ESAIM: M2AN 46 (2012) 1275–1319
DOI: 10.1051/m2an/2012005

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

THE ROLE OF THE PATCH TEST IN 2D ATOMISTIC-TO-CONTINUUM COUPLING METHODS *

Christoph Ortner[1]

Abstract. For a general class of atomistic-to-continuum coupling methods, coupling multi-body interatomic potentials with a P_1-finite element discretisation of Cauchy–Born nonlinear elasticity, this paper adresses the question whether *patch test consistency* (or, absence of *ghost forces*) implies a first-order error estimate. In two dimensions it is shown that this is indeed true under the following additional technical assumptions: (i) an energy consistency condition, (ii) locality of the interface correction, (iii) volumetric scaling of the interface correction, and (iv) connectedness of the atomistic region. The extent to which these assumptions are necessary is discussed in detail.

Mathematics Subject Classification. 65N12, 65N15, 70C20.

Received March 14, 2011. Revised October 21, 2011.
Published online March 27, 2012.

1. Introduction

Defects in crystalline materials interact through their elastic fields far beyond their atomic neighbourhoods. An accurate computation of such defects requires the use of atomistic models; however, the size of the atomistic systems that are often required to accurately represent the elastic far-field makes atomistic models infeasible or, at the very least, grossly inefficient. Indeed, atomistic accuracy is often only required in a small neighbourhood of the defect, while the elastic far field may be approximated using an appropriate continuum elasticity model.

Atomistic-to-continuum coupling methods (a/c methods) aim to exploit this fact by retaining atomistic models only in small neighbourhoods of defects, and coupling these neighbourhoods to finite element discretisations of continuum elasticity models; see Figures 1c, 1d. By employing a coarse discretisation of the continuum model, such a process can achieve a considerable reduction in computational complexity, however, some of the first a/c methods suffered from the so-called "ghost force problem": while homogeneous deformations are equilibria of both the pure atomistic and the pure continuum model, they are not equilibria of certain a/c models [35,38,50] due to spurious forces – the "ghost forces" – that can arise at the interface between the atomistic and continuum regions.

Keywords and phrases. Atomistic models, atomistic-to-continuum coupling, quasicontinuum method, coarse graining, ghost forces, patch test, consistency.

* *This work was supported by the EPSRC Critical Mass Programme "New Frontiers in the Mathematics of Solids" (OxMoS) and by the EPSRC grant EP/H003096 "Analysis of atomistic-to-continuum coupling methods".*

[1] Mathematics Institute, Zeeman Building, University of Warwick, CV4 7AL Coventry, UK. c.ortner@warwick.ox.ac.uk

Article published by EDP Sciences

© EDP Sciences, SMAI 2012

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

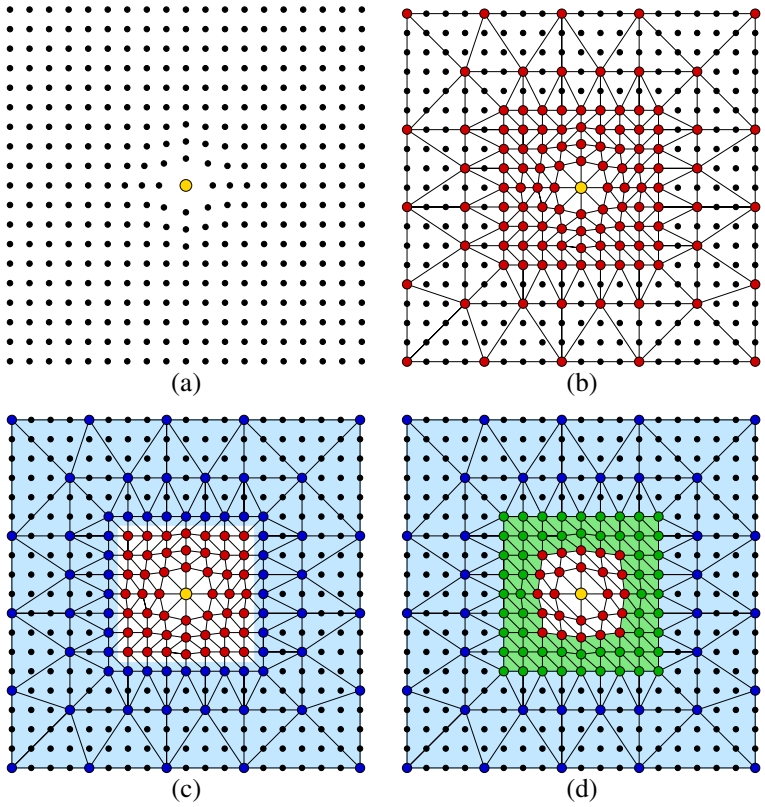


FIGURE 1. (a) A 2D atomistic configuration with an impurity that causes a large local deformation from the reference lattice $\varepsilon\mathbb{Z}^2$. (b) Triangulation of the deformed atomistic configuration to visualize the Galerkin projection described in Section 2.3.1. The positions of the large red atoms are free, while the positions of the small black atoms are constrained by the motion of the free atoms. (c) Visualization of the QCE method described in Section 2.3.2. The blue shaded region is the set $\Omega_{\rm c}^{\rm qce}$, the red atoms inside the white region are the set $\mathscr{L}_{\rm a}^{\rm qce}$ (both after deformation). (d) Visualization of an interface correction as described in Section 2.3.4. The blue shaded regions is the continuum region $\Omega_{\rm c}$, the green shaded region is the interface region $\Omega_{\rm i}$, and the white region is the atomistic region $\Omega_{\rm a}$. The mesh in $\Omega_{\rm i} \cup \Omega_{\rm a}$ is chosen so that it coincides with $\mathscr{T}_\varepsilon$ (*cf.* Sect. 3.1).

Much of the literature on a/c methods has focused on constructing a/c methods that did not exhibit, or reduced the effect of the "ghost forces" [5, 10, 18, 21, 23, 24, 34, 49–52, 56]. A straightforward solution was the introduction of *force-based* (*i.e.*, non-conservative) methods [10, 21, 26, 34, 50, 51]. The construction of accurate *energy-based* coupling mechanisms turned out to be more challenging. Several creative approaches providing partial solutions to the problem were suggested [18, 23, 49, 52], however, no general solution exists so far.

The inconsistency of early a/c methods is reminiscent of the inconsistency problems encountered in the early history of finite element methods. A simple criterion to test consistency of finite element methods is the *patch test* introduced by Bazeley *et al.* [6]; see also [7, 53]. The "ghost force problem" discussed above is precisely the failure of such a patch test. It is well known that, in general, the patch test is neither necessary nor sufficient for convergence of finite element methods; see, *e.g.*, [7, 53] where several variants of patch tests are discussed. The same is true for a/c methods: it was shown in [32] that a particular flavour of force-based a/c coupling typically has a consistency error of nearly 100%, even though it does pass the patch test.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Although a growing numerical analysis literature exists on the subject of a/c methods [10,12,14,16,29,31,37,40,42], it has so far focused primarily on one-dimensional model problems. (A notable exception is the work of Lu and Ming on force-based hybrid methods [31]. However, the techniques used therein require large overlaps and cannot accommodate sharp interfaces.) Only specific methods are analyzed; the question whether absence of "ghost forces" (or, *patch test consistency*, as we shall call it) in general guarantees substantially improved accuracy has neither been posed nor addressed so far. The purpose of the present paper is to fill precisely this gap. After introducing a general atomistic model and a general class of abstract a/c methods, and establishing the necessary analytical framework, it will be shown in Theorem 6.1, which is the main result of the paper, that in two dimensions patch test consistency together with additional technical assumptions implies first-order consistency of energy-based a/c methods.

1.1. Outline

Section 2 gives a detailed introduction to the construction of a/c methods. This section also develops a new notation that is well suited for the analysis of 2D and 3D a/c methods. In Section 2.3.3 we give a precise statement of the patch test consistency condition.

Section 3 contains a general framework for the *a priori* error analysis of a/c methods in $\mathrm{W}^{1,p}$-norms, similar to an error analysis of Galerkin methods with variational crimes. Several new technical results are presented in this section, such as the introduction of an oscillation operator to measure local smoothness of discrete functions (Section 3.2.1), an interpolation error estimate for piecewise affine functions (Sect. 3.2.2), and making precise the assumption made in much of the a/c numerical analysis literature that it is sufficient to estimate the modelling error without coarsening (Sect. 3.5). The main ingredient left open in this analysis is a stability assumption, which requires a significant amount of additional work and is beyond the scope of this paper.

Section 4 presents two 1D examples for modelling error estimates, which motivate the importance of patch test consistency, and discusses the modelling error estimates that can at best be expected if a method is patch test consistent.

Section 5 introduces the two main auxiliary results used in the 2D analysis of Section 6: (i) Shapeev's bond density lemma, which allows the translation between bond integrals and volume integrals; and (ii) a representation theorem for discrete divergence-free P_0-tensor fields.

Finally, Section 6 establishes the main result of this paper, Theorem 6.1: if an a/c method is patch test consistent and satisfies other natural technical assumptions, then it is also first-order consistent. The proof depends on a novel construction of stress tensors for atomistic models, related to the virial stress [2] (generalizing the 1D construction in [33,34]), and a corresponding construction for the stress tensor associated with the a/c energy. Moreover, we discuss in detail to what extent the technical conditions of Theorem 6.1 are required. For example, we show that, if the atomistic region is finite (*i.e.*, completely surrounded by the continuum region) then the condition of global energy consistency already follows from patch test consistency.

1.2. Sketch of the main result

In this section we discuss the main result in non-rigorous terms. Let $\mathscr{E}_{\mathrm{a}}$ be an atomistic energy functional and let $\mathscr{E}_{\mathrm{ac}}$ be an a/c energy functional, which uses the atomistic description in part of the computational domain, and couples it to a finite element discretisation of Cauchy–Born nonlinear elasticity (*cf.* Sect. 2.3.2 for its definition), with suitable interface treatment. Suppose that the following conditions are satisfied:

(i) $\mathscr{E}_{\mathrm{ac}}$ *is patch test consistent:* every homogeneous deformation is a critical point of $\mathscr{E}_{\mathrm{ac}}$ (*cf.* Sect. 2.1).
(ii) $\mathscr{E}_{\mathrm{ac}}$ *is globally energy consistent:* $\mathscr{E}_{\mathrm{ac}}$ is exact for homogeneous deformations (*cf.* Sect. 2.3.2).
(iii) *Locality and scaling of the interface correction:* the interface correction has (roughly) the same interaction range as the atomistic model, and has volumetric scaling (*cf.* Sect. 2.3.6).
(iv) The atomistic region Ω_{a} is connected.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

(v) *Stability:* for some $p \in [1,\infty]$, and for deformations y in a neighbourhood of the atomistic solution, the second variation $\delta^2\mathscr{E}_{\rm ac}(y)$ is stable, when understood as a linear operator from (discrete variants of) $\mathrm{W}^{1,p}$ to $\mathrm{W}^{-1,p}$ (*cf.* Sect. 3.3).
(vi) The finite element mesh in the continuum region is shape regular [8].
(vii) The solutions $y_{\rm a}$ and $y_{\rm ac}$ are "close" in the $\mathrm{W}^{1,\infty}$-norm.

Under assumptions (v)–(vii) we show in Section 3 that the error between the atomistic solution $y_{\rm a}$ and the a/c solution $y_{\rm ac}$ can be bounded by

$$\|\nabla y_{\rm a} - \nabla y_{\rm ac}\|_{\mathrm{L}^p(\Omega)} \lesssim \mathcal{E}_\varepsilon^{\rm model} + \mathcal{E}_h^{\rm ext} + \big\|\mathrm{h}\nabla^2 y_{\rm a}\big\|_{\mathrm{L}^p(\Omega_{\rm c})}, \tag{1.1}$$

where Ω is the computational domain, $\Omega_{\rm c}$ the continuum region, h a local mesh size function, $\mathcal{E}_h^{\rm ext}$ is the consistency error for the treatment of external forces, and $\mathcal{E}_\varepsilon^{\rm model}$ is the modelling error, which we describe next. Since $y_{\rm a}$ is not a continuous deformation, but a deformation of a discrete lattice, the various terms appearing above need to be interpreted with care. This is the focus of Section 3. For example, in the rigorous version of (1.1), we will replace $\nabla^2 y_{\rm a}$ by an oscillation of a suitably defined gradient.

The step outlined above does not distinguish different variants of a/c methods. The error introduced by the coupling mechanism and the continuum model is contained in the modelling error

$$\mathcal{E}_\varepsilon^{\rm model} = \big\|\delta\mathscr{E}_{\rm ac}(y_{\rm a}) - \delta\mathscr{E}_{\rm a}(y_{\rm a})\big\|_{\mathrm{W}_\varepsilon^{-1,p}},$$

where $\mathrm{W}_\varepsilon^{-1,p}$ is a suitably defined negative Sobolev norm on an atomistic grid. If $\mathscr{E}_{\rm ac}$ is not patch test consistent, then, typically, $\mathcal{E}_\varepsilon^{\rm model} = O(1)$. However, if conditions (i)–(iv) hold, and if the problem is set in either one or two space dimensions, then we will prove in Section 6 that

$$\mathcal{E}_\varepsilon^{\rm model} \lesssim \varepsilon \big\|\nabla^2 y_{\rm a}\big\|_{\mathrm{L}^p(\Omega_{\rm c}\cup\Omega_{\rm i})}, \tag{1.2}$$

where ε is the atomistic spacing and $\Omega_{\rm i}$ an interface region. A rigorous statement of (1.2) is given in Theorem 6.1, which is the main result of the paper.

1.3. Basic notational conventions

Vectors are denoted by lower case roman symbols, x, y, a, b, and so forth. Matrices are denoted by capital sans serif symbols $\mathsf{A}, \mathsf{B}, \mathsf{F}, \mathsf{G}$, and so forth. We will not distinguish between row and column vectors. If two vectors are multiplied, then it will be specified whether the operation is the dot product or the tensor product: for $a, b \in \mathbb{R}^k$, we define

$$a \cdot b := \sum_{j=1}^{k} a_j b_j, \quad \text{and} \quad a \otimes b := \big(a_i b_j\big)_{\substack{i=1,\dots,k\\ j=1,\dots,k}},$$

where, in $a \otimes b$, i is the row index and j the column index.

When matrix fields take the role of stress tensors, we will also call them tensor fields and use greek letters σ, Σ, and so forth.

The Euclidean norm of a vector and the Frobenius norm of a matrix are denoted by $|\cdot|$. The ℓ^p-norms, $p \in [1,\infty]$, of a vector or matrix are denoted by $|\cdot|_p$, and sometimes by $\|\cdot\|_{\ell^p}$. For $\varepsilon > 0$, the weighted ℓ^p-norms, on an index set $\mathscr{S}$, are defined as

$$\|a\|_{\ell^p_\varepsilon(\mathscr{S})} := \varepsilon^{1/p}\|a\|_{\ell^p(\mathscr{S})}.$$

The topological dual of a vector space $\mathscr{U}$ is denoted by $\mathscr{U}^*$, with duality pairing $\langle\cdot,\cdot\rangle$.

If $A \subset \mathbb{R}^d$ is a measurable set then $|A|$ denotes its d-dimensional volume. The symbol $\fint_A$ denotes $|A|^{-1}\int_A$, provided that $|A| > 0$. If A has a well-defined area or length, then these are denoted, respectively, by $\mathrm{area}(A)$ and $\mathrm{length}(A)$.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

For a measurable function $f : A \to \mathbb{R}$, $\|f\|_{\mathrm{L}^p(A)}$ denotes the standard L^p-norm. If $f : A \to \mathbb{R}^{k\times m}$, then $\|f\|_{\mathrm{L}^p(A)} := \| \, |f|_p \|_{\mathrm{L}^p(A)}$.

Partial derivatives with respect to a variable x_j, say, are denoted by $\partial/\partial x_j$. The Jacobi matrix of a differentiable function $f : A \to \mathbb{R}^k$ is denoted by ∂f. The symbol ∂ will also be redefined in some contexts, but used with essentially the same meaning as here.

When f is a deformation or a displacement, then we will also write $\nabla f = \partial f$. For $r \in \mathbb{R}^d$, the uni-directional derivative is denoted by

$$\nabla_r f(x) := \lim_{t \searrow 0} \frac{f(x+tr) - f(x)}{t},$$

whenever this limit exists. The symbol D_r denotes a finite difference operator, which will be defined in Section 2.1.2.

Additional notation will be defined throughout. A list of symbols, with references to their definitions is given in Appendix B.

2. Introduction to atomistic/continuum model coupling

In this section we introduce a general multi-body interaction model with periodic boundary conditions. Periodic boundary conditions are often used in molecular mechanics simulations as "artificial boundary conditions" modelling an infinite (or very large) crystal. For the purpose of the present work, they allow us to concentrate on the key issues of model coupling and coarse-graining, without addressing the difficult questions associated with the error committed by reducing an infinite to a finite domain, or with physical domain boundaries.

Next, we describe the construction of energy-based a/c methods that couple the atomistic interaction potential with a P_1-finite element discretization of the Cauchy–Born continuum model. We motivate the patch test ("ghost forces"), and why an interface correction is required to obtain accurate coupling schemes.

2.1. An atomistic model with periodic boundary condition

2.1.1. Periodic deformations

Let $d \in \{1, 2, 3\}$ denote the space dimension. We will, in subsequent sections, restrict our analysis to $d \in \{1, 2\}$, however, the introduction to a/c coupling methods is independent of the dimension.

For some $N \in \mathbb{N}$, and $\varepsilon := 1/N$, we define the periodic reference cell

$$\mathscr{L} := \varepsilon\big\{-N+1, \ldots, N\big\}^d.$$

The space of $2\mathbb{Z}^d$-periodic displacements of $\mathscr{L}^\# := \varepsilon\mathbb{Z}^d$ is given by

$$\mathscr{U} := \left\{ u : \mathscr{L}^\# \to \mathbb{R}^d \,\middle|\, u(x+\xi) = u(x) \text{ for all } \xi \in 2\mathbb{Z}^d, x \in \mathscr{L} \right\}.$$

A homogeneous deformation (or, *Bravais lattice*) of $\mathscr{L}^\#$ is a map $y_{\mathsf{A}} : \mathscr{L}^\# \to \mathbb{R}^d$, where $\mathsf{A} \in \mathbb{R}^{d\times d}$ and $y_{\mathsf{A}}(x) = \mathsf{A}x$ for all $x \in \mathscr{L}^\#$. We denote the space of *periodic deformations* of $\mathscr{L}^\#$ by

$$\mathscr{Y} := \left\{ y : \mathscr{L}^\# \to \mathbb{R}^d \,\middle|\, y - y_{\mathsf{A}} \in \mathscr{U} \text{ for some } \mathsf{A} \in \mathbb{R}^{d\times d} \right\},$$

and the space of deformations with prescribed *macroscopic strain* A by

$$\mathscr{Y}_{\mathsf{A}} := \{ y \in \mathscr{Y} \,|\, y - y_{\mathsf{A}} \in \mathscr{U} \}.$$

$\mathscr{Y}$ is a linear space, while $\mathscr{Y}_{\mathsf{A}}$ is an affine subspace of $\mathscr{Y}$.

For future reference we define notation that extends sets periodically: for any set $A \subset \mathbb{R}^d$ we define $A^\# := \bigcup_{\xi\in 2\mathbb{Z}^d}(\xi + A)$. This notation is consistent with the definition of $\mathscr{L}^\#$. If $\mathscr{A}$ is a family of sets, then $\mathscr{A}^\# := \{A^\# \,|\, A \in \mathscr{A}\}$.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

2.1.2. The atomistic energy

For a map $v : \mathscr{L}^{\#} \to \mathbb{R}^k$, $k \in \mathbb{N}$, and $r \in \mathbb{Z}^d \setminus \{0\}$, we define the finite difference operator

$$D_r v(x) := \frac{v(x + \varepsilon r) - v(x)}{\varepsilon}.$$

We assume that the *stored elastic energy* of a deformation $y \in \mathscr{Y}$ is given in the form

$$\mathscr{E}_{\rm a}(y) := \varepsilon^d \sum_{x \in \mathscr{L}} V\big(D_{\mathscr{R}} y(x)\big), \tag{2.1}$$

where $\mathscr{R} \subset \mathbb{Z}^d \setminus \{0\}$ is a finite *interaction range*, $D_{\mathscr{R}} y(x) := (D_r y(x))_{r \in \mathscr{R}}$, and where $V \in \mathrm{C}^2((\mathbb{R}^d)^{\mathscr{R}})$ is a multi-body interaction potential. Under these conditions, $\mathscr{E}_{\rm a} \in \mathrm{C}^2(\mathscr{Y})$.

The scaling of the lattice, of the finite difference operator, and of the energy were chosen to highlight the natural connection between molecular mechanics and continuum mechanics. For example, $\varepsilon^d \sum$ resembles an integral (or, a Riemann sum), while D_r resembles a directional derivative. It should be stressed, however, that ε is a fixed parameter of the problem, which is small but does *not* tend to zero.

The formulation (2.1) includes all commonly employed *classical* interatomic potentials (see, *e.g.*, [20, 47]): pair potentials of Lennard-Jones or Morse type, bond-angle potentials, embedded atom potentials, bond-order potentials, or any combination of the former, provided that they have a finite interaction range (see Sect. 2.2.2 for a discussion of the fact that the cut-off is taken in the reference configuration). The generality of the interaction potential also includes effective potentials obtained for in-plane or anti-plane deformations of 3D crystals.

No major difficulties should be expected in generalizing the analysis to infinite interaction ranges, provided the interaction strength decays sufficiently fast. A generalization of the analysis to genuine long-range interactions such as Coulomb interactions is not obvious.

2.1.3. Assumptions on the interaction potential

For $\mathbf{g} = (g_r)_{r \in \mathscr{R}} \in (\mathbb{R}^d)^{\mathscr{R}}$ we denote the first and second partial derivatives of V at $\mathbf{g}$, respectively, by

$$\partial_r V(\mathbf{g}) := \frac{\partial V(\mathbf{g})}{\partial g_r} \in \mathbb{R}^d, \quad \text{and} \quad \partial_{r,s} V(\mathbf{g}) := \frac{\partial^2 V(\mathbf{g})}{\partial g_r \partial g_s} \in \mathbb{R}^{d \times d}, \qquad \text{for } r, s, \in \mathscr{R}.$$

Throughout this work we assume the following *global* bound: there exist constants $M^{\rm a}_{r,s} \geq 0$, $r, s \in \mathscr{R}$, such that

$$\sup_{\mathbf{g} \in (\mathbb{R}^d)^{\mathscr{R}}} \big\|\partial_{r,s} V(\mathbf{g})\big\| \leq M^{\rm a}_{r,s}, \qquad \text{for all } r, s \in \mathscr{R}, \tag{2.2}$$

where $\|\cdot\|$ denotes the ℓ^2-operator norm. This assumption contradicts realistic interaction models and is made to simplify the notation; see Section 2.2.1 for further discussion of this issue.

In the subsequent analysis we will, in fact, never make direct use of the second partial derivatives $\partial_{r,s} V$, but only use the resulting Lipschitz property,

$$\big|\partial_r V(\mathbf{g}) - \partial_r V(\mathbf{h})\big| \leq \sum_{s \in \mathscr{R}} M^{\rm a}_{r,s} \, |g_s - h_s| \qquad \forall \mathbf{g}, \mathbf{h} \in (\mathbb{R}^d)^{\mathscr{R}}, \quad r \in \mathscr{R}. \tag{2.3}$$

The proof is straightforward. For future reference, we also define the constant

$$M^{\rm a} := \sum_{r,s \in \mathscr{R}} |r||s| M^{\rm a}_{r,s}. \tag{2.4}$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

2.1.4. The variational problem

Let $\mathscr{P}_{\rm a} \in {\rm C}^2(\mathscr{Y};\mathbb{R})$ be the *potential of external forces* modelling, for example, a substrate or an indenter. As explained in Section 2.2.3, one may also use such a potential to model simple point defects such as vacancies or impurities.

Given a potential of external forces $\mathscr{P}_{\rm a}$ and a macroscopic strain A, we consider the problem of finding local minimizers of the total energy $\mathscr{E}_{\rm a}^{\rm tot} := \mathscr{E}_{\rm a} + \mathscr{P}_{\rm a}$ in $\mathscr{Y}_{\mathsf{A}}$, in short,

$$y_{\rm a} \in \operatorname*{argmin}_{y \in \mathscr{Y}_{\mathsf{A}}} \mathscr{E}_{\rm a}^{\rm tot}(y), \tag{2.5}$$

where argmin denotes the set of *local* minimizers.

If $y_{\rm a}$ solves (2.5), then $y_{\rm a}$ is a *critical point* of $\mathscr{E}_{\rm a}^{\rm tot}$, that is,

$$\big\langle \delta\mathscr{E}_{\rm a}(y_{\rm a}) + \delta\mathscr{P}_{\rm a}(y_{\rm a}), u\big\rangle = 0 \qquad \forall u \in \mathscr{U}, \tag{2.6}$$

where, for a functional $\mathscr{E} \in {\rm C}^1(\mathscr{Y})$, we define the *first variation* of $\mathscr{E}$ at y, as

$$\langle \delta\mathscr{E}(y), z\rangle := \frac{\rm d}{{\rm d}t}\mathscr{E}(y+tz)\big|_{t=0} \qquad \text{for } y, z \in \mathscr{Y}.$$

If $\mathscr{E} \in {\rm C}^2(\mathscr{Y})$ then the second variation is defined analogously as

$$\langle \delta^2\mathscr{E}(y)z_1, z_2\rangle := \frac{\rm d}{{\rm d}t}\big\langle \delta\mathscr{E}(y+tz_1), z_2\big\rangle\big|_{t=0} \qquad \text{for } y, z_1, z_2 \in \mathscr{Y}.$$

The same notation will be used for functionals defined on different spaces.

2.1.5. The patch test for the atomistic model

The following proposition can be understood as the *patch test* for the atomistic model: *in the absence of external forces and defects a homogeneous lattice is always a critical point of* $\mathscr{E}_{\rm a}$.

Proposition 2.1. $\langle \delta\mathscr{E}_{\rm a}(y_{\mathsf{A}}), u\rangle = 0$ *for all* $u \in \mathscr{U}$ *and for all* $\mathsf{A} \in \mathbb{R}^{d\times d}$.

