Technische Universität München

Fakultät für Mathematik

Lehrstuhl für Angewandte Numerische Analysis

Lagrangian Methods for Constrained Non-Convex Minimizations and Applications in Fracture Mechanics

Marco Artina

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Martin Brokate
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Massimo Fornasier
	2.	O. Univ.-Prof. Dr. Karl Kunisch Karl-Franzens-Universität Graz/ Österreich (schriftliche Beurteilung)
	3.	Prof. Dr. Antonin Chambolle CMAP, Ecole Polytechnique Palaiseau Cedex/ Frankreich

Die Dissertation wurde am 31.08.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 30.11.2015 angenommen.

Abstract

We address the consistent numerical simulation of quasi-static evolutions of critical points of nonconvex and nonsmooth energy functionals.
We focus on the reliable numerical simulation of linearly constrained nonsmooth and nonconvex optimizations. We use these tools as building blocks for the construction of quasi-static evolutions. We apply the machinery on a model of cohesive fractures, giving an alternative and constructive proof of existence.
We consider models of brittle fractures and their adaptive anisotropic numerical approximations.

Acknowledgments

This work would not have been possible without the framework and the financial support of the IGDK 1754 (International Research Training Group), which I fell here greatly to acknowledge. In particular, I am very grateful to Prof. B. Vexler, who took the burden to coordinate together with Prof. K. Kunisch the entire graduate school. Within the IGDK I got exposed to several complementary research topics, which greatly widen my perspective beyond the scope of my thesis.

I would like to thank Prof. M. Fornasier for supervising me during these three years of Ph.D. studies. His support and advices have been fundamental for my personal and academic growth. He also guided me in all the cooperations and the developed work, but allowed me also all the independence I needed. He additionally transmitted me his passion, giving me the energy to step into a new topic and get fascinated by all the challenges it was revealing as soon as the work got more involved.

I am greatly in debt with Prof. K. Kunisch for his generous and patient collaboration to my thesis, especially during the visits I made at the university of Graz. His different perspective allowed me to develop alternative approaches to the problems addressed in the thesis, greatly widening its scope.

It is very important for me to thank Prof. S. Perotto and Prof. S. Micheletti (MOX - Politecnico di Milano) for their welcoming hospitality and for sharing with me their expertise on anisotropic finite elements, which allowed me to develop a significan part of my thesis.

Prof. P. Markowich, John von Neumann visiting Professor at our chair, who illustrated me different fields in which the knowledge I acquired during the studies at the university could be applied. I thank him for his commitment, the cooperation, and the fruitful discussions we had.

Francesco Solombrino, my mentor, my friend, my first Italian contact also outside the TUM after I moved to this city. I am truly thankful to him for helping me both on the academic and the human side. He never refused to spend some to listen to my problems and my doubts and he always gave me very fruitful answer that helped my to successfully conclude this thesis.

A special thank goes to Filippo Cagnetti. We worked together to develop a part of the project which ends with this thesis. He openly shared with me his huge expertise in co-

hesive fracture evolutions, and guided and inspired the idea of exploiting the algorithm we developed as a buiding block to constructively prove the existence of quasi-static evolutions for cohesive fractures.

Let me know spend a couple of words about my colleagues at M15 and IGDK, in particular, G. Albi, M. Bongini, J. Sigl, S. Peter, M. Keuthen, and K. Pieper. Together we built more than a research group, a real small community in which we discussed about problems, both personal and professional, elaborated new ideas, as, i.e., the students' workshop, and have very fruitful meetings where ideas and innovative solution have been created. Thank to everybody for being part of this group.

I thank my mother and my father. I know they trust me and this reassure me about my future choices and life.

My brother, Andrea, deserves a big sorry. I know that when I moved to Germany he felt a bit abandoned. I wish I managed to make him feeling that I was much more present in his life.

Veronica, thank you. During the last years she was always next to me, giving the moral and psychological support. She never refused to help me and she also played an important role in writing this work, giving me some tips to improve the style of the dissertation.

Last but not least, Matteo. I think that a normal thank is not enough. You decided to begin this adventure with me, you moved to Munich facing many challenging, starting from learning a new language and a new culture. Despite that, in less than one year you got more connection than what I did in three years. You made me appreciate this city and you helped me to build my very first house, the first place where I really fell home after I left my parent's house. In particular, I thank you for your comprehension in the last months. I know that when I am stressed I can be the worst person to live with, but you never made me feel uncomfortable.

Contents

1. Introduction to fracture mechanics **8**

1.1. Linear elasticity . . . 12

1.1.1. Equilibrium configuration . . . 15

1.1.2. Quasi-static evolution . . . 16

1.2. Fracture mechanics . . . 17

1.2.1. Griffith's Theory . . . 18

1.2.2. Equilibrium configuration and accessibility condition . . . 20

1.2.3. Francfort–Marigo brittle fracture model . . . 22

1.2.4. Cohesive fracture models . . . 25

2. Linearly constrained minimization of nonconvex and nonsmooth functionals **29**

2.0.5. General setting . . . 31

2.1. Nonconvex Augmented Lagrangian algorithm . . . 33

2.1.1. Augmented Lagrangian algorithm for convex problems . . . 34

2.1.2. The algorithm in the nonconvex case . . . 36

2.1.3. Analysis of convergence . . . 37

2.2. Nonconvex Split Bregman Iteration . . . 39

2.2.1. Split Bregman Iteration for convex functionals . . . 40

2.2.2. Split Bregman Iteration for nonconvex functionals . . . 47

2.2.3. Analysis of convergence . . . 48

3. Some applications of the algorithm **50**

3.1. Free-discontinuity problems . . . 50

3.1.1. The Mumford–Shah functional in image processing . . . 51

3.1.2. Quasi-static evolution of brittle fractures . . . 53

3.2. Truncated polynomial minimization . . . 54

3.2.1. Compressive sensing . . . 55

3.3. Numerical issues: smoothing and iterative thresholding procedures . . . 60

3.3.1. A smoothing procedure . . . 61

3.3.2. The application of the algorithm to coercive cases . . . 65

3.3.3. Iterative thresholding algorithms revisited . . . 67

3.4. Numerical Experiments . . . 72

3.4.1. Mumford–Shah functional minimization for image denoising . . . 73

3.4.2. Brittle fracture simulation 75
3.4.3. Damping noise-folding in compressive sensing 77

4. Quasi-static evolution of cohesive fracture models 86
4.1. The simulation of quasi-static evolution of a Barenblatt cohesive model . 87
4.1.1. Preliminary assumptions 89
4.1.2. Existence of approximable quasi-static evolution 91
4.1.3. An application to a Barenblatt cohesive fracture model 100
4.1.4. Recovering an approximable quasi-static evolution 109
4.1.5. Numerical experiments 115
4.2. The simulation of quasi-static evolution of a Dugdale cohesive model . . . 121
4.2.1. Discretization schemes and optimality conditions 123
4.2.2. Two efficient minimization algorithms 129
4.2.3. Numerical experiments 133
4.3. Conclusions 135

5. Variational formulation of brittle fractures 138
5.1. The Ambrosio–Tortorelli approximation of the Mumford–Shah functional 139
5.1.1. The finite element discretization 143
5.2. Anisotropic finite element setting and error estimator 145
5.2.1. The anisotropic background 146
5.2.2. An a posteriori error estimator 148
5.3. The numerical procedure 153
5.3.1. The minimization algorithm 154
5.3.2. The mesh adaptive procedure 155
5.3.3. The whole adaptive procedure 157
5.4. Numerical experiments 159
5.4.1. The antiplane experiments 160
5.4.2. Parameter sensitivity 168
5.4.3. The plane experiments 170
5.5. Conclusions 178

A. Complementary notions and definitions 179
A.1. Functional spaces and auxiliary definitions 179
A.1.1. Subdifferential and ν-convexity 181
A.2. Some notions about compressive sensing 183

B. A posteriori anisotropic error estimator of the plane displacement 185
B.1. The mathematical model of plane-strain fracture 185
B.2. A posteriori anisotropic error analysis 187

List of Figures **193**

Bibliography **195**

1. Introduction to fracture mechanics

Why does a dish used hundreds of times decide to break instantaneously despite it did not seem to be damaged? Why does a fracture start propagating in a glass or in a car's windscreen apparently without any accidental reason? Why did in 1967 the silver bridge connecting Point Pleasant, West Virginia with Kanauga, Ohio collapse [47]? How is it possible that the fuselage of an aircraft after years of safe flights suddenly breaks [62, 138, 212]? What is the reason behind the unexpected hull cracks of the US army SS John P. Gaines Liberty ship which sank on 24 November 1943 with the loss of 10 lives [145]?
Everyone, reading the first two questions, may have though: "well, it may happen. Eventually, you need to go to the shop and buy a new glass or dish". But the last three queries raise different emotions and we all hope that the next bridge, aircraft, or ship we may take have been built in a better manner than the ones that were involved in those catastrophic events.
Despite the apparent heterogeneity of the problems, they actually present an underlying phenomenon which is the direct responsible of all these failures: the appearance of a fracture in the material. Therefore, a study which aims to predict how a crack initiates and evolves in a body can be not only interesting for the design of more resistant everyday products, but can also help in reducing the possibility that a tragedy happens and thus saving also human lives.
The first studies about fractures date back to the 1920s, when the British Royal Aircraft Establishment commissioned to the engineer A. A. Griffith an investigation on the correlation between the metal surface treatments and structural failures of aircraft's bodies experienced during the World War I. This can be considered as the birth of a new branch of continuum mechanics: the fracture mechanics. Since then, many different models raised to predict when and where a fracture may appear and grow.
The multitude of models explaining the crack formation, may be distinguished and classified as follows.

- *Quasi-static evolution:* The external forces acting on the system is slowly changing in time and the body is assumed to be instantaneously at an equilibrium configuration ([78, 109, 110]).

- *Dynamic evolution:* The external action which causes the fracture is rapidly changing in time. The system is not in an equilibrium configuration and we may also

observe a wave propagation in the material due to the extreme variations of the forcing term ([44, 75, 150, 151, 153]).

- *Brittle fractures:* The specimen breaks instantaneously without any premonitory sign. After the deformation driven by the external force, the molecular bonds break at once and the two parts of the fractured body do not interact any more ([77, 78, 109, 110, 124, 175]).
- *Ductile fractures:* it can be observed a progressive weakening, up to the failure of the specimen, presenting a damaged area around the crack surface ([89, 115]).
- *Cohesive fractures:* A subclass of the family of brittle fractures, the crack forms in three phases. A deformation without fracture, a pre-fractured state in which a bridging force attracts the two lips of the crack, and the fractured phase ([26, 51, 52, 80, 96]).
- *Plastic deformation:* When an external force acts on the specimen, it deforms its shape irreversibly. In this case, not every deformation implies a loss of the structural integrity ([134, 156, 203]).
- *Microstructure-fractures:* The applied deformation creates a multitude of fractures on a molecular level and therefore invisible at the human eye. Despite the shape and the condition of the specimen seems to be perfect, the structural properties of the system is irreversibly compromised ([92, 93, 186]).
- *Thermal-fracturing:* The fracture is not caused by a force acting on the system but by a thermal shock. A fragile body immersed in a fluid at a very different temperature is subject to internal stress that may cause its fracturing. ([45, 187, 197, 205])
- *Fatigue:* The material is weakened by repeatedly applied loads. In this case the damage is progressive and localized and it occurs when a material is subjected to cyclic loading. The failure may appear also when the applied load is less than the maximum load the system was designed for ([35, 36, 144, 168, 206]).

The list above, although it does not pretend to give a complete overview of the different ways a body can break, gives an idea of the variety and complexity of the phenomena that fracture mechanics studies.
In this work, we focus on quasi-static evolutions and consider brittle and cohesive modes of fracture, with the specific aim of developing efficient algorithms to generate reliable numerical simulations of the crack initiation and propagation. In particular, the main challenges we deal with are: *1)* retrieve a physically sound quasi-static evolution of the system, and *2)* obtain a deformation of the specimen which is not biased by the numerical procedure and discretization adopted. The issues related to these points are

intrinsically connected with the nature of the problem, since, in general, we have to deal with the minimization of functionals that are nonconvex and present singularities.
To discuss the challenges of point *1)*, we observe that, in nature, often evolutions of physical phenomena proceed along energy local minimizers. Therefore, we designed a constructive procedure, an iterative algorithm, computing and selecting critical points, balancing proximity to the initial iteration and energy level. Used as a building block in quasi-static evolutions our algorithm promotes relatively smooth trajectories of critical points and physically sound solutions.
Point *2)* is conversely strictly related to the second characteristic of the fracture problem: the presence of singularities. Aiming to correctly detect a fracture path is indeed a non trivial problem from a numerical point of view, as we need to locate the place and the time, where and how a singularity appears and evolves in the domain. If, for example, we knew in advance where the crack may form, we could properly define the discretization of the domain in order to handle at best the singularity. Unfortunately, being interested in the case in which the fracture is free to form and evolve in the specimen, creating a discretization scheme which is both accurate and efficient is definitively a non trivial task.

With the present work, which is representing a self-contained scientific compendium of our papers [13, 14, 16, 15, 18, 19, 20], we make an attempt toward efficient and reliable methods which are able to capture and describe the quasi-static evolution of fractures, being aware that all the topics and arguments here presented can be extended in the future.
In the next sections of this chapter we introduce the reader to fracture mechanics. We first discuss linear elasticity which describes, in our setting, the deformation of the body in the unbroken part. After that, we give a historical overview of fracture models. Finally, we introduce the specific models that we will study in the next chapters.
Chapter 2 (see also [19]) deals with the general problem of linearly constrained minimization of nonconvex and non smooth objective functionals. We introduce two algorithms, being the second a generalization of the first one. Despite the very mild assumptions on the regularity of the functional, we prove unconditional convergence of the method to critical points.
Examples of a wide range of applications, even beyond fracture mechanics, in which these algorithms may be successfully used are reported in Chapter 3 (see also [18, 19]), where problems such as image denoising and reduction of the noise-folding phenomenon in compressive sensing are addressed.
Next, in Chapter 4 (see also [13, 20]), we start working with fracture models. In particular, we focus first on cohesive fractures. In the first part of the chapter, we show that the first algorithm presented in Chapter 2 not only can be effectively used as a building block to simulate reliably also this kind of problems, but also we show that it is possible

to retrieve a quasi-static evolution which corresponds to a physically admissible process and also converges to the continuous problem.
The second part is then devoted to the computational efficiency. Indeed, one of the drawback of the considered algorithm is that the computational times tend to be huge also for relatively simple problems. Therefore, we revise two well-known algorithms for convex optimization, the Split Bregman Iteration and the Primal Dual Active Set, and we show how they can be adjusted to solve our problem efficiently as well, despite the nonconvexity. The convergence of these new algorithms is still an open problem. However, we never experienced in the numerical simulation any failure of the convergence and we obtained agreement with the physical predictions.
The last chapter, Chapter 5 (see also [14, 16, 15]), is focused on brittle fractures, and faces the second aim of the project: how to discretize the domain in an efficient and reliable fashion, in order to detect the crack path in a quasi-static evolution. In particular, resolving to consider an anisotropic mesh adaptation scheme, we propose a procedure which not only correctly detects the crack path, but also produces a domain discretization with a minimal amount of degree of freedom.
We have also added two appendixes. In Appendix A we report some definitions which are fundamental for a deep understanding of the concepts described in the following but that we decided to isolate from the text to ease the reading of the individual chapters. Appendix B is an extension of Chapter 5. In particular, we obtain similar a posteriori error estimates for a different setting with respect to the one described in the chapter. Since the techniques adopted are actually really similar, we decided to move this computation to this appendix.
We additionally remark that all the chapters and the theoretical results are always supported by sections reporting numerically experiments designed to validate and challenge the introduced procedures.

1.1. Linear elasticity

Any physical body in a rest position, i.e., it does not interact with the external world, has a specific shape. As soon as some force starts interacting with it, its physical characteristics are consequently modified. These forces, which deform the specimen, may be distinguished in two categories: *volume* or *body forces*, like the gravity, which act on each particle of the specimen, and *surface* or *contact forces*, for example a load applied on a side of the body, which have a direct influence only on the surface of the mass.
The deformation the system undergoes may vary according to the material properties and the strength and the type of the action. Indeed, the body may break into parts, getting fractured; deform irreversibly, keeping the deformed shape also after the external action has ended; or deform reversibly, returning to its rest shape as soon as the force is removed. In particular, we define a material to be *elastic* if the last deformation occurs.
In this section, we give some information on linear elasticity which are fundamentals for a better understanding of the work, since the theory of fracture mechanics prescribes that the body behaves elastically before its failure. We refer to [122, 125, 148, 198, 208] for more extended treatises on linear elasticity. For a wider overview on elasticity, we mention [24], where the reader can find references and open problems, and [111, 166] and references within for a treatise on non-linear elasticity.
Let us suppose that our specimen is descibed by a simply connected domain $\Omega \subset \mathbb{R}^d$, with $d \in \{1, 2, 3\}$, whose boundary $\partial\Omega$, is Lipschitz continuous. We call $\mathbf{f} : \Omega \to \mathbb{R}^d$, the force deforming the body. The shape and the internal stresses of the domain will be modified, according to external force $\mathbf{f}$, until the system finds an *equilibrium state*. To describe the deformation of the body, we introduce the *displacement vector field* $\mathbf{u} : \Omega \to \mathbb{R}^d$ mapping the domain to its configuration. For the moment we do not allow for jumps in the displacement and therefore we assume $\mathbf{u}$ to be continuous and continuously differentiable and $\det(\partial x_i/\partial y_j)) > 0$ for all $i, j \in \{1, \dots, d\}$, where $y_i = \mathbf{u}(x_i)$. Moreover, denoting with $\mathbf{x} = (x_1, \dots, x_d)$ the Lagrangian coordinates of the points of the body in the initial configuration and with $\mathbf{y}$ the point coordinates of the deformed one, we have that

$$\mathbf{y} = \mathbf{x} + \mathbf{u}(\mathbf{x}).$$

Now, consider an infinitesimal volume dV. The action of the displacement $\mathbf{u}$ either moves the volume from $\mathbf{x}$ to $\mathbf{y}$ rigidly, or deforms it. We may distinguish two different kinds of deformations an elastic body can undergo: *stress* and *shear*. The first describes deformations orthogonal to the faces of dV, like compression and elongation. The second one entails distortions parallel to the faces, with a consequent modification of the right angles of the body, like torsion.
Additionally, let us assume stresses and strains to be small compared to the dimension of the specimen. This apparently strong assumption can be understood in the sense

of infinitesimal volume deformation analysis. Indeed, let us imagine to draw a straight line on Ω. Once the displacement $\mathbf{u}$ is applied, the straight line may be mapped into a smooth curve with radius of curvature k. If we take on the line a segment $d\mathbf{x}$ such that its length is way smaller than k, its deformed corresponding segment $d\mathbf{y}$ will be almost straight. Thus, independently of the nature of the deformation the body is subject to, we neglect all the deformations of higher order. We refer to [111, 183] and references within for a complete description of non-linear elasticity, where this assumption is not considered.
Now, we analyze in detail how the dimension of the infinitesimal volume dV is modified with respect to a given direction $\mathbf{e} \in \mathbb{R}^d$. Denoting with $\mathbf{x}_\delta := \mathbf{x} + \delta\mathbf{e}$ and consequently $\mathbf{y}_\delta := \mathbf{x} + \delta\mathbf{e} + \mathbf{u}(\mathbf{x} + \delta\mathbf{e})$ we have

$$\begin{aligned} d_{\mathbf{e}}(\mathbf{x}) &:= \lim_{\delta\to 0} \frac{\|\mathbf{y} - \mathbf{y}_\delta\| - \|\mathbf{x} - \mathbf{x}_\delta\|}{\|\mathbf{x} - \mathbf{x}_\delta\|} \\ &= \lim_{\delta\to 0} \frac{\|\mathbf{x} + \mathbf{u}(\mathbf{x}) - (\mathbf{x} + \delta\mathbf{e} + \mathbf{u}(\mathbf{x} + \delta\mathbf{e})\|}{\delta\mathbf{e}} - 1 \\ &= \lim_{\delta\to 0} \frac{\|\delta\mathbf{e} + \mathbf{u}(\mathbf{x} + \delta\mathbf{e}) - \mathbf{u}(\mathbf{x})\|}{\delta\mathbf{e}} - 1 = \|\mathbf{e} + \nabla\mathbf{u}(\mathbf{x}) \cdot \mathbf{e}\| - 1 \\ &= (1 + \mathbf{e}^t(\nabla\mathbf{u}(\mathbf{x}) + \nabla\mathbf{u}(\mathbf{x})^t + \nabla\mathbf{u}(\mathbf{x})^t \cdot \nabla\mathbf{u}(\mathbf{x}))\mathbf{e})^{1/2} - 1, \end{aligned} \tag{1.1}$$

where $\|\cdot\|$ represents the Euclidean norm. We define the *strain tensor* as

$$\boldsymbol{\epsilon}(\mathbf{u}) := \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^t + \nabla\mathbf{u}^t \cdot \nabla\mathbf{u}). \tag{1.2}$$

Considering (1.1) with respect to the canonical directions $\mathbf{e}_j$, with $j = 1, \dots, d$, we obtain a physical interpretation of the diagonal components of $\boldsymbol{\epsilon}$ by the following identity

$$d_{\mathbf{e}_j} = (1 + 2\mathbf{e}_j^t \cdot \epsilon(\mathbf{u}) \cdot \mathbf{e}_j)^{1/2} - 1 = (1 + 2\boldsymbol{\epsilon}_{jj}(\mathbf{u}))^{1/2} - 1.$$

Thanks to the first order assumption made above, we can drop the quadratic term in (1.2) introducing the *linearized strain tensor* $\boldsymbol{\epsilon} : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ defined by

$$\boldsymbol{\epsilon}(\mathbf{u}) := \frac{1}{2}(\nabla\mathbf{u}^t + \nabla\mathbf{u})$$

So far we studied the deformation described by a displacement $\mathbf{u}$ and we introduced a force $\mathbf{f} : \Omega \to \mathbb{R}^d$, which deforms the body, but without specifying the relation that connects them. This coupling is given by the *Hooke's constitutive law* that governs the connection between the *internal stresses* of the body $\boldsymbol{\sigma} \in \mathbb{R}^d \times \mathbb{R}^d$, generated by $\mathbf{f}$, and the displacement $\mathbf{u}$, describing the body deformation.

Definition 1.1 (Hooke's law)**.** *Let $\boldsymbol{\sigma}$ be the tensor of the internal stresses of the body, and $\boldsymbol{\epsilon}$ the linearized strain tensor. There exists a four-dimensional tensor $C \in \mathbb{R}^d \times \mathbb{R}^d \times$*

$\mathbb{R}^d \times \mathbb{R}^d$, *called* stiffness tensor*, containing all the information of the material elastic properties and such that*

$$\boldsymbol{\sigma} = C : \boldsymbol{\epsilon}. \tag{1.3}$$

The above identity is called Hooke's law*.*

Remark 1.2. *The diagonal components σ_{ii} represent the stress in the $i-th$ direction while the off-diagonal components give us the shear. Further, since we are considering only static equilibrium states, the sum of all the stress components acting on the d-directions with the total moment must be zero, implying the symmetry of the stress tensor. Therefore, if we consider the case $d = 3$, $\boldsymbol{\sigma}$ can be fully described by three components for the directional stress and only three components for the shear.*

Remark 1.3. *The structure of C may look complicated as it is a four-dimensional tensor, but it simplifies sensibly if we consider an* isotropic material*, that is a material with uniform properties in all directions. Indeed, in this case, the stiffness tensor is fully described only by two parameters, λ and μ, called* Lamé constants*. Thus, the Hooke's law* (1.3) *reduces to*

$$\boldsymbol{\sigma} = \lambda \boldsymbol{I} tr(\boldsymbol{\epsilon}) + 2\mu\boldsymbol{\epsilon}, \tag{1.4}$$

where $\boldsymbol{I}$ is the identity matrix and $tr(\boldsymbol{\epsilon})$ is the trace operator. For sake of simplicity, unless for some generalized case, we will consider only isotropic materials.

Let us consider $V \subseteq \Omega$ a small volume contained in Ω and let $\mathbf{f}|_V$ the restriction of $\mathbf{f}$ to V. This force generates a reaction inside the medium which can be observed by measuring the stress on the boundary ∂V. Recalling that we are analyzing only static situations and therefore the forces acting on the system must be balanced, we obtain

$$\int_V \mathbf{f}\, d\mathbf{x} = -\int_{\partial V} \boldsymbol{\sigma} \cdot \mathbf{n}\, d\mathbf{S},$$

where $\mathbf{n}$ is the outer unit normal vector to ∂V. Applying the Green's Theorem we get

$$\int_V \mathbf{f}\, d\mathbf{x} + \int_V \mathrm{div}(\boldsymbol{\sigma})\, d\mathbf{x} = 0. \tag{1.5}$$

Since (1.5) is valid for any choice of $V \subseteq \Omega$, we deduce that $\mathbf{f} = -\mathrm{div}(\boldsymbol{\sigma})$ almost everywhere.
This last equation, together with the Hooke's law (1.3) and additional boundary terms, gives us the following differential system:

$$\begin{cases} \mathrm{div}(C : \boldsymbol{\epsilon}(\mathbf{u})) = -\mathbf{f} & \text{on } \Omega \\ (C : \boldsymbol{\epsilon}(\mathbf{u})) \cdot \mathbf{n} = \boldsymbol{\sigma}^{\mathbf{n}} \cdot \mathbf{n} & \text{on } \partial\Omega_N \subseteq \partial\Omega \\ \mathbf{u} = \mathbf{g} & \text{on } \partial\Omega_D \subseteq \partial\Omega \end{cases} \tag{1.6}$$

where $\boldsymbol{\sigma}^{\mathbf{n}}$ is an external stress applied on $\partial\Omega_N$, on which we impose Neumann boundary conditions, $\mathbf{g}$ is an assigned displacement acting on $\partial\Omega_D$, on which we impose Dirichlet

conditions, $\partial\Omega_D \cap \partial\Omega_N = \emptyset$, and $\partial\Omega_D \cup \partial\Omega_N = \partial\Omega$.
In the next section, we introduce a variational approach to define stable configurations of the body. Indeed, instead of solving a differential problem, we consider the energetic level of the system and look for its minimum. In particular, this approach will be fundamental for fracture mechanics problems.

1.1.1. Equilibrium configuration

A body subject to an elastic deformation accumulates an elastic potential energy which can be expressed in terms of the displacement function

$$E_e(\mathbf{u}) := \int_\Omega W(\boldsymbol{\epsilon}(\mathbf{u}))\, d\mathbf{x} \tag{1.7}$$

where $W(\cdot)$ is the *energy density*. We assume that the function $W : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ is strictly convex, non-negative, $\mathcal{C}^1$, and homogeneous of degree $p > 1$. Moreover, being $\alpha, \beta > 0$, we require that W satisfies

$$\alpha|\boldsymbol{\epsilon}(\mathbf{u})|^p \leq W(\boldsymbol{\epsilon}(\mathbf{u})) \leq \beta(|\boldsymbol{\epsilon}(\mathbf{u})|^p + 1).$$

In particular, for linear elasticity, we may consider $p = 2$ and the elastic energy density can be explicitly written as

$$W(\boldsymbol{\epsilon}(\mathbf{u})) = \frac{1}{2}|(C : \boldsymbol{\epsilon}(\mathbf{u})) \cdot \boldsymbol{\epsilon}(\mathbf{u})|. \tag{1.8}$$

For sake of simplicity, we assume from now on the volume force $\mathbf{f}$ and the surface stress $\boldsymbol{\sigma}^{\mathbf{n}}$ to be zero. Therefore, the deformation of the medium is driven uniquely by the Dirichlet boundary conditions on the displacement. In the following we introduce the concept of equilibrium configuration. We refer to [90] for an easy one dimensional problem which may help for a better understanding of the setting.
We first define the space of admissible configurations. Defining $\mathfrak{D} \subseteq \mathbb{R}^d$ the set of admissible applicable displacement data, and being $\mathbf{g} : \mathbb{R}^d \to \mathfrak{D}$, we also introduce the *set of all possible admissible configuration* as

$$\mathfrak{U} := \{\mathbf{u} : \mathbf{u}(\mathbf{x}) \in \mathfrak{D}, \text{ for all } \mathbf{x} \in \partial\Omega_D\},$$

while for a given datum $\mathbf{g}$ the space of admissible configurations is $\mathfrak{U}_\mathbf{g} := \{\mathbf{u} : \mathbf{u}(\mathbf{x}) = \mathbf{g}(\mathbf{x}), \text{ for all } \mathbf{x} \in \partial\Omega_D\}$.

Definition 1.4 (Equilibrium configuration)**.** *Let $\boldsymbol{g} : \mathbb{R}^d \to \mathfrak{D}$. We say that $\boldsymbol{u} \in \mathfrak{U}_{\boldsymbol{g}}$ is an* equilibrium configuration *if the first variation of E_e*

$$\delta E_e(\boldsymbol{u}, \boldsymbol{\xi}) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(E_e(\boldsymbol{u} + \varepsilon\boldsymbol{\xi}) - E_e(\boldsymbol{u})) \tag{1.9}$$

is non-negative for all the perturbations $\boldsymbol{\xi} \in \mathfrak{U}_{\mathbf{0}}$.

Expanding the quantity in brackets in the right hand side of (1.9)

$$E_e(\mathbf{u}+\varepsilon\boldsymbol{\xi}) - E_e(\mathbf{u}) = \varepsilon \int_\Omega W'(\boldsymbol{\epsilon}(\mathbf{u})) \cdot \boldsymbol{\epsilon}(\boldsymbol{\xi})\, d\mathbf{x} + o(\varepsilon) \geq 0.$$

Since the above inequality must be satisfied for every perturbation $\boldsymbol{\xi} \in \mathfrak{U}_\mathbf{0}$, if it holds true for $\bar{\boldsymbol{\xi}} \in \mathfrak{U}_\mathbf{0}$ also $-\bar{\boldsymbol{\xi}}$ must be admissible, and therefore we deduce that the above integral must be zero at an equilibrium point. Note that by the definition of the energy density (1.8), we have

$$\int_\Omega (C : \boldsymbol{\epsilon}(\mathbf{u})) \cdot \boldsymbol{\epsilon}(\boldsymbol{\xi})\, d\mathbf{x} = 0,$$

which corresponds to the weak formulation of the first equation of (1.6).
Additionally, we say that $\bar{\mathbf{u}}$ is a *global minimizer* of E_e, if

$$E_e(\bar{\mathbf{u}}+\boldsymbol{\xi}) - E_e(\bar{\mathbf{u}}) \geq 0,$$

for all the perturbations $\boldsymbol{\xi} \in \mathfrak{U}_\mathbf{0}$. Combining this definition with (1.9) and the strict convexity of (1.7), we deduce that a global minimizer of the energy functional is the only equilibrium configuration of the system.
Thus, we can conclude that the following minimization problem

$$\arg\min_{\mathbf{u}\in\mathfrak{U}_\mathbf{g}} E_e(\mathbf{u}) \tag{1.10}$$

is equivalent to solve the weak formulation of the differential system (1.6).

Remark 1.5. *After the introduction of the elasticity equilibrium system* (1.6)*, we defined* $\boldsymbol{g}$ *as a function with values on the entire domain. Despite we merely need the external displacement to exist only on* $\partial\Omega_D$ *to write* (1.6) *and* (1.10)*, the characterization of the energy inequality* (1.11)*, which is fundamental to check and characterize the evolution of the system, needs the external displacement to be formally defined on the whole domain* Ω*.*

1.1.2. Quasi-static evolution

So far the assigned displacement $\mathbf{g}$ has been considered as a fixed quantity. We now introduce the concept of time evolution in the deformation process, starting with the definition of *load process*, which is a continuous function $\mathbf{g} : [0,T]\times\Omega \to \mathfrak{D}$, where the interval $[0,T] \subset \mathbb{R}$ is not necessarily identified with a physical time interval. Alternatively, we can say that the process which is deforming the specimen is rate-independent and phenomena like inertia and viscous dissipation are negligible in the following analysis.
For a given load process $t \mapsto \mathbf{g}(t) =: \mathbf{g}_t$, we say that the continuous map $t \mapsto \mathbf{u}(t) =: \mathbf{u}_t \in \mathfrak{U}_{\mathbf{g}_t}$ is a *deformation process* associated to $\mathbf{g}$, if it satisfies the following growth condition for all $t \in [0,T]$

$$E_e(\mathbf{u}(t)) \leq E_e(\mathbf{u}(0)) + \int_0^t \int_\Omega \boldsymbol{\epsilon}(\mathbf{u}(\tau)) \cdot \boldsymbol{\epsilon}(\dot{\mathbf{g}}(\tau))\, d\mathbf{x}\, d\tau, \tag{1.11}$$

where the superimposed dot denotes the differentiation with respect to t. The integral in (1.11) represents the work supplied by the external displacement to the system. Indeed, by the energy balance and Clausius–Duhem inequality [126, Eq. 81.7], it is not possible that the potential energy of the body at a time t is larger than the initial potential energy plus the energy contribution given by the force modeling the specimen. Alternatively, we say that inequality (1.11) regulates the energy growth of the system avoiding that, for any $\varepsilon > 0$ and any time interval $(t, t+\varepsilon)$, the variation of the energy exceeds the power supplied from the outside during the same time interval.
A given configuration $\tilde{\mathbf{u}}$ is said to be *accessible* if there exists a deformation process driven by a load process $t \mapsto \mathbf{g}_t$ such that $\tilde{\mathbf{u}} \in \mathfrak{U}_{\mathbf{g}_t}$ for some t, and inequality (1.11) is called *accessibility condition.*

Definition 1.6 (Quasi-static evolution)**.** *Let $\boldsymbol{g} : [0,T] \times \Omega \to \mathfrak{D}$ be a load process. The deformation process $t \mapsto \boldsymbol{u}_t$ describes a* quasi-static evolution *if every configuration $\boldsymbol{u}_t$ is also an equilibrium configuration.*

Remark 1.7. *Note that, since the energy functional* (1.7) *is strictly convex by definition, it is sufficient to solve, for each time $t \in [0,T]$, the convex minimization problem* (1.10)*.*

The definition of a quasi-static evolution effectively reflects the idea that a body subject to a varying external load, without any viscous and inertial effect, changes its shape following the path of least energy.

1.2. Fracture mechanics

Linear elasticity, thanks to its simplicity and regularity, is the most advocated model to describe the deformations that an elastic body undergoes when it is subject to a load process. A significant limitation of this model is that, to get a reliable description of the evolution, the external forces must be carefully calibrated so that the specimen will not break in the tests, since the model does not allow for failures. Indeed, if we pull a beam at both the extremities with constantly increasing force, we will observe the system energy growing as predicted by (1.7), but we will not be able to forecast at which instant the specimen will break. Unfortunately, there is no way to describe the rupture of the domain using only the theory introduced so far and therefore we look for a new model accounting for the formation of a fracture.
In real life problem, as we have already described at the beginning of the chapter,different kinds of material deformation can be observed. In particular, a crack may form in a *brittle* fashion, i.e. the material elongates elastically until, without any premonitory sign, breaks forming a sharp fracture surface, or, alternatively, *ductile* fractures can be observed after a progressive weakening of the specimen, resulting in a damaged area around the crack surface. In nature, many material have an intermediate behavior and they are

called *ductile-brittle*. In this case their rupture happens suddenly but only after a weakening ductile phase.
In particular, in this work we focus here on the first fracturing mode which is, due to its catastrophic character, the most fascinating and studied in fracture mechanics.
Looking at the deep origins of fracture mechanics, the first treaty on ruptures and material failure dates back to the seventeenth century, when Galileo Galilei wrote the book *Discorsi e dimostrazioni matematiche intorno a due nuove scienze* [114]. However, the appearance of a mathematical model describing the phenomenon of body fracture comes only in the beginning of the last century thanks to A. A. Griffith. The aeronautical engineer was commissioned during the World War I by the Royal Aircraft Establishment to study whether the surface treatment operations done on the body of the aircraft could influence the structural failure observed in the fleet. The result of this investigation was published in [124], where the very first theorem describing how a preexisting fracture in a body could spread and evolve appeared. Despite its innovation, the theory was still not complete and it had some limiting aspects that made this work not so popular for years, and in the meanwhile parallel new theories arose.
The main characteristic that A. A. Griffith and the later literature did not take into account was a proper description of the crack initiation, i.e. how does it happen that a fully intact material at some point starts breaking. This problem remained unsolved for almost eighty years till the 1998 when, in [110], G. Francfort and J.–J. Marigo, inspired by the image segmentation theory (see e.g. [31, 137]), proposed a model based on the Mumford–Shah functional [169] to overcome this issue.
An alternative modeling approach, was proposed by D. S. Dugdale [96] and, independently, a few years later by G. I. Barenblatt [26]. These and the subsequent models, called cohesive fracture models, introduce an intermediate phase in the evolution between the elastic elongation and the catastrophic failure, called indeed cohesive phase. The idea comes from the molecular attraction that tends to keep close to each other the lips of the growing crack, but does not introduce any permanent deformation in the material that returns to the rest configuration as soon as the cohesive force disappears.
In the next sections, we explore these models. First we start with the father of modern fracture mechanics theory, then we continue with an introduction to the Francfort–Marigo model, and finally we deal with cohesive fracture models.

1.2.1. Griffith's Theory

At the beginning of the twentieth century A. A. Griffith found the inspiration for devising the first fracture model from the *Theorem of Minimum Energy*:

> *"The equilibrium state of an elastic body, deformed by specified external forces, is such that the potential energy of the whole system is a minimum."* [191]

This theorem, which has actually also inspired Definition 1.6 of quasi-static evolution, does not deal with any criterion for the formation of a fracture. A. A. Griffith added a statement which introduces the possibility of a crack formation with the idea that, if an equilibrium is possible, it must be the one in which the medium is broken, if the system can pass from the unbroken status to the broken one via a continuous decrease of the potential energy. In other words, we say that a fracture appears if breaking is more convenient than increasing the elastic energy. This idea describes actually an evolution along global minimizers of the energy functional, which is the most studied approach, but, as we will comment later, recently the community seems to be in favor of evolutions along local and more physically sound minimizers.
How to design a principle which describes the opening of a fracture? How to describe the energy of a fracture? And, in particular, does a match between the criterion introduced and the physical phenomenon exist? The solution came thinking at a molecular level. Indeed, we know that each particle is linked to the neighboring ones by atomic bonds. In order to break these bonds, we need to spend some energy, which can be expressed as a fictitious surface energy in the continuous setting. We used the word *fictitious* because the fracture energy is not a physical quantity but just an artifact to describe the process. Once a body gets damaged, the fracture surface is traction free except for a very small area around the crack tip, where the atoms can still interact each others. Thus, the assumption that the failure is a traction free surface of the domain can be reasonable.
All these reasoning can be summarized by the *Griffith's Theorem* [124]:

Theorem 1.8 (Griffith's Theorem)**.** *In an elastic solid body deformed by specified forces applied at its surface, the sum of the potential energy on the applied forces and the strain energy of the body is diminished or unaltered by the introduction of a crack whose surfaces are traction free.*

Hence, according to this Theorem, we can characterize the *energy functional* of an elastic body which allows for fractures as the sum of the elastic energy E_e introduced in (1.7) and a surface fracture energy E_f defined on the $(d-1)$-dimensional hypersurface $\Gamma \subset \Omega$. This hypersurface, which is actually a surface if $d = 3$, a line when $d = 2$ and a point if $d = 1$, is traction free and reduces the internal stresses $\boldsymbol{\sigma}$ of the body:

$$E(\mathbf{u}, \Gamma) = E_e(\mathbf{u}) + E_f(\mathbf{u}, \Gamma). \tag{1.12}$$

As we already commented, the fracture energy measures the energy spent to break the molecular bonds, but since single particles are not taken into account, it is necessary to establish an alternative criterion to measure this quantity. The original idea was to relate the energy contribution given by the spreading of the crack to the surface of the fracture.
In the following section we redefine the concept of equilibrium state for the energy

functional (1.12), then we add some condition on the crack set Γ in order to extend the concept of accessible state to the empirical experience on fractures.

1.2.2. Equilibrium configuration and accessibility condition

Recalling Definition 1.9 we have that the first variation of the energy functional

$$\delta E\big((\mathbf{u},\Gamma),(\boldsymbol{\xi},\Xi)\big) := \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\big[E_e(\mathbf{u}+\varepsilon\boldsymbol{\xi}) - E_e(\mathbf{u}) + E_f(\mathbf{u}+\varepsilon\boldsymbol{\xi},\Gamma\cup\Xi) - E_f(\mathbf{u},\Gamma)\big] \quad (1.13)$$

must be non-negative for any admissible perturbation pair $(\boldsymbol{\xi},\Xi)$. In particular, Ξ may also not be related to the crack set Γ and add a fracture to the domain. Since now the domain can be fractured, we need to add an non-interpenetration condition

$$[\![\mathbf{u}]\!](\tilde{\mathbf{x}}) + [\![\boldsymbol{\xi}]\!](\tilde{\mathbf{x}}) \geq 0 \quad (1.14)$$

for all the newly created fracture points $\tilde{\mathbf{x}}$, where we denote with $[\![\cdot]\!]$ the jump amplitude between the lips of the fracture. We also recall that a perturbation, to be admissible, must not deform the domain

$$\int_{\Omega\setminus\Xi} \boldsymbol{\epsilon}(\boldsymbol{\xi})\,d\mathbf{x} + \int_{\Xi} [\![\boldsymbol{\xi}]\!]\,d\mathcal{H}^{d-1}(\Xi) = 0,$$

where $\mathcal{H}^{d-1}(\Xi)$ is the $(d-1)$-dimensional Hausdorff measure (see Appendix A). In Section 1.1.1 we observed that the global solution of the minimization problem (1.10) coincides with the unique equilibrium configuration of the purely elastic system. Unfortunately, in this new setting, condition (1.13) does not trivially coincide any more with the global minimum. Indeed, despite the elastic energy is convex, the fracture energy is in general nonconvex and nonsmooth and therefore it may exist more than one pair $(\mathbf{u},\Gamma)$ satisfying the equilibrium condition and not necessarily coinciding with the global minimizer of the energy functional (we refer to [90] for a one dimensional example clarifying this point).
For this reason, the definition of quasi-static evolution given in Section 1.1.2, which characterizes this phenomenon as an equilibrium process evolving exclusively along energy functional's global minimizers may not apply any longer. Indeed, due to the non uniqueness of the solution of the energy minimization problem, we need to redefine the concept of quasi-static evolution as a *stable equilibrium process* which is an equilibrium process whose configurations are stable.

Definition 1.9 (Stable configuration)**.** *Let $\boldsymbol{g}:[0,T]\times\Omega\to\mathfrak{D}$, $\boldsymbol{u}$ one of its admissible configuration and $\Gamma\subset\Omega$ the subset of the domain Ω where the displacement function is discontinuous. We say that the pair $(\boldsymbol{u},\Gamma)$ is a* stable configuration *if the following identity is satisfied*

$$\delta E\big((\boldsymbol{u},\Gamma),(\boldsymbol{\xi},\Xi)\big) = 0 \qquad \textit{for all admissible perturbations } (\boldsymbol{\xi},\Xi) \quad (1.15)$$

Now, considering that the energy functional is nonconvex both for brittle and cohesive fracture models, the most struggling question related to the problem (1.15) is *"Which one of the several critical points should be chosen?"*
Unfortunately, we are far from having a definitive and unique answer. As already mentioned, most of the studies followed Griffith's idea of minimum energy, selecting the global minimizer, see for example [76, 78, 109, 119, 158], [73] for the nonlinear elasticity case, and [74] for some qualitative observations.
But following the global minimizers does not agree with the physical evidences beyond elasticity theory. Indeed sometimes these evolutions require the crossing of energetic barriers which in reality are unfeasible. Therefore, more recently, a second point of view has been developed and it considers evolutions along critical points where a body should not move from its stable configuration if this implies the crossing of a relevant energetic barrier, see [77, 142, 143, 144, 154, 155, 209]. A comparison of these two philosophies has been done in [173]. Unfortunately, the theory of existence of such evolutions as well as their mathematical study is not so well established like for global minimizer evolution.
Additionally, we need to sightly modify the definition of deformation process adding a crack irreversibility criterion in order to avoid healing of the fracture. To proceed, we introduce the notion of compatible configuration.

Definition 1.10 (Compatible configuration)**.** *We say that a configuration* $(\boldsymbol{v}, \Phi)$ *is* compatible *with* $(\boldsymbol{u}, \Gamma)$ *if there exists a perturbation* $(\boldsymbol{\xi}, \Xi)$ *such that*

$$(\boldsymbol{v}, \Phi) = (\boldsymbol{u} + \boldsymbol{\xi}, \Gamma \cup \Xi).$$

This definition tells us that in order to be compatible, the fracture set Γ must be a subset of Φ.
Now, given a load process $t \mapsto \mathbf{g}_t$, the map $t \mapsto (\mathbf{u}_t, \Gamma_t)$ is a deformation process if the two following conditions are satisfied:

i) given an initial configuration $(\mathbf{u}_0, \Gamma_0) = (\mathbf{u}(0), \Gamma(0))$, for all $t \in [0, T]$ it holds

$$E(\mathbf{u}(t), \Gamma(t)) \leq E(\mathbf{u}(0), \Gamma(0)) + \int_0^t \int_{\Omega \setminus \Gamma(\tau)} \boldsymbol{\epsilon}(\mathbf{u}(\tau)) \cdot \boldsymbol{\epsilon}(\dot{\mathbf{g}}(\tau))\, d\mathbf{x}\, d\tau; \tag{1.16}$$

ii) for all $\tau > t$, $(\mathbf{u}(\tau), \Gamma(\tau))$ is compatible with $(\mathbf{u}(t), \Gamma(t))$.

Notice that point **ii)** is equivalent to require that $\Gamma(t) \subseteq \Gamma(\tau)$ for every $\tau > t$, which means that once the crack is created, it cannot heal in future times. This concept is defined as *irreversibility of fracture*, but it is also addressed in the literature with the name *unilateral condition* by some authors [43, 61, 149, 158], and it is fundamental for all the studies which does not consider a strictly increasing external displacement. Indeed, in [51, 52], where the applied load constantly increasing in time, this condition can be omitted.

1.2.3. Francfort–Marigo brittle fracture model

Griffith's model and the studies that followed, despite giving a sensible contribution for a better understanding of fracture evolution, still lacked the modeling the crack initiation. Some previously appeared studies [152, 193] made an attempt to overcome the problem, but the final solution was found only in 1998 when G. Francfort and J.–J. Marigo, presented a new model asserting that the work [124] was

> *"an unreliable instrument to predict crack initiation, crack path, and eventual crack jumps along the crack path."*[110]

In this section, we recall the arguments that the two authors proposed in [110] to introduce their model for brittle fracture, highlighting the limitations of the previous one.
Let us consider a domain $\Omega \subset \mathbb{R}^2$ with a straight fracture Γ of length $\mathcal{H}^1(\Gamma) = \ell$. Pulling apart the plate with an external displacement $\mathbf{g}_t$, Griffith's model prescribes a crack growth only if the internal stresses are of the order $1/\sqrt{\ell}$. Since we are interested in the crack initiation, we let $\ell \to 0$ and we observe that the stress needed to start the fracture grows to infinity, despite the evidence that this occurs at finite stress.
To resolve this inconsistency, the authors were inspired by the apparently unrelated field of image segmentation, but keeping Griffith's idea of energy balance between the elastic energy and a fictitious fracture energy. Thus, let us start introducing the two components of the energy functional: elastic and fracture energies. The elastic part is defined in (1.7), evaluating the energy density function in all the points of the domain except for the ones of the fracture set Γ. In general, any function W with the properties described in Section 1.2.2 can be chosen, but we stick to linear elasticity and isotropic materials for sake of simplicity. In particular, if the applied displacement on the boundary $\partial\Omega_D$ is on plane, the elastic energy density has the form (1.8), that for isotropic materials is

$$W(\boldsymbol{\epsilon}(\mathbf{u})) := \frac{1}{2}(C : \boldsymbol{\epsilon}(\mathbf{u})) \cdot \boldsymbol{\epsilon}(\mathbf{u}) = \frac{1}{2}(\lambda \mathbf{I} \operatorname{tr}(\boldsymbol{\epsilon}(\mathbf{u})) + 2\mu\boldsymbol{\epsilon}(\mathbf{u})) \cdot \boldsymbol{\epsilon}(\mathbf{u}). \tag{1.17}$$

Alternatively, if we consider an off-plane displacement, i.e. in a two-dimensional plate the deformation is orthogonal to the domain's plane, the elastic energy density simplifies to

$$W(u) := \frac{1}{2}|\nabla u|^2, \tag{1.18}$$

where the displacement $u : \Omega \to \mathbb{R}$ is now a scalar function.
The second element to introduce is the fracture energy. Similarly to Griffith, the authors considered a surface-like quantity proportional to the $(d-1)$-Hausdorff measure of the fracture set Γ

$$E_f(\Gamma) := \kappa \mathcal{H}^{d-1}(\Gamma), \tag{1.19}$$

where $\kappa > 0$ a material parameter. The peculiar property of this model is that the crack can assume any shape and extension. In particular, we look for fracture sets which are

contained in $\mathcal{K} \subset \Omega$, where

$$\mathcal{K} := \{\Gamma \subset \Omega : \Gamma \text{ has at most } m \geq 1 \text{ connected components}, \mathcal{H}^{d-1}(\Gamma) < \infty\}.$$

Indeed, despite the fracture set can assume any shape and thickness, the creation of a d-dimensional fracture set requires an infinite amount of energy if measured with the $(d-1)$-dimensional Hausdorff measure. Hence the properties of the fracture energy make the formation of a thick fracture extremely unlikely.
Adding the elastic and fracture energy together, we finally obtain the total energy

$$E(u, \Gamma) := \frac{1}{2} \int_{\Omega \setminus \Gamma} |\nabla u|^2 + \kappa \mathcal{H}^{d-1}(\Gamma) \tag{1.20}$$

that is, under proper boundary conditions, the quantity we need to minimize at each time $t \in [0, T]$ to get a quasi-static evolution.
This energy functional, despite being never used in continuum mechanics before [110], was not completely new. Indeed, it already appeared in the field of image segmentation [169]. Moreover, this minimization problem was already shown in [86] to be equivalent to a well-posed one-field minimization problem on $SBV(\Omega)$, the space of special bounded variation functions [7], and, thanks to other works in relatively broad fields [5, 73, 78, 109], this brittle fracture model was shown to have the desired properties in a pretty general setting, filling the gap between the seminal work and the theory of quasi-static fracture evolution. The situation is more involved if we consider the plane elastic energy density (1.17). In this case, the arguments used to show the well-posedness of the problem for the Mumford–Shah functional (1.20) are not valid and the displacement function $\mathbf{u}$ must be defined in a SBD space (see Appendix A for the definition of these function spaces).
Following the model (1.20), focusing for simplicity only on the off-plane case, and recalling that the distributional derivative can be decomposed as in (A.3), we denote with ∇u its absolutely continuous part and with $J(u)$ the set where the displacement is discontinuous and presents jumps. Notice that, we did not enforce Γ and the jump set of the displacement $J(u)$ to coincide. Thus, assuming $u \in SBV(\Omega)$ we formally redefine the energy functional as follows

$$E(u, \Gamma) = \begin{cases} \frac{1}{2} \int_{\Omega} |\nabla u|^2 + \kappa \mathcal{H}^{d-1}(\Gamma) & \text{if } \mathcal{H}^{d-1}(J(u) \setminus \Gamma) = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

We say that a given load process $g \in L^\infty(0, T; W^{1,\infty}(\Omega)) \cap W^{1,1}(0, T; H^1(\Omega))$ acting on $\Omega_D \subset \Omega$ is admissible if, for a given fracture set Γ, it verifies the conditions as defined in Section 1.2.2 and additionally:

- $g(\mathbf{x}) \in \mathfrak{D}$ for all $\mathbf{x} \in \Omega_D$;

- on $\Omega \cap \Gamma$ no boundary condition and no external displacement is imposed;
- on $\Omega \setminus \Omega_D$ no boundary condition is imposed on the displacement u.

Consequently, the set of all admissible configuration is

$$\mathfrak{U} := \{u \in SBV(\Omega) : u|_{\Omega_D} : \Omega_D \to \mathfrak{D}\},$$

while at a time $t \in [0,T]$ and for a given admissible applied displacement $g(t)$ the admissible configurations are characterized by

$$\mathfrak{U}_{g(t)} := \{u \in SBV(\Omega) : u|_{\Omega_D} = g(t)|_{\Omega_D}\}.$$

Thus, for every time $t \in [0,T]$ we need to find a globally minimizing pair for the energy functional

$$(u^*(t), \Gamma^*(t)) \in \underset{\substack{u \in \mathfrak{U}_{g(t)} \\ \Gamma \supseteq \bigcup_{s<t} \Gamma(s)}}{\arg\min} E(u, \Gamma). \tag{1.21}$$

The proof that the above problem has a solution which is a quasi-static evolution has been given in [78], where the mathematical properties of the model have been explored. Notice that for the moment we limit ourselves to evolutions along global minimizers in (1.21). In the following, we report the main result of that paper, providing all the feature of a quasi-static evolution for (1.21).

Theorem 1.11 (see [78]). *Let $g \in L^\infty(0,T; W^{1,\infty}(\Omega)) \cap W^{1,1}(0,T; H^1(\Omega))$ and let $\Gamma(0) = \Gamma_0 \in \mathcal{K}$. Then there exists a function $\Gamma : [0,T] \to K$ such that*

a) $\Gamma_0 \subset \Gamma(s) \subset \Gamma(t)$ for $0 \le s \le t \le T$;

b) $E(u(0), \Gamma(0)) \le E(u(0), \Gamma)$, for all $\Gamma \in \mathcal{K}$ with $\Gamma \supset \Gamma_0$;

c) for $0 \le t \le T$, $E(u(t), \Gamma(t)) \le E(u(t), \Gamma)$ for all $\Gamma \in \mathcal{K}$ with $\Gamma \supset \bigcup_{s<t} \Gamma(s)$;

d) $t \mapsto E(u(t), \Gamma(t))$ is absolutely continuous in $[0,T]$;

e) $\dfrac{d}{ds} E(u(t), \Gamma(s))\Big|_{s=t} = 0$, for almost every $t \in [0,T]$.

Moreover, for every function $\Gamma : [0,T] \to \mathcal{K}$, such that $(a) - (e)$ hold true, the following identity is satisfied

$$\frac{d}{dt} E(u(t), \Gamma(t)) = 2 \int_{\Omega \setminus \Gamma(t)} \nabla u(t) \nabla \dot{g}(t)\, d\boldsymbol{x}$$

for almost every $t \in [0,T]$, where $u(t)$ is a solution of the minimum problem (1.21) for Γ given, and $\dot{g}(t)$ is the time derivative of the function $g(t)$.

We mention that to achieve the result the authors introduce a time discretization step δ, find the pair $(u_\delta(t_k), \Gamma_\delta(t_k))$ for every t_i, with $0 \leq k \leq N_T$, such that $t_k - t_{k-1} = \delta$, and then they proved the convergence for $\delta \to 0$ to the continuous in time solution $(u(t), \Gamma(t))$. Moreover, we also remark that the assumption that the fracture set has at most m connected component was crucial for the proof of convergence by means of homogenization technique [22, 69, 81, 141, 170], despite G. Dal Maso and R. Toader observed that this assumption is sufficient only for $d \leq 2$.
A similar existence result, but for evolutions along critical point has appeared in [77]. In the paper, the authors consider at each time step, in the discrete-time formulation, local minimizers of the energy functional which are sufficiently close to the approximate solution obtained in the previous step. This has been done by introducing in the variational problem an additional term which penalizes the L^2-distance between the approximate solutions at two consecutive times. We mention here that in Chapter 4 we follow this balancing principle between proximity and energy minimization to define quasi-static evolutions in greater generality.

1.2.4. Cohesive fracture models

Griffith's model together with the new ideas of G. Francfort and J.–J. Marigo is not the unique approach to deal with brittle fractures. Indeed, in the second half of the last century, on both sides of the Atlantic ocean, during the cold war, two scientist, D. S. Dugdale in America [96] and G. I. Barenblatt in Russia [26], independently developed alternative models. Indeed, two molecules separated by a fracture but still sufficiently close each others present an attractive force, which is called *cohesive force*. The lack of this aspect in Griffith's theory motivated the authors to conceive new models. Hence, the elastic energy is not released instantaneously but more gradually as long as the distance between the lips of the fracture grows and the atomic attraction get weaker.
Even if the models are different, quasi-static evolutions of cohesive fractures require the minimization of the energy functional, which has once again two components: the elastic energy E_e and the cohesive fracture energy E_f. Thus, the equilibrium configuration that the body assumes at each time is given by a proper balance of these two terms. Despite the structural similarity, as already mentioned, models of cohesive fractures differ from the one described in [110, 124] by the crack formation process. The fracture is now described by a function depending on the amplitude of the jump of the displacement, denoted by $[\![u]\!]$.
Considering the non-interpenetration condition (1.14) the cohesive fracture function $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ is restricted only to the positive half-plane and is required to be monotonic and non decreasing. Moreover the following limits

$$\theta_a = \lim_{[\![u]\!] \to 0^+} \theta([\![u]\!]) \qquad \theta_r = \lim_{[\![u]\!] \to +\infty} \theta([\![u]\!])$$

define the activation energy θ_a and the rupture energy θ_r, where $\theta_a \leq \theta_r$. The activation energy is the energy needed to weaken the molecular bonds and initiate the fracture, while the rupture energy is spent to destroy all the connection between the two lips of the fracture, and the body can be considered broken. We additionally define $[\![u]\!]_r$ the minimal jump high needed to reach the rupture energy.

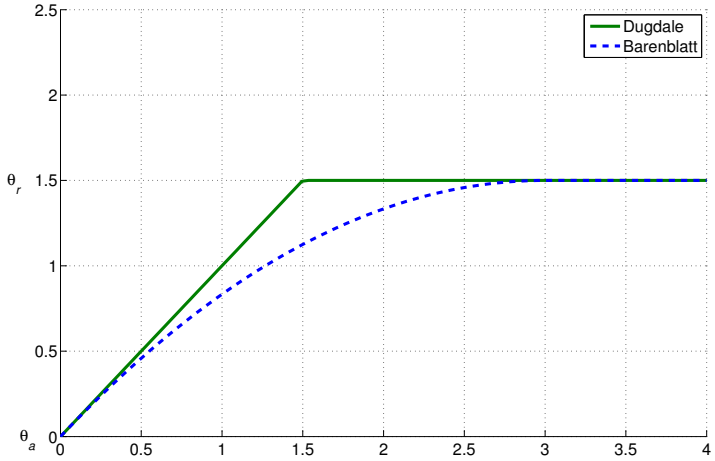


Figure 1.1.: The cohesive fracture energy functions of the two pioneering models: in green Dugdale function and in blue Barenblatt's

In the brittle fracture model $\theta_a = \theta_r$ and therefore we could distinguish the points of the domain into two class, fracture and non-fracture points. Conversely we now need to introduce the intermediate state called *bridging phase*, which describes the transition of the body between the elastic deformation and the rupture. *Non-fracture points* are all the point $x \in \Omega$ such that $[\![u]\!] = 0$, then, *pre-fracture points* are all the points where the bridging force is still acting, and finally *fracture points* are only the points $x \in \Omega$ where $[\![u(x)]\!] > [\![u]\!]_r$ and a fracture in the sense of Griffith is created. Additionally, by setting $\theta_a = 0$, as D. S. Dugdale and G. I. Barenblatt did, every pre-fracture point with jump amplitude 0 is also a non-fracture point. Thus, we do not make a mistake calling pre-fractured status every configuration in which the domain is not completely fractured.
In Figure 1.1, we report the cohesive fracture energy functions for the two pioneering models. For both the models the activation energy is set to zero. The main difference between the two models lies in the regularity of the function θ. Dugdale's function (in green) is piecewise linear, the cohesive force keeps constant during the bridging phase, while drops instantly to zero as soon as the critical value $[\![u]\!]_r$ is reached. Barenblatt's cohesive force derivative is instead continuous and monotonically decreasing; this implies

that the attraction between the two lips gets weaker the bigger their separation. It is also possible to design Barenblatt-like cohesive forces such that the critical value for the final rupture is set to infinity and therefore the bridging between the molecules of the body never disappears.

Remark 1.12. *Griffith's model can be seen as a discontinuous limit of a cohesive fracture model. Indeed, if we let* $[\![u]\!]_r \to 0$ *we have that, in the limit, the activation energy coincides with the rupture energy,* $\theta_a = \theta_r$*. Consequently, the cohesive transition phase vanishes and the brittle behavior of Griffith's model described in Section 1.2 is recovered. Thus, cohesive fracture models are often interpreted has a regularization approach to Griffith's brittle fracture model.*
As we will describe in Chapter 5, a direct minimization of the energy functional (1.20) *on a continuous domain* Ω *is extremely complicated and therefore most of the literature considering the Francfort–Marigo model introduces a smooth phase field approximation via the Ambrosio–Tortorelli functional [9] which* Γ*-approximates the Mumford–Shah functional [39, 40, 42, 48, 176] (see Appendix A for a definition of* Γ*-convergence). Despite this approach has been widely studied, the purely mathematical introduction of a phase field leaves open questions for a practical point of view: what is the physical meaning of the phase field function in the intermediate state? Where is the exact location of the fracture since width of the damaged area depends on the steepness of the phase field? Conversely, the approximation of brittle fractures via cohesive energy functionals offers not only a vast scenario of possible cohesive fracture functions which correspond to different bridging behavior, but also give a physical interpretation to the regularization introduced.*

Finally, to describe a quasi-static evolution, we need to solve for each time step the following minimization problem. Given a domain $\Omega \subset \mathbb{R}^d$ and a prescribed fracture set $\Gamma \subset \mathbb{R}^{d-1}$, a cohesive function $\theta : \mathbb{R}^+ \to \mathbb{R}^+$, a time interval $[0, T]$, and a time dependent load process $g \in L^\infty(0, T; W^{1,\infty}(\Omega)) \cap W^{1,1}(0, T; H^1(\Omega))$ acting on $\Omega_D \subset \Omega$ in the set of admissible displacements $\mathfrak{D}$, we need to solve, for each time $t \in [0, T]$ the following minimization problem:

$$u(t) \in \operatorname*{arg\,min}_{u=g(t) \text{ on } \Omega_D} \int_{\Omega\setminus\Gamma} (\nabla u)^2 \, d\mathbf{x} + \kappa \int_\Gamma \theta([\![u]\!]) \, d\mathcal{H}^{d-1}(\Gamma), \tag{1.22}$$

where κ is a material parameter as in (1.19). Notice that, for technical difficulties, we decided to fix the fracture set Γ. We may justify this assumption supposing that the specimen we are deforming was broken in Γ and then repaired. Hence, all the points in $\Omega \setminus \Gamma$ may be only non-fracture points, while the ones in Γ are either to pre-fracture or fracture points.
Although this minimization problem may appear easier than (1.21), it is far from being trivial. Indeed, the functional is still nonconvex and additionally issues regarding

the choice of the minimization algorithm may raise depending on the regularity of the function θ.

2. Linearly constrained minimization of nonconvex and nonsmooth functionals

The quasi-static evolution of fractures requires, as already described in the previous chapter, the minimization of an energy functional E (examples have been presented in the Introduction) at any time step t_k, with $k = 1, \ldots, N_T$, of a proper time discretization scheme of the time window $[0, T]$. Let us stress that in the first part of this chapter we report the theoretical results published in [19], while the second part contains unpublished results.

In the literature one can find efficient algorithmic solutions for linearly constrained *convex* and *nonsmooth* minimization, e.g., augmented Lagrangian methods [30, 112, 137, 185, 190], and for linearly constrained *nonconvex* minimization, such as sequentially quadratic programming (SQP) or (semi-smooth) Newton methods [182]. Unfortunately, in the latter cases only *smooth objective energies*, usually at least $\mathcal{C}^2$ functionals, can be addressed by algorithms, which are then guaranteed to converge *only locally* around the expected critical point. A more general setting is the one considered in [21], where a remarkable analysis of the convergence properties of descent methods for nonconvex optimization, also with constraints, has been carried out. A key role in the construction of the algorithm is played by a special condition, the so-called Kurdyka-Łojasiewicz inequality (see for instance [34]) allowing for a general convergence result in a nonsmooth nonconvex setting, but again under the assumption of a *good initial guess*. A certain smoothness, namely, *global* $\mathcal{C}^{1,1}$ regularity, is needed to remove the latter very restrictive assumption. Let us also stress that, although quite a mild condition from the point of view of the applications, the Kurdyka-Łojasiewicz inequality could *not* be verified even in the case of convex functions, as shown again in [34].

The above mentioned limitations of the currently available literature lead us to the motivation of this chapter, where we introduce two novel algorithms which aim to minimize nonconvex and nonsmooth linearly constrained functionals with minimal regularity assumptions. In order to guarantee the convergence of the algorithms to critical points of the objective functional we need to restrict ourselves to a discrete setting, introducing the Euclidean spaces $\mathcal{E} \simeq \mathbb{R}^n$ and $\mathcal{F} \simeq \mathbb{R}^m$ for $m \leq n$. This restriction may seem significant at a first look, but we may also consider among the spectra of possible application also discretization of infinite dimensional problems. Indeed, in such situations, the map u can be interpreted as an evolution of a physical system which can assume a contin-

uum of states and we first reduce the problem to an approximation where the evolution assumes at most a finite number n of states instead. In this setting falls exactly the discretization of continuum mechanics models as, for example, the fracture models introduced in Chapter 1. Besides continuum models, there are also physical systems, e.g., quantum systems [101, 113, 133, 146, 159, 204], that can indeed assume only a finite number of states and no approximation to a continuum of states is at all considered in first place. Moreover, if we allow ourselves to consider more general or abstract systems, perhaps related to social dynamics, u can easily model situations where a finite number of interacting agents are changing their states according to a socio-economical principle represented by a cost function $\mathcal{J}$ and are simultaneously bounded by certain linear rules encoded in the constraint pair (A, f) [1, 37, 56, 70, 188].

In the next sections we introduce the two algorithms (the first is also proposed in [19]) and their convergence properties. In Chapter 3 we show different application fields where such problem often occurs. Indeed, among the models that fit the general optimization problem, we present the Mumford–Shah functional in image processing, a new decoding strategy for reducing the noise-folding phenomenon in compressive sensing, and, naturally, the energy functionals driving the quasi-static evolution of fractures. In particular, we dedicate the entire Chapter 4 to the analysis of cohesive models [51, 52, 80]. This list is far from being complete, as we actually expect that the algorithms studied can have significant further numerical applications also in other problems involving nonsmooth and nonconvex energies with additional linear (boundary) conditions, like elasto-plastic evolutions [79, 165], atomic structure computations [23], and inverse gravimetric problems [29, 201].

An interesting feature of the proposed algorithms to linearly constrained nonsmooth and nonconvex minimization problems involving truncated polynomial energy terms is that the *inner loop* can be realized by means of an *iterative thresholding algorithm*. This technique has been firstly proposed in [108] to solve inverse free-discontinuity problems in one dimension, where no approximate smoothing of the energy was used, contrary to other previous approaches, e.g., based on graduated nonconvexity [31, 179, 177]. The extension we provide allows us now to similarly address problems, which are defined in any dimension, thanks to the appropriate handling of corresponding linear constraints and an eventual *very mild* smoothing.

Thresholding algorithms have by now a long history of successes, based on their extremely simple implementation, statistical properties, and, in the iterative case, strong convergence guarantees. We retrace briefly some of the relevant developments, without the intention of providing an exhaustive mention of the many contributions in this area. The terminology "thresholding" comes from image and signal processing literature, especially related to damping of wavelet coefficients in denoising problems, however the associated mathematical concept is the *Moreau proximity map* [68], well-known from convex optimization. The statistical theory of thresholding has been pioneered by D.

Dohono and I. Johnstone [95] in signal and image denoising and further and extensively explored in other work, e.g., [60]. *Iterative soft-thresholding algorithms* to numerically solve the minimization of convex energies, modeling inverse problems and formed by quadratic fidelity terms and ℓ_p-norm penalties, for $p \geq 1$, have been first proposed in [102]. Their strong convergence has been proven in the seminal work of I. Daubechies, M. Defrise, and C. De Mol [82]. The recent theory of *compressed sensing*, i.e., the universal and nonadaptive compressed acquisition of data [54, 94], stimulated also the research of iterative thresholding algorithms for nonconvex penalty terms, such as the ℓ_p-quasi-norms for $0 < p < 1$ (see e.g. [18]). Variational and convergence properties of *iterative firm-thresholding* algorithms, in particular the *iterative hard-thresholding*, have been recently studied in [33, 106]. Partially inspired by these latter achievements and the work of M. Nikolova [178] on the relationships between certain thresholding operators and discrete Mumford–Shah functionals, the results in [108] should be also considered as a contribution to the theory of thresholding algorithms in the new context of linearly constrained nonsmooth and nonconvex optimization.

This chapter is organized as follows. In the next section we introduce the general setting in which the algorithms are designed. The generality of this chapter is mainly motivated by giving the reader the largest overview of the potential of the algorithms, without limiting them to the case of fracture mechanics. Each of the two main Sections, 2.1 and 2.2, treats a different algorithm. In particular, the second algorithm can be considered a variation of the first one, appeared in [19], which can be convenient in some specific situations. For each algorithm, we proceed first analyzing its correspondent version created and successfully employed for convex problems, then we modify it in order to be able to deal with nonconvex functionals. Finally we prove its convergence to critical points of the nonconvex objective functional.

2.0.5. General setting

As already remarked, the algorithms can be applied not only to fracture mechanics problem and therefore we introduce the general setting in which the algorithm can be understood.
Let $\mathcal{J}\colon \mathcal{E} \to \mathbb{R}$ a lower semicontinuous nonconvex functional which we assume to be bounded from below. Notice that we may require in some section the convexity of the objective functional that will be denoted by $\widetilde{\mathcal{J}}$ to avoid any confusion. Since we will be concerned with the search of critical points, without any loss of generality we should suppose from now on that $\mathcal{J}(u) \geq 0$, for all $u \in \mathcal{E}$. We further consider a linear operator $A\colon \mathcal{E} \to \mathcal{F}$. Both the spaces $\mathcal{E}$ and $\mathcal{F}$ are endowed with an Euclidean norm, which we will denote in both cases by $\|\cdot\|$, since it will be always clear from the context in which space we are taking the norm. Dealing with finite dimensional spaces, it remains understood

that the only notion of convergence that we will use is the strong convergence in norm, since weak and strong topologies are in this case equivalent.
About the operator A, we should assume that it has *nontrivial kernel*, and is *surjective*. We should denote by $A^*\colon \mathcal{F} \to \mathcal{E}$ the adjoint operator of A. By our assumptions, for every $w \in \mathcal{F}$ we have that there exists $\delta > 0$ such that

$$\|A^* w\| \geq \delta \|w\| \,. \tag{2.1}$$

We consider $f \in \mathcal{F}$ and we are concerned with the problem of finding constrained critical points of $\mathcal{J}$ on the affine space $\mathcal{A}(f) := \{v \in \mathcal{E} : Av = f\}$. Since we deal with nonsmooth functionals, as usual, the notion of critical point is defined via the use of subdifferentiation, see Section A.1.1.
We recall, for sake of completeness, the definition of critical point.

Definition 2.1 (Critical point)**.** *Let $\mathcal{E}$ be an Euclidean space, $\mathcal{J}\colon \mathcal{E} \to \mathbb{R}$ a lower semi-continuous functional, and $v \in \mathcal{E}$. We say that v is a critical point of $\mathcal{J}$ if*

$$0 \in \partial \mathcal{J}(v) \,.$$

In the convex case this condition is sufficient to assure global minimality of v, otherwise it is only a necessary condition for local minimality.
In the following definition of constrained critical point the usual shorthand $\mathcal{J}(w + \ \cdot)$ is used to denote the functional $\xi \mapsto \mathcal{J}(w + \xi)$.

Definition 2.2 (Critical point on an affine space)**.** *Given a linear operator $A\colon \mathcal{E} \to \mathcal{F}$ with nontrivial kernel, and $f \in \mathcal{F}$, we say that w is a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f) = \{v \in \mathcal{E} : Av = f\}$ if $Aw = f$ and 0 is a critical point for the restriction to $\ker A$ of the functional $\mathcal{J}(w + \cdot)$.*

For $\mathcal{J}$ being a $\mathcal{C}^1$-perturbation of a convex function (in particular, with nonempty subdifferential at every point), the nonsmooth version of Lagrange multiplier Theorem assures that w is a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f)$ if and only if $Aw = f$ and

$$\partial \mathcal{J}(w) \cap \operatorname{ran}(A^*) \neq \emptyset \,, \tag{2.2}$$

where $\operatorname{ran}(A^*)$ is the range of the operator A^*, which is known to be the orthogonal complement of $\ker A$ in $\mathcal{E}$.
From now, about the function $\mathcal{J}$, we will make the following more specific assumptions:

(A0) the functional $v \mapsto \mathcal{J}(v) + |Av|^2$ is coercive;

(A1) $\mathcal{J}$ is ω-semi-convex, that is there exists $\omega > 0$ such that $\mathcal{J}(\cdot) + \omega \|\cdot\|^2$ is convex;

(A2) the subdifferential of $\mathcal{J}$ satisfies the following growth condition: there exist two nonnegative numbers K, L such that, for every $v \in \mathcal{E}$ and $\xi \in \partial\mathcal{J}(v)$

$$\|\xi\| \leq K\mathcal{J}(v) + L\,. \tag{2.3}$$

Remark 2.3. *We observe that*

(a) *condition (A1) is in fact met, for instance, by any $\mathcal{C}^1$ function in finite dimension with piecewise continuous and bounded second derivatives. However, let us stress that, conversely, ω-semi-convexity does not give any information on the smoothness of the function, other than local Lispchitzianity, hence, in finite dimension, its Fréchet-differentiability almost everywhere, by Rademacher's Theorem. We also recall that an ω-semi-convex function is a $\mathcal{C}^1$-perturbation of a convex function, therefore it has nonempty (and locally bounded) subdifferential at every point. If the subdifferential is uniformly bounded, then* (2.3) *is trivially satisfied;*

(b) *an ω-semi-convex function in finite dimension has a Fréchet differential almost everywhere, and, if* (2.3) *is satisfied only at points of differentiability, then it holds everywhere. This is true since it can be shown that the Fréchet subdifferential is contained in the so-called Clarke subdifferential, which is known to be at every $v \in \mathcal{E}$ the convex hull of limit points of differentials of $\mathcal{J}$ along sequences $v_n \to v$ (for these notions, see for instance [65, Chapter 2]). Therefore one needs not to calculate the subdifferential of $\mathcal{J}$ at non-differentiability points (which is in general quite a hard task) to check if the hypothesis is satisfied everywhere.*

Given $\omega > 0$, and $u \in \mathcal{E}$ we will denote

$$\mathcal{J}_{\omega,u}(v) := \mathcal{J}(v) + \omega\|v - u\|^2 \tag{2.4}$$

Notice that $\mathcal{J}_{\omega,u}$ is coercive whenever $\mathcal{J}$ is bounded from below. Thus assumption (A0) is guaranteed without any further restriction and it could be omitted. However, we leave it explicitly just to make clearer the requirement. We additionally observe that, if $\mathcal{J}$ satisfies (A1) we can always assume that ω is chosen in such a way that $\mathcal{J}_{\omega,u}$ is also ν-strongly convex with ν depending on $\mathcal{J}$ and ω, but not on u. Analogously, if (A1) and (A2) are satisfied, by using (A.8) it is easy to see that $\mathcal{J}_{\omega,u}$ satisfies (2.3) with two constants $\widetilde{K}, \widetilde{L}$ depending again on $\mathcal{J}$ and ω, but not on u.

2.1. Nonconvex Augmented Lagrangian algorithm

In this section we propose a *very general* and *simple* iterative algorithm to solve nonsmooth and nonconvex optimization problems with linear constraints that first appeared in [19]. In particular, for nonsmoothness we mean that we require our objective function

to be in general only a *locally* Lipschitz function, contrary to the much more restrictive $\mathcal{C}^2$, or $\mathcal{C}^{1,1}$ regularity requested by most of the above mentioned known methods for providing convergence guarantees, as in [21].
One of the most relevant features of our iteration is its *unconditionally guaranteed* convergence. By this we mean that the initial state does not need to be in a small neighborhood of a critical point which may not necessarily be the global minimum. Our algorithm may in fact be viewed as an appropriate combination of different techniques, resulting in a nested double loop iteration, where in the inner loop an augmented Lagrangian algorithm, or Bregman iteration, performs an adaptive finite number of iterations on a fixed *local* quadratic perturbation of the objective energy around the previous iteration, while the external loop performs an adaptation of the quadratic perturbation, similarly to SQP. Our analysis of convergence is confined to the setting of finite dimensional Euclidean spaces. Nevertheless, most of it could be done in the more general framework of (possibly infinite dimensional) Hilbert spaces since the only point where finite dimensionality is actually needed, is to recover strong compactness in the proof of Theorem 2.8.

2.1.1. Augmented Lagrangian algorithm for convex problems

We now recall some basic facts about Augmented Lagrangian iterations for constrained minimization of convex functionals. Here, we are given a *coercive convex functional* $\widetilde{\mathcal{J}}$. Then, for every $k \in \mathbb{N}$, $k \geq 1$, we define:

Algorithm 2.1 Augmentel Lagrangian algorithm for convex functionals

1: Take $q_0 \in \mathcal{F}$;
2: Initialize $k = 0$;
3: **while** $\|Av_k - f\| \neq 0$ **do**
4: $\quad v_{k+1} \in \arg\min_{v \in \mathcal{E}}(\widetilde{\mathcal{J}}(v) - \langle q_k, Av\rangle + \lambda\|Av - f\|^2)$;
5: $\quad q_{k+1} = q_k + 2\lambda(f - Av_{k+1})$;
6: $\quad k \leftarrow k + 1$;
7: **end while**

Convergence of the algorithm has been proved in [185], where it was called *Bregman Iteration*, and also in [112], being equivalent to the *Augmented Lagrangian Method* [137]. Precisely it has been shown that $\|Av_k - f\|$ decreases to 0 as k tends to $+\infty$, that the sequence v_k is compact and any limit point is a global minimum of $\widetilde{\mathcal{J}}$ under the constraint $Av = f$. Moreover, for every $k \geq 1$, $A^*q_k \in \partial\mathcal{J}(v_k)$. When $\widetilde{\mathcal{J}}$ is ν-strongly convex for some $\nu > 0$ we have also a quantitative estimate of the convergence of v_k to the unique (due to strict convexity) minimizer of the problem. We give a precise statement and a proof of this additional property, as it will be very useful later in the nonconvex case as well.

