# On the Hardness of Energy Minimisation for Crystal Structure Prediction

Duncan Adamson[1,2], Argyrios Deligkas[2], Vladimir V Gusev[2], and Igor Potapov[1]

[1] Univiersity of Liverpool, Department of Computer Science
[2] Leverhulme Research Centre for Functional Materials Design

**Abstract.** Crystal Structure Prediction (CSP) is one of the central and most challenging problems in materials science and computational chemistry. In CSP, the goal is to find a configuration of ions in 3D space that yields the lowest potential energy. Finding an efficient procedure to solve this complex optimisation question is a well known open problem in computational chemistry. Due to the exponentially large search space, the problem has been referred in several materials-science papers as "NP-Hard" without any formal proof. This paper fills a gap in the literature providing the first set of formally proven NP-Hardness results for a variant of CSP with various realistic constraints. In particular, this work focuses on the problem of *removal*: the goal is to find a substructure with minimal energy, by removing a subset of the ions from a given initial structure. The main contributions are NP-Hardness results for the CSP *removal* problem, new embeddings of combinatorial graph problems into geometrical settings, and a more systematic exploration of the energy function to reveal the complexity of CSP. These results contribute to the wider context of the analysis of computational problems for weighted graphs embedded into the 3-dimensional Euclidean space, where our NP-Hardness results holds for complete graphs with edges which are weighted proportional to the distance between the vertices.

## 1 Introduction

One of the central and most challenging problems in materials science and computational chemistry is the problem of predicting the structure of a crystal given the set of ions composing it [14]. The goal is to find a structure of ions that achieves the lowest energy. This problem, *Crystal Structure Prediction* (CSP), has remained open due to the complexity of solving it optimally [14] and the combinatorial explosion following a brute-force approach. Current approaches to this problem are based on heuristic techniques [9, 12], however they cannot guarantee optimality while remaining computationally demanding.

In generic formulations of CSP there are many degrees of freedom due to the numerous parameters: the number of ions, their positions, and the unique interactions between each type of ion. The search space remains exponential in size even for greatly simplified versions of CSP. Due to this, CSP has, incorrectly,

been referred to in several computational-chemistry papers as "NP-Hard and very challenging" [11]. However the argument that the search must be done in a set of exponential size implies NP-Hardness does not hold.

The two results which are often mentioned in context of the NP-Hardness of CSP are [3] and [13]. In [3], within the context of the Ising model, the authors show NP-Hardness in the model of placing $\pm 1$ charges on a graph with degree at most 6 taking into account only the local interactions between connected vertices. In [13], provides a reduction to TSP, showing the problem belongs to NP however not Hardness.

In this work, several variants of CSP are considered, providing alternative reasons for the hardness of closely related problems, focusing on the problem of *removal*. Inspiration comes from hard combinatorial problems in graph theory and proposes several new embeddings of NP-Hard graph problems into numerical versions of CSP which can be seen as an optimisation problem for weighted geometric graphs with a non-linear objective function. The input is a configuration of the ions, with the goal to remove a subset of the ions such that the interaction energy among the remaining atoms is minimised.The problem of removing vertices of a graph whose deletion results in a subgraph satisfying some specific property have been intensively studied in the combinatorial graph theory. [8] shows that for a large class of properties this problem is NP-Complete, extended in [16] and [15] to further properties showing NP-Completeness for bipartite graphs and for non-trivial hereditary properties.

The *removal* problem can be seen as a variant of combinatorial CSP problem, where the positions of the ions correspond to points in a discrete grid. The idea is to find an optimal structure by placing many copies of the ions used to build a new structure in unrealistic positions in the discrete space. Due to the nature of the energy function, when the goal is to minimise the potential energy, the excess ions must be removed. In this variant of the *removal* problem for which NP-Hardness is shown, the initial configuration (from where the ions are removed) is part of the input and has only vacant positions or positions with a single ions in the discrete three-dimensional-Euclidean space.

**Our contributions.** This work provides the first NP-Hardness results for CSP [7] with realistic constraints, providing new embeddings of combinatorial graph problems in geometrical settings, as well as exploring the energy function in a more systematic way that could reveal the computational complexity of CSP. Moreover, these results can be seen as part of a more general problem of removing vertices from a weighted graph embedded into 3D Euclidean space. Three versions of this problem are considered:

- **$k$-Charge Removal:** Remove exactly $k$ charges minimising the total energy;
- **Minimal At-Least-$k$-Charge Removal:** A generalisation of $k$-charge removal where the removed set is a *minimal* set of at least $k$ charges minimising the total energy;
- **At-Least-$k$-Charge Removal:** A generalisation of min-at-least-$k$-charge removal where the removed set is of least charges but not necessarily minimal, minimising the total energy.

