The Complexity of Periodic Energy Minimisation

Duncan Adamson
ICE-TCS, Department of Computer Science, Reykjavik University, Iceland

Argyrios Deligkas
Department of Computer Science, Royal Holloway, University of London, UK

Vladimir V. Gusev
Materials Innovation Factory, Department of Computer Science, University of Liverpool, UK

Igor Potapov
Department of Computer Science, University of Liverpool, UK

Abstract

The computational complexity of pairwise energy minimisation of N points in real space is a long-standing open problem. The idea of the potential intractability of the problem was supported by a lack of progress in finding efficient algorithms, even when restricted the integer grid approximation. In this paper we provide a firm answer to the problem on $\mathbb{Z}^d$ by showing that for a large class of pairwise energy functions the problem of periodic energy minimisation is NP-hard if the size of the period (known as a unit cell) is fixed, and is undecidable otherwise. We do so by introducing an abstraction of pairwise average energy minimisation as a mathematical problem, which covers many existing models. The most influential aspects of this work are showing for the first time: 1) undecidability of average pairwise energy minimisation in general 2) computational hardness for the most natural model with periodic boundary conditions, and 3) novel reductions for a large class of generic pairwise energy functions covering many physical abstractions at once. In particular, we develop a new tool of overlapping digital rhombuses to incorporate the properties of the physical force fields, and we connect it with classical tiling problems. Moreover, we illustrate the power of such reductions by incorporating more physical properties such as charge neutrality, and we show an inapproximability result for the extreme case of the 1D average energy minimisation problem.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness

Keywords and phrases Optimisation of periodic structures, tiling, undecidability, NP-hardness

Digital Object Identifier 10.4230/LIPIcs.MFCS.2022.8

Funding *Duncan Adamson*: supported by Icelandic Research Fund grant no. 217965.
Vladimir V. Gusev: The author is supported by the Leverhulme Trust via the Leverhulme Research Centre for Functional Materials Design.
Igor Potapov: partially supported by EP/R018472/1.

1 Introduction

Periodic structures and models with periodic boundary conditions appear both in nature and in mathematical interpretations of physical phenomena: spin systems in Ising models [8], Buckingham–Coulomb inter-atomic potential modelling crystal structures [32], Lennard-Jones potential in inter-molecular interaction [13]. Periodic boundary conditions are often used either to define, or to approximate, a large or infinite system from a small partition, known as a *unit cell*. The series of repeating unit cells in every dimension forms a periodic structure. A unit cell can be defined as a mapping from the points within a contiguous subspace of a lattice to a finite set of "colours", an abstraction that may be used to represent anything from ions to spin states. The main advantage of this model is that the periodic structure allows the properties of the effectively infinite global structure to be determined from the finite unit cell. This advantage has led to these structures attracting a great deal of attention in mathematics, physics, biology, chemistry and computer science [1, 12, 17, 18, 19, 20].

From the perspective of the physical sciences, one of the most fundamental properties is the *potential energy* of the structure, representing the sum of pairwise attractions between ions (or spins) within the unit cell. Usually, the energy landscape is a highly non-convex function with many local minima and saddle points. The fundamental optimisation problem, associated with predicting various physical phenomena in both periodic structures and a single unit cell, is known as the *cluster problem* [34]:

> "In which way can N points be occupied (in real space) so as to minimise the sum of their interactions?"

A lack of progress with the design of efficient algorithms for solving this fundamental optimisation problem led to various hypotheses about its intractability, which has not been formally addressed due to a large number of variants and complexity of practical details regarding the problem. However, very recently the first NP-hard result was shown for closely related *removal problem*: Given a cluster of N points, can a subset of them be removed to minimise the total energy of pairwise Buckingham–Coulomb interaction [3].

This paper builds upon these previous results, showing that hardness results apply even under the more realistic *periodic boundary conditions*. In the context of crystal structure prediction, this refers to the periodic structure of the crystal, meaning that the global structure of the crystal is represented by a repeating period. Previous work [3, 11] has focused on the interaction within the unit cell, while ignoring this periodic conditions.

In this paper, we propose a more universal approach and analyse the computational complexity of the original cluster problem for a large class of pairwise energy minimisation functions under more realistic periodic boundary conditions. We introduce an abstract class of r-distance *Common Minimum Value* functions ($\mathcal{CMV}(r)$), which capture typical properties of classical force-fields of pairwise interaction (attraction, repulsion) and incorporates the variable depth r of such interaction. Then we show that the cluster problem is NP-hard if the size of the unit cell is fixed and that it is undecidable otherwise for any function in the class $\mathcal{CMV}(r)$ defined on two or three dimensional grids ($\mathbb{Z}^2$ or $\mathbb{Z}^3$). Moreover, we show that particular known classical energy-interaction functions fit to this class and inherit the results on the computational complexity. In the case of 1D grids, we design a parameterised polynomial-time algorithm to solve the fixed period pairwise energy minimisation problem. Finally, we show that under the extra physical constraint of charge-neutrality (the total sum of charges/weights associated with points in the unit cell is zero) the problem still remains undecidable for 2 and 3 dimensions, and in dimension one it cannot be approximated within any constant factor unless $P = NP$.

Crystal Structure Prediction and Computational Complexity. Predicting crystal structures by computational methods without experimental input is the Holy Grail of crystallography and material science; it has remained a noted open problem for over 30 years [24]. In general, Crystal Structure Prediction (CSP) asks to identify the periodic crystal structure from a given set of *ions* – electromagnetically charged atoms – that minimise its potential energy based on some model of interaction.

A crystal is a structure defined by a repeating period called a unit cell. Informally, the unit cell can be thought of as a three dimensional box containing *ions*, see Figure 1. Each ion belongs to a class called a *species*, determining the properties of the ion. The unit cell acts as a periodic mapping from some set of ion species to the space $\mathbb{R}^3$, or in the discrete setting to a grid such as the integer grid $\mathbb{Z}^3$. In a discrete space, the unit cell can be represented as a necklace or bracelet [4, 2]. In the most general formulations of CSP the size and shape of the unit cell are unconstrained, however bounding the size and the shape is a common restriction for many cases of CSP [12].

Crystal structure prediction can be thought of as the problem of finding the "best" configuration of ions within a three-dimensional box, but it is also important for dimension one and two, see [12]. The quality of a configuration is determined by the average pairwise interaction between each ion in the structure. The pairwise interaction in turn is determined by an energy function, taking as input the distance between ions and a set of parameters based on the ion species. A positive interaction between ions corresponds to a force pushing the ions apart, while a negative interaction indicates a force of attraction. The goal of CSP is to find a stable structure, indicated by having the minimal average pairwise interaction [23].

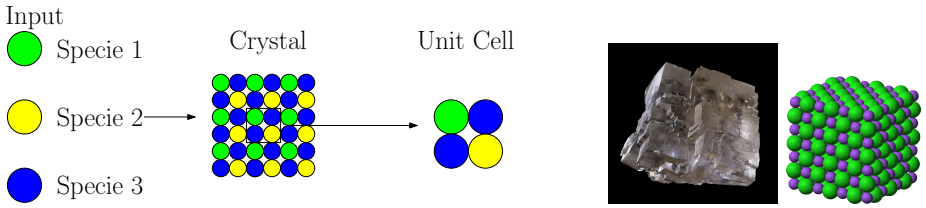

Figure 1 Left: A high level example of CSP. The input is a set of 3 species of ions (green, yellow, and blue), where each pair has an interaction computed by a given function determined by the distance and species. This set of species is transformed into a crystal structure (middle) defined by a unit cell (right). **Right:** Example of a Sodium Chloride crystal with the ionic structure. The middle and right pictures are shared under the creative commons licence.

