CRYSTALLIZATION IN THE HEXAGONAL LATTICE FOR IONIC DIMERS

MANUEL FRIEDRICH AND LEONARD KREUTZ

Abstract. We consider finite discrete systems consisting of two different atomic types and investigate ground-state configurations for configurational energies featuring two-body short-ranged particle interactions. The atomic potentials favor some reference distance between different atomic types and include repulsive terms for atoms of the same type, which are typical assumptions in models for ionic dimers. Our goal is to show a two-dimensional crystallization result. More precisely, we give conditions in order to prove that energy minimizers are connected subsets of the hexagonal lattice where the two atomic types are alternately arranged in the crystal lattice. We also provide explicit formulas for the ground-state energy. Finally, we characterize the net charge, i.e., the difference of the number of the two atomic types. Analyzing the deviation of configurations from the hexagonal Wulff shape, we prove that for ground states consisting of n particles the net charge is at most of order $\mathrm{O}(n^{1/4})$ where the scaling is sharp.

1. Introduction

A fundamental problem in crystallography is to understand why atoms or molecules typically arrange themselves into crystalline order at low temperature. The challenge to prove rigorously that global minimizers of certain configurational energies arrange in a periodic lattice is referred to as the *crystallization problem* [4]. Starting with seminal works in one dimension [16, 18, 27, 32], the last decades have witnessed a remarkable interest in this mathematical problem for systems consisting of *identical* particles.

At zero or very low temperature, atomic interactions are expected to be governed solely by the geometry of the atomic arrangement. Identifying configurations with the respective positions of identical atoms $\{x_1, \dots, x_n\}$, one is concerned with the minimization of a configurational energy $\mathcal{E}(\{x_1, \dots, x_n\})$ which comprises classical interaction potentials. Crystallization then consists in proving the periodicity of ground-state configurations of $\mathcal{E}$.

Despite of its paramount theoretical and applicative relevance, rigorous results on crystallization are scarce and even under simplified assumptions the problem presents many difficulties. In this paper, we contribute to these fundamental mathematical questions by investigating crystallization of particle systems consisting of *two different atomic types*, also called *dimers*.

It is well known that the geometry of molecular compounds often differs from that of their components. An example in this direction is sodium chloride (NaCl) which has a FCC structure (Face Centered Cubic), whereas sodium crystals are BCC (Body Centered Cubic) and chlorine crystals are orthorhombic. The goal of this contribution is to illustrate such a phenomenon in a simplified setting by deriving a two-dimensional crystallization result in the hexagonal lattice for dimers whose components instead would likely assemble themselves in a triangular lattice.

2010 *Mathematics Subject Classification.* 82D25.

Key words and phrases. Ionic dimers, ground-state, configurational energy minimization, crystallization, hexagonal lattice, net charge.

Similar to the problem for systems of identical particles, we follow the classical molecular-mechanical frame of configurational energy minimization. We identify configurations of n particles with their respective *positions* $\{x_1, \dots, x_n\} \subset \mathbb{R}^2$ and additionally with their *types* $\{q_1, \dots, q_n\} \in \{-1, 1\}^n$. Our goal is to determine global minimizers of a corresponding interaction energy $\mathcal{E}(\{(x_1, q_1), \dots, (x_n, q_n)\})$.

More specifically, the energy $\mathcal{E} = \mathcal{E}_{\rm a} + \mathcal{E}_{\rm r}$ is assumed to consist of two short-ranged two-body interaction potentials $\mathcal{E}_{\rm a}$ and $\mathcal{E}_{\rm r}$, where $\mathcal{E}_{\rm a}$ represents the interactions between atoms of different type and $\mathcal{E}_{\rm r}$ encodes the energy contributions of atoms having the same type. The potential $\mathcal{E}_{\rm a}$ is attractive-repulsive and favors atoms sitting at some specific reference distance, whereas $\mathcal{E}_{\rm r}$ is a pure repulsive term. The latter is supposed to be strong enough at short distances such that the hexagonal lattice, with the two atomic types alternating, is energetically preferred with respect to closer packed structures. (Note that weaker short-ranged repulsion may indeed favor crystallization in the square lattice, see [14].)

An example of a dimer satisfying qualitatively the assumptions on the attractive and repulsive potentials is boron nitride [20, 33], which, similar to graphene [17], is a one-atom thick layer of atoms arranged in a hexagonal lattice. Moreover, our choice of the interaction potentials is motivated by the modeling of ions in ionic compounds. Indeed, the two interaction energies may be interpreted as (simplified) Coulomb-interactions between ions of equal and opposite charge. In this spirit, we will frequently refer to the particles as *ions* and to the atomic types $\{q_1, \dots, q_n\} \in \{-1, 1\}^n$ as positive or negative *charges*.

The paper contains the following three main results:

(1) Theorem 2.1: Under suitable assumptions on the attractive and the repulsive potentials, we characterize the ground-state energy of n-particle configurations of ions in the plane. In particular, we quantify exactly the surface energy related to atoms at the boundary of the ensemble. Moreover, we show that in our modeling frame ground states are repulsion free, i.e., $\mathcal{E}_{\rm r}(\{(x_1, q_1), \dots, (x_n, q_n)\}) = 0$.

(2) Theorem 2.3: We characterize the geometry of ground states. We show that each global minimizer of the configurational energy is essentially a connected subset of the regular hexagonal lattice with the two atomic types alternating. This characterization holds except for possibly one defect occurring at the boundary which can be identified explicitly either as a single atom or as a simple octagonal cycle in the structure.

(3) Theorem 2.5: We consider the *net charge* defined by $\mathcal{Q}(\{(x_1, q_1), \dots, (x_n, q_n)\}) = \sum_{i=1}^n q_i$ or, in other words, the (signed) difference of the number of the two atomic types. We provide a fine asymptotic characterization of the net charge as the number of particles grows. More precisely, we show that its fluctuation can be at most of order $\mathrm{O}(n^{1/4})$, i.e., $|\mathcal{Q}(\{(x_1, q_1), \dots, (x_n, q_n)\})| \le cn^{1/4}$ for some constant $c > 0$ independent of n. By giving an explicit construction we further prove that the scaling is sharp.

To the best of our knowledge, the present paper provides a first rigorous mathematical crystallization result for two-dimensional molecular compounds. Let us mention that in case of different types of particles, simulations are abundant, but rigorous results seem to be limited to [3, 5, 28]. Our analysis is largely inspired by many contributions on identical particle systems. We include here only a minimal crystallization literature overview and refer the reader to the recent review [4] for a more general perspective.

Mathematical results on crystallization in one dimension, proving or disproving periodicity of ground-state configurations for different choices of the energy, are by now quite classical, see, e.g., [16, 18, 27, 32]. In two dimensions for a finite number of identical particles, ground

states have been proved to be subsets of the triangular lattice by Heitman, Radin, and Wagner [19, 26, 34] for sticky and soft, purely two-body interaction energies. Recently, a new perspective on this problem was given by De Luca and Friesecke [8] using a discrete Gauss-Bonnet method from discrete differential geometry. Including angular potentials favoring $2\pi/3$ bond-angles, which is typically the case for graphene, crystallization in the hexagonal lattice has been proved by Mainini and Stefanelli [24], see also [13] for a related classification of ground states. In a similar fashion, if $\pi/2$ bond-angles are favored, the square lattice can be recovered [23]. Under less restrictive assumptions on the potentials, various results have been obtained in the thermodynamic limit [9, 10], namely as the number of particles n tends to infinity, most notably the seminal paper by Theil [31]. The crystallization problem in three dimension appears to be very difficult and only few rigorous results [11, 12, 21, 30] are currently available.

In two dimensions, a fine characterization of ground-state geometries is possible: as n increases, one observes the emergence of a polygonal cluster, the *Wulff shape* [2]. For the triangular [29], the hexagonal [6], and the square lattice [23], ground states differ from the Wulff shape by at most $\mathrm{O}(n^{3/4})$ particles and at most by $\mathrm{O}(n^{1/4})$ in Hausdorff distance. This sharp bound is called the $n^{3/4}$ law [7].

Our general proof strategy follows the induction method on bond-graph layers developed in [19, 23, 24, 26]. In particular, our study is related to [24], where crystallization for identical particles in the hexagonal lattice has been investigated under three-body angular potentials. Actually, the ground-state energy in the present context coincides with the one obtained there.

Concerning the characterization of ground-state geometries, however, richer geometric structures with respect to the ones of [24] are possible since configurations are in general less rigid without angular potentials: simple examples show that ground states may contain distortions and defects at the boundary such as octagons. From a technical point of view, novel geometric arguments in the induction step are necessary, complementing the available approaches, see Remark 7.1 for more details. Let us mention that, in any case, certain quantitative requirements on the potentials are indispensable since other assumptions might favor an assemblence of the atoms in other periodic structures. (In particular, we refer to the forthcoming paper [14] for the square-lattice case.)

For the characterization of the net charge in terms of the sharp scaling $\mathrm{O}(n^{1/4})$, we use the fact that in ground states the two atomic types are alternately arranged in the hexagonal lattice. Morever, the argument fundamentally relies on the $n^{3/4}$-law [6] which states that ground states differ from the Wulff shape by at most $\mathrm{O}(n^{3/4})$ particles, or equivalently and more relevant in our context, by $\mathrm{O}(n^{1/4})$ in Hausdorff distance. Let us emphasize that in our model the net charge is not a priori given but a crucial part of the minimization problem. The investigation of ground-state geometries under preassigned net charge is a different challenging problem which we defer to future studies.

The article is organized as follows. In Section 2 we introduce the precise mathematical setting and state the main results. In Section 3 we discuss elementary geometric properties of ground states. In Section 4 we construct explicitly some configurations in order to provide sharp upper bounds for the ground-state energy and the (signed) net charge. In Section 5 we introduce the concept of boundary energy and prove corresponding estimates which are instrumental for the induction method used in the proof of Theorem 2.1 and 2.3. In Section 6 we give a lower bound for the ground-state energy matching the one of the configurations constructed in Section 4. In Section 7 we characterize the geometry of ground states and finally Section 8 is devoted to the proof of a $n^{1/4}$-law for the net charge of ground states.

2. Setting and main results

In this section we introduce our model, give some basic definitions, and present our main results.

2.1. **Configurations and interaction energy.** We consider two-dimensional particle systems consisting of two different atomic types and model their interaction by classical potentials in the frame of Molecular Mechanics [1, 15, 22].

Let $n \in \mathbb{N}$. We indicate the *configuration* of n particles by

$$C_n = \{(x_1, q_1), \dots, (x_n, q_n)\} \subset (\mathbb{R}^2 \times \{-1, 1\})^n,$$

identified with the respective *atom positions* $X_n = (x_1, \dots, x_n) \in \mathbb{R}^{2n}$ together with their *types* $Q_n = (q_1, \dots, q_n) \in \{-1, 1\}^n$. Referring to a model for ionic compounds, we will frequently call Q_n the *charges* of the atoms, $q = 1$ representing *cations* and $q = -1$ representing *anions*. Indeed, our choice of the empirical potentials (see below) is inspired by ions in ionic compounds which are primarily held together by the electrostatic forces between the net negative charge of the anions and net positive charge of the cations [25].

We define the phenomenological energy $\mathcal{E} : (\mathbb{R}^2 \times \{-1, 1\})^n \to \overline{\mathbb{R}}$ of a given configuration $\{(x_1, q_1), \dots, (x_n, q_n)\} \in (\mathbb{R}^2 \times \{-1, 1\})^n$ by

$$\mathcal{E}(C_n) = \frac{1}{2} \sum_{\substack{i \neq j \\ q_i = q_j}} V_{\mathrm{r}}(|x_i - x_j|) + \frac{1}{2} \sum_{\substack{i \neq j \\ q_i \neq q_j}} V_{\mathrm{a}}(|x_i - x_j|), \tag{1}$$

where $V_{\mathrm{r}}, V_{\mathrm{a}} : [0, +\infty) \to \overline{\mathbb{R}}$ are a *repulsive potential* and an *attractive-repulsive potential*, respectively. The factor $\frac{1}{2}$ accounts for the fact that every contribution is counted twice in the sum. The two potentials are pictured schematically in Fig. 1. Let $r_0 \in [1, (2\sin(\frac{\pi}{7}))^{-1})$ and note that $r_0 \leq \sqrt{2}$. The attractive-repulsive potential V_{a} satisfies

- [i] $V_{\mathrm{a}}(r) = +\infty$ for all $r < 1$,
- [ii] $V_{\mathrm{a}}(r) = -1$ if and only if $r = 1$ and $V_{\mathrm{a}}(r) > -1$ otherwise,
- [iii] $V_{\mathrm{a}}(r) \leq 0$ for all $r \geq 1$ with equality for all $r > r_0$.

The distance $r = 1$ represents the (unique) equilibrium distance of two atoms with opposite charge. The constraint $V_{\mathrm{a}}(r) = +\infty$ is usually called the *hard-interaction* assumption. The choice of V_{a} reflects a balance between a long-ranged Coulomb attraction and a short-ranged repulsive force when a pair of ions comes too close to each other. Assumption [iii] restricts the interaction range and ensures that the *bond graph* (see Section 2.2) is planar.

The repulsive potential V_{r} satisfies

- [iv] $V_{\mathrm{r}}(r) = +\infty$ for all $r < 1$ and $0 \leq V_{\mathrm{r}}(r) < +\infty$ for all $r \geq 1$,
- [v] V_{r} is non-increasing and convex for $r \geq 1$,
- [vi] $V_{\mathrm{r}}(\sqrt{2}r_0) > 3$,
- [vii] $V_{\mathrm{r}}(r) = 0$ iff $r \geq \sqrt{3}$.

Assumption [iv] is the hard-interaction assumption for the repulsive potential. The natural assumption [v] is satisfied for example for repulsive Coulomb interactions. We remark that some quantitative requirement of the form [vi] and [vii] are indispensable to obtain a crystallization result in the hexagonal lattice. Indeed, for vanishing V_{r}, ground states could be patches of the triangular lattice. Other quantitative assumptions on the repulsive potential will favor an

assemblence of the atoms in the square-lattice as we prove in [14]. Also note that two-body pair interactions for identical particle systems typically favor crystallization in the triangular lattice [26, 31], when the interaction is of attractive-repulsive-type (instead of pure repulsive-type). This reflects the aforementioned fact that the geometry of molecular compounds often differs from that of their components.

A main assumption is [vii], i.e., the repulsion vanishes at $\sqrt{3}$ (the distance between second neighbors in the hexagonal lattice). Note that, if $V_{\mathrm{r}}(\sqrt{3})$ is instead assumed to be positive and sufficiently large, then crystallization in the hexagonal lattice is not expected as, e.g., a one-dimensional chain of atoms is energetically favorable. For small, positive $V_{\mathrm{r}}(\sqrt{3})$ we still expect crystallization in the hexagonal lattice, but the analysis is much more demanding and beyond the scope of the present contribution.

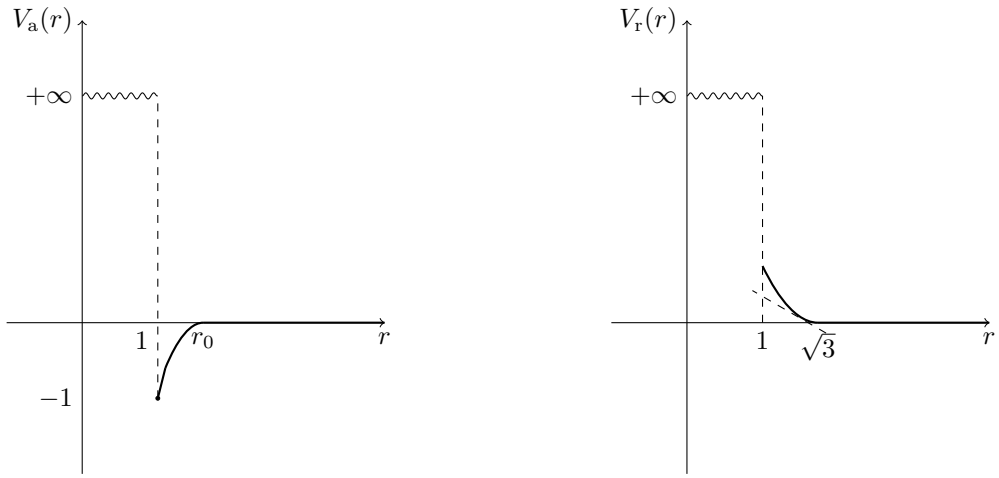

FIGURE 1. The potentials V_{a} and V_{r}.

Finally, we require the following *slope conditions*

[viii] $\quad V'_{\mathrm{r},-}(\sqrt{3}) < -\dfrac{3}{\pi}, \qquad \dfrac{1}{r-1}(V_{\mathrm{a}}(r) - V_{\mathrm{a}}(1)) > -2V'_{\mathrm{r},+}(1) \quad \text{for all } r \in (1, r_0],$

where the functions $V'_{\mathrm{r},+}, V'_{\mathrm{r},-}$ denote the right and left derivative, respectively. (They exist due to convexity of V_{r}.) These conditions are reminiscent of the *soft-interaction* assumption by RADIN [26] and the *slope condition* for an angular potential by MAININI AND STEFANELLI [24]. In particular, we assume that the repulsion grows linearly out of $\sqrt{3}$ and that, roughly speaking, the slope of V_{a} is steep enough compared to the slope of V_{r}. We highlight that without assumptions of this kind finite crystallization is presently not known, not even for identical particle systems. Assumption [viii] is only needed in Lemma 5.1 where the energy contribution of atoms at the boundary of the configuration is estimated. See also Remark 5.3 for a comparison of our model to [24]. Note that [viii] also restricts the choice of possible r_0. Assumptions [ii] and [vii] ensure that the (infinite) hexagonal lattice (see (6) below) is *stress free*. An assumption of this kind is necessary as otherwise for finite particle systems surface relaxation at the boundary of the configuration could occur, eventually ruling out finite crystallization in any lattice.

We remark that the assumptions are chosen here for the sake of maximizing simplicity rather than generality. Some conditions, e.g., about the hard-interaction assumption or the exact

value in [vi], could be weakened at the expense of more elaborated arguments. In particular, two different repulsive potentials for cations and anions could be considered as long as [iv]-[viii] hold for both potentials. In the following we assume that conditions [i]-[viii] are always satisfied.

