A thesis submitted for the degree

of Doctor of Philosophy in Mathematics

Quasiconvexity and PDE constraints:

analysis and applications

UNIVERSITY OF SUSSEX

By **Andreas Vikelis**

Under the supervision of Konstantinos Koumatos

March 2022

Abstract

In this thesis we focus on questions of stability, existence and uniqueness for PDE constrained problems in both dynamics and statics under appropriate convexity conditions. In the first part, we establish a Gårding-type inequality for quantities associated to $(\mathcal{A}, 0)$-quasiconvex functions, where $\mathcal{A}$ is a constant-coefficient, linear differential operator with constant rank. In dynamics, we initially apply a simplified version of our derived inequality to prove weak-strong uniqueness results for conservation laws possessing involutions i.e. a differential constraint $\mathcal{A}$ propagated by the initial data, provided that the system is endowed with an $\mathcal{A}$-quasiconvex entropy. In addition to this, combining our Gårding-type inequality with the relative entropy method, we establish a weak-strong uniqueness result for the hyperbolic system of adiabatic thermoelasticity under quasiconvexity conditions. In particular, we show that classical solutions of that system are unique within a suitable class of dissipative measure-valued solutions, provided that the internal energy is stronly $(\mathrm{curl}, 0)$-quasiconvex. In statics, we investigate the so-called Weierstrass problem of finding necessary and sufficient conditions for local minimisers. More precisely, we prove an $\mathcal{A}$-quasiconvexity based sufficiency theorem for local minimisers for general problems constrained by an operator $\mathcal{A}$. An additional contribution of our result is that we infer uniqueness of these local minimisers and quantify the difference in energy between them and arbitrary comperitors. In the second part, in the context of image processing, we study a class of PDE constrained variational problems whose regularising terms depend on the differential operator. We prove the lower semicontinuity of the functionals in question and existence of minimisers for the corresponding variational problems. Then, we embed the latter into a bilevel scheme in order to automatically compute the space-dependent regularisation parameters, and we establish existence of optima for the scheme. We finally substantiate its feasibility by numerical examples in image denoising.

In the memory of my grandparents

Vania Koukaki

and

Andreas Pittas

Acknowledgements

At this point I want to express my gratitude to all those who supported me during this difficult and long journey, from the beginning of my studies until today. Admittedly, completing a PhD in a pandemic is not the easiest thing to do and certainly requires the support and the understanding of the people surrounding you.

Firstly, I would like to thank Konstantinos Koumatos for being besides an excellent supervisor, an amazing person whose support could not have been more substantial during the last three years. From the academic point of view, he gave me very well-structured problems, which combined with his always "to the point" guidance, facilitated the successful completion of this thesis. He introduced me to various aspects of mathematical materials science and gave me the chance to meet a lot of high quality researchers and get involved in a variety of mathematical projects, while in parallel he left me my own space to develop other parts of my mathematical interests independently. On the other hand, on a personal level, he stood by me at every opportunity, showed great understanding in periods when I was psychologically quite unstable and was always present when I needed him.

Many thanks to my great collaborators Myrto Galanopoulou, Valerio Pagliari, Konstantinos Papafitsoros and Bogdan Raiţă for working with me and showing me how enjoyable mathematics can be especially when they are done collectively. Special thanks to Elisa Davoli, not only for her work in image reconstruction which constitutes the main source of inspiration for Chapter 5 of this thesis, but also because she introduced me to the Viennese mathematical community.

Specifically, Elisa brought me in contact with Ulisse Stefanelli, whom I would like to thank deeply for offering me the opportunity to be part of his group.

I also want to highlight the contribution of Kostas Papafitsoros and Bogdan Raiţă, not only to this particular thesis, but also to my development as a mathematician. The former was always there to guide me even from my bachelor thesis, while his work was what triggered my interest around mathematical image processing. The latter, besides a good friend, constitutes without exaggeration my second mentor. He was always there to discuss even my most naive questions and he helped me to better understand what he typically calls “linear operators business".

I am also thankful to Filippo Cagnetti and Stefano Spirito for their support with their very important recommendation letters, and my internal and external examiners Ali Taheri and Cleopatra Christoforou for spending time reading my thesis and making my Viva pleasant and constructive.

Moreover, I would like to thank Nikiforos Mimikos, on the one hand for his special and deep friendship for the last ten years, and on the other hand for always believing in me and giving me the cold shower when I needed it. He was the centrifugal force for my academic “career” so far, while I consider him one of the most special and talented people I know.

Above all, I would like to thank my mother Evelina, my father Giorgos , my brother Dimitris and my partner Magda for all their love and support all of these years. Without them I would not have achieved all these, but most importantly I would not be the person I am today.

Contents

List of Figures

Chapter 1

Introduction

In this thesis, using tools from the calculus of variations, we study mathematical problems where their solutions satisfy linear partial differential constaints. The problems we are interested in arise from physics, and more precisely continuum mechanics and image processing.

1.1 Continuum mechanics: dynamics and statics

In the setting of continuum mechanics and the theory of electromagnetism, often problems are constrained by linear partial differential equations (PDEs), that is their solutions are constrained to lie in the kernel of a certain differential operator $\mathcal{A}$. The prototypical example arises in elasticity. In elastostatics, one is concerned with the minimisation of the functional

$$\int_\Omega W(\nabla y) \tag{1.1}$$

and thus solutions $U = \nabla y$ are constrained by the operator $\mathcal{A} = \text{curl}$. Similarly, in dynamics, the equations of elasticity can be written in the form of the first-order system of conservation laws

$$\begin{aligned} \partial_t v - \text{div} DW(F) &= 0, \\ \partial_t F - \nabla v &= 0, \\ \operatorname{curl} F &= 0. \end{aligned}$$

Note that the last equation constrains solutions F to be gradients and it is satisfied as long as the initial data are curl-free. More generally, one may consider problems constrained by other differential operators $\mathcal{A}$, leading to the study of minimisation problems of the form

$$\mathcal{W}(U) = \int W(U), \quad \mathcal{A}U = 0 \tag{1.2}$$

and systems of conservation laws

$$\partial_t U + \operatorname{div} f(U) = 0 \tag{1.3}$$

with involutions $\mathcal{A}$, i.e. with the property that $\mathcal{A}U(t,\cdot) = 0$ whenever $\mathcal{A}U(0,\cdot) = 0$. PDE constrained problems of the above form, and others, have been studied extensively. Indeed, the theory of compensated compactness developed by Murat and Tartar originated within this $\mathcal{A}$-free context [87, 102, 103].

In particular, they understood that quadratic forms that are convex in certain directions associated to $\mathcal{A}$ are lower semicontinuous along $\mathcal{A}$-free, weakly converging sequences. This set of directions, $\Lambda_{\mathcal{A}}$, is referred to as the wave cone of $\mathcal{A}$, see Section 2.1, and contains the amplitudes along which ellipticity is lost. For example, for vectorial problems and $\mathcal{A} = \operatorname{curl}$, the wave cone consists of rank-one matrices and rank-one convexity becomes the relevant convexity condition. Note that rank-one convexity for quadratic forms is equivalent to the less transparent notion of quasiconvexity which is itself equivalent to the weak lower semicontinuity of (1.1), see [33].

Indeed, for problems of the form (1.2), an appropriate extension of quasiconvexity, called $\mathcal{A}$-quasiconvexity, was introduced by Dacorogna [32] and shown to be equivalent to the weak lower semicontinuity of (1.2), in [32, 52]. Following the work of Murat and Tartar, a developing body of literature has emerged on PDE constrained problems, including recent results on appropriate modifications of BV spaces, lower semicontinuity, Young measures, Sobolev-type inequalities etc [6, 7, 15, 46, 58, 80, 92, 93, 98].

In the context of dynamics, part of the thesis focuses on the system of adiabatic thermoelasticity in Lagrangian coordinates given by

$$\begin{aligned} \partial_t F_{i\alpha} - \partial_\alpha v_i &= 0 \\ \partial_t v_i - \partial_\alpha \Sigma_{i\alpha} &= 0 \\ \partial_t \left(\frac{1}{2}|v|^2 + e \right) - \partial_\alpha (\Sigma_{i\alpha} v_i) &= r, \end{aligned} \tag{1.4}$$

that describes the evolution of a thermomechanical process $\big(y(t,x), \eta(t,x)\big) \in \mathbb{R}^3 \times \mathbb{R}$ where the time variable $t \in \mathbb{R}^+$ and the spatial variable $x \in \mathbb{R}^3$. This is a first-order system and a solution to (1.4), consists of the deformation gradient $F = \nabla y \in \mathbb{M}^{3\times 3}$, the velocity $v = \partial_t y \in \mathbb{R}^3$ and the specific entropy η. The first equation is a compatibility relation between the partial derivatives of the motion, the second describes the balance of linear momentum, while the third equation stands for the balance of energy. One must include to (1.4) the constraint

$$\partial_\alpha F_{i\beta} = \partial_\beta F_{i\alpha}, \qquad i, \alpha, \beta = 1, 2, 3\,, \tag{1.5}$$

which guarantees that F is a deformation gradient associated to the motion $y(t,x)$. We note that relation (1.5) is an involution, namely, it is propagated from the initial data to the solution via $(1.4)_1$.

The remaining variables in (1.4) are the Piola-Kirchhoff stress $\Sigma_{i\alpha}$, the internal energy e, and the radiative heat supply r. Here, the referential heat flux $Q_\alpha = 0$ as this theory describes adiabatic processes, and it does not appear in the equations (1.4). The balance of entropy holds identically as an equality for strong solutions, that is

$$\partial_t \eta = \frac{r}{\theta(F,\eta)} \tag{1.6}$$

and it can be derived from system (1.4). By contrast, for weak solutions, (1.6) is replaced by the Clausius-Duhem inequality [28, 105, 37], according to the second law of thermodynamics, and it serves as a criterion of admissibility for thermodynamic processes that satisfy the balance laws of mass, momentum and energy. The system is closed through constitutive relations which, for smooth processes, are consistent with the Clausius-Duhem inequality and describe the material response. For thermoelastic materials under adiabatic conditions, the constitutive theory is determined from the thermodynamic potential of the internal energy depending solely on the deformation gradient F and the entropy η, via the relations

$$e = e(F,\eta), \quad \Sigma = \frac{\partial e}{\partial F} =: e_F, \quad \theta = \frac{\partial e}{\partial \eta} =: e_\eta, \tag{1.7}$$

for the stress Σ and the temperature θ. We refer the reader to [28, 105] for a detailed derivation of adiabatic thermoelasticity and its relation to other constitutive theories.

System (1.4) belongs to a general class of hyperbolic problems that are symmetrisable in the sense of Friedrichs and Lax [54], under appropriate hypotheses. It turns out that symmetrisability is guaranteed by the positivity of the matrix

$$\frac{1}{e_\eta}\begin{pmatrix} e_{FF} & 0 & e_{F\eta} \\ 0 & 1 & 0 \\ e_{F\eta} & 0 & e_{\eta\eta} \end{pmatrix} \tag{1.8}$$

which in turn amounts to $e(F,\eta)$ being strongly convex and $\theta(F,\eta) = \frac{\partial e(F,\eta)}{\partial \eta} > 0$. In subsection 4.1.2 we discuss the connection of thermoelasticity to the general theory of conservation laws for symmetrisable systems.

Convexity of $e(F,\eta)$ suffices to apply the standard theory of conservation laws to (1.4), however, the condition of convexity is too restrictive to encompass a large class of materials. A broader notion of convexity is polyconvexity, that is $e(F,\eta) = g(F, \operatorname{cof} F, \det F, \eta)$ for some convex function g. For polyconvex energies stability and weak-strong uniqueness results for system (1.4) have been obtained in [22, 23, 24]. Note that due to the presence of involutions (1.5), the positivity of the matrix appearing in (1.8) is indeed only required on the cone

$$\left\{(a \otimes n, v, \eta) : a,\, n,\, v \in \mathbb{R}^3,\, \eta \in \mathbb{R}\right\}$$

amounting to a notion of rank-one convexity for $e(F, \eta)$.

Nevertheless, as it was proved by Dafermos in [38], weak-strong uniqueness for hyperbolic systems of conservation laws with entropies which are convex on the wave cone, can be established under an extra assumption of small local oscilations on the weak solutions. In particular, Dafermos studied the system of conservation laws (1.3) endowed with involutions where $\mathcal{A} = \sum_\alpha A_\alpha \partial_\alpha$ was assumed to be a first order operator. He showed that if the involutions are complete , see [40] for the respective definition, system (1.3) becomes hyperbolic and he constructed a first order potential operator $\mathcal{B} = \sum_\alpha B_\alpha \partial_\alpha$ such that $U = \mathcal{B}W$ whenever $\mathcal{A}U = 0$. Through this potential $\mathcal{B}$, he extracted a Poincaré type inequality for $\mathcal{A}$-free functions which played a decisive role in the proof of his main tool: a Gårding-type inequality for the quantity

$$\begin{aligned} R(U|\bar{U}) &= R(U) - R(\bar{U}) - DR(\bar{U}) \cdot (U - \bar{U}) \\ &= \int_0^1 (1-t) D^2 R\left(\bar{U} + t(U - \bar{U})\right)\, dt (U - \bar{U}) \cdot (U - \bar{U}), \end{aligned} \tag{1.9}$$

associated to the $\Lambda_{\mathcal{A}}$-convex entropy R. Nevertheless, this Gårding inequality required that the weak solutions, assumed bounded and in the space BV, satisfy an extra assumption of small local oscillations. Then, naturally, it leads to stability and weak-strong uniqueness results for such entropic weak solutions. In [74] and the case of elasticity, it was understood that the crucial Gårding inequality and the subsequent weak-strong uniqueness result can be proved without the assumption of small oscillations, provided the entropy instead satisfies the stronger condition of quasiconvexity. Indeed, that nonlinear Gårding inequalities are connected to strong quasiconvexity and that they can be useful in the setting of the calculus of variations can already be found in [30]. We note that elements of the proof of the Gårding inequality in [74] appear in [30]. More recently, Kristensen and Campos Cordero in [79] have obtained a similar Gårding inequality in the curl-free setting following a different approach.

More generally, Gårding inequalities have been very important, for example, to establish existence, uniqueness and regularity for elliptic problems, see [1, 56, 57, 86, 88, 97]. Crucially, a Gårding-type inequality for the function in (1.9) also appeared in the resolution of the so-called Weierstrass problem in the vectorial calculus of variations, i.e. the problem of finding (quasiconvexity based) sufficient conditions for a map $\bar{y}$ to be a strong (or L^p) local minimiser of (1.1), see Section 4.2. This was indirectly employed in the original proof of [60] and more explicitly in the subsequent proofs in [31, 29] which seek the positivity of

$$\int W(\nabla\bar{y} + \nabla\varphi|\nabla\bar{y}),$$

related to a Gårding-type inequality for the function $W(\cdot|\cdot)$. In this context, this quantity is known as the Weierstrass excess or E-function, see [60, 29] for functionals depending on lower order terms.

1.1.1 Continuum mechanics: our contribution

Motivated by the problems presented above and in particular, in terms of notation, by the equations of adiabatic thermoelasticity, our work establishes a Gårding-type inequality for quantities of the form

$$e(z_1, z_2|\bar{z}_1, \bar{z}_2) := e(z_1, z_2) - e(\bar{z}_1, \bar{z}_2) - e_F(\bar{z}_1, \bar{z}_2)(z_1 - \bar{z}_1) - e_\eta(\bar{z}_1, \bar{z}_2)(z_2 - \bar{z}_2), \quad (1.10)$$

associated to $(\mathcal{A}, 0)$-quasiconvex functions $e : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ satisfying suitable growth and coercivity assumptions, see Section 3.3. We note here that for consistency, we adopt the notation of adiabatic thermoelasticity where by e_F and e_η we refer to the derivatives with respect to the first and the second variable respectively. As it has been probably already understood from the above discussion, these kind of quantities play an important role: on the one hand, they appear as the relative entropy in the theory of conservation laws and form the central object in the popular relative entropy method which allows to tackle problems of local stability and weak-strong uniqueness. On the other hand, in the calculus of variations, they correspond to the Weierstrass excess function which is crucial in the so-called Weierstrass problem of finding necessary and sufficient conditions for local minimisers.

More precisely, using ideas from the recently developed contributions of Raiță and Guerra in PDE contrained analysis [92, 63] and those from Campos-Cordero and Kristensen in the vectorial calculus of variations [31, 29], we prove a new Gårding-type inequality, Theorem 3.2:

$$\int_Q \left(|\varphi|^p + |\varphi|^2 + |\psi|^q + |\psi|^2\right) \lesssim \int_Q e(\bar{F} + \varphi, \bar{\eta} + \psi|\bar{F}, \bar{\eta}) + \|\varphi\|^p_{W^{-1,p}(Q)} + \|\varphi\|^2_{W^{-1,2}(Q)}, \quad (1.11)$$

for all $(\bar{F},\bar{\eta}) \in \mathcal{U}_K$, $\psi \in L^q(Q)$ with $\int_Q \psi = 0$ and all $\mathcal{A}$-free $\varphi \in L^p(Q)$ with $\int_Q \varphi = 0$. We refer the reader to Section 3.3 for the detailed setting and assumptions.

In dynamics, in [75], we employ a simplified version of our Gårding inequality to prove stability and weak-strong uniqueness results for dissipative solutions of general PDE-constrained systems of conservation laws of the form (1.3), under the assumption that the entropy is $\mathcal{A}$-quasiconvex. We note that no restrictions on the order of $\mathcal{A}$ are required and as in [38] weak solutions need only be bounded but with no additional assumptions on the local oscillations. Thus, we extend the ideas in [74] to the $\mathcal{A}$-free setting by exhibiting that in the general framework of [38] the assumption of small oscillations can be removed, assuming $\mathcal{A}$-quasiconvexity. In addition to this, in [55], we establish a novel uniqueness result for the hyperbolic system of adiabatic thermoelasticity (1.4) under quasiconvexity conditions, Theorem 4.3. In particular, we show that classical solutions of the system are unique within a suitable class of dissipative measure-valued solutions, provided that the internal energy e satisfies appropriate quasiconvexity assumptions, see Definition 3.1. The proof is achieved by combining the relative entropy method with our Gårding-type inequality (1.11). Our definition of quasiconvexity is associated with the symmetrisability of the system and hence guarantees its hyperbolicity.

In statics, we investigate energies of the form (1.2), i.e.

$$\mathcal{W}[U] := \int_Q W(U(x))dx, \tag{1.12}$$

for $\mathcal{A}$-free and zero average $U \in L^p(Q)$. In particular, Theorem 4.5 establishes uniqueness of local minimisers in the $W^{-1,p}$ topology for the above functionals, whenever W is strongly $\mathcal{A}$-quasivonvex, and we quantify the difference in energy between the minimiser and arbitrary competitors, i.e. we prove that if the Euler-Lagrange and the positivity of the second variation hold for some $\bar{U}$, then there exists $\varepsilon > 0$ such that

$$\mathcal{W}[U] - \mathcal{W}[\bar{U}] \gtrsim \int_Q \left(|U-\bar{U}|^p + |U-\bar{U}|^2\right), \tag{1.13}$$

for all $\mathcal{A}$-free, zero-average $U \in L^p(Q)$ with $\|U-\bar{U}\|_{W^{-1,p}(Q)} \leqslant \varepsilon$. The proof comes as a direct consequence of our main tool for the proof of (1.11), Theorem 3.1. As an example, we study the classical case $\mathcal{A} = \operatorname{curl}$ and bounded domains Ω with mixed boundary conditions and extend the results from [29], see Corollary 4.1. We note that the quantification of the difference in energy (1.13) was not known even in the classical case $\mathcal{A} = \operatorname{curl}$.

1.2 Image processing: bilevel schemes

In the second part of this thesis, Chapter 5, we study a bilevel training scheme for the automatic selection of spatially varying regularisation weights in the framework of variational image reconstruction. Specifically, given a suitably defined class Adm of admissible weights α, we look for solutions to the problem

$$\alpha^* \in \operatorname{argmin}\left\{F(u_\alpha) : \alpha \in \mathrm{Adm}\right\}, \tag{1.1}$$

where F is an assigned cost functional and u_α is an image reconstructed by minimising

$$I[u;\alpha] \coloneqq \Phi_g(u) + \mathcal{R}(u;\alpha). \tag{1.2}$$

Here, Φ_g is a fidelity term that penalises deviations of u from the datum g, whereas $\mathcal{R}(u;\alpha)$ is a regularisation functional whose strength can be tuned by an appropriate selection of the regularisation parameter α belonging to the admissible set Adm. The datum g is typically a corrupted version of some ground truth image u_{gt}. Often, one has

$$g = Tu_{\mathrm{gt}} + \eta,$$

with η denoting a random noise component and T being a bounded linear operator that corresponds to a certain image reconstruction problem. For instance, T is a blurring operator in the case of deblurring, a sub-sampled Fourier transform in magnetic resonance imaging (MRI), the Radon transform in tomography, or simply the identity in denoising tasks, on which we will be focusing here. The aim of solving a problem of the type (1.2) for suitable Φ_g, $\mathcal{R}$ and α is to obtain an output u which represents as well as possible the initial ground truth image u_{gt}.

Among classical regularisation functionals we find the total variation (TV) [95, 20], as well as higher order or anisotropic extensions of it. Particularly relevant for this work are the second order total variation (TV^2) [91, 11] and the total generalized variation (TGV) [13]. For a function $u \in L^1(\Omega)$, these functionals are defined by duality as follows:

$$\mathrm{TV}(u) = \sup\left\{\int_\Omega u\,\mathrm{div}\phi\,\mathrm{d}x : \ \phi \in \mathrm{C}_c^\infty(\Omega,\mathbb{R}^d),\ \|\phi\|_\infty \leqslant 1\right\}, \tag{1.3}$$

$$\mathrm{TV}^2(u) = \sup\left\{\int_\Omega u\,\mathrm{div}^2\phi\,\mathrm{d}x : \ \phi \in \mathrm{C}_c^\infty(\Omega,\mathbb{R}^{d\times d}),\ \|\phi\|_\infty \leqslant 1\right\}, \tag{1.4}$$

$$\mathrm{TGV}(u) = \sup\left\{\int_\Omega u\,\mathrm{div}^2\phi\,\mathrm{d}x : \ \phi \in \mathrm{C}_c^\infty(\Omega,\mathcal{S}^{d\times d}),\ \|\phi\|_\infty \leqslant \alpha_0,\ \|\mathrm{div}^2\phi\|_\infty \leqslant \alpha_1\right\}. \tag{1.5}$$

Here $\mathcal{S}^{d\times d}$ denotes the space of $d\times d$ symmetric matrices. Note that the scalar regularisation parameters $\alpha_0, \alpha_1 > 0$ are inserted within the definition of TGV, while the other functionals admit a single weighting parameter α acting in a multiplicative way, i.e., $\alpha\mathrm{TV}$ and $\alpha\mathrm{TV}^2$. If the supremum in (1.3) is finite, then we say that $u \in \mathrm{BV}(\Omega)$, the space of functions of bounded variation [5], and $\mathrm{TV}(u) = |\mathrm{D}u|(\Omega)$, where $|\mathrm{D}u|$ is the total variation measure associated with the distributional derivative $\mathrm{D}u \in \mathcal{M}(\Omega, \mathbb{R}^d)$. Similarly, if the right-hand side in (1.4) is finite, then $u \in \mathrm{BV}^2(\Omega)$, the space of functions of bounded second variation [91, 11], and $\mathrm{TV}^2(u) = |\mathrm{D}^2u|(\Omega)$, with $\mathrm{D}^2u \in \mathcal{M}(\Omega, \mathcal{S}^{d\times d})$. Finally, it turns out that if the supremum in (1.5) is finite, then $u \in \mathrm{BV}(\Omega)$ as well and

$$\mathrm{TGV}(u) = \min_{w\in\mathrm{BD}(\Omega)} \left\{ \alpha_1 \int_\Omega \mathrm{d}|\mathrm{D}u - w| + \alpha_0 \int_\Omega \mathrm{d}|\mathscr{E}w| \right\},$$

see [14, 12]. In the previous formula, $\mathrm{BD}(\Omega)$ is the space of functions of bounded deformation and $\mathscr{E}w$ denotes the symmetrised gradient of w. The advantage of higher order regularisers lies in their capability to reduce an undesirable artifact typical of TV, the so-called staircasing effect, that is, the creation of cartoon-like piecewise constant structures in the reconstruction [90].

Based on the concept of convex functions of Radon measures [48], variants of the above regularisers involving convex integrands have also been considered in the literature [91, 108, 70]. A widely used example is the one of Huber total variation TV_γ, which is defined for $u \in \mathrm{BV}(\Omega)$ as

$$\mathrm{TV}_\gamma(u) = \int_\Omega f_\gamma(\mathrm{dD}u) = \int_\Omega f_\gamma(\nabla u)\mathrm{d}x + \int_\Omega \mathrm{d}|\mathrm{D}^s u|, \tag{1.6}$$

with ∇u and $\mathrm{D}^s u$ denoting respectively the absolutely continuous and the singular part of $\mathrm{D}u$ with respect to the Lebesgue measure. The function $f_\gamma \colon \mathbb{R}^d \to \mathbb{R}$ is given for $\gamma \geq 0$ by

$$f_\gamma(z) = \begin{cases} |z| - \dfrac{1}{2}\gamma, & \text{if } |z| \geq \gamma, \\ \dfrac{1}{2\gamma}|z|^2, & \text{if } |z| < \gamma. \end{cases} \tag{1.7}$$

Note that $\mathrm{TV}_\gamma(u)$ can be equivalently defined via duality as

$$\mathrm{TV}_\gamma(u) = \sup \left\{ \int_\Omega u \,\mathrm{div}\phi \,\mathrm{d}x - \frac{1}{2}\gamma \int_\Omega |\phi|^2 \mathrm{d}x : \phi \in \mathrm{C}_c^\infty(\Omega, \mathbb{R}^d),\ \|\phi\|_\infty \leqslant 1 \right\}, \tag{1.8}$$

see [48]. This modification of TV, is typically considered in order to employ classical smooth numerical solvers for the solution of the minimisation problem (1.2). In this specific case, however, it also leads to a reduction of the staircasing effect by penalising small gradients with the Tikhonov term $(2\gamma)^{-1} \int_\Omega |\nabla u|^2 dx$, which promotes smooth reconstructions [70, 18].