Despite countless heuristics attempts such as quasi-random sampling [27, 29], basin hoping [6, 15, 16], simulated annealing [26, 30], swarm optimisation [9, 33], and genetic algorithms [14, 22, 25] the computational complexity of the CSP problem has not been investigated [10]. The recent interdisciplinary initiative to combine chemical knowledge with state of the art computer science techniques has lead to the first formalisation of CSP as a theoretical computer science problem [3, 6].

Our Contributions. This work introduces the average pairwise energy minimisation problem on $\mathbb{Z}^d$ as a generalisation of the physically motivated models and approximation of real space using the integer grid $\mathbb{Z}^d$. This problem can be seen as a variant on the well studied class of tiling problems [5, 7]. Rather than the "hard" constraints of a tiling problem, where tiles can only be placed adjacent to each other if they fulfil a set of strict conditions, our model uses "soft" constraints, giving an energy value to the interaction between each pair of vertices in the grid based on distance and the colour of the vertices.

Our main result is showing the average pairwise energy minimisation problem on $\mathbb{Z}^d$ to be NP-hard when the size of the unit cell is fixed and is undecidable otherwise. This strengthens the argument that CSP is intractable for a fixed-size unit cell and undecidable in general.

Our proof of both intractability and undecidability come by way of a series of reductions starting with the periodic tiling problem. This series of reductions is designed to enable us to more easily encode the concept of *orientation* into the pairwise interaction constraints that the average pairwise energy minimisation problem on $\mathbb{Z}^d$ uses. In the periodic tiling problem it is necessary for all tiles to have a shared orientation in order for the undecidability results to hold, however our physically motivated models determine interaction only by the colour of the vertices, and the distance between them. To encode this property we introduce two problems: the *k-unique radius tiling problem* (defined in Section 3.1), and the r-discretised rhombus all-distinct periodic complete assignment problem (defined in Section 3.2). Figure 2 provides a sketch of this process. We strengthen our results by showing that they hold under the further constraint of *charge neutrality*, an abstraction of the physical constraint that the periods of these structures must have an equal number of positive and negative charges.

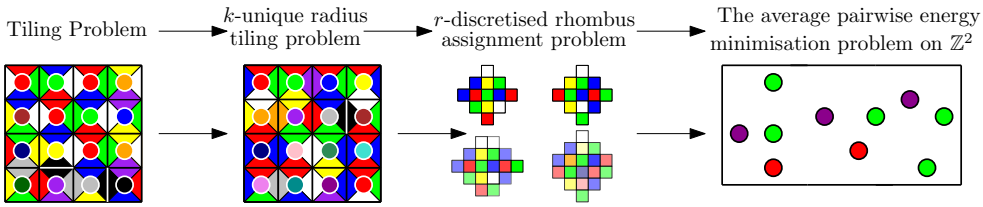

Figure 2 High level overview our series of reductions starting with the tiling problem.

2 Preliminaries

Informally, the average pairwise energy minimisation problem on $\mathbb{Z}^d$ can be thought of as the problem of determining a way of colouring the infinite grid $\mathbb{Z}^d$ with a *finite period*, while minimising the average pairwise interaction energy. The pairwise interaction energy between each pair of points on $\mathbb{Z}^d$ is determined by the colours of the points, and the distance between them. The period of a colouring is called the *unit cell*, which may equivalently be thought of as a mapping from the set of colours to the grid.

▶ **Definition 1** (Unit cell). *A* unit cell U *of size* $\vec{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{N}^d$ *is a periodic mapping from the integer grid* $\mathbb{Z}^d$ *to some set of colours* $\mathcal{C}$*, defined by a colouring on the* d *dimensional grid* $n_1 \times n_2 \times \ldots \times n_d$*. Given a vector* $\vec{y} \in \mathbb{Z}^d$*,* $U(\vec{y})$ *returns the colour at position* $(y_1 \bmod n_1, y_2 \bmod n_2, \ldots, y_d \bmod n_d)$ *on the grid defining* U*.*

The number of vertices in a unit cell U of size $\vec{n}$ is denoted by $|U|$, i.e. $|U| = n_1 \cdot n_2 \cdot \ldots \cdot n_d$. Similarly $\vec{x} \in U$ is used to denote that $\vec{x}$ is a position in the finite grid defining U. Where it is clear from context, given any vector $\vec{x} \in \mathbb{Z}^d$ the colour of the vertex at position $\vec{x}$ in the grid $\mathbb{Z}^d$ coloured by U is denoted $U(\vec{x})$, giving $U(\vec{x}) = U((x_1 \bmod n_1, x_2 \bmod n_2, \ldots, x_d \bmod n_d))$.

The goal of these colourings is to minimise the *average pairwise energy per vertex* of the coloured grid. The energy between two vertices represents the force between them, with a negative energy indicating attraction and a positive energy indicating repulsion. The pairwise energy between a pair of vertices in the grid is determined by a *pairwise energy function*.

This work considers parametric pairwise energy functions f of the form $f(\overline{\theta}_{(c_i,c_j)}, r)$ where $c_i, c_j \in \mathcal{C}$ are a pair of colours, $r \in \mathbb{R}$ is a euclidean distance and $\overline{\theta}_{(c_i,c_j)} \in \mathbb{R}^p$ is a vector of p parameters determined by the colours c_i and c_j. Further, this work assumes that the vector of parameters $\overline{\theta}_{(c_i,c_j)}$ are predefined for every pair of colours $c_i, c_j \in \mathcal{C}$. Each function returns a scalar real value, i.e. $f\colon \left(\overline{\theta}_{(c_i,c_j)} \in \mathbb{R}^p, r \in \mathbb{R}\right) \mapsto \mathbb{R}$.

▶ **Definition 2** (Average pairwise energy per vertex). *Given a unit cell* U *of size* $\vec{n}$ *colouring the grid* $\mathbb{Z}^d$*, the* average pairwise energy per vertex *is given by:*

$$\mathrm{AE}(U) = \frac{1}{|U|} \sum_{\vec{x} \in U} \sum_{\vec{y} \in \mathbb{Z}^d} f(\overline{\theta}_{(U(\vec{x}),U(\vec{y}))}, D(\vec{x}, \vec{y}))$$

where f *is the pairwise energy function,* $D(\vec{x}, \vec{y})$ *denotes the euclidean distance between* $\vec{x}$ *and* $\vec{y}$*, and* $\overline{\theta}_{(c_i,c_j)} \in \mathbb{R}^p$ *is a vector of* p *parameters.*

In this paper we assume that each energy function has a cut off distance, allowing the average pairwise energy per vertex to be compute efficiently. One further constraint that we introduce is that of *charge neutrality*. In this setting, every colour is associated with a integer charge. Given a unit cell U, the charge of $\vec{x} \in U$ is denoted $Q(U(\vec{x}))$. Note that the charge of any two points assigned the same colour are equal, i.e. if $U(\vec{x}) = U(\vec{y})$ then

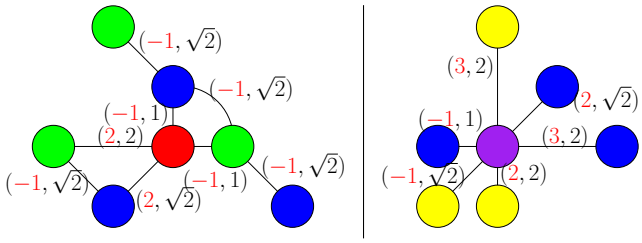

Figure 3 An overview of $\mathcal{CMV}(r)$. The energy and distance between each point is labelled as a pair (a, b) where a represents the energy and b represents the distance.