2.2. **Basic notions.** In this section we collect some basic notions. Consider a configuration $C_n \in (\mathbb{R}^2 \times \{-1, 1\})^n$ with finite energy consisting of the positions $X_n = (x_1, \dots, x_n) \in \mathbb{R}^{2n}$ and the charges $Q_n = (q_1, \dots, q_n) \in \{-1, 1\}^n$.

Neighborhood, bonds, angles: For $i \in \{1, \dots, n\}$ we define the *neighborhood of* $x_i \in \mathbb{R}^2$ by

$$\mathcal{N}(x_i) = (X_n \setminus \{x_i\}) \cap \{x \in \mathbb{R}^2 : \ |x - x_i| \leq r_0\}. \tag{2}$$

If $x_j \in \mathcal{N}(x_i)$, we say that x_i and x_j are *bonded*. We will say that x_i is k-bonded if $\#\mathcal{N}(x_i) = k$. Given $x_j, x_k \in \mathcal{N}(x_i)$, we define the *bond-angle* between x_j, x_i, x_k as the angle between the two vectors $x_k - x_i$ and $x_j - x_i$ (choose anti-clockwise orientation, for definiteness). In general, we say that it is an angle at x_i.

The bond graph: The set of atomic positions $X_n \subset \mathbb{R}^{2n}$ together with the set of bonds $\{\{x_i, x_j\} : x_j \in \mathcal{N}(x_i)\}$ forms a graph which we call the *bond graph*. Since for finite energy configurations we get $\operatorname{dist}(x_i, X_n \setminus \{x_i\}) \geq 1$ and $x_j \in \mathcal{N}(x_i)$ only if $|x_i - x_j| \leq r_0 < \sqrt{2}$ we have that their bond graph is a planar. Indeed, given a quadrangle with all sides and one diagonal in $[1, r_0]$, the second diagonal is at least $\sqrt{2} > r_0$. Note that assumption [iii] states that the attractive interactions are restricted to nearest neighbors in the bond graph only. If no ambiguity arises, the number of bonds in the bond graph will be denoted by b, i.e.,

$$b = \#\{\{x_i, x_j\} : x_j \in \mathcal{N}(x_i)\}.$$

We say a configuration is *connected* if each two atoms are joinable through a simple path in the bond graph. Any simple cycle of the bond graph is a *polygon*.

Acyclic bonds: We say that a bond is *acyclic* if it is not contained in any simple cycle of the bond graph. Among acyclic bonds we distinguish between *flags* and *bridges*. We call a bridge an acyclic bond which is contained in some simple path connecting two vertices which are included in two distinct cycles. All other acyclic bonds are called flags, see Fig. 2.

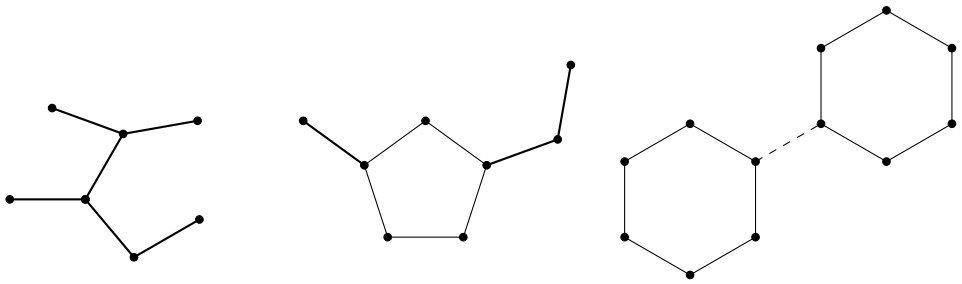

FIGURE 2. Examples of flags (bold) and a bridge (dashed)

Defects: We call an elementary polygon in the bond graph which is not a hexagon a *defect*. A configuration is said to be *defect-free* if all its elementary polygons are hexagons. We also introduce the *excess of edges* η by

$$\eta = \sum_{j \geq 6} (j - 6) f_j, \tag{3}$$

where f_j denotes the number of polygons with j vertices in the bond graph. The excess of edges is a tool to quantify the number of defects in the bond graph. Note that the summation in (3) runs over $j \geq 6$. This is due to the fact that we use this definition only for configurations whose bond graph contains only k-gons with $k \geq 6$, cf. Lemma 3.2. We remark that in [8] the *excess of edges* (with respect to the triangular lattice) is referred to as *defect measure*.

In the following we refer to C_n instead of X_n when speaking about its bond graph, acyclic bonds, or connectedness properties, when no confusion may occur.

Charges: We say that a configuration satisfying

$$\mathcal{N}(x_i) \cap \{x_j \in X_n : q_j = q_i\} = \emptyset \text{ for all } i \in \{1, \dots, n\} \tag{4}$$

has *alternating charge distribution*. Moreover, we call a configuration *repulsion-free* if $|x_i - x_j| \geq \sqrt{3}$ for all $x_i \neq x_j$ with $q_i = q_j$. The *net charge* of a configuration is defined as the (signed) difference of the number of the two atomic types, i.e.,

$$\mathcal{Q}(C_n) := \sum_{i=1}^{n} q_i. \tag{5}$$

Ground state: A configuration C_n is called *ground state* for the interaction energy (1) if for all $C_n' \subset (\mathbb{R}^2 \times \{-1, 1\})^n$ there holds

$$\mathcal{E}(C_n) \leq \mathcal{E}(C_n').$$

In other words, C_n minimizes (1) among all possible configurations consisting of n atoms.

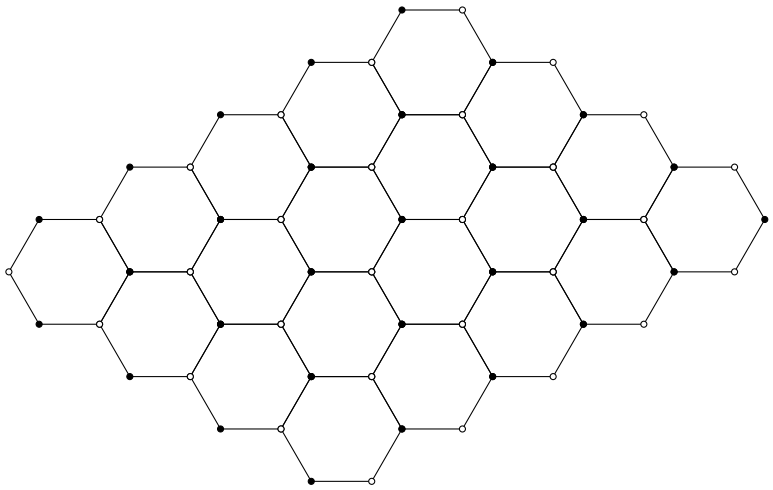

FIGURE 3. A configuration with atomic positions forming a subset of the hexagonal lattice. The configuration has alternating charge distribution where white indicates $q = 1$ and black indicates $q = -1$.

The hexagonal lattice: We define the *hexagonal lattice* by

$$\mathcal{L} := \{pv_1 + qv_2 + rv_0 : p, q \in \mathbb{Z}, r \in \{0, 1\}\}, \tag{6}$$

with $v_1 = (0, \sqrt{3})$, $v_2 = \frac{1}{2}(3, \sqrt{3})$ and $v_0 = (1, 0)$. We identify it with its bond graph defined above.

Note that the hexagonal lattice is planar, connected, and all edges have unit length. Since it is bipartite, we can associate to all positions $X_n \subset \mathcal{L}$ of atoms a competitor for the minimization of (1) by two-coloring (and hereby choosing the charge of) the atoms. The corresponding configuration C_n is pictured in Fig. 3. Then by assumption [ii], [iii], and [vii] we have

$$\mathcal{E}(C_n) = -b,$$

since all atoms of same charge have at least distance $\sqrt{3}$ and all atoms of X_n are bonded only to atoms of opposite charge. This means that for subsets of the hexagonal lattice the energy is computed (up to sign) by counting the number of bonds.

2.3. **Main results.** In this section we state our main results. We will derive a rigorous planar crystallization result in the spirit of [19, 23, 24, 26]. Afterwards, we will investigate the *net charge* of ground-state configurations. Our first result characterizes the energy of ground states and shows that all ground states are connected. To this end, we introduce the function

$$\beta(n) := \frac{3}{2}n - \sqrt{\frac{3}{2}n} \tag{7}$$

for $n \in \mathbb{N}$. By $\lfloor t \rfloor$ we denote the integer part of $t \in \mathbb{R}$.

Theorem 2.1 (Ground-state energy)**.** *Let $n \geq 1$ and let C_n be a ground state. Then C_n is connected and satisfies*

$$\mathcal{E}(C_n) = -b = -\lfloor \beta(n) \rfloor. \tag{8}$$

Remark 2.2. In view of assumptions [ii] and [vii], we have that $\mathcal{E}(C_n) \geq -b$ with equality if and only if the configuration is repulsion-free and all bonds have unit length. In particular, Theorem 2.1 implies that ground states satisfy both properties.

The next result states that ground states are essentially subsets of the hexagonal lattice with alternating charge distribution. Moreover, they satisfy some topological properties. Without further notice, all following statements regarding the geometry of ground states hold up to isometry.

Theorem 2.3 (Characterization of ground states)**.** *Let $n \geq 30$ and let C_n be a ground state. Then except for possibly one flag, C_n is a connected subset of the hexagonal lattice with alternating charge distribution. Moreover, C_n is defect-free and does not contain any bridges.*

Remark 2.4. (i) More precisely, we show that the bond graph of ground states consists only of hexagonal cycles except for at most two flags. Since ground states are repulsion-free and all bonds have unit length, at least one of the flags is contained in the hexagonal lattice. If two flags exist and are connected, we note that one of them can be rotated in a continuous way without changing the energy.

(ii) The ground states for $n \leq 29$ can also be characterized, but due to the smallness of the structures, more degeneracies can occur. In particular, for $n = 8, 9, 12, 15, 18, 21, 29$ there might be one octagon at the boundary. (We refer to Remark 7.11 for more details.) For $n < 10$ the structure can be much more flexible, see Fig. 8.

Our third main result addresses the net charge. Recall (5).

Theorem 2.5 (Net charge)**.** *The following properties for the net charge of ground states hold:*

(i) *(Charge control) There is a universal constant $c > 0$ such that for all $n \in \mathbb{N}$ and all ground states C_n the charge satisfies $|\mathcal{Q}(C_n)| \leq cn^{1/4}$.*
(ii) *(Sharpness of the $n^{1/4}$-scaling) There exists an increasing sequence of integers $(n_j)_j$ and ground states $(C_{n_j})_j$ such that*

$$\liminf_{j\to+\infty} n_j^{-1/4}|\mathcal{Q}(C_{n_j})| > 0.$$

The proof of our main results will be given in Sections 3-8. In particular, in Section 4 we construct explicitly configurations and give the upper bound for the ground state energy (Theorem 2.1) as well as the proof of Theorem 2.5(ii). In Section 6 we conclude the proof of Theorem 2.1 by providing the lower bound. Section 7 is then devoted to the characterization of ground-state geometries for $n \geq 30$ (Theorem 2.3). Finally, in Section 8 we prove Theorem 2.5(i).

3. Elementary geometric properties

In this section we provide some elementary geometric facts for ground states independently of $n \in \mathbb{N}$. In the proofs, we will use the following convention: we say that we *relocate* $(x, q) \in C_n$, and write $C_n - \{(x, q)\}$, by considering the configuration $(C_n \cup (x + \tau, q)) \setminus (x, q)$, where $\tau \in \mathbb{R}^2$ is chosen such that

$$\operatorname{dist}(X_n, x + \tau) \geq \sqrt{3}.$$

We relocate a set of atoms $A \subset C_n$ by relocating successively every $(x, q) \in A$ and write $C_n - A$. Note that $C_n - A$ still consists of n particles.

The first lemma addresses the neighbors. Recall (2) and (4).

Lemma 3.1 (Neighbors and charge distribution)**.** *Let C_n be a ground state. Then C_n has alternating charge distribution and*

$$\#\mathcal{N}(x_i) \leq 3 \text{ for all } i \in \{1, \dots, n\}. \tag{9}$$

Proof. Since C_n is a ground state, it holds that $\mathcal{E}(C_n) < +\infty$. Therefore, by assumption [i],[iv]

$$\operatorname{dist}(x_i, X_n \setminus \{x_i\}) \geq 1 \text{ for all } i \in \{1, \dots, n\}. \tag{10}$$

For brevity, we define $\mathcal{N}_{\rm rep}(x_i) = \mathcal{N}(x_i) \cap \{x_j \in X_n : q_j = q_i\}$ for all $i \in \{1, \dots, n\}$. We give the proof of the statement in two steps.

Claim 1: We have

$$\#\mathcal{N}_{\rm rep}(x_i) \leq 1 \text{ for all } i \in \{1, \dots, n\},$$

and if $\#\mathcal{N}(x_i) \leq 4$, then

$$\#\mathcal{N}_{\rm rep}(x_i) = 0.$$

Proof of Claim 1: First, (10) and $r_0 < (2\sin(\frac{\pi}{7}))^{-1}$ entail by an elementary geometric argument that

$$\#\mathcal{N}(x_i) \leq 6 \text{ for all } i \in \{1, \dots, n\}. \tag{11}$$

In fact, if $\#\mathcal{N}(x_i) \geq 7$, two neighbors of x_i would necessarily have distance smaller than 1. Now assume by contradiction that $\#\mathcal{N}_{\rm rep}(x_i) \geq 2$. Note that every bond between points of different charge contributes at least -1 to the energy by [ii]. This along with $V_{\rm r} \geq 0$ and the fact that

the energy per neighbor of same charge is larger than 3 (see [v] and [vi]) allows as to relocate (x_i, q_i): we obtain

$$\mathcal{E}(C_n - \{(x_i,q_i)\}) < \mathcal{E}(C_n) + \#\left(\mathcal{N}(x_i) \setminus \mathcal{N}_{\rm rep}(x_i)\right) - 3\#\mathcal{N}_{\rm rep}(x_i) \le \mathcal{E}(C_n) + 4 - 6 < \mathcal{E}(C_n).$$

This contradicts the fact that C_n is a ground state. The argument in the case $\#\mathcal{N}(x_i) \le 4$ and $\#\mathcal{N}_{\rm rep}(x_i) \ge 1$ is similar: due to the fact that every bond between points of different charge contributes at least -1 to the energy and the contribution of neighbors of the same charge is larger than 3, we have

$$\mathcal{E}(C_n - \{(x_i,q_i)\}) < \mathcal{E}(C_n) + \#\left(\mathcal{N}(x_i) \setminus \mathcal{N}_{\rm rep}(x_i)\right) - 3\#\mathcal{N}_{\rm rep}(x_i) \le \mathcal{E}(C_n) + 3 - 3 \le \mathcal{E}(C_n).$$

This contradiction shows that Claim 1 holds true.

Claim 2: $\#\mathcal{N}(x_i) \le 3$ for all $i \in \{1,\dots,n\}$.

Proof of Claim 2: Assume by contradiction that there exists $i \in \{1,\dots,n\}$ such that $\#\mathcal{N}(x_i) \ge 4$. By Claim 1 we may suppose that there exist $\{x_0,\dots,x_3\} \subset \mathcal{N}(x_i)$ with $q_j = -q_i$ $j = 0,\dots,3$. We let $\theta_j \in [0,2\pi)$ be the angle between x_j, x_i, x_{j+1}. (Here and in the following the indices have to be understood modulo 4.) We can choose $j_0 \in \{0,\dots,3\}$ such that

$$\theta_{j_0} + \theta_{j_0-1} \le \frac{1}{4}\sum_{j=0}^{3}(\theta_j + \theta_{j-1}) = \pi.$$

Note that $|x_{j_0} - x_{j_0+1}| \le 2r_0 \sin\left(\theta_{j_0}/2\right), |x_{j_0} - x_{j_0-1}| \le 2r_0 \sin\left(\theta_{j_0-1}/2\right)$. By concavity of $\sin(x)$ for $x \in [0,\pi]$ and the fact that $\sin(x)$ is increasing for $x \in [0, \frac{\pi}{2}]$, we have

$$\begin{aligned}|x_{j_0} - x_{j_0+1}| + |x_{j_0} - x_{j_0-1}| &\le r_0\left(2\sin\left(\frac{\theta_{j_0}}{2}\right) + 2\sin\left(\frac{\theta_{j_0-1}}{2}\right)\right)\\ &\le 4r_0 \sin\left(\frac{\theta_{j_0}+\theta_{j_0-1}}{4}\right) \le 4r_0\sin\left(\frac{\pi}{4}\right) = 2\sqrt{2}r_0.\end{aligned}$$

Since $V_{\rm r}$ is convex and non-increasing (see [v]), we find

$$V_{\rm r}\left(|x_{j_0} - x_{j_0+1}|\right) + V_{\rm r}\left(|x_{j_0} - x_{j_0-1}|\right) \ge 2V_{\rm r}\left(\frac{1}{2}\left(|x_{j_0} - x_{j_0+1}| + |x_{j_0} - x_{j_0-1}|\right)\right) \ge 2V_{\rm r}(\sqrt{2}r_0).$$

Using [ii], [vi], $V_{\rm r} \ge 0$, and (11) we finally observe

$$\mathcal{E}(C_n - \{(x_{j_0}, q_{j_0})\}) \le \mathcal{E}(C_n) - 2V_{\rm r}(\sqrt{2}r_0) + \#\mathcal{N}(x_{j_0}) \le \mathcal{E}(C_n) - 2V_{\rm r}(\sqrt{2}r_0) + 6 < \mathcal{E}(C_n).$$

This contradicts the fact that C_n is a ground state and concludes the proof of Claim 2.

Now (9) follows from Claim 2. The property of alternating charge (see (4)) follows from (9) and the second statement of Claim 1. This concludes the proof of Lemma 3.1. □

We now investigate simple cycles in the bond graph.

Lemma 3.2 (Polygons)**.** *Let C_n be a ground state. Then all polygons have at least 6 edges and the number of edges is always even.*

Proof. As C_n has alternating charge distribution, two successive vertices of a path of the bond graph of a ground state have to be of different charge. This prohibits cycles of odd length. In particular, this rules out triangles and pentagons. We are thus left to prove that there is no simple square.