In all these models, the choice of the weights in the regularisation term is crucial to establish an adequate balance between data fitting and denoising. On the one hand, the reconstructed u may remain too noisy or have many artifacts if the regularisation is too weak. On the other hand, a very strong regularisation may result in an unnatural smoothing effect. In the last years, bilevel minimisation methods have been employed to select these weights automatically. A subfamily of these methods assumes the existence of one or several training pairs (u_{gt}, g) consisting of the ground truth and its corrupted counterpart [19, 44, 43, 42, 84]. In these works the energy in the upper level problem (1.1) is usually given by

$$F_{\mathrm{PSNR}}(u) := \|u - u_{\mathrm{gt}}\|^2_{L^2(\Omega)}, \tag{1.9}$$

and its minimisation essentially corresponds to finding reconstructions that are closest to the ground truth in the least square sense. Since typical images generally feature both homogeneous regions and fine details, it is reasonable to assume that the optimal regularisation intensity is not uniform throughout the domain. This matter of fact has prompted researchers to consider bilevel schemes that output space-dependent weights, i.e., functions $\alpha\colon \Omega \to [0, +\infty)$ [26, 45]. A recent series of papers [68, 69, 66, 67] deals with schemes for TV and TGV that yield such weights without resorting to the ground truth. If the corrupted datum g is obtained by some additive Gaussian noise η with variance σ^2, this is achieved by the introduction of the statistics-based upper level objective

$$F_{\mathrm{stat}}(u) := \frac{1}{2}\int_\Omega \max(Ru - \overline{\sigma}^2, 0)^2 \mathrm{d}x + \frac{1}{2}\int_\Omega \min(Ru - \underline{\sigma}^2, 0)^2 \mathrm{d}x, \tag{1.10}$$

where $\underline{\sigma}^2 := \sigma^2 - \varepsilon$, $\overline{\sigma}^2 := \sigma^2 + \varepsilon$, and

$$Ru(x) := \int_\Omega w(x,y)(u-g)^2(y)\mathrm{d}y \quad \text{for } w \in L^\infty(\Omega \times \Omega), \quad \int_\Omega\int_\Omega w(x,y)\mathrm{d}x\mathrm{d}y = 1.$$

The idea is that if the reconstructed image u is close to u_{gt}, then it is expected that on average the value of $Ru(x)$ will be close to σ^2. This justifies the use of F_{stat}, since its minimisation forces the localised residuals Ru to fall within the tight corridor $[\underline{\sigma}^2, \overline{\sigma}^2]$.

Image processing: the contribution of this thesis

Our contribution in the field is connected to the aforementioned literature on several levels. Starting from an arbitrary l-th order, homogeneous, linear differential operator $\mathcal{B}$ between two finite dimensional Euclidean spaces $\mathbb{U}$ and $\mathbb{V}$, we introduce the general regulariser

$$\mathcal{R}(u;\alpha) = \sum_{i=1}^{l-1} \int_\Omega \alpha_i(x) f_i(x, \mathrm{d}\mathrm{D}^i u) + \int_\Omega \alpha_l(x) f_l(x, \mathrm{d}\mathcal{B}u), \tag{1.11}$$

$\alpha_i\colon \Omega \to [0, +\infty)$ being for $i = 1, \dots, l$ the spatially dependent weights. The functions f_i are of linear growth and convex in the second variable. We assume them to be Carathéodory integrands, or in other words, they explicitly depend on the spatial variable x in a measurable way. More details about the setting are to be found in Section 5.1. As a first contribution, we prove lower semicontinuity of the functional in (1.11) with respect to a suitable weak-$*$ convergence, a necessary step towards the existence of solutions of the corresponding variational image reconstruction problem (1.2). Secondly, we introduce and prove existence of solutions to the bilevel scheme, which provides an optimal spatially dependent weight and an associated reconstructed image.

Not much work has been done for functionals depending on general differential operators. One example comes from the recent preprint [41], where a bilevel scheme for first order differential operators $\mathcal{B}$ is developed. Interestingly, the authors identify classes of operators $\mathcal{B}$ such that the scheme outputs an optimal reconstructed image and an optimal $\mathcal{B}$ for the upper level problem. However, in their analysis one always obtains BV minimisers. In contrast, in our method the operator $\mathcal{B}$ is allowed to be *arbitrary*, see Theorems 5.4 and 5.6.

From the theoretical point of view, one of the main advantages of our approach is the fact that we can allow for *spatially dependent weights* and for *general convex integrands* in the regularisers. Our hypotheses on the convex integrands are optimal, due to the use of Young measures for oscillation and concentration, see Section 2.3. From an analytical point of view, our regularity assumptions on the weights are minimal, as can be seen from Section 5.1.3. In the future, we aim to develop our theory to include optimisation problems over linear PDE operators $\mathcal{B}$ that satisfy as few assumptions as possible. We expect that our lower semicontinuity and existense results, Theorems 5.4 and 5.6 respectively, will serve as preliminary work in this direction.

We conclude with a series of numerical examples that deal with versions of the Huber TV and TV^2 in which both the Huber and the regularisation parameter are spatially dependent. We devise a strategy to prefix the former in a sensible way, while the latter is computed automatically by the bilevel scheme. Even though the main purpose of these numerical examples is to support the applicability and versatility of the framework, we are able to draw two interesting conclusions. The first one is that the bilevel TV^2 with spatially varying weight, in combination with the statistics-based upper level objective F_{stat} is able to produce high quality reconstructions, even outperforming TGV, both in its scalar and spatially varying versions. The second one is that the introduction of the spatially varying Huber parameter can further enhance the detailed areas in the reconstructed images.

Chapter 2

General Preliminaries

In this chapter we summarise some general notions and tools which are widely used in this thesis. In particular, we present some important aspects from the theory of constant rank differential operators which is central in our work. After that, we proceed to a review of the theory of generalised Young measures, and then we mention some key results from the PDE constrained quasiconvex analysis. More specialised notions and tools that are used in this work, are mentioned in the respective chapters and sections of the thesis.

2.1 Constant rank linear operators

We first clarify the notation and then we present some important results which play a key role in the analysis of our work. For each d-multi index $j = (j_1, \ldots, j_d) \in \mathbb{N}^d$, let us consider a collection of linear operators $A_j \in \mathrm{Lin}(\mathbb{R}^N, \mathbb{R}^M)$. We define a homogeneous k-th order linear operator $\mathcal{A}$ by

$$\mathcal{A}\varphi := \sum_{|j|=k} A_j \partial^j \varphi, \ \varphi : \mathbb{R}^d \to \mathbb{R}^N, \tag{2.1}$$

where $|j| = \sum_i j_i$. We think of $\mathcal{A}$ as a polynomial in ∂ and so we write its principal symbol as

$$\mathbb{A} : \mathbb{R}^d \to \mathrm{Lin}(\mathbb{R}^N, \mathbb{R}^M), \quad \mathbb{A}(\xi) = (2\pi i)^k \sum_{|j|=k} A_j \xi^j.$$

The wave cone associated with $\mathcal{A}$ is denoted by

$$\Lambda_{\mathcal{A}} = \bigcup_{\xi \in S^{d-1}} \ker \mathbb{A}(\xi),$$

and contains the amplitudes $\lambda \in \mathbb{R}^N$ along which the system fails to be elliptic where ellipticity means that $\ker \mathbb{A}(\xi) = \{0\}$ for all $\xi \neq 0$. Indeed, it is straightforward to check that $\lambda \notin \Lambda_{\mathcal{A}}$ if and only if the operator $\mathcal{R}_\lambda(v) := \mathcal{A}(\lambda v)$ is elliptic, where $v \in C^\infty(Q; \mathbb{R})$.

Moreover, we assume throughout that the linear differential operator $\mathcal{A}$ has the constant rank property, i.e. there exists $r \in \mathbb{N}$ such that

$$\operatorname{rank} \mathbb{A}(\xi) = r \text{ for all } \xi \in S^{d-1}. \tag{2.2}$$

As we already mentioned in the introduction, the constant rank assumption, first introduced in the context of compensated compactness by Murat [87], ensures the smoothness of the projection mapping

$$\mathbb{P} : \mathbb{R}^d \setminus \{0\} \to \operatorname{Lin}(\mathbb{R}^N, \mathbb{R}^N), \;\; \xi \mapsto \operatorname{Proj}_{\ker \mathbb{A}(\xi)},$$

and thus makes tools of pseudo-differential calculus available. More precisely, the projection mapping $\mathbb{P}$ can be represented as

$$\mathbb{P}(\xi) = \operatorname{Id}_N - \mathbb{A}^\dagger(\xi)\mathbb{A}(\xi), \quad \text{for } \xi \in \mathbb{R}^d \backslash \{0\},$$

where by $\mathbb{A}^\dagger(\xi)$ we denote the pseudo-inverse of $\mathbb{A}(\xi)$. We recall the notion of Moore-Penrose generalised inverse (so-called pseudo-inverse): for a matrix $K \in \mathbb{R}^{N \times M}$, its pseudo-inverse $K^\dagger$ is the unique $M \times N$ matrix defined by the relations

$$KK^\dagger K = K, \;\; K^\dagger K K^\dagger = K^\dagger, \;\; (KK^\dagger)^* = KK^\dagger, \;\; (K^\dagger K)^* = K^\dagger K,$$

where we use the symbol * for the adjoint matrix. Using the above representation together with [47, Theorem 3], Raiţă in [92] gave a new characterisation for constant rank operators:

Theorem 2.1. *Let $\mathcal{A}$ be a linear homogeneous differential operator with constant coefficients of order $k \in \mathbb{N}$. Then $\mathcal{A}$ has constant rank if and only if there exists a linear homogeneous differential operator $\mathcal{B}$ with constant coefficients and order $l \in \mathbb{N}$ such that*

$$\operatorname{im} \mathbb{B}(\xi) = \ker \mathbb{A}(\xi) \text{ for all } \xi \in \mathbb{R}^d \setminus \{0\}.$$

We write, for some $B_j \in \operatorname{Lin}(\mathbb{R}^{M'}, \mathbb{R}^N)$,

$$\mathcal{B}\phi := \sum_{|j|=l} B_j \partial^j \phi, \;\; \phi : \mathbb{R}^d \to \mathbb{R}^{M'}. \tag{2.3}$$

We refer to the potential operator $\mathcal{B}$ simply as the potential of $\mathcal{A}$ although no meaningful notion of uniqueness is known, see [63] for a discussion. We mention here that the potential operator $\mathcal{B}$ is also of constant rank, see [92, Lemma 3].

For the sake of concreteness, we present some examples which are mainly motivated by material science and are also used in this thesis.

(i) Unconstrained case: If $\mathcal{A} = 0$ then it is straightforward to see that $\Lambda_{\mathcal{A}} = \mathbb{R}^N$. We

note that in that case $\mathcal{A}$-quasiconvexity, see Definition 2.3, is just the normal convexity;
(ii) Non-linear Elasticity: If $\mathcal{A} = \operatorname{curl}$ then curl-free vector fields can be written as gradients of some other vector fields, i.e. $\mathcal{B} = \nabla$, and the associated wave cone is the cone of rank-one matrices i.e. $\Lambda_{\mathcal{A}} = \{a \otimes b : a \in \mathbb{R}^N, b \in \mathbb{R}^d\}$. Similar arguments can be done for higher gradients or even when only some of the partial derivatives are considered, see [53, Example 3.10.];
(iii) Linear Elasticity: For $\mathcal{A} = \operatorname{curl}\operatorname{curl}$ then the potential operator is given as the symmetric gradient $\mathcal{E}$ so that $\Lambda_{\mathcal{A}} = \{\frac{1}{2}(a \otimes b + b \otimes a) : a \in \mathbb{R}^N, b \in \mathbb{R}^d\}$;
(iv) Thermoelasticity: In this case we observe a couple of constraints described via the operator $\mathcal{A} = (\operatorname{curl}, 0)$ where the associated wave cone is given by

$$\Lambda_{\mathcal{A}} = \Big\{(a \otimes n, \eta) : a \in \mathbb{R}^N,\, n \in \mathbb{R}^d \text{ and } \eta \in \mathbb{R}\Big\}.$$

We note that the above situation can be generalised by coupling any admissible constraints and hence obtaining a new operator.

Refer to reader to [63, 53, 81] for the details and also some more delicate examples.

2.1.1 Sobolev estimates

We first take a moment to clarify our setting. In a large part of this thesis and in particular in the analysis of our Gårding type inequality and its applications, for technical reasons we restrict attention to functions $\varphi : \mathbb{R}^d \to \mathbb{R}^N$ that are Δ-periodic where Δ denotes the unit cell of the lattice $\mathbb{Z}^d$, i.e.

$$\varphi(x + P) = \varphi(x) \ \ \forall\, P \in \mathbb{Z}^d.$$

These functions can be viewed as functions on the d-dimensional (flat) torus $\mathbb{T}^d$

$$\mathbb{T}^d := \Big\{(e^{2\pi i x_1}, \dots, e^{2\pi i x_d}) : (x_1, \dots, x_d) \in \mathbb{R}^d\Big\}$$

via the identification

$$\varphi_T(e^{2\pi i x_1}, \dots, e^{2\pi i x_d}) = \psi(x_1, \dots, x_d).$$

Thus, letting Q be (any translation of) the unit cube $(0,1)^d$, we may identify $L^p(\mathbb{T}^d)$ with $L^p(Q)$, understanding that the natural measure on $\mathbb{T}^d$ is the pushforward of the Lebesgue measure $\mathcal{L}^d$ on Q via the map $f : Q \to \mathbb{T}^d$,

$$f(x_1, \dots, x_d) = f(e^{2\pi i x_1}, \dots, e^{2\pi i x_d}).$$

Then, for any function $\varphi_T \in L^1(\mathbb{T}^d)$ and *any* unit cube Q (up to the obvious modifications to the map f) it holds that

$$\int_{\mathbb{T}^d} \varphi_T = \int_Q \varphi.$$

For the purposes of this thesis, we do not distinguish between φ_T and φ, or between $\mathbb{T}^d$ and Q and their respective measures, and we prefer to define our integrals and function spaces over Q with opposite boundaries glued, rather than $\mathbb{T}^d$. This is because we often consider functions defined on cubes which are seen as subsets of the unit cube Q, although Q itself is a unit cube with opposite sides identified. Consequently, we write $L^p(Q)$ instead of $L^p(\mathbb{T}^d)$ but also $C^k(Q)$ instead of $C^k(\mathbb{T}^d)$, i.e.

$$C^k(Q) := \left\{\varphi \in C^k(\mathbb{R}^d) : \partial^\alpha \varphi \ \Delta\text{-periodic for all } d\text{-multi-index } |\alpha| \leqslant k\right\}. \tag{2.4}$$

Henceforth, for a function $\varphi \in L^p(Q)$ we say that "$\mathcal{A}\varphi = 0$ in Q" in the sense of distributions on the torus, i.e.

$$-\int_Q \varphi \cdot \mathcal{A}^* v = 0 \ \text{ for all } \ v \in C^\infty(Q), \tag{2.5}$$

where $\mathcal{A}^*$ is the adjoint operator. We call $\mathcal{A}$-free any function satisfying (2.5). Taking this into account, we define the space

$$L^p_{\mathcal{A}}(Q) := \left\{\varphi \in L^p(Q) : \ \mathcal{A}\varphi = 0, \ \int_Q \varphi = 0\right\},$$

which contains all the $\mathcal{A}$-free and zero-average functions of $L^p(Q)$.

In this section, we present some fundamental estimates in Sobolev spaces for a class of primitive functions which we refer to as $\mathbb{B}^\dagger$-primitives, constructed in [92]. These estimates are necessary to replace Poincaré-type inequalities which we particularly require when introducing cut-offs in the analysis of our main inequality, section 3.3. We note that these estimates may fail for general primitives.

Remark 2.1. Throughout, $W^{l,q}(Q)$ denotes the closure of $C^\infty(Q)$ in the $W^{l,q}$ norm. Then, for $p = q/(q-1)$, the space $W^{-l,p}(Q)$ is its dual and its norm is equivalent to

$$\Big\|\mathcal{F}^{-1}\Big[\frac{\hat{\varphi}(\xi)}{(1+|\xi|^2)^{l/2}}\Big]\Big\|_{L^p(Q)}.$$

Note that when $\int_Q \varphi = 0$ this norm is equivalent to the norm

$$\Big\|\mathcal{F}^{-1}\Big[\frac{\hat{\varphi}(\xi)}{|\xi|^l}\Big]\Big\|_{L^p(Q)}$$

since the *Fourier* multipliers $(1+|\xi|^2)^{-l/2}$ and $|\xi|^{-l}$ are comparable for $\xi \in \mathbb{Z}^d \setminus \{0\}$. Here, $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, and for the Fourier coefficients we use the notation

$$\varphi(x) = \sum_{\xi \in \mathbb{Z}^d} \hat{\varphi}(\xi) e^{2\pi i \cdot x}, \quad \text{for } \ x \in Q, \, \varphi \in C^\infty(Q), \quad \text{where } \hat{\varphi}(\xi) := \int_Q \varphi(y) e^{-2\pi i \xi \cdot y} dy.$$

Before we proceed to the proof of the Sobolev estimates, we present a well-known result from Harmonic analysis which we use repeatedly, the Hörmander-Mikhlin multiplier theorem:

Theorem 2.2. *Let $m \in C^\infty(\mathbb{R}^d \setminus \{0\})$ be 0-homogeneous. Then*

$$\left\|\mathcal{F}^{-1}\left(m\hat{\varphi}\right)\right\|_{L^p(\mathbb{R}^d)} \leqslant \|\varphi\|_{L^p(\mathbb{R}^d)}, \quad \text{for all } \varphi \in L^p(\mathbb{R}^d).$$

In the sequel the symbol $\lesssim$ is used to denote that the inequality $\leqslant$ holds up to a positive constant i.e. for two quantities q_1 and q_2 we say that $q_1 \lesssim q_2$ if there exists a constant $C > 0$ such that $q_1 \leqslant C\, q_2$.

Lemma 2.1. *Let $\mathcal{A}$ and its potential $\mathcal{B}$ as in Theorem 2.1. Then for all $\varphi \in L^p_{\mathcal{A}}(Q)$, there exists $\phi \in W^{l,p}(Q)$ such that*

(i) $\varphi = \mathcal{B}\phi$;

(ii) $\|\phi\|_{L^p(Q)} \leqslant C\|\varphi\|_{W^{-l,p}(Q)}$;

(iii) $\|\phi\|_{W^{l,p}(Q)} \leqslant C\|\varphi\|_{L^p(Q)}$;

(iv) $\|\phi\|_{W^{l-i,p}(Q)} \leqslant C\|\varphi\|_{W^{-1,p}(Q)}$ *for all* $i = 1, .., l$.

We will call ϕ a $\mathbb{B}^\dagger$-primitive of φ.

Although (ii), (iii) and (iv) follow from the construction in [92], a proof is not explicitly given. Hence, for completeness, we provide a proof here.

Proof. We prove the result for $\varphi \in C^\infty(Q)$ and the general case follows by approximation. Indeed, (i) is known from [92, Lemma 2], where the primitive function $\phi \in C^\infty(Q)$ is constructed as

$$\phi(x) = \sum_{\xi \neq 0} \mathbb{B}^\dagger(\xi)\widehat{\varphi}(\xi)e^{2\pi i \xi \cdot x},$$

and $\mathbb{B}^\dagger(\cdot)$ is the pseudo-inverse of $\mathbb{B}(\cdot)$ which is itself smooth whenever $\mathcal{B}$ is, see [63]. This justifies our adopted terminology $\mathbb{B}^\dagger$-primitive.

For (ii), since $\mathbb{B}^\dagger(\cdot)$ is smooth ($\mathbb{B}(\cdot)$ is smooth by construction) and $(-l)$-homogeneous, the operator $\mathbb{B}^\dagger(\xi/|\xi|)$ is 0-homogeneous and smooth, and thus a Fourier multiplier, see [52, Proposition 2.13]. Hence, by the Hörmander-Mikhlin multiplier theorem, Theorem 2.2, and Remark 2.1

$$\|\phi\|_{L^p(Q)} = \left\|\mathcal{F}^{-1}\Big[\frac{1}{|\xi|^l}\mathbb{B}^\dagger(\frac{\xi}{|\xi|})\widehat{\varphi}(\xi)\Big]\right\|_{L^p} \lesssim \left\|\mathcal{F}^{-1}\Big[\frac{1}{|\xi|^l}\widehat{\varphi}(\xi)\Big]\right\|_{L^p} = \|\varphi\|_{W^{-l,p}(Q)}.$$

For (iii), by applying the Poincaré inequality for all the derivatives of ϕ, since $\int_Q \nabla^i \phi = \widehat{\nabla^i \phi}(0) = 0$, we have that $\|\nabla^{l-i}\phi\|_{L^p} \lesssim \|\nabla^l \phi\|_{L^p}$ for all i=0,..,l and so $\|\phi\|_{W^{l,p}} \lesssim \|\nabla^l \phi\|_{L^p}$. Then, by differentiating ϕ we obtain

$$\nabla^l \phi(x) = \sum_{\xi \neq 0} \mathbb{B}^\dagger(\xi)\widehat{\varphi}(\xi) e^{2\pi i \xi \cdot x} \otimes \xi^{\otimes l},$$

where $\mathbb{B}^\dagger(\xi) \otimes \xi^{\otimes l}$ is a 0-homogeneous multiplier of φ, since $\mathbb{B}^\dagger(\cdot)$ is $(-l)$-homogeneous. Hence, again by Theorem 2.2, we find that

$$\|\nabla^l \phi\|_{L^p(Q)} = \Big\| \mathcal{F}^{-1}\Big[\mathbb{B}^\dagger(\frac{\xi}{|\xi|})\widehat{\varphi}(\xi)\Big]\Big\|_{L^p} \lesssim \Big\| \mathcal{F}^{-1}\Big[\widehat{\varphi}(\xi)\Big]\Big\|_{L^p} = \|\varphi\|_{L^p(Q)} = \|\mathcal{B}\phi\|_{L^p(Q)}.$$

For (iv), by working similarly to (iii) we prove that

$$\|\nabla^{l-1}\phi\|_{L^p(Q)} \lesssim \Big\| \mathcal{F}^{-1}\Big[\frac{1}{|\xi|}\widehat{\varphi}(\xi)\Big]\Big\|_{L^p(Q)} = \|\varphi\|_{W^{-1,p}(Q)}$$

and since $\|\nabla^{l-i}\phi\|_{L^p} \lesssim \|\nabla^{l-1}\phi\|_{L^p}$ for i=1,..,l we conclude the proof. $\square$

2.2 $\mathcal{A}$-quasiconvexity

Here we recall the definition of $\mathcal{A}$-quasiconvexity and collect results that are used in the sequel. The following definition is due to Fonseca and Müller in [53].

Definition 2.3. A locally bounded, Borel function $W : \mathbb{R}^N \to \mathbb{R}$ is $\mathcal{A}$-quasiconvex at $\lambda \in \mathbb{R}^N$ if

$$\int_Q \big[W(\lambda + \varphi(x)) - W(\lambda)\big] dx \geqslant 0,$$

for all $\varphi \in C^\infty(Q)$ such that $\mathcal{A}\varphi = 0$ and $\int_Q \varphi = 0$.

It is proved in [92] that the above definition can equivalently be expressed over arbitrary domains and compactly supported test functions, i.e. it coincides with Dacorogna's definition of $\mathcal{A}$-$\mathcal{B}$ quasiconvexity [33] given below.

Definition 2.4. Let $\Omega \subseteq \mathbb{R}^d$ be a non-empty open subset. A locally bounded, Borel function $W : \mathbb{R}^N \to \mathbb{R}$ is $\mathcal{A}$-quasiconvex at $\lambda \in \mathbb{R}^N$ if and only if

$$\int_\Omega \big[W(\lambda + \mathcal{B}\phi(x)) - W(\lambda)\big] dx \geqslant 0, \text{ for all } \phi \in C_c^\infty(\Omega),$$

where $\mathcal{B}$ is the potential of $\mathcal{A}$.

Additionally, assuming that W has at most p-growth, i.e. $|W(z)| \leqslant c(1 + |z|^p)$, , using density results, the above definitions can also be expressed with test functions in $L^p(Q)$ and $W_0^{l,p}(\Omega)$, respectively, where $W_0^{l,p}(\Omega)$ denotes the closure of $C_c^\infty(\Omega)$ in the $W^{l,p}$-norm.

The results presented in this thesis, require strengthened versions of the quasiconvexity condition. Let $p \geqslant 2$ and for $k \in \mathbb{N}$ define the auxiliary function $V_p : \mathbb{R}^k \to \mathbb{R}$ as

$$V_p(z) := (|z|^2 + |z|^p)^{1/2}, \quad z \in \mathbb{R}^k. \tag{2.1}$$

If there exists a constant $c_0 > 0$ such that

$$\int_\Omega \big[W(\lambda + \mathcal{B}\phi(x)) - W(\lambda)\big]dx \geqslant c_0 \int_\Omega |V(\mathcal{B}\phi(x))|^2 dx, \tag{2.2}$$

for all $\phi \in W_0^{l,p}(\Omega)$, we say that W is strongly $\mathcal{A}$-quasiconvex at $\lambda \in \mathbb{R}^N$. Equivalently, W is strongly $\mathcal{A}$-quasiconvex at $\lambda \in \mathbb{R}^N$ if

$$\int_Q \big[W(\lambda + \varphi(x)) - W(\lambda)\big]dx \geqslant c_0 \int_Q |V_p(\varphi(x))|^2 dx,$$

for all $\varphi \in L^p_{\mathcal{A}}(Q)$. We say that W is (strongly) $\mathcal{A}$-quasiconvex, if it is (strongly) $\mathcal{A}$-quasiconvex at λ for all $\lambda \in \mathbb{R}^N$.

Note that $\mathcal{A}$-quasiconvex functions are not in general continuous as, unlike quasiconvex functions, they are not generally separately convex. However, the condition $\mathrm{span}\Lambda_{\mathcal{A}} = \mathbb{R}^N$ recovers this loss of separate convexity, see [63, Lemma 4.4], and then

$$|W(z_1) - W(z_2)| \leqslant C(1 + |z_1|^{p-1} + |z_2|^{p-1})|z_1 - z_2|, \text{ for all } z_1, z_2 \in \mathbb{R}^N.$$

We end this subsection with a remark on quadratic forms. It is well-known that for these functions rank-one convexity, see [34] for the respective definition, implies quasiconvexity. Here, repeating the arguments of [33], we extend the above implication in the case of $\mathcal{A}$-quasiconvexity. We first present a result which is crucial in the proof of the extension.

Lemma 2.2. *Let $M \in \mathbb{R}^{N\times N}$ be a symmetric matrix and define the function $f(\lambda) = M\lambda\cdot\lambda$, for all $\xi \in \mathbb{R}^N$. Then, if*

$$\int_Q f\big(\varphi(x)\big)dx \geqslant 0,$$

for all $\varphi \in C_c^\infty(Q)$, $\mathcal{A}$-free and zero-average, the function f is $\mathcal{A}$-quasiconvex.

