Multiscale Model. Simul.
Vol. 9, No. 3, pp. 1276–1300

AN EFFICIENT APPROACH TO THE NUMERICAL SOLUTION OF RATE-INDEPENDENT PROBLEMS WITH NONCONVEX ENERGIES*

SÖREN BARTELS† AND MARTIN KRUŽÍK‡

Abstract. We propose a new approach to the numerical treatment of non(quasi)convex rate-independent evolutionary problems. The main idea is to replace the original microscopic energy density with its polyconvexification. For this problem, first-order optimality conditions are derived and used in finding a discrete solution. The effectiveness of the method is illustrated with some numerical experiments.

Key words. nonconvexity, numerical solution, rate-independent problems

AMS subject classifications. 49J45

DOI. 10.1137/110821718

1. Introduction—the underlying mathematical model.

The aim of this contribution is to propose a computational method for solving non(quasi)convex vectorial and multidimensional variational problems. It is well known that if $V(y) := \int_\Omega W(\nabla y(x))\mathrm{d}x$, where $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain and $W:\mathbb{R}^{m\times n} \to \mathbb{R}$ is continuous but not quasiconvex [25], then V is not weakly sequentially lower semicontinuous on $W^{1,p}(\Omega;\mathbb{R}^m)$, $1 < p < +\infty$, and consequently V does not necessarily attain its minimum in this space. We recall that $W:\mathbb{R}^{m\times n} \to \mathbb{R}$ is quasiconvex if for all $\psi \in W_0^{1,\infty}(\Omega;\mathbb{R}^m)$ and all $F \in \mathbb{R}^{m\times n}$ it holds that

$$W(F)|\Omega| \le \int_\Omega W(F + \nabla\psi(x))\mathrm{d}x. \tag{1}$$

One way to overcome this difficulty is to replace W with its quasiconvex envelope QW defined as the pointwise supremum of all quasiconvex functions not greater than W; cf. [9]. This is, however, mostly a theoretical tool because the formula is generically not known in a closed form. Nevertheless, by the relaxation theorem [9] we have under standard polynomial growth conditions at infinity; i.e., we assume for some $0 < c < C$, $1 < p < +\infty$, and all $F \in \mathbb{R}^{m\times n}$,

$$c(-1 + |F|^p) \le W(F) \le C(1 + |F|^p), \tag{2}$$

that $\inf_{y\in\mathbb{Y}^p(\Omega;\mathbb{R}^m)} \int_\Omega W(\nabla y(x))\mathrm{d}x = \min_{y\in\mathbb{Y}^p(\Omega;\mathbb{R}^m)} \int_\Omega QW(\nabla y(x))\mathrm{d}x$, where $\mathbb{Y}^p(\Omega;\mathbb{R}^m)$ is a suitable subset of $W^{1,p}(\Omega;\mathbb{R}^m)$, which may include Dirichlet boundary conditions, for instance. In this paper, we suggest working with the polyconvex envelope PW instead. The polyconvex envelope is defined analogously to the quasiconvex one, and we say that W is polyconvex [3], [9] if there is a convex function $h:\mathbb{R}^\sigma \to \mathbb{R}$ such that

*Received by the editors January 20, 2011; accepted for publication (in revised form) July 7, 2011; published electronically September 29, 2011. This research was supported by the German Research Foundation (DFG) via SFB 611 "Singular phenomena and scaling in mathematical models."
http://www.siam.org/journals/mms/9-3/82171.html

†Institut für Numerische Simulation, Rheinische Friedrich-Wilhelms–Universität Bonn, Wegelerstraße 6, D-53115 Bonn, Germany (bartels@ins.uni-bonn.de).

‡Institute of Information Theory and Automation of the ASCR, Pod vodárenskou věží 4, 182 08 Prague, Czech Republic, and Czech Technical University, Faculty of Civil Engineering, Thákurova 7, 166 29 Prague, Czech Republic (kruzik@utia.cas.cz). The research of this author was supported by grants IAA100750802 (GAAV), 201/10/0357 (GAČR), and VZ6840770021(MŠMT ČR).

$W(F) = h(\mathbb{T}(F))$ for all $F \in \mathbb{R}^{m\times n}$. Here $\mathbb{T}(F)$ denotes the vector of all subdeterminants of F, i.e., the dimension of $\mathbb{T}(F)$ equals

$$\sigma := \sum_{k=1}^{\min(m,n)} \binom{m}{k}\binom{n}{k} = \binom{m+n}{n} - 1. \tag{3}$$

Polyconvexity implies quasiconvexity, and it is a stronger property if $\min(m, n) > 1$. The resulting variational problem $\min_{y\in\mathbb{Y}^p(\Omega;\mathbb{R}^m)} \int_\Omega PW(\nabla y(x))\mathrm{d}x$ is well-posed, and we have

$$\begin{aligned}\inf_{y\in\mathbb{Y}^p(\Omega;\mathbb{R}^m)} \int_\Omega W(\nabla y(x))\mathrm{d}x &= \min_{y\in\mathbb{Y}^p(\Omega;\mathbb{R}^m)} \int_\Omega QW(\nabla y(x))\mathrm{d}x\\ &\geq \min_{y\in\mathbb{Y}^p(\Omega;\mathbb{R}^m)} \int_\Omega PW(\nabla y(x))\mathrm{d}x,\end{aligned}$$

where the last inequality can be strict in particular cases. The advantage is that there are efficient numerical methods to evaluate $PW(F)$ [7], [29]. There is another widely used method to estimate QW, namely, the so-called rank-one convexification of W. The function W is called rank-one convex if it is convex along rank-one lines in $\mathbb{R}^{m\times n}$. Estimating the rank-one convex envelope of W by so-called laminates [27] in the context of elasticity is used, e.g., in [2], [15], [13], [14], [19]. In these cases, W is the stored energy density of a hyperelastic material. We recall that laminates are among experimentally observed material microstructures. Algorithms for the approximation of the rank-one convex envelope have been proposed and analyzed in [10], [6]. Unfortunately, they are extremely expensive and may lead to ill-posed, i.e., nonweakly lower semicontinuous, variational problems if $n \geq 2$ and $m \geq 3$ [28], [34].

The supremum definitions of QW or PW mentioned above are not very useful for numerical or analytical considerations. Much more suitable ways to evaluate and/or estimate them were developed in terms of parameterized (Young) measures.

1.1. Young measures and Young functionals. It is well known [4], [35] that if $\{z_k\}_{k\in\mathbb{N}} \subset L^p(\Omega;\mathbb{R}^{m\times n})$, $1 \leq p < +\infty$, is bounded, then there exists a subsequence (not relabeled) and a family $\nu = \{\nu_x\}_{x\in\Omega}$ of probability measures on Ω such that for all $g \in L^\infty(\Omega)$ and all $f \in C(\mathbb{R}^{m\times n})$ such that $\{f(z_k)\}_{k\in\mathbb{N}}$ is uniformly integrable it holds that

$$\lim_{k\to\infty} \int_\Omega f(z_k(x))g(x)\mathrm{d}x = \int_\Omega \int_{\mathbb{R}^{m\times n}} f(F)\mathrm{d}\nu_x(F)g(x)\mathrm{d}x. \tag{4}$$

Conversely, if $\nu = \{\nu_x\}_{x\in\Omega}$ is such that ν_x is for almost all $x \in \Omega$ a probability measure on $\mathbb{R}^{m\times n}$, $x \mapsto \int_{\mathbb{R}^{m\times n}} f(F)\mathrm{d}\nu_x(F)$ is measurable for all $f \in C_0(\mathbb{R}^{m\times n}) := \{g \in C(\mathbb{R}^{m\times n}); \lim_{|s|\to\infty} g(s) = 0\}$, and $\int_\Omega \int_{\mathbb{R}^{m\times n}} |F|^p \mathrm{d}\nu_x(F)\mathrm{d}x < +\infty$, then there exists a sequence $\{z_k\}_{k\in\mathbb{N}} \subset L^p(\Omega;\mathbb{R}^{m\times n})$ such that (4) holds. The family $\nu = \{\nu_x\}$ is called the L^p-Young measure and $\{z_k\}$ its generating sequence. We denote the set of L^p-Young measures by $\mathcal{Y}^p(\Omega;\mathbb{R}^{m\times n})$. It is well known that every Young measure ν as above can be generated by a sequence $\{z_k\}$ such that $\{f(z_k)\}$ is uniformly integrable for every continuous $f \in C^p(\mathbb{R}^{m\times n}) := \{f \in C(\mathbb{R}^{m\times n}); |f| \leq C(1 + |\cdot|^p), C > 0\}$.

We will be interested in Young measures generated by gradients, i.e., $z_k := \nabla y_k$ for some sequence $\{y_k\} \subset W^{1,p}(\Omega;\mathbb{R}^m)$. Such a Young measure will be referred to as a gradient Young measure. Fixing $1 \leq p < +\infty$, we denote the set of gradient Young measures generated by $\{\nabla y_k\}$ for $\{y_k\} \subset W^{1,p}(\Omega;\mathbb{R}^m)$ by $\mathbb{G}^p(\Omega;\mathbb{R}^{m\times n})$. Thus, if $\{y_k\} \subset$

$W^{1,p}(\Omega;\mathbb{R}^m)$ is bounded and $\{W(\nabla y_k)\}$ uniformly integrable, we have (up to a subsequence) that $\lim_{k\to\infty} V(y_k) = \int_\Omega \int_{\mathbb{R}^{m\times n}} W(F)\mathrm{d}\nu_x(F)$. Let us mention that if W is coercive with superlinear growth at infinity and $\{y_k\}$ is minimizing for V, then the uniform-integrability condition holds. In fact, if $p > 1$, then every gradient Young measure from $\mathbb{G}^p(\Omega;\mathbb{R}^{m\times n})$ can be generated by a bounded sequence $\{\nabla z_k\}_{k\in\mathbb{N}}$ such that $\{z_k\}_{k\in\mathbb{N}} \subset W^{1,p}(\Omega;\mathbb{R}^m)$ and $\{|\nabla z_k|^p\}_{k\in\mathbb{N}}$ is uniformly integrable [11]. The following well-known result of Kinderlehrer and Pedregal [17], [27] characterizes the set of gradient Young measures.

LEMMA 1.1. *Let $1 < p < +\infty$. A Young measure $\nu = \{\nu_x\}_{x\in\Omega}$ belongs to $\mathbb{G}^p(\Omega;\mathbb{R}^{m\times n})$ if and only if the following three conditions are satisfied simultaneously:*

(i) *there is $y \in W^{1,p}(\Omega;\mathbb{R}^m)$ such that for a.a. $x \in \Omega$*

$$\nabla y(x) = \int_{\mathbb{R}^{m\times n}} F\mathrm{d}\nu_x(F), \tag{5}$$

(ii) *for this y and all quasiconvex functions $v\colon\mathbb{R}^{m\times n} \to \mathbb{R}$, $|v| \le C(1 + |\cdot|^p)$, it holds that for a.a. $x \in \Omega$*

$$v(\nabla y(x)) \le \int_{\mathbb{R}^{m\times n}} v(F)\mathrm{d}\nu_x(F), \tag{6}$$

(iii) *it holds that*

$$\int_\Omega \int_{\mathbb{R}^{m\times n}} |F|^p \mathrm{d}\nu_x(F)\mathrm{d}x < +\infty. \tag{7}$$

Extending the validity of (ii) to all rank-one convex functions with p-growth at infinity defines a subset of $\mathbb{G}^p(\Omega;\mathbb{R}^{m\times n})$ called laminates [27].

In this paper, we propose a different approach, namely, to use a proper superset of $\mathbb{G}^p(\Omega;\mathbb{R}^{m\times n})$ by requiring that (6) holds only for all quasiaffine functions. We recall that v is quasiaffine if and only if it is an affine function of all subdeterminants of the matrix argument. This means that there is $\Xi \in \mathbb{R}^\sigma$ and $c \in \mathbb{R}$ such that $v(\cdot) \coloneqq \Xi \cdot \mathbb{T}(\cdot) + c$. In particular, if $m = n = 2$, then $v(F) \coloneqq A \cdot F + b \det F + c$ for some $A \in \mathbb{R}^{2\times 2}$ and $b, c \in \mathbb{R}$. If $n = 3$, then $v(F) \coloneqq A \cdot F + B \cdot \mathrm{cof} F + c \det F + d$ for some $A, B \in \mathbb{R}^{3\times 3}$ and $c, d \in \mathbb{R}$. Here "cof F" denotes the matrix composed of all 2×2 subdeterminants of F. We are going to deal with the following set of polyconvex Young measures [27], [31], which strictly contains all gradient Young measures.