Proposition 2.4. *Assume that $\widetilde{\mathcal{J}}$ is ν-strongly convex, let v_k and q_k the sequences generated by Algorithm 2.1, and let $\bar{v}$ the unique global minimizer of $\widetilde{\mathcal{J}}$ on the affine space $\{v \in \mathcal{E} : Av = f\}$. Then:*

(i) $(\|Av_k - f\|)_{k\in\mathbb{N}}$ *is a decreasing sequence;*

(ii) $\lim_{k\to+\infty} \|Av_k - f\| = 0$*;*

(iii) $\|v_k - \bar{v}\|^2 \leq \frac{1}{\nu}\|q_0 - \bar{q}\|\,\|Av_k - f\|$*, for all* $k \in \mathbb{N}$*,*

for every $\bar{q} \in \mathcal{F}$ such that $A^*\bar{q} \in \partial\widetilde{\mathcal{J}}(\bar{u})$.

Proof. Properties *(i)* and *(ii)* are proved in [185]. For the property *(iii)*, we first observe that such a $\bar{q}$ surely exists by (2.2). We define for all $k \geq 1$ the discrepancy $\Delta q_k := q_k - \bar{q}$, and we prove that $\|\Delta q_k\|$ is decreasing. We actually have, by elementary computations and using Algorithm 2.1, that

$$\begin{aligned}\|\Delta q_k\|^2 - \|\Delta q_{k-1}\|^2 &\leq 2\langle q_k - q_{k-1}, q_k - \bar{q}\rangle \\ &= 4\lambda\langle f - Av_k, q_k - \bar{q}\rangle \\ &= 4\lambda\langle \bar{v} - v_k, A^*q_k - A^*\bar{q}\rangle\,.\end{aligned}$$

Since $A^*q_k \in \partial\widetilde{\mathcal{J}}(v_k)$ and $A^*\bar{q} \in \partial\widetilde{\mathcal{J}}(\bar{u})$, the last term in the inequality is nonpositive by (A.10), therefore the claim follows. In particular

$$\|q_k - \bar{q}\| \leq \|q_0 - \bar{q}\|\,, \tag{2.5}$$

for all $k \geq 1$. Now, by (A.11), we have also

$$\nu\|v_k - \bar{v}\|^2 \leq \langle A^*q_k - A^*\bar{q}, v_k - \bar{v}\rangle = \langle q_k - \bar{q}, Av_k - f\rangle\,,$$

so that we conclude by the Cauchy-Schwarz inequality and (2.5). □

Replacing $\widetilde{\mathcal{J}}$ with $\mathcal{J}_{\omega,u}$ defined by (2.4), with an appropriate choice of ω, by the previous result, (2.1), and (2.3), we get the following corollary, whose rather immediate proof is therefore omitted.

Corollary 2.5. *Consider the function $\mathcal{J}_{\omega,u}$ defined by* (2.4)*, where ω is chosen in such a way that $\mathcal{J}_{\omega,u}$ is ν-strongly convex with ν not depending on u. Let $\bar{v}_u$ be the unique global minimizer of $\mathcal{J}_{\omega,u}$ on the affine space $\{v \in \mathcal{E} : Av = f\}$. Then there exist two positive constants C_1 and C_2 depending on A^*, $\mathcal{J}$, and ω, but not on u, such that*

$$\|v_{k,u} - \bar{v}_u\|^2 \leq [C_1(1 + \|q_0\|) + C_2\mathcal{J}_{\omega,u}(\bar{v}_u)]\,\|Av_{k,u} - f\|\,, \tag{2.6}$$

where $v_{k,u} := v_k$ is defined accordingly to Algorithm 2.1 for $\widetilde{\mathcal{J}} = \mathcal{J}_{\omega,u}$.

2.1.2. The algorithm in the nonconvex case

We now present the new algorithm for linearly constrained nonsmooth and nonconvex minimization, and discuss its convergence properties. We pick initial $v_0 \in \mathcal{E}$ and $q_0 \in \mathcal{F}$. Notice that there is no restriction to any specific neighborhood for the choice of the initial iteration. For a fixed scaling parameter $\lambda > 0$, and an *adaptively chosen* sequence of integers $(L_\ell)_{\ell\in\mathbb{N}}$, for every integer $\ell \geq 1$ we set (with the convention $L_0 = 0$):

Algorithm 2.2 Bregman Iteration for nonconvex functionals

1: Take $v_0 \in \mathcal{E}$, $q_0 \in \mathcal{F}$;
2: Initialize $\ell = 1$;
3: **while** $\|v_\ell - v_{\ell-1}\| \neq 0$ **do**
4: Set $v_{(\ell,0)} = v_{\ell-1}$, $q_{(\ell,0)} = q_{\ell-1}$;
5: Compute L_ℓ;
6: **for** $k = 1, \dots, L_\ell$ **do**
7: $v_{(\ell,k)} = \underset{v\in\mathcal{E}}{\arg\min}\ (\mathcal{J}_{\omega,v_{\ell-1}}(v) - \langle q_{(\ell,k-1)}, Av\rangle + \lambda\|Av - f\|^2)$;
8: $q_{(\ell,k)} = q_{(\ell,k-1)} + 2\lambda(f - Av_{(\ell,k)})$;
9: $k \leftarrow k + 1$;
10: **end for**
11: $v_\ell := v_{(\ell,L_\ell)}$, $q_\ell := q_{(\ell,L_\ell)}$;
12: $\ell \leftarrow \ell + 1$;
13: **end while**

Here, thanks to condition (A1), ω is chosen in such a way that $\mathcal{J}_{\omega,v_{\ell-1}}$ is ν-strongly convex, with ν independent of $v_{\ell-1}$, and the *finite* number of inner iterates L_ℓ calculated in line 5 of the algorithm is defined by the condition

$$(1 + \|q_{\ell-1}\|)\|Av_{(\ell,L_\ell)} - f\| \leq \frac{1}{\ell^\alpha}\,, \tag{2.7}$$

for a given parameter $\alpha > 1$.
Since the inner loops are simply the Augmented Lagrangian iterations for the functional $\mathcal{J}_{\omega,v_{\ell-1}}$, by Proposition 2.4 *(ii)* and (2.5) such an integer L_ℓ always exists. We also remark that by construction, for every $\ell \geq 1$ and $k = 1, \dots, L_\ell$, we have

$$A^* q_{(\ell,k)} \in \partial\mathcal{J}_{\omega,v_{\ell-1}}(v_{(\ell,k)})\,. \tag{2.8}$$

Moreover, for every $\ell \geq 1$, again by Proposition 2.4, $\|Av_{(\ell,k)} - f\|$ is nonincreasing in k. Let us also remark that Algorithm 2.2, which can also be viewed as an implementation of an implicit gradient descent with step $1/\omega$, is actually a natural generalization of Algorithm 2.1. Indeed, if $\mathcal{J} = \widetilde{\mathcal{J}}$ were actually convex, we could in fact choose $\omega = 0$, and Algorithm 2.2 would simply reduce to Algorithm 2.1.

2.1.3. Analysis of convergence

We now want to analyse the convergence properties of the Algorithm 2.2. To do that we will use the following basic calculus lemma.

Lemma 2.6. *Let $(a_\ell)_{\ell\in\mathbb{N}}$ a sequence of positive numbers, and let $(\delta_\ell)_{\ell\in\mathbb{N}}$ a positive decreasing sequence such that*

$$\sum_{\ell=0}^{\infty}\delta_\ell < +\infty\,.$$

If a_ℓ satisfies for every ℓ the inequality

$$a_\ell \le (1+\delta_{\ell-1})a_{\ell-1} + \delta_{\ell-1}\,, \tag{2.9}$$

then $(a_\ell)_{\ell\in\mathbb{N}}$ is a convergent sequence.

Proof. By the recurrence relation (2.9) we deduce

$$a_\ell \le \left[\prod_{k=0}^{\ell-1}(1+\delta_k)\right] a_0 + \sum_{\ell'=0}^{\ell-1}\left[\prod_{k=\ell'+1}^{\ell-1}(1+\delta_k)\right]\delta_{\ell'}. \tag{2.10}$$

Notice that

$$\begin{aligned}\log\left[\prod_{k=0}^{\infty}(1+\delta_k)\right] &= \sum_{k=0}^{\infty}\log(1+\delta_k)\\ &= \sum_{k=0}^{\infty}\left(\delta_k - \frac{1}{2\xi_k}\delta_k^2\right) < \infty,\end{aligned} \tag{2.11}$$

for suitable $\xi_k \in (1, 1+\delta_k)$, for $k\in\mathbb{N}$, hence

$$\prod_{k=0}^{\infty}(1+\delta_k) < \infty,$$

and, together with (2.10), we deduce that $(a_\ell)_{\ell\in\mathbb{N}}$ is actually uniformly bounded. Now, again by the recurrence relation (2.9), for $k' \le k$, we obtain

$$a_k = a_{k'} + \sum_{\ell=k'+1}^{k}(a_\ell - a_{\ell-1}) \le a_{k'} + \sum_{\ell=k'+1}^{k}\delta_{\ell-1}a_{\ell-1} + \sum_{\ell=k'+1}^{k}\delta_{\ell-1}.$$

Taking first the lim sup as $k\to+\infty$ and then the lim inf as $k'\to+\infty$ in the previous inequality, we conclude from the boundedness of $(a_\ell)_{\ell\in\mathbb{N}}$ and the convergence of the series $\sum_{\ell=0}^{\infty}\delta_\ell$ that $\limsup_{k\to+\infty} a_k \le \liminf_{k'\to+\infty} a_{k'}$, which implies the conclusion. □

In the following theorem we analyse the convergence properties of the proposed algorithm.

Theorem 2.7. *Assume that $\mathcal{J}$ satisfies (A1) and (A2), and let $(v_\ell)_{\ell\in\mathbb{N}}$ be the sequence generated by Algorithm 2.2. Then,*

(a) $(Av_\ell - f) \to 0$ *as* $\ell \to \infty$*;*

(b) $(v_\ell - v_{\ell-1}) \to 0$ *as* $\ell \to \infty$*.*

If in addition $\mathcal{J}$ is coercive on the affine space $\{v \in \mathcal{E} : Av = f\}$, then v_ℓ is bounded and $(\mathcal{J}(v_\ell))_{\ell\in\mathbb{N}}$ is a convergent sequence. More in general, if $\mathcal{J}$ only satisfies (A1) and (A2), the following implication holds:

$$\text{if } (v_\ell)_{\ell\in\mathbb{N}} \text{ is a bounded sequence, then } (\mathcal{J}(v_\ell))_{\ell\in\mathbb{N}} \text{ is convergent.} \tag{2.12}$$

Proof. Part *(a)* of the statement is a direct consequence of the construction of v_ℓ and Proposition 2.4 *(ii)*. We now set for every ℓ

$$\bar v_\ell := \operatorname*{arg\,min}_{Av=f} \mathcal{J}_{\omega,v_{\ell-1}}(v)\,. \tag{2.13}$$

Notice that by definition $\bar v_\ell$ coincides with the element $\bar v_u$ considered in Corollary 2.5 when $u = v_{\ell-1}$. Similarly the element v_ℓ given by Algorithm 2.2 coincides with the element $v_{k,u}$ considered in Corollary 2.5 when $k = L_\ell$ and $u = v_{\ell-1}$. Therefore (2.6) with $q_0 = q_{(\ell,0)}$ and (2.7) imply there exist two positive constants C_1 and C_2 independent of ℓ, such that

$$\|v_\ell - \bar v_\ell\|^2 \le [C_1 + C_2\mathcal{J}_{\omega,v_{\ell-1}}(\bar v_\ell)]\frac{1}{\ell^\alpha}\,. \tag{2.14}$$

By this latter estimate and the minimality of $\bar v_{\ell+1}$ we get

$$\begin{aligned}\mathcal{J}_{\omega,v_\ell}(\bar v_{\ell+1}) &= \mathcal{J}(\bar v_{\ell+1}) + \omega\|v_\ell - \bar v_{\ell+1}\|^2 \le \mathcal{J}(\bar v_\ell) + \omega\|v_\ell - \bar v_\ell\|^2\\ &\le \mathcal{J}(\bar v_\ell) + \frac{C_1\omega}{\ell^\alpha} + \frac{C_2\omega}{\ell^\alpha}\mathcal{J}_{\omega,v_{\ell-1}}(\bar v_\ell)\\ &\le \frac{C_1\omega}{\ell^\alpha} + \left(1 + \frac{C_2\omega}{\ell^\alpha}\right)\mathcal{J}_{\omega,v_{\ell-1}}(\bar v_\ell)\,.\end{aligned} \tag{2.15}$$

By Lemma 2.6 we eventually deduce that $(\mathcal{J}_{\omega,v_{\ell-1}}(\bar v_\ell))_{\ell\in\mathbb{N}}$ is a convergent sequence, in particular it is bounded. Therefore, there exists a constant C independent of ℓ such that, by (2.14),

$$\|v_{\ell+1} - \bar v_{\ell+1}\|^2 \le \frac{C}{(\ell+1)^\alpha}\,, \tag{2.16}$$

and, by (2.15), we have also

$$\mathcal{J}(\bar v_{\ell+1}) \le \mathcal{J}(\bar v_{\ell+1}) + \omega\|v_\ell - \bar v_{\ell+1}\|^2 \le \mathcal{J}(\bar v_\ell) + \frac{C}{\ell^\alpha}\,. \tag{2.17}$$

Again Lemma 2.6 entails now that

$$\mathcal{J}(\bar v_\ell) \text{ is a convergent sequence}\,, \tag{2.18}$$

so that, by (2.17) we get that $(v_\ell - \bar{v}_{\ell+1}) \to 0$ as ℓ goes to $+\infty$, and this vanishing convergence, combined with (2.16), gives part *(b)* of the statement.
Being $\mathcal{J}$ locally Lipschitz as it is an ω-semi-convex function, if v_ℓ is uniformly bounded, by (2.16) and (2.18) we immediately conclude that $(\mathcal{J}(v_\ell))_{\ell\in\mathbb{N}}$ is a convergent sequence. Moreover, if $\mathcal{J}$ is coercive on the affine space $\{v \in \mathcal{E} : Av = f\}$, then $\bar{v}_\ell$ is bounded by (2.18), and so is also $(v_\ell)_{\ell\in\mathbb{N}}$ by (2.16), as required. □

As a consequence we get our main result of this section. Whenever v_ℓ is bounded, every cluster point is a constrained critical point of $\mathcal{J}$ on the affine space $\{v \in \mathcal{E} : Av = f\}$. We again recall that boundedness of v_ℓ is guaranteed by Theorem 2.7 when $\mathcal{J}$ is assumed to be coercive on the above affine space.

Theorem 2.8. *Assume that $\mathcal{J}$ satisfies (A1) and (A2), and let $(v_\ell)_{\ell\in\mathbb{N}}$ be the sequence generated by Algorithm 2.2. If $(v_\ell)_{\ell\in\mathbb{N}}$ is bounded, every of its limit points is a constrained critical point of $\mathcal{J}$ on the affine space $\{v \in \mathcal{E} : Av = f\}$.*

Proof. Let $(q_\ell)_{\ell\in\mathbb{N}}$ be the sequence defined by Algorithm 2.2, and let $p_\ell := A^* q_\ell$, and $\hat{p}_\ell := p_\ell - 2\omega(v_\ell - v_{\ell-1})$. By (A.8) and (2.8), we have

$$\hat{p}_\ell \in \partial\mathcal{J}(v_\ell)\,, \tag{2.19}$$

and by the boundedness of $(v_\ell)_{\ell\in\mathbb{N}}$, (2.12), and (A2), we then get that $\hat{p}_\ell$ is bounded too. By Theorem 2.7, part *(b)*, we deduce that $p_\ell - \hat{p}_\ell \to 0$, which in particular gives

$$\lim_{\ell\to+\infty} \operatorname{dist}(\hat{p}_\ell\,, \operatorname{ran}(A^*)) = 0\,. \tag{2.20}$$

Now, if a subsequence $v_{\ell_j} \to v \in \mathcal{E}$, possibly taking a further subsequence we may assume that $\hat{p}_{\ell_j} \to \hat{p} \in \partial\mathcal{J}(v)$, where the last inclusion follows from (A.9) and (2.19). Moreover, since in finite dimension $\operatorname{ran}(A^*)$ is closed, by (2.20) $\hat{p} \in \operatorname{ran}(A^*)$. Since $Av = f$ by part *(a)* of Theorem 2.7, (2.2) yields now the desired conclusion. □

2.2. Nonconvex Split Bregman Iteration

An alternative algorithm to which the strategy described above can be applied is known as *Split Bregman Iteration* and it was introduced in [53]. In particular, this procedure has been designed to minimize convex functionals subject to linear constraints or being the combination of two convex term. Note that functional such as the ones considered in the previous section fall into this class of functionals. The idea of this method is to introduce a new variable and then perform the minimization of the functional in two separate instances, dealing either with the linear system or with the convex component. This ansatz can be computationally convenient in case the solution of the two separated problems can be performed more efficiently than a direct approach. This algorithm has

been already successfully applied to different problems, see e.g. [121, 189, 196, 213, 216], among them image processing, denoising, and segmentation [53, 120, 215], which actually was the problem that inspired the development of the algorithm.
In this section, following the scheme of Section 2.1, we introduce an extension of the Split Bregman Iteration via a vanishing quadratic regularization process as the one used by Algorithm 2.2. We additionally show that, if the functional $\mathcal{J}$ to be minimized is a $\mathcal{C}^1$ perturbation of a convex functional, i.e., it can be rewritten as $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$, being $\mathcal{J}_1$ is lower semicontinuous, and $\mathcal{J}_2$ a $\mathcal{C}^1$ regular, and is endowed with assumption (A1) and (A2), then the procedure we propose converges to a critical point of $\mathcal{J}$.

2.2.1. Split Bregman Iteration for convex functionals

In this section we introduce the Split Bregman Iteration for convex linearly constrained minimization problem in order to set all the preliminaries needed for a full understanding of the nonconvex case.
As before, we denote by $\tilde{\mathcal{J}}$ a convex and coercive functional (to be distinguished from the nonconvex $\mathcal{J}$), by $A : \mathcal{E} \to \mathcal{F}$ the linear operator, and by $f \in \mathcal{F}$ the fidelity term that must be matched by Av, with $v \in \mathcal{E}$. Now, following [53], we modify the minimization problem in Line 4 of Algorithm 2.1 adding a variable d as follows

$$\underset{\substack{v,d\in\mathcal{E}\\ d=v}}{\arg\min} \left(\tilde{\mathcal{J}}(v) - \langle q, Ad\rangle + \lambda \|Ad - f\|^2 \right). \tag{2.21}$$

Notice that the minimization problem (2.21) can now be naturally separated in two different subproblems since the convex term does not depend on d and, similarly, the variable v does not appear in the linear part. Moreover, since we added the constraint $v = d$ the two variable are effectively forced to coincide and, any solution of the new minimization problem is also a solution of the original one.
Using once again the Augmented Lagrangian approach, see [112, 137], given two constants $0 < \mu \leq 1$, $\lambda > 0$, we rewrite (2.21) as an unconstrained minimization problem:

$$\underset{v,d\in\mathcal{E}}{\arg\min} \left(\tilde{\mathcal{J}}(v) - \langle q, Ad\rangle + \lambda \|Ad - f\|^2 - \langle b, d - v\rangle + \mu \|u - d\|^2 \right), \tag{2.22}$$

where $b \in \mathcal{E}^*$, with $\mathcal{E}^*$ the dual space of $\mathcal{E}$, is the Lagrangian multiplier related to the constraint $v = d$.

Now, fixed the constant $0 < \delta_q < 2$, we can finally write the *Split Bregman Iteration* for convex functionals

Algorithm 2.3 Split Bregman Iteration for convex functionals

1: Take $q_0 \in \mathcal{F}$, $b_0 \in \mathcal{E}^*$, $v_0 \in \mathcal{E}$;
2: Initialize $k = 1$;
3: **while** $\|Av_{k-1} - f\| \neq 0$ **do**
4: $\quad d_k = \arg\min_d \lambda\|Ad - f\|^2 - \langle q_{k-1}, Ad\rangle + \mu\|d - v_{k-1}\|^2 + \langle b_{k-1}, d - v_{k-1}\rangle$;
5: $\quad v_k = \arg\min_v \widetilde{\mathcal{J}}(v) + \mu\|v - d_k\|^2 + \langle b_{k-1}, d_k - v\rangle$;
6: $\quad b_k = b_{k-1} + \mu(d_k - v_k)$;
7: $\quad q_k = q_{k-1} - \delta_q(Ad_k - f)$;
8: $\quad k - 1 \leftarrow k$;
9: **end while**

We now report the convergence result for the above algorithm presented in [53]. We additionally report the proof of the result, despite it has no new mathematical contribution, with the sole scope of specializing the notation to our specific setting of ν-convex functionals and with the idea of giving a complete and clear overview of the whole procedure.

Theorem 2.9. *Given $\bar{v}$ solution of the minimization problem* (2.13)*, where $\mathcal{J} = \tilde{\mathcal{J}}$. a $\nu-$strongly convex and coercive functional, and that $\mu, \lambda > 0$, and $0 < \delta_q < 2\lambda$. Then, the following properties for Algorithm 2.3 hold:*

$$\lim_{k\to+\infty} \|Av_k - f\| = 0, \qquad \lim_{k\to+\infty} \widetilde{\mathcal{J}}(v_k) = \widetilde{\mathcal{J}}(\bar{v}).$$

Furthermore,

$$\lim_{k\to+\infty} \|v_k - \bar{v}\| = 0 \qquad \lim_{k\to+\infty} \|d_k - \bar{d}\| = 0$$

whenever (2.13) *has a unique solution.*

Proof. First order optimality conditions give us the following:

$$\begin{cases} 0 = \lambda A^T(Ad_k - f) - A^T q_{k-1} + \mu(d_k - v_{k-1}) + b_{k-1} \\ 0 = b_k + \mu(v_k - d_k) - b_{k-1}, \qquad \text{with } b_k \in \partial\tilde{\mathcal{J}}(v_k) \\ b_k = b_{k-1} + \mu(d_k - v_k) \\ q_k = q_{k-1} - \delta_q(Ad_k - f). \end{cases} \tag{2.23}$$

Moreover, the above condition in the fixed point of (2.3) is

$$\begin{cases} 0 = \lambda A^T(A\bar{d} - f) - A^T\bar{q} + \mu(\bar{d} - \bar{v}) + \bar{b} \\ 0 = \bar{b} + \mu(\bar{v} - \bar{d}) - \bar{b}, & \text{with } \bar{b} \in \partial\widetilde{\mathcal{J}}(\bar{v}) \\ \bar{b} = \bar{b} + \mu(\bar{d} - \bar{v}) \\ \bar{q} = \bar{q} - \delta_q(A\bar{d} - f). \end{cases} \tag{2.24}$$

Let us now denote the errors by

$$v_k^e = v_k - \bar{v} \qquad d_k^e = d_k - \bar{d} \qquad q_k^e = q_k - \bar{q} \qquad b_k^e = b_k - \bar{b}.$$

Subtracting the first equation of (2.23) to the first (2.24), we obtain

$$0 = \lambda A^T(Ad_k^e) - A^T q_{k-1}^e + \mu(d_k^e - v_{k-1}^e) + b_{k-1}^e.$$

Now, taking the inner product of the left- and right-hand sides with respect to d_k^e, we have

$$0 = \lambda\|Ad_k^e\|^2 + \mu\|d_k^e\|^2 - \langle q_{k-1}^e, Ad_k^e\rangle - \mu\langle v_{k-1}^e, d_k^e\rangle + \langle b_{k-1}^e, d_k^e\rangle. \tag{2.25}$$

By doing the same operation on the second equations of (2.23) and (2.24), we get

$$0 = \mu\|v_k^e\|^2 + \langle b_k^e, v_k^e\rangle - \mu\langle d_k^e, v_k^e\rangle - \langle b_{k-1}^e, v_k^e\rangle. \tag{2.26}$$

The sum of (2.25) with (2.26) gives us

$$0 = \lambda\|Ad_k^e\|^2 + \langle b_k^e, v_k^e\rangle - \langle q_{k-1}^e, Ad_k^e\rangle + \langle b_{k-1}^e, d_k^e - v_k^e\rangle + \mu(\|d_k^e\|^2 + \|v_k^e\|^2 - \langle d_k^e, v_k^e + v_{k-1}^e\rangle) \tag{2.27}$$

Subtracting the third and fourth equations of (2.23) and (2.24) and squaring their norm we obtain respectively

$$\langle b_{k-1}^e, d_k^e - v_k^e\rangle = \frac{1}{2\mu}(\|b_k^e\|^2 - \|b_{k-1}^e\|^2) - \frac{\mu}{2}\|d_k^e - v_k^e\|^2 \tag{2.28}$$

$$-\langle q_{k-1}^e, Ad_k^e\rangle = \frac{1}{2\delta_q}(\|q_k^e\|^2 - \|q_{k-1}^e\|^2) - \frac{\delta_q}{2}\|Ad_k^e\|^2. \tag{2.29}$$

Now, we plug the last two equation into (2.27) and we get

$$\begin{aligned} &\frac{1}{2\delta_q}(\|q_{k-1}^e\|^2 - \|q_k^e\|^2) + \frac{1}{2\mu}(\|b_{k-1}^e\|^2 - \|b_k^e\|^2) = \left(\lambda - \frac{\delta_q}{2}\right)\|Ad_k^e\|^2 + \langle b_k^e, v_k^e\rangle \\ &+\mu(\|d_k^e\|^2 + \|v_k^e\|^2 - \langle d_k^e, v_k^e + v_{k-1}^e\rangle - \frac{1}{2}\|d_k^e - v_k^e\|^2). \end{aligned} \tag{2.30}$$

We now sum from 1 to K and with some algebraic operation we obtain

$$\begin{aligned} &\frac{1}{2\delta_q}(\|q_0^e\|^2 - \|q_K^e\|^2) + \frac{1}{2\mu}(\|b_0^e\|^2 - \|b_K^e\|^2) = \sum_{k=1}^{K}\left[\left(\lambda - \frac{\delta_q}{2}\right)\|Ad_k^e\|^2 + \langle b_k^e, v_k^e\rangle\right] \\ &+\frac{\mu}{2}\sum_{k=1}^{K}\|d_k^e - v_{k-1}^e\|^2 + \frac{\mu}{2}\left(\|v_K^e\|^2 - \|v_0^e\|^2\right). \end{aligned} \tag{2.31}$$

After some algebraic operations, we get

$$\frac{1}{2\delta_q}\|q_0^e\|^2+\frac{1}{2\mu}\|b_0^e\|^2+\frac{\mu}{2}\|v_0^e\|^2 \geq \sum_{k=1}^{K}\left[\left(\lambda-\frac{\delta_q}{2}\right)\|Ad_k^e\|^2+\langle b_k-\bar{b},v_k-\bar{v}\rangle+\frac{\mu}{2}\|d_k^e-v_{k-1}^e\|^2\right]. \tag{2.32}$$

Thanks to this latter inequality and the hypothesis made on λ and δ_q, we obtain

$$\lim_{k\to\infty}\|Ad_k-f\|=0. \tag{2.33}$$

$$\lim_{k\to\infty}\langle b_k-\bar{b},v_k-\bar{v}\rangle=0 \tag{2.34}$$

We now recall that $b_k\in\partial\tilde{\mathcal{J}}(v_k)$ for $k\geq 0$, $\bar{b}\in\partial\tilde{\mathcal{J}}(\bar{v})$, and the functional $\tilde{\mathcal{J}}$ is ν-convex, i.e.,

$$\langle b_k-\bar{b},v_k-\bar{v}\rangle\geq\nu\|v_k-\bar{v}\|^2,$$

which combined with (2.34) implies

$$\lim_{k\to\infty}\|v_k-\bar{v}\|^2\geq 0 \tag{2.35}$$

The latter limit and the convergence result stemmed from the last summand of inquality (2.32), together with the equality $\bar{d}=\bar{v}$, give us

$$\lim_{k\to\infty}\|d_k-v_{k-1}\|=0. \tag{2.36}$$

The limit $\lim_{k\to\infty}\|d_k-v_k\|=0$ is easily achieved combining (2.35) with (2.36).
We now focus on the convergence of $\tilde{\mathcal{J}}(v_k)$ to $\tilde{\mathcal{J}}(\bar{v})$ for k going to infinity. Recall that, for any convex function $\tilde{\mathcal{J}}$, the Bregman distance $B^{\cdot}_{\tilde{\mathcal{J}}}(\cdot,\cdot)$ satisfies

$$B^p_{\tilde{\mathcal{J}}}(u,z)+B^q_{\tilde{\mathcal{J}}}(z,u)=\langle q-p,u-z\rangle\qquad\forall p\in\partial\tilde{\mathcal{J}}(z),\forall q\in\partial\tilde{\mathcal{J}}(u).$$

This, together with (2.34), (2.36), and the nonnegativity of the Bregman distance, leads to

$$\lim_{k\to\infty}\tilde{\mathcal{J}}(v_k)-\tilde{\mathcal{J}}(\bar{v})-\langle v_k-\bar{v},\bar{b}\rangle=0. \tag{2.37}$$

Since $\lim_{k\to\infty}\|v_k-\bar{v}\|=0$ we get the limit

$$\lim_{k\to\infty}\tilde{\mathcal{J}}(v_k)=\tilde{\mathcal{J}}(\bar{v})$$

and this conclude the proof. □

This theorem gives us not only a proof of convergence of the Split Bregman iteration to a global minimizer of the ν-strongly convex functional $\tilde{\mathcal{J}}$, but we can also derive further convergence properties which will be useful for the nonconvex case.

Corollary 2.10. *Given $\bar{v}$ solution of the minimization problem* (2.13)*, where $\mathcal{J} = \tilde{\mathcal{J}}$. a $\nu-$strongly convex and coercive functional, and that $\mu, \lambda > 0$, and $0 < \delta_q < 2\lambda$. If $\{v_k\}_k$, $\{b_k\}_k$, $\{d_k\}_k$, and $\{q_k\}_k$ is the sequence generated by Algorithm 2.3, we have the following properties:*

(i) $\lim_{k\to+\infty} \|d_k - v_k\| = 0$

(ii) $\lim_{k\to+\infty} \|b_k - b_{k-1}\| = 0$

(iii) $\lim_{k\to+\infty} \|q_k - q_{k-1}\| = 0$

Proof. The proof of *(i)* follows immediately combining (2.36) with (2.35). *(ii)* is given by the third equation of (2.23) and *(i)*, while *(iii)* is derived from the fourth equation of (2.23) and (2.33). □

Analysis of convergence

One of the biggest drawback of the splitting approach is that we do not have anymore the following property:

$$v_k \in \mathrm{ran}(A^*) \qquad \forall k \in \mathbb{N}$$

which had been used in [19] to prove a speed of convergence of the inner loop of Algorithm 2.2 and hence to determine a concrete stopping criterion for the loop solving the convex minimization problem. Here, we give an alternative result relative to Algorithm 2.3 which is the first step towards a stopping criterion for the *nonconvex Split Bregman iteration*.

Proposition 2.11. *Let $\mathcal{J}$ be a $\nu-$convex functional and $\bar{v}$ the solution of the problem* (2.13)*. Then, being $\{v_k\}_k$, $\{b_k\}_k$, $\{d_k\}_k$, and $\{q_k\}_k$ the sequences generated by Algorithm 2.3, we have*

$$\begin{aligned}(\nu-\mu)\|v_k-\bar{v}\|^2 \le \lambda\|q_{k-1}-\bar{q}\|\|Ad_k-f\| + \mu\|d_k-\bar{d}\|\|v_k-v_{k-1}\| \\ + \|d_k-v_k\|(\|b_k-\bar{b}\| + \mu\|v_k-\bar{v}\| + \mu\|d_k-\bar{d}\|).\end{aligned}$$

Proof. Thanks to the $\nu-$convexity and by the fact that $\bar{v} = \bar{d}$ we obtain

$$\begin{aligned}\nu\|v_k-\bar{v}\|^2 &\le \langle v_k-\bar{v}, b_k-\bar{b}\rangle = \langle v_k-d_k, b_k-\bar{b}\rangle + \langle d_k-\bar{d}, b_k-\bar{b}\rangle \\ &= \langle v_k-d_k, b_k-\bar{b}\rangle + \langle d_k-\bar{d}, b_k-b_{k-1}\rangle + \langle d_k-\bar{d}, b_{k-1}-\bar{b}\rangle.\end{aligned} \tag{2.38}$$

We now plug (2.25) in (2.38) getting

$$\begin{aligned}\nu\|v_k-\bar v\|^2 &\le \langle v_k-d_k,b_k-\bar b\rangle+\langle d_k-\bar d,b_k-b_{k-1}\rangle-\lambda\|Ad_k-f\|^2\\ &\quad+\lambda\langle q_{k-1}-\bar q,Ad_k-f\rangle+\mu\langle v_{k-1}-\bar v,d_k-\bar d\rangle-\mu\|d_k-\bar d\|^2\\ &\le \langle v_k-d_k,b_k-\bar b\rangle+\langle d_k-\bar d,b_k-b_{k-1}\rangle+\lambda\langle q_{k-1}-\bar q,Ad_k-f\rangle\\ &\quad+\mu\langle v_{k-1}-\bar v,d_k-\bar d\rangle\\ &= \langle v_k-d_k,b_k-\bar b\rangle+\langle d_k-\bar d,b_k-b_{k-1}\rangle+\lambda\langle q_{k-1}-\bar q,Ad_k-f\rangle\\ &\quad-\mu\langle v_k-v_{k-1},d_k-\bar d\rangle+\mu\langle v_k-\bar v,d_k-\bar d\rangle.\end{aligned}\tag{2.39}$$

Now, we repeat the former procedure plugging (2.26) into (2.39), then we consider the third equation of (2.23) and we apply the Chauchy-Schwarz inequality obtaining

$$\begin{aligned}(\nu-\mu)\|v_k-\bar v\|^2 &\le \langle v_k-d_k,b_k-\bar b\rangle+\mu\langle d_k-\bar d,d_k-v_k\rangle+\lambda\langle q_{k-1}-\bar q,Ad_k-f\rangle\\ &\quad-\mu\langle v_k-v_{k-1},d_k-\bar d\rangle+\mu\langle d_k-v_k,v_k-\bar v\rangle\\ &\le \lambda\|q_{k-1}-\bar q\|\|Ad_k-f\|+\mu\|d_k-\bar d\|\|v_k-v_{k-1}\|\\ &\quad+\|d_k-v_k\|(\|b_k-\bar b\|+\mu\|v_k-\bar v\|+\mu\|d_k-\bar d\|)\end{aligned}\tag{2.40}$$

which completes the proof. □

Remark 2.12. *From inequality* (2.40) *we deduce that it is necessary to assume* $\nu>\mu$. *Luckily, this assumption does not affect the validity of the results in the nonconvex case.*

Notice that each summand of the right-hand side of (2.40) is given by the product of an element which by Theorem 2.9 goes to zero and a term which is bounded by the following

Corollary 2.13. *Let* $\mathcal{J}$ *be a* $\nu-$*strongly convex functional and* $\bar v$ *the solution of the problem* (2.21). *Then, being* $\{v_k\}_k$, $\{b_k\}_k$, $\{d_k\}_k$, *and* $\{q_k\}_k$ *the sequences generated by Algorithm 2.3 and assuming* $0<\mu<\nu$, *we have*

(i) $\|b_k-\bar b\|^2\le C_{11}\|q_0^e\|^2+C_{21}\|b_0^e\|^2+C_{31}\|v_0^e\|^2=K_b(q_0,b_0,v_0)^2$

(ii) $\|q_k-\bar q\|^2\le C_{12}\|q_0^e\|^2+C_{22}\|b_0^e\|^2+C_{33}\|v_0^e\|^2=K_q(q_0,b_0,v_0)^2$

(iii) $\|d_k-\bar d\|^2\le C_{13}\|q_0^e\|^2+C_{23}\|b_0^e\|^2+C_{33}\|v_0^e\|^2=K_d(q_0,b_0,v_0)^2$

Proof. We start from *(i)* and *(ii)* which can be proved using similar arguments. Let us take the third and the fourth equations of (2.23). The terms b_k and q_k can be rewritten

as

$$b_k = b_{k-1} + \mu(d_k - v_k) = b_0 + \mu \sum_{i=1}^{k} (d_i - v_i)$$

$$q_k = q_{k-1} - \delta_q(Ad_k - f) = q_0 + \delta_q \sum_{i=1}^{k} (f - Ad_i).$$

Therefore, using the triangle inequality and (2.32) we obtain

$$\|b_k - \bar{b}\|^2 \leq \|b_0 - \bar{b}\|^2 + \mu \sum_{i=1}^{k} \|d_i - v_i\|^2 \leq C_{11}\|q_0^e\|^2 + C_{21}\|b_0^e\|^2 + C_{31}\|v_0^e\|^2$$

$$\|q_k - \bar{q}\|^2 \leq \|q_0 - \bar{q}\|^2 + \delta_q \sum_{i=1}^{k} \|Ad_i - f\|^2 \leq C_{12}\|q_0^e\|^2 + C_{22}\|b_0^e\|^2 + C_{32}\|v_0^e\|^2.$$

A little bit of manipulation is needed to prove *(iii)*. Let us take (2.31), using *(i)* and *(ii)*, applying Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\mu\|d_k - \bar{d}\|^2 &\leq \frac{1}{2\delta_q}(\|q_{k-1}^e\|^2 - \|q_k^e\|^2) + \frac{1}{2\mu}(\|b_{k-1}^e\|^2 - \|b_k^e\|^2) - \left(\lambda - \frac{\delta_q}{2}\right)\|Ad_k^e\|^2 - \langle b_k^e, v_k^e\rangle \\
&\quad - \mu(\|v_k^e\|^2 - \langle d_k^e, v_k^e + v_{k-1}^e\rangle - \frac{\mu}{2\mu}\|d_k^e - v_k^e\|^2) \\
&\leq \frac{1}{2\delta_q}\|q_k^e\|^2 + \frac{1}{2\mu}\|b_k^e\|^2 \\
&\quad + \mu\langle d_k - v_k, v_k^e + v_{k-1}^e\rangle + \mu\langle v_k - \bar{v}, v_k^e + v_{k-1}^e\rangle + \frac{\mu}{2}\|d_k - v_k\|^2 \\
&\leq K_q(q_0, b_0, v_0)^2 + K_b(q_0, b_0, v_0)^2 + \mu|\langle d_k - v_k, v_k^e\rangle| + \mu|\langle d_k - v_k, v_{k-1}^e\rangle| \\
&\quad + \mu\langle v_k - \bar{v}, v_k^e\rangle + \mu\langle v_k - \bar{v}, v_{k-1}^e\rangle + \frac{\mu}{2}\|d_k - v_k\|^2 \\
&\leq K_q(q_0, b_0, v_0)^2 + K_b(q_0, b_0, v_0)^2 + \mu(\|d_k - v_k\|\|v_k - \bar{v}\| \\
&\quad + \|d_k - v_k\|\|v_{k-1} - \bar{v}\| + \|v_k - \bar{v}\|^2 + \|v_k - \bar{v}\|\|v_{k-1} - \bar{v}\| + \frac{1}{2}\|d_k - v_k\|^2).
\end{aligned}$$

Now, notice that every normed term in the inequality above is present also in (2.32), thus there exists three constants C_{13}, C_{23}, and C_{33} such that

$$\|d_k - \bar{d}\|^2 \leq C_{13}\|q_0^e\|^2 + C_{23}\|b_0^e\|^2 + C_{33}\|v_0^e\|^2$$

□

These results, together with Theorem 2.9 and the $\nu-$convexity property, tell us that there exist two constants $K_1, K_2 > 0$ such that

$$\|v_k - \bar{v}\|^2 \leq (K_1\|q_0^e\|^2 + K_2\|b_0^e\|^2)(\|v_k - d_k\| + \|v_k - v_{k-1}\| + \|Ad_k - f\|). \tag{2.41}$$

Remark 2.14. *To obtain the latter inequality we used the $\nu-$convexity as follows*

$$\nu\|v_0 - \bar{v}\|^2 \leq \langle v_0 - \bar{v}, b_0 - \bar{b}\rangle \leq \|v_0 - \bar{v}\|\|b_0 - \bar{b}\|.$$

2.2.2. Split Bregman Iteration for nonconvex functionals

Similarly to what we did for Algorithm 2.2, we now introduce the modification of Algorithm 2.3 to address problems of the type (2.13), where the functional $\mathcal{J}$ is as described at the beginning of the section. Recalling the technique used in Section 2.1, the algorithm nests the structure of the Split Bregman Iteration in a regularizing loop which updates a quadratic perturbation added to the functional. We prove not only that the perturbation term vanish at convergence, but also that the algorithm converges to critical points of $\mathcal{J}$.

Algorithm 2.4 Split Bregman Iteration for nonconvex functionals

Take $q_0 \in \mathcal{F}$, $b_0 \in \mathcal{E}^*$, $v_0 \in \mathcal{E}$;
Initialize $\ell = 1$;
while $\|v_\ell - v_{\ell-1}\| \neq 0$ **do**
  Set $v_{(\ell,0)} = v_{\ell-1}$, $q_{(\ell,0)} = q_{\ell-1}$;
  Set $d_{(\ell,0)} = d_{\ell-1}$, $b_{(\ell,0)} = b_{\ell-1}$;
  Compute L_ℓ;
  for $k = 0, \dots, L_\ell$ **do**
    $d_{(\ell,k+1)} = \arg\min_d \lambda\|Ad - f\|^2 - \langle q_{(\ell,k)}, Ad\rangle + \mu\|d - v_{(\ell,k)}\|^2 + \langle b_{(\ell,k)}, d - v_{(\ell,k)}\rangle$;
    $v_{(\ell,k+1)} = \arg\min_v \mathcal{J}_{\omega,v_{\ell-1}}(v) + \mu\|v - d_{(\ell,k+1)}\|^2 + \langle b_{(\ell,k)}, d_{(\ell,k+1)} - v\rangle$;
    $b_{(\ell,k+1)} = b_{(\ell,k)} + \mu(d_{(\ell,k+1)} - v_{(\ell,k+1)})$;
    $q_{(\ell,k+1)} = q_{(\ell,k)} - \delta_q(Ad_{(\ell,k+1)} - f)$;
    $k \leftarrow k + 1$;
  end for
  $v_\ell := v_{(\ell,L_\ell)}$, $q_\ell := q_{(\ell,L_\ell)}$;
  $d_\ell := d_{(\ell,L_\ell)}$, $b_\ell := b_{(\ell,L_\ell)}$;
  $\ell \leftarrow \ell + 1$;
end while

In the Algorithm, the functional

$$\mathcal{J}_{\omega,v_{\ell-1}}(\cdot) = \mathcal{J}(\cdot) + \omega\|\cdot - v_{\ell-1}\|^2 \qquad \forall \ell \in \mathbb{N} \tag{2.42}$$

is $\nu-$strongly convex for a proper choice of ω independent from $v_{\ell-1}$. Notice that this choice is always possible thanks to property (A1).
The adaptive finite number of iteration L_ℓ is determined upon the following condition:

$$\bar{K}(\|q_{\ell-1}\|^2 + \|b_{\ell-1}\|^2 + 1)(\|v_{(\ell,k)} - d_{(\ell,k)}\| + \|v_{(\ell,k)} - v_{(\ell,k-1)}\| + \|Ad_{(\ell,k)} - f\|) \leq \frac{1}{\ell^\alpha} \tag{2.43}$$

for a given parameter α and a constant $\bar{K}$.
Since the inner loops are the Split Bregman Iteration for the convex functional $\mathcal{J}_{\omega,v_{\ell-1}}(\cdot)$,

the results of Theorem 2.9, Proposition 2.11, and (2.41) tell us that such integer L_ℓ always exists.

Remark 2.15. *Notice that by construction, for every $\ell \geq 1$ and $k = 0, \dots, L_\ell$, we have*

$$b_{(\ell,k)} \in \partial \mathcal{J}_{\omega, v_{\ell-1}}(v_{(\ell,k)}).$$

Additionally, let us consider the first equation of (2.24) *for the convexified problem in the inner loop of Algorithm 2.4*

$$0 = \lambda A^T(A\bar{d} - f) - A^T\bar{q} + \mu(\bar{d} - \bar{v}) + \bar{b}.$$

Thanks to the results of Theorem 2.9 we get $A^T\bar{q} + \bar{b} = 0$ and since $\bar{b} \in \partial \mathcal{J}_{\omega, v_{\ell-1}}$, we deduce

$$A^T\bar{q} \in \partial \mathcal{J}_{\omega, v_{\ell-1}}. \tag{2.44}$$

We now introduce an estimate which is fundamental for the proof of convergence of the algorithm:

Corollary 2.16. *Consider the functional $\mathcal{J}_{\omega,u}$ defined by* (2.42)*, where ω is chosen in such a way that the functional is $\nu-$strongly convex with ν not depending on u. Let $\bar{v}_u$ be the unique global minimizer of $\mathcal{J}_{\omega,u}$ on the affine space $\{v \in \mathcal{E} : Av = f\}$. Then there exist two positive constants $\bar{K}_1$ and $\bar{K}_2$ depending on A^T, $\mathcal{J}$, and ω such that*

$$\begin{aligned}\|v_{k,u} - \bar{v}_u\|^2 \leq \Big[(\bar{K}_1 \left(1 + \|q_0\|^2 + \|b_0\|^2\right) + \bar{K}_2 \mathcal{J}_{\omega,u}(\bar{v}_u)\Big] \\ (\|v_k - d_k\| + \|v_k - v_{k-1}\| + \|Ad_k - f\|)\,,\end{aligned}$$

where $v_{k,u} := v_k$ is defined accordingly to (2.3) *for $\tilde{\mathcal{J}} = \mathcal{J}_{\omega,u}$.*

The proof of this Corollary can be easily achieved using the triangle inequality, property (A2), property (2.1), and Remark 2.15.

2.2.3. Analysis of convergence

In this section, we analyze the convergence properties of Algorithm 2.4 and we give a proof of convergence to critical point of the nonconvex functional $\mathcal{J}$ introduced in Section 2.2.
Recalling Lemma 2.6 and Theorem 2.7, we can now state the result which guarantees that, if v_ℓ is bounded, then the algorithm converges to a constrained critical point of the functional $\mathcal{J}$ on the affine space $\{v \in \mathcal{E} : Av = f\}$. Moreover, we stress that the boundedness of v_ℓ can be achieved adding an hypothesis of coercivity on $\mathcal{J}$.

Theorem 2.17. *Let $\mathcal{J}$ be a $\mathcal{C}^1$ perturbation of a convex functional satisfying (A1) and (A2). Additionally, let $\{v_\ell\}_\ell$ be the sequence generated by Algorithm 2.4. Then, each of its limit point is a constrained critical point in $\mathcal{J}$ on the affine space $\mathcal{A}(f) := \{v \in \mathcal{E} : Av = f\}$.*

Proof. Let us consider the first equation of the optimality conditions of our problem:

$$0 = \lambda A^T(Ad_\ell - f) - A^T q_{(\ell,L_\ell)} + \mu(d_\ell - v_{(\ell,L_\ell)}) + b_{(\ell,L_\ell)}.$$

By substituting respectively $q_{(\ell,L_\ell)}$ and $b_{(\ell,L_\ell)}$ with q_ℓ and b_ℓ given by Algorithm 2.4, we have

$$\begin{aligned} 0 &= \lambda A^T(Ad_\ell - f) - A^T q_\ell - \delta_q A^T(Ad_\ell - f) + \mu(d_\ell - v_{(\ell,L_\ell)}) + b_\ell - \mu(d_\ell - v_\ell) \\ &= (\lambda - \delta_q)A^T(Ad_\ell - f) - A^T q_\ell + b_\ell + \mu(v_\ell - v_{(\ell,L_\ell)}). \end{aligned}$$

Using the stopping criterion (2.43) of the inner loop of the algorithm we obtain

$$\|A^T q_\ell - b_\ell\| \leq \frac{C}{\ell^\alpha}$$

Therefore, we deduce that

$$\lim_{\ell\to\infty} \operatorname{dist}(b_\ell, \operatorname{ran}(A^*)) = 0. \tag{2.45}$$

Then, let us define $\hat{b}_\ell := b_\ell - \omega(v_\ell - v_{\ell-1})$. By the definition of subgradient, the particular choice of $\mathcal{J}$ as a $\mathcal{C}^1$ perturbation of a convex functional, and $b_\ell \in \partial\mathcal{J}_{\omega,v_{\ell-1}}(v_\ell)$ we have

$$\hat{b}_\ell \in \partial\mathcal{J}(v_\ell). \tag{2.46}$$

Since v_ℓ is bounded, $\{\mathcal{J}(v_\ell)\}_\ell$ converges, and (A2) holds, we also have that $\hat{b}_\ell$ is bounded. By Theorem 2.7 part *(b)*, we deduce that $\|\hat{b}_\ell - b_\ell\| \longrightarrow 0$ and in particular, by (2.45),

$$\lim_{\ell\to\infty} \operatorname{dist}(\hat{b}_\ell, \operatorname{ran}(A^*)) = 0.$$

Consider now take a subsequence $v_{\ell_j} \to v \in \mathcal{E}$. It is possible to extract a further subsequence such that $\hat{b}_{\ell_j} \to \hat{b} \in \partial\mathcal{J}(v)$, where the last inclusion comes again from Remark A.9 and (A.8). Moreover, since in finite dimension the range of A^* is closed, we deduce that $\hat{b} \in \operatorname{ran}(A^*)$. By part *(a)* of Theorem 2.7 which tells us $Av = f$ and by the definition of critical point of $\mathcal{J}$ on the affine space $\{v \in \mathcal{E} : Av = f\}$ the proof is concluded. □

We may now conclude that we have theoretical guarantees that the two algorithms proposed in this chapter converge to critical points of the linearly constrained nonconvex and nonsmooth minimization problem. In the next chapter we introduce some of the application where these procedures can be applied supported by some numerical experiments supporting the validity of the algorithms. The list of applications is far from being complete and we just aim to show that completely different problems from the most diverse disciplines can have common properties and can be addressed using the same ansatz opportunely tuned.

3. Some applications of the algorithm

Despite we are interested in applying Algorithms 2.2 and 2.4 in the specific field of fracture mechanics, we should not forget that they may be useful in many different fields. Therefore, we additionally tested the first of the two algorithms in different problems other than cohesive and brittle fractures. In this chapter we collect some of these tests that also appeared in [18, 19].
Let us first recall the general setting of the problem that will be then detailed for each test case. We consider two finite dimensional Euclidean spaces $\mathcal{E} \simeq \mathbb{R}^n$ and $\mathcal{F} \simeq \mathbb{R}^m$, with $n, m \in \mathbb{R}$, $m \leq n$, and a surjective linear constraint map $A : \mathcal{E} \to \mathcal{F}$. In addition let $\mathcal{K}$ a third finite dimensional Euclidean space and $T : \mathcal{E} \to \mathcal{K}$ a linear operator. For fixed $g \in \mathcal{K}$ and $f \in \mathcal{F}$, we consider functionals of the type

$$\mathcal{J}(v) = \|Tv - g\|^2 + \gamma \sum_{k=1}^{n} U_k(v_k), \tag{3.1}$$

of which we seek the critical points, subject to a linear constraint $Av = f$, where $\gamma > 0$ is a positive regularization parameter. Here $(v_k)_{k=1}^n$ are the components of the vector v with respect to a fixed basis in the space $\mathcal{E}$, and $U_k : \mathbb{R} \to \mathbb{R}_+$, for $k = 1, \ldots, n$, are scalar nonconvex maps.
Specifying each time the form of the operator, we introduce in the next section some of the problems, in which the algorithm can be efficaciously applied, divided in two main categories: free-discontinuity problems (Section 3.1) and truncated polynomial minimization (Section 3.2). In Section 3.3 we illustrate the different numerical issue we need to face in order to correctly and efficiently apply Algorithm 2.2. Finally in Section 3.4 we present the numerical experiments we implemented and comment the result obtained with the first of the two algorithms illustrated in the previous chapter.

3.1. Free-discontinuity problems

The terminology "*free-discontinuity problem*" was introduced by E. De Giorgi [84] to indicate a class of variational problems which consists in the minimization of a functional, involving both volume and surface energies, depending on a closed set $K \subset \mathbb{R}^d$, and a function u on $\mathbb{R}^d$ usually smooth outside of K. In particular,

- K is not fixed a priori and is an unknown of the problem;

- K is not a boundary in general, but a free-surface inside the domain of the problem.

3.1.1. The Mumford–Shah functional in image processing

The best-known example of a free-discontinuity problem is the one modeled by the Mumford–Shah functional [169], which in the general case is defined by

$$\mathcal{J}(u,K) := \int_{\Omega\setminus K} \left[|\nabla u(x)|^2 + \alpha(u(x)-g(x))^2\right] dx + \beta\mathcal{H}^{d-1}(K\cap\Omega).$$

The set Ω is a bounded open subset of $\mathbb{R}^d$, $\alpha,\beta>0$ are fixed constants, and $g\in L^\infty(\Omega)$. Inspired by image processing applications the dimension of the underlying Euclidean space $\mathbb{R}^d$ shall be $d=2$, although in principle the analysis can be conducted in any dimension. In fact, in the context of visual analysis, g is a given noisy image that we want to approximate by the minimizing function $u\in W^{1,2}(\Omega\setminus K)$; the set K is simultaneously used in order to *segment* the image into connected components. For a more specific overview on free-discontinuity problems, their analysis, and applications, we refer to [7].

Historically, one can think at the Mumford–Shah functional as the continuous version of a previous discrete formulation of the image segmentation problem proposed by D. Geman and S. Geman in [116]; see also the work of A. Blake and A. Zisserman in [31]. Let us recall this discrete approach.

Let $d=2$, $\Omega=[0,1]^2$, and let $u_{i,j}=u(h_i,h_j)$ be a discrete function defined on $\Omega_h := \Omega\cap h\mathbb{Z}^2$, for $h>0$. Define $W_r^2(t) := \min\{t^2,r^2\}$, $r>0$, to be the *truncated quadratic potential*, and

$$\mathcal{J}_h(u) := h^2 \sum_{(h_i,h_j)\in\Omega_h} \left[W^2_{\sqrt{\frac{\beta}{h}}}\left(\frac{u_{i+1,j}-u_{i,j}}{h}\right) + W^2_{\sqrt{\frac{\beta}{h}}}\left(\frac{u_{i,j+1}-u_{i,j}}{h}\right) + \alpha(u_{i,j}-g_{i,j})^2 \right]. \tag{3.2}$$

We shall now reformulate the minimization of this finite dimensional discrete problem into a linearly constrained minimization of a nonconvex functional of the discrete derivatives. For this purpose, we consider the derivative matrix $D_h:\mathbb{R}^{n^2}\to\mathbb{R}^{2n(n-1)}$ that maps the vector $(u_{j+(i-1)n}) := (u_{i,j})$ to the vector composed of the finite differences in the horizontal and vertical directions u_x and u_y respectively, given by

$$D_h u := \begin{bmatrix} u_x \\ u_y \end{bmatrix}, \quad \begin{cases} (u_x)_{j+n(i-1)} := (u_x)_{i,j} := \dfrac{u_{i+1,j}-u_{i,j}}{h}, & \substack{i=1,\dots,n-1 \\ j=1,\dots,n} \\[2ex] (u_y)_{j+(n-1)(i-1)} := (u_y)_{i,j} := \dfrac{u_{i,j+1}-u_{i,j}}{h}, & \substack{i=1,\dots,n \\ j=1,\dots,n-1} \end{cases}.$$

Note that its range $\operatorname{ran}(D_h)\subset\mathbb{R}^{2n(n-1)}$ is a (n^2-1)-dimensional subspace because $D_h c=0$ for constant vectors $c\in\mathbb{R}^{n^2}$. It is not difficult to show the representation of

any vector $u \in \mathbb{R}^{n^2}$ in terms of the following differentiation-integration formula, given by

$$u = D_h^\dagger D_h u + c,$$

where $D_h^\dagger$ is the pseudo-inverse matrix of D_h (in the Moore–Penrose sense); observing that $D_h^\dagger$ maps $\mathrm{ran}(D_h)$ injectively into $\mathbb{R}^{n^2}$. Also, c is a constant vector that depends on u, and the values of its entries coincide with the mean value $h^2 \sum_{(h_i,h_j)\in\Omega_h} u_{i,j}$ of u. Therefore, any vector u is uniquely identified by the pair $(D_h u, c)$.
Since constant vectors comprise the null space of D_h, the orthogonality relation

$$\langle D_h^\dagger D_h u, c\rangle = 0 \tag{3.3}$$

holds for any vector u and any constant vector c. Here the scalar product $\langle u, u'\rangle = \sum_{(h_i,h_j)\in\Omega_h} u_{i,j}u'_{i,j}$ is the standard Euclidean scalar product on $\mathbb{R}^{n^2}$, which induces the Euclidean norm $\|\cdot\|$.
Using the orthogonality property (3.3), denoting the mean value of g by c_g, we have that

$$\|u-g\|^2 = \|D_h^\dagger D_h u - D_h^\dagger D_h g + (c-c_g)\|^2 = \|D_h^\dagger D_h u - D_h^\dagger D_h g\|^2 + \|c-c_g\|^2 \tag{3.4}$$

Hence, with a slight abuse of notation, we can reformulate the original discrete functional (3.2) in terms of derivatives, and mean values, by

$$\mathcal{J}_h(v,c) = h^2\left[\alpha\|D_h^\dagger v - \tilde{g}\|^2 + \alpha\|c-c_g\|^2 + \sum_{i,j}\min\left\{|v_{i,j}|^2, \frac{\beta}{h}\right\}\right].$$

where $v = D_h u \in \mathbb{R}^{2n(n-1)}$, and $\tilde{g} = D_h^\dagger D_h g \in \mathbb{R}^{n^2}$. We assume that $c = c_g$ at any minimizer u, since the corresponding term in $\mathcal{J}_h$ does not depend on v. Additionally, we must minimize $\mathcal{J}_h(v,c)$ subject to the constraint $(D_h D_h^\dagger - I)v = 0$ in order to obtain vectors in $\mathbb{R}^{2n(n-1)}$ corresponding to derivatives of elements in $\mathbb{R}^{n^2}$. This $2n(n-1)$ linearly independent constraints actually correspond to a discrete curl-free condition on the vector v.
To summarize, we need to solve the following constrained optimization problem:

$$\text{Minimize} \quad \mathcal{J}_h(v) = h^2\big[\alpha\|Tv-\tilde{g}\|^2 + \sum_{i,j} W^2_{\sqrt{\frac{\beta}{h}}}(v_{i,j})\big] \qquad \text{subject to} \quad Av = 0, \tag{3.5}$$

for $T = D_h^\dagger$ and $A = I - D_h D_h^\dagger$. Actually the explicit use of the pseudo-inverse matrix $D^\dagger$ is only needed for determining the operator T, while the linear constraint $Av = 0$ is equivalent to a discrete curl-free condition on the vectors v, see [108] for details. Therefore it can simply be expressed in terms of a sparse linear system corresponding to

the discretization of the curl operator. Once the minimal derivative vector v is computed, we can assemble the minimal u by incorporating the mean value c_g of g as follows:

$$u = D_h^\dagger v + c_g.$$

We stress that when v is curl-free, a primitive u can be easily recovered, up to an arbitrary constant, by performing a line integration, so again this process does not require the explicit form of $D_h^\dagger$, see the details of (3.9) and (3.10) in next paragraph. Notice that the objective functional in the optimization problem (3.5) is precisely of the form (3.1), with the maps $U_k = h^2 W^2_{\sqrt{\frac{\beta}{h}}}$ for all $k = (i, j)$. As we shall discuss in Section 3.3, the functional $\mathcal{J}_h$ in (3.5) is in general neither coercive nor ω-semi-convex as required by the conditions of applicability of Algorithm 2.2. Nevertheless we show in Section 3.3.1 that manipulating the functional adding a mild regularization allows us to treat efficiently this family of problems. We provide in Remark 3.20 a possible guideline on the efficient implementation of the pseudoinverse matrix $D_h^\dagger$ and its adjoint $(D_h^\dagger)^*$, as it appears in the iterations of the inner loop of Algorithm 2.2 when applied to the minimization of the Mumford–Shah functional.

3.1.2. Quasi-static evolution of brittle fractures

Among the functionals that can be minimized through Algorithm 2.2 we find also energy functionals described by Griffith's Theorem 1.8. Thus, beside static models such as the Mumford–Shah functional minimization for image deblurring and denoising, quasi-static evolutions of brittle fractures modeled by (1.21)

$$(u^*(t), \Gamma^*(t)) \in \operatorname*{arg\,min}_{\substack{u \in \mathfrak{U}_{g(t)} \\ \Gamma \supseteq \bigcup_{s<t} = \Gamma(s)}} E(u, \Gamma) : \frac{1}{2} \int_{\Omega \setminus \Gamma} |\nabla u|^2 + \kappa \mathcal{H}^{d-1}(\Gamma),$$

can be addressed.
Following a similar discretization in space for $d = 2$ as done before, let $u_{i,j} = u(h_i, h_j)$ be a discrete function defined on $\Omega_h := \Omega \cap h\mathbb{Z}^2$, for $h > 0$. Define

$$E_h(u) := h^2 \sum_{(h_i,h_j) \in \Omega_h \setminus \Gamma_h} \left[W^2_{\sqrt{\frac{\beta}{h}}} \left(\frac{u_{i+1,j} - u_{i,j}}{h} \right) + W^2_{\sqrt{\frac{\beta}{h}}} \left(\frac{u_{i,j+1} - u_{i,j}}{h} \right) \right], \tag{3.6}$$

where Γ_h is the current fracture. Up to considering appropriate domain decompositions and without loss of generality we can assume $\Gamma_h = \emptyset$ and Ω_h be the reference domain for the optimization. Defining

$$v_{i,j} = \Big(\underbrace{\frac{u_{i+1,j} - u_{i,j}}{h}}_{:=(v_{i,j})_1}, \underbrace{\frac{u_{i,j+1} - u_{i,j}}{h}}_{:=(v_{i,j})_2} \Big),$$

we have that $v = (v_{i,j})_{(hi,hj)\in\Omega_h}$ fulfills again a curl-free condition as well as several linear constraints given by the discretization of the compatibility condition $u \in \mathfrak{U}_{g(t)}$. In particular, if we assume that g is locally constant with a jump at a given interface $\Gamma_D \subset \Omega$, we may have, as in the example shown in Figure 5.2, that

$$\begin{cases} v_{i,j} = 0, & \text{for all } (hi,hj) \in \mathrm{int}(\Omega_D)_h \setminus (\Gamma_D)_h \cap (\Omega_D)_h, \\ (v_{i,j})_2 = \dfrac{g_{i,j+1} - g_{i,j}}{h}, & \text{for all } (hi,hj) \in (\Gamma_D)_h \cap (\Omega_D)_h, \\ (v_{i,j})_2 = 0, & \text{for all } (hi,hj) \in (\partial(\Omega_D)_h \cap \Omega_h) \setminus (\Gamma_D)_h. \end{cases} \tag{3.7}$$

Then, we can rewrite the functional E_h, with a slight abuse of notation, as

$$E_h(v) := h^2 \sum_{i,j} W^2_{\sqrt{\beta/h}}(v_{i,j}) \tag{3.8}$$

to be minimized with respect to v under the constraint (3.7), which we compactly rewrite as $Av = f$. Let us stress now that g is actually a boundary datum, and provided the derivative field v of u we can recover u simply by line integration.
For a suitable coordinate system on the discrete domain, we define according to the line integration operator

$$u_{i,j} = \sum_{m=0}^{j} v_{i,m} + g_{0,m}, \quad \text{for all } (h_i, h_j) \in \Omega_h, \tag{3.9}$$

which can be expressed in the compact form by

$$u = Bv + g. \tag{3.10}$$

Notice that, while for the imaging problem we needed to consider the pseudo-inverse matrix $D^\dagger$ of the discrete differentiation operator D within the fidelity term $\|Tv - \tilde{g}\|$, here the $D^\dagger$ is of no use. We also mention that the minimization of (3.8) is again of the general type (3.1) for the choice of the maps $U_k = h^2 W^2_{\sqrt{\beta/h}}$ for all $k = (i,j)$. The issue of the coercivity of $\mathcal{J}_h$ on the affine space $\{Av = f\}$ is addressed in the work [13] on discrete rate-independent evolutions.
We shall show in Section 3.4 an application of our Algorithm 2.2 where we actually perform a simulation of a fracture in a one dimensional model, evolving through critical points, being a numerically robust and physically sound description of the happening of the fracture.

3.2. Truncated polynomial minimization

We should mention that, independently of the choice of the linear operators T and A, by [108, Theorem 2.3], the constrained minimization problem

$$\text{Minimize } \mathcal{J}_p(v) := \|Tv - g\|^2 + \gamma \sum_{i=1}^{n} W_r^p(v_i), \qquad \text{Subject to } Av = f, \tag{3.11}$$

where $W_r^p(v) := \min\{v^p, r^p\}$, has always global minimizers. Notice that the proof of existence of minimizers is far from being trivial (see Remark 3.1 below), since the problem is in general not coercive. Concerning uniqueness and stability of minimizers, we refer instead to the work of S. Durand and D. Nikolova [97, 98], about cases where T is injective on $\ker A$.

Remark 3.1. *The proof of existence of solutions of* (3.11) *is based on a special orthogonal decomposition of certain convex sets, see [108, Appendix, Section 8.1].*
Define $\bar{\mathcal{J}}_p(v) = \|Tv - g\|^2 + \gamma \sum_{i=1}^{n} c_i |v_i|^p$ *for* $c_1, \ldots, c_m$ *scalars; notice that we allow some of them to be negative or zero, as soon as* $\bar{\mathcal{J}}_p(v) \geq C_{\inf} > -\infty$ *for all* $v \in \mathcal{E}$*. Then for any constant* $C > 0$ *and any polyhedral convex set* $X \subset \mathcal{E}$*, there exists a linear subspace* $\mathcal{V} = \mathcal{V}_{X,C} \subset \mathcal{E}$*, such that the orthogonal projection* $X^\perp$ *of* X *onto* $\mathcal{V}^\perp$ *has the properties*

- $X = \{x = x^\perp \oplus tv : x^\perp \in X^\perp, v \in \mathcal{V}, t \in \mathbb{R}^+\}$,
- $M_C = X^\perp \cap \{v \in \mathcal{E} : \bar{\mathcal{J}}_p(v) \leq C\}$ *is compact, and*
- $\bar{\mathcal{J}}_p(\xi_t)$ *is constant along rays* $\xi_t = x^\perp \oplus tv$*, where* $x^\perp \in M_C$*,* $v \in \mathcal{V}$*, and* $t \in \mathbb{R}^+$*.*

For $\mathcal{I}_0 \subset \mathcal{I}$ *and* $\mathcal{U}_{\mathcal{I}_0} := \{v \in \mathcal{E} : |v_i| \leq r, i \in \mathcal{I}_0 \text{ and } |v_i| > r, i \in \mathcal{I} \setminus \mathcal{I}_0\}$*, in particular this result applies on* $X = \mathcal{F}(f) \cap \overline{\mathcal{U}_{\mathcal{I}_0}}$*, hence*

$$\text{Minimize } \bar{\mathcal{J}}_p(v) = \|Tv - g\|^2 + \gamma \sum_{i=1}^{n} c_i |t|^p, \text{ Subject to } Av = f \text{ and } v \in \mathcal{U}_{\mathcal{I}_0}, \tag{3.12}$$

has solutions in $\mathcal{E}$*, actually in the compact set* $M_{\bar{\mathcal{J}}_p(v^0)} = X^\perp \cap \{v \in \mathcal{E} : \bar{\mathcal{J}}_p(v) \leq \bar{\mathcal{J}}_p(v^0)\}$*, for any* $v^0 \in \mathcal{E}$*.*

3.2.1. Compressive sensing

Among the class of truncated polynomial minimization problem, we can find the decoding for compressed signals introduced in [18], in particular if the signal is corrupted by the noise. The main authorship of the cited work is S. Peter. We include in the following sections exclusively the parts of the paper where the author of this thesis played the role of main contributor.
Compressive sensing focuses on the robust recovery of nearly sparse vectors from the minimal amount of measurements obtained by a randomized linear process. So far, a vast literature appeared considering problems where deterministic or random noise is added after the measurement process, while it is not strictly related to the signal. One typically considers model problems of the type

$$y = Ax + w \tag{3.13}$$

where $x \in \mathbb{R}^N$ is a nearly sparse vector, $A \in \mathbb{R}^{m \times N}$ is the linear measurement matrix, $y \in \mathbb{R}^m$ is the result of the measurement, and w is a white noise vector affecting the measurements. However, in practice it is very uncommon to have a signal detected by a certain device, totally free from some external noise. Therefore, it is reasonable to consider the more realistic model

$$y = A(\bar{x} + n) + w,$$

instead of (3.13) where $\bar{x} \in \mathbb{R}^N$ is the noiseless signal and $n \in \mathbb{R}^N$ is the noise on the original signal.

The recent work [12, 210] shows how the measurement process actually causes the *noise-folding phenomenon*, which implies that the variance of the noise on the original signal is amplified by a factor of N/m, additionally contributing to the measurement noise, playing to our disadvantage in the recovery phase. More formally, if we add to the signal $\bar{x}$ a noise vector n composed by random entries with normal distribution $\mathcal{N}(0, \sigma_n)$, the measurement y given by

$$y = A(\bar{x} + n), \tag{3.14}$$

can be considered equivalently obtained by a measurement procedure of the form (3.13) where now the vector w is composed by i.i.d. Gaussian entries with normal distribution $\mathcal{N}(0, \frac{N}{m}\sigma_n^2)$. There are actually many different real life situations where the noise-folding phenomenon occurs and we report two examples in the following. In [12] it is described how the use of compressed sensing in the design of sub-Nyquist A/D converters may be affected by noise-folding. A second situation where noise-folding affects a real life application can be found in [3]. In this paper, the authors illustrate how the multitude of wireless communication standards such as cellular, digital radio and television broadcasting, GPS, WIFI, and Bluetooth may be affected by noise-folding if the decoding of the signal is not correctly performed, for example using a sub-sampling technique.

An approach to control the noise-folding is proposed in [12]. In this case, one may tune the linear measurement process in order to a priori filter the noise. However, this strategy requires to have a precise knowledge of the noise statistics and to design proper filters. Other related works [128, 129, 130] address the problem of designing adaptive measurements, called *distilled sensing*, in order to detect and locate the signal within white noise.

In [18] we perform a blind-to-statistic analysis of the signal reconstruction problem. Let us define for $r > \eta > 0$, $1 \leq k < m$, and $1 \leq p \leq 2$, the class of *sparse vectors affected by bounded noise*,

$$\mathcal{S}^p_{\eta,k,r} := \left\{ x \in \mathbb{R}^N | \#S_r(x) \leq k \text{ and } \sum_{i \in (S_r(x))^c} |x_i|^p \leq \eta^p \right\}, \tag{3.15}$$

where $S_r(x) := \{i \in \{1,\dots,N\} |\, |x_i| > r\}$ is the index support of the large entries exceeding in absolute value the threshold r. This class contains all vectors with at most $1 \leq k < m$ large entries exceeding the threshold r in absolute value, while the p-norm of the remaining entries stays below a certain noise level. Notice that vectors $x \in \mathcal{S}^p_{\eta,k,r}$ can be naturally decomposed in the noiseless (relevant) part $\bar{x} = x|_{S_r(x)}$ and the noise $n = x|_{S_r^c(x)}$. We additionally refer to Appendix A.2 for the definition of *Null-Space-Property* required to the linear operator A in the following

Theorem 3.2. *Let $A \in \mathbb{R}^{m\times N}$ have the $(2k,\gamma_{2k})$-Null-Space-Property, for $\gamma_{2k} < 1$, $1 \leq p \leq 2$, and $x, x' \in \mathcal{S}^p_{\eta,k,r}$ such that $Ax = Ax'$, and $0 \leq \eta < r$. Then*

$$\#(S_r(x)\Delta S_r(x')) \leq \frac{(2\gamma_{2k}\kappa_p\eta)^p}{(r-\eta)^p}, \tag{3.16}$$

where

$$\kappa_p := \kappa_p(N,k) := \begin{cases} 1, & p = 1, \\ \sqrt[q]{N-k}, & 1 < p \leq 2. \end{cases} \tag{3.17}$$

(Here we denote by "Δ" the set symmetric difference) If additionally

$$r > \eta(1 + 2\gamma_{2k}\kappa_p) \tag{3.18}$$

then $S_r(x) = S_r(x')$.

Proof. As $Ax = Ax'$, then $(x - x') \in \ker(A)$. By the $(2k,\gamma_{2k})$-NSP, Hölder's inequality, and the triangle inequality we have

$$\begin{aligned} \|(x-x')|_{S_r(x)\cup S_r(x')}\|_{\ell_p} &\leq \|(x-x')|_{S_r(x)\cup S_r(x')}\|_{\ell_1} \leq \gamma_{2k}\|(x-x')|_{(S_r(x)\cup S_r(x'))^c}\|_{\ell_1} \\ &\leq \gamma_{2k}\kappa_p\|(x-x')|_{(S_r(x)\cup S_r(x'))^c}\|_{\ell_p} \leq 2\gamma_{2k}\kappa_p\eta. \end{aligned} \tag{3.19}$$

Now we estimate the symmetric difference of the supports of the large entries of x and x' in absolute value as follows: if $i \in S_r(x)\Delta S_r(x')$, then either $|x_i| > r$ and $|x_i'| \leq \eta$ or $|x_i| \leq \eta$ and $|x_i'| > r$. This implies that $|x_i' - x_i| > (r-\eta)$. Thus we have $\|(x-x')|_{S_r(x)\Delta S_r(x')}\|^p_{ell_p} \geq (\#(S_r(x)\Delta S_r(x')))\,(r-\eta)^p$. Together with the non-negativity of $\|(x-x')|_{S_r(x)\cap S_r(x')}\|_{\ell_p}$, we obtain the chain of inequalities

$$\begin{aligned} (2\gamma_{2k}\kappa_p\eta)^p &\geq \|(x-x')|_{S_r(x)\cup S_r(x')}\|^p_{\ell_p} \\ &\geq \|(x-x')|_{S_r(x)\cap S_r(x')}\|^p_{\ell_p} + \|(x-x')|_{S_r(x)\Delta S_r(x')}\|^p_{\ell_p} \\ &\geq (\#(S_r(x)\Delta S_r(x')))\,(r-\eta)^p, \end{aligned}$$

and therefore we obtain (3.16).
Notice now that (3.16) and (3.18) imply $\mathbb{N} \ni \#(S_r(x)\Delta S_r(x')) < 1$ and $S_r(x)\Delta S_r(x') = \emptyset$. □

Unfortunately, none of the well known and commonly advocated decoder in this field can guarantee the decoded signal to be in the class $\mathcal{S}^p_{\eta,k,r}$. Therefore, to overcome the shortcomings of methods based exclusively on ℓ_1-minimizations in damping the noise-folding, and having a stable support recovery, we design a new decoding procedure with output in $\mathcal{S}^p_{\eta,k,r}$, which consequently allows us to have both these very desirable properties. Let us first introduce the following functional.

Definition 3.3 (Regularized selective p-potential)**.** *We define the* regularized truncated p-power function $W_r^{p,\epsilon}\colon \mathbb{R} \to \mathbb{R}_0^+$ *by*

$$W_r^{p,\epsilon}(t) = \begin{cases} t^p & 0 \le t < r-\epsilon, \\ \pi_p(t) & r-\epsilon \le t \le r+\epsilon, \\ r^p & t > r+\epsilon, \end{cases} \qquad t \ge 0, \tag{3.20}$$

where $0 < \varepsilon < r$*, and* $\pi_p(t)$ *is the third degree interpolating polynomial defined in Lemma 3.4 below. Moreover, we set* $W_r^{p,\epsilon}(t) = W_r^{p,\epsilon}(-t)$ *for* $t < 0$*. We call the functional* $\mathcal{SP}_r^{p,\epsilon}\colon \mathbb{R}^N \to \mathbb{R}_0^+$*,*

$$\mathcal{SP}_r^{p,\epsilon}(x) = \sum_{j=1}^{N} W_r^{p,\epsilon}(x_j), \quad r > 0, \quad 1 \le p \le 2, \tag{3.21}$$

the regularized selective p-potential functional*.*

Lemma 3.4. *Let* $0 < s_1 < s_2$ *and assume that*

$$\pi(t) := A(t-s_2)^3 + B(t-s_2)^2 + C,$$

is a third degree polynomial. Given $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ *and by setting*

$$\begin{cases} C = \gamma_3, \\ B = \dfrac{\gamma_1}{s_2 - s_1} - \dfrac{3(\gamma_3 - \gamma_2)}{(s_2 - s_1)^2}, \\ A = \dfrac{\gamma_1}{3(s_2 - s_1)^2} + \dfrac{2B}{3(s_2 - s_1)}, \end{cases} \tag{3.22}$$

then we have the following interpolation properties

$$\begin{cases} \pi(s_2) = \gamma_3, & \pi(s_1) = \gamma_2, \\ \pi'(s_2) = 0, & \pi'(s_1) = \gamma_1. \end{cases} \tag{3.23}$$

Proof. The equalities related to s_2 are straightforward, the others related to s_1 follow by simple direct computations:

$$\begin{aligned} \pi(s_1) &= -\frac{\gamma_1}{3}(s_2 - s_1) - \frac{2}{3}B(s_2 - s_1)^2 + B(s_2 - s_1)^2 + \gamma_3 \\ &= -\frac{\gamma_1}{3}(s_2 - s_1) + \frac{B}{3}(s_2 - s_1)^2 + \gamma_3 \\ &= -\frac{\gamma_1}{3}(s_2 - s_1) + \frac{\gamma_1}{3}(s_2 - s_1) - (\gamma_3 - \gamma_2) + \gamma_3 = \gamma_2, \end{aligned}$$

and

$$\pi'(s_1) = 3A(s_1 - s_2)^2 + 2B(s_1 - s_2) = \gamma_1 - 2B(s_1 - s_2) + 2B(s_1 - s_2) = \gamma_1.$$

□

Given $0 < \varepsilon < r$ for every $t \in [r - \varepsilon, r + \varepsilon]$ we define $\pi_p(t) = \pi(t)$ as in Lemma 3.4 for $s_1 = (r - \varepsilon)$, $s_2 = (r + \varepsilon)$, $\gamma_1 = p(r - \varepsilon)^{p-1}$, $\gamma_2 = (r - \varepsilon)^p$, and $\gamma_3 = r^p$. For example, for $p = 2$, we have

$$\pi_2(t) = \frac{[t + (r - \varepsilon)][\varepsilon(r + t) - (r - t)^2]}{4\varepsilon}, \quad t \in \mathbb{R}.$$

In the following, we state the main result of [18], which gives the guarantee that the result of the decoding process via the regularized selective p-potential is effectively an element of the class $\mathcal{S}^p_{\eta,k,r}$ implying that Theorem 3.2 holds.

