UNIVERSITY OF LIVERPOOL

LEVERHULME TRUST

The Leverhulme Research Centre for Functional Materials Design

Crystal Structure Prediction by Vertex Removal in Euclidean Space

Duncan Adamson
University of Liverpool, Department of Computer Science

April 3, 2019

CCDC diamond Hartree Centre Science & Technology Facilities Council Imperial College London ICNS King Abdullah University of Science and Technology University of Southampton

What is Crystal Structure Prediction?

Problem	*Crystal Structure Prediction* (CSP)
Input:	A set of ions, A, an area of space, C.
Output:	A structure, S, made by placing some copies of the ions in A in C, with a *neutral* charge *minimising* potential energy between the ions.

What is a *Crystal*?

- We consider crystals to be made up of unit cells.
- Each unit cell is the smallest repeating region of space within the crystal.

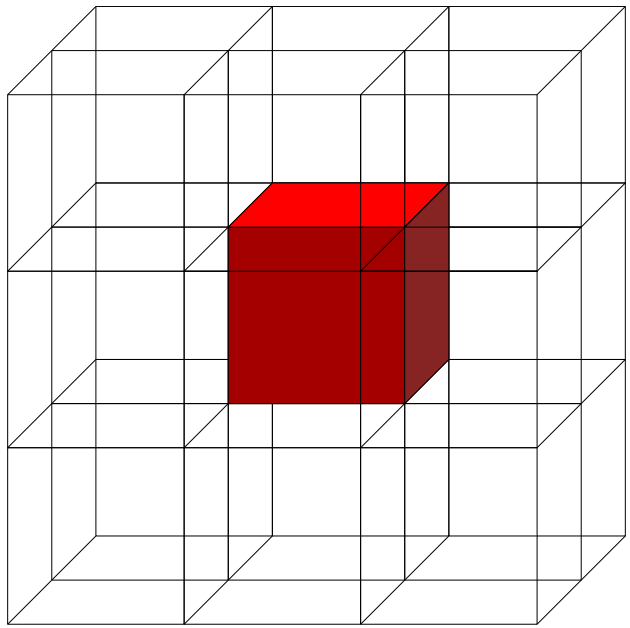

Figure 1: Unit cell highlighted in red, note any other box would be equivalent

What is a *Unit Cell*?

- Each unit cell is a collection of *Ions*.
- We assume each unit cell is *independent* of all other unit cells.
    - This means that we only consider the *interaction* of ions within the same cell.
- Every cell must have a total *Neutral* charge.

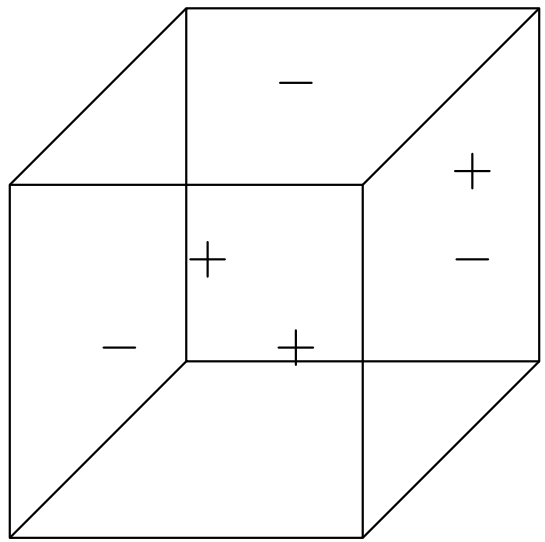

Figure 2: Example of a unit cell

What is an *Ion*?

- A point in space belonging to a *species*.
- The species determines its interaction with other ions, as well as its charge.
- We denote the charge of ion, i, q_i
- for a unit cell with the set of ions, S, we require the following to be satisfied:

$$\sum_{i \in S} q_i = 0$$

How do we determine *interaction*

- We define the pairwise interaction for any pair of ions by some parameterised function $U_{\theta_{i,j}}(r_{ij})$.
- r_{ij} is the distance between ions i and j
- The parameters for this function are determined by the species of ions i and j.
- A negative value for interaction means that the ions are trying to move closer together, which implies the crystal will be stronger.
- We can use this to represent the ions as a graph with weighted edges.