Proof. Let $y, z \in \mathscr{Y}$, then

$$\langle \delta\mathscr{E}_{\rm a}(y), z\rangle = \varepsilon^d \sum_{x\in\mathscr{L}} \sum_{r\in\mathscr{R}} \partial_r V\big(D_{\mathscr{R}} y(x)\big)\cdot D_r z(x). \tag{2.7}$$

Fix $\mathsf{A} \in \mathbb{R}^{d\times d}$ and $u \in \mathscr{U}$, then

$$\begin{aligned}\langle \delta\mathscr{E}_{\rm a}(y_{\mathsf{A}}), u\rangle &= \sum_{r\in\mathscr{R}} \varepsilon^d \sum_{x\in\mathscr{L}} \partial_r V\big(\mathsf{A}\mathscr{R}\big)\cdot D_r u(x) \\ &= \sum_{r\in\mathscr{R}} \partial_r V\big(\mathsf{A}\mathscr{R}\big)\cdot\Big\{\varepsilon^d \sum_{x\in\mathscr{L}} D_r u(x)\Big\} = 0,\end{aligned}$$

where we have used the fact that u is periodic. Above, and throughout, we use the notation $\mathsf{A}\mathscr{R} = (\mathsf{A}r)_{r\in\mathscr{R}} = D_{\mathscr{R}} y_{\mathsf{A}}(x)$. □

2.2. Remarks on the atomistic model

2.2.1. Invertibility of deformations

In Sections 2.1.2 and 2.1.3 we have assumed that $\mathscr{E}_{\rm a}$ is twice differentiable at *all* deformations $y \in \mathscr{Y}$, and that the second partial derivatives of the interaction potential are *globally bounded*.

However, realistic interaction potentials V take the value $+\infty$ if two atoms occupy the same position in space, and hence can only be differentiable at deformations that are one-to-one (*i.e.*, "true" deformations). With only minor additional technicalities such potentials can be admitted in the analysis. The global bound (2.2) would then be replaced by a local bound and certain explicit bounds on $D_{\mathscr{R}} y$; see, *e.g.*, [40, 41].

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

2.2.2. Reference cutoff

Another aspect of the atomistic energy (2.1), which makes it, at first glance, inappropriate for realistic applications is that the interaction potential V has a cut-off radius in the reference configuration. In atomistic models, atoms are unconstrained in their position and hence two atoms that are far apart in the reference configuration may be arbitrarily close, and hence interact, in the deformed configuration.

As a matter of fact, the reference cutoff in (2.1) is assumed only for the sake of brevity of the notation. One can, similarly as discussed in Section 2.2.1, take a more general form of the interaction potential that does not suffer from this drawback and make suitable assumptions on deformations under consideration that control the interaction neighbourhood.

2.2.3. Modelling crystal defects

Some simple crystal defects can be modelled via the potential of the external forces, $\mathscr{P}_{\rm a}$. The simplest example is an impurity, where a single atom is replaced with an atom from a different species. For (2.1) this means that the interaction potential is changed from $V(D_{\mathscr{R}}y(x))$ to $V^{\rm mod}(x; D_{\mathscr{R}}y(x))$ in a neighbourhood of the impurity. Alternatively, one may keep the original form of $\mathscr{E}_{\rm a}$ and define

$$\mathscr{P}_{\rm a}(y) = \varepsilon^d \sum_{x\in\mathscr{L}} \left[V^{\rm mod}(x; D_{\mathscr{R}}y(x)) - V(D_{\mathscr{R}}y(x))\right].$$

Similarly, a vacancy can be modelled by simply removing all interactions with a given atom. This would yield a difficulty with the "unused" degrees of freedom for the position of the vacancy atom, which could simply be removed from the system [41]. An interstitial (an additional atom) can also be modelled fairly easily, but one would need to augment the variable y with additional degrees of freedom for the position of the interstitial atom.

Dislocations, which possibly represent the most important class of crystal defects are, in general, more difficult to describe. In the atomistic minimization problem (2.5) they simply represent a special class of local minimizers, however, in the coupled atomistic/continuum models we discuss below most classes of dislocations are less straightforward to embed (but see [36, 55] for straightforward examples).

2.3. Construction of a/c coupling methods

The atomistic model problem (2.5) is a finite-dimensional optimisation problem and is therefore, in principle, solvable using standard optimisation algorithms. However, typical applications where atomistic models are employed require of the order 10^9 to 10^{12} atoms or more [35, 38]. It is therefore desirable to construct computationally efficient *coarse grained* models.

2.3.1. Galerkin projection

To motivate the idea of atomistic-to-continuum coupling we consider a crystal with a localized defect. Figure 1a shows a deformed 2D crystal with an impurity that repels its neighbouring atoms, causing a large local deformation. We observe that, except in a small neighbourhood of the defect, the atoms are arranged as a "smooth" deformation of the reference lattice $\mathscr{L}^{\#}$. It is therefore possible to approximate the atomistic configurations from a low-dimensional subspace constructed, for example, using a P_1-finite element method.

Let $\Omega := (-1, 1]^d$, and let $\mathscr{T}_h$ be a regular [8] triangulation of $\bar{\Omega}$, with vertices belonging to $\mathscr{L}^{\#}$, that can be extended periodically to a regular triangulation $\mathscr{T}_h^{\#}$ of $\mathbb{R}^d$. We make the convention that elements $T \in \mathscr{T}_h$ are closed sets. For $T \in \mathscr{T}_h$ we define $h_T := \mathrm{diam}(T)$, and for each $x \in \mathbb{R}^d$ we define $h(x) := \max\{h_T \,|\, T \in \mathscr{T}_h \text{ s.t. } x \in T\}$.

We define the P_1 finite element space

$$\mathrm{P}_1\left(\mathscr{T}_h^{\#}\right) := \left\{v_h : \mathscr{L}^{\#} \to \mathbb{R} \,\middle|\, v_h \text{ is piecewise affine w.r.t. } \mathscr{T}_h^{\#}\right\},$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

and we denote the spaces of piecewise affine displacements and deformations, respectively, by

$$\mathscr{U}_h := \mathscr{U} \cap \mathrm{P}_1\left(\mathscr{T}_h^{\#}\right)^d, \quad \mathscr{Y}_h := \mathscr{Y} \cap \mathrm{P}_1\left(\mathscr{T}_h^{\#}\right)^d, \quad \text{and} \quad \mathscr{Y}_{\mathsf{A},h} := \mathscr{Y}_{\mathsf{A}} \cap \mathrm{P}_1\left(\mathscr{T}_h^{\#}\right)^d.$$

For future reference, let $I_h : \mathscr{Y} \to \mathscr{Y}_h$ denote the nodal interpolation operator. We note that $I_h : \mathscr{Y}_{\mathsf{A}} \to \mathscr{Y}_{\mathsf{A},h}$ as well as $I_h : \mathscr{U} \to \mathscr{U}_h$. We also define $\mathrm{P}_1^{\#}(\mathscr{T}_h)$ to be the set of all $2\mathbb{Z}^d$-periodic functions $v_h \in \mathrm{P}_1(\mathscr{T}_h^{\#})$; *i.e.*, $\mathscr{U}_h = \mathrm{P}_1^{\#}(\mathscr{T}_h)^d$. Let $\mathscr{F}_h^{\#}$ denote the set of closed edges of the extended triangulation $\mathscr{T}_h^{\#}$, and let $\mathscr{F}_h$ denote the set of all edges $f \in \mathscr{F}_h^{\#}$ such that $\mathrm{area}(f \cap \Omega) \neq 0$. Finally, we introduce the spaces of piecewise constant functions $\mathrm{P}_0(\mathscr{T}_h)$, $\mathrm{P}_0(\mathscr{T}_h^{\#})$, and $\mathrm{P}_0^{\#}(\mathscr{T}_h)$, defined in a similar manner.

The *Galerkin approximation* of (2.5) is the coarse-grained minimization problem

$$y_{\mathrm{a},h} \in \operatorname*{argmin}_{y_h \in \mathscr{Y}_{\mathsf{A},h}} \mathscr{E}_{\mathrm{a}}^{\mathrm{tot}}(y_h), \tag{2.8}$$

where we recall that $\mathscr{E}_{\mathrm{a}}^{\mathrm{tot}} = \mathscr{E}_{\mathrm{a}} + \mathscr{P}_{\mathrm{a}}$.

For example, if we choose the triangulation $\mathscr{T}_h$ as in Figure 1b, then we obtain full atomistic resolution in the neighbourhood of the defect, while considerably reducing the overall number of degrees of freedom by coarsening the mesh away from the defect. In this way it is possible to obtain highly accurate approximations to nontrivial atomistic configurations. Under suitable technical assumptions it is not too difficult to employ the classical techniques of finite element error analysis in this context. Such analyses, including *a posteriori* grid generation, are given in [29,30,42].

Remark 2.2 (Regularity of atomistic solutions)**.** In order for the Galerkin projection method, or the subsequent a/c coupling methods, to be accurate we require "regularity" of atomistic solutions. Such a regularity theory does not exist at present, however, most numerical experiments performed on atomistic models for *simple lattices* indicate "smoothness" of atomistic deformations away from defects. The situation would be different for so-called *multi-lattices*, which require a homogenisation step and represent a far greater challenge.

2.3.2. Continuum region & Cauchy–Born approximation

The Galerkin projection (2.8) reduces the number of degrees of freedom considerably, however, the complexity of computing $\mathscr{E}_{\mathrm{a}}|_{\mathscr{Y}_h}$ is not reduced in the same manner. Due to the non-locality of the atomistic interaction $\mathscr{E}_{\mathrm{a}}|_{\mathscr{Y}_h}$ cannot be evaluated as easily as in the case of finite element methods for continuum mechanics. Several attempts have been made to use quadrature ideas to approximate $\mathscr{E}_{\mathrm{a}}$ and render (2.8) computationally efficient [19,22,25], however, it was shown in [32] that these approximations yield unacceptable consistency errors.

An alternative approach, proposed in [38], is to keep the full atomistic description for atomistically fine elements, while employing the *Cauchy–Born approximation* for coarser elements as well as an interface region. Following the terminology of [12] we call the resulting method the *QCE method* (the original energy-based quasicontinuum method).

To formulate this method we choose a set $\mathscr{L}_{\mathrm{a}}^{\mathrm{qce}} \subset \mathscr{L}$ of atoms that we wish to treat atomistically (the red atoms in Fig. 1c). Let $Q_\varepsilon(x) := x + \varepsilon(-\frac{1}{2}, \frac{1}{2}]^d$, and define

$$\Omega_{\mathrm{c}}^{\mathrm{qce}} := \Omega \setminus \bigcup_{x \in \mathscr{L}_{\mathrm{a}}^{\mathrm{qce}}} Q_\varepsilon(x)\,.$$

With this notation, the QCE energy functional is defined, for $y_h \in \mathscr{Y}_h$, as

$$\mathscr{E}_{\mathrm{qce}}(y_h) := \varepsilon^d \sum_{x \in \mathscr{L}_{\mathrm{a}}^{\mathrm{qce}}} V\big(D_{\mathscr{R}} y_h(x)\big) + \int_{\Omega_{\mathrm{c}}^{\mathrm{qce}}} W(\nabla y_h)\,\mathrm{d}x, \tag{2.9}$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

where $W : \mathbb{R}^{d\times d} \to \mathbb{R}$ is the *Cauchy–Born stored energy function*,

$$W(\mathsf{F}) := V(\mathsf{F}\mathscr{R}) = V\big((\mathsf{F}r)_{r\in\mathscr{R}}\big). \tag{2.10}$$

Note that $W(\mathsf{F})$ is the energy of a single atom in the Bravais lattice $\mathsf{F}\mathscr{L}^{\#}$. For future reference, we define $\partial W(\mathsf{F}) := (\partial_{\mathsf{F}_{ij}} W(\mathsf{F}))_{i,j=1}^{d}$.

One may readily check that the complexity of evaluating $\mathscr{E}_{\rm qce}$ (or its derivatives) is of the order $O(\#\mathscr{T}_h)$, that is, of the same order as the number of degrees of freedom. Moreover, it is easy to see that $\mathscr{E}_{\rm qce}(y_{\mathsf{A}}) = \mathscr{E}_{\rm a}(y_{\mathsf{A}})$ for all $\mathsf{A} \in \mathbb{R}^{d\times d}$. For future reference, we give a formal definition of this property.

Definition 2.3 (Global energy consistency)**.** We say that an energy functional $\mathscr{E} \in \mathrm{C}(\mathscr{Y})$ is *globally energy consistent* if

$$\mathscr{E}(y_{\mathsf{A}}) = \mathscr{E}_{\rm a}(y_{\mathsf{A}}) \qquad \forall \mathsf{A} \in \mathbb{R}^{d\times d}. \tag{2.11}$$

More generally one can show that $\mathscr{E}_{\rm qce}(y_h)$ is a good approximation to $\mathscr{E}_{\rm a}(y_h)$ if ∇y_h varies only moderately. Despite these facts, it turns out, as we discuss in Sections 2.3.3 and 4.2, that minimizers of $\mathscr{E}_{\rm qce}$ are poor approximations to minimizers of $\mathscr{E}_{\rm a}$.

Remark 2.4 (The Cauchy–Born model)**.** The Cauchy–Born model is a standard continuum model, for large deformations of single crystals. In the absence of defects, solutions of a pure Cauchy–Born model (no atomistic region) can provide excellent approximations to solutions of the atomistic model (2.5). For example, in [17] it is shown that, for smooth and small dead load external forces (forces that are independent of the atom positions), and under realistic stability assumptions on $\mathscr{E}_{\rm a}$, there exist solutions $y_{\rm a}$ of (2.5) and $y_{\rm c}$ of the Cauchy–Born model, such that

$$\left(\varepsilon^d \sum_{x\in\mathscr{L}} \sum_{j=1}^{d} \big|D_{e_j} y_{\rm a}(x) - D_{e_j} y_{\rm c}(x)\big|^2\right)^{1/2} \leq C\varepsilon^2,$$

where C depends on higher partial derivatives of V and on the regularity of $y_{\rm c}$. Hence, in the regime of "smooth elastic" deformations, the Cauchy–Born model can be considered an excellent approximation to the atomistic model (2.1).

2.3.3. The patch test

The *patch test* is often employed in the theory of finite element methods [6,7,53] as a simple test for consistency. The test also plays an important role in the design of a/c methods.

Definition 2.5 (Patch test consistency)**.** We say that an energy functional $\mathscr{E} \in \mathrm{C}^1(\mathscr{Y}_h)$ is *patch test consistent* if it satisfies

$$\big\langle \delta\mathscr{E}(y_{\mathsf{F}}), u_h \big\rangle = 0 \qquad \forall\, u_h \in \mathscr{U}_h, \quad \forall\, \mathsf{F} \in \mathbb{R}^{d\times d}. \tag{2.12}$$

The terminology "patch test consistency" is motivated by Proposition 2.1, where we have shown that the exact energy $\mathscr{E}_{\rm a}$ does satisfy the patch test (2.12).

However, the QCE energy $\mathscr{E}_{\rm qce}$ is *not* patch test consistent [50]. This result will be reviewed for a one-dimensional model problem in Section 4.2, where it will also be shown how failure of the patch test affects the consistency error.

Remark 2.6. In most of the a/c coupling literature the patch test is stated as the condition that

$$\left.\frac{\partial\mathscr{E}(y_h)}{\partial y_h(p)}\right|_{y_h = y_{\mathsf{F}}} = 0 \quad \text{for all finite element nodes } p.$$

It is straightforward to see that this condition is equivalent to the variational formulation given in (2.12).

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

2.3.4. Interface correction

In the engineering literature (see, *e.g.*, [35, 50]) the non-zero forces under homogeneous strains of patch test inconsistent a/c energies are usually dubbed "ghost forces". The discovery that $\mathscr{E}_{\rm qce}$ is not patch test consistent has resulted in a number of works constructing new a/c methods that removed or reduced the "ghost forces" [5, 18, 23, 24, 28, 49, 52, 54, 56]. In some cases, this is achieved through sacrificing a variational (*i.e.*, conservative, or, energy-based) formulation [10, 21, 26, 34, 46, 50, 51].

In the present paper we will focus only on energy-based a/c methods that are patch test consistent, *i.e.*, that remove the "ghost forces" altogether. None of the methods presently available in the literature have resolved this problem in its full generality, however, [18,23,49,52] present several interesting approaches and partial solutions. In the following, we present a generalization of the *geometrically consistent coupling method* [18], which is the most general approach but leaves some questions concerning its practical construction open.

Let $\mathscr{T}_h^{\rm c} \subset \mathscr{T}_h$ and $\Omega_{\rm c} = \cup \mathscr{T}_h^{\rm c}$. We assume that all atoms belonging to $\mathscr{L} \setminus {\rm int}(\Omega_{\rm c})$ are vertices of $\mathscr{T}_h$, and we define two sets $\mathscr{L}_{\rm a}^{\rm gc}, \mathscr{L}_{\rm i}^{\rm gc}$ such that

$$\mathscr{L}_{\rm a}^{\rm gc} \cup \mathscr{L}_{\rm i}^{\rm gc} = \mathscr{L} \setminus {\rm int}(\Omega_{\rm c}), \quad \text{and} \quad \mathscr{L}_{\rm a}^{\rm gc} \cap \mathscr{L}_{\rm i}^{\rm gc} = \emptyset.$$

For each $x \in \mathscr{L}_{\rm i}^{\rm gc}$, we define a modified interaction potential $\widetilde{V}(x;\cdot) \in {\rm C}^2((\mathbb{R}^d)^{\mathscr{R}})$, and we define the a/c energy functional as

$$\mathscr{E}_{\rm gc}(y_h) := \varepsilon^d \sum_{x \in \mathscr{L}_{\rm a}^{\rm gc}} V\big(D_{\mathscr{R}} y(x)\big) + \varepsilon^d \sum_{x \in \mathscr{L}_{\rm i}^{\rm gc}} \widetilde{V}\big(x; D_{\mathscr{R}} y(x)\big) + \int_{\Omega_{\rm c}} W(\nabla y_h)\,{\rm d}x. \tag{2.13}$$

For the modified potential $\widetilde{V}$ one takes a general ansatz with several free parameters, which are then fitted to remove or minimize the ghost force. For example, following the ideas of the quasinonlocal coupling method [52] and the geometrically consistent coupling method [18] one may define

$$\widetilde{V}\big(x; \mathbf{g}\big) = V\big((\tilde{g}_r)_{r \in \mathscr{R}}\big), \quad \text{where} \quad \tilde{g}_r = \sum_{s \in \mathscr{R}} C_{x,r,s} g_s.$$

Provided that the interface set $\mathscr{L}_{\rm i}^{\rm gc}$ is chosen appropriately, the parameters $C_{x,r,s}$ can then be determined analytically as in [18,45], or, as proposed in [44], numerically in a preprocessing step. The 2D numerical experiments performed in [44] suggest that it is always possible to determine parameters $C_{x,r,s}$ such that $\mathscr{E}_{\rm ac}$ becomes patch test consistent, however, a proof of this fact is still missing.

The purpose of the present work is to investigate the question whether patch test consistency is in fact a sufficient condition for first-order consistency of an a/c coupling method. If this would turn out to be false in general, then it would be necessary to develop new approaches for constructing accurate a/c methods.

2.3.5. General assumptions on the interface correction

For the subsequent analysis we assume an even more general form of the a/c functional than (2.13). We choose $\mathscr{T}_h^{\rm c}, \mathscr{T}_h^{\rm i}, \mathscr{T}_h^{\rm a} \subset \mathscr{T}_h$, mutually disjoint, such that $\mathscr{T}_h = \mathscr{T}_h^{\rm c} \cup \mathscr{T}_h^{\rm i} \cup \mathscr{T}_h^{\rm a}$, and we define the continuum, interface, and atomistic regions

$$\Omega_{\rm c} := \cup \mathscr{T}_h^{\rm c}, \quad \Omega_{\rm i} := \cup \mathscr{T}_h^{\rm i}, \quad \text{and} \quad \Omega_{\rm a} := \cup \mathscr{T}_h^{\rm a}.$$

(Note that $\Omega_{\rm c}, \Omega_{\rm i}, \Omega_{\rm a}$ are closed sets). Next, we define the set of all nodes $\mathscr{L}_{\rm a} \subset \mathscr{L}$ that interact with the atomistic region:

$$\mathscr{L}_{\rm a} := \Big\{ x \in \mathscr{L} \,\Big|\, (x, x + \varepsilon r) \cap \Omega_{\rm a}^{\#} \neq \emptyset \text{ for some } r \in \mathscr{R} \Big\},$$

where the ordered pair $(x, x') \in \mathscr{L}^{\#} \times \mathscr{L}^{\#}$ is called a *bond*; here, and throughout, the symbol (x, x') is understood as the closed segment ${\rm conv}\{x, x'\}$ whenever it is used as a set. To avoid interaction between $\mathscr{L}_{\rm a}$ and $\Omega_{\rm c}$, we assume throughout that

$$\big\{ x + tr \,\big|\, x \in \mathscr{L}_{\rm a}, t \in [0,1], r \in \mathscr{R} \big\} \subset \Omega_{\rm a} \cup \Omega_{\rm i}. \tag{2.14}$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Next, we define the set of interface bonds

$$\mathscr{B}_{\rm i} := \Big\{ b = (x, x+\varepsilon r) \,\Big|\, x \in \mathscr{L}, r \in \mathscr{R}, (x, x+\varepsilon r) \subset \Omega_{\rm i}^{\#} \Big\}.$$

Finally, we define an interface functional $\mathscr{E}_{\rm i} \in {\rm C}^2(\mathscr{Y})$ such that

$$\mathscr{E}_{\rm i}(y) = \varepsilon^d E_{\rm i}\Big(\big(D_r y(x); (x, x+\varepsilon r) \in \mathscr{B}_{\rm i}\big)\Big), \tag{2.15}$$

that is, the interface functional $\mathscr{E}_{\rm i}$ is given as a function of the finite differences $D_r y(x)$ of bonds $(x, x+\varepsilon r)$ that are contained in the interface region. Note also the volumetric scaling ε^d.

With this notation, we set

$$\mathscr{E}_{\rm ac}(y_h) := \varepsilon^d \sum_{x \in \mathscr{L}_{\rm a}} V\big(D_{\mathscr{R}} y_h(x)\big) + \int_{\Omega_{\rm c}} W(\nabla y_h)\,{\rm d}x + \mathscr{E}_{\rm i}(y_h). \tag{2.16}$$

The interface functional $\mathscr{E}_{\rm i}(y_h)$ specifies the different variants of a/c methods. It is easy to see that the functionals $\mathscr{E}_{\rm qce}$ and $\mathscr{E}_{\rm gc}$, discussed above, fit this framework (in the case of $\mathscr{E}_{\rm qce}$ we have to drop the assumption (2.14)).

If we define the total a/c energy as $\mathscr{E}_{\rm a}^{\rm tot} := \mathscr{E}_{\rm ac} + \mathscr{P}_{\rm ac}$, where $\mathscr{P}_{\rm ac}$ is a suitable a/c approximation to $\mathscr{P}_{\rm a}$, the a/c approximation to (2.5) is

$$y_{\rm ac} \in \operatorname*{argmin}_{y_h \in \mathscr{Y}_{{\rm A},h}} \mathscr{E}_{\rm ac}^{\rm tot}(y_h). \tag{2.17}$$

If $y_{\rm ac}$ solves (2.17), then it is a critical point of $\mathscr{E}_{\rm ac}^{\rm tot} = \mathscr{E}_{\rm ac} + \mathscr{P}_{\rm ac}$:

$$\big\langle \delta\mathscr{E}_{\rm ac}(y_{\rm ac}) + \delta\mathscr{P}_{\rm ac}(y_{\rm ac}), u_h \big\rangle = 0 \qquad \forall u_h \in \mathscr{U}_h. \tag{2.18}$$

2.3.6. The locality and scaling conditions

We define notation for first and second partial derivatives of $E_{\rm i}$ as follows: for $\mathbf{g} = (g_b)_{b \in \mathscr{B}_{\rm i}}$ let

$$\partial_b E_{\rm i}(\mathbf{g}) := \frac{\partial E_{\rm i}(\mathbf{g})}{\partial g_b}, \qquad \text{and} \qquad \partial_b \partial_{b'} E_{\rm i}(\mathbf{g}) := \frac{\partial^2 E_{\rm i}(\mathbf{g})}{\partial g_b \partial g_{b'}}.$$

We extend the definition periodically: if $b \in \mathscr{B}_{\rm i}$ and $\xi \in 2\mathbb{Z}^d$ then $\partial_{\xi+b} E_{\rm i} := \partial_b E_{\rm i}$, and we make a similar definition for the second partial derivatives. In our analysis we will require two crucial properties on $E_{\rm i}$, which we call the *locality and scaling conditions*:

The *locality condition*

$$\partial_{(x,x+\varepsilon r)} \partial_{(x',x'+\varepsilon s)} E_{\rm i}(y) = 0 \qquad \begin{array}{l} \text{for all bonds } (x, x+\varepsilon r), (x', x'+\varepsilon s) \in \mathscr{B}_{\rm i} \\ \text{such that } x \neq x', \end{array} \tag{2.19}$$

implies that the same bonds interact through $\mathscr{E}_{\rm i}$ as in the atomistic model. This condition can be weakened, by requiring that only bonds within an $O(\varepsilon)$ distance interact, however, such a more general condition would add additional notational complexity.

In the *scaling condition* we assume that there exist constants $M_{r,s}^{\rm i} \geq 0$, $r, s \in \mathscr{R}$, such that

$$\big\|\partial_{(x,x+\varepsilon r)} \partial_{(x,x+\varepsilon s)} E_{\rm i}(y)\big\| \leq \begin{cases} M_{r,s}^{\rm i}, & \forall (x, x+\varepsilon r), (x, x+\varepsilon s) \in \mathscr{B}_{\rm i}, \\ \frac{1}{2} M_{r,s}^{\rm i}, & \text{if length}\big(\partial\Omega_{\rm i}^{\#} \cap (x, x+\varepsilon r)\big) > 0. \end{cases} \tag{2.20}$$

This condition effectively yields an $O(1)$ Lipschitz bound for $\delta\mathscr{E}_{\rm i}$ in the function spaces we will use. The scaling aspect enters through an implicit assumption on the magnitude of the constants $M_{r,s}^{\rm i}$, namely, we will assume throughout that the constant

$$M^{\rm i} := \sum_{r \in \mathscr{R}} \sum_{s \in \mathscr{R}} |r||s| M_{r,s}^{\rm i} \tag{2.21}$$

is of the same order of magnitude as the constant $M^{\rm a}$ defined in (2.4).

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Remark 2.7. The factor $\frac{1}{2}$ for bonds on the boundary of the interface region is not strictly necessary, since it can be removed by simply replacing $M^{\mathrm{i}}_{r,s}$ with $2M^{\mathrm{i}}_{r,s}$, however, if stated as above it makes the statements of the results in Section 6 slightly sharper, and moreover simplifies the argument in (6.27).