DEFINITION 1.2. *Let $\min(m, n) < p < +\infty$. A Young measure $\nu = \{\nu_x\}_{x\in\Omega}$ is called polyconvex and belongs to the set $\mathbb{P}^p(\Omega;\mathbb{R}^{m\times n})$ if the following conditions are satisfied:*

(i) *there is $y \in W^{1,p}(\Omega;\mathbb{R}^m)$ such that for a.a. $x \in \Omega$*

$$\mathbb{T}(\nabla y(x)) = \int_{\mathbb{R}^{m\times n}} \mathbb{T}(F)\mathrm{d}\nu_x(F), \tag{8}$$

(ii) *it holds that*

$$\int_\Omega \int_{\mathbb{R}^{m\times n}} |F|^p \mathrm{d}\nu_x(F)\mathrm{d}x < +\infty. \tag{9}$$

There is a much more general approach to relaxation using the so-called Young functionals. We refer to [31] for a detailed exposition; cf. also [29], [30], [21]. The main idea is to choose a suitable subspace H of Carathéodory functions $h:\Omega\times\mathbb{R}^{m\times n}\to\mathbb{R}$ with at most p-growth; i.e., we consider $|h(x,F)|\le a_h(x)+b_h|F|^p$ if $a_h\in L^1(\Omega)$ and $b_h>0$. The space H is normed with the (semi)norm

$$|h|=\inf\{\|a\|_{L^1(\Omega)}+b;\ \forall(x,F)\in\Omega\times\mathbb{R}^{m\times n}:|h(x,F)|\le a(x)+b|F|^p\}.$$

The dual space to H, H^*, is equipped with the standard norm $\|\eta\|_{H^*}=\sup_{h\in H;|h|\neq 0}\langle\eta,h\rangle/|h|$. Following [31] we define the embedding $i:L^p(\Omega;\mathbb{R}^{m\times n})\to H^*$ by $\langle i(u),h\rangle:=\int_\Omega h(x,u(x))\mathrm{d}x$, $h\in H$.

The subset of H^*, called the set of Young functionals, is the following ("$w^*-\lim$" denotes the weak* limit):

$$Y^p_H(\Omega;\mathbb{R}^{m\times n}):=\{\eta\in H^*;\ \exists\{z_\alpha\}_\alpha\subset L^p(\Omega;\mathbb{R}^{m\times n})\,\text{a bounded net};\mathrm{w}^*-\lim\, i(z_\alpha)=\eta\}.$$

We then say that $\{z_\alpha\}$ generates η. It was proved in [31] that $Y^p_H(\Omega;\mathbb{R}^{m\times n})$ is a convex weakly* σ-compact subset of H^*. Moreover, if H contains a coercive function with p-growth and H is separable, then $Y^p_H(\Omega;\mathbb{R}^{m\times n})$ is closed locally weakly* sequentially compact. If η is generated by a net such that $\{|z_\alpha|^p\}$ is uniformly integrable, then we call η p-nonconcentrating. Every p-nonconcentrating Young functional can be represented by an L^p-Young measure ν (generally not unique) such that for all $h\in H$ $\langle\eta,h\rangle=\int_\Omega\int_{\mathbb{R}^{m\times n}}h(x,F)\mathrm{d}\nu_x(F)\mathrm{d}x$. Conversely, every L^p-Young measure defines a Young functional by the previous formula. Notice that the choice $H:=L^1(\Omega;C_0(\mathbb{R}^{m\times n}))$ identifies $Y^p(\Omega;\mathbb{R}^{m\times n})$ and $\mathcal{Y}^p(\Omega;\mathbb{R}^{m\times n})$. If the net in the definition of $Y^p_H(\Omega;\mathbb{R}^{m\times n})$ is a net of gradients of mappings from $W^{1,p}(\Omega;\mathbb{R}^m)$, then we call such a Young functional the gradient Young functional and their set is denoted $G^p_H(\Omega;\mathbb{R}^{m\times n})$, i.e.,

$$\begin{aligned}G^p_H(\Omega;\mathbb{R}^{m\times n}):=\{\eta\in H^*;\ \exists\{z_\alpha\}_\alpha\subset W^{1,p}(\Omega;\mathbb{R}^m)\,\text{a bounded net};\\ \mathrm{w}^*-\lim\, i(\nabla z_\alpha)=\eta\}.\end{aligned}$$

The interesting feature is that we have some freedom in the choice of H. Larger H makes the description of $G^p_H(\Omega;\mathbb{R}^{m\times n})$ more complicated (see Lemma 1.1 above) but allows us to evaluate the extension (relaxation) of a larger set of functionals to the space H^*. On the other hand, $G^p_H(\Omega;\mathbb{R}^{m\times n})$ can be easily characterized under crucial restrictions on H. The following result, which can be found in [30], demonstrates this statement. If $F\in\mathbb{R}^{m\times n}$, then $\mathbb{T}_s(F)$ denotes the vector of all subdeterminants of order $1\le s\le\min(m,n)$. Clearly, $\mathbb{T}_s(F)\in\mathbb{R}^{\sigma(s)}$, where $\sigma(s):=\binom{m}{s}\binom{n}{s}$. If U is a linear separable subspace of $C^p(\mathbb{R}^{m\times n})$, we recall that $C(\bar\Omega)\otimes U$ consists of all finite sums $\sum_j g_j(x)v_j(F)$, where $g_j\in C(\bar\Omega)$ and $v_j\in U$ for all j. Moreover, $[g\otimes v](x,F):=g(x)v(F)$ if $g\in C(\bar\Omega)$ and $v\in U$ and $\langle h\bullet\eta,g\rangle:=\langle\eta,g\cdot h\rangle$ for all $g\in C(\bar\Omega)$.

LEMMA 1.3. *Let* $\min(m,n)\le p$ *and* H *contain densely the space* $C(\bar\Omega)\otimes U$, *where* U *is a separable linear subspace of* $C^p(\mathbb{R}^{m\times n})$ *containing the function* $F\mapsto\mathbb{T}(F)$, *and let it hold that every function from* U *has its quasi-convex envelope polyconvex. Moreover, if for some* $\eta\in Y^p_H(\Omega;\mathbb{R}^{m\times n})$ p*-nonconcentrating, some* $y\in W^{1,p}(\Omega;\mathbb{R}^m)$, *and all* $1\le s\le\min(m,n)$

$$(1 \otimes \mathbb{T}_s) \bullet \eta = \mathbb{T}_s(\nabla y) \quad \text{in } L^{p/s}(\Omega; \mathbb{R}^{\sigma(s)}),$$

then $\eta \in G^p_H(\Omega; \mathbb{R}^{m\times n})$.

In particular, it means that only a finite number of conditions is needed to characterize gradient Young functionals in this situation. Young measures and Young functionals are an important tool in the mathematical treatment and relaxation of various nonconvex variational problems. A prominent example is the relaxation of energy functionals in the modeling of shape-memory materials [5], [22], [27], [31].

1.2. Shape-memory alloys. Shape-memory alloys (SMAs) have been the subject of intensive theoretical and experimental research during the past decades. Existing or potential applications can be found, for example, in medicine and mechanical or aerospace engineering. Shape-memory alloys are crystalline materials that exhibit specific hysteretic stress, strain, or temperature response; they have the ability to recover a trained shape after deformation and subsequent reheating. This is called the shape-memory effect. It is based on the ability of the alloy to rearrange atoms in different crystallographic configurations (in particular, with different symmetry groups). The stability depends on the temperature. Normally, at higher temperatures a high-symmetry (for example, cubic) lattice is stable, which is referred to as the austenite phase. At lower temperatures, a lattice of lower symmetry (for example, tetragonal, orthorhombic, monoclinic, or triclinic) becomes stable, called the martensite phase. Because of the loss of symmetry, this phase may occur in different variants. The number of variants M is the quotient of the order of the high-symmetry phase and the order of the low-symmetry group. So for a cubic high-symmetry phase, $M = 3$, 6, 12, or 4 for the tetragonal, orthorhombic, monoclinic, or triclinic martensites, respectively, mentioned above. The variants can be combined coherently with each other, forming so-called twins of two variants.

The mathematical and computational modeling of SMAs represents a tool for the theoretical understanding of phase transition processes in solids. Such an analysis may complement experimental results, predict the response of new materials, or facilitate the usage of SMAs in applications. SMAs are genuine multiscale materials and create a variety of challenges for mathematical modeling. We refer the reader to [32] for a survey of a wide menagerie of SMA models ranging from nano- to macroscales. In this article, we focus on a mesoscopic model in the framework of continuum mechanics. Beside the macroscopic deformation and its gradient, the model also involves the volume fractions of phases and variants and gradients of volume fraction. This seems a reasonable compromise, since it allows for the modeling scales of large single crystals or polycrystals.

Although the natural physical dimension is three, we will assume that our specimen occupies a bounded domain $\Omega \subset \mathbb{R}^n$ and the deformation y maps Ω to $\mathbb{R}^m$. This allows us to consider various variational problems. Nevertheless for shape-memory applications we obviously assume that $m = n = 3$. The stress-free parent austenite is a natural state of the material that makes it, in the context of continuum mechanics, a canonical choice for the reference configuration. As usual, $y\colon\Omega \to \mathbb{R}^m$ denotes the deformation and $u\colon\Omega \to \mathbb{R}^m$ the displacement, which are related to each other via the identity $y(x) = x + u(x)$, where $x \in \Omega$. Hence the deformation gradient is $F := \nabla y = \mathbb{I} + \nabla u$.

The total stored energy in the bulk occupying, in its reference configuration, the domain Ω is then

$$V(y) := \int_\Omega W(\nabla y(x))\mathrm{d}x. \tag{10}$$

A common variational principle in continuum mechanics is the minimization of the stored energy. Because of the coexistence of several variants at low temperature, W has multiple minima and thus a multiwell character. We consider an isothermal situation with several coexisting variants. Since W is a multiwell energy density, minimizing sequences of V tend to develop, in general, finer and finer spatial oscillations of their gradients. In other words, the deformation gradient often tends to develop fine spatial oscillations due to lack of (quasi-)convexity of the stored energy density. These oscillations are difficult to model in full detail, although some studies in this direction exist [1]. The oscillations correspond to the development of finer and finer microstructures when the stored energy is to be minimized. The minimum of V, under specific boundary conditions for y, is usually not attained in a space of functions. Therefore one needs to extend the notion of a solution. Young measures are here an appropriate tool. They are capable of recording, on a mesoscopic level, the limit information of the finer and finer oscillating deformation gradient as we move toward the macroscopic scale. This can be described, for a current macroscopic point $x \in \Omega$, by a probability measure ν_x on the set of deformation gradients, that is, matrices in $\mathbb{R}^{m\times n}$. The extension of V to Young measures then reads as [26]

$$\bar{V}(\nu) := \int_\Omega \int_{\mathbb{R}^{m\times n}} W(F)\mathrm{d}\nu_x(F)\mathrm{d}x. \tag{11}$$

1.2.1. Dissipation related to phase transitions. In order to describe dissipation due to transformations we adopt, following, e.g., [22], the standpoint that the amount of dissipated energy associated with a particular phase transition between austenite and a martensitic variant or between two martensitic variants can be described by a specific energy (of the dimension $\mathrm{J}/\mathrm{m}^3 = \mathrm{Pa}$). This viewpoint has been independently adopted in physics; see [16]. For an explicit definition of the transformation dissipation, we need to identify the particular phases or phase variants. In this behalf, we define a continuous mapping $\mathcal{L}:\mathbb{R}^{n\times n} \to \Delta$, where

$$\Delta := \left\{ \zeta \in \mathbb{R}^{1+M} \middle| \zeta_\ell \geq 0 \text{ for } \ell = 1, \ldots, M+1, \text{ and } \sum_{\ell=1}^{M+1} \zeta_\ell = 1 \right\}$$

is a simplex with $M+1$ vertices, with M being the number of martensitic variants. Here $\mathcal{L}$ is related to the material itself and thus has to be frame indifferent. We assume, beside $\zeta_\ell \geq 0$ and $\sum_{\ell=1}^{M+1} \zeta_\ell = 1$, that the coordinate ζ_ℓ of $\mathcal{L}(F)$ takes the value 1 if F is in the ℓth (phase) variant; that is, F is in a vicinity of ℓth well $\mathrm{SO}(n)\,U_\ell$ of W, which can be identified by the stretch tensor $F^\top F$ being close to $U_\ell^\top U_\ell$. If $\mathcal{L}(F)$ is not in any vertex of Δ, then it means that F is in the spinodal region where no definite phase or variant is specified. We assume, however, that the wells are sufficiently deep and the phases and variants are geometrically sufficiently far from each other that the tendency for minimization of the stored energy will essentially prevent F from ranging into the spinodal region. Thus, the concrete form of $\mathcal{L}$ is not important as long as $\mathcal{L}$ enjoys the properties listed above. We remark that $\mathcal{L}$ plays the role of what is often called vector of order parameters or a vector-valued internal variable.