Now assume by contradiction that the bond graph of a ground state contains a simple square. As two successive vertices of the square have different charge, the vertices connected by the diagonal have the same charge. Denote the lengths of the two diagonals by d_1 and d_2, respectively. An elementary computation shows

$$d_1^2 + d_2^2 \leq 4r_0^2. \tag{12}$$

In fact, denoting the vertices of the square by $x_1, \dots, x_4$ (counterclockwise) one can use the elementary expansion

$$|x_1 - x_3|^2 + |x_2 - x_4|^2 = |x_1 - x_2|^2 + |x_2 - x_3|^2 + |x_3 - x_4|^2 + |x_4 - x_1|^2 - |x_1 - x_2 + x_3 - x_4|^2$$

along with the fact that each bond length is smaller or equal to r_0. Now (12) together with Young's inequality gives

$$(d_1 + d_2)^2 \leq 2(d_1^2 + d_2^2) \leq 8r_0^2.$$

Since V_{r} is convex and non-increasing (see [v]), we obtain

$$V_{\mathrm{r}}(d_1) + V_{\mathrm{r}}(d_2) \geq 2V_{\mathrm{r}}\left(\frac{1}{2}(d_1 + d_2)\right) \geq 2V_{\mathrm{r}}(\sqrt{2}r_0).$$

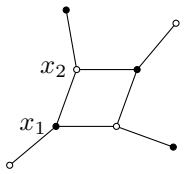

FIGURE 4. Relocating (x_1, q_1) and (x_2, q_2) strictly decreases the energy.

We relocate two neighboring particles in the square, denoted by (x_1, q_1) and (x_2, q_2). By (9) and the fact that x_1 and x_2 share a bond, we observe that hereby we remove at most 5 bonds between atoms of different charge, see Fig. 4. Using [ii], [vi], and $V_{\mathrm{r}} \geq 0$, we obtain

$$\mathcal{E}(C_n - \{(x_1, q_1), (x_2, q_2)\}) \leq \mathcal{E}(C_n) - 2V_{\mathrm{r}}(\sqrt{2}r_0) + 5 < \mathcal{E}(C_n) - 1 < \mathcal{E}(C_n).$$

This contradicts the fact that C_n is a ground state and concludes the proof. □

We close this section with some properties for repulsion-free configurations with bonds of unit length.

4. Upper bound on the ground-state energy

This section is devoted to the explicit construction of configurations C_n with alternating charge distribution which are subsets of the hexagonal lattice. These configurations provide a reference energy value for every n, namely $\mathcal{E}(C_n) = -\lfloor \beta(n) \rfloor$, cf. (7). This already gives the upper bound in (8). Moreover, we will construct explicitly configurations with net charge (see (5)) of the order $n^{1/4}$ which establishes Theorem 2.5(ii). We defer the lower bound on the ground-state energy, the upper bound on the net charge, and the characterization of the ground states to the subsequent sections.

By the special geometry of the hexagonal lattice, it is quite natural to give an interpretation of the two terms appearing in β. The leading order term of the energy is given by $-\frac{3}{2}n$. It is related to the fact that each atom which is not contained in the exterior face of the bond graph

is bonded to exactly three atoms of opposite charge and every bond connects two atoms. The repulsive term in the energy is zero for such configurations since the distance of each two atoms of the same charge is bigger or equal than $\sqrt{3}$. The additional lower order correction term in the energy is is due to the fact that a certain proportion of the atoms touching the exterior face of the bond graph is only bonded to two atoms of opposite charge. Their cardinality scales like $\sqrt{n}$.

4.1. **Special subsets of the hexagonal lattice.** We exhibit special configurations that are subsets of the hexagonal lattice with energy $-\lfloor\beta(n)\rfloor$. This provides an upper bound for the ground-state energy. We give a recursive construction for these geometries following the ideas in [24, Section 4,5].

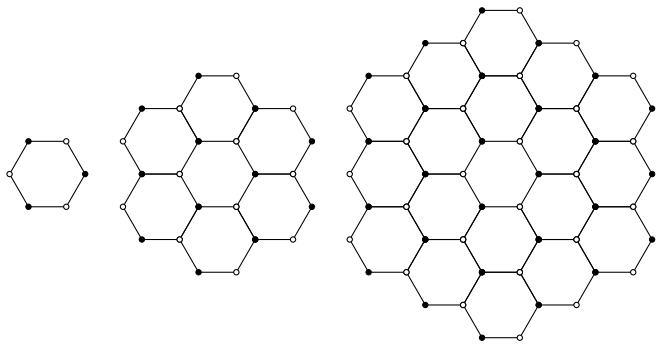

FIGURE 5. C_6, C_{24} and C_{54}.

First, we provide the construction for $n = 6k^2$, $k \in \mathbb{N}$. For $k = 1$ we define C_6 to be a regular hexagon $\{x_1, \dots, x_6\}$, where the points $x_1, \dots, x_6$ are arranged in a counter-clockwise sense as the vertices of the hexagon and with charges $q_i = (-1)^{i+1}, i = 1, \dots, 6$. Once we have constructed $C_{6(k-1)^2}$, we construct C_{6k^2} by attaching hexagons on all boundary sides of $C_{6(k-1)^2}$ such that every atom has two or three atoms of opposite charge in its neighborhood. (This is possible since the hexagonal lattice is bipartite.) These configurations are pictured in Fig. 5. Due to their symmetry, these configurations are called *daisies* and we will indicate their atomic positions also by $X_{6k^2}^{\text{daisy}}$. Note that daisies have net charge zero.

Now for $n \neq 6k^2$ we can assume that $n = 6k^2 + m$ for some $1 \leq m < 12k + 6$. We start from C_{6k^2} and add a new atom to the bond graph in such a way that it gets bonded to the leftmost among the uppermost atoms of C_{6k^2} and that it has distance larger or equal than $\sqrt{3}$ to all the other atoms. We choose the charge to be the opposite charge of the leftmost among the uppermost atoms. Then we add atoms in a counter-clockwise fashion such that the latest atom added is bonded to the atom added in the previous step and possibly to some other atom of C_{6k^2}. Moreover, its distance to all the other atoms is at least $\sqrt{3}$. We choose the charge to be the opposite of the charge of the atom added in the previous step. One can realize that this defines uniquely a procedure in order to add m atoms as shown in Fig. 6.

The first main result of this section is the following upper bound for the ground-state energy.

Proposition 4.1 (Upper bound for ground-state energy)**.** *For all $n \in \mathbb{N}$ we have*

$$\mathcal{E}(C_n) = -\lfloor\beta(n)\rfloor.$$

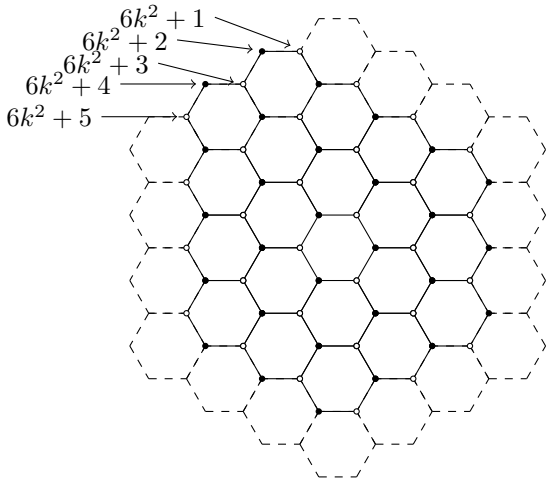

FIGURE 6. Construction of the $(6k^2+m)$-configuration for $k=3$ and $m=5$.

Proof. The proof follows as in [24, Proposition 4.1, Proposition 5.1], additionally noting that in the provided construction, all atoms in the bond graph are bonded to points of opposite charge only. □

4.2. **Daisies with an additional parallelogram.** Note that the above configurations have net charge in $\{-1,0,1\}$. Starting with a daisy and attaching a parallelogram in a suitable way, we can also construct configurations with energy $-\lfloor\beta(n)\rfloor$ having a net charge of order $n^{1/4}$. The following construction is inspired by the one of [6] applied in connection to the derivation of the so-called $n^{3/4}$-law.

We choose $k\in\mathbb{N}$ with $k\geq 252$, $r\in 6\mathbb{N}$ with $r\leq\sqrt{k/7}$, and let

$$n := 6k^2+2kr+\frac{1}{6}r^2+1 = 6\hat{k}^2+1, \tag{13}$$

where $\hat{k}=k+r/6$. We construct a configuration C_n as follows. We start from C_{6k^2} and add a new atom to the bond graph in such a way that it gets bonded to the leftmost among the uppermost atoms of C_{6k^2} and that it has distance greater or equal than $\sqrt{3}$ to all the other atoms. We choose the charge to be the opposite charge of the leftmost among the uppermost atoms. (Say, charge 1 for definiteness.) Then we add atoms in a counter-clockwise way similar to the previous construction until 'the line is completed'. Overall, we add $2k-1$ atoms by this procedure where k atoms have charge 1 and $k-1$ atoms have charge -1. We then repeat this construction by adding another row of atoms on the top where we need to add $2(k-1)-1$ atoms. We repeat this until we have added r rows of atoms corresponding to

$$\sum_{j=k-r+1}^{k}(2j-1) = 2kr-r^2$$

added atoms. Note that in each row the number of added atoms of charge 1 exceeds the number of added atoms of charge -1 by exactly one. The construction is sketched in Fig. 7.

Finally, in a last row we add $m := 7r^2/6 + 1$ atoms. Note that by the assumptions $k \geq 5$ and $r \leq \sqrt{k/7}$ we have

$$\frac{7}{6}r^2 + 1 \leq k/6 + 1 \leq \frac{k-r}{2},$$

i.e., the chain of atoms added in row r is 'long enough' so that m atoms can be added in the last row. Note that the resulting configuration consists of n atoms, see (13).

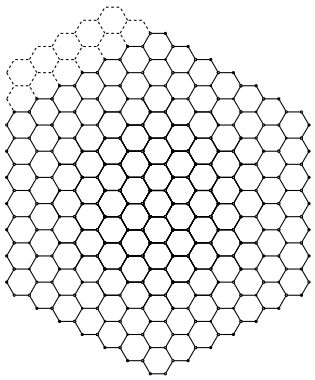

FIGURE 7. Construction of a daisy with an additional parallelogram. Two rows of atoms have been added where the first and the last atom in the added rows have charge +1. Thus, the net charge of the configuration is +2.

Concerning the energy, we observe that in the row where $2j-1$ atoms are added we add exactly $3j-2$ bonds to the bond graph for $k-r+1 \leq j \leq k$. In the ultimate row we add $1 + 3(m-1)/2 = 1 + 7r^2/4$ bonds. Consequently, using Proposition 4.1 the energy of C_n is given by

$$\begin{aligned}\mathcal{E}(C_n) &= \mathcal{E}(C_{6k^2}) - \sum_{j=k-r+1}^{k}(3j-2) - 1 - 7r^2/4 = -9k^2 + 3k - 3kr + 3r^2/2 + r/2 - 1 - 7r^2/4 \\ &= -\frac{3}{2}6(k+r/6)^2 + \sqrt{\frac{3}{2}6(k+r/6)^2} - 1 = -\lfloor \beta(6\hat{k}^2)\rfloor - 1 = -\lfloor \beta(n)\rfloor.\end{aligned}$$

We now determine the *net charge* of the configuration. Recall that the configuration C_{6k^2}, from which we started our construction, has net charge zero. As explained above, in each added row the number of added atoms of charge 1 exceeds the number of added atoms of charge -1 by exactly one, i.e., $\mathcal{Q}(C_n) = r + 1$.

We are now in the position to give the proof of Theorem 2.5(ii). To this end, consider the sequence of integers $n_j = 6(j + r_j/6)^2 + 1$, where $r_j = 6\lfloor \frac{1}{6}\sqrt{j/7}\rfloor$, and the ground states C_{n_j} constructed above. Note that by $r_j \leq j$ we get $n_j \leq 9j^2$. Thus, we calculate

$$\mathcal{Q}(C_{n_j}) = r_j + 1 \geq \sqrt{j/7} - 5 \geq \frac{1}{\sqrt{21}} n_j^{1/4} - 5.$$

This yields $\liminf_{j\to+\infty} n_j^{-1/4}|\mathcal{Q}(C_{n_j})| > 0$.

5. Boundary Energy

In this section we introduce the concept of boundary energy and prove a corresponding estimate which will be fundamental for the characterization of ground states and their energy in Section 6, 7.

Boundary atoms, boundary energy: Within the bond graph, we say that an atom is a *boundary atom* if it is not contained in the interior region of any simple cycle. Otherwise, we call it *bulk atom*. We denote the union of the boundary atoms by ∂X_n. A *boundary bond* is a bond containing a boundary atom. All other bonds are called *bulk bonds*. Similarly, a bond-angle will be called *boundary angle* or *bulk angle*, depending on whether the associated atom is a boundary atom or not.

We denote by d the number of boundary atoms of X_n. Given C_n, we define its *bulk*, denoted by C_n^{bulk}, as the sub-configuration obtained by dropping all boundary atoms (and the corresponding charges). Similarly, the particle positions are indicated by X_n^{bulk}. With the above definition, we have that the bulk is an $(n-d)$-atom configuration. We define $\mathcal{E}^{\mathrm{bulk}}(C_n)$ as the energy of C_n^{bulk}. We then have two contributions to the energy of C_n, namely $\mathcal{E}^{\mathrm{bnd}}$ and $\mathcal{E}^{\mathrm{bulk}}$, where

$$\mathcal{E}^{\mathrm{bnd}}(C_n) := \mathcal{E}(C_n) - \mathcal{E}^{\mathrm{bulk}}(C_n).$$

Maximal polygon: We introduce an additional notion in the case that C_n is connected and does not contain acyclic bonds. In this case, the bond graph is delimited by a simple cycle which we call the *maximal polygon*. We denote the atoms of the maximal polygon by $\{x_1, \dots, x_d\}$ and the interior angle at the point $x_i \in \partial X_n$ by θ_i. Moreover, we indicate by

$$I_2 = \{x_i \in \partial X_n : \#\mathcal{N}(x_i) = 2\},$$
$$I_3 = \{x_i \in \partial X_n : \#\mathcal{N}(x_i) = 3\},$$

the set of 2-bonded and 3-bonded boundary atoms, respectively. If C_n is a ground state, we note that $\#I_2 + \#I_3 = d$ by Lemma 3.1.

We now provide an estimate for the boundary energy. Its proof is inspired by [24, Lemma 6.2]. The precise estimates, however, deviate significantly from the study in [24] due to the presence of the repulsive potential V_{r} instead of an angular potential. We defer a discussion in that direction after the proof, see Remark 5.3.

Lemma 5.1 (Boundary energy)**.** *Let $n \geq 6$ and let C_n be a connected ground state with no acyclic bonds. Then*

$$\mathcal{E}^{\mathrm{bnd}}(C_n) \geq -\frac{3}{2}d + 3 \tag{14}$$

with equality only if the following conditions are satisfied:

$$\textit{All boundary bonds are of unit length,} \tag{15}$$

$$\#I_3 = \frac{1}{2}d - 3, \tag{16}$$

$$\theta_i = \frac{4\pi}{3} \ \textit{if } x_i \in I_3 \ \textit{and } \theta_i = \frac{2\pi}{3} \ \textit{if } x_i \in I_2. \tag{17}$$

Remark 5.2. (i) Note by Lemma 3.1 that C_n has alternating charge distribution and thus d is even. (ii) Observe by (16) that equality in (14) implies that $\frac{3}{2}d - 3$ bonds contribute to the boundary energy. Thus, equality in (14) together with [ii] and [vii] imply that for all boundary atoms x_i one has $\min\{|x_i - x_j| : \ j \in \{1, \dots, n\}, j \neq i, \ q_j = q_i\} \geq \sqrt{3}$.

Proof. Suppose that $\{x_1,\dots,x_d\}$ are ordered such that $x_i\in\mathcal{N}(x_{i+1})$, $i=1,\dots,d$. Here and in the following, we use the identification $x_{d+1}=x_1$ and $x_0=x_d$. For a 3-bonded atom x_i, denote by $x_i^b\in X_n\setminus\{x_{i-1},x_{i+1}\}$ the atom that is connected to x_i with the third bond. The boundary energy can be estimated by

$$\begin{aligned}\mathcal{E}^{\rm bnd}(C_n)\geq\sum_{i=1}^d\Big(\frac{1}{2}\big(V_{\rm a}(|x_i-x_{i+1}|)+V_{\rm a}(|x_i-x_{i-1}|)\big)+V_{\rm r}(|x_{i+1}-x_{i-1}|)\Big)\\+\sum_{x_i\in I_3}\Big(V_{\rm a}(|x_i-x_i^b|)+V_{\rm r}(|x_{i-1}-x_i^b|)+V_{\rm r}(|x_{i+1}-x_i^b|)\Big).\end{aligned}\tag{18}$$

Here, we split the interactions into two sums. The first sum represents the energy between successive boundary atoms. The second sum is a lower bound for the energy of 3-bonded atoms that may be bonded to other boundary atoms or to bulk atoms: note that we might double count (negative) attractive interaction if also x_i^b is a boundary atom. However, we never double count (positive) repulsive interaction. Indeed, suppose that $x_{i+k_i}-x_i^b$ and $x_{j+k_j}-x_j^b$ with $i\neq j$, $k_i,k_j\in\{-1,1\}$, represented the same bond, i.e., there exist x^1,x^2 with $\{x_{i+k_i},x_i^b,x_{j+k_j},x_j^b\}=\{x^1,x^2\}$. But then $\{x_i,x_j,x^1,x^2\}$ is a square in the bond graph which contradicts Lemma 3.2.

For a 3-bonded atom x_i, denote by $\theta_i^1,\theta_i^2\in[0,2\pi]$ the two angles forming θ_i enclosed by the three bonds at x_i. Finally, we define $\delta:=\frac{\#I_2+2\#I_3}{d}$ and note that $\delta\in[1,2]$ since $\#I_2+\#I_3=d$.

We will prove that

$$\begin{aligned}\mathcal{E}^{\rm bnd}(C_n)&\geq-\delta d+\sum_{x_i\in I_2}V_{\rm r}\Big(2\sin\Big(\frac{\theta_i}{2}\Big)\Big)+\sum_{x_i\in I_3}\sum_{j=1,2}V_{\rm r}\Big(2\sin\Big(\frac{\theta_i^j}{2}\Big)\Big)\\&\geq-\delta d+\delta dV_{\rm r}\Big(2\sin\Big(\frac{\pi(d-2)}{2\delta d}\Big)\Big),\end{aligned}\tag{19}$$

where the first inequality is strict if not all lengths of boundary bonds are equal to 1. We defer the proof of this estimate and first show that it implies the statement of the lemma.

