



Available online at www.sciencedirect.com

ScienceDirect

Comput. Methods Appl. Mech. Engrg. 312 (2016) 276–303

**Computer methods
in applied
mechanics and
engineering**

www.elsevier.com/locate/cma

A line search assisted monolithic approach for phase-field computing of brittle fracture

T. Gerasimov*, L. De Lorenzis

Institute of Applied Mechanics, Technische Universität Braunschweig, Bienroder Weg 87, 38106 Braunschweig, Germany

Available online 29 December 2015

Abstract

Phase-field modeling of fracture phenomena in solids is a very promising approach which has gained popularity within the last decade. However, within the finite element framework, already a two-dimensional quasi-static phase-field formulation is computationally quite demanding, mainly for the following reasons: (i) the need to resolve the small length scale inherent to the diffusive crack approximation calls for extremely fine meshes, at least locally in the crack phase-field transition zone, (ii) due to non-convexity of the related free-energy functional, a robust, but slowly converging *staggered* solution scheme based on algorithmic decoupling is typically used. In this contribution we tackle problem (ii) and propose a faster and equally accurate approach for quasi-static phase-field computing of (brittle) fracture using a *monolithic* solution scheme which is accompanied by a novel *line search* procedure to overcome the iterative convergence issues of non-convex minimization. We present a detailed critical evaluation of the approach and its comparison with the staggered scheme in terms of computational cost, accuracy and robustness.

© 2015 Elsevier B.V. All rights reserved.

Keywords: Phase-field; Fracture; Staggered; Monolithic; Line search

1. Introduction

The phase-field formulation of quasi-static brittle fracture is a non-linear problem with a strong coupling between the displacement field \mathbf{u} and the crack phase-field d to be solved for. At every given loading step, it stems from the minimization problem of a free-energy functional which is typically non-convex in (\mathbf{u}, d) . A widely adopted form of this functional reads

$$\mathcal{E}(\mathbf{u}, d) = \int_{\Omega} \left[(1-d)^2 \Psi_0^+(\boldsymbol{\varepsilon}(\mathbf{u})) + \Psi_0^-(\boldsymbol{\varepsilon}(\mathbf{u})) \right] \mathrm{d}\mathbf{x} + G_c \int_{\Omega} \left(\frac{1}{2\ell} d^2 + \frac{\ell}{2} |\nabla d|^2 \right) \mathrm{d}\mathbf{x}, \quad (1)$$

see e.g. the review paper [1]. In (1), the limiting values of d , namely, $d = 0$ and $d = 1$ represent the undamaged and fully broken material phases, respectively, Ψ_0^+ and Ψ_0^- are the so-called ‘tensile’ and ‘compressive’ parts of an additive decomposition of the elastic energy density function $\Psi_0 := \frac{1}{2} \boldsymbol{\varepsilon} : \mathbb{C} : \boldsymbol{\varepsilon} = \frac{1}{2} \lambda \text{tr}^2(\boldsymbol{\varepsilon}) + \mu \text{tr}(\boldsymbol{\varepsilon}^2)$, where, in turn, $\boldsymbol{\varepsilon}$ is

* Corresponding author.

E-mail address: t.gerasimov@tu-braunschweig.de (T. Gerasimov).

the second-order infinitesimal strain tensor, \mathbb{C} is the fourth-order elasticity tensor, and λ and μ are the Lamé constants. The decomposition of Ψ_0 is required in order to distinguish between fracture behavior in tension and compression, more precisely, to avoid crack growth and crack faces interpenetration in compression, and its various options are available in the literature, see [1] for an overview and [2–4] for the original ones. Finally, in (1), $0 < \ell \ll 1$ is the parameter that controls the width of the transition zone of d between the two states and G_c is the material fracture toughness.

In this contribution, we are concerned with devising an efficient computational scheme for solving the weak formulation for (1). As a starting point, we address the two possible options.

The first one is based on the idea of decoupling the weak formulation into the system of two equations which is then solved in a staggered manner: with one variable being frozen, one solves for the second one, which, in turn, is to be frozen to find an update for the first one, etc. The approach is therefore called *staggered*. The staggered solution scheme is the most commonly used, see [1–4], due to its robustness: the energy functional \mathcal{E} in (1) is convex with respect to each of its arguments \mathbf{u} and d separately. The phase-field formulation in [5–8] stemming from the Francfort and Marigo variational approach to fracture [9] and using an energy functional slightly different from (1) is solved with the alternate minimization procedure and backtracking algorithm, which is closely related to the staggered scheme. In [10], this formulation is computed with the staggered approach directly. The higher-order phase-field formulation from [11] is also solved in a staggered manner. Unfortunately, in related publications the question of efficiency of the staggered scheme is not explicitly addressed. We will show that the scheme may perform quite inefficiently: at a fixed loading step a significant amount of staggered iterations is typically required to achieve energy convergence, thus resulting in a high computational cost.

In this context, the alternative scheme implying the solution of the weak formulation for (1) for both variables simultaneously (more precisely, their increments) seems more promising. Such approach is called *monolithic*. Provided that iterative convergence is achieved, the monolithic solution scheme can be shown to be faster than the staggered one. However, due to non-convexity of the functional \mathcal{E} in (1) with respect to (\mathbf{u}, d) , the Newton–Raphson iterative process typically diverges, particularly in case of the so-called brutal crack evolution and in the post-peak regime of loading. To the best of our knowledge, only a few results related to some extent to the fully monolithic treatment of the problem at hand have been published so far, namely, [12] and [13]. In [12], the weak equation is augmented according to a dissipation-based arc-length procedure [14] and the resulting system is computed monolithically. Unfortunately, no discussion is provided on iterative convergence issues associated with the non-convexity of \mathcal{E} in this case. In [13], the latter issues are acknowledged and tackled through a ‘convexification’ trick, which implies the replacement of the unknown d in the elastic part of the derivative \mathcal{E}' by a fixed known extrapolation \tilde{d} and application of the monolithic scheme to the already modified equation. As mentioned by the authors, ‘this treatment is shown in subsequent examples to be numerically robust’, yet ‘the theoretical validity of the extrapolation remains an open question’.

In this contribution, we develop a procedure to restore iterative convergence of the fully monolithic scheme. This is achieved by implanting a new version of the so-called *line search* into the Newton–Raphson method. In the mechanics literature, this combination is sometimes called ‘damped Newton–Raphson method’ and is typically used in non-linear problems to facilitate and speed up convergence of the standard Newton–Raphson procedure, or of its variants (modified Newton’s method, quasi-Newton method, etc.), see e.g. [15–19]. In phase-field computing of fracture, we propose to activate our line search procedure when divergence of the Newton–Raphson iterative scheme is detected. We, moreover, allow for a negative search direction, which is not done in the standard procedure, yet will be argued to be necessary and feasible herein. The resulting approach is termed ‘line search assisted monolithic scheme’. It will be shown that the proposed line search overcomes iterative convergence issues and saves a significant amount of computational time in comparison with the staggered approach.

The paper is structured as follows. Section 2 introduces explicitly the weak formulation stemming from the minimization problem for \mathcal{E} in (1) equipped with the appropriate irreversibility condition for the fracture phase-field, as well as the benchmark problem chosen for testing the existing and newly proposed solution approaches. Section 3 is devoted to the staggered scheme and to the detailed evaluation of its efficiency. In Section 4, we first address the original monolithic scheme and illustrate the related convergence issues which appear already at early stages of the modeled crack propagation. Our line search procedure is then presented and a critical comparison of the results obtained with the two approaches follows. Conclusions and outlook finalize the paper.

Throughout the paper we use the abbreviations STAG, MON and MON-*ls* to denote, respectively, the staggered, the monolithic and the line search assisted monolithic approaches.

2. Phase-field formulation and benchmark problem

In this section we introduce the adopted phase-field formulation and the benchmark problem to be dissected.

2.1. Formulation

Let $\Omega \in \mathbb{R}^n$, $n = 2, 3$, be an open bounded domain which represents the configuration of an n -dimensional linearly elastic isotropic body in the absence of applied forces. Its boundary $\partial\Omega$ is decomposed, in general, into four non-overlapping sets such that $\partial\Omega = \Gamma_{D,0} \cup \Gamma_{D,1} \cup \Gamma_{N,0} \cup \Gamma_{N,1}$ with $\Gamma_{D,\cdot}$ and $\Gamma_{N,\cdot}$ as the portions where Dirichlet and Neumann boundary conditions are imposed, respectively. The second subscript of Γ indicates that the related condition is either homogeneous (0) or inhomogeneous (1). Let also $l = 0, 1, \dots$ be a loading (or, pseudo-time) step. We set

$$\mathbf{V}_0 := \{\mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v} = \mathbf{0} \text{ on } \Gamma_{D,0} \cup \Gamma_{D,1}\},$$

and

$$\mathbf{V}_1^l := \{\mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v} = \mathbf{0} \text{ on } \Gamma_{D,0}, \mathbf{v} = \bar{\mathbf{u}}^l \text{ on } \Gamma_{D,1}\},$$

with $\bar{\mathbf{u}}^l$ being prescribed, as the test and admissible displacement spaces of the problem, respectively. Herein, $\mathbf{H}^1(\Omega) := [H^1(\Omega)]^n$ with H^1 denoting the usual Sobolev space. Without loss of generality, we furthermore assume the absence of body forces and surface tractions on the rest of the boundary (this implies that $\Gamma_{N,0} = \partial\Omega/\Gamma_D$ and $\Gamma_{N,1} = \emptyset$).

According to a quasi-static approximation [9], the minimization of the functional \mathcal{E} in (1) at the loading step l consists of finding the pair (\mathbf{u}, d) such that

$$(\mathbf{u}, d) \in \underset{\substack{\mathbf{u} \in \mathbf{V}_1^l \\ d \in H^1(\Omega, [0,1]), d_{l-1} \leq d}}{\operatorname{argmin}} \mathcal{E}(\mathbf{u}, d),$$

where the condition $d_{l-1} \leq d$ enforces the irreversibility of the crack phase-field. Since this condition cannot be easily implemented, an alternative criterion was introduced by Bourdin in [6], where the irreversibility is given via an equality constraint. Thus, if at a loading step l the set

$$\text{CR}_{l-1} := \{\mathbf{x} \in \overline{\Omega} : d_{l-1} > 1 - \text{CRTOL}\},$$

where CRTOL is a small specified parameter, is non-empty, the irreversibility is enforced by setting $d := 1$ for all $\mathbf{x} \in \text{CR}_{l-1}$. The modified minimization problem for \mathcal{E} then reads: find

$$(\mathbf{u}, d) \in \underset{\substack{\mathbf{u} \in \mathbf{V}_1^l \\ d \in H^1(\Omega, [0,1]), d|_{\text{CR}_{l-1}} = 1}}{\operatorname{argmin}} \mathcal{E}(\mathbf{u}, d). \quad (2)$$

Note that $d|_{\text{CR}_{l-1}} = 1$ may be viewed as a Dirichlet boundary condition for d . Owing to this, the constraint in (2) can be relaxed by a penalization term, as proposed in [20], leading to the following modification of the original \mathcal{E} :

$$\begin{aligned} E(\mathbf{u}, d) = & \int_{\Omega} \left[(1-d)^2 \Psi_0^+(\boldsymbol{\varepsilon}(\mathbf{u})) + \Psi_0^-(\boldsymbol{\varepsilon}(\mathbf{u})) \right] \mathrm{d}\mathbf{x} + G_c \int_{\Omega} \left(\frac{1}{2\ell} d^2 + \frac{\ell}{2} |\nabla d|^2 \right) \mathrm{d}\mathbf{x} \\ & + \frac{1}{2\gamma} \int_{\text{CR}_{l-1}} (1-d)^2 \mathrm{d}\mathbf{x}, \end{aligned} \quad (3)$$

where γ is a (small) penalty constant.

Henceforth, we will be dealing with the functional E defined by (3) and the related (unconstrained) minimization problem: find

$$(\mathbf{u}, d) \in \underset{\substack{\mathbf{u} \in \mathbf{V}_1^l \\ d \in H^1(\Omega, [0,1])}}{\operatorname{argmin}} E(\mathbf{u}, d). \quad (4)$$

In Eqs. (3) we take

$$\Psi_0^\pm(\boldsymbol{\varepsilon}) := \frac{1}{2}\lambda\langle\text{tr}(\boldsymbol{\varepsilon})\rangle_\pm^2 + \mu\text{tr}(\boldsymbol{\varepsilon}_\pm^2),$$

which is the so-called ‘strain spectral decomposition’-based split of Ψ_0 , as proposed in [3,4]. Herein, $\boldsymbol{\varepsilon}_\pm := \sum_{I=1}^3 \langle\varepsilon_I\rangle_\pm \mathbf{n}_I \otimes \mathbf{n}_I$ with $\{\varepsilon_I\}_{I=1}^3$ and $\{\mathbf{n}_I\}_{I=1}^3$ as the principal strains and principal strain directions, respectively, and $\langle a \rangle_\pm := \frac{1}{2}(a \pm |a|)$.

With E as in (3), using the notion of the directional derivative of E ,

$$\begin{aligned} E'(\mathbf{u}, d; \mathbf{v}, w) &= \int_{\Omega} \left[(1-d)^2 \frac{\partial \Psi_0^+}{\partial \varepsilon}(\boldsymbol{\varepsilon}(\mathbf{u})) + \frac{\partial \Psi_0^-}{\partial \varepsilon}(\boldsymbol{\varepsilon}(\mathbf{u})) \right] : \boldsymbol{\varepsilon}(\mathbf{v}) \, dx \\ &\quad - \int_{\Omega} 2(1-d) \Psi_0^+(\boldsymbol{\varepsilon}(\mathbf{u})) w \, dx + G_c \int_{\Omega} \left(\frac{1}{\ell} dw + \ell \nabla d \cdot \nabla w \right) \, dx - \frac{1}{\gamma} \int_{\text{CR}_{l-1}} (1-d) w \, dx, \end{aligned} \quad (5)$$

problem (4) turns into the corresponding weak formulation: given a fixed $\bar{\mathbf{u}}^l$ on $\Gamma_{D,1}$, find a pair $(\mathbf{u}, d) \in \mathbf{V}_1^l \times H^1(\Omega, [0, 1])$ such that

$$E'(\mathbf{u}, d; \mathbf{v}, w) = 0, \quad \text{for all } (\mathbf{v}, w) \in \mathbf{V}_0 \times H^1(\Omega, [0, 1]). \quad (6)$$

Assuming more regularity for (\mathbf{u}, d) , we may integrate by parts in (6) and arrive at the strong formulation of the problem. This will reveal, in particular, the natural boundary condition for d on $\partial\Omega$, namely, $\frac{\partial d}{\partial \mathbf{n}} = 0$ on $\partial\Omega$, where \mathbf{n} is the outward unit normal vector to $\partial\Omega$. We do not report this formulation explicitly, since it is not used in the following.

Remark 1. In the constrained minimization problem (2) for \mathcal{E} , the condition $d|_{\text{CR}_{l-1}} = 1$ enforces irreversibility of only a fully developed crack (the phase-field pattern with $d = 1$), which is in a conceptual agreement with the original variational formulation of fracture by Francfort and Marigo [9]. Irreversibility of the phase-field patterns with $d < 1$, which from the mechanical standpoint may be viewed as a micro-damage formation or micro-cracking, is not accounted for. The same holds for the induced unconstrained minimization problem (4) for E we are working with in this paper. Implementation of a more general condition to account for irreversibility of $d < 1$ fields is feasible, yet it is out of our scope here.

Remark 2. In Miehe et al. [3,4], the irreversibility of the phase-field evolution stemming from the $\mathcal{E} \rightarrow \min$ problem is handled using the notion of a so-called history variable \mathcal{H} . This represents the maximal tensile energy Ψ_0^+ accumulated within loading history and enters the related weak equation $\mathcal{E}' = 0$ via replacement of Ψ_0^+ with \mathcal{H} . This, however, results in some inconsistency between the original energy functional \mathcal{E} and the corresponding modified weak equation.

2.2. Benchmark problem and computational data

The two-dimensional problem we consider throughout this paper is the so-called *shear test* for a square specimen with a preexisting crack, as sketched in the left plot of Fig. 1. It is a popular benchmark example in the literature on phase-field modeling of brittle fracture, see e.g. [1,4,11–13], since the problem setup is very simple, yet the failure pattern is not symmetric (the crack propagates towards the lower right corner of the specimen, see Fig. 2), being the result of a non-trivial combination and interaction of local tension–compression and loading–unloading processes within the specimen during shear. We find it a good candidate for the purposes we pursue in our study.

The specimen geometry, material and loading data are taken as in [4] and are as follows: $a = 1$ mm, $E = 210$ GPa, $\nu = 0.3$, $G_c = 2700$ N/m. We also set the length-scale parameter as $\ell = \frac{1}{100}a$. Plane-strain state is assumed. A monotonically increasing displacement loading $\bar{\mathbf{u}} = (\bar{u}, 0)$ is prescribed on the upper specimen edge, whereas the bottom edge is fixed. Finally, the parameters CRTOL and γ entering (3) are chosen to be 0.05 and 10^{-16} , respectively.