One challenge of the Euclidean graphs considered here is that these graphs are complete, with edges weighted proportional to the distance between the vertices. Many classical NP-Hard problems are much harder to embed into this setting. Even for some existing hardness results, in both the geometric and more restricted Euclidean setting, to bring these problems into a bounded number of dimensions often requires non-trivial technical proofs as dimension often is part of the input [2, 10]. Often these constructions utilise the results on geometric graphs embedded into the plane [5, 6], with many problems in this field open.

This work will be organised as follows: Section 2 provides relevant notation and definitions, Section 3 presents NP-Hardness for the general case of the problems under both energy function in $\mathcal{F}$ and the *Coulomb* (electrostatic) potential. Section 4 restricts the problem to only 2 species under the *Buckingham-Coulomb* (interatomic) potential, and is shown the remain NP-Hard in Theorem 5. The full version of this paper, containing the omitted proofs is available at arXiv [1].

| Theorem | Summary | Setting |
|---|---|---|
| Theorem 1 | NP-Completeness by reduction from the clique problem. | All problems, under any energy function in $\mathcal{F}$, charges of $\pm c$ for a given $c$ and an unbounded number of ion species. |
| Theorem 2 | NP-Completeness by extension of Theorem 1. | All problems, under any energy function in $\mathcal{F}$, any bounded set of charges and an unbounded number of ion species. |
| Theorem 3 | Reduction to max-weight-$k$-clique. | $k$-charge removal or minimal-at-least-$k$-charge removal under any computable energy function, charges of $\pm c$ for a given $c$, and a unbounded number of ion species. |
| Theorem 4 | NP-Completeness by reduction from the knapsack problem. | Minimal-at-least-$k$-charge removal and at-least-$k$-charge removal, under the Coulomb potential energy function, unbounded number of charges and unbounded number of ion species. |
| Theorem 5 | NP-Completeness by reduction from independent set on penny graphs. | All problems, under the Buckingham-Coulomb potential energy function, charges of $\pm 1$, and two species of ion. |

## 2 Notation and Definitions

**Unit Cell.** A *crystal* is a solid material whose ions, are arranged in a highly ordered arrangement, forming a *crystal structure* that extends in all directions. A crystal structure is described by its *unit cell*; a region of $\mathbb{R}^3$ bounded by a parallelepiped representing a period containing ions in a specific arrangement. The unit cells are stacked in $\mathbb{R}^3$ tiling the whole space forming a crystal. The unit cell is a parallelepiped alongside the *arrangement* of ions with their specie. Each unit cell contains a set of $n$ *ions* within the parallelepiped. Each ion, $i$, has a *specie*, e.g. Ti or Sr, and a non-zero charge $q_i$. The specie for an ion $i$

will be denoted $S(i)$. All unit cells are neutrally charged, i.e., $\sum_{1 \leq i \leq n} q_i = 0$. An arrangement defines a position for every ion in the unit cell.

**Energy.** The energy of a crystal is computed by summing the pairwise interactions between all pairs of ions. A positive value for the pairwise interaction means the two ions are repelling, while a negative value means they are attracting.Each pair of species has a unique set of parameters (called *force fields*) which are applied to the common energy function $U$ alongside the Euclidean distance between the ions. In general, energy is defined via series as a crystal is infinite.

In this paper interaction will be restricted to a single unit cell. The primary reason is that the energy between ions in different unit cells quickly converges, making the energy within a single unit cell a good approximation of the total.

Each arrangement has $n$ ions and a corresponding *potential energy* $PE$, calculated with respect to the given energy function $U$. The goal is to minimise the potential energy. Pairwise interaction between two ions $i$ and $j$ with respect to the energy function $U$ is $U(i, j)$, denoted $U_{ij}$ when it is clear from the context. The value of $U_{ij}$ is defined by the force field of the ions and the Euclidean distance between them, which is included as one of the parameters. The total potential energy for an arrangement of $n$ ions is given by $PE = \sum_{1 \leq i,j \leq n, i \neq j} U_{ij}$.

