THE UNIVERSITY OF
WARWICK

Original citation:
Ortner, Christoph and Shapeev, A. V.. (2013) Analysis of an energy-based atomistic/continuum approximation of a vacancy in the 2D triangular lattice. Mathematics of Computation, Volume 82 (Number 284). pp. 2191-2236.

Permanent WRAP url:
http://wrap.warwick.ac.uk/60455

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
© First published in Mathematics of Computation in Volume 82 (Number 284). pp. 2191-2236. published by the American Mathematical Society
http://dx.doi.org/10.1090/S0025-5718-2013-02687-7

A note on versions:
The version presented in WRAP is the published version or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: publications@warwick.ac.uk

warwickpublicationswrap
highlight your research

http://wrap.warwick.ac.uk/

MATHEMATICS OF COMPUTATION
Volume 00, Number 0, Pages 000–000
S 0025-5718(XX)0000-0

ANALYSIS OF AN ENERGY-BASED ATOMISTIC/CONTINUUM APPROXIMATION OF A VACANCY IN THE 2D TRIANGULAR LATTICE

C. ORTNER AND A. V. SHAPEEV

Abstract. We present an *a priori* error analysis of a practical energy based atomistic/continuum coupling method (A. V. Shapeev, Multiscale Model. Simul., 9(3):905–932, 2011) in two dimensions, for finite-range pair-potential interactions, in the presence of vacancy defects.

We establish first-order consistency and stability of the method, from which we obtain *a priori* error estimates in the H^1-norm and the energy in terms of the mesh size and the "smoothness" of the atomistic solution in the continuum region. From these error estimates we obtain heuristics for an optimal choice of the atomistic region and the finite element mesh, as well as convergence rates in terms of the number of degrees of freedom. Our analytical predictions are supported by extensive numerical tests.

1. Introduction

The purpose of this work is a rigorous study of a new computational multiscale method coupling an atomistic description of a defect to a continuum model of the elastic far field.

The accurate computational modelling of crystal defects requires an atomistic description of the defect core, as well as an accurate resolution of the elastic far field. Atomistic-to-continuum coupling methods (a/c methods) have been proposed to combine the accuracy of atomistic modelling with the efficiency of continuum mechanics (see [14, 21, 31, 32, 34] for selected references, and [19] for a recent overview).

The construction of accurate *energy-based* a/c methods has been proven particularly challenging, due to the so-called "ghost-forces" at the interface between the atomistic and continuum regions. This issue has been discussed at great length in [31, 4, 7, 20], and several interface corrections have been proposed to either remove or reduce the ghost forces [32, 7, 13, 34, 30, 12]. In general, the ghost-force removal problem remains unsolved.

A growing body of literature exists on the rigorous analysis of a/c methods (we refer to [30, 22, 18] for recent overviews), which has been largely restricted to one-dimensional model problems. We are currently aware of only two exceptions: (1) In

Received by the editor March 7, 2012.
2000 *Mathematics Subject Classification.* 65N12, 65N15, 70C20.
Key words and phrases. atomistic models, atomistic-to-continuum coupling, coarse graining.
This work was supported by the EPSRC Critical Mass Programme "New Frontiers in the Mathematics of Solids" (OxMoS), by the EPSRC grant "Analysis of atomistic-to-continuum coupling methods", and by the ANMC Chair at EPFL (Prof. Assyr Abdulle).

©XXXX American Mathematical Society

[22] it is shown that, in 2D, any a/c method that has no ghost forces is automatically first-order consistent. This work provides a general consistency analysis, but does not address stability of a/c methods. (2) In [17], a force-based a/c method with an overlap region is analyzed, in particular providing sharp stability estimates. However, the method proposed in [17] is not practical since it requires a prohibitively wide overlap region; moreover, the analytic methods employed cannot accommodate defects, or coarse-graining of the continuum region. Some of these challenges have been overcome in [16], where it is shown that a fairly narrow blending width is sufficient to ensure stability of the method, however, a complete analysis of a practical variant of force-based blending remains open.

In the present work, we give an *a priori* error analysis of a *practical* energy-based a/c method proposed by Shapeev [30]. The formulation of the method and its analysis are restricted to pair interactions in two dimensions, with periodic boundary conditions.

For the case of a vacancy defect our analysis requires an assumption on the magnitude of the deformation field generated by the defect, but is otherwise fully rigorous. For more general "vacancy sets", our consistency analysis remains fully rigorous, however, our stability analysis relies on a so-called vacancy stability index, which we estimate numerically (we give an analytical estimate for the single vacancy case).

Finally, we remark that the goal of a/c methods is to simulate far more complex situations than we can treat rigorously; we employ the example of a vacancy defect as the simplest non-trivial model problem.

1.1. **Outline.** In §2 we formulate an atomistic model for the 2D triangular lattice, with periodic boundary conditions, and two-body interactions. We also introduce a convenient notation for bonds.

In §3, we formulate the a/c method studied in this paper: the ECC method introduced in [30], but with periodic boundary conditions. This section contains all necessary results and notation required for an implementation of the a/c method, as well as a brief sketch of the proof of the *a priori* error estimate in order to motivate the subsequent analysis.

In §4 we collect auxiliary results, which are largely technical results for finite element spaces. We also introduce a new measure of "smoothness" of discrete functions.

In §5 we prove consistency error estimates in discrete variants of the $\mathrm{W}^{-1,p}$-norm, $p \in [1, \infty]$. Our estimates are stronger and require fewer technical assumptions than the general result given in [22].

In §6 we develop the stability analysis. We introduce a "vacancy stability index", which reduces the stability of a lattice with vacancies to stability of a lattice without defects. We provide numerical examples and one analytical estimation of stability indices.

In §7 we combine our previous steps to *a priori* error estimates in the H^1-norm and the energy. In §7.2 we translate these error estimates, which are stated in terms of the smoothness of the solution, into estimates in terms of degrees of freedom. This provides heuristics on how to optimally choose the atomistic region and the finite element mesh.

Finally, in §8, we present extensive numerical examples to confirm our analytical results, and to provide further discussions of points where our rigorous analysis is not sharp.

1.2. **Notation.** For $s, t \in \mathbb{R}$, we write $s \wedge t := \min\{s, t\}$.

The ℓ^p-norms in $\mathbb{R}^k$ are denoted by $|\cdot|_p$ and $|\cdot| := |\cdot|_2$. We do not normally distinguish between row and column vectors, but instead define three vector products: if $a, b \in \mathbb{R}^k$, then $a \cdot b := \sum_{j=1}^k a_j b_j$, and $a \otimes b := (a_i b_j)_{i,j=1}^k$, where i denotes the row index and j the column index. If $a, b \in \mathbb{R}^2$, then we also define $a \times b := a_1 b_2 - a_2 b_1$.

Matrices are usually denoted by sans serif symbols, $\mathsf{A}, \mathsf{B}, \mathsf{F}, \mathsf{G}$, and so forth. The set of $k \times k$ matrices with positive determinant is denoted by $\mathbb{R}^{k\times k}_+$. The set of rotations of $\mathbb{R}^2$ is denoted by $\mathrm{SO}(2)$. Throughout we will denote a rotation through angle $\pi/2$ by Q_4 and a rotation through angle $\pi/3$ by Q_6. If $\mathsf{G} \in \mathbb{R}^{k\times k}$, then $\|\mathsf{G}\|$ denotes its ℓ^2-operator norm, and $|\mathsf{G}|_p$ the $\ell^p(\mathbb{R}^{k\times k})$-norm. In particular, $|\mathsf{G}|$ is the Frobenius norm, with the associated inner product $\mathsf{F} : \mathsf{G}$. The symmetric component of a matrix $\mathsf{G} \in \mathbb{R}^{k\times k}$ is denoted by $\mathsf{G}^{\mathrm{sym}} := \frac{1}{2}(\mathsf{G} + \mathsf{G}^\top)$.

If $A \subset \mathbb{R}^k$ is (Lebesgue-)measurable, then $|A|$ denotes its measure. If $A \subset \mathbb{R}^2$ has Hausdorff dimension one, then we will denote its length by $\mathrm{length}(A)$. Volume integrals are denoted by dV, while surface (1D) integrals are denoted by $\mathrm{d}s$. For bonds, which are specific one-dimensional objects, it will be convenient to introduce a slightly different notation (see §2.2 and §3.2).

The interior and closure of a set $A \subset \mathbb{R}^k$ are denoted, respectively, by $\mathrm{int}(A)$ and $\mathrm{clos}(A)$. If $A \subset \mathbb{R}^2$ is understood as a one-dimensional object, then we will also use $\mathrm{int}(A)$ to denote its relative interior, but will normally specify this explicitly.

The Lebesgue norms $\|\cdot\|_{\mathrm{L}^p(A)}$ for one- or two-dimensional measurable sets A are defined in the usual way for scalar functions. If $w : A \to \mathbb{R}^k$ is measurable, then $\|w\|_{\mathrm{L}^p(A)} := \||w|_2\|_{\mathrm{L}^p(A)}$. If w is differentiable at a point x, then $\nabla w(x)$ denotes its Jacobi matrix. The symbol D is reserved for finite differences, and will be introduced in §2.2.

2. The Atomistic Model

2.1. **The triangular lattice with vacancy defects.** The 2D triangular lattice is the set

$$\mathbb{L}^\# := \mathsf{A}_6\mathbb{Z}^2, \qquad \text{where } \mathsf{A}_6 := \big[\mathsf{a}_1, \mathsf{a}_2\big] := \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix},$$

where a_i, $i = 1, 2$, are called the *lattice vectors*. We furthermore set $\mathsf{a}_3 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})^\top$ and $\mathsf{a}_{i+3} = -\mathsf{a}_i$ for $i \in \mathbb{Z}$, so that the set of *nearest-neighbour directions* is given by

$$\mathbb{L}_{\mathrm{nn}} := \big\{\mathsf{a}_j : j = 1, \dots, 6\big\} = \big\{\mathsf{Q}_6^{j-1}\mathsf{a}_1 : j = 1, \dots, 6\big\},$$

where $\mathsf{Q}_6 \in \mathrm{SO}(2)$ denotes the rotation through $\pi/3$. Finally, we denote the set of all *lattice directions* by $\mathbb{L}_* := \mathbb{L}^\# \setminus \{0\}$. The hexagonal symmetry of $\mathbb{L}^\#$ yields the following result, which decomposes the triangular lattice into lattice vectors of equal distance.

Lemma 2.1. *There exists a sequence $(r_n)_{n=1}^\infty \subset \mathbb{L}_*$ such that $\ell_n := |r_n|$ is monotonically increasing and the triangular lattice can be written as a union of disjoint sets $\mathbb{L}_* = \bigcup_{n=1}^\infty \big\{\mathsf{Q}_6^j r_n : j = 1, \dots, 6\big\}$.*

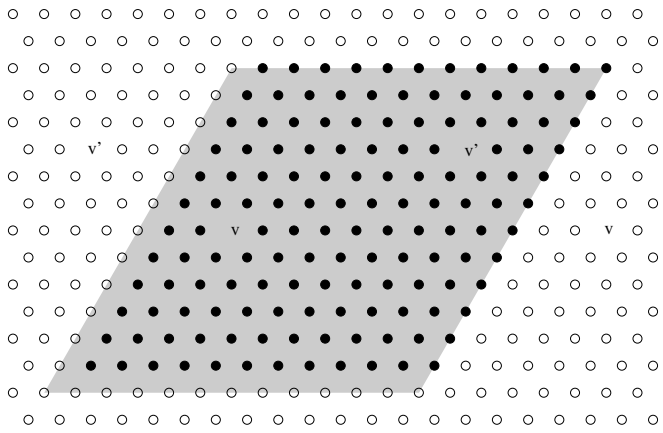


FIGURE 1. Lattice and the computational domain with $N = 12$ and two vacancies. The black disks denote the atoms belonging to the computational domain $\mathcal{L}$, the white disks denote the atoms belonging to $\mathcal{L}^\# \setminus \mathcal{L}$, and the vacancies are denoted by v and v'.

Lemma 2.1 motivates the splitting of lattice sums into hexagonally symmetric sets. In these calculations the following two identities will prove useful. Their proofs are given in Appendix A.

Lemma 2.2. *Let* $\mathsf{G} \in \mathbb{R}^{2\times 2}$, *and* $r \in \mathbb{R}^2$, $|r| = 1$; *then*

$$\sum_{j=1}^{6} \left|\mathsf{G}\mathsf{Q}_6^j r\right|^2 = 3|\mathsf{G}|^2, \qquad \textit{and} \tag{2.1}$$

$$\sum_{j=1}^{6} \left[(\mathsf{Q}_6^j r)^\top \mathsf{G}(\mathsf{Q}_6^j r)\right]^2 = \tfrac{3}{2}|\mathsf{G}^{\rm sym}|^2 + \tfrac{3}{4}|{\rm tr}\mathsf{G}|^2. \tag{2.2}$$

Throughout the paper we fix a periodicity parameter $N \in \mathbb{N}$. We say that a set $A \subset \mathbb{R}^2$ is N-periodic if $A + N\mathbb{L}^\# = A$. For any set $A \subset \mathbb{R}^2$ we denote its periodic continuation by $A^\# = A + N\mathbb{L}^\#$. If $\mathscr{A}$ is a family of sets, then we define $\mathscr{A}^\# = \{A^\# : A \in \mathscr{A}\}$.

We denote the continuous and discrete cells by

$$\Omega := \mathsf{A}_6(0, N]^2 \quad \text{and} \quad \mathbb{L} := \mathbb{L}^\# \cap \Omega.$$

We fix a set of *vacancy* sites $\mathbb{V} \subset \mathbb{L}$ and define the *discrete computational domain* as (*cf.* Figure 1)

$$\mathcal{L} := \mathbb{L} \setminus \mathbb{V}.$$

A *homogeneous deformation* of $\mathcal{L}^\#$ is a map $y_\mathsf{B} : \mathcal{L}^\# \to \mathbb{R}^2$ defined, for $\mathsf{B} \in \mathbb{R}^{2\times 2}_+$, as $y_\mathsf{B}(x) := \mathsf{B}x$, $x \in \mathcal{L}^\#$. The space of *periodic displacements* is denoted by

$$\mathscr{U} = \left\{u : \mathcal{L}^\# \to \mathbb{R}^2 : u(x + N\mathsf{a}_j) = u(x) \text{ for } x \in \mathcal{L}^\# \text{ and } j = 1, 2\right\}.$$

A map $y : \mathcal{L}^\# \to \mathbb{R}^2$ is said to be a *periodic deformation* with underlying macroscopic strain $\mathsf{B} \in \mathbb{R}^{2\times 2}_+$ if $y - y_\mathsf{B} \in \mathscr{U}$ and if y is *invertible*. To quantify the

invertibility condition we define

$$\mu_{\rm a}(y) = \inf_{x\neq x'\in\mathcal{L}^\#} \tfrac{|y(x')-y(x)|}{|x-x'|}$$

and denote

$$\mathscr{Y}_{\mathsf{B}} := \big\{ y : \mathcal{L}^\# \to \mathbb{R}^2 : y - y_{\mathsf{B}} \in \mathscr{U} \text{ and } \mu_{\rm a}(y) > 0 \big\}, \quad \text{and}$$
$$\mathscr{Y} := \textstyle\bigcup_{\mathsf{B}\in\mathbb{R}^{2\times 2}_+} \mathscr{Y}_{\mathsf{B}}.$$

2.2. **Bonds.** A *bond* is an ordered pair $(x, x') \in \mathbb{L}^\# \times \mathbb{L}^\#$, $x \neq x'$. When convenient we identify the bond $b = (x, x')$ with the line segment $\operatorname{conv}\{x, x'\}$, for example, to integrate over the segment, and correspondingly define $|b| := |x - x'|$. The set of bonds between atoms in the computational domain $\mathcal{L}$ and all other atoms is denoted by

$$\mathcal{B} := \big\{ (x, x') \in \mathcal{L} \times \mathcal{L}^\# : x \neq x' \big\}.$$

The *direction* of a bond b will be denoted by r_b, that is $b = (x, x + r_b)$ for some $x \in \mathbb{L}^\#$.

For a map $v : \mathbb{L}^\# \to \mathbb{R}^k$ and a bond $b = (x, x+r_b)$, we define the finite difference operators

$$D_b v := D_{r_b} v(x) := v(x + r_b) - v(x). \tag{2.3}$$

With this notation, we have $\mu_{\rm a}(y) = \min_{b\in\mathcal{B}} |D_b y|/|b|$.

We define the set of *all* bonds, including vacancy sites, by

$$\mathbb{B} := \big\{ (x, x + r) : x \in \mathbb{L}, r \in \mathbb{L}_* \big\}.$$

2.3. **The interaction potential.** Let $\varphi \in \mathrm{C}^2(0, +\infty)$ be an interaction potential, and let $\phi \in \mathrm{C}^2(\mathbb{R}^2 \setminus \{0\})$ be defined as $\phi(r) := \varphi(|r|)$. The *internal atomistic energy* of a deformation $y \in \mathscr{Y}$ is given by

$$\mathscr{E}_{\rm a}(y) = \sum_{b\in\mathcal{B}} \phi(D_b y) \tag{2.4}$$

$$= \sum_{x\in\mathcal{L}} \sum_{x'\in\mathcal{L}^\#\setminus\{x\}} \varphi\big(|y(x') - y(x)|\big). \tag{2.5}$$

It is crucial in our analysis that $\phi(r)$ and its derivatives decay rapidly as $|r| \to +\infty$. For example, our analysis is invalid for the slowly decaying Coulomb interactions. To avoid technicalities associated with the interaction decay altogether we assume throughout that there exists a cut-off radius $s_{\rm cut} > 0$ such that $\varphi(s) = 0$ for all $s \geq s_{\rm cut}$. The most commonly employed intermolecular potentials satisfy this property.

Despite the existence of a cut-off radius, we will need to quantify the decay within the interaction range. To that end we define

$$M_2 : (0, +\infty) \to [0, +\infty], \qquad M_2(s) = \sup_{\substack{r\in\mathbb{R}^2 \\ |r|\geq s}} \|\phi''(r)\|, \tag{2.6}$$

were $\phi'' : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^{2\times 2}$ denotes the 2nd Frechet derivative of ϕ, and $\|\cdot\|$ the operator norm of a matrix. We remark that, written in terms of φ, we have $M_2(s) = \sup_{t\geq s} \big(\big|\frac{\varphi''(t)}{t^2}\big|^2 + \big|\frac{\varphi'(t)}{t}\big|^2 \big)^{1/2}$.

Remark 2.1. (a) The more general form of the interaction potential admitted by (2.4) is useful since it includes plane-strain models of 3D crystals [33]. While our

consistency results remain valid for this general form of the interaction potential, the stability analysis relies more heavily on the specific form $\phi(r) = \varphi(|r|)$. Hence, for the purpose of the present paper, we understand (2.4) simply as a convenient replacement for the more conventional notation (2.5).

(b) External forces are often used to model, for example, a substrate or an indenter. In order to avoid the additional level of complexity they would introduce, we have decided against incorporating external forces. To obtain non-trivial solutions in our numerical experiments, we have instead allowed for defects in the atomistic lattice. □

2.4. **The variational problem.** The energy functional $\mathscr{E}_\mathrm{a}$ is twice continuously differentiable at every point $y \in \mathscr{Y}$. We understand the first variation $\delta\mathscr{E}_\mathrm{a}(y)$ as an element of $\mathscr{U}^*$, and the second variation $\delta^2\mathscr{E}_\mathrm{a}(y)$ as a linear operator from $\mathscr{U}$ to $\mathscr{U}^*$, formally defined by

$$\big\langle \delta\mathscr{E}_\mathrm{a}(y), u\big\rangle = \tfrac{\mathrm{d}}{\mathrm{d}t}\mathscr{E}_\mathrm{a}(y+tu)|_{t=0}, \quad \text{for } u \in \mathscr{U}, \text{ and}$$
$$\big\langle \delta^2\mathscr{E}_\mathrm{a}(y)u, v\big\rangle = \tfrac{\mathrm{d}}{\mathrm{d}t}\langle \delta\mathscr{E}_\mathrm{a}(y+tu), v\rangle|_{t=0}, \quad \text{for } u, v \in \mathscr{U}.$$

For some macroscopic strain $\mathsf{B} \in \mathbb{R}^{2\times 2}_+$, which shall be fixed throughout, the atomistic problem is to find

$$y_\mathrm{a} \in \operatorname{argmin} \mathscr{E}_\mathrm{a}(\mathscr{Y}_\mathsf{B}), \tag{2.7}$$

where "argmin" denotes the set of *local* minimizers. If $y_\mathrm{a} \in \mathscr{Y}_\mathsf{B}$ is a solution to (2.7), then it satisfies the first-order necessary optimality condition

$$\langle \delta\mathscr{E}_\mathrm{a}(y_\mathrm{a}), u\rangle = 0 \qquad \forall u \in \mathscr{U}. \tag{2.8}$$

3. A/C Coupling Method

The a/c method we present is motivated by the quasinonlocal QC method proposed in [32] and generalised in [7]. In the case of 1D second-neighbour pair interactions these methods take a particularly simple form amenable to rigorous analysis [4, 23, 27]. The generalisation to 2D finite range interactions we present here was first proposed by Shapeev [30]. Generalisations to 1D finite range interactions were independently developed in [15].

3.1. **Coarse-grained deformations and displacements.** The *atomistic region* is a closed polygonal set $\Omega_\mathrm{a} \subset \mathrm{int}(\Omega)$, and the *continuum region* is given by $\Omega_\mathrm{c} := \mathrm{clos}(\Omega \setminus \Omega_\mathrm{a}) \cap \Omega$. We assume throughout that all corners of Ω_a belong to $\mathcal{L}$, and that $\mathbb{V} \subset \mathrm{int}(\Omega_\mathrm{a})$.

Let $\mathcal{L}^\mathrm{c}_\mathrm{rep} \subset \mathcal{L} \cap \Omega_\mathrm{c}$ be a set of finite element nodes, or, in the language of the quasicontinuum method [21], *representative atoms*. We assume that the corners of the atomistic region belong to $\mathcal{L}^\mathrm{c}_\mathrm{rep}$. We also define $\mathcal{L}^\mathrm{a}_\mathrm{rep} = \mathcal{L} \cap \mathrm{int}(\Omega_\mathrm{a})$, and $\mathcal{L}_\mathrm{rep} := \mathcal{L}^\mathrm{a}_\mathrm{rep} \cup \mathcal{L}^\mathrm{c}_\mathrm{rep}$.

Let $\mathcal{T}^\mathrm{c}_h$ be a regular (and shape regular) triangulation of Ω_c with vertices belonging to $(\mathcal{L}^\mathrm{c}_\mathrm{rep})^\#$, which is extended periodically to a regular triangulation $(\mathcal{T}^\mathrm{c}_h)^\#$ of $\Omega^\#_\mathrm{c}$. An example of such a construction is displayed in Figure 2. We adopt the convention that lattice functions that are piecewise affine with respect to the triangulation $(\mathcal{T}^\mathrm{c}_h)^\#$ are understood as piecewise affine functions on *all of* $\Omega^\#_\mathrm{c}$, that is, they may be evaluated at any point $x \in \Omega^\#_\mathrm{c}$ and not only at lattice sites.

For each $T \in (\mathcal{T}^\mathrm{c}_h)^\#$ we define $h_T := \mathrm{diam}(T)$, and we define the mesh size function $h(x) := \max\{h_T : T \in (\mathcal{T}^\mathrm{c}_h)^\#, x \in T\}$, for $x \in \Omega^\#_\mathrm{c}$.

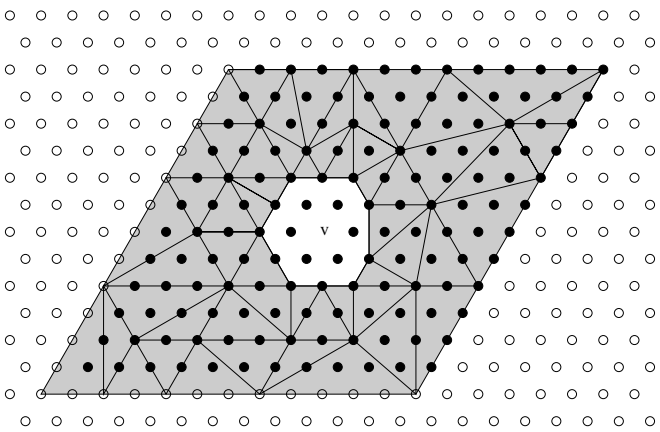


FIGURE 2. Example of a triangulation $\mathcal{T}_h^{\rm c}$ of the continuum region $\Omega_{\rm c}$ (shaded area), with nodes on $\partial\Omega_{\rm c}$ are such that the mesh can be extended periodically to a regular triangulation of $\Omega_{\rm c}^{\#}$. Note that the boundary of the atomistic region need not be aligned with nearest-neighbour directions.

Whenever we refer to the *shape regularity of* $\mathcal{T}_h^{\rm c}$ (and later $\mathcal{T}_h$), we mean the ratio between the largest and smallest angle between any two adjacent edges in $\mathcal{T}_h$. We will assume throughout that this is moderate.