For two states q_1 and q_2, with $q_j = (y_j, \nu_j, \lambda_j)$ for $j = 1, 2$, we now define the dissipation due to martensitic transformation that "measures" changes in the volume fraction $\lambda \in L^\infty(\Omega;\mathbb{R}^{M+1})$. This dissipation is given by

$$\mathcal{D}(q_1, q_2) := \int_\Omega |\lambda_1(x) - \lambda_2(x)|_{\mathbb{R}^{M+1}} \mathrm{d}x, \tag{12}$$

where

$$\lambda_j(x) := \int_{\mathbb{R}^{m\times n}} \mathcal{L}(F)\nu_{j,x}(\mathrm{d}F) \tag{13}$$

and $|\cdot|_{\mathbb{R}^{M+1}}$ is a norm on $\mathbb{R}^{M+1}$. As $\sum_{j=1}^{M+1}\lambda^j = 1$ we call λ the vector-valued volume fraction, as it gives us relative portions of variants at almost every $x \in \Omega$. In what follows, we will assume that the norm on $\mathbb{R}^{M+1}$ defining the dissipation in (12) is given as

$$|X|_{\mathbb{R}^{M+1}} := \sum_{i=1}^{M+1} c^i |X^i|, \qquad X = (X^1, \dots, X^{M+1}), \tag{14}$$

where $|\cdot|$ is the absolute value and $c^i > 0$ for all i. The physical meaning of c^i is the specific energy dissipated if X^i changes from zero to one (or vice versa).

1.2.2. Loading and boundary conditions. In experiments, a specimen occupying the region Ω will be subjected to external loads. To simplify our exposition, we consider only surface forces. We assume that there is a spatially constant tensor S taking values in $\mathbb{R}^{m\times n}$ such that the density of surface forces g applied on $\Gamma_1 \subset \partial\Omega$ is given by $g = S\varrho$, where ϱ is the unit outer normal vector to Γ_1. We also assume that we are given a set $\Gamma_0 \subset \partial\Omega$, where the $(n-1)$-dimensional Hausdorff measure of Γ_0 is positive. We consider Dirichlet boundary conditions $y = y_0$ on Γ_0 for some prescribed (time-dependent/independent) mapping y_0. As for the surface forces, we define a linear functional

$$L(y) := \int_{\Gamma_1} S\varrho \cdot y(x)\mathrm{d}A = \int_\Omega S \cdot \nabla y(x)\mathrm{d}x. \tag{15}$$

Below, we write $L = L(t, y)$ to indicate the possibility of temporally changing forces g and therefore also S.

1.3. Energetic solution. Combining the previous considerations, we arrive at the energy functional $\mathcal{I}$ [26] of the form

$$\mathcal{I}(t, q) := \int_\Omega \int_{\mathbb{R}^{m\times n}} (W(F) - S(t)\cdot F)\mathrm{d}\nu_x(F)\mathrm{d}x + \varepsilon\|\nabla\lambda\|_{L^2(\Omega;\mathbb{R}^{(1+M)\times n})}. \tag{16}$$

It is often convenient to write $\bar{V}$ with the argument q instead of ν:

$$\bar{V}(q) = \int_\Omega \int_{\mathbb{R}^{m\times n}} W(F)\mathrm{d}\nu_x(F)\mathrm{d}x + \varepsilon\|\nabla\lambda\|_{L^2(\Omega;\mathbb{R}^{(1+M)\times n})}, \tag{17}$$

where the $\nabla\lambda$-term is included to regularize the problem. It penalizes spatial jumps of the volume fraction λ and introduces a length scale to the problem depending on a parameter $\varepsilon > 0$. In particular, it allows us to pass to the limit in the dissipation term. To define an admissible set where we look for our solution triple $q = (y, \nu, \lambda)$, we put

$$y \in \mathbb{Y}^p(\Omega;\mathbb{R}^m) := \{y \in W^{1,p}(\Omega;\mathbb{R}^m) | y = 0 \text{ on } \Gamma_0\}, \tag{18}$$

where $\Gamma_0 \subset \partial\Omega$ with a positive surface measure, as described in subsection 1.2.2. Here y_0 is a time-dependent trace on Γ_0. We recall from that subsection that $\Gamma_0 \cap \Gamma_1 = \emptyset$. Then

we look for $q \in \mathcal{Q} := \mathbb{Y}^p(\Omega;\mathbb{R}^m) \times \mathbb{G}^p(\Omega;\mathbb{R}^{m\times n}) \times W^{1,2}(\Omega;\mathbb{R}^{M+1})$ and restrict the space further by imposing the admissibility condition

$$\mathbb{Q} := \{q \in \mathcal{Q} | \lambda = \mathcal{L} \bullet \nu \text{ and } \nabla y = \mathbb{I} \bullet \nu\}, \tag{19}$$

where, for almost all $x \in \Omega$, $[\mathcal{L} \bullet \nu](x) := \int_{\mathbb{R}^{m\times n}} \mathcal{L}(F)\nu_x(\mathrm{d}F)$; $\mathbb{I} \bullet \nu$ is defined analogously.

In what follows, we suppose that (2) holds. Following [12] we assume that there are constants $C_0, C_1 > 0$ such that

$$|\partial_t \mathcal{I}(t,q)| \le C_0(C_1 + \mathcal{I}(t,q)). \tag{20}$$

Let $\mathfrak{T} > 0$ denote the time horizon. We assume uniform continuity of $t \mapsto \partial_t \mathcal{I}(t,q)$ in the sense that there is $\omega : [0,\mathfrak{T}] \to [0,+\infty)$ nondecreasing such that for all $t_1, t_2 \in [0,\mathfrak{T}]$

$$|\partial_t \mathcal{I}(t_1,q) - \partial_t \mathcal{I}(t_2,q)| \le \omega(|t_1 - t_2|). \tag{21}$$

We also suppose that $q \mapsto \partial_t \mathcal{I}(t,q)$ is weakly continuous for all $t \in [0,\mathfrak{T}]$.

Remark 1.4. If we prescribe time-dependent boundary conditions $y_0(t) \in W^{1,p}(\Omega;\mathbb{R}^m)$ on Γ_0, then we write

$$\begin{aligned}\mathcal{I}(t,q) := \int_\Omega \int_{\mathbb{R}^{m\times n}} &(W(F + \nabla y_0(t,x)) - S(t) \\ &\cdot (F + \nabla y_0(t,x))\mathrm{d}\nu_x(F))\mathrm{d}x + \varepsilon \|\nabla \lambda\|_{L^2(\Omega;\mathbb{R}^{(1+M)\times n})}.\end{aligned} \tag{22}$$

This allows us to keep $\mathbb{Q}$ independent of time.

We seek to analyze the time evolution of a process $q(t) \in \mathbb{Q}$ during the time interval $[0,\mathfrak{T}]$. The following two properties are key ingredients of the so-called energetic solution introduced by Mielke and Theil [23]; see also [24].

(i) *Stability inequality.* For every $t \in [0,\mathfrak{T}]$ and every $\tilde{q} \in \mathbb{Q}$, it holds that

$$\mathcal{I}(t,q(t)) \le \mathcal{I}(t,\tilde{q}) + \mathcal{D}(q(t),\tilde{q}). \tag{23}$$

(ii) *Energy balance.* For every $0 \le t \le \mathfrak{T}$,

$$\mathcal{I}(t,q(t)) + \mathrm{Var}(\mathcal{D},q;[0,t]) = \mathcal{I}(0,q(0)) + \int_0^t \partial_t \mathcal{I}(\xi,q(\xi))\mathrm{d}\xi, \tag{24}$$

where

$$\mathrm{Var}(\mathcal{D},q;[s,t]) := \sup\left\{ \sum_{j=1}^N \mathcal{D}(q(t_{j-1}),q(t_j)) \,\middle|\, \{t_j\}_{j=0}^N \text{ is a partition of } [s,t] \right\}$$

is the *variation* of the dissipation.

Definition 1.5. *The mapping $q : [0,\mathfrak{T}] \to \mathbb{Q}$ is an energetic solution to the problem $(\mathcal{I},\mathcal{D})$ with the energy functional $\mathcal{I}$ as in* (16) *and the dissipation $\mathcal{D}$ as in* (15) *if the stability inequality* (23) *and energy balance* (24) *are satisfied for every $t \in [0,\mathfrak{T}]$.*

Further, we define the set of stable states at time $t \in [0,\mathfrak{T}]$ as

$$\mathbb{S}(t) := \{q \in \mathbb{Q};\ \forall \tilde{q} \in \mathbb{Q} : \mathcal{I}(t,q) \le \mathcal{I}(t,\tilde{q}) + \mathcal{D}(q,\tilde{q})\}.$$

In particular, we will always assume that the initial condition is stable, i.e., $q_0 \in \mathbb{S}(0)$. The following theorem regarding the existence of an energetic solution can be proved using a general strategy described in [12].

THEOREM 1.6. *Let $p > 1$, $S \in C^1([0,\mathfrak{T}];\mathbb{R}^{m\times n})$, and let assumptions* (2), (20), *and* (21) *hold. Then there is a process $q\colon[0,\mathfrak{T}] \to \mathbb{Q}$ with $q(t) = (y(t), \nu(t), \lambda(t))$ such that q is an energetic solution according to Definition* 1.5 *for a given stable initial condition $q_0 \in \mathbb{Q}$.*

The proof of this theorem proceeds via semidiscretizations in time for decreasing time steps, by a limit passage in the stability inequality (23) and in the energy equality (24); cf. [12]. We refer to [22] for the proof in case of relaxation using Young measures. The Tikhonoff compactness theorem is used to pass to the limiting Young measure.

In what follows, we are going to concentrate on a numerical solution to time-discrete approximations of energetic solutions. As it leads to a sequence of nonconvex global-minimization problems, it is a rather difficult task. Our new idea is to replace the set $\mathbb{G}^p(\mathbf{\Omega};\mathbb{R}^{m\times n})$ of gradient Young measures in the definition of $\mathbb{Q}$ by a larger set of polyconvex Young measures. This leads to a well-posed minimization problem for which we will derive necessary and sufficient optimality conditions if the deformation y is kept fixed; see Proposition 3.4. These are then used for efficient numerical solution in section 4, which is far more effective than existing computational strategies. This is demonstrated on computational examples.

2. Incremental problems. The proof of Theorem 1.6 relies on approximations by time-discrete (incremental) problems constructed for a given time step. These are minimization problems over spatial variables. Each minimization problem takes into account the solution obtained for the previous time step, while the initial condition serves as input for the first minimization problem. Details can be found, e.g., in [12]. In practical applications, finding global minimizers of incremental minimization problems is extremely important because global minimization is the key ingredient in the energetic-solution approach.

Thus, in a first step, a sequence of incremental problems is defined. We define a time discretization $0 = t_0 < \cdots < t_N = \mathfrak{T}$ with a time step $\tau := \max_i(t_i - t_{i-1})$. Let an initial state $\mathbb{S}(0) \ni q_0 =: q_0 \in \mathbb{Q}$ be given. For $1 \le j \le N$ we find $q_j \in \mathbb{Q}$ by solving

$$\text{minimize}\, \mathcal{I}(t_j, q) + \mathcal{D}(q_{j-1}, q), \quad \text{subject to } q \in \mathbb{Q}. \tag{25}$$

The existence of a solution to the time-step problem (25) follows in analogy to [12, Theorem 3.2] by the direct method of the calculus of variations. Notice that this is true even if $\varepsilon = 0$ and if we consider $\lambda \in L^2(\mathbf{\Omega};\mathbb{R}^{M+1})$ only.

We denote $q \in \mathbb{QP} := \mathbb{Y}^p(\mathbf{\Omega};\mathbb{R}^m) \times \mathbb{P}^p(\mathbf{\Omega};\mathbb{R}^{m\times n}) \times L^2(\mathbf{\Omega};\mathbb{R}^{M+1})$,

$$\mathbb{P} := \{q \in \mathbb{QP} | \lambda = \mathcal{L} \bullet \nu \text{ and } \nabla y = \mathbb{I} \bullet \nu\}; \tag{26}$$

moreover, clearly $\mathbb{Q} \subset \mathbb{P}$.

Then we can define the following incremental problem: For $1 \le j \le N$ we find $q_j \in \mathbb{P}$ by solving

$$\text{minimize}\, \mathcal{I}(t_j, q) + \mathcal{D}(q_{j-1}, q), \quad \text{subject to } q \in \mathbb{P}. \tag{27}$$

The existence of a solution to (27) follows again by the direct method of the calculus of variations. Moreover, as proved, e.g., in [12], the following two-sided energy estimate holds for $k \ge 1$:

$$\int_{t_{j-1}}^{t_j} \partial_t \mathcal{I}(s, q_j)\mathrm{d}s \le \mathcal{I}(t_j, q_j) + \mathcal{D}(q_{j-1}, q_j) - \mathcal{I}(t_{j-1}, q_{j-1}) \le \int_{t_{j-1}}^{t_j} \partial_t \mathcal{I}(s, q_{j-1})\mathrm{d}s. \tag{28}$$

Therefore, having q_{j-1} we look for $q_j \in \mathbb{P}$, which minimizes

$$\int_\Omega \int_{\mathbb{R}^{m\times n}} (W(F) - S(t_j)\cdot F)\mathrm{d}\nu_x(F)\mathrm{d}x + \sum_{i=1}^{M+1} \int_\Omega c^i |\lambda^i(x) - \lambda^i_{j-1}(x)|\mathrm{d}x. \tag{29}$$

This problem is nonsmooth, so using the Mosco transform we define an equivalent smooth problem (with inequality constraints), which includes $M+1$ auxiliary variables, namely,

$$\begin{aligned} \text{minimize } & \int_\Omega \int_{\mathbb{R}^{m\times n}} (W(F) - S(t_j)\cdot F)\mathrm{d}\nu_x(F)\mathrm{d}x + \sum_{i=1}^{M+1} \int_\Omega a^i(x)\mathrm{d}x \\ \text{subject to } & -a^i - c^i\mathcal{L}^i \bullet \nu \le -c^i\lambda^i_{j-1} \quad \forall\ 1 \le i \le M+1 \\ & -a^i + c^i\mathcal{L}^i \bullet \nu \le c^i\lambda^i_{j-1} \quad \forall\ 1 \le i \le M+1 \\ & (y, \nu, \lambda) \in \mathbb{P}, \qquad a^i \in L^2(\Omega). \end{aligned} \tag{30}$$

We invite the reader to verify that (29) and (30) are equivalent in the sense that minima are the same, and if (y, ν, λ) solves (29), then $(y, \nu, \lambda, \{c^i|\lambda^i - \lambda^i_{j-1}|\}_i)$ solves (30); conversely if $(y, \nu, \lambda, \{a^i\}_i)$ solves (30), then (y, ν, λ) solves (29). A proof can be also found in [18].