The finite element simulation setup is as follows. We employ the numerical package FreeFem++ [21]. Both the displacement field \mathbf{u} and phase-field d are approximated using P_1 -triangles. As the failure pattern is known a priori (from the discrete crack approach, see Fig. 2, as well as from earlier phase-field studies), our finite element mesh is

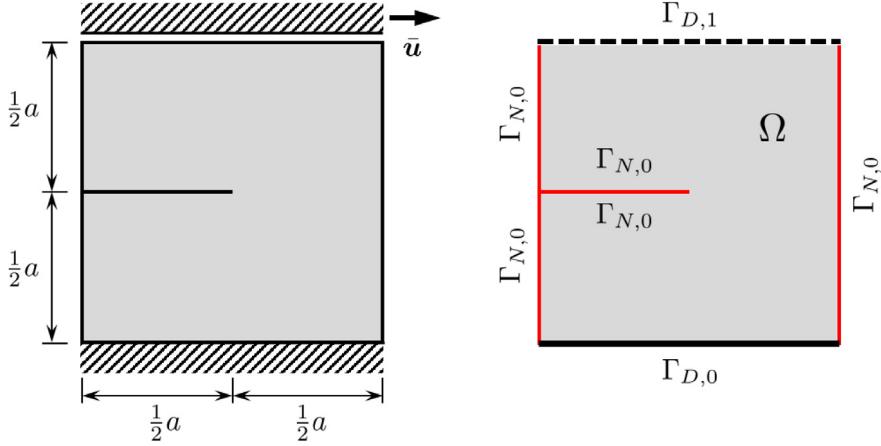


Fig. 1. Specimen geometry and loading conditions (left); notations adopted in Section 2.1 (right).

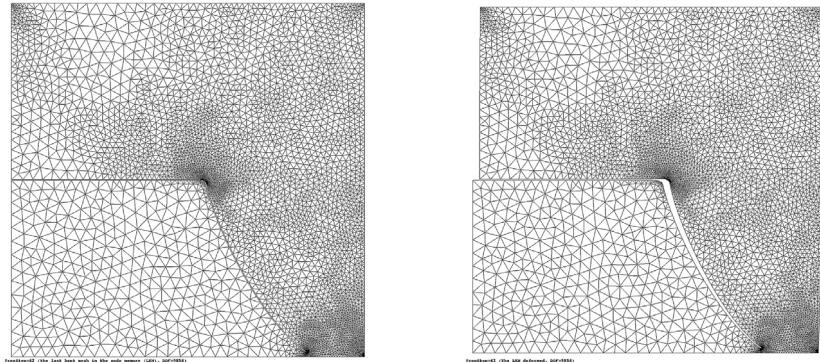


Fig. 2. Crack propagation path obtained with a discrete crack framework and error-controlled adaptive finite element computing (left); the corresponding deformed configuration (right).

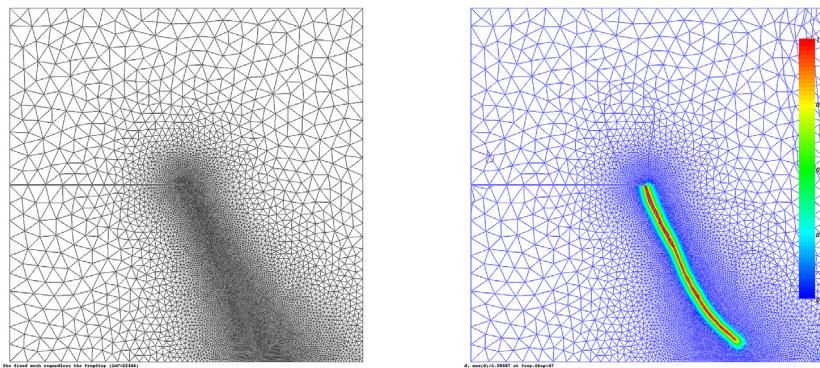


Fig. 3. Pre-adapted mesh for the phase-field computations (left); crack propagation path that simulated with the phase-field approach (right).

pre-adapted (refined) in the region where the phase-field evolution is expected, see the left plot in Fig. 3. The mesh contains 11,223 nodes. Importantly, the mesh size h in the localization zone (i.e. in the support) of d fulfills the property $h < \frac{\ell}{2}$, providing in this zone a smooth enough profile of d , see the right plot in Fig. 3. Finally, our computer characteristics are as follows: Processor Intel(R) Core(TM) i7-3770 OK, CPU 3.5 GHz, RAM 16.0 GB.

3. Staggered (STAG) approach

In this section, for the benchmark problem at hand, we perform a detailed evaluation of the staggered computational scheme for the solution of (6). This includes the study of the algorithmic aspects of staggered computations such as the convergence of the staggered iterative process, its stopping criterion and the related notion of tolerance, as well as the parametric study on the influence of loading increment size and tolerance on the efficiency of the approach (i.e., accuracy vs. computational cost). The presented results will motivate the need for an alternative to the staggered approach.

3.1. The procedure

Equation $E' = 0$ in (6) can be rewritten as the system of coupled equations

$$\begin{cases} Q_1(\mathbf{u}, d) := \int_{\Omega} \left[(1-d)^2 \frac{\partial \Psi_0^+}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}(\mathbf{u})) + \frac{\partial \Psi_0^-}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}(\mathbf{u})) \right] : \boldsymbol{\varepsilon}(\mathbf{v}) \, d\mathbf{x} = 0, \\ Q_2(\mathbf{u}, d) := G_c \int_{\Omega} \left(\frac{1}{\ell} dw + \ell \nabla d \cdot \nabla w \right) \, d\mathbf{x} - \int_{\Omega} 2(1-d) \Psi_0^+(\boldsymbol{\varepsilon}(\mathbf{u})) w \, d\mathbf{x} - \frac{1}{\gamma} \int_{\text{CR}_{l-1}} (1-d) w \, d\mathbf{x} = 0, \end{cases} \quad (7)$$

in the unknown fields \mathbf{u} and d . The first equation in this system is a non-linear elasticity problem and the second one is the so-called phase-field evolution equation. The boundary conditions for \mathbf{u} read

$$\begin{cases} \mathbf{u} = \mathbf{0} & \text{on } \Gamma_{D,0}, \\ \mathbf{u} = \bar{\mathbf{u}}^l & \text{on } \Gamma_{D,1}, \end{cases} \quad (8)$$

where, in our case, $\bar{\mathbf{u}}^l := (\bar{u}^l, 0)$ with $l \geq 0$ as the current loading step.

The staggered procedure is sketched as follows. Let $\bar{\mathbf{u}}^l$ and $\bar{\mathbf{u}}^{l+1}$ be the fixed displacement loadings at two consecutive loading steps l and $l+1$. Let also the solution pair (\mathbf{u}^l, d^l) to the system (7) at loading step l be known and we seek the solution pair $(\mathbf{u}^{l+1}, d^{l+1})$ at step $l+1$.

In general, there are two options for a staggered process initiation and arrangement depending on which of the two variables, \mathbf{u} or d , is to be ‘frozen’ first. In **Option 1**, which we employ in the simulations below, to initiate the process we set $\mathbf{u} := \mathbf{u}_l$ and solve the equation $Q_2(\mathbf{u}, d) = 0$ for d yielding d_0^{l+1} . Setting next $d := d_0^{l+1}$ and solving the equation $Q_1(\mathbf{u}, d) = 0$ for \mathbf{u} , we obtain \mathbf{u}_0^{l+1} . The result of this first staggered iteration (note that we, however, prefer to start counting iterations from 0) is the solution pair $(\mathbf{u}_0^{l+1}, d_0^{l+1})$. The step is repeated to obtain $(\mathbf{u}_1^{l+1}, d_1^{l+1})$, etc. An alternative **Option 2** is to start with $d := d^l$ and change the order of the equations to be solved accordingly. Regardless of the choice between these two options the staggered iterative process delivers at loading step $l+1$ the set $\{(\mathbf{u}_k^{l+1}, d_k^{l+1})\}$, $k = 0, 1, \dots$. In what follows, the subscript $l+1$ will be dropped.

It is worth noticing that the momentum equation in system (7) is non-linear due to non-linearity of $\sigma(\mathbf{u}, d) := (1-d)^2 \frac{\partial \Psi_0^+}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}(\mathbf{u})) + \frac{\partial \Psi_0^-}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}(\mathbf{u}))$. Hence, at every fixed staggered iteration k , with d_k being given, a Newton–Raphson iterative procedure to obtain the related solution \mathbf{u}_k is required. Since the original functional E in (3) is convex with respect to \mathbf{u} , the mentioned problem is well-posed, thus creating no practical difficulties for the Newton–Raphson method to converge.

3.2. Stopping criterion

Every iterative process has to be terminated when a prescribed tolerance for the converging quantity is achieved. For the phase-field model of brittle fracture, the choice of the quantity whose convergence has to be controlled within a staggered iterative process at a fixed loading step, as well as the related stopping criterion does not seem to be obvious. Some options were already dissected in e.g. [1] and [22]. Herein, we recall and expand the discussion in [1] for the sake of completeness and also to illustrate that convergence of the staggered scheme can be extremely slow.

Let us first introduce the notion of a residual for the staggered iterative process. Recall that according to **Option 1**, at a given staggered iteration k we first solve $Q_2(\mathbf{u}_{k-1}, d) = 0$ for d to obtain what is termed d_k . Then $Q_1(\mathbf{u}, d_k) = 0$ is solved for \mathbf{u} to obtain \mathbf{u}_k . The residual will then read

$$\text{Res}_k := E'(\mathbf{u}_k, d_k; \mathbf{v}, w) = Q_1(\mathbf{u}_k, d_k) + Q_2(\mathbf{u}_k, d_k) = Q_2(\mathbf{u}_k, d_k), \quad (9)$$

since the contribution of $Q_1(\mathbf{u}_k, d_k)$ in this case is negligible (namely, it is the residual of the Newton–Raphson process used for solving the nonlinear momentum equation $Q_1 = 0$). The staggered process is terminated when $\text{Res}_k < \text{Tol}$, where Tol is our user-prescribed tolerance.

For a multi-physics problem like the fracture phase-field formulation, which, in particular, stems from a non-convex minimization problem, the choice of Tol is rather difficult. It may e.g. hold that $\text{Tol} = 10^{-3}$ provides an accurate enough solution and decreasing this value to an ‘intuitively more appropriate’ $\text{Tol} = 10^{-8}$ will result in a tremendous increase of computational effort without an additional gain in accuracy.

To prevent such a situation in our case, we believe that the important quantity whose convergence within the staggered iterative process at a given loading step should be tracked along with the convergence of Res is the energy of the system. Indeed, we know that at a fixed loading step we look for the minimizer of the energy functional E in (3). Then, with $E_k := E(\mathbf{u}_k, d_k)$, we expect that the iteratively computed sequence $\{E_k\}$, $k = 0, 1, 2, \dots$ fulfills the following two properties:

- (a) it decays (to some constant value which is an approximation of the actual unknown $\min(E) =: E_{\min}$ at this loading step), and
- (b) it decays with an increasing slope (which should tend to 0 in the limit $k \rightarrow \infty$).

(a) and (b) constitute what we call (asymptotic) *convergence* of $\{E_k\}$, $k = 0, 1, 2, \dots$ within the staggered iterative process. Our practical observation is that such convergence always holds and the question is how to adequately monitor it. In other words, if (a) holds true, what quantity that traces the (b)-behavior of $\{E_k\}$ should we choose?

Let $N \geq 1$ be the fixed *current* staggered iteration and let $\{\tilde{E}_k\}_{k=0}^N = \{E_0, \dots, E_N\}$ be the computed sequence of energies such that (a) holds. A seemingly reasonable candidate to trace (b) is the quantity

$$\alpha_N := \arctan \left(\frac{E_{N-1} - E_N}{N - (N-1)} \right) \cdot \frac{180^\circ}{\pi}, \quad (10)$$

which controls at every given $N \geq 1$ the slope of segment $[E_{N-1}, E_N]$ with respect to segment $[N-1, N]$. However, as already outlined in [1], the defect of α_N is the significant difference of scales of segments $[E_{N-1}, E_N]$ and $[N-1, N]$, or, more generally, the difference between scales of $[E_0, E_N]$ and $[0, N]$. It can be e.g. of order 10^{2-3} , and what is also important, this can happen regardless of the size of the loading increment $\Delta \bar{u}$. As a result, the corresponding slope α_N may be small (of order 10^{-1-2}), already in the very beginning of the iterative process. This becomes misleading in terms of the magnitude of Res_N which, in turn, can be significantly large at this stage.

A remedy that we already proposed in [1] is based on an appropriate rescaling of the E_k - k diagram at a given fixed $N \geq 1$ to the unit square $[0, 1]^2$, as depicted in Fig. 4. More precisely, given $N \geq 1$ and having computed $\{\tilde{E}_k\}_{k=0}^N = \{\tilde{E}_0, \tilde{E}_1, \dots, \tilde{E}_N\}$, we introduce the normalized energies

$$\{\tilde{E}_k\}_{k=0}^N, \quad \tilde{E}_k := \frac{E_k - E_N}{E_0 - E_N}, \quad (11)$$

and the normalized iteration indexes

$$\{s_k\}_{k=0}^N, \quad s_k := \frac{k}{N}. \quad (12)$$

Note that with this definition, regardless of N it holds that $(\tilde{E}_0, s_0) = (1, 0)$ and $(\tilde{E}_N, s_N) = (0, 1)$. Using (11) and (12) we then define, for a given $N \geq 1$, the quantity

$$\beta_N := \arctan \left(\frac{\tilde{E}_{N-1} - \tilde{E}_N}{s_N - s_{N-1}} \right) \cdot \frac{180^\circ}{\pi}, \quad (13)$$

which controls at this N the slope of segment $[\tilde{E}_{N-1}, \tilde{E}_N]$ with respect to segment $[s_{N-1}, s_N]$, see Fig. 4. Note that at $N = 1$ it is always $\beta_1 = 45^\circ$.

In terms of non-normalized original quantities, Eq. (13) reads

$$\beta_N = \arctan \left(\frac{E_{N-1} - E_N}{E_0 - E_N} \cdot N \right) \cdot \frac{180^\circ}{\pi}. \quad (14)$$

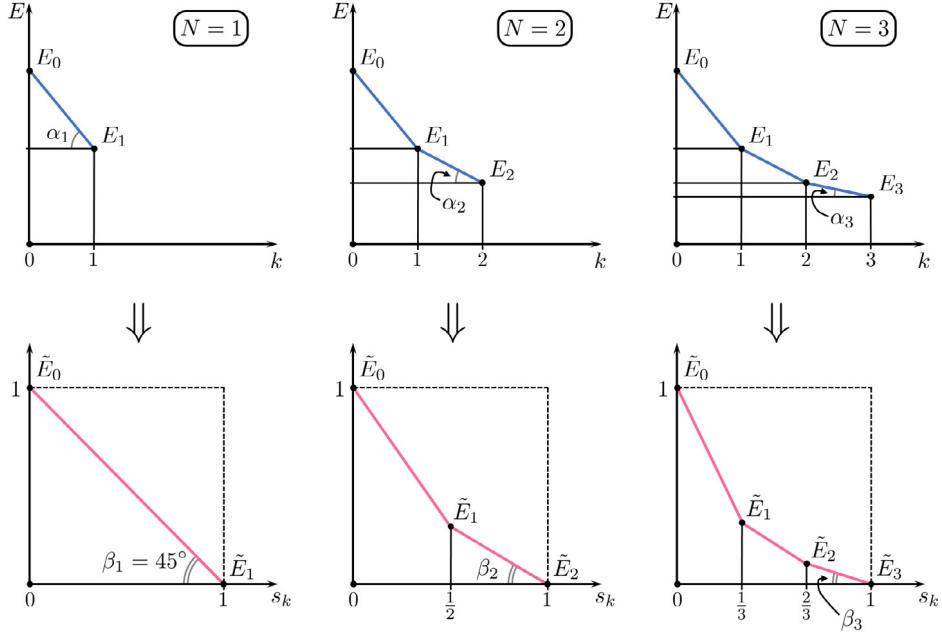


Fig. 4. Sketch of the normalization procedure (11)–(12): E_k - k diagram with the slope measure α_N for a fixed $N \in \{1, 2, 3\}$ (upper row) and the corresponding \tilde{E}_k - s_k diagram embedded in the unit square with β_N (lower row).

In the following, we thus monitor the convergence of the staggered iterative process through Res defined by (9) and in combination with the convergence of β from (14). Section 3.3 will show that setting in our simulations $\text{Tol} = 10^{-4}$ as a reference value is realistic, since this provides a value of approximately 2° for β to be achieved.

3.3. STAG computational results

Herein, we illustrate how the staggered approach works, particularly focusing on the use of the stopping criterion devised above and on the related computational cost.