Theorem 3.5. *Let $A \in \mathbb{R}^{m\times N}$ have the $(2k, \gamma_{2k})$-NSP, with $\gamma_{2k} < 1$, and $1 \le p \le 2$. Furthermore, we assume $x \in \mathcal{S}^p_{\eta,k,r+\epsilon}$, for $\epsilon > 0$, $0 < \eta < r + \epsilon$, with the property of having the minimal $\#S_{r+\epsilon}(x)$ within $\mathcal{F}(y)$, where $y = Ax$ is its associated measurement vector, i.e.,*

$$\#S_{r+\epsilon}(x) \le \#S_{r+\epsilon}(z) \text{ for all } z \in \mathcal{F}(y). \tag{3.24}$$

If x^* is such that

$$\mathcal{SP}^{p,\epsilon}_r(x^*) \le \mathcal{SP}^{p,\epsilon}(x), \tag{3.25}$$

and

$$|x^*_i| < r - \epsilon, \tag{3.26}$$

for all $i \in (S_{r+\epsilon}(x^*))^c$, then also $x^* \in \mathcal{S}^p_{\eta,k,r+\epsilon}$, implying noise-folding damping. Moreover, we have the support stability property

$$\#(S_{r+\epsilon}(x)\Delta S_{r+\epsilon}(x^*)) \le \frac{(2\gamma_{2k}\kappa_p\eta)^p}{(r + \epsilon - \eta)^p}. \tag{3.27}$$

Proof. Notice that we can equally rewrite the $\mathcal{SP}^{p,\epsilon}_r$ functional as

$$\mathcal{SP}^{p,\epsilon}_r(z) = r^p \#S_{r+\epsilon}(z) + \sum_{i\in(S_{r+\epsilon}(z))^c} |z_i|^p_\epsilon,$$

where $|t|^p_\epsilon := W^{p,\epsilon}_r(t)$ for $|t| \le r + \epsilon$. Here, by construction, we have $|t|^p_\epsilon \le |t|^p$. By the assumptions (3.25) and $x \in \mathcal{S}^p_{\eta,k,r+\epsilon}$, we obtain the estimates

$$\begin{aligned} r^p \#S_{r+\epsilon}(x^*) \le \mathcal{SP}^{p,\epsilon}_r(x^*) \le \mathcal{SP}^{p,\epsilon}_r(x) &= r^p \#S_{r+\epsilon}(x) + \sum_{i\in(S_{r+\epsilon}(x))^c} |x_i|^p_\epsilon \\ &\le r^p \#S_{r+\epsilon}(x) + \sum_{i\in(S_{r+\epsilon}(x))^c} |x_i|^p \le r^p \#S_{r+\epsilon}(x) + \eta^p, \end{aligned}$$

and thus $\#S_{r+\epsilon}(x^*) \leq \left(\frac{\eta}{r}\right)^p + \#S_{r+\epsilon}(x)$. As $\frac{\eta}{r} < 1$ by assumption, the minimality property (3.24) yields immediately

$$\#S_{r+\epsilon}(x^*) = \#S_{r+\epsilon}(x) \leq k. \tag{3.28}$$

Assumption (3.26) and again (3.25) yield

$$\begin{aligned} r^p \#S_{r+\epsilon}(x^*) + \sum_{i \in (S_{r+\epsilon}(x^*))^c} |x_i^*|^p &= r^p \#S_{r+\epsilon}(x^*) + \sum_{i \in (S_{r+\epsilon}(x^*))^c} |x_i^*|_\epsilon^p \\ \leq r^p \#S_{r+\epsilon}(x) + \sum_{i \in (S_{r+\epsilon}(x))^c} |x_i|_\epsilon^p &\leq r^p \#S_{r+\epsilon}(x) + \sum_{i \in (S_{r+\epsilon}(x))^c} |x_i|^p. \end{aligned}$$

By this latter inequality and (3.28) we obtain

$$\sum_{i \in (S_{r+\epsilon}(x^*))^c} |x_i^*|^p \leq \sum_{i \in (S_{r+\epsilon}(x))^c} |x_i|^p \leq \eta^p,$$

which implies $x^* \in \mathcal{S}^p_{\eta,k,r+\epsilon}$. We conclude (3.27) by an application of Theorem 3.2. □

Remark 3.6. *The best candidate x^* to fulfill condition* (3.25) *would be actually*

$$x^* := \operatorname*{arg\,min}_{z \in \mathcal{F}(y)} \mathcal{SP}_r^{p,\epsilon}(z) \tag{3.29}$$

because this will make (3.25) *automatically true, whichever x is. However* (3.29) *is a highly nonconvex problem whose solution is in general NP-hard [2]. The way we circumvent this drawback is computing x^* by means of a local minimization of $\mathcal{SP}_r^{p,\epsilon}$ in $\mathcal{F}(y)$ around a given vector x_0 which is considered as starting point for Algorithm 2.2. Ideally, the best choice for x_0 would be x itself, so that* (3.25) *may be fulfilled. As we do not dispose of the original vector x yet, a heuristic rule, which we will show to be very robust in our numerical simulations, is to choose the solution of the decoding process via the ℓ_1-minimization as x_0.*

3.3. Numerical issues: smoothing and iterative thresholding procedures

Due to their nonsmoothness and nonconvexity, for some problems, as for the Mumford–Shah functional $\mathcal{J}_p$ applied to image denoising, the linearly constrained minimization (3.11) was so far an open problem. Indeed, standard methods, such as SQP and Newton methods, do not apply, unless one provides a $\mathcal{C}^2$-regularization of the problem. In particular, it would be desirable that an appropriate algorithm performing such an optimization could retain both the simplicity of the thresholding iteration and its unconditional convergence properties, as given by [108, Theorem 4.8] in the case of unconstrained

minimization of the functional $\mathcal{J}_p$. Certainly the method of Algorithm 2.2 is a strong candidate, as the iterations of its inner loop actually requires only a unconstrained minimization, which can be again addressed by iterative thresholding, see Section 3.3.3 below. However, we encounter two major bottlenecks to the direct application of this algorithm to (3.11). The first problem is that $\mathcal{J}_p$ does not satisfy our main assumption (A1), i.e., it is not ω-semi-convex, as it is not a $\mathcal{C}^1$-perturbation of a convex functional. In fact the term W_r^p is too rough at the kink where the truncation applies. The second issue comes by the lack of coerciveness of $\mathcal{J}_p$ on the affine space $\mathcal{A}(f)$ in general, for a generic choice of T. A general convergence result (Theorem 3.14) will be therefore available only under an additional condition on T.

3.3.1. A smoothing procedure

In order to overcome the first of the two bottleneck described above, our analysis requires a smooth perturbation technique which is reminiscent of previous methods of continuation-based deterministic relaxation, such as the *graduated nonconvexity* (GNC) pioneered by A. Blake and A. Zisserman [31] in the context of the Mumford–Shah model, see also recent developments in [179, 177, 181, 180] and references therein. We briefly mention how this latter technique works. For a suitable parameter $\varepsilon \in [0,1]$, one considers a continuous family of smoother objectives $\mathcal{J}^\varepsilon$ such that $\lim_{\varepsilon\to 1} \mathcal{J}^\varepsilon = \mathcal{J}$ (at least pointwise), where $\mathcal{J}$ is the nonconvex energy to be minimized. Then one addresses the global minimization of $\mathcal{J}$ by iterated *local* minimizations along $\mathcal{J}^\varepsilon$ when ε is increasing from 0 to 1 with a strictly convex initial $\mathcal{J}^0$. More formally, we consider an increasing sequence $(\varepsilon_n)_{n\in\mathbb{N}}$, with $\varepsilon_0 = 0$ and $\lim_n \varepsilon_n = 1$ and the iterative algorithm

$$v^{n+1} = \arg \min_{v\in\mathcal{N}_{\varepsilon_n}(v^n)} \mathcal{J}^{\varepsilon_n}(v), \tag{3.30}$$

where $\mathcal{N}_{\varepsilon_n}(v^n)$ is a suitable neighborhood of the previous iteration v^n of size possibly depending on ε_n. While such semi-heuristic algorithms perform very well in practice, they usually do not provide any guarantee for global convergence and their applicability highly depends on the appropriate design of the approximating family $\{\mathcal{J}^\varepsilon : \varepsilon \in [0,1]\}$, depending on the particular application and form of $\mathcal{J}$. Our algorithm has instead *more general applicability* and *stronger convergence guarantees*, providing as a byproduct also some rigorous justification to those semi-heuristic methods.

In particular, in this section we show that the regularized truncated p-power introduced in Definition 3.3 can be considered as a smooth approximation of W_r^p and thus used to construct an appropriate perturbation $\mathcal{J}_p^\varepsilon$ of $\mathcal{J}_p$, which allows eventually for ω-semi-convexity, but does not modify essentially the minimizers of (3.11). Such modification does not affect the possibility of using thresholding functions in the numerical setting, although instead of the hard-type discontinuous thresholding encountered in the unconstrained case, as in [108, Proposition 4.3] and [108][Figure 2], our new thresholding

function will be a Lipschitz one, as an effect of the introduced regularization, in dependence of the choice of the parameters $\gamma, \varepsilon, r, \omega$ in appropriate ranges. We will see in Section 3.3.3 the usefulness of this feature in terms of guaranteed exponential convergence from the beginning of the iterations.
Recalling the notion of the regularized truncated p-power in Definition 3.3, we observe that the functional $W_r^{p,\varepsilon}$ is actually a $\mathcal{C}^1$-function of $\mathbb{R}$, for all $0 < \varepsilon < r$, and it can be considered as our desired smooth approximation of the function W_r^p for ε sufficiently small. The graphs of W_r^p and $W_r^{p,\epsilon}$ are shown in Figure 3.1 for $p = 2$, $r = 1$, and $\epsilon = 0.4$.

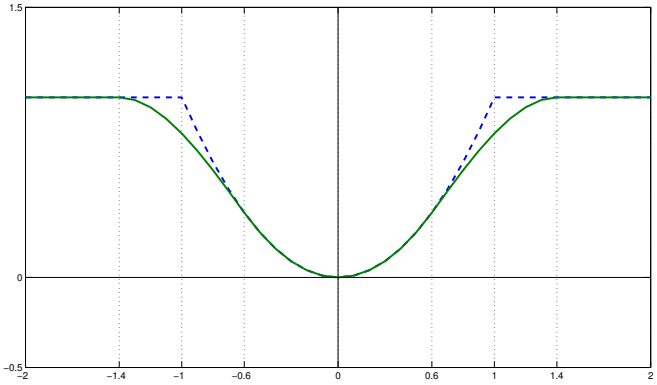


Figure 3.1.: Truncated quadratic potential W_r^p and its regularization $W_r^{p,\varepsilon}$, for $p = 2$, $r = 1$, and $\varepsilon = 0.4$.

Thus, we have

$$\mathcal{J}_p^\varepsilon(v) = \|Tv - g\|^2 + \gamma \sum_{i=1}^{n} W_r^{p,\varepsilon}(v_i). \tag{3.31}$$

About the existence of constrained minimizers of $\mathcal{J}_p^\varepsilon$ we have the following abstract result. We stress that the result holds for all inverse free-discontinuity problems, as it requires no further assumptions on T.

Theorem 3.7. *For $0 \leq \varepsilon < r$, the problem*

$$\textit{Minimize } \mathcal{J}_p^\varepsilon(v) = \|Tv - g\|^2 + \gamma \sum_{i=1}^{n} W_r^{p,\varepsilon}(v_i), \ \textit{Subject to } Av = f, \tag{3.32}$$

has solutions in $\mathcal{E}$. Actually, such minimal solutions can be taken in a compact set $M \subset \mathcal{E}$ independent of ε.

Proof. This proof uses a similar approach as for [108, Theorem 2.3]. Let us first consider a partition $\mathscr{P} = \{\mathcal{U}_{\mathcal{I}_0^j}\}_{j=1}^{2^n}$ of $\mathcal{E}$ indexed by all subsets $\mathcal{I}_0^j \subset \mathcal{I}$, as follows

$$\mathcal{U}_{\mathcal{I}_0^j} = \{v \in \mathcal{E} : |v_i| \leq r + \varepsilon, i \in \mathcal{I}_0^j, |v_i| > r + \varepsilon, i \in \mathcal{I} \setminus \mathcal{I}_0^j\}.$$

The minimization of $\mathcal{J}_p^\varepsilon$ over $\mathcal{A}(f) \cap \mathcal{U}_{\mathcal{I}_0^j}$ can be reformulated as

$$\begin{cases} \text{Minimize } \overline{\mathcal{J}_p^\varepsilon}(v) = \|Tv - g\|^2 + \gamma \sum_{i=1}^n c_i \overline{W_r^{p,\varepsilon}}(v_i), \text{ Subject to } v \in \mathcal{A}(f) \cap \mathcal{U}_{\mathcal{I}_0^j}, \\ c_i = 0 \text{ if } i \in \mathcal{I} \setminus \mathcal{I}_0^j \text{ and } c_i = 1 \text{ if } i \in \mathcal{I}_0^j, \end{cases} \tag{3.33}$$

where

$$\overline{W_r^{p,\varepsilon}}(t) = \begin{cases} t^p, & t \le r - \varepsilon, \\ \pi_p(t), & r - \varepsilon \le t \le r + \varepsilon, \\ |t - \varepsilon|^p & t \ge r + \varepsilon. \end{cases}$$

Showing that the minimization (3.33) has always a solution $v(\mathcal{I}_0^j)$ for all $j = 1, \dots, 2^n$, and such a minimizer belongs to a compact set M^j, independent of $\varepsilon \ge 0$, then

$$v^* = \arg\min_{j=1,\dots,2^n} \mathcal{J}_p^\varepsilon(v(\mathcal{I}_0^j)),$$

is actually a solution for (3.32) and it belongs to the compact set $M = \cup_{j=1}^{2^n} M^j$, independent of $\varepsilon \ge 0$. Hence, it is sufficient now to address (3.33). For that, we first show the following technical observation: if $x, v \in \mathcal{E}$ are fixed and $\overline{\mathcal{J}_p^\varepsilon}$ is bounded above and below on the ray $R_{x,v} = \{x + tv, t \ge 0\}$, then $\overline{\mathcal{J}_p^\varepsilon}$ is actually constant on $R_{x,v}$. In fact, let us consider the function

$$\mu(t) = \overline{\mathcal{J}_p^\varepsilon}(x + tv).$$

By the boundedness of $\overline{\mathcal{J}_p^\varepsilon}(x + tv)$, without loss of generality, we can assume that $0 \le \mu(t) \le 1$. Hence there exists a sequence $(t_k)_k \subset \mathbb{R}^+$ of points $t_k \to +\infty$ for $k \to \infty$ such that $\mu(t_k) \to \eta \in [0, 1]$ for $k \to \infty$. Moreover, by definition of $\overline{W_r^{p,\varepsilon}}$, for $t > 0$ sufficiently large we have actually the general expression $\mu(t) = P(t) + \gamma \sum_{i=1}^n c_i |x_i - \varepsilon + tv_i|^p$, where P is a polynomial of degree at most 2. Assume now, for instance, that $1 \le p \le 2$. As $\lim_{k\to\infty} \frac{\mu(t_k)}{t_k^2} = 0$ we deduce that all the coefficients in P of second degree are actually vanishing. In turn, then $0 = \lim_{k\to\infty} \frac{\mu(t_k)}{|t_k|^p}$ has the implication that for each i one of the coefficients c_i or d_i must vanish as well. Following in the same manner, we conclude that all linear coefficients in $\mu(t)$ also vanish, leaving only the possibility that $\mu(t)$ is a constant function. A similar approach can be conducted to prove the observation also for $p > 2$.

Notice now that $\overline{\mathcal{J}_p^\varepsilon}$ converges uniformly to $\bar{\mathcal{J}}_p$ on $\mathcal{U}_{\mathcal{I}_0^j}$ for $\varepsilon \to 0$, as defined in (3.12), or

$$|\overline{\mathcal{J}_p^\varepsilon}(v) - \bar{\mathcal{J}}_p(v)| \le \Gamma(\varepsilon), \quad \text{for all } v \in \mathcal{U}_{\mathcal{I}_0^j}, \tag{3.34}$$

for a continuous function $\Gamma(\varepsilon) = o(\varepsilon)$, $\varepsilon \to 0$. By Remark 3.1, for $X = \mathcal{A}(f) \cap \overline{\mathcal{U}_{\mathcal{I}_0^j}}$ and any $v^0 \in X$, there exists a linear subspace $\mathcal{V} \subset \mathcal{E}$, such that the orthogonal projection $X^\perp$ of X onto $\mathcal{V}^\perp$ has the properties

- $X = \{x = x^\perp \oplus tv : x^\perp \in X^\perp, v \in \mathcal{V}, t \in \mathbb{R}^+\}$,
- $M_C^j = X^\perp \cap \{v \in \mathcal{E} : \bar{\mathcal{J}}_p(v) \leq C\}$, for $C \geq \overline{\mathcal{J}_p^\varepsilon}(v^0) + \Gamma(\varepsilon)$ is compact, and
- $\bar{\mathcal{J}}_p(\xi_t)$ is constant along rays $\xi_t = x^\perp \oplus tv$, where $x^\perp \in M_C^j$, $v \in \mathcal{V}$, and $t \in \mathbb{R}^+$.

By the uniform estimate (3.34) and the last property, we deduce that $\overline{\mathcal{J}_p^\varepsilon}(\xi_t)$ is bounded from above and below by $\bar{\mathcal{J}}_p(x^\perp) \pm \Gamma(\varepsilon)$ on rays $\xi_t = x^\perp \oplus tv$, where $x^\perp \in M_C$, $v \in \mathcal{V}$, and $t \in \mathbb{R}^+$. Hence, we conclude that $\overline{\mathcal{J}_p^\varepsilon}(\xi_t)$ is also constant for $t \geq 0$. From (3.34), the set

$$X^\perp \cap \{v \in \mathcal{E} : \overline{\mathcal{J}_p^\varepsilon}(v) \leq \overline{\mathcal{J}_p^\varepsilon}(v^0)\},$$

is included in M_C^j, and

$$\inf_{v \in \mathcal{A}(f) \cap \mathcal{U}_{\mathcal{I}_0^j}} \overline{\mathcal{J}_p^\varepsilon}(v) = \inf_{v \in M_C^j} \overline{\mathcal{J}_p^\varepsilon}(v).$$

By compactness of $M^j = M_C^j$ and continuity of $\overline{\mathcal{J}_p^\varepsilon}$ we conclude the existence of minimizers in M^j. As pointed out above, this further implies the existence of minimal solutions in $M = \cup_{j=1}^{2^n} M^j$ of the original problem (3.32). Notice further that, by continuity of $\overline{\mathcal{J}_p^\varepsilon}(v^0) + \Gamma(\varepsilon)$ with respect to ε, the sets $M^j = M_C^j$ actually do not depend on $0 \leq \varepsilon$ as soon as $C \geq \max_{0<\varepsilon} \overline{\mathcal{J}_p^\varepsilon}(v^0) + \Gamma(\varepsilon)$ is large enough. □

Remark 3.8. *The previous result clarifies that, despite the fact that in general $\mathcal{J}_p^\varepsilon$ are not coercive functionals, up to restricting them to an appropriate compact set, independent of ε, they can be considered equi-coercive.*

Corollary 3.9. *The net of functionals $(\mathcal{J}_p^\varepsilon)_{0 \leq \varepsilon < r}$ Γ-converges to $\mathcal{J}_p$ on $\mathcal{A}(f)$. Moreover, if we consider the net of minimizers v_ε^* of $\mathcal{J}_p^\varepsilon$ in M for $0 \leq \varepsilon < r$, as constructed in Theorem 3.7 (which are actually minimizers of $\mathcal{J}_p^\varepsilon$ over $\mathcal{A}(f)$ as well), then the accumulation points of such a net are minimizers of $\mathcal{J}_p$.*

Proof. As $\mathcal{J}_p^\varepsilon$ converges uniformly to $\mathcal{J}_p$ on $\mathcal{A}(f)$, we deduce its Γ-convergence [72]. By Theorem 3.7 and compactness of M we conclude the convergence of minimizers. □

Proposition 3.10. *For all $0 < \varepsilon < r$, the functional $\mathcal{J}_p^\varepsilon$ satisfies the properties (A1) and (A2), i.e., it is ω-semi-convex, and* (2.3) *holds.*

Proof. The ω-semi-convexity follows from the piecewise continuity and boundedness of the second derivatives of $\mathcal{J}_p^\varepsilon$. Since $W_r^{p,\varepsilon}(t) \geq 0$ and $|(W_r^{p,\varepsilon})'(t)| \leq pr^{p-1}$ for every $t \in \mathbb{R}$, by means of the elementary inequality $a \leq \frac{1}{2}(a^2 + 1)$ we obtain

$$\begin{aligned} \|\nabla \mathcal{J}_p^\varepsilon(v)\| &\leq 2\|T^*(Tv - g)\| + \gamma\|((W_r^{p,\varepsilon})'(v_1), \dots, (W_r^{p,\varepsilon})'(v_n))\| \\ &\leq 2\|T^*\|\|Tv - g\| + \gamma n^{1/2} p r^{p-1} \\ &\leq \|T^*\|\|Tv - g\|^2 + \|T^*\| + \gamma n^{1/2} p r^{p-1} \\ &\leq \|T^*\| \mathcal{J}_p^\varepsilon(v) + \|T^*\| + \gamma n^{1/2} p r^{p-1}. \end{aligned} \tag{3.35}$$

Hence, for $K = \|T^*\|$ and $L = \|T^*\| + \gamma n^{1/2} p r^{p-1}$, we get that (2.3) holds for $\mathcal{J}_p^\varepsilon$. □

3.3.2. The application of the algorithm to coercive cases

As we clarified in the previous section, functionals of the type $\mathcal{J}_p^\varepsilon$, for $0 < \varepsilon < r$, satisfy the assumptions (A1) and (A2) for the applicability of Algorithm 2.2. In particular, when the algorithm is applied for $\mathcal{J} = \mathcal{J}_p^\varepsilon$, then by Theorem 2.7 the sequence $(v_\ell)_{\ell\in\mathbb{N}}$ generated by the algorithm has the properties

(a) $(Av_\ell - f) \to 0$ as $\ell \to \infty$;

(b) $(v_\ell - v_{\ell-1}) \to 0$ as $\ell \to \infty$.

However $\mathcal{J}_p^\varepsilon$ is unfortunately not necessarily coercive on $\mathcal{A}(f) = \{v \in \mathcal{E} : Av = f\}$, although it retains some coerciveness by considering suitable compact subsets M of competitors, see Theorem 3.7. Such information does not help to satisfy condition (A0), since there is no natural or simple way of restricting or projecting the iterations to such compact sets M. Hence, in order to apply Theorem 2.8, we need to explore the mechanism for which the iterations $(v_\ell)_{\ell\in\mathbb{N}}$ generated by the algorithm keep bounded. We show that this is the case where T is injective on $\ker(A)$, since this retrieves the coerciveness we need.
Let us first introduce some specific notation for the application of the Algorithm 2.2, in particular we denote

$$\mathcal{J}_{\omega,u}(v) := \mathcal{J}_{p,\omega,u}^\varepsilon(v) = \mathcal{J}_p^\varepsilon(v) + \omega\|v-u\|^2. \tag{3.36}$$

Lemma 3.11. *For all $0 < \varepsilon < r$, the sequence $(\|\nabla\mathcal{J}_p^\varepsilon(v_\ell)\|)_{\ell\in\mathbb{N}}$ is uniformly bounded, where the iterations $(v_\ell)_{\ell\in\mathbb{N}}$ are generated by Algorithm 2.2.*

Proof. As a consequence of (2.18) the sequence $(\|T\bar{v}_\ell\|)_\ell$, where $\bar{v}_\ell$ is defined in (2.13), is uniformly bounded. From (2.16), we have also that $(\|Tv_\ell\|)_\ell$ is uniformly bounded. As pointed out in (3.35) of Proposition 3.10 actually we have $\|\nabla\mathcal{J}_p^\varepsilon(v_\ell)\| \le 2\|T^*\|\|Tv_\ell - g\| + \gamma n^{1/2} p r^{p-1}$. Hence the sequence $(\|\nabla\mathcal{J}_p^\varepsilon(v_\ell)\|)_{\ell\in\mathbb{N}}$ is uniformly bounded. □

The next lemma will be crucial to show the convergence of the algorithm in our case.

Lemma 3.12. *For all $0 < \varepsilon < r$, the sequence $(A^*q_{\ell,L_\ell-1})_{\ell\in\mathbb{N}}$ generated by Algorithm 2.2 for $\mathcal{J} = \mathcal{J}_p^\varepsilon$ is uniformly bounded.*

Proof. By (2.8) we have

$$A^*q_\ell \in \nabla\mathcal{J}_{\omega,v_{\ell-1}}(v_\ell) = \nabla\mathcal{J}_p^\varepsilon(v_\ell) + 2\omega(v_\ell - v_{\ell-1})$$

As, by Lemma 3.11, $\nabla\mathcal{J}_p^\varepsilon(v_\ell)$ is uniformly bounded and $(v_\ell - v_{\ell-1}) \to 0$, for $\ell \to \infty$, we obtain that also A^*q_ℓ is uniformly bounded. By (2.2), we have also

$$A^*q_\ell = A^*q_{\ell,L_\ell-1} - 2\lambda A^*(Av_\ell - f),$$

from which, together with $(Av_\ell - f) \to 0$ for $\ell \to \infty$, we eventually deduce the uniform boundedness of $A^* q_{\ell,L_\ell-1}$ as well. □

Lemma 3.13. *Assume that T is injective on* $\ker(A)$*, or* $\ker(T) \cap \ker(A) = \{0\}$*. Then, for all $0 < \varepsilon < r$, the sequences $(v_\ell)_\ell$ generated by Algorithm 2.2 for $\mathcal{J} = \mathcal{J}_p^\varepsilon$ are uniformly bounded.*

Proof. Notice that, by its definition in Algorithm 2.2, v_ℓ necessarily solves the following linear system

$$(T^*T + \frac{1}{2}A^*A)v_\ell = \frac{1}{2}A^*(f + q_{\ell,L_\ell-1}) + \omega(v_{\ell-1} - v_\ell),$$

where the right-hand-side of this equality if uniformly bounded by Lemma 3.12 and Theorem 2.7 (b). Moreover, as $(Av_\ell - f) \to 0$ for $\ell \to \infty$, we can write that v_ℓ is solution of the system

$$\underbrace{\begin{bmatrix} (T^*T + \frac{1}{2}A^*A) \\ A \end{bmatrix}}_{:=G} v_\ell = w_\ell,$$

where the right-hand-side w_ℓ is actually uniformly bounded with respect to ℓ. Due to our assumption $\ker(T) \cap \ker(A) = \{0\}$, we obtain that $\ker(G) = \{0\}$ and

$$v_\ell = (G^*G)^{-1}G^*w_\ell, \qquad \text{for all } \ell \in \mathbb{N},$$

hence the uniform boundedness of $(v_\ell)_\ell$. □

We summarize this list of technical observations into the following convergence result.

Theorem 3.14. *Assume that T is injective on* $\ker A$*, or* $\ker T \cap \ker A = \{0\}$*. Then, for all $0 < \varepsilon < r$, the sequences $(v_\ell)_\ell$ generated by Algorithm 2.2 for $\mathcal{J} = \mathcal{J}_p^\varepsilon$ has at least one accumulation point, and every accumulation point is a constrained critical point of $\mathcal{J}_p^\varepsilon$ on the affine space $\mathcal{A}(f) = \{v \in \mathcal{E} : Av = f\}$.*

Proof. The result is obtained by a direct application of Theorem 2.8, after having recalled the boundedness of $(v_\ell)_\ell$, which results from Lemma 3.13. □

Remark 3.15. *The previous convergence result actually applies for the case of the Mumford–Shah functional, for which $T = D_h^\dagger$ and $A = I - D_h D_h^\dagger$, since $D_h^\dagger$ is in fact injective on* $\mathrm{ran}(D_h)$*, see Section 3.1.1.*

3.3.3. Iterative thresholding algorithms revisited

As already mentioned an iterative thresholding algorithm can be used for identifying local minimizers of the $\mathcal{J}_p$, see [108] for details. This algorithm is actually very attractive for its exceptional simplicity, and its ability of performing a *separation of components* at a finite number of iterations, leading eventually to a contractive iteration and its convergence.
In this section we design an iterative thresholding algorithm which can be applied also for linearly constrained problems of the type (3.11): namely, it is a very simple and efficient procedure for solving the inner loop minimization problems in our algorithm. Differently from the unconstrained case, however, the thresholding function we can use is a continuous one, so that we do not need to prove a result of separation of components after a finite number of iterations, and we gain additionally contractivity, unconditionally and from the beginning of the iteration.
For the sake of simplicity and without loss of generality, we consider the application of Algorithm 2.2 for $\lambda = 1/2$, and we define now the ν-strongly convex functional

$$\mathcal{J}_{\omega,u}(v,q) := \mathcal{J}^{\varepsilon}_{p,\omega,u}(v,q) = \mathcal{J}_{\omega,u}(v) + \frac{1}{2}\|Av - (f+q)\|^2. \tag{3.37}$$

Requiring ν-strong convexity is equivalent to the following lower bound on ω.

Lemma 3.16. *Define $\mathcal{J}_{\omega,u}(v,q)$ as in* (3.37)*. According to the notation introduced in Lemma 3.4, let B as in* (3.22) *for $s_1 = (r-\varepsilon)$, $s_2 = (r+\varepsilon)$, $\gamma_1 = p(r-\varepsilon)^{p-1}$, $\gamma_2 = (r-\varepsilon)^p$, and $\gamma_3 = r^p$. Then, for*

$$\omega > \gamma|B| = \gamma\left|\frac{p(r-\varepsilon)^{p-1}}{2\varepsilon} + \frac{3}{4\varepsilon^2}[(r-\varepsilon)^p - r^p]\right|, \tag{3.38}$$

$\mathcal{J}_{\omega,u}(v,q)$ is a ν-strongly convex function of v.

Proof. It obviously suffices to show that $\mathcal{J}_{\omega,u}(v)$ is ν-strongly convex, and since

$$\mathcal{J}_{\omega,u}(v) = \|Tv - g\|^2 + \gamma\sum_{i=1}^{n} W_r^{p,\varepsilon}(v_i) + \omega\sum_{i=1}^{n}(v_i - u_i)^2$$

it is enough to check that for every $s \in \mathbb{R}$ the real function

$$t \to \gamma W_r^{p,\varepsilon}(t) + \omega\sum_{i=1}^{n}(t-s)^2$$

is ν-strongly convex. But this function is piecewise $\mathcal{C}^2$ with bounded second derivatives, thus we must only check that for every t such that $|t| \notin \{r-\varepsilon, r+\varepsilon\}$, there exists $\nu > 0$ such that $\gamma(W_r^{p,\varepsilon})''(t) + 2\omega \geq \nu > 0$. By the explicit expression (3.20) of $W_r^{p,\varepsilon}(t)$, it all reduces to check that for every $t \in (r-\varepsilon, r+\varepsilon)$ one has

$$\gamma\pi_p''(t) + 2\omega \geq \nu > 0\,. \tag{3.39}$$

Now, in our case we have

$$B = \frac{p(r-\varepsilon)^{p-1}}{2\varepsilon} + \frac{3}{4\varepsilon^2}[(r-\varepsilon)^p - r^p], \tag{3.40}$$

$$A = \frac{p(r-\varepsilon)^{p-1}}{12\varepsilon^2} + \frac{B}{3\varepsilon}. \tag{3.41}$$

Since $(r-\varepsilon)^p - r^p \leq -\varepsilon p(r-\varepsilon)^{p-1}$ by convexity, we deduce from (3.40) that

$$B \leq -\frac{p(r-\varepsilon)^{p-1}}{4\varepsilon} < 0, \tag{3.42}$$

and therefore, from (3.41) we have also $A \leq 0$. But then for all $t \in (r-\varepsilon, r+\varepsilon)$, we deduce from these negativity relationships and again (3.41) that

$$\pi_p''(t) = 6A(t-(r+\varepsilon)) + 2B \geq 2B$$

so that (3.38) implies (3.39), with $\nu = 2(\omega - \gamma|B|)$, as required. □

We now define our thresholding function. We fix $0 < \varepsilon < r$, and for $W_r^{p,\varepsilon}(t)$ as in (3.20), B as in (3.40), and a positive parameter μ such that

$$\mu|B| < 1, \tag{3.43}$$

we consider

$$f_{\mu,r}^{\xi}(t) = (t-\xi)^2 + \mu W_r^{p,\varepsilon}(t) \tag{3.44}$$

with ξ a real number. By (3.43), arguing as in the proof of Lemma 3.16, we get ν-strong convexity of $f_{\mu,r}^{\xi}(t)$, therefore we can define a function $S_p^{\mu}(\xi)$ through

$$S_p^{\mu}(\xi) := \arg\min_{s\in\mathbb{R}} f_{\mu,r}^{\xi}(s)\,. \tag{3.45}$$

Then $S_p^{\mu}(\xi)$ satisfies the following properties.

Lemma 3.17. *For every $\xi \in \mathbb{R}$ and μ as in (3.43), the function $S_p^{\mu}(\xi)$ satisfies:*

(a) $S_p^{\mu}(\xi) = t$ if and only if $2(t-\xi) + \mu(W_r^{p,\varepsilon})'(t) = 0$.

(b) $S_p^{\mu}(\xi)$ is a strictly increasing function.

(c) $S_p^{\mu}(\xi)$ is Lipschitz continuous with

$$\mathrm{Lip}(S_p^{\mu}) \leq \frac{1}{1-\mu|B|}\,. \tag{3.46}$$

Proof. Part (a) of the statement is obvious by (3.44), (3.45) and the ν-strong convexity of $f^{\xi}_{\mu,r}(t)$.
To prove part (b), fix $\xi_1 < \xi_2 \in \mathbb{R}$ and correspondingly, let $t_1 := S^{\mu}_p(\xi_1)$ and $t_2 := S^{\mu}_p(\xi_2)$. We have $2(t_1 - \xi_1) + \mu(W^{p,\varepsilon}_r)'(t_1) = 0$. Assume by contradiction that $t_2 \leq t_1$. Since the function $t \mapsto 2(t - \xi_1) + \mu(W^{p,\varepsilon}_r)'(t)$ is strictly increasing by strong convexity, we get $2(t_2 - \xi_1) + \mu(W^{p,\varepsilon}_r)'(t_2) \leq 0$. Now $\xi_1 < \xi_2$ yields $2(t_2 - \xi_2) + \mu(W^{p,\varepsilon}_r)'(t_2) < 0$, in contradiction with part (a) of the statement.
To prove part (c), we fix ξ_1 and $\xi_2 \in \mathbb{R}$, and we can suppose without loss of generality that $\xi_1 < \xi_2$. Again, we define $t_1 := S^{\mu}_p(\xi_1)$, and $t_2 := S^{\mu}_p(\xi_2)$. From part (b), we have $t_1 < t_2$ and from part (a) we get that

$$2(t_1 - \xi_1) + \mu(W^{p,\varepsilon}_r)'(t_1) = 2(t_2 - \xi_2) + \mu(W^{p,\varepsilon}_r)'(t_2),$$

that is, since $\xi_1 < \xi_2$ and $t_1 < t_2$,

$$|t_2 - t_1| + \frac{\mu}{2}[(W^{p,\varepsilon}_r)'(t_2) - (W^{p,\varepsilon}_r)'(t_1)] = |\xi_2 - \xi_1|\,. \tag{3.47}$$

Now $(W^{p,\varepsilon}_r)'(t)$ is piecewise $\mathcal{C}^1$ with bounded derivative. Moreover, given B as in (3.40), arguing as in Lemma 3.16 we have $(W^{p,\varepsilon}_r)''(t) \geq 2B$ for every t such that $|t| \in (r-\varepsilon, r+\varepsilon)$. Since $(W^{p,\varepsilon}_r)''(t) \geq 0$ when $|t| \notin [r - \varepsilon, r + \varepsilon]$ and $B < 0$, we get that $(W^{p,\varepsilon}_r)''(t) \geq 2B$ for every t, with the only exceptions of the four points $t = r - \varepsilon$, $t = -r - \varepsilon$, $t = r + \varepsilon$, and $t = -r + \varepsilon$. Since $t_1 < t_2$, by the fundamental theorem of calculus we have

$$[(W^{p,\varepsilon}_r)'(t_2) - (W^{p,\varepsilon}_r)'(t_1)] \geq 2B(t_2 - t_1) = -2|B||t_2 - t_1|\,,$$

since $B < 0$. Using (3.47), this gives

$$(1 - \mu|B|)|t_2 - t_1| \leq |\xi_2 - \xi_1|,$$

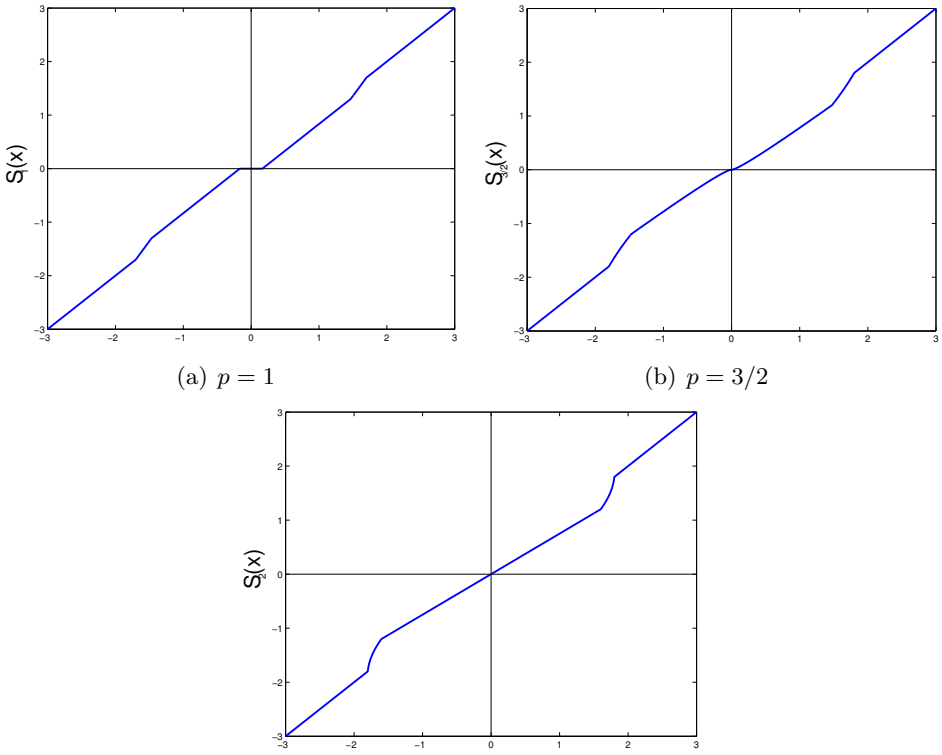
which concludes the proof. □

While this latter result states certain qualitative properties of S^{μ}_p for any $1 \leq p < \infty$, its expression for $p = 2$ can be easily obtained by solving a second degree polynomial equation:

$$S^{\mu}_2(\xi) = \begin{cases} \dfrac{\xi}{1+\mu} & |\xi| < (r-\varepsilon)(1+\mu) \\ \dfrac{4\varepsilon}{3\mu}\left(1 + \dfrac{\mu}{4\varepsilon}(2\varepsilon + r) - \sqrt{\dfrac{\Gamma(\xi)}{4}}\right) & (r-\varepsilon)(1+\mu) \leq |\xi| \leq r + \varepsilon \\ \xi & |\xi| > r + \varepsilon \end{cases} \tag{3.48}$$

where

$$\Gamma(\xi) = 4\left(1 + \left(\frac{\mu}{4\varepsilon}\right)^2 (2r + \varepsilon)^2 + \frac{\mu}{2\varepsilon}(r + 2\varepsilon) - \frac{3\mu}{2\varepsilon}\xi\right).$$



(a) $p = 1$ (b) $p = 3/2$

(c) $p = 2$

Figure 3.2.: The Lipschitz continuous thresholding functions S_1^μ, $S_{3/2}^\mu$, and S_2^μ, with parameters $r = 1.5$, $\mu = 5$, $\varepsilon = 0.3$.

We further report in Figure 3.2 the graphics of the thresholding function S_p^μ for $p \in \{1, 3/2, 2\}$, and parameters $r = 1.5$, $\mu = 5$, $\varepsilon = 0.3$.
We now get back to our functional $\mathcal{J}_{\omega,u}(v, q)$ defined in (3.37) and we further consider the associated surrogate functional,

$$\mathcal{J}^{surr}_{\omega,u}(v, q, w) := \mathcal{J}_{\omega,u}(v, q) + (\|v - w\|^2 - \|Tv - Tw\|^2) + (\|v - w\|^2 - \frac{1}{2}\|Av - Aw\|^2) + (\|v - w\|^2 - \omega\|v - w\|^2). \tag{3.49}$$

Up to rescaling of the intern parameter g, f, q, γ of $\mathcal{J}_{\omega,u}(v, q)$, we can assume without loss of generality, that $\|T\| < 1$, $\frac{1}{\sqrt{2}}\|A\| < 1$, and $\omega < 1$, while still keeping the lower bound on ω given by (3.38) which is necessary to ensure ν-strong convexity. Hence, we have

$$\mathcal{J}^{surr}_{\omega,u}(v, q, w) \geq \mathcal{J}_{\omega,u}(v, q), \tag{3.50}$$

where the equality occurs if and only if $w = v$.

Proposition 3.18. *Let $0 < \varepsilon < r$. Assume $\|T\| < 1$, $\frac{1}{\sqrt{2}}\|A\| < 1$, $\omega < 1$, and that ω and γ satisfy* (3.38)*. Then*

$$v^* = \arg\min_{v \in \mathcal{E}} \mathcal{J}_{\omega,u}(v, q) \tag{3.51}$$

if and only if v^* satisfies the following component-wise fixed-point equation: for $i = 1, \ldots, n$,

$$v_i^* = S_p\left(\frac{1}{3}\left\{[(I - T^*T) + (I - \frac{1}{2}A^*A) + (1-\omega)I]v^* + (T^*g + \frac{1}{2}A^*(f+q) + \omega u)\right\}_i\right), \tag{3.52}$$

where S_p is the thresholding function S_p^μ defined in Lemma 3.17 for $\mu = \gamma/3$.

Proof. Assume that v^* satisfies (3.51). From (3.50) we have the inequalities

$$\mathcal{J}^{surr}_{\omega,u}(v^*, q, v^*) = \mathcal{J}_{\omega,u}(v^*, q) \leq \mathcal{J}_{\omega,u}(v, q) = \mathcal{J}^{surr}_{\omega,u}(v, q, v) \leq \mathcal{J}^{surr}_{\omega,u}(v, q, v^*).$$

Hence we obtain also

$$v^* = \arg\min_{v \in \mathcal{E}} \mathcal{J}^{surr}_{\omega,u}(v, q, v^*). \tag{3.53}$$

We notice now by a direct computation that

$$\frac{1}{3}\mathcal{J}^{surr}_{\omega,u}(v, q, v^*) = \left\| v - \left(\frac{b^1 + b^2 + b^3}{3}\right)\right\|^2 + \frac{\gamma}{3}\sum_{i=1}^m W_r^{p,\varepsilon}(v_i) + C(b^1, b^2, b^3, \gamma), \tag{3.54}$$

where $b^1 = (I - T^*T)v^* + T^*g$, $b^2 = (I - \frac{1}{2}A^*A)v^* + \frac{1}{2}A^*(f+q)$, $b^3 = (I - \omega I)v^* + \omega u$, and $C(b^1, b^2, b^3, \gamma)$ is a term which does not depend on v. It now follows by the definition (3.45) of $S_p = S_p^\mu$ that v^* satisfies (3.52).
Conversely, by (3.54), if v^* satisfies (3.52), then it also satisfies (3.53). It follows that

$$0 \in \partial \mathcal{J}^{surr}_{\omega,u}(v^*, q, v^*) = \partial \mathcal{J}_{\omega,u}(v^*, q)$$

where the last equality trivially follows from (3.49). By convexity of $\mathcal{J}_{\omega,u}$, this implies (3.51). □

Looking at the fixed point equation (3.52), which characterizes the unique minimizer of $\mathcal{J}_{\omega,u}(v, q)$, it is natural to wonder whether the corresponding fixed-point iteration

$$v_i^{k+1} = S_p\left(\frac{1}{3}\left\{[(I - T^*T) + (I - \frac{1}{2}A^*A) + (1-\omega)I]v^k + (T^*g + \frac{1}{2}A^*(f+q) + \omega u)\right\}_i\right), \tag{3.55}$$

generates a sequence $(v^k)_{k \in \mathbb{N}}$ which converges to v^*.

Theorem 3.19. *Let $0 < \varepsilon < r$. Assume $\|T\| < 1$, $\frac{1}{\sqrt{2}}\|A\| < 1$, $\omega < 1$, and that ω and γ satisfy* (3.38). *Let*

$$v^* = \arg\min_{v \in \mathcal{E}} \mathcal{J}_{\omega,u}(v, q),$$

and consider the sequence v^k defined by the iteration (3.55). *Let $\delta := \dfrac{3-\omega}{3-\gamma|B|}$, where B is defined by* (3.40). *Then $\dfrac{2}{3} < \delta < 1$ and for every $n \in \mathbb{N}$ one has*

$$\|v^k - v^*\| \leq \frac{\delta^k}{1-\delta}\|v^1 - v^0\| \tag{3.56}$$

so that in particular $v^k \to v^*$ as n tends to $+\infty$.

Proof. By the assumptions $\omega < 1$ and (3.38), the bounds on δ are obvious. For every $k \geq 0$ one has $v^{k+1} = \mathbb{U}(v^k)$, where $\mathbb{U}$ is an operator having component-wise action defined by

$$[\mathbb{U}(v)]_i = S_p\left(\frac{1}{3}\left\{[(I - T^*T) + (I - \frac{1}{2}A^*A) + (1-\omega)I]v \right.\right. \\ \left.\left. +(T^*g + \frac{1}{2}A^*(f+q) + \omega u)\right\}_i\right), \tag{3.57}$$

where S_p is the function S_p^μ defined in Lemma 3.17 for $\mu = \gamma/3$. Using the hypotheses, it is easy to show that $\left\|\frac{1}{3}[(I - T^*T) + (I - \frac{1}{2}A^*A) + (1-\omega)I]\right\| \leq 1 - \frac{\omega}{3}$, therefore, using (3.46) for $\mu = \gamma/3$ we get

$$\mathrm{Lip}(\mathbb{U}) \leq \left(1 - \frac{\omega}{3}\right)\left(\frac{1}{1 - \gamma/3|B|}\right) = \delta\,;$$

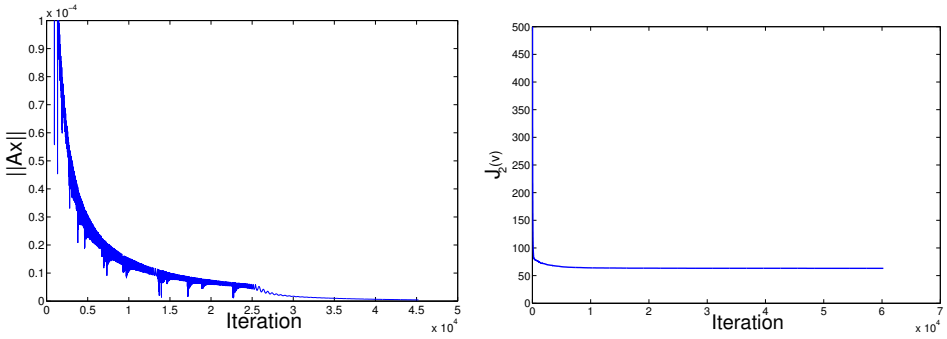
in particular, $\mathbb{U}$ is a contraction mapping, and we conclude by Banach fixed point Theorem. □

3.4. Numerical Experiments

In this section we report the results of numerical experiments performed to validate the behavior of the algorithm as predicted by our theoretical findings.
We first analyze the free discontinuity problems described in Section 3.1 as well as in [19], i.e., the minimization of the discrete Mumford–Shah functional in dimension two, and the discrete time quasi-static evolution of the Francfort–Marigo brittle fracture model in one dimension. Then, we describe in the performances of the Algorithm 2.2 for the minimization of the SLP-decoder introduced in [18] to recover compressed signals corrupted by noise. In particular, we will compare the decoder with respect to other well known techniques such as ℓ_1-minimization and its reweighted version.

3.4.1. Mumford–Shah functional minimization for image denoising

In all the simulations we used the iterative thresholding algorithm (3.55) in order to solve the convex optimizations of the inner loop. Our first experiment refers to the implementation of the algorithm for competitors u, being two dimensional arrays, of dimensions 25×25. The parameters chosen are $\gamma = 1.7 \times 10^{-1}$, $r = 3.5$, and $\varepsilon = 4.5 \times 10^{-3}$. Notice that ω is always explicitely fixed according to the formula $\omega > \gamma\left(\frac{1}{4} + \frac{r}{2\varepsilon}\right)$, as one can easily derive by combining (3.38) and (3.40), for $p = 2$.
In Figure 3.3 we show the dynamics of the discrepancy $\|Av_\ell\|$ to the realization of the linear constraint $Av = 0$, and of the energy $\mathcal{J}_p(v_\ell)$, depending on the iterations v_ℓ, for $\ell = 0, 1, 2, \ldots$. This simulation confirms that the algorithm tends to converge to a stationary point with energy level lower than the initial guess, and for which the constraint is numerically verified.

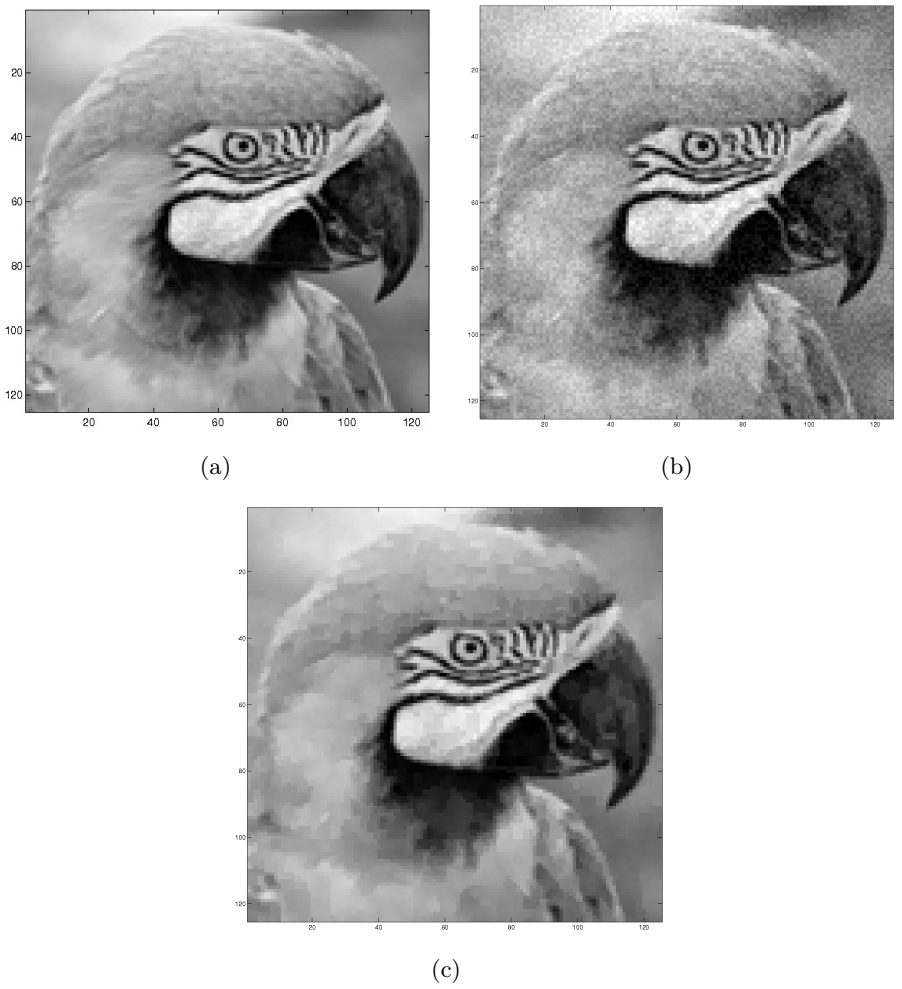


(a) (b)

Figure 3.3.: On the left subfigure we show the dynamics of the discrepancy $\|Av_\ell\|$ to the realization of the linear constraint $Av = 0$, and on the right subfigure the one of the energy $\mathcal{J}_p(v_\ell)$, depending on the iterations v_ℓ, for $\ell = 0, 1, 2, \ldots$

For a qualitative evaluation of the behavior of the algorithm, we report below an experiment on a denoising problem for an image of dimensions 125×125, see Figure 3.4, where the original image, the noisy version, and its denoised version after minimization are reported respectively in the subfigures (a), (b), and (c). The numerical experiments is conducted with 6% noise, and parameters $\gamma = 1.4 \times 10^{-1}$, $r = 2.8$, and $\varepsilon = 3.5 \times 10^{-3}$.

Remark 3.20. *As mentioned at beginning of this section, our numerical experiments are exclusively aimed at verifying the setting of the parameters and the convergence of the Algorithm 2.2, with no claim of optimal implementation. However, for the sake of completeness, we mention here how to treat the most demanding numerical issues. As the algorithm requires the applications of the matrices $D_h^\dagger$ and $(D_h^\dagger)^*$, one may wonder whether such matrices can be efficiently, stably computed and applied. In princi-*



(a) (b)

(c)

Figure 3.4.: Application of Algorithm 2.2 with inner loop realized by iterative thresholding (3.55) for a classical denoising problem.

ple, when enough memory is available there is no problem in computing such matrices in advance, also symbolically, and obtaining an iteration at machine precision. If the available memory is limited, one may avoid to attempt the explicit computation of such (pseudo)inverses, as it is also a good practice in numerical analysis. Rather one should use a preconditioned iterative method to approximate the results of their applications, i.e., $D_h^\dagger v$ *and* $(D_h^\dagger)^* u$*. For any matrix* X *the following identities hold:*

$$\begin{aligned} X^* X X^\dagger &= X^*, \\ X X^* (X^\dagger)^* &= X. \end{aligned}$$

In case of $X = D_h$ *the first identity gives us a method to compute* $D_h^\dagger v$*, as it is conse-*

quently sufficient to solve the linear system

$$\begin{cases} (D_h^* D_h)u = D_h^* v \\ c(u) = 0 \end{cases}, \tag{3.58}$$

and set $D_h^\dagger v = u$*, where* $c(u) = c_u$ *is the mean value of* u*. In order to see that the discrete system* (3.58) *can be efficiently and stably solved we need to highlight its relationships with a corresponding continuous system. In fact* (3.58) *actually can be simply interpreted as the discretization of the following continuous partial differential equation, which we write in its weak form*

$$\begin{cases} \displaystyle\int_\Omega \nabla u \cdot \nabla \varphi = \int_\Omega v \cdot \nabla \varphi \\ \displaystyle\int_\Omega u(x)dx = 0 \end{cases}$$

for all $\varphi \in H^1(\Omega)$*. Such an elliptic partial differential equation can be approached numerically very stably and efficiently by FEM or finite difference discretizations and solved by suitable preconditioned iterations, for instance by means of multigrid methods [127]. Similarly one can approach the computation of* $(D_h^\dagger)^* u$ *by defining the system*

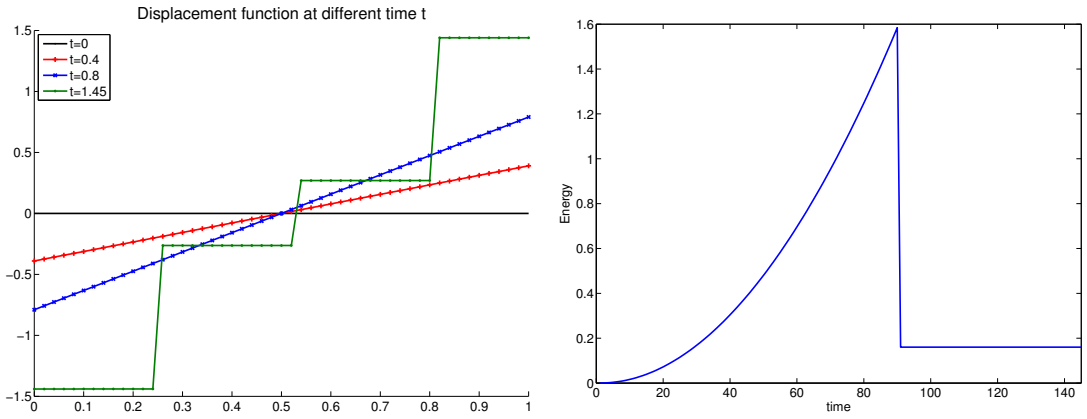
$$\begin{cases} (D_h D_h^*)v = D_h u, \\ Av = 0, \end{cases} \tag{3.59}$$

and setting $(D_h^\dagger)^* u = v$*, where* A *is again the discrete* curl *operator (notice that this matrix is sparse!). The efficient solution of the system* (3.59) *is again subordinated to the use of suitable preconditioners.*
Concerning the overall computational cost one may wonder whether the solution of two systems of (discretized) PDE such as (3.58) *and* (3.59) *is indeed an exceedingly large amount of effort. To this issue, let us respond that other well-known and established methods for the minimization of Mumford–Shah functional require also the solution of elliptic PDEs, for instance the L. Ambrosio and V. M. Tortorelli approach [9].*
In general, one can still object that the operator $T^*T = (D_h^\dagger)^* D_h^\dagger$ *as it appears in the iteration* (3.55) *is likely to be ill conditioned and this might affect negatively the convergence. However, as it is shown in Theorem 3.19, as soon as the operators* T *and* A *are properly rescaled and the parameters* ω *and* γ *are suitably set, the inner-loop* (3.55) *is guaranteed to converge with exponential rate.*

3.4.2. Brittle fracture simulation

We show in the following the result of the discrete time evolution of the Francfort–Marigo model for quasi-static brittle fracture evolution in one dimension adopting the discretization scheme presented in Section 3.1.2. Here we assume that Ω is the interval

$[0,1]$ and that the load is applied on the boundary $\Omega_D = \{0,1\}$. The load corresponds to a displacement $g(t,0) = -t$ and $g(t,1) = t$ at the boundary where $t \geq 0$ is the time variable. For the minimization of the functional (3.8) we again use Algorithm 2.2 with parameters $\gamma = 1$, $\varepsilon = 10^{-3}$, $r = 2$, and $\omega = \frac{1}{2}\left(\frac{1}{2} + \frac{r}{(N-1)\varepsilon}\right)$, where $N = 51$ is the number of space discretization points. The evolution proceeds with time steps of width $\Delta t = 0.01$. At every new time step, the algorithm is reinitiated and the initial guess correspond to the state of the gradient of the displacement at the previous time step.



(a) (b)

Figure 3.5.: Discrete time evolution of the Francfort–Marigo model of brittle fracture. In the subfigure (a) we show four stages of the beam displacement at different times, starting with an elastic evolution until crack formation. In the subfigure (b) we show the evolution of the energy (3.6), where the rupture time is highlighted also by the elastic energy collapse.

In Figure 3.5(a) we show four stages of the displacement u at the time $t = 0,\ 0.4,\ 0.8,\ 1.45$. As one can notice the beam is initially elastically deformed, then the crack appears at multiple positions at time $t = 0.9$, being a more favorable critical point of the energy. In Figure 3.5(b) we report the evolution of the energy (3.6) in time, where the failure time is highlighted also by the elastic energy collapse.
As clarified in Section 3.1.2, let us again stress that for this model there is no need of computing the action of the pseudoinverse matrix $D_h^\dagger$. The simulation of the entire evolution until the crack takes few minutes (a few seconds per time iteration) on a MacBook Pro equipped with a 2.6GHz Intel Core i7 processor, 8GB of RAM, 1600MHz DDR3 using a non optimized Matlab implementation.
In Figure 3.6 we show the computational time required at each discrete time and we observe how the algorithm needs to search longer for the new critical point, as soon as

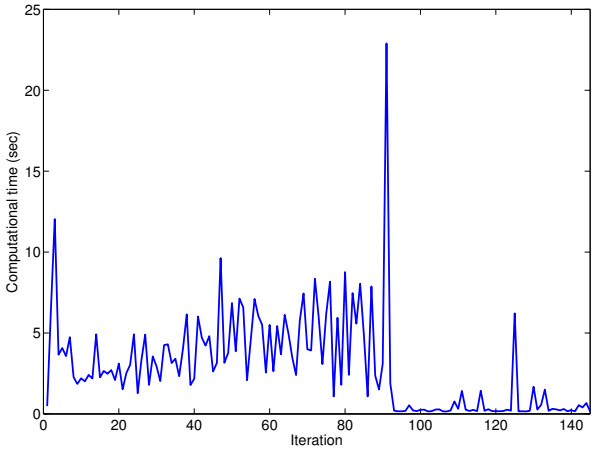


Figure 3.6.: Computational time in seconds for each discrete time step.

the physical phase transition from elastic evolution to fracture happens. The numerical results are consistent with the predicted analytic solution [18, 110], showing the robustness of Algorithm 2.2 towards the simulation of physical models.

3.4.3. Damping noise-folding in compressive sensing

In Section 3.2.1 we provided a result on support identification and damping noise-folding whose requirements cannot be unfortunately matched by standard decoders such has ℓ_1 and therefore we have introduced the regularized selective least p-powers (SLP) and this section aims to numerically show the enhancement of the decoding performances. Unfortunately, the numerical realization of the Algorithm 2.2 is computationally demanding as soon as the dimension N gets large. Therefore, we tested also the well-known iterative hard thresholding [32], which shows similar support identification properties as SLP while being very efficient in terms of computational time. Thus, with the following numerical simulations we additionally provide empirical confirmation of the theoretical observations in Section 3.2 and more widely analyzed in [18] and in the technical notes [17]. In particular, we observe that SLP and IHT, initialized by the ℓ_1-minimizer, are very robust and provide a significantly enhanced rate of recovery of the support of the unknown sparse vector as well as a better accuracy in approximating its large entries, with respect to the sole ℓ_1-minimization or its reweighted version, whenever limiting noise, i.e., $\eta \approx r$, is present on the signal. We refer to [18] for a generic overview of all these decoding procedures and in particular to [32] for a detailed description of the IHT, to [55, 171] for a in-depth analysis of the ℓ_1 decoders. We also consider as one of the test methods ℓ_1-minimization, where we substituted the equality constraint $Ax = y$ with an

inequality constraint which takes into account the noise level

$$\min \|z\|_{\ell_1} \text{ subject to } \|Az - y\|_{\ell_2} \leq \|An\|_{\ell_2} \leq \delta.$$

In the constraint we use the same parameter $\delta = \sigma^2(m + 2\sqrt{2m})$ as for the iteratively re-weighted ℓ_1-minimization (IRWℓ_1), indicated by the authors of [55, 171] as optimal. The stability parameter that avoids the denominator to be zero in the weight updating rule of IRWℓ_1 seems not to have a strong influence, and it is set to 0.1 in our experiments. We executed 8 iterations of IRWℓ_1 as a reasonable compromise between computational effort and accuracy. As we shall argue in detail below, the following numerical tests indicate that ℓ_1+IHT is much faster and slightly more robust than ℓ_1+SLP, and that both of them perform better than ℓ_1-minimization and IRWℓ_1 in terms of support recovery and accuracy in approximating the large entries in absolute value of the original signal. In order to fulfill the assumptions of our theoretical results, we use for the numerical experiments random matrices, satisfying the RIP with optimal constants with high probability. In particular all the tests presented in this section are realized with column-wise normalized i.i.d. Gaussian encoding matrices.

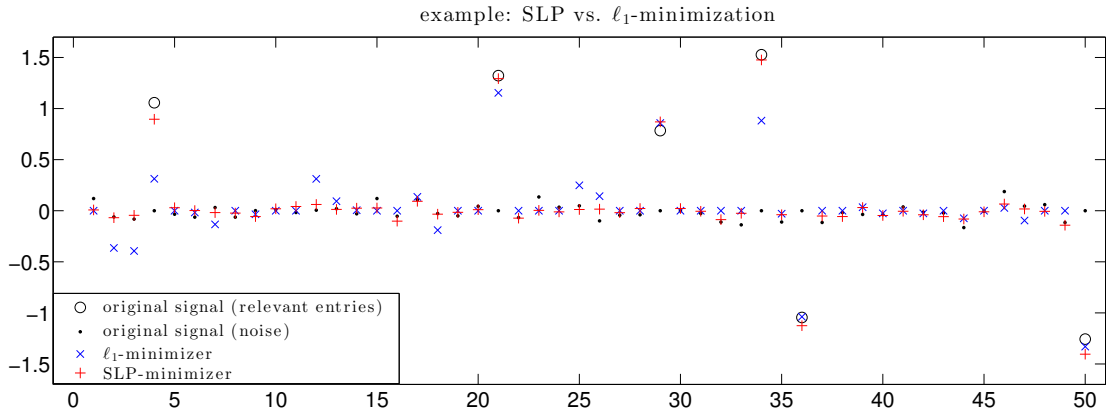


Figure 3.7.: Comparison of ℓ_1-minimizer and SLP-minimizer for one typical example. The two decoders are intended to recover the original signal (o), starting from the Gaussian random measurement of the noisy signal (o,·). The output of ℓ_1- and SLP-minimization is represented by (×) and (+) respectively.

Advantages of SLP with respect to ℓ_1-minimization We shall start the discussion on numerical experiments with a comparison between the ℓ_1-minimization and SLP-minimization for one typical example reported in Figure 3.7. For this experiment, we set $N = 50$, $m = 25$, $k = 6$, $r = 0.6$, $\eta = 0.59$, and choose the original noisy signal in the set $\mathcal{S}^2_{0.59,6,0.6}$ with all 6 relevant entries above $r = 0.6$ and the total norm of the noise $\|n\|_{\ell_2} = 0.5 < \eta$. Although the setting of the two methods is the same, the results

are different: the SLP-minimizer consists of 6 relevant values above $r = 0.6$ in absolute value and the norm of the remainder is $0.32 < \eta$, thus it is also in the class $\mathcal{S}^2_{0.59,6,0.6}$. The ℓ_1-minimizer consists of only 5 relevant entries above $r = 0.6$ in absolute value and the norm of the remainder is 0.75. Thus it is not an element of $\mathcal{S}^2_{0.59,6,0.6}$. Furthermore, it is evident that $|x_{\ell_1}(4)| < |x_{\ell_1}(12)|$ gives a wrong information about the location of the relevant entries, mismatching them with the noise. This phenomenon is due to the sparsefication of the noise in the recovery process. However, in this particular example, we were lucky to choose the right starting value for SLP. Due to its nonconvex character, in general SLP is computing a local minimizer, which might be far away from the original signal.

Choosing ℓ_1-minimization as a warm up As we mentioned, the Algorithm 2.2, which can be used to minimize locally the nonconvex functional $\mathcal{SP}^{p,\epsilon}_r$, finds only a critical point, so the condition $\mathcal{SP}^{2,\epsilon}_r(x^*) \leq \mathcal{SP}^{2,\epsilon}_r(x)$ (3.25) used in the proof of Theorem 3.5 may not be always valid. In order to enhance the chance that the resulting output of the algorithm satisfies this condition, the choice of an appropriate starting point is crucial. As we know that the ℓ_1-decoder provides us a global minimizer with at least some guarantees given by [18, Theorem 1], we use the result of this minimization process as a warm up to select the starting point. In the following, we denote with SLP which starts at $x_0 = 0$ and with ℓ_1+SLP the one starting from the ℓ_1-minimizer.
In Figure 3.8 we illustrate the robustness of ℓ_1+SLP (bottom left subfigure) in comparison to the ℓ_1-minimization based methods and SLP. Here SLP converged to a feasible critical point, but it is quite evident that the decoding process failed since the large entry at position 83 (signal) was badly recovered and even the entry at position 89 (noise) is larger. If we look at the ℓ_1-minimization result (top left subfigure) or the ℓ_1-minimization with inequality constraint (top right subfigure), the minimization process brings us close to the solution, but the results still significantly lack accuracy. By ℓ_1+SLP (center left subfigure) we obtain a good approximation of the relevant entries of the original signal and we get a significant correction and an improved recovery. Also IRWℓ_1 improves the result of ℓ_1-minimization significantly, but still approximates the large entries worse than ℓ_1+SLP. Although the difference is minor, we observe another important aspect of IRWℓ_1: the noise part is sparsely recovered, while ℓ_1+SLP distributes the noise in a more uniform way in a much smaller stripe around zero. This drawback of IRWℓ_1 can be crucial when it comes to the distinction of the relevant entries from noise. Let us stress once more that a correction after ℓ_1-minimization or IRWℓ_1 is necessary, because for these methods the noise part also is sparsely recovered. Indeed, these decoder has been designed to recover sparse signals and therefore also the noise tends to be recovered as a sparse contribution. Conversely, ℓ_1+SLP has been designed to distribute the noise in a more uniform way in a much smaller stripe around zero.

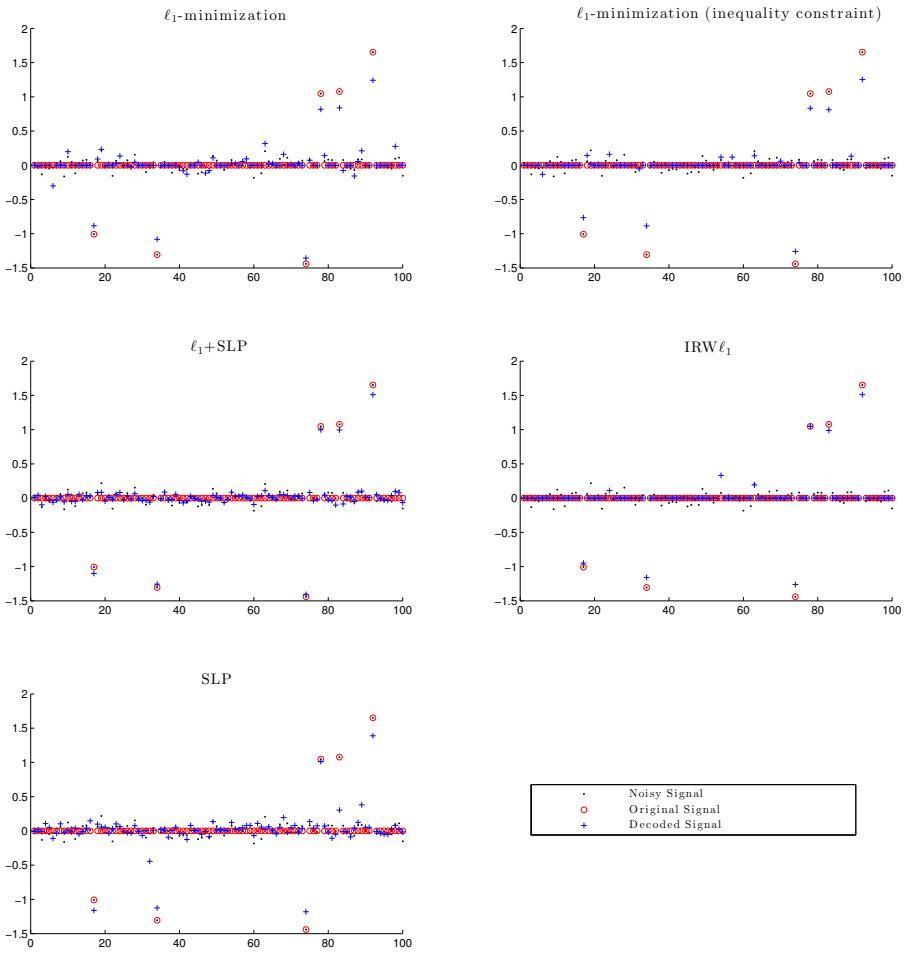


Figure 3.8.: The figure reports the results of five different decoding processes (+) of the same problem where the circles (o) represent the original signal and the points (·) represent the original signal corrupted by the noise.

Empirical statistics by extended computations The previously presented specific examples in support of our new decoding strategies are actually typical. In order to support this result with even more impressive and convincing evidences, we present some statistical data obtained by solving series of problems. We decided to fix the parameters in order to have the most coherent data to be analyzed; in particular, we set $N = 100$, $m = 40$, $r = 0.8$, $k = 1, \dots, 7$, and $\eta = 0.75$. The vector n is composed of random entries with normal distribution and then it is rescaled in order to have $\|n\|_{\ell_2} = \eta$. Figures 3.9, 3.10, and 3.11 report the results obtained considering 30 different i.i.d. Gaussian encoding matrices. In the following we use x^* generically for the decoded vector of any method.

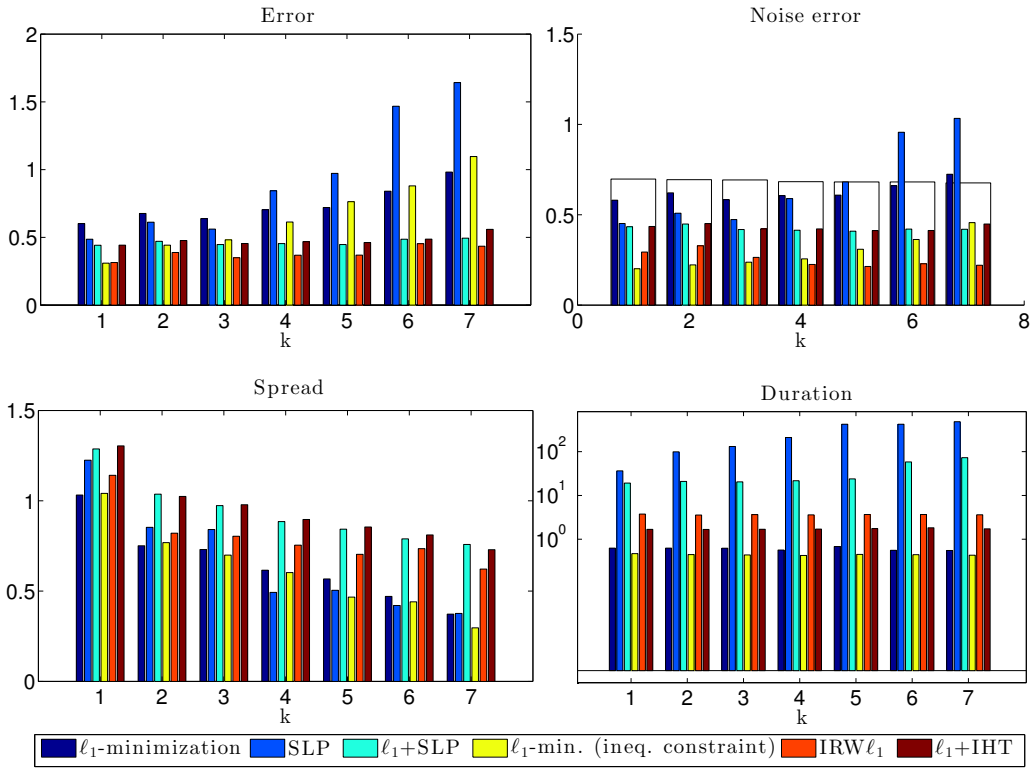


Figure 3.9.: The columns refer to the different results of ℓ_1-minimization (dark blue), SLP (blue), ℓ_1+SLP (cyan), ℓ_1-minimization with inequality constraint (yellow), IRWℓ_1 (orange), and ℓ_1+IHT (brown). In the *Noise error* subfigure the white column in the background represents the noise level. On the x-axis the different values of k are displayed and each column is the mean of the results given by 30 trials. The results were obtained by Gaussian matrices.

We start commenting the subfigures of Figure 3.9 clockwise. The first subfigure, on

the upper-left, represents the mean value of the error between the exact signal and the decoded one $\|x - x^*\|_{\ell_2}$. For ℓ_1+SLP, ℓ_1+IHT, and IRWℓ_1 the absolute ℓ_2-norm discrepancy between original and decoded vector is stable and independent of the (small) number k of large entries. These methods outperform ℓ_1-minimization; IRWℓ_1 performs slightly better. We also observe that the choice of the starting point is crucial for SLP and IHT.
The second subfigure is the mean value of the noise level $|\sigma_k(x)_{\ell_2} - \sigma_k(x^*)_{\ell_2}|$ and we can see exactly what we inferred looking at Figure 3.7: ℓ_1-minimization returns a larger noise level with respect to all the other methods, except SLP; and IRWℓ_1 has the best noise reduction property.
The third is the mean computational time, presented in logarithmic scale. All tests were implemented and run in Matlab in combination with CVX [71, 123], to solve the ℓ_1-minimization with equality and inequality constraint, its iteratively re-weighted version, and the QCQP on a MacBook Pro equipped with a 2.6GHz Intel Core i7 processor, 8GB of RAM, 1600MHz DDR3. We observe that SLP and ℓ_1+SLP are extremely slow. However, in comparison, the good starting point for SLP provides an advantage in terms of computational time. IHT has a computational complexity in between ℓ_1-minimization and IRWℓ_1.
The fourth plot reports the mean value of the discrepancy between noise level and the large entries of the signal, thus $\min_{i \in S_r(x)} |x_i^*| - \max_{i \in S_r(x)^c} |x_i^*|$. This plot shows how good the small entries are distinguished from the large ones in absolute value. We see that ℓ_1+SLP and ℓ_1+IHT perform best, which again is a result of their non-sparse noise recovery.