This paper will consider a general class of energy functions, called the *controllable* potential functions, denoted $\mathcal{F}$. All functions in $\mathcal{F}$ are computable in polynomial time for any input. Intuitively, for every $f \in \mathcal{F}$ there exists a set of force field parameters that counteract the distance parameter $r$. Formally, a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ belongs to $\mathcal{F}$ if and only if for any given $a \in \mathbb{R}$ and any fixed $r \in \mathbb{R}^+$ there exists a set $\{x_1 \dots x_{n-1}\} \in \mathbb{R}^{n-1}$ such that $f(x_1, \dots, x_{n-1}, r\} = a$.

The most popular function for CSP, which will be focused on in this paper, is the *Buckingham-Coulomb* potential [4], which is the sum of the Buckingham and Coulomb potentials. The Coulomb potential for a pair of ions $i, j$ is $U^C_{ij} = \frac{q_i q_j}{r_{ij}}$, where $r_{ij}$ is the Euclidean distance between the ions. The Buckingham potential for a pair of ions $i, j$, $U^B_{ij}$, is defined by four parameters. These are the distance and the three force field parameters, $A_{S(i),S(j)}$, $B_{S(i),S(j)}$, $C_{S(i),S(j)}$, which are dependent on the specie of the ions. It should be noted that all three parameters are positive values. The energy is calculated as $U^B_{ij} = \frac{A_{S(i),S(j)}}{e^{B_{S(i),S(j)} r_{ij}}} - \frac{C_{S(i),S(j)}}{r^6_{ij}}$. Therefore the Buckingham-Coulomb potential is given by:

$$U^{BC}_{ij} = U^B_{ij} + U^C_{ij} = \frac{A_{S(i),S(j)}}{e^{B_{S(i),S(j)} r_{ij}}} - \frac{C_{S(i),S(j)}}{r^6_{ij}} + \frac{q_i q_j}{r_{ij}}.$$

**Proposition 1.** *There exists a set of parameters for the Buckingham-Coulomb function such that it is in $\mathcal{F}$.*

**Crystals as geometric graphs.** Using the above definitions, it can be shown how crystals may be viewed as geometric graphs. Recall that each ion corresponds to a charged point in $\mathbb{R}^3$. Each ion is represented with a weighted vertex, also placed into $\mathbb{R}^3$ at the same position as the ion, giving a total of $n$ vertices. The vertex corresponding to the ion $i$, denoted $v_i$, is assigned a weight of $q_i$.

| Restriction | Summary |
|---|---|
| $k$-Charges | A *neutral* set of charges $R$ where $\left\lvert \sum_{v_i \in R^+} wt(v_i) \right\rvert = k$ |
| At-Least-$k$-charges | A neutral set of charges $R$ where $R \subseteq V$ and $\sum_{v_i \in R^+} wt(v_i) \geq k$. |
| Minimal-At-Least-$k$-charges | A *minimal* set of at-least-$k$-charges $R$ - where minimal means that there does not exist any neutral subset $R' \subset R$ where $R'$ is also a set of at least $k$-charges |

**Table 1.** Summary of restrictions for the charge removal problem.

$wt(v_i)$ will denote the weight of a given vertex $v_i$, i.e. $wt(v_i) = q_i$. For notation, $V^+$ will denote the set of vertices with a positive weight in $V$, and $V^-$ for the set of vertices with a negative weight in $V$. Between each pair of vertices there is an edge, weighted by the pairwise interaction of the corresponding ions $U_{ij}$. Note that $U_{ij}$ will be determined by the length of the edge, which will be a straight line in the space. The energy of a crystal graph $G = \{V, E\}$ can be computed as $PE = \sum_{\{v_i, v_j\} \in E} U_{ij}$. Geometric graphs created from a unit cell will be referred to as *crystal graphs*.

**The *Charge removal* Problem.** The Charge removal problem takes as input a crystal graph $G$ corresponding to a "dense" initial arrangement of ions, with the goal of removing some subset of vertices $X$. In the most general case this may be any subset, provided the final graph is *charge neutral*, meaning it satisfies $\sum_{v_i \in R} wt(v_i) = 0$. It will be assumed that the initial graph is charge neutral, and therefore that $X$ is also neutral. This work will consider three variants of this problem where there are further conditions on the set, summarised in Table 1. Note that the second of these, *At-Least-k-Charges*, becomes the general case when $k = 0$. This work considers three restrictions on the removed set, which are defined in Table 1. The base version of the problem is stated as:

*Instance:* A crystal graph $G$, with edges weighted by a given common energy function $U$.