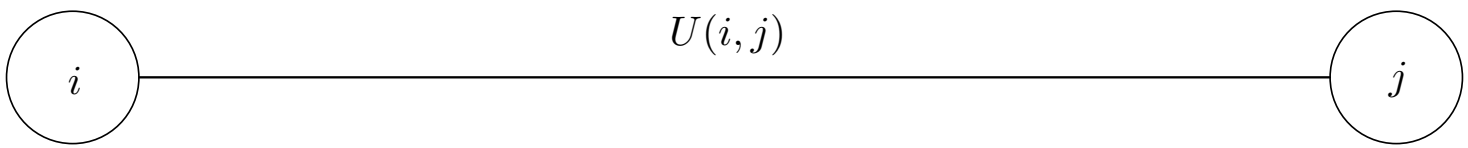

Figure 3: Interaction between two ions represented as a graph, each ion represents a vertex.

Representing the unit cell as a graph

- We can use this to represent the unit cell as a graph embedded into 3-dimensional space.
- Conversely, we can use this to create a complete graph where:
  - each ion is a vertex.
  - each edge has a weight equal to the interaction between the two ions it is connected to.

Representing the unit cell as a graph

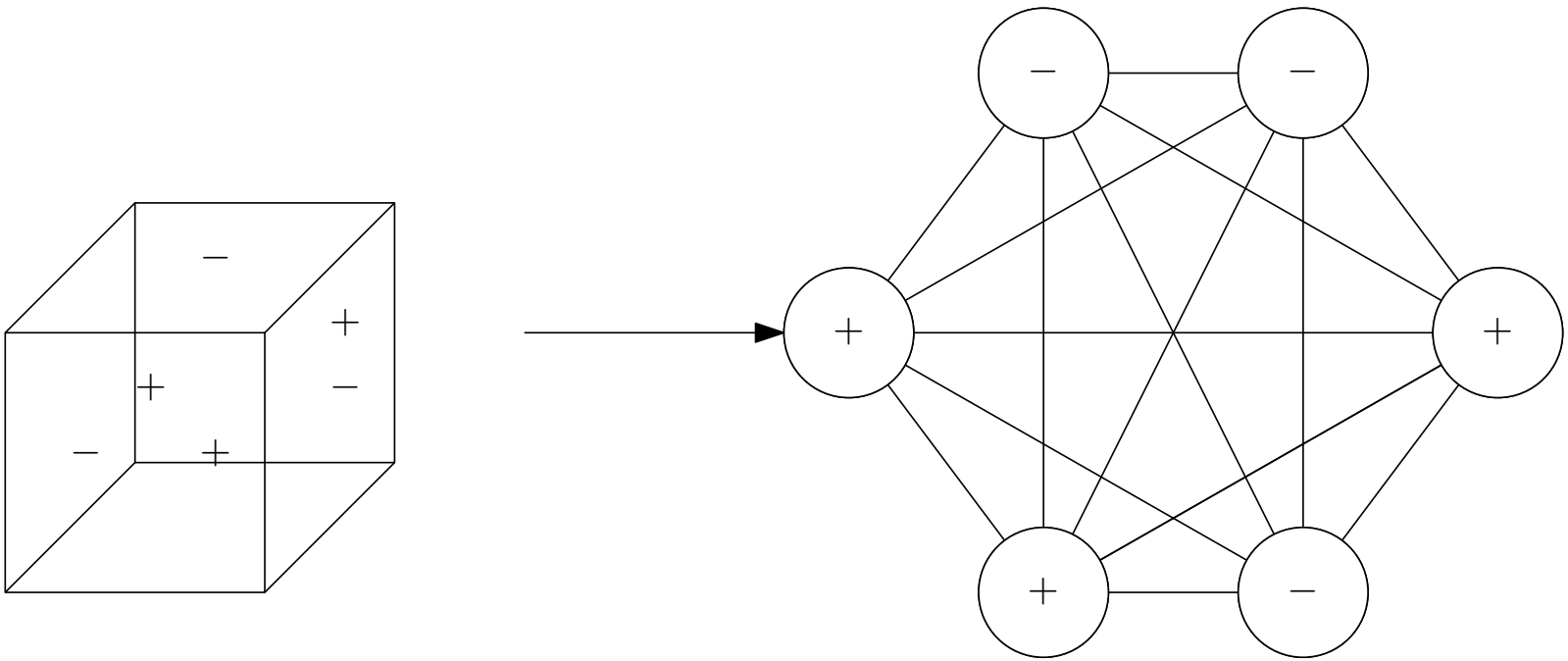

Figure 4: We can redraw our unit cell as a graph, using the ions as vertices and the interactions as weights on the edges.