It will be convenient to rewrite (30) in terms of Young functionals. In additional to the assumptions from Lemma 1.3, the underlying space H must also contain $(x, F) \mapsto 1 \otimes (W(F) + S(t_j)\cdot F)$ and $(x, F) \mapsto g(x)\mathcal{L}^i(F)$ for all $1 \le i \le M+1$ and $g \in L^1(\Omega)$. Let us denote $\Phi\colon H^* \times L^2(\Omega; \mathbb{R}^{M+1}) \to \mathbb{R}$, where

$$\Phi(\eta, a) \coloneqq \langle \eta, W - S(t_j)\cdot\rangle + \sum_{i=1}^{M+1} \int_\Omega a^i(x)\mathrm{d}x \tag{31}$$

and for $1 \le i \le M+1$ let $R^i_1, R^i_2\colon H^* \times L^2(\Omega; \mathbb{R}^{M+1}) \to L^1(\Omega)$ with $R^i_1(\eta, a) \coloneqq -a_i - c^i(1 \otimes \mathcal{L}^i) \bullet \eta$ and $R^i_2(\eta, a) \coloneqq -a_i + c^i(1 \otimes \mathcal{L}^i) \bullet \eta$. The problem (30) now reads as

$$\begin{aligned} \text{minimize } & \Phi(\eta, a) \\ \text{subject to } & (1 \otimes \mathbb{T}) \bullet \eta = \mathbb{T}(\nabla y) \\ & R^i_1(\eta, a) \le -c^i\lambda^i_{j-1} \quad \forall\ 1 \le i \le M+1 \\ & R^i_2(\eta, a) \le c^i\lambda^i_{j-1} \quad \forall\ 1 \le i \le M+1 \\ & (y, \eta, a) \in \mathbb{Y}^p(\Omega; \mathbb{R}^m) \times Y^p_H(\Omega; \mathbb{R}^{m\times n}) \times L^2(\Omega; \mathbb{R}^{M+1}). \end{aligned} \tag{32}$$

Notice that the constraint $(1 \otimes \mathbb{T}) \bullet \eta = \mathbb{T}(\nabla y)$ is nonconvex while the other constraints and Φ are all linear. Problem (32) is a two-scale problem. On the large scale we optimize the deformation $y \in \mathbb{Y}^p(\Omega; \mathbb{R}^m)$, while on the smaller scale, keeping ∇y fixed, we look for an optimal Young functional $\eta \in Y^p_H(\Omega; \mathbb{R}^m)$, which encodes material microstructures. We can see it as an example of a Stackelberg leadership game [33]. The deformation y is a leader, while particular η is a follower trying to minimize the energy if ∇y is fixed. In view of this, we can consider a special case of (32) when we keep ∇y

constant in $\mathbb{R}^{m\times n}$, so that we put $A := \nabla y$. Hence, if we denote $\Pi_s(\eta, a) = (1 \otimes \mathbb{T}_s) \bullet \eta$, where $\Pi_s : H^* \times L^2(\Omega;\mathbb{R}^{M+1}) \to L^{p/s}(\Omega;\mathbb{R}^{\sigma(s)})$ $(1 < s < p/\min(m,n))$ and $\Pi(\eta, a) := (\Pi_1(\eta, a), \dots, \Pi_{\min(m,n)}(\eta, a))$, $R_i(\eta, a) := (R_i^1(\eta, a), \dots, R_i^{M+1}(\eta, a))$ for $i = 1, 2$, and $\Lambda_{j-1} := (c^1\lambda_{j-1}^1, \dots, c^{M+1}\lambda_{j-1}^{M+1})$, we must solve the following *convex* problem:

$$\begin{aligned} &\text{minimize } \Phi(\eta, a)\\ &\text{subject to } \Pi(\eta, a) = \mathbb{T}(A)\\ &R_1(\eta, a) \le -\Lambda_{j-1}\\ &R_2(\eta, a) \le \Lambda_{j-1}\\ &(\eta, a) \in Y_H^p(\Omega;\mathbb{R}^{m\times n}) \times L^2(\Omega;\mathbb{R}^{M+1}), \end{aligned} \tag{33}$$

where the inequality constraints are understood componentwise.

3. Optimality conditions for (33). It is proved in [29, Lemma 3.3] that for any $A \in \mathbb{R}^{m\times n}$ there is $\eta \in Y_H^p(\Omega;\mathbb{R}^{m\times n})$ such that $\Pi_s(\eta, a) = \mathbb{T}_s(A)$ for all s admissible. Further, the range of R_1^i and R_2^i is the whole space $L^2(\Omega)$. Let us still denote $\mathbb{M}_1 = \{f \in L^2(\Omega;\mathbb{R}^{M+1}); f \le -\Lambda_{j-1}\}$ and $\mathbb{M}_2 = \{f \in L^2(\Omega;\mathbb{R}^{M+1}); f \le \Lambda_{j-1}\}$, and we write for $i = 1, 2$

$$(R_i)^{-1}(\mathbb{M}_i) := \{(\eta, a) \in Y_H^p(\Omega;\mathbb{R}^m) \times L^2(\Omega;\mathbb{R}^{M+1}); R_i(\eta, a) \in \mathbb{M}_i\}.$$

If $\mathcal{B}$ is a convex closed set in a linear space Z, we define the normal cone to $\mathcal{B}$ at $b \in \mathcal{B}$ as

$$N_{\mathcal{B}}(b) := \{\xi \in Z^*;\ \forall \tilde{b} \in \mathcal{B} : \langle \xi, \tilde{b} - b\rangle \le 0\}.$$

Moreover, we denote $\Pi_s^* : L^{p/(p-s)}(\Omega;\mathbb{R}^{\sigma(s)}) \to H^{**} \times L^2(\Omega;\mathbb{R}^{M+1})$ the adjoint operator to Π_s; similarly $[R_j^i]^* : L^2(\Omega) \to H^{**} \times L^2(\Omega;\mathbb{R}^{M+1})$ is adjoint to R_j^i.

The optimality conditions state (see [29], [36]) that there exist Lagrange multipliers $\mu_0 \in \Pi_{s=1}^{\min(m,n)} L^{p/(p-s)}(\Omega;\mathbb{R}^{\sigma(s)})$, $\mu_1 \in N_{\mathbb{M}_1}(R_1(\eta, a))$, $\mu_2 \in N_{\mathbb{M}_2}(R_2(\eta, a))$ such that

$$[\Pi^*]_a\mu_0 - [R_1^*]_a\mu_1 - [R_2^*]_a\mu_2 - \nabla_a\Phi(\eta, a) = 0, \tag{34}$$

$$[\Pi^*]_\eta\mu_0 - [R_1^*]_\eta\mu_1 - [R_2^*]_\eta\mu_2 - \nabla_\eta\Phi(\eta, a) \in N_{Y_H^p(\Omega;\mathbb{R}^m)}(\eta). \tag{35}$$

We start with a computation of $\nabla\Phi$.

LEMMA 3.1. *Let $W \in C^p(\Omega;\mathbb{R}^{m\times n})$. Then it holds for all $(\eta, a) \in H^* \times L^2(\Omega;\mathbb{R}^{M+1})$ that $\nabla\Phi(\eta, a) = (\nabla_\eta\Phi(\eta, a), \nabla_a\Phi(\eta, a)) = (W - S(t_j)\,\cdot, 1, \dots, 1) \in H \times L^2(\Omega;\mathbb{R}^{M+1})$.*

Proof. It is easy because Φ depends on (η, a) linearly. ∎

LEMMA 3.2. *If $1 \le s \le \min(m, n)$, then $[\Pi_s^*]\mu := ([\Pi_s^*]_\eta\mu, [\Pi_s^*]_a\mu) = (\mu \otimes \mathbb{T}_s, 0)$ for any $\mu \in L^{p/(p-s)}(\Omega;\mathbb{R}^{\sigma(s)})$.*

Proof. We have for arbitrary $(\eta, a) \in H^* \times L^2(\Omega;\mathbb{R}^{M+1})$ and $\mu \in L^{p/(p-s)}(\Omega;\mathbb{R}^{\sigma(s)})$

$$\begin{aligned} \langle [\Pi_s^*]\mu, (\eta, a)\rangle_{(H^{**}\times L^2, H^*\times L^2)} &= \langle \mu, \Pi_s(\eta, a)\rangle_{(L^{p/(p-s)}, L^{p/s})} = \langle \mu, (1 \otimes \mathbb{T}_s) \bullet \eta\rangle_{(L^{p/(p-s)}, L^{p/s})}\\ &\langle (1 \otimes \mathbb{T}_s) \bullet \eta, \mu\rangle_{(L^{p/s}, L^{p/(p-s)})} = \langle \eta, \mu \otimes \mathbb{T}_s\rangle_{(H^*, H)}. \end{aligned}$$ ∎

LEMMA 3.3. *It holds for every* $\mu \in L^2(\Omega)$ *and* $(\eta, a) \in H^* \times L^2(\Omega; \mathbb{R}^{M+1})$ *that* $[R_1^i]^*\mu = (-\mu \otimes c^i\mathcal{L}^i, -\mu)$ *and* $[R_2^i]^*\mu = (\mu \otimes c^i\mathcal{L}^i, -\mu)$.

Proof. We compute only the expression for $[R_1^i]^*\mu$ as the other can be obtained similarly. We have

$$\begin{aligned}\langle [R_1^i]^*\mu], (\eta, a)\rangle_{(H^{**}\times L^2, H^*\times L^2)} &= \langle \mu, R_1^i(\eta, a)\rangle_{(L^2,L^2)} = \langle \mu, -a^i - c^i(1 \otimes \mathcal{L}^i) \bullet \eta\rangle_{(L^2,L^2)} \\ &= \langle -a^i - c^i(1 \otimes \mathcal{L}^i) \bullet \eta, \mu\rangle_{(L^2,L^2)} = -\langle \eta, c^i\mu \otimes \mathcal{L}^i\rangle_{(H^*,H)} - \langle \mu, a^i\rangle_{(L^2,L^2)}. \qquad \square\end{aligned}$$

As all the expressions on the left-hand side of (35) take their values in H rather than in H^{**}, we search for the intersection of $N_{Y^p_H(\Omega;\mathbb{R}^{m\times n})}$ with H, i.e.,

$$\begin{aligned} N_{Y^p_H(\Omega;\mathbb{R}^{m\times n})}(\eta) \cap H &= \{h \in H;\ \forall \tilde{\eta} \in Y^p_H(\Omega; \mathbb{R}^{m\times n})\langle \tilde{\eta}, h\rangle \le \langle \eta, h\rangle\} \\ &= \left\{h \in H; \langle \eta, h\rangle = \sup_{u\in L^p(\Omega;\mathbb{R}^{m\times n})} \int_\Omega h(x, u(x))\mathrm{d}x\right\}. \end{aligned} \tag{36}$$

Defining the Hamiltonian

$$\mathcal{H}_\mu(t, x, F) := -W(F) + S(t) \cdot F + \mu_0(x) \cdot \mathbb{T}(F) + \sum_{i=1}^{M+1} c^i(\mu_1^i(x) - \mu_2^i(x))\mathcal{L}^i(F), \tag{37}$$

where $\mu = (\mu_0, \mu_1, \mu_2)$ is the triple of vector-valued Lagrange multipliers, the inclusion (35) reads that if (η, a) is a solution to (33), then there is a Lagrange multiplier $\mu = (\mu_0, \mu_1, \mu_2)$ such that

$$\langle \eta, \mathcal{H}_\mu(t_j, \cdot, \cdot)\rangle = \sup_{u\in L^p(\Omega;\mathbb{R}^{m\times n})} \int_\Omega \mathcal{H}_\mu(t_j, x, u(x))\mathrm{d}x.$$

As our problem is convex, this condition is also sufficient. Notice that (34) means that $\mu_1^i + \mu_2^i = 1$ for all $i = 1, \ldots, M+1$.