Proof. Let $\lambda \in \mathbb{R}^N$, $\varphi \in C_c^\infty(Q)$ $\mathcal{A}$-free and zero-average. Then

$$\begin{aligned}\int_Q f\big(\lambda + \varphi(x)\big)dx &= \int_Q M\big(\lambda + \varphi(x)\big)\cdot\big(\lambda + \varphi(x)\big)dx \\ &= \int_Q M\lambda\cdot\lambda dx + \int_Q M\varphi(x)\cdot\varphi(x)dx \\ &\geqslant \int_Q M\lambda\cdot\lambda dx = f(\lambda),\end{aligned}$$

where in the second equality we used the fact that $\int_Q \varphi = 0$. □

The above lemma together with the fact that $\Lambda_{\mathcal{A}}$-convex quadratic functions are non-negative on the wave cone $\Lambda_{\mathcal{A}}$ yields:

Lemma 2.3. *Let $M \in \mathbb{R}^{N\times N}$ be a symmetric matrix and define the function $f(\lambda) = M\lambda\cdot\lambda$, for all $\xi \in \mathbb{R}^N$, and assume in addition that f is convex in the wave cone $\Lambda_{\mathcal{A}}$. Then f is $\mathcal{A}$-quasiconvex.*

Proof. Since f is convex on the cone $\Lambda_{\mathcal{A}}$, for any $\xi \in \mathbb{R}^N$ it holds that

$$\begin{aligned}&\nabla_\xi^2 f(\xi)\lambda\cdot\lambda \geqslant 0\\ &\quad\Rightarrow M\lambda\cdot\lambda \geqslant 0, \quad \text{for all} \quad \lambda \in \Lambda_{\mathcal{A}}\backslash\{0\},\end{aligned}$$

and so $f(\lambda) \geqslant 0$ for all $\lambda \in \Lambda_{\mathcal{A}}\backslash\{0\}$. From Lemma 2.2, it is enough to show that

$$\int_Q f\big(\varphi(x)\big)dx \geqslant 0,$$

for all $\varphi \in C_c^\infty(Q)$, $\mathcal{A}$-free and zero-average. Indeed,

$$\int_Q f\big(\varphi(x)\big)dx = \int_Q M\varphi(x)\cdot\varphi(x)dx = \int_{\mathbb{R}^N} M\hat{\varphi}(\xi)\cdot\hat{\varphi}(\xi)d\xi \geqslant 0,$$

where the last equality follows from Plancherel's theorem. □

2.3 Young measures

In the sequel, we are interested in two classes of integrands. The first one $\mathbb{E}_p(\Omega,\mathbb{R}^N)$, which is mainly associated with Chapter 3, is more natural with respect to the assumed growth behaviour of our integrands, while the second one, denoted by $\mathbb{L}(\Omega,\mathbb{R}^N)$ and studied in Chapter 5 for integrands with linear growth, is larger from the respective class $\mathbb{E}(\Omega,\mathbb{R}^N) := \mathbb{E}_1(\Omega,\mathbb{R}^N)$ and contains integrands for which, it will turn out, that we are able to compute limits. We note that all of the following results hold the same if we replace $\mathbb{R}^N$ with any finite dimensional space $\mathbb{V}$, however we restrict ourselves to $\mathbb{R}^N$ for simplicity of the notation.

We want to understand the limiting behaviour of the sequence $(f(\cdot,\varphi_j))$ over weakly convergent sequences $(\varphi_j) \subset L^p(\Omega,\mathbb{R}^N)$ (or the sequence $(f(\cdot,\mu_j))$ for weakly-$*$ convergent sequences of measures $\mu_j \in \mathcal{M}(\Omega,\mathbb{R}^N)$ in the case $p=1$). To this end, we define the natural space

$$\mathbb{G}_p(\Omega,\mathbb{R}^N) := \left\{ f \in C(\Omega\times\mathbb{R}^N) :\ \|f\|_{\mathbb{G}_p} := \sup_{(x,z)} \frac{|f(x,z)|}{(1+|z|)^p} < \infty \right\},$$

which can be equivalently described using the sphere compactification operator $T_p : \mathbb{G}_p(\Omega, \mathbb{R}^N) \to C_b(\Omega \times B^N)$ (and its inverse T_p^{-1})

$$(T_p f)(x, \tilde{z}) := (1 - |\tilde{z}|)^p f\left(x, \frac{\tilde{z}}{1 - \tilde{z}}\right), \quad x \in \Omega,\ \tilde{z} \in B^N,$$
$$(T_p^{-1} g)(x, z) := (1 + |z|)^p g\left(x, \frac{z}{1 + z}\right), \quad x \in \Omega,\ z \in \mathbb{R}^N.$$

In particular it holds that $T_p : \mathbb{G}_p(\Omega, \mathbb{R}^N) \to C_b(\Omega \times B^N)$ is a linear isometric isomorphism and therefore, the same holds for the adjoint operator $T_p^* : C_b(\Omega \times B^N)^* \to \mathbb{G}_p(\Omega, \mathbb{R}^N)^*$, a property which reveals that the transformation mapping T_p captures the behaviour of the p-growth integrands. However, due to the fact that $\mathbb{G}_p(\Omega, \mathbb{R}^N)^* \simeq C_b(\Omega \times B^N)^*$ is the dual of a non-separable Banach space, by testing with integrands in $\mathbb{G}_p(\Omega, \mathbb{R}^N)$, we lack an appropriate compactness principle. In order to overcome this issue, we look at the subspace of $\mathbb{G}_p(\Omega, \mathbb{R}^N)$ which is isomorphic to the separable space $C(\bar{\Omega} \times \bar{B}^N)$, and we define the space of p-admissible intergrands $\mathbb{E}_p(\Omega, \mathbb{R}^N)$ as follows:

$$\begin{aligned}\mathbb{E}_p(\Omega, \mathbb{R}^N) &:= T_p^{-1}(C(\bar{\Omega} \times \bar{B}^N)) \\ &= \left\{ f \in C(\bar{\Omega} \times \mathbb{R}^N) : f_p^\infty(x, z) := \lim_{(x', z', t) \to (x, z, \infty)} \frac{f(x', tz')}{t^p} \in \mathbb{R},\ \text{for } (x, z) \in \bar{\Omega} \times \mathbb{R}^N \right\},\end{aligned}$$

where $f_p^\infty : \bar{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ is called the (strong) p-recession function of f. It is straightforward to see that the function f_p^∞ is itself a p-admissible integrand and positively p-homogeneous in z, i.e.

$$f_p^\infty \in \mathbb{E}_p(\Omega, \mathbb{R}^N), \quad \text{and} \quad f_p^\infty(x, tz) = t^p f_p^\infty(x, z) \ \text{ for } t \geqslant 0,\ x \in \bar{\Omega},\ z \in \mathbb{R}^N.$$

In the sequel, we write just $\mathbb{E}(\Omega, \mathbb{R}^N)$ and f^∞ for the case $p = 1$.

Using the above definitions, we are able to understand the space of Young measures as a subspace of $\mathbb{E}_p(\Omega, \mathbb{R}^N)^*$. To this end, we embed the space $L^p(\Omega, \mathbb{R}^N)$ (or $\mathcal{M}(\Omega, \mathbb{R}^N)$ in the case $p = 1$) into $\mathbb{E}_p(\Omega, \mathbb{R}^N)^*$ via

$$\textbf{p>1:} \quad \varepsilon_\varphi(f) := \int_\Omega f(x, \varphi(x)) dx, \quad \text{for } \varphi \in L^p(\Omega, \mathbb{R}^N),$$
$$\textbf{p=1:} \quad \varepsilon_\mu(f) := \int_\Omega f(x, \mathrm{d}\mu) = \int_\Omega f(x, \mu^a(x)) \mathrm{d}x + \int_\Omega f^\infty\left(x, \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu|}(x)\right) \mathrm{d}|\mu|(x),$$

where $\mu = \mu^a \mathcal{L}^n \llcorner \Omega + \mu^s$ is the Radon–Nikodym decomposition of $\mu \in \mathcal{M}(\Omega, \mathbb{R}^N)$. By the sequential Banach-Alaoglu theorem, we can infer that bounded L^p (or $\mathcal{M}$ in the case p=1) sequences are weakly-* compact in $\mathbb{E}^*$ under the above identification. In particular, if (φ_j) is bounded in $L^p(\Omega, \mathbb{R}^N)$, we know that, along a subsequence which is not relabeled, we have that $\varepsilon_{\varphi_j} \overset{*}{\rightharpoonup} \boldsymbol{\nu}$ in $\mathbb{E}_p(\Omega, \mathbb{R}^N)^*$.

We define $\sigma := (T_p^{-1})^*\boldsymbol{\nu} \in C(\overline{\Omega \times B^N})^* \backsimeq \mathcal{M}(\overline{\Omega \times B^N})$ and, for $f \in \mathbb{E}_p(\Omega, \mathbb{R}^N)$, we write

$$\begin{aligned}\langle\langle \boldsymbol{\nu}, f\rangle\rangle &:= \langle \boldsymbol{\nu}, f\rangle_{\mathbb{E},\mathbb{E}^*} = \langle \sigma, T_p f\rangle \\ &= \int_{\bar\Omega\times B^N} (1-|\tilde z|)^p f\left(x, \frac{\tilde z}{1-|\tilde z|}\right) \mathrm{d}\sigma(x,\tilde z) + \int_{\bar\Omega\times S^N} f_p^\infty(x,\tilde z)\mathrm{d}\sigma(x,\tilde z),\end{aligned}$$

where S^N denotes the unit sphere in $\mathbb{R}^N$. From this formula we derive two necessary conditions for the weak-$*$ limit $\boldsymbol{\nu}$, namely that $\sigma \geqslant 0$ in $\mathcal{M}(\overline{\Omega \times B^N})$ and

$$\int_\Omega \psi(x)\mathrm{d}x = \int_{\bar\Omega\times B^N} \psi(x)(1-|\tilde z|)^p \mathrm{d}\sigma(x,\tilde z) \text{ for all } \psi \in \mathrm{C}(\bar\Omega). \tag{2.1}$$

In particular, these two conditions characterise Young measures. We define:

Definition 2.5. A parametrized measure $\boldsymbol{\nu} = \big((\nu_x)_{x\in\Omega},\, \lambda,\, (\nu_x^\infty)_{x\in\bar\Omega}\big)$ is said to be a *Young measure* (or *generalized Young measure*) whenever

(a) $(\nu_x)_{x\in\Omega} \subset \mathcal{M}_1^+(\mathbb{R}^N)$ is weakly-* $\mathcal{L}^n$-measurable (the *oscillation measure*).

(b) $\lambda \in \mathcal{M}^+(\bar\Omega)$ (the *concentration measure*).

(c) $(\nu_x^\infty)_{x\in\bar\Omega} \subset \mathcal{M}_1^+(\mathbb{R}^N)$ is weakly-* λ-measurable (the *concentration-angle measure*).

(d) $\int_\Omega \int_{\mathbb{R}^N} |z|^p \mathrm{d}\nu_x(z)\mathrm{d}x < \infty$ (the *moment condition* holds).

Then $\boldsymbol{\nu}$ acts linearly on $\mathbb{E}_p(\Omega, \mathbb{R}^N)$ via

$$\langle\langle \boldsymbol{\nu}, f\rangle\rangle := \int_\Omega \int_{\mathbb{R}^N} f(x,\bullet)\mathrm{d}\nu_x \mathrm{d}x + \int_{\bar\Omega}\int_{S^N} f_p^\infty(x,\bullet)\mathrm{d}\nu_x^\infty \mathrm{d}\lambda(x), \quad \text{for } f \in \mathbb{E}_p(\Omega,\mathbb{R}^N).$$

We write $\mathrm{Y}^p(\Omega, \mathbb{R}^N)$ for the set of all such $\boldsymbol{\nu}$.

Moreover, the Young measures lie in $\mathbb{E}(\Omega,\mathbb{R}^N)^*$ and, more precisely, the inclusion $\mathrm{Y}^p(\Omega,\mathbb{R}^N) \subset \mathbb{E}(\Omega,\mathbb{R}^N)^*$ is strict. The theorem below can be seen as the fundamental theorem of Young measures.

Theorem 2.6. *We have that*

$$\mathrm{Y}^p(\Omega,\mathbb{R}^N) = T^*\{\sigma \in \mathcal{M}^+(\overline{\Omega\times B^N})\colon \text{ equation (2.1) holds}\}.$$

Using the above characterisation it has been proved that the space of Young measures $\mathrm{Y}^p(\Omega,\mathbb{R}^N)$ has the following closedness and compactness properties:

Theorem 2.7 (Compactness of Young measures)**.** *Let (φ_j) be a bounded sequence in $L^p(\Omega,\mathbb{R}^N)$. Then there exists $\boldsymbol{\nu} \in \mathrm{Y}^p(\Omega,\mathbb{R}^N)$ such that, along a subsequence, $\varepsilon_{\varphi_j} \overset{*}{\rightharpoonup} \boldsymbol{\nu}$ in $\mathbb{E}_p(\Omega,\mathbb{R}^N)^*$, i.e.,*

$$\lim_{j\to\infty}\int_\Omega f(x,\varphi_j(x))dx = \int_\Omega\int_{\mathbb{R}^N} f(x,z)\mathrm{d}\nu_x(z)\mathrm{d}x + \int_{\bar\Omega}\int_{S^N} f_p^\infty(x,z)\mathrm{d}\nu_x^\infty(z)\mathrm{d}\lambda(x), \tag{2.2}$$

for all $f \in \mathbb{E}_p(\Omega, \mathbb{R}^N)$*. In this case, we say that the sequence* (φ_j) generates the Young measure $\boldsymbol{\nu}$*. In addition to this,* $\mathrm{Y}^p(\Omega, \mathbb{R}^N)$ *is convex and weakly-$*$ closed in* $\mathbb{E}_p(\Omega, \mathbb{R}^N)^*$*.*

The above result holds also in the case $p = 1$, where we consider bounded sequences of measures in $\mathcal{M}(\Omega, \mathbb{R}^N)$. In our analysis in Chapter 5, we use the following consequence of Theorem 2.7:

Lemma 2.4. *Let* $(\mu_j) \subset \mathcal{M}(\bar{\Omega}, \mathbb{R}^N)$ *generate a Young measure* $\boldsymbol{\nu}$*. Then*

$$\mu_j \overset{*}{\rightharpoonup} \bar{\nu}_x \mathcal{L}^n \llcorner \Omega + \bar{\nu}_x^\infty \lambda \text{ in } \mathcal{M}(\bar{\Omega}, \mathbb{R}^N).$$

The limit measure is refered to as the barycentre *of* $\boldsymbol{\nu}$*.*

This follows simply by taking $f(x, z) = \psi(x) z_i$ for $\psi \in \mathrm{C}(\bar{\Omega})$ in Theorem 2.7, where we write

$$\bar{\nu}_x = \int_{\mathbb{R}^N} z \mathrm{d}\nu_x(z) \quad \text{and} \quad \bar{\nu}_x^\infty = \int_{B^N} z \mathrm{d}\nu_x^\infty(z)$$

for the expectations of the probability measures ν_x and ν_x^∞.

We next formalise an idea which helps us distinguish between the case of sequences produce only oscillations and those that produce also concentrations. In a way, as it becomes obvious from the result below, the recession integrand captures the concentration effects of the generating sequence and hence we expect the lack of it in the presence of suitable equiintegrability assumptions.

Proposition 2.1. *Let* $\Omega \subset \mathbb{R}^d$ *be bounded and open with* $\mathcal{L}^d(\partial\Omega) = 0$*,* $p \in [1, \infty)$*, and* $f\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ *be a measurable integrand such that* $f(x, \bullet)$ *is continuous for almost every* $x \in \Omega$ *(is a Carathéodory integrand). Let* $(\varphi_j) \subset L^p(\Omega, \mathbb{R}^N)$ *generate* $\boldsymbol{\nu} \in \mathrm{Y}^p(\Omega, \mathbb{R}^N)$ *be such that* $(f(\bullet, \varphi_j))_j$ *is p-equintegrable. Then*

$$\lim_{j \to \infty} \int_\Omega f(x, \varphi_j(x)) \mathrm{d}x = \int_\Omega \langle \nu_x, f(x, \bullet) \rangle \mathrm{d}x.$$

In particular, the barycentre $\bar{\nu}_x := \langle \nu_x, id \rangle$ *of the oscillating part of the generated Young measure* $\boldsymbol{\nu}$ *identifies the weak limit of the sequence* (φ_j)*, i.e.*

$$\varphi_j \rightharpoonup \bar{\nu}_x \quad \text{in } L^p(\Omega).$$

In Chapter 5, we focus on the case $p = 1$ and in particular we study Carathéodory integrands $f : \Omega \times \mathbb{R}^N \to \mathbb{R}$ with linear growth at infinity i.e.

$$f \text{ is } \mathcal{L}^d \times \mathcal{B}(\mathbb{R}^N)\text{-measurable such that } z \mapsto f(\cdot, z) \text{ continuous and } |f(\cdot, z)| \lesssim 1 + |z|. \tag{2.3}$$

More precisely we are interested in a smaller class of Carathéodory functions which is defined as

$$\mathbb{L}(\Omega, \mathbb{R}^N) := \{ f : \Omega \times \mathbb{R}^N \to \mathbb{R} : f \text{ satisfies (2.3) and } f^\infty \in C(\bar{\Omega} \times \mathbb{R}^N) \text{ exists} \},$$

and we want to study the limiting behaviour of the sequence ($\langle\langle \mu_j, f \rangle\rangle$) for $f \in \mathbb{L}(\Omega, \mathbb{R}^N)$ and bounded sequences of measures (μ_j). From the above discussion, we understand that the statement "(μ_j) generates $\boldsymbol{\nu} \in \mathrm{Y}(\Omega, \mathbb{R}^N)$" only implies the convergence $\langle\langle \mu_j, f \rangle\rangle \to \langle\langle \boldsymbol{\nu}, f \rangle\rangle$ for $f \in \mathbb{E}(\Omega, \mathbb{R}^N)$. However, the next result reveals that this convergence holds also for the class $\mathbb{L}(\Omega, \mathbb{R}^N)$.

Proposition 2.2. *[82, Proposition 2(i)] Let $\Omega \subset \mathbb{R}^d$ be bounded and open, and $f \in \mathbb{L}(\Omega, \mathbb{R}^N)$. Let $(\mu_j) \subset \mathcal{M}(\Omega, \mathbb{R}^N)$ generate $\boldsymbol{\nu} \in \mathrm{Y}(\Omega, \mathbb{R}^N)$. Then*

$$\lim_{j \to \infty} \int_\Omega f(x, \mathrm{d}\mu_j(x)) = \int_\Omega \langle \nu_x, f(x, \bullet) \rangle \mathrm{d}x + \int_{\bar{\Omega}} \langle \nu_x^\infty, f^\infty(x, \bullet) \rangle \mathrm{d}\lambda(x).$$

The idea behind the proof of the above result is to decompose $f \in \mathbb{L}(\Omega, \mathbb{R}^N)$ as $f(x,z) = h(x,z) + g(x,z)$ where

$$h(x,z) := f(x,z) - |z| f^\infty \left(x, \frac{z}{|z|} \right), \quad \text{and,} \quad g(x,z) := |z| f^\infty \left(x, \frac{z}{|z|} \right).$$

Now it is not hard to check that, due to the continuity and the 1-homogeneity of f^∞, $g \in \mathbb{E}(\Omega, \mathbb{R}^N)$ and hence $\langle\langle \mu_j, g \rangle\rangle \to \langle\langle \boldsymbol{\nu}, g \rangle\rangle = \langle\langle \boldsymbol{\nu}, f^\infty \rangle\rangle$. Regarding h, the observation that $h^\infty = f^\infty - f^\infty = 0$ allows us to deduce that $\langle\langle \mu_j, h \rangle\rangle \to \langle\langle \boldsymbol{\nu}, h \rangle\rangle = \langle\langle \boldsymbol{\nu}, f - f^\infty \rangle\rangle$, and hence conclude by putting the two convergences together. We refer the reader to [82, Proposotion 2(i)] for the detailed proof.

Before we proceed, we present some examples which help in the understanding of the difference between the oscillation and the concentration effects of the generating sequences, see [81] for the details of the corresponding proofs and more complex examples.

Example 1. (No concentration) Let $Q = (0,1)^n$ and $\varphi \in L^p_{\mathrm{loc}}(\mathbb{R}^d, \mathbb{R}^N)$ be Q-periodic. Set $\varphi_j(x) := \varphi(jx)$, so that, by using Riemann sums, we obtain that

$$\int_Q \psi(x) F(\varphi_j(x)) dx \to \int_Q \psi(x) F(\varphi(x)) dx,$$

for $\psi \in C_c(Q)$, $F \in C_c(\mathbb{R}^N)$. However, by using the formula (2.2) together with the fact that for these particular test functions $F_p^\infty = 0$, we get that $\nu_x = \varphi_\#(\mathcal{L}^n \lfloor Q)$. Now, by testing (2.2) with $\psi \otimes |\cdot|^p$ where $\psi \in C(\bar{Q})$, we see that the concentration measure λ is given by the weak-$*$ limit of the sequence $(|\varphi_j|^p - \langle \nu_x, |\cdot|^p \rangle)$. Since we already found the oscillation measure ν_x, we can conclude that $\lambda = 0$.

Example 2. (No oscillation) Consider now $Q = (-1, 1)$ and $\varphi_j = j\mathbb{1}_{(0,1/j)}$, where $\mathbb{1}_A$ denotes the characteristic function of the Borel set $A \subseteq Q$. Then again as in the previous example, for $\psi \in C_c(Q)$, $F \in C_c(\mathbb{R})$, we have that

$$\int_{-1}^{1} \psi(x)F(\varphi_j(x)) = \int_{-1}^{0} + \int_{1/j}^{1} \psi(x)F(0)dx + \int_{0}^{1/j} \psi(x)F(j)dx \to \int_{-1}^{1} \psi(x)F(0)dx,$$

i.e. $\nu_x = \delta_0$. Since we found the oscillation part, similarly with the previous example, considering the fact that $\varphi_j \mathcal{L}^n \lfloor_{(-1,1)} \overset{*}{\rightharpoonup} \delta_0$ in $\mathcal{M}([-1, 1])$, we infer that $\lambda = \delta_0$. Regarding the concentration angles, since the measure ν_x^∞ concentrates on the sphere, we know that $\nu_x^\infty = \alpha\delta_{-1} + \beta\delta_1$ for some α, $\beta > 0$ such that $\alpha + \beta = 1$. So, for

$$F(z) = \begin{cases} A, & z = -1 \\ B, & z = 1 \end{cases} \quad \text{so } \tilde{F}(z) = \begin{cases} -Az, & z \leqslant 0 \\ Bz & z > 0, \end{cases}$$

and $\psi \in C([-1, 1])$, if we calculate the limit we get that

$$\int_{-1}^{1} \psi(x)\tilde{F}(\varphi_j(x))dx \to \psi(0)B.$$

By using now the Young measure representation (2.2) and comparing the two limits we deduce that $\alpha = 0$ and so $\nu_0^\infty = \delta_1$. Remember that (2.2) gives us that

$$\int_{-1}^{1} \psi(x)\tilde{F}(\varphi_j(x)) \to \int_{-1}^{1} \psi(x)\langle \nu_x, \tilde{F}\rangle + \int_{-1}^{1} \psi(x)\langle \nu_x^\infty, F\rangle d\lambda = \psi(0)(\alpha A + \beta B).$$

Hence, to summarize we showed that for the weakly-$*$ convergent sequence of measures $\left(\varphi_j \mathcal{L}^n \lfloor_{(-1,1)}\right)$ it holds that

$$\varepsilon_{\varphi_j} \overset{*}{\rightharpoonup} \left((\delta_0)_{x\in(-1,1)}, \delta_0, (\delta_1)_{x\in[-1,1]}\right), \text{ in } \mathbb{E}_1(-1, 1)^*.$$

In the L^p case, where $1 < p < \infty$, we set $\tilde{\varphi}_j(x) = j^{1/p}\mathbb{1}_{(0,1/j)}$ and we see, similarly with the $p = 1$ case, that

$$\varepsilon_{v_j} \overset{*}{\rightharpoonup} \left((\delta_0)_{x\in(-1,1)}, \delta_0, (\delta_1)_{x\in[-1,1]}\right), \text{ in } \mathbb{E}_p(-1, 1)^*.$$

The only difference between the two cases is that in the case $1 < p < \infty$, the concentration is not visible in the weak limit, $\tilde{\varphi}_j \rightharpoonup 0$ in $L^p(-1, 1)$.

2.3.1 Jensen-type inequalities

One of the main interests of this thesis is the study of functionals of the form

$$\mathcal{F}[\varphi] := \int_\Omega f(x, \varphi(x))dx, \tag{2.4}$$

whenever the function φ is constrained by a constant rank differential operator $\mathcal{A}$, i.e. $\mathcal{A}\varphi = 0$. Young measures reveal the link between the lower semicontinuity of (2.4) and the convexity of the associated integrand $f(x,\cdot)$. One of the main contribution of Young measure theory is that they make sense to the limits $\mathcal{F}[\varphi_j]$, whenever f satisfies suitable growth conditions.

Let us first consider $f \in \mathbb{E}_p(\Omega,\mathbb{R}^N)$ and PDE-constrained, p-equiintegrable sequences $\varphi_j \rightharpoonup \varphi$ in $L^p_{\mathcal{A}}(\Omega,\mathbb{R}^N)$. Then, in view of Theorem 2.7, the question of weakly lower semicontinuity of $\mathcal{F}$ reduces to a question of some sort of convexity for fixed $x \in \Omega$ via the following:

$$\liminf_{j\to\infty} \mathcal{F}[\varphi_j] = \int_\Omega \langle \nu_x, f(x,\cdot)\rangle dx \geqslant \int_\Omega f(x,\langle \nu_x, \mathrm{id}\rangle)dx = \int_\Omega f(x,\varphi(x))dx = \mathcal{F}[\varphi].$$

The missing element which makes the above relation true, is the so-called Jensen-type inequality

$$\langle \nu_x, f(x,\cdot)\rangle \geqslant f(x,\langle \nu_x, \mathrm{id}\rangle), \qquad \text{for } \mathcal{L}^d\text{-a.e. } x \in \Omega. \tag{2.5}$$

As a first step to the understanding of the above inequality, it turns out that:

Theorem 2.8. *Let $f : \Omega \times \mathbb{R}^N$ be a Carathéodory function such that the mapping $z \mapsto f(x,z)$ is convex for $\mathcal{L}^d$-a.e. $x \in \Omega$. Then, it holds that*

$$\int_{\mathbb{R}^N} f(x,z)d\mu(z) \geqslant f\left(x, \int_{\mathbb{R}^N} z\, d\mu(z)\right), \qquad \text{for all } \mu \in \mathcal{M}^+_1(\mathbb{R}^N),$$

and $\mathcal{L}^d$-a.e. $x \in \Omega$.