We define the set of admissible coarse-grained displacements and deformations, respectively, as

$$\begin{aligned}
\mathscr{U}_h &= \big\{u_h \in \mathscr{U} :\ u_h \text{ is p.w. affine w.r.t. } \mathcal{T}_h^{\rm c}\big\},\\
\mathscr{Y}_{\mathsf{B},h} &= \big\{y_h \in \mathscr{Y} : y_h - y_{\mathsf{B}} \in \mathscr{U}_h \text{ and } \mu_{\rm c}(y_h) > 0\big\}, \quad \text{and}\\
\mathscr{Y}_h &= \textstyle\bigcup_{\mathsf{B}\in\mathbb{R}^{2\times 2}_+} \mathscr{Y}_{\mathsf{B},h},
\end{aligned}$$

where $\mu_{\rm c}$ is defined by

$$\mu_{\rm c}(y_h) := \inf_{\substack{x,x'\in\Omega_{\rm c}\\ x\neq x'}} \frac{|y_h(x)-y_h(x')|}{|x-x'|} \leq \operatorname*{ess.inf}_{x\in\Omega_{\rm c}} \min_{\substack{r\in\mathbb{R}^2\\ |r|=1}} |\nabla y_h(x) r|. \tag{3.1}$$

Note that, since a continuous interpolant of an invertible atomistic deformation need not necessarily be invertible, we are requiring a more stringent invertibility condition on coarse-grained deformations y_h.

Finally, we define the nodal interpolation operator $I_h : \mathscr{U} \to \mathscr{U}_h$ by

$$I_h u(x) = u(x) \qquad \forall x \in \mathcal{L}_{\rm rep},$$

extended to deformations by $I_h y - y_{\mathsf{B}} = I_h(y - y_{\mathsf{B}})$, for $y \in \mathscr{Y}_{\mathsf{B}}$.

3.2. **Bond integral formulation.** There are two steps in the construction of the a/c method. First, all bonds b that are entirely contained within the continuum region are replaced by line integrals. We collect these bonds into the set

$$\mathcal{B}_{\rm c} := \big\{b \in \mathcal{B} : \operatorname{int}(b) \subset \operatorname{int}(\Omega_{\rm c}^{\#})\big\},$$

where $\operatorname{int}(b)$ denotes the *relative interior* of a bond $b \in \mathbb{B}^{\#}$. The set of atomistic bonds is the complement $\mathcal{B}_{\rm a} := \mathcal{B} \setminus \mathcal{B}_{\rm c}$.

For any function v that is measurable on the segment $b = (x, x + r_b)$, we define the *bond integral*

$$\fint_b v \,\mathrm{db} = \fint_x^{x+r_b} v \,\mathrm{db} = \int_0^1 v(x + t r_b)\,\mathrm{d}t.$$

For any function $v_h \in \mathscr{U}_h \cup \mathscr{Y}_h$ the following one-sided directional derivatives are well-defined at almost every point of $b = (x, x + r)$:

$$\nabla_b v_h(x) := \nabla_r v_h(x) := \lim_{t \searrow 0} \frac{v_h(x + tr) - v_h(x)}{t}.$$

If x lies in the interior of an element T, then v_h is differentiable at x and hence $\nabla_r v_h(x) = (\nabla v_h(x))\, r$. If x lies on an edge or a vertex of the triangulation, the one-sided directional derivative of a continuous piecewise affine function is still well-defined. The directional derivative $\nabla_r v_h(x)$ is only undefined at points $x \in \partial\Omega_{\mathrm{a}}$ if r points to the interior of Ω_{a}. For future reference we note the following useful identity:

$$D_r y_h(x) = \fint_x^{x+r} \nabla_r y_h \,\mathrm{db}, \qquad \text{for } y \in \mathscr{Y}, x \in \mathbb{L}^{\#}, r \in \mathbb{L}_*. \tag{3.2}$$

Using this notation we see that if ∇y_h has small variation along the bond b, then $D_b y_h \approx \nabla_b y_h(x)$ for all $x \in \mathrm{int}(b)$, and hence we can make the following approximation:

$$\phi(D_b y_h) \approx \fint_b \phi(\nabla_b y_h) \,\mathrm{db}, \tag{3.3}$$

which naturally leads to the following definition of an a/c coupling method, which is labelled the ECC method in [30]:

$$\mathscr{E}_{\mathrm{ac}}(y_h) = \sum_{b \in \mathcal{B}_{\mathrm{a}}} \phi(D_b y_h) + \sum_{b \in \mathcal{B}_{\mathrm{c}}} \fint_b \phi(\nabla_b y_h) \,\mathrm{db}. \tag{3.4}$$

We will use this formulation of the a/c method in our analysis, however, it does not reduce the complexity of the energy evaluation, which is the purpose of the next section.

It is again easy to see that $\mathscr{E}_{\mathrm{ac}}$ is twice continuously differentiable in $\mathscr{Y}_{\mathsf{B},h}$, for all $\mathsf{B} \in \mathbb{R}^{2\times 2}_+$, and we define the first and second variations $\delta\mathscr{E}_{\mathrm{ac}}$ and $\delta^2\mathscr{E}_{\mathrm{ac}}$ analogously to $\delta\mathscr{E}_{\mathrm{a}}$ and $\delta^2\mathscr{E}_{\mathrm{a}}$ in §2.4.

3.3. **Practical formulation.** To make the a/c energy (3.4) "practical", we need to rewrite it in terms of volume integrals over the Cauchy–Born stored energy density. The tool for achieving this is the bond density lemma [30]. This result is false for general tetrahedra in 3D.

For any polygonal set $U \subset \mathbb{R}^2$ we define its *characteristic function*

$$\chi_U(x) = \lim_{t \to 0} \tfrac{|U \cap B_t(x)|}{|B_t(x)|} \qquad \text{for } x \in \mathbb{R}^2, \tag{3.5}$$

where $B_t(x)$ denotes the closed disk with radius t and centre x. From this definition it follows that, if $U_1, U_2 \subset \mathbb{R}^2$ are polygonal sets with $|\mathrm{int}(U_1) \cap \mathrm{int}(U_2)| = 0$, then $\chi_{U_1 \cup U_2} = \chi_{U_1} + \chi_{U_2}$.

The following result is a reformulation of [30, Lemma 4.4] for the triangular lattice with periodic boundary conditions.

Lemma 3.1 (Bond-Density Lemma). *Let $T \subset \mathrm{clos}(\Omega)$ be a non-degenerate triangle with vertices belonging to $\mathbb{L}^\#$, and let $r \in \mathbb{L}_*$, then*

$$\sum_{x\in\mathbb{L}} \fint_x^{x+r} \chi_{T^\#} \,\mathrm{db} = \frac{1}{\det \mathsf{A}_6}\,|T|.$$

Proof. A change of coordinates $x \mapsto \mathsf{A}_6 x$ in [30, Lemma 4.4] yields the following bond density formula for the full infinite lattice:

$$\sum_{x\in\mathbb{L}^\#} \fint_x^{x+r} \chi_{T} \,\mathrm{db} = \frac{1}{\det \mathsf{A}_6}\,|T|.$$

Splitting the lattice sum over copies of the cell $\mathbb{L}$, we obtain

$$\frac{1}{\det \mathsf{A}_6}\,|T| = \sum_{x\in\mathbb{L}^\#} \fint_x^{x+r} \chi_{T} \,\mathrm{db} = \sum_{x\in\mathbb{L}} \sum_{z\in N\mathbb{L}^\#} \fint_{x+z}^{x+z+r} \chi_{T} \,\mathrm{db}.$$

Upon shifting the integration variable by $-z$, we can rewrite this as

$$\frac{1}{\det \mathsf{A}_6}\,|T| = \sum_{x\in\mathbb{L}} \fint_x^{x+r} \sum_{z\in N\mathbb{L}^\#} \chi_{T-z} \,\mathrm{db} = \sum_{x\in\mathbb{L}} \fint_x^{x+r} \chi_{T^\#} \,\mathrm{db}. \qquad \square$$

Equipped with the periodic bond-density lemma, we derive a practical formulation of the a/c method (3.4). The proof of this result for Dirichlet boundary conditions is contained in [30]; the modifications for the periodic case are straightforward [25, App. A].

Theorem 3.2. *The energy $\mathscr{E}_{\mathrm{ac}}$, defined in* (3.4)*, can be rewritten as*

$$\mathscr{E}_{\mathrm{ac}}(y_h) = \sum_{b\in\mathcal{B}_{\mathrm{a}}} \phi(D_b y_h) + \int_{\Omega_{\mathrm{c}}} W(\nabla y_h)\,\mathrm{dV} + \Phi_{\mathrm{i}}(y_h), \quad \textit{where} \tag{3.6}$$

$$\Phi_{\mathrm{i}}(y_h) := -\sum_{b\in\mathbb{B}\setminus\mathcal{B}_{\mathrm{c}}} \fint_b \chi_{\Omega_{\mathrm{c}}^\#}\phi(\nabla_b y_h)\,\mathrm{db},$$

and where $W : \mathbb{R}^{2\times 2} \to \mathbb{R}\cup\{+\infty\}$ is the Cauchy–Born stored energy function,

$$W(\mathsf{F}) := \frac{1}{\det \mathsf{A}_6} \sum_{r\in\mathbb{L}_*} \phi(\mathsf{F}r).$$

Remark 3.1. While the bond-integral formulation (3.4) is easily extended to higher dimensions and to higher order finite element spaces, Theorem 3.2 holds only for piecewise affine trial functions in 2D. However, Shapeev [29] has developed an efficient algorithm for the evaluation of (3.4) in 3D as well. $\square$

3.4. **The coarse grained variational problem.** To apply the a/c method we compute

$$y_{\rm ac} \in \operatorname{argmin} \mathscr{E}_{\rm ac}(\mathscr{Y}_{\mathsf{B},h}). \tag{3.7}$$

If $y_{\rm ac} \in \mathscr{Y}_{\mathsf{B},h}$ is a solution to (3.7), then it satisfies the *first- and second-order necessary optimality conditions*

$$\big\langle \delta\mathscr{E}_{\rm ac}(y_{\rm ac}), u_h \big\rangle = 0 \qquad \forall u_h \in \mathscr{U}_h, \quad \text{and} \tag{3.8}$$

$$\big\langle \delta^2\mathscr{E}_{\rm ac}(y_{\rm ac})u_h, u_h \big\rangle \geq 0 \qquad \forall u_h \in \mathscr{U}_h. \tag{3.9}$$

Condition (3.9) is insufficient for error estimates; hence we will aim to prove the stronger *second-order sufficient optimality condition*

$$\big\langle \delta^2\mathscr{E}_{\rm ac}(y_{\rm ac})u_h, u_h \big\rangle \geq \gamma \|\nabla u_h\|^2_{\mathrm{L}^2(\Omega)} \qquad \forall u_h \in \mathscr{U}_h \tag{3.10}$$

for some $\gamma > 0$, where the norm $\|\nabla u_h\|^2_{\mathrm{L}^2(\Omega)}$ is yet to be defined for $u_h \in \mathscr{U}_h$. The choice of norm on the right-hand side of (3.10) is motivated by the fact that the equations (3.8) have a similar structure as finite element discretisations of second-order elliptic equations.

3.5. **Brief outline of the error analysis.** We give a brief sketch of the main result, Theorem 7.1, to motivate the subsequent technical details that we provide in §5–§7. We stress that this discussion is schematic, and that some steps are not properly defined at this point.

Let $y_{\rm a}$ be a solution of (2.7), and $y_{\rm ac}$ a solution of (3.7), and assume that $y_{\rm a}, y_{\rm ac}$, and $I_h y_{\rm a}$ are "close" in a sense to be made precise. Suppose, moreover, that (3.10) holds. Let $e_h := I_h y_{\rm a} - y_{\rm ac}$, then we can estimate

$$\begin{aligned}\gamma\|\nabla e_h\|^2_{\mathrm{L}^2(\Omega)} &\leq \big\langle \delta^2\mathscr{E}_{\rm ac}(y_{\rm ac})e_h, e_h\big\rangle \\ &\approx \big\langle \delta\mathscr{E}_{\rm ac}(I_h y_{\rm a}) - \delta\mathscr{E}_{\rm ac}(y_{\rm ac}), e_h\big\rangle = \langle \delta\mathscr{E}_{\rm ac}(I_h y_{\rm a}), e_h\rangle.\end{aligned}$$

The first inequality in the above estimate is the focus of the stability analysis in §6. The purpose of the consistency analysis §5 is to estimate

$$\big\langle \delta\mathscr{E}_{\rm ac}(I_h y_{\rm a}), e_h\big\rangle \leq \mathcal{E}^{\rm cons}\|\nabla e_h\|_{\mathrm{L}^2(\Omega)},$$

which immediately yields an *a priori* error estimate:

$$\|\nabla I_h y_{\rm a} - \nabla y_{\rm ac}\|_{\mathrm{L}^2(\Omega)} \lesssim \gamma^{-1}\mathcal{E}^{\rm cons}.$$

In §4 we will give an interpretation to $\nabla y_{\rm a}$, and establish interpolation error estimates, so that we can also estimate $\|\nabla y_{\rm a} - \nabla I_h y_{\rm a}\|_{\mathrm{L}^2(\Omega)}$. In §7, we will make the above arguments rigorous, and in addition establish an error estimate for the energy.

4. Auxiliary Results

4.1. **Extension to the vacancy set.** A substantial simplification of the subsequent analysis and notation can be achieved if we extend all function values to the vacancy set $\mathbb{V}$. Other approaches we have considered are significantly more technical and would yield only minor improvements. An altogether different approach might be required to extend the analysis to more general classes of defects.

We define the extension operator as the solution of a variational problem. Let the set of all displacement extensions be given by

$$\mathscr{U}_E := \big\{ v : \mathbb{L}^\# \to \mathbb{R}^2 : v(x + N\mathrm{a}_j) = v(x) \text{ for } x \in \mathbb{L}^\#, j = 1, 2\big\};$$

then, for $u \in \mathscr{U}$, we define

$$Eu := \operatorname*{argmin}_{\substack{v \in \mathscr{U}_E \\ v = u \text{ on } \mathcal{L}}} \Phi_{\mathbb{B}_{\rm nn}}(v), \quad \text{where} \quad \Phi_{\mathbb{B}_{\rm nn}}(v) := \sum_{b \in \mathbb{B}_{\rm nn}} \big| r_b \cdot D_b v \big|^2, \tag{4.1}$$

where

$$\mathbb{B}_{\rm nn} := \big\{ (x, x+r) : x \in \mathbb{L},\, r \in \mathbb{L}_{\rm nn} \big\}$$

is the set of nearest-neighbour bonds. This definition of the extension operator is motivated by the stability analysis, more precisely, the definition of the *vacancy stability index* in §6.1.

For the sake of simplicity of notation, we will identify $Ew \equiv w$, except where we need to strictly distinguish the original function w and its extension.

Proposition 4.1. *The variational problem* (4.1) *has a unique solution, that is, the extension operator E is a well-defined linear operator from $\mathscr{U}$ to $\mathscr{U}_E$.*

Proof. To prove that (4.1) has a unique solution it is sufficient to show that $\Phi_{\mathbb{B}_{\rm nn}}$ is a positive definite quadratic form on the affine subspace of $\mathscr{U}_E$ defined through the constraint $v = u$ on $\mathcal{L}$. The linearity is a straightforward consequence.

To establish this, we need to employ notation that will be properly defined in §4.2: let $\mathcal{T}_{\rm a}^{\#}$ denote the canonical triangulation of $\mathbb{L}^{\#}$, and, for each $v \in \mathscr{U}_E$, let $\bar{v}$ denote the corresponding continuous piecewise affine interpolant. In particular, we then have $D_b v = \nabla_b \bar{v}$ for all bonds $b \in \mathbb{B}_{\rm nn}$. Applying the bond density lemma, and (2.2), we obtain

$$\Phi_{\mathbb{B}_{\rm nn}}(v) = \int_\Omega \sum_{r \in \mathbb{L}_{\rm nn}} \big| r \cdot \nabla \bar{v} r \big|^2 \,{\rm dV} = \int_\Omega \Big\{ \tfrac{3}{2} \big| (\nabla \bar{v})^{\rm sym} \big|^2 + \tfrac{3}{4} \big| {\rm tr}(\nabla \bar{v}) \big|^2 \Big\} \,{\rm dV}.$$

Since $\bar{v}$ is fixed in the continuum region, Korn's inequality shows that $\Phi_{\mathbb{B}_{\rm nn}}$ is indeed coercive.

This proof shows that, in fact, E is defined through the solution of an isotropic linear elasticity problem, with boundary data provided on the edge of a suitably defined neighbourhood of the vacancy set. □

We extend the definition of E to include deformations $y \in \mathscr{Y}$, via $E(y_{\mathsf{B}} + u) = y_{\mathsf{B}} + Eu$ for all $\mathsf{B} \in \mathbb{R}^{2\times 2}_+$. We stress, however, that none of our results depend (explicitly or implicitly) on the extension of deformations. By contrast, the extension of displacements enters our analysis heavily.

4.2. **Micro-triangulation and extension of $\mathcal{T}_h^{\rm c}$.** The triangular lattice $\mathbb{L}^{\#}$ has a "canonical" triangulation $\mathcal{T}_{\rm a}^{\#}$, which is defined so that every nearest-neighbour bond is the edge of a triangle; see Figure 3. The subset of triangles $\tau \in \mathcal{T}_{\rm a}^{\#}$ that are contained in ${\rm clos}(\Omega)$ is denoted by $\mathcal{T}_{\rm a}$. We will assume throughout that the following assumption holds, but only cite it explicitly in the main results.

Assumption A. *The boundary of $\Omega_{\rm a}$ is aligned with edges of $\mathcal{T}_{\rm a}$ and the mesh size on $\partial\Omega_{\rm a}$ is equal to the lattice spacing.*

Assumption A implies that any microelement $\tau \in \mathcal{T}_{\rm a}$ must belong either entirely to $\Omega_{\rm a}$ or to $\Omega_{\rm c}$. This yields a natural extension $\mathcal{T}_h$ of $\mathcal{T}_h^{\rm c}$, which is obtained by adding all micro-elements $\tau \in \mathcal{T}_{\rm a}$, $\tau \subset \Omega_{\rm a}$, so that $\mathcal{T}_h$ and $\mathcal{T}_{\rm a}$ coincide in $\Omega_{\rm a}$. The requirement that the mesh size on $\partial\Omega_{\rm a}$ is equal to the lattice spacing implies that

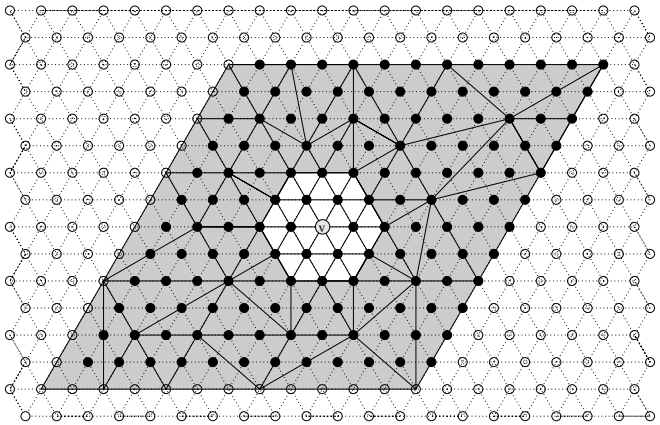


FIGURE 3. The micro-triangulation $\mathcal{T}_{\rm a}$ (dotted lines) and the extension $\mathcal{T}_h$ of the macro-triangulation to the atomistic domain. Note that in $\Omega_{\rm a}$, $\mathcal{T}_h$ coincides with $\mathcal{T}_{\rm a}$ and has no hanging nodes.

the extended mesh $\mathcal{T}_h$ *has no hanging nodes*. We emphasize that our subsequent analysis is valid only for this smaller family of meshes than the a/c coupling method was formulated for (compare Figure 2 with Figure 3).

The definitions of the element size h_T, the mesh size function $h(x)$, and the shape regularity, from §3.1, are extended to $\mathcal{T}_h$ and $\mathcal{T}_h^{\#}$.

For any lattice function $w : \mathbb{L}^{\#} \to \mathbb{R}^k$ we define the P1 micro-interpolant $\bar{w}$, that is, $\bar{w} \in \mathrm{W}^{1,\infty}_{\rm loc}(\mathbb{R}^2)^k$ and $\bar{w}(x) = w(x)$ on the lattice sites $x \in \mathbb{L}^{\#}$. In particular, the gradient $\nabla\bar{w}$, which is a piecewise constant function, is also well-defined.

Note that $y_h \in \mathscr{Y}_h$ is interpreted as the continuous P1 interpolant with respect to the mesh $\mathcal{T}_h$ (the *macro-interpolant*), while $\bar{y}_h$ is understood as the P1 interpolant with respect to the mesh $\mathcal{T}_{\rm a}$ (the *micro-interpolant*). In our analysis we will require some technical results comparing $\bar{y}_h$ and y_h. Lemma 4.2 gives a global comparison result, while a local variant is established in Lemma 4.5 below.

Lemma 4.2. *Let $y_h \in \mathscr{Y}_h$, and $p \in [1, \infty]$; then*

$$\|\nabla\bar{y}_h\|_{\mathrm{L}^p(\Omega)} \leq \bar{C}_\Omega \|\nabla y_h\|_{\mathrm{L}^p(\Omega)}, \tag{4.2}$$

where $\bar{C}_\Omega = \max(3^{(p-2)/(2p)}, 3^{(2-p)/(2p)}) \leq \sqrt{3}$.

Proof. The result follows from an argument analogous to the proof of [22, Lemma 2]. We present the details in [25, App. A]. □

4.3. **$\mathrm{W}^{2,\infty}$-conforming interpolants.** Smoothness of the atomistic solution in the continuum region is one of the key requirements for error estimates in a/c methods [6, 23]. In previous 1D analyses of a/c methods smoothness was measured via second- and third-order finite differences. A direct extension of this approach is technically and notationally demanding; hence, we propose to measure smoothness of discrete maps in terms of the smoothness of $\mathrm{W}^{2,\infty}$-conforming interpolants. In fact, it turns out that our analysis requires no explicit construction, and we therefore define the class of all $\mathrm{W}^{2,\infty}$-conforming interpolants of deformations $y \in \mathscr{Y}_{\mathsf{B}}$, $\mathsf{B} \in$

$\mathbb{R}^{2\times 2}_+$:

$$\Pi_2(y) := \big\{ \tilde{y} \in \mathrm{W}^{2,\infty}(\mathbb{R}^2)^2 : \tilde{y}(x) = y(x) \text{ for all } x \in \mathbb{L}^\#, \text{ and } \tilde{y}(x + N\mathrm{a}_j) = \mathsf{B}(N\mathrm{a}_j) + \tilde{y}(x) \text{ for all } x \in \mathbb{R}^2, j = 1, 2\big\}.$$

Lemma 4.3 (Interpolation Error Estimates). *Let $p \in [1,\infty]$, then there exists a constant $\tilde{C}_h$, depending only on p and on the shape regularity of $\mathcal{T}_h$, such that, for all $y \in \mathscr{Y}$,*

$$\big\|\nabla\tilde{y} - \nabla I_h y\big\|_{\mathrm{L}^p(T)} \le \tilde{C}_h h_T \big\|\nabla^2\tilde{y}\big\|_{\mathrm{L}^p(T)} \quad \forall T \in \mathcal{T}_h \quad \forall \tilde{y} \in \Pi_2(y). \tag{4.3}$$

Moreover, there exists a constant $\tilde{C}_\mathrm{a}$, depending only on p, such that

$$\big\|\nabla\tilde{y} - \nabla\bar{y}\big\|_{\mathrm{L}^p(\tau)} \le \tilde{C}_\mathrm{a} \big\|\nabla^2\tilde{y}\big\|_{\mathrm{L}^p(\tau)} \qquad \forall \tau \in \mathcal{T}_\mathrm{a} \quad \forall \tilde{y} \in \Pi_2(y). \tag{4.4}$$

Proof. Both estimates are standard [1]. The constant $\tilde{C}_\mathrm{a}$ is independent of the mesh quality since $\mathcal{T}_\mathrm{a}$ contains only a single element shape. □

Remark 4.1. We show in [25, Remark 4.1], that an explicit $\mathrm{W}^{2,\infty}$-interpolant $\tilde{y}_\mathrm{hct}$ can be constructed (using, e.g., the Hsieh–Clough–Tocher element) such that

$$c_1 \|\nabla^2\tilde{y}_\mathrm{hct}\|_{\mathrm{L}^p(\tau)} \le \big\|[\nabla\bar{y}]\big\|_{\mathrm{L}^p(\Gamma_\tau)} \le c_2 \|\nabla^2\tilde{y}_\mathrm{hct}\|_{\mathrm{L}^p(\omega_\tau)}, \tag{4.5}$$

where $c_i > 0$ are universal constants, Γ_τ is the union of all edges of the micro-triangulation touching τ, ω_τ the union of all micro-elements touching τ, and where $[\nabla\bar{y}]$ denotes the jump of $\nabla\bar{y}$ across micro-triangulation edges.

The inequalities in (4.5) show a local equivalence between second derivatives of "good" $\mathrm{W}^{2,\infty}$-conforming interpolants and jumps of $\nabla\bar{w}$, which one might consider the most natural measure of smoothness. □

4.4. **Notation for edges.** Several estimates in our consistency analysis will be phrased in terms of the jumps of ∇y_h, $y_h \in \mathscr{Y}_h$, across element edges, for which we now introduce the required notation: let $\mathcal{F}_h^\#$ denote the set of (closed) edges of the triangulation $\mathcal{T}_h^\#$, and let

$$\mathcal{F}_h := \big\{f \in \mathcal{F}_h^\# : \mathrm{int}(f) \subset \Omega\big\}, \quad \text{and} \quad \mathcal{F}_h^\mathrm{c} := \big\{f \in \mathcal{F}_h : f \not\subset \Omega_\mathrm{a}\big\},$$

where, here and throughout, $\mathrm{int}(f)$ denotes the relative interior of an edge f. That is, the set $\mathcal{F}_h$ includes one periodic copy of all element edges contained in Ω, and $\mathcal{F}_h^\mathrm{c}$ excludes all edges that are subsets of Ω_a.