$Q(U(\vec{x})) = Q(U(\vec{y}))$ for any pair of vectors $\vec{x}, \vec{y} \in U$. A unit cell U is *charge neutral* if and only if $\sum_{\vec{x} \in U} Q(U(\vec{x})) = 0$. In general we assume that the charge neutrality constraint can be ignored, effectively assuming that the charge is 0 for every colour.

2.1 The r-Distance Common Minimal Value Class

In this paper we restrict our energy functions to the class of r-distance common minimal value functions, denoted $\mathcal{CMV}(r)$, introduced in this section. This class of functions focuses on the interactions between vertices within a distance of r of each other, for some distance $r \in \mathbb{R}$. In order to simplify reasoning on the set $\mathcal{CMV}(r)$, it is assumed that given any distance $d > r$ the value of $f(\overline{\theta}_{(c_i,c_j)}, d) = 0$ for every pair of colours $c_i, c_j \in \mathcal{C}$. Let $R(r) = \{(i, j) \in \mathbb{Z}^2 \mid \sqrt{i^2 + j^2} \leq r, (i, j) \neq (0, 0)\}$ be the set of vertices on the integer grid $\mathbb{Z}^2$ within a distance of at most r of the central point $(0, 0)$. Further, let $d(r) = \{\sqrt{i^2 + j^2} \mid (i, j) \in R(r)\}$ be the set of possible distances between the central vertex and any vertex in $R(r)$. As an example, $R(2) = \{(2, 0), (1, 1), (1, 0), (1, -1), (0, 2), (0, 1), (0, -1), (0, -2), (-1, 1), (-1, 0), (-1, -1), (-2, 0)\}$ and $d(r) = \{1, \sqrt{2}, 2\}$. The goal of this class is to be able to "fix" the optimal distance between pair of colours as either being some distance in $d(r)$, or as being outside of $R(r)$ – in effect penalising two colours at a distance of r or less. To this end this work uses the idea of a *common minimal value*. Informally, the common minimal value can be thought of as some negative value M such that the smallest possible interaction between any pair of vertices is M. Further, the functions in this work restrict M to appear at most once in the set of possible distance between each colour, meaning that given some pair of colours c_i and c_j, there exists at most one distance $d \in d(r)$ such that $f(\theta_{(c_i,c_j)}, d) = M$. The following definition formalises the *r-distance common minimal value class*.

▶ **Definition 3** (Common Minimal Value Functions ($\mathcal{CMV}(r)$)). *Let $\overline{\theta}_{(c_i,c_j)} \in \mathbb{R}^p$ denote the vector of parameters assigned to some pair of colours $c_i, c_j \in \mathcal{C}$. The function $f(\overline{\theta}_{(c_i,c_j)}, d)$: $\mathbb{R}^{p+1} \to \mathbb{R}$ belongs to the* **class of common minimal value functions** *$\mathcal{CMV}(r)$ for $r \in \mathbb{R}$ if there exists a common minimum value $M \in \mathbb{R}$ for which the following hold:*

1. *For any two points at a distance $d > r$ and any pair of colours c_i,$c_j \in \mathcal{C}$ the value of $f(\overline{\theta}_{(c_i,c_j)}, d) = 0$;* **[Cut-off property]**.
2. *For any two colours c_i,$c_j \in \mathcal{C}$ it is possible to determine a vector $\overline{\theta}_{(c_i,c_j)}$ such that the energy between any pairs of points at any distance $d \in d(r)$ is $f(\overline{\theta}_{(c_i,c_j)}, d) > M$;* **[Separation property]**.
3. *For any two colours c_i,$c_j \in \mathcal{C}$ and any distance $d \in d(r)$ it is possible to determine a vector $\overline{\phi}_{(c_i,c_j)}$ such that the energy between any pair of points at distance d is $f(\overline{\phi}_{(c_i,c_j)}, d) = M$ and the energy between any pair of points at any distance $d' \in d(r), d' \neq d$ is $f(\overline{\phi}_{(c_i,c_j)}, d') > M$;* **[Optimal pairwise distance property]**.

An overview of these properties is given in Figure 3. These properties are used to encode the tiling problem into the average pairwise energy minimisation problem on $\mathbb{Z}^d$. The cut-off property (Property 1) ensures that there is no interaction between vertices over a certain cut off distance, allowing these interactions to be safely ignored. The separation property (Property 2) ensures that there exists a vector of parameters such that the corresponding colours must be placed further than r apart, or suffer a small energy penalty by having an interaction greater than M. Finally the optimal pairwise distance property (Property 3) ensures that there exists a vector of parameters such that the interaction of the corresponding colours is minimised at M at exactly one distance. The goal of these conditions is to be able to force a structure on the colouring based on the relative distances between colours. This allows the structure of the tiling problem to be utilised in the setting of the average pairwise energy minimisation problem on $\mathbb{Z}^d$.

2.2 The Pairwise Energy Minimisation Problem

This section introduces our central problem, the *average pairwise energy minimisation problem on* $\mathbb{Z}^d$. In this paper we consider two versions of this problem, depending on the constraints placed on the unit cell. In the most general case, the only constraint is that the average energy of the unit cell is below some bound g. In this paper, we limit the energy functions to the class $\mathcal{CMV}(r)$. The unit cell may be constrained by having the size given as part of the input. All values are given in binary as input to our problems.

▶ **Problem 1.** *The average pairwise energy minimisation problem on* $\mathbb{Z}^d$.

Input: A goal energy $g \in \mathbb{Q}$, a set of colours $\mathcal{C}$, a number of dimensions $d \in \mathbb{Z}$ an energy function $f \in \mathcal{CMV}$ and a set of $|\mathcal{C}|^2$ parameters $\overline{\theta}_{(c_i,c_j)} \in \mathbb{R}^p$.

Question: Does there exist a unit cell U of size $n_1 \times n_2 \times \ldots \times n_d$ for some $n_i \in \mathbb{N}^+$ where $\mathrm{AE}(U) \leq g$.

When the size of the unit cell is given as an input in the form of a vector of length d of the form $(n_1, n_2, \ldots, n_d)$, we refer to the problem as the **average pairwise energy minimisation problem on** $\mathbb{Z}^d$ **with a fixed period**. Here, *fixed period* refers to the size of the period being fixed as part of the input, in this case restricting the period to be of size $n_1 \times n_2 \times \ldots \times n_d$ for the given $n_1, n_2, \ldots, n_d$.

3 Undecidability for Unconstrained Period Size

We first look at the unbounded setting, where the size of the unit cell is not taken as part of the input. The main claim in this section is that Problem 1 is undecidable for any function in $\mathcal{CMV}(r)$ for $r \geq 2$. This section is split into three parts. First, we provide some background on the tiling problem that is used as the basis for this reduction. Second, we provide an auxiliary problem derived from the tiling problem to act as an intermediary step in proving the undecidablity of the average pairwise energy minimisation problem on $\mathbb{Z}^d$. Finally we prove the undecidablity of the average pairwise energy minimisation problem on $\mathbb{Z}^d$.

3.1 The Tiling Problem

In the tiling problem, we are given a set of *tiles*, square plates with a fixed orientation where each edge is coloured from some set of colours $\mathcal{C}$. The goal of a tiling problem is to completely cover the plane with tiles such that every pair of adjacent tiles is coloured the same along the shared edge. In this section, we introduce the further constraint that no two copies of the same tile may be within a distance of k or less of each other.