Even though the lower semicontinuity of $\mathcal{F}$ is equivalent to the convexity of $f(x,\cdot)$ in the unconstrained case $\mathcal{A} \equiv 0$, this is not the case when PDE-constraints are present. In particular, for the case $\mathcal{A} = \mathrm{curl}$ and hence, when Young measures generated by a sequence of gradients are considered, quasiconvexity suffices for weak lower semicontinuity [71, 72, 77, 100]. The extension of the latter to constraints given by constant rank operators is due to Fonseca and Müller [53]:

Theorem 2.9. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $1 < p < \infty$, and let $\mathcal{A}$ be defined by* (2.1) *and satisfy the constant rank assumption* (2.2)*. Then, ν is generated by a p-equiintegrable sequence $\varphi_j \in L^p_{\mathcal{A}}(\Omega,\mathbb{R}^N)$ if and only if*

(a) there exists $\varphi \in L^p_{\mathcal{A}}(\Omega,\mathbb{R}^N)$ such that $\varphi(x) = \langle \nu_x, \mathrm{id}\rangle$ for $\mathcal{L}^d$-a.e. $x \in \Omega$;

(b) ν has finite p-th moment, i.e. $\int_\Omega \langle \nu, |\cdot|^p\rangle dx < \infty$;

(c) the Jensen-type inequality

$$\langle \nu_x, f(\cdot)\rangle \geqslant f(\langle \nu_x, \mathrm{id}\rangle), \qquad \text{for } \mathcal{L}^d\text{-a.e. } x \in \Omega,$$

holds for all $\mathcal{A}$-quasiconvex $f \in C(\mathbb{R}^N)$ with p-growth, i.e. $|f(\cdot)| \lesssim 1 + |\cdot|^p$.

Case $p = 1$

In contrast with the case where $p > 1$, in the case $p = 1$ bounded sequences in L^1 do not necessarily have a weakly convergent subsequence and thus, as we saw in the previous subsection, in the case of linear growth integrands it would be more natural from the compactness point of view to consider weakly-$*$ convergent sequences of measures. In that case, the barycentre of ν, which due to Lemma 2.4 is

$$\bar{\nu} := \bar{\nu}_x \mathcal{L}^n \lfloor \Omega + \bar{\nu}_x^\infty \lambda,$$

retains information from the concentration part, something that technically can be seen from the fact that for $f(x, z) := \phi(x)z$ in Theorem 2.7, we have that $f_1^\infty = f \neq 0 = f_p^\infty$ where $p > 1$.

The analysis behind the study of weak lower semicontinuity in the presence of PDE-constraints for $p = 1$ and linear growth integrands is much more complicated, and its theory is based on very recent and powerful results that lead beyond the scope of this thesis, see [3, 82, 94].

Chapter 3

Gårding-type inequalities for quasiconvex integrands

3.1 Quasiconvexity

As we already discussed in the introduction, motivated by applications arising from materials science, we study problems where the solution is constrained by a constant rank differential operator $\mathcal{A}$, while we also consider the case where part of the solution remains unconstrained, and so in that case the constraints are enforced via operators of the form $(\mathcal{A}, 0)$. To this end, we study quasiconvex integrands where our notion of quasiconvexity at least guarantees convexity on the directions of the associated wave cone $\Lambda_{(\mathcal{A},0)}$. Even though our definition of quasiconvexity is associated with the operator $(\mathcal{A}, 0)$, we call it just quasiconvexity to distinguish it from the typical definition of $(\mathcal{A}, 0)$ quasiconvexity, see the discussion below Definition 3.1.

In particular, we say that a continuous function e is quasiconvex at $(\lambda_1, \lambda_2) \in \mathbb{R}^N \times \mathbb{R}$ if the inequality

$$\int_Q e(\lambda_1 + \varphi(x), \lambda_2 + \psi(x)) - e(\lambda_1, \lambda_2) - e_\eta(\lambda_1, \lambda_2)\psi(x)dx \geqslant 0,$$

holds for all $\psi \in C^\infty(Q)$ and all $\mathcal{A}$-free and zero-average $\varphi \in C^\infty(Q)$. The above definition, using the potential operator $\mathcal{B}$ of $\mathcal{A}$, can equivalently be expressed over arbitrary domains. In particular, for $\Omega \subseteq \mathbb{R}^d$ a non-empty open subset with $|\partial\Omega| = 0$, e is quasiconvex at $(\lambda_1, \lambda_2) \in \mathbb{R}^N \times \mathbb{R}$ if the inequality

$$\int_\Omega e(\lambda_1 + \mathcal{B}\phi(x), \lambda_2 + \psi(x)) - e(\lambda_1, \lambda_2) - e_\eta(\lambda_1, \lambda_2)\psi(x)dx \geqslant 0,$$

holds for all $\phi \in C_c^\infty(\Omega)$ and $\psi \in C^\infty(\Omega)$.

Henceforth, we assume that e has (p,q)-growth, i.e. $|e(z_1,z_2)| \leqslant c(1+|z_1|^p+|z_2|^q)$, and then by density we can also express the above definition with test functions in $W^{l,p}(Q)$ and $L^q(Q)$ or in $W_0^{l,p}(\Omega)$ and $L^q(\Omega)$ respectively.

The results presented in this chapter require a strengthened version of quasiconvexity which we now introduce. Remember that the auxiliary function $V_i : \mathbb{R}^k \to \mathbb{R}$ has been defined by

$$V_i(z) := (|z|^i + |z|^2)^{1/2}, \tag{3.1}$$

where $k=1$, $k=d$ or $k=N$ and $i \in \mathbb{N}$.

Definition 3.1. Let $\Omega \subseteq \mathbb{R}^d$ be a non-empty open subset with $|\partial\Omega| = 0$. A continuous function e is strongly quasiconvex at $(\lambda_1,\lambda_2) \in \mathbb{R}^N \times \mathbb{R}$ if the inequality

$$\int_\Omega e(\lambda_1+\mathcal{B}\phi, \lambda_2+\psi) - e(\lambda_1,\lambda_2) - e_\eta(\lambda_1,\lambda_2)\psi \geqslant c_0 \int_\Omega |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2,$$

holds for all $\phi \in W_0^{l,p}(\Omega)$ and $\psi \in L^q(\Omega)$. Equivalently, e is strongly quasiconvex at $(\lambda_1,\lambda_2) \in \mathbb{R}^N \times \mathbb{R}$ if

$$\int_Q e(\lambda_1+\mathcal{B}\phi, \lambda_2+\psi) - e(\lambda_1,\lambda_2) - e_\eta(\lambda_1,\lambda_2)\psi \geqslant c_0 \int_Q |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2,$$

for all $\phi \in W^{l,p}(Q)$ and $\psi \in L^q(Q)$. In addition, we say that e is (strongly) quasiconvex if it is (strongly) quasiconvex at (λ_1,λ_2) for all $(\lambda_1,\lambda_2) \in \mathbb{R}^N \times \mathbb{R}$.

At first sight and as we already mentioned above, our definition of quasiconvexity, Definition 3.1, differs from the classical definition of $(\mathcal{A},0)$-quasiconvexity associated with the cone $\Lambda_{(\mathcal{A},0)}$. We remark that, with respect to the latter definition, e is strongly $(\mathcal{A},0)$-quasiconvex if

$$\int_\Omega e(\lambda_1+\mathcal{B}\phi, \lambda_2+\psi) - e(\lambda_1,\lambda_2) \geqslant \tilde{c_0} \int_\Omega |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2, \tag{3.2}$$

for all $\phi \in W_0^{l,p}(\Omega)$ and $\psi \in L^q(\Omega)$ with $\int_\Omega \psi = 0$. In fact, as it can be seen from the corollary below, the two definitions are in a sense equivalent (the one direction is straightforward).

Corollary 3.1. *Let $e \in C^1(\mathbb{R}^N \times \mathbb{R})$ with (p,q)-growth i.e. $|e(z_1,z_2)| \lesssim 1+|z_1|^p+|z_2|^q$. Then, if e is strongly $(\mathcal{A},0)$-quasiconvex, there exists a constant $c_0>0$ (smaller than $\tilde{c_0}$) such that*

$$\int_\Omega e(\lambda_1+\mathcal{B}\phi, \lambda_2+\psi) - e(\lambda_1,\lambda_2) - e_\eta(\lambda_1,\lambda_2)\psi \geqslant c_0 \int_\Omega |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2 \tag{3.3}$$

for all $\phi \in W_0^{l,p}(\Omega)$ and $\psi \in L^q(\Omega)$, i.e. e is strongly quasiconvex.

Proof. Let $(\lambda_1,\lambda_2)\in\mathbb{R}^N\times\mathbb{R}$, $\phi\in W^{l,p}_0(\Omega)$, $\psi\in L^q(\Omega)$ and for $\psi|_\Omega := \fint_\Omega \psi$ we infer that

$$\begin{aligned}
&\int_\Omega e(\lambda_1+\mathcal{B}\phi,\lambda_2+\psi)-e(\lambda_1,\lambda_2)-e_\eta(\lambda_1,\lambda_2)\psi\\
&=\int_\Omega e(\lambda_1+\mathcal{B}\phi,\lambda_2+\psi|_\Omega+\psi-\psi|_\Omega)-e(\lambda_1,\lambda_2+\psi|_\Omega)\\
&+\int_\Omega e(\lambda_1,\lambda_2+\psi|_\Omega)-e(\lambda_1,\lambda_2)-e_\eta(\lambda_1,\lambda_2)\psi|_\Omega\\
&-\int_\Omega e_\eta(\lambda_1,\lambda_2)\psi+e_\eta(\lambda_1,\lambda_2)\psi|_\Omega := I_1+I_2+I_3.
\end{aligned}$$

From the $(\mathcal{A},0)$-quasiconvexity of e, see (3.2), we deduce that

$$I_1 \geqslant \tilde{c_0}\int_\Omega |V_p(\mathcal{B}\phi)|^2+|V_q(\psi-\psi|_\Omega)|^2.$$

Concerning the second term, we claim that

$$I_2 \geqslant \tilde{c_0}\int_Q |V_q(\psi|_\Omega)|^2, \tag{3.4}$$

and taking this into account, since $I_3=0$, we infer that

$$\begin{aligned}
\int_\Omega e(\lambda_1+\mathcal{B}\phi,\lambda_2+\psi)-e(\lambda_1,\lambda_2)-e_\eta(\lambda_1,\lambda_2)\psi \geqslant \tilde{c_0}\int_\Omega |V_p(\mathcal{B}\phi)|^2+|V_q(\psi-\psi|_\Omega)|^2\\
+\tilde{c_0}\int_\Omega |V_q(\psi|_\Omega)|^2.
\end{aligned} \tag{3.5}$$

Due to the convexity of the function $|\cdot|^q$ for $r\geqslant 2$, we have that

$$|\psi-\psi|_\Omega|^r+|\psi|_\Omega|^r \geqslant 2^{1-r}|\psi|^r,$$

and so, applying this into (3.5) for $r=2$ and $r=q$, we conclude the proof for $c_0 := \tilde{c_0}/2^{1-q}$.

It remains to prove our claim, inequality (3.4). In particular, we show that (3.2) implies that

$$e(\lambda_1,\beta)-e(\lambda_1,\alpha)-e_\eta(z_1,\alpha)(\beta-\alpha)\geqslant c_0|\alpha-\beta|^2\left(1+|\alpha-\beta|^{q-2}\right),$$

for all $z_1\in\mathbb{R}^N$ and α, $\beta\in\mathbb{R}$. To this end, for $\lambda\in(0,1)$ and $\alpha,\beta\in\mathbb{R}$ let

$$\psi(x)=\begin{cases}-(1-\lambda)(\alpha-\beta) & x\in\Omega_\alpha,\\ \lambda(\alpha-\beta) & x\in\Omega_\beta,\end{cases}$$

where Ω_α and Ω_β are two disjoint cubes such that $\mathcal{L}^d(\Omega_\alpha)=\lambda\mathcal{L}^d(\Omega)$ and $\mathcal{L}^d(\Omega_\beta)=(1-\lambda)\mathcal{L}^d(\Omega)$. Then, for $\phi=0$ and $\lambda_2:=(1-\lambda)\alpha+\lambda\beta$ in (3.2), we infer that

$$\begin{aligned}
c_0|\alpha-\beta|^2\lambda(1-\lambda)\Big[1+|\alpha-\beta|^{q-2}\big((1-\lambda)^{q-1}+\lambda^{q-1}\big)\Big]+e\big(\lambda_1,(1-\lambda)\alpha+\lambda\beta)\big)\\
\leqslant \lambda e(\lambda_1,\beta)+(1-\lambda)e(\lambda_1,\alpha).
\end{aligned}$$

Using the latter inequality, for $\beta_\lambda := (1-\lambda)\alpha + \lambda\beta$, we infer that

$$e(\lambda_1, \beta) = e\left(\lambda_1, \beta_\lambda + \frac{1-\lambda}{\lambda}(\beta_\lambda - \alpha)\right) \geqslant e(\lambda_1, \beta_\lambda) + \frac{1-\lambda}{\lambda}\left(e(\lambda_1, \beta_\lambda) - e(\lambda_1, \alpha)\right)$$
$$+ c_0|\alpha - \beta|^2(1-\lambda)\left[1 + |\alpha-\beta|^{q-2}\big((1-\lambda)^{q-1} + \lambda^{q-1}\big)\right].$$

Hence, since

$$\lim_{\lambda\to 0}\frac{1}{\lambda}\left(e(\lambda_1, \beta_\lambda) - e(\lambda_1, \alpha)\right) = \lim_{\lambda\to 0}\frac{1}{\lambda}\left[e(\lambda_1, \alpha + \lambda(\beta-\alpha)) - e(\lambda_1, \alpha)\right] = e_\eta(\lambda_1, \alpha)(\beta - \alpha),$$

we take the limit $\lambda \to 0^+$ to conclude the proof of the claim. $\square$

3.2 Decomposition Lemma

The proof of Theorem 3.2 is based on a decomposition lemma which splits a weakly converging sequence into $\mathcal{A}$-free oscillating and concentrating parts. This extends [31, Theorem 3.4] for the operator $\mathcal{B}$, rather than ∇, and finds its origins in the decomposition results of Kristensen [78], and Fonseca and Müller [52]. The former of these results is based on the Helmholtz Decomposition, a version of which in the $\mathcal{A}$-free setting can be found in [63]. Below, we instead use the construction of Fonseca and Müller [52, Lemma 2.14] but follow the structure of proof found in [31] to help the reader understand the connection and differences between the curl-free and $\mathcal{A}$-free cases.

Below we present a crucial result of Fonseca and Müller [52, Lemma 2.14] in which the constant rank property is essential and cannot be avoided.

Lemma 3.1. *Let $\mathcal{A}$ as in §2.1. For every $1 < p < +\infty$, there exists a linear and continuous projection operator $\mathcal{P} : L^p(Q) \to L^p(Q)$ and $C > 0$ such that*

$$\mathcal{A}(\mathcal{P}v) = 0, \quad \int_Q \mathcal{P}v = 0 \quad \textit{and} \quad \|v - \mathcal{P}v\|_{L^p(Q)} \leqslant C\|\mathcal{A}v\|_{W^{-l,p}(Q)},$$

for all $v \in L^p(Q)$ with $\int_Q v = 0$.

To reduce the number of indices in the proof of Lemma 3.2 we assume that the operator $\mathcal{A}$ has order 1 and its potential operator $\mathcal{B}$ has order $l \geqslant 1$. Nevertheless, the result holds in the general case where the operator $\mathcal{A}$ has order $k \geqslant 1$ and the proof remains essentially the same.

Lemma 3.2. *Let $2 \leqslant p < +\infty$ and $(\phi_j)_j \subset W^{l,2}(Q)$ such that $\mathcal{B}\phi_j \rightharpoonup \mathcal{B}\phi$ in $L^2(Q)$. Let also $(r_j)_j \subset (0,1)$ such that $(r_j\mathcal{B}\phi_j)_j$ bounded in $L^p(Q)$. Then, up to a subsequence, there exist sequences $(f_j)_j \subseteq W^{l,2}(Q)$ and $(b_j)_j \subseteq W^{l,2}(Q)$ such that*

(1) $\mathcal{B}f_j \rightharpoonup 0$ *and* $\mathcal{B}b_j \rightharpoonup 0$*;*

(2) $(|\mathcal{B}f_j|^2)_j$ *is equiintegrable;*

(3) $\mathcal{B}b_j \to 0$ *in measure;*

(4) $\mathcal{B}\phi_j = \mathcal{B}\phi + \mathcal{B}f_j + \mathcal{B}b_j$.

In addition, for a further subsequence, $(f_j)_j$ *and* $(b_j)_j$ *can be chosen so that*

(1′) $r_j\mathcal{B}f_j \rightharpoonup 0$ *and* $r_j\mathcal{B}b_j \rightharpoonup 0$ *in* $L^p(Q)$*;*

(2′) $(|r_j\mathcal{B}f_j|^p)_j$ *is equiintegrable;*

(3′) $r_j\mathcal{B}b_j \to 0$ *in measure.*

Proof. By extracting a subsequence, we may assume that $\mathcal{B}\phi_j \xrightarrow{Y} (\nu_x)_x$ and $r_j\mathcal{B}\phi_j \xrightarrow{Y} (\mu_x)_x$. The latter notation means that the sequences generate the respective Young measures and in particular that

$$G(\mathcal{B}\phi_j) \rightharpoonup \langle \nu_x, G\rangle = \int_{\mathbb{R}^N} G(z)\, d\nu_x(z) \text{ in } L^1(Q),$$

whenever $(G(\mathcal{B}\phi_j))$ is equiintegrable, see Section 2.3 for details on Young measures. We also observe that, by working with the sequence $\phi_j - \phi$ instead of ϕ_j, we may assume that $\phi = 0$. We split the proof into 4 steps.

<u>Step 1. Truncation</u>: Define, for $k \in \mathbb{N}$ and $z \in \mathbb{R}^N$, the truncation operator τ_k by

$$\tau_k(z) := \begin{cases} z, & |z| \leqslant k, \\ k\, z/|z|, & |z| > k. \end{cases}$$

Its straightforward to see that for fixed $k \in \mathbb{N}$ the sequence $\big(|\tau_k(\mathcal{B}\phi_j)|^2\big)_j$ is uniformly integrable, and hence, by Proposition 2.1 and the Monotone Convergence Theorem, we have that

$$\lim_{k\to\infty}\lim_{j\to\infty}\int_Q |\tau_k(\mathcal{B}\phi_j)|^2 dx = \lim_{k\to\infty}\int_Q \langle \nu_x, |\tau_k(\cdot)|^2\rangle dx = \int_Q \langle \nu_x, |\cdot|^2\rangle dx.$$

Moreover, if $1 \leqslant q < 2$, then

$$\begin{aligned}\int_Q |\tau_k(\mathcal{B}\phi_j) - \mathcal{B}\phi_j|^q = \int_{\{|\mathcal{B}\phi_j|>k\}} \left|1 - \frac{k}{|\mathcal{B}\phi_j|}\right|^q |\mathcal{B}\phi_j|^q &\leqslant \int_{\{|\mathcal{B}\phi_j|>k\}} 2^q \frac{|\mathcal{B}\phi_j|^2}{|\mathcal{B}\phi_j|^{2-q}} \\ &\leqslant k^{2-q} 2^q \int_Q |\mathcal{B}\phi_j|^2 dx \to 0,\end{aligned}$$

as $k \to \infty$. Summing up the above results, we may find a subsequence such that

$$\lim_{k\to\infty}\int_Q |\tau_k(\mathcal{B}\phi_{j_k})|^2 = \int_Q \langle |.|^2, \nu_x\rangle, \tag{3.1}$$

$$\lim_{k\to\infty}\int_Q |\tau_k(\mathcal{B}\phi_{j_k}) - \mathcal{B}\phi_{j_k}|^q = 0, \tag{3.2}$$

for $1 \leqslant q < 2$. Letting $v_k := \tau_k(\mathcal{B}\phi_{j_k})$, it then follows from (3.1), (3.2) that $(v_k)_k$ is 2-equiintegrable and generates $(\nu_x)_x$. From (3.2) and the continuity of the operator $\mathcal{A}$, it also follows that $\mathcal{A}v_k \to 0$ in $W^{-1,q}(Q)$.

Step 2. Decomposition: Since $v_k \in L^2(Q)$, we can extend it periodically to $\mathbb{R}^d$ and then apply Lemma 3.1 to infer that

$$v_k - \int_Q v_k = F_k + B_k$$

where $F_k := \mathcal{P}\Big(v_k - \int_Q v_k\Big)$, $B_k := v_k - \int_Q v_k - \mathcal{P}\Big(v_k - \int_Q v_k\Big)$.

Claim 1: $B_k \to 0$ in measure.

By Lemma 3.1 we infer that

$$\|B_k\|_{L^q(Q)} = \|v_k - \int_Q v_k - \mathcal{P}\Big(v_k - \int_Q v_k\Big)\|_{L^q(Q)} \leqslant C\|\mathcal{A}v_k\|_{W^{-1,q}(Q)} \to 0$$

for all $1 \leqslant q < 2$. Hence, $B_k \to 0$ in $L^q(Q)$ and so in measure.

Claim 2: $(|F_k|^2)_k$ is equiintegrable.

By Step 1, $\Big(v_k - \int_Q v_k\Big)_k$ is 2-equiintegrable, and hence for every $\varepsilon > 0$ and all $q > 2$ there exists a sequence $(W_k)_k$ such that

$$\|v_k - \int_Q v_k - W_k\|_{L^2(Q)} \leqslant \varepsilon/C$$

and $\sup_k \|W_k\|_{L^q(Q)} < +\infty$. This is an equivalent characterisation of equiintegrability, see [78]. Taking into account the properties of the projection $\mathcal{P}$, we infer that

$$\|F_k - \mathcal{P}(W_k)\|_{L^2} = \|\mathcal{P}\Big(v_k - \int_Q v_k - W_k\Big)\|_{L^2} \leqslant C\|v_k - \int_Q v_k - W_k\|_{L^2} \leqslant \varepsilon$$

and

$$\sup_k \|\mathcal{P}(W_k)\|_{L^q} \leqslant C \sup_k \|W_k\|_{L^q} < +\infty.$$

This concludes the proof of Claim 2.

Claim 3: F_k, $B_k \rightharpoonup 0$ in $L^2(Q)$.

Since $\mathcal{B}\phi_{j_k}$ has zero average, (3.2) and Claim 2 imply that

$$F_k - \mathcal{B}\phi_{j_k} = v_k - \int_Q v_k - \mathcal{B}\phi_{j_k} - B_k = v_k - \mathcal{B}\phi_{j_k} - \int_Q (v_k - \mathcal{B}\phi_{j_k}) - B_k \to 0$$

in measure. In addition, by (3.1), v_k is bounded in $L^2(Q)$ and by the continuity of $\mathcal{P}$, $(F_k)_k$ is also bounded in $L^2(Q)$ and $F_k - \mathcal{B}\phi_{j_k} \rightharpoonup 0$ in $L^2(Q)$. This proves the claim for

F_k, since $\mathcal{B}\phi_{j_k} \rightharpoonup 0$ in $L^2(Q)$. For $(B_k)_k$ the claim is immediate as it is bounded in $L^2(Q)$ and converges to 0 in measure.

Step 3. Concluding the L^2-decomposition: Since F_k is $\mathcal{A}$-free with zero average, from Lemma 2.1 (i), there exists a function $f_k \in W^{l,2}(Q)$ such that $F_k = \mathcal{B}f_k$. Set $b_k := \phi_{j_k} - f_k$. We thus conclude that

$$\mathcal{B}b_k = \mathcal{B}\phi_{j_k} - v_k + \int_Q v_k + B_k \to 0$$

in measure as, by Claim 1, $\mathcal{B}\phi_{j_k} - v_k \to 0$ in measure. Also, $\int_Q v_k \to 0$ since $\int_Q \mathcal{B}\phi_{j_k} = 0$ and (3.2) with $q = 1$, and $B_k \to 0$ by Claim 1. Thus,

$$\mathcal{B}\phi_{j_k} = \mathcal{B}f_k + \mathcal{B}b_k$$

satisfying (1)-(4).

Step 4. L^p-decomposition: This follows the arguments in [31] but we include it for completeness. Similarly to Step 1 we can extract a p-equiintegrable subsequence such that

$$\lim_{k\to\infty} \int_Q |\tau_k(r_{j_k}\mathcal{B}\phi_{j_k})|^p = \int_Q \langle |.|^p, \mu_x\rangle, \tag{3.3}$$

and with $v_k = \tau_k(\mathcal{B}\phi_{j_k})$, we infer that

$$|r_{j_k} v_k(x)| = |\tau_{kr_{j_k}}(r_{j_k}\mathcal{B}\phi_{j_k}(x))| \leqslant |\tau_k(r_{j_k}\mathcal{B}\phi_{j_k}(x))|,$$

since $r\tau_k(z) = \tau_{kr}(rz)$, $kr_{j_k} \leqslant k$ and $k \mapsto \tau_k(z)$ is non-decreasing in z. Hence, the sequence $(r_{j_k} v_k)_k$ is p-equiintegrable and bounded in $L^p(Q)$. From the linearity and continuity of the projection $\mathcal{P}$, we find that

$$\mathcal{P}\Big(r_{j_k} v_k - \int_Q r_{j_k} v_k\Big) = r_{j_k}\mathcal{P}\Big(v_k - \int_Q v_k\Big) = r_{j_k} F_k$$

and so $\|r_{j_k} F_k\|_{L^p(Q)} \lesssim \|r_{j_k} v_k\|_{L^p(Q)}$ which implies that the sequence $(r_{j_k} F_k)_k$ is also bounded in $L^p(Q)$. Hence, we can proceed as in Steps 2 and 3 and deduce that $r_{j_k}\mathcal{B}f_k$, $r_{j_k}\mathcal{B}b_k \rightharpoonup 0$ in $L^p(Q)$. Since $r_{j_k} \in (0,1)$, $(3')$ is a straightforward implication of (3). $\square$

Remark 3.1. We remark that the above decomposition applies to any $\mathcal{A}$-free and zero-average sequence $(\psi_j)_j \subseteq L^2(Q)$ with $\psi_j \rightharpoonup \psi$ in $L^2(Q)$. Indeed, by Lemma 2.1 (i), $\psi_j = \mathcal{B}\phi_j$, $\psi = \mathcal{B}\phi$ for some ϕ_j, $\phi \in W^{l,p}(Q)$. In addition, we can choose b_j to be a $\mathbb{B}^\dagger$-primitive and hence to satisfy the bounds of Lemma 2.1. Note that f_j is already chosen as a $\mathbb{B}^\dagger$-primitive.

Moreover, we note that a similar decomposition lemma can also be applied to functions ϕ_j which are defined on an open, bounded set $\Omega \subset \mathbb{R}^d$ with $\mathcal{L}^d(\partial\Omega) = 0$. We refer the reader to [62, Lemma 1.1] for details.

3.3 Gårding-type inequality

In this part of the thesis, and more precisely in Theorem 3.2, we prove the Gårding inequality (3.15) which plays a crucial role in the proof of our weak-strong uniqueness results, Theorems 4.2 and 4.3. Henceforth, we study functions $e : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ which satisfy the following assumptions:

(H_1) $e \in C^3(\mathbb{R}^N \times \mathbb{R})$;

(H_2) $c_2(|z_1|^p + |z_2|^q - 1) \leqslant e(z_1, z_2) \leqslant c_1(|z_1|^p + |z_2|^q + 1)$;

(H_3) $|e_F(z_1, z_2)| \lesssim 1 + |z_1|^{p-1} + |z_2|^{q\frac{p-1}{p}}$, and $|e_\eta(z_1, z_2)| \lesssim 1 + |z_1|^{p\frac{q-1}{q}} + |z_2|^{q-1}$.