3.3.1. Reference results

We prescribe the displacement loading as $\bar{u}^{l+1} := \bar{u}^0 + \Delta\bar{u} \cdot (l + 1)$, $l \geq 0$, where we set $\bar{u}^0 := 6 \cdot 10^{-3}$ mm and $\Delta\bar{u} := 0.3 \cdot 10^{-3}$ mm. We stop computations at loading step $l = 46$ yielding $\bar{u}^{47} := 20.1 \cdot 10^{-3}$ mm, at which the phase-field variable d , depicted in Fig. 3 (right), has evolved up to complete failure of the specimen. This is similar to the result we obtain with the discrete crack approach, see Fig. 2 for comparison. Note that compared with the considered loading range, the loading increment is of moderate size. In Section 3.3.2, larger values of $\Delta\bar{u}$ will be also taken to evaluate the effect of this alteration. Finally, at every fixed loading step the staggered iterative solution process is continued until the tolerance $\text{Tol} = 10^{-4}$ for Res is achieved.

The computed load–displacement curve is presented in the left plot of Fig. 5, where the diamond-shaped point corresponding to $\bar{u}^0 = 6 \cdot 10^{-3}$ mm indicates the switch from the pure linear elastic formulation to the full phase-field formulation. Note that the reason for the presence of an ascending branch which follows the softening post-peak response has been discussed in [1].

The framed region in the left plot of Fig. 5, enlarged in the right plot, is of our particular interest as it covers the transition between pre- and post-peak branches of the computed load–displacement curve. Therein, the step number relates to the subscript of \bar{u}^{l+1} . Step 11 is the last loading step of a quasi-linear regime, where the phase-field development is not yet pronounced so that $\max(d) < 1$. Yet, at this step we already observe the first local peak load on the load–displacement curve, with a slight softening behavior afterwards, until step 12. Step 14 then provides a global peak load after which a pronounced softening branch is obtained. For every mentioned step, we illustrate the convergence of $\{\text{Res}_k\}$, $\{E_k\}$ within the STAG iterative process, as well as the related quantity β . These plots are depicted in Fig. 6 (for step 11) and in Figs. 7–9 (for steps 12–13, 14–15, 16–17, respectively).

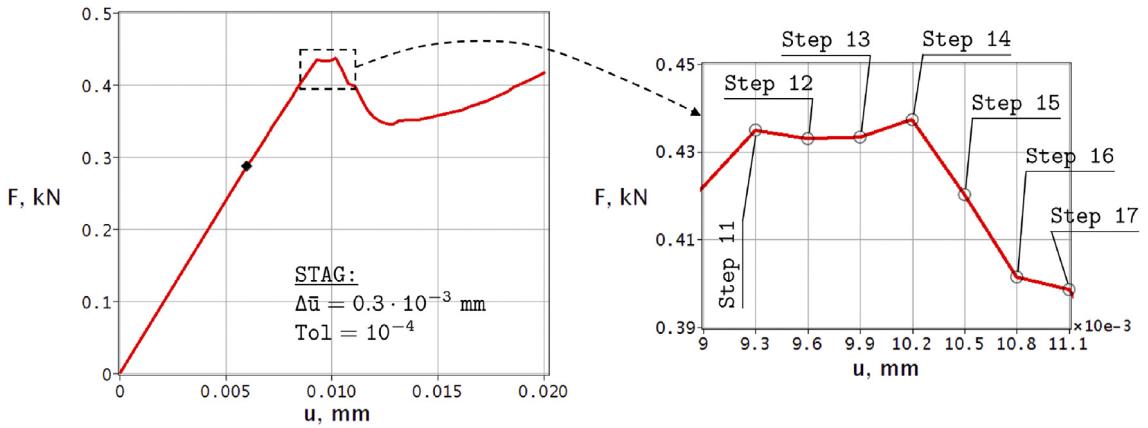


Fig. 5. Load–displacement curve obtained with the STAG approach in the range of loading $[0, 20.1 \cdot 10^{-3}$ mm] (left); zoom into the framed region of the left plot (right): the loading steps 11–17 are of particular interest.

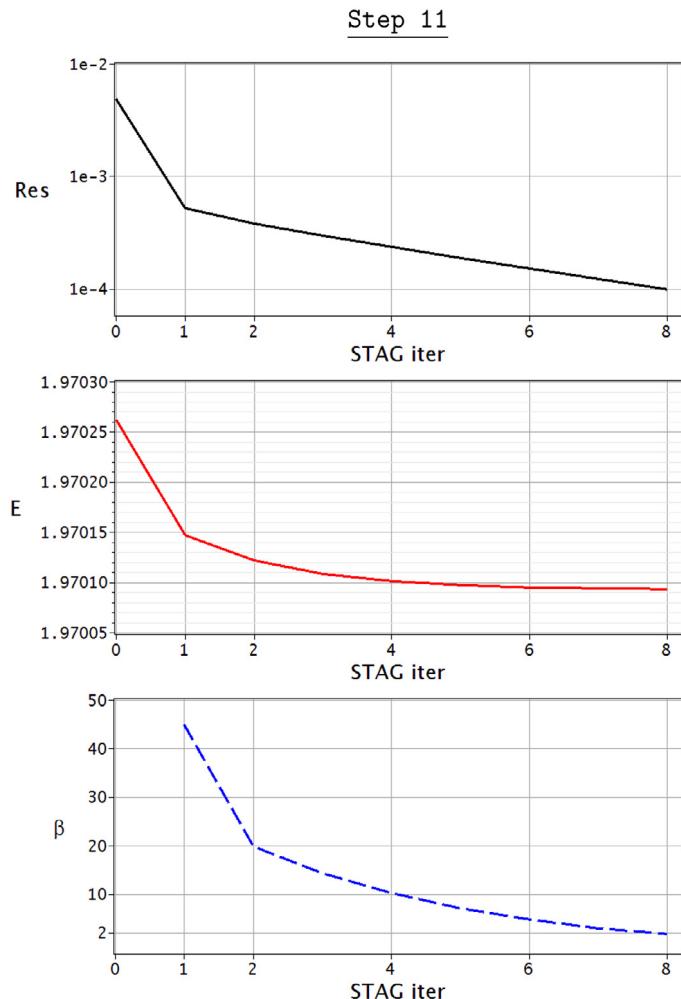


Fig. 6. STAG computations: iterative convergence of Res, E and β at step 11.

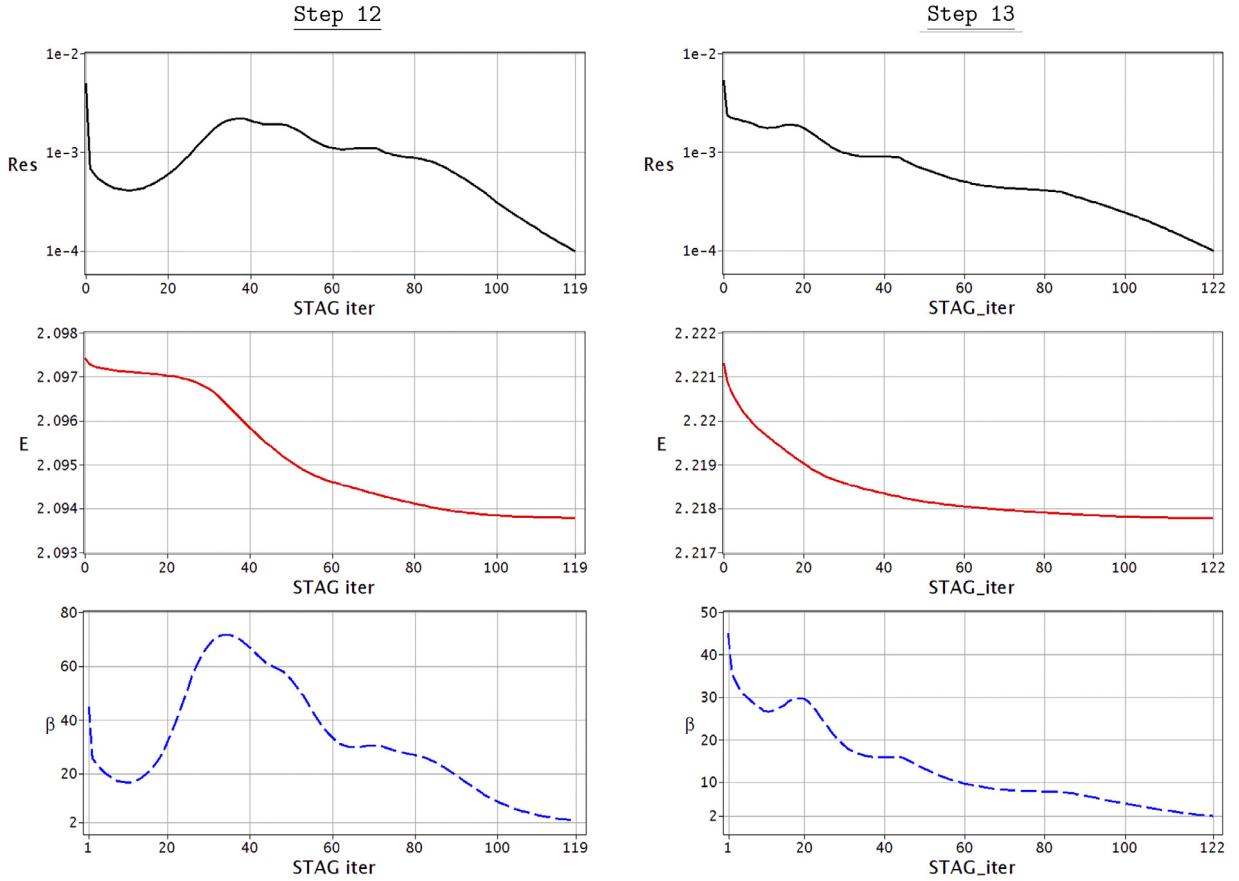


Fig. 7. STAG computations: iterative convergence of Res , E and β at steps 12 and 13.

Results in the figures lead to the following observations:

- (i) in general, a significant amount of staggered iterations is required in order to achieve asymptotic convergence of both $\{\text{Res}_k\}$ and $\{E_k\}$ at a particular fixed loading step. In the presented results, this is evident for all steps and, especially, for steps 15 and 16;
- (ii) tracing the quantity β along with the convergence of $\{\text{Res}_k\}$ is appropriate, as it assists in a clear choice of Tol (in this case, $\text{Tol} = 10^{-4}$ is seemingly good enough).

3.3.2. Parametric study

The importance of performing a sufficient number of staggered iterations for an accurate phase-field computation at a fixed loading step is highlighted in Fig. 10, where the effect of the number of iterations on the phase-field pattern for step 15 is illustrated. It is evident that the use of an inadequate stopping criterion may lead to a significant underestimation of the crack pattern.

The inaccuracy due to insufficient number of staggered iterations is also visible in the global load-displacement response, as illustrated in the left plot of Fig. 11. The corresponding curves are computed for the fixed increment $\Delta\bar{u} := 0.3 \cdot 10^{-3}$ mm and varying tolerance Tol . It is evident that increasing Tol from 10^{-4} (our reference value) to e.g. 10^{-2} the accuracy is sacrificed significantly. The related computational cost data, including the cumulative computational time, are reported in the right plot of Fig. 11 and the table below. Also, the evolution of the phase-field at three loading steps for $\text{Tol} = 10^{-4}$ and 10^{-2} is depicted in Fig. 12, showing a remarkable difference between the results obtained with the two different tolerances.

Finally, in Fig. 13 we illustrate the effect of the load step size: staggered computations with fixed reference $\text{Tol} = 10^{-4}$ and varying $\Delta\bar{u}$ are performed and the results are compared in terms of accuracy and cost. With an adequately small Tol , the size of $\Delta\bar{u}$ does not affect the accuracy significantly, yet the time resolution in capturing

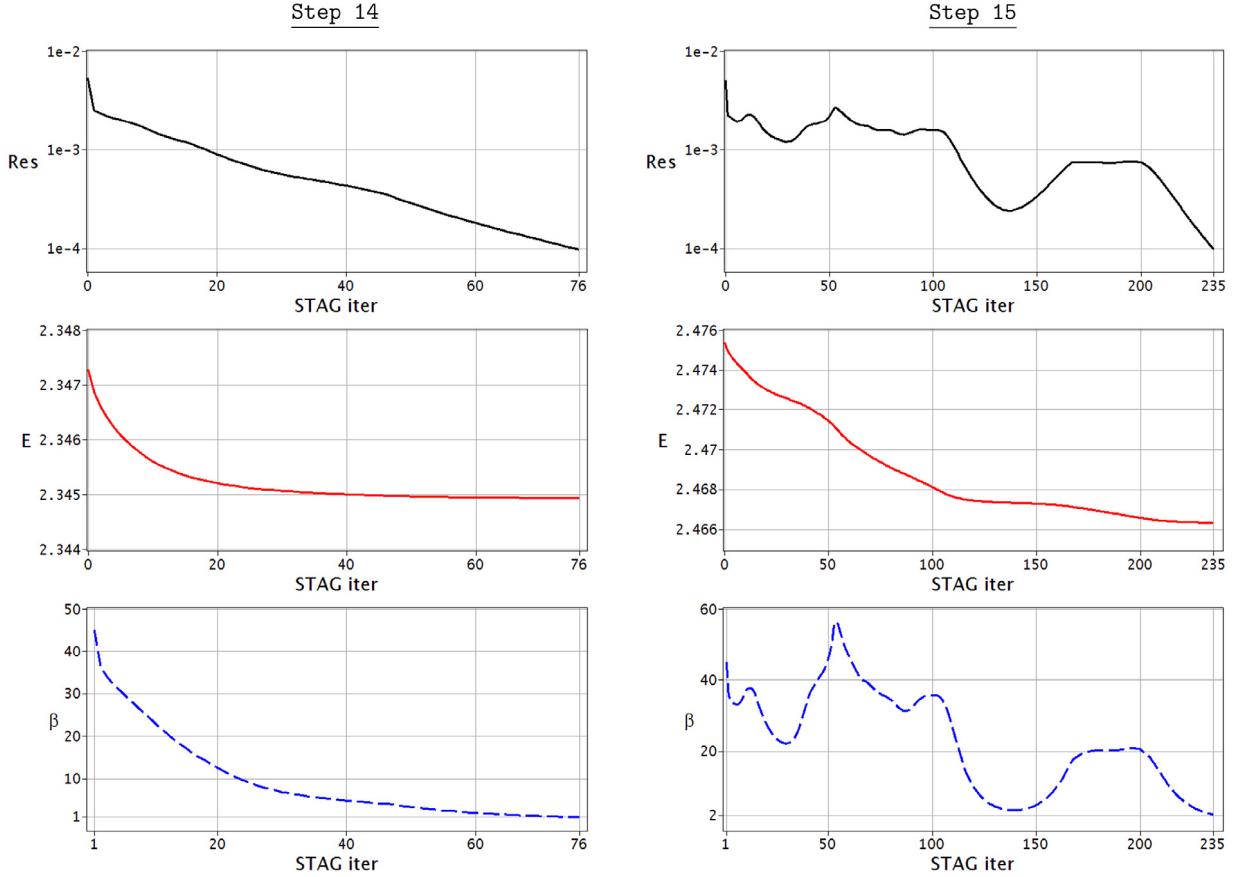


Fig. 8. STAG computations: iterative convergence of Res, E and β at steps 14 and 15.

the response is, of course, decreased. On the other hand, the computational cost decreases as the increment size is increased.

We conclude this section by summarizing our observations. The slow convergence of the energy within a staggered iterative process at a fixed loading step is the source of a high computational cost of the staggered scheme. This already holds for a rather simple 2d quasi-static phase-field formulation and is expected to become more pronounced in less trivial scenarios like e.g. multiple cracks evolution or 3d problems. Increasing the tolerance Tol of the staggered computations yields results whose accuracy can no longer be guaranteed. On the other hand, an increase of the loading increment size $\Delta\bar{u}$ does not have an impact on accuracy and enables a reduction in cost within a staggered approach. However, the computational cost, albeit reduced, remains quite onerous. Acceleration techniques in the spirit of [23] and [22] may, in general, appear appropriate in this case but should be reformulated in the present non-convex setting.

As follows, we investigate the monolithic treatment of the governing equation $E' = 0$ which promises to be faster than the staggered scheme and equally accurate.

4. A line search assisted monolithic (MON-ls) approach

In the monolithic scheme, at every fixed loading step equation $E' = 0$ in (6) is solved simultaneously for (\mathbf{u}, d) by the Newton–Raphson (N–R) method. Note that within the STAG approach, the N–R procedure is also employed but for solving solely the momentum equation; in other words, it is nested in every fixed staggered iteration, see Fig. 14.