Before discussing the new variations of the tiling problem, let us first present some notation. The edges of the tiles are labelled *East,West,North* and *South* such that the East edge is opposite the West edge, and the South edge is opposite the North edge. More precisely, given two tiles, v at position (x_1, y_1) and u at position (x_2, y_2) respectively such that $|x_1 - x_2| \leq 1$ and $|y_1 - y_2| \leq 1$, we say that:

	$x_1 < x_2$	$x_1 = x_2$	$x_1 > x_2$
$y_1 < y_2$	v is **North-West** of u	v is **North** of u	v is **North-East** of u
$y_1 = y_2$	v is **West** of u	v is u	v is **East** of u
$y_1 > y_2$	v is **South-West** of u	v is **South** of u	v is **South-East** of u

A tile t is represented by the four edges composing it. For notation let t_e be the colour of the tile t along edge e. A tile t can be represented as $t = \{t_{East}, t_{South}, t_{West}, t_{North}\}$. A *Tile Set* is a set of tiles with the edges coloured from some set of colours $\mathcal{C}$. It is assumed the tile set contains an infinite number of copies of each tile, allowing the complete plane to be covered with these tiles. The goal of the *Tiling problem* is to assemble copies of the tiles from a given tile set on an infinite plane ruled into squares of the size of one tile such that:

1. No tile is rotated or reflected.
2. A tile must be placed *exactly* over one square of the ruled integer plane.
3. The colour of adjacent edges must match.
4. Every square must be covered by one tile.

This problem is *solvable* for a given tile set if and only if such an assembly exists. An assembly is periodic if there exists some finite region of the plane that may be repeated so as to solve the tiling problem. The *Periodic Tiling Problem* asks if there is such a periodic assembly. Both the tiling problem and the periodic tiling problem are classical undecidable problems [5, 7]. Connections between this problem and chemistry are well established [28].

In this paper, we introduce the *k-unique radius* variant of the tiling problem to act as an intermediate problem between the general tiling problem and the average pairwise energy minimisation problem on $\mathbb{Z}^2$. The k-unique radius tiling problem is needed to help encode the notion of *orientation* into the average pairwise energy minimisation problem on $\mathbb{Z}^2$. In the tiling problem, it is integral that each tile is placed under the same orientation. This means that given two adjacent tiles, they must either touch West edge to East edge, or North edge to South edge. As our setting uses only the colours and distance between vertices to determine the pairwise energy, the concept of orientation is difficult to encode. Informally a tiling has a k-unique radius if and only if no two copies of a given tile are within a distance of at most k of each other. Let T be a tiling of $\mathbb{Z}^2$ such that $T(i, j)$ returns the tile at position (i, j). The tiling T has a k-unique radius if and only if for every $(i, j), (x, y) \in \mathbb{Z}^2$, where $D((i, j), (x, y)) \leq k$ the tile $T(i, j)$ is distinct from $T(x, y)$, i.e. $T(i, j) \neq T(x, y)$ where $D((i, j), (x, y))$ returns the distance between (i, j) and (x, y).

▶ **Problem 2.** *The periodic tiling problem with a k-unique radius.*

Input: A set of tiles, $\mathcal{T}$, and integer k
Question: Does there exist a periodic tiling of $\mathbb{Z}^2$ made from $\mathcal{T}$ such that given any tile t at position (x, y) there exists no other copy of t within a distance of k from (x, y)?

▶ **Proposition 4.** *The periodic k-unique radius tiling problem is undecidable for any $k \in \mathbb{N}$.*

Proof Sketch. The undecidability of the periodic k-unique radius tiling problem follows from the undecidability of the periodic domino problem [7]. The high level idea is to create a set of k^2 copies of each tile, labelled with $(x, y) \in [k]$. A set of additional colours are constructed

such that given two tiles $t_{(x,y)}$ and $s_{(a,b)}$, $t_{(x,y)}$ can be placed adjacent to $s_{(a,b)}$ if and only if the original tiles t and s can be placed adjacently, and (a, b) is adjacent to (x, y) on the $k \times k$ toroidal grid. In one direction, any valid tiling with the original set of tiles can be transformed into a valid tiling for the new set by choosing an arbitrary origin point, and replacing the tile t at position (x, y) with the tile $t_{(x \bmod k, y \bmod k)}$. In the other direction, any tiling using the new tiles can be transformed into a tiling of the original tiles by simply replacing each tile $t_{(x,y)}$ with the tile t from the original set. ◀

▶ **Problem 3.** *The fixed period k-unique tiling problem.*

Input:	A set of tiles, $\mathcal{T}$, integer k, and pair of lengths n_1, n_2.
Question:	Does there exist a periodic tiling of the plane of size $n_1 \times n_2$ over $\mathcal{T}$ where every tile within a distance of k for every other tile is distinct.

▶ **Proposition 5.** *The fixed period k-unique tiling problem is NP-hard.*

Proof. Following the same arguments as in Proposition 4, the fixed period tiling problem can be reduced to the fixed period k-unique tiling problem. As the fixed period tiling problem is known to be NP-hard [5, 7, 21], the fixed period k-unique tiling problem is NP-hard. ◀

3.2 Tiling with Overlapping Digitised Rhombuses

This section covers the problem of completely covering the integer grid $\mathbb{Z}^2$ using *overlapping digitised rhombuses*. Informally, a *digitised rhombus* with radius r can be thought of as a set of mono-chromatically coloured tiles organised as a rhombus from some set of colours $\mathcal{C}$.

▶ **Definition 6** (Digitised rhombus). *A **digitised rhombus** of radius r is the mapping from the grid $\{(x, y) \in \mathbb{Z}^2 \mid |x| + |y| \leq r\}$ to a set of colours $\mathcal{C}$.*

Given a rhombus R and position $(i, j) \in \{(x, y) \in \mathbb{Z}^2 \mid |x| + |y| \leq r\}$, $R_{i,j}$ is used to denote the colour mapped by R to position (i, j), i.e. the colour of the tile at position (i, j) in the rhombus. A rhombus is *distinctly coloured* if $R_{i,j} \neq R_{l,m}$ for every pair of positions $(i, j), (l, m) \in \{(x, y) \in \mathbb{Z}^2 \mid |x| + |y| \leq r\}$ where $(i, j) \neq (l, m)$.

We use a set of rhombuses $\mathcal{R}$ analogously to the set of tiles $\mathcal{T}$ used in tiling problems. Given the integer grid $\mathbb{Z}^2$, the assignment of a rhombus R to the position $(x, y) \in \mathbb{Z}^2$ is equivalent to colouring every vertex within a radius of r of (x, y) using R. For the remainder of this section, we focus on the Manhattan distance, defined as $D((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|$. Given a rhombus R assigned to position $(x, y) \in \mathbb{Z}^2$, the position (x', y') at a distance of no more than r from (x, y) is coloured $R_{x'-x,y'-y}$.

The focus in this work is on *overlapping rhombuses*. Given a pair of digitised rhombuses of radius $r \in \mathbb{N}$, R and S at positions (x_1, y_1) and (x_2, y_2), R overlaps S if the Manhattan distance between R and S is no more than $2r$. The *overlap* between R and S are the set of positions that are assigned colours by both R and S. This corresponds to the set of positions $\{(a, b) \in \mathbb{Z}^2 \mid D((a, b), (x_1, y_1)) \leq r$ and $D((a, b), (x_2, y_2)) \leq r\}$. Informally, R and S properly overlap when centred at (x_1, y_1) and (x_2, y_2) if every position in the overlap is assigned the same colour by both R and S. See Figure 4 for an example.