We collect all continuous functions $\bar{F} : \mathbb{R}^d \to \mathbb{R}^N$ and $\bar{\eta} : \mathbb{R}^d \to \mathbb{R}$ in the ball of $L^\infty(Q)$ of radius K, with uniform modulus of continuity ω, in the set

$$\mathcal{U}_K := \left\{(\bar{F}, \bar{\eta}) \in C_K(Q) : |\bar{F}(x) - \bar{F}(y)| + |\bar{\eta}(x) - \bar{\eta}(y)| \leqslant \omega(|x - y|), \forall x, y \in \overline{Q}\right\},$$

where $C_K(Q) := \left\{(\bar{F}, \bar{\eta}) \in C(Q; \mathbb{R}^N) \times C(Q; \mathbb{R}) : \|\bar{F}\|_{L^\infty(Q)} + \|\bar{\eta}\|_{L^\infty(Q)} \leqslant K\right\}$.

Next, for $(z_1, z_2) \in \mathbb{R}^N \times \mathbb{R}$ and $e \in C^3(\mathbb{R}^N \times \mathbb{R})$ which satisfies the growth conditions (H_2) and (H_3), we define the function

$$\tilde{e}(z_1, z_2) := e(z_1, z_2) - C_1|V_p(z_1)|^2 - C_2|V_q(z_2)|^2, \tag{3.1}$$

which is not hard to check that satisfies the same growth and coercivity conditions with e up to smaller positive constants. Note that for integrals without the z_2-dependence i.e. $W : \mathbb{R}^N \to \mathbb{R}$, the associated function $\tilde{W}$ is defined naturally as above without the last term on the RHS. Integrands of the latter form concern us in the subsection 4.1.1 and section 4.2. The corresponding Hessians of e and $\tilde{e}$ are denoted by L and $\tilde{L}$ respectively, i.e.

$$L(\lambda_1, \lambda_2)[(\xi_1, \xi_2), (\xi_1, \xi_2)] := e_{FF}(\lambda_1, \lambda_2)\xi_1\xi_1 + 2e_{\eta F}(\lambda_1, \lambda_2)\xi_1\xi_2 + e_{\eta\eta}(\lambda_1, \lambda_2)\xi_2\xi_2,$$

for all $(\xi_1, \xi_2) \in \mathbb{R}^N \times \mathbb{R}$ and $(\lambda_1, \lambda_2) \in \overline{B(0, K)} := \{\lambda \in \mathbb{R}^N \times \mathbb{R} : |\lambda| \leqslant K\}$. In the sequel, without loss of generality, we assume that $p \geqslant q$. In the opposite case, i.e. if $p < q$, the results below can be proved following the same strategy but with the respective adjustments in the proofs.

We next prove a series of results which lead to the proof of Theorem 3.1. Lemma 3.3 provides some properties of the relative function $e(\cdot|\cdot)$ and its proof can be found in the Appendix. For brevity, henceforth, constants shown to depend on K, e.g. $C(K)$, may also depend on the modulus of continuity ω but the latter dependence is omitted from the notation.

Lemma 3.3. *Let f (in place of e) satisfy (H_1), (H_2) and (H_3). Then the following hold:*

(a) There exists $C = C(f, K)$ such that for all $(\lambda_1, \lambda_2) \in \overline{B(0, K)}$

$$\begin{aligned}&|f(\lambda_1 + \xi_1, \lambda_2 + \xi_2|\lambda_1, \lambda_2) - f(\lambda_1 + z_1, \lambda_2 + z_2|\lambda_1, \lambda_2)|\\ &\leqslant C(|\xi_1| + |\xi_2| + |z_1| + |z_2| + |\xi_1|^{p-1} + |z_1|^{p-1} + |\xi_2|^{q\frac{p-1}{p}})|\xi_1 - z_1|\\ &+ C(|\xi_1| + |\xi_2| + |z_1| + |z_2| + |\xi_2|^{q-1} + |z_2|^{q-1} + |z_1|^{p\frac{q-1}{q}})|\xi_2 - z_2|.\end{aligned}$$

Additionally,

$$|f(\lambda_1 + \xi_1, \lambda_2 + \xi_2|\lambda_1, \lambda_2)| \leqslant C\big(|V_p(\xi_1)|^2 + |V_q(\xi_2)|^2\big).$$

(b) For every $\delta > 0$ there exists $R = R(\delta, f, K) > 0$ such that for all (λ_1, λ_2), $(\mu_1, \mu_2) \in \overline{B(0, K)}$ with $|(\lambda_1, \lambda_2) - (\mu_1, \mu_2)| < R$, it holds that

$$\begin{aligned}&|f(\lambda_1 + \xi_1, \lambda_2 + \xi_2|\lambda_1, \lambda_2) - f(\mu_1 + \xi_1, \mu_2 + \xi_2|\mu_1, \mu_2)|\\ &\qquad\leqslant \delta\big(|V_p(\xi_1)|^2 + |V_q(\xi_2)|^2\big).\end{aligned}$$

(c) There exist constants $d_1 = d_1(f, K)$, $d_2 = d_2(f, K)$ such that for all $(\lambda_1, \lambda_2) \in \overline{B(0, K)}$

$$f(\lambda_1 + \xi_1, \lambda_2 + \xi_2|\lambda_1, \lambda_2) \geqslant d_1\big(|\xi_1|^p + |\xi_2|^q\big) - d_2\big(|\xi_1|^2 + |\xi_2|^2\big).$$

Next, we prove two important properties of the function $\tilde{e}$. In the first lemma below, Lemma 3.4, we show that $\tilde{e}$ retains the key quasiconvexity property of e in $\overline{B(0, K)}$, and as a consequence of this result we next prove in Lemma 3.5 that the Hessian $\tilde{L}$ is positive for fixed $x_0 \in Q$.

Lemma 3.4. *Let e satisfy (H_1)-(H_3) be strongly quasiconvex. Then, the function $\tilde{e}$ is strongly quasiconvex at all $(\lambda_1, \lambda_2) \in \overline{B(0, K)}$ with constant $c_0/2$, i.e. for any $Q' \subseteq Q$ and all $|\lambda_1| + |\lambda_2| \leqslant K$*

$$\int_{Q'} \tilde{e}(\lambda_1 + \mathcal{B}\phi, \lambda_2 + \psi) - \tilde{e}(\lambda_1, \lambda_2) - \tilde{e}_\eta(\lambda_1, \lambda_2)\psi \geqslant c_0 \int_{Q'} |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2,$$

holds for all $\phi \in W_0^{l,p}(Q')$ and $\psi \in L^q(Q')$.

Proof. Let $Q' \subseteq Q$, $\phi \in W_0^{l,p}(Q')$, $\psi \in L^q(Q')$ and $(\lambda_1, \lambda_2) \in \overline{B(0, K)}$. Then, noting that $\int_{Q'} \mathcal{B}\phi = 0$,

$$\int_{Q'} \tilde{e}(\lambda_1 + \mathcal{B}\phi, \lambda_2 + \psi) - \tilde{e}(\lambda_1, \lambda_2) - \tilde{e}_\eta(\lambda_1, \lambda_2)\psi$$

$$
\begin{aligned}
&= \int_{Q'} e(\lambda_1 + \mathcal{B}\phi, \lambda_2 + \psi) - e(\lambda_1, \lambda_2) - e_\eta(\lambda_1, \lambda_2)\psi \\
&- C_1 \int_{Q'} |\lambda_1 + \mathcal{B}\phi|^p + |\lambda_1 + \mathcal{B}\phi|^2 - |\lambda_1|^p - |\lambda_1|^2 \\
&- C_2 \int_{Q'} |\lambda_2 + \psi|^q + |\lambda_2 + \psi|^2 - |\lambda_2|^q - |\lambda_2|^2 - q\lambda_2|\lambda_2|^{q-2}\psi - 2\lambda_2\psi \\
&=: I_1 + I_2 + I_3.
\end{aligned}
$$

By the quasiconvexity of e we infer that

$$
I_1 \geqslant c_0 \int_{Q'} |\mathcal{B}\phi|^p + |\mathcal{B}\phi|^2 + |\psi|^q + |\psi|^2,
$$

and for $f(\cdot) = |V_i(\cdot)|^2$ with $i = p$ and $i = q$ respectively in Lemma 3.3 (a), taking again into account that $\int_{Q'} \mathcal{B}\phi = 0$, we deduce that

$$
\begin{aligned}
I_2 &\geqslant -C_1 C \int_{Q'} |\mathcal{B}\phi|^p + |\mathcal{B}\phi|^2, \\
I_3 &\geqslant -C_2 C \int_{Q'} |\psi|^p + |\psi|^2.
\end{aligned}
$$

So, we may choose $C_1 \leqslant c_0/(2C)$ and $C_2 \leqslant c_0/(2C)$ to conclude the proof. $\square$

As a consequence of the above lemma, in the result below, we deduce the positivity of the Hessian $\tilde{L}$ for fixed $x_0 \in Q$.

Lemma 3.5. *Let e satisfy (H_1)-(H_3) be strongly quasiconvex and $Q' \subseteq Q$ and $x_0 \in Q'$. Then for all $\phi \in W_0^{l,p}(Q')$ and $\psi \in L^q(Q')$ it holds that*

$$
\int_{Q'} \tilde{L}(\bar{F}_0, \bar{\eta}_0)\big[(\mathcal{B}\phi(x), \psi(x)), (\mathcal{B}\phi(x), \psi(x))\big]\, dx \geqslant c_0 \int_{Q'} |\mathcal{B}\phi(x)|^2 + |\psi(x)|^2 dx,
$$

where $\bar{F}_0 = \bar{F}(x_0)$ and $\bar{\eta}_0 = \bar{\eta}(x_0)$.

Proof. The quasiconvexity of $\tilde{e}$, Lemma 3.4, says that $I(\phi, \psi) \geqslant I(0,0)$ for all $\phi \in W_0^{l,p}(Q')$ and $\psi \in L^q(Q')$, where

$$
\begin{aligned}
I(\phi, \psi) := &\int_{Q'} \tilde{e}(\bar{F}_0 + \mathcal{B}\phi, \bar{\eta}_0 + \psi) - \tilde{e}(\bar{F}_0, \bar{\eta}_0) - \tilde{e}_\eta(\bar{F}_0, \bar{\eta}_0)\psi \\
&- \frac{c_0}{2} \int_{Q'} |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2,
\end{aligned}
$$

and so we infer that $\frac{d^2}{d\varepsilon^2} I(\varepsilon\phi, \varepsilon\psi)\Big|_{\varepsilon=0} \geqslant 0$. However,

$$
\begin{aligned}
\frac{d}{d\varepsilon} I(\varepsilon\phi, \varepsilon\psi) := &\int_{Q'} \tilde{e}_F(\bar{F}_0 + \varepsilon\mathcal{B}\phi, \bar{\eta}_0 + \varepsilon\psi)\mathcal{B}\phi + \tilde{e}_\eta(\bar{F}_0 + \varepsilon\mathcal{B}\phi, \bar{\eta}_0 + \varepsilon\psi)\psi \\
&- \tilde{e}_\eta(\bar{F}_0, \bar{\eta}_0)\psi - \frac{c_0}{2} p\varepsilon^{p-1}|\mathcal{B}\phi|^p - c_0\varepsilon|\mathcal{B}\phi|^2 - \frac{c_0}{2} q\varepsilon^{q-1}|\psi|^q - c_0\varepsilon|\psi|^2,
\end{aligned}
$$

and so

$$
\begin{aligned}
\frac{d^2}{d\varepsilon^2} I(\varepsilon\phi, \varepsilon\psi) = \int_{Q'} & \tilde{e}_{FF}(\bar{F}_0 + \varepsilon\mathcal{B}\phi, \bar{\eta}_0 + \varepsilon\psi)\mathcal{B}\phi : \mathcal{B}\phi + \tilde{e}_{\eta\eta}(\bar{F}_0 + \varepsilon\mathcal{B}\phi, \bar{\eta}_0 + \varepsilon\psi)\psi \cdot \psi \\
& + 2\tilde{e}_{F\eta}(\bar{F}_0 + \varepsilon\mathcal{B}\phi, \bar{\eta}_0 + \varepsilon\psi)\mathcal{B}\phi \cdot \psi \\
& - \frac{c_0}{2} p(p-1)\varepsilon^{p-2}|\mathcal{B}\phi|^p - c_0|\mathcal{B}\phi|^2 - \frac{c_0}{2} q(q-1)\varepsilon^{q-2}|\psi|^q - c_0|\psi|^2.
\end{aligned}
$$

We conclude that

$$
0 \leqslant \frac{d^2}{d\varepsilon^2} I(\varepsilon\phi, \varepsilon\psi)\Big|_{\varepsilon=0} = \int_{Q'} \tilde{L}(\bar{F}_0, \bar{\eta}_0)\big[(\mathcal{B}\phi, \psi), (\mathcal{B}\phi, \psi)\big] - c_0|\mathcal{B}\phi|^2 - c_0|\psi|^2.
$$

□

We are now ready to prove a Gårding-type inequality for the delocalised version of the Hessian $\tilde{L}$, which is crucial for the contradiction argument of the proof of Theorem 3.1.

Proposition 3.1. *Let e satisfy (H_1)-(H_3) be strongly quasiconvex. Then, for every $\delta > 0$, there exist constants $c > 0$ and $C_{pen} := C(\delta) > 0$ such that*

$$
\int_Q \tilde{L}(\bar{F}(x), \bar{\eta}(x))\big[(\mathcal{B}\phi, \psi), (\mathcal{B}\phi, \psi)\big] \geqslant c(1-\delta)^2 \int_Q |\mathcal{B}\phi|^2 + |\psi|^2 - C_{pen} \sum_{i=1}^{l} \int_Q |\nabla^{l-i}\phi|^2
$$

for all $\phi \in W^{l,p}(Q)$, $\psi \in L^q(Q)$ and $(\bar{F}, \bar{\eta}) \in \mathcal{U}_K$.

Proof. Fix $\delta > 0$ and pick a finite cover $\{Q_i\}_i$, $Q_i := Q(x_i, r_i) \subseteq Q$ such that

$$
\begin{aligned}
&|\tilde{e}_{FF}(\bar{F}(x), \bar{\eta}(x)) - \tilde{e}_{FF}(\bar{F}(x_i), \bar{\eta}(x_i))| + |\tilde{e}_{\eta\eta}(\bar{F}(x), \bar{\eta}(x)) - \tilde{e}_{\eta\eta}(\bar{F}(x_i), \bar{\eta}(x_i))| \\
&+ 2|\tilde{e}_{F\eta}(\bar{F}(x), \bar{\eta}(x)) - \tilde{e}_{F\eta}(\bar{F}(x_i), \bar{\eta}(x_i))| \leqslant \frac{1}{2} c\delta(1-\delta).
\end{aligned}
$$

Note that since $(\bar{F}(x), \bar{\eta}(x)) \in \mathcal{U}_K$ (bounded with uniform modulus of continuity) and $\tilde{e} \in C^2$ the cover can be chosen uniformly.

Now choose a partition of unity $(\rho_i)_i$, $\rho_i \in C_c^\infty(Q_i)$ and $\sum_i \rho_i^2 = 1$. Given $\phi \in W^{l,p}(Q)$ and $\psi \in L^q(Q)$, we infer that

$$
\begin{aligned}
&\sum_i \int_{Q_i} \tilde{L}(\bar{F}, \bar{\eta})[(\rho_i\mathcal{B}\phi, \rho_i\psi), (\rho_i\mathcal{B}\phi, \rho_i\psi)] - \tilde{L}(\bar{F}_i, \bar{\eta}_i)[(\rho_i\mathcal{B}\phi, \rho_i\psi), (\rho_i\mathcal{B}\phi, \rho_i\psi)] \\
&\geqslant -\frac{c}{2}\delta(1-\delta) \int_Q \big(|\mathcal{B}\phi|^2 + |\mathcal{B}\phi||\psi| + |\psi|^2\big) \geqslant -c\delta(1-\delta) \int_Q \big(|\mathcal{B}\phi|^2 + |\psi|^2\big),
\end{aligned}
$$

and so,

$$
\begin{aligned}
&\int_Q \tilde{L}(\bar{F}, \bar{\eta})[(\mathcal{B}\phi, \psi), (\mathcal{B}\phi, \psi)] = \\
&\sum_i \int_{Q_i} \tilde{L}(\bar{F}, \bar{\eta})[(\rho_i\mathcal{B}\phi, \rho_i\psi), (\rho_i\mathcal{B}\phi, \rho_i\psi)] - \tilde{L}(\bar{F}_i, \bar{\eta}_i)[(\rho_i\mathcal{B}\phi, \rho_i\psi), (\rho_i\mathcal{B}\phi, \rho_i\psi)]
\end{aligned}
$$

$$+\sum_i \int_{Q_i} \tilde{L}(\bar{F}_i,\bar{\eta}_i)[(\rho_i\mathcal{B}\phi,\rho_i\psi),(\rho_i\mathcal{B}\phi,\rho_i\psi)] =: \sum_i I_i + \sum_i II_i.$$

Since we already bounded the term $\sum_i I_i$ from below, it remains to prove a similar bound for the second term. Since $\mathcal{B}(\rho_i\phi) = \rho_i\mathcal{B}\phi + \phi\otimes_{\mathcal{B}}\rho_i$, where $\phi\otimes_{\mathcal{B}}\rho_i := \sum_{j=1}^{l} B_j^L[\nabla^j\rho_i,\nabla^{l-j}\phi]$ and B_j^L are given by the Leibniz rule, we infer that

$$\begin{aligned}
II_i &= \int_{Q_i} \tilde{e}_{FF}(\bar{F}_i,\bar{\eta}_i)\mathcal{B}(\rho_i\phi) : \mathcal{B}(\rho_i\phi) + \tilde{e}_{\eta\eta}(\bar{F}_i,\bar{\eta}_i)(\rho_i\psi) : (\rho_i\psi) + 2\tilde{e}_{F\eta}(\bar{F}_i,\bar{\eta}_i)\mathcal{B}(\rho_i\phi) : (\rho_i\psi) \\
&+ \int_{Q_i} \tilde{e}_{FF}(\bar{F}_i,\bar{\eta}_i)(\phi\mathcal{B}\rho_i) : (\phi\otimes_{\mathcal{B}}\rho_i) - \int_{Q_i} \tilde{e}_{FF}(\bar{F}_i,\bar{\eta}_i)\mathcal{B}(\rho_i\phi) : (\phi\otimes_{\mathcal{B}}\rho_i) \\
&- \int_{Q_i} \tilde{e}_{F\eta}(\bar{F}_i,\bar{\eta}_i)(\phi\otimes_{\mathcal{B}}\rho_i) : (\rho_i\psi) =: \sum_{j=1}^{4} T_i^j.
\end{aligned}$$

In particular, for T_i^1 we apply Lemma 3.5 testing with $\rho_i\phi \in W_0^{l,p}(Q_i)$ and $\rho_i\psi \in L^q(Q_i)$ to infer that

$$T_i^1 \geqslant c_0 \int_{Q_i} \left(|\mathcal{B}(\rho_i\phi)|^2 + |\rho_i\psi|^2\right).$$

For the remaining three terms, by Young's inequality, we find that

$$\begin{aligned}
T_i^2 &\geqslant -c\left(\sup_j \|\nabla^j\rho_i\|_\infty\right)\sum_{j=1}^{l}\int_{Q_i}|\nabla^{l-j}\phi|^2, \\
T_i^3 &\geqslant -c(1-\delta)^2\int_{Q_i}|\mathcal{B}(\rho_i\phi)|^2 - 2C_1(\delta)\sum_{j=1}^{l}\int_{Q_i}|\nabla^{l-j}\phi|^2, \\
T_i^4 &\geqslant -c\frac{1+\delta^2}{2}\int_{Q_i}|\rho_i\psi|^2 - 2C_2(\delta)\sum_{j=1}^{l}\int_{Q_i}|\nabla^{l-j}\phi|^2.
\end{aligned}$$

Finally combining all the above and since

$$\begin{aligned}
|\mathcal{B}(\rho_i\phi)|^2 &= |\rho_i\mathcal{B}\phi + \phi\otimes_{\mathcal{B}}\rho_i|^2 \leqslant \frac{1}{2}|\rho_i\mathcal{B}\phi|^2 + \frac{1}{2}|\phi\otimes_{\mathcal{B}}\rho_i|^2, \\
|\mathcal{B}(\rho_i\phi)|^2 &= |\rho_i\mathcal{B}\phi + \phi\otimes_{\mathcal{B}}\rho_i|^2 \geqslant (1-\delta)\rho_i^2|\mathcal{B}\phi|^2 - C(\delta)|\phi\otimes_{\mathcal{B}}\rho_i|^2,
\end{aligned}$$

we conclude that

$$\int_Q \tilde{L}(\bar{F},\bar{\eta})[(\mathcal{B}\phi,\psi),(\mathcal{B}\phi,\psi)] \gtrsim (1-\delta)^2\int_Q |\mathcal{B}\phi|^2 + |\psi|^2 - C(\delta)\sum_{j=1}^{l}\int_Q|\nabla^{l-j}\phi|^2.$$

□

Following ideas of [31, 109], the following lemma shows that the function $\tilde{e}(\cdot|\cdot)$ has convex-like behaviour on the wave cone when it is integrated over cubes with sufficiently small radius and plays a crucial role in the proof of Proposition 3.2.

Lemma 3.6. *Let e satisfy (H_1)-(H_3) be strongly quasiconvex. Then, there exists $R = R(c_0, K) > 0$ such that for all $x_0 \in Q$ the inequality*

$$\int_{Q(x_0,r)} \tilde{e}(\bar{F}(x) + \mathcal{B}\phi, \bar{\eta}(x) + \psi | \bar{F}(x), \bar{\eta}(x)) \geqslant \frac{c_0}{4} \int_{Q(x_0,r)} |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2$$

holds for all $\phi \in W_0^{l,p}(Q(x_0,r))$ and $\psi \in L^q(Q(x_0,r))$ with $r \leqslant R$.

Proof. Observe that by Lemma 3.3 (b), letting $\delta = c_0/4$ we find $R = R(c_0, \tilde{e}, K)$ such that for all $(\bar{F}, \bar{\eta}) \in \mathcal{U}_K$ and whenever $|x - x_0| < R$

$$\left|\tilde{e}(\bar{F} + \mathcal{B}\phi, \bar{\eta} + \psi | \bar{F}, \bar{\eta}) - \tilde{e}(\bar{F}_0 + \mathcal{B}\phi, \bar{\eta}_0 + \psi | \bar{F}_0, \bar{\eta}_0)\right| \leqslant \frac{c_0}{4}\big(|V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2\big).$$

Note here that $\tilde{e}$ satisfies the growth conditions of Lemma 3.3. So for $\phi \in W_0^{l,p}(Q(x_0,r))$ and $\psi \in L^q(Q(x_0,r))$ with $r \leqslant R$ we infer that

$$\begin{aligned}\int_{Q(x_0,r)} \tilde{e}(\bar{F} + \mathcal{B}\phi, \bar{\eta} + \psi | \bar{F}, \bar{\eta}) \geqslant & \int_{Q(x_0,r)} \tilde{e}(\bar{F}_0 + \mathcal{B}\phi, \bar{\eta}_0 + \psi | \bar{F}_0, \bar{\eta}_0) \\ & - \frac{c_0}{4}\big(|V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2\big)\end{aligned}$$

$$\begin{aligned}= \int_{Q(x_0,r)} & \tilde{e}(\bar{F}_0 + \mathcal{B}\phi, \bar{\eta}_0 + \psi) - \tilde{e}(\bar{F}_0, \bar{\eta}_0) - \tilde{e}_\eta(\bar{F}_0, \bar{\eta}_0)\psi - \tilde{e}_F(\bar{F}_0, \bar{\eta}_0)\mathcal{B}\phi \\ & - \frac{c_0}{4}\big(|V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2\big) \geqslant \frac{c_0}{4} \int_{Q(x_0,r)} \big(|V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2\big),\end{aligned}$$

where in the last inequality we used Lemma 3.4 and that $\int_{Q(x_0,r)} \mathcal{B}\phi = 0$. ☐

We next prove a central result which can be seen as a limiting version of a Gårding inequality and replaces the quasiconvexity condition in the proof of Theorem 3.1.