Let $f \in \mathcal{F}_h^\#$, $f = T_+ \cap T_-$, $T_\pm \in \mathcal{T}_h$, and suppose that $w : \mathrm{int}(T_+) \cup \mathrm{int}(T_-) \to \mathbb{R}^k$ has well-defined traces $w^\pm$ from $T^\pm$, then we define the jump $[w](x) := |w_+(x) - w_-(x)|$ for all $x \in \mathrm{int}(f)$. Whenever we write $\int_{\mathcal{F}_h^\mathrm{c}}$, $\mathrm{L}^p(\mathcal{F}_h^\mathrm{c})$, *etc.*, we identify $\mathcal{F}_h^\mathrm{c}$ with the union of its elements.

In the next lemma we provide a tool to estimate jumps across edges in terms of smooth interpolants. The proof is given in Appendix A.

Lemma 4.4. *Let $y \in \mathscr{Y}$ and $f \in \mathcal{F}_h^\#$, $f = T_+ \cap T_-$ for $T_\pm \in \mathcal{T}_h$; then*

$$\big\|[\nabla I_h y]\big\|_{\mathrm{L}^p(f)} \le C_f \big\|h^{1/p'}\nabla^2\tilde{y}\big\|_{\mathrm{L}^p(T_+\cup T_-)} \qquad \forall \tilde{y} \in \Pi_2(y), \text{ and} \tag{4.6}$$

$$\big\|[\nabla I_h y]\big\|_{\mathrm{L}^p(\mathcal{F}_h^\mathrm{c})} \le C_f 3^{1/p} \big\|h^{1/p'}\nabla^2\tilde{y}\big\|_{\mathrm{L}^p(\Omega_\mathrm{c})} \qquad \forall \tilde{y} \in \Pi_2(y). \tag{4.7}$$

where C_f depends only on the shape regularity of $\mathcal{T}_h$.

4.5. **Micro- and macro-interpolants.** The following local version of Lemma 4.2 and its corollary, Lemma 4.6, are motivated by gradient jumps estimates of Lemma 4.4. The proof of Lemma 4.5 is given in Appendix A. We remark that the constant $\bar{C}_{\rm a}$ is fairly moderate as the discussion at the end of the proof shows.

Lemma 4.5. *Let $y_h \in \mathscr{Y}_h$, $\tau \in \mathcal{T}_{\rm a}$, and $p \in [1,\infty]$; then*

$$\|\nabla \bar{y}_h\|_{\mathrm{L}^p(\tau)} \le \bar{C}_{\rm a}\Big(\|\nabla y_h\|^p_{\mathrm{L}^p(\tau)} + \big\|[\nabla y_h]\big\|^p_{\mathrm{L}^p(\mathcal{F}_h^{\#}\cap \mathrm{int}(\tau))}\Big)^{1/p}, \tag{4.8}$$

where $\bar{C}_{\rm a}$ depends only on the shape regularity of $\mathcal{T}_h$.

Combining Lemma 4.5 and Lemma 4.3, we obtain the following result, which is a critical ingredient of the analysis in §5.1.

Lemma 4.6. *Let $y \in \mathscr{Y}$ and $p \in [1,\infty]$; then*

$$\big\|\nabla \bar{y} - \nabla \overline{I_h y}\big\|_{\mathrm{L}^p(\Omega_{\rm c})} \le \bar{C}_{I_h} \big\|h \nabla^2 \tilde{y}\big\|_{\mathrm{L}^p(\Omega_{\rm c})} \qquad \forall \tilde{y} \in \Pi_2(y), \tag{4.9}$$

where $\bar{C}_{I_h}$ depends only on the shape regularity of $\mathcal{T}_h$.

Proof. We cannot immediately use the interpolation error estimates (4.3) and (4.4) to estimate the term $\|\nabla(\bar{y} - \overline{I_h y})\|_{\mathrm{L}^p(\Omega)}$, due to the occurrence of $\overline{I_h y}$. Instead, we first fix a micro-element $\tau \subset \Omega_{\rm c}$, define $z(x) := (\nabla \bar{y}|_\tau) x$ for all $x \in \mathbb{R}^2$, and use (4.8) to estimate

$$\begin{aligned}
\big\|\nabla(\bar{y} - \overline{I_h y})\big\|^p_{\mathrm{L}^p(\tau)} &= \big\|\nabla \overline{I_h(y-z)}\big\|^p_{\mathrm{L}^p(\tau)} \\
&\le \bar{C}_{\rm a}^p \Big[\big\|\nabla I_h(y-z)\big\|^p_{\mathrm{L}^p(\tau)} + \big\|[\nabla I_h(y-z)]\big\|^p_{\mathrm{L}^p(\mathcal{F}_h^{\rm c}\cap \mathrm{int}(\tau))}\Big] \\
&= \bar{C}_{\rm a}^p \Big[\big\|\nabla(I_h y - \bar{y})\big\|^p_{\mathrm{L}^p(\tau)} + \big\|[\nabla I_h y]\big\|^p_{\mathrm{L}^p(\mathcal{F}_h^{\rm c}\cap \mathrm{int}(\tau))}\Big].
\end{aligned}$$

We will next sum this estimate for all $\tau \in \mathcal{T}_{\rm a}$. Using the fact that $\bar{y} = I_h y$ in $\Omega_{\rm a}$, as well as the interpolation error estimates (4.3) and (4.4), and the jump estimate (4.7), we obtain, for any $\tilde{y} \in \Pi_2(y)$,

$$\begin{aligned}
&\big\|\nabla(\bar{y} - \overline{I_h y})\big\|_{\mathrm{L}^p(\Omega)} \le \bar{C}_{\rm a}\Big[\big\|\nabla(I_h y - \bar{y})\big\|_{\mathrm{L}^p(\Omega_{\rm c})} + \big\|[\nabla I_h y]\big\|_{\mathrm{L}^p(\mathcal{F}_h^{\rm c})}\Big] \\
&\le \bar{C}_{\rm a}\Big[\big\|\nabla(I_h y - \tilde{y})\big\|_{\mathrm{L}^p(\Omega_{\rm c})} + \big\|\nabla(\tilde{y} - \bar{y})\big\|_{\mathrm{L}^p(\Omega_{\rm c})} + \big\|[\nabla I_h y]\big\|_{\mathrm{L}^p(\mathcal{F}_h^{\rm c})}\Big] \\
&\le \bar{C}_{\rm a}\Big[\tilde{C}_h \|h\nabla^2 \tilde{y}\|_{\mathrm{L}^p(\Omega_{\rm c})} + \tilde{C}_{\rm a} \|\nabla^2 \tilde{y}\|_{\mathrm{L}^p(\Omega_{\rm c})} + C_f 3^{1/p} \|h^{1/p'} \nabla^2 \tilde{y}\|_{\mathrm{L}^p(\Omega_{\rm c})}\Big].
\end{aligned}$$

Since $h \ge 1$, the stated result follows. □

5. Consistency

Recall from our preliminary discussion in §3.5 that the *total consistency error* associated with the atomistic solution $y^{\rm a}$ is

$$\big\|\delta\mathscr{E}_{\rm ac}(I_h y^{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} = \big\|\delta\mathscr{E}_{\rm ac}(I_h y^{\rm a}) - \delta\mathscr{E}_{\rm a}(y^{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} =: \mathcal{E}_p^{\rm cons}(y^{\rm a}),$$

where, for $\Psi \in \mathscr{U}_h^*$, the negative Sobolev norm is defined as

$$\|\Psi\|_{\mathrm{W}_h^{-1,p}} := \sup_{\substack{u_h \in \mathscr{U}_h \\ \|\nabla u_h\|_{\mathrm{L}^{p'}(\Omega)} = 1}} \big\langle \Psi, u_h \big\rangle.$$

In this section we prove the following estimate. Our assumption that ϕ has a finite cut-off radius (see §2.3) guarantees that $C^{\rm cons}$ is finite.

Theorem 5.1 (Consistency). *Suppose that Assumption A holds. Let $y \in \mathscr{Y}$ such that $\mu_{\rm a}(y) > 0$ and $\mu_{\rm c}(I_h y) > 0$. Then, for each $p \in [1,\infty]$, we have*

$$\mathcal{E}_p^{\rm cons}(y) \le C^{\rm cons} \inf_{\tilde y \in \Pi_2(y)} \|h\nabla^2 \tilde y\|_{\mathrm{L}^p(\Omega_{\rm c})}, \tag{5.1}$$

where $C^{\rm cons}$ depends only on $\mu_a(y_{\rm a})$, on $\mu_{\rm c}(I_h y)$, and on the shape regularity of $\mathcal{T}_h$.

Outline of the proof. To prove this result, we first split the consistency error into a *coarsening error* and a *modelling error*:

$$\begin{aligned}
\mathcal{E}_p^{\rm cons}(y) &= \big\|\delta\mathscr{E}_{\rm ac}(I_h y) - \delta\mathscr{E}_{\rm a}(y)\big\|_{\mathrm{W}_h^{-1,p}} \\
&\le \big\|\delta\mathscr{E}_{\rm ac}(I_h y) - \delta\mathscr{E}_{\rm a}(I_h y)\big\|_{\mathrm{W}_h^{-1,p}} + \big\|\delta\mathscr{E}_{\rm a}(I_h y) - \delta\mathscr{E}_{\rm a}(y)\big\|_{\mathrm{W}_h^{-1,p}}, \\
&=: \mathcal{E}_p^{\rm model}(y) + \mathcal{E}_p^{\rm coarse}(y).
\end{aligned}$$

We note that, since we estimate the modelling error at the interpolant $I_h y$, the mesh dependence is not entirely removed from $\mathcal{E}^{\rm model}$.

The estimate for the coarsening error is given in Lemma 5.4, and the estimate for the modelling error in Lemma 5.9, which together yield (5.1) with $C^{\rm cons} = C^{\rm coarse} + C^{\rm model}$. Note that we have ignored the improved mesh size dependence of the modelling error, and estimated $1 \le h$ to obtain $\mathcal{E}_p^{\rm model}(y) \le C^{\rm model}\|h\nabla^2\tilde y\|_{\mathrm{L}^p(\Omega_{\rm c})}$ for all $\tilde y \in \Pi_2(y)$. □

Remark 5.1. The details of the proof of Theorem 5.1 are technically involved, due to the relatively weak assumptions that we made on the mesh $\mathcal{T}_h$, as well as the fact that we insisted to estimate the consistency error in terms of $\|h\nabla^2\tilde y\|_{\mathrm{L}^p(\Omega_{\rm c})}$ only. A simplified argument can be given if weaker estimates are sufficient; see [25, App. B]. □

5.1. **Coarsening error.** The ingredients in the coarsening error estimate are a local Lipschitz bound on $\delta\mathscr{E}_{\rm a}$, and the interpolation error estimate of Lemma 4.6. We begin by stating a useful auxiliary result.

Lemma 5.2. *Let $r \in \mathbb{L}_*, q \in [1,\infty), u_h \in \mathscr{U}_h$ and $u \in \mathscr{U}$, then*

$$\sum_{x\in\mathbb{L}} \big|D_r u_h(x)\big|^q \le \sum_{x\in\mathbb{L}} \fint_x^{x+r} |\nabla_r u_h|^q \,\mathrm{db} = \tfrac{1}{\det \mathsf{A}_6}\|\nabla_r u_h\|_{\mathrm{L}^q(\Omega)}^q, \text{ and} \tag{5.2}$$

$$\sum_{x\in\mathbb{L}} \big|D_r u(x)\big|^q \le \sum_{x\in\mathbb{L}} \fint_x^{x+r} |\nabla_r \bar u|^q \,\mathrm{db} = \tfrac{1}{\det \mathsf{A}_6}\|\nabla_r \bar u\|_{\mathrm{L}^q(\Omega)}^q. \tag{5.3}$$

Proof. The result is a straightforward application of the bond density lemma. We give the proof for (5.2), since (5.3) is a particular case.

Jensen's inequality implies the inequality in (5.2):

$$\big|D_r u_h(x)\big|^q = \left|\fint_x^{x+r} \nabla_r u_h \,\mathrm{db}\right|^q \le \fint_x^{x+r} \big|\nabla_r u_h\big|^q \,\mathrm{db}.$$

Using the fact that that $\{\chi_T^{\#} : T \in \mathcal{T}_h\}$ is a partition of unity, continuity of $\nabla_r u_h$ across faces with direction r, and Lemma 3.1, we have

$$\begin{aligned}\sum_{x\in\mathbb{L}} \fint_x^{x+r} |\nabla_r u_h|^q \,\mathrm{db} &= \sum_{T\in\mathcal{T}_h} \sum_{x\in\mathbb{L}} \fint_x^{x+r} \chi_{T^\#} |\nabla_r u_h|^q \,\mathrm{db} \\ &= \sum_{T\in\mathcal{T}_h} |\nabla_r u_h|_T|^q \sum_{x\in\mathbb{L}} \fint_x^{x+r} \chi_{T^\#} \,\mathrm{db} \\ &= \frac{1}{\det \mathsf{A}_6} \sum_{T\in\mathcal{T}_h} |T| |\nabla_r u_h|_T|^q. \qquad \square\end{aligned}$$

The next auxiliary result is a Lipschitz bound on $\delta\mathscr{E}_{\mathrm{a}}$.

Lemma 5.3. *Let $y, z \in \mathscr{Y}$, and $\mu := \min\{\mu_{\mathrm{a}}(y), \mu_{\mathrm{a}}(z)\} > 0$; then*

$$\big|\big\langle \delta\mathscr{E}_{\mathrm{a}}(y) - \delta\mathscr{E}_{\mathrm{a}}(z), u_h\big\rangle\big| \le C_{\mathrm{L}} \big\|\nabla\bar{y} - \nabla\bar{z}\big\|_{\mathrm{L}^p(\Omega)} \|\nabla u_h\|_{\mathrm{L}^{p'}(\Omega)} \tag{5.4}$$

for all $u_h \in \mathscr{U}_h$, where $C_{\mathrm{L}} = C_{\mathrm{L}}(\mu) := \frac{1}{\det \mathsf{A}_6} \sum_{r\in\mathbb{L}_*} |r|^2 M_2(\mu|r|)$. The result remains true if u_h is replaced with $u \in \mathscr{U}$, and ∇u_h with $\nabla \bar{u}$.

Proof. Fix $u_h \in \mathscr{U}_h$ and $p \in (1,\infty)$; then

$$\begin{aligned}\big|\big\langle \delta\mathscr{E}_{\mathrm{a}}(y) - \delta\mathscr{E}_{\mathrm{a}}(z), u_h\big\rangle\big| &\le \sum_{b\in\mathcal{B}} \big|\phi'(D_b y) - \phi'(D_b z)\big| \, |D_b u_h| \\ &\le \sum_{b\in\mathcal{B}} M'_{|b|} \big|\tfrac{D_b y - D_b z}{|b|}\big| \, \big|\tfrac{D_b u_h}{|b|}\big|,\end{aligned}$$

where $M'_\rho = M_2(\mu\rho)\rho^2$. Let $w = y - z$, then, applying a Hölder inequality, we obtain that

$$\big|\big\langle \delta\mathscr{E}_{\mathrm{a}}(y) - \delta\mathscr{E}_{\mathrm{a}}(z), u_h\big\rangle\big| \le \left(\sum_{b\in\mathcal{B}} M'_{|b|} \big|\tfrac{D_b w}{|b|}\big|^p\right)^{1/p} \left(\sum_{b\in\mathcal{B}} M'_{|b|} \big|\tfrac{D_b u_h}{|b|}\big|^{p'}\right)^{1/p'}.$$

Each of the two groups can be estimated using Lemma 5.2, for example,

$$\begin{aligned}\sum_{b\in\mathcal{B}} M'_{|b|} \big|\tfrac{D_b w}{|b|}\big|^p &\le \sum_{b\in\mathbb{B}} M'_{|b|} \big|\tfrac{D_b w}{|b|}\big|^p = \sum_{r\in\mathbb{L}_*} M'_{|r|} |r|^{-p} \sum_{x\in\mathbb{L}} \big|D_r w(x)\big|^p \\ &\le \sum_{r\in\mathbb{L}_*} \tfrac{M'_{|r|}}{\det \mathsf{A}_6} |r|^{-p} \|\nabla_r \bar{w}\|_{\mathrm{L}^p(\Omega)}^p = \|\nabla\bar{w}\|_{\mathrm{L}^p(\Omega)}^p \sum_{r\in\mathbb{L}_*} \tfrac{M'_{|r|}}{\det \mathsf{A}_6}.\end{aligned}$$

By the same argument, using (5.2) instead of (5.3), we obtain

$$\sum_{b\in\mathcal{B}} M'_{|b|} \big|\tfrac{D_b u_h}{|b|}\big|^{p'} \le \tfrac{1}{\det \mathsf{A}_6} \sum_{r\in\mathbb{L}_*} M'_{|r|} \|\nabla u_h\|_{\mathrm{L}^{p'}(\Omega)}^{p'}.$$

This establishes (5.4) for $p \in (1,\infty)$. The cases $p \in \{1,\infty\}$ are obtained with minor modifications of the above argument. The proof for $u \in \mathscr{U}$ is analogous. $\square$

We can now formulate the coarsening error estimate.

Lemma 5.4. *Let $y \in \mathscr{Y}$ and $\mu := \min(\mu_{\mathrm{a}}(y), \mu_{\mathrm{a}}(I_h y)) > 0$; then,*

$$\mathcal{E}_p^{\mathrm{coarse}}(y) \le C^{\mathrm{coarse}} \big\|h\nabla^2 \tilde{y}\big\|_{\mathrm{L}^p(\Omega_{\mathrm{c}})}, \tag{5.5}$$

for all $p \in [1,\infty]$ and for all $\tilde{y} \in \Pi_2(y)$, where $C^{\mathrm{coarse}} = C_{\mathrm{L}}(\mu)\bar{C}_{I_h}$.

Proof. According to Lemma 5.3 we have

$$\big\langle \delta\mathscr{E}_{\rm a}(y) - \delta\mathscr{E}_{\rm a}(I_h y), u_h \big\rangle \le C_{\rm L} \big\| \nabla(\bar{y} - \overline{I_h y}) \big\|_{{\rm L}^p(\Omega)} \|\nabla u_h\|_{{\rm L}^{p'}(\Omega)}.$$

From Lemma 4.6 we obtain that

$$\|\nabla(\bar{y} - \overline{I_h y})\|_{{\rm L}^p(\Omega)} \le \bar{C}_{I_h} \big\| h\nabla^2 \tilde{y} \big\|_{{\rm L}^p(\Omega_{\rm c})} \qquad \forall \tilde{y} \in \Pi_2(y),$$

which yields (5.5) with $C^{\rm coarse} = C_{\rm L} \bar{C}_{I_h}$. □

Remark 5.2. With an alternative splitting of the consistency error (see, e.g., [27, 22]) it would have been necessary to estimate the coarsening error when $\mathscr{E}_{\rm a}$ is replaced with $\mathscr{E}_{\rm ac}$. In that case, we would have needed a Lipschitz estimate on $\delta\mathscr{E}_{\rm ac}$. Defining $\mathscr{E}_{\rm ac}(\bar{y})$ in a canonical way, our proof above is easily modified to yield

$$\begin{aligned} \big|\big\langle \delta\mathscr{E}_{\rm ac}(I_h y) - \delta\mathscr{E}_{\rm ac}(\bar{y}), u_h \big\rangle\big| \le \Big\{ &\sum_{b \in \mathcal{B}_{\rm a}} M'_{|b|} \fint_b \big| \nabla_b \overline{I_h y} - \nabla_b \bar{y} \big|^p \,{\rm db} \\ &+ \sum_{b \in \mathcal{B}_{\rm c}} M'_{|b|} \fint_b \big| \nabla_b I_h y - \nabla_b \bar{y} \big|^p \,{\rm db} \Big\}^{1/p} C_{\rm L}^{1/p'} \|\nabla u_h\|_{{\rm L}^{p'}}. \end{aligned}$$

The first group we can again convert into volume integrals and estimate using Lemma 4.6. However, the second group contains integrals over both macro- and micro-interpolants, and therefore cannot be converted into volume integrals using the bond density lemma.

However, as we show in [25, App. B], weaker (though technically less involved) estimates can be obtained in this way. □

5.2. **Modelling error.** For the majority of the modelling error analysis we can replace $I_h y$ by an arbitrary discrete deformation $y_h \in \mathscr{Y}_h$. Hence, we fix $y_h \in \mathscr{Y}_h$ such that $\mu := \min(\mu_a(y_h), \mu_{\rm c}(y_h)) > 0$. Moreover, we fix constants $a_r > 0$, $r \in \mathbb{L}_*$, which will be determined later, $a_b := a_{r_b}$ for all bonds $b \in \mathbb{B}$, and $M'_\rho := M_2(\mu\rho)\rho^2$ for all $\rho > 0$.

With this notation, and using (3.2), we have

$$\begin{aligned} \big\langle \delta\mathscr{E}_{\rm ac}(y_h) - \delta\mathscr{E}_{\rm a}(y_h), u_h \big\rangle &= \sum_{b \in \mathcal{B}_{\rm c}} \fint_b \phi'(\nabla_b y_h) \cdot \nabla_b u_h \,{\rm db} - \sum_{b \in \mathcal{B}_{\rm c}} \phi'(D_b y_h) \cdot D_b u_h \\ &= \sum_{b \in \mathcal{B}_{\rm c}} \fint_b \big[\phi'(\nabla_b y_h) - \phi'(D_b y_h)\big] \cdot \nabla_b u_h \,{\rm db} \\ &\le \sum_{b \in \mathcal{B}_{\rm c}} M_2(\mu|b|) \fint_b \big| \nabla_b y_h - D_b y_h \big| |\nabla_b u_h| \,{\rm db} \\ &= \sum_{b \in \mathcal{B}_{\rm c}} M'_{|b|} \fint_b \big(a_b^{-1} |b|^{-1} \big| \nabla_b y_h - D_b y_h \big| \big) \big(a_b |b|^{-1} |\nabla_b u_h| \big) \,{\rm db}. \end{aligned}$$

Following a similar procedure as in the proof of Lemma 5.3 (applying a Hölder inequality and Lemma 5.2), we obtain

$$\begin{aligned}
&\big\langle \delta\mathscr{E}_{\rm ac}(y_h) - \delta\mathscr{E}_{\rm a}(y_h), u_h\big\rangle \\
&\quad \le \Bigg(\sum_{b\in\mathcal{B}_{\rm c}} M'_{|b|}|b|^{-p} a_b^{-p} \fint_b \big|\nabla_b y_h - D_b y_h\big|^p \,{\rm d}b \Bigg)^{1/p} C_1^{1/p'} \|\nabla u_h\|_{{\rm L}^{p'}(\Omega)} \\
&\quad =: C_1^{1/p'} e(y_h) \|\nabla u_h\|_{{\rm L}^{p'}(\Omega)}, \qquad (5.6)
\end{aligned}$$

where $C_1 = \frac{1}{\det \mathsf{A}_6} \sum_{r\in\mathbb{L}_*} M'_{|r|} a_r^{p'}$, and where

$$\begin{aligned}
e(y_h)^p &:= \sum_{b\in\mathcal{B}_{\rm c}} M'_{|b|}|b|^{-p} a_b^{-p} e_b(y_h)^p, \\
e_b(y_h)^p &:= \fint_b \big|\nabla_b y_h - D_b y_h\big|^p \,{\rm d}b. \qquad (5.7)
\end{aligned}$$

We will estimate the terms $e_b(y_h)^p$ in terms of the jumps of ∇y_h. To that end, we define the jump sets

$$J(b) := \big\{ f \in \mathcal{F}_h : \#(f \cap {\rm int}(b)) = 1 \big\}. \qquad (5.8)$$

Faces parallel to b are ignored since the directional derivative $\nabla_{r_b} y_h$ is continuous across these faces. For each $f \in J(b)$ we define the weights

$$w_{b,f} := \begin{cases} 1, & \text{if } f \cap {\rm int}(b) \subset {\rm int}(f), \\ 1/2, & \text{otherwise;} \end{cases}$$

that is, $w_{b,f} = 1$ if b crosses f in its relative interior, and $w_{b,f} = 1/2$ if b crosses f at one of its endpoints. Finally we define the quantities

$$n_{\rm j}(b) := \sum_{f\in J(b)} w_{b,f}, \quad \text{and} \quad n_{\rm j}(r) := \max_{\substack{b\in\mathcal{B}_{\rm c} \\ r_b = r}} n_{\rm j}(b). \qquad (5.9)$$

Lemma 5.5. *Let $b \in \mathcal{B}_{\rm c}$, then*

$$e_b(y_h)^p \le n_{\rm j}(b)^{p-1} \sum_{f\in J(b)} w_{b,f} \big|[\nabla_b y_h]_f\big|^p. \qquad (5.10)$$

Proof. Define $\psi(t) = \nabla_b y_h(x + t r_b)$ and let $J_\psi \subset (0,1)$ be the set of jumps of ψ, then

$$e_b(y_h)^p = \int_0^1 \left| \psi(t) - \int_0^1 \psi(s)\,{\rm d}s \right|^p {\rm d}t. \qquad (5.11)$$

For any point $t \in (0,1) \setminus J_\psi$ we can estimate

$$\begin{aligned}
\left| \psi(t) - \int_0^1 \psi(s)\,{\rm d}s \right| &\le \int_0^1 \big|\psi(t) - \psi(s)\big|\,{\rm d}s \le \int_0^1 \int_{r\in(t,s)} |\psi'|\,{\rm d}r\,{\rm d}s \\
&\le \int_0^1 |\psi'|\,{\rm d}r = \sum_{r\in J_\psi} |\psi(r+) - \psi(r-)|,
\end{aligned}$$

where $|\psi'|\,\mathrm{d}r$ is understood as the measure that represents the distributional derivative of ψ. Inserting this estimate into (5.11), yields

$$e_b(y_h)^p \le \Big| \sum_{r\in J_\psi} |\psi(r+) - \psi(r-)| \Big|^p \le (\#J_\psi)^{p-1} \sum_{r\in J_\psi} \big|\psi(r+) - \psi(r-)\big|^p,$$

which translates directly into (5.10), in the case that b does not intersect any faces in their endpoints.