The necessity of this factor is related to the fact that we allow E_{i} to depend on bonds that lie on the boundary of $\Omega_{\mathrm{i}}^{\#}$; this is made clear in Proposition 6.5 where we construct a stress function for $\mathscr{E}_{\mathrm{ac}}$. Note that if we did not allow E_{i} to depend on these boundary bonds, then it would in fact be impossible to construct patch test consistent a/c methods for non-flat a/c interfaces.

3. A framework for the *a priori* error analysis of a/c methods

When analyzing the error of a numerical method, one should first of all determine the main quantities of interest. For a/c methods, one is usually interested in energy differences between homogeneous lattices and lattices with defects, or critical loads at which defects form or move (*i.e.*, bifurcation points). Since the present paper is mostly theoretical, we will simply focus on the error in the deformation gradient. We note, however, that many aspects of this analysis are crucial ingredients for the analysis of energy differences (see, *e.g.*, [41]) and would usually also enter an analysis of bifurcation points.

We assume from now on that $d \in \{1, 2\}$. To execute the abstract framework of this section also in 3D, several technical tools as well as the central consistency result need to be developed first.

3.1. Discrete and continuous functions

In the following analysis it will be important to extend the a/c functional $\mathscr{E}_{\mathrm{ac}}$ to all functions $y \in \mathscr{Y}$. To that end, we first define piecewise affine interpolants with respect to an atomistic mesh $\mathscr{T}_\varepsilon$.

We take a subdivision of the scaled unit cube $\varepsilon(0,1)^d$ into d-simplices (in 1D the interval $\varepsilon(0,1)$; in 2D two symmetric triangles; compare with the triangulation of the atomistic region in Fig. 1d), which we extend periodically to a triangulation $\mathscr{T}_\varepsilon^{\#}$ of $\mathbb{R}^2$ with vertex set $\mathscr{L}^{\#}$. The restriction of $\mathscr{T}_\varepsilon^{\#}$ to Ω is denoted by $\mathscr{T}_\varepsilon$. Each discrete function $v : \mathscr{L}^{\#} \to \mathbb{R}^k$ will from now on be identified in a canonical way with its continuous piecewise affine interpolant $v \in \mathrm{P}_1(\mathscr{T}_\varepsilon^{\#})^k$.

For future reference we denote the sets of edges of $\mathscr{T}_\varepsilon$ and $\mathscr{T}_\varepsilon^{\#}$, corresponding to the definitions of $\mathscr{F}_h$ and $\mathscr{F}_h^{\#}$ in Section 2.3.1, by $\mathscr{F}_\varepsilon$ and $\mathscr{F}_\varepsilon^{\#}$. Moreover, we define the subsets

$$\mathscr{T}_\varepsilon^{\mathrm{c}} := \big\{T \in \mathscr{T}_\varepsilon \,\big|\, T \subset \Omega_{\mathrm{c}}\big\}, \quad \mathscr{T}_\varepsilon^{\mathrm{i}} := \big\{T \in \mathscr{T}_\varepsilon \,\big|\, T \subset \Omega_{\mathrm{i}}\big\}, \quad \text{and} \quad \mathscr{T}_\varepsilon^{\mathrm{a}} := \big\{T \in \mathscr{T}_\varepsilon \,\big|\, T \subset \Omega_{\mathrm{a}}\big\}.$$

3.1.1. Ambiguity of continuous interpolants

If $y_h \in \mathscr{Y}_h$ then y_h can also be interpreted as a member of $\mathscr{Y}$ and therefore has two, possibly different, continuous interpolants. To distinguish them, we make the convention that the symbol y_h always denotes the interpolant in $\mathrm{P}_1(\mathscr{T}_h)^d$, while a symbol y always denotes the interpolant in $\mathrm{P}_1(\mathscr{T}_\varepsilon)^d$. If we wish to evaluate the $\mathrm{P}_1(\mathscr{T}_\varepsilon)^d$-interpolant of a function $y_h \in \mathscr{Y}_h$ then we will write $I_\varepsilon y_h$.

To compare a $\mathrm{P}_1(\mathscr{T}_\varepsilon)^d$-interpolant with a $\mathrm{P}_1(\mathscr{T}_h)^d$-interpolant, we use the following lemma. In 1D the result is easy to establish; in 2D it depends on a technical tool that we introduce in Section 5.1. The proof is given in the appendix.

Lemma 3.1. *Let $d \in \{1, 2\}$; then, for all $y_h \in \mathscr{Y}_h$ and $p \in [1, \infty]$, we have*

$$\|\nabla I_\varepsilon y_h\|_{\mathrm{L}^p(\Omega)} \leq \|\nabla y_h\|_{\mathrm{L}^p(\Omega)},$$

where we recall that we have defined $\|\nabla v\|_{\mathrm{L}^p} = \| \, |\nabla v|_p \|_{\mathrm{L}^p}$.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

3.1.2. Extension of the a/c energy

Before we extend the a/c energy $\mathscr{E}_{\rm ac}$ we make one last technical assumption, which considerably simplifies the subsequent analysis. We shall assume from now on, that

$$\mathscr{T}_h^{\rm a} \cup \mathscr{T}_h^{\rm i} \subset \mathscr{T}_\varepsilon.$$

If all atoms in $\mathscr{L} \cap (\Omega_{\rm a} \cup \Omega_{\rm i})$ are vertices of $\mathscr{T}_h^{\rm a} \cup \mathscr{T}_h^{\rm i}$, which is not uncommon, then this is no restriction.

With these conventions the a/c energy $\mathscr{E}_{\rm ac}$ defined in (2.16) can be defined canonically for functions $\tilde{y} = y + u_h \in \mathscr{Y} + \mathscr{U}_h$ by the same formula:

$$\mathscr{E}_{\rm ac}(\tilde{y}) = \int_{\Omega_{\rm c}} W(\nabla \tilde{y})\,{\rm d}x + \varepsilon^d \sum_{x \in \mathscr{L}_{\rm a}} V\big(D_{\mathscr{R}}\tilde{y}(x)\big) + \mathscr{E}_{\rm i}(\tilde{y}) \qquad \text{for } \tilde{y} \in \mathscr{Y} + \mathscr{U}_h.$$

We also assume that $\mathscr{P}_{\rm ac}$ has a suitable extension to $\mathscr{Y}$. It should be stressed that for general $y_h \in \mathscr{Y}_h$, $\mathscr{E}_{\rm ac}(y_h) \neq \mathscr{E}_{\rm ac}(I_\varepsilon y_h)$.

3.2. Measuring smoothness; an interpolation error estimate

The three main ingredients in the *a priori* error analysis of Galerkin-like approximations are (i) consistency, (ii) stability, and (iii) an interpolation error estimate. We begin by establishing the latter. To that end, we first need to find a convenient *measure of smoothness* for discrete functions $y \in \mathscr{Y}$.

3.2.1. Measuring smoothness in terms of local oscillation

There are several possibilities to measure the "smoothness" of a discrete function. The most obvious is possibly the use of higher order finite differences, *e.g.*, $D_{e_i} D_{e_j} y(x)$. If, in Section 3.1, we had chosen continuous interpolants belonging to $\mathrm{W}^{2,\infty}$, then we would be able to simply evaluate the second derivatives $\nabla^2 y$. However, since the interpolants we use are piecewise affine, the second derivative of y is the measure $[\![\nabla y]\!] \otimes \nu\,{\rm d}s\big|_{\mathscr{F}_\varepsilon^\#}$, where $[\![\nabla y]\!]$ denotes the jump of ∇y across an element edge, and ${\rm d}s$ the surface measure.

This last observation motivates the idea to measure smoothness of y by the local oscillation of ∇y. We define the *oscillation operator*, for measurable sets $\omega \subset \mathbb{R}^d$, and for $y \in \mathscr{Y}$, as

$$\operatorname{osc}(\nabla y; \omega) := \operatorname*{ess\,sup}_{x,x' \in \omega} \frac{\big|\nabla y(x) - \nabla y(x')\big|}{\varepsilon}. \tag{3.1}$$

The sets ω that arise naturally in our analysis will always have $O(\varepsilon)$ diameter, which is the reason for the ε^{-1}-scaling in the definition of osc.

Note, in particular, that if y were twice differentiable, and if $\operatorname{diam}(\omega) \leq C\varepsilon$, then we would obtain

$$\operatorname{osc}(\nabla y; \omega) \leq \frac{\operatorname{diam}(\omega)}{\varepsilon}\, \|\nabla^2 y\|_{\mathrm{L}^\infty(\omega)} \leq C \|\nabla^2 y\|_{\mathrm{L}^\infty(\omega)},$$

which further illustrates that the oscillation operator is a reasonable replacement for $\nabla^2 y$ to measure the local smoothness of a piecewise affine function.

3.2.2. Interpolation error estimate

The smoothness measure we defined in Section 3.2.1 yields a simple proof of an interpolation error estimate; see Appendix A.

Lemma 3.2. *Let $d \in \{1, 2\}$ and suppose that $\mathscr{T}_h^{\rm a} \cup \mathscr{T}_h^{\rm i} \subset \mathscr{T}_\varepsilon$; then there exists a constant C_I, which depends only on the shape regularity of $\mathscr{T}_h$, such that, for all $y \in \mathscr{Y}$, $p \in [1, \infty)$,*

$$\big\|\nabla(y - I_h y)\big\|_{\mathrm{L}^p(\Omega)} \leq C_I \left\{ \sum_{T \in \mathscr{T}_\varepsilon^{\rm c}} |T| \left[\mathrm{h}_T \operatorname{osc}(\nabla y; \omega_T^{\rm c})\right]^p \right\}^{1/p},$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

where, for $T \in \mathscr{T}_\varepsilon$,

$$\omega_T^{\rm c} := \Omega_{\rm c}^{\#} \cap \bigcup \big\{ T' \in \mathscr{T}_\varepsilon^{\#} \,\big|\, T \cap T' \neq \emptyset \big\}, \quad \textit{and} \quad {\rm h}_T := \max_{x \in T} |h(x)|. \tag{3.2}$$

Similarly, for $p = \infty$, *we have* $\big\|\nabla(y - I_h y)\big\|_{{\rm L}^\infty(\Omega)} \leq C_I \max_{T \in \mathscr{T}_\varepsilon^{\rm c}} \big[{\rm h}_T \, {\rm osc}(\nabla y; \omega_T^{\rm c}) \big]$.

3.3. The stability assumption

The stability of a/c methods relies, firstly, on the stability of atomistic models as well as their Cauchy–Born approximations. It requires a thorough understanding of the physics of a model and in particular more specific information about the interaction potential. Since the focus of the present work is the consistency of a/c methods, we will formulate stability as an *assumption*.

The simplest notion of stability one may use, which is also closely connected to local minimality, is coercivity of the second variation:

$$\langle \delta^2 \mathscr{E}_{\rm ac}^{\rm tot}(y_h) u_h, u_h \rangle \geq c_0 \|\nabla u_h\|_{{\rm L}^2(\Omega)}^2 \qquad \forall\, u_h \in \mathscr{U}_h, \tag{3.3}$$

where $c_0 > 0$, and $y_h \in \mathscr{Y}_h$ is a suitable deformation in a neighbourhood of the atomistic solution $y_{\rm a}$, *e.g.*, $y_h = I_h y_{\rm a}$. The choice of norm is motivated by the fact that the Cauchy–Born model, and hence the atomistic model, are closely related to second order elliptic differential equations.

Examples of sharp stability estimates for a/c methods in 1D can be found in [14, 27, 28, 40, 54]. For pair interactions in 2D the stability of Shapeev's method [49] is established in [41].

More generally, for some $p \in [1, \infty], p' = p/(p-1)$, we may assume an inf-sup condition of the form

$$\inf_{\substack{u_h \in \mathscr{U}_h \\ \|\nabla u_h\|_{{\rm L}^p(\Omega)} = 1}} \sup_{\substack{v_h \in \mathscr{U}_h \\ \|\nabla v_h\|_{{\rm L}^{p'}(\Omega)} = 1}} \big\langle \delta^2 \mathscr{E}_{\rm ac}^{\rm tot}(y_h) u_h, v_h \big\rangle \geq c_0, \tag{3.4}$$

for some constant $c_0 > 0$. The condition (3.4) is usually difficult to prove, especially for $p \neq 2$, and may indeed be false in general. We will only use it to demonstrate how such a stability result motivates consistency estimates in different negative norms. Examples of 1D inf-sup stability estimates for a/c methods can be found in [10, 15, 33, 42].

3.4. Outline of an *a priori* error analysis

The following outline of an *a priori* error analysis depends on a stability assumption that we will not prove. Moreover, since it primarily serves to motivate the *consistency problem*, and since a rigorous derivation would be more involved without yielding much additional insight, some steps will be kept vague. Most of these steps are easily made rigorous; the main assumption we make below, which is in fact very difficult to justify rigorously, is that $I_h y_{\rm a}$ and $y_{\rm ac}$ are "sufficiently close". See [33, 40, 42, 54] for similar analyses in 1D where all steps are rigorously justified, and [41, 43] for a similar semi-rigorous framework, where a proof of this step is replaced by an assumption.

Let $y_{\rm a}$ satisfy (2.6), $y_{\rm ac}$ satisfy (2.18), and suppose that the stability assumption (3.4) holds with $y_h = I_h y_{\rm a}$. Let $e_h := y_{\rm ac} - I_h y_{\rm a}$. Moreover, suppose that $\|\nabla I_h y_{\rm a} - \nabla y_{\rm ac}\|_{{\rm L}^\infty}$ is sufficiently small so that the following approximation can be made precise:

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

Taking the supremum over all $v_h \in \mathscr{U}_h$, and invoking the inf-sup condition (3.4) and the criticality condition (2.18), we obtain

$$c_0 \big\|\nabla e_h\big\|_{{\rm L}^p(\Omega)} \lesssim \big\|\delta(\mathscr{E}_{\rm ac} + \mathscr{P}_{\rm ac})(I_h y_{\rm a})\big\|_{{\rm W}_h^{-1,p}},$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

where we define

$$\|\Phi\|_{\mathrm{W}_h^{-1,p}} := \sup_{\substack{v_h \in \mathscr{U}_h \\ \|\nabla v_h\|_{\mathrm{L}^{p'}}=1}} \big\langle \Phi, v_h \big\rangle, \qquad \text{for } \Phi \in \mathscr{U}_h^*.$$

We split the consistency error $\|\delta(\mathscr{E}_{\rm ac} + \mathscr{P}_{\rm ac})(I_h y_{\rm a})\|_{\mathrm{W}_h^{-1,p}}$ into three separate contributions:

$$\begin{aligned} c_0 \big\|\nabla e_h\big\|_{\mathrm{L}^p(\Omega)} &\lesssim \big\|\delta(\mathscr{E}_{\rm ac} + \mathscr{P}_{\rm ac})(I_h y_{\rm a}) - \delta(\mathscr{E}_{\rm ac} + \mathscr{P}_{\rm ac})(y_{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} \\ &\quad + \big\|\delta\mathscr{E}_{\rm ac}(y_{\rm a}) - \delta\mathscr{E}_{\rm a}(y_{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} + \big\|\delta\mathscr{P}_{\rm ac}(y_{\rm a}) - \delta\mathscr{P}_{\rm a}(y_{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} \\ &=: \mathcal{E}_h^{\rm coarse} + \mathcal{E}_h^{\rm model} + \mathcal{E}_h^{\rm ext}. \end{aligned} \tag{3.5}$$

where we have used (2.6), and the extension of $\mathscr{E}_{\rm ac}$ and $\mathscr{P}_{\rm ac}$ for all deformations $y \in \mathscr{Y}$ constructed in Section 3.1.

The *coarsening error*, $\mathcal{E}_h^{\rm coarse}$, can be bounded by Lipschitz estimates for $\delta(\mathscr{E}_{\rm ac} + \mathscr{P}_{\rm ac})$ and an interpolation error estimate. Using our assumption that $\mathscr{T}_h^{\rm a} \cup \mathscr{T}_h^{\rm i} \subset \mathscr{T}_\varepsilon$ it is not difficult to derive Lipschitz estimates of the form

$$\mathcal{E}_h^{\rm coarse} \leq (M^{\rm a} + M_{\mathscr{P}_{\rm ac}})\big\|\nabla I_h y_{\rm a} - \nabla y_{\rm a}\big\|_{\mathrm{L}^p(\Omega)}, \tag{3.6}$$

where $M^{\rm a}$ is a Lipschitz constant for ∂W (*cf.* (2.10)), and $M_{\mathscr{P}_{\rm ac}}$ is a Lipschitz constant for $\delta\mathscr{P}_{\rm ac}$.

Combining (3.5), and (3.6), the inequality

$$\big\|\nabla y_{\rm a} - \nabla y_{\rm ac}\big\|_{\mathrm{L}^p} \leq \big\|\nabla y_{\rm a} - \nabla I_h y_{\rm a}\big\|_{\mathrm{L}^p} + \big\|\nabla e_h\big\|_{\mathrm{L}^p},$$

and the interpolation error estimate of Lemma 3.2, we arrive at the following basic error estimate

$$\big\|\nabla y_{\rm a} - \nabla y_{\rm ac}\big\|_{\mathrm{L}^p(\Omega)} \leq \frac{\mathcal{E}_h^{\rm model} + \mathcal{E}_h^{\rm ext}}{c_0} + \frac{c_1}{c_0}\left\{\sum_{T \in \mathscr{T}_\varepsilon^{\rm c}} [\mathrm{h}_T \operatorname{osc}(\nabla y_{\rm a}; \omega_T^{\rm c})]^p\right\}^{1/p}, \tag{3.7}$$

where $c_1 = C_I(c_0 + M^{\rm a} + M_{\mathscr{P}_{\rm ac}})$.

The *consistency error for the external forces*, $\mathcal{E}_h^{\rm ext}$, depends on the form of $\mathscr{P}_{\rm a}$ and $\mathscr{P}_{\rm ac}$ and cannot be discussed at this level of abstraction. The *modelling error*, $\mathcal{E}_h^{\rm model}$, is the focus of the remainder of the present paper.

Remark 3.3 (Choice of splitting)**.** Suppose, for simpliciy, that $\mathscr{P}_{\rm ac} = \mathscr{P}_{\rm a} = 0$. In a typical finite element error analysis of continuum mechanics problems one would usually choose a different splitting of the consistency error:

$$\big\|\delta\mathscr{E}_{\rm ac}(I_h y_{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} \leq \big\|\delta\mathscr{E}_{\rm ac}(I_h y_{\rm a}) - \delta\mathscr{E}_{\rm a}(I_h y_{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} + \big\|\delta\mathscr{E}_{\rm a}(I_h y_{\rm a}) - \delta\mathscr{E}_{\rm a}(y_{\rm a})\big\|_{\mathrm{W}_h^{-1,p}}.$$

This splitting was used in the analysis in [41] and led to a suboptimal estimate of the modelling error, since it still contains some coarsening error.

3.5. The modelling error

The main step that remains in obtaining an *a priori* error estimate from (3.7) is the estimation of the *modelling error*

$$\mathcal{E}_h^{\rm model} = \big\|\delta\mathscr{E}_{\rm ac}(y_{\rm a}) - \delta\mathscr{E}_{\rm a}(y_{\rm a})\big\|_{\mathrm{W}_h^{-1,p}} = \sup_{v_h \in \mathscr{U}_h \setminus \{0\}} \frac{\big\langle \delta\mathscr{E}_{\rm ac}(y_{\rm a}) - \delta\mathscr{E}_{\rm a}(y_{\rm a}), v_h \big\rangle}{\|\nabla v_h\|_{\mathrm{L}^{p'}(\Omega)}}.$$

Most of the numerical analysis literature on a/c methods estimates this modelling error only for the case when $\mathscr{T}_h = \mathscr{T}_\varepsilon$. In 1D it is easy to see that this is sufficient, since $I_\varepsilon v_h = v_h$ in that case; see also [43]. The following lemma provides the main technical step to explain why it is also sufficient in 2D to consider the case $\mathscr{T}_h = \mathscr{T}_\varepsilon$. Its proof uses arguments similar to those in the a posteriori error analysis of continuum finite element methods and is given in Appendix A.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Lemma 3.4. *Assume that* $\mathscr{T}_h^{\rm i} \cup \mathscr{T}_h^{\rm a} \subset \mathscr{T}_\varepsilon$. *Let* $\Phi \in \mathscr{U}^*$ *and* $\Phi_h \in \mathscr{U}_h^*$ *be given in the form*

$$\langle \Phi, u\rangle = \int_{\Omega_{\rm c}} \sigma : \nabla u \,{\rm d}x, \quad \text{and} \quad \langle \Phi_h, u_h\rangle = \int_{\Omega_{\rm c}} \sigma : \nabla u_h \,{\rm d}x,$$

for all $u \in \mathscr{U}$, $u_h \in \mathscr{U}_h$, *where* $\sigma \in {\rm P}_0^{\#}(\mathscr{T}_\varepsilon)^{d\times d}$; *then there exists a universal constant* C_M *such that, for all* $p \in [1,\infty)$,

$$\begin{aligned}
\left|\langle \Phi, I_\varepsilon u_h\rangle - \langle \Phi_h, u_h\rangle\right| &\le C_M \varepsilon \left(\sum_{T\in\mathscr{T}_\varepsilon^{\rm c}} |T|{\rm osc}(\sigma;\omega_T^{\rm c})^p\right)^{1/p} \big\|\nabla u_h\big\|_{{\rm L}^{p'}(\Omega_{\rm c})}, \quad \text{and} \\
\left|\langle \Phi, I_\varepsilon u_h\rangle - \langle \Phi_h, u_h\rangle\right| &\le C_M \varepsilon \left[\max_{T\in\mathscr{T}_\varepsilon^{\rm c}} {\rm osc}(\sigma;\omega_T^{\rm c})\right] \big\|\nabla u_h\big\|_{{\rm L}^1(\Omega_{\rm c})}.
\end{aligned} \tag{3.8}$$

The estimate (3.8), together with Lemma 3.1, implies the following theorem, where we use the notation

$$\|\Phi\|_{{\rm W}_\varepsilon^{-1,p}} := \sup_{\substack{v\in\mathscr{U}\setminus\{0\}\\ \|\nabla v\|_{{\rm L}^{p'}}=1}} \langle \Phi, v\rangle \qquad \text{for } \Phi \in \mathscr{U}^*. \tag{3.9}$$

Theorem 3.5. *Suppose that* $\mathscr{T}_h^{\rm a} \cup \mathscr{T}_h^{\rm i} \subset \mathscr{T}_\varepsilon$ *and that* $y \in \mathscr{Y}$; *then, for all* $p \in [1,\infty)$,

$$\big\|\delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y)\big\|_{{\rm W}_h^{-1,p}} \le M^{\rm a} C_M \varepsilon \left(\sum_{T\in\mathscr{T}_\varepsilon^{\rm c}} |T|{\rm osc}(\nabla y(T);\omega_T^{\rm c})^p\right)^{1/p} + \big\|\delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y)\big\|_{{\rm W}_\varepsilon^{-1,p}}, \tag{3.10}$$

with corresponding statement for $p = \infty$.

Proof. Since $\mathscr{E}_{\rm a}(I_\varepsilon y_h)$ uses only point values of $I_\varepsilon y_h$, which are the same as for y_h, we have

$$\langle \delta\mathscr{E}_{\rm a}(y), u_h\rangle = \langle \delta\mathscr{E}_{\rm a}(y), I_\varepsilon u_h\rangle \qquad \forall\, u_h \in \mathscr{U}_h.$$

Using this fact, we can estimate

$$\left|\langle \delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y), u_h\rangle\right| \le \left|\langle \delta\mathscr{E}_{\rm ac}(y), u_h\rangle - \langle \delta\mathscr{E}_{\rm ac}(y), I_\varepsilon u_h\rangle\right| + \left|\langle \delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y), I_\varepsilon u_h\rangle\right|.$$

Due to the assumption that $\mathscr{T}_h^{\rm a}\cup\mathscr{T}_h^{\rm c} \subset \mathscr{T}_\varepsilon$, the first group can be estimated using Lemma 3.4, with $\sigma = \partial W(\nabla y)$, which yields the first term in (3.10).

Using Lemma 3.1, the second group can be estimated by

$$\begin{aligned}
\left|\langle \delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y), I_\varepsilon u_h\rangle\right| &\le \big\|\delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y)\big\|_{{\rm W}_\varepsilon^{-1,p}} \big\|\nabla I_\varepsilon u_h\big\|_{{\rm L}^{p'}} \\
&\le \big\|\delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y)\big\|_{{\rm W}_\varepsilon^{-1,p}} \big\|\nabla u_h\big\|_{{\rm L}^{p'}}.
\end{aligned}$$

Taking the supremum over all $u_h \in \mathscr{U}_h$ with $\|\nabla u_h\|_{{\rm L}^{p'}} = 1$ yields the stated result. □

Applying Theorem 3.5 to the modelling error $\mathcal{E}_h^{\rm model}$, defined in (3.5), we obtain that

$$\mathcal{E}_h^{\rm model} \le \mathcal{E}_\varepsilon^{\rm model} + M^{\rm a} C_M \varepsilon \left(\sum_{T\in\mathscr{T}_\varepsilon^{\rm c}} |T|{\rm osc}(\nabla y(T);\omega_T^{\rm c})^p\right)^{1/p}, \tag{3.11}$$

where

$$\mathcal{E}_\varepsilon^{\rm model} := \big\|\delta\mathscr{E}_{\rm ac}(y_{\rm a}) - \delta\mathscr{E}_{\rm a}(y_{\rm a})\big\|_{{\rm W}_\varepsilon^{-1,p}}.$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Even though $\mathcal{E}_\varepsilon^{\rm model}$ is essentially an upper bound for $\mathcal{E}_h^{\rm model}$, it is usually easier to estimate. The *consistency problem* is to prove a sharp upper bound on $\mathcal{E}_\varepsilon^{\rm model}$.

In Section 4 we will discuss two simple 1D examples to determine what can be expected in more general situations. In Theorem 6.1 we will prove that for an a/c method that is patch test consistent, and satisfies various other technical conditions, one obtains

$$\mathcal{E}_\varepsilon^{\rm model} \leq C\,\varepsilon \left\{ \sum_{T \in \mathscr{T}_\varepsilon^{\rm c} \cup \mathscr{T}_\varepsilon^{\rm i}} |T| \big[{\rm osc}(\nabla y_{\rm a}; \omega_T) \big]^p \right\}^{1/p},$$

where C is a constant that is independent of $y_{\rm a}$, but does depend on the interface width, and $\omega_T \subset \Omega_{\rm c} \cup \Omega_{\rm i}$ is the interaction neighbourhood defined in (6.2).