First, introducing

$$\alpha(\delta):=\frac{\pi(d-2)}{2\delta d},$$

estimate (19) can be written as

$$\mathcal{E}^{\rm bnd}(C_n)\geq-\delta d+\delta dV_{\rm r}\Big(2\sin\Big(\frac{\pi(d-2)}{2\delta d}\Big)\Big)=\delta d\Big(V_{\rm r}\Big(2\sin\Big(\alpha(\delta)\Big)\Big)-1\Big).\tag{20}$$

We obtain (14) by minimizing the right hand side of (20) with respect to δ. To see this, set $\delta_0=\frac{3}{2}-\frac{3}{d}$. For $\delta\leq\delta_0$ we have

$$\alpha(\delta)\geq\alpha(\delta_0)=\frac{\pi}{3}.$$

By [vii] we get $V_{\rm r}(2\sin(\alpha(\delta)))=0$ for all $\delta\leq\delta_0$. Therefore, we find

$$\delta d\Big(V_{\rm r}\Big(2\sin\Big(\alpha(\delta)\Big)\Big)-1\Big)=-\delta d\geq-\delta_0 d=-\Big(\frac{3}{2}d-3\Big)\tag{21}$$

and we obtain estimate (14) for $\delta\leq\delta_0$. Now for $\delta>\delta_0$, we have $\alpha(\delta)<\alpha(\delta_0)$. By (v) we get

$$V_{\rm r}(2\sin(\alpha(\delta)))\geq V_{\rm r}(2\sin(\alpha(\delta_0)))+2V'_{\rm r,-}(2\sin(\alpha(\delta_0))(\sin(\alpha(\delta))-\sin(\alpha(\delta_0)).$$

Then by the fact that $\sin(\theta)$ is concave for $\theta \in [0,\pi]$, $V'_{\mathrm{r},-}(\sqrt{3}) < -3/\pi < 0$ by [viii], $\alpha(\delta_0) = \frac{\pi}{3}$, and $\alpha(\delta) - \alpha(\delta_0) < 0$ we derive

$$\begin{aligned} V_{\mathrm{r}}(2\sin(\alpha(\delta))) &\geq V_{\mathrm{r}}(\sqrt{3}) + 2V'_{\mathrm{r},-}(\sqrt{3})\cos(\alpha(\delta_0))\big(\alpha(\delta) - \alpha(\delta_0)\big) \\ &= V_{\mathrm{r}}(\sqrt{3}) + V'_{\mathrm{r},-}(\sqrt{3})\Big(\frac{\pi(d-2)}{2\delta d} - \frac{\pi}{3}\Big) > -\frac{3}{\pi}\Big(\frac{\pi(d-2)}{2\delta d} - \frac{\pi}{3}\Big) \\ &= \frac{1}{\delta d}\Big(\delta d - \frac{3}{2}d + 3\Big). \end{aligned} \tag{22}$$

Here, we also used that $\cos(\alpha(\delta_0)) = \frac{1}{2}$, $V_{\mathrm{r}}(\sqrt{3}) = 0$, and $V'_{\mathrm{r},-}(\sqrt{3}) < -3/\pi$. From the previous calculation and (20), estimate (14) follows also for $\delta > \delta_0$.

We now show that we have strict inequality in (14) if one of the conditions (15)-(17) is violated. First, if a boundary bond is not of unit length, we have strict inequality in (19) and thus also in (20). If (16) is violated, we find $\delta \neq \delta_0$ after a short computation. Then we obtain strict inequalities from (21) and (22), respectively. Finally, let use suppose that (17) is violated. We can assume that $\delta = \delta_0$ and (15)-(16) hold as otherwise the inequality in (14) is strict. If equality holds in (14), then equality also holds in (19). As $V_{\mathrm{r}}(2\sin(\alpha(\delta_0))) = 0$, this implies

$$\sum_{x_i \in I_2} V_{\mathrm{r}}\Big(2\sin\Big(\frac{\theta_i}{2}\Big)\Big) + \sum_{x_i \in I_3} \Big(V_{\mathrm{r}}\Big(2\sin\Big(\frac{\theta_i^1}{2}\Big)\Big) + V_{\mathrm{r}}\Big(2\sin\Big(\frac{\theta_i^2}{2}\Big)\Big)\Big) = 0.$$

In view of [vii], this gives

$$\theta_i \in [\tfrac{2\pi}{3}, \tfrac{4\pi}{3}] \quad \text{for all } x_i \in I_2, \qquad \theta_i^1, \theta_i^2 \in [\tfrac{2\pi}{3}, \tfrac{4\pi}{3}] \quad \text{for all } x_i \in I_3. \tag{23}$$

Under the assumption that (17) is violated, using (23) and $\theta_i = \theta_i^1 + \theta_i^2$ we find some $x_i \in I_3$ with $\theta_i > \frac{4\pi}{3}$ or some $x_i \in I_2$ with $\theta_i > \frac{2\pi}{3}$. Then (23) implies

$$\pi(d-2) = \frac{4\pi}{3}\Big(\frac{1}{2}d - 3\Big) + \frac{2\pi}{3}\Big(\frac{1}{2}d + 3\Big) = \frac{4\pi}{3}\#I_3 + \frac{2\pi}{3}\#I_2 < \sum_{x_i \in I_2} \theta_i + \sum_{x_i \in I_3} \theta_i = \pi(d-2),$$

where the last step follows from the fact that the maximal polygon has d vertices. This is a contradiction and shows strict inequality in (14) if (17) is violated.

To complete the proof, it remains to show (19). In the case of a 2-bonded x_i, define $r_i^1 = |x_i - x_{i-1}|$, $r_i^2 = |x_i - x_{i+1}|$. In the case of a 3-bonded x_i, define additionally $r_i^3 = |x_i - x_i^b|$. By the cosine rule we obtain

$$|x_{i+1} - x_{i-1}| = \ell(\theta_i, r_i^1, r_i^2), \quad |x_i^b - x_{i-1}| = \ell(\theta_i^1, r_i^1, r_i^3), \quad |x_i^b - x_{i+1}| = \ell(\theta_i^2, r_i^2, r_i^3), \tag{24}$$

where we have used the shorthand

$$\ell(\theta, r_1, r_2) = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta)}. \tag{25}$$

We want to prove that for every boundary atom x_i its contribution to the energy can be controlled by the energy contribution in a modified configuration which has the same angles but unit bond lengths instead of r_i^1, r_i^2. Recall that by [viii] we have for all $r \in (1, r_0]$

$$\frac{1}{2(r-1)}(V_{\mathrm{a}}(1) - V_{\mathrm{a}}(r)) < V'_{\mathrm{r},+}(1). \tag{26}$$

Note that $\ell(\theta, r_1, r_2) \geq 1, V'_{\mathrm{r},+}(r) \leq 0$ for $r \geq 1$ and $V'_{\mathrm{r},+}$ is monotone increasing due to the convexity assumption in (v) on V_{r}. Moreover, we have $\partial_{r_1}\ell(\theta, r_1, r_2) \leq 1$ by an elementary computation. (This can also be seen by a geometric argumentation: by changing the length of

r_1, the length change of ℓ is always smaller or equal to the length change of r_1, equal only if $\theta \in \{0, \pi\}$.) We therefore obtain

$$\frac{1}{2(r-1)}(V_{\rm a}(1) - V_{\rm a}(r)) \le V'_{{\rm r},+}(1) \le V'_{{\rm r},+}(\ell(\theta, s, r_2)) \le V'_{{\rm r},+}(\ell(\theta, s, r_2))\partial_{r_1}\ell(\theta, s, r_2).$$

Integrating this from 1 to r in the variable s and using the fundamental theorem of calculus, we get

$$\frac{1}{2}\left(V_{\rm a}(1) - V_{\rm a}(r)\right) \le \int_1^r V'_{{\rm r},+}(\ell(\theta, s, r_2))\partial_{r_1}\ell(\theta, s, r_2){\rm d}s = V_{\rm r}(\ell(\theta, r, r_2)) - V_{\rm r}(\ell(\theta, 1, r_2)).$$

Applying this estimate in the second as well as in the third component of $\ell(\theta, r_1, r_2)$ with r_1 and r_2 respectively, we derive

$$V_{\rm a}(1) + V_{\rm r}(\ell(\theta, 1, 1)) \le \frac{1}{2}\left(V_{\rm a}(r_1) + V_{\rm a}(r_2)\right) + V_{\rm r}(\ell(\theta, r_1, r_2)). \tag{27}$$

Now for all 2-bonded x_i, using (27) with r_i^1, r_i^2, and θ_i, we have

$$\frac{1}{2}\left(V_{\rm a}(r_i^1) + V_{\rm a}(r_i^2)\right) + V_{\rm r}(\ell(\theta_i, r_i^1, r_i^2)) \ge V_{\rm a}(1) + V_{\rm r}(\ell(\theta_i, 1, 1)). \tag{28}$$

On the other hand, for all 3-bonded x_i, using (27) twice, we have

$$\begin{aligned}
&\frac{1}{2}\left(V_{\rm a}(r_i^1) + V_{\rm a}(r_i^3)\right) + V_{\rm r}(\ell(\theta_i^1, r_i^1, r_i^3)) \ge V_{\rm a}(1) + V_{\rm r}(\ell(\theta_i^1, 1, 1)),\\
&\frac{1}{2}\left(V_{\rm a}(r_i^2) + V_{\rm a}(r_i^3)\right) + V_{\rm r}(\ell(\theta_i^2, r_i^2, r_i^3)) \ge V_{\rm a}(1) + V_{\rm r}(\ell(\theta_i^2, 1, 1)).
\end{aligned} \tag{29}$$

Using (24), (28)-(29), [ii], and $V_{\rm r} \ge 0$ we obtain by (18)

$$\begin{aligned}
\mathcal{E}^{\rm bnd}(C_n) &\ge \sum_{i=1}^d \Big(\frac{1}{2}\big(V_{\rm a}(|x_i - x_{i+1}|) + V_{\rm a}(|x_i - x_{i-1}|)\big) + V_{\rm r}(|x_{i+1} - x_{i-1}|)\Big)\\
&\quad + \sum_{x_i \in I_3}\Big(V_{\rm a}(|x_i - x_i^b|) + V_{\rm r}(|x_{i-1} - x_i^b|) + V_{\rm r}(|x_{i+1} - x_i^b|)\Big)\\
&= \sum_{x_i \in I_2}\Big(\frac{1}{2}\big(V_{\rm a}(r_i^1) + V_{\rm a}(r_i^2)\big) + V_{\rm r}(\ell(\theta_i, r_i^1, r_i^2))\Big) + \sum_{x_i \in I_3} V_{\rm r}(|x_{i+1} - x_{i-1}|)\\
&\quad + \sum_{x_i \in I_3}\Big(\frac{1}{2}\big(V_{\rm a}(r_i^1) + V_{\rm a}(r_i^2) + 2V_{\rm a}(r_i^3)\big) + V_{\rm r}(\ell(\theta_i^1, r_i^1, r_i^3)) + V_{\rm r}(\ell(\theta_i^2, r_i^2, r_i^3))\Big)\\
&\ge -(\#I_2 + 2\#I_3) + \sum_{i \in I_2} V_{\rm r}(\ell(\theta_i, 1, 1)) + \sum_{i \in I_3}\sum_{j=1,2} V_{\rm r}(\ell(\theta_i^j, 1, 1)).
\end{aligned} \tag{30}$$

For later purposes, we remark that this inequality is strict if one bond has not unit length. This follows from the strict inequality in (26).

Recall $\delta = \frac{\#I_2 + 2\#I_3}{d}$ and note that $\ell(\theta, 1, 1) = 2\sin(\theta/2)$ by (25). Using $\theta_i = \theta_i^1 + \theta_i^2$ for $x_i \in I_3$ and (30) we obtain

$$\begin{aligned}
\mathcal{E}^{\rm bnd}(C_n) &\ge -\delta d + \sum_{x_i \in I_2} V_{\rm r}(\ell(\theta_i, 1, 1)) + \sum_{x_i \in I_3}\sum_{j=1,2} V_{\rm r}(\ell(\theta_i^j, 1, 1))\\
&= -\delta d + \sum_{x_i \in I_2} V_{\rm r}\Big(2\sin\Big(\frac{\theta_i}{2}\Big)\Big) + \sum_{x_i \in I_3}\sum_{j=1,2} V_{\rm r}\Big(2\sin\Big(\frac{\theta_i^j}{2}\Big)\Big).
\end{aligned}$$

This yields the first inequality in (19). We note that this inequality is strict if one bond has not unit length since then (30) is strict.

The second inequality in (19) follows by a convexity argument: since $V_{\rm r}$ is convex and non-increasing by [v] and $\sin(\theta/2)$ is concave for $\theta \in [0, 2\pi]$, we have

$$\begin{aligned}\lambda V_{\rm r}\Big(2\sin\Big(\frac{\theta_1}{2}\Big)\Big) + (1-\lambda)V_{\rm r}\Big(2\sin\Big(\frac{\theta_2}{2}\Big)\Big) &\geq V_{\rm r}\Big(\lambda 2\sin\Big(\frac{\theta_1}{2}\Big) + (1-\lambda)2\sin\Big(\frac{\theta_2}{2}\Big)\Big)\\ &\geq V_{\rm r}\Big(2\sin\Big(\frac{\lambda\theta_1 + (1-\lambda)\theta_2}{2}\Big)\Big).\end{aligned}$$

Hence, $\theta \mapsto V_{\rm r}\Big(2\sin\Big(\frac{\theta}{2}\Big)\Big)$ is a convex function for $\theta \in [0, 2\pi]$. This together with the fact that $\#I_2 + 2\#I_3 = \delta d$ and

$$\pi(d-2) = \sum_{i=1}^{d}\theta_i = \sum_{x_i\in I_2}\theta_i + \sum_{x_i\in I_3}(\theta_i^1 + \theta_i^2)$$

yields the second inequality in (19). This concludes the proof. □

Remark 5.3. At this stage, let us highlight the difference of our analysis to [24]. In [24], an empirical angular potential is considered which penalizes deviations of the bond-angles from $\frac{2\pi}{3}$, modeling covalent bonding for carbon nanostructures. In our model for ionic compounds, the energy contribution can also be expressed in terms of the bond-angle. More precisely, by (25) we have energy contributions of the form

$$V_{\rm r}\big(\ell(\theta, r_1, r_2)\big) = V_{\rm r}\Big(\sqrt{r_1^2 + r_2^2 - 2r_1r_2\cos(\theta)}\Big).$$

In view of assumption [vii], only lengths $\ell(\theta, r_1, r_2) < \sqrt{3}$ are penalized. In particular, as $r_1, r_2 \geq 1$, bond-angles $\theta \geq \frac{2\pi}{3}$ never penalize the energy and, if the bond lengths r_1 and r_2 exceed one, also bond-angles less than $\frac{2\pi}{3}$ might not penalize the energy.

In principle, this implies more geometric flexibility of ground-state configurations with respect to [24]. This calls for refined arguments for controlling the boundary energy and, in particular, for characterizing the ground-state geometries in Section 7. Let us highlight that, in spite of the weaker penalization of bond-angles deviating from $\frac{2\pi}{3}$, it is still possible to prove that ground states assemble themselves in the hexagonal lattice.

Recall the excess of edges $\eta = \sum_{j\geq 6}(j-6)f_j$ introduced in (3), where f_j denotes the number of polygons with j vertices in the bond graph. By Lemma 3.2 for any ground state we have that $\eta \in 2\mathbb{N}$. Moreover, $\eta = 0$ if only if the bond graph consists of hexagons only. We now use this notion to estimate the cardinality of the bulk. It will turn out useful in Section 7 to exclude the existence of other polygons than hexagons.

Lemma 5.4 (Cardinality of the bulk)**.** *Suppose that C_n is a connected ground state and that it does not contain any acyclic bonds. Then*

$$n - d = 4b + 6 + \eta - 5n.$$

Proof. By f_j we denote the number of polygons in the bond graph with j vertices and set $f = \sum_{j\geq 3} f_j$. We have

$$\sum_{j\geq 3} jf_j = 2b - d,$$

since by the summation on the left all bonds contained in the maximal polygon are counted only once whereas all other bonds are counted twice. From Lemma 3.2 and the definition of η we obtain

$$6f = 2b - d - \eta.$$

Using this together with Euler's formula $n - b + f = 1$ (omitting the exterior face) we get

$$n - d = 4b + 6 + \eta - 5n.$$

□

6. Characterization of the ground-state energy

This section is devoted to the proof of Theorem 2.1. We only need to provide a lower bound on the ground-state energy since the upper bound has already been obtained by an explicit construction, see Proposition 4.1.

We state two algebraic lemmas that will be used in the sequel.

Lemma 6.1. *Let $j, n, m \in \mathbb{N}$ and let $x \in \mathbb{R}$ satisfy*

$$\frac{m}{4} - \frac{5}{4}n \geq x \geq -\frac{3}{2}n + j + \sqrt{\frac{3}{2}(-4x - 5n + m)}.$$

Then $x \geq -\frac{3}{2}n + j - 3 + \sqrt{\frac{3}{2}(-4j + m + n + 6)}$.

Proof. The proof is elementary: we note that the function

$$x \mapsto x + \frac{3}{2}n - j - \sqrt{\frac{3}{2}(-4x - 5n + m)}$$

is strictly increasing and vanishes for $x = -\frac{3}{2}n + j - 3 + \sqrt{\frac{3}{2}(-4j + m + n + 6)}$. □

We use the following properties of the function β which has been defined in (7).