▶ **Definition 7** (Overlapping rhombuses). *Let R and S be a pair of r-radius digitised rhombuses centred on positions (x_1, y_1) and (x_2, y_2) respectively. Rhombuses R and S **properly overlap** if and only if for every position $(i, j) \in \{(a, b) \in \mathbb{Z}^2 \mid D((a, b), (x_1, y_1)) \leq r$ and $D((a, b), (x_2, y_2)) \leq r\}$ it holds that $R_{i-x_1,j-y_1} = S_{i-x_2,j-y_2}$.*

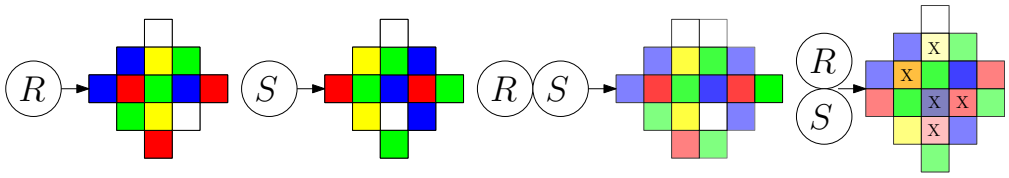

Figure 4 An example outlining how two rhombuses R and S may properly overlap (left) and when not (right). Note that X is used to denote a conflict between the two overlapping rhombuses.

Note that any two rhombuses at a distance greater than $2 \cdot r$ from each other properly overlap following Definition 7 as the set $\{(a, b) \in \mathbb{Z}^2 \mid D((a, b), (x_1, y_1)) \leq r \text{ and } D((a, b), (x_2, y_2)) \leq r\}$ is empty. An *assignment* of rhombuses $\mathcal{R}$ to the integer grid $\mathbb{Z}^2$ equates to a complete colouring of $\mathbb{Z}^2$ using the rhombuses in $\mathcal{R}$ as the colours. An assignment is valid if and only if a rhombus is centred on every vertex in $\mathbb{Z}^2$ and every pair of rhombuses properly overlap.

▶ **Definition 8** (Rhombus assignment). *An* ***assignment*** *A of rhombuses from the set $\mathcal{R}$ to $\mathbb{Z}^2$ is a mapping from $\mathbb{Z}^2$ to $\mathcal{R}$ such that $A : (x, y) \in \mathbb{Z}^2 \mapsto R$ for every $x, y \in \mathbb{Z}$. Let $A(x, y) : \mathbb{Z}^2 \mapsto \mathcal{R}$ return the rhombus assigned to position $(x, y) \in \mathbb{Z}^2$. An assignment A is* valid *if and only if $\forall (x_1, y_1), (x_2, y_2) \in \mathbb{Z}^2$, the rhombus $A(x_1, y_1)$ properly overlaps $A(x_2, y_2)$.*

▶ **Definition 9** (Periodic assignment). *An assignment A from $\mathbb{Z}^2$ to $\mathcal{R}$ is* ***periodic*** *if there exists a tuple $(a, b) \in \mathbb{Z}^2$ such that $A(x, y) = A(x \bmod a, y \bmod b)$ for every tuple $(x, y) \in \mathbb{Z}^2$.*

▶ **Problem 4.** *The r-discretised rhombus all-distinct periodic complete assignment problem.*

Input: A set $\mathcal{R}$ of r-radius digitised rhombuses
Question: Does there exist a valid periodic assignment of $\mathcal{R}$ to $\mathbb{Z}^2$?

▶ **Theorem 10.** *The r-discretised rhombus all-distinct periodic complete assignment problem is undecidable for any $r \geq 1$.*

Proof Sketch. This theorem is proven by transforming a set of tiles into a set of rhombuses. The set of rhombuses is constructed by taking the set of unique tilings on the grid $\{(x, y) \in \mathbb{Z}^2 \mid x^2 + y^2 \leq 1\}$, and creating a rhombus corresponding to each tiling. Each tile is represented in this model by a unique colour. ◀

The same construction as in Theorem 10 are used to derive an NP-completeness result for the fixed period r-discretised rhombus all-distinct periodic complete assignment problem. Observe that for a set of q rhombuses, and fixed period $n_1 \times n_2$, there are $q^{n_1 \cdot n_2}$ possible coverings, therefore a brute force algorithm can solve this problem in $O(q^{n_1 \cdot n_2})$-time, and therefore the problem belongs to NP. For hardness, Proposition 5 is used to establish the hardness of the fixed-period k-unique radius tiling problem, and by extension the hardness of the fixed period r-discretised rhombus all-distinct periodic complete assignment problem.

▶ **Corollary 11.** *The fixed period r-discretised rhombus all-distinct periodic complete assignment problem is NP-complete for any $r \geq 1$.*

3.3 Pairwise Energy Minimisation Problem on $\mathbb{Z}^2$ and $\mathbb{Z}^3$

With the undecidability of the r-discretised rhombus all-distinct periodic complete assignment problem established, the next step is to show how to reduce the r-discretised rhombus all-distinct periodic complete assignment problem to the average pairwise energy minimisation problem on $\mathbb{Z}^d$. In this section, the pairwise energy function is assumed to be a member of

the 2-distance common minimal value class, $\mathcal{CMV}(2)$. At a high level the reduction from the r-discretised rhombus all-distinct periodic complete assignment problem is done by encoding each rhombus as a colour, then tuning the parameters of the pairwise energy function so that a valid colouring of the grid corresponds to a valid assignment of rhombuses.

The main challenge of this encoding is due to the definition of $\mathcal{CMV}(2)$. Namely, for any energy function in $\mathcal{CMV}(2)$ the pairwise energy between vertices is determined solely by the distance between vertices, colour of each vertex, and some given vector of parameters. This means that given two adjacent vertices v_i and v_j, coloured c_i and c_j respectively, the energy between v_i and v_j is the same irrespective of the relative direction of each tile.

To simplify our reduction, we introduce some additional notation. Given two vertices v_i at position (x_1, y_1) and v_j at position (x_2, y_2), v_i is said to be *directly adjacent* to v_j if $|x_1 - x_2| + |y_1 - y_2| = 1$. Similarly v_i is *diagonally adjacent* to v_j if $|x_1 - x_2| = 1$ and $|y_1 - y_2| = 1$. Finally, v_i is *peripherally adjacent* if either $|x_1 - x_2| = 2$ and $y_1 = y_2$, or $x_1 = x_2$ and $|y_1 - y_2| = 2$. Further, in Proposition 12 we assume M to be the common minimal value for all functions in $\mathcal{CMV}(r)$.

▶ **Proposition 12.** *Let $\mathcal{R}$ be a set of distinctly coloured rhombuses and let $\mathcal{C}(\mathcal{R})$ be a set of $|\mathcal{R}|$ colours such that for every rhombus $r \in \mathcal{R}$, there exists some colour $C_r \in \mathcal{C}(\mathcal{R})$. For any pairwise energy function $f \in \mathcal{CMV}(2)$, and pair of rhombuses $i, j \in \mathcal{R}$ there exists some vector of parameters $\overline{\theta}_{(c_i,c_j)}$ such that $f(\overline{\theta}_{(c_i,c_j)}, r)$ satisfies:*

1. *If i and j properly overlap when i is centred at some position directly adjacent to j then $f(\overline{\theta}_{(c_i,c_j)}, 1) = M$, and $f(\overline{\theta}_{(c_i,c_j)}, r) > M$ for any $r > 1$.*
2. *If i and j properly overlap when i is centred at some position diagonally adjacent to j then $f(\overline{\theta}_{(c_i,c_j)}, \sqrt{2}) = M$, and $f(\overline{\theta}_{(c_i,c_j)}, r) > M$ for either $r = 1$ or $r = 2$.*
3. *If i and j properly overlap when i is centred at some position peripherally adjacent to j then $f(\overline{\theta}_{(c_i,c_j)}, 2) = M$, and $f(\overline{\theta}_{(c_i,c_j)}, r) > M$ for $r < 2$.*
4. *Otherwise $f(\overline{\theta}_{(c_i,c_j)}, r) > M$ for any distance r.*