Combined with (3.11) and (3.7), and using the fact that $\varepsilon \leq {\rm h}_T$, and that $\omega_T \supset \omega_T^{\rm c}$ for $T \in \mathscr{T}_\varepsilon^{\rm c}$, this bound yields

$$\big\| \nabla y_{\rm a} - \nabla y_{\rm ac} \big\|_{{\rm L}^p} \leq \frac{\mathcal{E}_h^{\rm ext}}{c_0} + \frac{c_2}{c_0} \left\{ \sum_{T \in \mathscr{T}_\varepsilon^{\rm c} \cup \mathscr{T}_\varepsilon^{\rm i}} |T| \Big[{\rm h}_T\, {\rm osc}(\nabla y_{\rm a}; \omega_T) \Big]^p \right\}^{1/p}, \tag{3.12}$$

where c_2 is a constant that is independent of $y_{\rm a}$. This estimate closely resembles a typical first-order *a priori* error estimate for a continuum mechanics finite element approximation; see also the interpretation given in Section 1.2.

It should be stressed again that (3.12) is not a rigorous error estimate, but depends on various assumptions made in the forgoing discussion, most prominently, the stability assumption (3.4), and the assumption that $\|\nabla I_h y_{\rm a} - \nabla y_{\rm ac}\|_{{\rm L}^\infty}$ is "sufficiently small".

Remark 3.6. The locality of the patches ω_T is crucial. If ${\rm diam}(\omega_T)$ is not of the order $O(\varepsilon)$, then it is possible that ${\rm osc}(\nabla y_{\rm a}; \omega_T) \gg 1$ even if $y_{\rm a}$ is globally smooth; see also Section 6.4.4.

4. Examples in 1D

In the present section we review the consistency analyses of specific a/c methods to point out the main features and to motivate what may be expected in the general case. Throughout this section we assume that $d = 1$, $\mathscr{R} = \{\pm 1, \pm 2\}$, and that V is given by

$$V(\{g_{\pm 1}, g_{\pm 2}\}) = \frac{1}{2}\big[\phi_1(g_1) + \phi_1(g_{-1}) + \phi_2(g_2) + \phi_2(g_{-2})\big],$$

where $\phi_1, \phi_2 \in {\rm C}^{2,1}(\mathbb{R})$ are, respectively, the first and second neigbhour interaction potentials, which are assumed to be symmetric about the origin. We assume that their derivatives ϕ_i' and ϕ_i'' have global Lipschitz constants m_i' and m_i''.

For the 1D analysis it is convenient to write $x_n = n\varepsilon$, $v_n = v(x_n)$, and to write all interactions in terms of the backward difference operator

$$v_n' = \frac{v_n - v_{n-1}}{\varepsilon}.$$

With this notation the atomistic energy can now be rewritten in the form

$$\mathscr{E}_{\rm a}(y) = \varepsilon \sum_{n=-N+1}^{N} \phi_1(y_n') + \varepsilon \sum_{n=-N+1}^{N} \phi_2(y_n' + y_{n+1}'), \tag{4.1}$$

where we note that $y_n' + y_{n+1}' = \varepsilon^{-1}(y_{n+1} - y_{n-1})$ describes a second neighbour bond.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

For future reference we also define the second and third finite differences

$$v_n'' = \frac{v_{n+1}' - v_n'}{\varepsilon}, \quad \text{and} \quad v_n'''(x) = \frac{v_{n+1}' - 2v_n' + v_{n-1}'}{\varepsilon^2}.$$

It is also worth pointing out that $v_n' = \nabla v(s)$ for all $s \in (x_{n-1}, x_n)$.

4.1. Consistency of the QNL method

We begin with a modelling error analysis for the *quasinonlocal coupling method* (QNL method) of Shimokawa *et al.* [52]. Some variants of the geometrically consistent coupling scheme [18] and Shapeev's method [49] reduce to the same method for 1D second neighbour interactions.

The following presentation follows largely [40], where the QNL method is defined as follows: let $\mathcal{N}_{\rm a} = \{-K, \dots, K\}$ for some $K \geq 1$ and $\mathcal{N}_{\rm c} = \{-N+1, \dots, N\} \setminus \mathcal{N}_{\rm a}$; then, for $y \in \mathcal{Y}$, the QNL energy is defined by

$$\mathcal{E}_{\rm qnl}(y) = \varepsilon \sum_{n=-N+1}^{N} \phi_1(y_n') + \varepsilon \sum_{n \in \mathcal{N}_{\rm a}} \phi_2(y_n' + y_{n+1}') + \varepsilon \sum_{n \in \mathcal{N}_{\rm c}} \frac{1}{2}\big[\phi_2(2y_n') + \phi_2(2y_{n+1}')\big]. \tag{4.2}$$

We observe that we have not modified the first neighbour interactions, but have "split" the non-local second neighbour interactions into local first neighbour interactions in the continuum region.

It is straightforward to rewrite $\mathcal{E}_{\rm qnl}$ in the form specified in (2.16), with

$$\Omega_{\rm c} = [\varepsilon(K+1), 1] \cup (-1, \varepsilon(-K-1)], \quad \mathscr{L}_{\rm a} = \varepsilon\{-K+1, \dots, K-1\},$$

and a suitably defined interface functional $\mathcal{E}_{\rm i}$; however, the form (4.2) is more convenient for the analysis.

The following modelling error estimate was established in [40], Theorem 3.1. Dobson and Luskin [12] treated a quadratic interaction case, using entirely different analytical techniques that gave an even more detailed analysis of the error; Ming and Yang [37] used related methods as [40], Theorem 3.1, but gave a qualitatively less precise estimate of the consistency error. An extension of the result to linear finite range pair interactions is given in [28].

Note also that it is shown in [12] that the consistency error of the QNL method in ℓ^p_ε-norms is of the order $O(1)$, that is, the usage of negative norms cannot be avoided.

We will discuss the estimate in detail in Section 4.3. The proof of the following result, which is included for convenience, serves as a first guidance on how one may approach proofs of consistency of a/c methods in more general situations.

Proposition 4.1 (Consistency of the QNL method)**.** *Let $y \in \mathcal{Y}$; then*

$$\big\|\delta\mathcal{E}_{\rm qnl}(y) - \delta\mathcal{E}_{\rm a}(y)\big\|_{\mathrm{W}_\varepsilon^{-1,p}} \leq \varepsilon m_2' \big\|y''\big\|_{\ell^p_\varepsilon(\{-K,K\})} + \varepsilon^2 m_2' \big\|y'''\big\|_{\ell^p_\varepsilon(\mathcal{N}_{\rm c}')} + \varepsilon^2 m_2'' \big\|y''\big\|^2_{\ell^{2p}_\varepsilon(\mathcal{N}_{\rm c})},$$

where $\mathcal{N}_{\rm c}' = \{-N+1, \dots, -K-1\} \cup \{K+2, \dots, N\}$,

Proof. Throughout the proof we will make use of the fact that the boundary conditions are periodic without comment, treating the boundary as if it belonged to the "interior" of the continuum region.

Since the first neighbour interactions as well as the second neighbour interactions in the atomistic region are treated identically in the atomistic model and the QNL method, we have

$$\big\langle \delta\mathcal{E}_{\rm a}(y) - \delta\mathcal{E}_{\rm qnl}(y), u\big\rangle = \varepsilon \sum_{n \in \mathcal{N}_{\rm c}} \big[\phi_2'(y_n' + y_{n+1}') \cdot (u_n' + u_{n+1}') - \phi_2'(2y_n') \cdot u_n' - \phi_2'(2y_{n+1}') \cdot u_{n+1}'\big].$$

Rearranging the sum in terms of the gradients u_n', and using $\mathcal{N}_{\rm a} = \{-K, \dots, K\}$, yields

$$\big\langle \delta\mathcal{E}_{\rm a}(y) - \delta\mathcal{E}_{\rm qnl}(y), u\big\rangle = \varepsilon \sum_{n=-N+1}^{N} \mathsf{R}_n \cdot u_n', \tag{4.3}$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

where $(\mathsf{R}_n)_{n=1}^N$ is defined as follows:

$$\mathsf{R}_n = \begin{cases} 0, & n \in \{-K+1, \dots, K\}, \\ \phi_2'(y_{n-1}' + y_n') - \phi_2'(2y_n'), & n = K+1, \\ \phi_2'(y_n' + y_{n+1}') - \phi_2'(2y_n'), & n = -K, \\ \phi_2'(y_n' + y_{n+1}') - 2\phi_2'(2y_n') + \phi_2'(y_{n-1}' + y_n'), & n \in \mathscr{N}_{\rm c}'. \end{cases}$$

At the interface, $n = K+1$, we have

$$\mathsf{R}_n \le m_2' |y_{n+1}' - y_n'| = \varepsilon m_2' |y_{n-1}''|,$$

with a similar estimate for $n = -K$. In the continuum region, $n \ge K+2$, or $n \le -K+1$, the terms R_n have second order structure, and a second order Taylor expansion yields

$$\begin{aligned} |\mathsf{R}_n| &\le |\phi_2''(2y_n')| |y_{n+1}' - 2y_n' + y_{n-1}'| + \frac{1}{2} m_2'' \Big[|y_{n+1}' - y_n'|^2 + |y_n' - y_{n_1}'|^2 \Big] \\ &= \varepsilon^2 m_2' |y_n'''| + \varepsilon^2 \frac{1}{2} m_2'' \Big[|y_n''|^2 + |y_{n-1}''|^2 \Big]. \end{aligned}$$

After inserting the two bounds into (4.3), and applying several weighted Hölder inequalities one obtains the stated estimate. □

4.2. Inconsistency of the QCE method

As in the previous section, let $\mathscr{N}_{\rm a} = \{-K, \dots, K\}$, $\mathscr{N}_{\rm c} = \{-N+1, \dots, N\} \setminus \mathscr{N}_{\rm a}$, and $\mathscr{L}_{\rm a} = \varepsilon \mathscr{N}_{\rm a}$; then, under the assumptions and notation set out at the beginning of Section 4, the QCE energy functional defined in (2.9) reads

$$\begin{aligned} \mathscr{E}_{\rm qce}(y) = \varepsilon \sum_{n \in \mathscr{N}_{\rm a}} \frac{1}{2} \Big[\phi_1(y_n') + \phi_1(y_{n+1}') + \phi_2(y_{n-1}' + y_n') + \phi_2(y_{n+1}' + y_{n+2}') \Big] \\ + \varepsilon \sum_{n \in \mathscr{N}_{\rm c}} \frac{1}{2} \Big[\phi_1(y_n') + \phi_1(y_{n+1}') + \phi_2(2y_n') + \phi_2(2y_{n+1}') \Big], \end{aligned} \tag{4.4}$$

noting that $W(\mathsf{F}) = \phi_1(\mathsf{F}) + \phi_2(2\mathsf{F})$.

The following result is a variant of [43], Theorem 3.2. Previous analyses of the QCE method [11, 12, 37] computed the consistency error contribution due to the "ghost forces" explicitly rather than estimating their $\mathrm{W}_\varepsilon^{-1,p}$-residual contribution.

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

$$\big\| \delta\mathscr{E}_{\rm qce}(y) - \delta\mathscr{E}_{\rm a}(y) \big\|_{\mathrm{W}_\varepsilon^{-1,p}} \le \varepsilon^{1/p} G + \varepsilon m_2' \|y''\|_{\ell^p_\varepsilon(\mathscr{N}_{\rm i})} + \varepsilon^2 m_2' \|y'''\|_{\ell^p_\varepsilon(\mathscr{N}_{\rm c}')} + \varepsilon^2 m_2'' \|y''\|^{2p}_{\ell^p_\varepsilon(\mathscr{N}_{\rm c} \cup \{-K, K\})},$$

where $\mathscr{N}_{\rm c}'$ is defined as in Proposition 4.1, $\mathscr{N}_{\rm i} = \{-K-1, -K, K, K+1\}$, and

$$G = \frac{1}{2} \big[|\phi_2'(2y_{-K-1}')| + |\phi_2'(2y_{-K+1}')| + |\phi_2'(2y_K')| + |\phi_2'(2y_{K+2}')| \big].$$

The estimate is sharp in the sense that, for some constant C, $\frac{1}{2} \le C \le 2$,

$$\big\| \delta\mathscr{E}_{\rm qce}(y_{\mathsf{A}}) \big\|_{\mathrm{W}_\varepsilon^{-1,p}} = \big\| \delta\mathscr{E}_{\rm qce}(y_{\mathsf{A}}) - \delta\mathscr{E}_{\rm a}(y_{\mathsf{A}}) \big\|_{\mathrm{W}_\varepsilon^{-1,p}} \ge C \varepsilon^{1/p} |\phi_2'(2\mathsf{A})| \qquad \forall \mathsf{A} \in \mathbb{R}.$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Proof. The first result can be proven in much the same way as Proposition 4.1, by rewriting the first variation $\delta\mathscr{E}_{\rm qce}$ in the form $\langle \delta\mathscr{E}_{\rm qce}(y), u\rangle = \varepsilon \sum_{n=-N+1}^{N} \mathsf{R}_n \cdot u'_n$, and carefully estimating the coefficients R_n. See [43] for the details of this computation.

To obtain the opposite estimate for $y = y_{\mathsf{A}}$, a brief computation gives

$$\big\langle \delta\mathscr{E}_{\rm qce}(y_{\mathsf{A}}) - \delta\mathscr{E}_{\rm a}(y_{\mathsf{A}}), u\big\rangle = \varepsilon\frac{1}{2}\phi'_2(2\mathsf{A})\big[u'_{-K-1} - u'_{-K+1} - u'_K + u'_{K+2}\big].$$

If we choose $u \in \mathscr{U}$ such that

$$u'_n = \operatorname{sign}\big(\phi'_2(2\mathsf{A})\big) \cdot \begin{cases} (4\varepsilon)^{-1/p'}, & n = -K-1, \quad K+2, \\ -(4\varepsilon)^{-1/p'}, & n = -K+1, \quad K, \\ 0, & \text{otherwise}, \end{cases}$$

then $\|\nabla u\|_{\mathrm{L}^{p'}(-1,1)} = \|u'\|_{\ell^{p'}_\varepsilon} = 1$, and we obtain that

$$\big\|\delta\mathscr{E}_{\rm qce}(y_{\mathsf{A}}) - \delta\mathscr{E}_{\rm a}(y_{\mathsf{A}})\big\|_{\mathrm{W}^{-1,p}_\varepsilon} \geq \big\langle \delta\mathscr{E}_{\rm qce}(y_{\mathsf{A}}) - \delta\mathscr{E}_{\rm a}(y_{\mathsf{A}}), u\big\rangle = \varepsilon^{1/p}\big[2\cdot 4^{-1/p'}\big|\phi'_2(2\mathsf{A})\big|\big]. \qquad \square$$

4.3. Discussion

This discussion of the 1D consistency error estimates largely follows the discussions in [40, 43].

We have estimated the modelling errors for two prototypical a/c methods. We see that the leading order terms in the upper bounds are $O(\varepsilon)$ and $O(\varepsilon^{1/p})$ for the QNL and QCE methods, respectively. However, a much finer distinction should be made.

First, we note that both methods reduce to the Cauchy–Born approximation in the continuum region, and the corresponding contributions are all of second order (see also Rem. 2.4; note also that this requires point symmetry of V, which we have not assumed in general in this paper).

Second, we see that the QCE method (and only the QCE method) has a zeroth-order term $G\varepsilon^{1/p}$ in the interface region. This term occurs since the QCE method is not patch test consistent, that is, homogeneous deformations are not equilibria of the QCE model:

$$\delta\mathscr{E}_{\rm qce}(y_{\mathsf{A}}) \neq 0.$$

The origin and effect of these "ghost forces" are discussed in more detail in [11, 12, 35, 37, 50].

We should call this term zeroth order for several reasons: firstly, it is clearly of zeroth order if $p = \infty$, in which case the consistency error is related to the error in the $\mathrm{W}^{1,\infty}$-norm. Secondly, the parameter ε is a constant of the problem and does *not* tend to zero. As a matter of fact, the accuracy of an a/c method should be related the smoothness of the solution (as opposed to the atomistic scale), and the term $G\varepsilon^{1/p}$ is independent of the magnitude of y'' in the interface region. The scaling $\varepsilon^{1/p}$ relates only to the *width* of the interface region.

Finally, it is worth remarking on the first-order consistency term in the interface region for the QNL method. The reason this term is of first order as opposed to second order is the loss of symmetry that is introduced by changing the interaction law at the interface between the atomistic and continuum regions. A recent result of Dobson [9] shows that no a/c method coupling an interatomic potential to the Cauchy–Born continuum model can achieve better than first-order accuracy in the interface region.

Note also, that the second finite differences y''_n can in fact be written in terms of the oscillation operator:

$$\big|y''_n\big| = \operatorname{osc}\big(\nabla y; [x_n - \varepsilon, x_n + \varepsilon]\big).$$

In our analysis in Section 6, we will ignore the possibility of proving a modelling error estimate that is of second order in the continuum region, but we will be satisfied with an estimate that is globally of first order.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Remark 4.3. In the present paper we are only concerned with controlling the error in an energy norm. Using only this measure of accuracy, "consistent" methods such as QNL have a clear advantage over "inconsistent" methods such as QCE. A legitimate question to ask, however, is whether this observations remains true if the error in different quantities of interest are measured, such as, a critical load at which a defect moves, or the error measured only in a neighbourhood of the defect (as opposed to globally). The experiments in [55] suggest that for some quantities of interest QCE gives very accurate answers, but in other cases can give even qualitatively wrong results. A more detailed investigation of this question is required.

5. Auxiliary results

5.1. The bond density lemma

The bond density lemma is a tool that allows a transition between integrals over bonds, and volume integrals. It was first derived in [49] for the construction of patch test consistent a/c methods for pair potential interactions. Related asymptotic results were used previously in Γ-convergence analyses of atomistic models [3]; the achievement of Shapeev [49] was to obtain a formula that is *exact* for any triangle.

For any triangle $T \subset \mathbb{R}^2$ with vertices belonging to $\mathscr{L}^\#$ we define the characteristic function

$$\chi_T : \mathbb{R}^2 \to \mathbb{R}, \qquad \chi_T(x) := \lim_{t \searrow 0} \frac{|T \cap B_t(x)|}{|B_t(x)|},$$

where $B_t(x)$ denotes the closed ball with centre x and radius t. Note that $\chi_T = 1$ in $\mathrm{int}(T)$, and $\chi_T = 1/2$ on the edges of T. The value of χ_T on the corners is not of importance.

Let $x, x' \in \mathbb{R}^2$ and let φ be a function that is measurable on the line segment (x, x'); then we define the line integral, or *bond integral*,

$$\fint_x^{x'} \varphi \,\mathrm{db} := \int_{t=0}^1 \varphi\big((1-t)x + tx'\big)\,\mathrm{d}t.$$

Lemma 5.1 (Bond density lemma ([49], Lem. 2))**.** *Let $T \subset \mathbb{R}^2$ be a non-degenerate triangle with vertices belonging to $\mathscr{L}^\# = \varepsilon\mathbb{Z}^2$, and let $r \in \mathbb{Z}^2$; then*

$$\varepsilon^2 \sum_{x \in \mathscr{L}^\#} \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db} = |T|. \tag{5.1}$$

Remark 5.2. In the above form, the bond density lemma is false in 3D, which is one of the reasons why the present work is restricted to 2D. Moreover, the condition that the vertices of T belong to lattice sites is also necessary. This is related to the assumption that the vertices of $\mathscr{T}_h$ belong to $\mathscr{L}^\#$, however, this is not crucial and could be removed with some additional work.

5.2. Discrete divergence-free tensor fields in 2D

Our second auxiliary result is the representation of discrete divergence-free P_0-tensor fields [4]. The extension of the result to the periodic setting is in principle straightforward. Nevertheless, we present a complete proof, following a recent construction given by Polthier and Preuß [48], since it motivates a crucial argument in Section 6.3.3. Since we will only use the atomistic finite element mesh $\mathscr{T}_\varepsilon$ from now on, we will formulate everything in terms of this mesh. However, all results hold for general periodic triangulations.

5.2.1. The Crouzeix–Raviart finite element space

The representation of discrete divergence-free tensor fields requires the use of the non-conforming Crouzeix–Raviart finite element space. Recall from Section 3.1, the definition of the sets of edges $\mathscr{F}_\varepsilon$ and $\mathscr{F}_\varepsilon^\#$. The

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

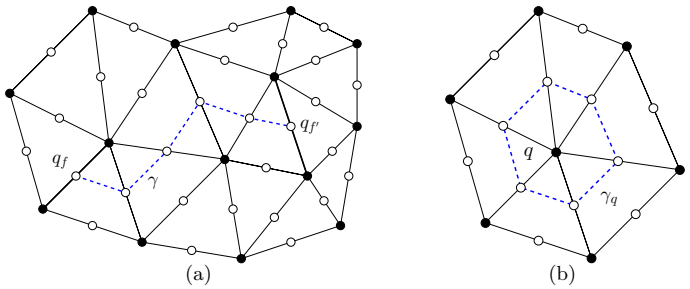


FIGURE 2. (a) Illustration of a piecewise affine path $\gamma \in \Gamma_{f,f'}$. (b) Illustration of the path γ_q used in the proof of Lemma 5.3.

Crouzeix–Raviart finite element space over $\mathscr{T}_\varepsilon^\#$ is defined as

$$\mathrm{N}_1(\mathscr{T}_\varepsilon^\#) = \Big\{ w : \cup_{T\in\mathscr{T}_\varepsilon^\#} \mathrm{int}(T) \to \mathbb{R} \,\Big|\, w \text{ is piecewise affine w.r.t. } \mathscr{T}_\varepsilon^\#, \text{ and continuous in edge midpoints } q_f, f \in \mathscr{F}_\varepsilon^\# \Big\}.$$

The degrees of freedom for functions $w \in \mathrm{N}_1(\mathscr{T}_\varepsilon^\#)$ are the values at edge midpoints, $w(q_f)$, $f \in \mathscr{F}_\varepsilon^\#$, and the corresponding nodal basis functions are denoted by ζ_f.

The Crouzeix–Raviart finite element space of periodic functions is defined as

$$\mathrm{N}_1^\#(\mathscr{T}_\varepsilon) = \left\{ w \in \mathrm{N}_1(\mathscr{T}_\varepsilon^\#) \,\middle|\, w(\xi + x) = w(x) \text{ for } \mathrm{d}x\text{-a.e. } x \in \mathbb{R}^2, \xi \in 2\mathbb{Z}^2 \right\}.$$

The periodic nodal basis functions are defined, for $f \in \mathscr{F}_\varepsilon$, by $\zeta_f^\# = \sum_{\xi\in 2\mathbb{Z}^2} \zeta_{\xi+f}$.

5.2.2. Path integrals

For two edges $f, f' \in \mathscr{F}_\varepsilon^\#$, let $\Gamma_{f,f'}$ denote the set of all piecewise affine paths from q_f to $q_{f'}$, crossing element edges only in edge midpoints; see Figure 2a for an example.

For any piecewise constant vector field $\sigma \in \mathrm{P}_0(\mathscr{T}_\varepsilon^\#)^2$ and for any path $\gamma \in \Gamma_{f,f'}$, $\gamma = \{x(t) \,|\, 0 \le t \le T\}$, we denote the standard path integral by

$$\int_\gamma \sigma \cdot \,\mathrm{d}x = \int_{t=0}^T \sigma \cdot \dot{x}(t)\,\mathrm{d}t.$$

For piecewise constant tensor fields $\sigma \in \mathrm{P}_0(\mathscr{T}_\varepsilon^\#)^{k\times 2}$ we define the path integral as

$$\int_\gamma \sigma \cdot \,\mathrm{d}x = \int_{t=0}^T \sigma\dot{x}(t)\,\mathrm{d}t.$$

Since functions $w \in \mathrm{N}_1(\mathscr{T}_\varepsilon^\#)^k$ have piecewise constant gradients ∇w, and since they are continuous in edge midpoints, it is easy to see that

$$\int_\gamma \nabla w \cdot \,\mathrm{d}x = w(q_{f'}) - w(q_f) \qquad \forall \gamma \in \Gamma_{f,f'}. \tag{5.2}$$

5.2.3. Discrete divergence-free tensor fields

The following lemma characterizes discrete divergence-free tensor fields.

Lemma 5.3. *A tensor field $\sigma \in \mathrm{P}_0^\#(\mathscr{T}_\varepsilon)^{k\times 2}$ satisfies*

$$\int_\Omega \sigma : \nabla u \,\mathrm{d}x = 0 \qquad \forall u \in \mathrm{P}_1^\#(\mathscr{T}_\varepsilon)^k$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

if and only if there exist a constant $\sigma_0 \in \mathbb{R}^{k\times 2}$ *and a function* $w \in \mathrm{N}_1^{\#}(\mathscr{T}_\varepsilon)^k$ *such that*

$$\sigma = \sigma_0 + \nabla w \mathsf{J}, \qquad \textit{where} \quad \mathsf{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathrm{SO}(2).$$

Sketch of the proof. The reverse direction, that any tensor field of the form $\sigma = \sigma_0 + \nabla w \mathsf{J}$ has zero discrete divergence, can be checked using a straightforward calculation, using (5.2).

Step 0. Outline: to simplify the notation, we define $\alpha = \sigma \mathsf{J}^T$ and, without loss of generality, we assume that $k = 1$. Initially, we treat α as a piecewise constant tensor field on all of $\mathbb{R}^2$, ignoring periodicity. We construct $w \in \mathrm{N}_1(\mathscr{T}_\varepsilon^{\#})$ by explicitly specifying $w(q_f)$ for all edges $f \in \mathscr{F}_\varepsilon^{\#}$. We will then show in the last step of the proof that w can be written as the sum of an affine function and a periodic function.

Step 1. Construction of w: fix a starting edge $\hat{f} \in \mathscr{F}_\varepsilon$ and define $w(q_{\hat{f}}) = 0$. For any edge $f \in \mathscr{F}_\varepsilon^{\#}$ let $\gamma \in \Gamma_{\hat{f},f}$ and define $w(q_f)$ via the path integral

$$w(q_f) := \int_\gamma \alpha \cdot \mathrm{d}x.$$

We need to show that this definition is independent of the path.

Let q be a vertex of the triangulation $\mathscr{T}_\varepsilon^{\#}$ and let φ_q be the corresponding nodal basis function with support ω_q; then a fairly straightforward calculation (see, *e.g.*, [48] for the details) shows that

$$0 = \int_{\omega_q} \sigma \cdot \nabla \varphi_q \,\mathrm{d}x = \frac{1}{2} \sum_{\substack{f \in \mathscr{F}_\varepsilon^{\#} \\ f \subset \mathrm{int}(\omega_q)}} \Big((\sigma \cdot \nu_f)^+ + (\sigma \cdot \nu_f)^- \Big) = \int_{\gamma_q} \alpha \cdot \mathrm{d}x,$$

where $\nu_f^{\pm}$ are the two unit normals to f, and γ_q is the piecewise affine path through edge midpoints circling q; *cf.* Figure 2b. Note that the rotation J in the definition of α comes from the fact that tangent vectors are rotated normal vectors.