*Goal:* The set of charges $R$ satisfying $P$ from $G$ such that $G' = G \setminus R$ created by the removal of $R$ from $G$ which minimises $\sum_{\{v_i, v_j\} \in E'} U_{ij}$.

From this problem, a decision version may be obtained by asking if there exists a removal that leaves $G'$ with no-more total energy than some goal $g$, i.e. $\sum_{v_i, v_j \in V', i \neq j} U_{ij} \leq g$. In the case there is some restriction on the output, there may also be additional input - in all the cases considered here this will be a natural $k$. In the remainder of this work, the problems under the restrictions in Table 1 and will be denoted as follows:

- The $k$-Charge-Removal Problem (K-CHARGE REMOVAL).
- The At-Least-$k$-Charge-Removal Problem (AT-LEAST-K-CHARGE REMOVAL).
- The Minimal At-Least-$k$-Charge Removal Problem (MINIMAL-AT-LEAST-K-CHARGE REMOVAL).

**Proposition 2.** *A solution to* K-CHARGE REMOVAL *or* AT-LEAST-K-CHARGE REMOVAL *can be verified in polynomial time.*

**Proposition 3.** *A set of $k$-charges may be verified as minimal in polynomial time if and only if the set of allowed values for charges is polynomially bounded.*

Proposition 2 follows from noting that for a given graph with precomputed weights for the edges, the requisite edges and vertices may be summed to verify that it is either a set of $k$ or of at-least-$k$ charges, and that the energy is bellow the required bound in the decision case. Proposition 3 is shown by reduction from the subset sum problem to the problem of verifying if the set is *minimal*, as defined in Table 1.

## 3 NP-Hardness for an unbounded number of ion species

This section will focus on results for an unbounded number of ion species. Theorems 1, 2 and 4 will show NP-Hardness for various settings via a series of reductions under the general class of potential function in the case of Theorems 2 and 2, and under the Coulomb energy in Theorem 4. Theorem 3 will show a novel way of encoding the problem into the well studied max-weight clique problem. While these results will apply to all restrictions, it should be noted that in the case the charges are not bounded, although MINIMAL-AT-LEAST-K-CHARGE REMOVAL will remain NP-Hard it will not be in NP.

**Theorem 1.** K-CHARGE REMOVAL, MINIMAL-AT-LEAST-K-CHARGE REMOVAL *and* AT-LEAST-K-CHARGE REMOVAL *are NP-Complete for energy functions in $\mathcal{F}$ for charges of $\pm c$, for any natural number $c$.*

**Theorem 2.** K-CHARGE REMOVAL *remains NP-Hard for set of allowed charges with unique magnitude and an energy function within $\mathcal{F}$.*

**Theorem 3.** K-CHARGE REMOVAL *can be reduced to* MAX-WEIGHT K-CLIQUE *in polynomial time, under the restriction that charges are limited $\pm c$ and the energy function is computable within polynomial time.*

Theorems 1 and 2 come by a reduction from the Max-Clique problem. Theorem 1 provides a construction for the decision version of the charge removal problem from an instance of Max-Clique using constant charges such that under any of the restrictions on the removed vertices a solution the the charge removal instance will imply a solution to the Max-Clique problem. This is extended in Theorem 2, where it is shown that this construction may be extended with a set of *dummy* vertices, the removal of which may be done at no cost while maintaining the total set a charge neutral. Theorem 3 provides a novel encoding of the charge removal problem into the well known maximum weight clique problem.

**Theorem 4.** AT-LEAST-K-CHARGE REMOVAL *remains NP-Hard when the energy function is limited to the Coulomb potential.*

Theorem 4 compliments Proposition 3 by showing that, even in the case the removal does not have to be verified as minimal, the complexity of finding a solution may still be NP-Hard for the Coulomb potential function.

## 4 NP-Hardness for a bounded number of species

In Section 3 NP-Hardness was shown for the case that there was an unbounded number of species, and NP-completeness in the case that there is a bounded number of charges. This will now be strengthened by considering instances with only two unique species. Only the Buckingham-Coulomb potential function with charges of $\pm 1$ will be considered in this section. All three problems will again be considered, noting that for charges of $\pm 1$ K-CHARGE REMOVAL is equivalent to MINIMAL-AT-LEAST-K-CHARGE REMOVAL. NP-Hardness will be shown by a reduction from INDEPENDENT-SET on penny graphs adapting it to the Euclidean settings of crystal graph of ions within a unit cell. The Independent Set problem, denoted INDEPENDENT-SET, takes as input a graph, $G$, and a natural number $k$. The goal is to find an *independent set*, i.e. a set of vertices such that no two are adjacent, of size $k$ in $G$, or report that one does not exist. Penny graphs are the class of graphs where each vertex may be drawn as a unit circle such that no two circles overlap, and an edge between two vertices exist if and only if the corresponding circles are tangent, i.e. they intersect at only a single point. Finding an independent set on this class of graphs is known to be NP-Hard [5].