Lemma 6.2. *The function $\beta : \mathbb{N} \to \mathbb{R}$ satisfies*

1) $\lfloor \beta(n-1) \rfloor + 1 \leq \lfloor \beta(n) \rfloor$.
2) $\lfloor \beta(m) \rfloor + \lfloor \beta(n-m) \rfloor + 1 \leq \lfloor \beta(n) \rfloor$ *for all $n \geq 12$, $n \geq m \geq 6$ and equality holds if and only if $n = 12$ and $m = 6$.*
3) $\lfloor \beta(n) \rfloor \geq \lfloor \beta(n-k) \rfloor + 2 + k$ *for all $n \geq 13$ and $n \geq k \geq 6$.*
4) $\lfloor \beta(n) \rfloor \geq \lfloor \beta(n-5) \rfloor + 7$ *for all $n \geq 13$ except for $n \in \{15, 18, 21, 29\}$.*

Proof. The proof of 1) and 2) is elementary and can be found in [24, Lemma 6.4, 6.5]. It relies on monotonicity and convexity properties of β. As a preparation for 3) and 4), we observe that for $n \geq 41$, $k = 5$ or for $n \geq 17$, $k = 6$ or for $n \geq 13$, $n \geq k \geq 7$ one has

$$\frac{3}{2}n - \sqrt{\frac{3}{2}n} \geq \frac{3}{2}(n-k) - \sqrt{\frac{3}{2}(n-k)} + 2 + k. \tag{31}$$

Indeed, after some manipulations, we see that this is equivalent to $\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n-k)} \geq 3k/(k-4)$. The latter holds true for $n \geq 41$, $k = 5$ or for $n \geq 17$, $k = 6$ or for $n \geq 13$, $n \geq k \geq 7$.

One can check directly that $\lfloor\beta(n)\rfloor = \lfloor\beta(n-6)\rfloor + 8$ for $n = 13, \dots, 16$, see Table 1. This together with (31) yields 3). Property 4) follows from (31) and an explicit computation for the cases $13 \leq n \leq 40$, see Table 2. □

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\lfloor\beta(n)\rfloor$	1	2	3	4	6	7	8	9	11	12	13	15	16	17
n	16	17	18	19	20	21	22	23	24	25	26	27	28	29
$\lfloor\beta(n)\rfloor$	19	20	21	23	24	25	27	28	30	31	32	34	35	36

TABLE 1. The function $\lfloor\beta(n)\rfloor$ for $2 \leq n \leq 29$. The table can be used to see that the configurations in Fig. 8, Fig. 9, and Fig. 12 are ground states.

n	13	14	15	16	17	18	19	20	21	22	23	24	25	26
$\gamma(n)$	7	7	6	7	7	6	7	7	6	7	7	7	7	7
n	27	28	29	30	31	32	33	34	35	36	37	38	39	40
$\gamma(n)$	7	7	6	7	7	7	7	7	7	7	7	7	7	7

TABLE 2. The function $\gamma(n) := \lfloor\beta(n)\rfloor - \lfloor\beta(n-5)\rfloor$ for $13 \leq n \leq 40$.

Before we proceed with the proof of Theorem 2.1, we will consider the cases $1 \leq n \leq 6$ which will serve as the induction base.

Lemma 6.3 (Cases $1 \leq n \leq 6$)**.** *For $1 \leq n \leq 6$ every ground state C_n is connected and satisfies $\mathcal{E}(C_n) = -b = \lfloor\beta(n)\rfloor$.*

Proof. Let $n \leq 5$. By Lemma 3.2 we have that the bond graph does not contain any polygons with less than or equal to 5 edges. Hence, the bond graph is cycle free and $b \leq n-1$. This provides the lower bound $\mathcal{E}(C_n) \geq -(n-1)$ by Remark 2.2. Clearly, one can construct configurations with n atoms and $n-1$ bonds having alternating charge distribution. Note also that $b = n-1 = -\lfloor\beta(n)\rfloor$ is only possible if the configuration is connected.

In the case $n = 6$, note that the number of polygons f satisfies $f \leq 1$ since by Lemma 3.2 we have that every polygon has at least 6 edges. By Euler's formula we get $6 - b + f \geq 1$ where the inequality is due to the fact that we may have more than one connected component. This implies $b \leq 6$ and thus $\mathcal{E}(C_6) \geq -b$ by Remark 2.2. Exactly a regular hexagon with alternating charge distribution and unit bond length yields a configuration with energy equal to $-6 = -\lfloor\beta(6)\rfloor$. This concludes the proof. □

We now proceed with the proof of Theorem 2.1. We follow the strategy devised in [24, Theorem 6.1] with the adaptions needed due to the presence of different atomic types and repulsive potentials instead of angle potentials. In contrast to [24], however, we split the proof of the characterization of the ground-state energy and the characterization of the ground states. Indeed, the latter is more involved in our setting and the investigation is deferred to Section 7.

Proof of Theorem 2.1. We start by noting that every ground state C_n has alternating charge distribution by Lemma 3.1. By Proposition 4.1 we have that the ground-state energy satisfies

$$\mathcal{E}(C_n) \leq -\lfloor \beta(n) \rfloor. \tag{32}$$

We proceed by induction. Suppose that the statement has been proven for all $m < n$ (for $1 \leq m \leq 6$ see Lemma 6.3). We first show connectedness of the ground state (Claim 1) and then the energy equality (Claim 2).

Claim 1: C_n is connected.

Proof of Claim 1: Assume by contradiction that C_n was not connected, i.e., C_n consists of two or more connected components. Let C'_m and C'_{n-m} be two sub-configurations consisting of m and $n-m$ atoms, respectively, which do not have any bonds between them. If $m \geq 6, n \geq 12$, we can apply the induction hypothesis, Lemma 6.2 2), and $V_{\rm r} \geq 0$ to get

$$\mathcal{E}(C_n) \geq \mathcal{E}(C'_m) + \mathcal{E}(C'_{n-m}) \geq -\lfloor \beta(n-m) \rfloor - \lfloor \beta(m) \rfloor > -\lfloor \beta(n) \rfloor.$$

If $m < 6$, we can apply Lemma 6.3 and Lemma 6.2 1) iteratively m times to get

$$\mathcal{E}(C_n) \geq -\lfloor \beta(n-m) \rfloor - \lfloor \beta(m) \rfloor = -\lfloor \beta(n-m) \rfloor - (m-1) > -\lfloor \beta(n) \rfloor.$$

The case $n < 12$ is already included in this argument since at least one connected component consists of less than 6 atoms. In view of (32), we obtain a contradiction to the fact that C_n is a ground state.

Claim 2: Energy equality $\mathcal{E}(C_n) = -b = -\lfloor \beta(n) \rfloor$.

Proof of Claim 2: We divide the proof into three steps. We first treat the case that C_n contains acyclic bonds (Step 1). Afterwards, we consider only configurations C_n without acyclic bonds and show $\mathcal{E}(C_n) = -b$ (Step 2) and $\mathcal{E}(C_n) = -\lfloor \beta(n) \rfloor$ (Step 3).

Step 1: C_n contains acyclic bonds: If there exist flags, we can find an atom x_i such that removing x_i removes exactly one flag. We can count the energy contribution of this flag by at least -1 and we estimate the energy of the rest of the configuration by induction. By Lemma 6.2 1) we get

$$\mathcal{E}(C_n) \geq -1 + \mathcal{E}(C_n \setminus \{(x_i, q_i)\}) \geq -1 - \lfloor \beta(n-1) \rfloor \geq -\lfloor \beta(n) \rfloor.$$

Equality also shows that $C_n \setminus \{(x_i, q_i)\}$ has $\lfloor \beta(n-1) \rfloor$ bonds by induction and C_n has $\lfloor \beta(n-1) \rfloor + 1 = \lfloor \beta(n) \rfloor$ bonds.

We now suppose that a bridge exists. Consider the two sub-configurations C'_m and C'_{n-m} which are connected by the bridge. By the definition of bridges we have that both C'_m and C'_{n-m} contain at least one simple cycle and therefore, by Lemma 3.2, $m, n-m \geq 6$. The energy contribution of the bridge is greater or equal to -1. Using the induction assumption and Lemma 6.2 2) we get

$$\mathcal{E}(C_n) \geq \mathcal{E}(C'_m) + \mathcal{E}(C'_{n-m}) - 1 \geq -\lfloor \beta(m) \rfloor - \lfloor \beta(n-m) \rfloor - 1 \geq -\lfloor \beta(n) \rfloor.$$

As before, equality also implies that C_n has $\lfloor \beta(m) \rfloor + \lfloor \beta(n-m) \rfloor + 1 = \lfloor \beta(n) \rfloor$ bonds.

Step 2: $\mathcal{E}(C_n) = -b$ for connected C_n with no acyclic bonds: Suppose by contradiction that the statement was false. Then there exist $x_1, x_2 \in X_n$ such that $q_1 = q_2$ and $|x_1 - x_2| < \sqrt{3}$ or $q_1 = -q_2$ and $1 < |x_1 - x_2| \leq r_0$. If $x_1 \in \partial X_n$ or $x_2 \in \partial X_n$, by using Lemma 5.1 and Remark 5.2(ii) we have the strict inequality

$$\mathcal{E}^{\rm bnd}(C_n) > -\frac{3}{2}d + 3$$

and by induction assumption we have

$$\mathcal{E}^{\mathrm{bulk}}(C_n) \geq -\lfloor \beta(n-d) \rfloor.$$

On the other hand, if we have $x_1, x_2 \notin \partial X_n$, we calculate by Lemma 5.1 and induction

$$\mathcal{E}^{\mathrm{bnd}}(C_n) \geq -\frac{3}{2}d + 3, \qquad \mathcal{E}^{\mathrm{bulk}}(C_n) > -\lfloor \beta(n-d) \rfloor.$$

Here, the estimate for the bulk part is strict. Indeed, since in this case C_n^{bulk}, which consists of $n-d$ particles, is not repulsion-free or has bonds longer than 1, by the induction assumption it cannot be a ground state, see Remark 2.2. In both cases it holds that

$$\mathcal{E}(C_n) > -\left\lfloor \frac{3}{2}n - \sqrt{\frac{3}{2}(n-d)} \right\rfloor + 3.$$

Since the right hand side is an integer, we obtain

$$-(\lfloor -\mathcal{E}(C_n) \rfloor + 1) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n-d)} + 3. \tag{33}$$

Recall that we assumed $\mathcal{E}(C_n) > -b$ by contradiction, which implies $-(\lfloor -\mathcal{E}(C_n) \rfloor + 1) \geq -b$. Now by Lemma 5.4 we obtain

$$n - d \geq 4(\lfloor -\mathcal{E}(C_n) \rfloor + 1) + 6 - 5n,$$

where we used that $\eta \geq 0$. Using the above inequality and (33) we obtain

$$-(\lfloor -\mathcal{E}(C_n) \rfloor + 1) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(4(\lfloor -\mathcal{E}(C_n) \rfloor + 1) + 6 - 5n)} + 3.$$

Now we can use Lemma 6.1 with $j = 3$, $m = 6$, and $x = -(\lfloor -\mathcal{E}(C_n) \rfloor + 1)$ to obtain

$$-(\lfloor -\mathcal{E}(C_n) \rfloor + 1) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}n}.$$

The last inequality implies $\mathcal{E}(C_n) > -\lfloor \beta(n) \rfloor$ contradicting (32).

Step 3: $\mathcal{E}(C_n) = -\lfloor \beta(n) \rfloor$ for connected C_n with no acyclic bonds: Due to (32), it suffices to prove $\mathcal{E}(C_n) \geq -\lfloor \beta(n) \rfloor$. Again we proceed by induction. By Lemma 5.4 and the induction assumption we obtain

$$\mathcal{E}^{\mathrm{bnd}}(C_n) \geq -\frac{3}{2}d + 3, \quad \mathcal{E}^{\mathrm{bulk}}(C_n) \geq -\lfloor \beta(n-d) \rfloor.$$

This gives

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n-d)} + 3.$$

By Lemma 5.4 and Step 2 we obtain $n - d \geq -4\mathcal{E}(C_n) + 6 - 5n$. This yields

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(-4\mathcal{E}(C_n) - 5n + 6)} + 3.$$

Applying Lemma 6.1 with $j = 3$, $m = 6$, and $x = \mathcal{E}(C_n)$ we obtain $\mathcal{E}(C_n) \geq -\beta(n)$. Finally, since $\mathcal{E}(C_n)$ is an integer due to Step 2, we conclude $\mathcal{E}(C_n) \geq -\lfloor \beta(n) \rfloor$. □

For later purposes, we observe that the calculations of Step 2 and Step 3 in the previous proof can be refined.

Lemma 6.4 (Refined energy inequality for η)**.** *Let $n \geq 6$ and let C_n be a ground state with no acyclic bonds. Then*

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n+\eta)}.$$

If $\mathcal{E}^{\rm bnd}(C_n) > -\frac{3}{2}d + 3$ or $\mathcal{E}^{\rm bulk}(C_n) > \lfloor \beta(n-d) \rfloor$, then

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n+\eta-4)} + 1.$$

Proof. By Lemma 5.1 and Theorem 2.1 applied on $C_n^{\rm bulk}$ we have

$$\mathcal{E}^{\rm bnd}(C_n) \geq -\frac{3}{2}d + 3, \qquad \mathcal{E}^{\rm bulk}(C_n) \geq -\lfloor \beta(n-d) \rfloor. \tag{34}$$

By Lemma 5.4 and Theorem 2.1 we get $n - d = -4\mathcal{E}(C_n) + 6 + \eta - 5n$. This together with the previous estimate yields

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(-4\mathcal{E}(C_n) - 5n + 6 + \eta)} + 3. \tag{35}$$

By applying Lemma 6.1 with $j = 3$ and $m = 6 + \eta$ we obtain

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n+\eta)}.$$

Finally, if $\mathcal{E}^{\rm bnd}(C_n) > -\frac{3}{2}d + 3$ or $\mathcal{E}^{\rm bulk}(C_n) > \lfloor \beta(n-d) \rfloor$, i.e., one inequality in (34) is strict, we can replace 3 by 4 in (35) since $\mathcal{E}(C_n)$ is an integer. Then applying again Lemma 6.1 with $j = 4$, $m = 6 + \eta$ we obtain $\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n+\eta-4)} + 1$. □

7. Characterization of ground states

In this section we characterize the ground states of (1). We do not provide a complete characterization for $n < 10$ since the system is highly flexible in those cases. Some of the ground states for $n < 10$ are pictured in Fig. 8. We will start by providing some geometric facts about ground states. Afterwards, we formulate and prove the first main result of the section which shows that ground states consist of hexagonal cycles except for possibly (at most) two flags or one octagon at the boundary, see Proposition 7.9. Finally, for $n \geq 30$, we will be able to prove that no octagons occur which will conclude the proof of Theorem 2.3. The proof will also show that, among the ground states for $10 \leq n \leq 29$, an octagon can only occur in the cases $n = 12, 15, 18, 21, 29$, see Remark 7.11 and Fig. 12.

Remark 7.1. We briefly remark that in the following our strategy deviates considerably from the one in [24] due to the different modeling assumptions concerning repulsive and angular potentials, see Remark 5.3 for details. On the one hand, for $n \leq 29$ indeed more flexible structures may occur which are not subsets of the hexagonal lattice. On the other hand, although we eventually will prove that for $n \geq 30$ ground states essentially assemble themselves in the hexagonal lattice, we need a different approach compared to [19, 23, 24, 26]: differently to the proof by induction performed there, we cannot use the property that ground states are subsets of the hexagonal lattice in the induction hypothesis (consider, e.g., the step from 29 to 30). Therefore, finer geometric considerations are necessary which are developed in two steps: first, we prove by induction that the breaking of the hexagonal symmetry due to presence of non-hexagonal polygons can only occur on the boundary. Then we show that for $n \geq 30$ the existence of such defects leads to an energy exceeding (8).

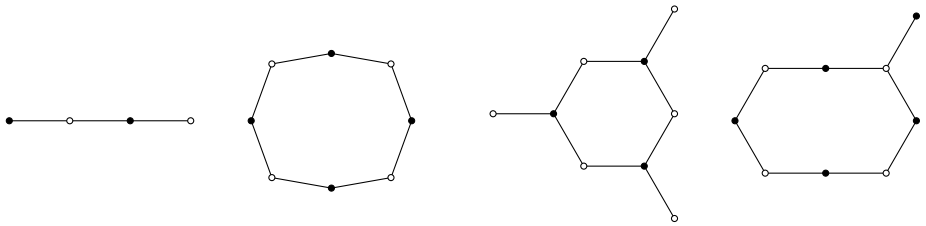

FIGURE 8. Some ground states for $n < 10$.

7.1. **Geometric properties of ground states.** In this section we collect some geometric properties of ground states. We start with an elementary property.

Lemma 7.2 (Bridges)**.** *Ground-states for $n \geq 13$ do not contain bridges.*

Proof. Suppose that a bridge exists. Consider the two sub-configurations C'_m and C'_{n-m} which are connected by the bridge. As the energy contribution of the bridge is greater or equal to -1, we get by Theorem 2.1, Lemma 6.2 2), the fact that $n \geq 13$, and $V_{\mathrm{r}} \geq 0$

$$\mathcal{E}(C_n) \geq -\lfloor \beta(m) \rfloor - \lfloor \beta(n-m) \rfloor - 1 > -\lfloor \beta(n) \rfloor.$$

This contradicts the fact that C_n is a ground state. □

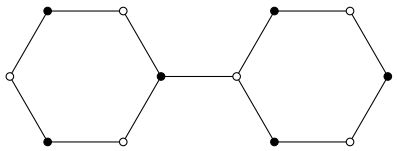

FIGURE 9. A ground state configuration for $n = 12$ containing a bridge.

Remark 7.3. Lemma 7.2 is sharp in the sense that for $n = 12$ there exists a ground state that contains a bridge connecting two hexagons (and then $\eta = 0$), cf. Table 1 and Figure 9. Also note that for $n \leq 11$ ground states cannot contain bridges as each polygon in the bond graph has at least 6 vertices by Lemma 3.2.

The next lemma states that the number of flags is at most two. Let us mention that this property also applies to the ground states of [24] although this has not been observed explicitly there.