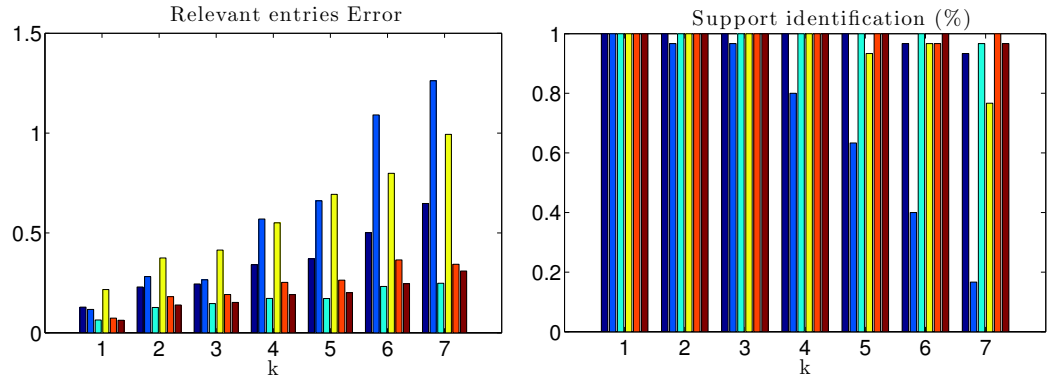


Figure 3.10.: The subfigures represent the error on the relevant entries and the support identification property by knowledge of k. For more details on the displayed data we refer to the caption of Figure 3.9. The results were obtained by Gaussian matrices.

In Figure 3.10 we report the histogram of the mean-value of the errors on the relevant

entries: the quantities on the left subfigure are computed as the mean values of $\|x|_{S_r(x)} - x|^*_{S_r(x)}\|_{\ell_2}$ where we suppose to know the number k of the nonzero entries of the original sparse signal. The right subfigure shows how often the k largest entries of x^* coincided with $S_r(x)$. Notice that there might be entries below the threshold r among the k largest entries of x^*. We conclude that, knowing the number of large entries, IRWℓ_1, ℓ_1-minimization, ℓ_1+SLP, and ℓ_1+IHT recover the support with nearly 100% success. In addition, ℓ_1+SLP approximates best the magnitudes of the relevant entries.

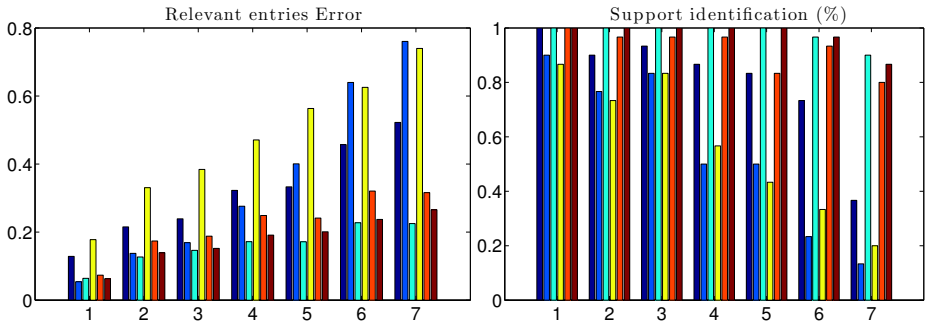


Figure 3.11.: The subfigures represent the error on the relevant entries and the support identification property by knowledge of r. For more details on the displayed data we refer to the caption of Figure 3.9. The results were obtained by Gaussian matrices.

In Figure 3.11 we compute again the mean-value of the relevant entries, but this time without the knowledge of k but the knowledge of r and therefore $S_r(x^*)$: the quantities on the left subfigure are the mean values of $\|x|_{S_r(x^*)} - x^*|_{S_r(x^*)}\|_{\ell_2}$. In the right subfigure we attribute a positive match in case $S_r(x^*) = S_r(x)$ so that the relevant entries of x^* coincide with the ones of the original signal. By our theory, we expect ℓ_1+SLP and ℓ_1+IHT to produce a high rate of success of correctly recovered support. Actually this is confirmed by the experiments: both methods do a very accurate recovery, as they give us almost always 100% of the correct result while the other methods perform worse.

Phase transition diagrams To give an even stronger support of the results in the previous paragraph, we extended the results of Figure 3.11 to a wider range of m and k. In Figure 3.12 we present phase transition diagrams of success rates in support recovery for ℓ_1-minimization, IRWℓ_1, ℓ_1+SLP, and ℓ_1+IHT in presence of nearly maximally allowed noise, i.e., $0.8 = r > \eta = 0.75$.
To produce phase transition diagrams, we varied the dimension of the measurement vector $m = 1, \dots, N$ with $N = 100$, and solved 20 different problems for all the admissible $k = \#S_r(x) = 1, \dots, m$. We colored black all the points (m, k), with $k \leq m$, which reported 100% of correct support identification, and we gradually reduce the tone up to

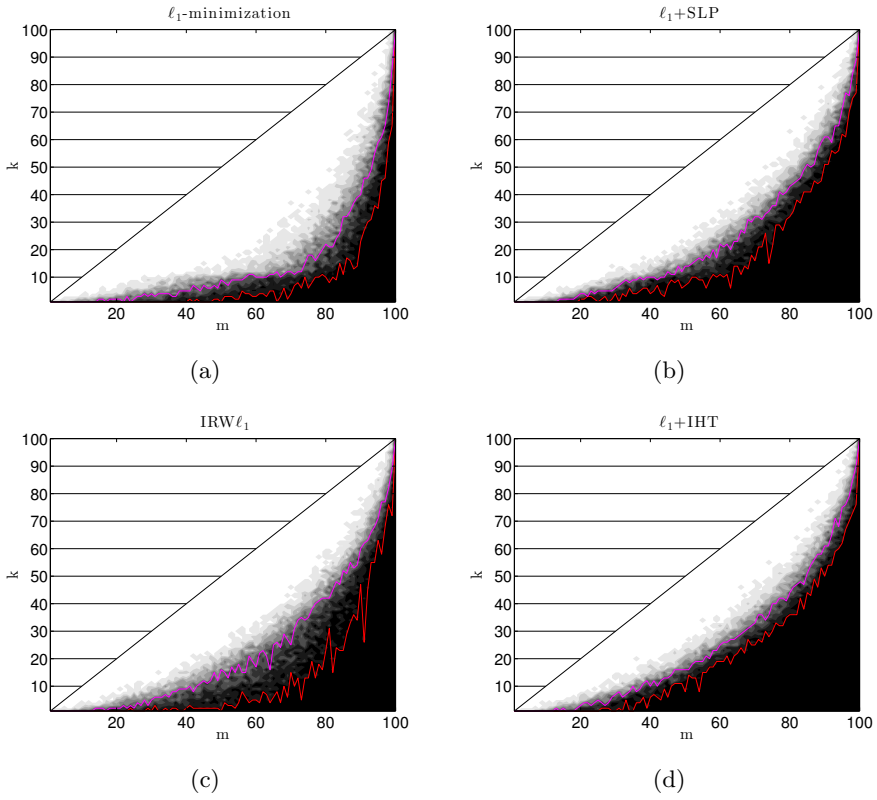
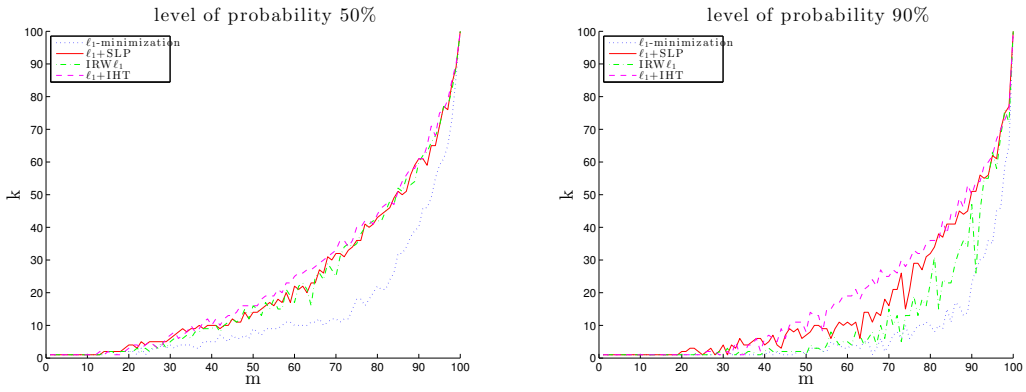


Figure 3.12.: Phase transition diagrams. The black area represents the couple (m, k) for which we had 100% of support recovery. The results of (a) ℓ_1-minimization, (b) ℓ_1+SLP, (c) IRWℓ_1, and (d) ℓ_1+IHT are reported. Note that the area for $k > m$ is not admissible. The red line shows the level bound of 90% of support recovery, and the magenta line 50% respectively.

white for the 0% result. The level bound of 50% and 90% is highlighted by a magenta and red line respectively. A visual comparison of the corresponding phase transitions confirms our previous expectations. In particular, ℓ_1+SLP and ℓ_1+IHT very significantly outperform ℓ_1-minimization in terms of correct support recovery. The difference of both methods towards IRWℓ_1 is less significant but still important. In Figure 3.13 we compare the level bounds of 50% and 90% among the four different methods. Observe that the 90% probability bound indicates the largest positive region for ℓ_1+IHT, followed by ℓ_1+SLP, and only eventually by IRWℓ_1, while the bounds are much closer to each other in the case of the 50% bound. Thus, surprisingly, ℓ_1+IHT works in practice even better than ℓ_1+SLP for some range of m, and offers the most stable support recovery results.



(a) (b)

Figure 3.13.: Comparison of phase transition diagrams for ℓ_1-minimization (dark blue, dotted), ℓ_1+SLP (red), IRWℓ_1 (green, dash-dotted), and ℓ_1+IHT (magenta, dashed). The level bound of 50% and 90% as it is displayed in Figure 3.12 is compared respectively in (a) and (b).

4. Quasi-static evolution of cohesive fracture models

In the previous chapter we have seen that the algorithms introduced in Chapter 2 can be used to solve minimization problem of various nature, from image denoising to fracture mechanics. The guarantee of convergence to critical points of the objective functional given by Algorithm 2.2 and Algorithm 2.4, together with the very mild regularity assumptions, is not only a purely mathematical achievement, but it may also have a nontrivial relevance in real life problem. In this framework we can naturally find the *quasi-static cohesive fracture evolution* introduced in Section 1.2.4.

In this chapter we present two different numerical approaches to the two classes of cohesive fractures: Barenblatt and Dugdale. In Section 4.1 we focus on Barenblatt cohesive energy functional and we show that by properly selecting the initial guess for Algorithm 2.2 at each time step of a given time discretization scheme, we are able to retrieve a physical evolution of a Barenblatt cohesive energy. Moreover, inspired by the seminal work of [51], we prove that not only we are able to select the critical point, which does not necessarily coincide with the global minimizer, that gives a physically sound evolution of the system, but also that the numerical and discrete, in time and space, solution obtained with the algorithm converges to the analytic and continuous, in time and space, one. We further specify that this first section is extracted from the preprint [13].

In the last part of the chapter, Section 4.2, we deal with a Dugdale energy with a particular focus on the computational efficiency. As we already observed for the compressive sensing case, the computational time of Algorithm 2.2 may blow up as soon as the system dimension grows. Therefore we propose alternative algorithms which perform efficiently and retrieve as well a feasible solution. Unfortunately, the algorithms presented in this section, despite being computationally very fast, do not have any convergence guarantee so far and thus in some application the use of these procedure may fail even if we did not experienced this undesired situation. Part of the material contained in this section will also appear, with some extensions, in [20], which is still in preparation.

4.1. The simulation of quasi-static evolution of a Barenblatt cohesive model

The reliability of a numerical simulation of a given process is based on the fact that for a discretization step going to zero the simulation coincides with solution of the continuous problem. In this section we analyze the discrete-to-continuous convergence of a quasi-static simulation of a Barenblatt cohesive fracture model obtained by using the Algorithm 2.2 presented in the Chapter 2. Since the convergence result is indeed valid in a more general setting than the fracture mechanics, we introduce the theoretical results considering the following assumptions. We additionally give advance notice that, as already done in Chapter 2, we need to restrict the domain of the problem to a finite dimensional space to guarantee the property of closure of the subdifferential of $\mathcal{J}$, which is indeed needed to prove Theorem 4.5, even if most of the results hold true also in a continuous setting.

Let us recall the two Euclidean spaces $\mathcal{E} \simeq \mathbb{R}^n$ and $\mathcal{F} \simeq \mathbb{R}^m$ for $m \leq n$. We assume we are given a functional $\mathcal{J} : \mathcal{E} \to [0, +\infty)$, a linear operator $A : \mathcal{E} \to \mathcal{F}$, and a time dependent forcing or constraint absolutely continuous function $f : [0, T] \to \mathcal{F}$, defined over a time interval $[0, T]$ with finite time horizon $T > 0$. A linearly constrained evolution of critical points relative to $\mathcal{J}$ and the linear constraint pair (A, f) is a bounded measurable time-dependent map $u : [0, T] \to \mathcal{E}$, with the following properties

$$Au(t) = f(t), \quad \mathrm{ran}(A^*) \cap \partial\mathcal{J}(u(t)) \neq 0, \text{ for almost every } t \in [0, T], \tag{4.1}$$

where $\partial\mathcal{J}$ is the subdifferential of $\mathcal{J}$, there exists a bounded measurable function $q : [0, T] \to \mathcal{F}$ such that $A^*q(t) \in \partial\mathcal{J}(u(t))$, and the inequality

$$\mathcal{J}(u(t)) \leq \mathcal{J}(u(0)) + \int_0^t \langle q(s), \dot{f}(s)\rangle_{\mathcal{F}} ds, \text{ for almost every } t \in [0, T], \tag{4.2}$$

holds, where the scalar product $\langle \cdot, \cdot \rangle_{\mathcal{F}}$ is the Euclidean one on $\mathcal{F}$. We additionally specify that in this section we always specify the space in which the norm is understood to ease the comprehension of the arguments.

We aim to set sufficiently general conditions on the functional $\mathcal{J}$ and the linear operator A which allow to establish existence of linearly constrained evolution of critical points u satisfying (4.1) and (4.2). For this reason we do not exclude that the convergence result presented in this section might be also applied to different models than Barenblatt cohesive fractures.

Beside the generality of the conditions on the energy function $\mathcal{J}$ and the linear operator A which we shall impose to derive the existence of evolutions of critical points, two additional features of our results should be emphasized:

(i) we stress that $u(t)$ is supposed to visit at different times t *critical points* of the cost function $\mathcal{J}$ over the affine space $\mathcal{A}(f(t)) := \{v \in \mathcal{E} : Av = f(t)\}$. This is

a rather general request compared to the more common requirement of being a global minimizer of $\mathcal{J}$ over the affine space $\mathcal{A}(f(t))$. Despite the more realistic and physically sound modeling of evolution along critical points, actually their analysis is usually more involved and may allow for solutions u measurable in time, but, unfortunately, not necessarily having any sort of smoothness [51, 77, 154, 155, 209]. While we shall be content with the generality of our approach, we have to live with the fact that our solutions may not be regular;

(ii) our approach for obtaining linearly constrained evolution of critical points u is *constructive*. In particular, as we will emphasize later, the functional $\mathcal{J}$ may have multiple feasible critical points at the same time t. Hence, in order to promote uniqueness of evolution, or even just their measurability, we need to design a proper selection principle. Accordingly, by properly choosing the starting guess of Algorithm 2.2 in such a way that the closest - in terms of Euclidean distance - critical point is selected to the one chosen at the previous instant of time, unless it is energetically convenient to perform a "jump" to another significantly different phase of the system. This corresponds to a rather common and well-established behavior of several physical (and non physical) systems [149, 184, 194, 195]. Summarizing, our evolution u is the result of a constructive machinery, which is designed to emulate physical principles, according to which a critical point is selected in terms of a balance between neighborliness (accounting the Euclidean path length between critical points) and energy convenience. In our view, this feature is really of great relevance, as it represents a blind black box out of which, through a process implementable by a computer, physically sound solutions come out;

(iii) as an important remark, we stress that a priori all of the constants appearing in the technical assumptions in Section 4.1.1 could depend on the dimension of the considered Euclidean spaces. Thus, our results can be applied to physical systems that can assume (a discrete or a continuum of) of infinitely many states, provided all the relevant estimates obtained are dimension free. For this reason, we explain very clearly which are the parameters that can affect the constants that come into (see Remark 4.12). Applying the method to cohesive fractures, we show how an infinite dimensional system can be studied with our method. This, in particular, gives an alternative proof of the existence of evolution of critical points for the cohesive fracture model firstly proven in [51]. In addition, we can provide numerical simulations (see Section 4.1.5).

Let us stress that similar approaches appeared in the recent literature, for instance [167, 174]. The main difference of the contribution of this chapter with respect to [174] in that we are mainly concerned with nonsmooth setting, while the cited work consider only functional with at least $\mathcal{C}^1$ regularity. Considering the abstract approach of [167],

which also applies to infinite dimensions, we instead provide a strategy that is suitable for numerical purposes and we investigate the relations between discretized and continuous problems.

4.1.1. Preliminary assumptions

Let $\mathcal{J} : \mathcal{E} \to [0, +\infty)$ be a (possibly nonconvex) functional. Consider the linear operator $A : \mathcal{E} \to \mathcal{F}$ and assume that it has nontrivial kernel and it is surjective. In particular, this implies that the adjoint operator $A^* : \mathcal{F} \to \mathcal{E}$ satisfies

$$|A^* q|_{\mathcal{E}} \geq \gamma |q|_{\mathcal{F}} \qquad \forall q \in \mathcal{F}, \tag{4.3}$$

for some $\gamma > 0$. Let us recall assumption (A0), (A1), and (A2) defined in Section 2.0.5

(A0) the functional $v \longmapsto \mathcal{J}(v) + |Av|^2_{\mathcal{F}}$ is coercive;

(A1) there exists $\omega > 0$ such that $v \longmapsto \mathcal{J}_\omega(v) := \mathcal{J}(v) + \omega |v|^2_{\mathcal{E}}$ is strictly convex;

(A2) there exist $K, L > 0$ such that, for every $v \in \mathcal{E}$, $\xi \in \partial \mathcal{J}(v) \Longrightarrow |\xi|_{\mathcal{E}'} \leq K \mathcal{J}(v) + L$,

which we assume hold true in this context.

Remark 4.1. *If $\mathcal{J}$ is globally Lipschitz continuous with Lipschitz constant L, then*

$$\xi \in \partial \mathcal{J}(v) \Longrightarrow |\xi|_{\mathcal{E}'} \leq L, \tag{4.4}$$

for every $v \in \mathcal{E}$.

Additionally, recall that in Remark 2.3 we observed that hypothesis (A1) implies that $\mathcal{J}$ is a smooth perturbation of a convex function and it follows that it is at least Lipschitz continuous. Thus, $\mathcal{J}$ is differentiable at almost every point, and that its Fréchet subdifferential is non empty at every point. Furthermore, it suffices to check condition (A2) only at differentiability points of $\mathcal{J}$, to ensure that it is satisfied at every point.
As an important remark, we stress that in principle all of the constants appearing in the technical assumptions are allowed to be depending on the dimension of the considered Euclidean spaces, and then so also the estimates we derive. A priori estimates when the dimension tends to infinity are to be obtained with different considerations, as we will show in the concrete example of Section 4.1.3.
We refer to Definition 2.2 for the notion of critical point. Additionally, it is well known that this condition is necessary for local minimality and it implies the following variational condition:

$$0 \leq \liminf_{\varepsilon \to 0} \frac{\mathcal{J}(v^* + \varepsilon w) - \mathcal{J}(v^*)}{\varepsilon}, \quad \forall w \in \ker(A). \tag{4.5}$$

Given $\omega > 0$ such that (A1) holds, and $u \in \mathcal{E}$, we denote, as before, by $\mathcal{J}_{\omega,u} : \mathcal{E} \to [0, \infty)$ the functional

$$\mathcal{J}_{\omega,u}(v) = \mathcal{J}(v) + \omega |v - u|^2_{\mathcal{E}}. \tag{4.6}$$

The time evolution of our system is driven by the time dependent forcing term $f : [0, T] \to \mathcal{F}$.

In order to prove the existence of an approximable discrete evolution (its definition is given in details in Definition 4.4), we shall assume a technical hypothesis which can replace in some cases requests of higher smoothness for the map $\mathcal{J}$. We warn nevertheless the reader that this requirement is automatically satisfied, for instance, when $\mathcal{J} \in \mathcal{C}^{1,1}$, independently of the given f. On the other hand, we will be able to retrieve this hypothesis also in some concrete examples, where $\mathcal{J} \notin \mathcal{C}^{1,1}$ (see Section 4.1.3). The hypothesis we make is then the following

(A3) There exists a positive constant $C_{\mathcal{J},\omega} > 0$. Given q_1 and q_2 such that

$$A^* q_i \in \partial \mathcal{J}_{\omega,u}(v_i), \quad i = 1, 2$$

then it holds

$$\langle q_1 - q_2, Av_1 - Av_2 \rangle_{\mathcal{F}} \leq C_{\mathcal{J},\omega} |v_1 - v_2|_{\mathcal{E}} \, |Av_1 - Av_2|_{\mathcal{F}} \tag{4.7}$$

for every $\overline{v} \in \mathcal{E}$.

Remark 4.2. *If $\mathcal{J} \in \mathcal{C}^{1,1}$, $\partial \mathcal{J}(v)$ is single valued at every $v \in \mathcal{E}$, and it coincides with the differential $D\mathcal{J}(v)$. Then, denoting by L the Lipschitz constant of $D\mathcal{J}$ and using* (4.3)*, one has*

$$|q_1 - q_2|_{\mathcal{F}} \leq \frac{1}{\gamma} |A^* q_1 - A^* q_2|_{\mathcal{E}} = \frac{1}{\gamma} |D\mathcal{J}(v_1) - D\mathcal{J}(v_2)|_{\mathcal{E}'} \leq \frac{L}{\gamma} |v_1 - v_2|_{\mathcal{E}}.$$

At this point, (4.7) *simply follows by the Cauchy-Schwarz inequality.*

Before stating our main result, we give the notion of dicrete and approximable quasi-static evolution, respectively.

Definition 4.3. *Let $v_0 \in \mathcal{E}$ be a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(0))$, and let $\delta > 0$. A* discrete quasi-static evolution *with time step δ, initial condition v_0, and constraint f is a right-continuous function $v_\delta : [0, T] \to \mathcal{E}$ such that*

- *$v_\delta(0) = v_0$;*
- *v_δ is constant in $[0, T] \cap [i\delta, (i + 1)\delta)$ for all $i \in \mathbb{N}_0$ with $i\delta \leq T$;*
- *$v_\delta(i\delta)$ is a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(i\delta))$ for every $i \in \mathbb{N}$ with $i\delta \leq T$.*

Definition 4.4. *Let $v_0 \in \mathcal{E}$ be a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(0))$. A bounded measurable function $v : [0,T] \to \mathcal{E}$ is said to be an* approximable quasi-static evolution *with initial condition v_0 and constraint f, if for every $t \in [0,T]$ there exists a sequence $\delta_k \to 0^+$ (possibly depending on t) and a sequence $\{v_{\delta_k}\}_{k\in\mathbb{N}}$ of discrete quasi-static evolutions with time step δ_k, initial condition v_0, and constraint f, such that*

$$\lim_{k\to+\infty} |v_{\delta_k}(t) - v(t)|_{\mathcal{E}} = 0. \tag{4.8}$$

4.1.2. Existence of approximable quasi-static evolution

We are now ready to state our main result,

Theorem 4.5. *Let* (4.3)*, (A0), (A1), (A2), and (A3) be satisfied. Additionally, consider $f \in W^{1,2}([0,T];\mathcal{F})$, and let v_0 be a critical point of $\mathcal{J}$ in the affine space $\mathcal{A}(f(0))$. Then, there exist bounded and measurable functions $v : [0,T] \to \mathcal{E}$ and $q : [0,T] \to \mathcal{F}$ such that:*

(a) v is an approximable quasi-static evolution with initial condition v_0 and constraint f;

(b) $A^*q(t) \in \partial\mathcal{J}(v(t))$ for every $t \in [0,T]$;

(c) The function $s \mapsto \langle q(s), \dot{f}(s)\rangle_{\mathcal{F}}$ belongs to $L^1(0,T)$ and for every $t \in [0,T]$ we have

$$\mathcal{J}(v(t)) \le \mathcal{J}(v_0) + \int_0^t \langle q(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds. \tag{4.9}$$

In order to prove the theorem, we need some preliminary and auxiliary result. We start by showing that, under suitable assumptions, an approximable quasi-static evolution is automatically an evolution of critical points of the energy functional $\mathcal{J}$.

Proposition 4.6. *Suppose that* (4.3)*, (A1), and (A2) are satisfied, and consider $f \in W^{1,2}([0,T];\mathcal{F})$. Let $v_0 \in \mathcal{E}$ be a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(0))$, and let $v : [0,T] \to \mathcal{E}$ be an approximable quasi-static evolution with initial condition v_0 and constraint f. Then, $v(t)$ is a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(t))$ for every $t \in [0,T]$.*

Proof. Let $\{v_{\delta_k}\}_{k\in\mathbb{N}}$ be as in (4.8), and let $t \in [0,T]$ be fixed. For every $k \in \mathbb{N}$, let $i_k \in \mathbb{N}$ be such that (to ease the notation, we do not stress the dependence of i_k on t)

$$i_k\delta_k \le t < (i_k+1)\delta_k.$$

From the definition of approximate quasi-static evolution we have $Av_{\delta_k}(t) = f(i_k\delta_k)$. Then, by continuity of f and (4.8) we obtain $Av(t) = f(t)$.

We thus need only to show that $\partial\mathcal{J}(v(t)) \cap \mathrm{ran}(A^*) \neq \emptyset$. By definition of constrained critical point, there exists $q_k \in \mathcal{F}$ (note that also q_k will, in general, depend on t) such that

$$A^* q_k \in \partial\mathcal{J}(v_{\delta_k}(t)) \qquad \text{for every } k \in \mathbb{N}. \tag{4.10}$$

From (A1) it follows that $\mathcal{J}$ is locally bounded and therefore, by (4.8), we have

$$\sup_{k\in\mathbb{N}} \mathcal{J}(v_{\delta_k}(t)) < +\infty\,.$$

Last relation, together with (4.3), (4.10), and (A2), gives

$$\sup_{k\in\mathbb{N}} |q_k|_{\mathcal{F}} \leq \frac{1}{\gamma} \sup_{k\in\mathbb{N}} |A^* q_k|_{\mathcal{E}} \leq \frac{1}{\gamma}(K \sup_{k\in\mathbb{N}} \mathcal{J}(v_{\delta_k}(t)) + L) < +\infty\,.$$

Thus, there exists $q \in \mathcal{F}$ such that, up to subsequences,

$$\lim_{k\to+\infty} |q_k - q|_{\mathcal{F}} = 0\,. \tag{4.11}$$

From (4.8), (4.10), and (4.11) we get, by the closure property of the subdifferential, that

$$A^* q \in \partial\mathcal{J}(v(t))\,,$$

as required. □

To construct an approximate quasi-static evolution, we first introduce an auxiliary minimum problem. Let $\delta \in (0,1)$ be a fixed time step, and let $i \in \mathbb{N}$ with $i\delta \leq T$. Suppose that $v^{i-1} \in \mathcal{E}$ is a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f((i-1)\delta))$. If property (A1) is satisfied, we define the sequence $\{v^i_j\}_{j\in\mathbb{N}_0}$ by setting $v^i_0 := v^{i-1}$ and

$$v^i_j := \operatorname*{arg\,min}_{Av=f(i\delta)} \{\mathcal{J}(v) + \omega|v - v^i_{j-1}|^2_{\mathcal{E}} \,:\, v \in \mathcal{E}\} \qquad \text{for every } j \in \mathbb{N}. \tag{4.12}$$

Remark 4.7. *Note that (A1) guarantees that the above minimum is unique.*

The following lemma gives some properties of the sequence $\{v^i_j\}_{j\in\mathbb{N}}$.

Lemma 4.8. *Let* (4.3)*, (A0), (A1), and (A2) be satisfied, and let $f \in W^{1,2}([0,T];\mathcal{F})$. Let $\delta \in (0,1)$ and let $i \in \mathbb{N}$ with $i\delta \leq T$. Suppose that v^{i-1} is a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f((i-1)\delta))$, and let $\{v^i_j\}_{j\in\mathbb{N}_0}$ be as in* (4.12)*. Then:*

1. *$\{\mathcal{J}(v^i_j)\}_{j\in\mathbb{N}}$ is a nonincreasing converging sequence;*

2. *$\{v^i_j\}_{j\in\mathbb{N}}$ is bounded and*
$$\lim_{j\to+\infty} |v^i_j - v^i_{j-1}|_{\mathcal{E}} = 0; \tag{4.13}$$

3. any limit point of $\{v_j^i\}_{j\in\mathbb{N}}$ is a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(i\delta))$.

Proof. For every $j \geq 2$ we have $Av_j^i = Av_{j-1}^i = f(i\delta)$, and therefore v_{j-1}^i is a competitor for the minimum problem in (4.12). Thus,

$$\mathcal{J}(v_j^i) \leq \mathcal{J}(v_{j-1}^i) - \omega|v_j^i - v_{j-1}^i|_{\mathcal{E}}^2, \qquad \text{for every } j \geq 2. \tag{4.14}$$

In particular, the sequence $\{\mathcal{J}(v_j^i)\}_{j\in\mathbb{N}}$ is nonincreasing. Since $\mathcal{J} \geq 0$, the limit

$$\lim_{j\to\infty} \mathcal{J}(v_j^i) =: C \geq 0 \tag{4.15}$$

exists and is nonnegative, and this shows *1.*. Let now $M \in \mathbb{N}$ with $M > 2$. Summing up relation (4.14) for $j = 2, \dots, M$ we obtain

$$\sum_{j=2}^{M} |v_j^i - v_{j-1}^i|_{\mathcal{E}}^2 \leq \frac{1}{\omega}(\mathcal{J}(v_1^i) - \mathcal{J}(v_M^i)).$$

Sending $M \to \infty$ we then have

$$\sum_{j=2}^{\infty} |v_j^i - v_{j-1}^i|_{\mathcal{E}}^2 \leq \frac{1}{\omega}(\mathcal{J}(v_1^i) - C) < \infty.$$

In particular, this shows that (4.13) holds true. Note now that, by (4.15), $\{\mathcal{J}(v_j^i)\}_{j\in\mathbb{N}}$ is bounded. Therefore, since $|Av_j^i|_{\mathcal{F}}^2 = |f(i\delta)|_{\mathcal{F}}^2$ for every $j \geq 1$, the sequence $\{\mathcal{J}(v_j^i) + |Av_j^i|_{\mathcal{F}}^2\}_{j\in\mathbb{N}}$ is bounded. By (A0), we have that $\{v_j^i\}_{j\in\mathbb{N}}$ is also bounded, and this concludes the proof of *2.*.

Let $v^i \in \mathcal{E}$ be a limit point of $\{v_j^i\}_{j\in\mathbb{N}}$. Up to subsequences, we can assume that

$$\lim_{j\to\infty} v_j^i = v^i, \qquad \text{in } \mathcal{E}.$$

First of all, note that $Av^i = f(i\delta)$. By (4.12), for every $j \in \mathbb{N}$ there exists $q_j^i \in \mathcal{F}$ such that

$$A^* q_j^i \in \partial\mathcal{J}(v_j^i) + 2\eta(v_j^i - v_{j-1}^i),$$

where we used (A.8). The previous relation can also be written as

$$A^* q_j^i = \xi_j^i + 2\eta(v_j^i - v_{j-1}^i), \tag{4.16}$$

for some $\xi_j^i \in \partial\mathcal{J}(v_j^i)$. Note that, since $\{v_j^i\}_{j\in\mathbb{N}}$ is bounded, by (A2) we also have that $\{\xi_j^i\}_{j\in\mathbb{N}}$ is bounded. Thanks to (4.13) and (4.3), this implies that $\{q_j^i\}_{j\in\mathbb{N}}$ is also bounded. Thus, up to subsequences, we can assume that

$$\lim_{j\to\infty} \xi_j^i = \xi^i \quad \text{ in } \mathcal{E}' \qquad \text{and} \qquad \lim_{j\to\infty} q_j^i = q^i \quad \text{ in } \mathcal{F},$$

for some $\xi^i \in \mathcal{E}'$ and $q^i \in \mathcal{F}$. Passing to the limit in (4.16), thanks to (4.13) we conclude that

$$A^* q^i = \xi^i.$$

By the closure property of subdifferentials we have $\xi^i \in \partial\mathcal{J}(v^i)$, and thus $\partial\mathcal{J}(v^i) \cap \mathrm{ran}(A^*) \neq \emptyset$. □

Remark 4.9. *Suppose that v^i and z^i are two limit points of the sequence $\{v^i_j\}_{j\in\mathbb{N}}$. By property (1) of the previous lemma, even if $v^i \neq z^i$ we have*

$$\mathcal{J}(v^i) = \mathcal{J}(z^i).$$

We state now a direct consequence of the previous lemma.

Corollary 4.10. *Let (4.3), (A0), (A1), and (A2) be satisfied, and $f \in W^{1,2}([0,T];\mathcal{F})$. Let $\delta \in (0,1)$ and let v_0 be a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(0))$. Set $v^0 := v_0$ and, for every $i \in \mathbb{N}$ with $i\delta \leq T$, let $\{v^i_j\}_{j\in\mathbb{N}_0}$ be defined by (4.12), and let v^i be a limit point of $\{v^i_j\}_{j\in\mathbb{N}_0}$. Then, the function $v_\delta : [0,T] \to \mathcal{E}$ defined as*

$$v_\delta(t) := v^i \quad \text{for every } t \in [0,T] \cap [i\delta, (i+1)\delta), \quad \text{for every } i \in \mathbb{N}_0 \text{ with } i\delta \leq T, \tag{4.17}$$

is a discrete quasi-static evolution with time step δ, initial condition v_0, and constraint f.

A key property of the discrete quasi-static evolution above is that it satisfies an approximate energy inequality, up to an error which vanishes with δ.

Theorem 4.11. *Let (4.3), (A0), (A1), (A2), and (A3) be satisfied, and additionally consider $f \in W^{1,2}([0,T];\mathcal{F})$. Let $\delta \in (0,1)$ and let v_0 be a critical point of $\mathcal{J}$ on the affine space $\mathcal{A}(f(0))$. Let $v_\delta : [0,T] \to \mathcal{E}$ be the discrete quasi-static evolution with time step δ, initial condition v_0, and constraint f given by (4.17). Then, there exist a piecewise constant right-continuous function $q_\delta : [0,T] \to \mathcal{F}$ and positive constants C_1, C_2 and C_3, independent of δ, such that*

(i) $A^* q_\delta(t) \in \partial\mathcal{J}(v_\delta(t))$ for every $t \in [0,T]$;

(ii) $|q_\delta(t)|_\mathcal{E} \leq C_1$ and $|v_\delta(t)|_\mathcal{F} \leq C_2$ for every $t \in [0,T]$;

(iii) for every $t \in [0,T]$

$$\mathcal{J}(v_\delta(t)) \leq \mathcal{J}(v_0) + \int_0^t \langle q_\delta(s), \dot{f}(s)\rangle_\mathcal{F}\, ds + C_3\sqrt{\delta}.$$

Remark 4.12. *More precisely, as it appears by a careful reading of the proof of Theorem 4.11, we have*

$$C_1 = C_1(\gamma, K, L, \omega, C_{\mathcal{J},\omega}, \|\dot f\|_{L^2((0,T);\mathcal{F})}) \ \text{ and } \ C_3 = C_3(\gamma, K, L, \omega, C_{\mathcal{J},\omega}, \|\dot f\|_{L^2((0,T);\mathcal{F})}).$$

Instead, concerning the constant C_2,

$$C_2 = C_2(\gamma, K, L, \omega, C_{\mathcal{J},\omega}, \|f\|_{L^\infty((0,T);\mathcal{F})}, \|\dot f\|_{L^2((0,T);\mathcal{F})}, \mathcal{J}).$$

The dependence on $\mathcal{J}$ above has to be intended in the sense of the coercivity assumption (A0). Therefore, if assumptions (4.3)*, (A1), (A2), and (A3) are satisfied by a family of functions $\{\mathcal{J}_\kappa\}_\kappa$, (with the same A, f, γ, ω, and $C_{\mathcal{J},\omega}$ for all κ) and, in addition, the functonals $v \longmapsto \mathcal{J}_\kappa(v) + |Av|^2_{\mathcal{F}}$ are equicoercive, then the constant C_2 is the same for all the family $\{\mathcal{J}_\kappa\}_\kappa$.*

Proof. Since v_δ is a discrete quasi-static evolution, for every $i \in \mathbb{N}$ with $i\delta \le T$ there exists $q^i \in \mathcal{F}$ such that $A^*q^i \in \partial\mathcal{J}(v_\delta(i\delta))$. Then, if we define $q_\delta : [0,T] \to \mathcal{F}$ as

$$q_\delta(t) := q^i \quad \text{for all } t \in [0,T] \cap [i\delta, (i+1)\delta), \quad \text{for all } i \in \mathbb{N}_0 \text{ with } i\delta \le T,$$

property (i) is satisfied. We now divide the remaining part of the proof into three steps.

Step 1: We show that there exists a constant M, depending only on ω and $C_{\mathcal{J},\omega}$, such that

$$\mathcal{J}(v_\delta(i\delta)) \le \mathcal{J}(v_\delta((i-1)\delta)) + \int_{(i-1)\delta}^{i\delta} \langle q^\delta(s), \dot f(s)\rangle_{\mathcal{F}}\, ds + M\delta \int_{(i-1)\delta}^{i\delta} |\dot f(s)|^2_{\mathcal{F}}\, ds, \quad (4.18)$$

for every $i \in \mathbb{N}$ with $i\delta \le T$.
To this aim, let $i \in \mathbb{N}$ with $i\delta \le T$ be fixed, and let $\{v^i_j\}_{j\in\mathbb{N}}$ be the sequence defined by (4.12). By property (A1), the functional $\mathcal{J}_{\omega,v^i_0}$ is strictly convex. Therefore, whenever $\xi \in \partial\mathcal{J}_{\omega,v^i_0}(v^i_1)$, we have

$$\mathcal{J}_{\omega,v^i_0}(v) \ge \mathcal{J}_{\omega,v^i_0}(v^i_1) + \langle \xi, v - v^i_1\rangle_{\mathcal{E}',\mathcal{E}} \qquad \text{for every } v \in \mathcal{E}.$$

In particular, choosing $v = v^i_0$ and recalling the definition of $\mathcal{J}_{\omega,v^i_0}$ we have

$$\mathcal{J}(v^i_0) \ge \mathcal{J}(v^i_1) + \omega|v^i_1 - v^i_0|^2_{\mathcal{E}} + \langle \xi, v^i_0 - v^i_1\rangle_{\mathcal{E}',\mathcal{E}} \qquad \text{for every } \xi \in \partial\mathcal{J}_{\omega,v^i_0}(v^i_1). \quad (4.19)$$

By (4.12), v^i_1 is the global minimizer of $\mathcal{J}_{\omega,v^i_0}$ on $\mathcal{A}(f(i\delta))$. Therefore, there exists $r^i \in \mathcal{F}$ such that $A^*r^i \in \partial\mathcal{J}_{\omega,v^i_0}(v^i_1)$ so that, by (4.19),

$$\mathcal{J}(v^i_0) \ge \mathcal{J}(v^i_1) + \omega|v^i_1 - v^i_0|^2_{\mathcal{E}} + \langle A^*r^i, v^i_0 - v^i_1\rangle_{\mathcal{E}',\mathcal{E}}.$$

Therefore, by the absolute continuity of f, and recalling that q_δ is constant in the interval $[(i-1)\delta, i\delta)$, we have

$$\begin{aligned}
\omega|v_1^i - v_0^i|_{\mathcal{E}}^2 + \mathcal{J}(v_1^i) - \mathcal{J}(v_0^i) &\le \langle A^* r^i, v_1^i - v_0^i\rangle_{\mathcal{E}',\mathcal{E}} \\
&= \langle A^* r^i - A^* q_\delta((i-1)\delta), v_1^i - v_0^i\rangle_{\mathcal{E}',\mathcal{E}} + \langle A^* q_\delta((i-1)\delta), v_1^i - v_0^i\rangle_{\mathcal{E}',\mathcal{E}} \\
&= \langle r^i - q_\delta((i-1)\delta), Av_1^i - Av_0^i\rangle_{\mathcal{F}} + \langle q^\delta((i-1)\delta), Av_1^i - Av_0^i\rangle_{\mathcal{F}} \\
&= \langle r^i - q_\delta((i-1)\delta), Av_1^i - Av_0^i\rangle_{\mathcal{F}} + \int_{(i-1)\delta}^{i\delta} \langle q^\delta((i-1)\delta), \dot f(s)\rangle_{\mathcal{F}}\, ds,
\end{aligned} \tag{4.20}$$

where we used the fact that $Av_1^i = f(i\delta)$ and $Av_0^i = f((i-1)\delta)$. Observe now that, by definition of q_δ, we have $A^* q_\delta((i-1)\delta) \in \partial\mathcal{J}(v_\delta((i-1)\delta))$. Thus, recalling that $v_\delta((i-1)\delta) = v^{i-1} = v_0^i$, we obtain

$$A^* q_\delta((i-1)\delta) \in \partial\mathcal{J}(v_0^i) = \partial\mathcal{J}_{\omega,v_0^i}(v_0^i).$$

Then, recalling that $A^* r^i \in \partial\mathcal{J}_{\omega,v_0^i}(v_1^i)$, by property (A3) we achieve

$$\langle r^i - q_\delta((i-1)\delta), Av_1^i - Av_0^i\rangle_{\mathcal{F}} \le C_{\mathcal{J},\omega}|v_1^i - v_0^i|_{\mathcal{E}}|Av_1^i - Av_0^i|_{\mathcal{F}}$$

which, together with (4.20), gives

$$\begin{aligned}
&\omega|v_1^i - v_0^i|_{\mathcal{E}}^2 + \mathcal{J}(v_1^i) - \mathcal{J}(v_0^i) \\
&\qquad \le C_{\mathcal{J},\omega}|v_1^i - v_0^i|_{\mathcal{E}}|f(i\delta) - f((i-1)\delta)|_{\mathcal{F}} + \int_{(i-1)\delta}^{i\delta} \langle q^\delta(s), \dot f(s)\rangle_{\mathcal{F}}\, ds.
\end{aligned}$$

Using Young's and Hölder's inequality, we get

$$\begin{aligned}
\mathcal{J}(v_1^i) &\le \mathcal{J}(v_0^i) + M|f(i\delta) - f((i-1)\delta)|_{\mathcal{F}}^2 + \int_{(i-1)\delta}^{i\delta} \langle q^\delta(s), \dot f(s)\rangle_{\mathcal{F}}\, ds \\
&= \mathcal{J}(v_0^i) + M\left|\int_{(i-1)\delta}^{i\delta} \dot f(s)\, ds\right|_{\mathcal{F}}^2 + \int_{(i-1)\delta}^{i\delta} \langle q^\delta(s), \dot f(s)\rangle_{\mathcal{F}}\, ds \\
&\le \mathcal{J}(v_0^i) + M\delta\int_{(i-1)\delta}^{i\delta} |\dot f(s)|_{\mathcal{F}}^2\, ds + \int_{(i-1)\delta}^{i\delta} \langle q^\delta(s), \dot f(s)\rangle_{\mathcal{F}}\, ds,
\end{aligned} \tag{4.21}$$

for a suitable constant $M > 0$ (depending only on ω and $C_{\mathcal{J},\omega}$), where we also used the fact that $f \in W^{1,2}([0,T];\mathcal{F})$. Recalling that $\mathcal{J}(v_j^i) \le \mathcal{J}(v_1^i)$ for all $j \ge 2$, by (4.17) and property (1) of Lemma 4.8, we have

$$\mathcal{J}(v_\delta(i\delta)) = \mathcal{J}(v^i) = \lim_{j\to\infty} \mathcal{J}(v_j^i) \le \mathcal{J}(v_1^i).$$

Taking into account last inequality, and recalling that $v_0^i = v^{i-1} = v_\delta((i-1)\delta)$, relation (4.21) gives (4.18).

Step 2: We prove (ii). Let $t \in [0,T]$ be fixed, and let $\bar{\imath} \in \mathbb{N}$ be such that

$$\bar{\imath}\delta \le t < (\bar{\imath}+1)\delta.$$

Adding up relation (4.18) for $i = 1, \dots, \bar{\imath}$, and recalling that $v_\delta(t) = v_\delta(\bar{\imath}\delta)$, we obtain

$$\mathcal{J}(v_\delta(t)) \le \mathcal{J}(v_0) + \int_0^{\bar{\imath}\delta} \langle q_\delta(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds + M\delta \int_0^T |\dot{f}(s)|_{\mathcal{F}}^2\, ds. \tag{4.22}$$

From (4.3) and (A2) we have

$$|q_\delta(s)|_{\mathcal{F}} \le \frac{1}{\gamma}|A^* q_\delta(s)|_{\mathcal{E}} \le \frac{K}{\gamma}\mathcal{J}(v_\delta(s)) + \frac{L}{\gamma} \qquad \text{for all } s \in [0,T]. \tag{4.23}$$

Therefore, (4.22) gives

$$\mathcal{J}(v_\delta(t)) \le \mathcal{J}(v_0) + \frac{K}{\gamma}\int_0^t \mathcal{J}(v_\delta(s))|\dot{f}(s)|_{\mathcal{F}}\, ds + \frac{L}{\gamma}\int_0^t |\dot{f}(s)|_{\mathcal{F}}\, ds + M\delta \int_0^T |\dot{f}(s)|_{\mathcal{F}}^2\, ds.$$

Thus, from Gronwall's inequality

$$\sup_{\substack{\delta\in(0,1)\\ t\in[0,T]}} \mathcal{J}(v_\delta(t)) \le K_1, \tag{4.24}$$

for some positive constant $K_1 = K_1(\gamma, K, L, \omega, C_{\mathcal{J},\omega}, \|\dot{f}\|_{L^2((0,T);\mathcal{F})})$. Then, from (4.23) and (4.24) we have

$$\sup_{\substack{\delta\in(0,1)\\ t\in[0,T]}} |q_\delta(t)|_{\mathcal{F}} \le C_1,$$

for some positive constant $C_1 = C_1(\gamma, K, L, \omega, C_{\mathcal{J},\omega}, \|\dot{f}\|_{L^2((0,T);\mathcal{F})})$. Thus,

$$\sup_{\substack{\delta\in(0,1)\\ t\in[0,T]}} \Big(\mathcal{J}(v_\delta(t)) + |Av_\delta(t)|_{\mathcal{E}}^2\Big) \le \sup_{\substack{\delta\in(0,1)\\ t\in[0,T]}} \Big(\mathcal{J}(v_\delta(t)) + \|f\|_{L^\infty((0,T);\mathcal{F})}^2\Big) < K_2,$$

for some positive constant $K_2 = K_2(\gamma, K, L, \omega, C_{\mathcal{J},\omega}, \|\dot{f}\|_{L^2((0,T);\mathcal{F})}, \|f\|_{L^\infty((0,T);\mathcal{F})})$. Taking into account (A0), last inequality implies that

$$|v_\delta(t)|_{\mathcal{E}} \le C_2, \qquad \text{for every } \delta \in (0,1) \text{ and } t \in [0,T],$$

with a constant C_2 that also depends on the coercivity of the function $v \mapsto \mathcal{J}(v) + |Av|_{\mathcal{E}}^2$, see Remark 4.12.

Step 3: We show (iii). From (4.22) and Hölder's inequality, taking into account that $\delta \in (0,1)$

$$\begin{aligned}
\mathcal{J}(v_\delta(t)) &\le \mathcal{J}(v_0) + \int_0^{\bar{\imath}\delta} \langle q_\delta(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds + M\delta \|\dot{f}\|^2_{L^2((0,T);\mathcal{F})} \\
&= \mathcal{J}(v_0) + \int_0^t \langle q^\delta(s), \dot{f}(s)\rangle_{\mathcal{F}} - \int_{\bar{\imath}\delta}^t \langle q^\delta(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds + M\delta \|\dot{f}\|^2_{L^2((0,T);\mathcal{F})} \\
&\le \mathcal{J}(v_0) + \int_0^t \langle q^\delta(s), \dot{f}(s)\rangle_{\mathcal{F}} + C_1\sqrt{\delta}\|\dot{f}\|_{L^2((0,T);\mathcal{F})} + M\sqrt{\delta}\|\dot{f}\|^2_{L^2((0,T);\mathcal{F})} \\
&= \mathcal{J}(v_0) + \int_0^t \langle q^\delta(s), \dot{f}(s)\rangle_{\mathcal{F}} + \sqrt{\delta}\left(C_1\|\dot{f}\|_{L^2((0,T);\mathcal{F})} + M\|\dot{f}\|^2_{L^2((0,T);\mathcal{F})}\right),
\end{aligned}$$

which gives (iii). □

Before giving the proof of Theorem 4.5, we need the following result (see [74, Lemma 3.6]).

Lemma 4.13. *Let X be a compact metric space. Let $p : [0,T] \to \mathbb{R}$, $p_k : [0,T] \to \mathbb{R}$ and $f_k : [0,T] \to X$ be measurable functions, for every $k \in \mathbb{N}$. For every $t \in [0,T]$ let us set*

$$\mathcal{I}(t) := \{x \in X : \exists\, k_j \to +\infty \text{ such that } x = \lim_{j\to+\infty} f_{k_j}(t) \text{ and } p(t) = \lim_{j\to\infty} p_{k_j}(t)\}.$$

Then

- *$\mathcal{I}(t)$ is closed for all $t \in [0,T]$;*
- *for every open set $U \subseteq X$ the set $\{t \in [0,T] : \mathcal{I}(t) \cap U \neq \emptyset\}$ is measurable.*

We conclude this section with the proof of Theorem 4.5. We acknowledge F. Cagnetti for being the main contributor to this result.

Proof of Theorem 4.5. We divide the proof into several steps.

Step 1: Proof of (a) and (b).

For every $\delta \in (0,1)$, let $q_\delta : [0,T] \to \mathcal{F}$ and $v_\delta : [0,T] \to \mathcal{E}$ be given by Theorem 4.11. Let $\Lambda \subset [0,T]$ be such that $\mathcal{L}^1(\Lambda) = 0$ and $\dot{f}(t)$ is well defined for every $t \in [0,T] \setminus \Lambda$. We fix a sequence $\{\delta_k\}_{k\in\mathbb{N}}$ such that $\delta_k \to 0^+$ and define

$$\eta_k(t) := \begin{cases} \langle q_{\delta_k}(t), \dot{f}(t)\rangle_{\mathcal{F}} & \text{for every } t \in [0,T] \setminus \Lambda, \\ 0 & \text{for every } t \in \Lambda, \end{cases}$$

and

$$\eta(t) := \limsup_{k\to\infty} \eta_k(t) \quad \text{for every } t \in [0,T].$$

By definition of θ, for every $t \in [0,T]$ we can extract a subsequence $\{\delta_{k_j}\}_{j\in\mathbb{N}}$ (possibly depending on t) such that

$$\eta(t) = \lim_{j\to\infty} \eta_{k_j}(t) \quad \text{for every } t \in [0,T].$$

By (ii) of Theorem 4.11, we have

$$|q_\delta(t)|_{\mathcal{F}} \le C_1, \quad |v_\delta(t)|_{\mathcal{E}} \le C_2, \quad \text{for every } \delta \in (0,1) \text{ and } t \in [0,T].$$

Thus, for every $t \in [0,T]$ we can extract a further subsequence (not relabelled) such that

$$\lim_{j\to\infty} v_{\delta_{k_j}}(t) = v(t) \quad \text{in } \mathcal{E} \quad \text{and} \quad \lim_{j\to\infty} q_{\delta_{k_j}}(t) = q(t) \quad \text{in } \mathcal{F},$$

for some $v(t) \in \mathcal{E}$ and $q(t) \in \mathcal{F}$ with $|q(t)|_{\mathcal{F}} \le C_1$ and $|v(t)|_{\mathcal{E}} \le C_2$. Let us now show that, for every $t \in [0,T]$, we can choose the subsequence $\{k_j\}_{j\in\mathbb{N}}$ in such a way that the maps $q : [0,T] \to \mathcal{F}$ and $v : [0,T] \to \mathcal{E}$ are measurable.
Let us denote by $B^{\mathcal{F}}_{C_1}$ ($B^{\mathcal{E}}_{C_2}$) the closed ball of $\mathcal{F}$ ($\mathcal{E}$) with center at the origin and radius C_1 (C_2). Applying Lemma 4.13 with $X = B^{\mathcal{F}}_{C_1} \times B^{\mathcal{E}}_{C_2}$, $f_k = (q_{\delta_k}, v_{\delta_k})$ and $p_k = \eta_k$, we have that

- $\mathcal{I}(t)$ is closed for all $t \in [0,T]$,
- for every open set $U \subseteq X$ the set $\{t \in [0,T] : \mathcal{I}(t) \cap U \neq \emptyset\}$ is measurable,

where the set $\mathcal{I}(t)$ is given by

$$\begin{aligned}\mathcal{I}(t) := \{(q(t),v(t)) \in B^{\mathcal{F}}_{C_1} \times B^{\mathcal{E}}_{C_2} : \exists\, k_j \to +\infty \text{ such that} \\ (q(t),v(t)) = \lim_{j\to+\infty} (q_{\delta_{k_j}}(t), v_{\delta_{k_j}}(t)) \text{ and } \eta(t) = \lim_{j\to\infty} \eta_{k_j}(t)\}.\end{aligned}$$

Thanks to [57, Theorem III.6], for every $t \in [0,T]$ we can select $(q(t),v(t)) \in B^{\mathcal{F}}_{C_1} \times B^{\mathcal{E}}_{C_2}$ such that $t \to (q(t),v(t))$ is measurable. Thus, (a) is proven. Finally, by repeating the arguments used in the proof of Proposition 4.6 we obtain (b).

Step 2: Proof of (c). Observe that, for every $t \in [0,T] \setminus \Lambda$,

$$\eta(t) = \limsup_{k\to\infty} \eta_k(t) = \lim_{j\to\infty} \eta_{k_j}(t) = \lim_{j\to\infty} \langle q_{\delta_{k_j}}(t), \dot f(t)\rangle_{\mathcal{F}} = \langle q(t), \dot f(t)\rangle_{\mathcal{F}}.$$

Let us now show that $\theta \in L^1(0,T)$. Since θ is the lim sup of measurable functions, we deduce that it is measurable. Moreover, we have

$$\int_0^T |\eta(t)|\,dt = \int_0^T |\langle q(t), \dot f(t)\rangle_{\mathcal{F}}|\,dt \le C_1 \int_0^T |\dot f(t)|_{\mathcal{F}}\,dt \le C_1\sqrt{T}\,\|\dot f\|_{L^2((0,T);\mathcal{F})}.$$

In order to get the energy inequality, recall that by (iii) of Theorem 4.11 we have, for every $j \in \mathbb{N}$,

$$\mathcal{J}(v_{\delta_{k_j}}(t)) \leq \mathcal{J}(v_0) + \int_0^t \langle q_{\delta_{k_j}}(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds + C_3\, \delta_{k_j}^{1/2}.$$

Taking the lim sup in j of the previous expression, using Fatou's Lemma

$$\begin{aligned}\mathcal{J}(v(t)) = \lim_{j\to\infty} \mathcal{J}(v_{\delta_{k_j}}(t)) &\leq \mathcal{J}(v_0) + \limsup_{j\to\infty} \int_0^t \langle q_{\delta_{k_j}}(s), \dot{f}(s)\rangle\, ds \\ &\leq \mathcal{J}(v_0) + \limsup_{k\to\infty} \int_0^t \langle q_{\delta_k}(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds \leq \mathcal{J}(v_0) + \int_0^t \limsup_{k\to\infty} \langle q_{\delta_k}(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds \\ &= \mathcal{J}(v_0) + \int_0^t \langle q(s), \dot{f}(s)\rangle_{\mathcal{F}}\, ds,\end{aligned}$$

so that (c) follows. □

Remark 4.14. *A careful inspection of the proof of Theorem 4.5 shows that*

$$|q(t)|_{\mathcal{E}} \leq C_1 \quad \text{and} \quad |v(t)|_{\mathcal{E}} \leq C_2 \qquad \text{for every } t \in [0,T],$$

where C_1 and C_2 are given by Theorem 4.11.

4.1.3. An application to a Barenblatt cohesive fracture model

In this section we introduce a specific Barenblatt cohesive fracture model that we use in the following to numerically support the theoretical results on approximable quasi-static evolution introduced in the previous section. In particular, we choose the fracture energy introduced by F. Cagnetti in [51] and we restrict to the case of the antiplanar shear. In the work cited above, the author studies the quasi-static evolution of the system along critical points of the energy functional which may not coincide with the global minimizer. Let us now introduce the setting and the notation we use in the following. The domain Ω is a bounded open set in $\mathbb{R}^d$ with Lipschitz boundary. We denote as usual with $u : \Omega \to \mathbb{R}$ the displacement function and we assume that the crack path in the reference configuration is contained in $\Gamma \cap \bar{\Omega}$, where $\Gamma \subset \mathbb{R}^d$ is a Lipschitz closed set such that $0 < \mathcal{H}^{d-1}(\Gamma \cap \bar{\Omega}) < \infty$ representing the prescribed crack path and $\Omega \setminus \Gamma = \Omega^+ \cup \Omega^-$, with Ω^+ and Ω^- disjoint open connected sets with Lipschitz boundary. In the following and exclusively for simplicity of definition of a space discretization, we shall consider Ω to be a cube in $\mathbb{R}^d$ for $d = 1$ and $d = 2$. We shall also consider a time dependent displacement $g \in H^1((0,T), H^1(\Omega))$ imposed on a fixed portion $\partial_D\Omega$ of the boundary of Ω, assumed to be well-separated from Γ and with intersections with $\partial\Omega^+$ and $\partial\Omega^-$ of positive $(d-1)$-dimensional measure.
We recall that the energy of the system related to cohesive fracture models is of the type

$$E(u) = \frac{1}{2}\int_{\Omega\setminus\Gamma} |\nabla u|^2\, dx + \kappa \int_\Gamma \theta(|u^+ - u^-|) d\mathcal{H}^{d-1}, \tag{4.25}$$

where κ is a material parameter and we set it to 1 in the following theoretical analysis. In particular, the displacement function $u \in H^1(\Omega \setminus \Gamma)$, while u^+ and u^- denote the trace on Γ of the restriction of u respectively to Ω^+ and to Ω^-, and $\theta : [0, \infty) \to [0, \infty)$ is the cohesive fracture function. Notice that, to ease the notation, we substitute $u^+ - u^-$ with the shorter form $[\![u]\!]$ if the quantity of interest is the jump amplitude. In particular, we recall that Barenblatt models prescribe in general, θ to be a $\mathcal{C}^1$, nondecreasing, bounded, concave fracture function with activation energy $\theta_a = 0$ and $\sigma := \theta'(0^+) \in (0, +\infty)$. For the specific model we choose, the function θ is defined as follows

$$\theta(s) = \begin{cases} -\dfrac{s^2}{2R} + s & \text{if } 0 \le s < R, \\ \dfrac{R}{2} & \text{if } s \ge R, \end{cases} \tag{4.26}$$

where $R > 0$ is a parameter which influences the crack formation mode.
Since we are interested in the study of a quasi-static evolution, we need to prove that an evolution of the system along configurations, which are also critical points of the energy functional, exists. One of the techniques used to prove its existence is called *singular perturbation method*. Let us introduce a generic time-dependent function $F(u, t)$ defined for u in a Banach space X for $t \in [0, T]$. An evolution of critical points is a function $u : [0, T] \to X$ satisfying

$$0 \in \partial_u F(u(t), t), \quad \text{for a.e. } t \in [0, T]. \tag{4.27}$$

Such evolution exists if for every $\varepsilon > 0$ it is possible to introduce an ε-gradient flow

$$-\varepsilon \dot{u}^\varepsilon \in \partial_u F(u^\varepsilon(t), t) \tag{4.28}$$

with initial datum $u^\varepsilon(t) = u_0$ where u_0 is a critical point of $F(\cdot, 0)$ and under suitable regularity assumptions, as $\varepsilon \to 0$, the solutions u^ε converge (in a sense to be specified) to a function u such that (4.27).
This method specifies for the case of cohesive fracture models as follows. We shall apply the previous scheme for $X = L^2(\Omega)$ and

$$F(u, t) = \begin{cases} E(u) & \text{for } u \in H^1(\Omega \setminus \Gamma), u = g(t) \text{ on } \partial_D \Omega, \\ +\infty & \text{otherwise in } L^2(\Omega), \end{cases}$$

where E is the functional in (4.25). We start observing that a minimizer of (4.25) at the time t is a weak solution ([51, Proposition 3.2]) of

$$\begin{cases} \Delta u(t) = 0 & \text{in } \Omega \setminus \Gamma, \\ u(t) = g(t) & \text{on } \partial_D \Omega, \\ \partial_\nu u(t) = 0 & \text{on } \partial\Omega \setminus \partial_D \Omega, \\ \partial_\nu u^+(t) = \partial_\nu u^-(t) & \text{on } \Gamma, \\ |\partial_\nu u| \le 1 & \text{on } \Gamma \setminus J_{u(t)}, \\ \partial_\nu u = \theta'([\![u(t)]\!]) \operatorname{sign}([\![u(t)]\!]) & \text{on } J_{u(t)}, \end{cases} \tag{4.29}$$

where ν denotes the inner unit normal to Ω, to Ω^+, and to Ω^-, and $J_u := \{\mathbf{x} \in \Gamma : [\![u(\mathbf{x})]\!] \neq 0\}$. Let u_0 be a critical point of (4.25) at the time $t = 0$. It turns out that a solution u^ε of (4.28) is given by a weak solution $u^\varepsilon \in H^1((0,T); L^2(\Omega)) \cap L^\infty(((0,T); H^1(\Omega \setminus \Gamma))$ of

$$\begin{cases} \Delta u^\varepsilon(t) = \varepsilon \dot{u}^\varepsilon(t) & \text{in } \Omega \setminus \Gamma, \\ u^\varepsilon(t) = g(t) & \text{on } \partial_D \Omega, \\ \partial_\nu u^\varepsilon(t) = 0 & \text{on } \partial\Omega \setminus \partial_D \Omega, \\ \partial_\nu u^\varepsilon(t)_{|\Omega^+} = \partial_\nu u^\varepsilon_{|\Omega^+}(t) & \text{on } \Gamma, \\ |\partial_\nu u^\varepsilon| \leq 1 & \text{on } \Gamma \setminus J_{u^\varepsilon(t)}, \\ \partial_\nu u^\varepsilon = \theta'([\![u^\varepsilon(t)]\!]) \operatorname{sign}([\![u^\varepsilon(t)]\!]) & \text{on } J_{u^\varepsilon(t)}. \end{cases} \tag{4.30}$$

The existence of a solution of (4.30) is proved ([51, Theorem 4.8]) by time discretization, solving suitable incremental minimum problems. Uniqueness is proven for sufficiently smooth θ. In [51, Proposition 4.13] it is shown that given a family $\{u^\varepsilon : \varepsilon \in (0,1)\}$ of solutions of (4.30) with initial condition u_0 and a boundary datum g, there exists a bounded measurable $u : [0,T] \to H^1(\Omega \setminus \Gamma)$ with $u(0) = u_0$ such that the following properties hold

- *approximability*: for every $t \in [0,T]$ there exists a sequence $\varepsilon_n \to 0$ such that
$$u^{\varepsilon_n}(t) \rightharpoonup u(t) \text{ weakly in } H^1(\Omega \setminus \Gamma);$$
- *stationarity*: for a.e. $t \in [0,T]$ the function $u(t)$ is a critical point for E at the time t, in particular, $u(t)$ is weak solution of (4.29) in the following sense: $u(t) = g(t)$ on $\partial_D\Omega$ and
$$\int_{\Omega\setminus\Gamma} \nabla u \cdot \nabla \psi dx + \int_\Gamma \left([\![\psi]\!] \theta'(|[\![u]\!]|) \operatorname{sign}([\![u]\!]) 1_{J_{u(t)}} + |[\![\psi]\!]| 1_{J^c_{u(t)}} \right) d\mathcal{H}^{d-1} \geq 0, \tag{4.31}$$
for all $\psi \in H^1_0(\Omega \setminus \Gamma, \partial_D\Omega)$;
- *energy inequality*: for every $t \in [0,T]$
$$E(u(t)) \leq E(u(0)) + \int_0^t \int_{\Omega\setminus\Gamma} \nabla u(s) \cdot \nabla \dot{g}(s) dx ds.$$

The discrete version of the model

In view of the applications of the theoretical results to the model just introduced, we need a finite dimensional version of the energy functional E in (4.25). To focus on the main ideas of our approach, we keep the formulation as clear as possible, considering a very simple geometry.
Let $d = 2$ and $\ell > 0$ be fixed, and define

$$\Omega := (0, 2\ell)^2, \qquad \Omega^- := (0,\ell) \times (0, 2\ell), \qquad \Omega^+ := (\ell, 2\ell) \times (0, 2\ell), \qquad \Gamma := \{\ell\} \times [0, 2\ell].$$

We will study a fracture evolution where the deformation is imposed on the set

$$\partial_D\Omega := (\{0\} \times [0, 2\ell]) \cup (\{2\ell\} \times [0, 2\ell]) .$$

Given a discretization parameter $N \in \mathbb{N}$, we set $h := \dfrac{\ell}{N}$, and we introduce a conform triangulation $\mathcal{T}_h$ of the set $\Omega \setminus \Gamma$, as in Figure 4.1.

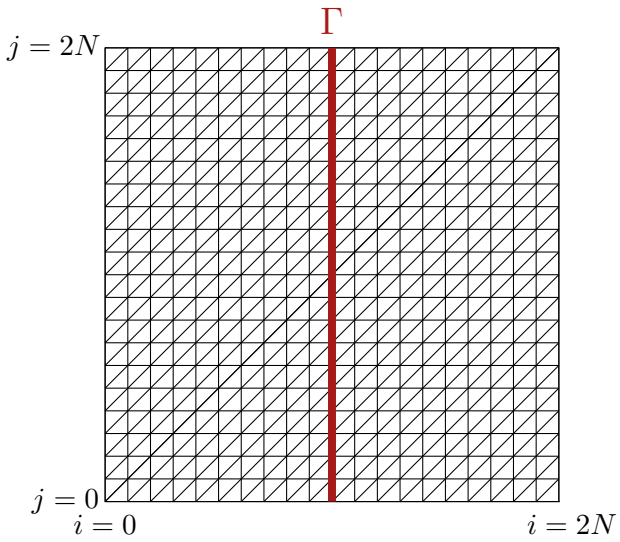


Figure 4.1.: Discrete geometry of the problem.

In particular, we define $\mathcal{E}_h$ as the finite dimensional space of continuous functions that are affine on each triangle belonging to $\mathcal{T}_h$. More precisely, we set

$$\mathcal{E}_h := \{u \in \mathcal{C}(\overline{\Omega} \setminus \Gamma) \cap H^1(\Omega \setminus \Gamma) : \quad \nabla_h u = \text{ const. a.e. on } T, \text{ for every } T \in \mathcal{T}_h\} ,$$

where the discrete gradient ∇_h is computed with a first order finite difference scheme [192]. Notice that the functions in the space $\mathcal{E}_h$ may be discontinuous in Γ and thus the interface between Ω^+ and Ω^- must be opportunely designed attributing two degrees of freedom to each point in Γ. Additionally, we define $\mathcal{E}_h^{reg}$ as the set of functions of $\mathcal{E}_h$ that do not jump across Γ:

$$\mathcal{E}_h^{reg} := \mathcal{E}_h \cap H^1(\Omega).$$

We endow $\mathcal{E}_h$ with the induced norm of $H^1(\Omega \setminus \Gamma)$

$$|u|^2_{\mathcal{E}_h} := \int_\Omega u^2 \, \mathrm{d}x + \int_{\Omega\setminus\Gamma} |\nabla u|^2 \, \mathrm{d}x \qquad\qquad u \in \mathcal{E}_h,$$

and we denote with $\langle u_1, u_2\rangle_{\mathcal{E}_h}$ the scalar product. Whenever we do not identify the dual space ${\mathcal{E}_h}'$ with $\mathcal{E}_h$, we will use the notation $\langle \cdot, \cdot \rangle_{{\mathcal{E}_h}',\mathcal{E}_h}$ for the duality pairing. Throughout

this section, we convene that the equality

$$\xi = v$$

where $\xi \in \mathcal{E}_h{}'$ and $v \in \mathcal{E}_h$, is meant in sense of the Riesz isometry. We define the discrete version of the energy functionals E and E_e respectively by

$$E_h := E\,|_{\mathcal{E}_h}, \qquad E_{e,h} := E_e\,|_{\mathcal{E}_h}\,.$$

We denote by A_h the operator which associates to every function of $\mathcal{E}_h$ its trace on $\partial_D\Omega$, and we set $\mathcal{F}_h := A_h(\mathcal{E}_h)$. Note that $\mathcal{F}_h$ is closed, since $\mathcal{E}_h$ is finite dimensional. Therefore, $\mathcal{F}_h$ endowed with the induced scalar product $\langle\cdot,\cdot\rangle_{\mathcal{F}_h}$ is a Hilbert subspace of $H^{1/2}(\partial_D\Omega)$.
We finally observe that, applying [217, Lemma 4.1.3] to our setting, we obtain the following version of Poincaré inequality:

$$\|u - u_D\|_{L^2(\Omega)} \le C\|\nabla u\|_{L^2(\Omega\setminus\Gamma)} \qquad \text{for every } u \in \mathcal{E}_h,$$

where the constant C depends on Ω and $\partial_D\Omega$, and

$$u_D := \left(\fint_{\partial_D\Omega} u^2\, d\mathcal{H}^1\right)^{1/2}.$$

The previous inequality in particular implies that

$$\|u\|_{L^2(\Omega)} \le u_D + C\|\nabla u\|_{L^2(\Omega\setminus\Gamma)} \le C\left(|A_h u|_{\mathcal{F}_h} + \|\nabla u\|_{L^2(\Omega\setminus\Gamma)}\right) \qquad \text{for every } u \in \mathcal{E}_h, \tag{4.32}$$

where with C we denote different constants, all depending on Ω and $\partial_D\Omega$. We conclude this subsection with an important remark that will be used later.

Remark 4.15. *Let $v \in \mathcal{E}_h$, $w \in \mathcal{E}_h{}^{reg}$, and let $\xi \in \partial E_h(v)$. Then, from the definition of subdifferential and direct computation, one can check that the action of ξ on w coincides with the action of the Fréchet differential $\partial E_{e,h}(v)$ on w. In formulas:*

$$\langle \xi, w\rangle_{\mathcal{E}_h{}',\mathcal{E}_h} = \langle \partial E_{e,h}(v), w\rangle_{\mathcal{E}_h{}',\mathcal{E}_h}.$$

We show now that the functional E_h satisfies the assumptions of Theorem 4.5.

Assumptions (A0)–(A3) are satisfied by E_h and A_h

First of all, we start by observing that condition (A0) is satisfied, by using standard arguments of calculus of variations.

Proposition 4.16. *The functional*

$$\mathcal{E}_h \ni v \longmapsto E_h(v) + |A_h v|^2_{\mathcal{E}_h}$$

is coercive.

Proof. Let $C > 0$ be fixed, and let $\{v_k\}_{k\in\mathbb{N}} \subset \mathcal{E}_h$ be a sequence such that

$$E_h(v_k) + |A_h v_k|^2_{\mathcal{E}_h} \leq C.$$

Then, recalling the expression of E_h and thanks to Poincaré inequality (4.32), we have

$$|v_k|^2_{\mathcal{E}_h} \leq C,$$

for some new constant, still denoted by C, depending on $\Omega, \partial_D\Omega$. Then, there exists a subsequence $\{v_{k_j}\}_{j\in\mathbb{N}}$ and a function $v \in \mathcal{E}_h$ such that

$$v_{k_j} \rightharpoonup v \qquad \text{weakly in } \mathcal{E}_h.$$

Since $\mathcal{E}_h$ is finite dimensional, this implies that

$$v_{k_j} \to v \qquad \text{in } \mathcal{E}_h,$$

and this concludes the proof. □

We now show that condition (A1) is satisfied.

Proposition 4.17. *There exists $\omega > 0$ such that the function*

$$u \longmapsto E_h(u) + \omega|u|^2_{\mathcal{E}_h}$$

is strictly convex.

Proof. We divide the proof into several steps.

Step 1. We show that there exists $\mu > 0$ such that the function $p_\mu : \mathbb{R} \to \mathbb{R}$ given by

$$p_\mu(s) := \theta(|s|) + \mu s^2, \qquad s \in \mathbb{R}, \tag{4.33}$$

is strictly convex. To this aim, we need to find μ such that the second distributional derivative p''_μ of p_μ is a positive Radon measure. Recalling the definition of θ, we have

$$p_\mu(s) = \begin{cases} |s| + \left(\mu - \dfrac{1}{2R}\right) s^2 & \text{if } 0 \leq |s| < R, \\ \dfrac{R}{2} + \mu s^2 & \text{if } |s| \geq R, \end{cases}$$

The distributional derivative p'_μ of p_μ is given by

$$p'_\mu(s) = \begin{cases} -1 + \left(2\mu - \dfrac{1}{R}\right) s & \text{if } -R < s < 0, \\ 1 + \left(2\mu - \dfrac{1}{R}\right) s & \text{if } 0 < s < R, \\ 2\mu s & \text{if } |s| \geq R. \end{cases}$$

Note that $p'_\mu \in L^1_{loc}(\mathbb{R})$. We can then calculate the second distributional derivative p''_μ of p_μ, which is the Radon measure in $\mathbb{R}$ given by

$$p''_\mu = \left(2\mu - \frac{1}{R}\right)\mathcal{L}^1\lfloor_{(-R,R)} + 2\mu\mathcal{L}^1\lfloor_{(-\infty,-R)\cup(R,\infty)} + 2\delta_0,$$

where δ_0 represents the Dirac measure concentrated at the origin. Note that

$$p''_\mu(B) \geq \left(2\mu - \frac{1}{R}\right)\mathcal{L}^1(B), \qquad \text{for every Borel set } B \subset \mathbb{R}.$$

Thus, if we choose μ such that

$$\mu > \frac{1}{2R}, \tag{4.34}$$

p''_μ is a positive Radon measure on $\mathbb{R}$, and p_μ is convex.

Step 2. We show that the functional $\overline{E}_h : \mathcal{E}_h \to [0,\infty)$ given by

$$\overline{E}_h(u) := E_h(u) + \mu \int_\Gamma |[\![u]\!]|^2 \, d\mathcal{H}^1,$$

is convex. By the previous step, the function $r_\mu : \mathcal{G}_h \to [0,\infty)$ defined as

$$r_\mu([\![u]\!]) := \int_\Gamma \theta(|[\![u]\!]|) \, d\mathcal{H}^1 + \mu \int_\Gamma |[\![u]\!]|^2 \, d\mathcal{H}^1$$

is convex, where $\mathcal{G}_h$ is the subset of $L^2(\Gamma)$ given by

$$\mathcal{G}_h := \{[\![u]\!] : u \in \mathcal{E}_h\}.$$

Note now that

$$\overline{E}_h(u) = E_{e,h}(u) + r_\mu(|[\![u]\!]|),$$

where E_e is the elastic energy. From the fact that $E_{e,h} : \mathcal{E}_h \to [0,\infty)$ is convex, we then obtain that also $\overline{E}_h : \mathcal{E}_h \to [0,\infty)$ is convex.

Step 3: conclusion. By [51, Lemma 5.3], there exists a constant $\overline{C} > 0$ such that

$$\int_\Gamma |[\![u]\!]|^2 \, d\mathcal{H}^1 \leq \overline{C}\, |u|^2_{\mathcal{E}_h}. \tag{4.35}$$

Taking $\omega > \mu\,\overline{C}$ we have

$$E_h(u) + \omega|u|^2_{\mathcal{E}_h} = \underbrace{E_h(u) + \mu \int_\Gamma |[\![u]\!]|^2 \, dx_2}_{\overline{E}_h(u)} + \mu \underbrace{\left(\overline{C}\, |u|^2_{\mathcal{E}_h} - \int_\Gamma |[\![u]\!]|^2 \, dx_2\right)}_{\widetilde{E}_h(u)} + (\omega - \mu\,\overline{C})|u|^2_{\mathcal{E}_h}.$$

We have already proven that $\overline{E}_h$ is convex. On the other hand, $\widetilde{E}_h$ is a quadratic form which is positive semidefinite by (4.35), and thus is convex. Since the remaining term $(\omega - \mu\,\overline{C})|u|^2_{\mathcal{E}_h}$ is strictly convex, this concludes the proof of (A1). □

Before passing to the proof of (A2) we need some preliminary results. First, we make a few remarks on the regularity of the elastic part and on the crack part of the energy functional.