Since all closed piecewise affine paths can be written as a sum over paths γ_q, this implies that $\int_\gamma \alpha \cdot \mathrm{d}x = 0$ for all closed piecewise affine paths γ, and in particular that the definition of $w(q_f)$ is independent of the choice of path, that is, w is well-defined.

Step 2. $\nabla w = \alpha$*:* from the definition of $w(q_f)$, $f \in \mathscr{F}_\varepsilon^{\#}$, it follows that, for $f, f' \subset T \in \mathscr{T}_\varepsilon^{\#}$,

$$\nabla w(T) \cdot (q_f - q_{f'}) = w(q_f) - w(q_{f'}) = \alpha(T) \cdot (q_f - q_{f'}),$$

which immediately implies that $\nabla w(T) = \alpha(T)$.

Step 3. Periodicity: we are only left to show that $w(x) = a \cdot x + w_1(x)$ for some $a \in \mathbb{R}^2$, $a = (a_1, a_2)$, and $w_1 \in \mathrm{N}_1^{\#}(\mathscr{T}_\varepsilon)$. Let $a_j = w(q_{\hat{f}} + 2e_j)$, $j = 1, 2$, and define $w_1(x) = w(x) - a \cdot x$. Fix $j \in \{1, 2\}$, let $f \in \mathscr{F}_\varepsilon^{\#}$ and let γ be a path from $q_{\hat{f}}$ to q_f. Since α is $2\mathbb{Z}^2$-periodic, and since $w_1(q_{\hat{f}} + 2e_j) = w_1(q_{\hat{f}}) = 0$, we have

$$w_1(q_f + 2e_j) = \int_{\gamma + 2e_j} \alpha \cdot \mathrm{d}x = \int_\gamma \alpha \cdot \mathrm{d}x = w_1(q_f).$$

This shows that w_1 is periodic and thus concludes the proof. □

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

6. A general consistency result in 2D

We are now finally in a position to make precise the statement that *patch test consistent a/c methods are first-order consistent*. Motivated by the example of the QNL consistency result (ignoring, as discussed in Sect. 4.3, the second-order consistency of the Cauchy–Born approximation), we would like to prove a result of the form

$$\left\|\delta\mathcal{E}_{\rm ac}(y)-\delta\mathcal{E}_{\rm a}(y)\right\|_{\mathrm{W}_\varepsilon^{-1,p}}\leq C\varepsilon\left\|\nabla^2 y\right\|_{\mathrm{L}^p(\Omega_{\rm c}\cup\Omega_{\rm i})}.$$

As explained in Section 3.2.1 we will use the oscillation operator (3.1) to replace the undefined second derivative.

For each element $T\in\mathscr{T}_\varepsilon^{\#}$, let $\omega_T^{\rm a}$ be the *interaction neighbourhood* of T in the atomistic model,

$$\omega_T^{\rm a}:=\left\{x+t_1r_1+t_2r_2\,\middle|\,x\in T,t_i\in[0,1],r_i\in\mathscr{R}\right\},\tag{6.1}$$

and ω_T its union with the set $\omega_T^{\rm c}$, restricted to the continuum and interface regions,

$$\omega_T:=\left(\omega_T^{\rm a}\cup\bigcup_{\substack{T'\in\mathscr{T}_\varepsilon^{\#}\\ T\cap T'\neq\emptyset}}T'\right)\setminus\Omega_{\rm a}^{\#}.\tag{6.2}$$

Note that assumption (2.14) implies that $\omega_T^{\rm a}\cup\omega_T^{\rm c}\subset\omega_T$ for all $T\in\mathscr{T}_\varepsilon^{\rm c}$.

Theorem 6.1 (First-order consistency of patch test consistent a/c methods)**.** *Suppose that $\mathcal{E}_{\rm ac}$ is patch test consistent (Sect. 2.3.3) and globally energy consistent (Sect. 2.3.2), that the locality condition (2.19) and the scaling condition (2.20) hold, that $\mathrm{int}(\Omega_{\rm a})$ is connected, that $\mathscr{T}_h^{\rm a}\cup\mathscr{T}_h^{\rm i}\subset\mathscr{T}_\varepsilon$, and that (2.14) holds. Then, for any $y\in\mathscr{Y}$, we have*

$$\left\|\delta\mathcal{E}_{\rm ac}(y)-\delta\mathcal{E}_{\rm a}(y)\right\|_{\mathrm{W}_\varepsilon^{-1,p}}\leq\varepsilon\left\{\sum_{T\in\mathscr{T}_\varepsilon^{\rm c}\cup\mathscr{T}_\varepsilon^{\rm i}}|T|\Big[M_T\,\mathrm{osc}(\nabla y;\omega_T)\Big]^p\right\}^{1/p},\tag{6.3}$$

where the oscillation measure osc *is defined in (3.1), the interaction neighbourhood ω_T is defined in (6.2), and the prefactors M_T are defined as follows:*

$$M_T=\begin{cases}0, & T\in\mathscr{T}_\varepsilon^{\rm a},\\ \big(M^{\rm i}+M^{\rm a}\big)\big(1+7\,\mathrm{width}(\Omega_{\rm i})\big), & T\in\mathscr{T}_\varepsilon^{\rm i},\\ M^{\rm a}+7\big(M^{\rm i}+M^{\rm a}\big)\mathrm{width}(\Omega_{\rm i}), & T\in\mathscr{T}_\varepsilon^{\rm c}.\end{cases}\tag{6.4}$$

In (6.4), the constants $M^{\rm a}$ and $M^{\rm i}$ are defined in (2.4) and (2.21), and the "interface width" $\mathrm{width}(\Omega_{\rm i})$ is given by

$$\mathrm{width}(\Omega_{\rm i}):=\max_{\substack{f\in\mathscr{T}_\varepsilon\\ f\subset\Omega_{\rm i}^{\#}}}\min_{\substack{f'\in\mathscr{T}_\varepsilon^{\#}\\ f'\subset\Omega_{\rm a}^{\#}}}\min_{\gamma\in\Gamma_{f,f'}}\frac{\mathrm{length}(\gamma)}{\varepsilon}.\tag{6.5}$$

Outline of the proof. We will construct "stress functions" $\Sigma_{\rm a}(y),\Sigma_{\rm ac}(y)\in\mathrm{P}_0(\mathscr{T}_\varepsilon)^{2\times2}$ such that

$$\big\langle\delta\mathcal{E}_{\rm a}(y),u\big\rangle=\int_\Omega\Sigma_{\rm a}(y):\nabla u\,\mathrm{d}x,\quad\text{and}\quad\big\langle\delta\mathcal{E}_{\rm ac}(y);u\big\rangle=\int_\Omega\Sigma_{\rm ac}(y):\nabla u\,\mathrm{d}x\qquad\forall u\in\mathscr{U}.$$

If we could prove an estimate of the form

$$\big|\Sigma_{\rm a}(y;T)-\Sigma_{\rm ac}(y;T)\big|\lesssim\varepsilon\,\mathrm{osc}(\nabla y;\omega_T),$$

then the result would follow immediately. It turns out that this is not possible.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Instead, we will use the fact that $\mathscr{E}_{\rm ac}$ is globally energy consistent and patch test consistent to construct a correction (*cf.* Cor. 6.7, Lem. 6.9, and Sect. 6.2.3)

$$\widehat{\Sigma}_{\rm ac}(y) = \Sigma_{\rm ac}(y) - \nabla\hat{\psi}(y)\mathsf{J},$$

which still represents the first variation $\delta\mathscr{E}_{\rm ac}$ (*cf.* Lem. 5.3), and which has the property that

$$\Sigma_{\rm a}(y_{\mathsf{F}};T) = \widehat{\Sigma}_{\rm ac}(y_{\mathsf{F}};T) = \partial W(\mathsf{F}) \qquad \forall\, T \in \mathscr{T}_\varepsilon^{\rm c} \cup \mathscr{T}_\varepsilon^{\rm i}, \quad \mathsf{F} \in \mathbb{R}^{2\times 2}.$$

In addition, we show that $\widehat{\Sigma}_{\rm ac}(y;T) = \Sigma_{\rm a}(y;T)$ for all $T \in \mathscr{T}_\varepsilon^{\rm a}$.

Lipschitz estimates for $\Sigma_{\rm a}$ and $\widehat{\Sigma}_{\rm ac}$, and careful modifications of the argument in the interface region, yield the following result (*cf.* Lem. 6.11):

$$\begin{aligned}\left|\widehat{\Sigma}_{\rm ac}(y;T) - \Sigma_{\rm a}(y;T)\right| &\leq \left|\widehat{\Sigma}_{\rm ac}(y;T) - \partial W(\nabla y(T))\right| + \left|\Sigma_{\rm a}(y;T) - \partial W(\nabla y(T))\right| \\ &\leq \varepsilon M_T \mathrm{osc}(\nabla y;\omega_T) \qquad \forall T \in \mathscr{T}_\varepsilon,\end{aligned}$$

and, in particular,

$$\begin{aligned}\left\langle \delta\mathscr{E}_{\rm ac}(y) - \delta\mathscr{E}_{\rm a}(y); u\right\rangle &= \sum_{T\in\mathscr{T}_\varepsilon} |T|\Big[\widehat{\Sigma}_{\rm ac}(y;T) - \Sigma_{\rm a}(y;T)\Big] : \nabla u(T) \\ &\leq \varepsilon \left\{ \sum_{T\in\mathscr{T}_\varepsilon^{\rm c}\cup\mathscr{T}_\varepsilon^{\rm i}} |T|\Big[M_T\,\mathrm{osc}(\nabla y;\omega_T)\Big]^p \right\}^{1/p} \|\nabla u\|_{\mathrm{L}^{p'}(\Omega)},\end{aligned}$$

which yields the stated first order consistency estimate. □

Remark 6.2. In 1D, a similar result can be proven using a similar framework but with significantly reduced technicalities. Note, in particular, that the 1D analogue of Lemma 5.3 is

$$\int_\Omega \sigma\cdot\nabla u\,\mathrm{d}x = 0 \quad \forall\, u \in \mathscr{U} \qquad \text{if and only if} \qquad \sigma \text{ is constant.}$$

Hence, the corrector function $\psi(\mathsf{F},\cdot)$ defined in Section 6.2.2 is always identically equal to zero, which removes the interface width dependence from the modelling error (*cf.* Sect. 6.3.3). Hence, if $d = 1$, we obtain (6.3) again but with modified prefactors

$$M_T^{\rm 1D} := \begin{cases} 0, & T \in \mathscr{T}_\varepsilon^{\rm a}, \\ M^{\rm i} + M^{\rm a}, & T \in \mathscr{T}_\varepsilon^{\rm i}, \\ M^{\rm a}, & T \in \mathscr{T}_\varepsilon^{\rm c}. \end{cases} \tag{6.6}$$

In one dimension, since $\psi \equiv 0$, and since symmetries are more easily exploited, it is not too difficult to prove second order consistency in the continuum region.

6.1. The atomistic stress function

A natural "weak" representation of $\delta\mathscr{E}_{\rm a}(y)$, $y \in \mathscr{Y}$, is given by

$$\langle \delta\mathscr{E}_{\rm a}(y), z\rangle = \varepsilon^2 \sum_{x\in\mathscr{L}} \sum_{r\in\mathscr{R}} \partial_r V(D_{\mathscr{R}} y(x))\cdot D_r z(x), \qquad \text{for } z \in \mathscr{Y}. \tag{6.7}$$

Using bond integrals we rewrite this in a form that will be useful for our subsequent analysis. We will then use the bond density lemma whenever we need to transition between bond integrals and volume integrals. This process yields a notion of stress for atomistic models, which is related to the virial stress (see [2] for a recent reference; this connection will be discussed in detail elsewhere). A variant of this result for pair interactions in 1D was developed in [33].

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Proposition 6.3. *Let $y, z \in \mathscr{Y}$, then*

$$\langle \delta \mathscr{E}_{\mathrm{a}}(y), z \rangle = \sum_{T \in \mathscr{T}_\varepsilon} |T|\, \Sigma_{\mathrm{a}}(y;T) : \nabla z(T) = \int_\Omega \Sigma_{\mathrm{a}}(y) : \nabla z \,\mathrm{d}x,$$

where the stress function $\Sigma_{\mathrm{a}}(y) \in \mathrm{P}_0^{\#}(\mathscr{T}_\varepsilon)^{2\times 2}$ is defined as follows:

$$\Sigma_{\mathrm{a}}(y;T) := \sum_{r \in \mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x \in \mathscr{L}^{\#}} \Big[\partial_r V\big(D_{\mathscr{R}} y(x)\big) \otimes r\Big] \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{d}b. \tag{6.8}$$

Proof. For the sake of brevity we will write $V_{x,r} = \partial_r V(D_{\mathscr{R}} y(x))$.

Recall that $T^{\#} = \bigcup_{\xi \in 2\mathbb{Z}^2} (\xi + T)$, for $T \in \mathscr{T}_\varepsilon$. It is easy to see that $\{\chi_{T^{\#}} \,|\, T \in \mathscr{T}_\varepsilon\}$ is a partition of unity for $\mathbb{R}^2$. Hence, using the identity

$$\fint_x^{x+\varepsilon r} \nabla_r z \,\mathrm{d}b = D_r z(x), \tag{6.9}$$

we can rewrite (6.7) as

$$\big\langle \delta \mathscr{E}_{\mathrm{a}}(y), z \big\rangle = \varepsilon^2 \sum_{x \in \mathscr{L}} \sum_{r \in \mathscr{R}} V_{x,r} \cdot \fint_x^{x+\varepsilon r} \left[\sum_{T \in \mathscr{T}_\varepsilon} \chi_{T^{\#}} \right] \nabla_r z \,\mathrm{d}b.$$

Interchanging the order of summation and using the fact that $\nabla_r z = (\nabla z) r$ holds db-a.e. (note that, if the bond is aligned with an element edge, then $\nabla_r z$ is continuous across that edge) yields

$$\langle \delta \mathscr{E}_{\mathrm{a}}(y), z \rangle = \sum_{T \in \mathscr{T}_\varepsilon} |T| \sum_{r \in \mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x \in \mathscr{L}} V_{x,r} \cdot \fint_x^{x+\varepsilon r} \chi_{T^{\#}} \nabla z(T^{\#}) r \,\mathrm{d}b.$$

The term $\nabla z(T^{\#}) r = \nabla z(T) r$ can be taken outside the bond integral, and hence, employing the identity

$$a \cdot (\mathsf{G} r) = (a \otimes r) : \mathsf{G}, \qquad \text{for } a, r \in \mathbb{R}^d, \mathsf{G} \in \mathbb{R}^{d\times d}, \tag{6.10}$$

yields

$$\begin{aligned} \langle \delta \mathscr{E}_{\mathrm{a}}(y), z \rangle &= \sum_{T \in \mathscr{T}_\varepsilon} |T| \left\{ \sum_{r \in \mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x \in \mathscr{L}} [V_{x,r} \otimes r] \fint_x^{x+\varepsilon r} \chi_{T^{\#}} \,\mathrm{d}b \right\} : \nabla z(T) \\ &=: \sum_{T \in \mathscr{T}_\varepsilon} |T|\, \Sigma_{\mathrm{a}}(y;T) : \nabla z(T). \end{aligned}$$

Finally, we use the fact that $D_{\mathscr{R}} y$ is $2\mathbb{Z}^2$-periodic to deduce that

$$\begin{aligned} \Sigma_{\mathrm{a}}(y;T) &= \sum_{r \in \mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x \in \mathscr{L}} [V_{x,r} \otimes r] \fint_x^{x+\varepsilon r} \chi_{T^{\#}} \,\mathrm{d}b \\ &= \sum_{r \in \mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x \in \mathscr{L}} [V_{x,r} \otimes r] \sum_{\xi \in 2\mathbb{Z}^2} \fint_x^{x+\varepsilon r} \chi_{\xi + T} \,\mathrm{d}b \\ &= \sum_{r \in \mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x \in \mathscr{L}^{\#}} [V_{x,r} \otimes r] \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{d}b. \end{aligned}$$

□

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

The terminology "stress function" for $\Sigma_{\rm a}$ is motivated by the fact that $\Sigma_{\rm a}(y)$ takes precisely the same role as the first Piola–Kirchhoff stress tensor in the continuum theory of elasticity.

In the next lemma we prove two useful properties of the atomistic stress function $\Sigma_{\rm a}$. We show that $\Sigma_{\rm a} = \partial W$ under locally homogeneous deformations and give a quantitative estimate for the discrepancy between $\Sigma_{\rm a}$ and ∂W.

Lemma 6.4. *The stress function $\Sigma_{\rm a}$ defined in (6.8) satisfies*

$$\Sigma_{\rm a}(y_{\mathsf{F}}; T) = \partial W(\mathsf{F}) \qquad \forall \mathsf{F} \in \mathbb{R}^{2\times 2}, \quad T \in \mathscr{T}_\varepsilon.$$

Moreover, we have the estimate

$$\left|\Sigma_{\rm a}(y; T) - \partial W\big(\nabla y(T)\big)\right| \leq \varepsilon M^{\rm a} \mathrm{osc}(\nabla y; \omega_T^{\rm a}) \qquad \forall y \in \mathscr{Y}, \quad T \in \mathscr{T}_\varepsilon, \tag{6.11}$$

where $M^{\rm a}$ is defined in (2.4) and $\omega_T^{\rm a}$ is defined in (6.1).

Proof.

Part 1: since $[\partial_r V(\mathsf{F}\mathscr{R}) \otimes r]$ is independent of x, we can apply the bond density lemma to the sum in curly brackets, to deduce that

$$\Sigma_{\rm a}(y_{\mathsf{F}}; T) = \sum_{r\in\mathscr{R}} \big[\partial_r V\big(\mathsf{F}\mathscr{R}\big) \otimes r\big] \left\{ \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}^\#} \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db} \right\} = \sum_{r\in\mathscr{R}} \big[\partial_r V\big(\mathsf{F}\mathscr{R}\big) \otimes r\big]\,.$$

Recalling that $W(\mathsf{F}) = V(\mathsf{F}\mathscr{R})$, it can be easily checked that the sum on the right-hand side of the second equality equals $\partial W(\mathsf{F})$.

Part 2: let $\mathsf{F} = \nabla y(T)$, $V_{x,r} = \partial_r V(D_{\mathscr{R}} y(x))$, and $V_{\mathsf{F},r} = \partial_r V(\mathsf{F}\mathscr{R})$. From part 1 we obtain that

$$\begin{aligned}
\left|\Sigma_{\rm a}(y; T) - \partial W(\mathsf{F})\right| &= \left| \sum_{r\in\mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}^\#} \Big[\big(V_{x,r} - V_{\mathsf{F},r}\big) \otimes r\Big] \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db} \right| \\
&\leq \sum_{r\in\mathscr{R}} |r| \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}^\#} \left|V_{x,r} - V_{\mathsf{F},r}\right| \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db}.
\end{aligned} \tag{6.12}$$

We use the Lipschitz property 2.3 to estimate

$$\begin{aligned}
\left|V_{x,r} - V_{\mathsf{F},r}\right| &\leq \sum_{s\in\mathscr{R}} M_{r,s}^{\rm a} \left|D_s y(x) - \mathsf{F}s\right| \\
&= \sum_{s\in\mathscr{R}} M_{r,s}^{\rm a} \left| \fint_x^{x+\varepsilon s} \Big(\nabla_s y - \mathsf{F}s\Big) \,\mathrm{db} \right| \\
&\leq \sum_{s\in\mathscr{R}} M_{r,s}^{\rm a} |s| \max_{x'\in(x,x+\varepsilon s)} \left|\nabla y(x') - \mathsf{F}\right|.
\end{aligned} \tag{6.13}$$

Since $(x, x+\varepsilon s) \subset \omega_T^{\rm a}$, and recalling that $\mathsf{F} = \nabla y(T)$, we can further estimate

$$\max_{x'\in(x,x+\varepsilon s)} \left|\nabla y(x') - \mathsf{F}\right| \leq \varepsilon \mathrm{osc}(\nabla y; \omega_T^{\rm a}). \tag{6.14}$$

We combine (6.14) with (6.13) and insert the resulting estimate into (6.12) to arrive at

$$\left|\Sigma_{\rm a}(y; T) - \partial W(\mathsf{F})\right| \leq \varepsilon \mathrm{osc}(\nabla y; \omega_T^{\rm a}) \sum_{r\in\mathscr{R}} \sum_{s\in\mathscr{R}} |r||s| M_{r,s}^{\rm a} \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}^\#} \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db}.$$

An application of the bond density lemma, and referring to the definition of $M^{\rm a}$ in (2.4), yields the stated result. □

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

6.2. The a/c stress function

We wish to derive a similar representation of $\delta\mathscr{E}_{\rm ac}$ in terms of a stress function $\Sigma_{\rm ac}$, as we did in Section 6.1 for $\delta\mathscr{E}_{\rm a}$. A straightforward calculation along the same lines as the proof of Proposition 6.3, recalling first the definition of the partial derivative $\partial_b E_{\rm i}$ from Section 2.3.6, yields the following result.

Proposition 6.5. *Suppose that* (2.14) *holds, then, for all* $y, z \in \mathscr{Y}$,

$$\big\langle \delta\mathscr{E}_{\rm ac}(y), z\big\rangle = \sum_{T\in\mathscr{T}_\varepsilon} |T|\, \Sigma_{\rm ac}(y;T) : \nabla z(T) = \int_\Omega \Sigma_{\rm ac}(y) : \nabla z \,{\rm d}x,$$

where

$$\Sigma_{\rm ac}(y;T) := \begin{cases} \Sigma_{\rm a}(y;T), & T\in\mathscr{T}_\varepsilon^{\rm a},\\ \partial W(\nabla y(T)), & T\in\mathscr{T}_\varepsilon^{\rm c},\\ \sum_{r\in\mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}_{\rm a}^\#} \big[\partial_r V\big(D_{\mathscr{R}} y(x)\big)\otimes r\big] \fint_x^{x+\varepsilon r} \chi_T \,{\rm d}b \\ \quad + \frac{\varepsilon^2}{|T|} \sum_{(x,x+\varepsilon r)\in\mathscr{B}_{\rm i}^\#} \big[\partial_{(x,x+\varepsilon r)} E_{\rm i}(y)\otimes r\big] \fint_x^{x+\varepsilon r} \chi_T^{\rm i} \,{\rm d}b, & T\in\mathscr{T}_\varepsilon^{\rm i}, \end{cases}$$

where $\chi_T^{\rm i}$ *is a modified characteristic function:*

$$\chi_T^{\rm i}(x) = \begin{cases} 1, & x\in\partial\Omega_{\rm i}^\#,\\ \chi_T(x), & \textit{otherwise.}\end{cases}$$

Proof. Employing again the notation $V_{x,r} = \partial_r V(D_{\mathscr{R}} y(x))$, the functional $\delta\mathscr{E}_{\rm ac}(y)$ can be written as

$$\big\langle \delta\mathscr{E}_{\rm ac}(y), z\big\rangle = \int_{\Omega_{\rm c}} \partial W(\nabla y) : \nabla z \,{\rm d}x + \varepsilon^2 \sum_{x\in\mathscr{L}_{\rm a}} \sum_{r\in\mathscr{R}} V_{x,r}\cdot D_r u(x) + \varepsilon^2 \sum_{(x,x+\varepsilon r)\in\mathscr{B}_{\rm i}} \partial_{(x,x+\varepsilon r)} E_{\rm i}(y)\cdot D_r u(x). \quad (6.15)$$

The first term on the right-hand side of (6.15) gives rise to the definition of $\Sigma_{\rm ac}(y;T)$ for $T\in\mathscr{T}_\varepsilon^{\rm c}$. (Note that (2.14) guarantees that the bonds in the second group in (6.15) do not contribute to $\Omega_{\rm c}$.)

After the same calculation as in the proof of Proposition 6.3, the second group on the right-hand side of (6.15) gives the definition of $\Sigma_{\rm ac}(y;T)$ for $T\in\mathscr{T}_\varepsilon^{\rm a}$, as well as the first group in the definition of $\Sigma_{\rm ac}(y;T)$ for $T\in\mathscr{T}_\varepsilon^{\rm i}$.

The crucial modification to the previous argument is that the modified characteristic functions $\chi_T^{\rm i}$, $T\in\mathscr{T}_\varepsilon^{\rm i}$ form a partition of unity for $\Omega_{\rm i}$ (except at a finite number of points, which do not contribute to bond integrals). Therefore, performing again a similar calculation as in the proof of Proposition 6.3 to "distribute" the third group on the right-hand side of (6.15) between interface elements only, we obtain the second group in the definition of $\Sigma_{\rm ac}(y;T)$ for $T\in\mathscr{T}_\varepsilon^{\rm i}$.

Note that if we hadn't made the modification to the characteristic function, then $\Sigma_{\rm ac}(y;T)$, $T\in\mathscr{T}_\varepsilon^{\rm a}\cup\mathscr{T}_\varepsilon^{\rm c}$ would contain contributions from $E_{\rm i}$. □

Since any discrete divergence-free tensor field may be added to $\Sigma_{\rm ac}$ and still yield a valid representation of $\delta\mathscr{E}_{\rm ac}$, it is not surprising that, in general, $\Sigma_{\rm ac}$ does not have the *necessary property* that $\Sigma_{\rm ac}(y_{\mathsf F};T) = \partial W(\mathsf F)$ for all $\mathsf F\in\mathbb{R}^{2\times 2}$. This can already be observed in the nearest-neighbour, flat interface constructions in [45]. Hence, we need to construct an alternative stress function $\widehat{\Sigma}_{\rm ac}$ representing $\delta\mathscr{E}_{\rm ac}$ that does have the desired properties. This construction will be undertaken in the remainder of this section, using the representation of discrete divergence-free vector fields as gradients of Crouzeix–Raviart functions discussed in Section 5.2.