Lemma 7.4 (Flags)**.** *Let $n \geq 10$ and let C_n be a ground state. Then the bond graph of C_n contains at most 2 flags.*

Proof. Assume by contradiction that there exist $j \geq 3$ flags. Using the fact that a flag contributes at least -1 to the energy and applying Theorem 2.1 on the sub-configuration obtained after removing the flags, we have $\mathcal{E}(C_n) \geq -j - \lfloor \beta(n-j) \rfloor$. By (7) we then get

$$\begin{aligned}\mathcal{E}(C_n) &\geq -\Big(\frac{3}{2}n - \frac{1}{2}j - \sqrt{\frac{3}{2}(n-j)}\Big) = -\Big(\frac{3}{2}n - \sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}n} - \frac{1}{2}j - \sqrt{\frac{3}{2}(n-j)}\Big) \\ &= -\Big(\frac{3}{2}n - \sqrt{\frac{3}{2}n} - \frac{1}{2}j + \frac{\frac{3}{2}j}{\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n-j)}}\Big).\end{aligned}$$

The function $f(j) := -\frac{1}{2}j + \frac{\frac{3}{2}j}{\sqrt{\frac{3}{2}n}+\sqrt{\frac{3}{2}(n-j)}}$ is non-increasing in j, non-positive, and we have that $f(3) \leq -1$ for $n \geq 16$. With the above estimate this implies

$$\mathcal{E}(C_n) > -\lfloor \beta(n) \rfloor$$

which leads to a contradiction to Theorem 2.1 in the cases $n \geq 16$. The cases $10 \leq n \leq 15$ can be checked directly by comparing the above formula $\mathcal{E}(C_n) \geq -j - \lfloor \beta(n-j) \rfloor, j \geq 3$, with $\lfloor \beta(n) \rfloor, n = 10, \dots, 15$, cf. Table 1. □

Lemma 7.4 is sharp in the sense that for $n = 9$ there exists a ground state that contains three flags in its bond graph, cf. Table 1 and the third configuration in Fig. 8.

Equilibrated atoms: We say an atom $x \in X_n$ is *equilibrated* if all bond-angles at x lie in $\{\frac{2\pi}{3}, \frac{4\pi}{3}\}$. By $\mathcal{A}$ we denote the atoms which are *not* equilibrated. By $\mathcal{A}_{\rm bulk} \subset \mathcal{A}$ we denote the bulk atoms which are not equilibrated. Note that if $\mathcal{A} = \emptyset$ and X_n is connected, then X_n is a subset of the hexagonal lattice $\mathcal{L}$. The following properties will be useful in the sequel.

Lemma 7.5 (Regular hexagons and bond-angles)**.** *Let C_n be a ground state. Then all hexagons are regular with unit bond length and alternating charge. All bond-angles θ satisfy $\frac{2\pi}{3} \leq \theta \leq \frac{4\pi}{3}$. If $x \in \mathcal{A}$, then x is 2-bonded and the bond angles lie in $(\frac{2\pi}{3}, \frac{4\pi}{3})$.*

Proof. By Theorem 2.1 and Remark 2.2 we have that all bonds in the bond graph are of unit length and that the configuration is repulsion-free. An additional necessary condition for equality is that all bond-angles θ satisfy

$$\frac{2\pi}{3} \leq \theta \leq \frac{4\pi}{3}. \tag{36}$$

In fact, suppose that x_1, x_0, x_2 form the angle θ. Since x_1, x_2 are neighbors of x_0, we have $q_1 = q_2$ by Lemma 3.1. The above mentioned necessary conditions imply $|x_1 - x_0| = |x_2 - x_0| = 1$ and $V_{\rm r}(|x_1 - x_2|) = 0$. The latter only holds if $|x_1 - x_2| \geq \sqrt{3}$ by (vii). Simple trigonometry then yields $\frac{2\pi}{3} \leq \theta \leq \frac{4\pi}{3}$

From this discussion we derive that the edges of each hexagon necessarily need to have length 1 and the interior angles are larger or equal to $2\pi/3$. As the sum of the interior angles in a planar hexagon sums to 4π, we get that each interior angle is $2\pi/3$, i.e., each hexagon is indeed a regular hexagon with unit bond length. The charge is alternating by Lemma 3.1.

Consider an atom x_i with a bond-angle θ_1 which satisfies $\frac{4\pi}{3} > \theta_1 > \frac{2\pi}{3}$. Suppose by contradiction that x_i had more than two bonds (i.e., three bonds, see Lemma 3.1). Summing up all the three bond-angles $\theta_1, \theta_2, \theta_3$ at x_i we get that $\min\{\theta_2, \theta_3\} < 2\pi/3$. This, however, contradicts (36). □

We observe that octagons contain non-equilibrated atoms. More precisely, we have the following statement.

Lemma 7.6 (Octagon)**.** *Let C_n be a ground state containing an octagon $\{x_0, \dots, x_7\}$ in the bond graph. Let θ_i, $i = 0, \dots, 7$, be the interior angles of the octagon. Then we have*

$$(i)\quad (\theta_i, \theta_{i+1}, \theta_{i+2}, \theta_{i+3}) \neq \left(\frac{2\pi}{3}, \frac{2\pi}{3}, \frac{2\pi}{3}, \frac{2\pi}{3}\right) \text{ for all } i = 0, \dots, 7,$$

$$(ii)\quad \theta_i < \frac{4\pi}{3} \text{ for all } i = 0, \dots, 7.$$

Here, the indices have to be understood mod 8.

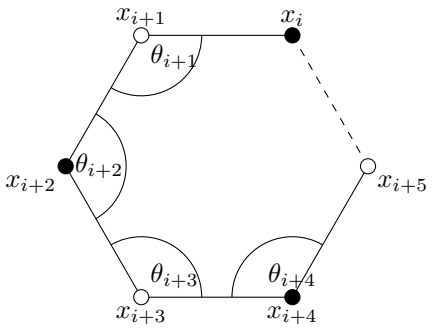

FIGURE 10. Part of an octagon with 4 consecutive angles $\theta_{i+1}, \dots, \theta_{i+4}$ equal to $\frac{2\pi}{3}$.

Proof. By Remark 2.2 and Lemma 7.5 we have that

$$\frac{2\pi}{3} \leq \theta_i \leq \frac{4\pi}{3} \text{ and } |x_{i+1} - x_i| = 1 \text{ for all } i = 0, \dots, 7. \tag{37}$$

Suppose by contradiction that (i) was wrong. If there are more than three consecutive angles of size $\frac{2\pi}{3}$, as indicated in Fig. 10, then the bond graph would contain an additional hexagon and at least one triangle or square. This is a contradiction to the fact that the bond graph contains an octagon.

We now show (ii). Assume by contradiction that, without restriction, $\theta_0 = \frac{4\pi}{3}$. Then by (37) and the fact that the interior angles of the octagon sum to 6π, there holds for all $i = 1, \dots, 7$

$$\theta_i + \frac{4\pi}{3} + 6\frac{2\pi}{3} \leq \sum_{j=0}^{7} \theta_j = 6\pi.$$

This implies $\theta_i \leq \frac{2\pi}{3}$ for all $i = 1, \dots, 7$. Again using (37) this yields $\theta_i = \frac{2\pi}{3}$ for all $i = 1, \dots, 7$ which contradicts (i). □

The following lemma investigates the properties of a configuration in which a non-equilibrated bulk atom is present. Roughly speaking, it states that the existence of such an atom induces the existence of more non-equilibrated atoms and a certain excess of edges η. Note that at this point our analysis deviates significantly from [24]: in a model with angle potentials favoring $\frac{2\pi}{3}$ angles, it is obvious that non-equilibrated atoms cannot exist in ground states.

Lemma 7.7 (Non-equilibrated atoms)**.** *Let C_n be a ground state with no acyclic bonds in the bond graph. Assume that $\#\mathcal{A}_{\text{bulk}} \geq 1$. Then one of the following holds:*

i) $\#\mathcal{A} \geq 2$ *and* $\eta \geq 6$,
ii) $\#(\mathcal{A} \setminus \mathcal{A}_{\text{bulk}}) \geq 1$ *and* $\eta = 4$.

Proof. We first prove that each polygon containing non-equilibrated atoms has at least eight vertices and contains at least two non-equilibrated atoms. Then we show that $\eta \geq 4$ and that in the case $\eta < 6$, we have $\#(\mathcal{A} \setminus \mathcal{A}_{\text{bulk}}) \geq 1$. The statement clearly follows from these claims.

Claim 1: Each polygon containing a non-equilibrated atom has at least eight vertices.

Proof of Claim 1: Due to Lemma 7.5, both angles at non-equilibrated atoms lie in $(\frac{2\pi}{3}, \frac{4\pi}{3})$ and therefore each polygon containing a non-equilibrated atom is not a hexagon. Thus, it has at least eight vertices by Lemma 3.2.

Claim 2: Each polygon contains either no or at least two non-equilibrated atoms.

Proof of Claim 2: Consider a polygon with k edges which contains a non-equilibrated atom with interior angle $\theta_1 \neq \frac{2\pi}{3}, \frac{4\pi}{3}$. By Lemma 3.2 we have that $k \in 2\mathbb{N}$. Assume by contradiction that all the other angles $\theta_2, \dots, \theta_k$ are either $\frac{2\pi}{3}$ or $\frac{4\pi}{3}$. We have

$$\sum_{j=1}^{k} \theta_j = \pi(k-2).$$

Since we assumed that θ_j, $j \geq 2$, are integer multiples of $\frac{2\pi}{3}$, we have

$$\theta_1 + k' \frac{2\pi}{3} = \pi(k-2),$$

where $k' \in \mathbb{N}$ is given by $k' = \#\{j : \theta_j = \frac{2\pi}{3}\} + 2\#\{j : \theta_j = \frac{4\pi}{3}\}$. This implies

$$k = \frac{\theta_1}{\pi} + 2 + \frac{2}{3}k'.$$

Since both $k, k' \in \mathbb{N}$ and $\frac{2\pi}{3} < \theta_1 < \frac{4\pi}{3}$ by Lemma 7.5, there exists only a solution to the equation if $\theta_1 = \pi$ and $k' \in 3\mathbb{N}$. This implies that k is odd: a contradiction.

Claim 3: We have $\eta \geq 4$. If $\eta < 6$, then $\#(\mathcal{A} \setminus \mathcal{A}_{\rm bulk}) \geq 1$.

Proof of Claim 3: Since $\#\mathcal{A}_{\rm bulk} \geq 1$, there exists a non-equilibrated bulk atom. As C_n does not have acyclic bonds, we observe that this atom is a vertex of at least two polygons. Claim 1 then yields that there have to be at least two polygons with at least eight vertices, i.e., $\eta \geq 4$, $\eta \in 2\mathbb{N}$.

It remains to show that, if $\eta = 4$, then $\#(\mathcal{A} \setminus \mathcal{A}_{\rm bulk}) \geq 1$. Assume by contradiction that $\eta = 4$ and $\mathcal{A} \setminus \mathcal{A}_{\rm bulk} = \emptyset$. As $\eta = 4$, the two non-hexagons have to be octagons. By Claim 1, the assumption $\mathcal{A} \setminus \mathcal{A}_{\rm bulk} = \emptyset$, and the fact that C_n has no acyclic bonds, we find that all non-equilibrated atoms are contained in both octagons.

Denote the interior angles in the first octagon different from $\{\frac{2\pi}{3}, \frac{4\pi}{3}\}$ by $\alpha_1, \dots, \alpha_k$, where $1 \leq k \leq 8$. Similarly, the angles in the second octagon different from $\{\frac{2\pi}{3}, \frac{4\pi}{3}\}$ are denoted by $\beta_1, \dots, \beta_k$, where without restriction α_i and β_i lie at the same atom. Note that $\beta_i = 2\pi - \alpha_i$ for $i = 1, \dots, k$ as non-equilibrated atoms are 2-bonded. Due to Lemma 7.6(ii), all other interior angles of the octagons are $\frac{2\pi}{3}$. Thus, by the interior angle sum of the octagons we obtain

$$\sum_{j=1}^{k} \alpha_j + (8-k)\frac{2\pi}{3} = 6\pi, \quad \sum_{j=1}^{k} \beta_j + (8-k)\frac{2\pi}{3} = 6\pi.$$

Using $\sum_{j=1}^{k} \beta_j = 2\pi k - \sum_{j=1}^{k} \alpha_j$ and summing the two equations, we obtain the unique solution $k = 2$. In particular, this implies $\#\mathcal{A} = 2$. Denote the two atoms in $\mathcal{A}$ by x_1 and x_2. From Lemma 7.6(i) we get that the atoms x_1 and x_2 lie 'on opposite sides' of the octagons, i.e., the shortest path in the bond graph connecting x_1 and x_2 has length 4. Then we also see that $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = \pi$ by a simple geometric argument. (An octagon with this geometry is depicted in the rightmost configuration in Fig. 8.) Finally, this yields that the two octagons are identical up to an isometry. This, however, contradicts the fact that both non-equilibrated atoms x_1 and x_2 are contained in both octagons. Thus, $\eta = 4$ and $\mathcal{A} \setminus \mathcal{A}_{\rm bulk} = \emptyset$ is not possible. This concludes the proof. □

Based on Lemma 7.7, we now show that non-equilibrated bulk atoms cannot exist in ground states with no acyclic bonds.

Lemma 7.8 (Non-equilibrated bulk atoms)**.** *Let $n \geq 1$ and let C_n be a ground state with no acyclic bonds. Then $\mathcal{A}_{\text{bulk}} = \emptyset$.*

Proof. We prove the statement by induction. We first note that the statement is true for $1 \leq n \leq 9$. In fact, in this case the bond graph contains at most one polygon by Lemma 3.2 and Lemma 7.5. This implies $\mathcal{A}_{\text{bulk}} \subset X_n^{\text{bulk}} = \emptyset$. Let $n \geq 10$. We assume that the result has been proven for $m < n$ and proceed to show the statement for n. Assume by contradiction that $\mathcal{A}_{\text{bulk}} \neq \emptyset$. By Lemma 7.7 there are two cases to consider:

i) $\#(\mathcal{A} \setminus \mathcal{A}_{\text{bulk}}) \geq 1$ and $\eta = 4$.
ii) $\#\mathcal{A} \geq 2$ and $\eta \geq 6$.

Proof for Case i)*:* By Lemma 5.1 we obtain the strict inequality

$$\mathcal{E}^{\text{bnd}}(C_n) > -\frac{3}{2}d + 3.$$

By Lemma 6.4 (with strict inequality) this gives

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n + \eta - 4)} + 1. \tag{38}$$

Note that $\eta = 4$ and therefore we have

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}n} + 1 > -\lfloor \beta(n) \rfloor \tag{39}$$

which contradicts Theorem 2.1 and the fact that C_n is a ground state.

Proof for Case ii): We are now in the case that $\#\mathcal{A} \geq 2$ and $\eta \geq 6$. By the previous case we can assume that $\mathcal{A} \setminus \mathcal{A}_{\text{bulk}} = \emptyset$. This implies $\#\mathcal{A}_{\text{bulk}} \geq 2$. After removing the boundary, we can suppose that

$$\mathcal{E}(C_n^{\text{bulk}}) = -\lfloor \beta(n - d) \rfloor. \tag{40}$$

Indeed, if $\mathcal{E}(C_n^{\text{bulk}}) > -\lfloor \beta(n-d) \rfloor$, we derive by Lemma 6.4 (with strict inequality) that (38) holds. Since $\eta \geq 4$, we get a contradiction exactly as in Case i), see (39). Hence, C_n as well as C_n^{bulk} are ground states. We now distinguish the following cases:

a) C_n^{bulk} contains at least two flags,
b) C_n^{bulk} contains a bridge,
c) C_n^{bulk} contains less than two flags and no bridge.

Proof for Case a): We use Lemma 5.1 to obtain

$$\mathcal{E}^{\text{bnd}}(C_n) \geq -\frac{3}{2}d + 3. \tag{41}$$

Using the fact that a flag contributes at least -1 to the energy and applying Theorem 2.1 on the sub-configuration obtained after removing exactly two flags from C_n^{bulk}, we obtain

$$\mathcal{E}^{\text{bulk}}(C_n) = \mathcal{E}(C_n^{\text{bulk}}) \geq -2 - \lfloor \beta(n - d - 2) \rfloor \geq -\frac{3}{2}(n-d) + \sqrt{\frac{3}{2}(n - d - 2)} + 1. \tag{42}$$

Combining (41)-(42) we obtain

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n - d - 2)} + 4.$$

By Lemma 5.4 and $b = -\mathcal{E}(C_n)$ we obtain $\mathcal{E}(C_n) \geq -\frac{3}{2}n + 4 + \sqrt{\frac{3}{2}(-4\mathcal{E}(C_n) + 4 + \eta - 5n)}$. Lemma 6.1 for $j = 4$, $m = 4 + \eta$, and $x = \mathcal{E}(C_n)$ yields

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n + 10 + \eta - 16)} + 1 \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}n} + 1 > -\lfloor \beta(n) \rfloor,$$

where we used $\eta \geq 6$. This contradicts Theorem 2.1 and the fact that C_n is a ground state.

Proof for Case b): In view of Remark 7.3, C_n^{bulk} can only contain a bridge if $n - d = 12$ and C_n^{bulk} consists of two regular hexagons connected with a bridge. This contradicts $\#\mathcal{A}_{\mathrm{bulk}} \geq 2$.

Proof for Case c): Denote by $l \in \{0, 1\}$ the number of flags of C_n^{bulk} and let C^*_{n-d-l} be the configuration which arises by removing l atoms from C_n^{bulk} such that C^*_{n-d-l} has no acyclic bonds. Observe that $\mathcal{E}(C^*_{n-d-l}) - l \leq \mathcal{E}(C_n^{\mathrm{bulk}})$. Then also C^*_{n-d-l} is a ground state since otherwise $\mathcal{E}(C_n^{\mathrm{bulk}}) > -\lfloor \beta(n-d) \rfloor$ by Lemma 6.2 1) which contradicts (40). As $\#\mathcal{A}_{\mathrm{bulk}} \geq 2$ and $l \leq 1$, C^*_{n-d-l} contains a non-equilibrated atom. Thus, due to the fact that all hexagons in the bond graph are regular (see Lemma 7.5), we have that $\eta^* = \eta(C^*_{n-d-l}) \geq 2$, where η^* denotes the excess of C^*_{n-d-l}. By the induction assumption we have that C^*_{n-d-1} has no non-equilibrated bulk atom and thus has a non-equilibrated boundary atom. Thus, strict inequality holds in (14) for C^*_{n-d-l}, see Lemma 5.1. Therefore, by Lemma 6.4 (with strict inequality) applied for C^*_{n-d-l} and $\eta^* \geq 2$ we obtain

$$\mathcal{E}^{\mathrm{bulk}}(C_n) = \mathcal{E}(C_n^{\mathrm{bulk}}) \geq \mathcal{E}(C^*_{n-d-l}) - l \geq -\frac{3}{2}(n - d - l) + \sqrt{\frac{3}{2}(n - d - l - 2)} + 1 - l.$$

Using Lemma 5.1 for C_n and summing $\mathcal{E}^{\mathrm{bnd}}(C_n)$ and $\mathcal{E}^{\mathrm{bulk}}(C_n)$, we derive

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + 4 + \sqrt{\frac{3}{2}(n - d - 2 - l)} + \frac{l}{2}.$$

By Lemma 5.4 and $\eta \geq 6$ we obtain

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + 4 + \frac{l}{2} + \sqrt{\frac{3}{2}(-4\mathcal{E}(C_n) + 10 - 5n - l)}.$$

Lemma 6.1 for $j = 4 + \frac{l}{2}$, $m = 10 - l$, and $x = \mathcal{E}(C_n)$ yields

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + 1 + \frac{l}{2} + \sqrt{\frac{3}{2}(n - 3l)} = -\frac{3}{2}n + \sqrt{\frac{3}{2}n} + 1 + l\Big(\frac{1}{2} - \frac{9}{2(\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n - 3l)})}\Big).$$

This estimate can be used to calculate $\mathcal{E}(C_n) \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}n} + 1 > -\lfloor \beta(n) \rfloor$ for all $n \geq 16$ when $l = 1$ or for all $n \geq 10$ when $l = 0$. In the cases $10 \leq n \leq 15$, $l = 1$, one can use $\mathcal{E}(C_n) \geq -\frac{3}{2}n + 1 + \frac{1}{2} + \sqrt{\frac{3}{2}(n-3)}$ and compare this estimate directly with $\lfloor \beta(n) \rfloor$ to obtain $\mathcal{E}(C_n) > -\lfloor \beta(n) \rfloor$, cf. Table 1. In every case, this yields a contradiction to the fact that C_n is a ground state. □

7.2. **Characterization of ground states: proof of Theorem 2.3.** In this section we prove Theorem 2.3.