If b does intersect certain faces in endpoints then one replaces the path $\{x + tr_b : t \in (0,1)\}$ by two paths that "circle" around the endpoints, each weighted with a factor $1/2$. □

Recall the detail of the definition of $\mathcal{F}_h^{\rm c}$ from §4.4. Since only bonds $b \in \mathcal{B}_{\rm c}$ contribute to the consistency error, it follows that only jumps across faces $f \in \mathcal{F}_h^{\rm c}$ occur in the following estimate. Interchanging the order of summation, we obtain

$$\begin{aligned} e(y_h)^p &\le \sum_{b\in\mathcal{B}_{\rm c}} M'_{|b|} |b|^{-p} a_b^{-p} n_{\rm j}(b)^{p-1} \sum_{f\in J(b)} w_{b,f} \big|[\nabla_b y_h]_f\big|^p \\ &\le \sum_{r\in\mathbb{L}_*} M'_{|r|} |r|^{-p} a_r^{-p} n_{\rm j}(r)^{p-1} \sum_{\substack{b\in\mathcal{B}_{\rm c}\\ r_b = r}} \sum_{f\in J(b)} w_{b,f} \big|[\nabla_b y_h]_f\big|^p \\ &= \sum_{r\in\mathbb{L}_*} M'_{|r|} |r|^{-p} a_r^{-p} n_{\rm j}(r)^{p-1} \sum_{f\in\mathcal{F}_h^{\rm c}} n_{\rm cross}(f,r) \big|[\nabla_r y_h]_f\big|^p, \end{aligned} \tag{5.12}$$

where $n_{\rm cross}(f,r)$ is the (weighted) number of bonds b with direction r and crossing the face f; more precisely,

$$n_{\rm cross}(f,r) := \sum_{\substack{b\in\mathcal{B}_{\rm c}, r_b = r\\ f\in J(b)}} w_{b,f}.$$

Lemma 5.6. *Let $f \in \mathcal{F}_h^{\rm c}$ and $r \in \mathbb{L}_*$; then*

$$n_{\rm cross}(f,r) \le \tfrac{1}{\det \mathsf{A}_6} 2|r|\,\mathrm{length}(f). \tag{5.13}$$

Proof. Let $f = \{z + ts : t \in [0,1]\}$, and define the parallelogram

$$P = \{z + t_1 s + t_2 r : t_1 \in [0,1], t_2 \in (-1,1)\},$$

then we have

$$n_{\rm cross}(f,r) = \sum_{\substack{b\in\mathcal{B}_{\rm c}, r_b = r\\ f\in J(b)}} \fint_b \chi_P \,\mathrm{db} \le \sum_{x\in\mathbb{L}^\#} \fint_x^{x+r} \chi_P \,\mathrm{db} = \tfrac{1}{\det \mathsf{A}_6} |P|,$$

where, in the last equality, we have used the fact that P is the union of two triangles, which implies that the bond density lemma holds for P as well. To obtain the result we simply note that $|P| \le 2|r|\mathrm{length}(f)$. □

Estimate (5.13) and $|[\nabla_r y_h]_f| \le |r||[\nabla y_h]_f|$ yield

$$e(y_h)^p \le C_2 \Big(\sum_{f\in\mathcal{F}_h^{\rm c}} h_f \big|[\nabla y_h]_f\big|^p \Big), \tag{5.14}$$

where $C_2 = \frac{1}{\det \mathsf{A}_6} \sum_{r\in\mathbb{L}_*} 2M'_{|r|} |r| a_r^{-p} n_{\rm j}(r)^{p-1}$ and $h_f := \mathrm{length}(f)$.

Choosing the constants a_r so that $C_1 = C_2$,

$$2|r|a_r^{-p} n_{\rm j}(r)^{p-1} = a_r^{p'} = (2|r|)^{1/p} n_{\rm j}(r)^{1/p'},$$

we obtain

$$C_1 = C_2 = \tfrac{1}{\det \mathsf{A}_6} 2^{1/p} \sum_{r \in \mathbb{L}_*} M_2(\mu|r|)|r|^{2+1/p} n_{\rm j}(r)^{1/p'}. \tag{5.15}$$

To obtain a more explicit constant, we estimate $n_{\rm j}(r)$ next. The following result is unsurprising, but its proof rather technical; hence we have postponed it to Appendix A.

Lemma 5.7. *There exists a constant $C_{n_{\rm j}}$, which depends only on the shape regularity of $\mathcal{T}_h$, such that*

$$n_{\rm j}(b) \le C_{n_{\rm j}}(|b| + 1) \qquad \forall b \in \mathcal{B}_{\rm c}. \tag{5.16}$$

Combining (5.14), (5.15), and (5.16), we deduce the following intermediate result, which is interesting in its own right, since it could serve as a basis for *a posteriori* error estimates.

Lemma 5.8. *Let $y_h \in \mathscr{Y}_h$, $\mu := \min(\mu_{\rm a}(y_h), \mu_{\rm c}(y_h)) > 0$; then*

$$\big\langle \delta\mathscr{E}_{\rm ac}(y_h) - \delta\mathscr{E}_{\rm a}(y_h), u_h \big\rangle \le C_1^{\rm model} \big\| [\nabla y_h] \big\|_{{\rm L}^p(\mathcal{F}_h^{\rm c})} \|\nabla u_h\|_{{\rm L}^{p'}(\Omega)} \tag{5.17}$$

for all $u_h \in \mathscr{U}_h$, where $C_1^{\rm model} = C' \sum_{r \in \mathbb{L}_*} M_r(\mu|r|)|r|^3$ and C' depends only on the shape regularity of $\mathcal{T}_h$.

Applying Lemma 4.4 to estimate $\|[\nabla y_h]\|_{{\rm L}^p(\mathcal{F}_h^{\rm c})}$ in (5.17), we obtain the final modelling error estimate.

Lemma 5.9 (Modelling Error). *Let $y \in \mathscr{Y}$ and suppose that*

$$\mu := \min(\mu_{\rm a}(I_h y), \mu_{\rm c}(I_h y)) > 0;$$

then

$$\mathcal{E}_p^{\rm model}(y) \le C^{\rm model} \big\| h^{1/p'} \nabla^2 \tilde{y} \big\|_{{\rm L}^p(\Omega_{\rm c})}, \tag{5.18}$$

where $C^{\rm model} = C \sum_{r \in \mathbb{L}_*} M_2(\mu|r|)|r|^3$ and C depends only on the shape regularity of $\mathcal{T}_h$.

Remark 5.3. At first glance it may seem that the terms $\|[\nabla y_h]\|_{{\rm L}^p(\mathcal{F}_h^{\rm c})}$ in (5.17) and $\|h^{1/p'} \nabla^2 \tilde{y}\|_{{\rm L}^p(\Omega_{\rm c})}$ in (5.18) are not scale invariant. This is, however, deceiving. Since the atomic scale is 1 in our case, one should read (5.18) as

$$\big\| h^{1/p'} \nabla^2 \tilde{y} \big\|_{{\rm L}^p(\Omega_{\rm c})} = \big\| 1^{1/p} h^{1/p'} \nabla^2 \tilde{y} \big\|_{{\rm L}^p(\Omega_{\rm c})},$$

which is again scale invariant if 1 is scaled in the same way as h. Indeed, it can be checked that, had we formulated the entire analysis with scaled quantities $x \to \varepsilon x$, $y \to \varepsilon y$, and $\sum \to \varepsilon^2 \sum$, then we would have obtained $\|\varepsilon^{1/p} h^{1/p'} \nabla^2 \tilde{y}\|_{{\rm L}^p(\Omega_{\rm c})}$. □

6. Stability

6.1. **Main results.** The most natural notion of stability for variational problems is positivity of the second variation (at certain deformations of interest).

We first state the main stability result for the case of a homogeneous deformation and $\mathbb{V} = \emptyset$. This serves as reference point and motivation for the general stability result below, which has a more involved formulation.

Theorem 6.1. *Suppose that Assumption A holds, and that $\mathbb{V} = \emptyset$. Let $\mathsf{B} \in \mathbb{R}^{2\times 2}_+$ with singular values $0 < m \le M$; then*

$$\big\langle \delta^2 \mathscr{E}_{\rm ac}(y_{\mathsf{B}}) u_h, u_h \big\rangle \ge \gamma_{\rm hom} \|\mathsf{B}^\top \nabla u_h\|^2_{\L^2(\Omega)} \qquad \forall u_h \in \mathscr{U}_h,$$

where $\gamma_{\rm hom} = \gamma_{\rm hom}(m,M) := \min(\frac{3}{4}c + \frac{9}{4}c^\perp, \frac{9}{4}c + \frac{3}{4}c^\perp)$,

$$\begin{aligned} c^{(\perp)} = c^{(\perp)}(m,M) &:= \tfrac{1}{\det \mathsf{A}_6} \textstyle\sum_{n=1}^\infty c_n^{(\perp)}, \qquad \text{and} \\ c_n &:= \begin{cases} \min_{s\in[m,M]} \frac{\varphi''(s)}{s^2}, & n = 1, \\ 0 \wedge \min_{s\in[m,M]} \frac{\ell_n^2 \varphi''(s\ell_n)}{s^2}, & n > 1, \end{cases} \\ c_n^\perp &:= \begin{cases} \min_{s\in[m,M]} \frac{\varphi'(s)}{s^3}, & n = 1, \\ 0 \wedge \min_{s\in[m,M]} \frac{\ell_n \varphi'(s\ell_n)}{s^3}, & n > 1. \end{cases} \end{aligned} \tag{6.1}$$

Theorem 6.1 is a special case of Theorem 6.2 below, restricted to homogeneous lattices without defects. A direct proof can be given by first specializing the definition of $\mathcal{H}(y_h)$ in (6.11) to $y_h = y_{\mathsf{B}}$ and $\mathbb{V} = \emptyset$, and then applying Lemma 6.4, with $\overline{\mathcal{H}}$ replaced with $\mathcal{H}(y_{\mathsf{B}})$.

Our generalisation to defects uses the concept of a *vacancy stability index*. Using the extension operator $E : \mathscr{U} \to \mathscr{U}_E$ (see §4.1) we define

$$\kappa(\mathbb{V}) := \max\Big\{ k > 0 \,:\, \sum_{b\in\mathcal{B}_{\rm nn}} \big|r_b \cdot D_b u\big|^2 \ge k \sum_{b\in\mathbb{B}_{\rm nn}} \big|r_b \cdot D_b E u\big|^2 \text{ for } u \in \mathscr{U} \Big\}. \tag{6.2}$$

We present numerically computed lower bounds on $\kappa(\mathbb{V})$ in Table 1, and in §6.7 rigorously prove the bound $\kappa(\mathbb{V}) \ge 2/7$ if $\mathbb{V}$ consists of separated vacancies.

Remark 6.1 (Optimality of the extension operator). Recall the definition of $\Phi_{\mathbb{B}_{\rm nn}}$ from §4.1, and let $\Phi_{\mathcal{B}_{\rm nn}}$ be defined analogously, then (6.2) can be rewritten as

$$\kappa(\mathbb{V}) = \max\Big\{ k > 0 \,:\, \Phi_{\mathcal{B}_{\rm nn}}(u) \ge k \Phi_{\mathbb{B}_{\rm nn}}(Eu) \text{ for all } u \in \mathscr{U} \Big\}.$$

Since Eu is chosen to minimize the value of $\Phi_{\mathbb{B}_{\rm nn}}(Eu)$, it gives the largest possible stability index among all possible extensions.

Moreover, we can characterise $\kappa(\mathbb{V})$ in terms of an operator norm of E. Let $\mathscr{U}$ be equipped with the norm $\sqrt{\Phi_{\mathcal{B}_{\rm nn}}}$ and $\mathscr{U}_E$ with the norm $\sqrt{\Phi_{\mathbb{B}_{\rm nn}}}$, then $\kappa(\mathbb{V}) = \inf_{u\in\mathscr{U}\setminus\{0\}} \frac{\Phi_{\mathcal{B}_{\rm nn}}(u)}{\Phi_{\mathbb{B}_{\rm nn}}(Eu)} = \|E\|^{-2}_{L(\mathscr{U},\mathscr{U}_E)}$. □

Separation distance	4	8	12
$\mathbb{V} = \emptyset$	1		
Vacancies	0.28	0.39	0.41
Divacancies	0.16	0.26	0.29

TABLE 1. Numerically computed lower bounds on vacancy stability indices (rounded down to two significant digits) when $\mathbb{V}$ consists of either single vacancies, or divacancies separated by "separation distance".

To state the main stability result, we define a family of regions in the space of deformations: for $0 < m \le M$ and $\Delta > 0$ let

$$\begin{aligned}\mathscr{S}_{\mathsf{B},h}(m, M, \Delta) := \big\{ y_h \in \mathscr{Y}_{\mathsf{B},h} : \quad & \mu_{\rm a}(y_h) \ge m \text{ and } \mu_{\rm c}(y_h) \ge m; \\ & |D_b y_h| \le M|b| \ \forall b \in \mathcal{B}_{\rm a} \text{ and } \|\nabla y_h|_T\| \le M \ \forall T \in \mathcal{T}_h^{\rm c}; \\ |\mathsf{B}^{-1} D_b y_h - r_b| \le \Delta |b| \ \forall b \in \mathcal{B}_{\rm a} & \text{ and } \|\mathsf{B}^{-1}\nabla y_h|_T - \mathbb{1}\| \le \Delta \ \forall T \in \mathcal{T}_h^{\rm c}\big\}.\end{aligned}$$

Theorem 6.2. *Suppose that Assumption A holds and suppose that $y_h \in \mathscr{S}_{\mathsf{B},h}(m, M, \Delta)$ for $0 < m \le M$ and $0 \le \Delta \le \sqrt{\kappa(\mathbb{V})}/2$; then*

$$\big\langle \delta^2 \mathscr{E}_{\rm ac}(y_h) u_h, u_h \big\rangle \ge \gamma \|\mathsf{B}^\top \nabla u_h\|^2_{\mathrm{L}^2(\Omega)} \qquad \text{for all } u_h \in \mathscr{U}_h,$$

where $\gamma = \gamma(m, M, \Delta, \kappa(\mathbb{V}))$ is defined by (6.32)*, and by* (6.35)–(6.38)*.*

Theorem 6.2 is our main stability result for nonlinear deformations with defects. Its proof is contained in §6.2–§6.6. In §6.2 we estimate $\delta^2\mathscr{E}_{\rm ac}$ below by an operator $\mathcal{H}$ that localizes all finite differences of test functions. In §6.3 we use a perturbation argument to reduce the problem to a homogeneous deformation with a defect. In §6.4 we employ the definition of the stability index to reduce the problem to one without defect, which is then analyzed in §6.5. In the perturbation argument we introduce several free parameters, which are finally optimized in §6.6.

In §6.8 we will investigate the range of parameters for which $\gamma > 0$, and in particular show numerically that non-trivial solutions exist to which Theorem 6.2 applies.

6.2. **Stability proof 1: a general lower bound.** The representation (3.4) of the a/c energy $\mathscr{E}_{\rm ac}$ yields the following expression for the second variation $\delta^2\mathscr{E}_{\rm ac}$:

$$\begin{aligned}\langle \delta^2 \mathscr{E}_{\rm ac}(y_h) u_h, u_h \rangle = & \sum_{b \in \mathcal{B}_{\rm a}} D_b u_h^\top \phi''(D_b y_h) D_b u_h \\ & + \sum_{b \in \mathcal{B}_{\rm c}} \fint_b \nabla_b u_h^\top \phi''(\nabla_b y_h) \nabla_b u_h \, {\rm d}b,\end{aligned} \tag{6.3}$$

for all $u_h \in \mathscr{U}_h$, where we recall that $\phi''(r)$ is understood as the Hessian matrix of ϕ. A straightforward calculation shows that ϕ'' can be written, in terms of φ' and φ'', as

$$\phi''(r) = \varphi''(|r|) \tfrac{r}{|r|} \otimes \tfrac{r}{|r|} + \tfrac{\varphi'(|r|)}{|r|} \big(\mathbb{1} - \tfrac{r}{|r|} \otimes \tfrac{r}{|r|} \big). \tag{6.4}$$

We will use the fact that $\frac{r}{|r|} \otimes \frac{r}{|r|}$ is the orthogonal projection onto the space $\mathrm{span}\{r\}$ and that $(\mathbb{1} - \frac{r}{|r|} \otimes \frac{r}{|r|})$ is the orthogonal projection onto $\mathrm{span}\{r\}^\perp$. Recalling the notation $a \times b = (\mathsf{Q}_4 a) \cdot b$, where Q_4 denotes a rotation through angle $\pi/2$, we have

$$h^\top \big(\tfrac{r}{|r|} \otimes \tfrac{r}{|r|}\big) h = \big|h \cdot \tfrac{r}{|r|}\big|^2, \quad \text{and} \quad h^\top \big(\mathbb{1} - \tfrac{r}{|r|} \otimes \tfrac{r}{|r|}\big) h = \big|h \times \tfrac{r}{|r|}\big|^2.$$

Hence, we can rewrite (6.3) as

$$\begin{aligned} &\langle \delta^2 \mathscr{E}_{\rm ac}(y_h) u_h, u_h \rangle \\ &= \sum_{b \in \mathcal{B}_{\rm a}} \Big\{ \tfrac{\varphi''(|D_b y_h|)}{|D_b y_h|^2} |D_b y_h \cdot D_b u_h|^2 + \tfrac{\varphi'(|D_b y_h|)}{|D_b y_h|^3} |D_b y_h \times D_b u_h|^2 \Big\} \\ &\quad + \sum_{b \in \mathcal{B}_{\rm c}} \fint_b \Big\{ \tfrac{\varphi''(|\nabla_b y_h|)}{|\nabla_b y_h|^2} |\nabla_b y_h \cdot \nabla_b u_h|^2 + \tfrac{\varphi'(|\nabla_b y_h|)}{|\nabla_b y_h|^3} |\nabla_b y_h \times \nabla_b u_h|^2 \Big\} \, \mathrm{db}. \end{aligned} \tag{6.5}$$

Next, we construct a relatively crude lower bound on the Hessian $\delta^2 \mathscr{E}_{\rm ac}$, which will nevertheless be sufficient to obtain stability estimates in a range of interesting deformations. Our goal is to "localise" the finite differences $D_b u_h$ occurring in the Hessian representation (6.5), and to render the scalar coefficients hexagonally symmetric.

Since $y_h \in \mathscr{S}_{\mathsf{B},h}(m, M, \Delta)$, we can estimate the coefficients in (6.5) by

$$\tfrac{\varphi''(|D_b y_h|)}{|D_b y_h|^2} \geq C_{|b|} \quad \text{and} \quad \tfrac{\varphi'(|D_b y_h|)}{|D_b y_h|^3} \geq C^\perp_{|b|}, \quad \text{for } b \in \mathcal{B}_{\rm a},$$

with analogous estimates for $b \in \mathcal{B}_{\rm c}$, where

$$\begin{aligned} C_\rho &:= \begin{cases} \min_{s \in [m,M]} \frac{\varphi''(\rho s)}{(\rho s)^2}, & \rho = 1, \\ 0 \wedge \min_{s \in [m,M]} \frac{\varphi''(\rho s)}{(\rho s)^2}, & \rho > 1, \end{cases} \qquad \text{and} \\ C^\perp_\rho &:= \begin{cases} \min_{s \in [m,M]} \frac{\varphi'(\rho s)}{(\rho s)^3}, & \rho = 1, \\ 0 \wedge \min_{s \in [m,M]} \frac{\varphi'(\rho s)}{(\rho s)^3}, & \rho > 1. \end{cases} \end{aligned} \tag{6.6}$$

We note that these lower bounds are independent of y_h, and moreover, all coefficients for non-nearest neighbour bonds are non-positive.

With this notation, we obtain from (6.5) that

$$\begin{aligned} \langle \delta^2 \mathscr{E}_{\rm ac}(y_h) u_h, u_h \rangle \geq &\sum_{b \in \mathcal{B}_{\rm a}} \Big\{ C_{|b|} |D_b y_h \cdot D_b u_h|^2 + C^\perp_{|b|} |D_b y_h \times D_b u_h|^2 \Big\} \\ &+ \sum_{b \in \mathcal{B}_{\rm c}} \fint_b \Big\{ C_{|b|} |\nabla_b y_h \cdot \nabla_b u_h|^2 + C^\perp_{|b|} |\nabla_b y_h \times \nabla_b u_h|^2 \Big\} \, \mathrm{db}. \end{aligned} \tag{6.7}$$

We now observe that we have constructed the extended mesh $\mathcal{T}_h$ in such a way that in the atomistic region every nearest-neighbour bond $b \in \mathcal{B}_{\rm nn}$ lies on the edge of a triangle. As a result we have the identity

$$D_b u_h = \nabla_b u_h(x) \qquad \text{for all } x \in \mathrm{int}(b), \text{ for all } b \in \mathcal{B}_{\rm a} \cap \mathcal{B}_{\rm nn}, \tag{6.8}$$

which we will use heavily throughout. In particular, this implies that

$$\begin{aligned} &\sum_{b \in \mathcal{B}_{\rm nn} \cap \mathcal{B}_{\rm a}} \Big\{ C_1 |D_b y_h \cdot D_b u_h|^2 + C_1^\perp |D_b y_h \times D_b u_h|^2 \Big\} \\ &\qquad = \sum_{b \in \mathcal{B}_{\rm nn} \cap \mathcal{B}_{\rm a}} \fint_b \Big\{ C_1 |D_b y_h \cdot \nabla_b u_h|^2 + C_1^\perp |D_b y_h \times \nabla_b u_h|^2 \Big\} \, \mathrm{db}. \end{aligned} \tag{6.9}$$

Our second observation is that, since $C_{|b|}, C^{\perp}_{|b|} \leq 0$ for $b \in \mathcal{B}_{\rm a} \setminus \mathcal{B}_{\rm nn}$ we can use (3.2) and Jensen's inequality to estimate

$$\begin{aligned}
&\sum_{b\in\mathcal{B}_{\rm a}\setminus\mathcal{B}_{\rm nn}} \Big\{C_{|b|}|D_b y_h \cdot D_b u_h|^2 + C^{\perp}_{|b|}|D_b y_h \times D_b u_h|^2\Big\} \\
&= \sum_{b\in\mathcal{B}_{\rm a}\setminus\mathcal{B}_{\rm nn}} \Big\{C_{|b|}\big|D_b y_h \cdot \fint_b \nabla_b u_h \,{\rm d}b\big|^2 + C^{\perp}_{|b|}\big|D_b y_h \times \fint_b \nabla_b u_h \,{\rm d}b\big|^2\Big\} \\
&\geq \sum_{b\in\mathcal{B}_{\rm a}\setminus\mathcal{B}_{\rm nn}} \fint_b \Big\{C_{|b|}\big|D_b y_h \cdot \nabla_b u_h\big|^2 + C^{\perp}_{|b|}\big|D_b y_h \times \nabla_b u_h\big|^2\Big\} \,{\rm d}b. \qquad (6.10)
\end{aligned}$$

Inserting (6.9) and (6.10) into (6.7) we obtain the following estimate:

$$\begin{aligned}
\big\langle \delta^2 \mathscr{E}_{\rm ac}(y_h) u_h, u_h\big\rangle &\geq \langle \mathcal{H}(y_h) u_h, u_h\rangle \\
&:= \sum_{b\in\mathcal{B}_{\rm c}} \fint_b \Big\{C_{|b|}|\nabla_b y_h \cdot \nabla_b u_h|^2 + C^{\perp}_{|b|}|\nabla_b y_h \times \nabla_b u_h|^2\Big\} \,{\rm d}b \qquad (6.11)\\
&\quad + \sum_{b\in\mathcal{B}_{\rm a}} \fint_b \Big\{C_{|b|}|D_b y_h \cdot \nabla_b u_h|^2 + C^{\perp}_{|b|}|D_b y_h \times \nabla_b u_h|^2\Big\} \,{\rm d}b,
\end{aligned}$$

where $C_{|b|}, C^{\perp}_{|b|}$ are defined in (6.6).

6.3. **Stability proof 2: the perturbation argument.** In the next step, we will estimate the effect of replacing $D_r y_h$ and $\nabla_r y_h$ with $\mathsf{B}r$. To that end, the following lemma will be helpful.

Lemma 6.3. *Suppose that $y_h \in \mathscr{S}_{\mathsf{B},h}(m, M, \Delta)$; then, for all $g \in \mathbb{R}^2, x \in \Omega, r \in \mathbb{R}^2$, and for all possible choices of $\alpha > 0$,*

$$\big|\,|\nabla_r y_h(x)\cdot g|^2 - |\mathsf{B}r\cdot g|^2\big| \leq \alpha\big|\mathsf{B}r\cdot g\big|^2 + \big(1+\tfrac{1}{\alpha}\big)\Delta^2|r|^2|\mathsf{B}^\top g|^2. \qquad (6.12)$$

Similarly, for all $g \in \mathbb{R}^2, x \in \mathcal{L}, r \in \mathbb{L}_*$, and $\alpha > 0$, we have

$$\big|\,|D_r y_h(x)\cdot g|^2 - |\mathsf{B}r\cdot g|^2\big| \leq \alpha\big|\mathsf{B}r\cdot g\big|^2 + \big(1+\tfrac{1}{\alpha}\big)\Delta^2|r|^2|\mathsf{B}^\top g|^2. \qquad (6.13)$$

The same inequalities hold if "$\cdot$" is replaced with "$\times$".