Remark 4.18. *Note that $E_{e,h} \in \mathcal{C}^{1,1}(\mathcal{E}_h)$. In particular, $\partial E_{e,h} : \mathcal{E}_h \to \mathcal{E}_h'$ is a single-valued Lipschitz function with Lipschitz constant 1. Indeed, we have*

$$\langle \partial E_{e,h}(w), v\rangle_{\mathcal{E}_h',\mathcal{E}_h} = \int_{\Omega\setminus\Gamma} \nabla w \cdot \nabla v \, dx \qquad \text{for every } w, v \in \mathcal{E}_h. \tag{4.36}$$

Then, for every $w_1, w_2, v \in \mathcal{E}_h$

$$\begin{aligned}
|\partial E_{e,h}(w_1) - \partial E_{e,h}(w_2)|_{\mathcal{E}_h'} &= \sup\left\{ \int_{\Omega\setminus\Gamma} (\nabla w_1 - \nabla w_2)\cdot \nabla v \, dx, v \in \mathcal{E}_h \text{ with } |v|_{\mathcal{E}_h} = 1 \right\} \\
&\le \sup\left\{ \|\nabla w_1 - \nabla w_2\|_{L^2(\Omega\setminus\Gamma)} \|\nabla v\|_{L^2(\Omega\setminus\Gamma)}, v \in \mathcal{E}_h \text{ with } |v|_{\mathcal{E}_h} = 1 \right\} \\
&\le \sup\left\{ |w_1 - w_2|_{\mathcal{E}_h} |v|_{\mathcal{E}_h}, \quad v \in \mathcal{E}_h \text{ with } |v|_{\mathcal{E}_h} = 1 \right\} \\
&\le |w_1 - w_2|_{\mathcal{E}_h}.
\end{aligned}$$

Since $\partial E_{e,h}(0) = 0$, this implies

$$|\partial E_{e,h}(w)|_{\mathcal{E}_h'} \le |w|_{\mathcal{E}_h}, \qquad \text{for every } w \in \mathcal{E}_h. \tag{4.37}$$

Remark 4.19. *From the previous remark, it also follows that*

$$|\partial E_{e,h}(w_1) - \partial E_{e,h}(w_2)|_{\mathcal{E}_h'} \le \|\nabla w_1 - \nabla w_2\|_{L^2(\Omega\setminus\Gamma)} \qquad \text{for every } w_1, w_2 \in \mathcal{E}_h.$$

Remark 4.20. *The functional $G_h : \mathcal{E}_h \to [0,\infty)$ defined as*

$$G_h(v) := \int_\Gamma \theta(|[\![v]\!]|)\, d\mathcal{H}^1.$$

is globally Lipschitz continuous. Indeed, for every $v_1, v_2 \in \mathcal{E}_h$ we have

$$\begin{aligned}
|G_h(v_1) - G_h(v_2)| &\le \int_\Gamma |\theta(|[\![v_1]\!]|) - \theta(|[\![v_2]\!]|)| \, d\mathcal{H}^1 \le \int_\Gamma ||[\![v_1]\!]| - |[\![v_2]\!]|| \, d\mathcal{H}^1 \\
&\le \int_\Gamma |[\![v_1]\!] - [\![v_2]\!]| \, d\mathcal{H}^1 = \int_\Gamma |[\![v_1 - v_2]\!]| \, d\mathcal{H}^1 \\
&\le \left(H^1(\Gamma)\right)^{1/2} \|[\![v_1 - v_2]\!]\|_{L^2(\Gamma)} \le \overline{C} \left(H^1(\Gamma)\right)^{1/2} |v_1 - v_2|_{\mathcal{E}_h},
\end{aligned}$$

where $\overline{C}$ is given by (4.35)*, and we used the fact that $\|\theta'\|_{L^\infty([0,\infty))} = 1$.*

Next proposition shows condition (A2).

Proposition 4.21. *E_h satisfies condition (A2).*

Proof. Note that

$$E_h(u) = E_{e,h}(u) + G_h(u),$$

where $G_h : \mathcal{E}_h \to [0, \infty)$ is defined in Remark 4.20. Let now $v \in \mathcal{E}$. By (A.8), every $\xi \in \partial \mathcal{J}(v)$ can be written as

$$\xi = \xi_1 + \xi_2,$$

where $\xi_1 \in \partial E_{e,h}(v)$, and $\xi_2 \in \partial G_h(v)$. By Remark 4.19 we have

$$|\xi_1|_{\mathcal{E}_h'} \leq \|\nabla v\|_{L^2(\Omega\setminus\Gamma)} \leq 1 + \|\nabla v\|^2_{L^2(\Omega\setminus\Gamma)} \leq 2(1 + E_h(v)).$$

On the other hand, thanks to Remark 4.20 G_h is globally Lipschitz continuous with Lipschitz constant $\overline{C}\left(H^1(\Gamma)\right)^{1/2}$. Therefore, by (4.4)

$$|\xi_2|_{\mathcal{E}_h'} \leq \overline{C}\left(H^1(\Gamma)\right)^{1/2}.$$

Thus,

$$|\xi|_{\mathcal{E}_h'} \leq |\xi_1|_{\mathcal{E}_h'} + |\xi_2|_{\mathcal{E}_h'} \leq 2E_h(v) + 2 + \overline{C}\left(H^1(\Gamma)\right)^{1/2}.$$

□

Next lemma will be used to prove (A3), and gives a bound on the norm of a regular critical point, in terms of its trace on $\partial_D\Omega$.

Lemma 4.22. *Let $w \in {\mathcal{E}_h}^{reg}$, let $f \in \mathcal{F}_h$ be such that $A_h w = f$, and suppose $\partial E_{e,h}(w) \in \mathrm{ran}(A^*)$. Then, there exists a positive constant $C = C(\Omega, \gamma)$ such that*

$$|w|_{\mathcal{E}_h} \leq C|f|_{\mathcal{F}_h}. \tag{4.38}$$

Proof. By (4.32) and (4.36)

$$|w|^2_{\mathcal{E}_h} = \|w\|^2_{L^2(\Omega)} + \|\nabla w\|^2_{L^2(\Omega\setminus\Gamma)} \leq C\left(|A_h w|^2_{\mathcal{F}_h} + \|\nabla w\|^2_{L^2(\Omega\setminus\Gamma)}\right) \tag{4.39}$$

$$= C\left[|f|^2_{\mathcal{F}_h} + \langle \partial E_{e,h}(w), w\rangle_{{\mathcal{E}_h}',\mathcal{E}_h}\right], \tag{4.40}$$

where C denotes different constants, depending only on Ω and $\partial_D\Omega$. Let now $q \in \mathcal{F}_h$ be such that $A_h^* q = \partial E_{e,h}(w)$. Then,

$$|\partial E_{e,h}(w)|_{\mathcal{E}_h'} = |A_h^* q|_{\mathcal{E}_h} \geq \gamma |q|_{\mathcal{F}_h}.$$

Thus, taking into account (4.37) we have

$$\begin{aligned}\langle \partial E_{e,h}(w), w\rangle_{{\mathcal{E}_h}',\mathcal{E}_h} = \langle A_h^* q, w\rangle_{\mathcal{E}_h} = \langle q, A_h w\rangle_{\mathcal{F}_h} = \langle q, f\rangle_{\mathcal{F}_h} \leq |q|_{\mathcal{F}_h}|f|_{\mathcal{F}_h} \\ \leq \frac{1}{\gamma}\,|\partial E_{e,h}(w)|_{\mathcal{E}_h'}\,|f|_{\mathcal{F}_h} \leq \frac{1}{\gamma}\,|w|_{\mathcal{E}_h}\,|f|_{\mathcal{F}_h}.\end{aligned}$$

Using (4.39), we obtain

$$|w|^2_{\mathcal{E}_h} \leq C\left[|f|^2_{\mathcal{F}_h} + \frac{1}{\gamma}\,|w|_{\mathcal{E}_h}\,|f|_{\mathcal{F}_h}\right].$$

From the previous relation, the thesis follows using Young inequality. □

We can finally prove (A3).

Proposition 4.23. *Let $v_1, v_2, \overline{v} \in \mathcal{E}_h$ and $q_1, q_2, f_1, f_2 \in \mathcal{F}_h$ be such that*

$$A_h v_i = f_i, \qquad A_h^* q_i \in \partial (E_h)_{\omega,\overline{v}}(v_i), \qquad i = 1, 2,$$

where ω is given by Proposition 4.17. Then, there exists $C > 0$, depending only on Ω and $\partial_D \Omega$ such that

$$\langle q_1 - q_2, A_h v_1 - A_h v_2 \rangle_{\mathcal{F}_h} \leq C |v_1 - v_2|_{\mathcal{E}_h} |A_h v_1 - A_h v_2|_{\mathcal{F}_h}.$$

Proof. Let w be the unique solution of the following minimization problem:

$$w = \operatorname*{arg\,min}_{v \in \mathcal{E}_h^{reg}} \{ E_{e,h}(v) : A_h(v) = f_1 - f_2 \}. \tag{4.41}$$

By Remark 4.15 we have

$$\langle \xi_i, w \rangle_{\mathcal{E}_h', \mathcal{E}_h} = \langle \partial E_{e,h}(v_i), w \rangle_{\mathcal{E}_h', \mathcal{E}_h} \qquad i = 1, 2.$$

Now, by definition of $(E_h)_{\omega,\overline{v}}$ and (A.8), there exist $\xi_i \in \partial E_h(v_i)$, with $i = 1, 2$ such that

$$A_h^* q_i - 2\omega(v_i - \overline{v}) = \xi_i, \qquad i = 1, 2.$$

Subtracting term by term we obtain

$$A_h^*(q_1 - q_2) - 2\omega(v_1 - v_2) = \xi_1 - \xi_2.$$

Thus, thanks to Remark 4.18

$$\begin{aligned}
&\langle q_1 - q_2, A_h v_1 - A_h v_2 \rangle_{\mathcal{F}_h} = \langle q_1 - q_2, f_1 - f_2 \rangle_{\mathcal{F}_h} = \langle q_1 - q_2, A_h(w) \rangle_{\mathcal{F}_h} \\
&= \langle A_h^*(q_1 - q_2), w \rangle_{\mathcal{E}_h} = \langle \xi_1 - \xi_2, w \rangle_{\mathcal{E}_h', \mathcal{E}_h} + 2\omega \langle v_1 - v_2, w \rangle_{\mathcal{E}_h} \\
&= \langle \partial E_{e,h}(v_1) - \partial E_{e,h}(v_2), w \rangle_{\mathcal{E}_h', \mathcal{E}_h} + 2\omega \langle v_1 - v_2, w \rangle_{\mathcal{E}_h} \\
&\leq (1 + 2\eta) |w|_{\mathcal{E}_h} |v_1 - v_2|_{\mathcal{E}_h} \leq C(1 + 2\eta) |f_1 - f_2|_{\mathcal{F}_h} |v_1 - v_2|_{\mathcal{E}_h},
\end{aligned}$$

where we also used the fact that w satisfies the assumptions of Lemma 4.22 with $f = f_1 - f_2$. □

4.1.4. Recovering an approximable quasi-static evolution

In this section, whose main contributors are F. Cagnetti and F. Solombrino, co-authors of [13], we show now that the existence of a quasi-static evolution for the functional E, in the sense of [51] and Section 4.1.3, can be recovered from a discrete quasi-static evolution for E_h, when the parameter h controlling the mesh size tends to 0. More precisely we can prove the following version of [51, Theorem 4.4], as a consequence of Theorem 4.5.

Theorem 4.24. *Let $g \in W^{1,2}([0,T], H^1(\Omega))$, and let u_0 be a critical point of E at time 0 with $u_0 = g(0)$ on $\partial_D\Omega$. Then, there exists a bounded measurable function $u : [0,T] \to H^1(\Omega \setminus \Gamma)$ with $u(0) = u_0$ such that the following properties are satisfied:*

(a) approximability*: for every $t \in [0,T]$ there exists a sequence $h_j \to 0^+$ such that*

$$u_{h_j}(t) \rightharpoonup u(t) \text{ weakly in } H^1(\Omega \setminus \Gamma)$$

where, for every $j \in \mathbb{N}$, u_{h_j} is an approximable quasi-static evolution of E_{h_j} with initial condition u_0 and constraint f;

(b) stationarity*: for a.e. $t \in [0,T]$ the function $u(t)$ is a critical point of E at time t;*

(c) energy inequality*: for every $t \in [0,T]$*

$$E(u(t)) \le E(u(0)) + \int_0^t \int_{\Omega\setminus\Gamma} \nabla u(s) \cdot \nabla \dot{g}(s)\, dxds. \tag{4.42}$$

Before proving Theorem 4.24, we need to introduce a finite dimensional setting. We set

$$D := \left\{ h > 0 : h = \frac{\ell}{N} \text{ for some } N \in \mathbb{N} \right\}.$$

By [199], a sequence $\{g_h\}_{h\in D} \subset W^{1,2}([0,T], H^1(\Omega))$ such that $g_h \in W^{1,2}([0,T], \mathcal{E}_h)$ for every $h \in D$ exists and

$$g_h \overset{h\to 0^+}{\longrightarrow} g \qquad \text{in } W^{1,2}([0,T], H^1(\Omega)). \tag{4.43}$$

For every $t \in [0,T]$, we define $f_h(t) := A_h g_h(t)$. Before applying Theorem 4.5 to E_h, we need the following remark.

Remark 4.25. *By a careful reading of Section 4.1.2, one can see that the proof of Theorem 4.5 can be repeated even if the initial condition v_0 is not a critical point of the energy functional. Thus, we can apply Theorem 4.5 obtaining, for every $h \in D$, a function $u_h : [0,T] \to H^1(\Omega \setminus \Gamma)$ such that*

$$u_h(0) = u_0, \qquad \text{and} \qquad u_h(t) \in \mathcal{E}_h \quad \text{for every } t \in (0,T], \tag{4.44}$$

and all the other properties of Theorem 4.5 are satisfied. Condition (a') of the next theorem has to be intended in this sense.

We now apply Theorem 4.5 to E_h, obtaining the following result.

Theorem 4.26. *Let $h \in D$ be fixed, let $g_h \in W^{1,2}([0,T]; \mathcal{E}_h)$ be given by (4.43), and let $u_0 \in H^1(\Omega\setminus\Gamma)$ be a critical point of E with boundary condition $u_0 = g(0)$ on $\partial_D\Omega$. Then, there exists a measurable bounded mapping $u_h \colon [0,T] \to H^1(\Omega \setminus \Gamma)$ satisfying (4.44) such that*

(a') $u_h(\cdot)$ *is an approximable quasi-static evolution with initial condition* u_0 *and constraint* f_h*;*

(b') *stationarity: for every* $t \in (0,T]$ *we have*

$$\int_{\Omega\setminus\Gamma} \nabla u_h \cdot \nabla\psi dx + \int_\Gamma \Big([\![\psi]\!]\theta'(|[\![u_h(t)]\!]|)\,\mathrm{sign}([\![u_h(t)]\!])1_{J_{u_h(t)}} + |[\![\psi]\!]|1_{J^c_{u_h(t)}} \Big)\, d\mathcal{H}^{d-1} \geq 0, \tag{4.45}$$

for all $\psi \in \mathcal{E}_h$ *with* $\psi = 0$ *on* $\partial_D\Omega$*.*

(c') energy inequality*: The function* $s \longmapsto \displaystyle\int_{\Omega\setminus\Gamma} \nabla u_h(s)\cdot\nabla\dot g_h(s)\,dx$ *belongs to* $L^1(0,T)$ *and*

$$E_h(u_h(t)) \leq E(u_0) + \int_0^t \int_{\Omega\setminus\Gamma} \nabla u_h(s)\cdot\nabla\dot g_h(s)\,dxds \qquad \text{for every } t\in[0,T].$$

(d') Uniform bound*: There exists a constant* $\overline{C}_2$*, independent of* h*, such that*

$$\|u_h(t)\|_{\mathcal{E}_h} \leq \overline{C}_2 \qquad \textit{for every } t\in[0,T]. \tag{4.46}$$

Proof. As proven in the previous subsection, assumptions (A0)–(A2) are satisfied. Therefore, properties (a)–(c) of Theorem 4.5 hold true. In particular, (a) implies (a'). By (b) of Theorem 4.5, there exists a bounded measurable function $q_h : (0,T] \to \mathcal{F}_h$ such that

$$A^* q_h(t) \in \partial E_h(u_h(t)) \qquad \text{for every } t\in(0,T]. \tag{4.47}$$

Let now $t \in (0,T]$ be fixed. Thanks to (4.5), we have

$$0 \leq \liminf_{\varepsilon\to 0^+} \frac{E_h(u_h(t)+\varepsilon w) - E_h(u_h(t))}{\varepsilon} \qquad \text{for every } w\in\ker(A_h^*).$$

A careful inspection of the proof of [51, Proposition 3.1] shows that last inequality implies (b'). Let us now show (c'). From (c) of Theorem 4.5, the function $s \mapsto \langle q_h(s), \dot f_h(s)\rangle_{\mathcal{F}_h}$ belongs to $L^1(0,T)$, and for every $t\in[0,T]$ we have

$$E_h(u_h(t)) \leq E(u_0) + \int_0^t \langle q_h(s), \dot f_h(s)\rangle_{\mathcal{F}_h}\, ds.$$

Recalling that $f_h(s) = A_h g_h(s)$ and that the linear operator A_h is independent of time, we have

$$\begin{aligned} E_h(u_h(t)) &\leq E(u_0) + \int_0^t \langle q_h(s), \dot f_h(s)\rangle_{\mathcal{F}_h}\, ds \\ &= E(u_0) + \int_0^t \langle q_h(s), A_h\dot g_h(s)\rangle_{\mathcal{F}_h}\, ds \\ &= E(u_0) + \int_0^t \langle A_h^* q_h(s), \dot g_h(s)\rangle_{\mathcal{E}_h}\, ds. \end{aligned}$$

By (4.47), for every $s \in (0,T)$ we have $A^*q_h(s) \in \partial E_h(u_h(s))$. Since $\dot g_h(s) \in \mathcal{E}_h^{\rm reg}$ for every $s \in (0,T)$, by Remark 4.15 we have

$$\langle A_h^* q_h(s), \dot g_h(s)\rangle_{\mathcal{E}_h} = \int_{\Omega\setminus\Gamma} \nabla u_h(s)\cdot\nabla \dot g_h(s)\,dx \qquad \text{for every } s\in(0,T).$$

Therefore,

$$E_h(u_h(t)) \le E(u_0) + \int_0^t \int_{\Omega\setminus\Gamma} \nabla u_h(s)\cdot\nabla \dot g_h(s)\,dx\,ds,$$

which gives (c). Finally, property (d') directly follows from Remark 4.14 and Remark 4.12. □

We can now pass to the limit as $h \to 0^+$.

Proof of Theorem 4.24. Let $t \in [0,T]$. We will use argument similar to those used in the proof of Theorem 4.5.

Step 1: Proof of (a) and (c).

First of all, we fix the subsequence $\{h_k\}_{k\in\mathbb{N}}$ given by

$$h_k := \frac{\ell}{2^k}, \qquad\qquad k\in\mathbb{N},$$

so that

$$\mathcal{E}_{h_l} \subset \mathcal{E}_{h_m} \quad \text{for every } m > l. \tag{4.48}$$

By (4.43), we have

$$\dot g_{h_k} \to \dot g \quad \text{strongly in } L^2([0,T];H^1(\Omega)) \quad \text{as } k\to\infty.$$

Thus, there exists a set $\Lambda_2 \subset [0,T]$ with $\mathcal{L}^1(\Lambda_2)=0$ such that $\dot g(t)$ is well defined for every $t\in[0,T]\setminus\Lambda$ and

$$\dot g_{h_k}(t) \to \dot g(t) \quad \text{strongly in } H^1(\Omega) \quad \text{for every } t\in[0,T]\setminus\Lambda \quad \text{as } k\to\infty. \tag{4.49}$$

For every $h\in D$, let $u_h : [0,T] \to H^1(\Omega\setminus\Gamma)$ be given by Theorem 4.26. We define

$$\eta_k(t) := \begin{cases} \displaystyle\int_{\Omega\setminus\Gamma} \nabla u_{h_k}(t)\cdot\nabla \dot g_{h_k}(t)\,dx & \text{for every } t\in[0,T]\setminus\Lambda,\\ 0 & \text{for every } t\in\Lambda, \end{cases}$$

and

$$\eta(t) := \limsup_{k\to\infty} \eta_k(t) \quad \text{for every } t\in[0,T].$$

By definition of θ, for every $t \in [0,T]$ we can extract a subsequence $\{h_{k_j}\}_{j\in\mathbb{N}}$ (possibly depending on t) such that

$$\eta(t) = \lim_{j\to\infty} \eta_{k_j}(t) \quad \text{for every } t \in [0,T].$$

By (4.46), for every $t \in [0,T]$ we can extract a further subsequence (not relabelled) such that

$$u_{h_{k_j}}(t) \rightharpoonup u(t) \quad \text{weakly in } H^1(\Omega\setminus\Gamma) \quad \text{as } j\to\infty. \tag{4.50}$$

for some $u(t) \in H^1(\Omega\setminus\Gamma)$ with $\|u(t)\|_{H^1(\Omega\setminus\Gamma)} \le \overline{C}_2$. By repeating what was done in the proof of Theorem 4.5, we can show that the subsequence $\{k_j\}_{j\in\mathbb{N}}$ can be chosen in such a way that the map $u : [0,T] \to \mathcal{H}^1(\Omega\setminus\Gamma)$ is measurable, and this shows (a).
Let us now show the energy inequality. By (4.49) and (4.50) we have that, for every $t \in [0,T]\setminus\Lambda$,

$$\eta(t) = \limsup_{k\to\infty} \eta_k(t) = \lim_{j\to\infty} \eta_{k_j}(t) = \lim_{j\to\infty} \int_{\Omega\setminus\Gamma} \nabla u_{h_{k_j}}(t)\cdot\nabla \dot g_{h_{k_j}}(t)\,dx = \int_{\Omega\setminus\Gamma} \nabla u(t)\cdot\nabla\dot g(t)\,dx.$$

In order to prove that $\theta \in L^1(0,T)$, we first observe that, θ is the lim sup of measurable functions, it is measurable. Moreover, we have

$$\begin{aligned}\int_0^T |\eta(t)|\,dt &= \int_0^T \left|\int_{\Omega\setminus\Gamma} \nabla u(t)\cdot\nabla\dot g(t)\,dx\right| dt \le \int_0^T \|\nabla u(t)\|_{L^2(\Omega\setminus\Gamma)}\|\nabla\dot g(t)\|_{L^2(\Omega\setminus\Gamma)}\,dt\\ &\le \overline{C}_2 \int_0^T \|\nabla\dot g(t)\|_{L^2(\Omega\setminus\Gamma)}\,dt \le \overline{C}_2\sqrt{T}\,\|\dot g\|_{L^2((0,T);H^1(\Omega))}.\end{aligned}$$

By (c') of Theorem 4.26 we have, for every $j\in\mathbb{N}$ and for every $t\in[0,T]$

$$\begin{aligned}E(u_{h_{k_j}}(t)) = E_{h_{k_j}}(u_{h_{k_j}}(t)) &\le E(u_0) + \int_0^t \int_{\Omega\setminus\Gamma} \nabla u_{h_{k_j}}(s)\cdot\nabla\dot g_{h_{k_j}}(s)\,dxds\\ &= E(u_0) + \int_0^t \int_{\Omega\setminus\Gamma} \nabla u_{h_{k_j}}(s)\cdot\nabla\dot g_{h_{k_j}}(s)\,dxds.\end{aligned}$$

Note that the energy $E(\cdot)$ is lower semicontinuos w.r.t. weak convergence in $H^1(\Omega\setminus\Gamma)$. Therefore, taking the limsup in j of the previous expression and using Fatou's Lemma

$$\begin{aligned}E(u(t)) &\le \liminf_{j\to\infty} E(u_{h_{k_j}}(t)) \le E(u_0) + \limsup_{j\to\infty} \int_0^t \int_{\Omega\setminus\Gamma} \nabla u_{h_{k_j}}(s)\cdot\nabla\dot g_{h_{k_j}}(s)\,dxds\\ &\le E(u_0) + \limsup_{k\to\infty} \int_0^t \int_{\Omega\setminus\Gamma} \nabla u_{h_k}(s)\cdot\nabla\dot g_{h_k}(s)\,dxds\\ &\le E(u_0) + \int_0^t \limsup_{k\to\infty} \int_{\Omega\setminus\Gamma} \nabla u_{h_k}(s)\cdot\nabla\dot g_{h_k}(s)\,dx\,ds\\ &= E(u_0) + \int_0^t \int_{\Omega\setminus\Gamma} \nabla u(s)\cdot\nabla\dot g(s)\,dx\,ds,\end{aligned}$$

so that (c) follows.
We finally prove the stationarity. Let $\psi \in H^1(\Omega \setminus \Gamma)$ with $\psi = 0$ on $\partial_D\Omega$. Then [199], we can find a sequence $\{\psi_{h_{k_j}}\}_{j\in\mathbb{N}}$ such that

$$\psi_{h_{k_j}} \to \psi \quad \text{strongly in } H^1(\Omega \setminus \Gamma) \quad \text{as } j \to \infty$$

and $\psi_{h_{k_j}} \in \mathcal{E}_{h_{k_j}}$ with $\psi_{h_{k_j}} = 0$ on $\partial_D\Omega$, for every $j \in \mathbb{N}$. Note that, by (4.48), we have

$$\psi_{h_{k_l}} \in \mathcal{E}_{h_{k_j}} \quad \text{with } \psi_{h_{k_l}} = 0 \text{ on } \partial_D\Omega \quad \text{for every } j > l.$$

Therefore, by (4.45)

$$\begin{aligned}\int_{\Omega\setminus\Gamma} &\nabla u_{h_{k_j}}(t) \cdot \nabla \psi_{h_{k_l}}\, dx \\ &\geq \int_\Gamma \left(-[\![\psi_{h_{k_l}}]\!]\theta'(|[\![u_{h_{k_j}}(t)]\!]|)\,\mathrm{sign}([\![u_{h_{k_j}}(t)]\!])1_{J_{u_{h_{k_j}}}(t)} - |[\![\psi_{h_{k_l}}]\!]|1_{J^c_{u_{h_{k_j}}}(t)} \right) d\mathcal{H}^{d-1},\end{aligned} \tag{4.51}$$

for every $j > l$. By (4.50) we have

$$\lim_{j\to\infty} \int_{\Omega\setminus\Gamma} \nabla u_{h_{k_j}}(t) \cdot \nabla \psi_{h_{k_l}}\, dx = \int_{\Omega\setminus\Gamma} \nabla u(t) \cdot \nabla \psi_{h_{k_l}}\, dx. \tag{4.52}$$

Define now, for every $t \in [0,T]$ and for every $j > l$, the function $f_j(t) : \Gamma \to \mathbb{R}$ as

$$f_j(t) := -[\![\psi_{h_{k_l}}]\!]\theta'(|[\![u_{h_{k_j}}(t)]\!]|)\,\mathrm{sign}([\![u_{h_{k_j}}(t)]\!])1_{J_{u_{h_{k_j}}}(t)} - |[\![\psi_{h_{k_l}}]\!]|1_{J^c_{u_{h_{k_j}}}(t)}.$$

We want to prove that for every $t \in [0,T]$

$$\liminf_{j\to\infty} f_j(t) \geq -[\![\psi_{h_{k_l}}]\!]\theta'(|[\![u(t)]\!]|)\,\mathrm{sign}([\![u(t)]\!])1_{J_{u(t)}} - |[\![\psi_{h_{k_l}}]\!]|1_{J^c_{u(t)}} \qquad \mathcal{H}^1\text{-a.e. in } \Gamma. \tag{4.53}$$

Up to extracting a further subsequence, we can assume that

$$\liminf_{j\to\infty} f_j(t) = \lim_{j\to\infty} f_j(t) \qquad \mathcal{H}^1\text{-a.e. in } \Gamma, \tag{4.54}$$

and

$$\lim_{j\to\infty} [\![u_{h_{k_j}}(t)]\!] = [\![u(t)]\!] \qquad \mathcal{H}^1\text{-a.e. in } \Gamma. \tag{4.55}$$

Now, let us fix $x \in \mathcal{J}_{u(t)}$ such that (4.54) and (4.55) hold true. Then, for $j \in \mathbb{N}$ large enough we have

$$x \in J_{u_{h_{k_j}}(t)} \qquad \text{and} \qquad \mathrm{sign}([\![u_{h_{k_j}}(t)]\!](x)) = \mathrm{sign}([\![u(t)]\!](x)).$$

Therefore,

$$\begin{aligned}
&\liminf_{j\to\infty} f_j(t)(x) = \lim_{j\to\infty} f_j(t)(x)\\
&= \lim_{j\to\infty} -[\![\psi_{h_{k_l}}]\!](x)\theta'(|[\![u_{h_{k_j}}(t)]\!](x)|)\,\mathrm{sign}([\![u_{h_{k_j}}(t)]\!](x))1_{J_{u_{h_{k_j}}(t)}}(x) - |[\![\psi_{h_{k_l}}]\!](x)|1_{J^c_{u_{h_{k_j}}(t)}}(x)\\
&= -[\![\psi_{h_{k_l}}]\!](x)\theta'(|[\![u(t)]\!](x)|)\,\mathrm{sign}([\![u(t)]\!](x))1_{J_{u(t)}}(x) - |[\![\psi_{h_{k_l}}]\!](x)|1_{J^c_{u(t)}}(x) \qquad (4.56)
\end{aligned}$$

for $\mathcal{H}^1$-a.e. $x \in \Gamma \cap J_{u(t)}$. If, instead, $x \in J^c_{u(t)}$, then recalling that $0 \le \theta' \le 1$ we have

$$\begin{aligned}
&\liminf_{j\to\infty} f_j(t)(x) = \lim_{j\to\infty} f_j(t)(x)\\
&= \lim_{j\to\infty} -[\![\psi_{h_{k_l}}]\!](x)\theta'(|[\![u_{h_{k_j}}(t)]\!](x)|)\,\mathrm{sign}([\![u_{h_{k_j}}(t)]\!](x))1_{J_{u_{h_{k_j}}(t)}}(x) - |[\![\psi_{h_{k_l}}]\!](x)|1_{J^c_{u_{h_{k_j}}(t)}}(x)\\
&\ge -|[\![\psi_{h_{k_l}}]\!](x)| = -|[\![\psi_{h_{k_l}}]\!](x)|1_{J^c_{u(t)}}(x). \qquad (4.57)
\end{aligned}$$

Combining (4.56) and (4.57) we obtain (4.53). Thanks to (4.52) and (4.53) we can pass to the limit in (4.51), obtaining

$$\begin{aligned}
&\int_{\Omega\setminus\Gamma} \nabla u(t)\cdot\nabla\psi_{h_{k_l}}\,dx = \lim_{j\to\infty}\int_{\Omega\setminus\Gamma} \nabla u_{h_{k_j}}(t)\cdot\nabla\psi_{h_{k_l}}\,dx\\
&\ge \liminf_{j\to\infty}\int_\Gamma \left(-[\![\psi_{h_{k_l}}]\!]\theta'(|[\![u_{h_{k_j}}(t)]\!]|)\,\mathrm{sign}([\![u_{h_{k_j}}(t)]\!])1_{J_{u_{h_{k_j}}(t)}} - |[\![\psi_{h_{k_l}}]\!]|1_{J^c_{u_{h_{k_j}}(t)}}\right)d\mathcal{H}^{d-1}\\
&\ge \int_\Gamma \liminf_{j\to\infty}\left(-[\![\psi_{h_{k_l}}]\!]\theta'(|[\![u_{h_{k_j}}(t)]\!]|)\,\mathrm{sign}([\![u_{h_{k_j}}(t)]\!])1_{J_{u_{h_{k_j}}(t)}} - |[\![\psi_{h_{k_l}}]\!]|1_{J^c_{u_{h_{k_j}}(t)}}\right)d\mathcal{H}^{d-1}\\
&\ge \int_\Gamma \left(-[\![\psi_{h_{k_l}}]\!]\theta'(|[\![u(t)]\!]|)\,\mathrm{sign}([\![u(t)]\!])1_{J_{u(t)}} - |[\![\psi_{h_{k_l}}]\!]|1_{J^c_{u(t)}}\right)d\mathcal{H}^{d-1}.
\end{aligned}$$

Finally, passing to the limit as $l\to\infty$ we have

$$\int_{\Omega\setminus\Gamma}\nabla u(t)\cdot\nabla\psi\,dx \ge \int_\Gamma \left(-[\![\psi]\!]\theta'(|[\![u(t)]\!]|)\,\mathrm{sign}([\![u(t)]\!])1_{J_{u(t)}} - |[\![\psi]\!]|1_{J^c_{u(t)}}\right)d\mathcal{H}^{d-1},$$

and we conclude.

□

In our specific case, by Proposition 4.17 and direct calculation we have that

$$\omega = \frac{1}{2R} + \max\{4, 4\sqrt{\ell}\}.$$

4.1.5. Numerical experiments

The scope of this section is to practically show that the procedure illustrated in the previous sections can be effectively implemented and produces the desired quasi-static evolution, according to the one described in [51]. We refer the reader to Section 4.1.3 for the notations used here.

Numerical simulations in one dimension

We first analyze the results obtained for a one dimensional problem, when $\Omega \subset \mathbb{R}$. Despite its simplicity, the one dimensional setting allows to give a detailed comparison between numerical results and analytic predictions, since in this case the explicit solutions of (4.29) are known. We consider the following geometry:

$$\Omega = [0, 2\ell], \qquad \ell = 0.5, \qquad \Gamma = \{\ell\}, \qquad \partial_D\Omega = \{0, 2\ell\}.$$

We follow the evolution in the time interval $[0, T] = [0, 1]$, and the external load applied to the endpoints $\partial_D\Omega = \{0, 2\ell\}$ is given by

$$g(t)(x) = 2(x - \ell)t, \qquad \text{for every } x \in [0, 1] \text{ and } t \in [0, 1].$$

We uniformly discretize the domain into $2N = 80$ intervals, so that the spatial discretization step is given by $h = \ell/N$. Finally, we choose a time step $\delta = 0.02$, so that the total evolution is concluded after 50 time steps. In our specific case, by Proposition 4.17 and a direct calculation we have that the parameter η in condition (A1) can be taken as

$$\omega = \frac{1}{2R} + \max\{4, 4\sqrt{\ell}\},$$

where the constant R is the one appearing in the definition of the function g, see (4.26). From a practical viewpoint, the computational time needed to solve the minimization problem (4.12) could grow without any control. Hence, for any $i \in \{0, \dots, 50\}$ and $j \in \mathbb{N}$ fixed, we stop the minimization loop as soon as $\|Av - g(i\delta)\| < 10^{-6}$. That is, v_j^i is chosen in such a way that $\|Av_j^i - g(i\delta)\| < 10^{-6}$. Instead, we stop the external loop (that is, the limit of v_j^i as $j \to \infty$), as soon as $\|v_j^i - v_{j-1}^i\| < 10^{-13}$. The main reason for these choices is that the quasi-static evolution generated by the algorithm is extremely sensitive to any perturbation. Thus, a larger time step δ, or a too badly approximated critical point at each time step, could lead to nonphysical results.
We observe that the analytic evolutions discussed in [51, Section 9] depend on the size of the parameter R. Therefore, in order to compare our results with those in [51], we distinguish two cases.

Case $R \geq 2\ell$. When R is chosen large with respect to the size 2ℓ of the elastic body, the evolution found by numerical simulations evolves along *global* minimizers of the energy, and we can observe the three phases of the cohesive fracture formation: non-fractured, pre-fractured (that is when the opening of the crack is smaller than R and cohesive forces appear), and completely fractured (when the opening of the crack is larger than R and the cohesive forces disappear), see Figure 4.2.
Note that in the time interval $[0, 0.5]$ the evolution follows the elastic deformation. After $t = 0.5$ a fracture appears, since the elastic deformation is not any more a critical point of the energy functional (see [51, Section 9]). Then, the pre-fracture phase starts, showing

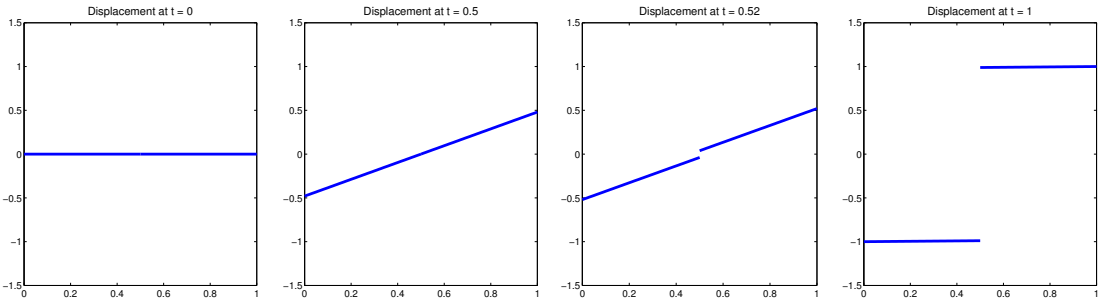


Figure 4.2.: The evolution of the quasi-static cohesive fracture for $R \geq 2\ell$ at time instances $t = 0, 0.5, 0.52, 1$.

a bridging force acting on the two lips of the crack. At time $t = 1$ the cohesive energy reaches its maximum, and the body is completely fractured. It is worth observing that this evolution coincides with the one analytically calculated in [51, Section 9].

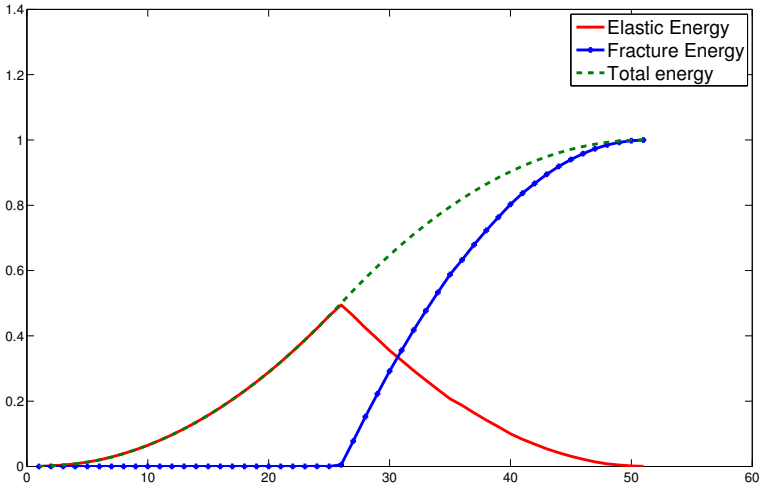


Figure 4.3.: The total, fracture, and elastic energy evolution of the quasi-static cohesive fracture for $R \geq 2\ell$.

We can also investigate what happens from the energy point of view, see Figure 4.3. We have a smooth transition between the different phases, and the total energy has a nondecresing profile. The beginning of the pre-fractured phase can be observed at the 25^{th} time step (i.e. at time $t = 0.5$), when the elastic energy (in red) starts decreasing and the crack energy (in blue) starts increasing. The final phase of complete rupture is then attained at the final time step $t = 1$. Although we focused on the time interval $[0, 1]$, one could check that the three energy profiles remain constant for $t > 1$.

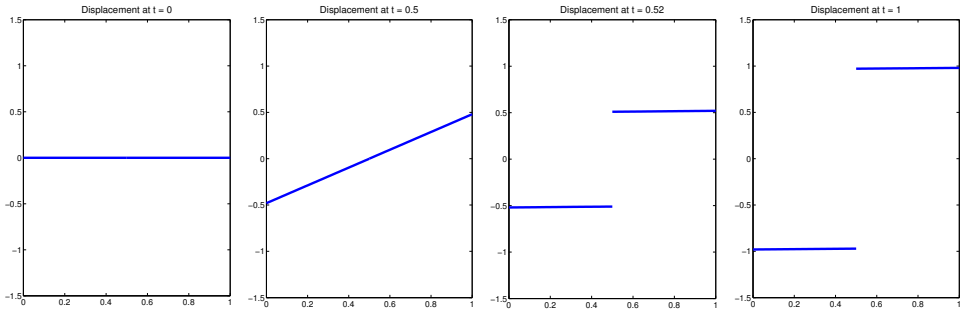


Figure 4.4.: The evolution of the quasi-static cohesive fracture for $R < 2\ell$ at time instances $t = 0, 0.5, 0.52, 1$.

Case $R < 2\ell$. The evolution of the system changes radically when $R < 2\ell$. In this case (see Figure 4.4) the failure happens instantaneously, without a bridging phase, and thus the body exhibits what in literature is known as *brittle* behavior. More precisely, in the time interval $[0, 0.5]$ the evolution follows again the elastic deformation, and a crack appears at $t = 0.5$. However, immediately after $t = 0.5$ the body is completely fractured, and no cohesive forces appear. It is important to observe that in this case we actually observe an evolution along critical points that are *not global minimizers*. Indeed, the evolution is elastic until $t = 0.5$, although it would be energetically convenient to completely break the body at some earlier time $\bar{t} < 0.5$ (see [51, Section 9] for a detailed description of all critical points). Thus, we see that the algorithm chooses the critical point which is the closest to the initial configuration, even if other options are available, which are more convenient from an energetic point of view. This evolution is particularly supported by the idea that in nature a body does not completely change its configuration crossing high energetic barriers if a stable configuration can be found with less energetic effort.

Also in this case, we can observe the evolution from the energetic viewpoint, see Figure 4.5. At time $t = 0.5$, when the elastic deformation ceases to be a critical point, the domain breaks and the total energy decreases up to the value of $R/2$, so that no bridging force is keeping the two lips together. As we already observed, the evolution along global minimizers would instead lead to a fracture way before the critical load is reached.

Again, the evolution found with our numerical simulation coincides with that one given in [51, Section 9]. In particular, our simulations agree with the *crack initiation criterion* (see [51, Theorem 4.6]), which states that a crack appears only when the maximum sustainable stress along Γ is reached. In this case, this happens at $t = 0.5$, when the slope of the elastic evolution reaches the value $\theta'(0) = 1$.

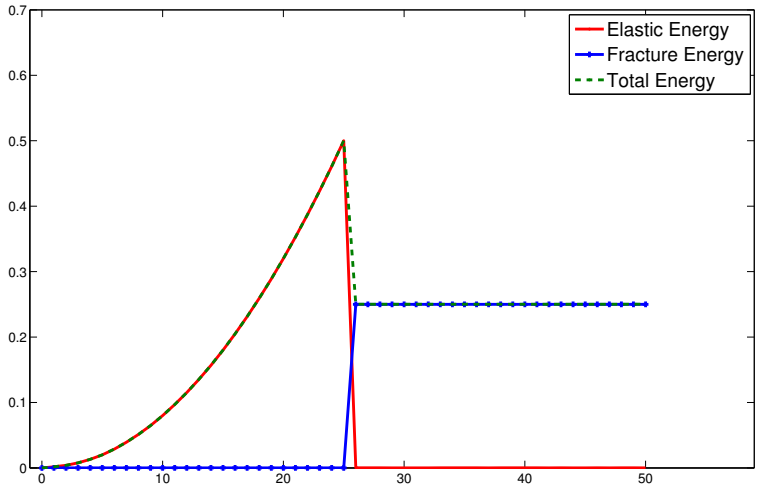


Figure 4.5.: The total, fracture, and elastic energy evolution of the quasi-static cohesive fracture for $R < 2\ell$.

Numerical simulations in two dimensions

Having a first analytic validation of the numerical minimization procedure, we can now challenge the algorithm in the simulation of two dimensional evolutions. We now consider the domain introduced in Section 4.1.3 setting $\ell = 0.5$, $2N = 8$, and $\kappa = 1/2$. Within this choice, the crack initiation time is reduced exactly of a factor $1/2$, allowing us to speed up the failure process. Since all the computations are performed on a MacBook Pro equipped with a 2.6GHz Intel Core i7 processor, 8GB of RAM, 1600MHz DDR3, the two dimensional simulations are performed only for a qualitative purpose. Indeed, we are mainly interested in showing that our algorithm produces physically sound evolutions also in dimension 2, and when the external displacement g is non-trivial. The very sparse discretization of the domain Ω is due to the fact that the minimization in (4.12) requires a huge computational effort, both in terms of time and memory. Indeed, in order to implement more realistic experiments, with a finer discretization, we would need to modify the architecture of the minimization algorithm, in such a way that it may run on parallel cores.
We perform two different series of experiments, one with boundary datum

$$g_1(t)(\mathbf{x}) = 2(x_1 - \ell)t, \qquad \text{for every } t \in [0,1] \text{ and } \mathbf{x} \in \Omega,$$

see Figure 4.6, and the other one with boundary datum

$$g_2(t)(\mathbf{x}) = 2t\cos\left(2\frac{x_2 - \ell}{\ell}\right)(x_1 - \ell), \qquad \text{for every } t \in [0,1] \text{ and } \mathbf{x} \in \Omega,$$

see Figure 4.7. Here, we denoted by $\mathbf{x} = (x_1, x_2)$ the generic point of $\Omega = (0,1) \times (0,1)$.

We now need to reduce the tolerance of the termination condition of the outer loop of the Algorithm, setting it to $5 \cdot 10^{-14}$. Indeed, we experimented that for bigger values of this tolerance some instabilities in the solution were introduced, leading to an asymmetric evolution, also in the case of g_1 as external displacement, where we expect an invariant behavior with respect to the space variable x_2.

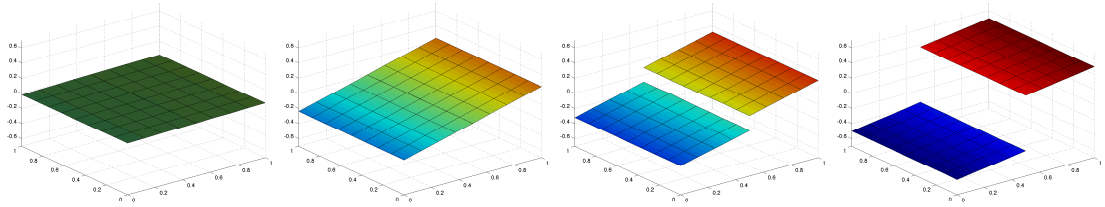


Figure 4.6.: The evolution of the quasi-static cohesive fracture for $R \geq 2\ell$ at time instances $t = 0, 0.24, 0.32, 0.5$ with external displacement g_1.

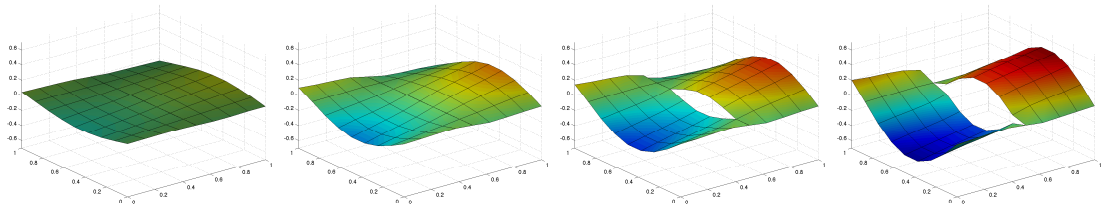


Figure 4.7.: The evolution of the quasi-static cohesive fracture for $R \geq 2\ell$ at time instances $t = 0.1, 0.24, 0.34, 0.5$ with external displacement g_2.

Case $R \geq 2\ell$. In Figure 4.6 and 4.7 we report 4 different instances of the evolution for the two different boundary data, when $R \geq 2\ell$. When the external displacement is g_1, which is constant with respect to the second coordinate x_2, we observe that the evolution is also constant with respect to x_2. For both boundary data, the failure of the body undergoes the three phases of deformation, as it happened in the one dimensional case.

Case $R < 2\ell$. When the boundary daum is g_1, see Figure 4.8, the specimen breaks in a brittle fashion, without showing any cohesive intermediate phase. This simulation is actually an evidence that the algorithm still characterizes the correct critical points, following the principle that the domain should not fracture as long as a non-fractured configuration is still a critical point. We conclude commenting the simulation where the boundary datum is g_2 with $R < 2\ell$, see Figure 4.9. By setting a displacement highly varying with respect to the x_2 coordinate, we observe that the different phases of the fracture formation can coexist. At time $t = 0.24$ the domain still presents no fracture, as expected by the previous numerical experiments. Then, at $t = 0.34$, a pre-fracture

appears, but only at those points where the external load is bigger, i.e. around $x_2 = \ell$. In fact, even at the final time $t = 1$, the domain is not completely fractured.
Note that, when the boundary datum is g_1, the evolution coincides with the one obtained analytically [51, Section 9]. In particular, the fracture appears at $t = 0.25$, when the slope of the elastic evolution reaches the value $\kappa\theta'(0) = 1/2$ and thus the *crack initiation criterion is satisfied*.

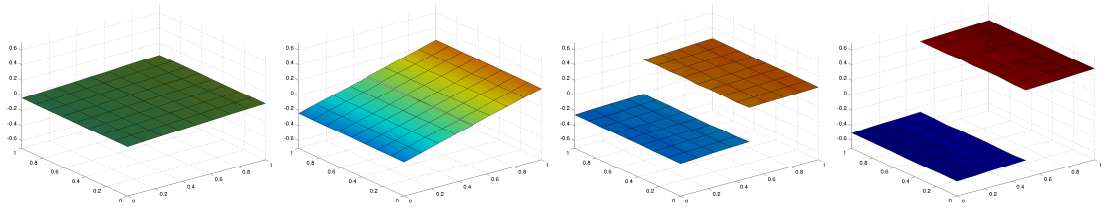


Figure 4.8.: The evolution of the quasi-static cohesive fracture for $R < 2\ell$ at time instances $t = 0.04, 0.24, 0.26, 0.5$ with external displacement g_1.

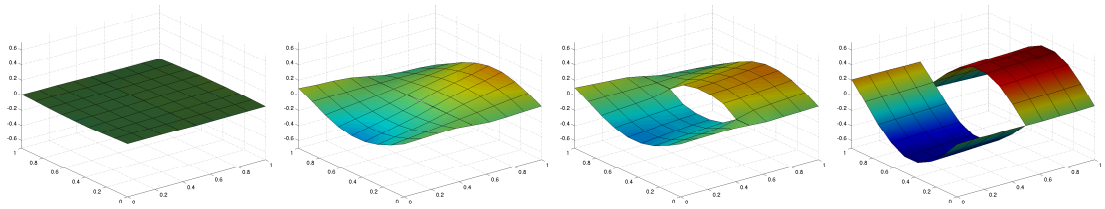


Figure 4.9.: The evolution of the quasi-static cohesive fracture for $R < 2\ell$ at time instances $t = 0.02, 0.22, 0.24, 0.5$ with external displacement g_2.

4.2. The simulation of quasi-static evolution of a Dugdale cohesive model

In this section we focus on the cohesive fracture model proposed by D. S. Dugdale in [96]. The model designed by the American scientist differs from the one proposed by G. I. Barenblatt for the regularity of the cohesive function θ and consequently for the mode of transition between the three phases of the cohesive fracture formation. By looking at Figure 1.1 we notice that the main difference with the model considered in the previous section is the evolution of the bridging force, which is given by $\theta'(|[\![u]\!]|)$. In this case, indeed, the cohesive attraction between the lips of the crack keeps constant with respect to the jump amplitude until the rupture value θ_r is reached. At this point the bridging force vanishes instantaneously and the domain becomes completely fractured.
Let us once again introduce all the elements that we need for the rest of the section.

$\Omega \subset \mathbb{R}^d$, with $d = 1, 2$, is a Lipschitz domain and $u : \Omega \to \mathbb{R}$, is the displacement function, recalling that we are in a off-plane setting. The deformation of the domain is driven by an external force which we express in term of an external displacement function $g : \Omega \times [0, T] \to \mathbb{R}$. Notice that we defined the function g on the whole domain since this property is required to check the energy inequality (1.16) and thus to prove the existence of quasi-static solution, see [51], but from a practical viewpoint we require that the displacement u coincides with the external deformation,

$$u|_{\partial\Omega_D} = g|_{\partial\Omega_D},$$

only on $\partial\Omega_D$, the subset of $\partial\Omega$, on which Dirichlet boundary conditions are prescribed. Also in this case we consider $\Gamma \subset \Omega$ to be the $(d-1)$-dimentional hypersurface in Ω prescribed crack path, assuming an a priori knowledge on where the specimen will break. Accordingly, the admissible configurations of the system are functions in $H^1(\Omega\backslash\Gamma)$. Since we are in a quasi-static setting, we introduce the usual time discretization $0 = t_0 < t_1 < \cdots < t_T = T$ and look for the equilibrium configurations which also are minimizers of the energy of the system. This means that for every $i \in \{0, \ldots, T\}$ we need to minimize the energy of the system

$$E(u) := \frac{1}{2}\int_{\Omega\backslash\Gamma} |\nabla u|^2 dx + \kappa \int_\Gamma \theta(|[\![u]\!]|)\, d\mathcal{H}^1, \tag{4.58}$$

with respect to a given boundary datum g

$$u^* \in \operatorname*{arg\,min}_{u = g(t_i)\, on\, \partial\Omega_D} E(u) \tag{4.59}$$

where $\kappa > 0$ is a material parameter.
In particular, the specific Dugdale cohesive fracture energy we choose is

$$\theta(s) = \min(s, 1). \tag{4.60}$$

In the following sections we introduce two different discretization schemes which may be adopted for a numerical implementation of the problem. We focus in particular only on the one dimensional problem because, thanks to its simplicity, it is possible to have a deeper and better understanding of the phenomenon we are describing. The modification of the procedure to higher dimensional problems requires indeed the addition of linear constraints and more complicated handling of the fracture set without increasing significantly the mathematical issues of the problem. In Section 4.2.2 we present two different and efficient algorithms to describe the quasi-static evolution of the system along minimizers of the energy functional. Finally, in Section 4.2.3 we present some numerical experiments showing the resulting evolutions given by the two algorithm presented. We will show that not only we obtain a physically sound solution, but also we outperform the computational times that we would obtain with Algorithm 2.2 introduced before.

4.2.1. Discretization schemes and optimality conditions

Let us assume that the one dimensional domain Ω is the segment $[0, 2\ell]$, with $\ell > 0$, and that $\Gamma = \{\ell\}$ is its midpoint. We need to discretize the domain subdividing it in $2N$ intervals of the same length to introduce a numerical scheme of the minimization of the energy functional. The Dirichlet datum is applied on $\Omega_D = \{0, 2\ell\}$ and the external displacement driving the quasi-static evolution is defined by

$$g(x,t) = xt. \tag{4.61}$$

As in the previous sections, we approximate the displacement function with a function u_h that is piecewise linear on $\Omega \setminus \Gamma$ and has two degrees of freedom on Γ to represent correctly the two lips of the fracture, denoting with u_N^- the degree on $\Omega^- := [0, \ell]$ and u_N^+ the one on $\Omega^+ := [\ell, 2\ell]$. The discrete energy functional corresponding to (4.58) is

$$E_h(u_h) = \frac{1}{2}\sum_{i=1}^{2N} \frac{N}{\ell}|u_i - u_{i-1}|^2 + \kappa\theta(|[\![u_N]\!]|), \tag{4.62}$$

where if $i \leq N$ we identify $u_N = u_N^-$ while for $i > N$, $u_N = u_N^+$. Notice that the jump of the displacement is not taken into account in the sum, and the gradient of u is approximated with finite differences of the first order.
Hence, to obtain a quasi-static evolution for the Dugdale cohesive model we need to solve the following minimization problem at each time step t_i, with $i \in \{1, \dots, T\}$,

$$u_h^* \in \underset{u=g \text{ on } \Omega_D}{\arg\min} E_h(u_h). \tag{4.63}$$

In order to solve the problem above, we follow two different approaches. In the first we make a change of variable rewriting the problem uniquely in terms of the displacement derivative $z_h = \nabla_h u_h$ while the second approach deals with the problem (4.63). The first ansatz eases sensibly the computation of the optimality conditions of problem (4.63) but, as a drawback, as soon as we consider problems in more than one dimension, additional linear constraints must be added to the minimization problem to guarantee the minimizer z_h^* corresponds to the derivative of a displacement functional u_h. Conversely, the direct approach requires a bigger effort from an analytic viewpoint, but avoids the numerical issues arising in the first approach.
Let us discuss in detail the change of variable. We set

$$\begin{aligned} z_i &= u_{i+1} - u_i, && \text{for } 0 \leq i < N \\ z_N &= u_N^+ - u_N^- \\ z_i &= u_i - u_{i-1}, && \text{for } N+1 \leq i \leq 2N, \end{aligned} \tag{4.64}$$

where we consider $u_N = u_N^-$ in the first line and $u_N = u_N^+$ in the third one. Notice that the vector $z_h \in \mathbb{R}^{2N+1}$ since $u_h \in \mathbb{R}^{2(N+1)}$ due to the presence of the two degrees of freedom in Γ. The energy functional (4.62) becomes

$$E_h(z_h) = \frac{1}{2} \sum_{\substack{i=0 \\ i \neq N}}^{2N} \frac{N}{\ell} |z_i|^2 + \kappa\theta(|z_N|). \tag{4.65}$$

We now rewrite the Dirichlet boundary conditions as a linear constraint depending on the new variable. Observing that

$$g(2\ell, \cdot) - g(0, \cdot) = u_{2N} - u_0 = \sum_{i=1}^{2N} (u_i - u_{i-1}) + (u_N^+ - u_N^-) = \sum_{i=0}^{2N} z_i,$$

we can define the time dependent linear-constraint's right-hand-side as $f(t) := g(2\ell, t) - g(0, t)$, and consequently the new minimization problem at time t_i, with $i \in \{0, \dots, T\}$, becomes

$$z_h^* \in \underset{a^t z_h = f(t_i)}{\arg\min}\ E_h(z_h), \tag{4.66}$$

where $a = (1 \cdots 1)^t$ is a vector in $\mathbb{R}^{2N+1}$. Unless differently specified, we work always with a fixed time step and since all the procedures described do not depend on the time, we drop the time dependence of the linear constraint f to ease the notation.
To enforce the linear constraint in the minimization process, we adopt the strategy of adding it to the energy functional as a penalization term. Hence, the new unconstrained minimization problem is

$$z_h^* \in \underset{z_h \in \mathbb{R}^{2N+1}}{\arg\min}\ E_{h,\gamma}(z_h) := \frac{1}{2} \|A z_h - f_h\|^2 + \kappa\theta(|z_N|), \tag{4.67}$$

where the linear operator $A : \mathbb{R}^{2N+1} \to \mathbb{R}^{2N+1}$ is

$$A = \begin{bmatrix} \frac{N}{\ell} I_N & \gamma a \end{bmatrix}^t, \tag{4.68}$$

where the matrix $I_N \in \mathbb{R}^{2N+1 \times 2N}$ is the identity that has the N-th column removed. The vector $f_h \in \mathbb{R}^{2N+1}$ then defined has

$$f_h = \left(\underbrace{0 \quad \cdots \quad 0}_{2N} \quad f \right)^t.$$

We can now discuss the optimality conditions for the problem (4.67).
First, notice that for the components z_i, with $i \neq N$, the fracture energy is not defined and therefore, for these components, the minimization problem reduces to finding the global minimizer of the elastic energy, which is convex and regular. Hence, we mainly

focus on the N-th component of z_h, where the fracture term plays a role and the functional is both nonconvex and nonsmooth.
Let us suppose $\bar{z}_h$ to be a solution of (4.67). Then, we define

$$\bar{f}_i := f - A\bar{z}_h + A_i\bar{z}_i \qquad i = 0, \dots, 2N$$

where A_i is the i-th column of the matrix A. Thanks to this definition, we can perform a component-wise minimization

$$z_i^* = \arg\min_{z_i \in \mathbb{R}} E_i(z_i) \qquad i = 0, \dots, 2N, \tag{4.69}$$

where the function $E_i : \mathbb{R} \to \mathbb{R}$ is defined as

$$E_i(z_i) := \begin{cases} \dfrac{1}{2}\|A_i z_i - \bar{f}_i\|^2 & i \neq N \\ \dfrac{1}{2}\|A_N z_N - \bar{f}_N\|^2 + \kappa\theta(|z_N|) & i = N \end{cases}$$

Proposition 4.27. *Let z_h, A_i, E_i,$\bar{f}_i$, and θ as already defined, with $i = 0, \dots, 2N$. Then, the optimality conditions for the problems* (4.69) *are*

$$\begin{cases} z_i = \dfrac{A_i^t \bar{f}_i}{|A_i|^2} & i \neq N \\ z_N = G(\bar{f}_N) \end{cases},$$

where

$$G(\bar{f}_N) = \begin{cases} \dfrac{A_N^t \bar{f}_N}{|A_N|^2} & A_N^t \bar{f}_N \geq |A_N|^2 + \dfrac{\kappa}{2} \\ \dfrac{A_N^t \bar{f}_N - \kappa}{|A_N|^2} & \kappa \leq A_N^t \bar{f}_N \leq |A_N|^2 + \dfrac{\kappa}{2} \\ 0 & A_N^t \bar{f}_N \leq \kappa \end{cases}.$$

Proof. In case $i \neq N$ the optimality conditions follow from direct computation since the functional for those components is differentiable and convex. Conversely, if $i = N$, the energy functional is nonconvex and not differentiable everywhere. Since $E_{h,\gamma}(z_N)$ is not differentiable only for $z_N = 0, 1$, we proceed computing $(E_{h,\gamma})'_N = 0$ where the functional is regular:

$$\begin{cases} A_N^t(A_N z_N - \bar{f}_N) = 0 & z_N > 1 \\ A_N^t(A_N z_N - \bar{f}_N) + \kappa = 0 & 0 < z_N < 1 \end{cases}.$$

From this, we get the following solutions

$$z_N = \begin{cases} \dfrac{A_N^t \bar{f}_N}{|A_N|^2} & A_N^t \bar{f}_N > |A_N|^2 \\ \dfrac{A_N^t \bar{f}_N - \kappa}{|A_N|^2} & \kappa < A_N^t \bar{f}_N < |A_N|^2 + \kappa \end{cases}.$$

Comparing the value of the energy functional in the points where the solution is not unique and in the points of non-differentiability we get that the global minimum of the energy functional in z_N is given by $G(\bar{f}_N)$. □

Remark 4.28. *Notice that, thanks to the change of variable, we minimize the energy of the system with respect to the gradient of the displacement. It is then necessary to reconstruct by integration the variable u_h. In case $d > 1$, not every solution of* (4.66) *corresponds to a gradient of a displacement function u_h. Therefore, an additional linear constraint which guarantees that the result of the minimization problem is also a derivative of a displacement vector must be added to the minimization problem. In particular, a sufficient condition for a given function z_h to be the discrete derivative of a function u_h is to be rotor-free, see [108] for details. Moreover, since on the fracture set Γ the displacement function u_h is non-regular, the above mentioned condition should not be enforced between the two lips of the fracture set.*

As we commented in Remark 4.28 and in the introduction of this section, it may be in practice not convenient to perform the change of variable (4.64) for $d > 1$ and with nontrivial geometries. Therefore, in the following we analyze the optimality condition also of the problem (4.63), where we directly minimize the displacement function. We reserve to the reader the possibility of choosing between the two approaches, accordingly to the one that best fits to his setting.

First, we notice that we can enforce the boundary condition via direct assignation of the values $u_0 = g(0)$ and $u_{2N} = g(2\ell)$. Then, introducing the finite difference operators $\overline{D}_e : \mathbb{R}^{2(N+1)} \to \mathbb{R}^{2N+1}$ for the calculation of the discrete elastic energy and $D_f : \mathbb{R}^{2(N+1)} \to \mathbb{R}$ to compute the jump between the two lips of the fracture, rescaling the first with respect to discretization step $\frac{\ell}{N}$, i.e. $D_e := \frac{\ell}{N}\overline{D}_e$, we can write the energy functional as

$$E_h(u_h) = \frac{1}{2}\frac{N}{\ell}\|D_e u_h - \tilde{g}\|^2 + \kappa\theta(|D_f u_h|), \tag{4.70}$$

where $\tilde{g}_0 = g(0, t_i)$, $\tilde{g}_{2N} = g(2\ell, t_i)$, with $i \in \{0, \dots, T\}$ (as before we drop the time dependence in the notation), while the remaining components of $\tilde{g}$ are set to zero.
Now, we study the optimality conditions for the problem

$$u_h^* \in \operatorname*{arg\,min}_{u_h \in \mathbb{R}^{2(N+1)}} \frac{1}{2}\frac{N}{\ell}\|D_e u_h - \tilde{g}\|^2 + \kappa\theta(|D_f u_h|). \tag{4.71}$$

Let $\bar{u}_h$ be a solution of the problem above. Then we can define

$$\begin{aligned}
\bar{g}_i &= \tilde{g} - D_e\bar{u}_h + (D_e)_i\bar{u}_i \qquad i = 0, \dots, 2N, i \neq N \\
\bar{g}_{N^-} &= \tilde{g} - D_e\bar{u}_h + (D_e)_{N^-}\bar{u}_N^- \\
\bar{g}_{N^+} &= \tilde{g} - D_e\bar{u}_h + (D_e)_{N^+}\bar{u}_N^+
\end{aligned}$$

where $(D_e)_i$ is the column of the matrix D_e which is coupled with u_i in the scalar product, and similarly $(D_e)_{N^\pm}$ are the columns coupled respectively with u_N^+ and u_N^-. Rewriting the problem (4.71) component-wise we obtain

$$u_i^* = \begin{cases} \min\limits_{u_i} \dfrac{1}{2}\dfrac{N}{\ell}|(D_e)_i u_i - \bar{g}_i|^2 & i \neq N, \\ \min\limits_{u_i} \dfrac{1}{2}\dfrac{N}{\ell}|(D_e)_i u_i - \bar{g}_i|^2 + \kappa\theta(|u_N^+ - u_N^-|) & i = N. \end{cases} \tag{4.72}$$

Notice that, since we set two degrees of freedom on the fracture set Γ, we need to solve in total $2N + 2$ problems: $2N$ are given by the first line of (4.72) and the last two, relative to the u_N^- and u_N^+ components are given by the second line on (4.72). To ease the notation, we use the index N^- to refer to the minimization problem to determine u_N^- and N^+ for the one relative to u_N^+.

Proposition 4.29. *Let u, $(D_e)_i$, $\bar{g}_i$, with $i = 0, \dots, 2N$, and θ as already defined, $\Lambda_N = (D_e)^t_{N^-}(D_e)_{N^-} = (D_e)^t_{N^+}(D_e)_{N^+}$, and $\tilde{\kappa} = \dfrac{\ell}{N}\kappa$. Then, the optimality conditions for the problem* (4.72) *are, for $i \neq N^\pm$,*

$$u_i = \frac{(D_e)_i^t \bar{g}_i}{|(D_e)_i|^2},$$

while for $i = N^\pm$

$$\begin{cases} u_i = \dfrac{(D_e)_i^t \bar{g}_i}{\Lambda_N} & |(D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-}| > \Lambda_N + \tilde{\kappa}, \\ u_{N^-} = \dfrac{(D_e)^t_{N^-}\bar{g}_{N^-} + \tilde{\kappa}}{\Lambda_N} & 2\tilde{\kappa} \leq |(D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-}| \leq \Lambda_N + \tilde{\kappa}, \\ u_{N^+} = \dfrac{(D_e)^t_{N^+}\bar{g}_{N^+} - \tilde{\kappa}}{\Lambda_N} & 2\tilde{\kappa} \leq |(D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-}| \leq \Lambda_N + \tilde{\kappa}, \\ u_i = \dfrac{(D_e)^t_{N^+}\bar{g}_{N^+} + (D_e)^t_{N^-}\bar{g}_{N^-}}{2\Lambda_N} & |(D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-}| < 2\tilde{\kappa}. \end{cases}$$

Proof. First, we focus on the case $i \neq N^\pm$. Notice that for these indexes the fracture energy is not defined and therefore the computation of the optimality conditions are trivially obtained using convex optimization arguments. Thus, we analyze in detail only the case where the functional is nonconvex, for $i = N^\pm$.
In the following, we will assume $u_N^+ \geq u_N^-$ just for sake of simplicity. The same procedure applies for $u_N^+ < u_N^-$ by changing opportunely the signs.
Taking the first derivative with respect to $u_N^\pm$ of the second equation of (4.72) where the

energy functional is regular enough, we obtain that $(u^*)_N^{\pm}$ satisfies

$$\begin{cases} \frac{N}{\ell}\left(\Lambda_N(u^*)_N^{\pm} - (D_e)_{N^{\pm}}^t \bar{g}_{N^{\pm}}\right) \pm \kappa = 0 & \text{if } 0 < |D_f u^*| < 1 \\ \Lambda_N(u^*)_N^{\pm} - (D_e)_{N^{\pm}}^t \bar{g}_{N^{\pm}} = 0 & \text{if } |D_f u^*| > 1 \end{cases} . \tag{4.73}$$

Considering the first equation of (4.73) for u_N^+ we have

$$u_N^+ = \frac{1}{\Lambda_N}(-\tilde{\kappa} + (D_e)_{N^+}^t \bar{g}_{N^+}). \tag{4.74}$$

Similarly, for u_N^- we get

$$u_N^- = \frac{1}{\Lambda_N}(\tilde{\kappa} + (D_e)_{N^-}^t \bar{g}_{N^-}). \tag{4.75}$$

Thus, we can compute the jump amplitude by subtracting (4.75) to (4.74), obtaining

$$D_f u = u_N^+ - u_N^- = \frac{1}{\Lambda_N}\left(-2\tilde{\kappa} + (D_e)_{N^+}^t \bar{g}_{N^+} - (D_e)_{N^-}^t \bar{g}_{N^-}\right).$$

We have that $D_f u_h \in (0,1)$ if and only if

$$0 < \frac{1}{\Lambda_N}\left(-2\tilde{\kappa} + (D_e)_{N^+}^t \bar{g}_{N^+} - (D_e)_{N^-}^t \bar{g}_{N^-}\right) < 1.$$

Doing some algebraic calculation we obtain

$$2\tilde{\kappa} < (D_e)_{N^+}^t \bar{g}_{N^+} - (D_e)_{N^-}^t \bar{g}_{N^-} < \Lambda_N + 2\tilde{\kappa}. \tag{4.76}$$

Notice that (4.74) and (4.75) are not necessary conditions for a global minimum. Indeed, they only characterize stationary points when (4.76) holds.
Doing the same steps for the second equation of (4.73), we find the following critical points

$$u_j = \frac{(D_e)_j^t \bar{f}_j}{\Lambda_N} \qquad \text{if } (D_e)_{N^+}^t \bar{g}_{N^+} - (D_e)_{N^-}^t \bar{g}_{N^-} > \Lambda_N \qquad j = N^{\pm}. \tag{4.77}$$

Now, we can proceed by direct comparison to select which one of the critical point is actually a minimizer of the energy functional. Notice that

$$\tilde{u}_N\pm = \frac{1}{\Lambda_N}(D_e)_{N^{\pm}}^t \bar{g}_{N^{\pm}} (D_e)_{N^{\pm}}$$

are respectively orthogonal projection of $\bar{g}_{N^+}$ and $\bar{g}_{N^-}$ onto $(D_e)_{N^+}$ and $(D_e)_{N^-}$. Let us now define

$$E_N = \frac{1}{2}\frac{N}{\ell}\|(D_e)_{N^-} u_N^- - \bar{g}_{N^-}\|^2 + \frac{1}{2}\frac{N}{\ell}\|(D_e)_{N^+} u_N^+ - \bar{g}_{N^+}\|^2 + \kappa\theta(|[\![u_N]\!]|), \tag{4.78}$$

the contribution to the energy functional given by the components $N^{\pm}$, and substitute first $u_N^{\pm}$ with the value of the critical points (4.74) and (4.75) determined by the first equation of (4.73). We get

$$E_{N,1} = \frac{1}{2}\frac{N}{\ell}\left(\frac{2\tilde{\kappa}^2}{\Lambda_N} + \|P^{\perp}_{(D_e)_{N^-}}\bar{g}_{N^-}\|^2 + \|P^{\perp}_{(D_e)_{N^+}}\bar{g}_{N^+}\|^2\right) + \kappa\frac{1}{\Lambda_N}\left(-2\tilde{\kappa} + (D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-}\right).$$

Now, analogously, we use the critical points (4.77) derived by the second equation of (4.73) obtaining

$$E_{N,2} = \frac{1}{2}\frac{N}{\ell}\left(\|P^{\perp}_{(D_e)_{N-1}}\bar{f}_{N-1}\|^2 + \|P^{\perp}_{(D_e)_N}\bar{f}_N\|^2\right) + \kappa.$$

Thus, we can determine which critical point found for the interval

$$\Lambda_N < (D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-} < \Lambda_N + 2\tilde{\kappa}$$

is also a minimum of the energy functional by comparing $E_{N,1}$ and $E_{N,2}$. Therefore, we can deduce that (4.74) and (4.75) are minimzers of the energy functional if and only if $E_{N,1} < E_{N,2}$, or equivalently

$$\frac{1}{2}\frac{N}{\ell}\left(\frac{2\tilde{\kappa}^2}{\Lambda_N} + \|P^{\perp}_{(D_e)_{N^-}}\bar{g}_{N^-}\|^2 + \|P^{\perp}_{(D_e)_{N^+}}\bar{g}_{N^+}\|^2\right) + \kappa\frac{1}{\Lambda_N}\left(-2\tilde{\kappa} + (D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-}\right) < \frac{1}{2}\frac{N}{\ell}\left(\|P^{\perp}_{(D_e)_{N^-}}\bar{g}_{N^-}\|^2 + \|P^{\perp}_{(D_e)_{N^+}}\bar{g}_{N^+}\|^2\right) + \kappa.$$

Subtracting the terms $P^{\perp}_{\cdot}$ and recalling the definition of $\tilde{\kappa}$, we obtain

$$(D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-} < \tilde{\kappa} + \Lambda_N.$$

So far, we have not analyzed the case $(D_e)^t_{N^+}\bar{g}_{N^+} - (D_e)^t_{N^-}\bar{g}_{N^-} < 2\tilde{\kappa}$ yet. The candidates critical points of the energy functional are the points where the fracture energy is not differentiable, i.e. $|D_f u_h| \in \{0, 1\}$. Since, for $|D_f u_h| = 1$, $\partial E_h(u_h) = \emptyset$, we find that the unique minimum is obtained for $|D_f u_h| = 0$ and thus $u_{N^+} = u_{N^-}$. □

4.2.2. Two efficient minimization algorithms

In this section we first show how the splitting approach described in Section 2.2 can by used for this specific problem. By adding a new variable we decompose the nonconvex minimization problem (4.63) into the coupling of a convex minimization and a function

assignment. This procedure not only finds the critical points we seek in a computationally efficient way, but it is also preparatory for the second algorithm we propose.
Exploiting the properties of the Lagrangian multiplier, we introduce the second procedure called Primal Dual Active Set (PDAS), see [135, 136, 139] for the general theory in the convex case and the proof of convergence in some particular cases. This ansatz removes the added variable keeping nonetheless the advantages of the splitting method but reducing sensibly the computational times, as it is shown in Section 4.2.3.
Since the total energy E_h is composed by two summands, the elastic energy, which is convex and smooth, and the fracture energy, which is a nonconvex and nonsmooth term, the use of the Split Bregman Iteration could be advantageous from a computational point of view. Indeed, since dealing with the convex term is not a demanding numerical issue and many different and efficient techniques can be applied, a good strategy is to isolate the nonconvex term and deal with it separately. For sake of simplicity we will present the two strategies only for the problem (4.66), since the procedure is the same also for (4.63), but, as we already have seen in the previous section, the notation might become complicated and confusing. We refer to [20], which is still in preparation, for all the details about the algorithms for problem (4.63).
Let us introduce the new variable $v \in \mathbb{R}$ which replaces z_N in the fracture energy

$$E_{h,\gamma}(z_h) = \frac{1}{2}\|Az_h - f_h\|^2 + \kappa\theta(|v|).$$

Now, in order to enforce the new variable to coincide with the jump in the displacement, we add a linear constraint to the minimization problem

$$(z_h^*, v^*) \in \underset{\substack{v=z_N \\ z_h\in\mathbb{R}^{2N+1}, v\in\mathbb{R}}}{\arg\min} \quad \frac{1}{2}\|Az_h - f_h\|^2 + \kappa\theta(|v|). \tag{4.79}$$

Additionally, for $\lambda \in \mathbb{R}$, the Lagrangian equation associated to (4.79) is

$$\mathcal{L}(z_h, v, \lambda) = \frac{1}{2}\|Az_h - f_h\|^2 + \kappa\theta(|v|) + \frac{A_N^t A_N}{2}|z_N - v|^2 + \lambda(z_N - v). \tag{4.80}$$

Since we are interested in finding critical points of the previous equation, we can proceed introducing Algorithm 2.3 for our specific problem and adding the termination condition $z_N = v$. This condition guarantees that the critical point obtained is also a critical point of the total energy $E_{h,\gamma}$. For a fixed time step t_i, with $i \in \{0, \dots, T\}$ the procedure is the following

Algorithm 4.1 Split Bregman Iteration to find and equilibrium configuration of (4.79) at time t_i

1: $\lambda^0 \in \mathbb{R}$, $v^0 \in \mathbb{R}$, $\mathbf{0} = z_h^0 \in \mathbb{R}^{2N+1}$;
2: Compute the forcing term f_h according to $g(t_i)$;
3: Initialize $k = 0$;

4: **while** $\|Az_h^k - f_h\| \neq 0$ and $|z_N^k - v^k| \neq 0$ **do**
5: $\quad z_h^{n+1} = \min_z \mathcal{L}(z, v^n, \lambda^n)$;
6: $\quad v^{n+1} = \min_v \mathcal{L}(z^{n+1}, v, \lambda^n)$;
7: $\quad \lambda^{n+1} = \lambda^n + A_N^t A_N (z_N^{n+1} - v^{n+1})$;
8: $\quad k \leftarrow k + 1$;
9: **end while**

Proposition 4.30. *The* `while`*-loop of Algorithm 4.1 can be equivalently written as*

$$\begin{cases} A^t(Az_h^{n+1} - f_h) + e_N \left(|A_N|^2 (z_N^{n+1} - v^n) + \lambda^n \right) = 0 \\ v^{n+1} = \phi(z_N^{n+1}, \lambda^n) \\ \lambda^{n+1} = \lambda^n + |A_N|^2 (z_N^{n+1} - v^{n+1}) \end{cases} \tag{4.81}$$

where e_N is the N-th vector of the canonical basis of $\mathbb{R}^{2N+1}$ and

$$\phi(z_N, \lambda) = \begin{cases} \dfrac{\lambda + |A_N|^2 z_N}{|A_N|^2} & \text{if } \lambda + |A_N|^2 z_N \geq |A_N|^2 + \dfrac{\kappa}{2}, \\ \dfrac{\lambda + |A_N|^2 z_N - \kappa}{|A_N|^2} & \text{if } \kappa < \lambda + |A_N|^2 z_N < |A_N|^2 + \dfrac{\kappa}{2}, \\ 0 & \text{if } \lambda + |A_N|^2 z_N < \kappa. \end{cases} \tag{4.82}$$

Proof. First, notice that the Lagrangian function is convex and smooth with respect to the variable z_h. Thus the problem of line 5 in Algorithm 4.1 has a unique solution given by the first equation of (4.81). Now, we need to focus only on the second minimization problem. Following the same approach to find the optimality conditions used in Proposition 4.27, we compute the first derivative of (4.80) with respect to v and then, by direct comparison, we can establish which of the critical points found is effectively a minimum of the functional. □

Remark 4.31. *For the minimization of the energy functional $E_{h,\gamma}$, we decided to use a more direct approach, avoiding the introduction of a quadratic perturbation as in Algorithm 2.4. Indeed, for this specific Dugdale cohesive fracture functional, the functional is not regular enough and we should have adopted a smoothing technique. As shown in Proposition 4.30, v is explicitly determined by ϕ and therefore no regularization is required. Although there is no proof of convergence for Algorithm 4.1 yet, we know that, if it converges, we retrieve a critical point of the objective functional and it has been successfully used in different applications [137].*

In Section 4.2.3 we show that this algorithm not only is, in practice, always converging, but also computationally really efficient. Despite the result already obtained would be

satisfying, we can extract additional information from the procedure described above in order to tackle the problem with an even more efficient approach: the PDAS.