Proposition 3.2. *Let e satisfy (H_1)-(H_3) be strongly quasiconvex and $(\bar{F}_k, \bar{\eta}_k)_k \subseteq \mathcal{U}_K$, $(\phi_k)_k \subseteq W^{l,p}(Q)$, $(\psi_k)_k \subseteq L^q(Q)$ and $(a_k)_k \subseteq \mathbb{R}$ such that*

- $a_k^{-1} V_p(\nabla^{l-i}\phi_k) \to 0$ *strongly in $L^2(Q)$ for all $i = 1, .., l$,*
- $\big(a_k^{-1} V_p(\mathcal{B}\phi_k)\big)_k$ *is bounded in $L^2(Q)$,*
- $\big(a_k^{-1} V_q(\psi_k)\big)_k$ *is bounded in $L^2(Q)$.*

Then,

$$\begin{aligned}\liminf_k \frac{c_0}{4} a_k^{-2} \int_Q & |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 \\ & \leqslant \liminf_k a_k^{-2} \int_Q \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k | \bar{F}_k, \bar{\eta}_k).\end{aligned}$$

Proof. Since $\left(a_k^{-2}|V_p(\mathcal{B}\phi_k)|^2 + a_k^{-2}|V_q(\psi_k)|^2\right)_k$ is bounded in $L^1(Q)$ we may assume that (up to a subsequence)

$$a_k^{-2}|V_p(\mathcal{B}\phi_k)|^2 + a_k^{-2}|V_q(\psi_k)|^2\mathcal{L}^d\lfloor_{\mathbb{T}^d} \overset{*}{\rightharpoonup} \mu, \quad \text{in } \mathcal{M}(\mathbb{T}^d) = \Big(C(\mathbb{T}^d)\Big)^*.$$

Since μ is a positive measure, there can be at most a countable number of hyperplanes parallel to the coordinate axes which admit positive μ-measure. Hence, we can extract a finite cover of Q by non overlapping cubes $Q_{r_j} := Q(x_j, r_j)$ with the property that $r_j < R$, so that Lemma 3.6 applies and

$$\mu(\partial Q(x_j, r_j)) = 0. \tag{3.2}$$

Next, consider cut-off functions $\rho_j \in C_c^\infty(Q(x_j, r_j))$ such that for $\lambda \in (0,1)$

$$\mathbb{1}_{Q(x_j,\lambda r_j)} \leqslant \rho_j \leqslant \mathbb{1}_{Q(x_j,r_j)}, \ \|\nabla^i \rho_j\|_{L^\infty(Q)} \leqslant \frac{C}{(1-\lambda)^i},$$

for $i = 1, .., l$. For simplicity, we denote $Q_{r_j} := Q(x_j, r_j)$ and $Q_{\lambda r_j} := Q(x_j, \lambda r_j)$. We now apply Lemma 3.6 for the functions $\rho_j\phi_k \in W_0^{l,p}(Q_{r_j})$ and $\rho_j\psi_k \in L^q(Q_{r_j})$ to find that

$$\frac{c_0}{4}\int_{Q_{r_j}} |V_p(\mathcal{B}(\rho_j\phi_k))|^2 + |V_q(\rho_j\psi_k)|^2 \leqslant \int_{Q_{r_j}} \tilde{e}(\bar{F}_k + \mathcal{B}(\rho_j\phi_k), \bar{\eta}_k + \rho_j\psi_k|\bar{F}_k, \bar{\eta}_k),$$

where $(\bar{F}_k, \bar{\eta}_k) \in \mathcal{U}_K$. Thus by Lemma 3.3 (a) and for $C = C(\tilde{e}, K)$, it holds that

$$\begin{aligned}
&\frac{c_0}{4}\int_{Q_{\lambda r_j}} |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 + \frac{c_0}{4}\int_{Q_{r_j}\setminus Q_{\lambda r_j}} |V_p(\mathcal{B}(\rho_j\phi_k))|^2 + |V_q(\rho_j\psi_k)|^2 \\
&\leqslant \int_{Q_{\lambda r_j}} \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k, \bar{\eta}_k) + \int_{Q_{r_j}\setminus Q_{\lambda r_j}} \tilde{e}(\bar{F}_k + \mathcal{B}(\rho_j\phi_k), \bar{\eta}_k + \rho_j\psi_k|\bar{F}_k, \bar{\eta}_k) \\
&\leqslant \int_{Q_{\lambda r_j}} \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k, \bar{\eta}_k) + C\int_{Q_{r_j}\setminus Q_{\lambda r_j}} |V_p(\mathcal{B}(\rho_j\phi_k))|^2 + |V_q(\rho_j\psi_k)|^2 \\
&\leqslant \int_{Q_{\lambda r_j}} \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k, \bar{\eta}_k) + C\int_{Q_{r_j}\setminus Q_{\lambda r_j}} |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 \\
&\qquad + \sum_{i=1}^{l}\Big|V_p\Big(\frac{\nabla^{l-i}\phi_k}{(1-\lambda)^i}\Big)\Big|^2,
\end{aligned}$$

where in the last inequality we used the definition of the cut-offs and the fact that $\mathcal{B}(\rho_j\phi_k) = \phi_k \otimes_{\mathcal{B}} \rho_j + \rho_j\mathcal{B}\phi_k$. We observe that the second term on the left hand side is positive and so by summing over j we infer that

$$\begin{aligned}
&\frac{c_0}{4}\int_Q |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 - \frac{c_0}{4}\sum_j \int_{Q_{r_j}\setminus Q_{\lambda r_j}} |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 \\
&\leqslant \int_Q \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k, \bar{\eta}_k) - \sum_j \int_{Q_{r_j}\setminus Q_{\lambda r_j}} \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k, \bar{\eta}_k) \\
&+ C\sum_j \int_{Q_{r_j}\setminus Q_{\lambda r_j}} |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 + \sum_{i=1}^{l}\Big|V_p\Big(\frac{\nabla^{l-i}\phi_k}{(1-\lambda)^i}\Big)\Big|^2.
\end{aligned}$$

Using again Lemma 3.3 (a) we deduce that

$$\frac{c_0}{4}\int_Q |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 \leqslant \int_Q \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k, \bar{\eta}_k)$$
$$+ C\sum_j \int_{Q_{r_j}\setminus Q_{\lambda r_j}} |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 + \sum_{i=1}^{l}\Big|V_p\Big(\frac{\nabla^{l-i}\phi_k}{(1-\lambda)^i}\Big)\Big|^2.$$

Multiplying with a_k^{-2} and taking the limit over k we conclude that

$$\liminf_k \frac{c_0}{4}a_k^{-2}\int_Q |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 \leqslant a_k^{-2}\liminf_k \int_Q \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k, \bar{\eta}_k)$$
$$+ C\sum_j \mu\big(\bar{Q}_{r_j}\setminus Q_{\lambda r_j}\big).$$

Finally, send $\lambda \to 1$ to complete the proof via (3.2). $\square$

The Gårding inequality, Theorem 3.2, follows as a consequence of Theorem 3.1 below which forms the core of this section. Before we proceed to the theorem, we define the auxiliary mapping $\|\cdot\|_{W^{-1,(2,p)}} : L^p(Q) \to \mathbb{R}$ (though not a norm) by

$$\|\varphi\|_{W^{-1,(2,p)}} := \big(\|\varphi\|^2_{W^{-1,2}(Q)} + \|\varphi\|^p_{W^{-1,p}(Q)}\big)^{1/2}. \tag{3.3}$$

Theorem 3.1. *Let e satisfy (H_1)-(H_3) be strongly quasiconvex. Then, there exists $\epsilon_0 > 0$ and constants $\tilde{C}_0 = \tilde{C}_0(e,K) > 0$, $\tilde{C}_1 = \tilde{C}_1(e,K) > 0$ such that for all $(\bar{F},\bar{\eta}) \in \mathcal{U}_K$, $\psi \in L^q(Q)$ with $\int_Q \psi = 0$ and all $\varphi \in L^p_{\mathcal{A}}(Q)$ with $\|\varphi\|_{L^p(Q)} < \epsilon_0$ it holds that*

$$\int_Q \Big(|V_p(\varphi(x))|^2 + |V_q(\psi(x))|^2\Big)dx$$
$$\leqslant \tilde{C}_0\int_Q e(\bar{F}(x)+\varphi(x), \bar{\eta}(x)+\psi(x)|\bar{F}(x),\bar{\eta}(x))dx + \tilde{C}_1\|\varphi\|^2_{W^{-1,(2,p)}}.$$

Proof. It is enough to show the existence of some $\epsilon_0 > 0$ such that for all $(\bar{F},\bar{\eta}) \in \mathcal{U}_K$, $\psi \in L^q(Q)$ with zero average and $\varphi \in L^p_{\mathcal{A}}(Q)$ with $\|\varphi\|_{W^{-1,p}(Q)} < \epsilon_0$ it holds that

$$\int_Q \tilde{e}(\bar{F}+\mathcal{B}\phi, \bar{\eta}+\psi|\bar{F},\bar{\eta}) + \frac{C_{pen}}{2}\sum_{j=1}^{l}\int_Q |\nabla^{l-j}\phi|^2 \geqslant 0, \tag{3.4}$$

where ϕ is the $\mathbb{B}^\dagger$-primitive of φ whose existence is guaranteed by Lemma 2.1.

Indeed, due to the convexity of the functions $f_p(\cdot) := |V_p(\cdot)|^2$ and $f_q(\cdot) := |V_q(\cdot)|^2$, we see that (3.4) implies that there exists $C > 0$ such that

$$C\int_Q \big(|V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2\big) \leqslant C_1\int_Q f_p(\bar{F}+\mathcal{B}\phi|\bar{F}) + C_2\int_Q f_q(\bar{\eta}+\psi|\bar{\eta})$$
$$\leqslant \int_Q e(\bar{F}+\mathcal{B}\phi, \bar{\eta}+\psi|\bar{F},\bar{\eta}) + \frac{C_{pen}}{2}\sum_{j=1}^{l}\int_Q |\nabla^{l-j}\phi|^2.$$

This would conclude the proof of Theorem 3.1 since by Lemma 2.1 (iv),

$$\sum_{j=1}^{l}\int_Q |\nabla^{l-j}\phi|^2 \leqslant \|\phi\|^2_{W^{l-1,2}(Q)} + \|\phi\|^p_{W^{l-1,p}(Q)} \leqslant C\|\mathcal{B}\phi\|^2_{W^{-1,(2,p)}}.$$

We proceed to prove (3.4) by contradiction. Suppose that (3.4) fails. Then, there exist $(\bar{F}_k,\bar{\eta}_k)_k \subseteq \mathcal{U}_K$, $(\psi_k)_k \subseteq L^q(Q)$ zero-average and pairs $(\phi_k,\varphi_k) \subseteq W^{l,p}(Q)\times L^p(Q)$ with $\|\varphi_k\|_{W^{-1,p}(Q)} \to 0$ and $(\bar{F}_k,\bar{\eta}_k) \to (\bar{F},\bar{\eta})$ in $C^0(Q)$ such that

$$\int_Q \tilde{e}(\bar{F}_k+\mathcal{B}\phi_k,\bar{\eta}_k+\psi_k|\bar{F}_k,\bar{\eta}_k) + \frac{C_{pen}}{2}\sum_{j=1}^{l}\int_Q |\nabla^{l-j}\phi|^2 < 0, \tag{3.5}$$

where ϕ_k is the $\mathbb{B}^\dagger$-primitive of φ_k. Note that, again by Lemma 2.1 (iv), we extract the strong convergence $\|\phi_k\|_{W^{l-1,p}} \lesssim \|\varphi_k\|_{W^{-1,p}} \to 0$.

We split the proof into 5 steps. Let

$$\beta_k^\eta := \Big(\|\mathcal{B}\phi_k\|^q_{L^q(Q)} + \|\psi_k\|^q_{L^q(Q)}\Big)^{1/q},\ \ \beta_k^F := \Big(\|\mathcal{B}\phi_k\|^p_{L^p(Q)} + \|\psi_k\|^q_{L^q(Q)}\Big)^{1/p}$$
$$\text{and}\quad \alpha_k := \Big(\|\mathcal{B}\phi_k\|^2_{L^2(Q)} + \|\psi_k\|^2_{L^2(Q)}\Big)^{1/2}.$$

Step 1: We show that $\|\mathcal{B}\phi_k\|_{L^p(Q)} \to 0$, $\|\psi_k\|_{L^q(Q)} \to 0$ as $k\to\infty$ and

$$\Lambda^F := \sup_k \frac{(\beta_k^F)^p}{\alpha_k^2} < \infty, \qquad \Lambda^\eta := \sup_k \frac{(\beta_k^\eta)^q}{\alpha_k^2} < \infty. \tag{3.6}$$

Indeed, by the coercivity of $\tilde{e}$ and Young's inequality we infer that

$$\begin{aligned}&\tilde{e}(\bar{F}_k+\mathcal{B}\phi_k,\bar{\eta}_k+\psi_k|\bar{F}_k,\bar{\eta}_k) = \tilde{e}(\bar{F}_k+\mathcal{B}\phi_k,\bar{\eta}_k+\psi_k) - \tilde{e}(\bar{F}_k,\bar{\eta}_k)\\ &- \tilde{e}_F(\bar{F}_k,\bar{\eta}_k)\mathcal{B}\phi_k - \tilde{e}_\eta(\bar{F}_k,\bar{\eta}_k)\psi_k \geqslant -C(\delta) + (\frac{c}{2^p}-\delta)|\mathcal{B}\phi_k|^p + (\frac{c}{2^q}-\delta)|\psi_k|^q,\end{aligned}$$

and for $\delta>0$ small enough we conclude that

$$0 \overset{(3.5)}{>} \int_Q \tilde{e}(\bar{F}_k+\mathcal{B}\phi_k,\bar{\eta}_k+\psi_k|\bar{F}_k,\bar{\eta}_k) \geqslant -C + c\int_Q |\mathcal{B}\phi_k|^p + |\psi_k|^q.$$

The above inequality tell us that the sequences $(\mathcal{B}\phi_k)_k$ and $(\psi_k)_k$ are bounded in $L^p(Q)$ and $L^q(Q)$ respectively. We now apply Proposition 3.2 for the sequences $(\bar{F}_k,\bar{\eta}_k)_k$, $(\mathcal{B}\phi_k)_k$, $(\psi_k)_k$ and $a_k=1$ to find that

$$\liminf_k \frac{c_0}{4}\int_Q |V_p(\mathcal{B}\phi_k)|^2 + |V_q(\psi_k)|^2 \leqslant \liminf_k \int_Q \tilde{e}(\bar{F}_k+\mathcal{B}\phi_k,\bar{\eta}_k+\psi_k|\bar{F}_k,\bar{\eta}_k) < 0,$$

and so, up to a subsequence, $\|\mathcal{B}\phi_k\|_{L^p(Q)} \to 0$ and $\|\psi_k\|_{L^q(Q)} \to 0$. Note that, in order to apply Proposition 3.2, we used the fact that $\phi_k \to 0$ in $W^{l-1,p}(Q)$.

For the rest of Step 1 we recall Lemma 3.3 (c) which tell us that

$$0 \overset{(3.5)}{\geqslant} d_1\int_Q \Big(|\mathcal{B}\phi_k|^p + |\psi_k|^q\Big) - d_2\int_Q \Big(|\mathcal{B}\phi_k|^2 + |\psi_k|^2\Big),$$

$$0 \overset{(3.5)}{\geqslant} d_3 \int_Q \Big(|\mathcal{B}\phi_k|^q + |\psi_k|^q \Big) - d_4 \int_Q \Big(|\mathcal{B}\phi_k|^2 + |\psi_k|^2 \Big),$$

and so by dividing both inequalities by α_k^2 we conclude the proof of this step. In the second inequality we used that

$$\tilde{e}(F, \eta) \gtrsim -1 + |F|^p + |\eta|^q \gtrsim -1 + |F|^q + |\eta|^q,$$

since $p \geqslant q$ and so $|F|^q \leqslant 1 + |F|^p$.

<u>Step 2</u>: Following the strategy of [31], [29], [61] we decompose the normalised sequences

$$s_k := \frac{\phi_k}{\alpha_k} \quad \text{and} \quad c_k := \frac{\psi_k}{\alpha_k}.$$

Since $\|\mathcal{B}s_k\|^2_{L^2(Q)} + \|c_k\|^2_{L^2(Q)} = 1$ we find $s \in W^{1,2}(Q)$ and $c \in L^2(Q)$ with $\int_Q c = 0$ such that $\mathcal{B}s_k \rightharpoonup \mathcal{B}s$ and $c_k \rightharpoonup c$ in $L^2(Q)$. Moreover, $\int_Q \mathcal{B}s_k = 0$ and $\mathcal{A}(\mathcal{B}s_k) = 0$. Setting

$$M_k^F := 2^{-\frac{p-2}{2p}} \frac{\alpha_k}{\beta_k^F} \quad \text{and} \quad M_k^\eta := 2^{-\frac{q-2}{2q}} \frac{\alpha_k}{\beta_k^\eta},$$

we also infer that $\|M_k^F \mathcal{B}s_k\|_{L^p(Q)}$, $\|M_k^\eta c_k\|_{L^q(Q)} \leqslant 1$ and M_k^F, $M_k^\eta \in (0,1]$. The first two bounds come directly from the definition of the sequences s_k and c_k, while for the rest we have that

$$\frac{(M_k^F)^p}{2^{-\frac{p-2}{2}}} = \frac{\alpha_k^p}{(\beta_k^F)^p} = \frac{\Big(\|\mathcal{B}\phi_k\|^2_{L^2} + \|\psi_k\|^2_{L^2}\Big)^{p/2}}{\|\mathcal{B}\phi_k\|^p_{L^p} + \|\psi_k\|^q_{L^q}} \leqslant 2^{(p-2)/2} \frac{\|\mathcal{B}\phi_k\|^p_{L^2} + \|\psi_k\|^p_{L^2}}{\|\mathcal{B}\phi_k\|^p_{L^p} + \|\psi_k\|^q_{L^q}} \leqslant 2^{\frac{p-2}{2}},$$

for k large enough. The last inequality comes from the fact that, since $p \geqslant q$ and $\|\psi_k\|_{L^q} \to 0$, we may find a subsequence (denoted again by ψ_k) such that $\|\psi_k\|^p_{L^q} \leqslant \|\psi_k\|^q_{L^q}$. Similarly, we see that

$$\frac{(M_k^\eta)^q}{2^{-\frac{q-2}{2}}} = \frac{\alpha_k^q}{(\beta_k^\eta)^q} = \frac{\Big(\|\mathcal{B}\phi_k\|^2_{L^2} + \|\psi_k\|^2_{L^2}\Big)^{q/2}}{\|\mathcal{B}\phi_k\|^q_{L^q} + \|\psi_k\|^q_{L^q}} \leqslant 2^{(q-2)/2} \frac{\|\mathcal{B}\phi_k\|^q_{L^2} + \|\psi_k\|^q_{L^2}}{\|\mathcal{B}\phi_k\|^q_{L^q} + \|\psi_k\|^q_{L^q}} \leqslant 2^{\frac{q-2}{2}}.$$

According to the decomposition lemmas [51, Lemma 8.13] and Lemma 3.2, we find $\mathbb{B}^\dagger$-primitives g_k, $b_k \in W^{1,2}(Q)$ and functions G_k, $B_k \in L^2(Q)$ such that (up to common subsequences)

(a) $\mathcal{B}s_k = \mathcal{B}s + \mathcal{B}g_k + \mathcal{B}b_k$;

(b) $\mathcal{B}g_k$, $\mathcal{B}b_k \rightharpoonup 0$ in $L^2(Q)$ and $M_k^F \mathcal{B}g_k$, $M_k^F \mathcal{B}b_k \rightharpoonup 0$ in $L^p(Q)$;

(c) $\big(|\mathcal{B}g_k|^2\big)_k$ and $\big(|M_k^F \mathcal{B}g_k|^p\big)_k$ are equiintegrable;

(d) $\mathcal{B}b_k \to 0$ and $M_k^F \mathcal{B}b_k \to 0$ in measure, and

(a') $c_k = c + G_k + B_k$;

(b') G_k, $B_k \rightharpoonup 0$ in $L^2(Q)$ and $M_k^\eta G_k$, $M_k^\eta B_k \rightharpoonup 0$ in $L^q(Q)$;

(c') $\big(|G_k|^2\big)_k$ and $\big(|M_k^F G_k|^q\big)_k$ are equiintegrable;

(d') $B_k \to 0$ and $M_k^\eta B_k \to 0$ in measure.

We define

$$f_k(x) := \alpha_k^{-2}\Big(\tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} s_k, \bar{\eta}_k + \alpha_k c_k | \bar{F}_k, \bar{\eta}_k) - \tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} b_k, \bar{\eta}_k + \alpha_k B_k | \bar{F}_k, \bar{\eta}_k)\Big)$$

and then we deduce that

$$\begin{aligned}
&\int_Q f_k(x) + \alpha_k^{-2}\tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} b_k, \bar{\eta}_k + \alpha_k B_k | \bar{F}_k, \bar{\eta}_k) + \frac{C_{pen}}{2}\sum_{i=1}^{l} |\nabla^{l-i} s_k|^2 \\
&\quad = \alpha_k^{-2}\int_Q \tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} s_k, \bar{\eta}_k + \alpha_k c_k | \bar{F}_k, \bar{\eta}_k) + \frac{C_{pen}}{2}\sum_{i=1}^{l} |\nabla^{l-i} \alpha_k s_k|^2 \\
&\quad = \alpha_k^{-2}\int_Q \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k | \bar{F}_k, \bar{\eta}_k) + \frac{C_{pen}}{2}\sum_{i=1}^{l} |\nabla^{l-i} \phi_k|^2 \overset{(3.4)}{<} 0.
\end{aligned}$$

This shows that

$$\int_Q f_k(x) + \alpha_k^{-2}\tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} b_k, \bar{\eta}_k + \alpha_k B_k | \bar{F}_k, \bar{\eta}_k) + \frac{C_{pen}}{2}\sum_{i=1}^{l} |\nabla^{l-i} s_k|^2 < 0. \tag{3.7}$$

Step 3: In this step we show that the contribution of the concentrating part must be nonnegative in the limit due to quasiconvexity. In particular, we prove that

$$\liminf_k \alpha_k^{-2}\int_Q \tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} b_k, \bar{\eta}_k + \alpha_k B_k | \bar{F}_k, \bar{\eta}_k) \geqslant 0.$$

To this aim we apply Proposition 3.2 for the sequences $(\bar{F}_k, \bar{\eta}_k)_k$, $(\alpha_k b_k)_k$, $(\alpha_k B_k)_k$ and $(\alpha_k)_k$, after noting that

$$\alpha_k^{p-2} = \frac{(\beta_k^F)^p}{\alpha_k^2}\frac{\alpha_k^p}{(\beta_k^F)^p} \leqslant 2^{\frac{p-2}{2}}\Lambda^F |M_k^F|^p,$$

where, by Step 1, $\Lambda^F := \sup_k (\beta_k^F)^p/\alpha_k^2$ is finite. Thus, again due to the control of the full Sobolev norm of the $\mathbb{B}^\dagger$-primitives b_k, Lemma 2.1, we infer that

$$\alpha_k^{-2}|V_p(\alpha_k \nabla^{l-i} b_k)|^2 = |\nabla^{l-i} b_k|^2 + \alpha_k^{p-2}|\nabla^{l-i} b_k|^p \lesssim |\nabla^{l-i} b_k|^2 + \Lambda^F |M_k^F \nabla^{l-i} b_k|^p \to 0,$$

in $L^1(Q)$, for $i = 1, .., l$. Also,

$$\sup_k \int_Q \alpha_k^{-2}|V_p(\alpha_k \mathcal{B} b_k)|^2 \lesssim \sup_k \int_Q |\mathcal{B} b_k|^2 + \Lambda^F \sup_k |M_k^F \mathcal{B} b_k|^p < \infty,$$

$$\sup_k \int_Q \alpha_k^{-2} |V_q(\alpha_k B_k)|^2 \lesssim \sup_k \int_Q |B_k|^2 + \Lambda^\eta \sup_k |M_k^\eta B_k|^q < \infty,$$

and so Proposition 3.2 says that

$$\begin{aligned} 0 \leqslant \liminf_k \frac{c_0}{4} \alpha_k^{-2} \int_Q |V_p(\alpha_k \mathcal{B} b_k)|^2 + |V_q(\alpha_k B_k)|^2 \\ \leqslant \liminf_k \alpha_k^{-2} \int_Q \tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} b_k, \bar{\eta}_k + \alpha_k B_k | \bar{F}_k, \bar{\eta}_k). \end{aligned}$$

Combining this with (3.7) we have that

$$\liminf_k \int_Q f_k(x) + \frac{C_{pen}}{2} \sum_{i=1}^{l} |\nabla^{l-i} s_k|^2 < 0. \tag{3.8}$$

Step 4: Next, consider the $(\mathcal{A}, 0)$-2-Young measure generated by the sequence $(\mathcal{B} s_k, c_k)$, say $\nu = (\nu_x)_{x \in Q}$, and recall that $(\bar{F}_k, \bar{\eta}_k) \to (\bar{F}, \bar{\eta})$ in $C^0(Q)$. In this step we show that

$$\frac{1}{2} \int_Q \big\langle \nu_x, \tilde{L}(\bar{F}(x), \bar{\eta}(x)) \, [\Lambda, \Lambda] \big\rangle dx \leqslant \liminf_k \int_Q f_k(x),$$

where for simplicity we use the notation $\Lambda := (\lambda_F, \lambda_\eta)$.

We first prove the equiintegrability of f_k. Recalling that $\mathcal{B} s_k - \mathcal{B} b_k := \mathcal{B} s + \mathcal{B} g_k$ and $c_k - B_k = c + G_k$, by Lemma 3.3 (a), we find that

$$|f_k| = \frac{|\tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} s_k, \bar{\eta}_k + \alpha_k c_k | \bar{F}_k, \bar{\eta}_k) - \tilde{e}(\bar{F}_k + \alpha_k \mathcal{B} b_k, \bar{\eta}_k + \alpha_k B_k | \bar{F}_k, \bar{\eta}_k)|}{\alpha_k^2}$$

$$\begin{aligned} \leqslant \alpha_k^{-2} \big(|\alpha_k \mathcal{B} s_k| + |\alpha_k c_k| + |\alpha_k \mathcal{B} b_k| + |\alpha_k B_k| \\ + |\alpha_k \mathcal{B} s_k|^{p-1} + |\alpha_k \mathcal{B} b_k|^{p-1} + |\alpha_k B_k|^{q \frac{p-1}{p}} \big) \alpha_k |\mathcal{B} s + \mathcal{B} g_k| \end{aligned}$$

$$\begin{aligned} + \alpha_k^{-2} \big(|\alpha_k \mathcal{B} s_k| + |\alpha_k c_k| + |\alpha_k \mathcal{B} b_k| + |\alpha_k B_k| \\ + |\alpha_k c_k|^{q-1} + |\alpha_k B_k|^{q-1} + |\alpha_k \mathcal{B} b_k|^{p \frac{q-1}{q}} \big) \alpha_k |c + G_k| \end{aligned}$$

$$\begin{aligned} = & \big(|\mathcal{B} s_k| + |c_k| + |\mathcal{B} b_k| + |B_k| + \alpha_k^{p-2} |\mathcal{B} s_k|^{p-1} + \alpha_k^{p-2} |\mathcal{B} b_k|^{p-1} \big) |\mathcal{B} s + \mathcal{B} g_k| \\ & + \big(|\mathcal{B} s_k| + |c_k| + |\mathcal{B} b_k| + |B_k| + \alpha_k^{q-2} |c_k|^{q-1} + \alpha_k^{q-2} |B_k|^{q-1} \big) |c + G_k| \\ & + \alpha_k^{-2} |\alpha_k B_k|^{q \frac{p-1}{p}} \alpha_k |\mathcal{B} s + \mathcal{B} g_k| + \alpha_k^{-2} |\alpha_k \mathcal{B} b_k|^{p \frac{q-1}{q}} \alpha_k |c + G_k| =: I_1 + I_2 + I_3. \end{aligned}$$

Regarding I_k, $k = 1, 2$, for a given set $A \subset Q$, we apply Young's inequality and we integrate both sides to infer that

$$\int_A I_k \leqslant C\delta + C(\delta) \int_A |\mathcal{B} s + \mathcal{B} g_k|^2 + \alpha_k^{p-2} |\mathcal{B} s + \mathcal{B} g_k|^p + |c + G_k|^2 + \alpha_k^{q-2} |c + G_k|^q,$$

where we used the boundness of the sequences $(\mathcal{B}s_k)$, $(\mathcal{B}b_k)$, (c_k), (B_k) and $(\alpha_k^{p-2}|\mathcal{B}s_k|^p)$, $(\alpha_k^{p-2}|\mathcal{B}b_k|^p)$, $(\alpha_k^{q-2}|c_k|^q)$, $(\alpha_k^{q-2}|B_k|^q)$ in $L^2(Q)$ and $L^1(Q)$ respectively. For I_3, again by Young's inequality, we see that

$$
\begin{aligned}
I_3 &= \alpha_k^{-2}\big(|\alpha_k B_k|^{q\frac{p-1}{p}}\alpha_k|\mathcal{B}s+\mathcal{B}g_k| + |\alpha_k\mathcal{B}b_k|^{p\frac{q-1}{q}}\alpha_k|c+G_k| \big)\\
&\leqslant \alpha_k^{-2}\big(\delta\alpha_k^q|B_k|^q + C(\delta)\alpha_k^p|\mathcal{B}s+\mathcal{B}g_k|^p + \delta\alpha_k^p|\mathcal{B}b_k|^p + C(\delta)\alpha_k^q|c+G_k|^q\big)\\
&= \delta\alpha_k^{q-2}|B_k|^q + C(\delta)\alpha_k^{p-2}|\mathcal{B}s+\mathcal{B}g_k|^p + \delta\alpha_k^{p-2}|\mathcal{B}b_k|^p + C(\delta)\alpha_k^{q-2}|c+G_k|^q.
\end{aligned}
$$