6.2.1. *Consequences of global energy consistency*

Recall from (2.11) that a functional $\mathscr{E}\in{\rm C}^1(\mathscr{Y})$ is called globally energy consistent if $\mathscr{E}(y_{\mathsf F}) = \mathscr{E}_{\rm a}(y_{\mathsf F})$ for all matrices $\mathsf F\in\mathbb{R}^{2\times 2}$. The following lemma establishes a simple but crucial consequence of this property.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Lemma 6.6. *Suppose that $\mathscr{E} \in \mathrm{C}^1(\mathscr{Y})$ is globally energy consistent, then*

$$\langle \delta\mathscr{E}(y_{\mathsf{F}}), y_{\mathsf{G}} \rangle = |\Omega|\, \partial W(\mathsf{F}) : \mathsf{G} \qquad \forall\, \mathsf{F}, \mathsf{G} \in \mathbb{R}^{2\times 2}.$$

Proof. From the assumption of global energy consistency, and (2.10), we obtain that

$$\mathscr{E}(y_{\mathsf{F}}) = \mathscr{E}_{\mathrm{a}}(y_{\mathsf{F}}) = |\Omega| W(\mathsf{F}) \qquad \forall \mathsf{F} \in \mathbb{R}^{2\times 2}.$$

Since $y_{\mathsf{F}} + t y_{\mathsf{G}} = y_{\mathsf{F}+t\mathsf{G}}$ this implies that

$$\big\langle \delta\mathscr{E}(y_{\mathsf{F}}), y_{\mathsf{G}} \big\rangle = |\Omega| \lim_{t\to 0} \frac{W(\mathsf{F}+t\mathsf{G}) - W(\mathsf{F})}{t} = |\Omega| \partial W(\mathsf{F}) : \mathsf{G}. \qquad \square$$

If we apply the foregoing lemma to an a/c functional $\mathscr{E}_{\mathrm{ac}}$ we obtain the following corollary.

Corollary 6.7. *Suppose that $\mathscr{E}_{\mathrm{ac}}$ is globally energy consistent, then*

$$⨍_{\Omega} \Sigma_{\mathrm{ac}}(y_{\mathsf{F}})\, \mathrm{d}x = \partial W(\mathsf{F}) \qquad \forall \mathsf{F} \in \mathbb{R}^{2\times 2}.$$

Proof. This result follows simply from Lemma 6.6 and the fact that, for all $\mathsf{G} \in \mathbb{R}^{2\times 2}$,

$$\big\langle \delta\mathscr{E}_{\mathrm{ac}}(y_{\mathsf{F}}), y_{\mathsf{G}} \big\rangle = \int_{\Omega} \Sigma_{\mathrm{ac}}(y_{\mathsf{F}}) : \mathsf{G}\, \mathrm{d}x = \left(⨍_{\Omega} \Sigma_{\mathrm{ac}}(y_{\mathsf{F}})\, \mathrm{d}x \right) : \mathsf{G}. \qquad \square$$

6.2.2. Consequences of patch test consistency

The examples given in [45] show that patch test consistency does not necessarily imply that $\Sigma_{\mathrm{ac}}(y_{\mathsf{F}}; T) = \partial W(\mathsf{F})$. In the current paragraph we characterise the discrepancy between $\Sigma_{\mathrm{ac}}(y_{\mathsf{F}}; T)$ and $\partial W(\mathsf{F})$.

First, we show that the test functions $u_h \in \mathscr{U}_h$ in the patch test (2.12) may be replaced by arbitrary displacements $u \in \mathscr{U}$.

Lemma 6.8. *Suppose that $\mathscr{E}_{\mathrm{ac}}$ is patch test consistent and that $\mathscr{T}_h^{\mathrm{a}} \cup \mathscr{T}_h^{\mathrm{i}} \subset \mathscr{T}_\varepsilon$; then we also have*

$$\big\langle \delta\mathscr{E}_{\mathrm{ac}}(y_{\mathsf{F}}), u \big\rangle = 0 \qquad \forall\, u \in \mathscr{U}, \quad \mathsf{F} \in \mathbb{R}^{2\times 2}. \tag{6.16}$$

Proof. Fix $u \in \mathscr{U}$; then, using the assumption that $\mathscr{T}_h^{\mathrm{a}} \cup \mathscr{T}_h^{\mathrm{i}} \subset \mathscr{T}_\varepsilon$, we have

$$\int_{\Omega_{\mathrm{c}}} \nabla u\, \mathrm{d}x = \int_{(\partial\Omega_{\mathrm{c}}^{\#})\cap\Omega} u \otimes \nu\, \mathrm{d}s = \int_{(\partial\Omega_{\mathrm{c}}^{\#})\cap\Omega} I_h u \otimes \nu\, \mathrm{d}s = \int_{\Omega_{\mathrm{c}}} \nabla I_h u\, \mathrm{d}x.$$

Since $\Sigma_{\mathrm{ac}}(y_{\mathsf{F}}) = \partial W(\mathsf{F})$ is constant in Ω_{c}, and since $I_h u = u$ in $\Omega_{\mathrm{i}} \cup \Omega_{\mathrm{a}}$, we can therefore deduce that

$$\big\langle \delta\mathscr{E}_{\mathrm{ac}}(y_{\mathsf{F}}), u \big\rangle = \int_{\Omega} \Sigma_{\mathrm{ac}}(y_{\mathsf{F}}) : \nabla u\, \mathrm{d}x = \int_{\Omega} \Sigma_{\mathrm{ac}}(y_{\mathsf{F}}) : \nabla I_h u\, \mathrm{d}x = \big\langle \delta\mathscr{E}_{\mathrm{ac}}(y_{\mathsf{F}}), I_h u \big\rangle = 0.$$

The penultimate equality requires some justification, but follows quite easily from the particular form of $\mathscr{E}_{\mathrm{ac}}$ assumed in (2.16) and the assumption that $\mathscr{T}_h^{\mathrm{a}} \cup \mathscr{T}_h^{\mathrm{i}} \subset \mathscr{T}_\varepsilon$. $\square$

Lemma 6.9. *Suppose that $\mathscr{E}_{\mathrm{ac}}$ is patch test consistent and globally energy consistent and that $\mathscr{T}_h^{\mathrm{i}} \cup \mathscr{T}_h^{\mathrm{a}} \subset \mathscr{T}_\varepsilon$; then, for each $\mathsf{F} \in \mathbb{R}^{2\times 2}$, there exists a function $\psi(\mathsf{F}; \cdot) \in \mathrm{N}_1^{\#}(\mathscr{T}_\varepsilon)^2$ such that*

$$\Sigma_{\mathrm{ac}}(y_{\mathsf{F}}; T) = \partial W(\mathsf{F}) + \nabla\psi(\mathsf{F}; T)\mathsf{J} \qquad \forall\, T \in \mathscr{T}_\varepsilon,$$

where J is a rotation matrix defined in Lemma 5.3. Moreover, if $\mathrm{int}(\Omega_{\mathrm{a}})$ is connected, then we may choose $\psi(\mathsf{F}) = 0$ in Ω_{a}.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Proof. If $\mathcal{E}_{\rm ac}$ is patch test consistent then, according to Lemma 6.8,

$$\int_\Omega \Sigma_{\rm ac}(y_{\mathsf{F}}) : \nabla u \, {\rm d}x = \langle \delta \mathcal{E}_{\rm ac}(y_{\mathsf{F}}), u \rangle = 0 \qquad \forall u \in \mathscr{U}.$$

Hence, according to Lemma 5.3, there exists a constant $\Sigma_0 \in \mathbb{R}^{2\times 2}$, a vector-valued Crouzeix–Raviart function $\psi = \psi(\mathsf{F};\cdot) \in {\rm N}_1^\#(\mathscr{T}_\varepsilon)^2$, and a rotation matrix J, such that

$$\Sigma_{\rm ac}(y_{\mathsf{F}}; T) = \Sigma_0 + \nabla\psi(T)\mathsf{J} \qquad \forall T \in \mathscr{T}_\varepsilon.$$

Using global energy consistency of $\mathcal{E}_{\rm ac}$ and Corollary 6.7 we obtain that

$$\partial W(\mathsf{F}) = \fint_\Omega \Sigma_{\rm ac}(y_{\mathsf{F}}) \, {\rm d}x = \Sigma_0 + \fint_\Omega \nabla\psi \mathsf{J} \, {\rm d}x.$$

If $\int_\Omega \nabla\psi \, {\rm d}x = 0$, then $\Sigma_0 = \partial W(\mathsf{F})$ and hence the result follows.

To prove this, we integrate by parts separately in each element:

$$\int_\Omega \nabla\psi \, {\rm d}x = \sum_{T\in\mathscr{T}_\varepsilon} \int_{\partial T} \psi \otimes \nu \, {\rm d}s = \sum_{f\in\mathscr{F}_\varepsilon} \int_f \left(\psi^+ \otimes \nu^+ + \psi^- \otimes \nu^-\right) {\rm d}s = 0,$$

where, in the last equality, we used the fact that $\int_f (\psi^+ - \psi^-)\, {\rm d}s = 0$ for all edges f, since ψ is continuous in the edge midpoints.

Finally, since $\Sigma_{\rm ac}(y_{\mathsf{F}}; T) = \partial W(\mathsf{F})$ for all $T \in \mathscr{T}_\varepsilon^{\rm a}$ (*cf.* Lems. 6.5 and 6.4), it follows that $\nabla\psi = 0$ in $\Omega_{\rm a}$. Hence, if ${\rm int}(\Omega_{\rm a})$ is connected, then we can shift ψ by a constant so that $\psi = 0$ in $\Omega_{\rm a}$. □

6.2.3. The modified a/c stress function

We wish to construct a modified a/c stress function $\widehat{\Sigma}_{\rm ac}$ that can be used to represent $\delta\mathcal{E}_{\rm ac}$, and satisfies the crucial property that $\widehat{\Sigma}_{\rm ac}(y_{\mathsf{F}}; T) = \partial W(\mathsf{F})$ for all $\mathsf{F} \in \mathbb{R}^{2\times 2}$, $T \in \mathscr{T}_\varepsilon^{\rm i}$.

To this end, we generalize the Crouzeix–Raviart function $\psi(\mathsf{F})$ defined in Lemma 6.9 to arbitrary deformations $y \in \mathscr{Y}$. Since we will use later on that the modified function $\hat{\psi}$ vanishes in $\Omega_{\rm a}$, we require from now on that ${\rm int}(\Omega_{\rm a})$ is connected so that we can choose $\psi(\mathsf{F},\cdot) = 0$ in $\Omega_{\rm a}$ for all $\mathsf{F} \in \mathbb{R}^{2\times 2}$.

For each $y \in \mathscr{Y}$ and each face $f \in \mathscr{F}_\varepsilon^\#$, $f = T_1 \cap T_2$, we define the patch $\omega_f = (T_1 \cup T_2) \setminus \Omega_{\rm a}^\#$, and the deformation gradient averages

$$\mathsf{F}_f(y) := \begin{cases} \fint_{\omega_f} \nabla y \, {\rm d}x, & \text{if } |\omega_f| > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Note that ω_f was defined in such a way that $\omega_f \subset \omega_{T_1} \cap \omega_{T_2}$. The value $\mathsf{F}_f(y) = 0$ for $f \subset \Omega_{\rm a}^\#$ is of no importance, and could have been replaced by any other value.

With this notation, and recalling the definitions of the edge midpoints q_f and the periodic nodal basis functions $\zeta_f^\#$ from Section 5.2.1, we can define

$$\hat{\psi}(y;\cdot) = \sum_{f\in\mathscr{F}_\varepsilon} \psi\big(\mathsf{F}_f(y); q_f\big)\zeta_f^\#. \tag{6.17}$$

Note, in particular, that $\hat{\psi}(y_{\mathsf{F}}) = \psi(\mathsf{F})$ for all $\mathsf{F} \in \mathbb{R}^{2\times 2}$. It is therefore natural to define the modified stress function

$$\widehat{\Sigma}_{\rm ac}(y; T) := \Sigma_{\rm ac}(y; T) - \nabla\hat{\psi}(y; T)\mathsf{J}, \qquad \text{for } T \in \mathscr{T}_\varepsilon. \tag{6.18}$$

In the following lemma we establish some elementary properties of $\widehat{\Sigma}_{\rm ac}$.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Lemma 6.10. *Suppose that $\mathscr{E}_{\rm ac}$ is energy and patch test consistent, that* $\mathrm{int}(\Omega_{\rm a})$ *is connected, and that $\mathscr{T}_h^{\rm a} \cup \mathscr{T}_h^{\rm i} \subset \mathscr{T}_\varepsilon$; then the modified a/c stress function $\widehat{\Sigma}_{\rm ac}$, defined in (6.18), has the following properties:*

$$\big\langle \delta\mathscr{E}_{\rm ac}(y), z\big\rangle = \int_\Omega \widehat{\Sigma}_{\rm ac}(y) : \nabla z \,\mathrm{d}x \qquad \forall\, y, z \in \mathscr{Y}; \tag{6.19}$$

$$\widehat{\Sigma}_{\rm ac}(y_{\mathsf{F}}; T) = \partial W(\mathsf{F}) \qquad \forall\, \mathsf{F} \in \mathbb{R}^{2\times 2}, \quad T \in \mathscr{T}_\varepsilon^{\rm c} \cup \mathscr{T}_\varepsilon^{\rm i}; \quad \textit{and} \tag{6.20}$$

$$\widehat{\Sigma}_{\rm ac}(y; T) = \Sigma_{\rm a}(y; T) \qquad \forall\, y \in \mathscr{Y}, \quad T \in \mathscr{T}_\varepsilon^{\rm a}. \tag{6.21}$$

Proof. To prove (6.19) let $z = y_{\mathsf{B}} + u$ for some $\mathsf{B} \in \mathbb{R}^{2\times 2}$ and $u \in \mathscr{U}$; then

$$\int_\Omega \widehat{\Sigma}_{\rm ac}(y) : \nabla z \,\mathrm{d}x = \langle \delta\mathscr{E}_{\rm ac}(y), z\rangle - \int_\Omega \big(\nabla\hat{\psi}\mathsf{J}\big) : \big(\mathsf{B} + \nabla u\big)\,\mathrm{d}x.$$

Since $\int_\Omega (\nabla\hat{\psi}\mathsf{J}) : \nabla u \,\mathrm{d}x = 0$ by Lemma 5.3, and since $\int_\Omega \nabla\hat{\psi}\,\mathrm{d}x = 0$ (see the proof of Lemma 6.9), the representation (6.19) follows.

Property (6.20) follows from Lemma 6.9 and the fact that $\hat{\psi}(y_{\mathsf{F}}) = \psi(\mathsf{F})$:

$$\widehat{\Sigma}_{\rm ac}(y_{\mathsf{F}}; T) = \Sigma_{\rm ac}(y_{\mathsf{F}}; T) - \nabla\psi(\mathsf{F}; T)\mathsf{J} = \partial W(\mathsf{F}) \quad \forall T \in \mathscr{T}_\varepsilon.$$

Property (6.21) follows from the fact that we constructed $\hat{\psi}(y)$ to be zero in $\Omega_{\rm a}$ for all $y \in \mathscr{Y}$, and from Proposition 6.5, which states that $\Sigma_{\rm ac}(y; T) = \Sigma_{\rm a}(y; T)$ for all $T \in \mathscr{T}_\varepsilon^{\rm a}$. □

6.3. The Lipschitz property

The final remaining ingredient for the proof of first-order consistency, is a Lipschitz property for $\widehat{\Sigma}_{\rm ac}$, similar to the Lipschitz property (6.11) of $\Sigma_{\rm a}$. In order to ensure that there are no modelling error contributions from the atomistic region it turns out to be most convenient to work directly with the stress difference

$$\mathsf{R}(y; T) := \widehat{\Sigma}_{\rm ac}(y; T) - \Sigma_{\rm a}(y; T). \tag{6.22}$$

From (6.21) we immediately obtain that

$$\mathsf{R}(y; T) = 0 \qquad \forall\, T \in \mathscr{T}_\varepsilon^{\rm a}. \tag{6.23}$$

In the remainder of the section we will estimate $\mathsf{R}(y; T)$ for $T \in \mathscr{T}_\varepsilon^{\rm c} \cup \mathscr{T}_\varepsilon^{\rm i}$. To motivate the following result we note that, from Lemma 6.4 and from (6.20) we see that

$$\mathsf{R}(y_{\mathsf{F}}; T) = 0 \qquad \forall\, \mathsf{F} \in \mathbb{R}^{2\times 2}, \quad T \in \mathscr{T}_\varepsilon^{\rm c} \cup \mathscr{T}_\varepsilon^{\rm i}; \tag{6.24}$$

hence, a suitable Lipschitz estimate for R yields the following result.

Lemma 6.11. *Suppose that all conditions of Lemma 6.10 hold and, in addition, that $\mathscr{E}_{\rm i}$ satisfies the locality and scaling conditions (2.19) and (2.20); then*

$$\big|\mathsf{R}(y; T)\big| \le \varepsilon M_T \,\mathrm{osc}(\nabla y; \omega_T) \qquad \forall y \in \mathscr{Y}, \quad T \in \mathscr{T}_\varepsilon, \tag{6.25}$$

where M_T is defined in (6.4).

The proof of this central lemma is split over the following paragraphs.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

6.3.1. Estimates in the continuum region

Let $T \in \mathscr{T}_\varepsilon^{\mathrm{c}}$, then

$$\begin{aligned}\big|\mathsf{R}(y;T)\big| &= \Big|\big[\Sigma_{\mathrm{ac}}(y;T) - \nabla\hat{\psi}(y;T)\mathsf{J}\big] - \Sigma_{\mathrm{a}}(y;T)\Big| \\ &\leq \big|\partial W(\nabla y) - \Sigma_{\mathrm{a}}(y;T)\big| + \big|\nabla\hat{\psi}(y;T)\mathsf{J}\big| \\ &\leq \varepsilon M^{\mathrm{a}}\mathrm{osc}\big(\nabla y;\omega_T\big) + \big|\nabla\hat{\psi}(y;T)\big|, \end{aligned} \tag{6.26}$$

where, in the last inequality, we used (6.11) and the fact that $\omega_T^{\mathrm{a}} \subset \omega_T$ for $T \in \mathscr{T}_\varepsilon^{\mathrm{c}}$. We still need to estimate $\nabla\hat{\psi}(y;T)$, which we postpone until Section 6.3.3.

6.3.2. Estimates in the interface region

Let $T \in \mathscr{T}_\varepsilon^{\mathrm{i}}$, and let $V_{x,r} = \partial_r V(D_{\mathscr{R}} y(x))$, then

$$\begin{aligned}\mathsf{R}(y;T) &= \Sigma_{\mathrm{ac}}(y;T) - \Sigma_{\mathrm{a}}(y;T) - \nabla\hat{\psi}(y;T)\mathsf{J} \\ &= \sum_{r\in\mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}_{\mathrm{a}}^{\#}} \Big[V_{x,r}\otimes r\Big] \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db} + \frac{\varepsilon^2}{|T|} \sum_{\substack{b\in\mathscr{B}_{\mathrm{i}}^{\#} \\ b=(x,x+\varepsilon r)}} \Big[\partial_b E_{\mathrm{i}}(y)\otimes r\Big] \fint_x^{x+\varepsilon r} \chi_T^{\mathrm{i}} \,\mathrm{db} \\ &\quad - \sum_{r\in\mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}^{\#}} \Big[V_{x,r}\otimes r\Big] \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db} - \nabla\hat{\psi}(y;T)\mathsf{J},\end{aligned}$$

which, after combining the first and third group, becomes

$$\begin{aligned}\mathsf{R}(y;T) &= \frac{\varepsilon^2}{|T|} \sum_{\substack{b\in\mathscr{B}_{\mathrm{i}}^{\#} \\ b=(x,x+\varepsilon r)}} \Big[\partial_b E_{\mathrm{i}}(y)\otimes r\Big] \fint_x^{x+\varepsilon r} \chi_T^{\mathrm{i}} \,\mathrm{db} - \sum_{r\in\mathscr{R}} \frac{\varepsilon^2}{|T|} \sum_{x\in\mathscr{L}^{\#}\setminus\mathscr{L}_{\mathrm{a}}^{\#}} \Big[V_{x,r}\otimes r\Big] \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db} - \nabla\hat{\psi}(y;T)\mathsf{J} \\ &=: \mathsf{R}^{(1)}(y;T) - \mathsf{R}^{(2)}(y;T) - \nabla\hat{\psi}(y;T)\mathsf{J}.\end{aligned}$$

We will again postpone the estimation of $\nabla\hat{\psi}(T)$ to Section 6.3.3, and focus on the terms $\mathsf{R}^{(1)}(y;T)$ and $\mathsf{R}^{(2)}(y;T)$.

Let $\mathsf{F} = \nabla y(T)$. Using the *locality and scaling conditions* (2.19) and (2.20), we can estimate

$$\begin{aligned}\big|\mathsf{R}^{(1)}(y;T) - \mathsf{R}^{(1)}(y_{\mathsf{F}};T)\big| &\leq \frac{\varepsilon^2}{|T|} \sum_{\substack{b\in\mathscr{B}_{\mathrm{i}}^{\#} \\ b=(x,x+\varepsilon r)}} \Big|\partial_b E_{\mathrm{i}}(y) - \partial_b E_{\mathrm{i}}(y_{\mathsf{F}})\Big| |r| \fint_x^{x+\varepsilon r} \chi_T^{\mathrm{i}} \,\mathrm{db} \\ &\leq \frac{\varepsilon^2}{|T|} \sum_{\substack{b\in\mathscr{B}_{\mathrm{i}}^{\#} \\ b=(x,x+\varepsilon r)}} \sum_{\substack{s\in\mathscr{R} \\ (x,x+\varepsilon s)\in\mathscr{B}_{\mathrm{i}}^{\#}}} M_{r,s}^{\mathrm{i}} \big|D_s y(x) - \mathsf{F}s\big| |r| \fint_x^{x+\varepsilon r} \chi_T \,\mathrm{db}. \end{aligned} \tag{6.27}$$

In the transition from the first to the second line we have used the fact that on bonds that lie on the boundary of $\Omega_{\mathrm{i}}^{\#}$ the constant $M_{r,s}^{\mathrm{i}}$ is replaced by $\frac{1}{2}M_{r,s}^{\mathrm{i}}$, which effectively replaces χ_T^{i} by χ_T. Bounding $|D_s y(x) - \mathsf{F}s|$ by the local oscillation, and applying the bond density lemma (see Lem. 6.4 for a similar calculation), we deduce that

$$\big|\mathsf{R}^{(1)}(y;T) - \mathsf{R}^{(1)}(y_{\mathsf{F}};T)\big| \leq \varepsilon \sum_{r\in\mathscr{R}} \sum_{s\in\mathscr{R}} |r||s| M_{r,s}^{\mathrm{i}} \,\mathrm{osc}(\nabla y;\omega_T) = \varepsilon M^{\mathrm{i}} \mathrm{osc}(\nabla y;\omega_T). \tag{6.28}$$

Following closely the proof of (6.11), we obtain a similar estimate for $\mathsf{R}^{(2)}$:

$$\big|\mathsf{R}^{(2)}(y;T) - \mathsf{R}^{(2)}(y_{\mathsf{F}};T)\big| \leq \varepsilon \sum_{r\in\mathscr{R}} \sum_{s\in\mathscr{R}} |r||s| M_{r,s}^{\mathrm{a}} \,\mathrm{osc}(\nabla y;\omega_T) = \varepsilon M^{\mathrm{a}} \mathrm{osc}(\nabla y;\omega_T). \tag{6.29}$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Note that it is enough to measure the oscillation over ω_T (which does not intersect with $\Omega_{\rm a}$), since $(x, x+\varepsilon s)\cap \Omega_{\rm a} = \emptyset$ for all $x \in \mathscr{L}^{\#} \setminus \mathscr{L}^{\#}_{\rm a}, s \in \mathscr{R}$.

Combining (6.28) and (6.29), and using the fact that $\mathsf{R}(y_{\mathsf{F}}; T) = 0$, we conclude that

$$\begin{aligned}\big|\mathsf{R}(y;T)\big| &= \big|\mathsf{R}(y;T) - \mathsf{R}(y_{\mathsf{F}};T)\big| \\ &\le \big|\mathsf{R}^{(1)}(y;T) - \mathsf{R}^{(1)}(y_{\mathsf{F}};T)\big| + \big|\mathsf{R}^{(2)}(y;T) - \mathsf{R}^{(2)}(y_{\mathsf{F}};T)\big| + \big|\nabla\hat{\psi}(y;T) - \nabla\hat{\psi}(y_{\mathsf{F}};T)\big| \\ &\le \varepsilon\big(M^{\rm i} + M^{\rm a}\big)\mathrm{osc}(\nabla y;\omega_T) + \big|\nabla\hat{\psi}(y;T) - \nabla\psi(\mathsf{F};T)\big|. \end{aligned} \tag{6.30}$$

6.3.3. Estimates on ψ and on $\hat{\psi}$

To finalize the estimates in Sections 6.3.1 and 6.3.2 we are left to establish a Lipschitz property for $\hat{\psi}$. The following result is a fundamental technical lemma that will allow us to achieve this. Its proof is deceptively simple, however, it uses implicitly many of the foregoing calculations. Moreover, some questions left open by Theorem 6.1 may be answered through a better understanding of this step.

Lemma 6.12. *Suppose that the conditions of Lemma 6.11 hold; then, for all $f \in \mathscr{F}^{\#}_{\varepsilon}$, $f \subset \Omega_{\rm c}\cup\Omega_{\rm i}$, and for all $\mathsf{F}, \mathsf{G} \in \mathbb{R}^{2\times 2}$, we have*

$$\big|\psi(\mathsf{F};q_f) - \psi(\mathsf{G};q_f)\big| \le \varepsilon\big(M_{\rm a} + M_{\rm i}\big)\mathrm{width}(\Omega_{\rm i})\big|\mathsf{F}-\mathsf{G}\big|,$$

where $\mathrm{width}(\Omega_{\rm i})$ *is defined in* (6.5).