Let us assume that Algorithm 4.1 converges and let v^* and z_h^* be its output. Since the termination condition has been satisfied we have that $v^* = z_N^*$. Hence, we can substitute z_N with v in (4.82) obtaining

$$\begin{cases} \lambda^* = 0 & \text{if } \lambda^* + |A_N|^2 z_N^* \geq |A_N|^2 + \dfrac{\kappa}{2}, \\ \lambda^* = \kappa & \text{if } \kappa < \lambda^* + |A_N|^2 z_N^* < |A_N|^2 + \dfrac{\kappa}{2}, \\ \begin{cases} |\lambda^*| \leq \kappa \\ z_N^* = 0 \end{cases} & \text{if } \lambda^* + |A_N|^2 z_N^* < \kappa, \end{cases} \tag{4.83}$$

while the first equation becomes

$$A^t(Az_h^* - f_h) + e_N \lambda^* = 0. \tag{4.84}$$

Now, the solution of the minimization problem (4.66) is given by the couple (z_h^*, λ^*) which is uniquely determined by solving (4.83) and (4.84). Notice that v^* is indeed not necessary to find the equilibrium configurations which describe the quasi-static evolution of the system. Inspired by this observation we can write the Primal Dual Active Set algorithm for to determine the quasi-static evolution of Dugdale cohesive fracture model. For each time step t_i, with $i \in \{0, \dots, T\}$ do:

Algorithm 4.2 Primal Dual Active Set to find and equilibrium configuration of (4.66) at time t_i

$\lambda^0 \in \mathbb{R}$, $\mathbf{0} = z_h^0 \in \mathbb{R}^{2N+1}$;
Compute the forcing term f_h according to $g(t_i)$;
Initialize $k = 0$;
while $(z^n, \lambda^n) \neq (z^{n-1}, \lambda^{n-1})$ **do**
    if $\lambda^n + |A_N|^2 z_N^n < \kappa$ **then**
        $z_N^{n+1} = 0$;
        Solve $A^t(Az^{n+1} - f_h) + e_N \lambda^{n+1} = 0$ for z_N^{n+1} fixed;
        to compute λ^{n+1} and the remaining $2N$ components of z_h^{n+1};
    else
        if $\kappa < \lambda^n + |A_N|^2 z_N^n < |A_N|^2 + \dfrac{\kappa}{2}$ **then**
            $\lambda^{n+1} = \kappa$;
        else
            $\lambda^{n+1} = 0$;
        end if

Solve $A^t(Az^{n+1} - f_h) + e_N\lambda^{n+1} = 0$ for λ^{n+1} fixed to compute z_h^{n+1};
end if
$k \leftarrow k + 1$;
end while

Unfortunately, in our specific case does not exist a proof of convergence. Indeed, some attempts have been done in [139], but convergence gurarantees are obtained only under stronger regularity assumptions, which our energy functional does not fulfill. In Section 4.2.3 we present some numerical tests where we show that in practice Algorithm 4.2 always converged to a minimizer of (4.66) requiring much less computational times that for Algorithm 4.1.

Remark 4.32. *Despite in this section we only refereed to the one-dimensional setting, the same procedure can be also applied in higher dimensions. Indeed, it is not strictly necessary that the variable v is a scalar and thus we may consider also lines and surfaces as fracture sets Γ respectively in the two and three-dimensional settings.*

4.2.3. Numerical experiments

In this section we present numerical experiments to validate the algorithms proposed in the last section. The characteristics that the procedures must exhibit in order to be considered as reliable are:

- convergence of the algorithm to a minimizer of the energy functional;
- the quasi-static evolution corresponds with the expectation and the energy is monotonically non decreasing in time;
- competitive computational times with respect to Algorithm 2.2.

Let us consider a pure shear off-plane deformation of a beam of length $2\ell = 1$ with a prescribed crack $\Gamma = \{0.5\}$. We discretize it by uniform partitioning into $2N$ intervals, with $N = 100$. The time step in the time discretization of the time window $[0, T]$, with $T = 3$, is set to $dt = 0.01$, and the external force $g : \mathbb{R} \times [0, T] \to \mathbb{R}$ monotonically increasing in time

$$g(\mathbf{x}, t) = xt.$$

The parameters of the energy functional $E_{h,\gamma}$ are set to $\gamma = 50$, $\kappa = 1$, and the tolerance for the stopping criterion of the algorithms is 10^{-15}.
In Figure 4.10, 4.11, and 4.12 we report three time frames of the evolution obtained with Algorithm 4.1.
The beam effectively evolves presenting the three phases that we expect from a cohesive fracture model:

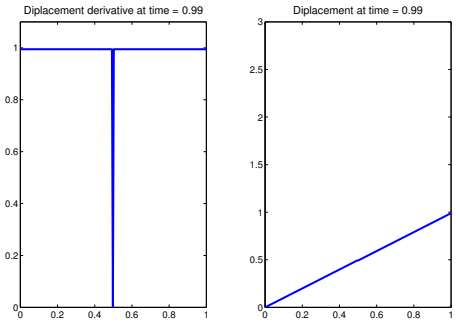


Figure 4.10.: A time step of the evolution of the pure off-plane shear of a Dugdale cohesive model for $t = 0.99$. On the left the gradient of the displacement ∇u and on the right the displacement u.

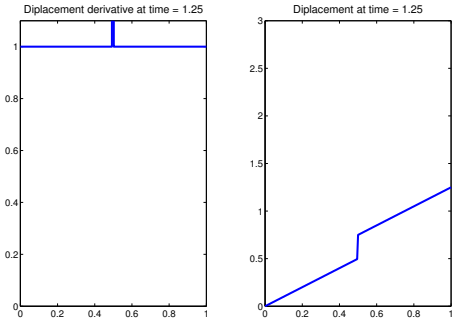


Figure 4.11.: A time step of the evolution of the pure off-plane shear of a Dugdale cohesive model for $t = 1.25$. On the left the gradient of the displacement ∇u and on the right the displacement u

- *Figure 4.10:* The beam is purely elastically deformed. Notice that the jump amplitude is zero and the gradient of the displacement is constant on $\Omega \setminus \Gamma$;
- *Figure 4.11:* The beam presents a pre-fracture. The elastic deformation is still present (the gradient of the displacement is constant and equal 1) but the two lips of the fracture do not touch each-other any longer;
- *Figure 4.12:* The beam is fractured. The bridging force is vanished and the two parts of the beam are free to move. The elastic energy is zero.

Looking at Figure 4.13, we can analyze the evolution of the energy. After a purely elastic deformation where the energy grows quadratically, at time $t = 1$, following the crack initiation principle as prescribed by [51], the pre-fractured phase starts. Since the function θ grows linearly, the elastic energy keeps constant till the second significant

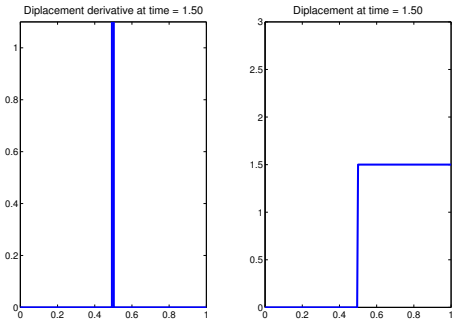


Figure 4.12.: A time step of the evolution of the pure off-plane shear of a Dugdale cohesive model for $t = 1.50$. On the left the gradient of the displacement ∇u and on the right the displacement u

time, $t = 1.5$, when the bridging force instantaneously vanish. Compared with the Barenblatt model, see Figure 4.3, the transition between the different evolution phases is clearer in this case.
Comparing the energy evolution given by the two algorithms, see Figure 4.13, we notice that there is no difference between the top and the bottom graph and we can state that the quasi-static evolution simulated by the two algorithms coincides.
The last check to validate the procedures involves the computational efficiency. In Figure 4.14, we report the number of iteration needed for convergence of the algorithms at each time step. It can be noticed that the PDAS (in red) requires in general two or three iteration of the minimization loop, which is considerably less than the forty-five required by Algorithm 4.1 (in blue). Naturally, the number of iteration has a direct influence on the overall computational times. Indeed, the time required on a MacBook Pro equipped with a 2.6GHz Intel Core i7 processor, 8GB of RAM, 1600MHz DDR3 to compute a whole quasi-static evolution in the time window $t \in [0, 3]$ with Algorithm 4.1 is 32.12 seconds, while we have a sensible reduction using Algorithm 4.2, which decreases the time needed to 4 seconds.

4.3. Conclusions

In this chapter we analyzed the two different cohesive fracture models and we proposed three different algorithms which are able to successfully retrieve quasi-static evolutions for these models. In particular, thanks to the quadratic perturbation added the Barenblatt model, Algorithm 2.2 selects the critical points of the energy functional which correspond to the ones of a physical quasi-static evolution. Additionally we proved that the discrete in time and space solution we obtain from the algorithm converges to the

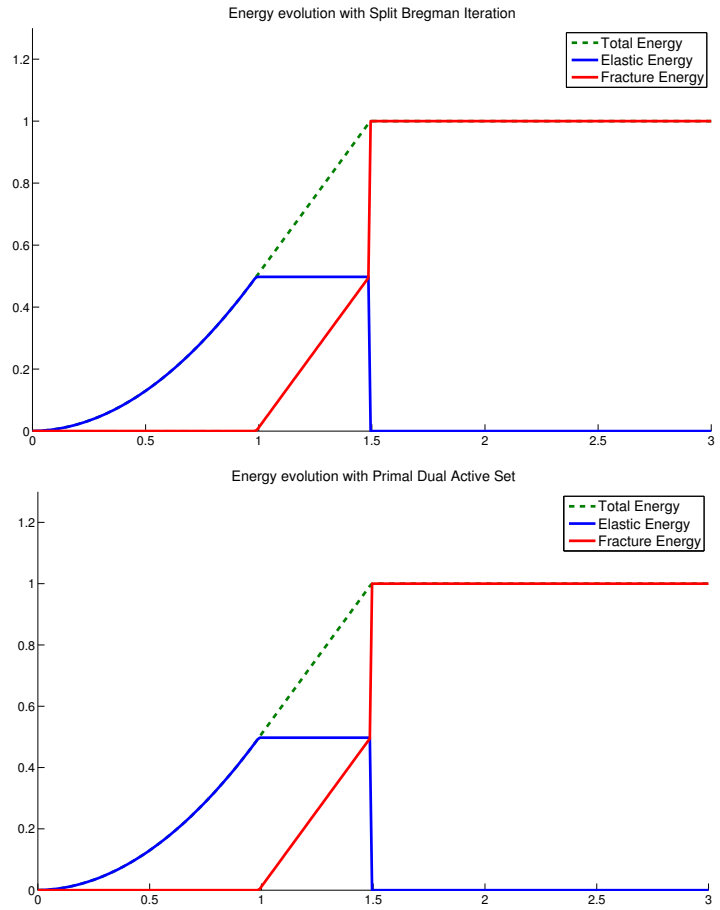


Figure 4.13.: The energy evolution of the system. The evolution obtained with Algorithm 4.1 (top) coincides with the one given by Algorithm 4.2 (bottom). The green line is the total energy of the system, in red the fracture energy and in blue the elastic one.

continuous in time and space analytic solution. Unfortunately the performances of the method are not competitive for a real life application since a simulation of the one dimensional problem showed in Section 4.1.5 requires more than one hour on a MacBook Pro equipped with a 2.6GHz Intel Core i7 processor, 8GB of RAM, 1600MHz DDR3 while the two dimensional problems needed circa twelve hours of computation to produce the whole simulation.
This computational cost issue stimulated us to investigate two different procedures which could have been more efficient. Facing the drawback of loosing the convergence guarantees, we focused on a model which prescribes evolution along local minimizers of the energy functional. We showed that the algorithms always converge, in practice, for the

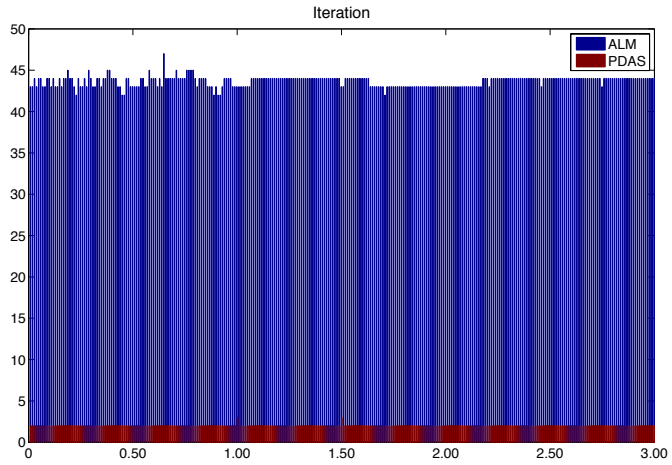


Figure 4.14.: In blue (ALM) the number of iterations at each time step needed for the convergence of Algorithm 4.1 while in red (PDAS) the ones needed for Algorithm 4.2

experiments we considered and they produce quasi-static evolutions which follow the expected course with computational times that are way smaller than the ones needed for Algorithm 2.2. Even if more numerical tests should be performed, it might be possible to extend the convergence theory also to the class of nonconvex functionals studied in this Chapter. An other interesting challenge would be to properly tune the efficient algorithms in order to select the expected critical points also for Barenblatt problems which do not necessarily the global minimizer of the energy function and thus creating a competitive strategy for real life problems.

5. Variational formulation of brittle fractures: reliable numerical simulation by anisotropic mesh adaptation

In Section 1.2.3 we introduced one of the most advocated models for quasi-static brittle fracture evolution that was presented by G. Francfort and J.–J. Marigo in [110]. This model is particularly relevant because it is able to predict complex crack paths, without making a priori assumptions on their possible propagation. Despite the success of this model for its mathematical well-posedness and, at the same time, its rather general framework, mechanical engineers and physicists of solids tend to favor more realistic models, where a smoother process towards fracture is considered and a minimal cohesion between the surfaces of the crack is not negligible. From this perspective, the approximation made by L. Ambrosio and V. M. Tortorelli in [10] of the energy functional driving the quasi-static evolution of the Francfort–Marigo model is very interesting because the crack is identified by a smooth phase field $v : \Omega \to [0, 1]$ instead of a sharp lower dimensional set. For this reason, in this chapter, as well as in [14, 16, 15], we consider exclusively a numerical analysis and simulation of a quasi-static evolution based on the Ambrosio–Tortorelli functional. In particular, our main focus is not on the detection of the critical point of the energy functional, which was the main argument of the last chapter, but on the proper fracture path detection, dropping the hypothesis of a prescribed crack path.

In this chapter, for sake of simplicity, we focus only on the pure shear off-plane displacement (1.18) if not differently specified. The analogous analysis relative to the alternative deformation modes, (1.17), is moved to Appendix B and it can also be found in [16]. In detail, we proceed as follows. In the next section we introduce the Ambrosio–Tortorelli functional, then in Section 5.1.1 we introduce the specific finite element scheme that we adopted. Section 5.2 is dedicated to the introduction of anisotropic domain discretization and anisotropic finite elements. Indeed, to efficiently compute the crack path of the quasi-static evolution we resolved to this particular technique that, as we show in the numerical experiment in Section 5.4, guarantees not only a reliable simulation but also advantages from a computational point of view. In order to guarantee the independence of the result from the discretization we choose, in Section 5.2.2 we derive a posteriori error estimator on which we base a mesh adaptation procedure illustrated in Section 5.3

together with the algorithm to tackle the minimization of the Ambrosio–Tortorelli functional. As anticipated, Section 5.4 reports several numerical experiments performed to qualitatively validate the procedure proposed in this chapter. The test are divided in three parts. The first intends to check whether the mesh adaptation process let the fracture free to detect an expected physical path according to the properties of the domain. The second tests the parameter sensitivity of the algorithm analyzing the changes in the output for different choices of the parameters. Finally the third one concerns numerical experiments in the in-plane strain and shear. In this particular case, it can be appreciated not only the crack initiation process, but also the extreme precision to detect the correct angle of propagation of the fracture depending on the direction in which the displacement is applied.

5.1. The Ambrosio–Tortorelli approximation of the Mumford–Shah functional

In the specific off-plane case we split the set Ω_D into two subsets Ω_{D^+} and Ω_{D^-} such that $\Omega_D = \Omega_{D^+} \cup \Omega_{D^-}$ and $\Omega_{D^+} \cap \Omega_{D^-} = \emptyset$, and we define the external load $g : [0, T] \to \mathcal{F}$ as

$$g(t) = \begin{cases} t & \text{on } \Omega_{D^+} \\ -t & \text{on } \Omega_{D^-} \\ 0 & \text{elsewhere} \end{cases} .$$

The space of the admissible configuration is then

$$\mathfrak{U}_t := \{u \in H^1(\Omega) : u|_{\Omega_D} = g(t)|_{\Omega_D}\}$$

while the Ambrosio–Tortorelli energy functional $E_\varepsilon : H^1(\Omega) \times H^1(\Omega; [0, 1]) \to \mathbb{R}$ which substitutes the Mumford–Shah functional in (1.21) is

$$E_\varepsilon(u, v) = \int_\Omega (v^2 + \eta)|\nabla u|^2 \, d\mathbf{x} + \kappa \int_\Omega \left[\frac{1}{4\varepsilon}(1 - v)^2 + \varepsilon|\nabla v|^2\right] d\mathbf{x}, \tag{5.1}$$

where $0 < \eta \ll \varepsilon \ll 1$, $\kappa > 0$, $v \in H^1(\Omega; [0, 1]) = \{v \in H^1(\Omega) : 0 \leq v \leq 1, \text{a.e. in } \Omega\}$, and $u \in \mathfrak{U}_t$. The first integral of (5.1) accounts for the elastic energy E_e of the body and the energy density coincides with the one defined in (1.18) except for a weighting factor $(v^2 + \eta)$, while the second one substitutes the fracture energy defined in (1.19).

Remark 5.1. *Let us observe that if $v \equiv 1$ the second integral vanishes and only the elastic energy is contributing to the system energy, while the points of the fracture set are identified by the subset of the domain where v is close to zero. Moreover, the transition area $0 < v < 1$ has thickness of order ε. This can be observed looking at the behavior of the functional for $\varepsilon \to 0$. The limit case requires $v \to 1$ almost everywhere and*

thus the transition area reduces to a 0 *and the set where* $v = 0$ *becomes a* $(d-1)$*-dimensional hypersurface. Thus, in practice,* v *can be considered as a phase field for the crack interface [38, 164].*

In [118] A. Giacomini proved that the evolution given by the minimization of this functional Γ-converges in $L^1(\Omega)\times L^1(\Omega)$ to the quasi-static evolution of the Francfort–Marigo model, as $\delta, \varepsilon \to 0$. The proof is built upon the original result of convergence made by L. Ambrosio and V.M. Tortorelli in [9] for the approximation of the Mumford–Shah functional [169]. Moreover, in [10], the proof of the existence of minimizers for (5.1) is provided for all ε, $\eta > 0$. Alternative Γ-approximations results are addressed, e.g., in [28, 103].
Now, to describe the quasi-static evolution for the time discretization introduced in the previous section, we proceed as follows. At the first time step $t = t_0$ we need to solve

$$(u(t_0), v(t_0)) \in \begin{array}{c} \arg\min \\ u \in \mathfrak{U}_{t_0}, \\ v \in H^1(\Omega;[0,1]) \end{array} E_\varepsilon(u, v),$$

whereas, for subsequent times $t = t_k$, for $k = 1, \dots, N_T$, we seek a pair $(u(t_k), v(t_k))$ such that

$$(u(t_k), v(t_k)) \in \begin{array}{c} \arg\min \\ u\in\mathfrak{U}_{t_k}, \\ v\in H^1(\Omega;[0,1]), v\le v(t_{k-1}) \end{array} E_\varepsilon(u, v), \tag{5.2}$$

where the condition $v \le v(t_{k-1})$ enforces the irreversibility of the crack. Indeed, the constraint of v being a non-increasing function in time and identifying the fracture set with

$$CR_k := \{\mathbf{x} \in \bar{\Omega} \,|\, v(t_{k-1}) < \texttt{CRTOL}\}, \tag{5.3}$$

where `CRTOL` is a small constant, it is naturally enforced the condition **(a)** of Theorem 1.11 that avoids healing of the crack in the future time steps.
The minimization process (5.2) requires minimizing a functional subject to constraints on both u and v. Accordingly with the procedure proposed in [14, 16, 15], we propose a minimization process where the constraints are relaxed through suitable penalty terms. We made this choice in order to avoid the selection of special function spaces and simply pose the problem as an unconstrained minimization in $H^1(\Omega)$ for both u and v, with a consequent simplification of the numerical implementation.
Before introducing the penalized functional, we exploit the definition of the crack set (5.3) to write the inequality constraint

$$v(t_k) \le v(t_{k-1}), \tag{5.4}$$

in an alternative way such that the fracture irreversibility is maintained and at the same time expressed as an equality constraint. This change is done mainly because (5.4) is not

easily implementable. In particular, we follow an alternative criterion, first introduced by B. Bourdin in [39]. Let CR_{k-1} the fracture set as defined in (5.3) at time $t = t_{k-1}$, then we can enforce condition **(a)** of Theorem 1.11 also by the following condition

$$v(\mathbf{x}, t_i) = 0 \qquad \forall \mathbf{x} \in CR_{k-1} \quad \text{and} \quad \forall i \,:\, k \leq i \leq N_T. \tag{5.5}$$

Moving from this idea, we propose minimizing the following penalized functional

$$\begin{aligned} E^{\gamma}_{\varepsilon,k}(u,v) \;&= \int_\Omega (v^2+\eta)|\nabla u|^2 \, d\mathbf{x} + \kappa \int_\Omega \left[\frac{1}{4\varepsilon}(1-v)^2 + \varepsilon|\nabla v|^2\right] d\mathbf{x} \\ &+ \frac{1}{\gamma_A}\int_{\Omega_D} (g(t_k)-u)^2 \, d\mathbf{x} + \frac{1}{\gamma_B}\int_{CR_{k-1}} v^2 \, d\mathbf{x}, \end{aligned} \tag{5.6}$$

where γ_A and γ_B are the two (small) penalty constants. Hence, setting $CR_{-1} = \emptyset$, the new optimization problem is

$$(u(t_k), v(t_k)) \in \underset{\substack{u\in H^1(\Omega),\\ v\in H^1(\Omega;[0,1])}}{\arg\min} \; E^{\gamma}_{\varepsilon,k}(u,v), \qquad \text{for } k = 0, \ldots, N_T, \tag{5.7}$$

where, differently from the previous models [39, 48], there is no more distinction between the first minimization step and the following ones as the last penalty term vanishes at $t = 0$. Notice that the function v must take values only in $[0, 1]$ but this constraint is not effectively enforced during the minimization process. Indeed, as shown in Proposition 5.4 using a truncation argument, any local minimizer (u, v) of (5.6) in the $H^1(\Omega) \times H^1(\Omega)$ topology is such that $0 \leq v \leq 1$ a.e. in Ω and thus there is no need for additional constraints. Moreover, adding the penalty terms we do not affect the Γ-convergence result of the functional. Being the additional terms clearly continuous, convex, and always non-negative, the proof of the convergence of the minimizers of (5.7) to ones fulfilling (5.5) instead of (5.4) in (5.2) for $\gamma_A, \gamma_B \to 0$, follows from the Γ-convergence theory [72].

In order to ease the notation, we assume from now on unless differently specified that $\kappa = 1$ and we define $\alpha = (4\varepsilon)^{-1}$. Throughout the remaining part of this section, we mimic the analysis in E. Süli et al. [48] by suitably modifying it to deal with functional (5.6).

Thanks to the boundedness of v we can eventually restrict its space to $L^\infty(\Omega) \cap H^1(\Omega)$ and introduce the following

Proposition 5.2. *The functional $E^{\gamma}_{\varepsilon,k}$ is Fréchet-differentiable in $H^1(\Omega) \times (H^1(\Omega) \cap L^\infty(\Omega))$.*

Proof. The proof follows directly from Proposition 1.1 in [48]. In particular, the differentiability of the additional penalty terms is trivial and thus the penalty terms does not change the regularity of the functional $E^{\gamma}_{\varepsilon,k}$. □

We introduce now the Fréchet derivative of $E^{\gamma}_{\varepsilon,k}(w,z)$ in the direction (φ,ψ), i.e.,

$$\begin{aligned}(E^{\gamma}_{\varepsilon,k})'(w,z;\varphi,\psi) &= 2\left(\int_{\Omega}(z^2+\eta)\nabla w\cdot\nabla\varphi\,d\mathbf{x}+\frac{1}{\gamma_A}\int_{\Omega_D}(w-g(t_k))\varphi\,d\mathbf{x}\right)\\ &\quad+2\left(\int_{\Omega}\left[z\psi|\nabla w|^2+\alpha(z-1)\psi+\varepsilon\nabla z\cdot\nabla\psi\right]d\mathbf{x}+\frac{1}{\gamma_B}\int_{CR_{k-1}}z\psi\,d\mathbf{x}\right)\\ &=: 2a_{\gamma_A}(z;w,\varphi)+2b_{\gamma_B}(w;z,\psi),\end{aligned}\tag{5.8}$$

where we have split the derivative in two parts; the first one, a_{γ_A}, associated with the derivative in the direction φ, and the second one, b_{γ_B}, related to the direction ψ. Accordingly, we can characterize the critical points of $E^{\gamma}_{\varepsilon,k}$ as follows

Definition 5.3. *The pair $(u,v)\in H^1(\Omega)\times(H^1(\Omega)\cap L^\infty(\Omega))$ is a* critical point *of $E^{\gamma}_{\varepsilon,k}$ if $(E^{\gamma}_{\varepsilon,k})'(u,v;\varphi,\psi)=0$ for all $\varphi\in H^1(\Omega)$ and for all $\psi\in(H^1(\Omega)\cap L^\infty(\Omega))$.*

As anticipated above, thanks to the following proposition, we can get rid of the constraint on v.

Proposition 5.4. *If $(u,v)\in H^1(\Omega)\times(H^1(\Omega)\cap L^\infty(\Omega))$ is a critical point of $E^{\gamma}_{\varepsilon,k}$, then $0\le v(\boldsymbol{x})\le 1$ for a.e. $\boldsymbol{x}\in\Omega$.*

Proof. Following the argument of Proposition 1.3 of [48], suppose that (u,v) is a critical point of $E^{\gamma}_{\varepsilon,k}$ and that Ω_1 and Ω_2 are the two subsets of Ω such that $\Omega_1=\{\mathbf{x}\in\Omega\,|\,v(\mathbf{x})>1\}$, $\Omega_2=\{\mathbf{x}\in\Omega\,|\,v(\mathbf{x})<0\}$, and $|\Omega_1\cup\Omega_2|>0$. Since (u,v) is a critical point of $E^{\gamma}_{\varepsilon,k}$, we have

$$b_{\gamma_B}(u;v,\psi)=0\qquad\forall\psi\in H^1(\Omega)\cap L^\infty(\Omega).$$

Then, if we choose

$$\psi(\mathbf{x})=\begin{cases}1-v(\mathbf{x}) & \mathbf{x}\in\Omega_1\\ -v(\mathbf{x}) & \mathbf{x}\in\Omega_2\\ 0 & \text{elsewhere,}\end{cases}$$

we obtain

$$\begin{aligned}b_{\gamma_B}(u;v,\psi) &= \int_{\Omega_1}\left[v(1-v)|\nabla u|^2-\alpha(v-1)^2-\varepsilon|\nabla v|^2\right]d\mathbf{x}\\ &\quad-\int_{\Omega_2}\left[v^2|\nabla u|^2+\alpha(v-1)v+\varepsilon|\nabla v|^2\right]d\mathbf{x}\\ &\quad-\frac{1}{\gamma_B}\int_{CR_{k-1}\cap\Omega_2}v^2\,d\mathbf{x}+\frac{1}{\gamma_B}\int_{CR_{k-1}\cap\Omega_1}v(1-v)\,d\mathbf{x}=0.\end{aligned}\tag{5.9}$$

The left-hand side of (5.9) consists of four negative terms, leading to a contradiction. □

5.1.1. The finite element discretization

In this section we introduce the discrete counterpart of the minimization problem (5.7) in a finite element setting. Thus, fixing $d = 2$ and eventually restricting Ω to be a polygonal domain, we denote with $\{\mathcal{T}_h\}_{h>0}$ a family of meshes of the domain $\overline{\Omega}$, with N_h the index set of the vertices of $\mathcal{T}_h$ and $\mathfrak{e}_h$ the skeleton of $\mathcal{T}_h$. Henceforth we assume that the boundary of Ω_D coincides with the union of consecutive edges in $\mathfrak{e}_h$. Finally, we associate $\{\mathcal{T}_h\}_{h>0}$ with the space X_h of the continuous piecewise linear finite elements [63] and we denote by $E_{h,k}(u_h, v_h)$ the discrete correspondent of $E^{\gamma}_{\varepsilon,k}(u, v)$ in (5.6), defined as follows

$$\begin{aligned} E_{h,k}(u_h, v_h) = \quad & \int_\Omega \Big[\Big(P_h(v_h^2) + \eta \Big) |\nabla u_h|^2 \, d\mathbf{x} + \alpha P_h((1 - v_h)^2) + \varepsilon |\nabla v_h|^2 \Big] \, d\mathbf{x} \\ & + \frac{1}{\gamma_A} \int_{\Omega_D} P_h \left((g_h(t_k) - u_h)^2 \right) \, d\mathbf{x} + \frac{1}{\gamma_B} \int_{CR_{k-1}} P_h \left(v_h^2 \right) \, d\mathbf{x}, \end{aligned} \tag{5.10}$$

where $P_h : C^0(\overline{\Omega}) \to X_h$ is the Lagrangian interpolant onto the space X_h, and $g_h \in X_h$ is a suitable discrete approximation of g. In particular, we pick $g_h(t_k)$, for any $k \in \{0, \dots, N_T\}$, such that

$$\int_\Omega g_h(t_k) w_h \, d\mathbf{x} = \int_\Omega g(t_k) w_h \, d\mathbf{x} \qquad \forall w_h \in X_h, \tag{5.11}$$

i.e., $g_h(t_k)$ is the $L^2(\Omega)$-projection of $g(t_k)$ onto X_h. Moreover, we point out that the action of the operator P_h is equivalent to a mass lumping [207] and it allows us to extend Proposition 5.4 to the critical points of $E_{h,k}$ as well.
In the sequel, we assume that the off-diagonal entries of the stiffness matrix $[k_{ij}]$ associated with the space X_h be non-positive, i.e.,

$$k_{ij} := \int_\Omega \nabla \xi_i \cdot \nabla \xi_j \, d\mathbf{x} \leq 0 \qquad \forall i \neq j \in N_h, \tag{5.12}$$

where $\{\xi_l\}_{l=1}^{\#N_h}$ denotes the finite element basis of X_h. This condition is related to discrete maximum principle as discussed, for instance, in [64, 147, 202].
We are now able to define the discrete analogue to (5.7) which is

$$(u_h(t_k), v_h(t_k)) \in \underset{\substack{\widehat{u}_h \in X_h, \\ \widehat{v}_h \in X_h}}{\arg\min} \; E_{h,k}(\widehat{u}_h, \widehat{v}_h), \qquad \text{for } k = 0, \dots, N_T,$$

while the discrete version of the Definition 5.3 of critical point is

Definition 5.5. *The pair* $(u_h, v_h) \in X_h \times X_h$ *is a* critical point *of the discrete energy functional* $E_{h,k}$ *if* $(E_{h,k})'(u_h, v_h; \varphi_h, \psi_h) = 0$ *for all* $(\varphi_h, \psi_h) \in X_h \times X_h$*, where*

$$
\begin{aligned}
(E_{h,k})'(u_h, v_h; \varphi_h, \psi_h) =& 2\bigg(\int_\Omega (P_h(v_h^2) + \eta) \nabla u_h \cdot \nabla \varphi_h \, d\boldsymbol{x} + \frac{1}{\gamma_A} \int_{\Omega_D} P_h((u_h - g_h(t_k))\varphi_h) \, d\boldsymbol{x} \bigg) \\
&+ 2\bigg(\int_\Omega \Big[P_h(v_h\psi_h) |\nabla u_h|^2 + \alpha P_h((v_h - 1)\psi_h) + \varepsilon \nabla v_h \cdot \nabla \psi_h \Big] \, d\boldsymbol{x} \\
&+ \frac{1}{\gamma_B} \int_{CR_{k-1}^h} P_h\left(v_h \psi_h\right) \, d\boldsymbol{x} \bigg) =: 2a_{\gamma_A}^h(v_h; u_h, \varphi_h) + 2b_{\gamma_B}^h(u_h; v_h, \psi_h).
\end{aligned}
$$

Also Proposition 5.4 can be adapted to the discrete case, suitably relying on assumption (5.12) and on the properties of P_h, as follows

Proposition 5.6. *Let* $(u_h, v_h) \in X_h \times X_h$ *be a critical point of* $E_{h,k}$*, then* $0 \le v_h \le 1$ *for all* $x \in \Omega$*.*

Proof. The proof generalizes Proposition 2.2 in [48], by properly including the term $\int_{CR_{k-1}} P_h(v_h^2) \, d\mathbf{x}$. By mimicking in a discrete setting the proof of Proposition 5.4, we suppose, by contradiction, that there exist two index sets $J_1, J_2 \subset N_h$ where $v_i > 1$ for all $i \in J_1$ and $v_j < 0$ for all $j \in J_2$, where we let $v_i = v_h(\mathbf{x}_i, t_k)$.
Consider $j \in J_2$ such that $v_j \le v_i$, for all $i \in N_h$ and let Δ_j be the patch of elements associated with $\mathbf{x}_j$ with $M_j = \{i \in N_h \, : \, \mathbf{x}_i \in \Delta_j\}$.
Now, if we choose as a test function the hat function ξ_j associated with $\mathbf{x}_j$, from the equality $b_{\gamma_B}^h(u_h; v_h, \psi_h) = 0$ we have

$$
\begin{aligned}
\varepsilon \int_{\Delta_j} \nabla v_h \cdot \nabla \xi_j \, d\mathbf{x} = & \; - \int_{\Delta_j} \left[P_h(v_h \xi_j) |\nabla u_h|^2 - \alpha P_h((v_h - 1)\xi_j) \right] d\mathbf{x} \\
& - \frac{1}{\gamma_B} \int_{CR_{k-1} \cap \Delta_j} P_h\left(v_h \xi_j\right) \, d\mathbf{x} \\
> & \; - v_j \sum_{K \in \Delta_j} |\nabla u_h|_K|^2 \frac{|K|}{3} - \alpha (v_j - 1) \frac{|\Delta_j|}{3} > 0,
\end{aligned}
\tag{5.13}
$$

where the last inequality is obtained considering that $v_j < 0$ and that $P_h\left(v_h \xi_j\right)$ is a non-positive function on the set $CR_{k-1} \cap \Delta_j$. Since $v_h = \sum_{i \in N_h} v_i \xi_i$, and the sum of the rows of the stiffness matrix is zero, we have

$$
\varepsilon \int_{\Delta_j} \nabla v_h \cdot \nabla \xi_j \, d\mathbf{x} = \varepsilon \sum_{i \in M_j} k_{ji} v_i = \varepsilon \sum_{i \in M_j} k_{ji}(v_i - v_j) + \varepsilon \sum_{i \in M_j} k_{ji} v_j = \varepsilon \sum_{i \in M_j} k_{ji}(v_i - v_j).
$$

Thus, using assumption (5.12) and the hypothesis $v_j \le v_i$, we have that

$$
\varepsilon \int_{\Delta_j} \nabla v_h \cdot \nabla \xi_j \, d\mathbf{x} \le 0
$$

in contradiction with (5.13).
Similarly, we can proceed to contradict the existence of nodes in J_1. Consider $j \in J_1$ such that $v_j \geq v_i$ for all $i \in N_h$ and let Δ_j be the patch of elements associated with $\mathbf{x}_j$. Now, choosing again as test function the hat function ξ_j associated with $\mathbf{x}_j$, from the equality $b^h_{\gamma_B}(u_h; v_h, \psi_h) = 0$ we have

$$\begin{aligned} \varepsilon \int_{\Delta_j} \nabla v_h \cdot \nabla \xi_j \, d\mathbf{x} = & -\int_{\Delta_j} [P_h(v_h \xi_j)|\nabla u_h|^2 - \alpha P_h((v_h - 1)\xi_j)] \, d\mathbf{x} \\ & -\frac{1}{\gamma_B} \int_{CR_{k-1} \cap \Delta_j} P_h\left(v_h \xi_j\right) \, d\mathbf{x} \\ < & -v_j \sum_{K \in \Delta_j} |\nabla u_h|_K|^2 \frac{|K|}{3} - \alpha(v_j - 1)\frac{|\Delta_j|}{3} < 0, \end{aligned} \tag{5.14}$$

where the last inequality is obtained considering that $v_j > 1$ and that $P_h\left(v_h \xi_j\right)$ is a positive function on the set $CR_{k-1} \cap \Delta_j$. Following a similar reasoning as before and thanks to the hypothesis $v_j \geq v_i$, we have that $\varepsilon \displaystyle\int_{\Delta_j} \nabla v_h \cdot \nabla \xi_j \, d\mathbf{x} \geq 0$ in contradiction with (5.14). □

Remark 5.7. *Following [48], we replace also the definition of the crack set* (5.3) *by a discrete version*

$$CR^h_{k-1} = \bigcup_{e \in \mathfrak{e}^{CR}_h} \overline{e}, \quad \textit{where} \quad \mathfrak{e}^{CR}_h := \{e \in \mathfrak{e}_h \ : \ v_h(\boldsymbol{x}, t_{k-1}) \leq \texttt{CRTOL}, \ \forall \boldsymbol{x} \in \overline{e}\},$$

which enjoys one-dimensional features. Although this definition introduces an approximation error whose study is beyond the scope of this work, the reason of introducing such approximation is the avoidance of improper crack propagation. Indeed, as already observed in [42], the choice of a too big `CRTOL` *may cause a widening of the fracture. In the same spirit, we consider to be crack edges only the ones which are entirely contained in the crack set CR_{k-1} to avoid edges not aligned with the fracture to be considered part of the crack set causing consequently the crack widening phenomenon.*

5.2. Anisotropic finite element setting and error estimator

In this section we face the problem of designing an accurate, robust, and efficient mesh adaptation procedure that, combined with a proper minimization algorithm, allows us to successfully *break the dilemma "the grid follows the fracture or the fracture follows the grid"* with a proper balance between accuracy and complexity. Indeed, previous accurate numerical simulations of the Francfort–Marigo model have been performed either on very fine *ad hoc* meshes [39, 42] or towards adaptive discretization of free-discontinuity functionals of Mumford–Shah type [41, 48, 49, 59, 172].

In particular, the strategy we propose in this section follows and combines both strategies suggested in the papers [41, 48]; we merge the minimization of the Ambrosio–Tortorelli model with an adaptive *anisotropic* discretization by exploiting the fact that the considered model exhibits solutions with very steep features close to the crack. Moreover, being the crack a $(d-1)$-dimensional object evolving along a specific direction, the choice of an anisotropic adaptation can be really advantageous compared with a more standard isotropic approach.
Before introducing the a posteriori estimator for $(E_{h,k})'(u_h, v_h; \varphi, \psi)$, we lay down the anisotropic background which is fundamental for a deep understanding of this chapter.

5.2.1. The anisotropic background

We shortly introduce the anisotropic setting described in [88, 161] that we will use for our mesh adaptation. In these works, the directional anisotropic information is derived from the spectral properties of the standard affine map $T_K : \widehat{K} \to K$ with

$$\mathbf{x} = T_K(\widehat{\mathbf{x}}) = M_K \widehat{\mathbf{x}} + \mathbf{t}_K$$

between the equilateral reference triangle $\widehat{K}$ inscribed in the unit circle and the generic triangle K of the mesh $\mathcal{T}_h$, with $M_K \in \mathbb{R}^{2\times 2}$, $\mathbf{t}_K \in \mathbb{R}^2$, $\mathbf{x} \in K$, $\widehat{\mathbf{x}} \in \widehat{K}$.

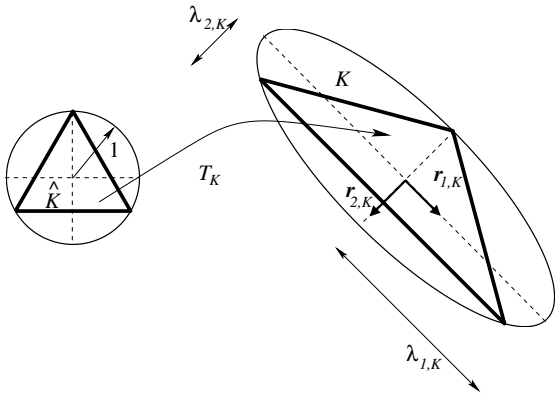


Figure 5.1.: Geometric quantities associated with the map T_K

We introduce the polar decomposition of the Jacobian M_K, i.e., $M_K = B_K Z_K$, where $B_K, Z_K \in \mathbb{R}^{2\times 2}$ are respectively a symmetric positive definite and an orthogonal matrix. The first matrix models the deformation of K, while Z_K rotates it rigidly. Then we consider the eigenvalue factorization of B_K as $B_K = R_K^T \Lambda_K R_K$, with $R_K^T = [\mathbf{r}_{1,K}, \mathbf{r}_{2,K}]$ and $\Lambda_K = \mathrm{diag}(\lambda_{1,K}, \lambda_{2,K})$. In particular, the eigenvectors $\mathbf{r}_{1,K}$, $\mathbf{r}_{2,K}$ give the directions of the semi-axes of the ellipse circumscribed to K, while the eigevalues $\lambda_{1,K}$, $\lambda_{2,K}$ measure the length of these semi-axes (see Figure 5.1). We also define the aspect ratio of the

element K by $s_K = \lambda_{1,K}/\lambda_{2,K}$. With a view to an anisotropic control of the mesh, we introduce the quasi-interpolant Clément operator $I_h : L^2(\Omega) \to X_h$ [66]. We recall the following anisotropic estimate for the interpolation error.

Lemma 5.8. *Let $w \in H^1(\Omega)$. If the cardinality $\#\Delta_K \leq \mathcal{N}$ for some $\mathcal{N} \in \mathbb{N}$, and $\mathrm{diam}(T_K^{-1}(\Delta_K)) \leq C_\Delta \simeq O(1)$, where $\Delta_K = \{T \in \mathcal{T}_h : T \cap K \neq \emptyset\}$, then there exist constants $C_s = C_s(\mathcal{N}, C_\Delta)$, with $s = 0, 1, 2$, such that, for any $K \in \mathcal{T}_h$, it holds*

$$\|w - I_h(w)\|_{H^s(K)} \leq C_s \left(\frac{1}{\lambda_{2,K}}\right)^s \left[\sum_{i=1}^{2} \lambda_{i,K}^2 (\boldsymbol{r}_{i,K}^T G_{\Delta_K}(w) \boldsymbol{r}_{i,K})\right]^{1/2}, \tag{5.15}$$

with $s = 0, 1$, and

$$\|w - I_h(w)\|_{L^2(\partial K)} \leq C_2 \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{1/2} \left[\sum_{i=1}^{2} \lambda_{i,K}^2 (\boldsymbol{r}_{i,K}^T G_{\Delta_K}(w) \boldsymbol{r}_{i,K})\right]^{1/2}, \tag{5.16}$$

where $h_K = \mathrm{diam}(K)$,

$$G_{\Delta_K}(w) = \sum_{T \in \Delta_K} G_T(w) \tag{5.17}$$

is a symmetric positive semi-definite matrix with

$$G_T(w) := \begin{bmatrix} \displaystyle\int_T \left(\frac{\partial w}{\partial x_1}\right)^2 d\boldsymbol{x} & \displaystyle\int_T \frac{\partial w}{\partial x_1}\frac{\partial w}{\partial x_2}\, d\boldsymbol{x} \\ \displaystyle\int_T \frac{\partial w}{\partial x_1}\frac{\partial w}{\partial x_2}\, d\boldsymbol{x} & \displaystyle\int_T \left(\frac{\partial w}{\partial x_2}\right)^2 d\boldsymbol{x} \end{bmatrix}, \tag{5.18}$$

for any $T \in \mathcal{T}_h$.

Proof. See [104, 105] for the details. □

We observe that the geometrical hypotheses in Lemma 5.8 do not limit explicitly the anisotropic features (stretching factor and orientation) of each element, but rather they ensure some smoothness in the variation of the anisotropic features [163].
Let us recall an equivalence result between the standard $H^1(\Delta_K)$-seminorm and its anisotropic correspondent:

Lemma 5.9. *Let $w \in H^1(\Omega)$ and $K \in \mathcal{T}_h$. For any $\beta_1, \beta_2 > 0$, it holds*

$$\min\{\beta_1, \beta_2\} \leq \frac{\beta_1(\boldsymbol{r}_{1,K}^T G_{\Delta_K}(w)\boldsymbol{r}_{1,K}) + \beta_2(\boldsymbol{r}_{2,K}^T G_{\Delta_K}(w)\boldsymbol{r}_{2,K})}{|w|^2_{H^1(\Delta_K)}} \leq \max\{\beta_1, \beta_2\}, \tag{5.19}$$

where $G_{\Delta_K}(\cdot)$ is defined as in (5.17).

Proof. See [160] for the details. □

5.2.2. An a posteriori error estimator

We can now state the main result of this section which represents the anisotropic analogue of Proposition 3.1 in [48].

Proposition 5.10. *Let $(u_h, v_h) \in X_h \times X_h$ be the critical point of $E_{h,k}$ according to Definition 5.5. Then it holds*

$$|(E^{\gamma}_{\varepsilon,k})'(u_h, v_h; \varphi, \psi)| \leq C \sum_{K \in \mathcal{T}_h} \left\{ \rho^A_K(v_h, u_h)\, \omega_K(\varphi) + \rho^B_K(u_h, v_h)\, \omega_K(\psi) \right\} \quad \forall \varphi, \psi \in H^1(\Omega), \tag{5.20}$$

where $C = C(\mathcal{N}, C_\Delta)$, while

$$\begin{aligned}
\rho^A_K(v_h, u_h) = {} & \frac{1}{2} \|[\![\nabla u_h]\!]\|_{L^\infty(\partial K)} \, \|v_h^2 + \eta\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}} \right)^{\frac{1}{2}} \\
& + \|2v_h(\nabla v_h \cdot \nabla u_h)\|_{L^2(K)} + \frac{\delta_{K,\Omega_D}}{\gamma_A} \Big(\|u_h - g_h(t_k)\|_{L^2(K)} \\
& + \|g_h(t_k) - g(t_k)\|_{L^2(K)} \Big) + \frac{1}{\lambda_{2,K}} \Big[\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)} \, \|\nabla u_h\|_{L^2(K)} \\
& + \frac{|K|^{1/2}\, h_K^2}{\gamma_A} \, |u_h - g_h(t_k)|_{W^{1,\infty}(K)} \Big] \\
\rho^B_K(u_h, v_h) = {} & \|(|\nabla u_h|^2 + \alpha) v_h - \alpha\|_{L^2(K)} + \frac{\varepsilon}{2} \, \|[\![\nabla v_h]\!]\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}} \right)^{1/2} \\
& + \frac{\delta_{K,CR_{k-1}}}{\gamma_B} \, \|v_h\|_{L^2(K)} + \frac{h_K^2}{\lambda_{2,K}} \Big[\| \, |\nabla u_h|^2 + \alpha \|_{L^2(K)} \\
& + \frac{|K|^{1/2}\, \delta_{K,CR_{k-1}}}{\gamma_B} \Big] \, |v_h|_{W^{1,\infty}(K)} \\
\omega_K(w) = {} & \Big[\sum_{i=1}^{2} \lambda^2_{i,K} (\boldsymbol{r}^T_{i,K} G_{\Delta_K}(w) \boldsymbol{r}_{i,K}) \Big]^{1/2} \qquad \forall w \in H^1(\Omega),
\end{aligned}$$

where

$$[\![w_h]\!] := \begin{cases} |[\nabla w_h \cdot \nu]| & \text{on } \mathfrak{e}_h \\ |\nabla w_h \cdot \nu| & \text{on } \mathfrak{e}_h \cap \partial\Omega \end{cases} \tag{5.21}$$

denotes the absolute value of the jump of the normal derivative, with ν the unit normal vector to the generic edge in $\mathfrak{e}_h$, g_h is chosen as in (5.11) and $\delta_{K,\varpi}$ is such that $\delta_{K,\varpi} = 1$ if $K \cap \varpi \neq \emptyset$ and $\delta_{K,\varpi} = 0$ otherwise.

Proof. Since (u_h, v_h) is a critical point of $E_{h,k}$, we have

$$a^h_{\gamma_A}(v_h; u_h, \varphi_h) = 0 \quad \forall \varphi_h \in X_h, \qquad b^h_{\gamma_B}(u_h; v_h, \psi_h) = 0 \quad \forall \psi_h \in X_h. \tag{5.22}$$

Moreover, from (5.8), for any pair $(\varphi,\psi)\in H^1(\Omega)\times H^1(\Omega)$, it holds

$$|(E^{\gamma}_{\varepsilon,k})'(u_h,v_h;\varphi,\psi)|\le 2|a_{\gamma_A}(v_h;u_h,\varphi)|+2|b_{\gamma_B}(u_h;v_h,\psi)|. \tag{5.23}$$

Let us deal with the two terms above, separately. We start from $|a_{\gamma_A}(v_h;u_h,\varphi)|$. Thanks to (5.22), we have

$$\begin{aligned}|a_{\gamma_A}(v_h;u_h,\varphi)| &\le |a_{\gamma_A}(v_h;u_h,\varphi-\varphi_h)|\\ &+ |a_{\gamma_A}(v_h;u_h,\varphi_h)-a^h_{\gamma_A}(v_h;u_h,\varphi_h)| \quad \forall\varphi\in H^1(\Omega),\quad \forall\varphi_h\in X_h.\end{aligned} \tag{5.24}$$

Concerning the first term on the right-hand side of (5.24), we get

$$\begin{aligned}&|a_{\gamma_A}(v_h;u_h,\varphi-\varphi_h)|\\
&= \Big|\sum_{K\in\mathcal{T}_h}\Big\{\int_K (v_h^2+\eta)\nabla u_h\cdot\nabla(\varphi-\varphi_h)\,d\mathbf{x}+\frac{1}{\gamma_A}\int_K (u_h-g(t_k))(\varphi-\varphi_h)\,\chi_{\Omega_D}\,d\mathbf{x}\Big\}\Big|\\
&= \Big|\sum_{K\in\mathcal{T}_h}\Big\{\int_K -2v_h(\nabla v_h\cdot\nabla u_h)(\varphi-\varphi_h)\,d\mathbf{x}+\int_{\partial K}(v_h^2+\eta)\nabla u_h\cdot\nu(\varphi-\varphi_h)ds\\
&\quad +\frac{1}{\gamma_A}\int_K [(u_h-g_h(t_k))+(g_h(t_k)-g(t_k))](\varphi-\varphi_h)\,\chi_{\Omega_D}\,d\mathbf{x}\Big\}\Big|\\
&\le \sum_{K\in\mathcal{T}_h}\Big\{\|2v_h(\nabla v_h\cdot\nabla u_h)\|_{L^2(K)}\,\|\varphi-\varphi_h\|_{L^2(K)}+\frac{1}{2}\int_{\partial K}[\![\nabla u_h]\!]\,|v_h^2+\eta|\,|\varphi-\varphi_h|ds\\
&\quad +\frac{1}{\gamma_A}\Big(\|(u_h-g_h(t_k))\,\chi_{\Omega_D}\|_{L^2(K)}+\|(g_h(t_k)-g(t_k))\,\chi_{\Omega_D}\|_{L^2(K)}\Big)\|(\varphi-\varphi_h)\,\chi_{\Omega_D}\|_{L^2(K)}\Big\}\\
&\le \sum_{K\in\mathcal{T}_h}\Big\{\|2v_h(\nabla v_h\cdot\nabla u_h)\|_{L^2(K)}\,\|\varphi-\varphi_h\|_{L^2(K)}\\
&\quad +\frac{1}{2}\|[\![\nabla u_h]\!]\|_{L^\infty(\partial K)}\|v_h^2+\eta\|_{L^2(\partial K)}\|\varphi-\varphi_h\|_{L^2(\partial K)}+\frac{1}{\gamma_A}\Big(\|(u_h-g_h(t_k))\,\chi_{\Omega_D}\|_{L^2(K)}\\
&\quad +\|(g_h(t_k)-g(t_k))\,\chi_{\Omega_D}\|_{L^2(K)}\Big)\|(\varphi-\varphi_h)\,\chi_{\Omega_D}\|_{L^2(K)}\Big\},\end{aligned} \tag{5.25}$$

after splitting the integrals on the mesh elements, exploiting integration by parts, Hölder and Cauchy–Schwarz inequalities, and definition (5.21). Hereafter, χ_{ϖ} denotes the characteristic function of the set ϖ. Picking $\varphi_h=I_h(\varphi)$ and thanks to Lemma 5.8 with $s=0$, we obtain

$$\begin{aligned}&|a_{\gamma_A}(v_h;u_h,\varphi-\varphi_h)|\\
&\le C\sum_{K\in\mathcal{T}_h}\Big\{\|2v_h(\nabla v_h\cdot\nabla u_h)\|_{L^2(K)}+\frac{1}{2}\|[\![\nabla u_h]\!]\|_{L^\infty(\partial K)}\|v_h^2+\eta\|_{L^2(\partial K)}\left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{\frac{1}{2}}\\
&\quad +\frac{\delta_{K,\Omega_D}}{\gamma_A}\Big(\|u_h-g_h(t_k)\|_{L^2(K)}+\|g_h(t_k)-g(t_k)\|_{L^2(K)}\Big)\Big\}\left[\sum_{i=1}^{2}\lambda_{i,K}^2(\mathbf{r}_{i,K}^T G_{\Delta_K}(\varphi)\mathbf{r}_{i,K})\right]^{\frac{1}{2}}.\end{aligned} \tag{5.26}$$

Let us now deal with the second term on the right-hand side of (5.24). We anticipate the auxiliary result

$$|w_h\varphi_h|_{H^2(K)} \le 2\,|w_h|_{W^{1,\infty}(K)}\,\|\nabla\varphi_h\|_{L^2(K)} \qquad \forall w_h,\varphi_h \in X_h, \quad \forall K\in\mathcal{T}_h, \tag{5.27}$$

which can be proved by straightforward calculus. Now, employing standard inequalities (Hölder, Cauchy–Schwarz) together with the definition of $g_h(t_k)$ and the standard isotropic estimate for the L^2-norm of the interpolation error associated with P_h, we get

$$\begin{aligned}
|a_{\gamma_A}(v_h;u_h,\varphi_h) - a^h_{\gamma_A}(v_h;u_h,\varphi_h)| &\le \Big|\int_\Omega [v_h^2 - P_h(v_h^2)]\nabla u_h\cdot\nabla\varphi_h\,d\mathbf{x}\Big| \\
&\quad + \frac{1}{\gamma_A}\Big|\int_{\Omega_D}[(u_h - g_h(t_k))\varphi_h - P_h((u_h-g_h(t_k))\varphi_h)]\,d\mathbf{x}\Big| \\
&\quad + \frac{1}{\gamma_A}\Big|\int_{\Omega_D}(g_h(t_k)\varphi_h - g(t_k)\varphi_h)\,d\mathbf{x}\Big| \\
&\le C\sum_{K\in\mathcal{T}_h}\Big\{\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)}\,\|\nabla u_h\|_{L^2(K)}\,\|\nabla\varphi_h\|_{L^2(K)} \\
&\quad + \frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|(u_h-g_h(t_k))\varphi_h|_{H^2(K)}\Big\},
\end{aligned} \tag{5.28}$$

where the constant C does not depend on the aspect ratio s_K of K. Then we employ (5.27) together with estimate (5.15) with $s=1$ and Lemma 5.9 with $\beta_1 = \lambda_{1,K}^2$, $\beta_2 = \lambda_{2,K}^2$, to obtain

$$\begin{aligned}
&|a_{\gamma_A}(v_h;u_h,\varphi_h) - a^h_{\gamma_A}(v_h;u_h,\varphi_h)| \\
&\le C\sum_{K\in\mathcal{T}_h}\Big\{\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)}\,\|\nabla u_h\|_{L^2(K)}\,\|\nabla\varphi_h\|_{L^2(K)} \\
&\qquad + \frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|u_h - g_h(t_k)|_{W^{1,\infty}(K)}\,\|\nabla\varphi_h\|_{L^2(K)}\Big\} \\
&\le C\sum_{K\in\mathcal{T}_h}\Big\{\Big(\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)}\,\|\nabla u_h\|_{L^2(K)} + \frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|u_h - g_h(t_k)|_{W^{1,\infty}(K)}\Big) \\
&\qquad \Big(\|\nabla\varphi_h - \nabla\varphi\|_{L^2(K)} + \|\nabla\varphi\|_{L^2(K)}\Big)\Big\} \\
&\le C\sum_{K\in\mathcal{T}_h}\Big\{\Big(\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)}\,\|\nabla u_h\|_{L^2(K)} + \frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|u_h - g_h(t_k)|_{W^{1,\infty}(K)}\Big) \\
&\qquad \frac{1}{\lambda_{2,K}}\Big[\sum_{i=1}^{2}\lambda_{i,K}^2(\mathbf{r}_{i,K}^T G_{\Delta_K}(\varphi)\mathbf{r}_{i,K})\Big]^{1/2}\Big\}.
\end{aligned} \tag{5.29}$$

Therefore, collecting (5.26) and (5.29), we are able to bound the first term on the right-hand side of (5.23), as

$$|a_{\gamma_A}(v_h;u_h,\varphi)| \le C\sum_{K\in\mathcal{T}_h}\rho_K^A(v_h,u_h)\,\omega_K^A(\varphi).$$

Let us consider now the second term on the righ-hand side of (5.23). In the same way as in (5.24) and thanks to (5.22), we have

$$\begin{aligned}|b_{\gamma_B}(u_h;v_h,\psi)| \le\ & |b_{\gamma_B}(u_h;v_h,\psi-\psi_h)| \\ & +|b_{\gamma_B}(u_h;v_h,\psi_h)-b^h_{\gamma_B}(u_h;v_h,\psi_h)| \qquad \forall\psi\in H^1(\Omega),\ \forall\psi_h\in X_h.\end{aligned} \tag{5.30}$$

We tackle the first term $|b_{\gamma_B}(u_h;v_h,\psi-\psi_h)|$. Rewriting the integrals on Ω over the mesh elements, integrating by parts, and thanks to the Cauchy–Schwarz inequality and definition (5.21), we obtain

$$\begin{aligned}&|b_{\gamma_B}(u_h;v_h,\psi-\psi_h)| \\ =\ & \Big|\sum_{K\in\mathcal{T}_h}\Big\{\int_K\left[((|\nabla u_h|^2+\alpha)v_h-\alpha)(\psi-\psi_h)+\varepsilon\nabla v_h\cdot\nabla(\psi-\psi_h)\right]d\mathbf{x} \\ & +\frac{1}{\gamma_B}\int_K v_h(\psi-\psi_h)\,\chi_{CR_{k-1}}\,d\mathbf{x}\Big\}\Big| \\ \le\ & \sum_{K\in\mathcal{T}_h}\Big\{\|(|\nabla u_h|^2+\alpha)v_h-\alpha\|_{L^2(K)}\|\psi-\psi_h\|_{L^2(K)}+\Big|\varepsilon\int_{\partial K}(\psi-\psi_h)\nabla v_h\cdot\nu\,ds\Big| \\ & +\frac{1}{\gamma_B}\|v_h\,\chi_{CR_{k-1}}\|_{L^2(K)}\,\|(\psi-\psi_h)\,\chi_{CR_{k-1}}\|_{L^2(K)}\Big\} \\ \le\ & \sum_{K\in\mathcal{T}_h}\Big\{\|(|\nabla u_h|^2+\alpha)v_h-\alpha\|_{L^2(K)}\|\psi-\psi_h\|_{L^2(K)}+\frac{\varepsilon}{2}\,\|[\![\nabla v_h]\!]\|_{L^2(\partial K)}\|\psi-\psi_h\|_{L^2(\partial K)} \\ & +\frac{1}{\gamma_B}\|v_h\,\chi_{CR_{k-1}}\|_{L^2(K)}\,\|(\psi-\psi_h)\,\chi_{CR_{k-1}}\|_{L^2(K)}\Big\}.\end{aligned} \tag{5.31}$$

We now choose $\psi_h=I_h(\psi)$ and use Lemma 5.8 to get

$$\begin{aligned}&|b_{\gamma_B}(u_h;v_h,\psi-\psi_h)| \\ &\le C\sum_{K\in\mathcal{T}_h}\Big\{\|(|\nabla u_h|^2+\alpha)v_h-\alpha\|_{L^2(K)}+\frac{\varepsilon}{2}\,\|[\![\nabla v_h]\!]\|_{L^2(\partial K)}\left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{1/2} \\ &\quad+\frac{\delta_{K,CR_{k-1}}}{\gamma_B}\,\|v_h\|_{L^2(K)}\Big\}\left[\sum_{i=1}^{2}\lambda^2_{i,K}(\mathbf{r}^T_{i,K}G_{\Delta_K}(\psi)\mathbf{r}_{i,K})\right]^{1/2}.\end{aligned} \tag{5.32}$$

We estimate now the second term on the right-hand side of (5.30). By mimicking the arguments employed in (5.28)–(5.29), we obtain the following bound:

$$\begin{aligned}&|b_{\gamma_B}(u_h;v_h,\psi_h)-b^h_{\gamma_B}(u_h;v_h,\psi_h)| \\ &\le\Big|\int_\Omega(v_h\psi_h-P_h(v_h\psi_h))(|\nabla u_h|^2+\alpha)\,d\mathbf{x}\Big|+\frac{1}{\gamma_B}\int_{CR_{k-1}}(v_h\psi_h-P_h(v_h\psi_h))\,d\mathbf{x}\Big|\end{aligned}$$

$$
\begin{aligned}
&\le C \sum_{K\in\mathcal{T}_h} \Big\{ \|v_h\psi_h - P_h(v_h\psi_h)\|_{L^2(K)} \Big[\| \, |\nabla u_h|^2 + \alpha \|_{L^2(K)} + \frac{|K|^{1/2}\,\delta_{K,CR_{k-1}}}{\gamma_B} \Big] \Big\} \\
&\le C \sum_{K\in\mathcal{T}_h} \Big\{ \Big[\| \, |\nabla u_h|^2 + \alpha \|_{L^2(K)} + \frac{|K|^{1/2}\,\delta_{K,CR_{k-1}}}{\gamma_B} \Big] h_K^2 \, |v_h\psi_h|_{H^2(K)} \Big\} \\
&\le C \sum_{K\in\mathcal{T}_h} \Big\{ \Big[\| \, |\nabla u_h|^2 + \alpha \|_{L^2(K)} + \frac{|K|^{1/2}\,\delta_{K,CR_{k-1}}}{\gamma_B} \Big] h_K^2 \, |v_h|_{W^{1,\infty}(K)} \|\nabla\psi_h\|_{L^2(K)} \Big\} \\
&\le C \sum_{K\in\mathcal{T}_h} \Big\{ \Big[\| \, |\nabla u_h|^2 + \alpha \|_{L^2(K)} + \frac{|K|^{1/2}\,\delta_{K,CR_{k-1}}}{\gamma_B} \Big] h_K^2 \, |v_h|_{W^{1,\infty}(K)} \\
&\qquad\qquad \frac{1}{\lambda_{2,K}} \Big[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(\psi) \mathbf{r}_{i,K}) \Big]^{1/2} \Big\}.
\end{aligned}
\tag{5.33}
$$

Inequalities (5.32) and (5.33) allow us to control the second term on the right-hand side of (5.23), i.e.,

$$
|b_{\gamma_B}(u_h; v_h, \psi)| \le C \sum_{K\in\mathcal{T}_h} \rho_K^B(u_h, v_h)\, \omega_K^B(\psi).
$$

Estimate (5.20) now follows in a straightforward way. □

Remark 5.11. *To link result* (5.20) *to the theory in [48], we seek a bound of the dual norm* $\|(E_{\varepsilon,k}^{\gamma})'(u_h, v_h)\|_{(H^1(\Omega)\times H^1(\Omega))^*}$*, where* $(H^1(\Omega)\times H^1(\Omega))^*$ *is the dual space of* $H^1(\Omega)\times H^1(\Omega)$*. Thanks to Lemma 5.9 and the discrete Cauchy–Schwarz inequality, we have*

$$
\begin{aligned}
\|(E_{\varepsilon,k}^{\gamma})'(u_h, v_h)\|_{(H^1(\Omega)\times H^1(\Omega))^*} &= \sup_{(\varphi,\psi)\in H^1(\Omega)\times H^1(\Omega)} \frac{|(E_{\varepsilon,k}^{\gamma})'(u_h, v_h; \varphi, \psi)|}{\Big[\|\varphi\|_{H^1(\Omega)}^2 + \|\psi\|_{H^1(\Omega)}^2 \Big]^{1/2}} \\
&\le C \left[\Big(\sum_{K\in\mathcal{T}_h} \lambda_{1,K}^2 \big(\rho_K^A(v_h, u_h)\big)^2 \Big)^{1/2} + \Big(\sum_{K\in\mathcal{T}_h} \lambda_{1,K}^2 \big(\rho_K^B(u_h, v_h)\big)^2 \Big)^{1/2} \right] \\
&\qquad \left[\frac{\big(\|\varphi\|_{H^1(\Omega)} + \|\psi\|_{H^1(\Omega)}\big)}{(\|\varphi\|_{H^1(\Omega)}^2 + \|\psi\|_{H^1(\Omega)}^2)^{1/2}} \right],
\end{aligned}
$$

i.e.,

$$
\begin{aligned}
&\|(E_{\varepsilon,k}^{\gamma})'(u_h, v_h)\|_{(H^1(\Omega)\times H^1(\Omega))^*} \\
&\quad \le C \left[\Big(\sum_{K\in\mathcal{T}_h} \lambda_{1,K}^2 \big(\rho_K^A(v_h, u_h)\big)^2 \Big)^{1/2} + \Big(\sum_{K\in\mathcal{T}_h} \lambda_{1,K}^2 \big(\rho_K^B(u_h, v_h)\big)^2 \Big)^{1/2} \right],
\end{aligned}
\tag{5.34}
$$

with $C = C(\mathcal{N}, C_\Delta)$*. Nevertheless, the right-hand side of* (5.34) *turns out to be a very poor error estimator in terms of driving efficient anisotropic mesh adaptation. Thus we cannot pursue the approach in [48] and use an error estimator completely independent of the choice of directions* φ *and* ψ*.*

Since estimate (5.20) holds for any pair of test functions $(\varphi, \psi) \in H^1(\Omega) \times H^1(\Omega)$, we make a specific choice, which allows us to write the left-hand side of (5.20) in terms of an energy estimate.

Corollary 5.12. *We have that*

$$E^{\gamma}_{\varepsilon,k}(u,v) - E^{\gamma}_{\varepsilon,k}(u_h,v_h) = \frac{1}{2}(E^{\gamma}_{\varepsilon,k})'(u_h,v_h;u-u_h,v-v_h) + R,$$

with R a third-order remainder in $u-u_h$, $v-v_h$.

Proof. The result follows by picking $\varphi = u-u_h$ and $\psi = v-v_h$ and using the arguments in Proposition 2.1 in [27]. □

As a consequence, the error estimator that we propose is

$$\eta(u_h,v_h) = \sum_{K\in\mathcal{T}_h} \eta_K(u_h,v_h), \tag{5.35}$$

with $\eta_K(u_h,v_h) = \rho^A_K(v_h,u_h)\,\omega^R_K(u-u_h) + \rho^B_K(u_h,v_h)\,\omega^R_K(v-v_h)$, where

$$\omega^R_K(z) = \Big[\sum_{i=1}^{2} \lambda^2_{i,K}(\mathbf{r}^T_{i,K} G^R_{\Delta_K}(z)\mathbf{r}_{i,K})\Big]^{1/2} \quad \text{with} \quad z = u-u_h, v-v_h,$$

with $G^R_{\Delta_K}(z)$ the matrix G_{Δ_K} defined as in Lemma 5.8 applied to the recovered gradient from z_h [218, 161, 162]. In particular, due to the dependence of the weights ω^R_K on the first-order derivatives of the error, we approximate the entries of matrix (5.18) by computable quantities, resorting to the well-known Zienkiewicz–Zhu recovery procedure, as detailed in [88, formula (33) and Remark 9].

5.3. The numerical procedure

The introduction of the Ambrosio–Tortorelli approximation of the functional (1.20) simplified the issue of performing a numerical simulation of the quasi-static evolution of a brittle fracture substituting the minimization on the set Γ with its phase field v. Although this approach eases the process, it is far from being trivial. Indeed, the functional (5.6) is nonconvex due to the presence of the term $v^2|\nabla u|^2$. Therefore, it is not possible, in general, to construct an algorithm with polynomial complexity guaranteeing convergence to the global minimizers. Indeed, the methods in the literature in general only ensure convergence to local minima (see [19] and references therein) and we cannot aim for better numerical results.

5.3.1. The minimization algorithm

The algorithm we resolve to minimize (5.6) is the *alternate minimization algorithm* proposed in [42]. The idea for designing this procedure is to exploit the convexity of the functional with respect to the two separate variables. Indeed, finding a minimizer of the energy functional with respect either to u or to v, keeping the other variable fixed, is relatively trivial and any convex minimization strategy can be adopted. Thus for a given a termination tolerance $0 < \texttt{VTOL} \ll 1$, the algorithm is the following:

Algorithm 5.1 The alternate minimization algorithm proposed in [42]

1. Set $k = 0$;
2. If $k = 0$, set $v^1 = 1$; else $v^1 = v(t_{k-1})$.
3. Set $i = 1$; $err = 1$;

while $err \geq \texttt{VTOL}$ **do**

4. $u^i = \underset{z \in H^1(\Omega)}{\arg\min}\, E^{\gamma}_{\varepsilon,k}(z, v^i)$;
5. $v^{i+1} = \underset{z \in H^1(\Omega)}{\arg\min}\, E^{\gamma}_{\varepsilon,k}(u^i, z)$;
6. $err = \|v^{i+1} - v^i\|_{L^\infty(\Omega)}$;
7. $i \leftarrow i + 1$;

end while

8. $u(t_k) = u^{i-1}$; $v(t_k) = v^i$;
9. $k \leftarrow k + 1$;
10. if $k > F$, stop; else goto 2.

Notice that Steps 4. and 5. involve respectively the convex minimizations in the variables u and v. In the literature, several examples of implementations of this algorithm are available (see, e.g., [39, 40, 42, 48]), while the first proof, although incomplete, of convergence of the algorithm to a local minimizer of the objective functional appeared in [39, Theorem 1]. The proof was then completed only a few years later in the work [48, Theorems 4.1 and 4.2].

With a view to the numerical implementation, we will consider the discrete counterpart of Algorithm 5.1. Since, in general, we expect the crack propagation to be a strongly anisotropic process, characterized by very steep gradients of both the fields u and v, we will resort to a finite element discretization based on anisotropic adapted meshes, driven by the a posteriori error estimator derived in Proposition 5.10. The challenge is to properly merge the minimization algorithm with an anisotropic adaptive procedure, as shown in the next section.

5.3.2. The mesh adaptive procedure

Following [160, 161, 162], we use a metric-based mesh adaptive approach (see, e.g., [117]). In particular, for a fixed accuracy tolerance `ADTOL`, we "predict" the optimal mesh with the least number of elements.

A metric is a symmetric positive-definite tensor field $\mathcal{M} : \Omega \to \mathbb{R}^{2\times 2}$ which, for any $\mathbf{x} \in \Omega$, provides the sizes that the optimal mesh should have along all the directions around $\mathbf{x}$. In practice, we approximate $\mathcal{M}$ via a piecewise constant metric on a given mesh $\mathcal{T}_h$, i.e., $\mathcal{M}|_K = \mathcal{M}_K = \mathcal{R}_K^T \mathcal{L}_K^{-2} \mathcal{R}_K$, for any $K \in \mathcal{T}_h$, where the matrices $\mathcal{R}_K$ and $\mathcal{L}_K$ share the same structure as R_K and Λ_K in Section 5.2, respectively.

Actually, there exists a strict link between metrics and meshes. We can associate with an assigned mesh $\mathcal{T}_h$, a corresponding piecewise-constant metric identified by $M_K = R_K^T \Lambda_K^{-2} R_K$, for any $K \in \mathcal{T}_h$, where matrices R_K and Λ_K are exactly the same as in Section 5.2.1. Vice versa, for a given metric $\mathcal{M}$, we can build a mesh, say $\mathcal{T}_{\mathcal{M}}$, such that $\mathcal{M}_K \equiv M_K$, for any $K \in \mathcal{T}_{\mathcal{M}}$ (for all the details, we refer, for instance, to [160, 161]).

The procedure we follow is first to derive a metric moving from the a posteriori error estimator (5.35) and then to generate the new mesh induced by this metric via a metric-based mesh generator. In particular, we exploit the function `adaptmesh` in `FreeFem++` [132].

In the spirit of a standard predictive approach, the metric $\mathcal{M}$ is obtained via an iterative procedure. At each iteration, say i, we deal with three quantities:

i) the actual mesh $\mathcal{T}_h^{(i)}$;

ii) the new metric $\mathcal{M}^{(i+1)}$ computed on $\mathcal{T}_h^{(i)}$;

iii) the updated mesh $\mathcal{T}_h^{(i+1)}$ induced by $\mathcal{M}^{(i+1)}$.

The most tricky step is the prediction of the new metric out of the estimator $\eta(u_h, v_h) := C \sum_{K\in\mathcal{T}_h} \{\rho_K^A(v_h, u_h)\, \omega_K(u - u_h) + \rho_K^B(u_h, v_h)\, \omega_K(v - v_h)\}$. For this purpose, we suitably rewrite the local estimator as

$$\eta_K(u_h, v_h) := \mu_K \left\{ \overline{\rho}_K^A(v_h, u_h)\, \overline{\omega}_K^R(u - u_h) + \overline{\rho}_K^B(u_h, v_h)\, \overline{\omega}_K^R(v - v_h) \right\}, \tag{5.36}$$

where $\mu_K = |\widehat{K}| \left(\lambda_{1,K}\lambda_{2,K}\right)^{3/2}$ gathers all the area $|K|$ information,

$$\overline{\rho}_K^A(v_h, u_h) = \frac{\rho_K^A(v_h, u_h)}{\left(|\widehat{K}|\lambda_{1,K}\lambda_{2,K}\right)^{1/2}}, \quad \overline{\rho}_K^B(v_h, u_h) = \frac{\rho_K^B(v_h, u_h)}{\left(|\widehat{K}|\lambda_{1,K}\lambda_{2,K}\right)^{1/2}},$$

are approximately pointwise values (at least for a sufficiently fine mesh), while the new weights

$$\overline{\omega}_K^R(z) = \left[s_K\, \mathbf{r}_{1,K}^T\, \overline{G}_{\Delta_K}^R(z)\, \mathbf{r}_{1,K} + \frac{1}{s_K}\, \mathbf{r}_{2,K}^T\, \overline{G}_{\Delta_K}^R(z)\, \mathbf{r}_{2,K} \right]^{1/2} \quad \text{with} \quad z = u - u_h, v - v_h,$$

collect the anisotropic information associated with K, with

$$\overline{G}^R_{\Delta_K}(\cdot) = G^R_{\Delta_K}(\cdot)/(|\widehat{K}|\,\lambda_{1,K}\lambda_{2,K}).$$

Following [162, Section 4], we properly merge the two terms in (5.36) to deal with a single metric. This yields

$$\eta_K(u_h, v_h) = \mu_K \Upsilon_K$$

with

$$\Upsilon_K = \Big[s_K\, \mathbf{r}^T_{1,K}\, \Gamma_K\, \mathbf{r}_{1,K} + \frac{1}{s_K}\, \mathbf{r}^T_{2,K}\, \Gamma_K\, \mathbf{r}_{2,K} \Big]^{1/2}, \tag{5.37}$$

where the local matrix

$$\Gamma_K = \left[\overline{\rho}^A_K(v_h, u_h)\right]^2 \overline{G}^R_{\Delta_K}(u - u_h) + \left[\overline{\rho}^B_K(v_h, u_h)\right]^2 \overline{G}^R_{\Delta_K}(v - v_h) \tag{5.38}$$

merges the anisotropic information provided by u and v suitably weighted via the local residuals. In this way, we are able to grasp all the directional features induced by u and v, thus avoiding the metric intersection issue.
Now to minimize the number of mesh elements, we equivalently maximize the area of each element K with the *equidistribution* constraint, i.e., for each element $K \in \mathcal{T}_h^{(j+1)}$, $\eta_K(u_h, v_h) = \mu_K\, \Upsilon_K = \texttt{TOL}/\#\mathcal{T}_h^{(j)}$, where $\texttt{TOL}$ and $\#\mathcal{T}_h^{(j)}$ are the fixed global tolerance and the number of mesh elements in $\mathcal{T}_h^{(j)}$, respectively. The maximization is achieved by minimizing the quantity Υ_K with respect to s_K and $\mathbf{r}_{1,K}$, i.e., by solving elementwise the constrained minimization problem

$$\min_{s_K \geq 1, \mathbf{r}_{i,K}\cdot\mathbf{r}_{j,K} = \delta_{ij}} \Upsilon_K(\mathbf{r}_{1,K}, s_K), \tag{5.39}$$

δ_{ij} being the Kronecker symbol. Notice that all the quantities involved in (5.38) are computed on the background grid $\mathcal{T}_h^{(j)}$. On the other hand, the aspect ratio s_K and the unit vectors $\mathbf{r}_{i,K}$ in (5.37) represent our actual unknowns.
According to [162, Proposition 4.2], we can state the desired minimization result as

Proposition 5.13. *Let $\{\boldsymbol{\gamma}_{i,K}, g_{i,K}\}$ be the eigenvector-eigenvalue pair of Γ_K with $g_{1,K} \geq g_{2,K} > 0$. Then the minimum (5.39) is obtained for the choices*

$$\boldsymbol{r}_{1,K} = \boldsymbol{\gamma}_{2,K} \quad \textit{and} \quad s_K = \left(\frac{g_{1,K}}{g_{2,K}}\right)^{1/2}, \tag{5.40}$$

yielding the value $\left(2\sqrt{g_{1,K} g_{2,K}}\right)^{1/2}$ for Υ_K.