Proof. We verify the bound (6.12) by a straightforward algebraic manipulation (suppressing the argument x), using $\|\mathsf{B}^{-1}\nabla y_h - \mathbb{1}\| \leq \Delta$:

$$\begin{aligned}
\big|\,|\nabla_r y_h\cdot g|^2 - |\mathsf{B}r\cdot g|^2\big| &= \big|\,(|\nabla_r y_h\cdot g| + |\mathsf{B}r\cdot g|)\,(|\nabla_r y_h\cdot g| - |\mathsf{B}r\cdot g|)\big| \\
&\leq (|(\nabla_r y_h - \mathsf{B})r\cdot g| + 2|\mathsf{B}r\cdot g|)\,|(\nabla_r y_h - \mathsf{B})r\cdot g| \\
&\leq 2|\mathsf{B}r\cdot g|\Delta|r||\mathsf{B}^\top g| + \Delta^2|r|^2|\mathsf{B}^\top g|^2.
\end{aligned}$$

Applying a weighted Cauchy inequality $2ab \leq \alpha a^2 + \alpha^{-1}b^2$ we obtain (6.12). The proofs of (6.13), and of the inequalities where "$\cdot$" is replaced with "$\times$" are analogous. □

Applying Lemma 6.3 to the operator $\mathcal{H}(y_h)$, defined in (6.11), we obtain, for constants $\alpha_{|b|}, \alpha^\perp_{|b|} > 0$,

$$\begin{aligned}(6.14)\qquad \langle \mathcal{H}(y_h)u_h, u_h\rangle &\geq \langle \mathcal{H}(y_{\mathsf{B}})u_h, u_h\rangle \\ &\quad - \sum_{b\in\mathcal{B}} \fint_b \Big\{ \alpha_{|b|}|C_{|b|}|\big|\mathsf{B}r_b\cdot\nabla_b u_h\big|^2 + \alpha^\perp_{|b|}|C^\perp_{|b|}|\big|\mathsf{B}r_b\times\nabla_b u_h\big|^2\Big\}\,\mathrm{db} \\ &\quad - \Delta^2\sum_{b\in\mathbb{B}}|b|^2\Big\{\big(1+\tfrac{1}{\alpha_{|b|}}\big)|C_{|b|}| + \big(1+\tfrac{1}{\alpha^\perp_{|b|}}\big)|C^\perp_{|b|}|\Big\}\fint_b \big|\mathsf{B}^\top\nabla_b u_h\big|^2\,\mathrm{db},\end{aligned}$$

for all $y_h \in \mathscr{S}_{\mathsf{B},h}(m,M,\Delta)$ and $u_h \in \mathscr{U}_h$. Note that, in the third term, we have estimated the sum over $\mathcal{B}$ below by the sum over $\mathbb{B}$. We also remark that, for the time being, we retain maximal flexibility in our choice of the constants $\alpha_{|b|}$ and $\alpha^\perp_{|b|}$. We will (partially) optimize over all possible choices in the last step of our proof.

From here on, to simplify the notation, we define the transformed displacement

$$v_h := \mathsf{B}^\top u_h.$$

This means that we can replace $(\mathsf{B}r_b\cdot\nabla_b u_h)$ by $(r_b\cdot\nabla_b v_h)$, and so forth.

Since the algebraic structure of the first and second term in (6.14) is identical it is natural to combine them. Hence, we define

$$\begin{aligned}(6.15)\qquad \langle\tilde{\mathcal{H}}u_h, u_h\rangle &:= \sum_{b\in\mathcal{B}}\fint_b\Big\{\tilde{C}_{|b|}\big|r_b\cdot\nabla_b v_h\big|^2 + \tilde{C}^\perp_{|b|}\big|r_b\times\nabla_b v_h\big|^2\Big\}\,\mathrm{db},\\ \langle\tilde{\mathscr{L}}u_h, u_h\rangle &:= \sum_{b\in\mathbb{B}}|b|^2\Big\{\big(1+\tfrac{1}{\alpha_{|b|}}\big)|C_{|b|}| + \big(1+\tfrac{1}{\alpha^\perp_{|b|}}\big)|C^\perp_{|b|}|\Big\}\fint_b\big|\nabla_b v_h\big|^2\,\mathrm{db},\end{aligned}$$

where $\tilde{C}^{(\perp)}_\rho := C^{(\perp)}_\rho - \alpha^{(\perp)}_\rho|C^{(\perp)}_\rho|$. (Here and throughout the superscript $(\perp)$, e.g., in $C^{(\perp)}_\rho$, is used to refer simultaneously to C_ρ or $C^\perp_\rho$.)

Employing the periodic bond-density lemma, the decomposition of the triangular lattice described in Lemma 2.1, and the definition of the constants $c^{(\perp)}_n := \ell_n^4 C^{(\perp)}_{\ell_n}$, the operator $\tilde{\mathscr{L}}$ can be rewritten as

$$\begin{aligned}(6.16)\qquad \langle\tilde{\mathscr{L}}u_h, u_h\rangle &= (\tilde{L}+\tilde{L}^\perp)\|\nabla v_h\|^2_{\mathrm{L}^2(\Omega)}, \quad \text{where}\\ \tilde{L}^{(\perp)} &= \tfrac{3}{\det\mathsf{A}_6}\textstyle\sum_{n=1}^\infty\big(1+1/\alpha^{(\perp)}_{\ell_n}\big)|c^{(\perp)}_n|.\end{aligned}$$

In summary, we have obtained that, if $y_h \in \mathscr{S}_{\mathsf{B},h}(m,M,\Delta)$, then

$$(6.17)\qquad \langle\delta^2\mathscr{E}_{\mathrm{ac}}(y_h)u_h, u_h\rangle \geq \langle\tilde{\mathcal{H}}u_h, u_h\rangle - \Delta^2(\tilde{L}+\tilde{L}^\perp)\|\nabla v_h\|^2_{\mathrm{L}^2},$$

for all $u_h \in \mathscr{U}_h$, where $\tilde{\mathcal{H}}$ is defined in (6.21), and and $\tilde{L}^{(\perp)}$ in (6.16).

6.4. **Stability proof 3: extension to $\mathbb{B}$.** In the next step, we apply the extension operator (see §4.1) and the definition of the stability index $\kappa := \kappa(\mathbb{V})$ (see §6.1).

Distinguishing whether $\tilde{C}_1$ is positive or negative, and using the definition of κ in the first case, we obtain

$$\sum_{b\in\mathcal{B}_{\mathrm{nn}}}\tilde{C}_1\big|r_b\cdot D_b v_h\big|^2 \geq \begin{cases}\kappa\sum_{b\in\mathbb{B}_{\mathrm{nn}}}\tilde{C}_1\big|r_b\cdot D_b v_h\big|^2, & \tilde{C}_1\geq 0,\\ \sum_{b\in\mathbb{B}_{\mathrm{nn}}}\tilde{C}_1\big|r_b\cdot D_b v_h\big|^2, & \tilde{C}_1<0,\end{cases}$$

which can be rewritten as

$$(6.18)\qquad \sum_{b\in\mathcal{B}_{\mathrm{nn}}}\tilde{C}_1\fint_b\big|r_b\cdot\nabla_b v_h\big|^2\,\mathrm{db} \geq \min(\tilde{C}_1,\kappa\tilde{C}_1)\sum_{b\in\mathbb{B}_{\mathrm{nn}}}\fint_b\big|r_b\cdot\nabla_b v_h\big|^2\,\mathrm{db}.$$

For the "perpendicular" nearest-neighbour terms the same argument (we now need to use (6.2) with $u = \mathsf{Q}_4^\top \mathsf{B}^\top u_h = \mathsf{Q}_4^\top v_h$) yields

$$\sum_{b\in\mathcal{B}_{\rm nn}} \tilde{C}_1^\perp \fint_b |r_b \times \nabla_b v_h|^2 \,{\rm db} \geq \min(\tilde{C}_1^\perp, \kappa\tilde{C}_1^\perp) \sum_{b\in\mathbb{B}_{\rm nn}} \fint_b |r_b \times \nabla_b v_h|^2 \,{\rm db}. \tag{6.19}$$

Since the non-nearest-neighbour terms in $\tilde{\mathcal{H}}$ are non-positive, we have

$$\sum_{b\in\mathcal{B}\setminus\mathcal{B}_{\rm nn}} \fint_b \tilde{C}_{|b|} |r_b \cdot \nabla_b v_h|^2 \,{\rm db} \geq \sum_{b\in\mathbb{B}\setminus\mathbb{B}_{\rm nn}} \fint_b \tilde{C}_{|b|} |r_b \cdot \nabla_b v_h|^2 \,{\rm db}, \quad \text{and}$$
$$\sum_{b\in\mathcal{B}\setminus\mathcal{B}_{\rm nn}} \fint_b \tilde{C}^\perp_{|b|} |r_b \times \nabla_b v_h|^2 \,{\rm db} \geq \sum_{b\in\mathbb{B}\setminus\mathbb{B}_{\rm nn}} \fint_b \tilde{C}^\perp_{|b|} |r_b \times \nabla_b v_h|^2 \,{\rm db}.$$

Hence, defining the constants

$$\overline{C}_\rho^{(\perp)} := \begin{cases} \min(\tilde{C}_\rho^{(\perp)}, \kappa\tilde{C}_\rho^{(\perp)}), & \rho = 1, \\ \tilde{C}_\rho^{(\perp)}, & \rho > 1, \end{cases} \tag{6.20}$$

we arrive at (recall that $v_h = \mathsf{B}^\top u_h$)

$$\begin{aligned} \big\langle \tilde{\mathcal{H}} u_h, u_h \big\rangle &\geq \langle \overline{\mathcal{H}} u_h, u_h \rangle \\ &:= \sum_{b\in\mathbb{B}} \fint_b \Big\{ \overline{C}_{|b|} |r_b \cdot \nabla_b v_h|^2 + \overline{C}^\perp_{|b|} |r_b \times \nabla_b v_h|^2 \Big\} \,{\rm db} \quad \forall u_h \in \mathscr{U}_h. \end{aligned} \tag{6.21}$$

We note that $\overline{\mathcal{H}}$ depends only on m, M, Δ, κ.

6.5. **Stability proof 4: homogeneous lattice.** Combining (6.21) and (6.17), we have that, for all $y_h \in \mathscr{S}_{\mathsf{B},h}(m, M, \Delta)$, $u_h \in \mathscr{U}_h$,

$$\big\langle \delta^2 \mathscr{E}_{\rm ac}(y_h) u_h, u_h \big\rangle \geq \big\langle \overline{\mathcal{H}} u_h, u_h \big\rangle - \Delta^2 (\tilde{L} + \tilde{L}^\perp) \|\mathsf{B}^\top \nabla u_h\|^2_{{\rm L}^2(\Omega)}. \tag{6.22}$$

Lemma 6.4. *The operator $\overline{\mathcal{H}}$, defined in* (6.21)*, satisfies*

$$\langle \overline{\mathcal{H}} u_h, u_h \rangle \geq \overline{\gamma} \|\mathsf{B}^\top \nabla u_h\|^2_{{\rm L}^2(\Omega)} \qquad \forall u_h \in \mathscr{U}_h, \tag{6.23}$$

where $\overline{\gamma} := \min(\frac{3}{4}\bar{c} + \frac{9}{4}\bar{c}^\perp, \frac{9}{4}\bar{c} + \frac{3}{4}\bar{c}^\perp)$ and $\bar{c}^{(\perp)} := \frac{1}{\det \mathsf{A}_6} \sum_{n=1}^\infty \ell_n^4 \overline{C}^{(\perp)}_{\ell_n}$.

The proof of Lemma 6.4 is given at the end of the present subsection.

Remark 6.2. The estimate (6.23) is sharp in the sense that, if $\mathcal{T}_h = \mathcal{T}_{\rm a}$, then

$$\lim_{N\to\infty} \inf_{u\in\mathscr{U}} \frac{\langle \overline{\mathcal{H}}_{\mathsf{B}} u, u \rangle}{\|\mathsf{B}^\top \nabla \bar{u}\|^2} = \overline{\gamma}. \tag{6.24}$$

This statement follows immediately from the proof of Lemma 6.4. □

Application of the bond-density lemma yields

$$\begin{aligned} \langle \overline{\mathcal{H}} u_h, u_h \rangle &= \sum_{T\in\mathcal{T}_h} |T| \Big\{ \sum_{r\in\mathbb{L}_*} \frac{\overline{C}_{|r|}}{\det \mathsf{A}_6} \big| r \cdot \nabla_r v_h|_T \big|^2 + \sum_{r\in\mathbb{L}_*} \frac{\overline{C}^\perp_{|r|}}{\det \mathsf{A}_6} \big| r \times \nabla_r v_h|_T \big|^2 \Big\} \\ &=: \sum_{T\in\mathcal{T}_h} |T| \big\{ H_T[v_h] + H_T^\perp[v_h] \big\}. \end{aligned} \tag{6.25}$$

Let $\mathsf{G} := \nabla v_h = \mathsf{B}^\top \nabla u_h$, and $\mathsf{G}_T := \nabla v_h|_T$, then we can rewrite $H_T[v_h]$, using Lemma 2.1, in the form

$$H_T[u_h] = \sum_{r\in\mathbb{L}_*} \tfrac{\overline{C}_{|r|}}{\det \mathsf{A}_6} \big[r^\top \mathsf{G}_T r\big]^2 = \sum_{n=1}^{\infty} \tfrac{\overline{C}_{\ell_n}}{\det \mathsf{A}_6} \sum_{j=1}^{6} \big[(\mathsf{Q}_4^j r_n)^\top \mathsf{G}_T (\mathsf{Q}_4^j r_n)\big]^2. \tag{6.26}$$

Exploiting the hexagonal symmetry of the inner sum, using (2.2), and recalling the definition of $\bar{c}$ from Lemma 6.4, we obtain

$$H_T[v_h] = \tfrac{1}{\det \mathsf{A}_6} \big\{\textstyle\sum_{n=1}^{\infty} \ell_n^4 \overline{C}_{\ell_n}\big\} |\mathsf{G}_T|_{\rm el}^2 = \bar{c} |\mathsf{G}_T|_{\rm el}^2, \tag{6.27}$$

where $|\mathsf{G}|_{\rm el} := \frac{3}{2}|\mathsf{G}^{\rm sym}|^2 + \frac{3}{4}|{\rm tr}\mathsf{G}|^2$ (*cf.* (2.2)). Replacing r with $\mathsf{Q}_4 r$ in the above computations yields

$$H_T^\perp[v_h] = \tfrac{1}{\det \mathsf{A}_6} \big\{\textstyle\sum_{n=1}^{\infty} \ell_n^4 \overline{C}_{\ell_n}^\perp\big\} \big|\mathsf{Q}_4 \mathsf{G}_T\big|_{\rm el}^2 = \bar{c}^\perp \big|\mathsf{Q}_4 \mathsf{G}_T\big|_{\rm el}^2. \tag{6.28}$$

Lemma 6.5. *Let $|\cdot|_{\rm el}$ be defined as in* (2.2)*, and let $\mathsf{G} \in \mathbb{R}^{2\times 2}$, then*

$$\begin{aligned} |\mathsf{G}|_{\rm el}^2 &= \tfrac{3}{4}|\mathsf{G}|^2 + \tfrac{3}{2}(\mathsf{G}_{11}+\mathsf{G}_{22})^2 - \tfrac{3}{2}\det \mathsf{G}, \quad \textit{and} \\ \big|\mathsf{Q}_4\mathsf{G}\big|_{\rm el}^2 &= \tfrac{3}{4}|\mathsf{G}|^2 + \tfrac{3}{2}(\mathsf{G}_{12}-\mathsf{G}_{21})^2 - \tfrac{3}{2}\det \mathsf{G}, \end{aligned}$$

and in particular,

$$\begin{aligned} \bar{c}|\mathsf{G}|_{\rm el}^2 + \bar{c}^\perp |\mathsf{Q}_4\mathsf{G}|_{\rm el}^2 = {} & \tfrac{3}{4}(\bar{c}+\bar{c}^\perp)|\mathsf{G}|^2 + \tfrac{3}{2}\bar{c}|\mathsf{G}_{11}+\mathsf{G}_{22}|^2 \\ & + \tfrac{3}{2}\bar{c}^\perp|\mathsf{G}_{12}-\mathsf{G}_{21}|^2 - \tfrac{3}{2}(\bar{c}+\bar{c}^\perp)\det \mathsf{G}. \end{aligned} \tag{6.29}$$

Proof. The first identity can be verified by a straightforward algebraic manipulation. The second identity is an immediate consequence of the first. The third identity follows by combining the first two. □

Proof of Lemma 6.4. We define the fourth-order tensor $\mathbb{C}$, using summation convention, by $\mathbb{C}_{i\alpha}^{j\beta}\mathsf{G}_{i\alpha}\mathsf{G}_{j\beta} := \bar{c}|\mathsf{G}|_{\rm el}^2 + \bar{c}^\perp\big|\mathsf{Q}_4\mathsf{G}\big|_{\rm el}^2$.

The Legendre–Hadamard condition (see, e.g., [10]) states that

$$\inf_{\substack{v\in \mathrm{H}^1_{\#}(\Omega)^2 \\ \|\nabla v\|_{\mathrm{L}^2}=1}} \int_\Omega \mathbb{C}_{i\alpha}^{j\beta} (\nabla v)_{i\alpha} (\nabla v)_{j\beta} \,\mathrm{dV} = \min_{\substack{w,k\in\mathbb{R}^2 \\ |w|=|k|=1}} \mathbb{C}_{i\alpha}^{j\beta} w_i w_j k_\alpha k_\beta =: \overline{\gamma}.$$

Thus, we have reduced the task to testing $\mathbb{C}$ with rank-1 matrices $w\otimes k$. Using the definition of $\mathbb{C}$, identity (6.29), and noting that $\det(w\otimes k) = 0$, we obtain

$$\mathbb{C}_{i\alpha}^{j\beta} w_i w_j k_\alpha k_\beta = \tfrac{3}{4}(\bar{c}+\bar{c}^\perp)|w|^2|k|^2 + \tfrac{3}{2}\bar{c}(w\cdot k)^2 + \tfrac{3}{2}\bar{c}^\perp (w\times k)^2. \tag{6.30}$$

If $\bar{c} \geq \bar{c}^\perp$ then (6.30) is minimised for $w \perp k$, and

$$\overline{\gamma} = \tfrac{3}{4}(\bar{c}+\bar{c}^\perp) + \tfrac{3}{2}\bar{c}^\perp = \tfrac{3}{4}\bar{c} + \tfrac{9}{4}\bar{c}^\perp.$$

If $\bar{c} \leq \bar{c}^\perp$ then (6.30) is minimised for $w = k$, and

$$\overline{\gamma} = \tfrac{3}{4}(\bar{c}+\bar{c}^\perp) + \tfrac{3}{2}\bar{c} = \tfrac{9}{4}\bar{c} + \tfrac{3}{4}\bar{c}^\perp.$$

Combining the two cases gives the stated result. □

6.6. **Stability proof 5: optimizing the parameters.** Combining Lemma 6.4 with (6.22), we obtain the stability estimate

$$\langle \delta^2 \mathscr{E}_{\rm ac}(y_h) u_h, u_h \rangle \geq \gamma \|\mathsf{B}^\top \nabla u_h\|_{\mathrm{L}^2}^2, \text{ where } \gamma = \overline{\gamma} - \Delta^2(\tilde{L} + \tilde{L}^\perp), \tag{6.31}$$

for all $y_h \in \mathscr{S}_{\mathsf{B},h}(m, M, \Delta)$ and $u_h \in \mathscr{U}_h$. The constant γ still depends on the free parameters $\alpha_{\ell_n}^{(\perp)}$. Ideally, we would like to optimize γ over all possible choices, however, the double-minimization problem in the definition of γ makes this impractical. We will choose the parameters so that they are optimal in the case, which is the most important in our numerical computations. For the following discussion, recall the definition of $c_n^{(\perp)}$ from (6.1) and let $\alpha_n^{(\perp)} := \alpha_{\ell_n}^{(\perp)}$.

We begin by noting that γ can be rewritten in the form

$$\gamma = \min\big(\gamma_1 + \gamma_1^\perp, \gamma_2 + \gamma_2^\perp\big), \tag{6.32}$$

where

$$\gamma_1 := \tfrac{3}{4}\bar{c} - \Delta^2 \tilde{L}, \qquad \gamma_1^\perp := \tfrac{9}{4}\bar{c}^\perp - \Delta^2 \tilde{L}^\perp,$$
$$\gamma_2 := \tfrac{9}{4}\bar{c} - \Delta^2 \tilde{L}, \qquad \gamma_2^\perp := \tfrac{3}{4}\bar{c}^\perp - \Delta^2 \tilde{L}^\perp.$$

Near global minima of $\mathscr{E}_{\rm ac}$ we expect that $c_1 > 0$ and $c_1^\perp \approx 0$, which suggests to optimise the parameters $\alpha_n^{(\perp)}$ for the case $\gamma = \gamma_1 + \gamma_1^\perp$.

Recalling from (6.16) the definition of $\tilde{L}$, and recalling that $c_n \leq 0$ for $n \geq 2$, we can rewrite γ_1 in the form

$$\begin{aligned}
\gamma_1 &= \tfrac{1}{\det \mathsf{A}_6}\Big(\min\Big\{ \tfrac{3}{4}(c_1 - \alpha_1 |c_1|), \tfrac{3}{4}\kappa(c_1 - \alpha_1|c_1|)\Big\} - 3\big(1 + \tfrac{1}{\alpha_1}\big)\Delta^2 |c_1|\Big) \\
&\quad + \tfrac{1}{\det \mathsf{A}_6} \sum_{n=2}^\infty \big(\tfrac{3}{4} + \tfrac{3}{4}\alpha_n + 3\big(1 + \tfrac{1}{\alpha_n}\big)\Delta^2\big) c_n \\
&=: \tfrac{1}{\det \mathsf{A}_6}\psi_1(\alpha_1) + \tfrac{1}{\det \mathsf{A}_6} \sum_{n=2}^\infty \big(\tfrac{3}{4} + \tfrac{3}{4}\alpha_n + 3\big(1 + \tfrac{1}{\alpha_n}\big)\Delta^2\big) c_n.
\end{aligned} \tag{6.33}$$

We see immediately that $\alpha_n = 2\Delta$ is optimal for $n \geq 2$. For $n = 1$, the situation is more complicated and we treat it separately in the following lemma. We omit the straightforward proof and refer to [25, Lemma 6.7] for the details of the argument.

Lemma 6.6. *Suppose that $\Delta \leq \sqrt{\kappa}/2$; then*

$$\max_{\alpha_1 > 0} \psi_1(\alpha_1) = \min\big\{(\tfrac{3}{4}\kappa - 3\sqrt{\kappa}\Delta - 3\Delta^2)c_1, (\tfrac{3}{4} + 3\Delta + 3\Delta^2)c_1\big\}, \tag{6.34}$$

which is attained for $\alpha_1 = 2\Delta/\sqrt{\kappa}$ if $c_1 > 0$ and for $\alpha_1 = 2\Delta$ if $c_1 \leq 0$.

If we insert $\alpha_n = 2\Delta$ for $n \geq 2$, and the value for α_1 for which (6.34) is attained, into (6.33), then we obtain

$$\begin{aligned}
\gamma_1 = \tfrac{1}{\det \mathsf{A}_6} &\min\big\{(\tfrac{3}{4}\kappa - 3\sqrt{\kappa}\Delta - 3\Delta^2)c_1, (\tfrac{3}{4} + 3\Delta + 3\Delta^2)c_1\big\} \\
&+ \tfrac{1}{\det \mathsf{A}_6}\textstyle\sum_{n=2}^\infty \big(\tfrac{3}{4} + 3\Delta + 3\Delta^2\big) c_n.
\end{aligned} \tag{6.35}$$

Using analogous arguments, we choose $\alpha_n^\perp = 2\Delta/\sqrt{3}$ for $n \geq 2$ and for $n = 1$ if $c_1^\perp \leq 0$; and $\alpha_1^\perp = 2\Delta/\sqrt{3\kappa}$ if $c_1^\perp > 0$ (note that under the assumption $\Delta \leq \sqrt{\kappa}/2$ we also get $\alpha_1^\perp \leq 1$). Inserting these values into $\gamma_1^\perp$, we obtain

$$\begin{aligned}
\gamma_1^\perp = \tfrac{1}{\det \mathsf{A}_6} &\min\big\{(\tfrac{9}{4}\kappa - 3\sqrt{3\kappa}\Delta - 3\Delta^2)c_1^\perp, (\tfrac{9}{4} + 3\sqrt{3}\Delta + 3\Delta^2)c_1^\perp\big\} \\
&+ \tfrac{1}{\det \mathsf{A}_6}\textstyle\sum_{n=2}^\infty \big(\tfrac{9}{4} + 3\sqrt{3}\Delta + 3\Delta^2\big) c_n^\perp.
\end{aligned} \tag{6.36}$$

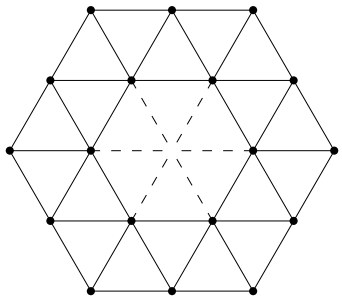


FIGURE 4. Neighbourhood of a void to illustrate the proof of Theorem 6.7. The bonds $\mathcal{B}_1$ are dashed, the bonds $\mathcal{B}_2$ are solid.