If $\nu \in \mathcal{Y}^p(\Omega; \mathbb{R}^{m\times n})$ is a Young measure representing η (we write $\eta \cong \nu$) in the sense that $\langle \eta, h\rangle = \int_\Omega \int_{\mathbb{R}^{m\times n}} h(, x, F)\mathrm{d}\nu_x(F)\mathrm{d}x$ for all $h \in H$, the previous equality reads

$$\int_\Omega \int_{\mathbb{R}^{m\times n}} \mathcal{H}_\mu(t_j, x, F)\mathrm{d}\nu_x(F)\mathrm{d}x = \sup_{u\in L^p(\Omega;\mathbb{R}^{m\times n})} \int_\Omega \mathcal{H}_\mu(t_j, x, u(x))\mathrm{d}x. \tag{38}$$

The maximum principle (38) can be localized to a pointwise one using the arguments of [29, Theorem 3.2].

PROPOSITION 3.4. *Let* $p > \min(m, n)$ *and* (2) *hold. If* $(\eta, a) \in Y^p_H(\Omega; \mathbb{R}^{m\times n}) \times L^2(\Omega; \mathbb{R}^{M+1})$ *solves* (33) *and* η *is represented by* $\nu \in \mathcal{Y}^p(\Omega; \mathbb{R}^{m\times n})$, *then there exist* $\mu_0 \in \Pi_{s=1}^{\min(m,n)} L^{p/(p-s)}(\Omega; \mathbb{R}^{\sigma(s)})$ *and* $\mu_1, \mu_2 \in L^2(\Omega; \mathbb{R}^{M+1})$ *such that for almost all* $x \in \Omega$

$$\max_{F\in\mathbb{R}^{m\times n}} \mathcal{H}_\mu(t_j, x, F) = \int_{\mathbb{R}^{m\times n}} \mathcal{H}_\mu(t_j, x, F)\mathrm{d}\nu_x(F). \tag{39}$$

Conversely, if (39) *holds for some* $\nu \in \mathcal{Y}^p(\Omega; \mathbb{R}^{m\times n})$, *the constraints in* (33) *are satisfied and* $\mu_1^i + \mu_2^i = 1$, $\mu_1^i, \mu_2^i \ge 0$ *for all* $i = 1, \ldots, M+1$, *then* $(\eta, a) \cong (\nu, a)$ *is a solution to* (33).

Remark 3.5. If there is no dissipation, i.e., $\mathcal{L} = 0$, and no loading, i.e., $S = 0$, the Hamiltonian in (37) is the same as in [7].

In particular, having a solution to the problem (33) we also know the values of Lagrange multipliers μ, say, μ^*. Then the solution to (33) is the same as the solution to the following minimization problem involving a homogeneous (i.e., independent of $x \in \Omega$) Young measure ν:

$$\begin{aligned} \text{minimize} \int_{\mathbb{R}^{m\times n}} \Big(W(F) - S(t_j)\cdot F - \mu_0^*\cdot \mathbb{T}(F) + \sum_{i=1}^{M+1} (\mu_2^{i*} - \mu_1^{i*})\mathcal{L}^i(F) \Big) \mathrm{d}\nu(F) \\ \text{subject to } \mathbb{T}\bullet\nu = \mathbb{T}(A) \\ \nu \in \mathcal{Y}^p(\Omega;\mathbb{R}^{m\times n}), \qquad \bar{\nu} = A. \end{aligned} \tag{40}$$

In each time step, (40) is a smooth convex minimization problem. The optimal value of the objective function is the polyconvex envelope of $F \mapsto W(F) - S(t_j)\cdot F - \mu_0^*\cdot\mathbb{T}(F) + \sum_{i=1}^{M+1}(\mu_2^{i*} - \mu_1^{i*})\mathcal{L}^i(F)$ evaluated at the point $A \in \mathbb{R}^{m\times n}$ [9], [31].

In what follows, we show how to use advantageously the maximum principle in a numerical-solution strategy. This approach was first suggested and tested in a static scalar one-dimensional example in [8] and further used in micromagnetics calculations in [18], [20].

4. Numerical approximations.

We are about to discuss spatial discretization of incremental problems (27) and describe an efficient strategy leading to their solutions. To keep the explanation as simple as possible, we confine ourselves to the case of two wells, i.e., $M = 1$, and instead of surface forces we drive the evolution of our specimen by time-dependent boundary conditions; i.e., we set $S = 0$. It is straightforward to generalize the method to a more general scenario.

4.1. Problem description.

Recall that the volume fraction is defined by

$$\lambda(x) = \int_{\mathbb{R}^{m\times n}} \mathcal{L}(F)\mathrm{d}\nu_x(F)$$

for almost every $x \in \Omega$. We have by the definition of $\mathcal{L}$ that $\sum_{j=1}^{M+1}\lambda^j = \sum_{j=1}^{M+1}\mathcal{L}^j = 1$. Therefore, if $M = 1$, i.e., in case of the so-called double-well problem, we have that $\lambda^2 := 1 - \lambda^1$. Thus, having two pairs $\lambda_1 := (\lambda^1, 1-\lambda^1)$ and $\lambda_2 := (\lambda^2, 1-\lambda^2)$, we have by (12) that $\mathcal{D}(q_1, q_2) := \int_\Omega (c^1 + c^2)|\lambda^1 - \lambda^2|\mathrm{d}x$. To simplify the notation, we set $c_{\mathcal{D}} := c^1 + c^2$. The dissipation functional then reads

$$\mathcal{D}(q_1, q_2) = c_{\mathcal{D}} \int_\Omega |\lambda^1 - \lambda^2|\mathrm{d}x.$$

4.2. General discrete problem.

Given a finite element partition $\mathcal{T}$ of Ω with $\mathrm{diam}(T) \le h$ for all $T \in \mathcal{T}$ and a set $\mathcal{A}_{d,r} \subset \mathbb{R}^{m\times n}$, $\mathcal{A}_{d,r} \subset d\mathbb{Z}^{m\times n} \cap B_r^\infty(0)$, where $B_r^\infty(0)$ is a ball in $\mathbb{R}^{m\times n}$ centered at zero and with the radius r. We consider a Young measure $\nu = \{\nu_x\}_{x\in\Omega}$ where $\nu_x|_T := \sum_{A\in\mathcal{A}_{d,r}}\theta_{T,A}\delta_A$, with $\{\theta_{T,A}\}_{A\in\mathcal{A}_{d,r}}$ coefficients of a convex combination, and δ_A the Dirac mass supported at A. The deformation y is approximated by a continuous, elementwise affine map y_h defined by its nodal values at the set of nods $\mathcal{N}_h$. A typical time step t_j of the discrete version of (30) consists in solving the following optimization problem:

$$\text{Find } ((y_h(z))_{z\in\mathcal{N}_h}, (a_h|_T)_{T\in\mathcal{T}}, (\theta_{T,A})_{T\in\mathcal{T},A\in\mathcal{A}_{d,r}}), \text{ which minimizes}$$

$$\sum_{T\in\mathcal{T}} |T| \sum_{A\in\mathcal{A}_{d,r}} \theta_{T,A} W(A) + \sum_{T\in\mathcal{T}} c_{\mathcal{D}} |T| a_h|_T$$

$$\text{subject to } y_h|_{\Gamma_{\mathrm{D}}} = y_{D,h}(t_j,\cdot),\ \theta_{T,A} \geq 0,\ \sum_{A\in\mathcal{A}_{d,r}} \theta_{T,A} = 1, \quad \text{and}$$

$$\mathbb{T}(\nabla y_h|_T) = \sum_{A\in\mathcal{A}_{d,r}} \theta_{T,A} \mathbb{T}(A),$$

$$-a_h|_T/c_{\mathcal{D}} - \sum_{A\in\mathcal{A}_{d,r}} \theta_{T,A} \mathcal{L}^1(A) \leq -\lambda^1_{h,j-1}|_T,$$

$$-a_h|_T/c_{\mathcal{D}} + \sum_{A\in\mathcal{A}_{d,r}} \theta_{T,A} \mathcal{L}^1(A) \leq \lambda^1_{h,j-1}|_T$$

$$\text{for all } T \in \mathcal{T}.$$

Notice that the constraint with the left-hand side $\mathbb{T}(\nabla y_h|_T)$ is nonlinear if $\min(m,n) \geq 2$. Otherwise, this defines a linear program that we solve iteratively with an active set strategy based on the maximum principle of section 3. Given the (approximate) solution of this problem, the next time step t_{j+1} is the same problem with $y_{D,h}(t_j)$ replaced by $y_{D,h}(t_{j+1})$ and $\lambda_{h,j-1}$ replaced by

$$\lambda^1_{h,j}|_T = c_{\mathcal{D}} \sum_{A\in\mathcal{A}_{d,r}} \theta_{T,A} \mathcal{L}^1(A)$$

for all $T \in \mathcal{T}$.

4.3. Simplification through enforced homogeneity. To simplify the calculations we consider solutions of the model problem that are spatially homogeneous and the deformation is entirely defined through the affine boundary data $F_{\mathrm{D}}(t)$; i.e., the deformation $y(t,x) = F_{\mathrm{D}}(t)x$ for all $x \in \Omega$ is fully prescribed in the entire evolution. Hence the minimization problem in the jth time step reduces to

$$\text{minimize} \int_{\mathbb{R}^{m\times n}} W(F)\mathrm{d}\nu(F) + c_{\mathcal{D}} a \quad \text{among } (a,\nu) \in \mathbb{R}\times\mathbb{P}_{\mathrm{hom}}$$

$$\text{subject to} \int_{\mathbb{R}^{m\times n}} \mathbb{T}(F)\mathrm{d}\nu(F) = \mathbb{T}(F_{\mathrm{D}}(t_j))$$

$$-a/c_{\mathcal{D}} - \int_{\mathbb{R}^{m\times n}} \mathcal{L}^1(F)\mathrm{d}\nu(F) \leq -\lambda^1_{j-1}, \qquad -a/c_{\mathcal{D}} + \int_{\mathbb{R}^{m\times n}} \mathcal{L}^1(F)\mathrm{d}\nu(F) \leq \lambda^1_{j-1}.$$

4.4. Discretization of spatially homogeneous problems. We choose a finite subset $\mathcal{A}_{d,r} \subset \mathbb{R}^{m\times n}$ and discretize the convex set of (homogeneous) polyconvex Young measures $\mathbb{P}_{\mathrm{hom}} = \{\nu \in \mathbb{P}^p(\Omega;\mathbb{R}^{m\times n}); \nu_x = \nu_y \forall x,y \in \Omega\}$ by

$$\mathbb{P}_{\mathrm{hom}} \subset YM^{\mathrm{hom}}_{\mathcal{A}_{d,r}} = \left\{ \nu_{d,r} = \sum_{A\in\mathcal{A}_{d,r}} \theta_A \delta_A : \theta_A \geq 0,\ \sum_{A\in\mathcal{A}_{d,r}} \theta_A = 1 \right\}.$$

We will identify a discrete Young measure with its convex coefficients $\theta = (\theta_A)_{A\in\mathcal{A}_{d,r}}$. This choice leads to the following discretization of the homogeneous vectorial problem described above:

$$\text{minimize} \sum_{A\in\mathcal{A}_{d,r}} \theta_A W(A) + c_{\mathcal{D}} a \quad \text{among } (a,\nu_{d,r}) \in \mathbb{R}\times YM^{\rm hom}_{\mathcal{A}_{d,r}}$$

$$\text{subject to } \mathbb{T}(F_{\rm D}(t_j)) = \sum_{A\in\mathcal{A}_{d,r}} \theta_A \mathbb{T}(A),$$

$$-a/c_{\mathcal{D}} - \sum_{A\in\mathcal{A}_{d,r}} \theta_A \mathcal{L}^1(A) \le -\lambda^1_{j-1}, \qquad -a/c_{\mathcal{D}} + \sum_{A\in\mathcal{A}_{d,r}} \theta_A \mathcal{L}^1(A) \le \lambda^1_{j-1}.$$

We subsequently set $\lambda^1_j = \sum_{A\in\mathcal{A}_{d,r}} \theta_A \mathcal{L}^1(A)$.