Combing the above we conclude that

$$
\begin{aligned}
\int_A f_k \leqslant C\delta + C(\delta)&\int_A |\mathcal{B}s+\mathcal{B}g_k|^2 + \alpha^{p-2}|\mathcal{B}s+\mathcal{B}g_k|^p\\
&+ C(\delta)\int_A |c+G_k|^2 + \alpha^{q-2}|c+G_k|^q.
\end{aligned}
$$

All sequences appearing on the right hand side are equiintegrable so, choosing $\delta > 0$ appropriately, we see that f_k is equiintegrable. Then, for $\varepsilon > 0$ fixed, we can find $m_\varepsilon > 0$ such that

$$
\int_Q f_k = \int_{A_k^c\cup B_k^c} f_k + \int_{A_k\cap B_k} f_k > -\varepsilon + \int_{A_k\cap B_k} f_k, \tag{3.9}
$$

for all $m \geqslant m_\varepsilon$, where

$$
\begin{aligned}
A_k &= \big\{x\in Q:\ (|\mathcal{B}s_k|^2+|c_k|^2)^{1/2} < m\big\},\\
B_k &= \big\{x\in Q:\ (|\mathcal{B}b_k|^2+|B_k|^2)^{1/2} < m\big\}.
\end{aligned}
$$

This indeed follows from the fact that $\mathcal{B}b_k$, $B_k \to 0$ in measure and that

$$
\lim_{R\to\infty}\sup_k \Big|\{x\in Q:\ (|\mathcal{B}s_k(x)|^2+|c_k(x)|^2)^{1/2} > R\}\Big| = 0,
$$

where the latter holds from Chebyshev's inequality. Now, by choosing m_ε larger if necessary, we assume that

$$
\Big|\int_Q \big\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\,[\Lambda,\Lambda]\mathbb{1}_{\mathbb{R}^{d\times d}\times\mathbb{R}\setminus B(0,m)}(\Lambda)\big\rangle\Big| < \varepsilon, \quad \text{for all } m\geqslant m_\varepsilon, \tag{3.10}
$$

where $\mathbb{1}_A$ denotes the indicator function of a set $A\subset\mathbb{R}^N\times\mathbb{R}$. Indeed, due to the fact that $\int_Q\langle\nu_x,|\Lambda|^2\rangle < \infty$, Young's inequality and dominated convergence give us

$$
\begin{aligned}
\Big|\int_Q \big\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\,[\Lambda,\Lambda]\mathbb{1}_{\mathbb{R}^{d\times d}\times\mathbb{R}\setminus B(0,m)}(\Lambda)\big\rangle\Big| &\leqslant C\int_Q \big|\langle \nu_x, |\Lambda|^2\mathbb{1}_{\mathbb{R}^{d\times d}\times\mathbb{R}\setminus B(0,m)}(\Lambda)\,\rangle\big|\\
&= C\int_Q \big|\langle\nu_x,|\Lambda|^2\,\rangle - \langle\nu_x,|\Lambda|^2\mathbb{1}_{B(0,m)}(\Lambda)\,\rangle\big| \longrightarrow 0
\end{aligned}
$$

and so,

$$
\int_Q \big\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\,\big[\Lambda,\Lambda\big]\big\rangle = \int_Q \big\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\,\big[\Lambda,\Lambda\big]\,\mathbb{1}_{B(0,m)}(\Lambda)\big\rangle
$$

$$+\int_Q \left\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\left[\Lambda,\Lambda\right]\mathbb{1}_{\mathbb{R}^{d\times d}\times\mathbb{R}\setminus B(0,m)}(\Lambda)\right\rangle \leqslant \int_Q \left\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\left[\Lambda,\Lambda\right]\mathbb{1}_{B(0,m)}(\Lambda)\right\rangle + \varepsilon, \tag{3.11}$$

for all $m \geqslant m_\varepsilon$. However, $\mathbb{1}_{B(0,m)}$ is lower semicontinuous and hence, for all $x \in Q$, the function

$$\Lambda \mapsto \tilde{L}(\bar{F},\bar{\eta})[\,\Lambda,\Lambda\,]\mathbb{1}_{B(0,m)}(\Lambda)$$

is also lower semicontinuous. Then, since $(\mathcal{B}s_k, c_k)$ generates $(\nu_x)_x$ we infer that

$$\begin{aligned}\int_Q \left\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\left[\Lambda,\Lambda\right]\mathbb{1}_{B(0,m)}(\Lambda)\right\rangle &\leqslant \liminf_k \int_{A_k} \tilde{L}(\bar{F},\bar{\eta})[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,]\\ &= \liminf_k \int_{A_k} \tilde{L}(\bar{F}_k,\bar{\eta}_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,]\end{aligned}$$

where the last equality follows from the strong convergence $(\bar{F}_k,\bar{\eta}_k) \to (\bar{F},\bar{\eta})$ in $C^0(Q)$. Going back to (3.11) we conclude that

$$\int_Q \left\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\left[\Lambda,\Lambda\right]\right\rangle \leqslant \liminf_k \int_{A_k} \tilde{L}(\bar{F}_k,\bar{\eta}_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,] + \varepsilon, \tag{3.12}$$

for all $m \geqslant m_\varepsilon$. We now claim that

$$\frac{1}{2}\liminf_k \int_{A_k} \tilde{L}(\bar{F}_k,\bar{\eta}_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,] = \liminf_k \int_{A_k\cap B_k} f_k, \tag{3.13}$$

for all $m \geqslant m_\varepsilon$. Indeed, we observe that

$$\begin{aligned}f_k = \int_0^1 (1-t)\Big(\,&\tilde{L}(\bar{F}_k + t\alpha_k\mathcal{B}s_k, \bar{\eta}_k + t\alpha_k c_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,]\\ &- \tilde{L}(\bar{F}_k + t\alpha_k\mathcal{B}b_k, \bar{\eta}_k + t\alpha_k B_k)[\,(\mathcal{B}b_k,B_k),(\mathcal{B}b_k,B_k)\,]\Big)dt.\end{aligned}$$

Then, since $\int_0^1 (1-t)dt = 1/2$, we infer that

$$\begin{aligned}\mathbb{1}_{A_k\cap B_k} f_k = \mathbb{1}_{A_k\cap B_k}\int_0^1 (1-t)\Big(\,&\tilde{L}(\bar{F}_k + t\alpha_k\mathcal{B}s_k, \bar{\eta}_k + t\alpha_k c_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,]\\ &- \tilde{L}(\bar{F}_k,\bar{\eta}_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,]\Big)dt\\ &+ \mathbb{1}_{A_k}\frac{1}{2}\tilde{L}(\bar{F}_k,\bar{\eta}_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,]\\ &- \mathbb{1}_{A_k}\frac{1}{2}\tilde{L}(\bar{F}_k,\bar{\eta}_k)[\,(\mathcal{B}s_k,c_k),(\mathcal{B}s_k,c_k)\,]\,(1-\mathbb{1}_{B_k})\\ &- \mathbb{1}_{A_k\cap B_k}\int_0^1 (1-t)\tilde{L}(\bar{F}_k + t\alpha_k\mathcal{B}b_k, \bar{\eta}_k + t\alpha_k B_k)[\,(\mathcal{B}b_k,B_k),(\mathcal{B}b_k,B_k)\,]\,dt\\ &=: I_1^k + I_2^k + I_3^k + I_4^k,\end{aligned}$$

and so it is enough to prove that

$$\lim_{k\to\infty}\int_Q I_1^k = \lim_{k\to\infty}\int_Q I_3^k = \lim_{k\to\infty}\int_Q I_4^k = 0.$$

Recall that $\alpha_k \to 0$ and $(\bar{F}_k, \bar{\eta}_k) \to (\bar{F}, \bar{\eta})$ in $C^0(Q)$. Thus, for I_1^k and since we are in the set A_k, we find that

$$\Big| \tilde{L}(\bar{F}_k + t\alpha_k \mathcal{B}s_k, \bar{\eta}_k + t\alpha_k c_k)[\,(\mathcal{B}s_k, c_k), (\mathcal{B}s_k, c_k) - \tilde{L}(\bar{F}_k, \bar{\eta}_k)[\,(\mathcal{B}s_k, c_k), (\mathcal{B}s_k, c_k)\,] \Big|$$
$$\leqslant C(K)\alpha_k m^3 \to 0, \;\; k \to \infty,$$

and thus, $\int_Q I_1^k \to 0$ by dominated convergence. As for I_3^k, again since $\tilde{L}$ is continuous, $\|(\bar{F}, \bar{\eta})\|_{L^\infty(Q)} \leqslant K$ and we are in A_k, we get that

$$|I_3^k| \leqslant C(W,K)m^2\,(1 - \mathbb{1}_{B_k}) = C(W,K)m^2 \mathbb{1}_{B_k^c}.$$

Hence, $\int_Q I_3^k \to 0$ as $\mathcal{B}b_k \to 0$ and $B_k \to 0$ in measure. We note here that

$$\{x \in Q : |\mathcal{B}b_k(x)|^2 + |B_k(x)|^2 \geqslant \varepsilon^2\}$$
$$\subseteq \{x \in Q : |\mathcal{B}b_k(x)|^2 \geqslant \varepsilon^2/2\} \cup \{x \in Q : |B_k(x)|^2 \geqslant \varepsilon^2/2\}.$$

Lastly, for I_4^k, as we are in B_k and $t \in (0,1)$, we get that

$$(\bar{F}_k, \bar{\eta}_k) + t\alpha_k(\mathcal{B}b_k,\, B_k) \to (\bar{F}, \bar{\eta}), \;\; k \to \infty,$$

uniformly and thus

$$|I_4^k| \leqslant C(W,K)(|\mathcal{B}b_k|^2 + |B_k|^2) \leqslant C(W,K)m(|\mathcal{B}b_k| + |B_k|) \to 0$$

in measure. In particular, restricting to B_k, $\int_Q I_4^k \to 0$ by dominated convergence. Finally combining (3.12), (3.13) and (3.9) we infer that

$$\frac{1}{2}\int_Q \big\langle \nu_x, \tilde{L}(\bar{F}, \bar{\eta})\,[\Lambda, \Lambda]\big\rangle \leqslant \liminf_k \int_{A_k \cap B_k} f_k + \varepsilon/2$$
$$< \liminf_k \int_Q f_k + 3\varepsilon/2,$$

for all $m \geqslant m_\varepsilon$. Since ε is arbitrary and the dependence on m_ε has been removed, we take the limit $\varepsilon \to 0$ to deduce that

$$\frac{1}{2}\int_Q \big\langle \nu_x, \tilde{L}(\bar{F}, \bar{\eta})\,[\Lambda, \Lambda]\big\rangle \leqslant \liminf_k \int_Q f_k.$$

Combining the above inequality with (3.8), and since $s_k \to s$ in $W^{l-1,2}(Q)$, we conclude that

$$\frac{1}{2}\int_Q \big\langle \nu_x, \tilde{L}(\bar{F}, \bar{\eta})\,[\Lambda, \Lambda]\big\rangle + C_{pen}\sum_{i=1}^{l} |\nabla^{l-i}s|^2 < 0. \tag{3.14}$$

Step 5: In this step we show how (3.14) leads to a contradiction. By Lemma 3.5, the function

$$h(x,\xi_1,\xi_2) := \tilde{L}(\bar{F}(x),\bar{\eta}(x))[\,(\xi_1,\xi_2),(\xi_1,\xi_2)\,]$$

is strongly $(\mathcal{A},0)$-quasiconvex for each $x \in Q$. Hence, since $(\mathcal{B}s_k, c_k)$ generates the $(\mathcal{A},0)$-2-Young measure $(\nu_x)_x$ and $h(x,\xi_1,\xi_2)$ grows quadratically in (ξ_1,ξ_2), Jensen's inequality for $(\mathcal{A},0)$-quasiconvex functions, Theorem 2.9(c), says that, for a.e. $x \in Q$,

$$\tilde{L}(\bar{F},\bar{\eta})[\,(\mathcal{B}s,c),(\mathcal{B}s,c)\,] \leqslant \big\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\,\big[(\lambda_F,\lambda_\eta),(\lambda_F,\lambda_\eta)\big]\big\rangle.$$

Adding $C_{pen}\sum_{i=1}^{l}|\nabla^{l-i}s|^2$ on both sides and integrating over Q, we infer that

$$\begin{aligned}\int_Q C_{pen}|s|^2 + &\tilde{L}(\bar{F},\bar{\eta})[\,(\mathcal{B}s,c),(\mathcal{B}s,c)\,]\\ &\leqslant \int_Q C_{pen}\sum_{i=1}^{l}|\nabla^{l-i}s|^2 + \big\langle \nu_x, \tilde{L}(\bar{F},\bar{\eta})\,\big[(\lambda_F,\lambda_\eta),(\lambda_F,\lambda_\eta)\big]\big\rangle \overset{(3.14)}{\leqslant} 0.\end{aligned}$$

However, by Proposition 3.1, since $\mathcal{B}s \in W^{l,p}(Q)$ and $c \in L^q(Q)$, we know that

$$0 \geqslant \int_Q \tilde{L}(\bar{F}(x),\bar{\eta}(x))\big[(\mathcal{B}s,c),(\mathcal{B}s,c)\big] + C_{pen}\sum_{i=1}^{l}|\nabla^{l-i}s|^2 \geqslant c\int_Q |\mathcal{B}s|^2 + |c|^2.$$

In particular, we see that $\mathcal{B}s = 0$ and $c = 0$, and so $s = \mathcal{F}^{-1}(\mathcal{B}^\dagger(\cdot)) \star \mathcal{B}s = 0$. We may thus apply Proposition 3.2 for the sequences $(\bar{F}_k,\bar{\eta}_k)_k$, $(\alpha_k s_k)_k$, $(\alpha_k c_k)_k$ and $(\alpha_k)_k$. Recall that $a_k^{p-2} \leqslant 2^{\frac{p-2}{2}}\Lambda^F|M_k^F|^p < \infty$ and hence

$$\alpha_k^{-2}|V_p(\alpha_k\nabla^{l-i}s_k)|^2 = |\nabla^{l-i}s_k|^2 + \alpha_k^{p-2}|\nabla^{l-i}s_k|^p \lesssim |\nabla^{l-i}s_k|^2 + \Lambda^F|M_k^F\nabla^{l-i}s_k|^p \to 0,$$

in $L^1(Q)$. Also,

$$\begin{aligned}\sup_k \int_Q \alpha_k^{-2}|V_p(\alpha_k\mathcal{B}s_k)|^2 &\lesssim \sup_k\int_Q |s_k|^2 + \Lambda^F \sup_k |M_k^F s_k|^p < \infty,\\ \sup_k \int_Q \alpha_k^{-2}|V_q(\alpha_k c_k)|^2 &\lesssim \sup_k\int_Q |c_k|^2 + \Lambda^\eta \sup_k |M_k^\eta c_k|^q < \infty,\end{aligned}$$

and recalling that $\alpha_k s_k = \phi_k$ and $\alpha_k c_k = \psi_k$, Proposition 3.2 yields

$$\begin{aligned}0 < \frac{c_0}{4} &= \liminf_{k\to\infty} \frac{\tilde{c}}{4}\int_Q |\mathcal{B}s_k|^2 + |c_k|^2\\ &\leqslant \liminf_{k\to\infty} \frac{\tilde{c}}{4}\int_Q |\mathcal{B}s_k|^2 + |c_k|^2 + \alpha_k^{p-2}|\mathcal{B}s_k|^p + \alpha_k^{q-2}|c_k|^q\\ &\leqslant \liminf_{k\to\infty} \alpha_k^{-2}\int_Q \tilde{e}(\bar{F}_k + \alpha_k\mathcal{B}s_k, \bar{\eta}_k + \alpha_k c_k|\bar{F}_k,\bar{\eta}_k)\\ &= \liminf_{k\to\infty} \alpha_k^{-2}\int_Q \tilde{e}(\bar{F}_k + \mathcal{B}\phi_k, \bar{\eta}_k + \psi_k|\bar{F}_k,\bar{\eta}_k) + \frac{C_{pen}}{2}\sum_{i=1}^{l}\int_Q |\nabla^{l-i}s|^2 \overset{(3.4)}{\leqslant} 0.\end{aligned}$$

But $c_0 > 0$, concluding the proof. □

We are now ready to prove our main result, the Gårding inequality. Note here that it suffices to show that the Gårding inequality of Theorem 3.1 remains valid for test functions ϕ such that $\|\phi\|_{L^p(Q)} > \varepsilon_0$.

Theorem 3.2. *Let e satisfy (H_1)-(H_3) be strongly quasiconvex. Then, there exist constants $C_0 = C_0(e,K) > 0$, $C_1 = C_1(e,K) > 0$ such that for all $(\bar{F},\bar{\eta}) \in \mathcal{U}_K$, $\psi \in L^q(Q)$ with $\int_Q \psi = 0$ and all $\varphi \in L^p_{\mathcal{A}}(Q)$, it holds that*

$$\begin{aligned}\int_Q &\Big(|V_p(\varphi(x))|^2 + |V_q(\psi(x))|^2\Big)dx \\ &\leqslant C_0 \int_Q e(\bar{F}(x)+\varphi(x),\bar{\eta}(x)+\psi(x)|\bar{F}(x),\bar{\eta}(x))dx + C_1\|\varphi\|^2_{W^{-1,(2,p)}}. \end{aligned} \tag{3.15}$$

Proof. We claim that for all $\varepsilon > 0$, all $(\bar{F},\bar{\eta}) \in \mathcal{U}_K$, $\psi \in L^q(Q)$ zero-average and all $\varphi \in L^p_{\mathcal{A}}(Q)$ with $\|\varphi\|_{W^{-1,p}(Q)} \geqslant \varepsilon$ it holds that

$$\int_Q \Big(|V_p(\varphi)|^2 + |V_q(\psi)|^2\Big) \leqslant C_0(\varepsilon)\int_Q e(\bar{F}+\varphi,\bar{\eta}+\psi|\bar{F},\bar{\eta}) + C_1(\varepsilon)\|\varphi\|^2_{W^{-1,(2,p)}},$$

where C_0 and C_1 also depend on ε. By Lemma 2.1 (i), we find $\phi \in W^{l,p}(Q)$ such that $\varphi = \mathcal{B}\phi$ and by the assumed coercivity of e, its smoothness and the fact that $(\bar{F},\bar{\eta}) \in \mathcal{U}_K$, we estimate by Young's inequality

$$\begin{aligned} e(\bar{F}+\mathcal{B}\phi,\bar{\eta}+\psi|\bar{F},\bar{\eta}) &\geqslant c\left(-1+|\bar{F}+\mathcal{B}\phi|^p + |\bar{\eta}+\psi|^q\right) \\ &\quad - C(\delta)|e_1(\bar{F},\bar{\eta})|^{p'} - \delta|\mathcal{B}\phi|^p - C(\delta)|e_2(\bar{F},\bar{\eta})|^{q'} - \delta|\psi|^q \\ &\geqslant -C(\delta) + (2^{1-p}c-\delta)|\mathcal{B}\phi|^p + (2^{1-q}c-\delta)|\psi|^q. \end{aligned}$$

Since $p \geqslant q$, choose $\delta = 2^{-p}c$ to find that

$$e(\bar{F}+\mathcal{B}\phi,\bar{\eta}+\psi|\bar{F},\bar{\eta}) \geqslant -C + 2^{-p}c|\mathcal{B}\phi|^p + 2^{-q}c|\psi|^q. \tag{3.16}$$

Note that since $\|\mathcal{B}\phi\|_{W^{-1,p}(Q)} \geqslant \varepsilon$, it follows that

$$C \leqslant \frac{C}{\varepsilon^p}\|\mathcal{B}\phi\|^p_{W^{-1,p}(Q)},$$

so that, integrating (3.16) over Q with $|Q| = 1$, we infer that

$$2^{-p}c\int_Q |\mathcal{B}\phi|^p + 2^{-q}c\int_Q |\psi|^q \leqslant \int_Q e(\bar{F}+\mathcal{B}\phi,\bar{\eta}+\psi|\bar{F},\bar{\eta}) + \frac{C}{\varepsilon^p}\|\mathcal{B}\phi\|^p_{W^{-1,p}(Q)}. \tag{3.17}$$

But, by the virtue of the compact embedding $L^p(Q) \hookrightarrow W^{-1,p}(Q)$, we have that

$$1 \leqslant \frac{1}{\varepsilon^p}\|\mathcal{B}\phi\|^p_{W^{-1,p}(Q)} \lesssim \frac{1}{\varepsilon^p}\|\mathcal{B}\phi\|^p_{L^p(Q)}$$

and so

$$\int_Q |V_p(\mathcal{B}\phi)|^2 = \|\mathcal{B}\phi\|^2_{L^2(Q)} + \|\mathcal{B}\phi\|^p_{L^p(Q)} \leqslant 1 + 2\|\mathcal{B}\phi\|^p_{L^p(Q)} \lesssim \left(\frac{1}{\varepsilon^p}+2\right)\|\mathcal{B}\phi\|^p_{L^p(Q)},$$

and similarly

$$\int_Q |V_q(\psi)|^2 \lesssim \frac{1}{\varepsilon^p}\|\mathcal{B}\phi\|^p_{L^p(Q)} + \|\psi\|^q_{L^q(Q)}.$$

So combining the above two results,

$$\begin{aligned}\int_Q |V_p(\mathcal{B}\phi)|^2 + \int_Q |V_q(\psi)|^2 &\lesssim \left(\frac{1}{\varepsilon^p} + 1\right)\|\mathcal{B}\phi\|^p_{L^p(Q)} + \|\psi\|^q_{L^q(Q)} \\ &\leqslant C(\varepsilon)\left(\|\mathcal{B}\phi\|^p_{L^p(Q)} + \|\psi\|^q_{L^q(Q)}\right).\end{aligned}$$

Together with (3.17) we infer that

$$C^{-1}(\varepsilon)\int_Q |V_p(\mathcal{B}\phi)|^2 + |V_q(\psi)|^2 \leqslant \int_Q e(\bar{F} + \mathcal{B}\phi, \bar{\eta} + \psi|\bar{F}, \bar{\eta}) + \frac{C}{\varepsilon^p}\|\mathcal{B}\phi\|^p_{W^{-1,p}(Q)},$$

which is the desired inequality. However, Theorem 3.1 says that there exists $\varepsilon_0 > 0$ such that whenever $\|\varphi\|_{W^{-1,p}(Q)} \leqslant \varepsilon_0$ it holds that

$$\int_Q |V_p(\varphi)|^2 + |V_q(\psi)|^2 \leqslant \tilde{C}_0 \int_Q e(\bar{F} + \varphi, \bar{\eta} + \psi|\bar{F}, \bar{\eta}) + \tilde{C}_1\|\varphi\|^2_{W^{-1,(2,p)}}.$$

Choosing $\varepsilon = \varepsilon_0$ we conclude the proof. $\square$

Chapter 4

Applications in dynamics and statics

4.1 Dynamics: conservation laws with involutions

This section is devoted to applications of Theorem 3.2 in dynamics. In particular, we study systems of conservation laws for which the corresponding entropy is not convex, and hence additional structure in such systems compensate this lack of convexity. This special structure is enforced via PDE constraints propagated by the initial data, and lead to more general notions of convexity, which in our case are different versions of quasiconvexity. In the first part of the section we use a simplified version of Theorem 3.2 to prove stability and weak-strong uniqueness for general systems of conservation laws with involutions, typical examples are the (non-)linear elasticity and Maxwell equations. In the second part, the combination of our Gårding-type inequality, Theorem 3.2, with the so-called relative entropy method leads to weak-strong uniqueness results for a class of dissipative measure-valued solutions for the system of adiabatic thermoelasticity.

We note that the relative entropy method provides a general structure upon which one can compare two solutions in various frameworks [35, 36, 49] as it can be interpreted as a "metric" measuring the distance between two solutions. The calculation provides means to control the norm of the difference of two solutions by the initial data, and the most important ingredient is the convexity of the entropy. Our Gårding inequality, Theorem 3.2, compensates this lack of convexity.

4.1.1 General systems of conservations laws with involutions

Here, we study local stability and weak-strong uniqueness properties for general systems of conservation laws (4.1) possessing involutions (4.3) and an $\mathcal{A}$-quasiconvex entropy. In particular, for $T > 0$ and Q the flat torus, we examine the system

$$\begin{aligned} \partial_t U(t,x) + \mathrm{div}_x f(U(t,x)) = 0, \quad (t,x) \in (0,T) \times Q \\ U(0,x) = U^0(x), \quad x \in Q \end{aligned} \tag{4.1}$$

for the unknown periodic function $U : (0,T) \times Q \to \mathbb{R}^N$ with

$$\int_Q U(t,x)\, dx = 0, \text{ for all } 0 < t \leqslant T. \tag{4.2}$$

In (4.1), the flux function $f = (f_{i\alpha})_{(i,\alpha)\in\mathbb{R}^{N\times d}} : \mathbb{R}^N \to \mathbb{R}^{N\times d}$ is a given C^3 mapping. We say that system (4.1) possesses an involution if there exists a linear differential operator $\mathcal{A}$ with the property that

$$\mathcal{A}U^0 = 0 \;\Rightarrow\; \mathcal{A}U(t,\cdot) = 0 \quad \text{for all } t \in (0,T). \tag{4.3}$$

Note that in continuum mechanics, systems like (4.1), are typically supplemented with an inequality of the form

$$\partial_t \eta + \mathrm{div}_x q \leqslant 0, \tag{4.4}$$

known as the Clausius-Duhem inequality, expressing the second law of thermodynamics in this context. Mathematically, $\eta : \mathbb{R}^N \to \mathbb{R}$ is referred to as an entropy and $q : \mathbb{R}^N \to \mathbb{R}^d$ as an entropy flux and are assumed to satisfy

$$\frac{\partial q_\alpha}{\partial U_i} = \frac{\partial \eta}{\partial U_j}\frac{\partial f_{j\alpha}}{\partial U_i}. \tag{4.5}$$

In particular,

$$\frac{\partial f_{j\alpha}}{\partial U_i}\frac{\partial^2 \eta}{\partial U_k \partial U_j} = \frac{\partial^2 \eta}{\partial U_i \partial U_j}\frac{\partial f_{j\alpha}}{\partial U_k} \tag{4.6}$$

and thus Lipschitz solutions to (4.1) satisfy (4.4) as an equality.