A tedious but straightforward computation, for which we skip the details, shows that, if $\gamma_2, \gamma_2^\perp$ are defined by (6.32), then the above choices for $\alpha_n^{(\perp)}$ yield

$$\begin{aligned}\gamma_2 &:= \tfrac{1}{\det \mathsf{A}_6} \min\big\{(\tfrac{9}{4}\kappa - 6\sqrt{\kappa}\Delta - 3\Delta^2)c_1, (\tfrac{9}{4} + 6\Delta + 3\Delta^2)c_1\big\} \\ &\quad + \tfrac{1}{\det \mathsf{A}_6}\textstyle\sum_{n=2}^{\infty}\big(\tfrac{9}{4} + 6\Delta + 3\Delta^2\big)c_n, \text{ and} \qquad (6.37)\\ \gamma_2^\perp &:= \tfrac{1}{\det \mathsf{A}_6} \min\big\{(\tfrac{3}{4}\kappa - 2\sqrt{3\kappa}\Delta - 3\Delta^2)c_1^\perp, (\tfrac{3}{4} + 2\sqrt{3}\Delta + 3\Delta^2)c_1^\perp\big\} \\ &\quad + \tfrac{1}{\det \mathsf{A}_6}\textstyle\sum_{n=2}^{\infty}\big(\tfrac{3}{4} + 2\sqrt{3}\Delta + 3\Delta^2\big)c_n^\perp. \qquad (6.38)\end{aligned}$$

Conclusion of the proof of Theorem 6.2. Combining (6.31) and (6.32) with the choice of $\alpha_n^{(\perp)}$ made above, we obtain the statement of Theorem 6.2. □

6.7. **Stability index of separated vacancies.** In Table 1 we have provided numerical (i.e., non-rigorous) lower bounds for vacancy stability indices. In this section, we prove that $\kappa(\mathbb{V}) \geq 2/7$ if $\mathbb{V}$ consists only of single vacancy sites, which are separated by a short distance.

Theorem 6.7. *Suppose that* $\mathbb{V}$ *satisfies the* separation condition

$$x_1 \in \mathbb{V}, x_2 \in \mathbb{V}^\# \setminus \{x_1\} \quad \Rightarrow \quad |x_1 - x_2| \geq 4, \qquad (6.39)$$

then $\kappa(\mathbb{V}) \geq \frac{2}{7}$.

Proof. We define the alternative extension operator (*cf.* Figure 4)

$$(\tilde{E}w)(x) := \frac{1}{6}\sum_{r \in \mathbb{L}_{\rm nn}} w(x+r) \qquad \forall x \in \mathbb{V}^\#. \qquad (6.40)$$

Using the notation introduced in Figure 4 we aim to prove that

$$\sum_{b \in \mathcal{B}_2} |r_b \cdot D_b u|^2 \geq \frac{2}{7}\sum_{b \in \mathcal{B}_1 \cup \mathcal{B}_2} \big|r_b \cdot D_b \tilde{E}u\big|^2 \qquad \forall u \in \mathscr{U}, \qquad (6.41)$$

Before we prove (6.41), let us discuss why this establishes the result. Firstly, (6.41) and the separation condition (6.39) imply that

$$\sum_{b \in \mathcal{B}} |r_b \cdot D_b u|^2 \geq \kappa \sum_{b \in \mathbb{B}} \big|r_b \cdot D_b \tilde{E}u\big|^2 \qquad \forall u \in \mathscr{U}. \qquad (6.42)$$

Since the actual extension operator minimizes the right-hand side, we can replace $\tilde{E}$ with E in (6.42), and hence obtain the result.

Proof of (6.41)*:* We begin by noting that 18 vertices of $\mathcal{T}_h$ are involved in (6.41), which correspond to 36 degrees of freedom for a transformed displacement u. We construct a basis of the space of these degrees of freedom $\{w^{(k,j)} : -2 \le k \le 3, 1 \le j \le 6\}$ as follows: Firstly, we require that all basis functions satisfy the symmetry

$$w^{(k,j)}(\mathsf{Q}_6\xi) = e^{\mathrm{i}k\arg(\xi)}\mathsf{Q}_6 w^{(k,j)}(\xi). \tag{6.43}$$

Secondly, we define $c_k := \cos(k\pi/6)$, and $s_k := \sin(k\pi/6)$, and prescribe the nodal values

$$\begin{aligned}
w^{(k,1)}(1,0) &= (-\sqrt{3},0), & w^{(k,4)}(\tfrac{3}{2},\tfrac{\sqrt{3}}{2}) &= -\sqrt{3}\mathrm{i}\, e^{\mathrm{i}k\frac{\pi}{6}}(s_1,c_1),\\
w^{(k,2)}(1,0) &= (0,3\mathrm{i}), & w^{(k,5)}(2,0) &= (\sqrt{3},0),\\
w^{(k,3)}(\tfrac{3}{2},\tfrac{\sqrt{3}}{2}) &= 3\, e^{\mathrm{i}k\frac{\pi}{6}}(c_1,-s_1), & w^{(k,6)}(2,0) &= (0,3\mathrm{i}).
\end{aligned}$$

Finally, for all remaining vertices ξ we define $w^{(k,j)}(\xi) = (0,0)$.

Consider the two quadratic forms

$$a[u] = \sum_{b\in\mathcal{B}_2} |r_b\cdot D_b u|^2, \qquad \text{and} \qquad b[u] = \sum_{b\in\mathcal{B}_1\cup\mathcal{B}_2} |r_b\cdot D_b u|^2.$$

The corresponding "stiffness matrices" with respect to the basis $\{w^{(k,j)}\}$ have a block-diagonal structure: if $u = \sum_{k=-2}^{3}\sum_{j=1}^{6} U_{k,j} w^{(k,j)}$ then

$$a[u] = \sum_{k=-2}^{3}\sum_{j,j'=1}^{6} A^{(k)}_{j,j'} U_{k,j} U_{k,j'} \quad \text{and} \quad b[u] = \sum_{k=-2}^{3}\sum_{j,j'=1}^{6} B^{(k)}_{j,j'} U_{k,j} U_{k,j'}$$

with the blocks

$$A^{(k)} = \begin{pmatrix}
4+c_{2k} & s_{2k} & c_k & s_k & 2 & 0\\
s_{2k} & 2-c_{2k} & -s_k & c_k & 0 & 0\\
c_k & -s_k & 1 & 0 & 0 & 0\\
s_k & c_k & 0 & 5 & 2s_k & 2c_k\\
2 & 0 & 0 & 2s_k & 3 & 0\\
0 & 0 & 0 & 2c_k & 0 & 1
\end{pmatrix}, \quad \text{and}$$

$$B^{(k)} = A^{(k)} + \begin{pmatrix}
2-\frac{1}{2}\big(1-(-1)^k\big)(1+c_{2k}) & \frac{1}{6}\big(1-(-1)^k\big)s_{2k} & 0 & 0 & 0 & 0\\
\frac{1}{6}\big(1-(-1)^k\big)s_{2k} & \frac{1}{18}\big(1-(-1)^k\big)(1+c_{2k}) & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

We need to find a maximal positive κ such that $A^{(k)} \ge \kappa B^{(k)}$, in the sense of Hermitian matrices, for all k. Such a constant exists if $\mathrm{Ker}A^{(k)} \subset \mathrm{Ker}B^{(k)}$ for all k. An explicit constant κ can be obtained if we can find minimal constants $\lambda^{(k)}$ such that, for some vector $v^{(k)} \notin \mathrm{Ker}A^{(k)}$,

$$A^{(k)}v^{(k)} = \lambda^{(k)}(B^{(k)} - A^{(k)})v^{(k)}.$$

In that case we would obtain $\kappa = \lambda/(1+\lambda)$, where $\lambda = \min_k \lambda^{(k)}$. We perform these calculations separately for $k = 0, \pm1, \pm2, 3$.

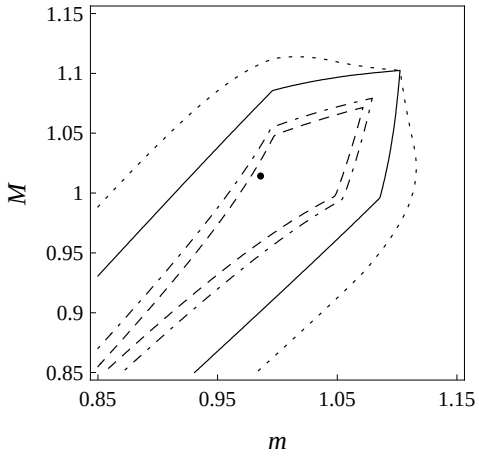


FIGURE 5. Regions of stability in (m, M) parameter space. Dotted line: (m, M)-region such that y_{B} is stable in the full atomistic model for all B with singular values $\{m, M\}$. Solid line: 0-level set of $\gamma(m, M, 0, 1)$. Dot-dashed line: 0-level set of $\gamma(m, M, 0, 2/7)$. Dashed line: zero level set of $\gamma(m, M, 0.02, 2/7)$, where $2/7$ corresponds to a vacancy. The black dot corresponds to the numerical solution described in §8.1.

Case $k = 0$*:* $\mathrm{Ker}(A^{(0)}) = \mathrm{Ker}(B^{(0)}) = \mathrm{span}\{v_0\}$, with $v_0 = (0, 1, 0, -1, 0, 2)$; therefore we add $v_0 \otimes v_0$ to $A^{(0)}$ to make it strictly positive definite and solve

$$0 = \det\left(v_0 \otimes v_0 + A^{(0)} - \lambda(B^{(0)} - A^{(0)})\right) = 72(4 - 3\lambda),$$

to obtain that $\lambda^{(0)} = \frac{4}{3}$.

Case $k = \pm 1$*:* $\mathrm{Ker}(A^{(0)}) = \mathrm{Ker}(B^{(0)}) = \mathrm{span}\{v_0\}$, with $v_0 = (\mp 1, \sqrt{3}, \pm\sqrt{3}, -1, \pm 1, \sqrt{3})$; therefore we add $v_0 \otimes v_0$ to $A^{(\pm 1)}$ and solve

$$0 = \det(v_0 \otimes v_0 + A^{(\pm 1)} - \lambda(B^{(\pm 1)} - A^{(\pm 1)})) = 24(24 - 5\lambda),$$

from where we find $\lambda^{(\pm 1)} = \frac{24}{5}$.

Case $k = \pm 2$*:* In this case $\mathrm{Ker}A^{(\pm 2)} = \mathrm{Ker}B^{(\pm 2)} = \{0\}$; hence we solve

$$0 = \det(A^{(2)} - \lambda(B^{(2)} - A^{(2)})) = 6(2 - 5\lambda),$$

to obtain that $\lambda^{(\pm 2)} = \frac{2}{5}$.

Case $k = 3$*:* In this case $\mathrm{Ker}A^{(3)} = \mathrm{Ker}B^{(3)} = \{0\}$; hence we solve

$$0 = \det(A^{(3)} - \lambda(B^{(3)} - A^{(3)})) = 4(9 - 11\lambda),$$

to obtain that $\lambda^{(3)} = \frac{9}{11}$.

Conclusion: The smallest of the eigenvalues is given by

$$\lambda = \min_{k=-2,\dots,3} \lambda^{(k)} = \tfrac{2}{5},$$

which gives the coercivity constant $\kappa = \frac{\lambda}{1+\lambda} = \frac{2}{7}$. □

6.8. **Sharpness of the stability estimate.** To understand whether Theorem 6.2 is sharp, we consider a homogeneous deformation $y_h = y_{\mathsf{B}}$. If B is a multiple of the identity (exact triangular lattice), then it is not too difficult to see analytically that our estimate cannot be sharp, that is, the actual region of stability of $\delta^2\mathscr{E}_{\rm ac}$ is a strict superset of the region where $\gamma_{\rm hom}$ is positive (see [25, Section 6.6]).

However, the gap for exact triangular lattices is small as the numerical experiment shown in Figure 5 demonstrates. By contrast, if B contains a non-negligible shear component, then our estimates are not very sharp. In Figure 5 we plot the zero level line of γ in (m, M) parameter space for $\kappa \in \{2/7, 1\}$, and for $\Delta \in \{0, 0.02\}$. The case $\kappa = 2/7, \Delta = 0.02$ corresponds to our numerical experiment in §8.1.

7. A Priori Error Estimates

7.1. **Main result.** Having established consistency and stability of the a/c method introduced in §3, we can now prove an *a priori* error estimate. For the statement of the following result recall the definitions of $\mathscr{S}_{\mathsf{B},h}(m, M, \Delta)$ from §6.1, $\Pi_2(y)$ from §4.3, and γ from (6.32).

Theorem 7.1. *Suppose that Assumption A holds. Let $\mathsf{B} \in \mathbb{R}^{2\times 2}_+$, and let $y_{\rm a} \in \mathscr{Y}_{\mathsf{B}}$ be a solution of* (2.8) *and $y_{\rm ac} \in \mathscr{Y}_{\mathsf{B},h}$ a solution of* (3.7)*, such that the following* stability assumption *holds:*

There exist $0 < m \le M$ and $\Delta > 0$ such that $\gamma = \gamma(m, M, \Delta, \kappa(\mathbb{V}))$ is positive and such that

$$(1-t)y_{\rm ac} + tI_h y_{\rm a} \in \mathscr{S}_{\mathsf{B},h}(m, M, \Delta) \qquad \forall t \in [0,1]. \tag{7.1}$$

Then, there exist constants c_1, c_2, which depend only on the shape regularity of $\mathcal{T}_h$, on m, and on $\mu_{\rm a}(y_{\rm a})$, such that

$$\left\|\nabla \bar{y}_{\rm a} - \nabla y_{\rm ac}\right\|_{\mathrm{L}^2(\Omega)} \le \frac{c_1}{\gamma} \inf_{\tilde{y}_{\rm a} \in \Pi_2(y_{\rm a})} \left\|h\nabla^2 \tilde{y}_{\rm a}\right\|_{\mathrm{L}^2(\Omega_{\rm c})}, \quad \text{and} \tag{7.2}$$

$$\left|\mathscr{E}_{\rm a}(y_{\rm a}) - \mathscr{E}_{\rm ac}(y_{\rm ac})\right| \le \frac{c_2}{\gamma^2} \inf_{\tilde{y}_{\rm a} \in \Pi_2(y_{\rm a})} \left\|h\nabla^2 \tilde{y}_{\rm a}\right\|^2_{\mathrm{L}^2(\Omega_{\rm c})}. \tag{7.3}$$

Remark 7.1 (The Stability Assumption). The main assumption in Theorem 7.1 that we have not justified rigorously is the stability condition (7.1). It is a natural assumption since it requires, essentially, that $y_{\rm ac}$ belongs to the same basin of stability as $y_{\rm a}$.

One would prefer to be able to prove (7.1) rigorously, however, short of proving the existence of atomistic and a/c solutions $y_{\rm a}, y_{\rm ac}$ such that

$$\|\nabla \bar{y}_{\rm a} - \nabla y_{\rm ac}\|_{\mathrm{L}^\infty} + \|\nabla \bar{y}_{\rm a} - \nabla I_h y_{\rm a}\|_{\mathrm{L}^\infty} \quad \text{is "sufficiently small"}, \tag{7.4}$$

one cannot hope to remove it, except by postulating even stronger requirements, e.g., phrasing (7.4) itself as an assumption.

A rigorous estimate on $\|\nabla \bar{y}_{\rm a} - \nabla I_h y_{\rm a}\|_{\mathrm{L}^\infty}$ requires a regularity theory for atomistic systems with defects, and we are currently unaware of any results in this direction.

A rigorous estimate on $\|\nabla \bar{y}_{\rm a} - \nabla y_{\rm ac}\|_{\mathrm{L}^\infty}$ could, in principle, be achieved using the inverse function theorem [26, 18, 23], but requires stability of $\delta^2\mathscr{E}_{\rm ac}(I_h y_{\rm a})$ as an operator from (discrete variants of) $\mathrm{W}^{1,\infty}$ to $\mathrm{W}^{-1,\infty}$. For the discretized Laplace operator such results are classical for quasiuniform meshes [28] and have recently

been extended to locally refined meshes [3]. These results give legitimate hope that assumption (7.1) might be (partially) removed with substantial additional work. □

Proof. 1. Error in the H^1*-norm.* Let $e_h = I_h y_\mathrm{a} - y_\mathrm{ac}$, then there exists $\theta_h \in \mathrm{conv}\{I_h y_\mathrm{a}, y_\mathrm{ac}\}$ such that

$$\begin{aligned}\big\langle \delta^2 \mathscr{E}_\mathrm{ac}(\theta_h) e_h, e_h \big\rangle &= \int_0^1 \big\langle \delta^2 \mathscr{E}_\mathrm{ac}(y_\mathrm{ac} + t e_h) e_h, e_h \big\rangle \,\mathrm{d}t \\ &= \big\langle \delta \mathscr{E}_\mathrm{ac}(I_h y_\mathrm{a}) - \delta \mathscr{E}_\mathrm{ac}(y_\mathrm{ac}), e_h \big\rangle.\end{aligned}$$

Using the stability assumption (7.1) to bound $\big\langle \mathscr{E}_\mathrm{ac}(\theta_h) e_h, e_h \big\rangle$ from below, and the fact that $\langle \delta \mathscr{E}_\mathrm{ac}(y_\mathrm{ac}), e_h \rangle = 0$, we obtain

$$m^2 \gamma \|\nabla e_h\|^2_{\mathrm{L}^2(\Omega)} \le \big\langle \delta \mathscr{E}_\mathrm{ac}(I_h y_\mathrm{a}), e_h \big\rangle.$$

We employ the consistency result, Theorem 5.1, to estimate

$$m^2 \gamma \|\nabla e_h\|^2_{\mathrm{L}^2(\Omega)} \le C^\mathrm{cons} \inf_{\tilde{y}_\mathrm{a} \in \Pi_2(y_a)} \big\| h \nabla^2 \tilde{y} \big\|_{\mathrm{L}^2(\Omega_\mathrm{c})} \|\nabla e_h\|_{\mathrm{L}^2}, \tag{7.5}$$

where C^cons depends on $\mu_a(y_\mathrm{a})$ and $\mu_\mathrm{c}(I_h y_\mathrm{a})$.

Employing the interpolation error bounds (4.3) and (4.4) to estimate

$$\begin{aligned}&\|\nabla \bar{y}_\mathrm{a} - \nabla y_\mathrm{ac}\|_{\mathrm{L}^2} \le \|\nabla \bar{y}_\mathrm{a} - \nabla I_h y_\mathrm{a}\|_{\mathrm{L}^2(\Omega)} + \|\nabla e_h\|_{\mathrm{L}^2(\Omega)} \\ &\le \inf_{\tilde{y}_\mathrm{a} \in \Pi_2(y_\mathrm{a})} \Big[\|\nabla \bar{y}_\mathrm{a} - \nabla \tilde{y}_\mathrm{a}\|_{\mathrm{L}^2(\Omega_\mathrm{c})} + \|\nabla \tilde{y}_\mathrm{a} - \nabla I_h y_\mathrm{a}\|_{\mathrm{L}^2(\Omega_\mathrm{c})} \Big] + \|\nabla e_h\|_{\mathrm{L}^2(\Omega)} \\ &\le \inf_{\tilde{y}_\mathrm{a} \in \Pi_2(y_\mathrm{a})} \big\| (\tilde{C}_\mathrm{a} + \tilde{C}_h h) \nabla^2 \tilde{y}_\mathrm{a} \big\|_{\mathrm{L}^2(\Omega_\mathrm{c})} + \|\nabla e_h\|_{\mathrm{L}^2(\Omega)},\end{aligned}$$

applying (7.5), and noting that $h \ge 1$, we obtain (7.2) with constant $c_1 = m^{-2}(C^\mathrm{cons} + \gamma(\tilde{C}_\mathrm{a} + \tilde{C}_h))$, which depends indeed only on the shape regularity of $\mathcal{T}_h$, on $\mu_\mathrm{a}(y_\mathrm{a})$, and on $\mu_c(I_h y_\mathrm{a}) \ge m$.

2. Error in the energy. To estimate the error in the energy we split

$$\begin{aligned}|\mathscr{E}_\mathrm{a}(y_\mathrm{a}) - \mathscr{E}_\mathrm{ac}(y_\mathrm{ac})| &\le |\mathscr{E}_\mathrm{a}(y_\mathrm{a}) - \mathscr{E}_\mathrm{a}(I_h y_\mathrm{a})| + |\mathscr{E}_\mathrm{a}(I_h y_\mathrm{a}) - \mathscr{E}_\mathrm{ac}(I_h y_\mathrm{a})| \\ &\quad + |\mathscr{E}_\mathrm{ac}(I_h y_\mathrm{a}) - \mathscr{E}_\mathrm{ac}(y_\mathrm{ac})| \\ &=: \mathrm{E}_1 + \mathrm{E}_2 + \mathrm{E}_3,\end{aligned}$$

and estimate the three terms E_j, $j = 1, 2, 3$, separately.

2.1. The term E_1*.* Since $y_\mathrm{a} \in \mathscr{Y}_\mathsf{B}$, and $\delta \mathscr{E}_\mathrm{a}(y_\mathrm{a}) = 0$, we can estimate

$$\begin{aligned}\big|\mathscr{E}_\mathrm{a}(I_h y_\mathrm{a}) - \mathscr{E}_\mathrm{a}(y_\mathrm{a})\big| &= \Big| \big\langle \delta \mathscr{E}_\mathrm{a}(y_\mathrm{a}), I_h y_\mathrm{a} - y_\mathrm{a} \big\rangle \\ &\qquad + \int_0^1 \big\langle \delta \mathscr{E}_\mathrm{a}\big((1-t) y_\mathrm{a} + t I_h y_\mathrm{a}\big) - \delta \mathscr{E}_\mathrm{a}(y_\mathrm{a}), I_h y_\mathrm{a} - y_\mathrm{a} \big\rangle \,\mathrm{d}t \Big| \\ &\le \int_0^1 \Big| \big\langle \delta \mathscr{E}_\mathrm{a}\big((1-t) y_\mathrm{a} + t I_h y_\mathrm{a}\big) - \delta \mathscr{E}_\mathrm{a}(y_\mathrm{a}), I_h y_\mathrm{a} - y_\mathrm{a} \big\rangle \Big| \,\mathrm{d}t\end{aligned}$$

For each $t \in [0, 1]$ we use Lemma 5.3 (replacing u_h with u in its formulation) to further estimate

$$\big| \big\langle \delta \mathscr{E}_\mathrm{a}\big((1-t) y_\mathrm{a} + t I_h y_\mathrm{a}\big) - \delta \mathscr{E}_\mathrm{a}(y_\mathrm{a}), I_h y_\mathrm{a} - y_\mathrm{a} \big\rangle \big| \le t C_\mathrm{L} \big\| \nabla \bar{y}_\mathrm{a} - \nabla \overline{I_h y_\mathrm{a}} \big\|^2_{\mathrm{L}^2},$$

where $C_{\rm L}$ depends on $\mu_{\rm a}(y_{\rm a})$ and $\mu_{\rm a}(I_h y_{\rm a}) \geq \min\{\mu_{\rm a}(y_{\rm a}), m\}$, and apply Lemma 4.6, to obtain

$$\begin{aligned}\big|\mathscr{E}_{\rm a}(I_h y_{\rm a}) - \mathscr{E}_{\rm a}(y_{\rm a})\big| &\leq \max_{t\in[0,1]} \big|\big\langle \delta\mathscr{E}_{\rm a}\big((1-t)y_{\rm a} + tI_h y_{\rm a}\big) - \delta\mathscr{E}_{\rm a}(y_{\rm a}), I_h y_{\rm a} - y_{\rm a}\big\rangle\big| \\ &\leq C_1 \inf_{\tilde{y}_{\rm a}\in\Pi_2(y_{\rm a})} \big\|h\nabla^2\tilde{y}_{\rm a}\big\|^2_{{\rm L}^2(\Omega_{\rm c})}, \end{aligned} \tag{7.6}$$

where C_1 depends only on $\mu_{\rm a}(y_{\rm a})$ and on m.

2.2 The term E_3. The term E_3 can be estimated in a similar manner as E_1. Following closely the proof of the Lipschitz estimate for $\delta\mathscr{E}_{\rm a}$, Lemma 5.3, one can prove that, if $y_h^{(j)} \in \mathscr{Y}_h$, $j = 1, 2$, then

$$\big|\big\langle \delta\mathscr{E}_{\rm ac}(y_h^{(1)}) - \delta\mathscr{E}_{\rm ac}(y_h^{(2)}), u_h\big\rangle\big| \leq C_{\rm L}\|\nabla y_h^{(1)} - \nabla y_h^{(2)}\|_{{\rm L}^2(\Omega)}\|\nabla u_h\|_{{\rm L}^2(\Omega)}$$

for $u_h \in \mathscr{U}_h$, where $C_{\rm L} = C_{\rm L}(\min\{\mu_{\rm c}(y_h^{(1)}), \mu_{\rm c}(y_h^{(2)})\})$. Repeating the first part of the argument in step 2.1, and using the H^1-norm error estimate (7.2), we obtain

$$\begin{aligned}|\mathscr{E}_{\rm ac}(I_h y_{\rm a}) - \mathscr{E}_{\rm ac}(y_{\rm ac})| &\leq C_3'\big\|\nabla I_h y_{\rm a} - \nabla y_{\rm ac}\big\|^2_{{\rm L}^2} \\ &\leq C_3 \inf_{\tilde{y}_{\rm a}\in\Pi_2(y_{\rm a})} \big\|h\nabla^2\tilde{y}_{\rm a}\big\|^2_{{\rm L}^2(\Omega_{\rm c})}, \end{aligned} \tag{7.7}$$

where C_3' and C_3 depend on m and on the shape regularity of $\mathcal{T}_h$, and C_3 depends also on γ.

2.3. The term E_2. Estimating this term requires more work. In Lemma 7.2 below, we prove that

$$\big|\mathscr{E}_{\rm a}(I_h y_{\rm a}) - \mathscr{E}_{\rm ac}(I_h y_{\rm a})\big| \leq C_2 \inf_{\tilde{y}\in\Pi_2(y_{\rm a})} \big\|h^{1/2}\nabla^2\tilde{y}\big\|^2_{{\rm L}^2(\Omega_{\rm c})}, \tag{7.8}$$

where C_2 depends on $\mu_{\rm c}(I_h y_{\rm a}) \geq m$, and on the shape regularity of $\mathcal{T}_h$.