Some Notation and Terminology

- A *Structure* refers to a set containing all the ions within a given unit cell.
- Give a structure, S, we use S^+ to denote the set of ions with a positive charge, and S^- for the ions with a negative charge.
- We use $|S^+|$ to denote the sum of the magnitude of the positive charges, $|S^+| = \sum_{i \in S^+} q_i$
- We use $|S^-|$ to denote the sum of the magnitude of the negative charges, $|S^-| = \sum_{i \in S^-} -(q_i)$.

Crystal Structure Prediction by Vertex Removal

- We can use this to define our problem: Crystal Structure Prediction by K-Ion (Vertex) Removal.
- We take as input some highly dense initial structure, $S = S^+ \cup S^-$, within our unit cell, and an integer, k.
- Our goal is to remove some substructure of S, $S = -S'^+ \cup S'^-$, such that:

$$|S'^+| \leq |S^+| - k$$
$$|S'^-| \leq |S^-| - k$$
$$|S'^+| = |S'^-|$$

- We also want our solution to be *minimal*, in that there is no substructure, $S'' \subset S'$, that also satisfies the above.

Crystal Structure Prediction by Vertex Removal

Problem	*K-Vertex Removal* (KIR).
Input	A structure of ions, S, a pairwise energy function, U, and an integer k.
Output	A substructure, $S' \subseteq S$, formed by a minimal removal of k charges from S with minimal total energy with respect to U

Decision problem

- For NP-Completeness, we need to reformulate this as a decision problem.
- We do this by adding a goal energy, g, which is the maximum allowed energy.
- Given an instance of KIR, we report yes if there is a substructure with total energy less than or equal to g, or no otherwise.

The Energy function

- The given energy function determines what will and won't be a good solution.
- We will be considering a general class of functions for which this problem is NP and APX complete, which we call the *Crystalline* class of functions, $\mathcal{F}$.

$$\forall f \in \mathcal{F}, \exists a, b \in \mathbb{R}, a > b \text{ s.t.}$$
$$\forall r \in \mathbb{R}^+ \exists \theta_{\mathbf{ar}}, \theta_{\mathbf{br}} \in \mathbb{R}^n \text{ s.t.} f_{\theta_{ar}}(r) = a, f_{\theta_{br}} = b$$

Buckingham-Coulomb potential energy

- The *Buckingham-Coulomb* potential is used frequently in computational chemistry for determining the energy between ions.
- In this function we use the charge of the ions, as well as 3 *force field* parameters, determined by the species of the ions, as our parameters.
- These are A_{ij}, B_{ij} and C_{ij}.
- The energy function is:

$$U_{\{A_{ij},B_{ij},C_{ij},q_i,q_j\}}(r_{ij}) = \frac{A_{ij}}{e^{B_{ij}r_{ij}}} - \frac{C_{ij}}{r_{ij}^6} + \frac{q_i q_j}{r_{ij}}$$

Clique to K-Vertex removal, for $U \in \mathcal{F}$

- Assume we have some instance of the Clique problem, where we have a graph, G, and wish to find a clique of size k.
- We claim that we can reduce this problem to KIR, making the latter NP and APX complete.
- We will do this by constructing a structure such that we will be left with only the ions corresponding to vertices in G in a clique of size k.
- The main idea is to create 2 ions for each vertex in G, labelled with their corresponding vertex.
- We will assume that our energy function is some arbitary function in $\mathcal{F}$

Parameters

- We assign parameters so that the energy between a pair of ions, i, j corresponding to v_i and v_j, is as follows:

$$U_{\theta_{ij}}(r_{ij}) = \begin{cases} b & \text{if } (v_i, v_j) \in E, \text{ or } v_i = v_j \\ a & \text{otherwise} \end{cases}$$

- We know from our definition of $\mathcal{F}$ that we can always achieve this.