Notice that the minimization problem (5.39) is not a computational overhead, since it can be solved analytically via (5.40). Moreover, we observe that the optimal weight Υ_K does not depend any more on the aspect ratio.

Finally, the optimal metric $\mathcal{M}^{(j+1)}$ is obtained by exploiting the equidistribution constraint, i.e., by solving the equations

$$|\widehat{K}|\left(\lambda_{1,K}\lambda_{2,K}\right)^{3/2}\left(2\sqrt{g_{1,K}g_{2,K}}\right)^{1/2}=\frac{\texttt{TOL}}{\#\mathcal{T}_h^{(j)}}\quad\text{and}\quad\frac{\lambda_{1,K}}{\lambda_{2,K}}=s_K=\left(\frac{g_{1,K}}{g_{2,K}}\right)^{1/2}. \tag{5.41}$$

System (5.41) provides us with the distinct values

$$\lambda_{1,K}=\left(\frac{1}{|\widehat{K}|\sqrt{2}}\left(\frac{g_{1,K}}{g_{2,K}^2}\right)^{1/2}\frac{\texttt{TOL}}{\#\mathcal{T}_h^{(j)}}\right)^{1/3},\quad\lambda_{2,K}=\left(\frac{1}{|\widehat{K}|\sqrt{2}}\left(\frac{g_{2,K}}{g_{1,K}^2}\right)^{1/2}\frac{\texttt{TOL}}{\#\mathcal{T}_h^{(j)}}\right)^{1/3}. \tag{5.42}$$

Eventually, the optimal metric $\mathcal{M}^{(j+1)}$ is characterized by $\mathbf{r}_{1,K}$ in (5.40), $\lambda_{1,K}$ and $\lambda_{2,K}$ in (5.42), with $\mathbf{r}_{2,K}\perp\mathbf{r}_{1,K}$.

5.3.3. The whole adaptive procedure

The next step is to design a strategy to efficiently combine the discrete counterpart of Algorithm 5.1 with the mesh adaptive procedure. In particular, we propose two algorithms, which are different in the way the minimization and the mesh adaptivity are interlaced. For both algorithms, we denote by $\mathcal{T}_h$ the mesh used to start up the mesh adaptive procedure.

The first algorithm, which is a variant of [48, `ALGORITHM 1`], applies the mesh adaptation after convergence of the minimization algorithm on both u_h and v_h. In particular, after fixing a termination tolerance `VTOL` $\ll 1$ for the minimization algorithm, a relative tolerance `MESHTOL` $\ll 1$ on the change of the mesh cardinality, and `ADTOL` $\ll 1$ which fixes the accuracy on the functional (5.20), the algorithm is the following:

Algorithm 5.2 Optimize-then-Adapt

Set $k=0$, $j=0$, $\mathcal{T}_h^{(0)}=\mathcal{T}_h$;
If $k=0$, set $v_h^1=1$; else $v_h^1=v_h(t_{k-1})$;
Set $j=0$; errmesh$=1$;
while errmesh $\geq$ `MESHTOL` **do**
  Set $i=1$; $err=1$;
  while err $\geq$ `VTOL` **do**
    $u_h^i=\arg\min_{z_h\in X_h^{(j)}} E_{h,k}(z_h,v_h^i)$;
    $v_h^{i+1}=\arg\min_{z_h\in X_h^{(j)}} E_{h,k}(u_h^i,z_h)$;
    err $=\|v_h^{i+1}-v_h^i\|_{L^\infty(\Omega)}$;
    $i\leftarrow i+1$;
  end while

12: Compute the new metric $\mathcal{M}^{(j+1)}$ based on u_h^{i-1} and v_h^i with TOL = ADTOL;
13: Build the adapted mesh $\mathcal{T}_h^{(j+1)}$;
14: errmesh = $|\#\mathcal{T}_h^{(j+1)} - \#\mathcal{T}_h^{(j)}|/\#\mathcal{T}_h^{(j)}$;
15: Set $v_h^1 = \Pi_{j\to j+1}(v_h^i)$;
16: $j \leftarrow j+1$;
17: **end while**
18: $u_h(t_k) = \Pi_{j-1\to j}(u_h^{i-1})$; $v_h(t_k) = \Pi_{j-1\to j}(v_h^i)$; $\mathcal{T}_h^k = \mathcal{T}_h^{(j)}$;
19: Set $\mathcal{T}_h^{(0)} = \mathcal{T}_h^k$;
20: $k \leftarrow k+1$;
21: if $k > F$, stop; else goto 2.

The convergence of the mesh adaptivity is checked by monitoring the variation of the number of elements during the adaptivity process. Although this check is not rigorously sound, in practice it provides an effective stopping criterion.
An interpolation step between two successive adapted meshes is also employed before restarting any new optimization or time loop. This is carried out by a suitable interpolation operator, $\Pi_{n\to n+1}(w_h)$, which maps a finite element function w_h defined on $\mathcal{T}_h^n$ onto the new mesh $\mathcal{T}_h^{n+1}$.
This algorithm performs well if the tip of the fracture moves sufficiently slow in time. Indeed, since the coupling between optimization and adaptation is not so tight, a time-adaptivity could be desirable to restrain a fast fracture evolution. Nevertheless, time adaptivity is not able to contain the final evolution steps when the actual fracture lead to a sudden failure of the material which splits it into two separate parts. This limit can be ascribed also to the deficiency of the employed quasi-static model, which clearly fails in describing very fast dynamics.
To dampen the crack propagation, we propose a second algorithm, which introduces a tighter alternation of the optimization and mesh adaptation phases. The meaning of all the involved parameters is the same as in Algorithm 5.2.

Algorithm 5.3 Optimize-and-Adapt

1: Set $k = 0$, $\mathcal{T}_h^{(1)} = \mathcal{T}_h$;
2: If $k = 0$, set $v_h^1 = 1$; else $v_h^1 = v_h(t_{k-1})$;
3: Set $j = 0$; errmesh= 1; err= 1;
4: **while** errmesh $\geq$ MESHTOL | err $\geq$ VTOL **do**
5: Set $i = 1$; err=1;
6: **while** err $\geq$ VTOL & $i \leq$ nMIN **do**
7: $u_h^i = \arg\min_{z_h \in X_h^{(j)}} E_{h,k}(z_h, v_h^i)$;

8: $v_h^{i+1} = \arg\min_{z_h \in X_h^{(j)}} E_{h,k}(u_h^i, z_h)$;
9: err $= \|v_h^{i+1} - v_h^i\|_{L^\infty(\Omega)}$;
10: $i \leftarrow i + 1$;
11: **end while**
12: Compute the new metric $\mathcal{M}^{(j+1)}$ based on u_h^{i-1} and v_h^i;
13: Build the adapted mesh $\mathcal{T}_h^{(j+1)}$;
14: errmesh $= |\#\mathcal{T}_h^{(j+1)} - \#\mathcal{T}_h^{(j)}|/\#\mathcal{T}_h^{(j)}$;
15: Set $v_h^1 = \Pi_{j \to j+1}(v_h^i)$;
16: $j \leftarrow j + 1$;
17: **end while**
18: $u_h(t_k) = \Pi_{j-1 \to j}(u_h^{i-1})$; $v_h(t_k) = \Pi_{j-1 \to j}(v_h^i)$; $\mathcal{T}_h^k = \mathcal{T}_h^{(j)}$;
19: Set $\mathcal{T}_h^{(0)} = \mathcal{T}_h^k$;
20: $k \leftarrow k + 1$;
21: if $k > F$, stop; else goto 2.

The main difference with respect to the Algorithm 5.2 is that, through `nMIN`, the minimizer of the energy functional $E_{h,k}$ is not necessarily found after the inner **while** loop and which can be recovered by setting `nMIN` $= \infty$. Alternatively setting `nMIN` $= 1$, we alternate optimization and mesh adaptation in only one **while** loop but, in such a case, the crack evolution may be biased by the mesh which is adapted to too badly optimized fields, u_h, v_h. These values of `nMIN` represent two extreme choices. In general, we may pick any intermediate value, e.g., `nMIN` $= 7$ in the section below.

5.4. Numerical experiments

Goal of this section is to assess the robustness of the algorithms proposed in the previous section on some benchmark problems. In particular, to have a comparison solution, we choose the test-cases proposed in the literature. The next section is devoted to the comparison with the two experiments appeared in [48], while the next one reports a study made on the parameter sensitivity of the two algorithm appeared in [14]. The last section of the numerical experiments addresses the plane strain scenario described in Appendix B. In particular, we challenged the algorithms and the mesh adaptation strategy with the benchmark tests published in [39].
Compared with previous attempt of simulating quasi-static brittle fracture evolution, our works are innovative under various aspect. Although in the literature adaptive anisotropic meshes are considered (see [41]), on the one hand, we deal with the Ambrosio–Tortorelli functional instead of the Mumford–Shah functional, and on the other hand, while in [41] a heuristic Hessian-based approach is employed to drive the mesh adapta-

tion, we resort to a metric-based procedure hinging on a sound error estimator. Moreover, the main improvements with respect to [48] are both in considering anisotropic meshes, in contrast to exclusively isotropic refinement, and the design of a different algorithm which adapts the mesh at each minimization step. These apparently minor changes lead to considerable improvements both in terms of accuracy and computational costs as shown in the next section. In particular, the proposed a posteriori error estimator has two main properties: the automatically generated meshes are very fine and strongly anisotropic in a thin neighborhood of the crack, whereas they show highly isotropic behavior in a neighborhood of the crack tip. As a consequence, the resulting discretization follows very closely the propagation of the fracture, which is not significantly influenced by the discretization, delivering a physically sound prediction of the crack path, with a reasonable computational effort. These features resulted to be fundamental for the success in the plane strain experiments. Indeed, in this setting, one of the two experiments has been designed to test the capability of the algorithms to detect the proper direction of the propagation by direct comparison with theoretical predictions.

5.4.1. The antiplane experiments

Let us identify a brittle material of rectangular shape, $\Omega = (0, 2) \times (0, 2.2)$, containing a slit along $\{1\} \times [1.5, 2.2]$ (see Figure 5.2, left), that we approximate with a very thin gap $2 \cdot 10^{-5}$ thick. We apply the antiplane displacement $g(t) = -t$ on $\Omega_{D^-} = (0, 1) \times (2, 2.2)$, $g(t) = t$ on $\Omega_{D^+} = (1, 2) \times (2, 2.2)$. We perform two test cases to assess whether the fracture changes direction if the domain exhibits a weak inset, such as a hole. Indeed, for both the experiments the computational domain is the same and they differ only for the presence of a circular hole of radius 0.2, centered at $(0.3, 0.3)$ (see Figure 5.2, right).

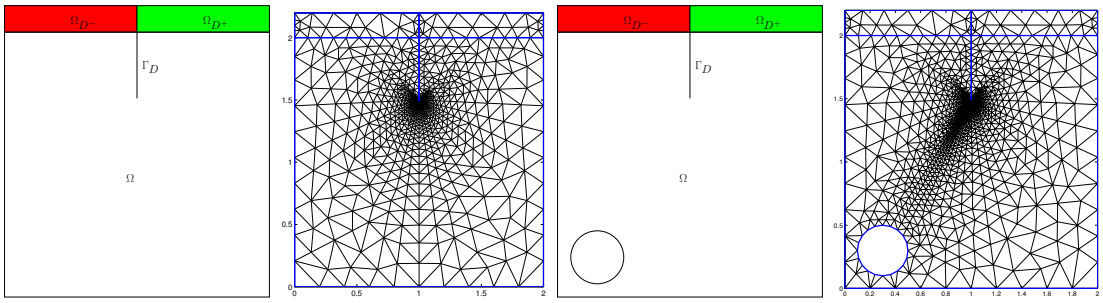


Figure 5.2.: Domain and initial mesh for the straight crack (left pair), and for the curved crack (right pair)

The straight crack

Let us start the analysis of the numerical test cases with the domain without the hole. In this circumstances, due to the perfect symmetry of this problem, we expect that the fracture does not bend but that it goes straight down starting from the tip of the slit. As an initial grid, we pick the uniform unstructured mesh in Figure 5.2, left. We consider a time window $[0, 1.5]$ sufficiently wide to contain the whole phenomenon. Concerning the parameters involved in both the algorithms, we choose the ones in Table 5.1.

Table 5.1.: The straight crack: parameters involved in Algorithms 5.2 and 5.3

$\varepsilon = 2 \cdot 10^{-2}$	$\eta = 10^{-5}$	$\gamma_A = \gamma_B = 10^{-5}$	$\Delta t = 10^{-2}$
CRTOL$= 3 \cdot 10^{-4}$	VTOL$= 2 \cdot 10^{-3}$	MESHTOL$= 10^{-2}$	ADTOL$= 10^{-2}$

Figure 5.3 compares the crack path yielded by the two algorithms. Notice that the final part of the crack delivered by Algorithm 5.3 is slightly straighter and more regular. This is likely due to the fact that Algorithm 5.2 is more sensitive to the possible coarseness of the mesh ahead of the tip. As a consequence, when the crack reaches the final stage, it tries to enter a region where the mesh has not been modified yet. Conversely, the tighter interplay between optimization and mesh adaptation in Algorihm 5.3 lets the crack find an already properly adapted mesh. An additional difference is the time when the breakdown is detected, i.e., $t = 1.36$ for Algorihm 5.2 and $t = 1.33$ in the case of Algorihm 5.3, compared with $t = 1.24$ in [48]. Indeed, since in the first algorithm we do not update the mesh during the minimization process, it can happen that the crack growing is slowed down in order to find a good compromise between the actual mesh and the fracture evolution. We additionally observe that for both Algorithms 5.2 and 5.3, the time of initiation of the fracture actually occurs later, i.e., at time $t = 0.35$, than the experiments in [48], where $t = 0.25$. We ascribe this discrepancy to the finite-width representation of the initial crack path via the vertical slit, while in [48] this is modeled via an actual one-dimensional manifold. Concerning the computational effort, the run time of Algorithms 5.2 and 5.3 is respectively $1,541.30s$ and $1,639.29s$.
Figure 5.4 shows the adatpted mesh $\mathcal{T}_h^{algo2}$ and $\mathcal{T}_h^{algo3}$ obtained by the two algorithms at the final time. The meshes, consisting of $38,299$ and $33,927$ elements respectively, exhibit really stretched elements which closely follow the crack path, whereas the mesh is very coarse in the unfractured domain, i.e., where $v_h \simeq 1$. The maximum aspect ratio is $s_K = 2,154.3$ for $\mathcal{T}_h^{algo2}$ and $s_K = 1,891.5$ for $\mathcal{T}_h^{algo3}$. The close up in Figure 5.4 at time $t = 1.21$ highlights the strongly anisotropy of the mesh far from the crack tip. We observe instead that the triangles closer to the tip are still rather isotropic. This should guarantee that the next advancing step of the crack is not biased by the directionality

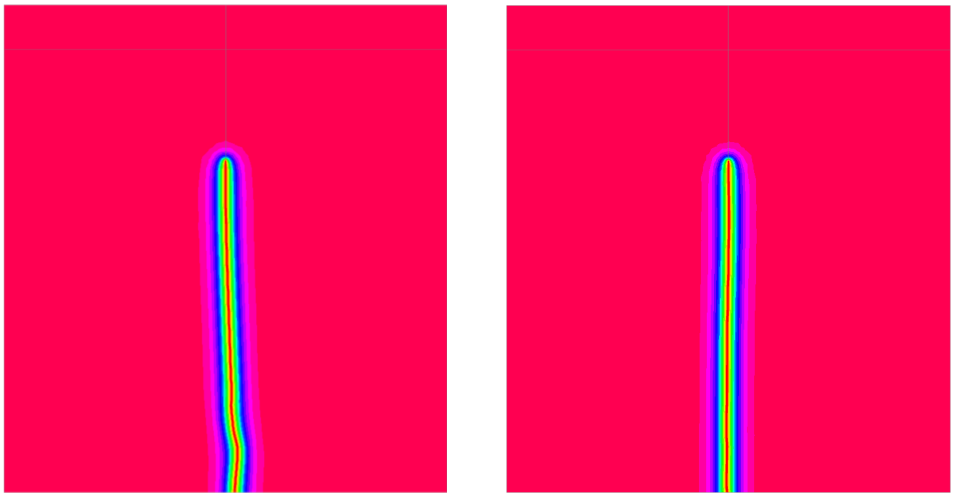


Figure 5.3.: The straight crack: v-field at the final time yielded by Algorihm 5.2 (left) and Algorihm 5.3 (right)

of the elements. After [41, 59], there has been the perception that anisotropic mesh adaptation may influence the propagation of the fracture, in particular its initiation [58]. However, it seems that the numerical procedure that we propose is in practice robust and stable thanks to its automatic capability of yielding a rounded tip.
Figure 5.5 provides the v-field and the final adapted mesh for an isotropic adaptation, obtained by enforcing $s_K = 1$ for all $K \in \mathcal{T}_h$ in Algorithm 5.2. The crack is detected also in this case even though the required number of elements is far larger, i.e., $78,025$ triangles versus $38,299$. Moreover, a slightly wavier path is exhibited with respect to Figure 5.3, left.

The curved crack

In this section we analyze the behavior of the crack in case of an asymmetric domain. The presence of the hole introduces an element of weakness in the material. As a consequence, due to energy arguments, we expect that the fracture bends its path towards the hole instead of proceedings along a straight line. As observed in [48], this test case is more challenging than the previous one. Therefore, we choose a tighter tolerance, i.e., $\texttt{REFTOL} = 10^{-3}$. The simulated crack path is very stable with respect to the choice of the parameters, as long as they are not larger than those in Table 5.1. These parameters have been properly tuned thanks to an extensive sensitivity analysis reported in Section 5.4.2 and appeared in [14].
In Figure 5.6 we show the v-field at the final time yielded by the two algorithms. In both cases, the crack enters the hole. As already observed in the previous test case,

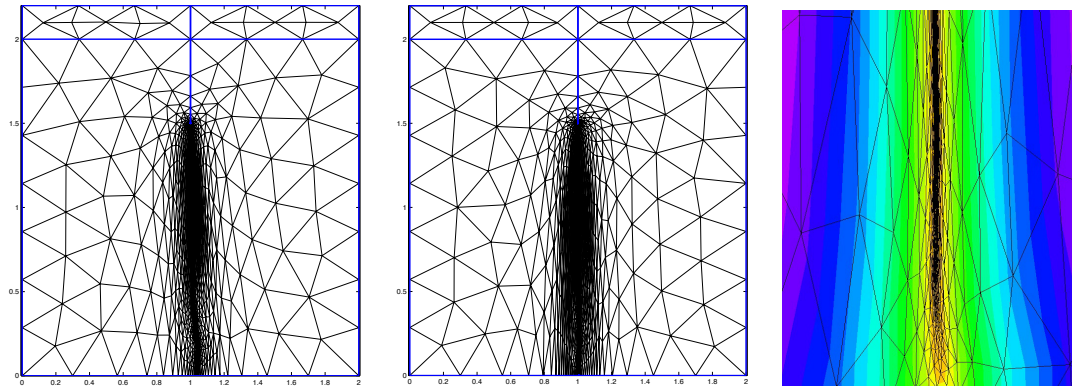


Figure 5.4.: The straight crack: final anisotropic adapted mesh provided by Algorithm 5.2 (left); final anisotropic adapted mesh (center) and zoom in (right) delivered by Algorithm 5.3

Algorithm 5.2 leads a more "bumpy" crack path ahead of the hole (compare Figure 5.6, left with Figure 5.6, center, and the corresponding zooms in).
Figure 5.7 displays the u-field superposed to the adapted meshes at $t = 1.37$ (left) and at $t = 1.43$ (right) in the case of Algorithm 5.3. A very steep ridge is evident where the tearing apart is exerted. The mesh in both cases follows very closely the crack propagation.
A top view of the final adapted meshes generated via Algorithm 5.2 and 5.3 is provided in Figure 5.8, together with a detail of the second mesh. Notice that the anisotropic adaptive procedure is able to detect the presence of a very fine structure inside the crack in correspondence with the ridges. Moreover, the cardinality of the two meshes is very different: Algorithm 5.2 employs $48,599$ elements in contrast to Algorithm 5.3 which demands only $15,987$ triangles. The maximum aspect ratio is $s_K = 1,525.3$ for $\mathcal{T}_h^{algo2}$ and $s_K = 1,469.9$ for $\mathcal{T}_h^{algo3}$.
Figure 5.9 shows four snapshots close to the material failure by comparing four successive iterations of Algorithm 5.2 (top) with Algorithm 5.3 (bottom). In the case of Algorithm 5.2, the crack, after entering the hole, reaches a region where the mesh is still coarse. Afterwards, the grid is correctly refined. On the contrary, the strategy optimize-and-adapt detects a sharper path already on entering the hole. Moreover, before converging to the failure of the material, two possible paths, energetically equivalent, pop out past the hole.
Figure 5.10, left provides the time evolution of the energy, $E_e + E_f$ in (5.1). The energy constantly increases. During the very first phase, when the crack has not started yet, the dominant contribution to the energy is E_e. Successively, after the onset of the crack propagation at $t = 0.35$, the fictitious energy E_f contributes to the whole energy. After the breakdown of the material ($t = 1.43$), the energy decreases suddenly since the elastic energy abruptly vanishes. On physical grounds we would expect the energy

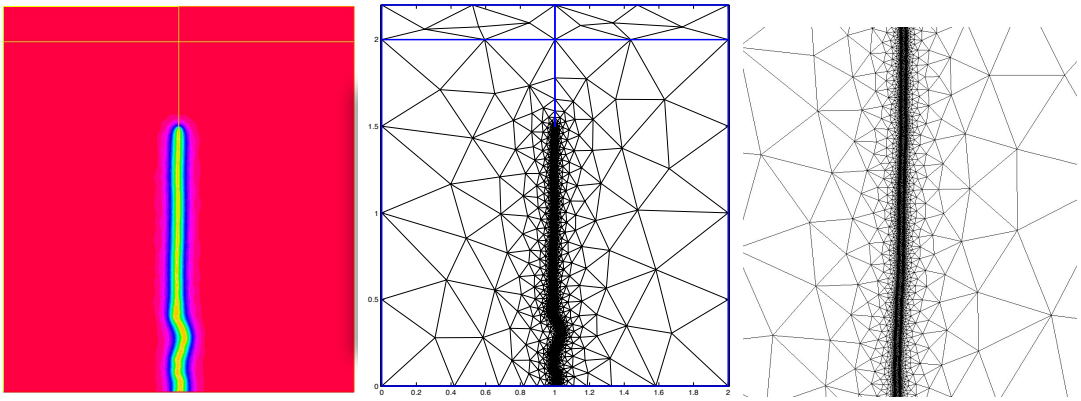


Figure 5.5.: The straight crack: v-field (left), final adapted mesh (center), and zoom in (right) at the final time in the case of the isotropic counterpart of Algorithm 5.2.

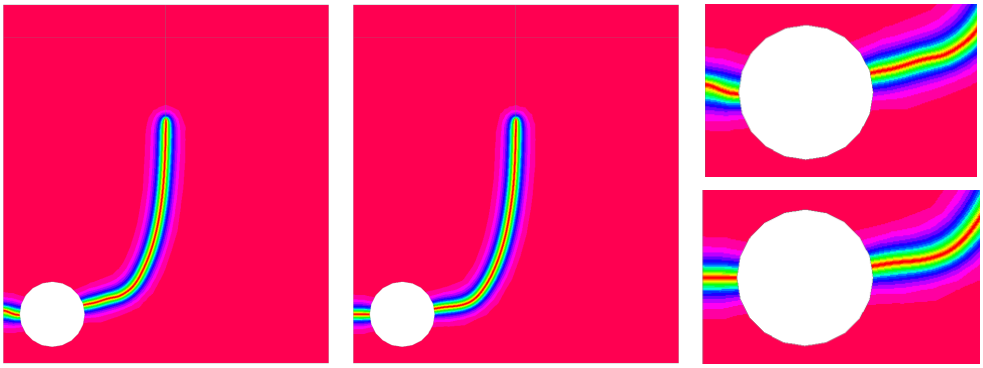


Figure 5.6.: The curved crack: v-field at the final time provided by Algorithm 5.2 (left) and by Algorithm 5.3 (center); zooms in around the hole for Algorithm 5.2 (top-right) and for Algorithm 5.3 (bottom-right)

to go to zero after the breakdown. However, the fictitious energy leaves a trace which never disappears. In Figure 5.10, right, we compare the trend of the cardinality of the anisotropic meshes associated with the two adaptive algorithms along with the isotropic counterpart of Algorithm 5.3. It is evident the saving brought by Algorithm 5.3 and the strong increase exhibited by Algorithm 5.2 in the very final phase.
Moreover, the isotropic variant of Algorithm 5.3 generates a larger number of elements throughout all the time evolution. In particular, at the final time, the isotropic mesh consists of $131,367$ triangles, i.e., about a factor 8 with respect to Algorithm 5.3.

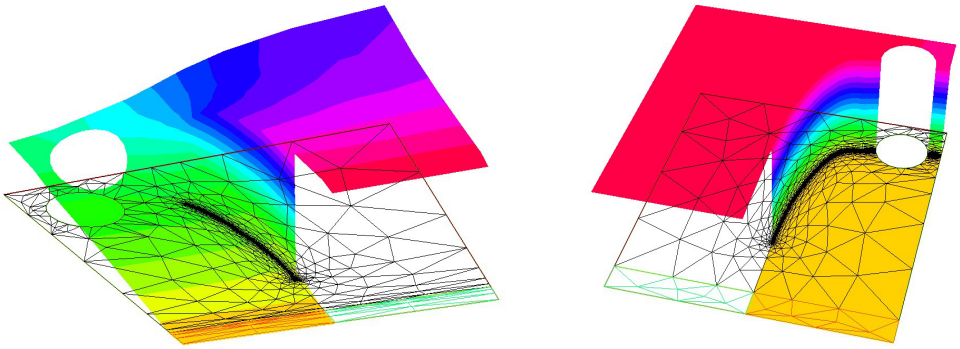


Figure 5.7.: The curved crack: u-field and adapted mesh at $t = 1.37$ (left) and at $t = 1.43$ (right) provided by Algorithm 5.3

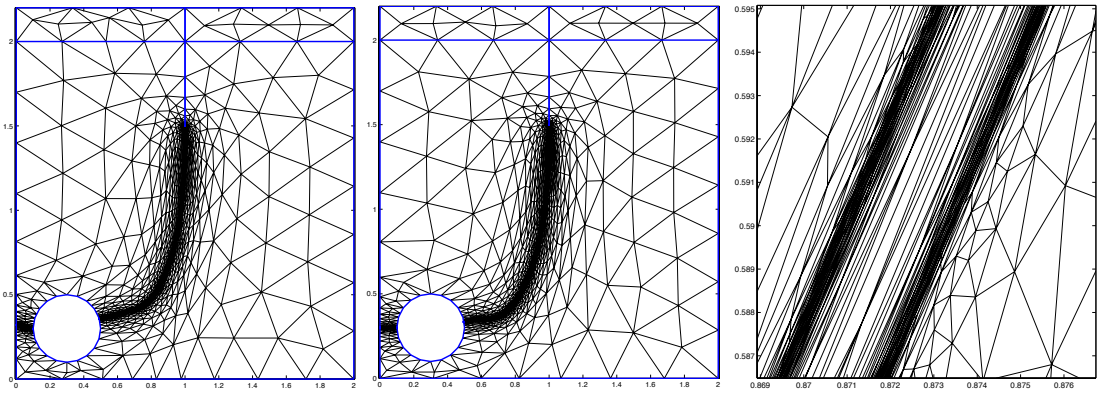


Figure 5.8.: The curved crack: final anisotropic adapted mesh provided by Algorithm 5.2 (left); final anisotropic adapted mesh (center) and zoom in (right) delivered by Algorithm 5.3

Finally in Figure 5.11 we collect the results obtained through Algorithm 5.3 when we enforce an isotropic mesh adaptation, i.e., $s_K = 1$ for all $K \in \mathcal{T}_h$. We first recognize the different path undertaken by the crack, namely, the crack leaves the hole on the bottom instead on the left. However this different path could be plausible from a physical point of view since both the paths are energetically equivalent. On the other hand, physical experiments select the one in Figure 5.6 as the most likely (see [42, 172, 48]). The alternative path in Figure 5.11 suggests that a more thorough numerical investigation should be carried out in order to properly calibrate the algorithm parameters for isotropic meshes.

As a last check, we quantify the computational performance of the two proposed algo-

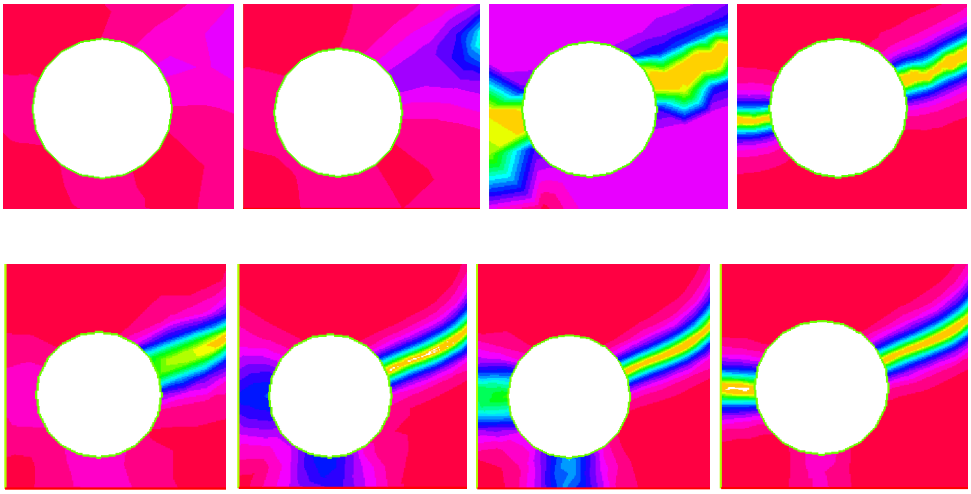


Figure 5.9.: The curved crack: successive iterations of Algorithm 5.2 (top) and Algorithm 5.3 (bottom) in the breakdown phase

rithms with respect to isotropic adaptation. In particular, for both Algorithm 5.2 and 5.3, we report the CPU times, in seconds, at each time level, associated with the optimization and adaptation phases (see Fig. 5.12 and Table 5.2). The CPU times have been recorded by using a MacBook Pro equipped with a 2.6GHz Intel Core i7 processor, 8GB of RAM, 1600MHz DDR3. As expected, the isotropic procedure is more demanding, requiring more than twice the time for the anisotropic adaptation. Concerning the computational effort characterizing the two phases, while in Algorithm 5.2 they are quite comparable, the adaptation phase of Algorithm 5.3 is about 1.8 times more expensive than the optimization phase. Algorithm 5.2 slightly outperforms Algorithm 5.3 in the anisotropic case, while the two algorithms take about the same total CPU time. Qualitatively, the trend of the bars is more varying in the adaptation phase of Algorithm 5.3.

Table 5.2.: The curved crack: cumulative CPU time [s] involved in the minimization and adaptation phases for both Algorithms 5.2 and 5.3

	Algorithm 2			Algorithm 3		
	Minimization	Adaptation	Total	Minimization	Adaptation	Total
Aniso	217	245	462	192	344	536
Iso	646	605	1,251	452	765	1,217

Remark 5.14. *Condition* (5.12) *essentially ensures a discrete min-max principle for*

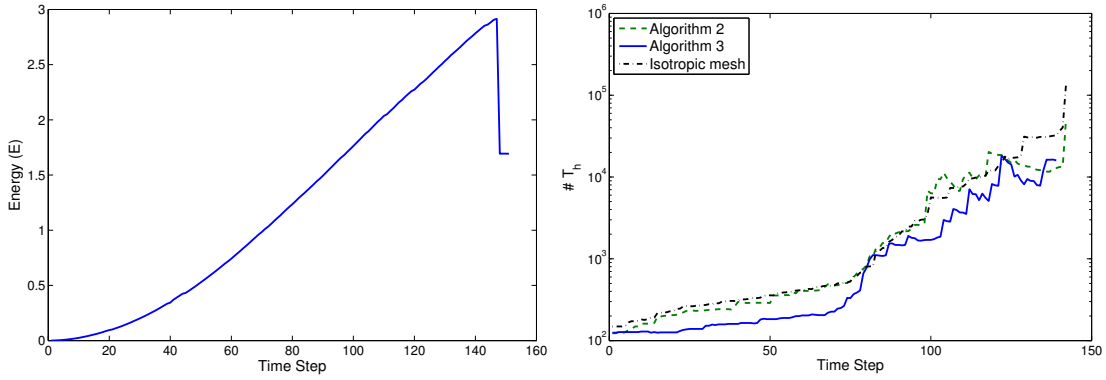


Figure 5.10.: The curved crack: energy (left) and cardinality (right) evolution for the anisotropic meshes yielded by the two Algorithms and for the isotropic counterpart of Algorithm 5.3

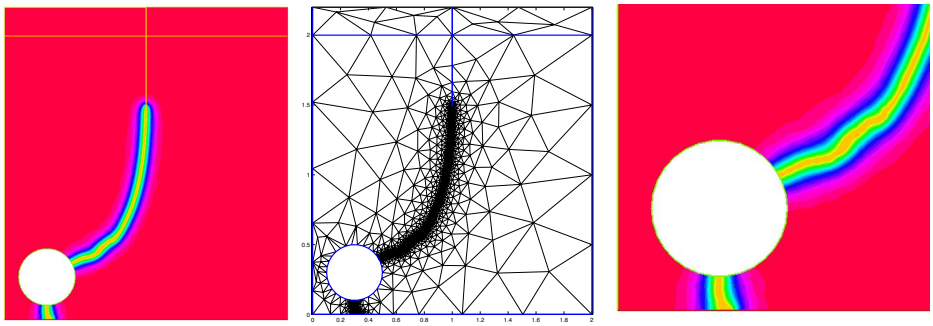


Figure 5.11.: The curved crack: v-field (left) at the final time, final adapted mesh (center) and zoom in (right) on the hole in the case of the isotropic counterpart of Algorithm 5.3

v_h, *i.e.*, $0 \leq v_h \leq 1$. *According to [211], a sufficient condition to guarantee* (5.12) *is that the mesh* $\mathcal{T}_h$ *is of Delaunay type (plus an additional constraint on the boundary elements).*
In general, this is not the case of an anisotropic grid. For this reason, we have numerically checked the possible violation of relation $0 \leq v_h \leq 1$. *The minimum value is* $4.68 \cdot 10^{-7}$, *whereas the maximum value* 1.00155 *is reached only at a single time level. The average of the maximum values of* v_h *over the time levels is* 1.0001. *Likely, this value can be related to the selected tolerances.*

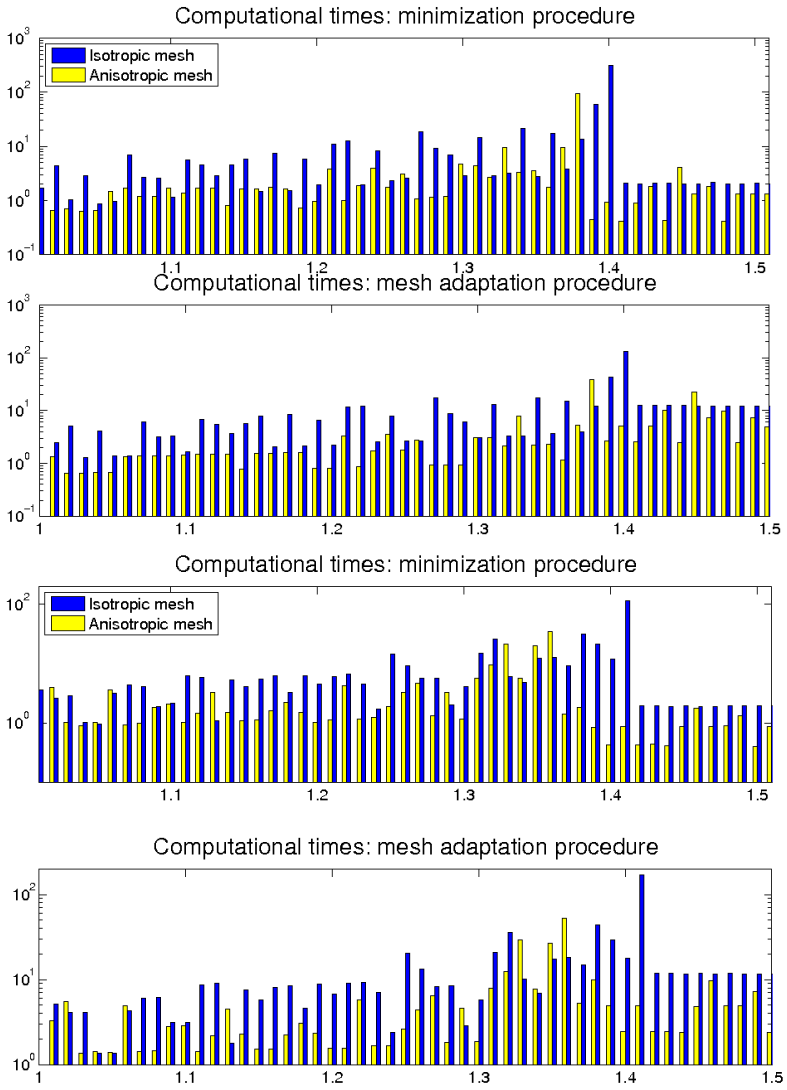


Figure 5.12.: The curved crack: CPU time [s] involved in the minimization and adaptation phases at each time level, for both Algorithms 5.2 (left) and 5.3 (right)

5.4.2. Parameter sensitivity

In this section we carry out a sensitivity analysis of the parameters characterizing the Algorithms introduced in Section 5.3 according to [14]. In particular, we refer only to the Algorithm optimize-and-adapt for sake of synthesis since the numerical results we achieved are similar. To test the change in the response of the algorithm depending on its main parameters, we considered the most challenging of the two experiments presented

in the previous section, i.e. the one in which the domain is weakened by a hole (see Figure 5.2 right). For the rest of the section, we refer to the values of Table 5.1 as default parameters which yield to the crack path reported in Figure 5.6 and the mesh described in Figure 5.8.

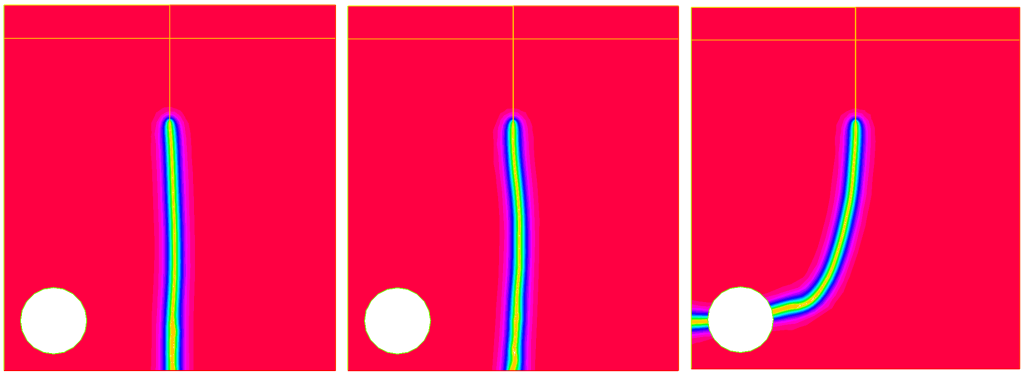


Figure 5.13.: Sensitivity to the penalty constants: colour plot of the v_h-field for $\gamma_A = \gamma_B = 10^{-4}$ (left), $\gamma_A = \gamma_B = 5 \cdot 10^{-5}$ (center), $\gamma_A = \gamma_B = 10^{-5}$ (right)

The first series of tests check on the sensitivity to the penalty constants $\gamma_A = \gamma_B$, by choosing three pairs of values, i.e., 10^{-4}, $5 \cdot 10^{-5}$, 10^{-5}. From Figure 5.13, it is evident that the higher the values of these constants, the larger is the deviation of the crack path with respect to the one assumed as default. In particular, with the first two choices the crack even misses the hole. The two meshes yielding the straight path consist of fewer elements ($12,027$ and $12,628$) than the default mesh in Figure 5.6.

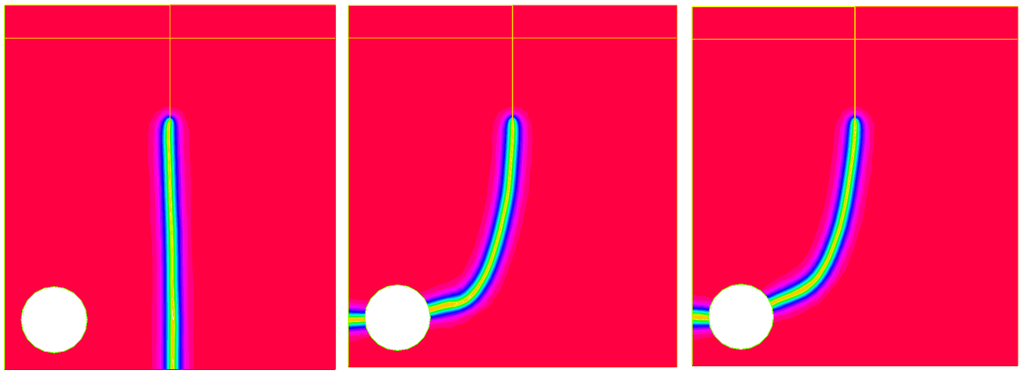


Figure 5.14.: Sensitivity to `REFTOL`: colour plot of the v_h-field for `REFTOL`$= 10^{-1}$ (left), `REFTOL`$= 10^{-3}$ (center), `REFTOL`$= 10^{-4}$ (right)

The second trial of checks deals with the sensitivity to the tolerance `REFTOL` involved in the mesh adaptation procedure. We choose both a larger and a smaller value with

respect to the default, namely REFTOL$= 10^{-1}$ and REFTOL$= 10^{-4}$. The associated v_h-field are displayed in Figure 5.14. The largest value leads to a wrong path detection with only $8,547$ triangles, whereas the choice REFTOL$= 10^{-4}$ identifies essentially the same path as the default one, but with an excessive number of elements $(23,521)$. Thus, it seems that a too small tolerance just increases the computational effort without improving the crack path tracking.

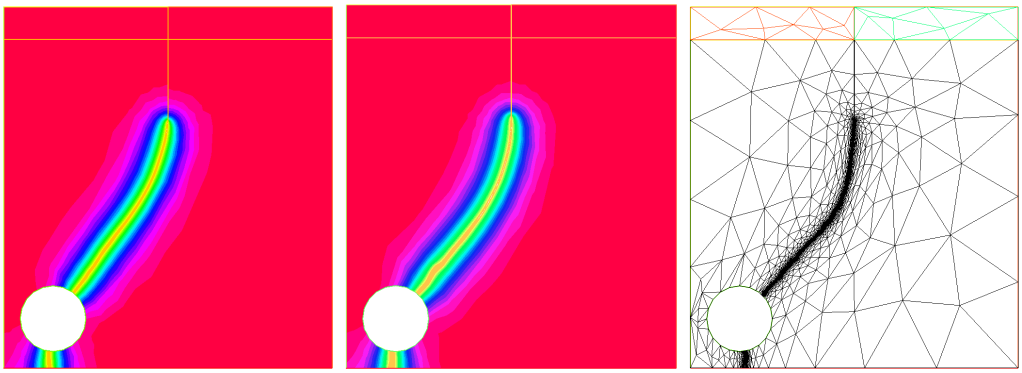


Figure 5.15.: Colour plot of the v_h-field for REFTOL$= 10^{-1}$ (left), REFTOL$= 10^{-2}$ (center), and adapted mesh for $t = 1.43$ and REFTOL$= 10^{-2}$ (right)

The last batch of tests assesses the behavior of the optimize-and-adapt algorithm for a different value of ε, i.e., $\varepsilon = 5 \cdot 10^{-2}$. We observe that ε controls the width of the crack. As expected, the larger value of ε widens the crack boundaries (compare the thickness of the crack in Figures 5.14 and 5.15). Moreover, also the crack trajectory changes considerably. For $\varepsilon = 5 \cdot 10^{-2}$ the crack suddenly turns left entering directly the hole, independently of the two chosen tolerances REFTOL$= 10^{-1}, 10^{-2}$. Although from a physical viewpoint the behavior seems correct, the bending of the actual path occurs too early and the crack leaves the hole downward instead to the left. A cross-comparison between Figures 5.14 and 5.15 leads to argue that for $\varepsilon = 5 \cdot 10^{-2}$ the value of REFTOL is not so crucial in identifying the actual path of the crack.
The assessment above seems to confirm that there is an actual sensitivity of the crack behavior to the parameters involved in both the energy functional $E_{h,k}$ and in the algorithms.

5.4.3. The plane experiments

The last tests we challenged the mesh adaptation strategy for the detection of the fracture crack path described in Section 5.3 are inspired by [42, 91] where the plane displacement is considered. In this case, the energy density has the form (1.17) and therefore the a posteriori error estimators must be adapted to this different setting. In particular, we have to deal with three different residuals: two are given by the displacement vector $\mathbf{u}$

(one for each component) and the last is given by the phase field v. This analysis has appeared in [16] and reported, for sake of completeness, in Appendix B. These tests play a key role in validating the reliability and the applicability of anisotropic mesh adaptation in the context of quasi-static crack path detection. Indeed, for assessing the quality of our results we can count on previous precise studies of the behaviour of the fracture, both from numerical and physical viewpoints [11, 42].
The numerical experiments in the following show that the proposed method is very stable and it allows us to reproduce all the previously obtained predictions on fracture development, in particular its directionality features. Additionally, we expect that our method, based on an extremely careful tuning of the anisotropic adaptation, outperforms significantly the ones used to achieve similar degrees of accuracy in previous studies. Unfortunately, in B. Bourdin et al., A. Chambolle et al., G. Del Piero et al. [42, 59, 91], where these benchmark test case are presented, there is no accurate description of the computational results and therefore we are obliged to extrapolate our positive expectation from the very fine meshes showed in the corresponding numerical sections.

Traction of a Fiber-Reinforced Matrix

We consider the rectangular domain $\Omega = (0, 3) \times (0, 3.5)$ in Fig. 5.16, comprising a nonelastic circular fiber of radius 0.5 centered at $(1.5, 1.5)$, denoted by Ω_F, and in the time span $t \in [0, 0.5]$, uniformly partitioned with a total number of $F = 50$ time steps.

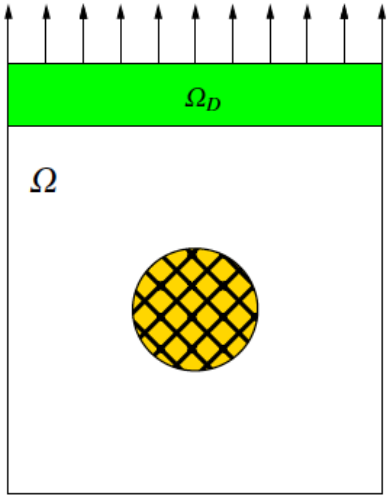


Figure 5.16.: Geometric configurations for the traction of a fiber-reinforced matrix

On the subdomain $\Omega_D = (0, 3) \times (3, 3.5)$ we enforce the load $\mathbf{g}$, with $\mathbf{g}|_{\Omega_D} = (0, t)^T$. The fiber Ω_F is held fixed, i.e. $\mathbf{u}_h \equiv 0$, a uniform vertical displacement is induced by $\mathbf{g}|_{\Omega_D}$

on the top side of the matrix while the rest of the domain $\Omega \setminus \Omega_D$ is traction-free. As a function of time, at the beginning the matrix behaves elastically; then an asymmetric crack suddenly develops and eventually cuts the matrix in two parts. The parameters involved in (B.3) are set to

$$\varepsilon = 10^{-1}, \quad \eta = 10^{-3}, \quad \gamma_A = \gamma_B = 10^{-7}, \quad \lambda = \frac{\mathbb{E}\nu}{(1+\nu)(1-2\nu)} \quad \mu = \frac{\mathbb{E}}{2(1+\nu)},$$

where $\mathbb{E} = 30$ is Young's modulus and $\nu = 0.18$ is the Poisson coefficient.
The values of the tolerances required by Algorithm 5.3 are

$$\texttt{VTOL} = 5 \cdot 10^{-3}, \quad \texttt{CRTOL} = \texttt{REFTOL} = 10^{-3}, \quad \texttt{MESHTOL} = 10^{-2}.$$

Figure 5.17 shows the v_h-field at three time levels as well as the associated anisotropic adapted mesh. At time $t = 0.25$ a crack on top of the fiber is created, and starts propagating slowly and symmetrically with respect to the fiber. At time $t = 0.35$ the symmetry is broken and the crack splits the matrix on one side only. Afterwards, at time $t = 0.39$, the domain is thoroughly split into two parts.
This behavior is not essentially affected by ε. Actually a reduction of this parameter by one order of magnitude yields the results in Fig. 5.18, which shares the same pattern as in Fig. 5.17, although with a sharper crack. In all cases the adapted meshes are very fine close to the fracture and in the area of higher stress. Moreover the correct path of the crack is detected in a very efficient way, i.e., with quite few elements. In particular, in Fig. 5.17 and 5.18 (bottom-right), the meshes consist only of $1,810$ and $12,381$ elements, respectively. The maximum aspect ratio of the three meshes in Fig. 5.17 is 16, 32 and 109.
Figure 5.19 shows the time evolution of the energy. The dashed line is associated with the elastic energy, while the dash-dotted line represent fictitious crack energy. The black line is the sum of these two contributions. Theoretically, we expect the elastic energy to disappear after the collapse of the domain. On the contrary, a residual energy remains, due to the regularization parameter η in the model. Moreover, three sudden increases of the crack energy occur: the first at time $t = 0.24$, when a finite-length crack appears on top of the fiber; the second at time $t = 0.37$, when the domain breaks on one side; and the last takes place when the domain breaks down, at $t = 0.39$. This behavior is qualitatively comparable with the ones in Fig. 4 in [91] and in Fig. 3 in [42]. This corroborates the fact that anisotropic meshes do not affect the crack dynamics.

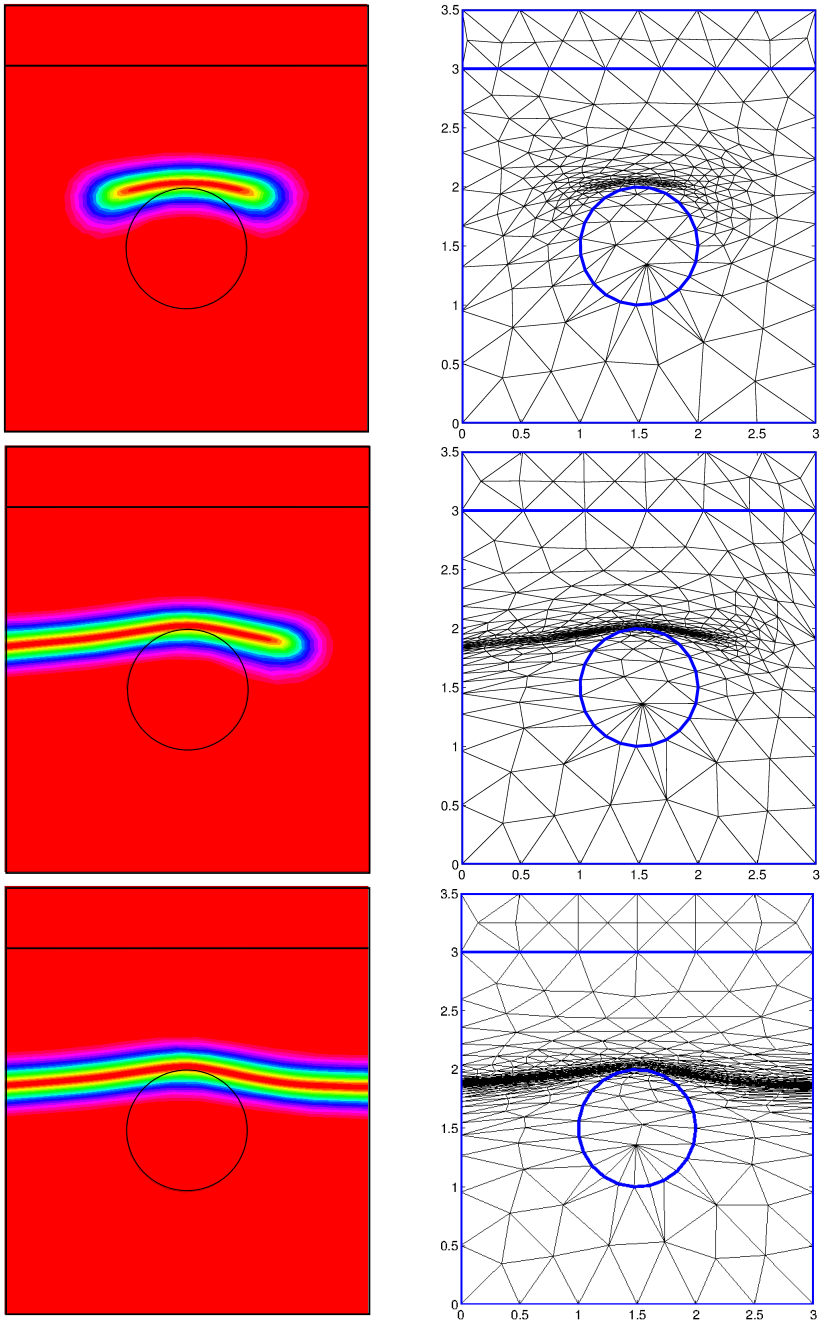


Figure 5.17.: Traction of a fiber-reinforced matrix. Time evolution of the v_h-field (left): $t = 0.25$ (top), $t = 0.35$ (center), and $t = 0.39$ (bottom); corresponding adapted meshes (right) with $\varepsilon = 10^{-1}$

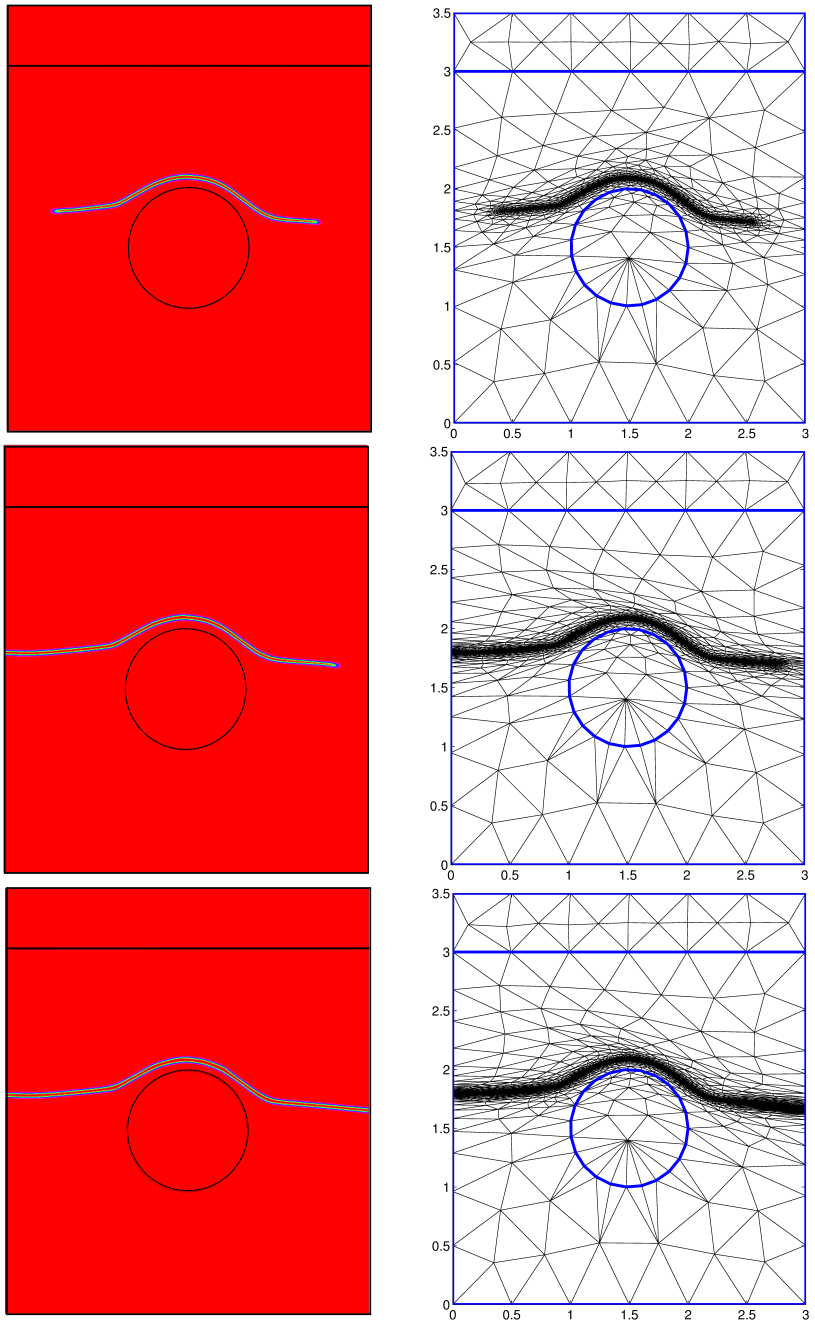


Figure 5.18.: Traction of a fiber-reinforced matrix. Time evolution of the v_h-field (left): $t = 0.30$ (top), $t = 0.38$ (center), and $t = 0.40$ (bottom); corresponding adapted meshes (right) with $\varepsilon = 10^{-2}$

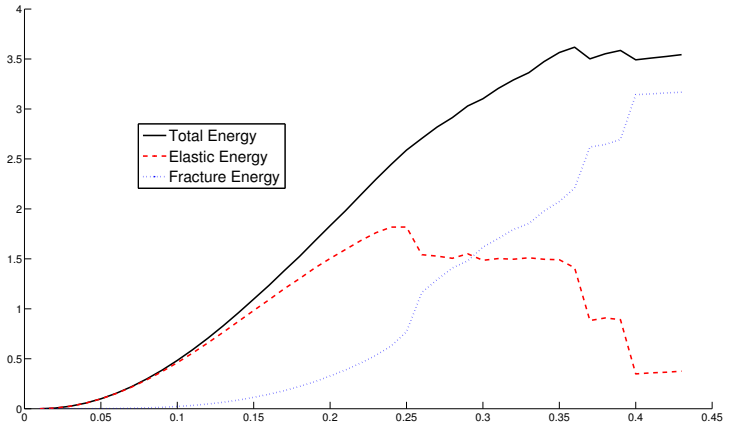


Figure 5.19.: Traction of a fiber-reinforced matrix. Time evolution of the energy

Crack Branching

The domain for the second test case is the cracked rectangular elastic sample shown in Fig. 5.20.

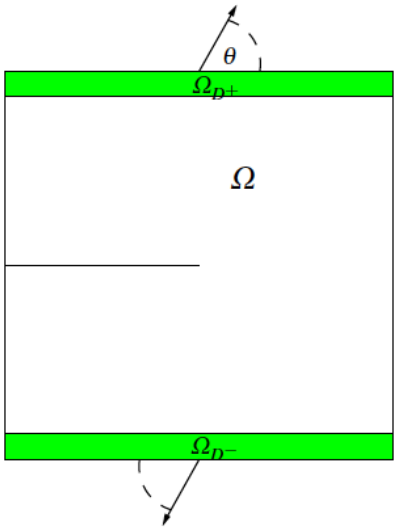


Figure 5.20.: Geometric configurations for the crack branching of the directional detection test

The initial crack is horizontal and parallel to the upper end lower sides of the sample, while a displacement field of increasing magnitude and fixed orientation, θ, to the x_1-axis, is applied to the horizontal sides. The later crack evolution is monitored for several

values of θ. The final time is set to $T = 0.2$, and the total number of uniform time steps is $F = 20$. The final time is chosen when the crack is about to turn towards the bottom right corner of the domain. The key issues of this problem is the correct prediction of the actual branching angle of the crack, in particular when the applied displacement field is not orthogonal to the domain border. For this purpose we resort to a suitable mesh adaptation strategy. In particular, we identify Ω with the square domain $(-1.5, 1.5)^2$, $\Omega_D = \Omega_{D^-} \cup \Omega_{D^+}$ with $\Omega_{D^-} = (-1.5, 1.5) \times (-1.5, -1.3)$ and $\Omega_{D^+} = (-1.5, 1.5) \times (1.3, 1.5)$, and the external displacement $\mathbf{g}$ is

$$\mathbf{g}(t) = \begin{cases} (t\cos(\theta), t\sin(\theta)) & \text{on } \Omega_{D^+} \\ (-t\cos(\theta), -t\sin(\theta)) & \text{on } \Omega_{D^-} \\ 0 & \text{elsewhere} \end{cases} \tag{5.43}$$

and the model parameters are

$$\varepsilon = 10^{-2}, \quad \eta = 10^{-5}, \quad \gamma_A = \gamma_B = 10^{-5}, \quad \lambda = \frac{\mathbb{E}\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{\mathbb{E}}{2(1+\nu)},$$

with $\mathbb{E} = 45$ and $\nu = 0.18$. The tolerances of Algorithm 5.3 are

$$\texttt{VTOL} = 10^{-4}, \quad \texttt{CRTOL} = 3 \cdot 10^{-4}, \quad \texttt{REFTOL} = 10^{-3}, \quad \texttt{MESHTOL} = 10^{-2}.$$

Figure 5.21 gathers the v_h-field and the corresponding anisotropic adapted mesh at the final time, for several orientations θ. The cardinality of the meshes in Fig. 5.21 is $2,941$, $1,268$, $1,652$, $1,302$, $1,570$, $3,804$, in top-down order. Notice that the mesh adaptive procedure identifies the configurations associated with $\theta = \pi/2$ and $\theta = 0$ as being the most challenging. In all cases the mesh closely matches the crack path, with a very thin thickness of the adapted area. The anisotropic features of the meshes are highlighted by the values of the maximum aspect ratio, which varies between 28, for $\theta = \pi/20$, and 384, for $\theta = 0$. Moreover, when $\theta = 0$, in contrast to [42], where it appears an unphysical symmetric crack branching, we obtain a crack which moves straight a very short distance, before turning downwards but with a slightly smaller angle than expected. In practice, we are able to predict reliably the crack branching for $\theta \gtrsim 3^\circ$.
Figure 5.22 shows the branching angle as a function of the orientation θ. This angle has been computed by picking the angle at which the distribution of the unit vectors, $\mathbf{r}_{1,K}$, gathered in bins of 20 angles each, over the rectangle $[0, 0.08] \times [-0.08, 0]$ is a maximum. On comparing our results with the ones in [42], we observe a good agreement, with the additional capability of correctly simulating the physical behavior for $3^\circ \lesssim \theta \lesssim 7^\circ$, by enlarging the range of reliability of the numerical tool in [42] where $\theta \gtrsim 7^\circ$.

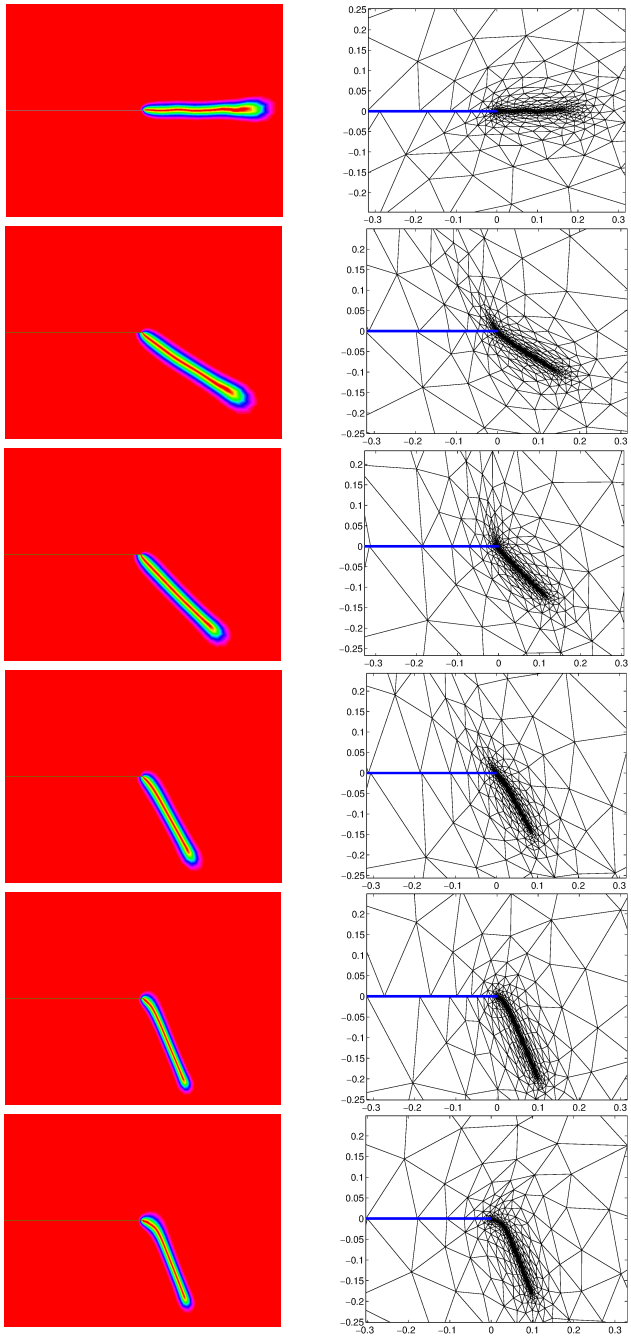


Figure 5.21.: Crack branching. Distribution of the v_h-field around the tip of the initial crack (left) and final adapted mesh (right) for $\theta = \pi/2, \pi/4, \pi/6, \pi/20, \pi/60, 0$, top-down

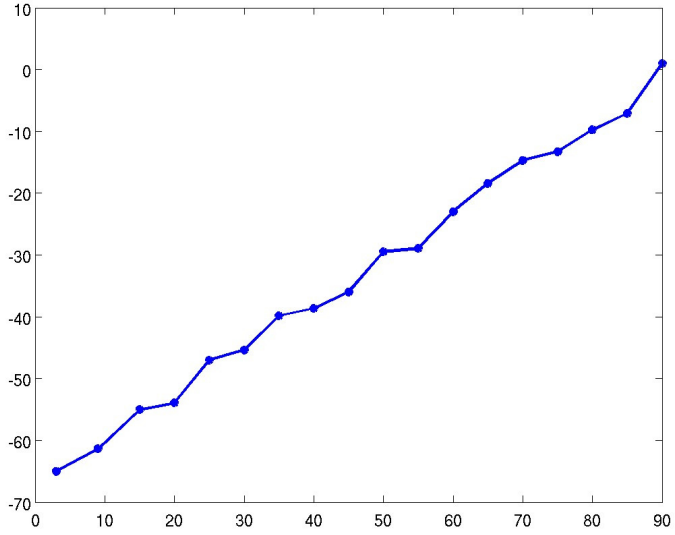


Figure 5.22.: Crack branching. Branching angle as a function of the impressed displacement orientation

5.5. Conclusions

In this chapter as well as in [14, 16, 15], we have shown that the proposed Algorithm 5.2 and 5.3 correctly identify the physical crack path, under reasonable choices of the physical and algorithmic parameters, aware also of the theoretical limits of the adopted mechanical model.
The first two experiments performed successfully the test on the crack path identification with the correct detection of the expected crack path, where the weakness and the characteristics of the domain influence the fracture evolution. The second set of numerical test showed also the stability of the algorithms with respect to the parameters proposed and confirmed that the values used for the different challenges are optimal in the sense of efficiency and accuracy of the results. The third group of experiments, in particular in the crack branching test case, additionally confirmed that the mesh adaptation procedure does not bias the crack propagation.
We can therefore conclude that the proposed anisotropic mesh adaptation is able to properly detect the expected behavior of the crack path and additionally turns out to be a stable, robust and efficient numerical procedure.
Finally, although in this work we deal with a specific case of linear elasticity constitutive law, we do believe that it is possible to extend the a posteriori analysis to a more general model, for instance, the one recently introduced in [50].

A. Complementary notions and definitions

In this section we collect definitions and notions which are necessary for the complete understanding of all the arguments of the work. However, since the content of this appendix lies outside the purpose of this work, we prefer to collect it here.
In the next section we define the functional spaces in which we generally define the displacement function. Additionally we define the concepts of Hausdorff measure, Γ-convergence, and ν-convexity.
Section A.2 introduces the fundamental notions of compressive sensing. Here we define the Null Space Property and Restricted Isometry Property on which the results described in Section 3.2.1 are based.

A.1. Functional spaces and auxiliary definitions

The Hausdorff measure

The concept of Hausdorff measures was first introduced by F. Hausdorff in [131], and it is used for a class of outer measures on subsets of a generic metric space (X, d), or for their restrictions to the corresponding measurable sets.

Definition A.1. *Let (X, d) be a metric space. For any $\Lambda \subset X$, $\delta \in (0, \infty]$ and $\alpha \in [0, \infty)$, we consider the outer measure*

$$\mathcal{H}^{\alpha}_{\delta}(\Lambda) := \omega_{\alpha} \inf \left\{ \sum_{i=1}^{\infty} (\operatorname{diam} \Lambda_i)^{\alpha} : \Lambda \subset \bigcup_i \Lambda_i \text{ and } \operatorname{diam}(\Lambda_i) < \delta \right\}, \tag{A.1}$$

where $\omega_{\alpha} > 0$ is a positive factor.
The map $\delta \mapsto \mathcal{H}^{\alpha}_{\delta}(\Lambda)$ is monotone nonincreasing. We can define the Hausdorff α-dimensional measure *of Λ as*

$$\mathcal{H}^{\alpha}(\Lambda) := \lim_{\delta \to 0} \mathcal{H}^{\alpha}_{\delta}(\Lambda). \tag{A.2}$$

Remark A.2. *If $\alpha = d \in \mathbb{N}$, the normalization constant ω_d coincides with the Lebesgue measure of the unit ball in $\mathbb{R}^d$. With this choice the d-dimensional Hausdorff measure on the euclidean space $\mathbb{R}^d$ coincides with the Lebesgue measure. In particular, if $d = 1$, $\mathcal{H}^1(\Lambda)$ measure the length of the set Λ.*

BV, SBV and SBD function spaces

The concept of Bounded Variation function and the space of Bounded Variation was first introduced by C. Jordan in [140]. Since the theory on BV functions is pretty vast, we refer to [100, Chapter 5] for the general theory.

Definition A.3. *$BV(\Omega)$ is the space of all scalar functions with bounded variation on $\Omega \subset \mathbb{R}^d$, i.e., the space of all $u : \Omega \to \mathbb{R}$ whose distributional derivative Du is (represented by) a measure in $\mathcal{M}(\Omega, \mathbb{R}^d)$.*

If we consider a function $u \in BV(\Omega)$, recalling Radon–Nykodym decomposition and considering that $|Du|$ cannot charge any $(d-1)$-negligible set, we can decompose Du into the sum of three mutually singular measures, see [4],

$$Du = \nabla u + D_c u + D_J u.$$

In particular, ∇u is the absolutely continuous part of Du with respect to Lebesgue measures, $D_J u$ is the jump part of Du, and $D_c u$ is the Cantor part. The jump part of the distributional derivative is a singular measure of the form $D_J u = f\mathcal{H}^{d-1}(J(u))$ where $f \in L^1(\Omega|_{J(u)}, \mathbb{R}^d)$.
Thus the distributional derivative may be rewritten as

$$Du = \nabla u \mathcal{L}(\Omega) + D_c u + [\![u]\!]\mathcal{H}^{d-1}(J(u)), \tag{A.3}$$

where $\mathcal{L}(\Omega)$ is the Lebesgue measure of Ω.
The space of the Special Bounded Variation functions ($SBV(\Omega)$) was introduced by E. De Giorgi and L. Ambrosio [85] to provide a weak formulation for some variational problems with free discontinuity. The theory has been developed by L. Ambrosio in [4, 5, 6] to find solution of a large class of problems, using compactness and lower semicontinuity results.

Definition A.4. *The space $SBV(\Omega)$ of special functions with bounded variation is defined as the subspace of all $u \in BV(\Omega)$ without Cantor part, i.e., the space of all u such that $D_c u = 0$. Then for every $u \in SBV(\Omega)$, (A.3) becomes*

$$Du = \nabla u \mathcal{L}(\Omega) + [\![u]\!]\mathcal{H}^{d-1}(J(u)), \tag{A.4}$$

where $\mathcal{L}(\Omega)$ is the Lebesgue measure of Ω.

In the context of fracture mechanics, the spaces of Bounded Deformation ($BD(\Omega)$) and Special Bounded Deformation ($SBD(\Omega)$) have been introduced, see [110, 203] and we may heuristically say that these spaces are the equivalent of BV and SBV for vector valued functions. More precisely:

Definition A.5. *$BD(\Omega)$ is the space of all scalar functions with bounded deformation on $\Omega \subset \mathbb{R}^d$, i.e., the space of all $\boldsymbol{u} : \Omega \to \mathbb{R}^n$ whose symmetrized gradient $\epsilon(\boldsymbol{u})$ is (represented by) a Radon measure.*

As before, we can decompose the symmetrized gradient into three parts, an absolutely continuous part $\epsilon_a(\mathbf{u})$, a jump part $[\![\mathbf{u}]\!]\mathcal{H}^{d-1}(J(\mathbf{u}))$, and a Cantor part $\epsilon_c(\mathbf{u})$, and, similarly to Definition A.4, we have that

Definition A.6. *The space $SBD(\Omega)$ of special functions with bounded deformation is defined as the subspace of all $\boldsymbol{u} \in BD(\Omega)$ without Cantor part, i.e., the space of all u such that $\epsilon_c(\boldsymbol{u}) = 0$. Then, for every $\boldsymbol{u} \in SBD(\Omega)$, the symmetrized gradient can be written as*

$$\epsilon(\boldsymbol{u}) = \epsilon_a(\boldsymbol{u})\mathcal{L}(\Omega) + [\![\boldsymbol{u}]\!]\mathcal{H}^{d-1}(J(\boldsymbol{u})). \tag{A.5}$$

Γ-convergence

The notion of Γ-convergence has been introduced by E. De Giorgi and T. Franzoni in [87], and it is intrinsically connected with applications to problems in the calculus of variations, see [46, 72] for a comprehensive introduction to the subject. Here we give a particularly general definition of Γ-convergence in a metric space.

Definition A.7. *Let $X = (X; d)$ be a metric space, and for every $h \in \mathbb{N}$ let $F_h : X \to [0; +\infty]$ be a function defined on X. We say that the sequence $\{F_h\}_h$ Γ-converges in $x_0 \in X$ to the value $r \in [0; +\infty]$ (and we write $r = \Gamma\text{-}\lim_h F_h(x_0)$) if we have:*

(i) for every sequence $\{x_h\}_h$ such that $d(x_h; x_0) \to 0$ we have

$$r \leq \liminf_h F_h(x_h);$$

(ii) there exists a sequence $\{\bar{x}_h\}_h$ such that $d(\bar{x}_h; x_0) \to 0$, and we have

$$r \geq \limsup_h F_h(x_h).$$

If the Γ-limit $\Gamma\text{-}\lim_h F_h(x)$ exists for all $x \in X$, and the function $F : X \to [0; +\infty]$ satisfies $F(x) = \Gamma\text{-}\lim_h F_h(x)$ for all $x \in X$, then we say that the sequence $\{F_h\}_h$ Γ-converges to F (on X) and we write $F = \Gamma\text{-}\lim_h F_h$.

A.1.1. Subdifferential and ν-convexity

The subdifferential may be seen as the generalization of the notion of differential for non-regular function. Indeed, the subdifferential of a given functional $\mathcal{J}$ is single-valued precisely at (Fréchet) differentiability points, where it coincides with the differential, but can be in general multivalued, or even empty. It is well-known (see, for instance [8, Chapter 1]) that it is a closed convex set.

Definition A.8. *Let $\mathcal{E}$ be an Euclidean space, $\mathcal{J}\colon \mathcal{E} \to \mathbb{R}$ a lower semicontinuous functional, and $v \in \mathcal{E}$. We say that $\xi \in \mathcal{E}' \simeq \mathcal{E}$ belongs to the* subdifferential $\partial\mathcal{J}(v)$ *of $\mathcal{J}$ at v if and only if*

$$\liminf_{w \to v} \frac{\mathcal{J}(w) - (\mathcal{J}(v) + \langle \xi, w - v \rangle)}{\|w - v\|} \geq 0\,. \tag{A.6}$$

In the special case of a convex functional $\tilde{\mathcal{J}}$, it is nonempty at every point and it can be shown (see [8, Proposition 1.4.4]) that Definition A.8 of subdifferential coincides with the more classical one, generally found on convex nonsmooth analysis:

$$\xi \in \partial\tilde{\mathcal{J}}(v) \text{ if and only if } \tilde{\mathcal{J}}(w) - (\tilde{\mathcal{J}}(v) + \langle \xi, w - v \rangle) \geq 0 \tag{A.7}$$

for every $w \in \mathcal{E}$.
In the case of a $\mathcal{C}^1$ perturbation of a lower semincontinuous functional, that is $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ where $\mathcal{J}_1$ is lower semicontinuous, and $\mathcal{J}_2$ is of class $\mathcal{C}^1$, it follows from the definition that if $\partial\mathcal{J}_1(v)$ is nonempty, then $\partial\mathcal{J}(v) \neq \emptyset$ and the decomposition

$$\partial\mathcal{J}(v) = \partial\mathcal{J}_1(v) + D\mathcal{J}_2(v)\,, \tag{A.8}$$

holds true. Here D denotes the Fréchet differential of $\mathcal{J}_2$ at v. In particular, $\mathcal{C}^1$-perturbations of lower semicontinuous convex functionals have nonempty subdifferential at every point. We collect in the following remark some useful properties of the subdifferential that are used in Chapter 2.