**Sketch of the construction of the K-CHARGE REMOVAL instance:** Starting with an instance of INDEPENDENT-SET on a maximum degree 3 planar graph, containing the graph and a natural number $k$ a penny graph, $G$, is created using Theorem 1.2 from Cerioli et al, using a radius of $\frac{n}{2}$ for the pennies. Graphs created in this manner will be denoted *long orthogonal penny graphs*. The K-CHARGE REMOVAL instance is created by placing a positive ion above the centre of each penny, and a negative ion bellow.

**Ion species:** The positive and negative species are assigned charges of magnitude 1. From these species there are parameters for the interaction between two ions of the positive specie, two ions of the negative specie, and between one ion of the positive specie and one of the negative specie. For brevity, 1 and 2 will denote the positive and negative specie respectively. Under this construction, the interaction between the two ions of the positive specie is the same as between two ions of the negative specie. Therefore the parameters that may be set are $A_{11}, B_{11}, C_{11}, A_{12}, B_{12}$, and $C_{12}$.

**Notation:** Let $k' = n - k$, being the number of charges that are required to be removed to be left with an independent set of size $k$.charge Note that as the charge of each ion has a magnitude of one, a removal of $k'$ can only be achieved by removing $k'$ positive and $k'$ negative ions. The goal energy for the construction is set as $g = (k-1)(\frac{A_{12}}{e^{B_{12}}} - C_{12} - 1)$. To simplify the equations regarding the interaction between planes, let $\widehat{r}$ denote $\sqrt{r^2+1}$. An independent set is *left* if the ions left after a removal of $k'$ charges have labels corresponding to an independent set in $G$. To ensure that an independent set is left of size $k$ if and only if one exists, the following three inequalities must be satisfied:

$$\frac{A_{11}}{e^{B_{11}n}} - \frac{C_{11}}{n^6} + \frac{1}{n} + \frac{A_{12}}{e^{B_{12}\widehat{n}}} - \frac{C_{12}}{\widehat{n}^6} - \frac{1}{\widehat{n}} \geq \left| \frac{A_{12}}{e^{B_{12}}} - C_{12} - 1 \right| \tag{1}$$

$$n^2 \left| \frac{A_{11}}{e^{B_{11}r}} - \frac{C_{11}}{r^6} + \frac{1}{r} + \frac{A_{12}}{e^{B_{12}\widehat{r}}} - \frac{C_{12}}{\widehat{r}^6} - \frac{1}{\widehat{r}} \right| \leq \left| \frac{A_{12}}{e^{B_{12}}} - C_{12} - 1 \right|, \quad r \geq \sqrt{2}n \tag{2}$$

$$\frac{A_{11}}{e^{B_{11}r}} - \frac{C_{11}}{r^6} + \frac{1}{r} + \frac{A_{12}}{e^{B_{12}\widehat{r}}} - \frac{C_{12}}{\widehat{r}^6} - \frac{1}{\widehat{r}} > 0, \quad r \geq \sqrt{2}n \tag{3}$$

Theorem 5 formally states the correctness of this reduction, via Lemmas 1 - 4.

**Lemma 1.** *Inequalities (1) and (2) are sufficient to ensure that an independent set is left if one exists.*

**Lemma 2.** *There exists, for any structure created from a long orthogonal penny graph, some parameters such that Inequalities (1, 2) and (3) are satisfied.*

**Lemma 3.** *Given $k$ pairs, the energy will be less than $(k-1)(\frac{A_{12}}{e^{B_{12}}} - C_{12} - 1)$ only if the pairs correspond to an independent set of size $k$, for $\frac{A_{12}}{e^{B_{12}}} - C_{12} - 1 < 0$.*

**Lemma 4.** *It is always preferable to remove pairs from the construction from a long orthogonal penny graph under Inequalities (1- 3).*

Lemmas 1 and 2 show that the inequalities ensure that leaving an independent set is preferable, and are satisfiable for the Buckingham-Coulomb potential. Lemma 3 provides bounds, which may be calculated exactly using the construction provided by Lemma 2. Lemma 4 proves that when removing either member of a pair vertices, it is always preferable to select the the other member for removal.