4.5. Efficient solution via active set strategy. The discrete problem is a linear optimization problem that can be solved directly with standard algorithms. Realizing that only a small number of coefficients θ_A will be nonvanishing, it is desirable to employ an iterative scheme in which a large number of vanishing or small coefficients will not be incorporated in the approximating problems. The key to such an iterative method is the optimality condition

$$\max_{F\in\mathcal{A}_{d,r}} (\mathbb{T}(F)\cdot\mu_0 - W(F) + c_{\mathcal{D}}(\mu^1-\mu^2)\mathcal{L}^1(F)) = \sum_{A\in\mathcal{A}_{d,r}} (\theta_A\mathbb{T}(A)\cdot\mu_0 - W(A) + c_{\mathcal{D}}(\mu^1-\mu^2)\mathcal{L}^1(A)),$$

where μ_0 is the Lagrange multiplier related to the equality constraints involving $\mathbb{T}(F_{\rm D}(t_j))$ and (μ_1,μ_2) are the multipliers related to the inequality constraints involving a. Notice that we write μ_1, μ_2 instead of μ^1_1 and μ^1_2 to simplify the notation. It shows that only those atoms $A\in\mathcal{A}_{d,r}$ can have a convex coefficient θ_A different from zero, for which the function on the left-hand side assumes its maximum. Given a guess or a good approximation of the Lagrange multipliers and some tolerance $\varepsilon>0$, those atoms A are activated within an iterative strategy for which the maximum is attained up to the tolerance ε. This defines the new set $\mathcal{A}_{d,r}$. If the solution of the (reduced) linear program satisfies the maximum principle (up to a small tolerance chosen equal to the grid size d) for the full set of atoms, then the solution is accepted and otherwise the activation parameter is enlarged and a new (presumably larger) active set is computed based on the new approximate multipliers. Since the optimality conditions are necessary and sufficient, the iterative strategy converges. To obtain accurate initial guesses, this strategy is combined with a multilevel scheme in which the discretization parameter d is gradually decreased. In our implementation the (reduced) optimization problems were solved with the MATLAB routine `linprog`. The precise scheme for the solution of one time step reads as follows. As we are interested in x-independent problems, we drop the explicit dependence of the Hamiltonian on $x\in\Omega$.

Algorithm ($A^{\rm hom}_{\rm active\ set}$).
Input: Parameters $0<d_{\rm final}\le r$, number of refinement levels $J\ge 0$ such that $2^J d_{\rm final}\le r$, time step t_j, vector $\lambda_{j-1}\in\mathbb{R}^{M+1}$.

(1) Set $d=2^J d_{\rm final}$, $\mu_0=0$, $\mu_1=0$, $\mu_2=0$, and $\varepsilon_{mp}=d/2$.

(2) Define

$$\mathcal{A}_{d,r} = \left\{ A\in d\mathbb{Z}^{m\times n}\cap B^\infty_r(0) : \mathcal{H}_\mu(t_j,A) \ge \sum_{B\in\mathcal{A}_{d,r}} \theta_B \mathcal{H}_\mu(t_j,B) - \varepsilon_{mp} \right\},$$

where $\mathcal{H}_\mu(t_j, A) = \mathbb{T}(A) \cdot \mu_0(t_j) - W(A) + c_\mathcal{D}(\mu_1(t_j) - \mu_2(t_j)) \cdot \mathcal{L}^1(A)$.

(3) Add further elements $A \in \mathbb{R}^{m\times n}$ to ensure feasibility and solve the linear program described in section 4.4 with the set $\mathcal{A}_{d,r}$. This provides updates of the multipliers μ_0, μ_1, μ_2.

(4) If there exists $A \in d\mathbb{Z}^{m\times n} \cap B_r^\infty(0)$ with

$$\mathcal{H}_\mu(t_j, A) > \sum_{B\in\mathcal{A}_{d,r}} \theta_B \mathcal{H}_\mu(t_j, B),$$

then set $\varepsilon_{mp} = 2\varepsilon_{mp}$ and go to (2).

(5) If $d > d_{\text{final}}$, set $d = d/2$ and $\varepsilon_{mp} = d/2$ and go to (2).

We consider the following specification of the model problem.

Example 4.1. Let $m = n = 2$, $\mathfrak{T} = 1$, $\lambda_0 = 0$,

$$W(F) = \min\{|F^TF - F_1^TF_1|^2/2, |F^TF - F_2^TF_2|^2/2\}$$

for $F_1 = \operatorname{diag}(\delta, 1/\delta)$, $F_2 = \operatorname{diag}(1/\delta, \delta)$, and $F_D(t) = (1-t)F_1 + tF_2$. The function $\mathcal{L}^1$ is for given $\epsilon > 0$ chosen as

$$\mathcal{L}^1(F) = f_C([F^TF]_{11})$$

with

$$f_C(z) = \begin{cases} 0, & z \geq \delta^2 - \epsilon, \\ (z - (\delta^2 - \varepsilon))/(1/\delta^2 + 2\epsilon - \delta^2), & 1/\delta^2 + \epsilon \leq z \leq \delta^2 - \epsilon, \\ 1, & z \leq 1/\delta^2 + \epsilon. \end{cases}$$

The value of ϵ relates to the elastic region of the material. Indeed, starting the evolution with the Young measure supported in one of the energy wells of W, the parameter ϵ determines "how far" the support can move without any dissipation. For the experiments we choose $\epsilon = 1/20$ and $\delta = \sqrt{5/4}$.

We employed $d = 1/20$, $\tau = 1/40$, $r = 2$, and

$$c_\mathcal{D} = 1, 1/10, 1/100$$

in our experiments. The multilevel strategy always started with the coarse grid defined by $d = 1$. In Figure 1 we displayed for the time steps $t_j = j/40$ for $j = 0, 10, 20, 30, 40$ (from left to right) and the dissipation coefficients $c_\mathcal{D} = 1, 1/10$, and $1/100$ (from top to bottom) the deformed body $F_{\mathrm{D}}(t)\Omega$ for $\Omega = (0,1)^2$ and, indicated by the gray shading, the volume fraction $\lambda(t_j)$. We see that the specimen does not transform for the large dissipation constant but does for the other values of $c_\mathcal{D}$. This is what we observe in physical experiments because large dissipation, i.e., large $c_\mathcal{D}$, makes the transformation more difficult, or not possible at all.

In Figure 2 we plotted the functions

$$\begin{aligned} &t_j \mapsto W(F_{\mathrm{D}}(t_j)), \\ &t_j \mapsto I(t_j) = \sum_{A\in d\mathbb{Z}^{2\times 2}} \theta_A^j W(A) + c_\mathcal{D} a, \\ &t_j \mapsto \mathcal{D}(\delta q_j) = c_\mathcal{D} a_j = c_\mathcal{D}|\lambda_j - \lambda_{j-1}|, \qquad \lambda_j = \sum_{A\in\mathcal{A}_{d,r}} \theta_{A,j}\mathcal{L}^1(A) \end{aligned}$$

for $c_\mathcal{D} = 1$, $1/10$, and $1/100$.

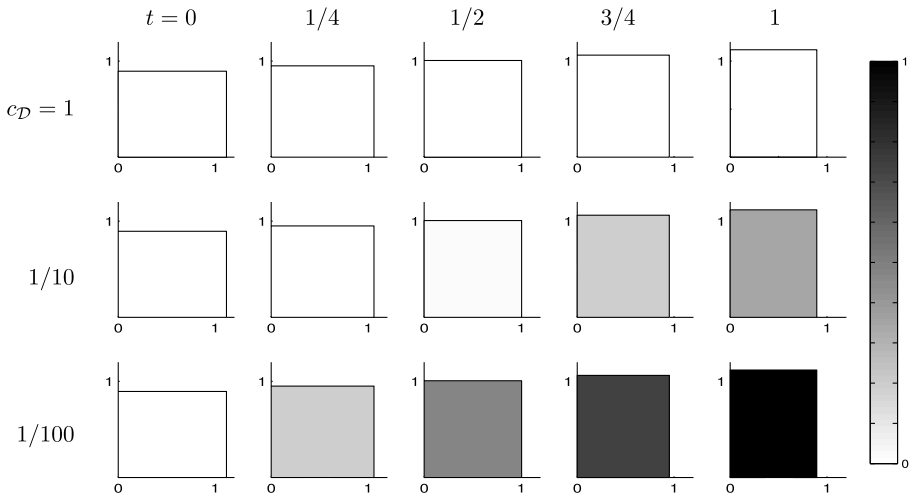

FIG. 1. *Transformation of a specimen for different dissipation strengths. The upper row shows the deformed body colored by the volume fraction* $\lambda^1(t_j)$ *for* $t_j = j/40$, $j = 0, 10, 20, 30, 40$, *and* $c_{\mathcal{D}} = 1$. *The second and third rows show the related quantities for* $c_{\mathcal{D}} = 1/10$ *and* $c_{\mathcal{D}} = 1/100$, *respectively.*

For $c_{\mathcal{D}} = 1/100$ we plotted in Figure 3 the relative number of activated atoms on the finest level corresponding to $d = 1/10$ for each time step. Notice that we have here for $d = 1/10$ and $r = 2$

$$\operatorname{card}(d\mathbb{Z}^{2\times 2} \cap B_r^\infty(0)) = 2825761.$$

We thus obtain an average reduction to about 0.1% of the theoretical number of atoms.

4.6. Scalar experiment with spatial dependence. We consider the following specification of the model problem.

Example 4.2. Let $\Omega = (0,1)^2$, $\mathfrak{T} = 1$, $\Gamma_{\mathrm{D}} = \partial\Omega$, $u_D(x,t) = \sin(2\pi t)x_1$, $\lambda_0(x) = 1/2$,

$$W(F) = \min\{|F - F_1|^2/2 + c_1, |F - F_2|^2/2 + c_2\}$$

for $F_1 = (1,0)^T$, $F_2 = (-1,0)^T$, $c_1 = 1/5$, and $c_2 = 0$.

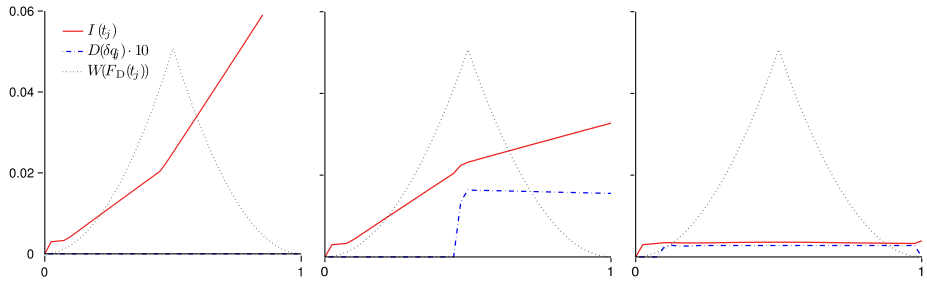

FIG. 2. *Total energy, dissipation, and elastic energy for the macroscopic deformation for* $c_{\mathcal{D}} = 1, 1/10,$ *and* $1/100$ *(from left to right).*

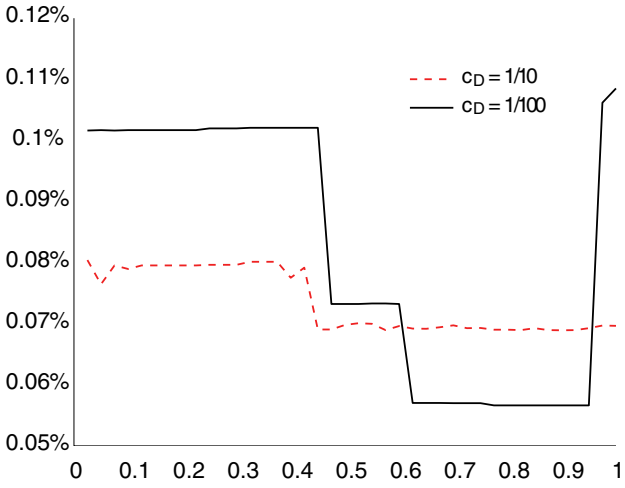

FIG. 3. *Relative number of activated atoms during the evolution for different dissipation constants.*

For a triangulation of Ω consisting of 32 triangles, the set $\mathcal{A}_{d,r}$ defined through $d = 2^{-4}$ and $r = 2$, the time-step size $\tau = 1/80$, and the choices of constants

$$c_{\mathcal{D}} = 1 \quad \text{and} \quad c_{\mathcal{D}} = 1/10,$$

Figures 4 and 5, respectively, show snapshots of the evolution for $t = j/20$ with $j = 4$, 20, 40, 60, 80.

Figure 6 displays the energy and the dissipation contribution, i.e., the quantities

$$\begin{aligned} E(t_j) &= \int_\Omega \int_{\mathbb{R}^2} W(A)\nu_{j,d,h}(\mathrm{d}A) + D(t_j) \\ &= \sum_{T\in\mathcal{T}} |T| \sum_{A\in\mathcal{A}_{d,r}} \theta^j_{T,A} W(A) + D(t_j), \qquad D(t_j) = c_{\mathcal{D}} \int_\Omega a_{h,j}\mathrm{d}x = c_{\mathcal{D}} \sum_{T\in\mathcal{T}} |T| a_T(t_j) \end{aligned}$$

as functions of $t \in [0,1]$ for $c_{\mathcal{D}} = 1$ and $c_{\mathcal{D}} = 1/10$.