Proof. Recall that all functions in $\mathcal{CMV}(2)$ must have some vector of parameters $\overline{\theta}_{(c_i,c_j,d)} \in \mathbb{R}^p$ for every $d \in \{1, \sqrt{2}, 2\}$ such that $f(\overline{\theta}_{(c_i,c_j,d)}, d) = M$, and for every other distance distance $d' \in d(r)$ where $d' \neq d$ the value of the energy function $f(\overline{\theta}_{(c_i,c_j,d)}, d') > M$ by the optimal pairwise distance property (Property 3) of Definition 3. Therefore Conditions 1, 2, and 3 in the statement can be satisfied by choosing the appropriate vector of parameters for the distances of $1, \sqrt{2}$ and 2 respectively. Further, by the separation property (Property 2) there exists some vector of parameters $\overline{\theta}_{(c_i,c_j)} \in \mathbb{R}^p$ such that for every distance $d \in d(r)$ the energy $f(\overline{\theta}_{(c_i,c_j)}, d) > M$, satisfying Condition 4 above. ◀

Setting the parameter vectors so as to satisfy the conditions given in Proposition 12, Lemma 13 shows that a valid assignment of $\mathcal{R}$ to $\mathbb{Z}^2$ can be used to construct a valid colouring of $\mathbb{Z}^2$ using $\mathcal{C}(\mathcal{R})$. Lemma 14 shows that given such a colouring of $\mathbb{Z}^2$ using $\mathcal{C}(\mathcal{R})$, there must exist a valid assignment from $\mathbb{Z}^2$ to $\mathcal{R}$.

▶ **Lemma 13.** *Let A be a valid assignment of the set of distinctly coloured 2-radius rhombuses $\mathcal{R}$ to $\mathbb{Z}^2$ with a period of $n_1 \times n_2$. Given such an assignment there exists a periodic colouring of $\mathbb{Z}^2$ using the set of colours $\mathcal{C}(\mathcal{R})$ with an average energy per vertex of $12 \cdot M$.*

Proof Sketch. Observe that following Proposition 12 the interaction between any pair of colours within a distance of 2 is M. As A is a valid assignment of Rhombuses, the interaction between each point within a distance of at most 2 is M, giving an average energy of $12 \cdot M$. ◀

▶ **Lemma 14.** *Let U be a unit cell of size $n_1 \times n_2$ colouring $\mathbb{Z}^2$ with the colour set $\mathcal{C}(\mathcal{R})$ such that the average energy per vertex is $12 \cdot M$. Given such a unit cell there exists a valid assignment of the set of distinctly coloured 2-radius rhombuses $\mathcal{R}$ to $\mathbb{Z}^2$.*

Proof Sketch. The key observation behind this lemma is that given some vertex v at position (x, y) coloured with $c \in \mathcal{C}(\mathcal{R})$, there are exactly 4 colours $c_1, c_2, c_3, c_4 \in \mathcal{C}(\mathcal{R})$ that can be used to colour the vertices directly adjacent to v. Further, for the colour c_1 there exists exactly 1 colour that can be at a distance of 2 from a vertex coloured c_1 and a distance of 1 from a vertex coloured c. Therefore the local neighbourhood of each vertex must be coloured in such a way that the corresponding rhombuses correspond to a correct assignment. By extension, a valid colouring for the graph with an average energy per vertex of $12 \cdot M$ must correspond to a correct assignment of rhombuses to the plane $\mathbb{Z}^2$. ◀

▶ **Theorem 15.** *The average pairwise energy minimisation problem on $\mathbb{Z}^d$ is undecidable for any function in the 2-distance common minimal value class, and $d \in \{2, 3\}$.*

Proof. From Lemmas 13 and 14, there exists a valid colouring of $\mathbb{Z}^2$ with an average energy per vertex of $12M$ of $\mathcal{C}(\mathcal{R})$ if and only if there exists a valid assignment of $\mathcal{R}$ to $\mathbb{Z}^2$. As the r-discretised rhombus all-distinct periodic complete assignment problem is undecidable, the average pairwise energy minimisation problem on $\mathbb{Z}^2$ is undecidable.

To show the undecidability in 3D, consider the 3D tiling problem, where each tile is a 3D block with each face coloured. This problem is shown to be undecidable by reduction from the tiling problem. Let each block have a top and bottom face, along with the $North, West, South$ and $East$ faces. Given a set of tiles $\mathcal{T}$ a block b is constructed for each tile $t \in \mathcal{T}$ such that the colour of $North, West, South$ and $East$ faces of b match the corresponding colours of t, and the top and bottom faces of b are coloured with some universal colour c. See Figure 5 for an example. Observe that each plane of any valid tiling of these blocks on $\mathbb{Z}^3$ corresponds to a valid tiling of $\mathcal{T}$. As in the 2D setting, this problem can be restricted with the k-unique radius property. Similarly, k-unique radius tilings with 3D blocks can be converted into an all-distinct discretised rhombohedron in the same manner as the 2D case.

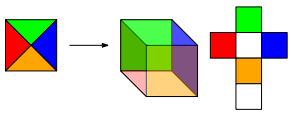

Figure 5 The transformation from a tile (left) to a 3 dimensional block (middle) and to an unfolded representation (right). The top and bottom faces are coloured with the same new colour.

An average pairwise energy minimisation problem on $\mathbb{Z}^d$ instance is constructed from these rhombohedrons in the same manner as in the 2D case. Note that an all discretised rhombohedron containing all points within a distance of 2 has 33 blocks. In this case, the average energy per vertex is $32 \cdot M$ if and only if there exists a valid tiling of $\mathbb{Z}^3$ of the original set of blocks. In one direction, if there exists such a tiling then the corresponding unit cell has an average energy per vertex of $32 \cdot M$. In the other direction, the same arguments as in the 2D case may be applied to show that any colouring with an average energy per vertex of $32 \cdot M$ corresponds to a valid tiling. ◀

Observe that the number of possible solution to the average pairwise energy minimisation problem on $\mathbb{Z}^2$ with a fixed period is at most $q^{n_1 \cdot n_2}$, where q is the number of tiles and (n_1, n_2) the size of the unit cell. Therefore, this problem is in NP. In the other direction the same arguments from Theorem 15 alongside Corollary 11 show the problem to be NP-hard.

▶ **Corollary 16.** *The fixed period average pairwise energy minimisation problem on $\mathbb{Z}^d$ is NP-complete for any function in the 2-distance common minimal value class, and $d \in \{2, 3\}$.*

4 Physically motivated pairwise energy functions

In this section, we apply the results from our abstract model to the problem of crystal structure prediction. In order to do so, we claim that the the *Buckingham-Coulomb* [32] and *Ising* [8] energy functions belong to $\mathcal{CMV}(2)$. The main results in this section, focus on *charge neutrality* in the context of the Buckingham-Coulomb potential, showing that the average pairwise energy minimisation problem on $\mathbb{Z}^d$ remains undecidable even with this restriction, and that the average pairwise energy minimisation problem on $\mathbb{Z}^d$ with a fixed period becomes hard to approximate within any positive factor in the 1D case. These results are strengthened in Section 4.2 by providing a parameterised algorithm for the average pairwise energy minimisation problem on $\mathbb{Z}$.

First, we outline the properties used by the Buckingham-Coulomb and 2-radius n-vector Ising energy functions to show that they belong to the class $\mathcal{CMV}(2)$.