Proof. Fix some $f' \in \mathscr{F}^{\#}_{\varepsilon}$, $f' \subset \Omega_{\rm a}$; then, for any connecting path $\gamma \in \Gamma_{f',f}$ we have

$$\begin{aligned}\psi(\mathsf{F};q_f) - \psi(\mathsf{G};q_f) &= \int_\gamma \Big(\nabla\psi(\mathsf{F}) - \nabla\psi(\mathsf{G})\Big)\cdot \mathrm{d}x \\ &= \int_\gamma \Big(\big[\partial W(\mathsf{F}) - \Sigma_{\rm ac}(y_{\mathsf{F}})\big] - \big[\partial W(\mathsf{G}) - \Sigma_{\rm ac}(y_{\mathsf{G}})\big]\Big)\cdot \mathrm{d}x.\end{aligned}$$

Since $\Sigma_{\rm ac}(y_{\mathsf{B}};T) = \partial W(\mathsf{B})$ for all $T \in \mathscr{T}^{\rm c}_{\varepsilon}\cup\mathscr{T}^{\rm a}_{\varepsilon}$, $\mathsf{B}\in\mathbb{R}^{2\times 2}$, the integrand vanishes in $\Omega_{\rm a}\cup\Omega_{\rm c}$. Hence, it follows that

$$\big|\psi(\mathsf{F};q_f) - \psi(\mathsf{G};q_f)\big| \le \mathrm{length}(\gamma\cap\Omega_{\rm i}) \max_{T\in\mathscr{T}^{\rm i}_{\varepsilon}} \Big|\big[\Sigma_{\rm ac}(y_{\mathsf{F}};T) - \Sigma_{\rm ac}(y_{\mathsf{G}};T)\big] - \big[\partial W(\mathsf{F}) - \partial W(\mathsf{G})\big]\Big|.$$

Following closely the calculations in Section 6.3.2 we can deduce that

$$\big|\psi(\mathsf{F};q_f) - \psi(\mathsf{G};q_f)\big| \le \mathrm{length}(\gamma\cap\Omega_{\rm i})\,\big(M_{\rm a}+M_{\rm i}\big)\,\big|\mathsf{F}-\mathsf{G}\big|.$$

Since we are free to choose the path γ, we can choose it so that $\mathrm{length}(\gamma\cap\Omega_{\rm i})$ is minimized, which yields the stated result. □

Let $T \in \mathscr{T}^{\rm i}_{\varepsilon}\cup\mathscr{T}^{\rm c}_{\varepsilon}$, let $\mathsf{F} = \nabla y(T)$, and recall that ζ_f are the Crouzeix–Raviart nodal basis functions associated with edge midpoints q_f; then, using Lemma 6.12 we obtain

$$\begin{aligned}\big|\nabla\hat{\psi}(y;T) - \nabla\psi(\mathsf{F};T)\big| &\le \sum_{\substack{f\in\mathscr{F}^{\#}_{\varepsilon}\\ f\subset\partial T}} \big|\psi(\mathsf{F}_f(y);q_f) - \psi(\mathsf{F};q_f)\big|\big|\nabla\zeta_f(T)\big| \\ &\le \big(M_{\rm a}+M_{\rm i}\big)\,\mathrm{width}(\Omega_{\rm i})\left\{\max_{\substack{f\in\mathscr{F}^{\#}_{\varepsilon}\\ f\subset\partial T}}\big|\mathsf{F}_f(y)-\mathsf{F}\big|\right\}\left\{\varepsilon\sum_{\substack{f\in\mathscr{F}^{\#}_{\varepsilon}\\ f\subset\partial T}}\big|\nabla\zeta_f(T)\big|\right\}.\end{aligned}$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

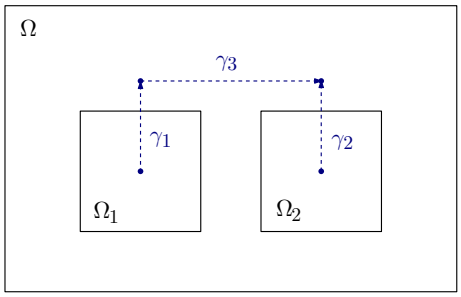


FIGURE 3. Atomistic region with two components to visualise the argument given in Section 6.4.2.

A direct calculation yields

$$\varepsilon \sum_{\substack{f\in\mathscr{F}_\varepsilon^{\#}\\ f\subset\partial T}} \left|\nabla\zeta_f(T)\right| = 2+2+2\sqrt{2} \le 7.$$

From the definitions of $\mathsf{F}_f(y)$ and F it follows that

$$\left|\nabla\hat{\psi}(y;T) - \nabla\psi(\mathsf{F};T)\right| \le \varepsilon 7\big(M_{\rm a}+M_{\rm i}\big)\mathrm{width}(\Omega_{\rm i})\,\mathrm{osc}(\nabla y;\omega_T) \quad \forall T\in\mathscr{T}_\varepsilon^{\rm i}\cup\mathscr{T}_\varepsilon^{\rm c}. \tag{6.31}$$

Proof of Lemma 6.11. Combining (6.31) with (6.26) and noting that $\nabla\psi(\mathsf{F}) = 0$ for $T\in\mathscr{T}_\varepsilon^{\rm c}$ (*cf.* (6.20) and the fact that $\Sigma_{\rm ac}(y_{\mathsf{F}}) = \partial W(\mathsf{F})$ in $\Omega_{\rm c}$), and also combining (6.31) with (6.30), we finally arrive at the result of Lemma 6.11. □

6.4. Remarks on the assumptions of Theorem 6.1

In this section, we construct simple examples to discuss the various assumptions of Theorem 6.1. We will show that most of them are also necessary.

6.4.1. Technical conditions

The assumption that $\mathscr{T}_h^{\rm a}\cup\mathscr{T}_h^{\rm i}\subset\mathscr{T}_\varepsilon$, and the assumption (2.14), were made for the sake of convenience of the analysis and simplicity of presentation. Dropping this assumption is not straightforward, but it is reasonable to expect that a careful analysis should allow to do so.

The same statement applies to the assumptions made on the interaction potential; this was already discussed in Section 2.1.3.

6.4.2. Connectedness of $\Omega_{\rm a}$

The assumption that $\mathrm{int}(\Omega_{\rm a})$ is connected is more problematic; at this point it is unclear whether or not it can be removed in general. A more detailed analysis of the functions $\psi(\mathsf{F},\cdot)$, $\mathsf{F}\in\mathbb{R}^{2\times 2}$, defined in Section 6.2.2 is required to understand this issue. There are, however, at least two special cases where one can attempt to remove it with relatively little effort:

- *Well separated components:* if $\Omega_{\rm a}$ has several connected components, which are separated by an $O(1)$ distance, then one can localize the consistency error estimate to each of the components and obtain a qualitatively similar result as Theorem 6.1.
- *Specific a/c methods:* suppose that the GCC method described in Section 2.3.4 is used to construct a patch test consistent coupling scheme, with parameters $C_{x,r,s}$. Suppose, moreover, that $\Omega_{\rm a}$ has two connected components, Ω_1 and Ω_2, each of which have a portion of the boundary with the same orientation (say, normal e_1), as displayed in Figure 3.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

It is then reasonable to assume that the parameters $C_{x,r,s}$ have the same value in those parts of the interface surrounding Ω_1 and Ω_2, which would imply that

$$\int_{\gamma_1} \Sigma_{\rm ac}(y_{\mathsf{F}}) \cdot \mathrm{d}x = \int_{\gamma_2} \Sigma_{\rm ac}(y_{\mathsf{F}}) \cdot \mathrm{d}x.$$

Moreover, since $\Sigma_{\rm ac}(y_{\mathsf{F}}) = \partial W(\mathsf{F})$ in the continuum region, we would obtain that

$$\int_{\gamma_1 \cup \gamma_3 \cup \gamma_2'} \big(\Sigma_{\rm ac}(y_{\mathsf{F}}) - \partial W(\mathsf{F})\big) \cdot \mathrm{d}x = 0,$$

where γ_2' denotes the curve γ_2 with reversed orientation.
This shows that it is possible to choose $\psi(\mathsf{F};\cdot) = 0$ in both components of $\Omega_{\rm a}$, and as a consequence, Theorem 6.1 would remain true.

A related issue is the dependence of the modelling error estimate (6.3) on $\mathrm{width}(\Omega_{\rm i})$, which comes solely from the Lipschitz estimate on $\mathsf{F} \mapsto \psi(\mathsf{F};\cdot)$; *cf.* Section 6.3.3. Hence, a better understanding of this function may also allow a finer analysis of this undesirable dependence.

6.4.3. The global energy consistency condition

Global energy consistency is a natural and convenient condition that yields the important intermediate result (Cor. 6.7) that

$$\fint_{\Omega} \Sigma_{\rm ac}(y_{\mathsf{F}})\,\mathrm{d}x = \partial W(\mathsf{F}) \qquad \forall\, \mathsf{F} \in \mathbb{R}^{2\times 2}. \tag{6.32}$$

Note also that (6.32) implies $\mathscr{E}_{\rm ac}(y_{\mathsf{F}}) = \mathscr{E}_{\rm a}(y_{\mathsf{F}}) + c$ for all $\mathsf{F} \in \mathbb{R}^{2\times 2}$, where c is a fixed constant that is independent of F; that is, (6.32) is practically equivalent to global energy consistency.

In some important situations patch test consistency already implies (6.32). The following result gives such a result for finite atomistic regions.

Proposition 6.13. *Suppose that $\Omega_{\rm a} \cup \Omega_{\rm i} \subset \mathrm{int}(\Omega)$; then patch test consistency (2.12) of $\mathscr{E}_{\rm ac}$ implies (6.32).*

Proof. According to Lemmas 6.5 and 6.9 we have

$$\big\langle \delta\mathscr{E}_{\rm ac}(y_{\mathsf{F}}), u\big\rangle = \int_{\Omega_{\rm c}} \partial W(\mathsf{F}) : \nabla u\,\mathrm{d}x + \int_{\Omega_{\rm a}\cup\Omega_{\rm i}} \Sigma_{\rm ac}(y_{\mathsf{F}}) : \nabla u\,\mathrm{d}x.$$

Let $\mathsf{G} \in \mathbb{R}^{2\times 2}$ and choose any $u \in \mathscr{U}$ such that $\nabla u = \mathsf{G}$ in $\Omega_{\rm i} \cup \Omega_{\rm a}$; this is possible due to the assumption that $\Omega_{\rm a} \cup \Omega_{\rm i} \subset \mathrm{int}(\Omega)$. Integrating by parts twice, letting ν denote the unit outward normal to $\Omega_{\rm i} \cup \Omega_{\rm a}$, and noting that the portions of the surface integrals along $\partial\Omega$ cancel each other out, yields

$$\begin{aligned}
\big\langle \delta\mathscr{E}_{\rm ac}(y_{\mathsf{F}}), u\big\rangle &= \int_{\Omega_{\rm c}} \partial W(\mathsf{F}) : \nabla u\,\mathrm{d}x + \int_{\Omega_{\rm a}\cup\Omega_{\rm i}} \Sigma_{\rm ac}(y_{\mathsf{F}}) : \mathsf{G}\,\mathrm{d}x \\
&= -\int_{\partial(\Omega_{\rm a}\cup\Omega_{\rm i})} \partial W(\mathsf{F}) : (u\otimes\nu)\,\mathrm{d}s + \int_{\Omega_{\rm a}\cup\Omega_{\rm i}} \Sigma_{\rm ac}(y_{\mathsf{F}}) : \mathsf{G}\,\mathrm{d}s \\
&= -\int_{\Omega_{\rm a}\cup\Omega_{\rm i}} \partial W(\mathsf{F}) : \nabla u\,\mathrm{d}x + \int_{\Omega_{\rm a}\cup\Omega_{\rm i}} \Sigma_{\rm ac}(y_{\mathsf{F}}) : \mathsf{G}\,\mathrm{d}s \\
&= \int_{\Omega_{\rm a}\cup\Omega_{\rm i}} \big[\Sigma_{\rm ac}(y_{\mathsf{F}}) - \partial W(\mathsf{F})\big] : \mathsf{G}\,\mathrm{d}x.
\end{aligned}$$

Since $\mathscr{E}_{\rm ac}$ is patch test consistent, the last term vanishes, and hence the result follows. □

It turned out to be difficult to devise a counterexample, which clearly demonstrates that absence of (6.32) can yield an inconsistent method. A more thorough investigation of this condition is still required.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

6.4.4. The locality condition

To show that the locality condition (2.19) (or a variant thereof) is necessary we assume, without loss of generality, that N is even and define a functional $\mathscr{J} \in \mathrm{C}^2(\mathscr{Y})$, $\mathscr{J} = \varepsilon^2 J$,

$$J(y) = \left| \sum_{x \in \mathscr{L}^{\mathrm{i}}_{-,+}} D_{e_1} y(x) \right|^2 + \left| \sum_{x \in \mathscr{L}^{\mathrm{i}}_{+,+}} D_{e_1} y(x) \right|^2 - \left| \sum_{x \in \mathscr{L}^{\mathrm{i}}_{-,-}} D_{e_1} y(x) \right|^2 - \left| \sum_{x \in \mathscr{L}^{\mathrm{i}}_{+,-}} D_{e_1} y(x) \right|^2,$$

where,

$$\mathscr{L}^{\mathrm{i}}_{-,\pm} = \big\{ x \in \mathscr{L} \,\big|\, x_1 \leq 0, x_2 = \pm 1/2 \big\}, \quad \text{and}$$
$$\mathscr{L}^{\mathrm{i}}_{+,\pm} = \big\{ x \in \mathscr{L} \,\big|\, x_1 > 0, x_2 = \pm 1/2 \big\}.$$

From the definition it is obvious that $J(y_{\mathsf{F}}) = 0$ for all $\mathsf{F} \in \mathbb{R}^{2\times 2}$. Moreover, using summation by parts along the two lines $\mathscr{L}^{\mathrm{i}}_{-,+} \cup \mathscr{L}^{\mathrm{i}}_{+,+}$ and $\mathscr{L}^{\mathrm{i}}_{-,-} \cup \mathscr{L}^{\mathrm{i}}_{+,-}$, it is easy to check that the patch test (2.12) holds. Finally, $\mathscr{J}$ satisfies the scaling condition,

$$\partial_{(x,x+\varepsilon r)} \partial_{(x',x'+\varepsilon r)} J(y) = \begin{cases} 2\,\mathsf{I}, & \text{if } r = e_1 \text{ and } x, x' \in \mathscr{L}^{\mathrm{i}}_{a,b}, \quad a,b \in \{+,-\}, \\ 0, & \text{otherwise.} \end{cases}$$

However, $\mathscr{J}$ clearly violates the locality condition.

We may think of $\mathscr{J}$ as an a/c functional for $\mathscr{E}_{\mathrm{a}} = 0$, or, alternatively, as an additional contribution that can be added to any a/c functional whose interface satisfies $\mathscr{L}^{\mathrm{i}}_{-,+} \cup \mathscr{L}^{\mathrm{i}}_{-,-} \cup \mathscr{L}^{\mathrm{i}}_{+,+} \cup \mathscr{L}^{\mathrm{i}}_{+,-} \subset \mathrm{int}(\Omega^{\#}_{\mathrm{i}})$.

Let $y \in \mathscr{Y}_{\mathsf{A}}$ be "smooth" but not affine in the upper half plane $\{x \in \mathscr{L} \,|\, x_2 \geq 0\}$, and let $y = y_{\mathsf{A}}$ on the lower interface $\mathscr{L}^{\mathrm{i}}_{-,-} \cup \mathscr{L}^{\mathrm{i}}_{+,-}$; then, testing $\delta \mathscr{J}(y)$ with the unique displacement $u \in \mathscr{U}$ such that

$$\begin{aligned} u(x) &= y(x) - \mathsf{A}x && \text{for } x \in \{(0,1/2),(1,1/2)\}; \\ x_1 &\mapsto u(x_1, 1/2) && \text{is affine in } [-1,0] \text{ and in } [0,1]; \text{ and} \\ D_{e_2} u(x) &= 0 && \text{for all } x \in \mathscr{L}; \end{aligned}$$

then we obtain, after a brief computation,

$$\big\|\delta \mathscr{J}(y)\big\|_{\mathrm{W}_{\varepsilon}^{-1,2}} \geq \frac{\langle \delta \mathscr{J}(y), u \rangle}{\|\nabla u\|_{\mathrm{L}^2(\Omega)}} = \left(\left| \varepsilon \sum_{x \in \mathscr{L}^{\mathrm{i}}_{-,+}} \big(D_{e_1} y(x) - \mathsf{A} e_1 \big) \right|^2 + \left| \varepsilon \sum_{x \in \mathscr{L}^{\mathrm{i}}_{+,+}} \big(D_{e_1} y(x) - \mathsf{A} e_1 \big) \right|^2 \right)^{1/2}.$$

This final estimate is scaled like a surface integral and is clearly a zeroth order term if y is smooth but $y(0,1/2) \neq y(1,1/2)$. This shows that the locality condition (2.19) (or a variant thereof) is indeed necessary to obtain a first-order consistency estimate.

6.4.5. The scaling condition

It is fairly clear that the modelling error estimate can be arbitrarily large without the scaling condition (2.20). We nevertheless briefly discuss a simple example with a natural interpretation.

Using a similar argument as in the previous paragraph, we define a functional $\mathscr{J} \in \mathrm{C}^2(\mathscr{Y})$, $\mathscr{J} = \varepsilon^2 J$,

$$J(y) = \beta \sum_{x \in \mathscr{L}^{\mathrm{i}}_{+}} \big| D_{e_1} y(x) \big|^2 - \beta \sum_{x \in \mathscr{L}^{\mathrm{i}}_{-}} \big| D_{e_1} y(x) \big|^2,$$

where $\beta > 0$ is a constant, and where

$$\mathscr{L}^{\mathrm{i}}_{\pm} = \big\{ x \in \mathscr{L} \,\big|\, x_2 = \pm 1/2 \big\}.$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

It is easy to see that J is patch test consistent, that $J(y_{\mathsf{F}}) = 0$ for all $\mathsf{F} \in \mathbb{R}^{2\times 2}$, and that it satisfies the locality condition.

Let $y \in \mathscr{Y}$ such that $y(x) = \mathsf{A}x$ for $x_2 = -1/2$, then testing $\delta\mathscr{J}(y)$ with the unique displacement $u \in \mathscr{U}$ such that

$$\begin{aligned} u(x) &= y(x) - \mathsf{A}x, \quad \text{for } x \in \mathscr{L}^{\mathrm{i}}_+, \\ D_{e_2}u(x) &= 0, \quad \text{for all } x \in \mathscr{L}, \end{aligned}$$

we obtain

$$\left\|\delta\mathscr{J}(y)\right\|_{\mathrm{W}_\varepsilon^{-1,2}} \geq \frac{\left\langle \delta\mathscr{J}(y), u\right\rangle}{\|\nabla u\|_{\mathrm{L}^2}} = \beta\varepsilon\left[\varepsilon \sum_{x\in\mathscr{L}^{\mathrm{i}}_+} \left|D_{e_1}y(x) - \mathsf{A}e_1\right|^2\right]^{1/2}. \tag{6.33}$$

If y is "smooth" but not affine, then the term in square brackets is of the order $O(1)$. By choosing β arbitrarily large, the modelling error can be made arbitrarily large as well. In particular, the choice $\beta = 1/\varepsilon$ would give a seemingly natural surface scaling to the interface functional, and in this case we would obtain an $O(1)$ modelling error.

7. Conclusion

A fairly complete consistency analysis of general patch test consistent a/c coupling methods in (one and) two space dimensions was developed in this paper. The main result is the first order modelling error estimate, Theorem 6.1. An undesirable assumption in this result is the connectedness of $\mathrm{int}(\Omega_{\mathrm{a}})$. To remove this assumption a finer analysis of the corrector functions $\psi(\mathsf{F}, \cdot)$ defined in Section 6.2.2 is required. At this point one cannot exclude the possibility that a/c methods exist for which this assumption is in fact necessary.

Many open problems remain to be answered. First and foremost, one ought to answer the question whether a/c methods satisfying all the conditions of Theorem 6.1 always exist (that is, for general a/c interface geometries). In [44], we present a general construction (a variant on the geometrically consistent coupling method [18]), in which we compute the geometry reconstruction parameters numerically in a preprocessing step. This process appears to work in practise, however, at this point we have no proof that the linear system that is solved to that end actually has a solution. Indeed, if it should turn out that in certain cases the "ghost forces" cannot be completely eliminated, then an extension of Theorem 6.1 estimating the contribution of the "ghost force" to the modelling error is highly desirable since such a result would provide the correct quantity that needs to be minimized. It is by no means clear that minimizing the "ghost force" itself is the best possible target. A similar analysis would also be useful for estimating the modelling error of blending methods [54].

It should be conceptually straightforward, though technically more demanding, to generalize all results to higher order finite element methods in the continuum region, however, it would then also be desirable to obtain the second order modelling error estimate in the continuum region. Such a result seems difficult to obtain without a more detailed understanding of the corrector functions $\psi(\mathsf{F}, \cdot)$.

An immediate question is whether a variant of the main result is still valid in 3D. This is by no means clear at this point. From a technical point of view, we require generalizations of the two main technical tools: the bond density lemma (Sect. 5.1) and the characterisation of discrete divergence-free P_0-tensor fields (Sect. 5.2). While the bond density lemma as stated in this paper is false in 3D, one can establish variants that are potentially useful for a 3D analysis (work in progress). Generalising the explicit construction of Section 5.2 is entirely open at this point.

Another important and difficult question is the extension to multi-lattices where the Cauchy–Born model is obtained through a homogenization procedure [1,13].

Finally, it should be stressed, that Theorem 6.1 is a general abstract result, and as such can undoubtedly be improved upon when a specific coupling method is analyzed. It may be possible for specific methods to obtain more information about the corrector functions $\psi(\mathsf{F}, \cdot)$, and hence obtain a better estimate on the dependence

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

of the modelling error on the interface width. For example, the consistency analysis in [41] requires no corrector functions at all, and in the consistency analysis of nearest-neighbour interactions [44] the corrector function vanishes in the continuum region. The proof of Theorem 6.1 may, however, serve as a general guidance for modelling error estimates in specific cases.

Finally, the stability of a/c methods in 2D/3D is largely open at this point.

Acknowledgements. I thank B. Langwallner, X.H. Li, M. Luskin, E. Süli, A. Shapeev and L. Zhang for their comments on a draft of this manuscript, which have greatly helped to improve its quality. E. Süli pointed out to me the literature on the patch test. Early sketches of some of the technical results presented in Section 3 were developed in discussion with A. Shapeev during our work on [41]. The representation of discrete divergence-free P_0-tensor fields discussed in Section 5.2 was pointed out to me by L. Zhang. We have used a variant in our explicit construction of consistent a/c methods in [45].

Appendix A. Proofs of Section 3

Proof of Lemma 3.1. For $d = 1$, since $I_\varepsilon y_h = y_h$, the result is trivial; hence assume that $d = 2$. Assume also that $p < \infty$. Since the norms involved are effectively weighted ℓ^p-norms, one can obtain the case $p = \infty$ as the limit $p \nearrow \infty$.

In this proof we will in fact use a periodic version of (5.1), which is a simple consequence of (5.1) (see also [41]):

$$|T| = \varepsilon^2 \sum_{x\in\mathscr{L}} \fint_x^{x+\varepsilon r} \chi_{T\#}\,\mathrm{db}.$$

With our definition of the L^p-norms for matrix-valued functions, we have

$$\big\|\nabla I_\varepsilon y_h\big\|^p_{\mathrm{L}^p(\Omega)} = \sum_{j=1}^2 \int_\Omega \big|\nabla I_\varepsilon y_h e_j\big|_p^p\,\mathrm{d}x = \sum_{j=1}^2 \int_\Omega \big|\nabla_{e_j} I_\varepsilon y_h\big|_p^p\,\mathrm{d}x.$$

Using the periodic bond density lemma, and the fact that $\{\chi_{T\#} \,|\, T \in \mathscr{T}_\varepsilon\}$ is a partition of unity for $\mathbb{R}^2$, we have

$$\begin{aligned}
\int_\Omega \big|\nabla_{e_j} I_\varepsilon y_h\big|_p^p\,\mathrm{d}x &= \sum_{T\in\mathscr{T}_\varepsilon} |T|\big|\nabla_{e_j} I_\varepsilon y_h(T)\big|_p^p \\
&= \sum_{T\in\mathscr{T}_\varepsilon} \big|\nabla_{e_j} I_\varepsilon y_h(T)\big|_p^p \varepsilon^2 \sum_{x\in\mathscr{L}} \fint_x^{x+\varepsilon e_j} \chi_{T\#}\,\mathrm{db} \\
&= \varepsilon^2 \sum_{x\in\mathscr{L}} \sum_{T\in\mathscr{T}_\varepsilon} \fint_x^{x+\varepsilon e_j} \big|\nabla_{e_j} I_\varepsilon y_h\big|_p^p \chi_{T\#}\,\mathrm{db} \\
&= \varepsilon^2 \sum_{x\in\mathscr{L}} \fint_x^{x+\varepsilon e_j} \big|\nabla_{e_j} I_\varepsilon y_h\big|_p^p\,\mathrm{db}. \qquad \text{(A.1)}
\end{aligned}$$

We have also used the fact that $\nabla_{e_j} I_\varepsilon y_h$ is continuous across edges that have direction e_j.

Due to the specific choice of the triangulation $\mathscr{T}_\varepsilon$ it follows that $\nabla_{e_j} I_\varepsilon y_h$ is constant along each bond $(x, x+\varepsilon e_j)$, and hence

$$\fint_x^{x+\varepsilon e_j} \big|\nabla_{e_j} I_\varepsilon y_h\big|_p^p\,\mathrm{db} = \big|D_{e_j} I_\varepsilon y_h(x)\big|_p^p = \left|\fint_x^{x+\varepsilon e_j} \nabla_{e_j} y_h\,\mathrm{db}\right|_p^p \le \fint_x^{x+\varepsilon e_j} \big|\nabla_{e_j} y_h\big|_p^p\,\mathrm{db},$$

where we employed Jensen's inequality in the last step.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

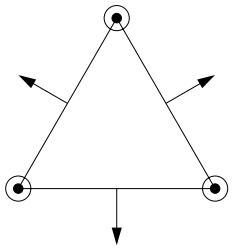


FIGURE 4. Illustration of the degrees of freedom in the C^1-conforming Clough–Tocher element: black dots denote point values, circles denote gradient values, arrows denote directional derivatives.

Inserting this estimate into (A.1), and reversing the argument in (A.1), we arrive at

$$\begin{aligned}\int_\Omega \left|\nabla_{e_j} I_\varepsilon y_h\right|_p^p \mathrm{d}x &\leq \varepsilon^2 \sum_{x\in\mathscr{L}} \fint_x^{x+\varepsilon e_j} \left|\nabla_{e_j} y_h\right|_p^p \mathrm{db} \\ &= \sum_{T\in\mathscr{T}_h} \varepsilon^2 \sum_{x\in\mathscr{L}} \fint_x^{x+\varepsilon e_j} \left|\nabla_{e_j} y_h\right|_p^p \chi_{T^\#} \mathrm{db} \\ &= \sum_{T\in\mathscr{T}_h} |T| \left|\nabla_{e_j} y_h(T)\right|_p^p = \left\|\nabla_{e_j} y_h\right\|_{\mathrm{L}^p(\Omega)}^p. \end{aligned}$$

□

Remark A.1. From the foregoing proof, it follows that

$$\left\|\nabla y\right\|_{\mathrm{L}^p(\Omega)} = \left(\varepsilon^2 \sum_{j=1}^{2} \sum_{x\in\mathscr{L}} \left|D_{e_j} y(x)\right|_p^p\right)^{1/p} \qquad \text{for } y \in \mathscr{Y}.$$

Proof of Lemma 3.2. To simplify the notation, we define the scalar function $z = y_i$ for some fixed i. Moreover, we prove the result only for $d = 2$; for $d = 1$ the result follows from the interpolation error estimates established in [42].