**Theorem 5.** K-CHARGE REMOVAL*,* MINIMAL-AT-LEAST-K-CHARGE REMOVAL *and* AT-LEAST-K-CHARGE REMOVAL *are NP-Complete when limited to only two species of ion and restricted to the Buckingham-Coulomb potential function.*

*Proof.* Lemma 1 shows that, under Inequalities (1) and (2), the optimal solution will be to leave an independent set. Lemma 2 provides a construction such that the inequalities are satisfiable. Lemma 3 shows the upper bound is reachable if and only if an independent set has been left. It follows from Lemma 4 that it is preferable to remove a set of pairs over any other set of charges. Therefore there will be a satisfiable instance of K-CHARGE REMOVAL or any generalisation if and only if the instance of INDEPENDENT SET on a max degree 3 planar graph is satisfiable. Conversely if the INDEPENDENT SET instance is satisfiable, the corresponding K-CHARGE REMOVAL instance can be satisfied by leaving the vertices corresponding to the independent set in the penny graph construction. Hence under these restriction all three problems will be NP-Complete. □

**Conclusions and future work:** Motivated by analyses of computational complexity for CSP Problem we defined a class of functions for which the $k$-charge removal problem is NP-Complete in general. We have also shown that the problem remains NP-Complete under both the restriction that we have only two

species of ions and the Buckingham-Coulomb energy function and the restriction we only use the Coulomb potential on an unbounded number of ion species. One obvious question would be if approximation results can be gained for this problem. From a chemistry stand point, while we have made progress towards physical constructions there is still a lot that could be done. As such investigation into the restrictions of having more realistic physical values remains an important unexplored direction. Another question would be if we can investigate the convergence of these interactions, particularly the Coulomb potential, over a periodic structure to more fully understand the energy function.

## References


1. D. Adamson, A. Deligkas, V. Gusev, and I. Potapov. On the Hardness of Energy Minimisation for Crystal Structure Prediction. *arXiv.org/abs/1910.12026*.
2. AA Ageev, AV Kelmanov, and AV Pyatkin. Np-hardness of the euclidean max-cut problem. In *Doklady Mathematics*, volume 89, pages 343–345, 2014.
3. F. Barahona. On the computational complexity of ising spin glass models. *Journal of Physics A: Mathematical and General*, 15(10):3241–3253, oct 1982.
4. R. A. Buckingham. The classical equation of state of gaseous helium, neon and argon. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1938.
5. M. R. Cerioli, L. Faria, T. O. Ferreira, and F. Protti. A note on maximum independent sets and minimum clique partitions in unit disk graphs and penny graphs: complexity and approximation. *RAIRO*, 45(3):331–346, 2011.
6. B. N. Clark, C. J. Colbourn, and D. S. Johnson. Unit disk graphs. *Discrete Mathematics*, 86(1):165 – 177, 1990.
7. C Collins, MS Dyer, MJ Pitcher, GFS Whitehead, M Zanella, P Mandal, JB Claridge, GR Darling, and MJ Rosseinsky. Accelerated discovery of two crystal structure types in a complex inorganic phase field. *Nature*, 546(7657):280, 2017.
8. M. Krishnamoorthy and N. Deo. Node-deletion np-complete problems. *SIAM Journal on Computing*, 8(4):619–625, 1979.
9. A. O. Lyakhov, A. R. Oganov, and Mario V. How to predict very large and complex crystal structures. *Computer Physics Communications*, 181(9):1623 – 1632, 2010.
10. M. Mahajan, P. Nimbhorkar, and K. Varadarajan. The planar k-means problem is np-hard. In *WALCOM: Algorithms and Computation*, pages 274–285, 2009.
11. A. R. Oganov. Crystal structure prediction: reflections on present status and challenges. *Faraday Discuss.*, 211:643–660, 2018.
12. 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.
13. 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, jun 1985.
14. S. M Woodley and R. Catlow. Crystal structure prediction from first principles. *Nature materials*, 7(12):937, 2008.
15. M. Yannakakis. Node-and edge-deletion np-complete problems. In *STOC 1978*, pages 253–264, 1978.
16. M. Yannakakis. Node-deletion problems on bipartite graphs. *SIAM Journal on Computing*, 10(2):310–327, 1981.