The direct monolithic treatment of the strongly non-linear formulation $E' = 0$ seems more natural than the staggered one and can also be expected to be faster. We will show this using the loading history from Section 3.3.1 where the STAG reference results were obtained. However, it will also be shown that starting from some loading step the MON approach does not achieve convergence of the N–R iterative procedure. To overcome this issue, we propose

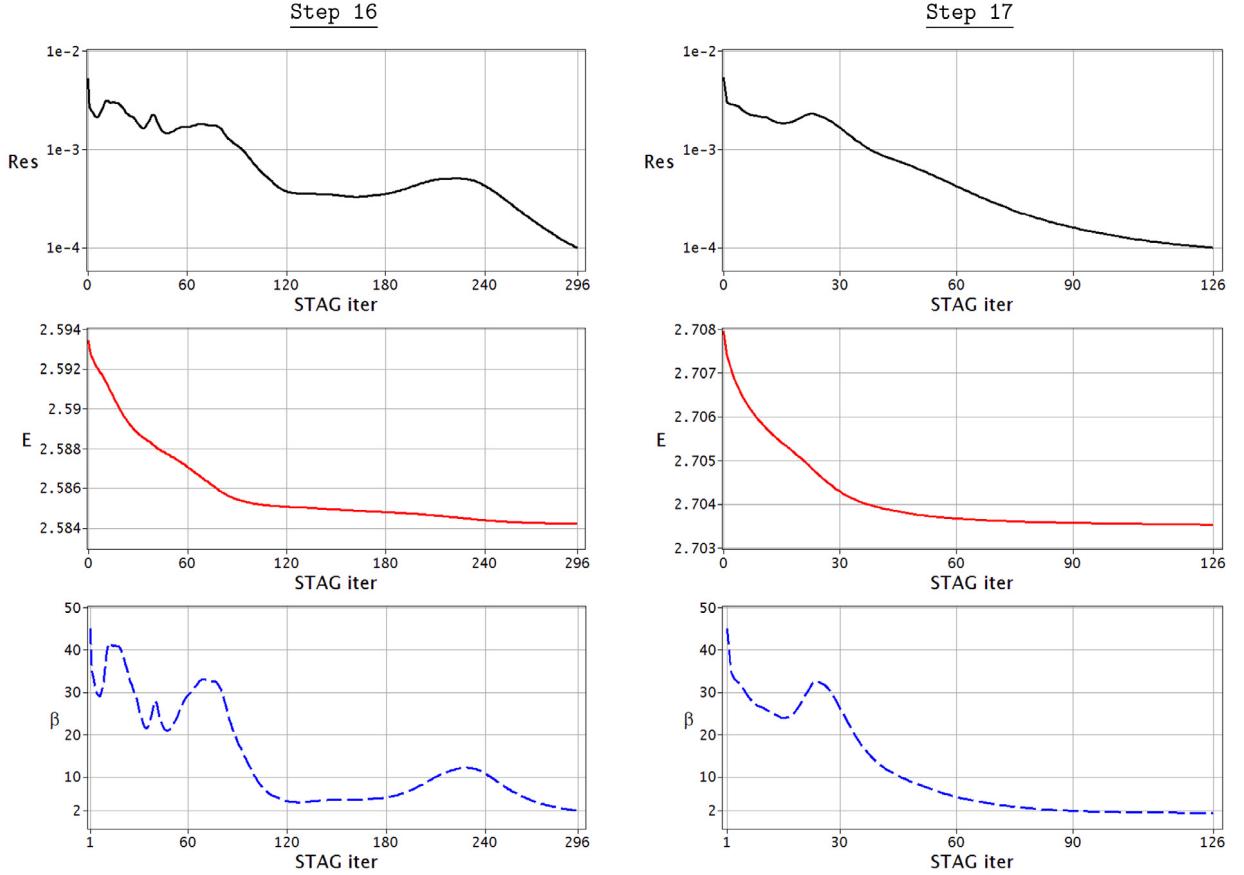


Fig. 9. STAG computations: iterative convergence of Res, E and β at steps 16 and 17.

a novel line search procedure that is specifically tailored to the considered phase-field formulation. The results and computational cost of the STAG and the MON-Is approaches will be compared.

4.1. Newton–Raphson procedure

At every loading step $l \geq 0$, that is, with a given fixed $\bar{\mathbf{u}} = (\bar{\mathbf{u}}^l, 0)$ on $\Gamma_{D,1}$, we look for a solution (\mathbf{u}, d) to Eq. (6). A linearization procedure and the consequent N–R iterative solution process apply: at every iteration $i \geq 0$, starting from a known solution pair (\mathbf{u}_i^l, d_i^l) , the equation

$$E'(\mathbf{u}_i^l + \Delta\mathbf{u}_i^l, d_i^l + \Delta d_i^l; \mathbf{v}, w) \approx E'(\mathbf{u}_i^l, d_i^l; \mathbf{v}, w) + \Phi(\Delta\mathbf{u}_i^l, \Delta d_i^l; \mathbf{u}_i^l, d_i^l; \mathbf{v}, w) = 0, \quad (15)$$

is solved for the unknown correction increment $(\Delta\mathbf{u}_i^l, \Delta d_i^l)$ assuming $(\mathbf{v}, w) \in \mathbf{V}_0 \times H^1(\Omega)$ to be arbitrary. Φ is a functional which is linear in $\Delta\mathbf{u}_i^l$ and Δd_i^l , reading explicitly as

$$\begin{aligned} \Phi(\Delta\mathbf{u}_i^l, \Delta d_i^l; \mathbf{u}_i^l, d_i^l; \mathbf{v}, w) &:= \int_{\Omega} \left\{ \boldsymbol{\epsilon}(\Delta\mathbf{u}_i^l) : \left[(1 - d_i^l)^2 \frac{\partial^2 \Psi_0^+}{\partial \varepsilon^2}(\boldsymbol{\epsilon}(\mathbf{u}_i^l)) + \frac{\partial^2 \Psi_0^-}{\partial \varepsilon^2}(\boldsymbol{\epsilon}(\mathbf{u}_i^l)) \right] : \boldsymbol{\epsilon}(\mathbf{v}) \right. \\ &\quad - 2(1 - d_i^l) \Delta d_i^l \frac{\partial \Psi_0^+}{\partial \varepsilon}(\boldsymbol{\epsilon}(\mathbf{u}_i^l)) : \boldsymbol{\epsilon}(\mathbf{v}) - 2(1 - d_i^l) \frac{\partial \Psi_0^+}{\partial \varepsilon}(\boldsymbol{\epsilon}(\mathbf{u}_i^l)) : \boldsymbol{\epsilon}(\Delta\mathbf{u}_i^l) w \\ &\quad \left. + 2\Delta d_i^l \Psi_0^+(\boldsymbol{\epsilon}(\mathbf{u}_i^l)) w + G_c \left(\frac{1}{\ell} \Delta d_i^l w + \ell \nabla(\Delta d_i^l) \cdot \nabla w \right) \right\} dx + \frac{1}{\gamma} \int_{CR_{i-1}} \Delta d_i^l w dx. \end{aligned} \quad (16)$$

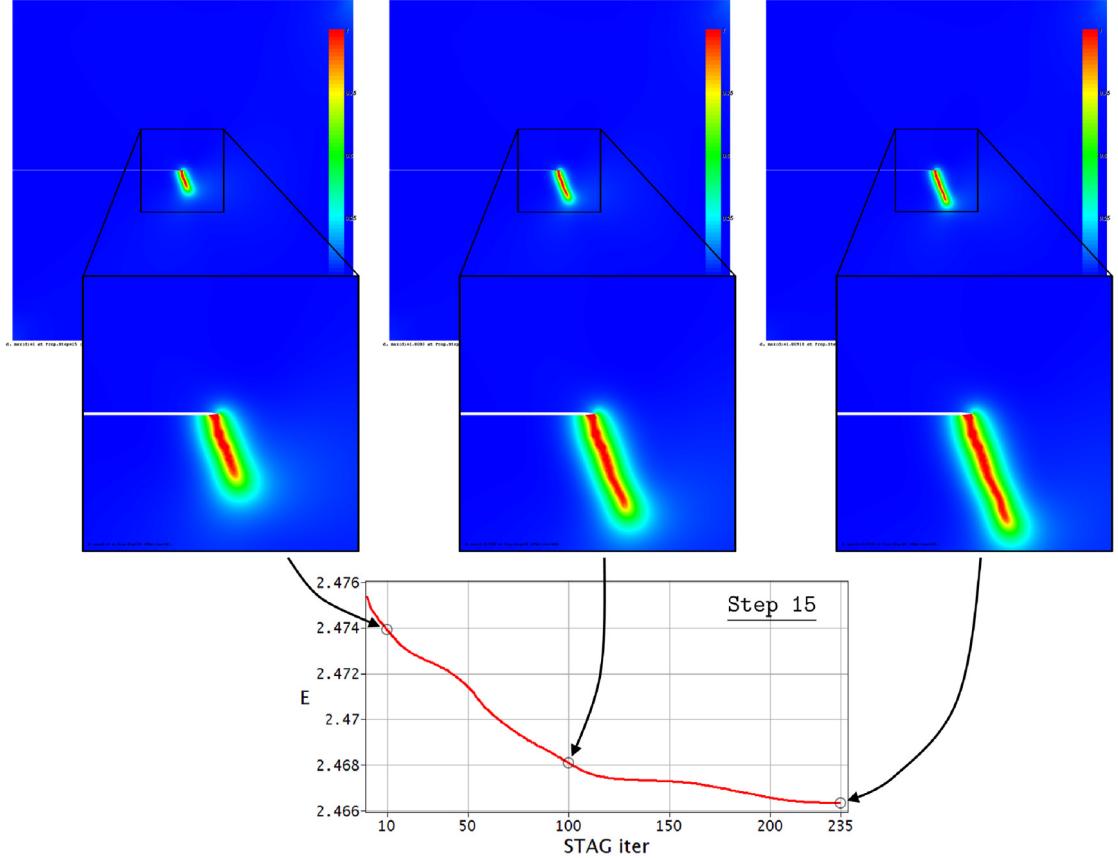


Fig. 10. Step15: crack phase-field at a staggered iteration $k \in \{10, 100, 235\}$ corresponding to the energy levels E_{10}, E_{100}, E_{235} ; a significant difference in the profile of d is visible.

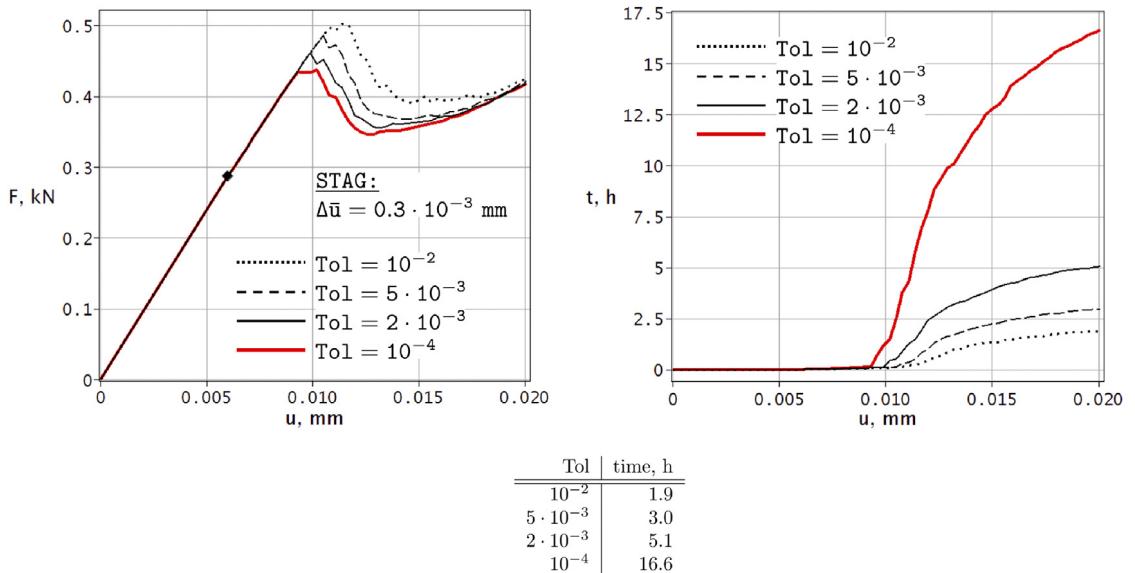
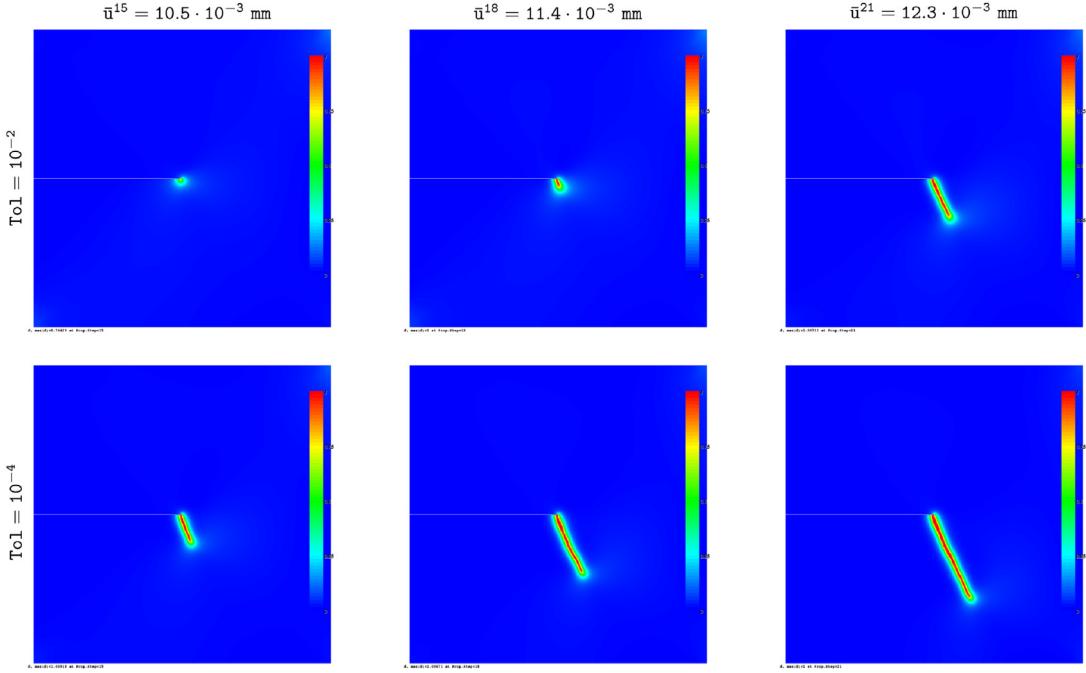
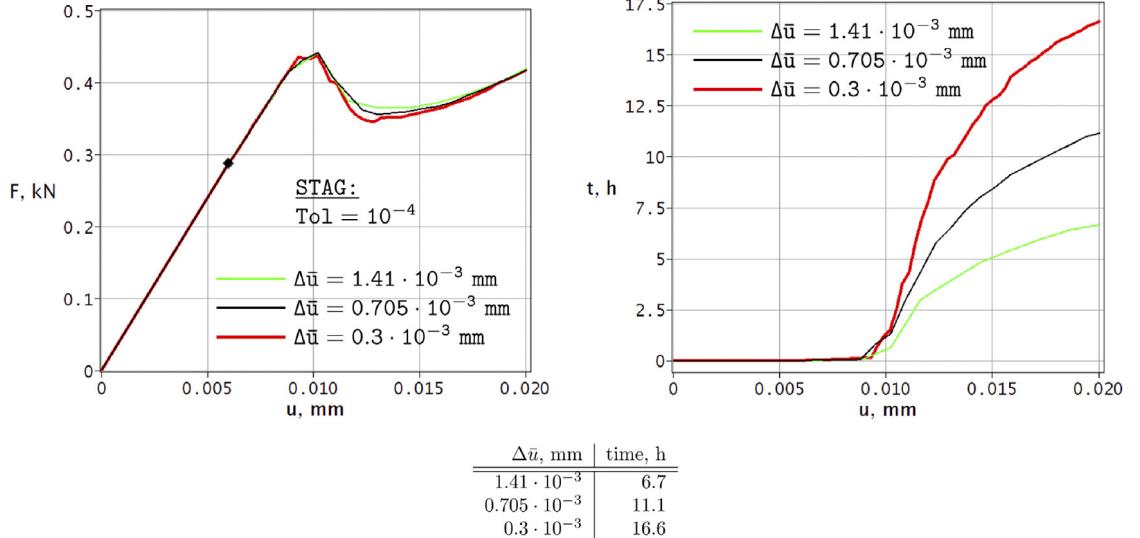


Fig. 11. Load–displacement curves and related time–displacement diagram obtained with the staggered approach for fixed $\Delta\bar{u}$ and varying Tol; the cumulative computational time is reported in the tables.

Fig. 12. Crack phase-field at applied displacements \bar{u}^{15} , \bar{u}^{18} , \bar{u}^{21} for the different values of Tol .Fig. 13. Load–displacement curves and related time–displacement diagram obtained with the staggered approach for fixed Tol and varying $\Delta \bar{u}$; the cumulative computational time is reported in the tables.