Figure 6 illustrates hysteresis effects that occur in the evolution defined by the rate-independent process. We displayed the spatial averages of the (spatially constant) quantities $\partial_1 y_{h,j}$ and $\sigma_{h,j}\varrho$ on Γ_D (ϱ is the outer unit normal to the boundary), where

$$\sigma_{h,j} = \sum_{A\in\mathcal{A}_{d,r}} \theta_{A,j} DW(A).$$

The validity of a fully discrete analogue of the semidiscrete two-sided energy estimate (28)

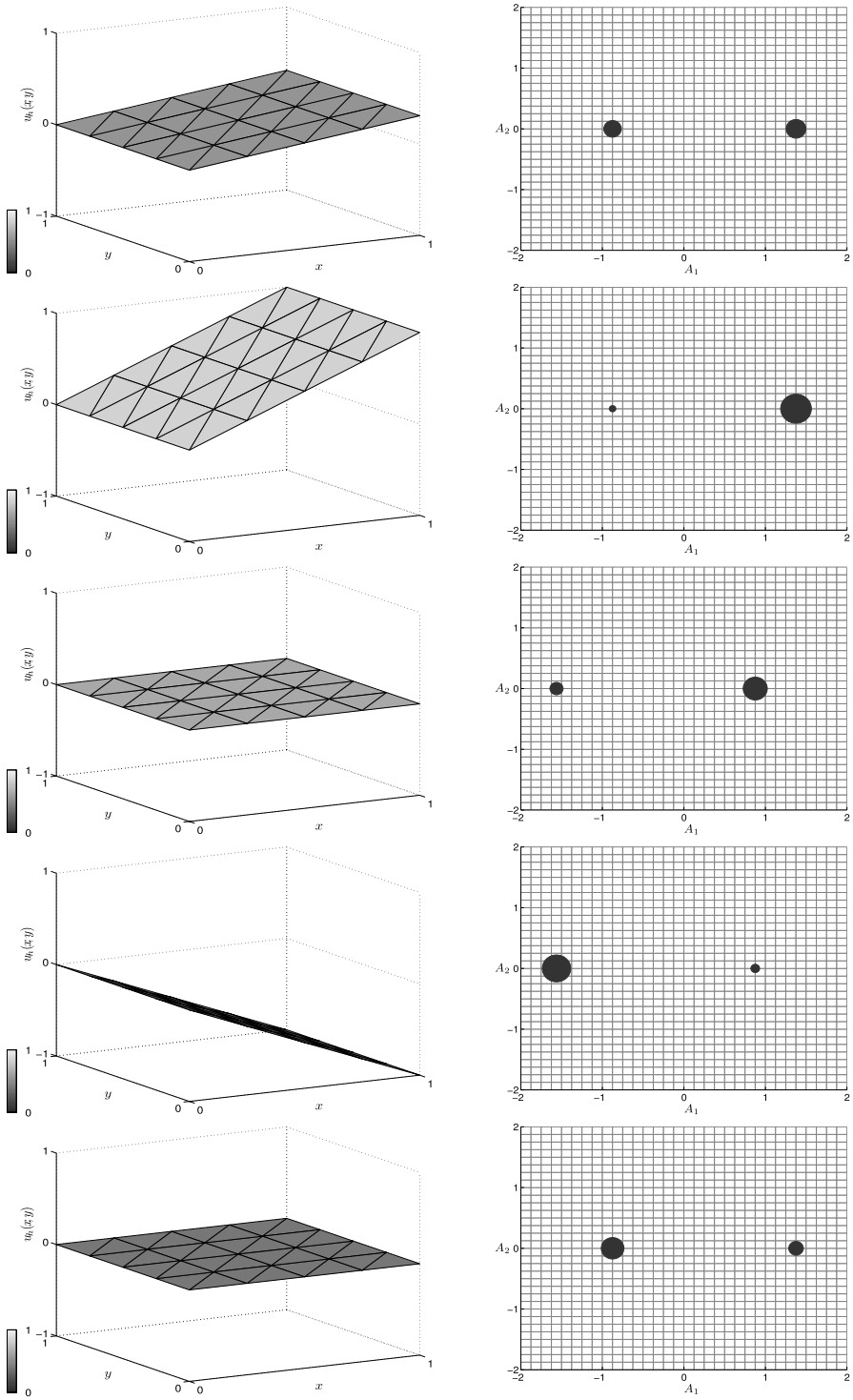

FIG. 4. *Scalar displacement $u_{h,j}$ (left) and discrete Young measure $\nu_{j,h}$ (right) for $j = 4, 20, 40, 60, 80$ in Example 4.3 with $c_{\mathcal{D}} = 1$. The displacement is colored by the quantity λ_h, and the sizes of the dots in the grid in the right plots are proportional to the volume fraction associated with a grid point. The grid indicates every second atom.*

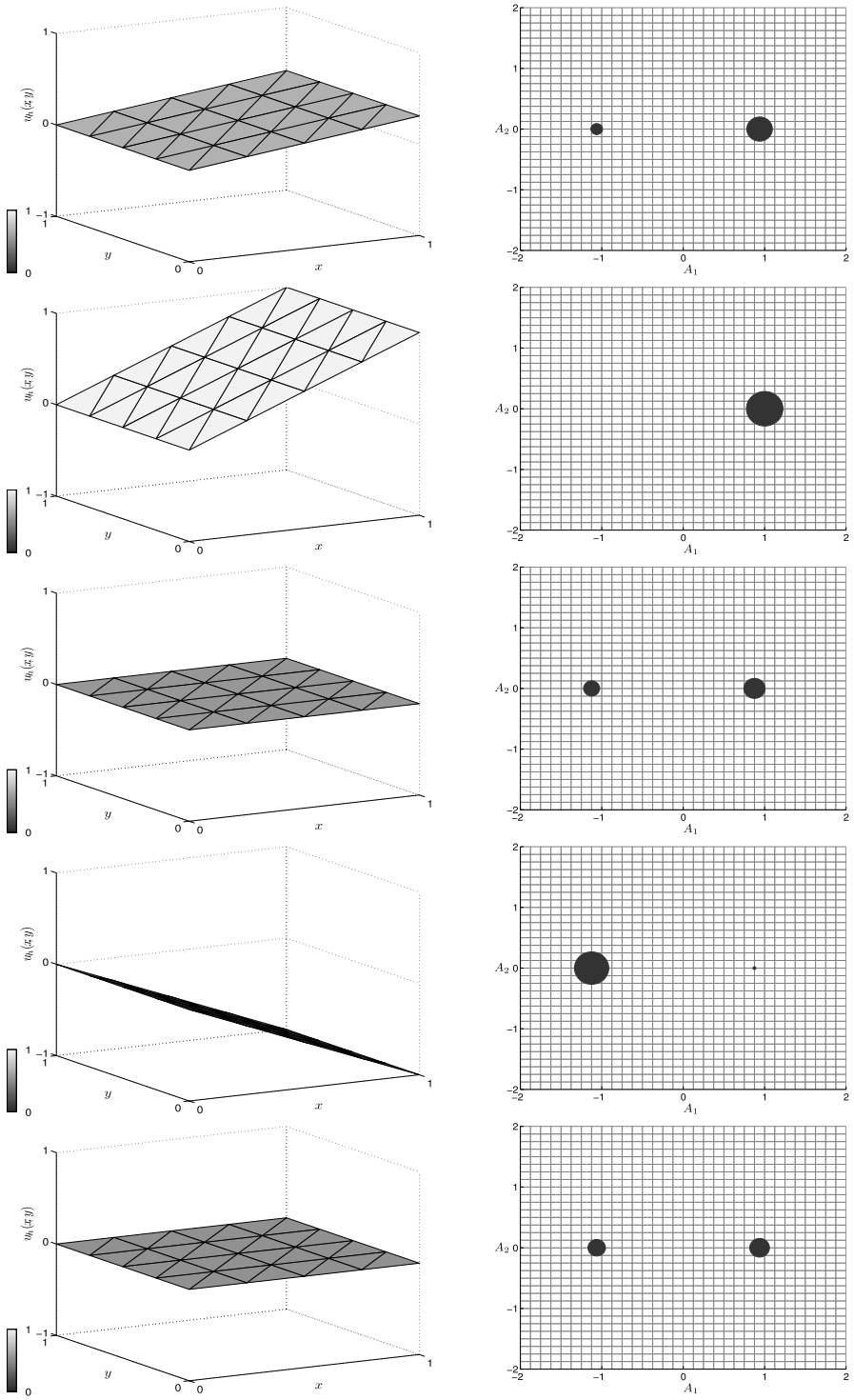

FIG. 5. *Scalar displacement $u_{j,h}$ (left) and discrete Young measure $\nu_{j,h}$ (right) for $j = 4, 20, 40, 60, 80$ in Example 4.3 with $c_{\mathcal{D}} = 1/10$. The displacement is colored by the quantity λ_h, and the sizes of the dots in the grid in the right plots are proportional to the volume fraction associated with a grid point. The grid indicates every second atom.*

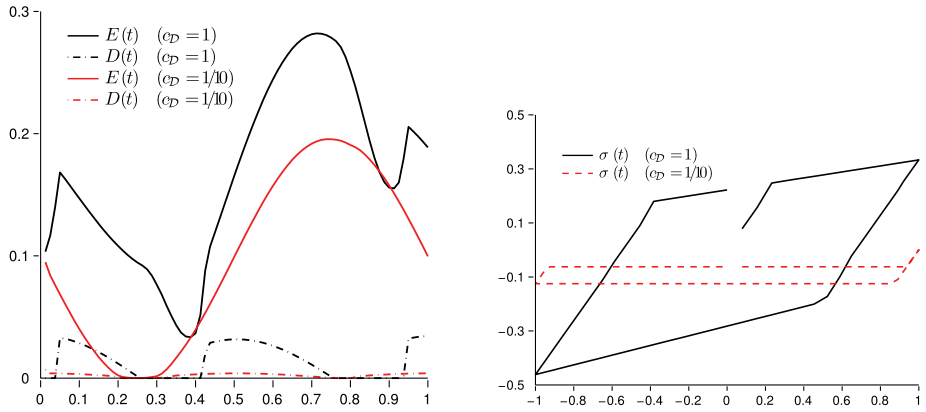

FIG. 6. *Total energy and dissipation as functions of* $t \in [0,1]$ *in Example* 4.3 *with* $c_{\mathcal{D}} = 1$ *and* $c_{\mathcal{D}} = 1/10$ *(left). Stress versus strain for* $t \in [0,1]$ *in Example* 4.3 *with* $c_{\mathcal{D}} = 1$ *and* $c_{\mathcal{D}} = 1/10$ *(right).*

$$
\begin{aligned}
m(t_j) &:= \int_{\Gamma_D} \sigma_{h,j} \cdot \varrho(u_{D,j} - u_{D,j-1}) \mathrm{d}s \\
\leq \Xi(t_j) &:= \int_\Omega \int_{\mathbb{R}^2} W(s) \nu_{j,x}(\mathrm{d}s)\mathrm{d}x + c_{\mathcal{D}} \int_\Omega |\lambda_j^1 - \lambda_{j-1}^1| \mathrm{d}x - \int_\Omega \int_{\mathbb{R}^2} W(s) \nu_{j-1,x}(\mathrm{d}s)\mathrm{d}x \\
&\leq \int_{\Gamma_D} \sigma_{h,j-1} \cdot \varrho(u_{D,j} - u_{D,j-1}) \mathrm{d}s =: M(t_j)
\end{aligned}
$$

is graphically analyzed in Figure 7. We observe that $m(t_j)$ and $M(t_j)$ are close together but that $m(t_j)$ is not below $M(t_j)$ and $\Xi(t_j)$ is not in between. Therefore, a fully discrete two-sided energy estimate can be expected to hold only with error terms related to the spatial discretization.

4.6.1. Scalar, inhomogeneous example. We consider the following specification of the model problem that leads to an inhomogeneous solution.

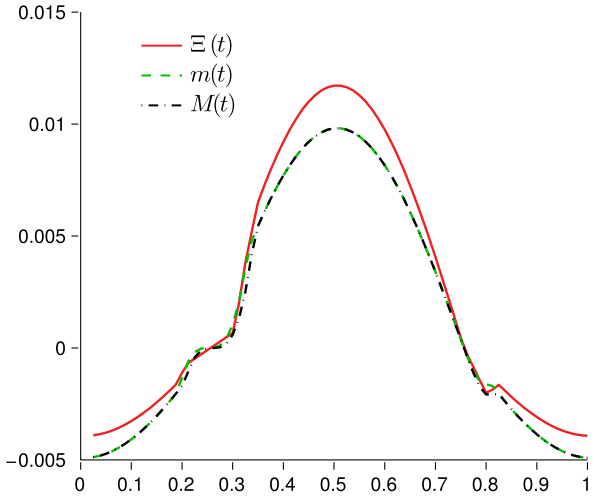

FIG. 7. *Experimental two-sided energy estimate in Example* 4.3 *for* $c_{\mathcal{D}} = 1$.

