. 105-122.

96. Bertsimas, D. and J.N. Tsitsiklis, *Introduction to linear optimization*. Vol. 6. 1997: Athena Scientific Belmont, MA.

97. Fischer, C.C., et al., *Predicting crystal structure by merging data mining with quantum mechanics.* Nat Mater, 2006. **5**(8): p. 641-6.

98. Yang, K., et al., *A search model for topological insulators with high-throughput robustness descriptors.* Nat Mater, 2012. **11**(7): p. 614-9.

99. Rong, Z., et al., *Materials Design Rules for Multivalent Ion Mobility in Intercalation Structures.* Chemistry of Materials, 2015. **27**(17): p. 6016-6021.

100. Liu, M., et al., *Spinel compounds as multivalent battery cathodes: a systematic evaluation based on ab initio calculations.* Energy & Environmental Science, 2015. **8**(3): p. 964-974.

101. Jain, A., Y. Shin, and K.A. Persson, *Computational predictions of energy materials using density functional theory.* Nature Reviews Materials, 2016. **1**: p. 15004.

102. van de Walle, A. and G. Ceder, *Automating first-principles phase diagram calculations.* Journal of Phase Equilibria, 2002. **23**(4): p. 348-359.

103. De Fontaine, D., *Configurational thermodynamics of solid solutions.* Solid state physics, 1979. **34**: p. 73-274.

104. De Fontaine, D., *Cluster approach to order-disorder transformations in alloys.* Solid state physics, 1994. **47**: p. 33-176.

105. Artrith, N. and A. Urban, *An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO 2.* Computational Materials Science, 2016. **114**: p. 135-150.

106. Artrith, N., B. Hiller, and J. Behler, *Neural network potentials for metals and oxides–First applications to copper clusters at zinc oxide.* physica status solidi (b), 2013. **250**(6): p. 1191–1203.

107. Nahas, S., et al., *First-principles cluster expansion study of functionalization of black phosphorene via fluorination and oxidation.* Physical Review B, 2016. **93**(16): p. 165413.

108. Predith, A., et al., *Ab initio prediction of ordered ground-state structures in ZrO 2-Y 2 O 3.* Physical Review B, 2008. **77**(14): p. 144104.

109. Herder, L.M., J.M. Bray, and W.F. Schneider, *Comparison of cluster expansion fitting algorithms for interactions at surfaces.* Surface Science, 2015. **640**: p. 104-111.

110. Ceder, G. and A. Van der Ven, *Phase diagrams of lithium transition metal oxides: investigations from first principles.* Electrochimica Acta, 1999. **45**(1): p. 131-150.

111. Andersen, E.D., C. Roos, and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization.* Mathematical Programming, 2003. **95**(2): p. 249-277.

112. Ceder, G., *Alloy theory and its applications to long-period superstructure ordering in metallic alloys and high-temperature superconductors*. 1991, California Univ., Berkeley, CA (United States).

113. Hawkins, D.M., *The problem of overfitting.* J Chem Inf Comput Sci, 2004. **44**(1): p. 1-12.

114. Candès, E.J., J. Romberg, and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information.* IEEE Transactions on information theory, 2006. **52**(2): p. 489-509.

115. Winston, W.L., M. Venkataramanan, and J.B. Goldberg, *Introduction to mathematical programming*. Vol. 1. 2003: Thomson/Brooks/Cole Duxbury; Pacific Grove, CA.

116. Kim, S.-J., et al. *An efficient method for compressed sensing*. in *2007 IEEE International Conference on Image Processing*. 2007. IEEE.

117. Urban, A., D.-H. Seo, and G. Ceder, *Computational understanding of Li-ion batteries*. npj Computational Materials, 2016. **2**: p. 16002.

118. Goodenough, J.B. and K.S. Park, *The Li-ion rechargeable battery: a perspective*. J Am Chem Soc, 2013. **135**(4): p. 1167-76.

119. Hong, J., et al., *Review—Lithium-Excess Layered Cathodes for Lithium Rechargeable Batteries*. Journal of The Electrochemical Society, 2015. **162**(14): p. A2447-A2467.

120. Rozier, P. and J.M. Tarascon, *Review—Li-Rich Layered Oxide Cathodes for Next-Generation Li-Ion Batteries: Chances and Challenges*. Journal of The Electrochemical Society, 2015. **162**(14): p. A2490-A2499.

121. Urban, A., J. Lee, and G. Ceder, *The Configurational Space of Rocksalt ‐ Type Oxides for High ‐ Capacity Lithium Battery Electrodes*. Advanced Energy Materials, 2014. **4**(13).

122. Lee, J., et al., *Unlocking the potential of cation-disordered oxides for rechargeable lithium batteries*. Science, 2014. **343**(6170): p. 519-22.

123. Mizushima, K., et al., *LixCoO2 (0< x<-1): A new cathode material for batteries of high energy density*. Materials Research Bulletin, 1980. **15**(6): p. 783-789.

124. Glazier, S.L., et al., *Characterization of Disordered Li (1+ x) Ti2 x Fe (1–3 x) O2 as Positive Electrode Materials in Li-Ion Batteries Using Percolation Theory*. Chemistry of Materials, 2015. **27**(22): p. 7751-7756.

125. Lee, J., et al., *A new class of high capacity cation-disordered oxides for rechargeable lithium batteries: Li–Ni–Ti–Mo oxides*. Energy & Environmental Science, 2015. **8**(11): p. 3255-3265.

126. Bongers, P., *Structure and Magnetic Properties of Several Complex Oxides of the Transition Elements*. University of Leiden thesis, 1957.

127. Lecerf, A. *Sur Quelques Proprietes Chimiques Des Oxydes TiO Et Ti2O3-Preparation Et Etudes De Nouveaux Composes Ternaires Oxygenes Du Titane Trivalent*. in *ANNALES DE CHIMIE FRANCE*. 1962.

128. Hoffmann, A., *Crystal Chemistry of Lithium Ferrite*. Naturwissenschaften, 1938. **26**(26): p. 431.

129. Posnjak, E. and T.F. Barth, *A New Type of Crystal Fine-Structure: Lithium Ferrite (Li 2 O· Fe 2 O 3)*. Physical Review, 1931. **38**(12): p. 2234.

130. Perdew, J.P., K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*. Physical review letters, 1996. **77**(18): p. 3865.

131. Anisimov, V.V., J. Zaanen, and O.K. Andersen, *Band theory and Mott insulators: Hubbard U instead of Stoner I*. Phys Rev B Condens Matter, 1991. **44**(3): p. 943-954.

132. Jain, A., et al., *A high-throughput infrastructure for density functional theory calculations*. Computational Materials Science, 2011. **50**(8): p. 2295-2310.

133. Hart, G.L. and R.W. Forcade, *Algorithm for generating derivative structures*. Physical Review B, 2008. **77**(22): p. 224115.

134. Bishop, C.M., *Pattern recognition*. Machine Learning, 2006. **128**.