Boundary k-gon: We say that a k-gon in the bond graph is a *boundary* k-gon, whenever it shares at least one edge with the unbounded face.

The following proposition is the main ingredient for the proof of Theorem 2.3.

Proposition 7.9. *Let $n \geq 10$ and let C_n be a ground state. Then the bond graph consists only of hexagonal cycles except for at most two flags and at most one boundary octagon. The bond graph cannot contain both flags and an octagon at the same time.*

Once this proposition is proven, for the proof of Theorem 2.3 it remains to show that in the case $n \geq 30$ no octagons may occur. In fact, Theorem 2.3 then follows from Theorem 2.1, Remark 2.4, Lemma 3.1, and Lemma 7.2. For $n = 9$, a ground state may contain an octagon and a flag at the same time, see the rightmost configuration in Fig. 8. In this sense, the assumption $n \geq 10$ in Proposition 7.9 is sharp.

Proof of Proposition 7.9. Let C_n be a ground state. Recall by Theorem 2.1 that $\mathcal{E}(C_n) = -b = -\lfloor \beta(n) \rfloor$ and that X_n is connected. We divide the proof into several steps. First, we prove that ground states contain only hexagonal cycles if $\mathcal{A} = \emptyset$ (Claim 1). Then, in the case $\mathcal{A} \neq \emptyset$, we prove that at most one boundary octagon may exist (Claim 2). Finally, we show that the existence of a non-hexagonal cycle excludes the existence of flags (Claim 3). The statement follows from Claim 1 - Claim 3 and Lemma 7.4. The claims are proven by contradiction.

Claim 1: If $\mathcal{A} = \emptyset$, C_n is defect-free.

Proof of Claim 1: Suppose that there exists a k-gon, $k \geq 8$. Since X_n is connected and $\mathcal{A} = \emptyset$, we have that X_n is a connected subset of the hexagonal lattice. We observe that then our energy coincides with the one considered in [24]. We can repeat the argument in the proof of [24, Proposition 6.7], i.e., we can move boundary atoms inside this k-gon and observe that one can strictly lower the energy. This contradicts the fact that C_n is a ground state.

Claim 2: For every ground state without flags there exists at most one k-gon, $k \geq 8$. If it exists, it has to be a boundary octagon.

Proof of Claim 2: In view of Claim 1, we can suppose $\mathcal{A} \neq \emptyset$ and $\eta \geq 2$. As $\eta \geq 2$, C_n contains no bridge (see Lemma 7.2, Remark 7.3) and thus no acyclic bonds. By Lemma 7.8 we have that $\mathcal{A}_{\text{bulk}} = \emptyset$ and therefore $\mathcal{A} \setminus \mathcal{A}_{\text{bulk}} \neq \emptyset$. Applying Lemma 5.1 and Lemma 6.4 (with strict inequality) we obtain that

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + 1 + \sqrt{\frac{3}{2}(n + \eta - 4)}.$$

In the case $\eta \geq 4$ we obtain a contradiction to the fact that C_n is a ground state. Therefore, we can assume that $\eta = 2$, i.e., there exists exactly one octagon. We apply Lemma 7.5 to find that hexagons are regular which implies $\mathcal{A}$ has to be contained in the octagon of the bond graph. By Lemma 7.8 we observe $\mathcal{A} \subset \partial X_n$. Since atoms in $\mathcal{A}$ are 2-bonded (see Lemma 7.5), we get that the octagon shares at least two edges with the unbounded face. This implies that the octagon is a boundary octagon.

Claim 3: A ground state cannot contain both a k-gon, $k \geq 8$, and a flag.

Proof of Claim 3: Assume by contradiction that there exists a k-gon and l flags in the bond graph. (As before in Claim 2, by Remark 7.3 there are no bridges.) Using the fact that a flag contributes at least -1 to the energy and removing the flags we obtain a sub-configuration C_{n-l} satisfying $\mathcal{E}(C_n) \geq \mathcal{E}(C_{n-l}) - l$. Then also C_{n-l} is a ground state since otherwise $\mathcal{E}(C_n) > -\lfloor \beta(n) \rfloor$ by Lemma 6.2 1) which contradicts Theorem 2.1. As C_{n-l} has no acyclic bonds, we can use Lemma 7.8 to find $\mathcal{A}_{\text{bulk}} = \emptyset$ and therefore $\mathcal{A} \setminus \mathcal{A}_{\text{bulk}} \neq \emptyset$, where $\mathcal{A}_{\text{bulk}}, \mathcal{A}$ correspond to configuration C_{n-l}. Applying Lemma 5.1 and Lemma 6.4 on C_{n-l} (with strict inequality for

$\eta \geq 2$), we obtain

$$\mathcal{E}(C_n) \geq \mathcal{E}(C_{n-l}) - l \geq -l - \frac{3}{2}(n-l) + 1 + \sqrt{\frac{3}{2}(n-l-2)}.$$

Setting $j = l + 2$, we find

$$\mathcal{E}(C_n) \geq -\Big(\frac{3}{2}n - \sqrt{\frac{3}{2}n} - \frac{1}{2}j + \frac{\frac{3}{2}j}{\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n-j)}}\Big).$$

Recall that $j \geq 3$. At this point, we can follow verbatim the proof of Lemma 7.4 to get the contradiction $\mathcal{E}(C_n) > -\lfloor \beta(n) \rfloor$ for each $n \geq 10$. This contradicts Theorem 2.1 and the fact that C_n is a ground state. □

As a final preparation for the proof of Theorem 2.3, we need the following elementary geometric lemma.

Lemma 7.10. *Let C_n be a ground state with $\eta = 2$ that does not contain any acyclic bonds. Let $\{x_0, \dots, x_7\}$ be the octagon in the bond graph. Set $X^3 := \{x_i \in \{0, \dots, 7\} : x_i$ is 3-bonded$\}$. We have the following:*

i) *If $\#X^3 \leq 3$, then X^3 is connected,*
ii) *If $\#X^3 \in \{4, 5\}$, then there exist $j_1, j_2, j_3 \in \{0, \dots, 7\}$ such that $X_n \setminus \{x_{j_1}, x_{j_2}, x_{j_3}\}$ is not connected, $x_{j_1}, x_{j_2}, x_{j_3}$ are 2-bonded, and*

$$\mathcal{E}\big(C_n \setminus \{(x_{j_1}, q_{j_1}), (x_{j_2}, q_{j_2}), (x_{j_3}, q_{j_3})\}\big) \leq \mathcal{E}(C_n) + 5. \tag{43}$$

iii) *If $\#X^3 \geq 6$, then $C_n^{\rm bulk}$ is not connected.*

Proof. Let $\{x_0, \dots, x_7\}$ be the octagon in the bond graph. All following statements are seen mod 8 with respect to the numeration of i. Assume that $x_{i+1}, x_{i-1} \in \mathcal{N}(x_i)$ for all $i = 0, \dots, 7$ and $x_i \notin \mathcal{N}(x_j)$ for all $j \notin \{i-1, i+1\}$. Let θ_i, $i = 0, \dots, 7$, be the interior angles of the octagon. We start with three preliminary observations.

Claim 1: If x_i is 2-bonded, then $x_i \in \partial X_n$.

Proof of Claim 1: Consider a 2-bonded x_i and assume by contradiction that $x_i \in X_n^{\rm bulk}$. By Lemma 7.6(ii) we have that $\theta_i < \frac{4\pi}{3}$ and therefore the other angle α at x_i satisfies $\alpha > \frac{2\pi}{3}$. Since x_i is a bulk atom, Lemma 7.5 and $\alpha > \frac{2\pi}{3}$ imply that x_i is contained in the octagon and in another polygon which is not a hexagon since for hexagons all interior angles are equal to $\frac{2\pi}{3}$. This contradicts $\eta = 2$.

Claim 2: If x_i is 3-bonded, then x_{i-1} or x_{i+1} is also 3-bonded.

Proof of Claim 2: Assume by contradiction that there exists a 3-bonded x_i such that x_{i-1} and x_{i+1} are 2-bonded. Since the bond graph of C_n does not contain acyclic bonds, one of the atoms x_{i+1} or x_{i-1} has to be a bulk atom. This, however, contradicts Claim 1.

Claim 3: If x_{i-1}, x_i, x_{i+1} are 3-bonded, then x_{i-2} and x_{i+2} are 2-bonded.

Proof of Claim 3: If there were four consecutive 3-bonded atoms, there would be four consecutive interior angles of size $\frac{2\pi}{3}$, see Lemma 7.5. This, however, contradicts Lemma 7.6(i).

We now proceed with the proof of the statement. Set $k = \#X^3$. First, by Lemma 7.5, Lemma 7.6(ii), and the fact that $\sum_i \theta_i = 6\pi$ we have that $k \leq 6$. If $k = 0$, there is nothing to prove. $k = 1$ is not possible due to Claim 2. If $k = 2$, again due to Claim 2, the 3-bonded atoms have

to be bonded. If $k = 3$, the 3-bonded atoms necessarily need to be of the form x_{i-1}, x_i, x_{i+1} for some $i = 0, \dots, 7$. Otherwise, we have a contradiction to Claim 2. This proves i).

Now suppose that $k \in \{4, 5\}$. By Lemma 7.5 we have that, if x_i is 3-bonded, then $\theta_i = \frac{2\pi}{3}$. By Lemma 7.6(i) the 3-bonded atoms cannot be of the form $\{x_i, \dots, x_{i+k-1}\}$. Hence, there exist $0 \le i_1 < i_1 + 1 < j_1 < i_2 < i_2 + 2 < j_2$ such that $x_{i_1}, x_{i_1+1}, x_{i_2}, x_{i_2+1}$ are 3-bonded and $x_{j_1}, x_{j_2}, x_{j_2+1}$ are 2-bonded. By Claim 1 we get that, if x_j is 2-bonded, then $x_j \in \partial X_n$. Therefore, x_{i_1} and x_{i_2} are connected only through paths going through x_{j_1} or x_{j_2} as otherwise one of the two atoms x_{j_1} and x_{j_2} would not be contained in the boundary. Thus, the set $X_n \setminus \{x_{j_1}, x_{j_2}, x_{j_2+1}\}$ is not connected. Since $x_{j_1}, x_{j_2}, x_{j_2+1}$ are 2-bonded and x_{j_2} is bonded to x_{j_2+1}, we remove exactly 5 bonds. As each bond contributes at least -1 to the energy, this yields (43). This proves ii).

Finally, suppose $k = 6$. By Claim 3 we have that there exists $i \in \{0, \dots, 7\}$ such that $x_i, x_{i+1}, x_{i+2}, x_{i+4}, x_{i+5}, x_{i+6}$ are 3-bonded. Since no acyclic bonds are present in the bond graph, we have that $x_{i+1}, x_{i+5} \in X_n^{\text{bulk}}$. Moreover, arguing as in case ii), x_{i+1} is not connected with x_{i+5} in X_n^{bulk}. This proves iii). □

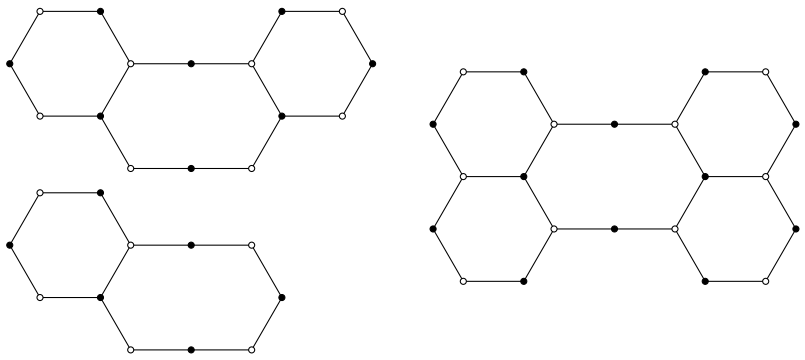

FIGURE 11. Illustration of the three cases in Lemma 7.10.

We are now in a position to prove Theorem 2.3.

Proof of Theorem 2.3. As observed below Proposition 7.9, it suffices to check that for $n \ge 30$ the bond graph of a ground state C_n does not contain an octagon. Assume by contradiction that $n \ge 30$ and that the bond graph contains an octagon. By Proposition 7.9 we know that there do not exist any flags in the bond graph and the bond graph contains only one (boundary) octagon and hexagons otherwise. By Lemma 7.2 the bond graph does not contain any acyclic bonds. Denote by $\{x_0, \dots, x_7\}$ the octagon in the bond graph. We need to consider the three cases

$$\text{(a)} \ \ \#X^3 \le 3, \qquad \text{(b)} \ \ \#X^3 \in \{4, 5\}, \qquad \text{(c)} \ \ \#X^3 \ge 6,$$

where $X^3 := \{x_i \in \{0, \dots, 7\} : x_i \text{ is 3-bonded}\}$. (See Fig. 11 for an illustration.)

Proof of Case (a)*:* Since $\#X^3 \le 3$, by Lemma 7.10 we have that there are k 2-bonded atoms with $k \ge 5$ and they form a connected set. Hence, removing these 2-bonded atoms, we remove exactly $k + 1$ bonds. Estimating the energy of every bond by -1 we get by Theorem 2.1

$$\mathcal{E}(C_n) \ge -(k + 1) - \lfloor \beta(n - k) \rfloor.$$

This implies $\mathcal{E}(C_n) > \lfloor \beta(n) \rfloor$. Indeed, this follows from Lemma 6.2 3)-4) and the fact that $k \ge 5$, $n \ge 30$. This gives a contradiction in Case (a).

Proof of Case (b)*:* By the assumption and Lemma 7.10 we have that there exist $j_1, j_2, j_3 \in \{0, \dots, 7\}$ such that

$$\mathcal{E}\big(C_n \setminus \{(x_{j_1}, q_{j_1}), (x_{j_2}, q_{j_2}), (x_{j_3}, q_{j_3})\}\big) \leq \mathcal{E}(C_n) + 5$$

and $X_n \setminus \{x_{j_1}, x_{j_2}, x_{j_3}\}$ is not connected. Denote by $n_1, n_2 \in \mathbb{N}, n_1 + n_2 = n - 3$ the cardinality of the two connected components of $X_n \setminus \{x_{j_1}, x_{j_2}, x_{j_3}\}$ which do not have any bonds between them. Since the bond graph of C_n does not contain any acyclic bonds, as explained at the beginning of the proof, we have $n_1, n_2 \geq 6$. By Lemma 6.2 2) and Theorem 2.1 we obtain

$$\mathcal{E}(C_n) \geq \mathcal{E}\big(C_n \setminus \{(x_i, q_i) : \ i = i_1, i_2, i_3\}\big) - 5 \geq -\lfloor \beta(n_1) \rfloor - \lfloor \beta(n_2) \rfloor - 5 > -\lfloor \beta(n-3) \rfloor - 4.$$

By Theorem 2.1 we have that $\mathcal{E}(C_n)$ is an integer. This implies

$$\begin{aligned} \mathcal{E}(C_n) &\geq -\lfloor \beta(n-3) \rfloor - 3 \geq -\frac{3}{2}n + \sqrt{\frac{3}{2}(n-3)} + \frac{3}{2} \\ &= -\frac{3}{2}n + \sqrt{\frac{3}{2}n} + \frac{3}{2} - \frac{9}{2\Big(\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n-3)}\Big)}. \end{aligned} \tag{44}$$

It is elementary to check that for $n \geq 16$ we have

$$\frac{3}{2} - \frac{9}{2\Big(\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n-3)}\Big)} \geq 1.$$

This is a contradiction to the fact that C_n is a ground state.

Proof of Case (c)*:* By the assumption and Lemma 7.10 we have that C_n^{bulk} is not connected. As the bond graph contains an octagon, there exists a non-equilibrated atom. Lemma 7.8 implies $\mathcal{A}_{\mathrm{bulk}} = \emptyset$ and thus a boundary atom is not equilibrated. Then by Lemma 5.1 we get

$$\mathcal{E}^{\mathrm{bnd}}(C_n) > -\frac{3}{2}d + 3.$$

As C_n^{bulk} is not connected, C_n^{bulk} cannot be a ground state. Applying Theorem 2.1 to C_n^{bulk} we thus obtain

$$\mathcal{E}^{\mathrm{bulk}}(C_n) > -\lfloor \beta(n-d) \rfloor.$$

By Theorem 2.1, we have $\mathcal{E}(C_n) = -b$ and therefore $\mathcal{E}^{\mathrm{bulk}}(C_n)$ and $\mathcal{E}^{\mathrm{bnd}}(C_n)$ are integers. Hence, we obtain

$$\mathcal{E}(C_n) \geq -\lfloor \beta(n-d) \rfloor - \frac{3}{2}d + 5.$$

Using Lemma 5.4 and $\eta = 2$ we get

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + 5 - \sqrt{\frac{3}{2}(-4\mathcal{E}(C_n) - 5n + 8)}.$$

This together with Lemma 6.1 applied for $j = 5$, $m = 8$, and $x = \mathcal{E}(C_n)$ leads to

$$\mathcal{E}(C_n) \geq -\frac{3}{2}n + 2 + \sqrt{\frac{3}{2}(n-6)} = -\frac{3}{2}n + \sqrt{\frac{3}{2}n} + 2 - \frac{9}{\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n-6)}}. \tag{45}$$

For $n \geq 17$ we have

$$2 - \frac{9}{\sqrt{\frac{3}{2}n} + \sqrt{\frac{3}{2}(n-6)}} \geq 1$$

which leads to a contradiction to the fact that C_n is a ground state. □

Remark 7.11. Inspection of the previous proof shows that among $13 \leq n \leq 29$ only for $n = 15, 18, 21, 29$ boundary octagons may occur. Indeed, in Case (a) this follows from Lemma 6.2 4), see particularly Table 2. In Case (b) and Case (c) we obtain a contradiction for each $13 \leq n \leq 29$: in Case (b), we necessarily have $n \geq 16$ (see upper left configuration in Fig. 11) and thus a contradiction in (44). In Case (c), we necessarily have $n \geq 22$ (see rightmost configuration in Fig. 11) and thus a contradiction in (45).