Step 1. $\mathrm{W}^{2,\infty}$*-interpolant:* we first define a $\mathrm{W}^{2,\infty}$-interpolant $\tilde{z}$ of z, using the C^1-conforming Hsieh–Clough–Tocher (HCT) element [8]; see Figure 4.

Let $T \in \mathscr{T}_\varepsilon^{\mathrm{c}}$ and let Q_T denote the set of vertices of T, and F_T the set of edges of T.

For each vertex $q \in Q_T$, we define the point value $\tilde{z}(q) = z(q)$, and the gradient value by $\nabla\tilde{z}(q) = \fint_{\omega_q^{\mathrm{c}}} \nabla z \,\mathrm{d}x$, where

$$\omega_q^{\mathrm{c}} = \bigcup \left\{T' \in (\mathscr{T}_\varepsilon^{\mathrm{c}})^\# \,\middle|\, q \in T'\right\}.$$

Similarly, for each edge $f \in F_T$, $f = T \cap T'$, with midpoint q_f, we define the patch $\omega_f^{\mathrm{c}} = \Omega_{\mathrm{c}}^\# \cap (T \cup T')$, and the directional derivative $\nabla_\nu\tilde{z}(q_f) = \fint_{\omega_f^{\mathrm{c}}} \nabla_\nu z \,\mathrm{d}x$.

Let φ_q be the nodal basis function associated with the point value at a vertex q, ϕ_f the nodal basis function associated with the normal derivative at an edge f, and let $\Phi_{q,\alpha}$ be the nodal basis function associated with the α-component of the derivative $\nabla z(q)$.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Step 2. Estimating $z - \tilde{z}$: fix $T \in \mathscr{T}_\varepsilon^{\mathrm{c}}$, $x \in T$, and define $\mathsf{F} = \nabla z(T)$, then we have

$$\begin{aligned}\big|\nabla z(T) - \nabla \tilde{z}(x)\big| &= \big|\mathsf{F} - \nabla \tilde{z}(x)\big| \\ &\leq \left|\sum_{f \in F_T} \big(\mathsf{F}\cdot \nu_f - \nabla_{\nu_f} \tilde{z}(q_f)\big) \otimes \nabla \varphi_f(x)\right| + \left|\sum_{\substack{q \in Q_T \\ \alpha \in \{1,2\}}} \big(\mathsf{F}_\alpha - \partial_{x_\alpha} \tilde{z}(q)\big) \otimes \nabla \Phi_{q,\alpha}\right|.\end{aligned}$$

Since all elements $T \in \mathscr{T}_\varepsilon$ are translated, scaled, and possibly reflected, copies of the reference triangle $\hat{T} = \mathrm{conv}\{(0,0),(1,0),(0,1)\}$, it follows that the HCT nodal basis functions are given (up to translations and reflections) by

$$\varphi_f(x) = \varepsilon \hat{\varphi}_f\big(\varepsilon^{-1} x\big), \quad \text{and} \quad \hat{\Phi}_{q,\alpha}(x) = \varepsilon \Phi_{q,\alpha}\big(\varepsilon^{-1} x\big).$$

Note in particular, that the gradients of these nodal basis functions are scale invariant, that is,

$$\|\nabla \varphi_f\|_{\mathrm{L}^\infty} \leq C \quad \text{and} \quad \|\nabla \Phi_{q,\alpha}\|_{\mathrm{L}^\infty} \leq C,$$

where C is a fixed constant that is independent of ε.

From the construction of $\tilde{z}$ it is easy to see that, for $f \in F_q, q \in Q_T, \alpha \in \{1,2\}$,

$$\big|\mathsf{F}\nu_f - \nabla_{\nu_f} \tilde{z}(q_f)\big| \leq \varepsilon\, \mathrm{osc}(\nabla z; \omega_T^{\mathrm{c}}), \quad \text{and} \quad \big|\mathsf{F}_\alpha - \partial_{x_\alpha} \tilde{z}(q)\big| \leq \varepsilon\, \mathrm{osc}(\nabla z; \omega_T^{\mathrm{c}});$$

and hence we obtain

$$\big\|\nabla z(T) - \nabla \tilde{z}\big\|_{\mathrm{L}^p(T)} \leq C_1 \varepsilon\, |T|^{1/p}\, \mathrm{osc}(\nabla z; \omega_T^{\mathrm{c}}), \tag{A.2}$$

for some generic constant C_1.

Step 3. Interpolation error: using standard interpolation error estimates [8], we obtain

$$\big\|\nabla \tilde{z} - \nabla I_h \tilde{z}\big\|_{\mathrm{L}^p(\Omega_{\mathrm{c}})} \leq C_I' \big\|h \nabla^2 \tilde{z}\big\|_{\mathrm{L}^p(\Omega_{\mathrm{c}})}.$$

Let $T \in \mathscr{T}_\varepsilon$ and $\mathsf{F} = \nabla z(T)$, then application of an inverse inequality, and (A.2) yield

$$\big\|\nabla^2 \tilde{z}\big\|_{\mathrm{L}^p(T)} = \big\|\nabla^2(\tilde{z} - z)\big\|_{\mathrm{L}^p(T)} \leq C_2 \varepsilon^{-1} \big\|\nabla \tilde{z} - \mathsf{F}\big\|_{\mathrm{L}^p(T)} \leq C_2\, |T|^{1/p}\, \mathrm{osc}(\nabla z; \omega_T^{\mathrm{c}}). \tag{A.3}$$

Finally, since $\tilde{z}(x) = z(x)$ for all $x \in \mathscr{L}^\#$, it follows that $I_h z = I_h \tilde{z}$, and hence we can estimate

$$\begin{aligned}\big\|\nabla z - \nabla I_h z\big\|_{\mathrm{L}^p(\Omega)} &= \big\|\nabla z - \nabla I_h z\big\|_{\mathrm{L}^p(\Omega_{\mathrm{c}})} \\ &\leq \big\|\nabla z - \nabla \tilde{z}\big\|_{\mathrm{L}^p(\Omega_{\mathrm{c}})} + \big\|\nabla \tilde{z} - \nabla I_h \tilde{z}\big\|_{\mathrm{L}^p(\Omega_{\mathrm{c}})}.\end{aligned}$$

Employing (A.2) and (A.3), we obtain the stated result. □

Proof of Lemma 3.4. For each $f \in \mathscr{F}_\varepsilon$, let $f = T_- \cap T_+$, $T_\pm \in \mathscr{T}_\varepsilon$, let $\nu_\pm$ denote the corresponding unit outward normals, and $\omega_f' = T_+ \cup T_-$.

We integrate by parts in each element $T \in \mathscr{T}_\varepsilon$ and use the fact that $I_\varepsilon u_h = u_h$ in $\Omega_{\mathrm{i}} \cup \Omega_{\mathrm{a}}$ to obtain

$$\begin{aligned}\big|\langle \Phi, I_\varepsilon u_h\rangle - \langle \Phi_h, u_h\rangle\big| &= \left|\sum_{T \in \mathscr{T}_\varepsilon} \int_T \sigma(T) : \nabla\big(I_\varepsilon u_h - u_h\big)\, \mathrm{d}x\right| \\ &= \left|\sum_{\substack{f \in \mathscr{F}_\varepsilon \\ f \not\subset \Omega_{\mathrm{i}} \cup \Omega_{\mathrm{a}}}} \int_f \big(\sigma(T_+)\nu_+ + \sigma(T_-)\nu_-\big) \cdot \big(I_\varepsilon u_h - u_h\big)\, \mathrm{d}s\right| \\ &\leq \sum_{\substack{f \in \mathscr{F}_\varepsilon \\ f \not\subset \Omega_{\mathrm{i}} \cup \Omega_{\mathrm{a}}}} \varepsilon \mathrm{osc}(\sigma; \omega_f') \int_f \big|I_\varepsilon u_h - u_h\big|\, \mathrm{d}s.\end{aligned} \tag{A.4}$$

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

Let $v := I_\varepsilon u_h - u_h$. An application of [39], Lemma 6.6, yields the trace inequality

$$\|v\|_{\mathrm{L}^1(f)} \leq \varepsilon^{-1}\|v\|_{\mathrm{L}^1(\omega'_f)} + \|\nabla v\|_{\mathrm{L}^1(\omega'_f)}. \tag{A.5}$$

Furthermore, since v is Lipschitz continuous and $v(p) = 0$ on every vertex of the triangulation $\mathscr{T}_\varepsilon$, we can use [39], Lemma 6.8, to deduce that

$$\|v\|_{\mathrm{L}^1(\omega'_f)} \leq \sqrt{2}\varepsilon\|\nabla v\|_{\mathrm{L}^1(\omega'_f)}. \tag{A.6}$$

Combining (A.6), (A.5), and (A.4), applying two Hölder inequalities, and estimating the overlaps between the patches ω'_f, we deduce that

$$\begin{aligned}\left|\langle \Phi, I_\varepsilon u_h\rangle - \langle \Phi_h, u_h\rangle\right| &\leq (1+\sqrt{2})\varepsilon \sum_{\substack{f\in\mathscr{F}_\varepsilon\\ f\not\subset\Omega_{\mathrm{i}}\cup\Omega_{\mathrm{a}}}} \mathrm{osc}(\sigma;\omega'_f)|\omega'_f|^{1/p}\big\|\nabla v\big\|_{\mathrm{L}^{p'}(\omega'_f)}\\ &\leq C_1\varepsilon\left(\sum_{T\in\mathscr{T}_h^{\mathrm{c}}} |T|\mathrm{osc}(\sigma;\omega_T^{\mathrm{c}})^p\right)^{1/p}\big\|\nabla v\big\|_{\mathrm{L}^{p'}(\Omega_{\mathrm{c}})}.\end{aligned}$$

An application of Lemma 3.1 yields the stated result. □

Appendix B. List of symbols

$a\cdot b$, $a\otimes b$	vector dot product and tensor product; Section 1.3
$\lvert\cdot\rvert, \lvert\cdot\rvert_p$	ℓ^p-norms; Section 1.3
$\lVert\cdot\rVert_{\ell^p_\varepsilon}$	weighted ℓ^p-norms; Section 1.3
$\mathscr{L}, \mathscr{L}^\#$	Lattice and lattice domain; Section 2.1.1
y_{A}	homogeneous deformation; Section 2.1.1
$\mathscr{U}, \mathscr{Y}, \mathscr{Y}_{\mathsf{A}}$	spaces of periodic displacements and deformations, Section 2.1.1
$A^\#, \mathscr{A}^\#$	periodic extension of a set or family of sets, Section 2.1.1
$\mathscr{R}$	interaction range, Section 2.1.2
$D_r, D_{\mathscr{R}}$	finite difference operator and stencil, Section 2.1.2
∇_r, ∇	directional derivative, deformation or displacement gradient, Section 1.3
∂v	Jacobi matrix of vector valued function, Section 1.3
$\mathscr{E}_{\mathrm{a}}$	atomistic energy, Section 2.1.2
V	atomistic interaction potential, Section 2.1.2
$\mathscr{P}_{\mathrm{a}}$	external potential in atomistic model, Section 2.1.4
$\delta\mathscr{E}$, $\delta^2\mathscr{E}$, $\langle\cdot,\cdot\rangle$	first and second variations, abstract duality pairing, Section 2.1.4
$\partial_r V, \partial_{r,s} V$	first and second partial derivatives of V, Section 2.1.3
$M^{\mathrm{a}}_{r,s}, M^{\mathrm{a}}$	bounds on $\partial_{r,s}V$ and Lipschitz constant for $\delta\mathscr{E}_{\mathrm{a}}$, Section 2.1.3
$\mathscr{T}_h, \mathscr{T}_h^\#, \mathscr{F}_h, \mathscr{F}_h^\#$	triangulations, and edge sets, Section 2.3.1
$h_T, h(x)$	mesh size functions, Section 2.3.1
$\mathrm{P}_1, \mathrm{P}_0, \mathrm{P}_1^\#, \mathrm{P}_0^\#$	finite element spaces, Section 2.3.1
$\mathscr{U}_h, \mathscr{Y}_h, \mathscr{Y}_{\mathsf{A},h}$	finite element spaces, Section 2.3.1
I_h	nodal interpolation operator for $\mathrm{P}_1(\mathscr{T}_h)$, Section 2.3.1
W, ∂W	Cauchy–Born stored energy function and stress, Section 2.3.2
int	interior of a set; Section 2.3.4
conv	convex hull of a set; Section 2.3.4
$\mathscr{E}_{\mathrm{ac}}$, $\mathscr{P}_{\mathrm{ac}}$	a/c energy and external potential, Section 2.3.5
$\Omega_{\mathrm{a}}, \Omega_{\mathrm{c}}, \Omega_{\mathrm{i}}$	atomistic, continuum, and interface region, Section 2.3.5
$\mathscr{T}_h^{\mathrm{a}}, \mathscr{T}_h^{\mathrm{c}}, \mathscr{T}_h^{\mathrm{i}}$	atomistic, continuum, and interface triangulations, Section 2.3.5

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

$\mathscr{L}_{\rm a}$	set of atomistic sites in a/c method, Section 2.3.5
$(x, x'), (x, x+\varepsilon r)$	bonds, Section 2.3.5
$\mathscr{B}_{\rm i}$, $\mathscr{B}_{\rm i}^{\#}$	set of interface bonds, Section 2.3.5
$\mathscr{E}_{\rm i}$, $E_{\rm i}$	interface functional, Section 2.3.5
$\partial_b E_{\rm i}, \partial_{(x,x+\varepsilon r)} E_{\rm i}$	scaled first partial derivatives of $E_{\rm i}$, Section 2.3.6
$M^{\rm i}_{r,s}$, $M^{\rm i}$	bounds on second partial derivatives of $E_{\rm i}$, Section 2.3.6
$\mathscr{T}_\varepsilon, \mathscr{T}_\varepsilon^{\#}, \mathscr{F}_\varepsilon, \mathscr{F}_\varepsilon^{\#}$	atomistic triangulation and edge sets, Section 3.1
osc	oscillation operator, Section 3.1
I_ε	nodal interpolation operator for $\mathrm{P}_1(\mathscr{T}_\varepsilon)$, Section 3.1
$\omega_T^{\rm c}$	patch used in the interpolation error estimate, Section 3.2.2
h_T	modified mesh size function, Section 3.2.2
$\|\cdot\|_{\mathrm{W}_h^{-1,p}}$	$\mathrm{W}^{1,p'}$-dual norm on $\mathrm{P}_1(\mathscr{T}_h)^*$, Section 3.4
$\|\cdot\|_{\mathrm{W}_\varepsilon^{-1,p}}$	$\mathrm{W}^{1,p'}$-dual norm on $\mathrm{P}_1(\mathscr{T}_\varepsilon)^*$, Section 3.5
ϕ_1, ϕ_2	first and second neighbour potential, Section 4
$x_n, v_n, v_n', v_n'', v_n'''$	notation for 1D grid functions, Section 4
χ_T	characteristic function used in bond density lemma, Section 5.1
$\chi_T^{\rm i}$	characteristic function used to define $\Sigma_{\rm ac}$, Proposition 6.5, page 1303
$\fint_x^{x'} f\,\mathrm{db}, \fint_x^{x+\varepsilon r} f\,\mathrm{db}$	bond integrals, Section 5.1
N_1, $\mathrm{N}_1^{\#}$	Crouzeix–Raviart finite element spaces, Section 5.2.1
q_f, ζ_f, $\zeta_f^{\#}$	midpoint of an edge f and associated nodal basis, Section 5.2.1
$\int_\gamma \sigma \cdot \mathrm{d}x$	path integral, Section 5.2.2
J	rotation about $\pi/2$, Lemma 5.3, page 1297
$\omega_T^{\rm a}, \omega_T$	atomistic interaction neighbourhoods, Section 6
M_T	prefactors in modelling error estimate, equation (6.4), page 1299
$\mathrm{width}(\Omega_{\rm i})$	width of $\Omega_{\rm i}$, equation (6.5), page 1299
$\Sigma_{\rm a}$	atomistic stress function, equation (6.8), page 1301
$V_{x,r}, V_{\mathsf{F},r}$	alternative notation for $\partial_r V(D_{\mathscr{R}} y(x))$ and for $\partial_r V(\mathsf{F}\mathscr{R})$
$\Sigma_{\rm ac}$	a/c stress function, Proposition 6.5, page 1303
$\psi(\mathsf{F};\cdot)$	corrector function for $\Sigma_{\rm ac}(y_{\mathsf{F}})$, Lemma 6.9, page 1304
$\hat{\psi}(y;\cdot)$	corrector function for $\Sigma_{\rm ac}(y)$, equation (6.17), page 1305
$\widehat{\Sigma}_{\rm ac}$	modified a/c stress function, equation (6.18), page 1305
$\mathsf{R}(y;T)$	stress error, equation (6.22), page 1306

References

[1] A. Abdulle, P. Lin and A. Shapeev, Homogenization-based analysis of quasicontinuum method for complex crystals. arXiv:1006.0378.

[2] N.C. Admal and E.B. Tadmor, A unified interpretation of stress in molecular systems. *J. Elasticity* **100** (2010) 63–143.

[3] R. Alicandro and M. Cicalese, A general integral representation result for continuum limits of discrete energies with superlinear growth. *SIAM J. Math. Anal.* **36** (2004) 1–37 (electronic).

[4] D.N. Arnold and R.S. Falk, A uniformly accurate finite element method for the Reissner-Mindlin plate. *SIAM J. Numer. Anal.* **26** (1989) 1276–1290.

[5] S. Badia, M. Parks, P. Bochev, M. Gunzburger and R. Lehoucq, On atomistic-to-continuum coupling by blending. *Multiscale Model. Simul.* **7** (2008) 381–406.

[6] G.P. Bazeley, Y.K. Cheung, B.M. Irons and O.C. Zienkiewicz, Triangle elements in plate bending: conforming and nonconforming solutions, in *Proc. Conf. Matrix Meth. Struc. Mech.* Wright Patterson AFB, Ohio (1966).

[7] T. Belytschko, W.K. Liu and B. Moran, *Nonlinear finite elements for continua and structures*. John Wiley & Sons Ltd., Chichester (2000).

[8] P.G. Ciarlet, The finite element method for elliptic problems. *Classics in Appl. Math.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA **40** (2002). Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[9] M. Dobson, There is no pointwise consistent quasicontinuum energy. arXiv:1109.1897.

[10] M. Dobson and M. Luskin, Analysis of a force-based quasicontinuum approximation. *ESAIM: M2AN* **42** (2008) 113–139.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

[11] M. Dobson and M. Luskin, An analysis of the effect of ghost force oscillation on quasicontinuum error. *ESAIM: M2AN* **43** (2009) 591–604.
[12] M. Dobson and M. Luskin, An optimal order error analysis of the one-dimensional quasicontinuum approximation. *SIAM J. Numer. Anal.* **47** (2009) 2455–2475.
[13] M. Dobson, R. Elliot, M. Luskin and E. Tadmor, A multilattice quasicontinuum for phase transforming materials: cascading cauchy born kinematics. *J. Computer-Aided Mater. Design* **14** (2007) 219–237.
[14] M. Dobson, M. Luskin and C. Ortner, Accuracy of quasicontinuum approximations near instabilities. *J. Mech. Phys. Solids* **58** (2010) 1741–1757.
[15] M. Dobson, M. Luskin and C. Ortner, Stability, instability, and error of the force-based quasicontinuum approximation. *Arch. Rational Mech. Anal.* **197** (2010) 179–202.
[16] W. E and P. Ming, Analysis of the local quasicontinuum method, in Frontiers and prospects of contemporary applied mathematics. *Ser. Contemp. Appl. Math. CAM* **6** (2005) 18–32.
[17] W. E and P. Ming, Cauchy-Born rule and the stability of crystalline solids: static problems. *Arch. Rational Mech. Anal.* **183** (2007) 241–297.
[18] W. E, J. Lu and J.Z. Yang, Uniform accuracy of the quasicontinuum method. *Phys. Rev. B* **74** (2006) 214115.
[19] B. Eidel and A. Stukowski, A variational formulation of the quasicontinuum method based on energy sampling in clusters. *J. Mech. Phys. Solids* **57** (2009) 87–108.
[20] M. Finnis, Interatomic Forces in Condensed Matter. *Oxford Series on Materials Modelling* **1** (2003).
[21] J. Fish, M.A. Nuggehally, M.S. Shephard, C.R. Picu, S. Badia, M.L. Parks, and M. Gunzburger, Concurrent AtC coupling based on a blend of the continuum stress and the atomistic force. *Comput. Methods Appl. Mech. Eng.* **196** (2007) 4548–4560.
[22] M. Gunzburger and Y. Zhang, A quadrature-rule type approximation to the quasi-continuum method. *Multiscale Model. Simul.* **8** (2009/2010) 571–590.
[23] M. Iyer and V. Gavini, A field theoretical approach to the quasi-continuum method. *J. Mech. Phys. Solids* **59** (2011) 1506–1535.
[24] P.A. Klein and J.A. Zimmerman, Coupled atomistic-continuum simulations using arbitrary overlapping domains. *J. Comput. Phys.* **213** (2006) 86–116.
[25] J. Knap and M. Ortiz, An analysis of the quasicontinuum method. *J. Mech. Phys. Solids* **49** (2001) 1899–1923.
[26] S. Kohlhoff and S. Schmauder, A new method for coupled elastic-atomistic modelling, in *Atomistic Simulation of Materials: Beyond Pair Potentials*, edited by V. Vitek and D.J. Srolovitz. Plenum Press, New York (1989) 411–418.
[27] X.H. Li and M. Luskin, An analysis of the quasi-nonlocal quasicontinuum approximation of the embedded atom model. To appear in *Int. J. Multiscale Comput. Eng.,* `arXiv:1008.3628.`
[28] X.H. Li and M. Luskin, A generalized quasi-nonlocal atomistic-to-continuum coupling method with finite range interaction. To appear in *IMA J. Numer. Anal.*, `arXiv:1007.2336.`
[29] P. Lin, Theoretical and numerical analysis for the quasi-continuum approximation of a material particle model. *Math. Comp.* **72** (2003) 657–675.
[30] P. Lin, Convergence analysis of a quasi-continuum approximation for a two-dimensional material without defects. *SIAM J. Numer. Anal.* **45** (2007) 313–332 (electronic).
[31] J. Lu and P. Ming, Convergence of a force-based hybrid method for atomistic and continuum models in three dimension. `arXiv:1102.2523.`
[32] M. Luskin and C. Ortner, An analysis of node-based cluster summation rules in the quasicontinuum method. *SIAM J. Numer. Anal.* **47** (2009) 3070–3086.
[33] 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) 83–121.
[34] C. Makridakis, C. Ortner and E. Süli, Stress-based atomistic/continuum coupling: a new variant of the quasicontinuum approximation. *Int. J. Multiscale Comput. Eng.* forthcoming.
[35] R.E. Miller and E.B. Tadmor, The quasicontinuum method: overview, applications and current directions. *J. Computer-Aided Mater. Design* **9** (2003) 203–239.
[36] R.E. Miller and E.B. Tadmor, A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. *Model. Simul. Mater. Sci. Eng.* **17** (2009).
[37] P. Ming and J.Z. Yang, Analysis of a one-dimensional nonlocal quasi-continuum method. *Multiscale Model. Simul.* **7** (2009) 1838–1875.
[38] M. Ortiz, R. Phillips and E.B. Tadmor, Quasicontinuum analysis of defects in solids. *Philos. Mag. A* **73** (1996) 1529–1563.
[39] C. Ortner, *Analysis of the Quasicontinuum Method*. Ph.D. thesis, University of Oxford (2006).
[40] C. Ortner, *A priori* and *a posteriori* analysis of the quasinonlocal quasicontinuum method in 1D. *Math. Comp.* **80** (2011) 1265–1285.
[41] C. Ortner and A.V. Shapeev, Analysis of an energy-based atomistic/continuum coupling approximation of a vacancy in the 2d triangular lattice. To appear in *Math. Comp.*, `arXiv1104.0311`.
[42] C. Ortner and E. Süli, Analysis of a quasicontinuum method in one dimension. *ESAIM: M2AN* **42** (2008) 57–91.
[43] C. Ortner and H. Wang, *A priori* error estimates for energy-based quasicontinuum approximations of a periodic chain. *Math. Models Methods Appl. Sci.* **21** (2011) 2491–2521.
[44] C. Ortner and L. Zhang, work in progress.
[45] C. Ortner and L. Zhang, Construction and sharp consistency estimates for atomistic/continuum coupling methods with general interfaces: a 2d model problem. `arXiv:1110.0168`.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press

[46] M.L. Parks, P.B. Bochev and R.B. Lehoucq, Connecting atomistic-to-continuum coupling and domain decomposition. *Multiscale Model. Simul.* **7** (2008) 362–380.

[47] D. Pettifor, *Bonding and structure of molecules and solids*. Oxford University Press (1995).

[48] K. Polthier and E. Preuß, Identifying vector field singularities using a discrete Hodge decomposition, in *Visualization and mathematics III*, *Math. Vis.* Springer, Berlin (2003) 113–134.

[49] A.V. Shapeev, Consistent energy-based atomistic/continuum coupling for two-body potentials in one and two dimensions. *Multiscale Model. Simul.* **9** (2011) 905–932.

[50] 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** (1999) 611–642.

[51] L.E. Shilkrot, R.E. Miller and W.A. Curtin, Coupled atomistic and discrete dislocation plasticity. *Phys. Rev. Lett.* **89** (2002) 025501.

[52] 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** (2004) 214104.

[53] G. Strang and G. Fix, *An Analysis of the Finite Element Method*. Wellesley-Cambridge Press (2008).

[54] B. Van Koten and M. Luskin, Development and analysis of blended quasicontinuum approximations. To appear in *SIAM J. Numer. Anal.*, `arXiv:1008.2138.`

[55] B. Van Koten, Z.H. Li, M. Luskin and C. Ortner, A computational and theoretical investigation of the accuracy of quasicontinuum methods, in Numerical Analysis of Multiscale Problems, edited by I. Graham, T. Hou, O. Lakkis and R. Scheichl. *Springer Lect. Notes Comput. Sci. Eng.* **83** (2012).

[56] S.P. Xiao and T. Belytschko, A bridging domain method for coupling continua with molecular dynamics. *Comput. Methods Appl. Mech. Eng.* **193** (2004) 1645–1669.

https://doi.org/10.1051/m2an/2012005 Published online by Cambridge University Press