The Buckingham-Coulomb Potential. The Buckingham-Coulomb energy between a pair of vertices coloured i and j at a distance of $r_{i,j}$ is given by the equation $BC(i, j, r_{i,k}) = \frac{A_{i,j}}{e^{B_{i,j} \cdot r_{i,j}}} - \frac{C_{i,j}}{r_{i,j}^6} + \frac{q_i \cdot q_j}{r_{i,j}}$ where $A_{i,j}, B_{i,j}$ and $C_{i,j}$ are a set of force field parameters, determined by the colours, q_i is the charge of colour i, and q_j is the charge of colour j. Here, we assume that rather than the integer grid, the Buckingham-Coulomb potential is performed on the set of points $\{(10 \cdot x_1, 10 \cdot x_2, \ldots, 10 \cdot x_d) \mid (x_1, x_2, \ldots, x_d) \in \mathbb{Z}^d\}$. In order to show that The Buckingham-Coulomb potential belongs to the class $\mathcal{CMV}(2)$, it is necessary to show that there exists a vector of parameters for each distance $d \in [10, \sqrt{200}, 20]$ such that (1) $BC(i, j, d) = M$, (2) $BC(i, j, d') > M$ for every $d' \in [10, \sqrt{200}, 20]$ where $d' \neq d$ and (3) there exists a vector of parameters such that $BC(i, j, d') > M$ for every $d' \in [10, \sqrt{200}, 20]$. Here $M = -1$ and the cutoff distance is set to 2. Conditions (1) and (2) are satisfied by using the faster convergence of the term $\frac{A_{i,j}}{e^{B_{i,j} \cdot r_{i,j}}}$ to 0 than the term $\frac{C_{i,j}}{r_{i,j}^6}$. Condition (3) can be satisfied by setting $A_{i,j}$ to a sufficiently large value, while setting $C_{i,j}$ to 0.

The 2-Radius n-Vector Ising Model. The second energy function we look at is a generalisation of the n-vector Ising model [31]. In the n-vector Ising model, each colour $c \in \mathcal{C}$ corresponds to a unit vector $\vec{c}$. Given a pair of adjacent vertices v and u, coloured c_v and c_u respectively, the energy between v and u is given by the dot product of the vectors, $\vec{c}_v \cdot \vec{c}_u$. In the 2-radius n-vector Ising model, each colour corresponds to a triple of n-length unit vectors. For notation, let $c[i]$ be the i^{th} vector in the triple corresponding to colour c. Given a pair of vertices v and u, coloured c_v and c_u respectively, the energy between v and u is given by $c_v[i] \cdot c_u[i]$ where i is 1 if v and u are at a distance of 1, 2 if v and u are at a distance of $\sqrt{2}$ or 3 if v and u are at a distance of 2. The value of each vector is chosen such that the product of any pair of vertices at distance d is either M, where M is some minimum value, or 0.

▷ Claim 17. The **Buckingham-Coulomb potential** and **2-radius n-vector Ising model** belong to $\mathcal{CMV}(2)$.

4.1 Charge Neutrality

In this section we focus on charge neutrality constraint. Recall that the charge of each colour, denoted $Q(c)$ is an integer value, and that a unit cell is charge neutral if $\sum_{\vec{x} \in U} Q(U(\vec{x})) = 0$.

▶ **Corollary 18.** *The Charge-Neutral average pairwise energy minimisation problem on $\mathbb{Z}^d$ with non-zero charges is undecidable for the Buckingham-Coulomb Potential for $d \in \{2, 3\}$.*

Theorem 19 compliments the proof of NP-hardness from Corollary 16 by showing that the charge-neutral fixed period pairwise energy minimisation problem is NP hard both to solve and to approximate within any constant factor for the Buckingham-Coulomb potential in 1D.

▶ **Theorem 19.** *The charge-neutral fixed period pairwise energy minimisation problem in 1D for the Buckingham-Coulomb potential can not be approximated within any constant factor in polynomial time unless $P = NP$.*

Proof Sketch. This theorem is proven via a reduction from the k-independent set problem. The high level idea is to construct a colour for each vertex in the input graph with a positive charge of $+1$, and a single negative ion of charge $-k$. The energy function is determined such that the interaction between any two colours representing adjacent vertices in the input graph is arbitrarily high, while the pairwise interaction between the positive and negative ions set to -1. This ensures that a valid solution to the charge-neutral fixed period pairwise energy minimisation problem can only be found if there exists an independent set. ◀

4.2 A Parameterised Algorithm for the 1D Setting

In this section we compliment the hardness results by providing a parameterised algorithm for solving the average pairwise energy minimisation problem on $\mathbb{Z}^d$ in 1D. Our algorithm provides solution in $O(n^3 \cdot q^{3 \cdot d})$ where n is the size of the unit cell, q is the number colours, and d is the cut off distance.

Construction. Given an instance of the fixed period pairwise energy minimisation problem for the 1D grid with length n, and set of colours $\mathcal{C}$, a graph G is constructed. Let $\mathcal{V}(d, \mathcal{C}) = \{(x_1, x_2, \ldots, x_{d+1}) \mid x_1, x_2, \ldots, x_{d+1} \in \mathcal{C}\}$. For notation given $l \in \mathcal{V}(d, \mathcal{C})$, l_i is used to denote the colour of the i^{th} position of l, i.e. given $l = (1, 2, 1, 2)$, $l_2 = 2$ while $l_3 = 1$. For every $i \in [n]$ and $l \in \mathcal{V}(d, \mathcal{C})$ the vertex $v_{i,l}$ is constructed and added to the set V of vertices. Given a pair vertices $v_{i,l}, v_{j,k} \in V$, the edge $v_{i,l}, v_{j,k}$ is added to the set of edges E if and only if $i + 1 = j$ and $l_2, l_3, \ldots, l_{d+1} = k_1, k_2, \ldots, k_d$. The weight of $(v_{i,l}, v_{j,k})$, denoted $w(v_{i,l}, v_{j,k})$, equals to $\sum_{i=1}^{d} f(i, \overline{\theta}_{(k_1, k_{i+1})})$. This means that each edge $(v_{i,l}, v_{j,k})$ corresponds to the pairwise interaction energy between k_1 and each subsequent vertex in k. In order to account for the energy from the first vector, an additional set of q^{d+1} vertices labelled v_l for every $l \in \mathcal{V}(d, \mathcal{C})$. The vertex v_l has only a single edge connecting it to $v_{1,l}$, weighted as before. Hence by constructing a path of length n starting at some vertex v_l and ending at the vertex $v_{n,l}$ the weight of the path with correspond to the total pairwise energy of the corresponding unit cell. Thus by finding such a path with minimum energy the solution to the fixed period pairwise energy minimisation problem may be found. Using the above construction, the solution to the fixed period pairwise energy minimisation problem instance is found by determining the shortest path from each vertex of the form v_l to the vertex $v_{n,l}$ for every $l \in \mathcal{V}(d, \mathcal{C})$. Note that this graph can be constructed in $O(((n+1) \cdot q^{d+1})^2)$ time for any energy function that can be computed in constant time, by simply constructing the full set of $(n + 1) \cdot q^{d+1}$ vertices (corresponding to each position in the grid and list of d colours), and computing the energy between them using the energy function.