For $n = 10, 11$, the presence of an octagon is excluded by Proposition 7.9 and the fact that each hexagon can share at most 2 atoms with an octagon (see Lemma 7.5 and Lemma 7.6). Consequently, for $n \leq 29$, ground states may contain a boundary octagon only for $n = 8, 9, 12, 15, 18, 21, 29$. This is indeed possible as shown in Fig. 8 and Fig. 12, cf. Table 1.

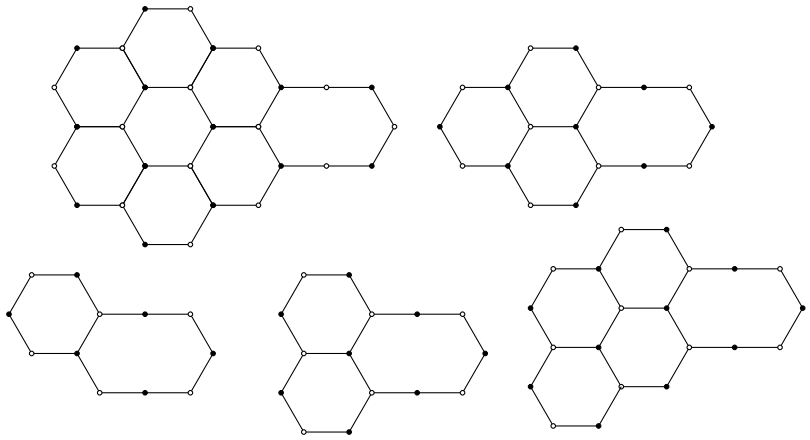

FIGURE 12. The ground states containing an octagon for $n = 12, 15, 18, 21, 29$ (up to isometry and changing of the charges).

8. Characterization of the net charge

This final section is devoted to the proof of Theorem 2.5(i). We recall that part (ii) of the statement has already been addressed by an explicit construction in Section 4.2.

We start with some preliminary definitions. First, recall the definition of the hexagonal lattice $\mathcal{L}$ in (6). Set $u_1 = (1, 0)$, $u_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$, $u_3 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$, $u_4 = (-1, 0)$.

Definition 8.1 (Zig-zag paths)**.** (i) A tuple $(p_1, \dots, p_m) \subset \mathcal{L}$ is called a *zig-zag path* if there exists $k \in \{1, 2, 3\}$ such that

$$\begin{aligned} &(a) \quad p_j - p_{j-1} \in \{u_k, u_{k+1}\} \quad \text{for all } j \in \{2, \dots, m\}, \\ &(b) \quad p_{j+1} - p_j \neq p_j - p_{j-1} \quad \text{for all } j \in \{2, \dots, m-1\}. \end{aligned} \tag{46}$$

(ii) We say that two zig-zag paths have the same *orientation* if the same $k \in \{1, 2, 3\}$ appears in (46).

(iii) Given a configuration C_n with $X_n \subset \mathcal{L}$, we say the zig-zag path $(p_1, \dots, p_m) \subset \mathcal{L}$, $m \geq 3$, is a *bridging zig-zag path* for C_n if $p_1, p_m \in X_n$ and $p_2, \dots, p_{m-1} \notin X_n$.

Let C_n be a ground state for $n \geq 30$ with no acyclic bonds. By Theorem 2.3 we get that X_n is defect-free and satisfies $X_n \subset \mathcal{L}$ (up to isometry). Let $(p_1, \dots, p_m)$ be a bridging zig-zag

path for C_n. Since X_n is connected and defect-free, we find that $\mathcal{L} \setminus (X_n \cup \bigcup_{j=2}^{m-1} p_j)$ consists of two (one of them possibly empty) connected components. Exactly one of these components is bounded which we denote by $\mathcal{B}(X_n; (p_1, \dots, p_m))$.

Lemma 8.2 (Bridging zig-zag paths of ground states)**.** *Let $n \geq 30$. Each ground state with no acyclic bonds has at most one bridging zig-zag path.*

Proof. As a preparation, we observe the following: suppose that there exists a bridging zig-zag path for a ground state C_n. Then we can choose a bridging zig-zag path $(p_1, \dots, p_m) \subset \mathcal{L}$ for C_n such that $\mathcal{B}(X_n; (p_1, \dots, p_m)) = \emptyset$. To see this, we proceed as follows. Take an arbitrary bridging zig-zag path $(\bar{p}_1, \dots, \bar{p}_{\bar{m}})$ and consider $\bar{\mathcal{B}} := \mathcal{B}(X_n; (\bar{p}_1, \dots, \bar{p}_{\bar{m}}))$. If $\bar{\mathcal{B}} = \emptyset$, we have concluded. If $\bar{\mathcal{B}} \neq \emptyset$, we can consider another bridging zig-zag path $(\tilde{p}_1, \dots, \tilde{p}_{\tilde{m}})$ for C_n having the same orientation as $(\bar{p}_1, \dots, \bar{p}_{\bar{m}})$ and satisfying $\tilde{p}_2, \dots, \tilde{p}_{\tilde{m}-1} \in \bar{\mathcal{B}}$. Define $\tilde{\mathcal{B}} := \mathcal{B}(X_n; (\tilde{p}_1, \dots, \tilde{p}_{\tilde{m}}))$ and note that $\tilde{\mathcal{B}} \subset \bar{\mathcal{B}}$, $\#\tilde{\mathcal{B}} < \#\bar{\mathcal{B}}$. This construction can be iterated and after a finite number of iteration steps we find a bridging zig-zag path $(p_1, \dots, p_m) \subset \mathcal{L}$ for C_n such that $\mathcal{B}(X_n; (p_1, \dots, p_m)) = \emptyset$.

Now suppose by contradiction that there was a ground state C_n, $n \geq 30$, with no acyclic bonds which contains two bridging zig-zag paths. Consider a bridging zig-zag path $(p_1, \dots, p_m)$ satisfying $\mathcal{B}(X_n; (p_1, \dots, p_m)) = \emptyset$ and define $X_n' = X_n \cup \bigcup_{j=2}^{m-1} p_j$. Clearly, we can assign charges to the atoms $p_2, \dots, p_{m-1}$ to obtain a configuration C_n' with alternating charge distribution. Note that C_n' consists of $n + m - 2$ atoms. We now estimate the energy of C_n'.

First, we observe that between the atoms of the path $(p_1, \dots, p_m)$ there are $m - 1$ bonds. Since $\mathcal{B}(X_n; (p_1, \dots, p_m)) = \emptyset$, either each $p_2, p_4, \dots, p_{2\lfloor (m-1)/2 \rfloor}$ or each $p_3, p_5, \dots, p_{2\lfloor m/2 \rfloor - 1}$ is bonded to an atom of C_n. (The latter set is empty if $m = 3$.) Thus, we obtain

$$\mathcal{E}(C_n') \leq \mathcal{E}(C_n) - (m-1) - \min\{\lfloor (m-1)/2 \rfloor, \lfloor m/2 \rfloor - 1\} \leq \mathcal{E}(C_n) - \frac{3}{2}m + \frac{5}{2}.$$

In particular, as C_n was supposed to be a ground state, this implies $\mathcal{E}(C_n') \leq -\lfloor \beta(n) \rfloor - \frac{3}{2}m + \frac{5}{2} \leq -\lfloor \beta(n+m-2) \rfloor$. Here the second inequality is elementary to check. This shows that C_n' is a ground state.

As C_n has two bridging zig-zag paths, there is at least one bridging zig-zag path for C_n'. We now repeat the above procedure. Choose $(p_1', \dots, p_{m'}')$ with $\mathcal{B}(X_n'; (p_1', \dots, p_{m'}')) = \emptyset$ and define a configuration C_n'' with $\#X_n'' = n + m + m' - 4$, with alternating charge distribution and consisting of the atoms $X_n' \cup \bigcup_{j=2}^{m'-1} p_j'$. Arguing as before, we calculate

$$\mathcal{E}(C_n'') \leq \mathcal{E}(C_n') - \frac{3}{2}m' + \frac{5}{2} \leq \mathcal{E}(C_n) - \frac{3}{2}(m + m') + 5.$$

Since C_n was supposed to be a ground state, this implies

$$\begin{aligned} \mathcal{E}(C_n'') &\leq -\left\lfloor \frac{3}{2}n - \sqrt{\frac{3}{2}n} \right\rfloor - \frac{3}{2}(m+m') + 5 \\ &\leq -\left\lfloor \frac{3}{2}(n+m+m'-4) - \sqrt{\frac{3}{2}(n+m+m'-4)} \right\rfloor - \frac{1}{2}. \end{aligned}$$

This implies $\mathcal{E}(C_n'') < -\lfloor \beta(n+m+m'-4) \rfloor$ which contradicts Theorem 2.1. This concludes the proof. □

Recall the construction of daisies in Section 4.1. From [6] we obtain the following result.

Proposition 8.3 (Deviation from Wulff-shape)**.** *Let $n \geq 30$ and let C_n be a ground state with no acyclic bonds. Then, possibly after translation, we find two daisies $X^{\mathrm{daisy}}_{6k_1^2} \subset \mathcal{L}$ and $X^{\mathrm{daisy}}_{6k_2^2} \subset \mathcal{L}$ with $X^{\mathrm{daisy}}_{6k_1^2} \subset X_n \subset X^{\mathrm{daisy}}_{6k_2^2}$ such that*

$$0 < k_2 - k_1 \leq cn^{1/4},$$

where $c > 0$ is a universal constant independent of n and C_n.

Proof. From Theorem 2.3 and the fact that C_n does not have acyclic bonds we get that the ground state C_n is a subset of the hexagonal lattice. Moreover, C_n is repulsion-free, see Remark 2.2. Thus, the energy of a ground state coincides with the one in [6], see [6, Equation (6)]. The claim then follows from [6, Theorem 1.2]. □

We are now in the position to prove Theorem 2.5.

Proof of Theorem 2.5. As discussed at the beginning of the section, it remains to prove part (i) of the statement. In view of Theorem 2.3 and Remark 2.4(i), it suffices to treat the case that C_n does not have acyclic bonds. We apply Proposition 8.3 to find two daisies with $X^{\mathrm{daisy}}_{6k_1^2} \subset X_n \subset X^{\mathrm{daisy}}_{6k_2^2}$. It is elementary to see that $X^{\mathrm{daisy}}_{6k_2^2} \setminus X^{\mathrm{daisy}}_{6k_1^2}$ can be written as the union of $6(k_2 - k_1)$ zig-zag paths as introduced in Definition 8.1. Then by Lemma 8.2 we get that $X_n \setminus X^{\mathrm{daisy}}_{6k_1^2}$ can be written as the union of at most $6(k_2 - k_1) + 1$ zig-zag paths. Recall that C_n has alternating charge distribution and therefore the net charge of each zig-zag path is in $\{-1, 0, 1\}$. Also recall from Section 4.1 that daisies always have net charge zero. This implies that the net charge of the configuration C_n satisfies

$$|\mathcal{Q}(C_n)| \leq 6(k_2 - k_1) + 1.$$

The statement follows from the fact that $k_2 - k_1 \leq cn^{1/4}$, see Proposition 8.3. □

Acknowledgements

M. F. acknowledges support from the Alexander von Humboldt Stiftung. L. K. acknowledges support from the Austrian Science Fund (FWF) project P 29681, and from the Vienna Science and Technology Fund (WWTF), the City of Vienna, and the Berndorf Private Foundation through Project MA16-005. The authors would like to thank Ulisse Stefanelli for turning their attention to this problem.

References

[1] N.L. Allinger. *Molecular structure: understanding steric and electronic effects from molecular mechanics*. John Wiley & Sons (2010).

[2] Y. Au Yeung, G. Friesecke, B. Schmidt. *Minimizing atomic configurations of short range pair potentials in two dimensions: crystallization in the Wulff-shape*. Calc. Var. Partial Differential Equations **44** (2012), 81–100.

[3] L. Bétermin, H. Knüpfer, F. Nolte. *Crystallization of one-dimensional alternating two-component systems*. Preprint at arXiv:1804.05743.

[4] X. Blanc, M. Lewin. *The crystallization conjecture: a review*. EMS Surv. Math. Sci. **2** (2015), 255–306.

[5] D. C. Brydges, P. A. Martin. *Coulomb systems at low density: A review*. J. Stat. Phys. **96** (1999), 1163–1330.

[6] E. Davoli, P. Piovano, U. Stefanelli. *Wulff shape emergence in graphene*. Math. Models Methods Appl. Sci. **26** (2016), 12:2277–2310.

[7] E. Davoli, P. Piovano, U. Stefanelli. *Sharp $n^{3/4}$ law for the minimizers of the edge-isoperimetric problem in the triangular lattice*. J. Nonlin. Sci. **27** (2017), 627–660.
[8] L. De Luca, G. Friesecke. *Crystallization in two dimensions and a discrete GaussBonnet Theorem*. J. Nonlinear Sci. **28** (2017), 69–90.
[9] W. E, D. Li. *On the crystallization of 2D hexagonal lattices*. Comm. Math. Phys. **286** (2009), 1099–1140.
[10] B. Farmer, S. Esedoḡlu, P. Smereka. *Crystallization for a Brenner-like potential*. Comm. Math. Phys. **349** (2017), 1029–1061.
[11] L. Flatley, M. Taylor, A. Tarasov, F Theil. *Packing twelve spherical caps to maximize tangencies*. J. Comput. Appl. Math. **254** (2013),220–225.
[12] L. Flatley, F. Theil. *Face-centered cubic crystallization of atomistic configurations*. Arch. Ration. Mech. Anal. **218** (2015), 363–416.
[13] M. Friedrich, U. Stefanelli. *Graphene ground states*. Z. Angew. Math. Phys. **69** (2018): 70.
[14] M. Friedrich, L. Kreutz. *Crystallization in the square lattice for ionic compounds*. In preparation.
[15] G. Friesecke, F. Theil. *Molecular geometry optimization, models*. In the Encyclopedia of Applied and Computational Mathematics, B. Engquist (Ed.), Springer, 2015.
[16] C. S. Gardner, C. Radin. *The infinite-volume ground state of the Lennard-Jones potential*. J. Stat. Phys. **20** (1979), 719–724.
[17] A. K. Geim, K. S. Novoselov. *The rise of graphene*. Nat. Mater. **6** (2007), 183–191.
[18] G. C. Hamrick and C. Radin. *The symmetry of ground states under perturbation*. J. Stat. Phys. **21** (1979), 601–607.
[19] R. Heitman, C. Radin. *Ground states for sticky disks*. J. Stat. Phys. **22** (1980), 3:281–287.
[20] K. K. Kim, A. Hsu, X. Jia, S. M. Kim, Y. Shi, M. Dresselhaus, T. Palacios, J. Kong. *Synthesis and characterization of hexagonal boron nitride film as a dielectric layer for graphene devices*. Acs Nano **6** (2012), 8583–8590.
[21] G. Lazzaroni, U. Stefanelli. *Chain-like minimizers in three dimensions*. Submitted, 2018. Preprint at http://cvgmt.sns.it/paper/3418/.
[22] E. G. Lewars. *Computational Chemistry*. 2nd edition, Springer, 2011.
[23] E. Mainini, P. Piovano, U. Stefanelli. *Finite crystallization in the square lattice*. Nonlinearity **27** (2014), 717–737.
[24] E. Mainini, U. Stefanelli. *Crystallization in carbon nanostructures*. Comm. Math. Phys. **328** (2014), 545–571.
[25] L. Pauling. *The nature of the chemical bond and the structure of molecules and crystals: an introduction to modern structural chemistry*. Ithaca, New York: Cornell University Press 1960.
[26] C. Radin. *The ground state for soft disks*. J. Stat. Phys. **26** (1981), 2:365–373.
[27] C. Radin. *Classical ground states in one dimension*. J. Stat. Phys. **35** (1983), 109–117.
[28] C. Radin. *Crystals and quasicrystals: a continuum model*. Comm. Math. Phys. **105** (1986), 385–390.
[29] B. Schmidt. *Ground states of the 2D sticky disc model: fine properties and $N^{3/4}$ law for the deviation from the asymptotic Wulff-shape*. J. Stat. Phys. **153** (2013), 727–738.
[30] A. Sütő. *From bcc to fcc: Interplay between oscillation long-range and repulsive short range forces*. Phys. Rev. B **74** (2006), 104117.
[31] F. Theil. *A proof of crystallization in two dimensions*. Comm. Math. Phys. **262** (2006), 209–236.
[32] W. J. Ventevogel, B. R. A. Nijboer. *On the configuration of systems of interacting atom with minimum potential energy per atom*. Phys. A **99** (1979), 565–580.
[33] J. Wang, F. Ma, M. Sun. *Graphene, hexagonal boron nitride, and their heterostructures: properties and applications*. Rsc Adv. **7** (2017), 16801–16822.
[34] H. J. Wagner. *Crystallinity in two dimensions: a note on a paper of C. Radin*. J. Stat. Phys. **33**, (1983), 523–526.

(Manuel Friedrich) Applied Mathematics Münster, University of Münster, Einsteinstrasse 62, 48149 Münster, Germany.

E-mail address: manuel.friedrich@uni-muenster.de

(Leonard Kreutz) Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria.

E-mail address: leonard.kreutz@univie.ac.at