2.4. Conclusion. Combining (7.6), (7.7), and (7.8) yields the energy error estimate (7.3) and concludes the proof of the theorem. □

Lemma 7.2. *Let $y \in \mathscr{Y}$ such that $\mu_{\rm c}(I_h y) > 0$, then*

$$\big|\mathscr{E}_{\rm a}(I_h y) - \mathscr{E}_{\rm ac}(I_h y)\big| \leq C_2^E \inf_{\tilde{y}\in\Pi_2(y)} \big\|h^{1/2}\nabla^2\tilde{y}\big\|^2_{{\rm L}^2(\Omega_{\rm c})}, \tag{7.9}$$

where $C_2^E = c_2' \sum_{r\in\mathbb{L}_*} M_2(\mu_{\rm c}(I_h y)|r|)|r|^4$, and c_2' depends only on the shape regularity of $\mathcal{T}_h$.

Proof. First note that the difference $\mathscr{E}_{\rm a}(y_h) - \mathscr{E}_{\rm ac}(y_h)$ depends only on continuum bonds:

$$\mathscr{E}_{\rm a}(y_h) - \mathscr{E}_{\rm ac}(y_h) = \sum_{b\in\mathcal{B}_{\rm c}} \left\{\phi(D_b y_h) - \fint_b \phi(\nabla_b y_h)\,{\rm d}b\right\}.$$

For each $b \in \mathcal{B}_{\rm c}$, we have

$$\begin{aligned}\phi(\nabla_b y_h) &= \phi(D_b y_h) + \phi'(D_b y_h)\cdot(\nabla_b y_h - D_b y_h) \\ &\quad + \int_0^1 \Big[\phi'\big(t\nabla_b y_h + (1-t)D_b y_h\big) - \phi'(D_b y_h)\Big]\,{\rm d}t\cdot(\nabla_b y_h - D_b y_h).\end{aligned}$$

Since $\phi'(D_b y_h)$ is a constant on the bond b and using (3.2) and a Lipschitz bound for ϕ' inside the integral over t, we obtain

$$\left| \fint_b \left[\phi(\nabla_b y_h) - \phi(D_b y_h) \right] \mathrm{db} \right| \leq \tfrac{1}{2} M_{|b|} \fint_b \left| \nabla_b y_h - D_b y_h \right|^2 \mathrm{db},$$

where $M_{|b|} = M_2(\mu_{\mathrm{c}}(y_h)|b|)$.

Summing over all bonds $b \in \mathcal{B}_{\mathrm{c}}$ yields the estimate

$$\left| \mathscr{E}_{\mathrm{a}}(y_h) - \mathscr{E}_{\mathrm{ac}}(y_h) \right| \leq \frac{1}{2} \sum_{b \in \mathcal{B}_{\mathrm{c}}} M_{|b|} \fint_b \left| \nabla_b y_h - D_b y_h \right|^2 \mathrm{db}, \tag{7.10}$$

which is identical to $e(y_h)^2$ defined in (5.7), with $p = 2$ and $a_b = 1$. Hence, we can use (5.14) and (5.16) to obtain

$$\left| \mathscr{E}_{\mathrm{a}}(y_h) - \mathscr{E}_{\mathrm{ac}}(y_h) \right| \leq C_1^E \left\| [\nabla y_h] \right\|_{\mathrm{L}^2(\Omega_{\mathrm{c}})}^2, \tag{7.11}$$

where $C_1^E = c_1' \sum_{r \in \mathbb{L}_*} M_2(\mu_{\mathrm{c}}(y_h)|r|)|r|^4$, and c_1' depends only on the shape regularity of $\mathcal{T}_h$.

The estimate (7.9) follows immediately from Lemma 4.4. □

7.2. **Optimal mesh design.** In this section we develop heuristics on the choice of atomistic region sizes and coarsening rates of the finite element mesh, in order to obtain error estimates in terms of the number of degrees of freedom. For the sake of generality we will slightly deviate from the assumptions and results of our analysis. Throughout this section, we will liberally make use of the symbols $\lesssim$ and $\eqsim$ to indicate bounds up to constants that are independent of the mesh parameters (but may depend on the shape regularity).

Recall that Ω has diameter $O(N)$, and consider an atomistic region of diameter $O(K)$ such that $\frac{K}{N} \leq C < 1$ (i.e., the atomistic region does not occupy most of the domain Ω), with a defect in the centre of the atomistic region. We conjecture that (7.2) holds for general $p \in [1, \infty]$,

$$\left\| \nabla \bar{y}_{\mathrm{a}} - \nabla y_{\mathrm{ac}} \right\|_{\mathrm{L}^p(\Omega)} \lesssim \inf_{\tilde{y}_{\mathrm{a}} \in \Pi_2(y_{\mathrm{a}})} \left\| h \nabla^2 \tilde{y}_{\mathrm{a}} \right\|_{\mathrm{L}^p(\Omega_{\mathrm{c}})}. \tag{7.12}$$

We assume that, for some "good" interpolant $\tilde{y}_{\mathrm{a}}$ (e.g., the HCT interpolant discussed in [25, Remark 4.1]) we have the decay property

$$\left| \nabla^2 \tilde{y}_{\mathrm{a}}(x) \right| \eqsim r^{-\beta}, \tag{7.13}$$

where $\beta > 0$, and where r denotes the distance from the defect. For example, it can be observed numerically that $\beta = 2$ for a dislocation [9], and $\beta = 3$ for a vacancy (§8).

Suppose that mesh $\mathcal{T}_h$ has mesh size function $h(r) \eqsim h_K (r/K)^{\alpha}$, where $h_K \geq 1$ and $\alpha > 0$ are the refinement parameters that we want to optimize. We have shown (7.2) only under the assumption that $h = 1$ on $\partial\Omega_{\mathrm{a}}$, which would require us to choose $h_K \eqsim 1$, however, for the sake of argument, we might assume that (7.12) still holds for more general h_K. Our analysis below shows that $h_K \eqsim 1$ is in fact a quasi-optimal choice.

In terms of the parameters K, h_K, α, (7.12) can be rewritten as

$$\left\| \nabla \bar{y}_{\mathrm{a}} - \nabla y_{\mathrm{ac}} \right\|_{\mathrm{L}^p(\Omega)} \lesssim \left(\int_K^N \left(h_K \left(\tfrac{r}{K} \right)^{\alpha} r^{-\beta} \right)^p r \, \mathrm{d}r \right)^{1/p} =: \mathrm{Err}, \tag{7.14}$$

	Parameter Regime	Err	DoF
1.	$\beta > 1$ and $p > \frac{2}{\beta-1}$	$\mathrm{DoF}^{1/p-\beta/2}$	K^2
2.	$\beta > 1$ and $p = \frac{2}{\beta-1}$	$\mathrm{DoF}^{-1/2}(\log \frac{N}{K})^{1/2+1/p}$	$K^2 \log \frac{N}{K}$
3.	$\beta \leq 1$ or $p < \frac{2}{\beta-1}$	$\mathrm{DoF}^{-1/2} N^{1/2+1/p-\beta/2}$	$K^2 \big(\frac{N}{K}\big)^{2-2\alpha}$

TABLE 2. Convergence rates in terms of degrees of freedom for optimised choices of the atomistic region and finite element mesh. In all cases $\alpha = \beta p/(2+p)$ and $h_K \eqsim 1$ are quasi-optimal, leaving the atomistic domain size, K, as the remaining free parameter. All quantities are understood as approximate orders of magnitude.

and the number of degrees of freedom approximated by

$$K^2 + \int_K^N \frac{1}{h(r)^2} r \,\mathrm{d}r \eqsim K^2 + \int_K^N \frac{r}{h_K^2 (r/K)^{2\alpha}} \mathrm{d}r =: \mathrm{DoF}. \tag{7.15}$$

In the following paragraphs we will obtain heuristic optimal choices for the mesh parameters, α and h_K, in terms of K, p, and β. It turns out that $\alpha = \beta p/(2+p)$ and $h_K \eqsim 1$ are always quasi-optimal. The remaining results are summarized in Table 2. In the case $p = 2$ and $\beta = 3$ (vacancy), for which the error estimate (7.12) was rigorously proved, we obtain $\mathrm{Err} \eqsim \mathrm{DoF}^{-1}$.

1. Equidistribution principle: We begin by applying the error equidistribution principle to obtain the optimal value for α (see [2, Sec. 5] for the case $p = 2$, which is readily generalized).

Consider a vertex q at distance r from the defect, with local mesh size $h(q) \equiv h(r)$. The error contribution of a degree of freedom associated with this vertex can be approximately estimated as

$$\big|h(r)\nabla^2 \tilde{y}_{\mathrm{a}}\big|^p h(r)^2 \eqsim \big(\tfrac{r}{K}\big)^{2\alpha} \big(h_K \big(\tfrac{r}{K}\big)^{\alpha} r^{-\beta}\big)^p h_K^2 = r^{\alpha(2+p)-\beta p} K^{-\alpha(2+p)} h_K^{p+2}.$$

According to the equidistribution principle, this quantity should be independent of r, i.e., $\alpha(2+p) - \beta p = 0$, which implies $\alpha = \frac{p}{2+p}\beta$.

We now consider three cases: $\alpha > 1$, $\alpha = 1$, and $\alpha < 1$. If $\beta > 1$ then these three cases correspond, respectively, to $p > \frac{2}{\beta-1}$, $p = \frac{2}{\beta-1}$, and $p < \frac{2}{\beta-1}$. If $\beta \leq 1$ then $\alpha < 1$ always holds.

2. Case 1: $\alpha > 1 \Leftrightarrow (\beta > 1$ and $p > \frac{2}{\beta-1})$: In this case, since $2 - 2\alpha < 0$, the approximate number of degrees of freedom is given by

$$\mathrm{DoF} \eqsim K^2 + \frac{N^{2-2\alpha} - K^{2-2\alpha}}{h_K^2 (2-2\alpha)} \eqsim K^2 + h_K^{-2} K^2 \eqsim K^2.$$

The error can be estimated as

$$\begin{aligned} \mathrm{Err} &= \tfrac{1}{p(\beta-\alpha)-2} h_K K^{2/p-\beta} \Big(1 - \big(\tfrac{K}{N}\big)^{p(\beta-\alpha)-2}\Big)^{1/p} \\ &\eqsim h_K K^{2/p-\beta} \eqsim h_K \mathrm{DoF}^{1/p-\beta/2}. \end{aligned} \tag{7.16}$$

Since the estimate for DoF does not depend on h_K, the optimal choice for h_K is $h_K \eqsim 1$, and the resulting convergence rate is $\mathrm{Err} \eqsim \mathrm{DoF}^{1/p-\beta/2}$.

Remark 7.2. (a) In the present case one can show directly (without using the equidistribution principle) that $h_K \eqsim 1$ and any α such that $1 < \alpha < \beta - \frac{2}{p}$, including $\alpha = \frac{p}{p+2}\,\beta$, are quasi-optimal, i.e., the error for this choice differs from the error for the best choice by at most a constant factor. This constant, however, tends to infinity as α tends to 1 or to $\beta - \frac{2}{p}$.

(b) Dropping the error equidistribution assumption and allowing $\alpha = 1$, while still assuming $p > 2/(\beta - 1)$, yields

$$\mathrm{Err} \eqsim \mathrm{DoF}^{1/p-\beta/2} \big(\log \tfrac{N}{K}\big)^{\beta/2-1/p}, \tag{7.17}$$

which is suboptimal in comparison with (7.16), but still acceptable if $\log \frac{N}{K}$ is moderate. For instance, in §8 we used $4 \le K \le 64$, $N = 128$, $\beta = 3$, and $p = 2$, in which case the error estimate is at most 4 times larger than for the optimal mesh.

The advantage of the choice $\alpha = 1$ is that it is easier to construct such a mesh: e.g., for a hexagonal region one can consider a mesh $\mathcal{T}_h$ consisting of hexagonal layers (i.e., hexagonal rings), each of the 6 sides of the layer is refined M times, so that the typical size of a triangle at distance r is $h_T \eqsim \frac{r}{M}$; see Figure 6(a). The condition $h_K \eqsim 1$ corresponds to $M \eqsim K$. □

Case 2: $\alpha = 1 \Leftrightarrow (\beta > 1$ *and* $p = \frac{2}{\beta-1})$*:* In this case, we obtain $h(r) \eqsim r h_K/K$, and hence the error and the number of degrees of freedom can be estimated as

$$\mathrm{Err} \eqsim h_K K^{-1} \big(\log \tfrac{N}{K}\big)^{1/p}, \quad \text{and} \quad \mathrm{DoF} \eqsim K^2 + \log \tfrac{N}{K} h_K^{-2} K^2.$$

For fixed Err, we choose K and h_K to minimize DoF by solving the corresponding constrained minimization problem in two variables (a slightly tedious but straightforward computation). We obtain for the optimal choices of K and h_K that $K\mathrm{Err} \eqsim (\log \frac{N}{K})^{1/p}$, and hence $h_K \eqsim 1$. Inserting these into the above expression for DoF one obtains

$$\mathrm{Err} \eqsim \mathrm{DoF}^{-1/2} \big(\log \tfrac{N}{K}\big)^{1/2+1/p} \quad \text{and} \quad \mathrm{DoF} \eqsim K^2 \log \tfrac{N}{K}.$$

Case 3: $\alpha < 1 \Leftrightarrow (\beta \le 1$ *or* $p < \frac{2}{\beta-1})$*:* In this case we obtain the following estimates on Err and DoF:

$$\mathrm{Err} \eqsim h_K K^{-p\beta/(2+p)} N^{2/p-2\beta/(2+p)} = h_K K^{-\alpha} N^{2(1-\alpha)/p}, \quad \text{and}$$
$$\mathrm{DoF} \eqsim K^2 + h_K^{-2} K^{2p\beta/(2+p)} N^{2-2p\beta/(2+p)} = K^2 + h_K^{-2} K^{2\alpha} N^{2-2\alpha}.$$

Solving again the constrained optimization problem of minimizing DoF subject to keeping Err fixed, we obtain $K\mathrm{Err} \eqsim K^{1-\alpha} N^{(1-\alpha)/p}$, which yields once again $h_K \eqsim 1$,

$$\mathrm{Err} \eqsim \mathrm{DoF}^{-1/2} N^{1/2+1/p-\beta/2}, \quad \text{and} \quad \mathrm{DoF} \eqsim K^2 \big(\tfrac{N}{K}\big)^{2-2\alpha}.$$

8. Numerical Examples

We conducted several numerical experiments to confirm the convergence rates obtained in §7.2, and to experimentally verify stability of the a/c method near bifurcation points, where our stability analysis does not apply.

In all tests, the region of periodicity is a hexagon centered at the origin with each side of the length $N = 128$, with a defect placed near the origin (*cf.* Fig. 6).

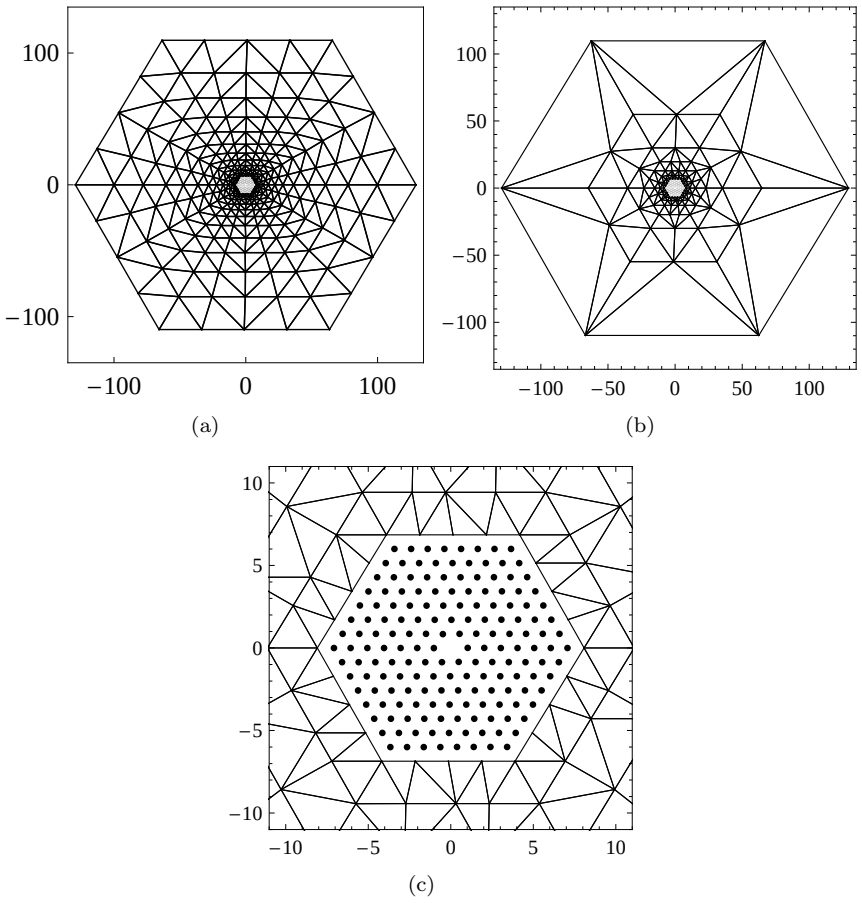
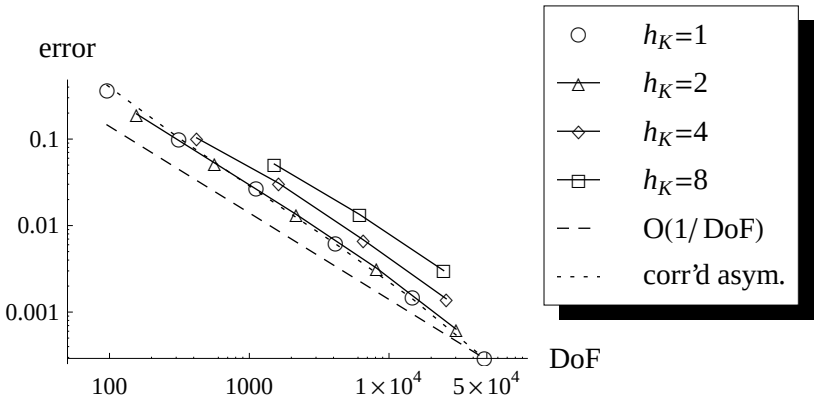


FIGURE 6. (a) Radially refined mesh, $\alpha = 1$; (b) Algebraically refined mesh, $\alpha = 3/2$; (c) Closeup of the atomistic region, $K = 8$ and $h_K = 2$.

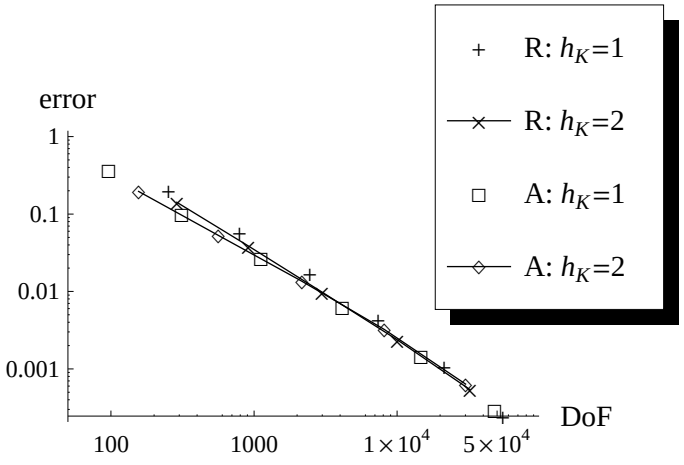
The different shape of the computational domain does not affect the results of the analysis.

The atomistic region forms a smaller hexagon, also centered at the origin, with side lengths K (*cf.* Fig. 6(c)). In the continuum region, either an algebraically refined mesh with $h(r) \approx h_K(r/K)^{3/2}$ (where r is the distance from the defect) or a radial mesh with $h(r) \approx h_K(r/K)$ is used (*cf.* Fig. 6). The parameter $\alpha = \frac{3}{2}$ is optimal for $\beta = 3$ and $p = 2$ (see Table 2).

In all experiments, the interaction potential is a Lennard-Jones potential with cut-off distance 3.1, measured in the reference hexagonal configuration. A nonlinear conjugate gradient solver with linesearch, and Laplace preconditioner to accelerate convergence, is used to find a stable equilibrium of the atomistic system.



(a)



(b)

FIGURE 7. Error of the computed solutions as a function of the number of degrees of freedom (DoF) in the vacancy example (§8.1): (a) Comparison of choices of h_K; (b) Comparison of algebraically refined ("A") and radial ("R") meshes.

8.1. **Vacancy.** We consider an example with a single vacancy defect. The macroscopic strain B is chosen as

$$\mathsf{B} = \begin{pmatrix} 1.01 & 0.01 \\ 0 & 0.99 \end{pmatrix}.$$

In Figure 7(a) we plot the relative error, $\frac{\|\nabla \bar{y}_{\mathrm{a}} - \nabla y_{\mathrm{ac}}\|_{\mathrm{L}^2(\Omega)}}{\|\nabla \bar{y}_{\mathrm{a}} - \nabla y_{\mathsf{B}}\|_{\mathrm{L}^2(\Omega)}}$ against the number of degrees of freedom (DoF). We observe first-order convergence, for the optimal choices $h_K = 1$ or $h_K = 2$, which is in full agreement with predictions made

in §7.2, and indicates that the error estimates obtained in the present paper are qualitatively sharp.

It is also interesting to compare the algebraically refined mesh with $\alpha = \frac{3}{2}$ and the radial mesh with $\alpha = 1$. The error for these two meshes is plotted in Figure 7(b). We observe that there is only a negligible difference in the error. This is in correspondence with the estimate (7.17): the effect of the term $\log \frac{N}{K}$ can only be observed only for large ratios N/K.

8.2. **Collapsed Cavity.** The second test case is a collapsed cavity defect, as considered in [30]. This defect is formed by removing eight atoms and applying a macroscopic compression to force the cavity to collapse and form two edge dislocations (see Figure 8(a) and [30] for a detailed test case description). Since they have opposite Burgers' vectors we obtain again $\beta = 3$ for the analysis in §7.2.

The results, presented in Figure 8(b) are similar to the single vacancy case, the main difference being that one requires larger K to represent the defect and that, for fixed (K, h_K), the error is higher than for the single vacancy case due to a slightly "stronger" defect. In particular, we observe again a first-order convergence in DoF.

8.3. **Stability Test for a Vacancy.** In addition to investigating the error in the a/c method, in terms of the number of degrees of freedom, we also conduct a series of numerical tests to explore the stability region of the a/c method (3.4). We used only radial mesh refinement in these tests.

Our first test case we set $\mathbb{V} = \{0\}$, and

$$\mathsf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1+t \end{pmatrix}, \quad 0 \le t.$$

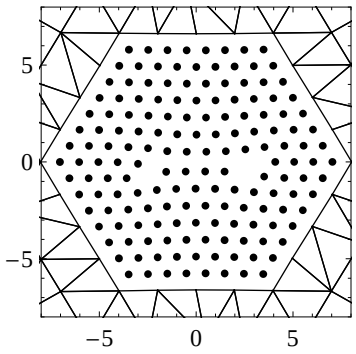
For increasing values of t the atomistic and a/c solutions are computed using Newton's method taking the previous critical point as the initial guess. The lowest eigenvalue of $\delta^2\mathscr{E}_{\rm a}$ (respectively, $\delta^2\mathscr{E}_{\rm ac}$), ignoring the two zero eigenvalues corresponding to translations, is used to determine whether the computed solution is a stable equilibrium, and thus determine the critical parameter $t_{\rm a}$ (respectively, $t_{\rm ac}$) at which the solution becomes unstable.

The results of the experiment are displayed in Table 3. We observe at least a quadratic convergence rate $|t_{\rm a} - t_{\rm ac}| \lesssim \mathrm{DoF}^{-2}$, and in particular, that the a/c method is stable up to this bifurcation point.

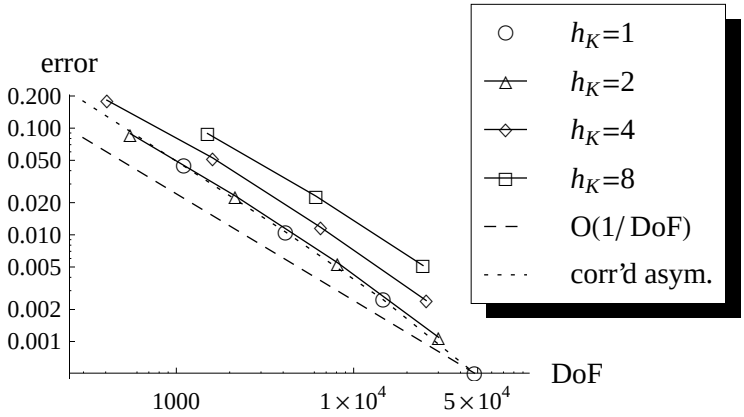
8.4. **Stability Test for a Bravais Lattice.** Our second stability test is conducted with a two parameter family of the macroscopic strains

$$\mathsf{B} = \begin{pmatrix} 1+s & 0.1 \\ 0 & 1+t \end{pmatrix} \tag{8.1}$$

for a lattice with no defects. In the (s,t)-plane we compare two regions of stability: the region of the stability of the atomistic model (as $N \to \infty$; *cf.* [11]), and the region of stability of the a/c method for $K = 16$ and $h_K = 2$. The results are shown in Figure 9. We observe that the stability region of the a/c method contains the stability region for the atomistic model, but that they are comparable up to numerical errors.



(a) Illustration of the defect.



(b) Graph of error.

FIGURE 8. Error of the computed solutions for the collapsed cavity test (§8.2) as a function of the number of degrees of freedom (DoF) for various choices of h_K.