Finally, the boundary conditions to be fulfilled are

$$\begin{cases} \mathbf{u}_i^l + \Delta \mathbf{u}_i^l = \mathbf{0} & \text{on } \Gamma_{D,0}, \\ \mathbf{u}_i^l + \Delta \mathbf{u}_i^l = \bar{\mathbf{u}} = (\bar{u}^l, 0) & \text{on } \Gamma_{D,1}. \end{cases} \quad (17)$$

For $i = 0$, the pair (\mathbf{u}_0^l, d_0^l) in (15)–(17) is the initial guess, whose choice should be specified. A standard option is to use the solution pair $(\mathbf{u}^{l-1}, d^{l-1})$ to $E' = 0$ found at the previous loading step, setting directly $(\mathbf{u}_0^l, d_0^l) :=$

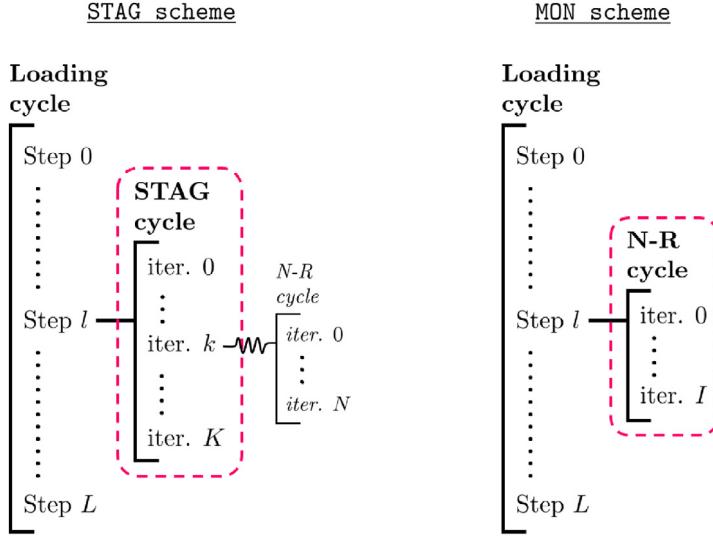


Fig. 14. Sketch of the STAG and the MON approaches: the core iterative process of computing (\mathbf{u}, d) at every fixed loading step l within either approach is framed.

$(\mathbf{u}^{l-1}, d^{l-1})$. In this case, \mathbf{u}_0^l is such that

$$\begin{cases} \mathbf{u}_0^l = \mathbf{0} & \text{on } \Gamma_{D,0}, \\ \mathbf{u}_0^l = (\bar{\mathbf{u}}^{l-1}, 0) & \text{on } \Gamma_{D,1}, \end{cases}$$

holds. Then, as follows from (17), the corresponding boundary conditions for $\Delta\mathbf{u}_0^l$ will be

$$\begin{cases} \Delta\mathbf{u}_0^l = \mathbf{0} & \text{on } \Gamma_{D,0}, \\ \Delta\mathbf{u}_0^l = (\bar{\mathbf{u}}^l - \bar{\mathbf{u}}^{l-1}, 0) & \text{on } \Gamma_{D,1}, \end{cases}$$

and for $\Delta\mathbf{u}_i^l$, $i \geq 1$, they are

$$\begin{cases} \Delta\mathbf{u}_i^l = \mathbf{0} & \text{on } \Gamma_{D,0}, \\ \Delta\mathbf{u}_i^l = \mathbf{0} & \text{on } \Gamma_{D,1}. \end{cases} \quad (18)$$

For our further ‘line search’ purposes (particularly, to enable the possibility of introducing a negative line search parameter, see Section 4.2), it proves convenient that $\Delta\mathbf{u}_i^l$ satisfies (18) for all i , including $i = 0$. To achieve this, a ‘pre-initial’ guess $(\mathbf{u}_{00}^l, d_{00}^l) := (\mathbf{u}^{l-1}, d^{l-1})$ is devised and the problem (15)–(17) is solved for $(\mathbf{u}_{00}^l + \Delta\mathbf{u}_{00}^l, d_{00}^l + \Delta d_{00}^l)$ to extract the increments $(\Delta\mathbf{u}_{00}^l, \Delta d_{00}^l)$. These are now used to produce an actual initial guess

$$(\mathbf{u}_0^l, d_0^l) := (\mathbf{u}_{00}^l + \Delta\mathbf{u}_{00}^l, d_{00}^l + \Delta d_{00}^l),$$

which provides the desired property. In what follows, the subscript l is dropped.

We finalize the algorithmic description of the *standard* N–R iterative process above by defining its stopping criterion. At every iteration $i \geq 0$, for the computed $(\Delta\mathbf{u}_i, \Delta d_i)$ and the obtained update $(\mathbf{u}_i + \Delta\mathbf{u}_i, d_i + \Delta d_i)$, the residual is naturally defined via the quantity

$$\text{Res}_i := E'(\mathbf{u}_i + \Delta\mathbf{u}_i, d_i + \Delta d_i; \mathbf{v}, w), \quad (19)$$

and the N–R process is stopped when $\text{Res}_i < \text{Tol}$ is achieved. In the numerical experiments to follow, the reference value for Tol is prescribed to 10^{-4} as in the staggered computations.

Figs. 15 and 16 depict the results obtained with the MON scheme using the loading setup from Section 3.3.1 and their comparison with the corresponding reference STAG results. The load–displacement curves up to step 11, as well as the phase-field profiles at step 11 are identical in both cases. Yet, the MON approach loses iterative convergence of

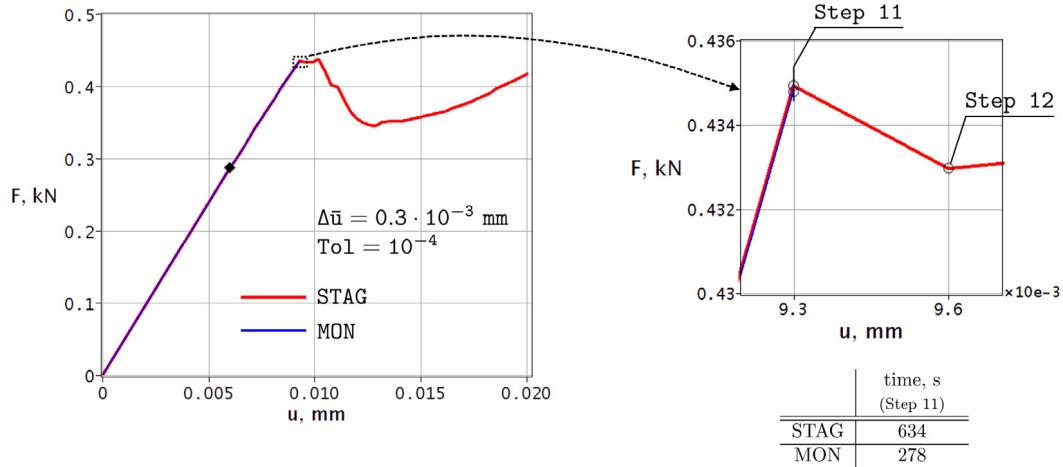


Fig. 15. Load-displacement curves obtained with the STAG scheme and the MON scheme which uses the *standard* N-R procedure: step 11 is the last loading step when the MON approach achieves iterative convergence; the table provides the computational time up to step 11.

the N-R procedure at step 12, which can be explained as follows. A drop in the recovered reactions F on the reference STAG curve for the two consequent loading steps 11 and 12 in Fig. 15 is reflected through a so-called ‘brutal’ crack phase-field evolution — a jump from a state with ‘slightly’ degraded material properties with $\max(d) < 1$, to a state with a well-pronounced degradation of stiffness due to $\max(d) = 1$, see the related STAG plot in Fig. 16. Within the MON computing, non-convexity of the governing free-energy functional E seems to be not well-pronounced until step 11 (when $\max(d) < 1$ and the material behavior is almost linear), thus creating no essential difficulties for the standard N-R procedure to converge. However, when d evolves to its highest value at step 12, non-convexity of E becomes prominent thus making the application of the standard N-R method insufficient.

Our final comment here is that until step 11, the MON approach turns out to be superior than the STAG one in computational time, see the table in Fig. 15. At the moment, this is just an observation supporting our guess that the MON scheme should, in general, have a better computational efficiency provided the N-R iterative process converges.

4.2. Our version of the line search procedure

Given a fixed loading step and the N-R iterative process, the convergence issues of the N-R procedure imply an increase of the residual E' between *two consequent* iterations. A simultaneous increase can be registered for the energy E . Once occurred, the trend will typically continue at following N-R iterations (no self-stabilization is observed), thus producing no meaningful solution at this loading step.

We propose to use the line search procedure to recover convergence of the N-R process. As was already mentioned, in the mechanics literature this procedure is typically employed to *accelerate* convergence of the standard N-R procedure, or its variants (modified Newton’s method, quasi-Newton method, etc.), see e.g. [15–19]. In contrast, in our phase-field formulation, line search will be a tool to *achieve* convergence in the first place, while dealing with a problem of non-convex minimization.

The basic idea of the line search technique is illustrated in Fig. 17. Having at hand at iteration i both the solution \mathbf{u}_i and the computed correction increment $\Delta\mathbf{u}_i$ such that $\Psi(\mathbf{u}_i) > \Psi(\mathbf{u}_i + \Delta\mathbf{u}_i)$ holds, with Ψ as the (energy) functional to be minimized, one rescales $\Delta\mathbf{u}_i$ by a multiplier $\alpha \in (0, 1)$ called line search parameter, in order to arrive at the lowest energy $\Psi(\mathbf{u}_i + \alpha\Delta\mathbf{u}_i)$ along the search direction $\Delta\mathbf{u}_i$. With an optimal value $\alpha^* \in (0, 1]$ derived, the residual at $\mathbf{u}_i + \alpha^*\Delta\mathbf{u}_i$ reduces as well, thus improving the convergence rate at the given N-R iteration. Finally, $\mathbf{u}_i + \alpha^*\Delta\mathbf{u}_i$ is taken as guess for the $(i+1)$ -th iteration.

In our fracture phase-field case, the situation is more delicate and the sketched line search procedure has to be adjusted accordingly. Let $i \geq 0$ be the N-R iteration at which (\mathbf{u}_i, d_i) and $(\Delta\mathbf{u}_i, \Delta d_i)$ are given and fixed. Recall that the divergence of the N-R procedure at i is indicated by the increase of energy at the updated solution $(\mathbf{u}_i + \Delta\mathbf{u}_i, d_i + \Delta d_i)$, that is,

$$E(\mathbf{u}_i, d_i) < E(\mathbf{u}_i + \Delta\mathbf{u}_i, d_i + \Delta d_i).$$

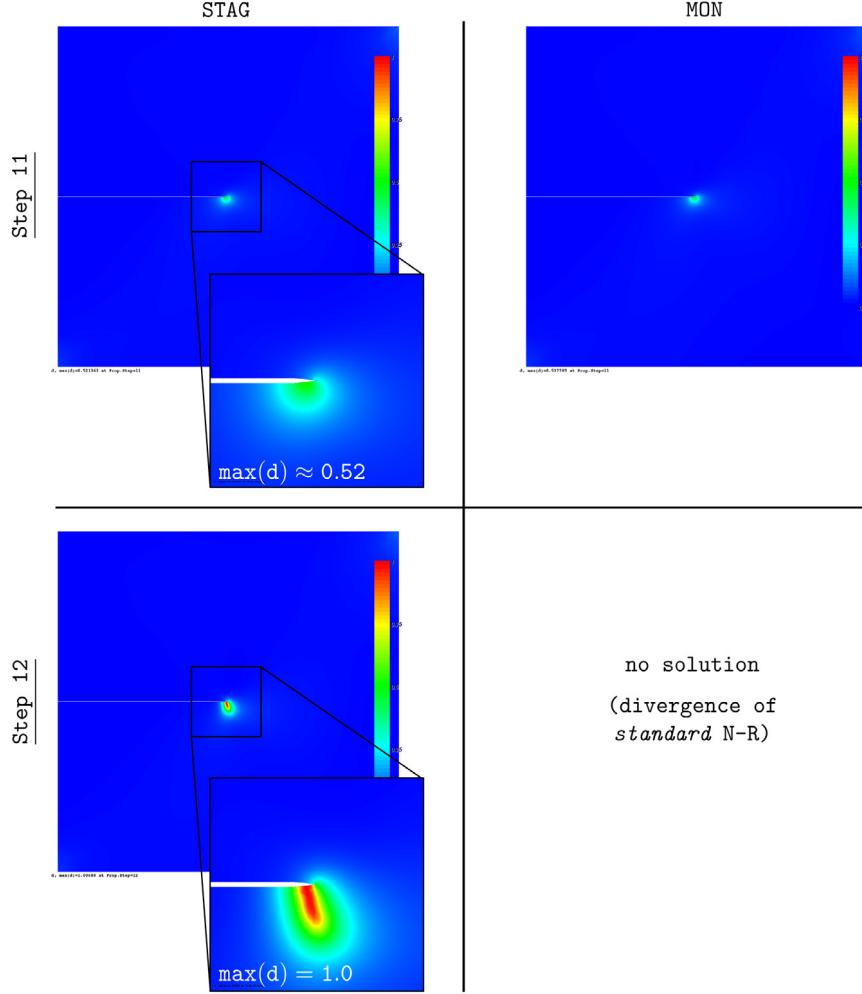


Fig. 16. The phase-field profile obtained with the STAG and MON approaches: identical results at step 11 are recovered, but no MON solution for d is obtained at step 12, where, according to the STAG results, the phase-field is already well-developed, i.e. $\max(d) = 1$.

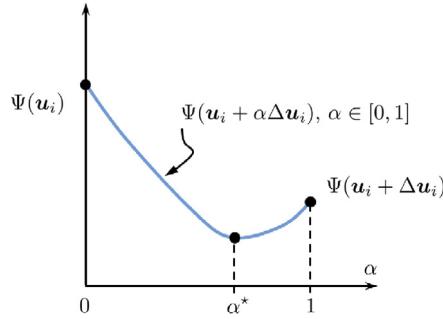


Fig. 17. Line search idea in the ‘classical case’ with the main goal to improve the residual convergence rate.

As above, a homotopy $E(\mathbf{u}_i + \alpha \Delta \mathbf{u}_i, d_i + \alpha \Delta d_i)$ with $\alpha \in [0, 1]$ can be introduced such that

$$E(\mathbf{u}_i + \alpha \Delta \mathbf{u}_i, d_i + \alpha \Delta d_i) = \begin{cases} E(\mathbf{u}_i, d_i) & \text{if } \alpha = 0, \\ E(\mathbf{u}_i + \Delta \mathbf{u}_i, d_i + \Delta d_i) & \text{if } \alpha = 1, \end{cases} \quad (20)$$

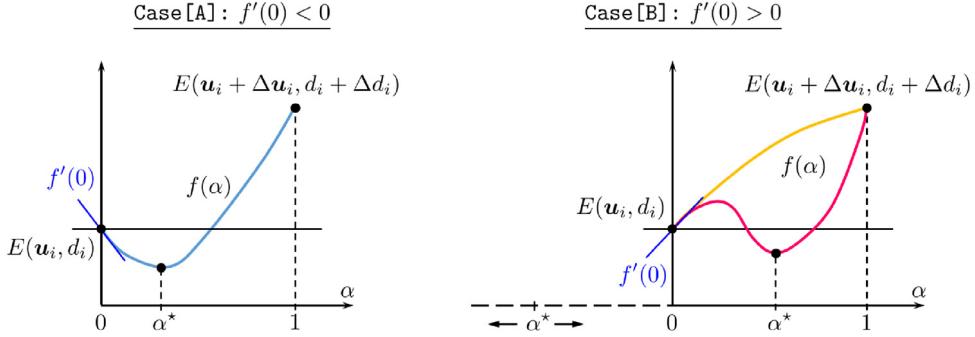


Fig. 18. Sketch of our line search procedure: depending on the sign of $f'(\alpha)$ at $\alpha = 0$, the need of looking for a negative line search parameter is illustrated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

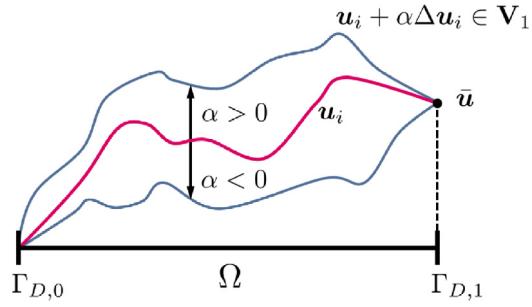


Fig. 19. Due to the imposed property $\Delta u_i = \mathbf{0}$ on $\Gamma_{D,0}$ and particularly on $\Gamma_{D,1}$ for every N-R iteration $i \geq 0$, the line search variation $u_i + \alpha \Delta u_i$ remains in the admissible displacement space regardless of the sign of α .

holds in the limiting cases, and the line search idea will be to find $\alpha^* \in (0, 1)$ that provides the required decrease of energy:

$$E(u_i, d_i) > E(u_i + \alpha^* \Delta u_i, d_i + \alpha^* \Delta d_i). \quad (21)$$

We may call the line search related to the above homotopy a one-parametric line search, as it assumes using the same line search parameter for the two directions Δu_i and Δd_i , which are different by nature. The more general, two-parametric version, where two different parameters are associated to Δu_i and Δd_i is not further investigated here.