Remark A.9. *If $\mathcal{J}$ is a $\mathcal{C}^1$-perturbation of a convex function, it can be proved that the subdifferential enjoys the following closure property:*

$$\xi_n \in \partial\mathcal{J}(v_n), \quad v_n \to v, \quad \xi_n \to \xi \text{ implies } \xi \in \partial\mathcal{J}(v) \text{ and } \mathcal{J}(v_n) \to \mathcal{J}(v)\,. \tag{A.9}$$

The subdifferential of a convex function $\tilde{\mathcal{J}}$ is known to be a monotone operator [99], that is, for every v and $w \in \mathcal{E}$

$$\xi \in \partial\tilde{\mathcal{J}}(v) \text{ and } \omega \in \partial\tilde{\mathcal{J}}(w) \text{ implies } \langle \xi - \omega, v - w \rangle \geq 0\,. \tag{A.10}$$

We conclude this section introducing the concecpt of *ν-strongly convex* function.

Definition A.10. *A function is* ν-strongly convex *if there exists $\nu > 0$ such that*

$$\xi \in \partial\tilde{\mathcal{J}}(v) \text{ and } \omega \in \partial\tilde{\mathcal{J}}(w) \text{ implies } \langle \xi - \omega, v - w \rangle \geq \nu\|v - w\|^2\,. \tag{A.11}$$

It is well-known that this is equivalent to saying that $\tilde{\mathcal{J}}(\cdot) - \dfrac{\nu}{2}\|\cdot\|^2$ is convex.

A.2. Some notions about compressive sensing

In this section we report some definitions in the context of compressive sensing that can be useful for a better comprehension of Section 3.2.1.
In compressive sensing, we call the rank m matrix A the *encoder* which maps the N-dimensional signal x into the *measurement vector* $y \in \mathbb{R}^m$ of dimension $m \ll N$. In practice, we do not know x and wonder if it is possible to recover it somehow robustly by an efficient nonlinear *decoder* $\Delta\colon \mathbb{R}^m \to \mathbb{R}^N$. The theory only works under the assumption that the signal x is *sparse* or at least *compressible*. Let us explain formally this terminology.

Definition A.11 (k-sparse vector)**.** *Let $k \in \mathbb{N}^+$, $k \leq N$. We call the vector $x \in \mathbb{R}^N$ k-*sparse *if $x \in \Sigma_k := \left\{z \in \mathbb{R}^N | \#\operatorname{supp}(z) \leq k\right\}$, where* $\operatorname{supp}(z) := \{i \in \{1, \dots, N\} | z_i \neq 0\}$ *denotes the* support *of z.*

In applications, signals are often not exactly sparse but at least compressible, see, e.g., [157]. We define compressibility in terms of the *best k-term approximation error* with respect to the ℓ_p-norm, given by

$$\|x\|_{\ell_p} = \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p}, \quad 1 \leq p < \infty.$$

Definition A.12 (Best k-term approximation)**.** *Let x be an arbitrary vector in $\mathbb{R}^N$. We denote the* best k-term approximation *of x by*

$$x_{[k]} := \operatorname*{arg\,min}_{z \in \Sigma_k} \|x - z\|_{\ell_p}, \qquad 1 \leq p < \infty,$$

and the respective best k-term approximation error *of x by*

$$\sigma_k(x)_{\ell_p} := \min_{z \in \Sigma_k} \|x - z\|_{\ell_p} = \left\|x - x_{[k]}\right\|_{\ell_p}.$$

Remark A.13. *The best k-term approximation error is the minimal distance of x to a k-sparse vector. Informally, vectors having a relatively small best k-term approximation error are considered to be* compressible*.*

A desirable property of an encoder/decoder pair (A, Δ) is given by the following stability estimate, called *instance optimality*:

$$\|x - \Delta(Ax)\|_{\ell_p} \leq C\sigma_k(x)_{\ell_p}, \tag{A.12}$$

for all $x \in \mathbb{R}^N$, with a positive constant C independent of x, and k the closest possible to m [67]. This would in particular imply that by means of Δ we are able to recover a k-sparse signal x exactly, since in this case $\sigma_k(x)_{\ell_p} = 0$. It turns out that the existence

of such a pair restricts the range of k to be maximally of the order of $\frac{m}{\log \frac{N}{m} + 1}$. We refer to [25, 67, 94] for more details. Actually, the above mentioned condition (A.12) can be realized in practice, at least for $p = 1$, by pairing the ℓ_1-*minimization* (see (A.13) below) as the decoder with an encoder A, which has the so-called *Null Space Property* of optimal order k. (For realizations of the instance optimality in other ℓ_p-norms, for instance for $p = 2$, one needs more restrictive requirements, see [214].)

Definition A.14 (Null Space Property)**.** *A matrix $A \in \mathbb{R}^{m\times N}$ has the* Null Space Property of order k and for positive constant $\gamma_k > 0$ *if $\|z|_\Lambda\|_{\ell_1} \leq \gamma_k \|z|_{\Lambda^c}\|_{\ell_1}$, for all $z \in ker(A)$ and all $\Lambda \subset \{1, \ldots, N\}$ such that $\#\Lambda \leq k$. We abbreviate this property with the writing (k, γ_k)-*NSP*.*

The Null Space Property states that the kernel of the encoding matrix A contains no vectors where some entries have a significantly larger magnitude with respect to the others. In particular, no compressible vector is contained in the kernel. This is a natural requirement since otherwise no decoder would be able to distinguish a sparse vector from zero.

Lemma A.15. *Let $A \in \mathbb{R}^{m\times N}$ have the (k, γ_k)-NSP, with $\gamma_k < 1$, and define $\mathcal{F}(y) := \{z \in \mathbb{R}^N | Az = y\}$, the set of feasible vectors for the measurement vector $y \in \mathbb{R}^m$. Then the decoder*

$$\Delta_1(y) := \arg\min_{z\in\mathcal{F}(y)} \|z\|_{\ell_1}, \tag{A.13}$$

which we call ℓ_1-minimization*, performs*

$$\|x - \Delta_1(y)\|_{\ell_1} \leq C\sigma_k(x)_{\ell_1}, \tag{A.14}$$

for all $x \in \mathcal{F}(y)$ and the constant $C := \dfrac{2(1+\gamma_k)}{1-\gamma_k}$.

This result is by now well-known, see, e.g., [200] for a proof. Unfortunately the NSP is hard to verify in practice. Therefore one can introduce another property, called the *Restricted Isometry Property*, which implies the NSP, see [83] for a proof of this latter statement. Being a spectral concentration property, the Restricted Isometry Property is particularly suited to be verified with high probability by certain random matrices.

Definition A.16 (Restricted Isometry Property)**.** *A matrix $A \in \mathbb{R}^{m\times N}$ has the* Restricted Isometry Property (RIP) of order K with constant $0 < \delta_K < 1$ *if*

$$(1 - \delta_K)\|z\|_{\ell_2} \leq \|Az\|_{\ell_2} \leq (1 + \delta_K)\|z\|_{\ell_2},$$

*for all $z \in \Sigma_K$. We refer to this property by (K, δ_K)-*RIP*.*

Lemma A.17. *Let $k, h \in \mathbb{N}^+$ and $K = k+h$. Assume that $A \in \mathbb{R}^{m\times N}$ has (K, δ_K)-RIP. Then A has (k, γ_k)-NSP, where $\gamma_k := \sqrt{\dfrac{k}{h}\dfrac{1+\delta_K}{1-\delta_K}}$.*

The proof of this latter result can be found, for instance, in [107].

B. A posteriori anisotropic error estimator of the plane displacement

In this chapter we develop a reliable a posteriori anisotropic first order estimator for the numerical simulation of the Ambrosio–Tortorelli functional in case of brittle material subject to plane-strain. The numerical simulation of fracture initiation and evolution in this framework has been first addressed by B. Bourdin et al. [42]. In this work the authors used an extremely fine discretization to be able to capture the fracture path and its expected directional developments independently of the intrisic anisotropies of the a priori prescribed mesh. Despite the reliability of the result they obtained, the cost of an extremely fine discretization to render the material numerically homogenous is enormous, leading to the quest for possible alternative techniques based on adaptive strategies, which can break the ambiguity of *"the crack following the mesh or the mesh following the crack"*, as already pointed out in the previous chapters.

B.1. The mathematical model of plane-strain fracture

Recalling the content of Chapter 5, the correspondent of (5.1) for plane displacements is

$$E^{\mathrm{P}}_{\varepsilon,k}(\mathbf{u},v)=\frac{1}{2}\int_{\Omega}(v^2+\eta)W(\epsilon(\mathbf{u}))\,d\mathbf{x}+\frac{1}{2}\int_{\Omega}\left[\alpha(1-v)^2+\varepsilon|\nabla v|^2\right]d\mathbf{x}, \tag{B.1}$$

where $\Omega\subset\mathbb{R}^2$, and$0<\eta\ll\varepsilon\ll 1$ and $\alpha=1/(4\varepsilon)$ are the regularizing constants. In setting, the energy density function W is given by (1.17) and thus we consider only isotropic elastic material characterized by the Lamé constants λ and μ.
Introducing the time partitioning $0=t_0<\cdots<t_{N_T}=T$ of the time window $[0,T]$ as before, let $\mathbf{g}:\Omega\times[0,T]\to\mathcal{E}$ be the displacement assigned over a subset $\Omega_D\subset\Omega$ which drives the fracture onset, i.e.,

$$\mathbf{g}(\mathbf{x},t)=\begin{cases}\mathbf{g}_D(t) & \text{if } \mathbf{x}\in\Omega_D,\\ \mathbf{0} & \text{elsewhere .}\end{cases}$$

As usual, we denote by $\mathfrak{U}_k(\mathbf{g})=\{\mathbf{u}\in[H^1(\Omega)]^2\,:\,\mathbf{u}(\mathbf{x})=\mathbf{g}(\mathbf{x},t_k)\ \forall\mathbf{x}\in\Omega_D\}$ the space of the admissible configuration at time $t=t_k$.
As in Section 1.2.2, to describe the quasi-static approximation, we need to minimize the

energy functional $E^{\mathrm{P}}_{\varepsilon,k}$ in (B.1), finding the pair $(\mathbf{u}(t_k), v(t_k))$ such that

$$(\mathbf{u}(t_k), v(t_k)) \in \underset{\substack{\mathbf{u}\in\mathfrak{U}_k(\mathbf{g})\\ v\in H^1(\Omega;[0,1]),v|_{CR_{k-1}}=0}}{\arg\min} E^{\mathrm{P}}_{\varepsilon,k}(\mathbf{u}, v), \tag{B.2}$$

where $CR_{k-1} = \{\mathbf{x} \in \overline{\Omega} : v(t_{k-1}) < \texttt{CRTOL}\}$ as defined in (5.3). For simplicity we denote hereafter $\mathbf{g}(\mathbf{x}, t)$ with $\mathbf{g}(t)$.
Following Chapter 5, we relax the constraint in (B.2) with two penalization terms which lead us to rewrite the plane-strain Ambrosio–Tortorelli elasticity functional as

$$\begin{aligned} E^{\mathrm{P},\gamma}_{\varepsilon,k}(\mathbf{u}, v) = \; & \frac{1}{2}\int_\Omega (v^2+\eta)W(\epsilon(\mathbf{u}))\,d\mathbf{x} + \frac{1}{2}\int_\Omega \left[\alpha(v-1)^2 + \varepsilon|\nabla v|^2\right] d\mathbf{x} \\ & + \frac{1}{2\gamma_A}\int_{\Omega_D} |\mathbf{u}-\mathbf{g}(t_k)|^2\,d\mathbf{x} + \frac{1}{2\gamma_B}\int_{CR_{k-1}} v^2\,d\mathbf{x}, \end{aligned} \tag{B.3}$$

where γ_A and γ_B are the penalty constants. Henceforth we always deal with this functional instead of (B.1). We are dealing now with an unconstrained minimization process. At each time level we seek the pair $(\mathbf{u}(t_k), v(t_k))$ such that

$$(\mathbf{u}(t_k), v(t_k)) \in \underset{(\mathbf{u},v)\in[H^1(\Omega)]^2\times H^1(\Omega;[0,1])}{\arg\min} E^{\mathrm{P},\gamma}_{\varepsilon,k}(\mathbf{u}, v). \tag{B.4}$$

Since the penalized constraints are clearly continuous, convex, and always non-negative, the proof of the convergence of the minimizers of (B.4) to the minimizers of (B.2), for $\gamma_A, \gamma_B \to 0$, follows from Γ-convergence arguments (see [72]).
Mimicking the proof in [48] for the anti-plane case, we can prove that the functional $E^{\mathrm{P},\gamma}_{\varepsilon,k}$ is Fréchet-differentiable in $[H^1(\Omega)]^2 \times (H^1(\Omega)\cap L^\infty(\Omega))$. In particular, the Fréchet derivative of $E^{\mathrm{P},\gamma}_{\varepsilon,k}$ along direction $(\mathbf{w}, z)$ is

$$\begin{aligned} \left(E^{\mathrm{P},\gamma}_{\varepsilon,k}(\mathbf{u}, v; \mathbf{w}, z)\right)' \;=\; & \underbrace{\int_\Omega (v^2+\eta)\sigma(\mathbf{u}) : \epsilon(\mathbf{w})\,d\mathbf{x} + \frac{1}{\gamma_A}\int_{\Omega_D} (\mathbf{u}-\mathbf{g}(t_k))\cdot\mathbf{w}\,d\mathbf{x}}_{=a(v;\mathbf{u},\mathbf{w})} \\ & + \underbrace{\int_\Omega \left[v\,z\sigma(\mathbf{u}) : \epsilon(\mathbf{u}) + \alpha(v-1)z + \varepsilon\nabla v\cdot\nabla z\right] d\mathbf{x} + \frac{1}{\gamma_B}\int_{CR_{k-1}} v\,z\,d\mathbf{x}}_{=b(\mathbf{u};v,z)}, \end{aligned} \tag{B.5}$$

where we substituted the elastic energy density with its explicit form $W(\epsilon(\mathbf{u})) = \sigma(\mathbf{u}) : \epsilon(\mathbf{u})$, where σ is the stress tensor defined by the Hooke's law (1.4) for isotropic materials. We recall the definition of critical points of $E^{\mathrm{P},\gamma}_{\varepsilon,k}$:

Definition B.1. *The pair* $(\boldsymbol{u}, v) \in [H^1(\Omega)]^2 \times (H^1(\Omega)\cap L^\infty(\Omega))$ *is a* critical point *of* $E^{P,\gamma}_{\varepsilon,k}$ *if* $\left(E^{P,\gamma}_{\varepsilon,k}(\boldsymbol{u}, v; \boldsymbol{w}, z)\right)' = 0$ *for all* $\boldsymbol{w} \in [H^1(\Omega)]^2$ *and for all* $z \in (H^1(\Omega)\cap L^\infty(\Omega))$.

Following Proposition 5.4, we can prove that condition $0 \le v \le 1$ is automatically guaranteed for any critical point.

B.2. A posteriori anisotropic error analysis

This section provides the discrete approximation of the functional $E^{\mathrm{P},\gamma}_{\varepsilon,k}$ and, using tools of the anisotropic theory introduced in Section 5.2.1, we derive a posteriori error estimator to drive the anisotropic mesh adaptation described in Section 5.3.

Discretization of $E^{\mathbf{P},\gamma}_{\varepsilon,k}$

We introduce the discrete counterpart of the minimization problem (B.4) in a finite element setting. Thus, we denote by $\{\mathcal{T}_h\}_{h>0}$ a family of conforming meshes of $\overline{\Omega}$, and let N_h be the index set of the vertices of $\mathcal{T}_h$, and $\mathfrak{e}_h$ the skeleton of $\mathcal{T}_h$. Henceforth, we assume that the boundary of Ω_D coincides with the union of consecutive edges in $\mathfrak{e}_h$. We associate with $\mathcal{T}_h$ the space X_h of continuous piecewise linear finite elements [63]. We denote by $E^{\mathrm{P}}_{h,k}(\mathbf{u}_h, v_h)$ the discrete correspondent of $E^{\mathrm{P},\gamma}_{\varepsilon,k}(\mathbf{u}, v)$ in (B.3), with $\mathbf{u}_h = (u_{h,1}, u_{h,2})^T \in [X_h]^2$ and $v_h \in X_h$, given by

$$\begin{aligned} E^{\mathrm{P}}_{h,k}(\mathbf{u}_h, v_h) &= \frac{1}{2}\int_\Omega \Big[\Big(P_h(v_h^2) + \eta\Big)\,\sigma(\mathbf{u}_h) : \epsilon(\mathbf{u}_h) + \alpha P_h((v_h-1)^2) + \varepsilon|\nabla v_h|^2\Big]\, d\mathbf{x} \\ &+ \frac{1}{2\gamma_A}\sum_{i=1}^{2}\int_{\Omega_D} P_h\left((u_{h,i} - g_{h,i}(t_k))^2\right)\, d\mathbf{x} + \frac{1}{2\gamma_B}\int_{CR_{k-1}} P_h\left(v_h^2\right)\, d\mathbf{x}, \end{aligned} \tag{B.6}$$

where $P_h : C^0(\overline{\Omega}) \to X_h$ is the Lagrangian interpolant onto the space X_h, with $\mathbf{g}_h(t_k) = (g_{h,1}(t_k), g_{h,2}(t_k))^T \in [X_h]^2$ a suitable discrete approximation of $\mathbf{g}(t_k)$. In particular we pick $\mathbf{g}_h(t_k)$ such that

$$\int_{\Omega_D} \mathbf{g}_h(t_k)\cdot\mathbf{w}_h\, d\mathbf{x} = \int_{\Omega_D} \mathbf{g}(t_k)\cdot\mathbf{w}_h\, d\mathbf{x} \qquad \forall \mathbf{w}_h \in [X_h]^2, \tag{B.7}$$

i.e., $\mathbf{g}_h(t_k)$ is the $L^2(\Omega_D)$-projection of $\mathbf{g}(t_k)$ onto $[X_h]^2$. The action of the operator P_h is equivalent to a mass lumping [207].
The discrete analogue to (B.4) consists of finding the pair $(\mathbf{u}_h(t_k), v_h(t_k))$ such that

$$(\mathbf{u}_h(t_k), v_h(t_k)) \in \underset{(\mathbf{u}_h, v_h)\in[X_h]^2\times X_h}{\arg\min} E^{\mathrm{P}}_{h,k}(\mathbf{u}_h, v_h).$$

Definition B.1 can be also provided in the discrete case.

Definition B.2. *The pair* $(\boldsymbol{u}_h, v_h) \in [X_h]^2 \times X_h$ *is a* critical point *of* $E^P_{h,k}$ *if, for all*

$(\boldsymbol{w}_h, z_h) \in [X_h]^2 \times X_h$, $\left(E^P_{h,k}(\boldsymbol{u}_h, v_h; \boldsymbol{w}_h, z_h)\right)' = 0$, *where*

$$\left(E^P_{h,k}(\boldsymbol{u}_h, v_h; \boldsymbol{w}_h, z_h)\right)' =$$

$$\underbrace{\int_\Omega (P_h(v_h^2)+\eta)\sigma(\boldsymbol{u}_h) : \epsilon(\boldsymbol{w}_h)\, d\boldsymbol{x} + \frac{1}{\gamma_A}\sum_{i=1}^{2}\int_{\Omega_D} P_h\left((u_{h,i} - g_{h,i}(t_k))w_{h,i}\right)\, d\boldsymbol{x}}_{=a_h(v_h;\boldsymbol{u}_h,\boldsymbol{w}_h)}$$

$$+\underbrace{\int_\Omega \Big[P_h(v_h z_h)\sigma(\boldsymbol{u}_h) : \epsilon(\boldsymbol{u}_h) + \alpha P_h\big((v_h-1)z_h\big) + \varepsilon \nabla v_h \cdot \nabla z_h\Big] d\boldsymbol{x} + \frac{1}{\gamma_B}\int_{CR_{k-1}} P_h(v_h z_h) d\boldsymbol{x}}_{=b_h(\boldsymbol{u}_h;v_h,z_h)}$$

is the Fréchet derivative of $E^P_{h,k}$.

Thanks to the mass lumping associated with P_h and to the assumption

$$k_{ij} = \int_\Omega \nabla\xi_i \cdot \nabla\xi_j \, d\mathbf{x} \le 0 \qquad \forall i \neq j \in N_h,$$

about the stiffness matrix K, with $\{\xi_l\}_{l=1}^{\#N_h}$ the basis of X_h, the property $0 \le v_h \le 1$, related to the discrete maximum principle (see, e.g., [64, 147, 202]), can be assessed for any critical point v_h of (B.6).

A posteriori error estimator

The following proposition provides a variant on the anti-plane case addressed in Chapter 5

Proposition B.3. *Let* $(\boldsymbol{u}_h, v_h) \in [X_h]^2 \times X_h$ *be a critical point of* $E^P_{h,k}$ *according to Definition B.2. Then, for any pair of functions* $(\boldsymbol{w}, z) \in [H^1(\Omega)]^2 \times H^1(\Omega)$, *with* $\boldsymbol{w} = (w_1, w_2)^T$, *it holds*

$$\left|\left(E^{P,\gamma}_{\epsilon,k}(\boldsymbol{u}_h, v_h; \boldsymbol{w}, z)\right)'\right| \le C \sum_{K\in\mathcal{T}_h}\Big\{\sum_{i=1}^{2}\rho^A_{i,K}(v_h, \boldsymbol{u}_h)\,\omega_K(w_i) + \rho^B_K(\boldsymbol{u}_h, v_h)\,\omega_K(z)\Big\}, \quad \text{(B.8)}$$

where $C = C(\mathcal{N}, C_\Delta)$, *the residuals* $\rho^A_{i,K}(v_h, \boldsymbol{u}_h)$ *and* $\rho^B_K(\boldsymbol{u}_h, v_h)$ *are*

$$\rho^A_{i,K}(v_h, \boldsymbol{u}_h) = \|2v_h\sigma_i(\boldsymbol{u}_h)\cdot\nabla v_h\|_{L^2(K)} + \frac{1}{\lambda_{2,K}}\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)}\,\|\sigma_i(\boldsymbol{u}_h)\|_{L^2(K)}$$

$$+\frac{1}{2}\|[\![\sigma_i(\boldsymbol{u}_h)]\!]\|_{L^\infty(\partial K)}\,\|v_h^2+\eta\|_{L^2(\partial K)}\left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{1/2} + \frac{|K|^{1/2}\,h_K^2}{\lambda_{2,K}\,\gamma_A}|u_{h,i} - g_{h,i}(t_k)|_{W^{1,\infty}(K)}$$

$$+\frac{\delta_{K,\Omega_D}}{\gamma_A}\left(\|u_{h,i} - g_{h,i}(t_k)\|_{L^2(K)} + \|g_{h,i}(t_k) - g_i(t_k)\|_{L^2(K)}\right),$$

$$\rho_K^B(\boldsymbol{u}_h, v_h) = \|(\sigma(\boldsymbol{u}_h) : \epsilon(\boldsymbol{u}_h) + \alpha)v_h - \alpha\|_{L^2(K)} + \frac{\varepsilon}{2}\,\|[\![\nabla v_h]\!]\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{1/2}$$
$$+\frac{\delta_{K,CR_{k-1}}}{\gamma_B}\,\|v_h\|_{L^2(K)} + \frac{h_K^2}{\lambda_{2,K}}\left[\|\sigma(\boldsymbol{u}_h) : \epsilon(\boldsymbol{u}_h) + \alpha\|_{L^2(K)} + \frac{|K|^{1/2}\delta_{K,CR_{k-1}}}{\gamma_B}\right]|v_h|_{W^{1,\infty}(K)},$$

with $\boldsymbol{u}_h = (u_{h,1}, u_{h,2})^T$*, the weights are*

$$\omega_K(\xi) = \Big[\sum_{i=j}^{2} \lambda_{j,K}^2 (\boldsymbol{r}_{j,K}^T G_{\Delta_K}(\xi)\boldsymbol{r}_{j,K})\Big]^{1/2} \qquad \forall \xi \in H^1(\Omega),$$

where

$$[\![\sigma_i(\boldsymbol{u}_h)]\!] = \begin{cases} [\sigma_i(\boldsymbol{u}_h)\cdot\boldsymbol{n}]_e & e \in \mathfrak{e}_h \cap \Omega \\ 2(\sigma_i(\boldsymbol{u}_h)\cdot\boldsymbol{n})|_e & e \in \mathfrak{e}_h \cap \partial\Omega \end{cases}, \quad [\![\nabla v_h]\!] = \begin{cases} [\nabla v_h \cdot \boldsymbol{n}]_e & e \in \mathfrak{e}_h \cap \Omega \\ 2(\nabla v_h \cdot \boldsymbol{n})|_e & e \in \mathfrak{e}_h \cap \partial\Omega \end{cases} \tag{B.9}$$

denote the generalized jump of the i-th component of the normal Cauchy stress tensor and of the normal derivative of v_h*, respectively, with* $[\cdot]_e$ *the standard jump across* e*,* $\boldsymbol{n}$ *the unit normal vector to the generic edge in* $\mathfrak{e}_h$*,* $\sigma_i(\boldsymbol{u}_h)$ *the i-th column of* σ*,* $\boldsymbol{g}_h$ *is chosen as in* (B.7)*, and* $\delta_{K,\varpi} = 1$ *if* $K \cap \varpi \neq \emptyset$ *and* $\delta_{K,\varpi} = 0$ *otherwise, with* $\varpi \subset \Omega$*.*

Proof. Since $(\mathbf{u}_h, v_h)$ is a critical point of $E_{h,k}^{\mathrm{P}}$, we have that

$$a_h(v_h; \mathbf{u}_h, \mathbf{w}_h) = 0 \quad \forall \mathbf{w}_h \in [X_h]^2, \qquad b_h(\mathbf{u}_h; v_h, z_h) = 0 \quad \forall z_h \in X_h. \tag{B.10}$$

Moreover, from (B.5), for any pair $(\mathbf{w}, z) \in [H^1(\Omega)]^2 \times H^1(\Omega)$, it holds

$$|(E_{\epsilon,k}^{\mathrm{P},\gamma}(\mathbf{u}_h, v_h; \mathbf{w}, z))'| \leq |a(v_h; \mathbf{u}_h, \mathbf{w})| + |b(\mathbf{u}_h; v_h, z)|. \tag{B.11}$$

Now we analyze the two terms in (B.11) separately, starting from $|a(v_h; \mathbf{u}_h, \mathbf{w})|$. Thanks to (B.10), for any $\mathbf{w} \in [H^1(\Omega)]^2$ and $\mathbf{w}_h \in [X_h]^2$, we have that

$$|a(v_h; \mathbf{u}_h, \mathbf{w})| \leq |a(v_h; \mathbf{u}_h, \mathbf{w} - \mathbf{w}_h)| + |a(v_h; \mathbf{u}_h, \mathbf{w}_h) - a_h(v_h; \mathbf{u}_h, \mathbf{w}_h)|. \tag{B.12}$$

Let us focus on the first term on the right-hand side of (B.12). After splitting the integrals on the mesh elements, and by exploiting integration by parts, we get

$$\begin{aligned} |a(v_h; \mathbf{u}_h, \mathbf{w} - \mathbf{w}_h)| &= \Big|\sum_{K\in\mathcal{T}_h} \Big\{ \int_K (v_h^2 + \eta)\sigma(\mathbf{u}_h) : \epsilon(\mathbf{w} - \mathbf{w}_h)\, d\mathbf{x} \\ &\quad + \frac{1}{\gamma_A}\int_K (\mathbf{u}_h - \mathbf{g}(t_k))\cdot(\mathbf{w} - \mathbf{w}_h)\,\chi_{\Omega_D}\, d\mathbf{x}\Big\}\Big| \\ &= \Big|\sum_{K\in\mathcal{T}_h} \Big\{ \int_K -2v_h\,\sigma(\mathbf{u}_h)\,(\mathbf{w} - \mathbf{w}_h)\cdot\nabla v_h\, d\mathbf{x} + \int_{\partial K} (v_h^2 + \eta)\sigma(\mathbf{u}_h)\,(\mathbf{w} - \mathbf{w}_h)\cdot\mathbf{n}\, ds \\ &\quad + \frac{1}{\gamma_A}\int_K [(\mathbf{u}_h - \mathbf{g}_h(t_k)) + (\mathbf{g}_h(t_k) - \mathbf{g}(t_k))]\cdot(\mathbf{w} - \mathbf{w}_h)\,\chi_{\Omega_D}\, d\mathbf{x}\Big\}\Big|, \end{aligned}$$

where χ_{ϖ} denotes the characteristic function of the generic set $\varpi \subset \Omega$. To preserve the directional information, we now deal with the terms on the right-hand side component-wise. For this purpose, we define

$$a(v_h; \mathbf{u}_h, \mathbf{w} - \mathbf{w}_h) = \sum_{i=1}^{2} a_i(v_h; \mathbf{u}_h, w_i - w_{h,i}),$$

with $\mathbf{w}_h = (w_{h,1}, w_{h,2})^T$, and

$$\begin{aligned}
&a_i(v_h; \mathbf{u}_h, w_i - w_{h,i}) = \\
&\sum_{K \in \mathcal{T}_h} \Big\{ \int_K -2v_h \sigma_i(\mathbf{u}_h) \cdot \nabla v_h (w_i - w_{h,i})\, d\mathbf{x} + \int_{\partial K} (v_h^2 + \eta) \sigma_i(\mathbf{u}_h) \cdot \mathbf{n} (w_i - w_{h,i})\, ds \\
&+ \frac{1}{\gamma_A} \int_K \left[(u_{h,i} - g_{h,i}(t_k)) + (g_{h,i}(t_k) - g_i(t_k)) \right] (w_i - w_{h,i})\, \chi_{\Omega_D}\, d\mathbf{x} \Big\}.
\end{aligned}$$

Thanks to Hölder and Cauchy–Schwarz inequalities and definition (B.9), we obtain

$$\begin{aligned}
&|a_i(v_h; \mathbf{u}_h, w_i - w_{h,i})| \le \sum_{K \in \mathcal{T}_h} \Big\{ \|2v_h \sigma_i(\mathbf{u}_h) \cdot \nabla v_h\|_{L^2(K)}\, \|w_i - w_{h,i}\|_{L^2(K)} \\
&+ \frac{1}{2} \|[\![\sigma_i(\mathbf{u}_h)]\!]\|_{L^\infty(\partial K)} \|v_h^2 + \eta\|_{L^2(\partial K)}\, \|w_i - w_{h,i}\|_{L^2(\partial K)} + \frac{1}{\gamma_A} \|(w_i - w_{h,i})\, \chi_{\Omega_D}\|_{L^2(K)} \\
&\Big(\|(u_{h,i} - g_{h,i}(t_k)) \chi_{\Omega_D}\|_{L^2(K)} + \|(g_{h,i}(t_k) - g_i(t_k)) \chi_{\Omega_D}\|_{L^2(K)} \Big) \Big\}.
\end{aligned}$$

Picking $w_{h,i} = \mathcal{C}_h(w_i)$ and thanks to Lemma 5.8, we obtain

$$\begin{aligned}
&|a_i(v_h; \mathbf{u}_h, w_i - w_{h,i})| \le C \sum_{K \in \mathcal{T}_h} \Big\{ \|2v_h \sigma_i(\mathbf{u}_h) \cdot \nabla v_h\|_{L^2(K)} \\
&+ \frac{1}{2} \|[\![\sigma_i(\mathbf{u}_h)]\!]\|_{L^\infty(\partial K)} \|v_h^2 + \eta\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K} \lambda_{2,K}} \right)^{1/2} + \frac{\delta_{K,\Omega_D}}{\gamma_A} \\
&\left(\|u_{h,i} - g_{h,i}(t_k)\|_{L^2(K)} + \|g_{h,i}(t_k) - g_i(t_k)\|_{L^2(K)} \right) \Big\} \left[\sum_{j=1}^{2} \lambda_{j,K}^2 (\mathbf{r}_{j,K}^T G_{\Delta_K}(w_i) \mathbf{r}_{j,K}) \right]^{1/2}.
\end{aligned} \tag{B.13}$$

Now we deal with the second term on the right-hand side of (B.12), that we bound as

$$\begin{aligned}
&|a(v_h; \mathbf{u}_h, \mathbf{w}_h) - a_h(v_h; \mathbf{u}_h, \mathbf{w}_h)| \le \Big| \int_\Omega [v_h^2 - P_h(v_h^2)] \sigma(\mathbf{u}_h) : \epsilon(\mathbf{w}_h)\, d\mathbf{x} \Big| \\
&+ \frac{1}{\gamma_A} \Big| \int_{\Omega_D} (I - P_h)((\mathbf{u}_h - \mathbf{g}_h(t_k)) \cdot \mathbf{w}_h)\, d\mathbf{x} \Big| + \frac{1}{\gamma_A} \Big| \int_{\Omega_D} (\mathbf{g}_h(t_k) - \mathbf{g}(t_k)) \cdot \mathbf{w}_h\, d\mathbf{x} \Big|.
\end{aligned} \tag{B.14}$$

We anticipate the auxiliary result which is based on the equivalence of norms on a finite-dimensional space,

$$|\varphi_h \psi_h|_{H^2(K)} \le 4\, |\varphi_h|_{W^{1,\infty}(K)}\, \|\nabla \psi_h\|_{L^2(K)} \qquad \forall \varphi_h, \psi_h \in X_h, \quad \forall K \in \mathcal{T}_h, \tag{B.15}$$

which follows by straightforward calculus. Using the definition (B.7) of $\mathbf{g}_h(t_k)$, the last term in (B.14) turns out to be zero. Considering again (B.14) componentwise, employing Hölder and Cauchy–Schwarz inequalities together with the standard isotropic estimate for the L^2-norm of the interpolation error associated with P_h, we get

$$|a_i(v_h;\mathbf{u}_h,\mathbf{w}_h)-a_{i,h}(v_h;\mathbf{u}_h,\mathbf{w}_h)|\le C\sum_{K\in\mathcal{T}_h}\Big\{\frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|(u_{h,i}-g_{h,i}(t_k))w_{h,i}|_{H^2(K)}$$
$$+\|v_h^2-P_h(v_h^2)\|_{L^\infty(K)}\,\|\sigma_i(\mathbf{u}_h)\|_{L^2(K)}\,\|\nabla w_{h,i}\|_{L^2(K)}\Big\},$$

where the constant C does not depend on the aspect ratio s_K of K. Then, we employ (B.15) together with estimate (5.15) and Lemma 5.9 with $\beta_1=\lambda_{1,K}^2$, $\beta_2=\lambda_{2,K}^2$, to obtain

$$\begin{aligned}
&|a_i(v_h;u_h,\mathbf{w}_h)-a_{i,h}(v_h;u_h,\mathbf{w}_h)|\le C\sum_{K\in\mathcal{T}_h}\Big\{\Big(\frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|u_{h,i}-g_{h,i}(t_k)|_{W^{1,\infty}(K)}\\
&+\|v_h^2-P_h(v_h^2)\|_{L^\infty(K)}\,\|\sigma_i(\mathbf{u}_h)\|_{L^2(K)}\Big)\|\nabla w_{h,i}\|_{L^2(K)}\Big\}\\
\le\;& C\sum_{K\in\mathcal{T}_h}\Big\{\Big(\frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|u_{h,i}-g_{h,i}(t_k)|_{W^{1,\infty}(K)}+\|v_h^2-P_h(v_h^2)\|_{L^\infty(K)}\,\|\sigma_i(\mathbf{u}_h)\|_{L^2(K)}\Big)\\
&\Big(\|\nabla w_{h,i}-\nabla w_i\|_{L^2(K)}+\|\nabla w_i\|_{L^2(K)}\Big)\Big\}\\
\le\;& C\sum_{K\in\mathcal{T}_h}\Big\{\Big(\frac{|K|^{1/2}\,h_K^2}{\gamma_A}\,|u_{h,i}-g_{h,i}(t_k)|_{W^{1,\infty}(K)}+\|v_h^2-P_h(v_h^2)\|_{L^\infty(K)}\,\|\sigma_i(\mathbf{u}_h)\|_{L^2(K)}\Big)\\
&\frac{1}{\lambda_{2,K}}\Big[\sum_{j=1}^{2}\lambda_{j,K}^2(\mathbf{r}_{j,K}^T G_{\Delta_K}(w_i)\mathbf{r}_{j,K})\Big]^{1/2}\Big\}.
\end{aligned}\tag{B.16}$$

Therefore, collecting (B.13) and (B.16), we are able to bound componentwise the first term on the right-hand side of (B.11), as

$$|a(v_h;\mathbf{u}_h,\mathbf{w})|\le C\sum_{K\in\mathcal{T}_h}\sum_{i=1}^{2}\rho_{i,K}^A(v_h,\mathbf{u}_h)\,\omega_K^A(w_i).$$

The estimate of the second term on the right-hand side of (B.11) can be carried out exactly as the corresponding one in the proof of Proposition 5.10, after replacing $|\nabla u_h|^2$ with $\sigma(\mathbf{u}_h):\epsilon(\mathbf{u}_h)$. This yields

$$|b(\mathbf{u}_h;v_h,z)|\le C\sum_{K\in\mathcal{T}_h}\rho_K^B(\mathbf{u}_h,v_h)\,\omega_K(z).$$

□

To make estimate (B.8) useful in practice, we have to pick the pair of functions $(\mathbf{w},z)$. Mimicking the considerations in Section 5.2.2 , we choose $\mathbf{w}=\mathbf{u}-\mathbf{u}_h$ and $z=v-v_h$.

This leads us to define the error estimator

$$\eta = \sum_{K \in \mathcal{T}_h} \eta_K(\mathbf{u}_h, v_h),$$

with

$$\eta_K(\mathbf{u}_h, v_h) = \sum_{i=1}^{2} \rho^A_{i,K}(v_h, \mathbf{u}_h)\, \omega_K(u_i - u_{h,i}) + \rho^B_K(\mathbf{u}_h, v_h)\, \omega_K(v - v_h), \tag{B.17}$$

where

$$\omega^R_K(z) = \Big[\sum_{i=1}^{2} \lambda^2_{i,K} (\mathbf{r}^T_{i,K} G^R_{\Delta_K}(z) \mathbf{r}_{i,K}) \Big]^{1/2} \quad \text{with} \quad z = u_1 - u_{h,1}, u_2 - u_{h,2}, v - v_h,$$

Remark B.4. *The numerical procedure to deal with this problem coincides with the one presented in Chapter 5. The sole difference with respect to the previously introduced strategy is relative to the re-arranged local estimator, that in this case is*

$$\eta_K(\boldsymbol{u}_h, v_h) = \mu_K \Big\{ \sum_{i=1}^{2} \overline{\rho}^A_{i,K}(v_h, \boldsymbol{u}_h)\, \overline{\omega}_K(u_{h,i}) + \overline{\rho}^B_K(\boldsymbol{u}_h, v_h)\, \overline{\omega}_K(v_h) \Big\}. \tag{B.18}$$

Notice that, since the displacement has two components, the information for the local estimator is retrieved by three residuals equally considered: two related to the displacement (one for each direction) and one to the phase field variable.

We refer to Section 5.4.3 for some numerical experiments performed on benchmark test cases appeared in [42].

List of Figures

1.1. Cohesive fracture energy functions 26

3.1. Truncated p-potential W_r^p and its regularization $W_r^{p,\varepsilon}$ 62
3.2. The Lipschitz continuous thresholding functions S_1^μ, $S_{3/2}^\mu$, and S_2^μ 70
3.3. The dynamics of Algorithm 2.2 73
3.4. Image denoising 74
3.5. Discrete time evolution of the Francfort–Marigo model of brittle fracture . 76
3.6. Computational times for the simulation of the 1D Francfort-Marigo model 77
3.7. Comparison of ℓ_1-minimizer and SLP-minimizer for one typical example . 78
3.8. The results of five different decoding processes 80
3.9. Decoder performances' statistics 81
3.10. The error on the relevant entries and the support identification property by knowledge of k 82
3.11. The error on the relevant entries and the support identification property by knowledge of r 83
3.12. Phase transition diagrams 84
3.13. Comparison of phase transition diagrams 85

4.1. Discrete geometry of the problem. 103
4.2. One dimensional cohesive crack evolution $R \geq 2\ell$ 117
4.3. One dimensional cohesive crack energy growth with $R \geq 2\ell$ 117
4.4. One dimensional cohesive crack evolution $R \leq 2\ell$ 118
4.5. One dimensional cohesive crack energy growth with $R \leq 2\ell$ 119
4.6. Two dimensional cohesive crack evolution $R \geq 2\ell$ (1) 120
4.7. Two dimensional cohesive crack evolution $R \geq 2\ell$ (2) 120
4.8. Two dimensional cohesive crack evolution $R \leq 2\ell$ (1) 121
4.9. Two dimensional cohesive crack evolution $R \leq 2\ell$ (2) 121
4.10. Elastic deformation of a beam 134
4.11. Pre-fractured deformation of a beam 134
4.12. Fractured deformation of a beam 135
4.13. Energy evolution of the system 136
4.14. Iterations at each time step needed for the convergence of Algorithm 4.1 and Algorithm 4.2 137

5.1. Geometric quantities associated with the map T_K . . . 146
5.2. Domain and initial mesh . . . 160
5.3. The straight crack: v-field at the final time . . . 162
5.4. The straight crack: final anisotropic adapted mesh . . . 163
5.5. The straight crack: isotropic mesh . . . 164
5.6. The curved crack: v-field at the final time . . . 164
5.7. The curved crack: u-field and adapted mesh . . . 165
5.8. The curved crack: final anisotropic adapted mesh . . . 165
5.9. The curved crack: the istant before the failure . . . 166
5.10. The curved crack: energy and cardinality evolution . . . 167
5.11. The curved crack: v-field at the final time . . . 167
5.12. The curved crack: CPU time . . . 168
5.13. Sensitivity to the penalty constants . . . 169
5.14. Sensitivity to `REFTOL` . . . 169
5.15. Sensitivity to ε . . . 170
5.16. Geometric configurations for the traction of a fiber-reinforced matrix . . . 171
5.17. Traction of a fiber-reinforced matrix: $\varepsilon = 10^{-1}$ . . . 173
5.18. Traction of a fiber-reinforced matrix: $\varepsilon = 10^{-2}$ . . . 174
5.19. Fiber-reinforced matrix: time evolution of the energy . . . 175
5.20. Geometric configurations for the crack branching . . . 175
5.21. Crack branching. Distribution of the v_h-field for different θ . . . 177
5.22. Crack branching: angle as a function of the displacement orientation . . . 178

Bibliography

[1] G. Albi, M. Herty, and L. Pareschi. Kinetic description of optimal control problems and applications to opinion consensus. *Commun. Math. Sci.*, to appear, 2014.

[2] B. Alexeev and R. Ward. On the complexity of Mumford-Shah-type regularization, viewed as a relaxed sparsity constraint. *IEEE Trans. Image Process.*, 19(10):2787–2789, 2010.

[3] B. Almeroth, S. Krone, and G. Fettweis. Analyzing the signal-to-noise ratio of direct sampling receivers. In *ICC'13*, pages 4561–4565. IEEE, 2013.

[4] L. Ambrosio. A compactness theorem for a new class of functions of bounded variation. *Boll. Unione Mat. Ital.*, VII(4):857–881, 1989.

[5] L. Ambrosio. Variational problems in SBV and image segmentation. *Acta Appl. Math.*, 17(1):1–40, 1989.

[6] L. Ambrosio. Existence theory for a new class of variational problems. *Arch. Rational Mech. Anal.*, 111(4):291–322, 1990.

[7] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free-Discontinuity Problems.* Oxford Mathematical Monographs. Oxford: Clarendon Press. xviii, 2000.

[8] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures. 2nd ed.* Basel: Birkhäuser, 2008.

[9] L. Ambrosio and V. M. Tortorelli. Approximation of functional depending on jumps by elliptic functional via γ-convergence. *Comm. Pure Appl. Math.*, 43(8):999–1036, 1990.

[10] L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Boll. Unione Mat. Ital.*, 6(1):105–123, 1992.

[11] M. Amestoy. *Propagations de Fissures en Élasticité Plane.* Thèse d'Etat, Paris, 1987.

[12] E. Arias-Castro and Y. C. Eldar. Noise folding in compressed sensing. *IEEE Signal Process. Lett.*, pages 478–481, 2011.

[13] M. Artina, F. Cagnetti, M. Fornasier, and F. Solombrino. Linearly constrained evolutions of critical points and an application to cohesive fractures. *arXiv preprint arXiv:1508.02965*, 2015.

[14] M. Artina, M. Fornasier, S. Micheletti, and S. Perotto. Anisotropic adaptive meshes for brittle fractures: Parameter sensitivity. In A. Abdulle, S. Deparis, D. Kressner, F. Nobile, and M. Picasso, editors, *Numerical Mathematics and Advanced Applications - ENUMATH 2013*, volume 103 of *Lect. Notes Comput. Sci. Engrg.*, pages 293–301. Springer International Publishing, 2015.

[15] M. Artina, M. Fornasier, S. Micheletti, and S. Perotto. Anisotropic mesh adaptation for crack detection in brittle materials. *SIAM J. Sci. Comput.*, 37(4):B633–B659, 2015.

[16] M. Artina, M. Fornasier, S. Micheletti, and S. Perotto. The benefits of anisotropic mesh adaptation for brittle fractures under plane-strain conditions. In S. Perotto and L. Formaggia, editors, *New Challenges in Grid Generation and Adaptivity for Scientific Computing*, volume 5 of *SEMA SIMAI Springer Series*. Springer International Publishing, 2015.

[17] M. Artina, M. Fornasier, and S. Peter. Damping Noise-Folding and Enhanced Support Recovery in Compressed Sensing - Extended Technical Report. *ArXiv e-prints*, Nov. 2014.

[18] M. Artina, M. Fornasier, and S. Peter. Damping noise-folding and enhanced support recovery in compressed sensing. *IEEE Trans. Signal Process.*, to appear, 2015.

[19] M. Artina, M. Fornasier, and F. Solombrino. Linearly constrained nonsmooth and nonconvex minimization. *SIAM J. Optim.*, 23(3):1904–1937, 2013.

[20] M. Artina and K. Kunisch. Efficient numerical procedures to simulate the quasistatic evolution of dugdale cohesive models. *preprint*, 2015.

[21] H. Attouch, J. Bolte, and B. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized gauss-seidel methods. *Math. Program.*, 137:91–129, 2013.

[22] H. Attouch and C. Picard. Comportement limite de problemes de transmission unilateraux a travers des grilles de forme quelconque. *Rend. Sem. Mat. Politec. Torino*, 45:71–85, 1987.

[23] Y. Au-Yeung, G. Friesecke, and B. Schmidt. Minimizing atomic configurations of short range pair potentials in two dimensions: crystallization in the Wulff shape. *Calc. Var. PDE*, 44(1-2):81–100, 2012.

[24] J. M. Ball. Some open problems in elasticity. In *Geometry, mechanics, and dynamics*, pages 3–59. Springer, New York, 2002.

[25] R. G. Baraniuk, M. Davenport, R. A. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constr. Approx.*, 28(3):253–263, 2008.

[26] G. I. Barenblatt. The mathematical theory of equilibrium cracks in brittle fracture. *Adv. in Appl. Mech.*, 7(55-129):104, 1962.

[27] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.

[28] G. Bellettini and A. Coscia. Discrete approximation of a free discontinuity problem. *Numer. Funct. Anal. Optim.*, 15(3-4):201–224, 1994.

[29] H. Bertete-Aguirre, E. Cherkaev, and M. Oristaglio. Non-smooth gravity problem with total variation penalization functional. *Geophys. J. Internat.*, 149(2):499–507, 2002.

[30] D. Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic Press, New York,, 1982.

[31] A. Blake and A. Zisserman. *Visual reconstruction*, volume 2. MIT press Cambridge, 1987.

[32] T. Blumensath and M. E. Davies. Iterative thresholding for sparse approximations. *J. Fourier Anal. Appl.*, 14(5-6):629–654, 2008.

[33] T. Blumensath and M. E. Davies. Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.*, 27(3):265–274, 2009.

[34] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet. Charachterization of Łojasiewicz inequalities. Subgradient flows, talweg, convexity. *Trans. Amer. Math. Soc.*, 362:3319–3363, 2010.

[35] E. Bonetti and G. Schimperna. Local existence for frémond's model of damage in elastic materials. *Contin. Mech. Thermodyn.*, 16(4):319–335, 2004.

[36] E. Bonetti, G. Schimperna, and A. Segatti. On a doubly nonlinear model for the evolution of damaging in viscoelastic materials. *J. Differential Equations*, 218(1):91–116, 2005.

[37] M. Bongini, M. Fornasier, O. Junge, and B. Scharf. Sparse control of alignment models in high dimension. *Networks and Heterogeneous Media*, to appear, 2014.

[38] M. J. Border, C. V. Verhoosel, M. A. Scott, T. Hughes, and C. Landis. A phase-field description of dynamic brittle fracture. *Comput. Methods Appl. Mech. Engrg.*, 217-220:77–95, 2012.

[39] B. Bourdin. Numerical implementation of the variational formulation for quasi-static brittle fracture. *Interfaces Free Bound.*, 9(3):411–430, 2007.

[40] B. Bourdin. The variational formulation of brittle fracture: numerical implementation and extensions. In *IUTAM Symposium on Discretization Methods for Evolving Discontinuities*, pages 381–393. Springer, New York, 2007.

[41] B. Bourdin and A. Chambolle. Implementation of an adaptive finite-element approximation of the Mumford-Shah functional. *Numer. Math.*, 85(4):609–646, 2000.

[42] B. Bourdin, G. A. Francfort, and J.-J. Marigo. Numerical experiments in revisited brittle fracture. *J. Mech. Phys. Solids*, 48(4):797 – 826, 2000.

[43] B. Bourdin, G. A. Francfort, and J.-J. Marigo. The variational approach to fracture. *J. Elasticity*, 91(1-3):5–148, 2008.

[44] B. Bourdin, C. J. Larsen, and C. L. Richardson. A time-discrete model for dynamic fracture based on crack regularization. *Internat. J. Fracture*, 168(2):133–143, 2011.

[45] B. Bourdin, J.-J. Marigo, C. Maurini, and P. Sicsic. Morphogenesis and propagation of complex cracks induced by thermal shocks. *Phys. Rev. Lett.*, 112(1):014301, 2014.

[46] A. Braides. Γ-*Convergence for Beginners*, volume 22 of *Oxford Lecture Ser. Math. Appl.* Oxford University Press, Oxford, 2002.

[47] S. G. Bullard, B. J. Gromek, M. Fout, and R. Fout. *The Silver Bridge Disaster of 1967.* Arcadia Publishing, 2012.

[48] S. Burke, C. Ortner, and E. Süli. An adaptive finite element approximation of a variational model of brittle fracture. *SIAM J. Numer. Anal.*, 48(3):980–1012, 2010.

[49] S. Burke, C. Ortner, and E. Süli. Adaptive finite element approximation of the Francfort-Marigo model of brittle fracture. In *Approximation and Computation*, volume 42 of *Springer Optim. Appl.*, pages 297–310. Springer, New York, 2011.

[50] S. Burke, C. Ortner, and E. Süli. An adaptive finite element approximation of a generalized Ambrosio-Tortorelli functional. *Math. Models Methods Appl. Sci.*, 23(9):1663–1697, 2013.

[51] F. Cagnetti. A vanishing viscosity approach to fracture growth in a cohesive zone model with prescribed crack path. *Math. Models Methods Appl. Sci.*, 18(7):1027–1071, 2008.

[52] F. Cagnetti and R. Toader. Quasistatic crack evolution for a cohesive zone model with different response to loading and unloading: a Young measures approach. *ESAIM, Control Optim. Calc. Var.*, 17(1):1–27, 2011.

[53] J.-F. Cai, S. Osher, and Z. Shen. Split Bregman methods and frame based image restoration. *Multiscale Model. Simul.*, 8(2):337–369, 2009/10.

[54] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.*, 59(8):1207 – 1223, 2006.

[55] E. J. Candès, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, 2008.

[56] M. Caponigro, M. Fornasier, B. Piccoli, E. Trélat, et al. Sparse stabilization and control of the cucker-smale model. *Math. Control Relat. Fields*, 4(4):447âĂŞ–466, 2013.

[57] C. Castaing and M. Valadier. *Convex Analysis and Measurable Multifunctions.* Lecture Notes in Mathematics, 580. Springer-Verlag, Berlin-New York, 1977.

[58] A. Chambolle. Personal communication.

[59] A. Chambolle and G. Dal Maso. Discrete approximation of the Mumford-Shah functional in dimension two. *M2AN, Math. Model. Numer. Anal.*, 33(4):651–672, 1999.

[60] A. Chambolle, R. A. DeVore, N. Lee, and B. J. Lucier. Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Trans. Image Process.*, 7(3):319–335, 1998.

[61] A. Chambolle, A. Giacomini, and M. Ponsiglione. Crack initiation in brittle materials. *Arch. Rational Mech. Anal.*, 188(2):309–349, 2008.

[62] D. Chen and J. Schijve. Bulging of fatigue cracks in a pressurized aircraft fuselage. Technical report, Delft University of Technology, 1991.

[63] Ph. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North–Holland, Amsterdam, 1978.

[64] Ph. G. Ciarlet and P.-A. Raviart. Maximum principle and uniform convergence for the finite element method. *Comput. Methods Appl. Mech. Engrg.*, 2(1):17–31, 1973.

[65] F. H. Clarke. *Optimization and nonsmooth analysis.* New York: John Wiley, and sons, 1983.

[66] Ph. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.*, 2:77–84, 1975.

[67] A. Cohen, W. Dahmen, and R. A. DeVore. Compressed sensing and best k-term approximation. *J. Amer. Math. Soc.*, 22(1):211–231, 2009.

[68] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.*, 4(4):1168–1200, 2005.

[69] G. Cortesiani. Asymptotic behaviour of a sequence of neumann problems. *Comm. Partial Differential Equations*, 22(9-10):1691–1729, 1997.

[70] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Trans. Autom. Control*, 52(5):852–862, 2007.

[71] I. CVX Research. CVX: Matlab software for disciplined convex programming, version 2.0, 2012.

[72] G. Dal Maso. *An introduction to* Γ*-convergence.* Birkhäuser, Basel, 1993.

[73] G. Dal Maso, G. A. Francfort, and R. Toader. Quasistatic crack growth in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 176(2):165–225, 2005.

[74] G. Dal Maso, A. Giacomini, and M. Ponsiglione. A variational model for quasistatic crack growth in nonlinear elasticity: some qualitative properties of the solutions. *Boll. Unione Mat. Ital.*, 9(2):371–390, 2009.

[75] G. Dal Maso and C. J. Larsen. Existence for wave equations on domains with arbitrary growing cracks. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 22(3):387–408, 2011.

[76] G. Dal Maso and G. Lazzaroni. Quasistatic crack growth in finite elasticity with non-interpenetration. In *Ann. Inst. H. Poincare (C) Anal. Non Lineaire*, volume 27, pages 257–290. Elsevier, 2010.

[77] G. Dal Maso and R. Toader. A model for the quasi-static growth of brittle fractures based on local minimization. *Math. Models Methods Appl. Sci.*, 12(12):1773–1799, 2002.

[78] G. Dal Maso and R. Toader. A model for the quasi-static growth of brittle fractures: Existence and approximation results. *Arch. Ration. Mech. Anal.*, 162(2):101–135, 2002.

[79] G. Dal Maso and R. Toader. Quasistatic crack growth in elasto-plastic materials: The two-dimensional case. *Arch. Ration. Mech. Anal.*, 196(3):867–906, 2010.

[80] G. Dal Maso and C. Zanini. Quasi-static crack growth for a cohesive zone model with prescribed crack path. *Proc. R. Soc. Edinb., Sect. A, Math.*, 137(2):253–279, 2007.

[81] A. Damlamian. Le problème de la passoire de Neumann. *Rend. Sem. Mat. Univ. Politec. Torino*, 43(3):427–450 (1986), 1985.

[82] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.*, 57(11):1413–1457, 2004.

[83] M. Davenport. The RIP and the NSP, Apr. 2011.

[84] E. De Giorgi. Free-discontinuity problems in calculus of variations. In R. Dautray, editor, *Frontiers in pure and applied mathematics, a collection of papers dedicated to J.-L. Lions on the occasion of his* 60^{th} *birthday*, pages 55–62. North Holland, 1991.

[85] E. De Giorgi and L. Ambrosio. Un nuovo tipo di funzionale del calcolo delle variazioni. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.(8)*, 82:199–210, 1988.

[86] E. De Giorgi, M. Carriero, and A. Leaci. Existence theorem for a minimum problem with free discontinuity set. *Arch. Rational Mech. Anal.*, 108(4):195–218, 1989.

[87] E. De Giorgi and T. Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.(8)*, 58(6):842–850, 1975.

[88] L. Dedè, S. Micheletti, and S. Perotto. Anisotropic error control for environmental applications. *Appl. Numer. Math.*, 58(9):1320–1339, 2008.

[89] G. Del Piero. One-dimensional ductile-brittle transition, yielding, and structured deformations. In *IUTAM Symposium on Variations of Domain and Free-Boundary Problems in Solid Mechanics*, pages 203–210. Springer, 1999.

[90] G. Del Piero. A variational approach to fracture and other inelastic phenomena. *J. Elasticity*, 112:3–77, 2013.

[91] G. Del Piero, G. Lancioni, and R. March. A variational model for fracture mechanics: numerical experiments. *J. Mech. Phys. Solids*, 55:2513–2537, 2007.

[92] G. Del Piero and D. R. Owen. Integral-gradient formulae for structured deformations. *Arch. Rational Mech. Anal.*, 131(2):121–138, 1995.

[93] G. Del Piero and L. Truskinovsky. Elastic bars with cohesive energy. *Cont. Mech. Thermodyn.*, 21(2):141–171, 2009.

[94] D. L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, 2006.

[95] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[96] D. S. Dugdale. Yielding of steel sheets containing slits. *J. Mech. Phys. Solids*, 8(2):100–104, 1960.

[97] S. Durand and M. Nikolova. Stability of the minimizers of least squares with a non-convex regularization. I: Local behavior. *Appl. Math. Optimization*, 53(2):185–208, 2006.

[98] S. Durand and M. Nikolova. Stability of the minimizers of least squares with a non-convex regularization. II: Global behavior. *Appl. Math. Optimization*, 53(3):259–277, 2006.

[99] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems. Translated by Minerva Translations, Ltd., London.* Studies in Mathematics and its Applications. Vol. 1. Amsterdam - Oxford: North-Holland Publishing Company; New York: American Elsevier Publishing Company, Inc., 1976.

[100] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* CRC press, 2015.

[101] M. D. Feit, J. A. Fleck, and A. Steiger. Solution of the schrödinger equation by a spectral method. *J. Comput. Phys.*, 47(3):412–433, 1982.

[102] M. A. T. Figueiredo and R. D. Nowak. Wavelet-based image estimation: An empirical Bayes approach using Jeffrey's noninformative prior. *IEEE Trans. Image Process.*, 10(9):1322–1331, 2001.

[103] M. Focardi. On the variational approximation of free-discontinuity problems in the vectorial case. *Math. Models Methods Appl. Sci.*, 11(04):663–684, 2001.

[104] L. Formaggia and S. Perotto. New anisotropic a priori error estimates. *Numer. Math.*, 89(4):641–667, 2001.

[105] L. Formaggia and S. Perotto. Anisotropic error estimates for elliptic problems. *Numer. Math.*, 94:67–92, 2003.

[106] M. Fornasier and H. Rauhut. Iterative thresholding algorithms. *Appl. Comput. Harmon. Anal.*, 25(2):187–208, 2008.

[107] M. Fornasier and H. Rauhut. Compressive Sensing. In O. Scherzer, editor, *Handbook of Mathematical Methods in Imaging*, pages 187–228. Springer, 2011.

[108] M. Fornasier and R. Ward. Iterative thresholding meets free-discontinuity problems. *Found. Comput. Math.*, 10(5):527–567, 2010.

[109] G. A. Francfort and C. J. Larsen. Existence and convergence for quasi-static evolution in brittle fracture. *Comm. Pure Appl. Math.*, 56(10):1465–1500, 2003.

[110] G. A. Francfort and J.-J. Marigo. Revisiting brittle fracture as an energy minimization problem. *J. Mech. Phys. Solids*, 46(8):1319–1342, 1998.

[111] G. A. Francfort and A. Mielke. Existence results for a class of rate-independent material models with nonconvex elastic energies. *J. Reine Angew. Math.*, 2006(595):55–91, 2006.

[112] K. Frick and O. Scherzer. Regularization of ill-posed linear equations by the non-stationary augmented Lagrangian method. *J. Integral Equations Appl.*, 22(2):217–257, 2010.

[113] G. Friesecke and B. D. Goddard. Asymptotics-based CI models for atoms: properties, exact solution of a minimal model for Li to Ne, and application to atomic spectra. *Multiscale Model. Simul.*, 7(4):1876–1897, 2009.

[114] G. Galilei. *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*. Elsevier, 1636.

[115] M. G. Geers, R. A. Engelen, and R. Ubachs. On the numerical modelling of ductile damage with an implicit gradient-enahced formulation. *Rev. Européenne Élém. Finis*, 10(2/3/4):173, 2001.

[116] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell*, 6:721–741, 1984.

[117] P.-L. George and H. Borouchaki. *Delaunay Triangulation and Meshing. Application to Finite Elements*. Edition Hermès, Paris, 1998.

[118] A. Giacomini. Ambrosio-Tortorelli approximation of quasi-static evolution of brittle fractures. *Calc. Var. Partial Differential Equations*, 22(2):129–172, 2005.

[119] A. Giacomini and M. Ponsiglione. Non-interpenetration of matter for sbv deformations of hyperelastic brittle materials. *Proc. Roy. Soc. Edinburgh: Sec. A*, 138(05):1019–1041, 2008.

[120] T. Goldstein, X. Bresson, and S. Osher. Geometric applications of the split bregman method: segmentation and surface reconstruction. *J. Sci. Comput.*, 45(1-3):272–293, 2010.

[121] T. Goldstein and S. Osher. The split bregman method for l1-regularized problems. *SIAM J. Imaging Sci.*, 2(2):323–343, 2009.

[122] P. L. Gould. *Introduction to linear elasticity*. Springer, 1994.

[123] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008.

[124] A. A. Griffith. The phenomena of rupture and flow in solids. *Phil. Trans. R. Soc. Lond. Series A*, 221:163–198, 1921.

[125] M. E. Gurtin. The linear theory of elasticity. In *Linear Theories of Elasticity and Thermoelasticity*, pages 1–295. Springer, 1973.

[126] M. E. Gurtin, E. Fried, and L. Anand. *The mechanics and thermodynamics of continua*. Cambridge University Press, 2010.

[127] W. Hackbusch. *Multigrid methods and applications*, volume 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[128] J. Haupt, R. Baraniuk, R. Castro, and R. Nowak. Compressive distilled sensing: Sparse recovery using adaptivity in compressive measurements. In *Proceedings of the 43rd Asilomar conference on Signals, systems and computers*, Asilomar'09, Piscataway, NJ, USA, 2009. IEEE Press.

[129] J. Haupt, R. Baraniuk, R. Castro, and R. Nowak. Sequentially designed compressed sensing. In *Statistical Signal Processing Workshop (SSP)*, pages 401–404, 2012.

[130] J. Haupt, R. Castro, and R. Nowak. Distilled sensing: Adaptive sampling for sparse detection and estimation. *IEEE Trans. Inf. Theory*, 57(9):6222–6235, 2011.

[131] F. Hausdorff. Dimension und äußeres maß. *Math. Ann.*, 79(1-2):157–179, 1918.

[132] F. Hecht. New development in FreeFem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.

[133] F. Henneke and M. Liebmann. A generalized suzuki–trotter type method in optimal control of coupled schrödinger equations. *J. Optim. Theory Appl.*, 2015.

[134] R. Hill. *The mathematical theory of plasticity*, volume 11. Oxford university press, 1998.

[135] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth newton method. *SIAM J. Optim.*, 13(3):865–888, 2002.

[136] K. Ito and K. Kunisch. The primal-dual active set method for nonlinear optimal control problems with bilateral constraints. *SIAM J. Control Optim.*, 43(1):357–376, 2004.

[137] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications.* Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2008.

[138] D. Y. Jeong, O. Orringer, and G. C. Sill. Strain energy density approach to stable crack extension under net section yielding of aircraft fuselage. *Theoret. Appl. Fracture Mech.*, 22(2):127–137, 1995.

[139] Y. Jiao, B. Jin, and X. Lu. A primal dual active set with continuation algorithm for the ℓ^0-regularized optimization problem. *Appl. Comput. Harmon. Anal.*, 2014.

[140] C. Jordan. Sur la série de fourier. *CR Acad. Sci. Paris*, 92(5):228–230, 1881.

[141] E. Y. Khruslov. The asymptotic behavior of solutions of the second boundary value problem under fragmentation of the boundary of the domain. *Mat. Sb.*, 148(4):604–621, 1978.

[142] D. Knees and A. Mielke. Energy release rate for cracks in finite-strain elasticity. *Math. Methods Appl. Sci.*, 31(5):501–528, 2008.

[143] D. Knees, A. Mielke, and C. Zanini. On the inviscid limit of a model for crack propagation. *Math. Models Methods Appl. Sci.*, 18(09):1529–1569, 2008.

[144] D. Knees, R. Rossi, and C. Zanini. A vanishing viscosity approach to a rate-independent damage model. *Math. Models Methods Appl. Sci.*, 23(04):565–616, 2013.

[145] H. Kobayashi and H. Onoue. Brittle fracture of liberty ships. *Failure Knowledge Database*, 100, 1943.

[146] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133, 1965.

[147] S. Korotov, M. Křížek, and P. Neittaanmäki. Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle. *Math. Comp.*, 70(233):107–119, 2001.

[148] L. D. Landau and E. Lifshitz. Theory of elasticity, vol. 7. *Course of Theoretical Physics*, 3, 1986.

[149] C. J. Larsen. Epsilon-stable quasi-static brittle fracture evolution. *Comm. Pure Appl. Math.*, 63(5):630–654, 2010.

[150] C. J. Larsen. A new variational principle for cohesive fracture and elastoplasticity. *Mech. Res. Comm.*, 58:133–138, 2014.

[151] C. J. Larsen and V. Slastikov. Dynamic cohesive fracture: models and analysis. *Math. Models Methods Appl. Sci.*, 24(9):1857–1875, 2014.

[152] B. R. Lawn and A. G. Evans. A model for crack initiation in elastic/plastic indentation fields. *J. Mater. Sci.*, 12(11):2195–2199, 1977.

[153] G. Lazzaroni, R. Bargellini, P.-E. Dumouchel, and J.-J. Marigo. On the role of kinetic energy during unstable propagation in a heterogeneous peeling test. *Internat. J. Fracture*, 175(2):127–150, 2012.

[154] G. Lazzaroni and R. Toader. A model for crack propagation based on viscous approximation. *Math. Models Methods Appl. Sci.*, 21(10):2019–2047, 2011.

[155] G. Lazzaroni and R. Toader. Some remarks on the viscous approximation of crack growth. *Discrete Contin. Dyn. Syst. Ser. S*, 6:131–146, 2013.

[156] J. Lubliner. *Plasticity theory.* Macmillan publishing company, 1990.

[157] S. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way.* Academic Press, 3rd edition, 2008.

[158] J.-J. Marigo. Initiation of cracks in Griffith's theory: an argument of continuity in favor of global minimization. *J. Nonlinear Sci.*, 20(6):831–868, 2010.

[159] D. A. Mazziotti. *Advances in Chemical Physics, Reduced-Density-Matrix Mechanics: With Application to Many-Electron Atoms and Molecules*, volume 134. John Wiley & Sons, 2007.

[160] S. Micheletti and S. Perotto. Reliability and efficiency of an anisotropic Zienkiewicz–Zhu error estimator. *Comput. Methods Appl. Mech. Engrg.*, 195(9):799–835, 2006.

[161] S. Micheletti and S. Perotto. Output functional control for nonlinear equations driven by anisotropic mesh adaption: the Navier-Stokes equations. *SIAM J. Sci. Comput.*, 30(6):2817–2854, 2008.

[162] S. Micheletti and S. Perotto. The effect of anisotropic mesh adaptation on PDE-constrained optimal control problems. *SIAM J. Control. Optim.*, 49(4):1793–1828, 2011.

[163] S. Micheletti, S. Perotto, and M. Picasso. Stabilized finite elements on anisotropic meshes: a priori error estimates for the advection–diffusion and the Stokes problems. *SIAM J. Numer. Anal.*, 41(3):1131–1162, 2003.

[164] C. Miehe, M. Hofacker, and W. Fabian. A phase field model rate-independent crack propagation: Robust algorithmic implementation based on operator splits. *Comput. Methods Appl. Mech. Engrg.*, 199(45-48):2765–2778, 2010.

[165] A. Mielke. Evolution of rate-independent inelasticity with microstructure using relaxation and young measures. In C. Miehe, editor, *IUTAM Symposium on Computational Mechanics of Solid Materials at Large Strains*, volume 108 of *Solid Mechanics and Its Applications*, pages 33–44. Springer Netherlands, 2003.

[166] A. Mielke. Evolution of rate-independent systems. In *Evolutionary equations. Vol. II*, Handb. Differ. Equ., pages 461–559. Elsevier/North-Holland, Amsterdam, 2005.

[167] A. Mielke, R. Rossi, and G. Savaré. Nonsmooth analysis of doubly nonlinear evolution equations. *Calc. Var. Partial Differential Equations*, 46(1-2):253–310, 2013.

[168] A. Mielke and T. Roubíček. Rate-independent damage processes in nonlinear elasticity. *Math. Models Methods Appl. Sci.*, 16(2):177–209, 2006.

[169] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42(5):577–685, 1989.

[170] F. Murat. The neumann sieve. non linear variational problems (isola d'elba, 1983). *Res. Notes in Math*, 127:27–32, 1985.

[171] D. Needell. Noisy signal recovery via iterative reweighted l1-minimization. In *Proceedings of the 43rd Asilomar conference on Signals, systems and computers*, Asilomar'09, pages 113–117, Piscataway, NJ, USA, 2009. IEEE Press.

[172] M. Negri. A finite element approximation of the Griffith's model in fracture mechanics. *Numer. Math.*, 95(4):653–687, 2003.

[173] M. Negri. A comparative analysis on variational models for quasi-static brittle crack propagation. *Ad. Calc. Var.*, 3(2):149–212, 2010.

[174] M. Negri. Quasi-static rate-independent evolutions: characterization, existence, approximation and application to fracture mechanics. *ESAIM Control Optim. Calc. Var.*, 20(4):983–1008, 2014.

[175] M. Negri and C. Ortner. Quasi-static crack propagation by griffith's criterion. *Math. Models Methods Appl. Sci.*, 18(11):1895–1925, 2008.

[176] M. Negri and M. Paolini. Numerical minimization of the Mumford-Shah functional. *Calcolo*, 38(2):67–84, 2001.

[177] M. Nikolova. Markovian reconstruction using a GNC approach. *IEEE Trans. Image Process.*, 8(9):1204–1220, 1999.

[178] M. Nikolova. Thresholding implied by truncated quadratic regularization. *IEEE Trans. Signal Process.*, 48:3437–3450, 2000.

[179] M. Nikolova, J. Idier, and A. Mohammad-Djafari. Inversion of large-support ill-posed linear operators using a piecewise Gaussian MRF. *IEEE Trans. Image Process.*, 7(4):571–585, 1998.

[180] M. Nikolova, M. Ng., and C. Tam. Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. *IEEE Trans. Image Process.*, 19, 2010.

[181] M. Nikolova, M. Ng., S. Zhang, and W.-K. Ching. Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. *SIAM J. Imaging Sci.*, 1:2–25, 2008.

[182] J. Nocedal and S. J. Wright. *Numerical optimization. 2nd ed.* New York, NY: Springer, 2006.

[183] R. W. Ogden and Y. B. Fu. *Nonlinear elasticity: theory and applications.* Cambridge University Press, 2001.

[184] M. Ortiz and E. Repetto. Nonconvex energy minimization and dislocation structures in ductile single crystals. *J. Mech. Phys. Solids*, 47(2):397 – 462, 1999.

[185] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Model. Simul.*, 4(2):460–489, 2005.

[186] D. R. Owen. Balance laws and a dissipation inequality for general constituents undergoing disarrangements and mixing. *ZAMM-Z. Angew. Math. Mech.*, 88(5):365–377, 2008.

[187] K. Pham, J.-J. Marigo, and C. Maurini. The issues of the uniqueness and the stability of the homogeneous response in uniaxial tests with gradient damage models. *J. Mech. Phys. Solids*, 59(6):1163–1190, 2011.

[188] B. Piccoli, N. P. Duteil, and B. Scharf. Optimal control of a collective migration model. *arXiv preprint arXiv:1503.05168*, 2015.

[189] G. Plonka and J. Ma. Curvelet-wavelet regularized split bregman iteration for compressed sensing. *Int. J. Wavelets, Multiresolut. Inf. Process.*, 9(01):79–110, 2011.

[190] V. T. Polyak and N. Y. Tret'yakov. The method of penalty estimates for conditional extremum problems. *Z. Vychisl. Mat. I Mat. Fiz.*, 13:34–36, 1973.

[191] J. H. Poynting and J. J. Thomson. *A Text-book of Physics: Properties of matter. 1927*, volume 1. C. Griffin & Company, 1913.

[192] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23. Springer Science & Business Media, 2008.

[193] R. A. Schapery. A theory of crack initiation and growth in viscoelastic media. *Internat. J. Fracture*, 11(1):141–159, 1975.

[194] S. Serfaty. Local minimizers for the Ginzburg–Landau energy near critical magnetic field: Part i. *Commun. Contemp. Math.*, 01(02):213–254, 1999.

[195] S. Serfaty. Local minimizers for the Ginzburg–Landau energy near critical magnetic field: Part ii. *Commun. Contemp. Math.*, 1(03):295–333, 1999.

[196] S. Setzer, G. Steidl, and T. Teuber. Deblurring poissonian images by split bregman techniques. *J. Vis. Commun. Image Represent.*, 21(3):193–199, 2010.

[197] Y. Shao, Y. Zhang, X. Xu, Z. Zhou, W. Li, and B. Liu. Effect of crack pattern on the residual strength of ceramics after quenching. *J. Amer. Ceram. Soc.*, 94(9):2804–2807, 2011.

[198] W. S. Slaughter. *The linearized theory of elasticity*. Springer Science & Business Media, 2002.

[199] G. D. Smith. *Numerical solution of partial differential equations: finite difference methods*. Oxford university press, 1985.

[200] M. Stojnic, W. Xu, A. S. Avestimehr, and B. Hassibi. Compressed sensing of approximately sparse signals. In *ISIT*, pages 2182–2186, 2008.

[201] V. N. Strakhov and M. A. Brodsky. On the uniqueness of the inverse logarithmic potential problem. *SIAM J. Appl. Math.*, 46(2):324–344, 1986.

[202] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*, volume 212. Prentice-Hall Englewood Cliffs, NJ, 1973.

[203] P. Suquet. Sur les équations de la plasticité: existence et régularité des solutions. *J. Mécanique*, 20(1):3–39, 1981.

[204] A. Szabo and N. S. Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.

[205] J. W. Tester, H. D. Murphy, C. O. Grigsby, R. M. Potter, B. A. Robinson, et al. Fractured geothermal reservoir growth induced by heat extraction. *SPE reservoir engineering*, 4(01):97–104, 1989.

[206] M. Thomas and A. Mielke. Damage of nonlinearly elastic materials at small strain–existence and regularity results–. *ZAMM-Z. Angew. Math. Mech.*, 90(2):88–112, 2010.

[207] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25. Springer Verlag, Berlin, 1997.

[208] S. Timoshenko and J. N. Goodier. Theory of elasticity. *New York*, 412, 1951.

[209] R. Toader and C. Zanini. An artificial viscosity approach to quasistatic crack growth. *Discrete Contin. Dyn. Syst. Ser. S*, 2(9):1–35, 2009.

[210] J. Treichler, M. A. Davenport, and R. G. Baraniuk. Application of compressive sensing to the design of wideband signal acquisition receivers. In *6th U.S. / Australia Joint Workshop on Defense Applications of Signal Processing (DASP)*, Lihue, Hawaii, Sept. 2009.

[211] R. Vanselow. About delaunay triangulations and discrete maximum principles for the linear conforming FEM applied to the Poisson equation. *Appl. Math.*, 46(1):13–28, 2001.

[212] L. Wang, F. Brust, and S. Atluri. The elastic-plastic finite element alternating method (EPFEAM) and the prediction of fracture under WFD conditions in aircraft structures. *Comput. Mech.*, 20(3):199–212, 1997.

[213] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.*, 1(3):248–272, 2008.

[214] P. Wojtaszczyk. Stability and instance optimality for Gaussian measurements in compressed sensing. *Found. Comput. Math.*, 10(1):1–13, 2010.

[215] C. Wu, X.-C. Tai, et al. Augmented lagrangian method, dual methods, and split bregman iteration for rof, vectorial tv, and high order models. *SIAM J. Imaging Sci.*, 3(3):300–339, 2010.

[216] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for \ell_1-minimization with applications to compressed sensing. *SIAM J. Imaging Sci.*, 1(1):143–168, 2008.

[217] W. P. Ziemer. *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, volume 120. Springer Science & Business Media, 2012.

[218] O. C. Zienkiewicz and J. Z. Zhu. The superconvergent patch recovery and a posteriori error estimates. part 1: The recovery technique. *Int. J. Numer. Methods Engrg.*, 33(7):1331–1364, 1992.