K	DoF	$t_{\mathrm{ac}}, t_{\mathrm{a}}$	a
4	288	0.06104434	
8	912	0.05962851	2.15
16	2976	0.05950837	2.19
32	9984	0.05949904	2.53
64	32256	0.05949861	2.57
exact	105338	0.05949859	

TABLE 3. Stability test described in §8.3. K and $h_K = 2$ are the mesh parameters, DoF the number of degrees of freedom, $t_{\mathrm{ac}}, t_{\mathrm{a}}$ the computed critical parameters, and a the estimated convergence rate: $|t_{\mathrm{ac}} - t_{\mathrm{a}}| \approx \mathrm{DoF}^a$.

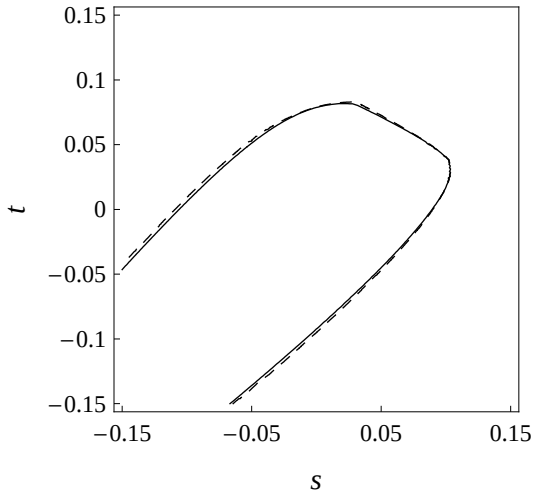


FIGURE 9. Stability regions of the atomistic model (solid line) and the a/c method for $K = 16$ and $M = 8$ (dashed line). The axis variables, s and t, are the parameters for the macroscopic strain (8.1). One can observe that the stability region of the a/c method contains the stability region for the atomistic model, and that the discrepancy is "small".

We believe that the minor visual difference between the two regions is caused by a finite size of the domain and the discretization of the continuum region. It would require extensive calculations to verify that the stability region of the a/c method indeed converges to the stability region of the atomistic model as DoF $\to \infty$.

CONCLUSION

We have presented a comprehensive *a priori* error analysis of a practical energy based atomistic/continuum coupling method, recently proposed in [30]. The method and the analysis are valid in two dimensions, for pair-potential interactions, and in the presence of simple defects.

The main theoretical question left open in our analysis is whether the a/c method is stable up to bifurcation points. This is a question first posed in [5] as a fundamental ingredient in understanding a/c methods. Our numerical experiments in §8.3 and §8.4 indicate that the error in the stability regions between the atomistic model and the a/c method is indeed "small", however, establishing such a result rigorously appears to be challenging.

Among the other interesting questions motivated by our analysis are: (1) Rigorously establishing the stability assumption (7.1), for example, following the discussion in Remark 7.1. (2) Developing a regularity theory for crystal defects, to make the analysis in §7.2 rigorous. In particular, this would allow for optimal *a priori* mesh refinement and remove the need for mesh adaptivity. (3) Extending the analysis to other classes of defects. While treating impurities should be straightforward with the present techniques, other defects with zero Burgers vector such as interstitials, or dislocation dipoles, require a more advanced account of stability.

An extension to dislocations would in addition require a more general consistency analysis as dislocations do not have an underlying reference configuration, which is a Bravais lattice.

Appendix A. Proofs of Some Auxiliary Results

Proof of Lemma 2.2. 1. Proof of (2.1)*:* The first result is motivated by the observation that the quadratic form $a[r] := \sum_{j=1}^{6} \left|\mathsf{G}\mathsf{Q}_6^j r\right|^2$ has hexagonal symmetry, that is, $a[\mathsf{Q}_6 r] = a[r]$ for all $r \in \mathbb{R}^2$. Suppose that a is represented by the symmetric matrix $\mathsf{A} \in \mathbb{R}^{2\times 2}$, $a[r] = r^\top \mathsf{A} r$, then $\mathsf{Q}_6^\top \mathsf{A}\mathsf{Q}_6 = \mathsf{A}$. By equating the entries in this matrix one obtains that A must in fact be a multiple of the identity. In particular, this implies that $a[r] = |r| a[e_1]$, and a direct computation yields (2.1).

2. Proof of (2.2)*:* The second result is motivated by the observation that the map $\mathsf{G} \mapsto \sum_{j=1}^{6} \left[(\mathsf{Q}_6^j r)^\top \mathsf{G} (\mathsf{Q}_6^j r)\right]^2$ defines a fourth-order tensor with hexagonal symmetry, and the usual major and minor symmetries. It is well-known that such a tensor is isotropic and must therefore take the form given in (2.2), though with still undermined Lamé parameters, which can be computed by judicious testing. In [25, App. A] we present a proof by a direct algebraic computation. □

The following well-known trace identity (see, e.g., in the proof of Lemma 2 in [24]) is used in the proof of Lemma 4.4 and Lemma 4.5.

Lemma A.1. *Let f be a face of a non-degenerate simplex $T \subset \mathbb{R}^d$, q_f the corner of T not contained in f, and $|f|$ the $(d-1)$-dimensional area of f; then*

$$\frac{|T|}{|f|} \int_f w \,\mathrm{d}s = \int_T w \,\mathrm{dV} + \frac{1}{2} \int_T (x - q_f) \cdot \nabla w \,\mathrm{dV} \quad \forall w \in \mathrm{W}^{1,1}(T). \tag{A.1}$$

Proof of Lemma 4.4. Let $y_h = I_h y$ and $\tilde{y} \in \Pi_2(y)$. Since $\tilde{y} \in \mathrm{C}^1(\mathbb{R}^d)$, we have the following estimate,

$$\begin{aligned} h_f [\nabla y_h]_f &= \left| \int_f \left(\nabla (y_h - \tilde{y})^+ - \nabla (y_h - \tilde{y})^- \right) \mathrm{d}s \right| \\ &\le \left| \int_f \nabla (y_h - \tilde{y})^+ \,\mathrm{d}s \right| + \left| \int_f \nabla (y_h - \tilde{y})^- \,\mathrm{d}s \right|. \end{aligned}$$

We deduce from (A.1), choosing $w = \nabla(y_h - \tilde{y})$ and $T = T_\pm$, that

$$\frac{|T_\pm|}{h_f} \left| \int_f \nabla (y_h - \tilde{y})^\pm \,\mathrm{d}s \right| \le \left\| \nabla y_h - \nabla \tilde{y} \right\|_{\mathrm{L}^1(T_\pm)} + \tfrac{1}{2} h_{T_\pm} \left\| \nabla^2 \tilde{y} \right\|_{\mathrm{L}^1(T_\pm)}.$$

Note, moreover, that $|T_\pm|/h_f \ge \frac{1}{C_f'} h_T$, where C_f' depends only on the shape regularity of $T_\pm$.

Recalling that $y_h = I_h y$, we can use Lemma (4.3) to deduce that

$$\frac{h_{T^\pm}}{C_f'} \left| \int_f \nabla (y_h - \tilde{y})^\pm \,\mathrm{d}s \right| \le \left(\tilde{C}_h + \tfrac{1}{2} \right) h_{T^\pm} \left\| \nabla^2 \tilde{y} \right\|_{\mathrm{L}^1(T_\pm)},$$

which immediately yields (4.6) for $p = 1$:

$$\left\| [\nabla y_h]_f \right\|_{\mathrm{L}^1(f)} \le C_f' \left(\tilde{C}_h + \tfrac{1}{2} \right) \left\| \nabla^2 \tilde{y} \right\|_{\mathrm{L}^1(T_+ \cup T_-)}. \tag{A.2}$$

Using similar calculations it is also easy to prove, for $p = \infty$:

$$\big|[\nabla y_h]_f\big| \leq 2\tilde{C}_h \big\|h\nabla^2 \tilde{y}\big\|_{\mathrm{L}^\infty(T_+ \cup T_-)}.$$

Applying the Riesz–Thorin interpolation theorem, we obtain (4.6) for all p. The estimate (4.7) is an immediate consequence of (4.6). □

Lemma A.2. *Let $f \in \mathcal{F}_\mathrm{a}$, $f \subset \tau \in \mathcal{T}_\mathrm{a}$ and let $w : \tau \to \mathbb{R}^k$ be piecewise constant with respect to the mesh $\mathcal{T}_h$; then*

$$|\tau| \left| \int_f w \,\mathrm{d}s \right| \leq \big\|w\big\|_{\mathrm{L}^1(\tau)} + \tfrac{1}{2}\big\|[w]\big\|_{\mathrm{L}^1(\mathcal{F}_h^\# \cap \mathrm{int}(\tau))}.$$

Proof. Assume, first, that $w_\varepsilon \in \mathrm{W}^{1,1}(\tau)^k$, then (A.1) implies

$$|\tau| \left| \int_f w_\varepsilon \,\mathrm{d}s \right| \leq \int_\tau |w_\varepsilon| \,\mathrm{dV} + \frac{1}{2} \int_\tau |\nabla w_\varepsilon| \,\mathrm{dV}.$$

Since $\mathrm{W}^{1,1}(\tau)^k$ is dense in $\mathrm{BV}(\mathrm{int}(\tau))^k$ (which contains piecewise constant functions) in the strict topology [8, Sec. 5.2.2], it follows that

$$|\tau| \left| \int_f w \,\mathrm{d}s \right| \leq \int_\tau |w| \,\mathrm{dV} + \frac{1}{2} |D'w|(\mathrm{int}(\tau))$$

as well, where $|D'w|$ denotes the total variation measure of w. Using integration by parts it is straightforward to show that

$$|D'w|(\mathrm{int}(\tau)) := \sup_{\substack{\psi \in \mathrm{C}_0^1(\tau)^{k\times 2} \\ |\psi| \leq 1}} \int_\tau w \cdot \mathrm{div}\psi \,\mathrm{dV} \leq \big\|[w]\big\|_{\mathrm{L}^1(\mathcal{F}_h^\# \cap \mathrm{int}(\tau))}. \qquad \square$$

Proof of Lemma 4.5. Fix an edge $f \in \mathcal{F}_\mathrm{a}$, $f \subset \tau$, $f = (q, q+\mathrm{a}_j)$; then, using Lemma A.2, we have

$$\begin{aligned} \big|(\nabla \bar{y}_h|_\tau)\mathrm{a}_j\big| &= \big|D_{\mathrm{a}_j} y_h(q)\big| = \left| \int_f \nabla y_h \mathrm{a}_j \,\mathrm{d}s \right| \\ &\leq |\tau|^{-1} \Big[\|\nabla y_h \mathrm{a}_j\|_{\mathrm{L}^1(\tau)} + \tfrac{1}{2}\big\|[\nabla y_h \mathrm{a}_j]\big\|_{\mathrm{L}^1(\mathcal{F}_h^\# \cap \mathrm{int}(\tau))} \Big]. \end{aligned}$$

There exists a constant C_3, depending only on the shape regularity of $\mathcal{T}_h$, such that $\mathrm{length}(\mathcal{F}_h^\# \cap \mathrm{int}(\tau)) \leq C_3$; hence, Hölder's inequality yields

$$\big|(\nabla \bar{y}_h|_\tau)\mathrm{a}_j\big| \leq |\tau|^{1/p'-1} \|\nabla y_h \mathrm{a}_j\|_{\mathrm{L}^p(\tau)} + \tfrac{1}{2} C_3^{1/p'} |\tau|^{-1} \big\|[\nabla y_h \mathrm{a}_j]\big\|_{\mathrm{L}^p(\mathcal{F}_h^\# \cap \mathrm{int}(\tau))}.$$

Summing over $j = 1, 2, 3$, applying Lemma 2.2, (2.1), and noting that all constants can be bounded independently of p, we obtain the result.

We also remark that, for $p = 2$, a careful computation yields

$$\|\nabla \bar{y}_h\|^2_{\mathrm{L}^2(\tau)} \leq 2\|\nabla y_h\|^2_{\mathrm{L}^2(\tau)} + \tfrac{2}{3^{1/4}} C_3 \big\|[\nabla y_h]\big\|^2_{\mathrm{L}^2(\mathcal{F}_h^\# \cap \mathrm{int}(\tau))}. \qquad \square$$

Proof of Lemma 5.7. We will prove a stronger statement, that (5.16) is true for any segment $b = (x, x + r)$, $x \in \mathbb{R}^2$, $r \in \mathbb{L}_*$. We hence extend the definitions of $J(b)$ and $n_\mathrm{j}(b)$ canonically to all such segments b.

Throughout this proof, we denote the set of vertices of $\mathcal{T}_h$ by $\mathcal{V}_h$. An inequality $\lesssim$ denotes a bound up to a constant that may only depend on the mesh regularity.

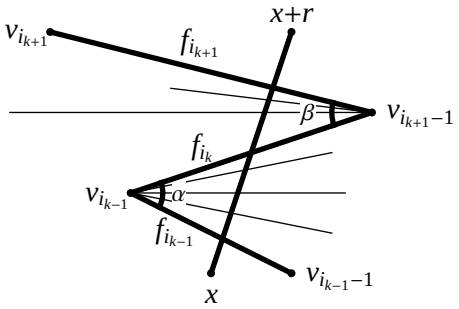


FIGURE 10. Illustration of counting the number of faces crossing a bond $b = (x, x + r)$. The bond b and the faces $f_{i_{k-1}}$, f_{i_k} and $f_{i_{k+1}}$ are bold lines. The rest of the faces $f \in J(b)$ are normal lines.

We will first reduce the statement to the case $\mathrm{int}(b) \cap \mathcal{V}_h = \emptyset$ (where $\mathrm{int}(b)$ always denotes the relative interior) and $n_{\mathrm{j}}(b) \neq 0$, and then estimate the lengths between points of intersections of b with $f \in J(b)$ and compare these lengths to $|b|$.

Case 1. ($\mathrm{int}(b) \cap \mathcal{V}_h \neq \emptyset$*)* Denote $x_0 = x$, $x_n = x + r$ and let $\mathrm{int}(b) \cap \mathcal{V}_h = \{x_1, \dots, x_{n-1}\}$, $n > 1$, where $x_1, \dots, x_n$ are sorted by increasing distance to x. Since any two points in $\mathcal{V}_h$ have at least distance 1, $n \leq |b|$.

If (5.16) holds for all $b_i = (x_{i-1}, x_i)$ $(i = 1, \dots, n)$ then we can estimate $n_{\mathrm{j}}(b)$ by respective contributions of b_i and contributions of those $f \in \mathcal{F}_h$ that contain any of points x_i. We will show that $n_{\mathrm{j}}(b_i) \lesssim |b_i| + 1$ (it falls under Case 2), and hence we can estimate

$$n_{\mathrm{j}}(b) \lesssim n + \sum_{i=1}^{n} n_{\mathrm{j}}(b_i) \lesssim n + \sum_{i=1}^{n} (|b_i| + 1) = 2n + |b| \leq 3|b| + 2,$$

which proves (5.16) for b.

Case 2.1. ($\mathrm{int}(b) \cap \mathcal{V}_h = \emptyset$ *and* $n_{\mathrm{j}}(b) = 0$*)* The estimate (5.16) is trivial in this case.

Case 2.2. ($\mathrm{int}(b) \cap \mathcal{V}_h = \emptyset$ *and* $n_{\mathrm{j}}(b) \neq 0$*)* In this case, $n_{\mathrm{j}} := n_{\mathrm{j}}(b)$ is simply the number of faces that cross b. Let $J(b) = \{f_1, \dots, f_m\}$, where f_i are sorted by increasing distance of $f_i \cap b$ to x. We need to prove that $n_{\mathrm{j}} \lesssim |b| + 1$. Any two faces, f_i and f_{i+1}, share exactly one common vertex $v_i \in \mathcal{V}_h$, $i = 1, \dots, n_{\mathrm{j}} - 1$. We also denote by v_0 the vertex of f_1 other than v_1, and by $v_{n_{\mathrm{j}}}$ the vertex of $f_{n_{\mathrm{j}}}$ other than $v_{n_{\mathrm{j}}-1}$.

It is of course possible that v_i coincides with v_{i+1} for some $i = 1, \dots, n_{\mathrm{j}} - 2$. Hence, denote the indices i of unique vertices v_i as

$$\mathcal{I} = \big\{ i \in \{1, \dots, n_{\mathrm{j}} - 2\} \,:\, v_i \neq v_{i+1} \big\} \cup \{0, n_{\mathrm{j}} - 1, n_{\mathrm{j}}\},$$

and let $\mathcal{I} = \{i_1, \dots, i_K\}$, where i_k is an increasing sequence.

If $K = 2$, then $n_{\mathrm{j}} = 1$. If $K = 3$ then n_{j} is bounded by the number of faces touching the vertex v_{i_2}, which is bounded by a constant depending only on the shape regularity of $\mathcal{T}_h$. Hence, we assume in the following that $K \geq 4$.

Split all faces in $J(b)$ into groups of faces between $f_{i_{k-1}}$ and $f_{i_{k+1}}$ $(k = 2, 4, \dots, 2\lfloor \frac{K}{2} \rfloor)$ and, if K is odd, the faces between $f_{i_{K-1}}$ and f_{i_K}. The number of faces in each

group is bounded by a finite number that depends only on the shape regularity of $\mathcal{T}_h$. To estimate the number of groups, notice that the distance between $b \cap f_{i_{k-1}}$ and $b \cap f_{i_{k+1}}$ can be bounded below in the following way (see illustration on Figure 10):

$$\begin{aligned}\operatorname{dist}(b \cap f_{i_{k-1}}, b \cap f_{i_{k+1}}) &\geq \operatorname{dist}(f_{i_{k-1}}, f_{i_{k+1}}) \\ &= \min\{\operatorname{dist}(v_{i_{k-1}-1}, f_{i_{k+1}}), \operatorname{dist}(v_{i_{k-1}}, f_{i_{k+1}})\} \\ &\geq \min\{\operatorname{dist}(v_{i_{k-1}-1}, f_{i_k}), \quad \operatorname{dist}(v_{i_{k-1}}, f_{i_{k+1}})\},\end{aligned}$$

Denote α and β to be angles formed by, respectively, the vertices $v_{i_{k-1}-1}, v_{i_{k-1}}, v_{i_{k+1}-1}$ and $v_{i_{k-1}}, v_{i_{k+1}-1}, v_{i_{k+1}}$ (*cf.* Fig. 10); then we have

$$\operatorname{dist}(b \cap f_{i_{k-1}}, b \cap f_{i_{k+1}}) \geq \min\{|f_{i_{k-1}}| \sin\alpha, |f_{i_k}| \sin\beta\} \geq \min\{\sin\alpha, \sin\beta\},$$

which is bounded below by a positive number that depends only on the shape regularity of $\mathcal{T}_h$. Thus, the number of such groups, $\lfloor \frac{K}{2} \rfloor$, is bounded by a constant multiple of $|b|$.

This finally establishes the estimate $n_{\mathrm{j}}(b) = \#(J(b)) \lesssim |b| + 1$. □

References

1. P. G. Ciarlet, *The finite element method for elliptic problems*, Classics in Applied Mathematics, vol. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, Reprint of the 1978 original. MR MR1930132
2. L. Demkowicz, Ph. Devloo, and J. T. Oden, *On an h-type mesh-refinement strategy based on minimization of interpolation errors*, Comput. Methods Appl. Mech. Engrg. **53** (1985), no. 1, 67–89. MR 812590 (87a:65177)
3. A. Demlow, D. Leykekhman, A. H. Schatz, and L. B. Wahlbin, *Best approximation property in the W_∞^1 norm on graded meshes*, Preprint.
4. M. Dobson and M. Luskin, *An optimal order error analysis of the one-dimensional quasicontinuum approximation*, SIAM Journal on Numerical Analysis **47** (2009), no. 4, 2455–2475.
5. M. Dobson, M. Luskin, and C. Ortner, *Accuracy of quasicontinuum approximations near instabilities*, J. Mech. Phys. Solids **58** (2010), no. 10, 1741–1757. MR 2742030
6. ———, *Stability, instability, and error of the force-based quasicontinuum approximation*, Arch. Ration. Mech. Anal. **197** (2010), no. 1, 179–202. MR 2646818
7. W. E, J. Lu, and J. Z. Yang, *Uniform accuracy of the quasicontinuum method*, Phys. Rev. B **74(21)** (2006), 214115.
8. L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992. MR 1158660 (93f:28001)
9. F. C. Frank and J. H. van der Merwe, *One-dimensional dislocations. I. static theory*, Proc. R. Soc. London **A198** (1949), 205–216.
10. M. Giaquinta, *Introduction to regularity theory for nonlinear elliptic systems*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1993. MR 1239172 (94g:49002)
11. T. Hudson and C. Ortner, *On the stability of bravais lattices and their cauchy-born approximations*, M2AN Math. Model. Numer. Anal. **46** (2012), 81–110.
12. Mrinal Iyer and Vikram Gavini, *A field theoretical approach to the quasi-continuum method*, J. Mech. Phys. Solids **59** (2011), no. 8, 1506–1535. MR 2848058
13. P. A. Klein and J. A. Zimmerman, *Coupled atomistic-continuum simulations using arbitrary overlapping domains*, J. Comput. Phys. **213** (2006), no. 1, 86–116. MR 2203436 (2006i:74031)
14. S. Kohlhoff and S. Schmauder, *A new method for coupled elastic-atomistic modelling*, Atomistic Simulation of Materials: Beyond Pair Potentials (V. Vitek and D. J. Srolovitz, eds.), Plenum Press, New York, 1989, pp. 411–418.
15. X. H. Li and M. Luskin, *A generalized quasi-nonlocal atomistic-to-continuum coupling method with finite range interaction*, IMA Journal of Numerical Analysis, to appear., doi:10.1093/imanum/drq049.
16. X. H. Li, M. Luskin, and C. Ortner, *Positive-definiteness of the blended force-based quasicontinuum method*, arXiv:1112.2528.

17. J. Lu and P. Ming, *Convergence of a force-based hybrid method for atomistic and continuum models in three dimension*, arXiv:1102.2523.
18. C. Makridakis, C. Ortner, and E. Süli, *A priori error analysis of two force-based atomistic/continuum models of a periodic chain*, Numer. Math. **119** (2011), no. 1, 83–121. MR 2824856
19. R. Miller and E. Tadmor, *A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods*, Modelling Simul. Mater. Sci. Eng. **17** (2009).
20. P. Ming and J. Z. Yang, *Analysis of a one-dimensional nonlocal quasi-continuum method*, Multiscale Modeling & Simulation **7** (2009), no. 4, 1838–1875.
21. M. Ortiz, R. Phillips, and E. B. Tadmor, *Quasicontinuum analysis of defects in solids*, Philosophical Magazine A **73(6)** (1996), 1529–1563.
22. C. Ortner, *The role of the patch test in 2D atomistic-to-continuum coupling methods*, arXiv:1101.5256v2.
23. C. Ortner, *A priori and a posteriori analysis of the quasinonlocal quasicontinuum method in 1D*, Math. Comp. **80** (2011), no. 275, 1265–1285. MR 2785458
24. C. Ortner and D. Praetorius, *On the convergence of adaptive nonconforming finite element methods for a class of convex variational problems*, SIAM J. Numer. Anal. **49** (2011), no. 1, 346–367.
25. C. Ortner and A. V. Shapeev, *Analysis of an Energy-based Atomistic/Continuum Coupling Approximation of a Vacancy in the 2D Triangular Lattice*, arXiv:1104.0311v1.
26. C. Ortner and E. Süli, *Analysis of a quasicontinuum method in one dimension*, M2AN Math. Model. Numer. Anal. **42** (2008), no. 1, 57–91. MR MR2387422
27. C. Ortner and H. Wang, *A priori error estimates for energy-based quasicontinuum approximations of a periodic chain*, Math. Models Methods Appl. Sc. **21** (2011), 2491–2521.
28. R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp. **38** (1982), no. 158, 437–445. MR 645661 (83e:65180)
29. A. V. Shapeev, *Consistent energy-based atomistic/continuum coupling for two-body potentials in three dimensions*, arXiv:1108.2991.
30. A. V. Shapeev, *Consistent energy-based atomistic/continuum coupling for two-body potentials in one and two dimensions*, Multiscale Model. Simul. **9** (2011), no. 3.
31. V. B. Shenoy, R. Miller, E. B. Tadmor, D. Rodney, R. Phillips, and M. Ortiz, *An adaptive finite element approach to atomic-scale mechanics–the quasicontinuum method*, J. Mech. Phys. Solids **47(3)** (1999), 611–642.
32. T. Shimokawa, J. J. Mortensen, J. Schiotz, and K. W. Jacobsen, *Matching conditions in the quasicontinuum method: Removal of the error introduced at the interface between the coarse-grained and fully atomistic region*, Phys. Rev. B **69(21)** (2004), 214104.
33. B. Van Koten, X. H. Li, M. Luskin, and C. Ortner, *A computational and theoretical investigation of the accuracy of quasicontinuum methods*, Numerical Analysis of Multiscale Problems (Ivan Graham, Tom Hou, Omar Lakkis, and Rob Scheichl, eds.), Springer Lecture Notes in Computational Science and Engineering, vol. 83, Springer, 2012.
34. S. P. Xiao and T. Belytschko, *A bridging domain method for coupling continua with molecular dynamics*, Comput. Methods Appl. Mech. Engrg. **193** (2004), no. 17-20, 1645–1669. MR 2069430

Mathematics Institute, Zeeman Building, University of Warwick, Coventry CV4 7AL, UK

E-mail address: c.ortner@warwick.ac.uk

A. V. Shapeev, Section of Mathematics, Swiss Federal Institute of Technology (EPFL), Station 8, CH-1015, Lausanne, Switzerland

Current address: School of Mathematics, 206 Church St. SE, University of Minnesota, Minneapolis, MN 55455, US

E-mail address: ashapeev@umn.edu