▶ **Theorem 20.** *There exists an algorithm to solve the fixed period pairwise energy minimisation problem in $O(n^3 \cdot q^{3 \cdot (d+1)})$ time for any function in $\mathcal{CMV}(d)$.*

Proof. Using the construction above, the solution to the corresponding fixed period pairwise energy minimisation problem instance can be found by using an efficient algorithm for solving the all pairs shortest path problem. Note the graph can be constructed in $O(((n+1)\cdot q^{d+1})^2) \approx O(n^2 \cdot q^{2(d+1)})$ time, assuming that the energy function can be evaluated in constant time. Using the Floyd–Warshall algorithm, the paths may be found in $O(|V|^3)$ time. Note that the number of vertices equals $(n+1)\cdot q^{d+1}$ giving a total complexity of $O\left((n+1)^3 \cdot q^{3\cdot(d+1)}\right) \approx O\left(n^3 \cdot q^{3\cdot(d+1)}\right)$. ◀

References

1. A. Adamatzky. On Diversity of Configurations Generated by Excitable Cellular Automata with Dynamical Excitation Intervals. *International Journal of Modern Physics C*, 23(12):1250085, November 2012. doi:10.1142/S0129183112500854.
2. D. Adamson. Ranking binary unlabelled necklaces in polynomial time. *24th International Conference on Descriptional Complexity of Formal systems (DFCS), Lecture Notes in Computer Science, Springer*, 2022.
3. D. Adamson, A. Deligkas, V. V. Gusev, and I. Potapov. On the hardness of energy minimisation for crystal structure prediction. *Fundamenta Informaticae*, 184:1–23, February 2022.
4. Duncan Adamson, Vladimir V. Gusev, Igor Potapov, and Argyrios Deligkas. Ranking Bracelets in Polynomial Time. In Paweł Gawrychowski and Tatiana Starikovskaya, editors, *32nd Annual Symposium on Combinatorial Pattern Matching (CPM 2021)*, volume 191 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 4:1–4:17, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CPM.2021.4.
5. C. Allauzen and B. Durand. Tiling problems. In The Classical Decision Problems*, Perspectives in Mathematical Logic*, pages 407–420. Springer, 2001.
6. D. Antypov, A. Deligkas, V.V. Gusev, M. J. Rosseinsky, P. G. Spirakis, and M. Theofilatos. Crystal Structure Prediction via Oblivious Local Search. In *SEA 2020*, volume 160 of *LIPIcs*, pages 21:1–21:14, 2020.
7. R. Berger. *The undecidability of the domino problem.* Number 66 in memoirs of the american mathematical society. American Mathematical Soc., 1966.
8. A. Blanca, R. Gheissari, and E. Vigoda. Random-Cluster Dynamics in Z^2: Rapid Mixing with General Boundary Conditions. In Dimitris Achlioptas and László A. Végh, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2019, September 20-22, 2019, Massachusetts Institute of Technology, Cambridge, MA, USA*, volume 145 of *LIPIcs*, pages 67:1–67:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.APPROX-RANDOM.2019.67.
9. S. T. Call, D. Y. Zubarev, and A. I. Boldyrev. Global minimum structure searches via particle swarm optimization. *Journal of computational chemistry*, 28(7):1177–1186, 2007.
10. R. Catlow and S. M. Woodley. Crystal structure prediction from first principles. *Nature materials*, 7(12):937, 2008.
11. B. A. Cipra. The Ising model is NP-complete. *SIAM News*, 33(6):1–3, 2000.
12. C. Collins, M.S. Dyer, M.J. Pitcher, G.F.S. Whitehead, M. Zanella, P. Mandal, J.B. Claridge, G.R. Darling, and M.J. Rosseinsky. Accelerated discovery of two crystal structure types in a complex inorganic phase field. *Nature*, 546(7657):280, 2017.
13. V. K de Souza and D. J Wales. The potential energy landscape for crystallisation of a Lennard-Jones fluid. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(7):074001, July 2016. doi:10.1088/1742-5468/2016/07/074001.
14. D. M. Deaven and K. M. Ho. Molecular geometry optimization with a genetic algorithm. *Physical review letters*, 75(2):288, 1995. Added up to here.
15. M. S. Dyer, C. Collins, D. Hodgeman, P. A. Chater, A. Demont, S. Romani, R. Sayers, M. F. Thomas, J. B. Claridge, G. R. Darling, and M. J. Rosseinsky. Computationally assisted identification of functional inorganic materials. *Science*, 340(6134):847–852, 2013.

16 S. Goedecker. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of chemical physics*, 120(21):9911–9917, 2004.

17 L. A. Goldberg and H. Guo. The Complexity of Approximating complex-valued Ising and Tutte partition functions. *Comput. Complex.*, 26(4):765–833, 2017. doi:10.1007/s00037-017-0162-2.

18 L. A. Goldberg and M. Jerrum. Approximating Pairwise Correlations in the Ising Model. *ACM Trans. Comput. Theory*, 11(4):23:1–23:20, 2019. doi:10.1145/3337785.

19 M. Hill, S. Stepney, and F. Wan. Penrose Life: Ash and Oscillators. In Mathieu S. Capcarrère, Alex Alves Freitas, Peter J. Bentley, Colin G. Johnson, and Jon Timmis, editors, *Advances in Artificial Life, 8th European Conference, ECAL 2005, Canterbury, UK, September 5-9, 2005, Proceedings*, volume 3630 of *Lecture Notes in Computer Science*, pages 471–480. Springer, 2005. doi:10.1007/11553090_48.

20 J. Kari and E. Moutot. Decidability and Periodicity of Low Complexity Tilings. In C. Paul and M. Bläser, editors, *37th International Symposium on Theoretical Aspects of Computer Science, STACS 2020, March 10-13, 2020, Montpellier, France*, volume 154 of *LIPIcs*, pages 14:1–14:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.STACS.2020.14.

21 H. R. Lewis. Complexity of solvable cases of the decision problem for the predicate calculus. In *19th Annual Symposium on Foundations of Computer Science (FOCS 1978)*, pages 35–47, 1978. doi:10.1109/SFCS.1978.9.

22 D. C. Lonie and E. Zurek. XtalOpt: An open-source evolutionary algorithm for crystal structure prediction. *Computer Physics Communications*, 182(2):372–387, 2011.

23 A. O. Lyakhov, A. R. Oganov, and M. Valle. How to predict very large and complex crystal structures. *Computer Physics Communications*, 181(9):1623–1632, 2010.

24 J. Maddox. Crystals from first principles. *Nature*, 335(6187):201–201, 1988. doi:10.1038/335201a0.

25 A. R. Oganov and C. W. Glass. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *The Journal of chemical physics*, 124(24), 2006.

26 J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal, and V. Caignaert. Prediction of crystal structures from crystal chemistry rules by simulated annealing. *Nature*, 346(6282):343–345, 1990.

27 R.J. Pickard, C. J .and Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

28 Z. Rotman and E. Eisenberg. Finite-temperature liquid-quasicrystal transition in a lattice model. *Phys. Rev. E*, 83:011123, January 2011. doi:10.1103/PhysRevE.83.011123.

29 M. U. Schmidt and U. Englert. Prediction of crystal structures. *Journal of the Chemical Society, Dalton Transactions*, 10:2077–2082, 1996.

30 J. C. Schön and M. Jansen. First step towards planning of syntheses in solid-state chemistry: determination of promising structure candidates by global optimization. *Angewandte Chemie International Edition in English*, 35(12):1286–1304, 1996.

31 H. E. Stanley. Dependence of Critical Properties on Dimensionality of Spins. *Phys. Rev. Lett.*, 20:589–592, March 1968.

32 D. J. Wales. Exploring Energy Landscapes. *Annual Review of Physical Chemistry*, 69(1):401–425, 2018. PMID: 29677468. doi:10.1146/annurev-physchem-050317-021219.

33 Y. Wang, J. Lv, L. Zhu, S. Lu, K. Yin, Q. Li, H. Wang, L. Zhang, and Y. ma. Materials discovery via CALYPSO methodology. *Journal of physics. Condensed matter : an Institute of Physics journal*, 27:203203, April 2015.

34 L. T. Wille and J. Vennik. Computational complexity of the ground-state determination of atomic clusters. *Journal of Physics A: Mathematical and General*, 18(8):L419–L422, June 1985.