In practice, the homotopy using the interval $[0, 1]$ for the parameter α appears insufficient, as elaborated in the following. If for the sake of convenience, we introduce function

$$f(\alpha) := E(u_i + \alpha \Delta u_i, d_i + \alpha \Delta d_i),$$

such that

$$f'(\alpha) = E'(\mathbf{u}_i + \alpha \Delta \mathbf{u}_i, d_i + \alpha \Delta d_i; \Delta \mathbf{u}_i, \Delta d_i),$$

two distinct situations can be recovered. If $f'(\alpha)$ at $\alpha = 0$ is negative, as sketched in the left plot of Fig. 18, the existence of $\alpha^* \in (0, 1)$ such that condition (21) is fulfilled can be proven explicitly (analytically). We term this situation Case [A] and note that it relates to the classical line search scenario. If $f'(\alpha)$ at $\alpha = 0$ has a positive sign – this alternative situation is termed Case [B] – the existence of $\alpha^* \in (0, 1)$ cannot be guaranteed, see the right plot in Fig. 18. In particular, a positive α^* such that (21) holds may still exist, such as for the ‘oscillating’ $f(\alpha)$ in red, or may not, such as for the monotonically increasing $f(\alpha)$ in yellow. In either case, it seems appropriate to look also in the negative direction of α .

Before we proceed with the idea of searching also for a negative parameter α^* and give the related technical details, let us bring one extra argument to justify the general *admissibility* of such a search. Indeed, we may notice that we ‘designed’ the initial guess in Section 4.1 such that for all N-R iterations the increment $\Delta u_i = \mathbf{0}$ on $\Gamma_{D,0}$ and $\Gamma_{D,1}$. This, in turn, yields that the line search variation $u_i + \alpha \Delta u_i$ remains in the admissible displacement space

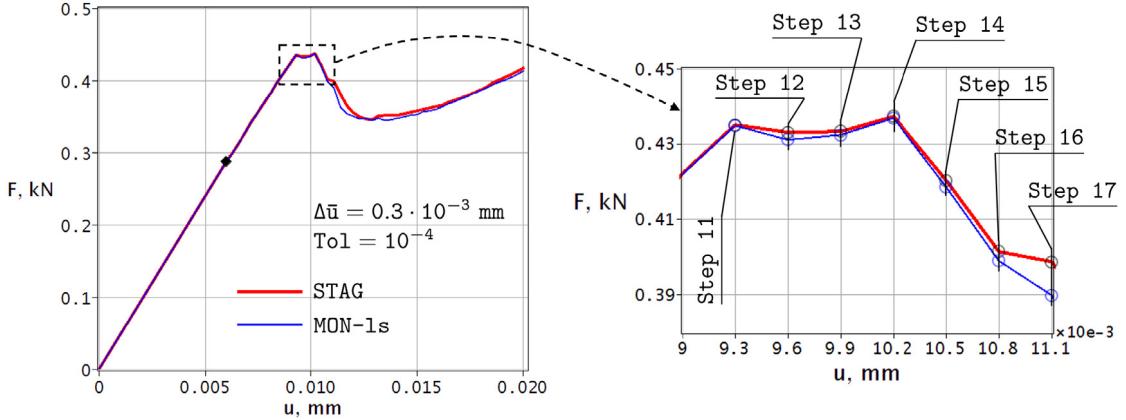


Fig. 20. The STAG and MON-ls load–displacement curves in the entire range of loading (left); zoom into the framed region of the left plot (right): the loading steps of particular interest.

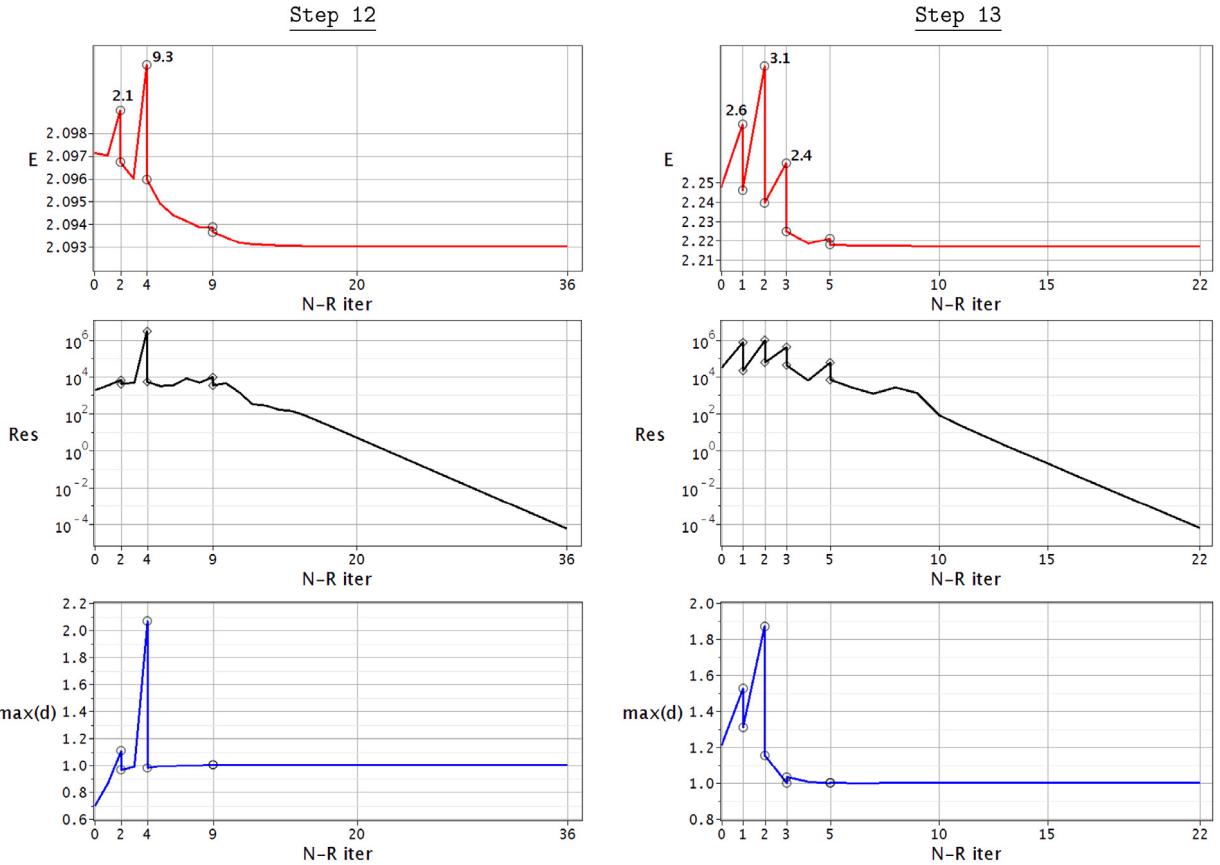


Fig. 21. MON-ls computations: convergence of E , Res and $\max(d)$ at loading steps 12 and 13.

∇_1 regardless of the sign of α , see the sketch in Fig. 19. Note that this also holds for the variation $d_i + \alpha \Delta d_i$. More precisely, with d_i and Δd_i both belonging to $H^1(\Omega)$, $d_i + \alpha \Delta d_i \in H^1(\Omega)$ will be fulfilled naturally regardless of the sign of α .

It is clear that the search interval in the positive direction of α will be $[0, 1]$ by default. Once Case [B] is detected, the question is how far in the negative direction of α should we go. In the numerical experiments to follow we restrict

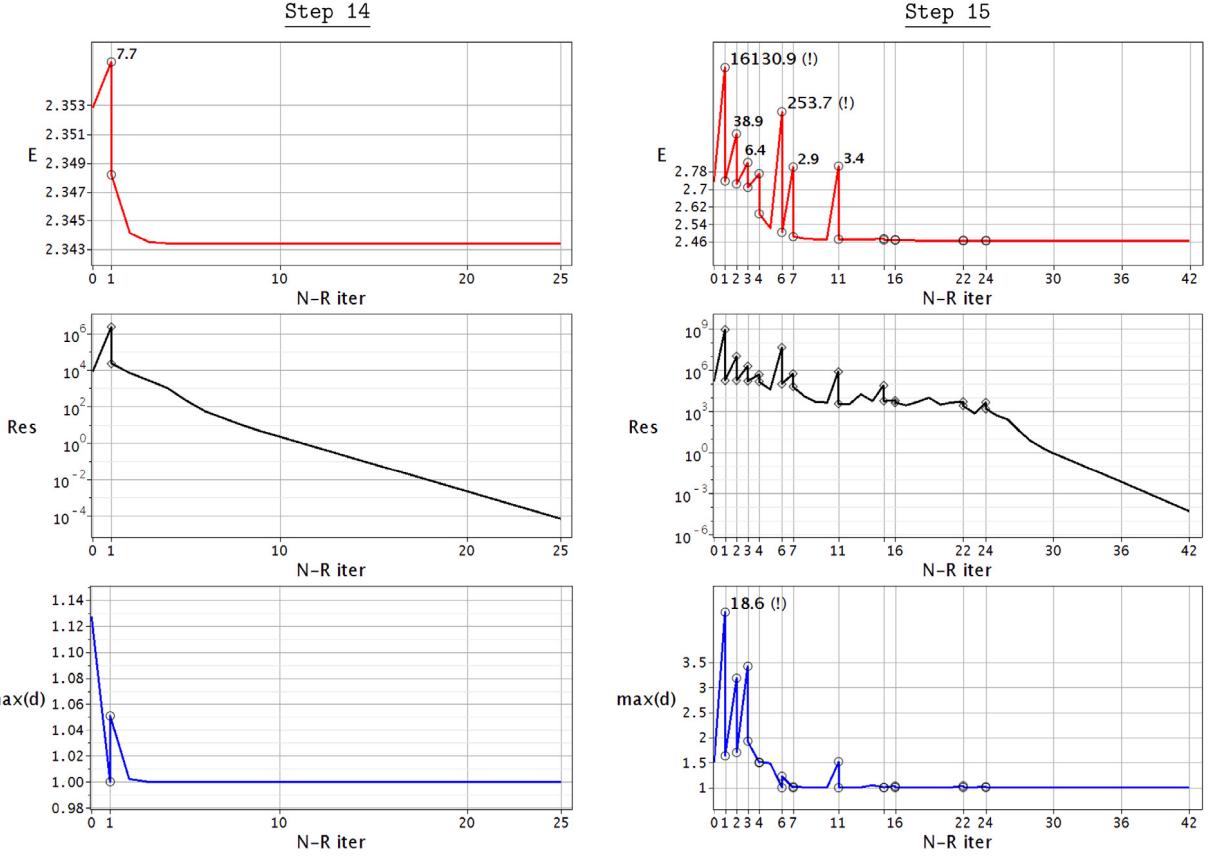


Fig. 22. MON-Is computations: convergence of E , Res and $\max(d)$ at loading steps 14 and 15.

ourselves to the interval $[-1, 0]$ and show (for the particular mechanical setup at hand) that we are always able to find $\alpha^* \in (-1, 0)$ in Case [B].

Note that allowing $\alpha = -1$ in the constructed homotopy $E(\mathbf{u}_i + \alpha \Delta \mathbf{u}_i, d_i + \alpha \Delta d_i)$, does not mean that we are sent back to the energy level $E(\mathbf{u}_{i-1}, d_{i-1})$ achieved at the previous $(i-1)$ th iteration, as it may seem at first glance. That is,

$$E(\mathbf{u}_i + \alpha \Delta \mathbf{u}_i, d_i + \alpha \Delta d_i) \neq E(\mathbf{u}_{i-1}, d_{i-1}) \quad \text{if } \alpha = -1.$$

Indeed, $(\mathbf{u}_i - \Delta \mathbf{u}_i, d_i - \Delta d_i) \neq (\mathbf{u}_{i-1}, d_{i-1})$, since $(\mathbf{u}_i, d_i) = (\mathbf{u}_{i-1} + \Delta \mathbf{u}_{i-1}, d_i + \Delta d_{i-1})$ and $(\Delta \mathbf{u}_i, \Delta d_i) \neq (\Delta \mathbf{u}_{i-1}, \Delta d_{i-1})$.

We finalize this subsection by re-defining the notion of the residual in (19) introduced for a standard N-R iteration. This definition has to be complemented in order to account for the update $(\mathbf{u}_i + \alpha^* \Delta \mathbf{u}_i, d_i + \alpha^* \Delta d_i)$ due to the line search and we have

$$\text{Res}_i := \begin{cases} E'(\mathbf{u}_i + \Delta \mathbf{u}_i, d_i + \Delta d_i; \mathbf{v}, w), & \text{line search is OFF} \\ & \text{(not required),} \\ E'(\mathbf{u}_i + \alpha^* \Delta \mathbf{u}_i, d_i + \alpha^* \Delta d_i; \mathbf{v}, w), & \text{line search is ON.} \end{cases} \quad (22)$$

Also, we will use (22) in the stopping criterion $\text{Res}_i < \text{Tol} = 10^{-4}$.

4.2.1. Computation of α^*

We finally comment on how α^* fulfilling (21) is computed in the two possible cases. In Case [A], the interval $[0, 1]$ is discretized in 10 uniform segments $[\alpha_k, \alpha_{k+1}]$, with $\alpha_k := \frac{1}{10}k$ and $k = 0, \dots, 10$, whereas in Case [B], the interval $[-1, 1]$ is discretized in 20 uniform segments $[\alpha_k, \alpha_{k+1}]$, with $\alpha_k := -1 + \frac{1}{10}k$ and $k = 0, \dots, 20$. Regardless of

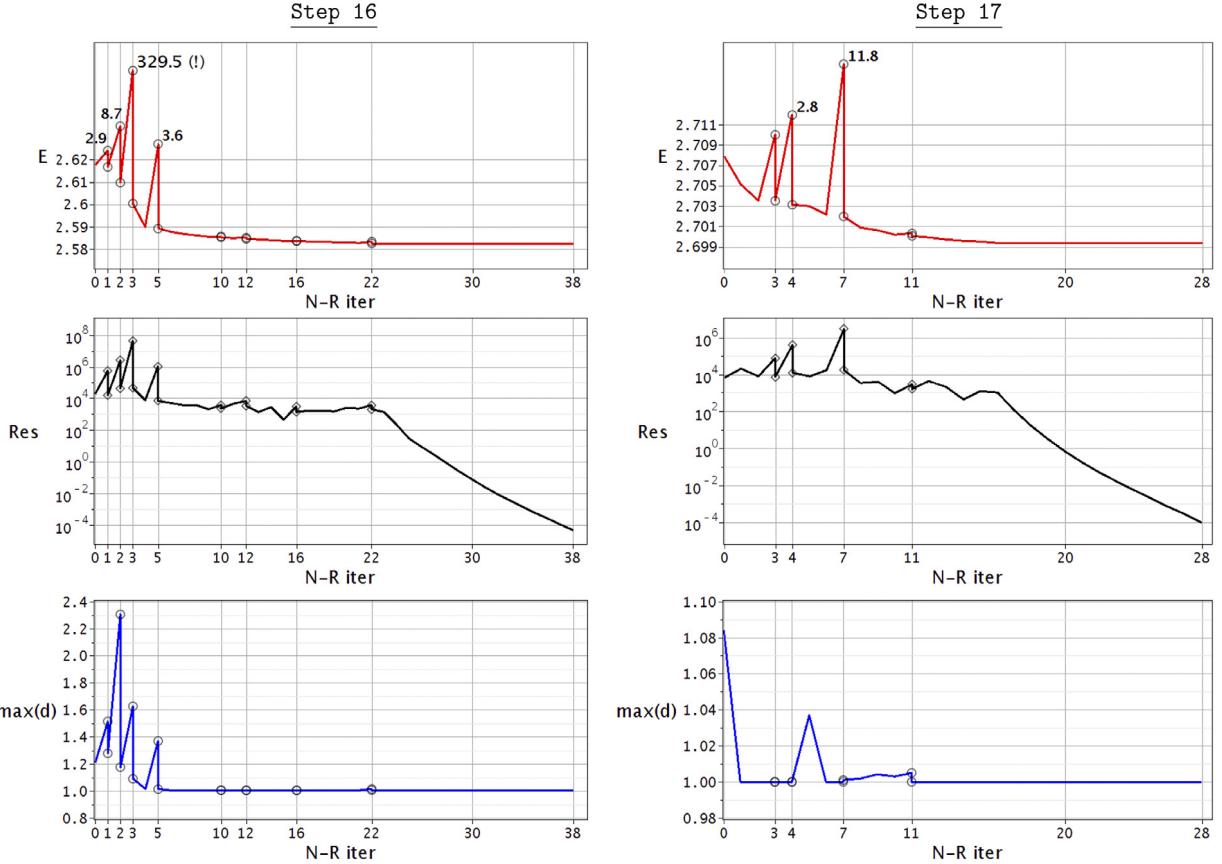


Fig. 23. MON-Is computations: convergence of E , Res and $\max(d)$ at loading steps 16 and 17.

the case, the values of f and f' are computed for all $\alpha = \alpha_k$, what is then used to recover the cubic polynomial interpolation of f in every $[\alpha_k, \alpha_{k+1}]$. In this way, globally in the corresponding intervals $[0, 1]$ and $[-1, 1]$, we construct a C^1 -approximation of f , which we denote as f_I . The candidate for α^* is then computed explicitly from equation $f'_I = 0$ along with the conditions $f''_I \geq 0$ and $0 < f_I < f(0) := E(\mathbf{u}_i, d_i)$. Also, since the solution to $f'_I = 0$ that satisfies the above properties may be not unique yielding $\{\alpha_s^*\}$, $s \geq 1$, as an ‘admissible’ set, we eventually choose $\alpha^* := \min_{\alpha_s^*} \{f_I(\alpha_s^*)\}$. This will be demonstrated later.