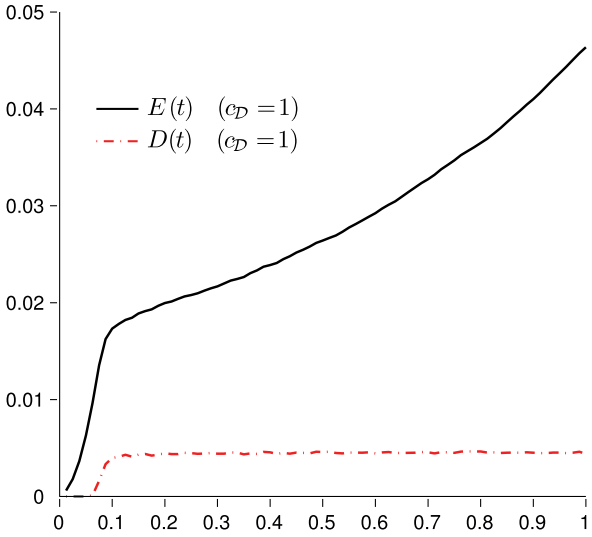

FIG. 8. *Total energy and dissipation as functions of* $t \in [0,1]$ *in Example* 4.3.

Example 4.3. Let $n=2$, $m=1$, $\Omega=(0,1)^2$, $\mathfrak{T}=1$, $\Gamma_{\mathrm{D}}=\partial\Omega$, $\lambda_0(x)=1/2$, and the displacement

$$u_D(t,x)=\begin{cases} -3(z-z_b)^5/128-(z-z_b)^3/3 & \text{for } z\le z_b,\\ (z-z_b)^3/24+(z-z_b) & \text{for } z\ge z_b \end{cases}$$

for $z=x\cdot F_0$, $z_b=1/2$, $F_0=(\cos\phi,\sin\phi)$, $\phi=\pi/6$, and

$$W(F)=|F-F_0|^2|F+F_0|^2.$$

For a triangulation of Ω consisting of 128 triangles, the set $\mathcal{A}_{d,r}$ defined through $d=2^{-6}$ and $r=3/2$, the time-step size $\tau=1/20$, and the choice

$$c_{\mathcal{D}}=1,$$

Figure 9 shows snapshots of the evolution for $t=j/20$ with $j=1$, 5, 10, 20.

Figure 8 displays the energy and the dissipation contribution, i.e., the quantities

$$E(t_j)=\sum_{T\in\mathcal{T}}|T|\sum_{A\in\mathcal{A}_{d,r}}\theta^j_{T,A}\,W(A)+D(t_j),\qquad D(t_j)=c_{\mathcal{D}}\sum_{T\in\mathcal{T}}|T|a_T(t_j)$$

as functions of $t\in[0,1]$.

We found out that the proposed method works effectively. In fact, if we consider Lemma 1.1 (ii) not only polyconvex functions but also for a finite number of quasiconvex ones, the related maximum principle provides an improved lower bound of the quasiconvex envelope. A further step might be to investigate whether the polyconvex Young measure obtained by our algorithm is a laminate. These aspects are left for future research.

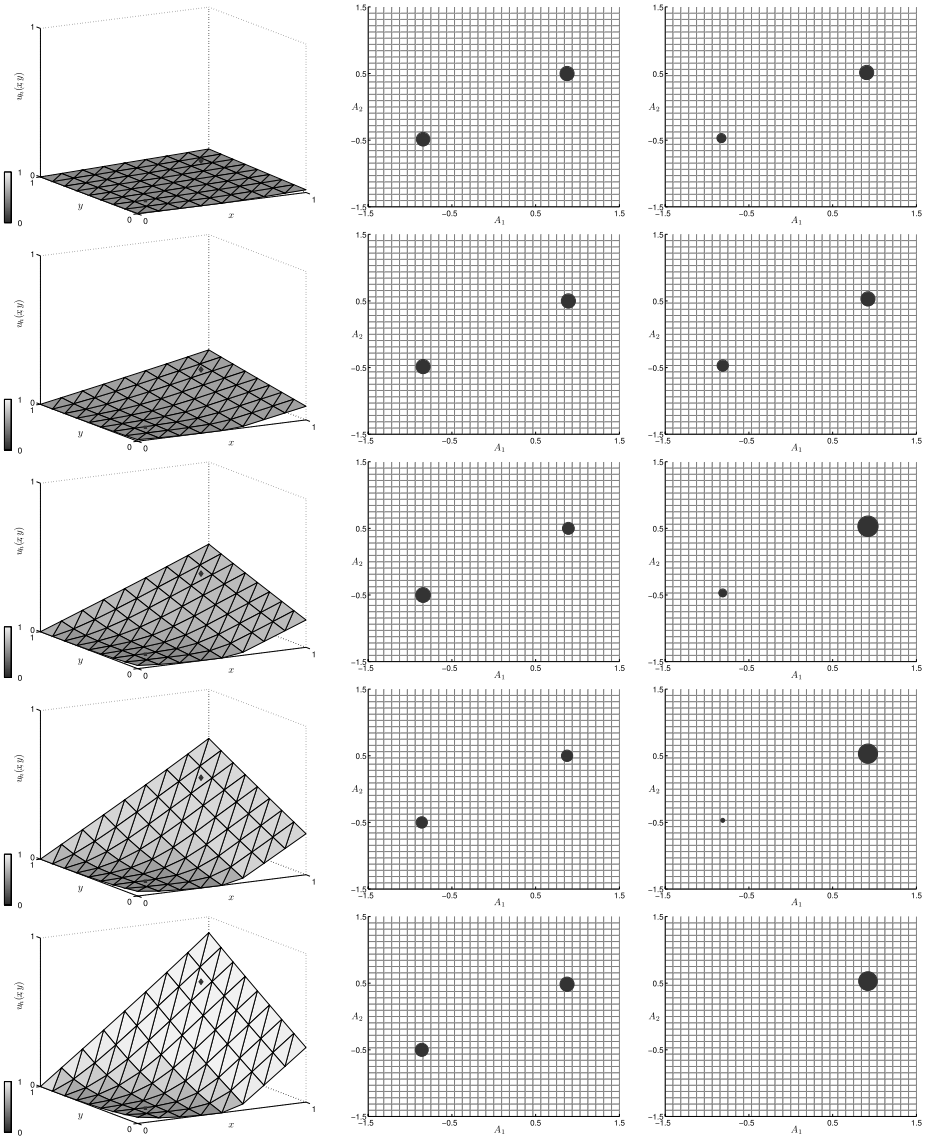

FIG. 9. *Scalar displacement $u_{j,h}$ (left) and discrete Young measure $\nu_{j,h}$ related to the elements indicated by the filled square (middle) and the filled diamond (right) for $j = 4, 20, 40, 60, 80$ in Example 4.3 with $c_{\mathcal{D}} = 1$. The displacement is colored by the quantity λ_h; the sizes of the dots represent the volume fractions of every sixth atom.*

Acknowledgments. We thank the anonymous referee for extremely careful reading of the manuscript and for the valuable comments and suggestions that significantly improved our paper.

REFERENCES

[1] M. ARNDT, *Upscaling from Atomistic Models to Higher Order Gradient Continuum Models for Crystalline Solids*. Ph.D. thesis, Institut für Numerische Simulation, Universität Bonn, Bonn, Germany, 2004.

[2] S. Aubri, M. Fago, and M. Ortiz, *A constrained sequential-lamination algorithm for the simulation of sub-grid microstructure in martensitic materials*, Comput. Methods Appl. Mech. Engrg., 192 (2003), pp. 2823–2843.
[3] J. M. Ball, *Convexity conditions and existence theorems in nonlinear elasticity*, Arch. Ration. Mech. Anal., 63 (1977), pp. 337–403.
[4] J. M. Ball, *A version of the fundamental theorem for Young measures*, in PDEs and Continuum Models of Phase Transition, Lecture Notes in Physics 344, M. Rascle, D. Serre, and M. Slemrod, eds., Springer, Berlin, 1989, pp. 207–215.
[5] J. M. Ball and R. D. James, *Fine phase mixtures as minimizers of energy*, Arch. Ration. Mech. Anal., 100 (1988), pp. 13–52.
[6] S. Bartels, *Linear convergence in the approximation of rank-one convex envelopes*, M2AN Math. Model. Numer. Anal., 38 (2004), pp. 811–820.
[7] S. Bartels, *Reliable and efficient approximation of polyconvex envelopes*, SIAM J. Numer. Anal., 43 (2005), pp. 363–385.
[8] C. Carstensen and T. Roubíček, *Numerical approximation of Young measures in nonconvex variational problems*, Numer. Math., 84 (2000), pp. 395–415.
[9] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Springer, Berlin, 1989.
[10] G. Dolzmann and N. Walkington, *Estimates for numerical approximations of rank one convex envelopes*, Numer. Math., 85 (2000), pp. 647–663.
[11] I. Fonseca, S. Müller, and P. Pedregal, *Analysis of concentration and oscillation effects generated by gradients*, SIAM J. Math. Anal., 29 (1998), pp. 736–756.
[12] G. Francfort and A. Mielke, *Existence results for a class of rate-independent material models with nonconvex elastic energies*, J. Reine Angew. Math., 595 (2006), pp. 55–91.
[13] S. Govindjee, A. Mielke, G. J. Hall, and C. Miehe, *The free-energy of mixing for n-variant martensitic phase transformations using quasi-convex analysis*, J. Mech. Phys. Solids, 50 (2002), pp. 1897–1922.
[14] S. Govindjee, A. Mielke, G. J. Hall, and C. Miehe, *The free-energy of mixing for n-variant martensitic phase transformations using quasi-convex analysis, erratum and correct reprinting*, J. Mech. Phys. Solids, 51 (2003), pp. 763 & I–XXVI.
[15] K. Hackl and U. Hoppe, *On the calculation of microstructure for inelastic materials using relaxed energies*, in IUTAM Symposium on Compututational Mechanics of Solid Materials at Large Strains, Ch. Miehe, ed., Kluwer, Dordrecht, 2003, pp. 77–86.
[16] Y. Huo and I. Müller, *Nonequilibrium thermodynamics of pseudoelasticity*, Contin. Mech. Thermodyn., 5 (1993), pp. 163–204.
[17] D. Kinderlehrer and P. Pedregal, *Gradient Young measures generated by sequences in Sobolev spaces*, J. Geom. Anal., 4 (1994), pp. 59–90.
[18] M. Kružík, *Maximum principle based algorithm for hysteresis in micromagnetics*, Adv. Math. Sci. Appl., 13 (2003), pp. 461–485.
[19] M. Kružík, A. Mielke, and T. Roubíček, *Modelling of microstructure and its evolution in shape memory-alloy single-crystals, in particular in CuAlNi*, Meccanica, 40 (2005), pp. 389–418.
[20] M. Kružík and A. Prohl, *Young measure approximation in micromagnetics*, Numer. Math., 90 (2001), pp. 291–307.
[21] M. Kružík and T. Roubíček, *Optimization problems with concentration and oscillation effects: Relaxation theory and numerical approximation*, Numer. Funct. Anal. Optim., 20 (1999), pp. 511–530.
[22] A. Mielke and T. Roubíček, *A rate-independent model for inelastic behavior of shape-memory alloys*, Multiscale Model. Simul., 1 (2003), pp. 571–597.
[23] A. Mielke and F. Theil, *A mathematical model for rate-independent phase transformations with hysteresis*, in Models of Continuum Mechanics in Analysis and Engineering, H.-D. Alber, R. Balean, and R. Farwig, eds., Shaker Verlag, Aachen, 1999, pp. 117–129.
[24] A. Mielke, F. Theil, and V. I. Levitas, *A variational formulation of rate-independent phase transformations using an extremum principle*, Arch. Ration. Mech. Anal., 162 (2002), pp. 137–177.
[25] C. B. Morrey, Jr., *Quasi-convexity and the lower semicontinuity of multiple integrals*, Pacific J. Math., 2 (1952), pp. 25–53.
[26] P. Pedregal, *On the numerical analysis of non-convex variational problems*, Numer. Math., 74 (1996), pp. 325–336.
[27] P. Pedregal, *Parametrized Measures and Variational Principles*, Birkhäuser, Basel, 1997.
[28] P. Pedregal, *Rank-One Convexity Implies Quasiconvexity for Two-Component Maps*, manuscript.
[29] T. Roubíček, *Relaxation of vectorial variational problems*, Math. Bohem., 120 (1995), pp. 411–430.
[30] T. Roubíček, *Effective characterization of generalized Young measures generated by gradients*, Boll. Unione Mat. Ital., 9-B (1995), pp. 755–779.

[31] T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus.*, W. de Gruyter, Berlin, 1997.
[32] T. Roubíček, *Models of microstructure evolution in shape memory materials.* in Nonlinear Homogenization and Its Applications to Composites, Polycrystals and Smart Materials, NATO Sci. Ser. II 170, P. Ponte Castaneda, J. J. Telega, and B. Gambin, eds., Kluwer, Dordrecht, 2004, pp. 269–304.
[33] H. von Stackelberg, *The Theory of the Market Economy*, Oxford University Press, Oxford, 1952.
[34] V. Šverák, *Rank-one convexity does not imply quasiconvexity*, Proc. Roy. Soc. Edinburgh Sect. A, 120 (1992), pp. 293–300.
[35] L. C. Young, *Generalized curves and existence of an attained absolute minimum in the calculus of variations*, Comptes Rendus de la Société et des Lettres de Varsovie, Classe III, 30 (1937), pp. 212–234.
[36] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.