The described procedure can be obviously modified in several ways. A uniform discretization of the search intervals $[0, 1]$ and $[-1, 1]$ into 10 and 20 segments, respectively, is heuristic and can be changed considering that the computation of $f(\alpha_k)$ and $f'(\alpha_k)$ at many interpolation points may be quite expensive. In the classical line search procedure depicted in Fig. 17, $\Psi(\mathbf{u}_i + \alpha \Delta \mathbf{u}_i)$ is typically interpolated or extrapolated in $[0, 1]$ at only a few points and an approximate candidate for α^* can be extracted using e.g. a so-called *regula falsi* (or secant) procedure. Given the strong non-linearity and non-convexity of the functional E we deal with, we did not focus on efficiency but rather on accuracy in the evaluation of α^* , in order to demonstrate the applicability of the proposed line search idea.

4.3. MON-Is computational results

In this section we demonstrate the proposed line search procedure. The loading setup is the same as in Section 3.3.1 for which the reference STAG results were obtained and, to recall, $\text{Tol} = 10^{-4}$. A detailed comparison of both approaches in terms of accuracy and computational cost is given in the following.

Fig. 20 depicts the results of STAG and MON-Is computations and the zoom into the region of interest. First, it can be seen that the first peak load at step 11 is successfully overcome and the MON-Is computations continue to achieve convergence. The two curves are very close.

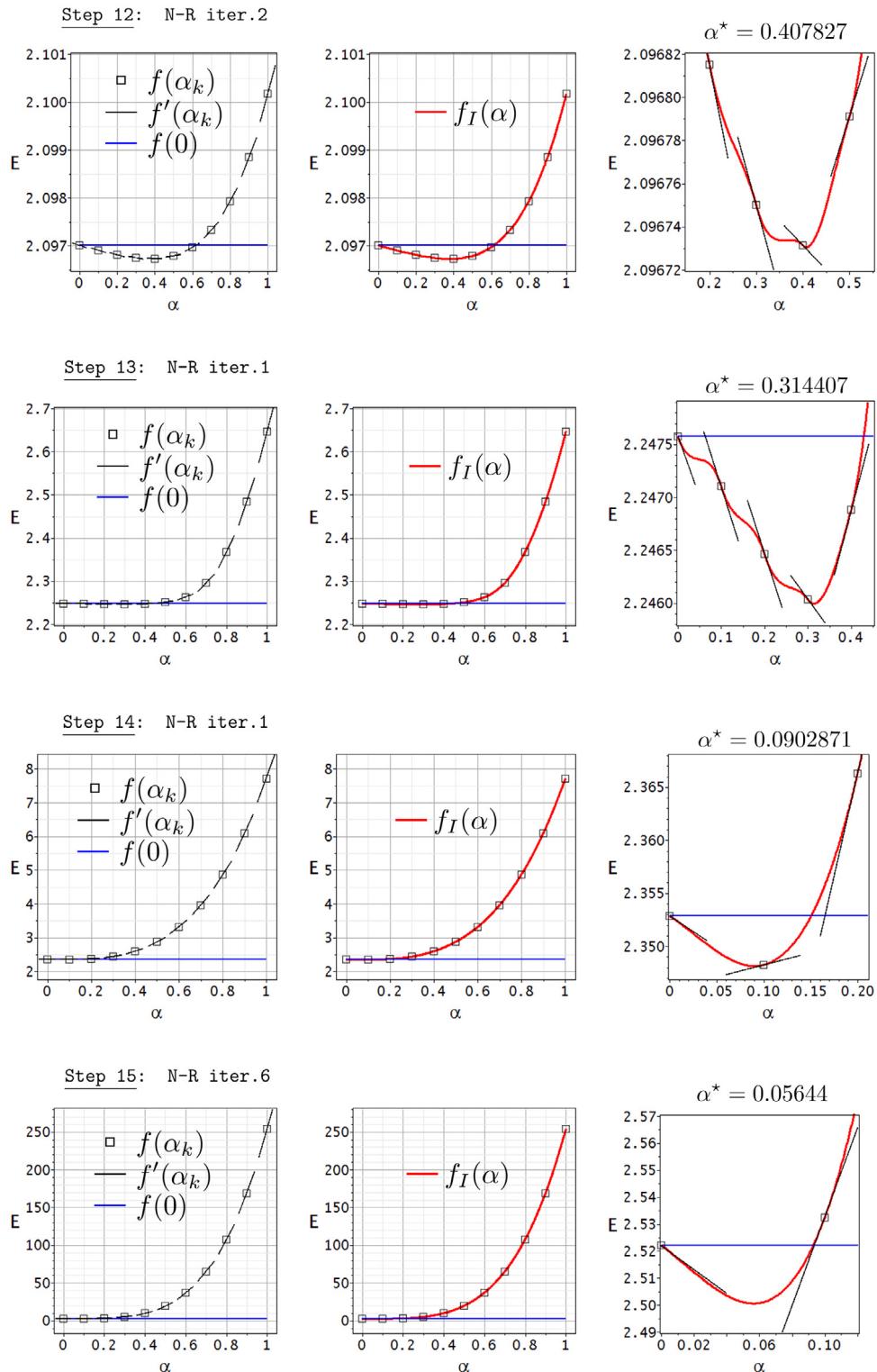


Fig. 24. Detection of α^* using f_I in Case [A] ($f'(0) < 0$); note that e.g. in step 12 a candidate from the interval $\alpha \in (0, 1)$ is not unique.

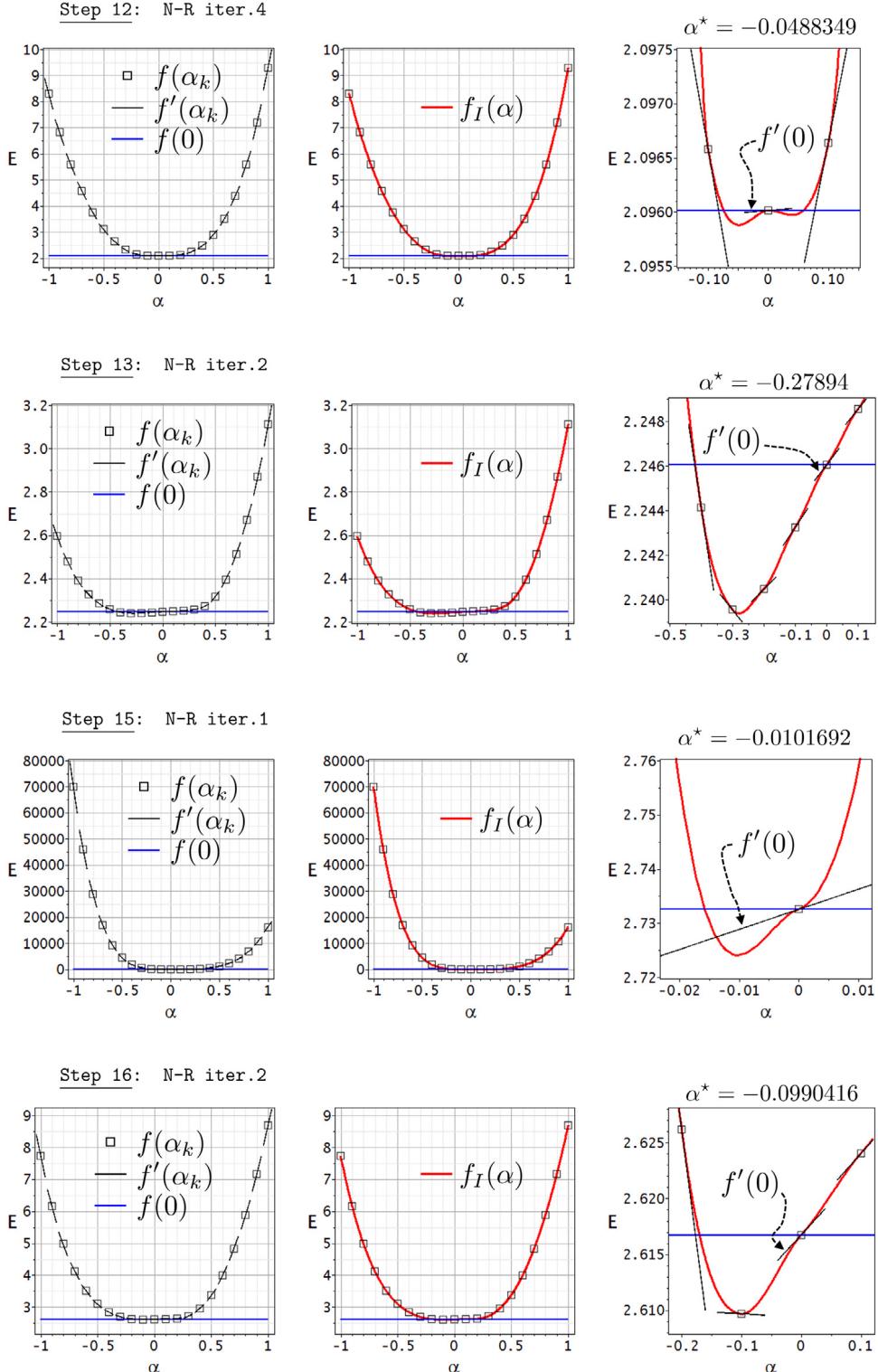


Fig. 25. Detection of α^* using f_I in Case [B] ($f'(0) > 0$): the condition motivates the need to seek for the candidate also in the interval $\alpha \in (-1, 0)$; note that in step 12 a positive α^* also exists (yet, it is not optimal), whereas in steps 13, 15 and 16 α^* is uniquely negative.

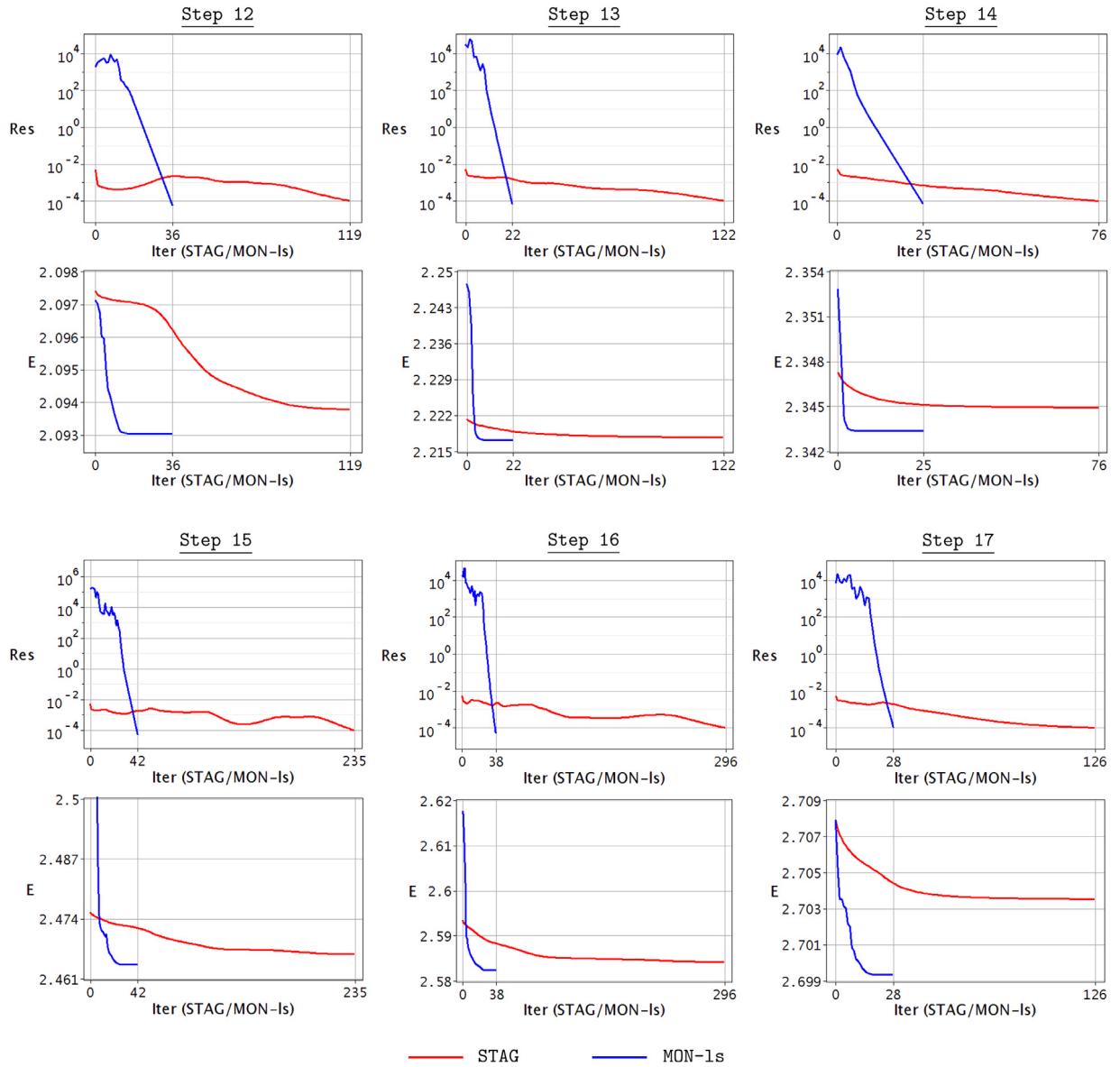


Fig. 26. Comparison of residual and energy convergence with the STAG and the MON-Is procedures at steps 12–17.

The following observations emerge from steps 12–17 for which the convergence of quantities E , Res and $\max(d)$ within the N–R iterative process is illustrated in Figs. 21–23:

- the proposed line search procedure is capable of recovering convergence in extremely critical situations $E_i < E_{i+1}$ as can be seen e.g. at step 12 (iteration 4), step 13 (iteration 2), step 14 (iteration 1), step 15 (particularly, iterations 1, 2, 6), step 16 (iteration 2, 3) and step 17 (iteration 7). With a suitable α^* , the desired trend $E_i > E_{i+1}$ is recovered even after very large values of E ;
- when an increase of energy at a N–R iteration is registered (which triggers the activation of the line search procedure), a similar increase typically occurs for Res. In case of d (the latter is depicted in terms of $\max(d)$) an extremely critical situation occurs, namely, $\max(d) > 1$. The proposed line search yields, for all controlled and depicted quantities, a return back to ‘normal’ behavior;
- if a natural (i.e. not line search assisted) decrease of energy from iteration i to $i + 1$ occurs, $E_i > E_{i+1}$, a slight increase of the corresponding residual, i.e. $\text{Res}_i < \text{Res}_{i+1}$, may be registered (this is quite natural, considering that

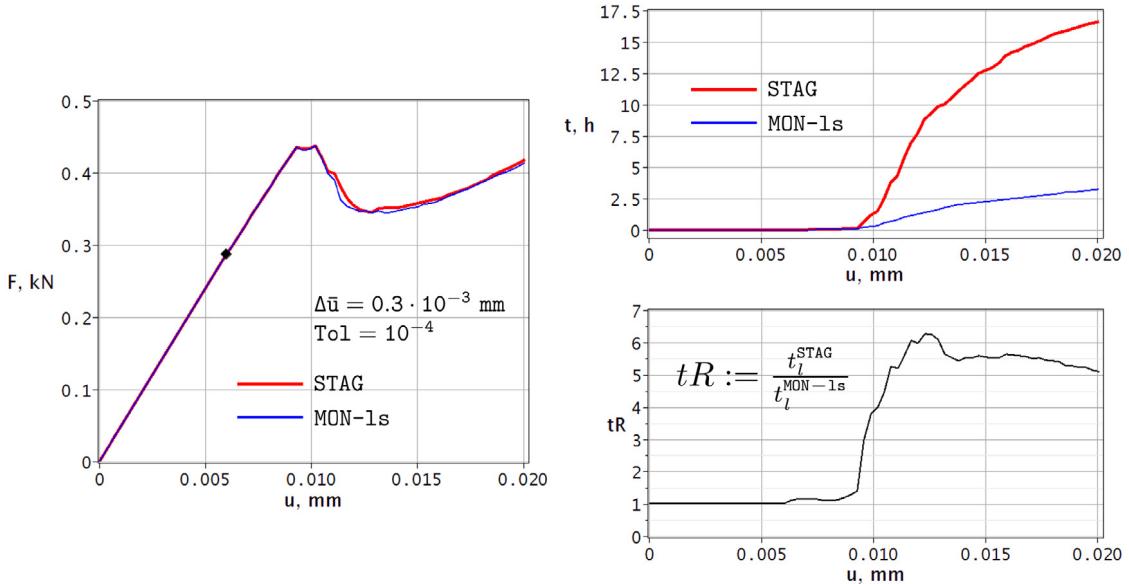


Fig. 27. STAG and MON-Is load–displacement curves in the entire range of loading and related time–displacement diagram for the comparison of cumulative computational cost within each approach.

we deal with non-convex minimization), as can be seen e.g. at steps 12, 13, 15–17. Note that in this situation we do not activate the line search as the desired condition $E_i > E_{i+1}$ has been already fulfilled;

- it is typical that after a series of iterations with an active line search, the standard N–R takes over and leads to the desired behavior for all depicted quantities: a pronounced asymptotic convergence of E , a monotonic decrease of Res and a stabilization of d ($\max(d) \approx 1$);
- It can be noticed that the residual convergence rate is not asymptotically quadratic as in the standard N–R procedure. Such deteriorated rate is most likely due to the strong non-convexity of E ;
- it may take a significant number of iterations (up to 40, as e.g. at steps 15 and 16) until the corresponding tolerance $Tol = 10^{-4}$ for Res is achieved. However, as will be shown below, convergence of Res and E within the MON-Is iterative process still appears to be much faster than with the STAG one, at a fixed loading step.

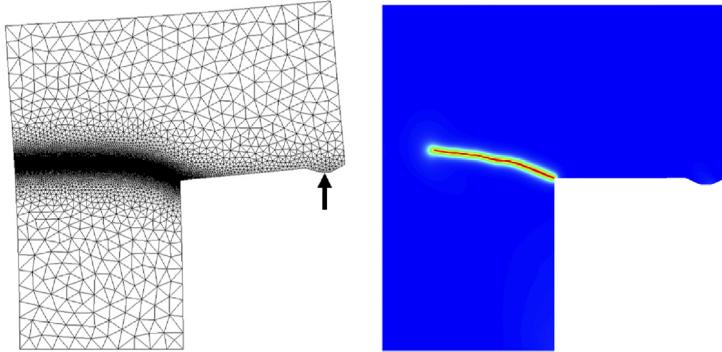
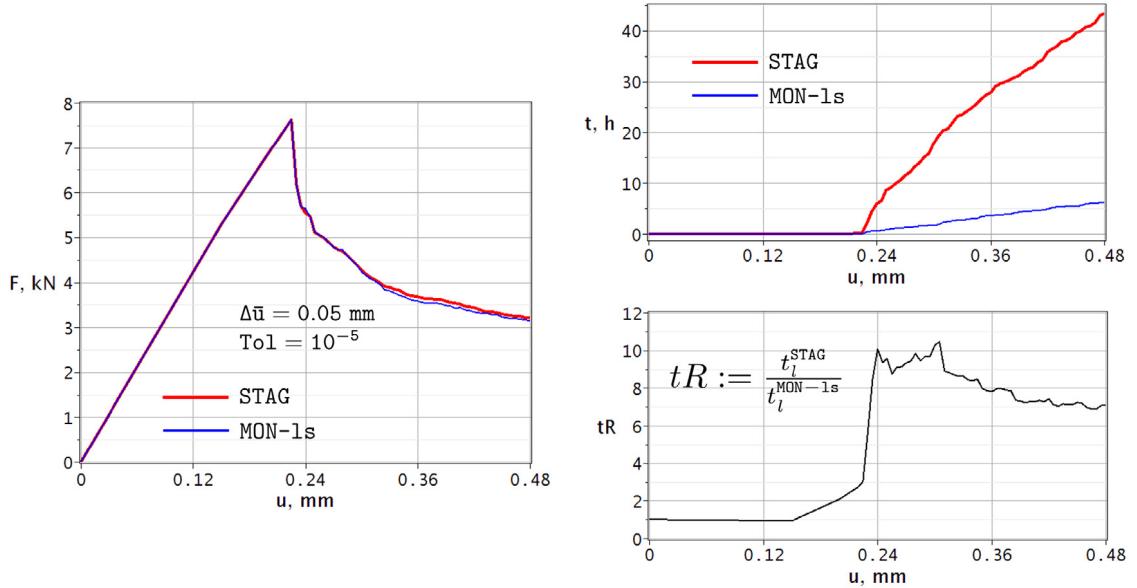
Figs. 24 and 25 illustrate the evaluation of α^* for some pairs ‘Step # – N–R iteration #’ from Figs. 21–23 according to the algorithm depicted in Section 4.2.1. The figures are arranged with respect to Cases [A] and [B]. Note that in some cases more than one candidate for α^* can be obtained.

Fig. 26 compares the convergence of residual Res and energy E within the STAG and MON-Is iterative processes at loading steps 12–17. The corresponding results in either approach were already presented in Figs. 7–9 and Figs. 21–23, respectively, and are now merged for every mentioned step. It can be observed that even though the MON-Is process typically starts with higher residual and energy levels, it converges to the same level of Res and (approximately) the same minimum value of E faster than the STAG one.

Finally, Fig. 27 illustrates and compares the computational time for the STAG and the MON-Is approaches. The time–displacement diagram clearly indicates that the ratio in cumulative computational time between the two approaches increases. With the same loading setup and under comparable accuracy requirements expressed through the corresponding tolerance Tol the latter scheme appears far superior to the former one, at least for the benchmark example considered.

4.4. Two additional examples at glance

We finalize this section by briefly reporting the results obtained with the proposed MON-Is approach for (a) an L-shaped panel and (b) a notched plate with hole. Geometric and material data, as well as loading setup are taken from [1]. Our goal here is to confirm the good performance of the approach and to show its superiority with respect to the STAG one. The corresponding loading setup and the phase-field evolution pattern are depicted in Figs. 28 and 30.

Fig. 28. *L*-shaped panel: sketch of the loading setup and final crack pattern.Fig. 29. *L*-shaped panel: STAG and MON-1s load–displacement curves in the entire range of loading and related time–displacement diagram for the comparison of cumulative computational cost within each approach.

In either case, we sketch the deformed configuration of the specimen before onset of fracture. Figs. 29 and 31 report and compare the load–displacement curves, including the loading increment $\Delta\bar{u}$ and Tol parameters, as well as the time–displacement diagrams obtained both for STAG and MON-1s approaches in each case.

5. Conclusions and outlook

In this paper, we proposed a line-search assisted monolithic approach for the solution of the coupled system of non-linear equations arising from the phase-field brittle fracture formulation. The need for a monolithic solution scheme was motivated based on a thorough analysis of the performance of the staggered scheme on the chosen benchmark. The staggered scheme (or alternate minimization procedure) is very robust due to the convexity of the free energy functional with respect to each of the two unknown fields separately. However, the convergence of energy requires a significant amount of iterations.

Relaxation of the convergence tolerance, while reducing the computational cost, may lead to significant inaccuracies in the computation of the phase field as well as of the global response. Also, due to the nested iteration loops the staggered procedure appears unsuitable for the implementation of fully adaptive refinement strategies, see [24] and [20].

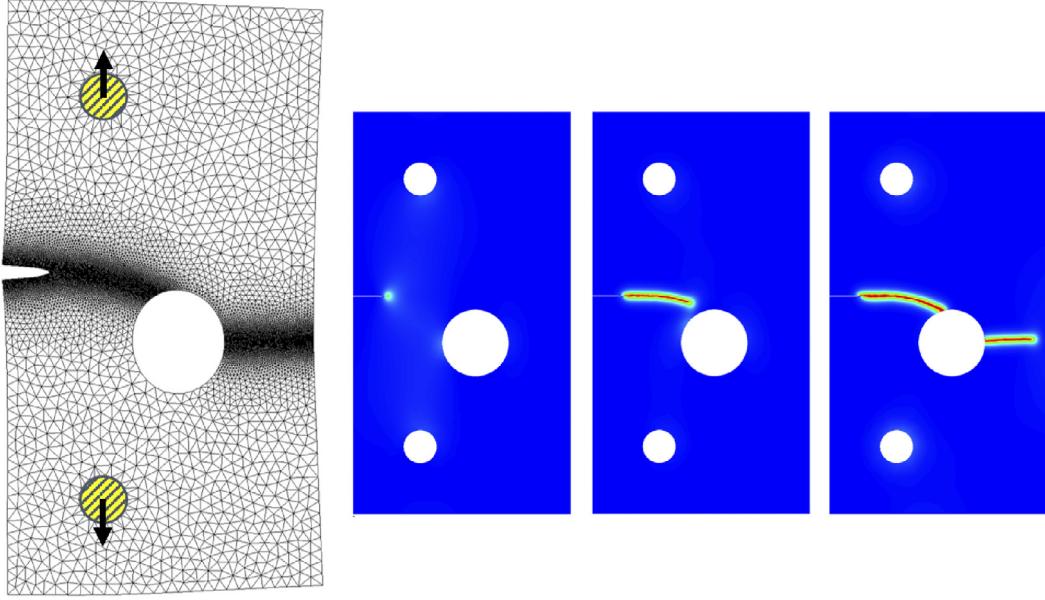


Fig. 30. Notched plate with hole: sketch of the loading setup and stages of crack propagation.

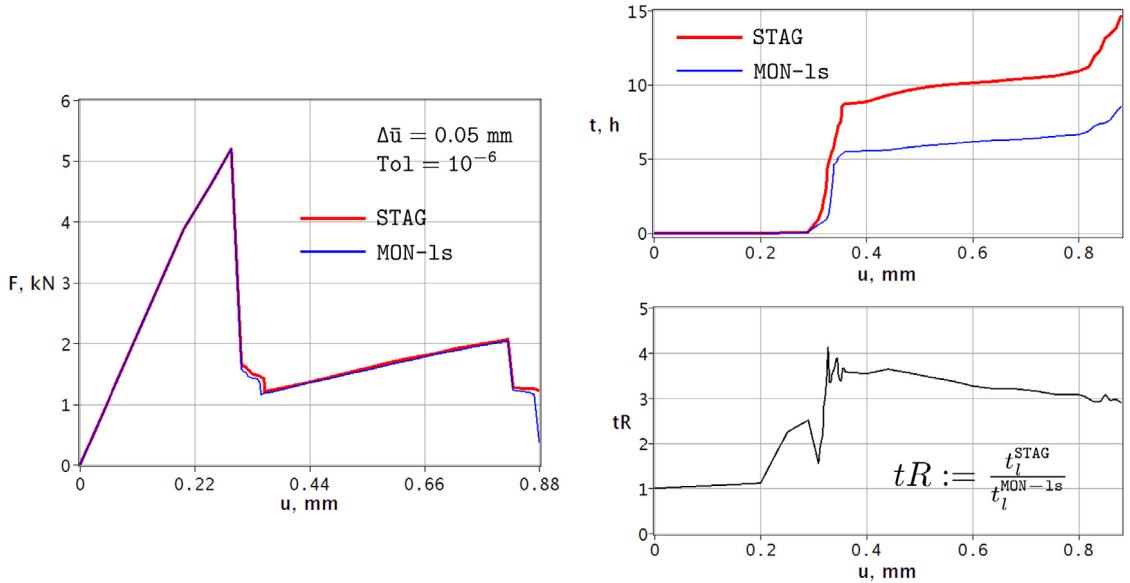


Fig. 31. Notched plate with hole: STAG and MON-1s load–displacement curves in the entire range of loading and related time–displacement diagram for the comparison of cumulative computational cost within each approach.

On the other hand, the standard monolithic solution scheme based on the Newton–Raphson iterative procedure suffers from lack of robustness, with loss of iterative convergence already at early stages of the crack development. To solve this issue, we proposed a non-standard line search approach which is activated whenever an energy increase between subsequent iterations is detected. A single line search parameter is determined for both the displacement and the phase field increments by minimizing the energy. Unlike in standard line search procedures, in our approach the line search parameter is allowed to take negative values, which not only does not compromise the admissibility of the solution but also proved to play an important role for the robustness of the procedure. Minimization of the energy is conducted through the evaluation of a scalar function and of its derivative at a user-defined number of points

and through subsequent interpolation. The proposed method was shown to be able to recover iterative convergence in situations where catastrophic divergence would have occurred, and to enable an accurate computation of the benchmark example with different sizes of the load increment, at a significantly reduced computational cost with respect to the staggered scheme.

Further extensions of the present research might include additional verification with more complex (e.g. 3d) examples, the adoption of a two-parametric line search procedure, as well as the combination with fully adaptive mesh refinement schemes, the latter being the primary motivation of this work.

Acknowledgment

This research was funded by the European Research Council, ERC Starting Researcher Grant INTERFACES, Grant Agreement No. 279439.

References

- [1] M. Ambati, T. Gerasimov, L. De Lorenzis, A review on phase-field models of brittle fracture and a new fast hybrid formulation, *Comput. Mech.* 55 (2015) 383–405.
- [2] H. Amor, J.-J. Marigo, C. Maurini, Regularized formulation of the variational brittle fracture with unilateral contact: Numerical experiments, *J. Mech. Phys. Solids* 57 (2009) 1209–1229.
- [3] C. Miehe, F. Welschinger, M. Hofacker, Thermodynamically consistent phase-field models of fracture: variational principles and multi-field FE implementations, *Internat. J. Numer. Methods Engrg.* 83 (2010) 1273–1311.
- [4] C. Miehe, M. Hofacker, F. Welschinger, A phase field model for rate-independent crack propagation: robust algorithmic implementation based on operator splits, *Comput. Methods Appl. Mech. Engrg.* 199 (2010) 2765–2778.
- [5] B. Bourdin, G.A. Francfort, J.-J. Marigo, Numerical experiments in revisited brittle fracture, *J. Mech. Phys. Solids* 48 (2000) 797–826.
- [6] B. Bourdin, Numerical implementation of the variational formulation for quasi-static brittle fracture, *Interfaces Free Bound.* 9 (2007) 411–430.
- [7] B. Bourdin, G.A. Francfort, J.-J. Marigo, The variational approach to fracture, *J. Elasticity* 91 (2008) 5–148.
- [8] A. Mesgarnejad, B. Bourdin, M.M. Khonsari, Validation simulations for the variational approach to fracture, *Comput. Methods Appl. Mech. Engrg.* 290 (2015) 420–437.
- [9] G.A. Francfort, J.-J. Marigo, Revisiting brittle fractures as an energy minimization problem, *J. Mech. Phys. Solids* 46 (1998) 1319–1342.
- [10] C. Kuhn, R. Müller, A continuum phase field model for fracture, *Eng. Fract. Mech.* 77 (2010) 3625–3634.
- [11] M.J. Borden, T.J.R. Hughes, C.M. Landis, C.V. Verhoosel, A higher-order phase-field model for brittle fracture: Formulation and analysis within the isogeometric analysis framework, *Comput. Methods Appl. Mech. Engrg.* 273 (2014) 100–118.
- [12] J. Vignollet, S. May, R. de Borst, C.V. Verhoosel, Phase-field models for brittle and cohesive fracture, *Mechanica* 49 (2014) 2587–2601.
- [13] T. Heister, M.F. Wheeler, T. Wick, A primal–dual active set method and predictor–corrector mesh adaptivity for computing fracture propagation using a phase-field approach, *Comput. Methods Appl. Mech. Engrg.* 290 (2015) 466–495.
- [14] C. Verhoosel, J.J.C. Remmers, M. Gutierrez, A dissipation-based arc-length method for robust simulation of brittle and ductile failure, *Internat. J. Numer. Methods Engrg.* 77 (2009) 1290–1321.
- [15] M.A. Crisfield, Non-Linear Finite Element Analysis of Solids and Structure, in: Essentials, vol. 1, John Wiley & Sons, 2000.
- [16] O.C. Zienkiewicz, R.L. Taylor, The Finite Element Method for Solid and Structural Mechanics, Elsevier, Butterworth–Heinemann, 2005.
- [17] J. Bonet, R.D. Wood, Nonlinear Continuum Mechanics for Finite Element Analysis, Cambridge University Press, 2008.
- [18] P. Wriggers, Nonlinear Finite Element Methods, Springer-Verlag, 2008.
- [19] R. de Borst, M.A. Crisfield, J.J.C. Remmers, C.V. Verhoosel, Non-Linear Finite Element Analysis of Solids and Structure, John Wiley & Sons, 2012.
- [20] M. Artina, M. Fornasier, S. Micheletti, S. Perotto, Anisotropic mesh adaptation for crack detection in brittle materials, *SIAM J. Sci. Comput.* (37) (2015) 633–659.
- [21] F. Hecht, A. Leharic, O. Pironneau, FreeFem++: Language for Finite Element Method and Partial Differential Equations (PDE), Université Pierre et Marie, Laboratoire Jacques-Louis Lions, <http://www.freefem.org/fem++/>.
- [22] P.E. Farrell, C. Maurini, Linear and nonlinear solvers for variational phase-field models of brittle fracture, submitted for publication.
- [23] P. Erbts, A. Düsler, Accelerated staggered coupling schemes for problems of thermoelasticity at finite strains, *Comput. Math. Appl.* 64 (2012) 2408–2430.
- [24] S. Burke, C. Ortner, E. Süli, An adaptive finite element approximation of a variational model of brittle fracture, *SIAM J. Numer. Anal.* 48 (2010) 980–1012.