Typical examples include the equations of elasticity and electromagnetism, see [39]. Indeed, the equations of motion of a hyperelastic body in the absence of external forces take the form $y_{tt} = \mathrm{div} DW(\nabla y)$ where W denotes the stored energy function. Upon the change of variables $v = y_t$ and $F = \nabla y$, we obtain the system

$$\begin{aligned} \partial_t v - \mathrm{div}_x DW(F) &= 0, \\ \partial_t F - \nabla v &= 0, \\ \mathrm{curl}\, F &= 0. \end{aligned}$$

The second equation shows that $\mathcal{A} = \mathrm{curl}$ is an involution. Similarly, in linear elasticity, the equations take the form

$$\partial_t u - \mathrm{div}_x CE = 0,$$

$$\partial_t E - \mathcal{E}(u) = 0,$$
$$\operatorname{curl}\operatorname{curl} E = 0,$$

where $2\mathcal{E}(u) = \nabla u + (\nabla u)^T$ and $\mathcal{A} = \operatorname{curl}\operatorname{curl}$ is an involution whose kernel consists of symmetric gradients. Also, note that a natural assumption on the quadratic form $CE \cdot E$ is convexity on the wave cone of the operator curl curl which, by Lemma 2.3, is equivalent to curl curl-quasiconvexity. Moreover, the equations of electromagnetism in the absence of charges and currents become

$$\partial_t B + \operatorname{curl} E = 0,$$
$$\partial_t D - \operatorname{curl} H = 0,$$
$$\operatorname{div} B = \operatorname{div} D = 0,$$

where B is the magnetic induction, D is the electric displacement, and E, H are, respectively, the electric and magnetic fields. Typically, Maxwell's equations are assumed linear, however, there are relevant nonlinear theories, see [27], [96], [39], with the so-called Maxwell's equations in the Born-Infeld medium being the most known. The reader is referred to [17] for a mathematical treatment.

Entropies in physical systems are often convex which, combined with (4.6), renders the system symmetrisable upon the change of variables $U \to D\eta(U)$ and hence locally well-posed, see [39]. At the same time, inequality (4.4) restricts admissible solutions and may rule out unphysical solutions.

On the other hand, it is also known that convexity of the entropy may be ruled out as a consequence of physical invariance. This is precisely the case in nonlinear elasticity due to frame-indifference [39], and in electromagnetism due to Lorentz invariance [96]. However, the presence of involutions may compensate this loss of convexity, but only in the directions where the operator $\mathcal{A}$ has elliptic behaviour. Essentially, the "bad" behaviour is expected to occur in the directions of the wave cone $\Lambda_{\mathcal{A}}$, and convexity along these directions, i.e. $\Lambda_{\mathcal{A}}$-convexity, may be enough to partially recover results ensured by convexity.

Indeed, Dafermos in [38] examined such systems endowed with a $\Lambda_{\mathcal{A}}$-convex entropy and, under additional assumptions on the involutions $\mathcal{A}$, recovered hyperbolicity. Moreover, he showed that local stability and weak-strong uniqueness results can also be recovered within a class of BV weak solutions, if they satisfy an assumption of *small local oscillations*, required to prove a Gårding-type inequality for $\Lambda_{\mathcal{A}}$-convex functions. In this section, we show that in fact this assumption is redundant when the entropy is $\mathcal{A}$-quasiconvex. In this sense, $\mathcal{A}$-quasiconvexity captures the structure of these systems and arises as a natural convexity condition.

We note that Maxwell's equations do not generally fall under this setting. For vector fields B, $D : \mathbb{R}^3 \to \mathbb{R}^3$, the wave cone of $\mathcal{A} = \operatorname{div}$ is the entire space $\mathbb{R}^6$ and thus $\mathcal{A}$-quasiconvexity and $\Lambda_{\mathcal{A}}$-convexity reduce to convexity. However, when B, $D : \mathbb{R}^2 \to \mathbb{R}^3$, the wave cone is strictly smaller than $\mathbb{R}^6$. Still, it is a matter of tedious computations to show that the entropy at least for the Born-Infeld medium is not even $\Lambda_{\mathcal{A}}$-convex and thus, unlike polyconvex elasticity, not convex in the null-Lagrangians of $\mathcal{A} = \operatorname{div}$. Nevertheless, similar to polyconvex elasticity, the system can be extended to an enlarged system that admits a convex entropy, see [39, 96].

Weak-strong uniqueness

In the sequel, we assume that an entropy-entropy flux pair exists satisfying (4.5) and that η satisfies the assumptions:

(h1) $\eta \in C^3(\mathbb{R}^N)$;

(h2) $|\eta(z)| \leqslant c_1(1+|z|^p)$ and $|D\eta(z)| \leqslant c(1+|z|^{p-1})$;

(h3) $c_2(|z|^p - 1) \leqslant \eta(z)$;

(h4) η is strongly $\mathcal{A}$-quasiconvex;

Remark 4.1. Recall that, as discussed in §2.2, if $\Lambda_{\mathcal{A}}$ spans $\mathbb{R}^N$, the growth on $D\eta$ in (h2) follows from (h1) and the growth of η in (h2).

We note that the assumptions $(h1)$, $(h2)$ and $(h3)$ are the same with the respective assumptions $(H1)$, $(H2)$ and $(H3)$ in section 3.3, if in the latter we remove the z_2-dependance. In addition to this, the same holds for the quasiconvexity assumption $(h4)$, inequality (2.2), and the respective quasiconvexity assumption in section 3.1, Definition 3.1. This observation is crucial in the sequel, since it allow us to apply Theorem 3.2 for the entropy η which is crucial for the proof of our weak-strong uniqueness result. In this case, inequality (3.15) takes the form

$$\int_Q |V_p(\varphi(x))|^2 dx \leqslant C_0 \int_Q \eta(\bar{F}(x) + \varphi(x)|\bar{F}(x))dx + C_1 \|\varphi\|^2_{W^{-1,(2,p)}}, \tag{4.7}$$

for all $\bar{F} \in \mathcal{F}_K$ and all $\mathcal{A}$-free functions $\varphi \in L^p(Q)$ with $\int_Q \varphi = 0$. Here

$$\mathcal{F}_K := \{\bar{F} \in C(Q) : \|\bar{F}\|_{L^\infty(Q)} \leqslant K,\, |\bar{F}(x) - \bar{F}(y)| \leqslant \omega(|x-y|),\, \forall\, x, y \in \overline{Q}\},$$

constitutes the projection of the set $\mathcal{U}_K$, defined in section 3.3, on $\mathbb{R}^N$.

Moreover, as in [38], we assume that weak solutions are bounded. We refer the reader to Remark 4.2 following the proof for a discussion on these assumptions.

Definition 4.1. Let $U \in L^\infty((0,T) \times Q)$. We say that the function U is a dissipative weak solution to (4.1) with initial data U^0 if

$$\int_Q \phi_i(0,\cdot)U_i^0 + \int_0^T \int_Q \partial_t \phi_i \cdot U_i + \int_0^T \int_Q \partial_\alpha \phi_i \cdot f_{i\alpha}(U) = 0 \tag{4.8}$$

for any $\phi \in C_c^1([0,T), C^1(Q))$ and i=1,..,N, and the dissipation inequality

$$\int_Q \theta(0)\eta(U^0) + \int_0^T \int_Q \dot{\theta}\, \eta(U) \geqslant 0 \tag{4.9}$$

holds for any nonnegative test function $\theta \in C_c^1([0,T))$.

Recall that Lipschitz solutions $\bar{U} \in W^{1,\infty}([0,T] \times \overline{Q})$ satisfy (4.9) as an equality, that is

$$\int_Q \theta(0)\eta(\bar{U}^0) + \int_0^T \int_Q \dot{\theta}\, \eta(\bar{U}) = 0. \tag{4.10}$$

Moreover, note that if $\int_Q U^0 = 0$ then also $\int_Q U(t,\cdot) = 0$ for a.e. $t \in (0,T)$. This follows by testing (4.8) with $\phi(t,x) = \theta(t)$ where θ localises at a fixed time, as in (4.13). The main theorem of this section now follows, cf. [38, Theorem 4.1].

Theorem 4.2. *Let $\bar{U} \in W^{1,\infty}([0,T]\times\overline{Q})$ and $U \in L^\infty((0,T)\times Q)$ be, respectively, a strong and a dissipative weak solution of* (4.1) *emanating from the zero-average initial data $\bar{U}^0$, $U^0 \in L^\infty(Q)$. Assume that U and $\bar{U}$ satisfy the PDE constraint $\mathcal{A}\bar{U} = \mathcal{A}U = 0$, and that the entropy η satisfies (h1)-(h4). Then, there exist constants C_1, $C_2 > 0$ such that for almost all $t \in (0,T)$*

$$\int_Q |V_p(U(t,\cdot) - \bar{U}(t,\cdot))|^2 \leqslant C_1 \int_Q |V_p(U^0 - \bar{U}^0)|^2\, e^{C_2 t},$$

see (2.1) *for the definition of V_p.*

Proof. Let U and $\bar{U}$ as in the statement and test the equations (4.8) with the function $\phi(t,x) = \theta(t) D\eta(\bar{U}(t,x))$, where $\theta \in C_c^1([0,T))$. Note that this is an appropriate test function by density. Subtracting the equations for U from the equations for $\bar{U}$, we infer that

$$\begin{aligned}
&\int_Q \theta(0)\, D_j\eta(\bar{U}^0)\,(U_j^0 - \bar{U}_j^0) + \int_0^T \int_Q \dot{\theta}\, D_j\eta(\bar{U})\,(U_j - \bar{U}_j) \\
&= -\int_0^T \int_Q \theta \left\{ \partial_\alpha D_k\eta(\bar{U}) \big(f_{k\alpha}(U) - f_{k\alpha}(\bar{U})\big) + \partial_t D_j\eta(\bar{U})\,(U_j - \bar{U}_j) \right\},
\end{aligned}$$

where ∂_α, ∂_t and D_j stand for the operators $\frac{\partial}{\partial x_\alpha}$, $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial U_j}$ respectively. By (4.6), we observe that

$$\begin{aligned}\partial_t D_j \eta(\bar{U}) = \frac{\partial \bar{U}_k}{\partial t} \frac{\partial \eta(\bar{U})}{\partial U_j \partial U_k} &\overset{(4.1)}{=} -\partial_\alpha \bar{U}_i \frac{\partial f_{k\alpha}(\bar{U})}{\partial U_i} \frac{\partial^2 \eta(\bar{U})}{\partial U_j \partial U_k} \\ &= -\partial_\alpha \bar{U}_i \frac{\partial^2 \eta(\bar{U})}{\partial U_i \partial U_k} \frac{\partial f_{k\alpha}(\bar{U})}{\partial U_j} = -\left[\partial_\alpha D_k \eta(\bar{U})\right] D_j f_{k\alpha}(\bar{U})\end{aligned}$$

and thus

$$\begin{aligned}&\int_Q \theta(0)\, D_j \eta(\bar{U}^0)\,(U_j^0 - \bar{U}_j^0) + \int_0^T \int_Q \dot{\theta}\, D_j \eta(\bar{U})\,(U_j - \bar{U}_j) \\ &= -\int_0^T \int_Q \theta \left[\partial_\alpha D_k \eta(\bar{U})\right] f_{k\alpha}(U|\bar{U}) =: \mathcal{R},\end{aligned} \tag{4.11}$$

where $f_{k\alpha}(U|\bar{U}) := f_{k\alpha}(U) - f_{k\alpha}(\bar{U}) - D_j f_{k\alpha}(\bar{U})\,(U_j - \bar{U}_j)$ is the relative flux. This complies with the notation in the previous section as $U = \bar{U} + (U - \bar{U})$. By using the definition of the relative entropy, testing with $\theta \in C_c^1([0,T))$ and integrating in space and time, we get that

$$\begin{aligned}&\int_0^T \int_Q \dot{\theta}\, \eta(U|\bar{U}) + \int_Q \theta(0)\, \eta(U^0|\bar{U}^0) \\ &= \int_0^T \int_Q \dot{\theta} \eta(U) + \int_Q \theta(0) \eta(U^0) - \left(\int_Q \dot{\theta} \eta(\bar{U}) + \int_Q \theta(0) \eta(\bar{U}^0)\right) \\ &- \int_0^T \int_Q \dot{\theta}\, D_j \eta(\bar{U})\,(U_j - \bar{U}_j) - \int_Q \theta(0)\, D_j \eta(\bar{U}^0)\,(U_j^0 - \bar{U}_j^0) \\ &\overset{(4.11)}{\geqslant} -\mathcal{R},\end{aligned} \tag{4.12}$$

where the quantity in the second line, due to the relations (4.9) and (4.10), is non-negative. Recall that the relative entropy is given by

$$\eta(U|\bar{U}) = \eta(U) - \eta(\bar{U}) - D_j \eta(\bar{U})(U_j - \bar{U}_j).$$

We next follow a standard argument to localise in time. Let $(\theta_m)_{m\in\mathbb{N}} \subset C_c^\infty([0,T))$ be a bounded sequence approximating the function

$$\theta(\tau) = \begin{cases} 1, & \tau \in [0,t) \\ (t-\tau)/\epsilon + 1, & \tau \in [t, t+\epsilon) \\ 0, & \tau \in [t+\epsilon, T) \end{cases} \tag{4.13}$$

such that $(\theta_m)_m$ is nonincreasing and $\dot{\theta}_m(\tau) \to \dot{\theta}(\tau)$ for all $\tau \neq t, t+\epsilon$. Note that $\dot{\theta}_m \leqslant 0$ and so testing (4.12) with θ_m we find that

$$\int_0^T \int_Q |\dot{\theta}_m(\tau)|\, \eta(U(\tau,x)|\bar{U}(\tau,x))\, dx d\tau \leqslant \mathcal{R} + \int_Q \theta_m(0) \eta(U^0(x)|\bar{U}^0(x))\, dx. \tag{4.14}$$

Since $U \in L^\infty((0,T)\times Q)$ and $\partial_\alpha D_k \eta(\bar{U})$ is bounded, for the locally Lipschitz function $f_{k\alpha}$, we compute from (4.11) that

$$|\mathcal{R}| \leqslant C \int_0^T \int_Q |\theta|\, |U-\bar{U}|^2.$$

As U is bounded, taking the limit $m \to \infty$ in (4.14) by dominated convergence, gives

$$\frac{1}{\epsilon}\int_t^{t+\epsilon}\int_Q \eta(U|\bar{U}) \leqslant C\int_0^{t+\epsilon}\int_Q |U-\bar{U}|^2 + \int_Q \eta(U^0|\bar{U}^0).$$

Then, sending $\epsilon \to 0$, we get that for almost all $t \in (0,T)$,

$$\int_Q \eta(U|\bar{U}) \leqslant C\int_0^t\int_Q |U-\bar{U}|^2 + \int_Q \eta(U^0|\bar{U}^0).$$

Note that since η satisfies (h1)-(h4), by Lemma 3.3 (a), $\eta(U^0|\bar{U}^0) \lesssim |V_p(U^0-\bar{U}^0)|^2$ and thus, integrating (4.7) (Theorem 3.2) in time, we deduce that for almost all $t \in (0,T)$ and up to a suitable constant

$$\int_Q |V_p(U-\bar{U})|^2 \lesssim \int_0^t\int_Q |U-\bar{U}|^2 + \int_Q |V_p(U^0-\bar{U}^0)|^2 + \|U-\bar{U}\|^2_{W^{-1,(2,p)}}, \tag{4.15}$$

where $\|\cdot\|_{W^{-1,(2,p)}}$ is the auxiliary mapping defined in (3.3). In order to apply Grönwall's inequality and conclude the proof, it remains to estimate the last term on the right-hand side of (4.15). Similarly to Dafermos in [38], for $r \in \{2,p\}$, we infer that, since $L^r(Q)$ embeds into $W^{-1,r}(Q)$,

$$\|U(t,\cdot)-\bar{U}(t,\cdot)\|_{W^{-1,r}} \lesssim \|U^0-\bar{U}^0\|_{L^r} + \int_0^t \|\partial_t\{U(s,\cdot)-\bar{U}(s,\cdot)\}\|_{W^{-1,r}}ds.$$

By taking into account (4.1) we deduce the bound

$$\begin{aligned}\|\partial_t\{U(s,\cdot)-\bar{U}(s,\cdot)\}\|_{W^{-1,r}(Q)} &\leqslant \|\partial_\alpha f_{i\alpha}(U)-\partial_\alpha f_{i\alpha}(\bar{U})\|_{W^{-1,r}(Q)}\\ &\leqslant C\|f_{i\alpha}(U)-f_{i\alpha}(\bar{U})\|_{L^r(Q)}\\ &\leqslant C\|U(s,\cdot)-\bar{U}(s,\cdot)\|_{L^r(Q)},\end{aligned} \tag{4.16}$$

where the last inequality follows from the fact that f is locally Lipschitz and U is bounded. Finally, by Hölder's inequality, we infer that

$$\|U(t,\cdot)-\bar{U}(t,\cdot)\|^r_{W^{-1,r}} \lesssim \|U^0-\bar{U}^0\|^r_{L^r} + T^{r-1}\int_0^t \|U(s,\cdot)-\bar{U}(s,\cdot)\|^r_{L^r}ds.$$

Returning to (4.15) and applying the above bound for $r=2$ and $r=p$ we arrive at

$$\int_Q |V_p(U-\bar{U})|^2 \lesssim \int_0^t\int_Q |V_p(U-\bar{U})|^2 + \int_Q |V_p(U^0-\bar{U}^0)|^2.$$

An application of Grönwall's inequality completes the proof. □

Remark 4.2. Note that the L^∞ bounds on weak solutions are needed in the estimate (4.16). Moreover, we note that the assumed growths on η do not e.g. directly apply to elasticity where $\eta(v,F) = \frac{1}{2}|v|^2 + W(F)$. However, as $|v|^2$ is convex, it is immediate to deduce the result assuming (h1)-(h4) on W [74].

Remark 4.3. **(L^p bounds and elliptic estimates)** Here, we propose an additional structure to the PDE that allows one to prove weak-strong uniqueness for merely L^p bounded weak solutions of (4.1). For simplicity, we focus on $p = 2$. For $p > 2$ similar arguments were used in [74] where we refer the reader. To be more precise, instead of the PDE constraint (4.3) we assume that

$$\partial_t U(t,x) - \mathcal{B}g(U(t,x)) = 0, \tag{4.17}$$

where $\mathcal{B}$ is a first-order elliptic potential operator of $\mathcal{A}$ and $g : \mathbb{R}^N \to \mathbb{R}^N$ is globally Lipschitz. Then, the estimates of Lemma 2.1 on the primitives follow directly due to the ellipticity of $\mathcal{B}$ and in particular due to the fact that $\mathbb{B}^*(\xi)\mathbb{B}(\xi)$ is invertible for all $\xi \in \mathbb{R}^d \setminus \{0\}$. We claim that these assumptions suffice to bound the term $\|U - \bar{U}\|^2_{W^{-1,2}}$ and replace estimate (4.16) without any L^∞ assumptions.

Indeed, for zero-average $U \in L^\infty(0,T;L^2(Q))$ and $W \in L^\infty(0,T;W^{1,2}(Q))$ a primitive of U, equation (4.17) implies that

$$\int_0^T \int_Q (W - \bar{W})\mathcal{B}^*\psi_t + \int_0^T \int_Q \left(g(U) - g(\bar{U})\right) \mathcal{B}^*\psi = 0,$$

for all $\psi \in C_c^\infty((0,T);C^\infty(Q))$. Now, by testing the above equation with $\psi = \mathcal{B}h$ where, for $\phi \in C_c^\infty((0,T);C^\infty(Q))$ with zero average, h is the unique solution of the elliptic system

$$-\mathcal{B}^*\mathcal{B}h = \phi, \quad \int_Q h = 0, \tag{4.18}$$

we infer that

$$\int_0^T \int_Q (W - \bar{W})_t \phi - \int_0^T \int_Q \left(g(U) - g(\bar{U})\right) \phi = 0. \tag{4.19}$$

Note that we have moved the time derivative on $(W - \bar{W})$. This is indeed possible since by (4.17) and the fact that g is globally Lipschitz, $U_t \in L^\infty(0,T;H^{-1}(Q))$. In particular, $\mathcal{B}W_t \in L^\infty(0,T;H^{-1}(Q))$ and by ellipticity of $\mathcal{B}$, we infer that $W_t \in L^\infty(0,T;L^2(Q))$. We may now test (4.19) with the function $\phi = W - \bar{W}$, while localising in time, to get that by Young's inequality and the Lipschitz condition on g,

$$\begin{aligned}\int_Q |W - \bar{W}|^2 &\lesssim \int_Q |W^0 - \bar{W}^0|^2 + \int_0^t \int_Q |g(U) - g(\bar{U})|^2 + \int_0^t \int_Q |W - \bar{W}|^2, \ t \in (0,T) \\ &\lesssim \int_Q |W^0 - \bar{W}^0|^2 + \int_0^t \int_Q |U - \bar{U}|^2 + \int_0^t \int_Q |W - \bar{W}|^2, \ t \in (0,T).\end{aligned}$$

Then, the above estimate inserted in (4.15) and Grönwall's inequality allows us to complete the proof. This setting may seem restrictive but it is, up to suitable modifications, the case of elasticity where $(\mathcal{B}, \mathcal{A}) = (\nabla, \mathrm{curl})$ and (4.18) reduces to the Poison equation which provides the appropriate elliptic estimates. We present the above remark as we find its generalisation to the $\mathcal{A}$-free setting interesting.

4.1.2 Quasiconvex adiabatic thermoelasticity

We turn our attention to the system of adiabatic thermoelasticity which, recalling from the introduction, is given by

$$\begin{aligned} \partial_t F_{i\alpha} - \partial_\alpha v_i &= 0 \\ \partial_t v_i - \partial_\alpha \Sigma_{i\alpha} &= 0 \\ \partial_t \left(\frac{1}{2}|v|^2 + e \right) - \partial_\alpha (\Sigma_{i\alpha} v_i) &= r. \end{aligned}$$

We remind that smooth solutions of (1.4) satisfy the entropy production identity

$$\partial_t \eta = \frac{r}{\theta},$$

which is replaced by the inequality $\partial_t \eta \geqslant \frac{r}{\theta}$ for weak solutions. The latter serves as an admissibility condition for our system. We additionally assume that the differential constraint $\mathrm{curl} F = 0$ is an involution i.e.

$$\mathrm{curl} F(0, \cdot) = 0 \Rightarrow \mathrm{curl} F(t, \cdot) = 0 \quad \text{for any } \ t \in (0, T), \tag{4.20}$$

which enforces F to be a deformation gradient as long as the solution exists.

System (1.4) belongs to a general class of symmetrisable hyperbolic systems of conservation laws describing the evolution of a function $U : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^N$ and have the form

$$\partial_t A(U) + \partial_\alpha f_\alpha(U) = 0 \tag{4.21}$$

where A and $f_\alpha : \mathcal{O} \subset \mathbb{R}^N \to \mathbb{R}^N$, $\alpha = 1, \ldots, d$, are smooth functions describing fluxes. Here the matrix $A(U)$ is globally invertible on the domain of definition $\mathcal{O} \ni U$ and $\nabla A(U)$ is nonsingular. System (4.21) is endowed with an entropy - entropy flux pair $\eta, q_\alpha : \mathbb{R}^N \to \mathbb{R}$ if any smooth solution $U(t, x) \in C^1(\mathbb{R}^N)$ of (4.21) satisfies the additional conservation law of entropy

$$\partial_t \eta(U) + \partial_\alpha q_\alpha(U) = 0\,. \tag{4.22}$$

This is equivalent to the existence of a multiplier $G : \mathbb{R}^N \to \mathbb{R}^N$ which is a smooth function of the solution $G = G(U)$ satisfying the relations

$$\begin{aligned} G \cdot \nabla A &= \nabla \eta \\ G \cdot \nabla f_\alpha &= \nabla q_\alpha \end{aligned} \tag{4.23}$$

or equivalently the relations

$$\begin{aligned} \nabla G^T \nabla A &= \nabla A^T \nabla G \\ \nabla G^T \nabla f_\alpha &= \nabla f_\alpha{}^T \nabla G \,. \end{aligned} \tag{4.24}$$

In particular, whenever (4.23) is satisfied by a smooth solution of (4.21), then it also satisfies the entropy identity (4.22).

Suppose now that (4.21) is endowed with a smooth entropy pair $\eta - q_\alpha$, that is for some multiplier $G(U)$ relations (4.24) are satisfied. We can rewrite (4.21) for smooth solutions, in the form of an equivalent system with symmetric coefficients:

$$(\nabla G^T \nabla A)\partial_t U + (\nabla G^T \nabla f_\alpha)\partial_\alpha U = 0. \tag{4.25}$$

The hypothesis

$$\nabla G^T \nabla A > 0$$

guarantees that the system (4.21) is symmetrisable in the sense of Friedrichs and Lax [54], it has real eigenvalues and it is hyperbolic. Moreover, it induces a relative entropy identity and therefore a notion of stability for the system (4.21), see [54, 25]. Using (4.22), it can be equivalently expressed in the form

$$\nabla^2 \eta - \sum_{k=1} G^k \nabla^2 A^k > 0 \,. \tag{4.26}$$

For weak solutions the entropy pair $\eta - q_\alpha$ gives rise to a notion of admissibility. The function $U \in L^1_{loc}(\mathbb{R}^N)$ is an entropy weak solution if it satisfies, in the sense of distributions, (4.21) and the entropy inequality

$$\partial_t \eta(U) + \partial_\alpha q_\alpha(U) \leq 0. \tag{4.27}$$

Adiabatic thermoelasticity (1.4) fits into the general form of system (4.21), by setting

$$U = (F, v, \eta) \quad A(U) = \big(F, v, \tfrac{1}{2}|v|^2 + e(F, \eta)\big).$$

The positivity condition $\theta = \partial e / \partial \eta > 0$ for the temperature guarantees that $A(U)$ is invertible and $\nabla A(U)$ is nonsingular. By construction of the theory, there is a multiplier $G(U)$ that leads to the entropy pair $\check{\eta}(U)$ - $\check{q}_\alpha(U)$ with

$$\check{\eta}(U) := -\eta, \quad \check{q}_\alpha(U) := 0, \quad G(U) = \tfrac{1}{\theta(F,\eta)}\left(\frac{\partial e}{\partial F}(F, \eta), v, -1\right).$$

We may compute that

$$\nabla^2 \check{\eta}(U) - \sum_{k=1} G^k(U) \nabla^2 A^k(U) = \frac{1}{e_\eta} \begin{pmatrix} e_{FF} & 0 & e_{F\eta} \\ 0 & 1 & 0 \\ e_{F\eta} & 0 & e_{\eta\eta} \end{pmatrix}$$

and thus the condition of symmetrisability (4.26) amounts to $e(F, \eta)$ strongly convex and $\theta(F, \eta) = \frac{\partial e(F,\eta)}{\partial \eta} > 0$. Convexity of $e(F, \eta)$ suffices to apply the standard theory of conservation laws to (1.4) and, in that case, the entropy admissibility inequality (4.27) amounts to the growth of the physical entropy.

However, as we already discussed in the introduction, the requirement of convex internal energy is too restrictive due to frame indifference, and so we here assume a weaker assumption which is associated with the symmetrasibility of the system and hence with the positivity of the matrix (1.8) in specific directions. The latter is guaranteed assumping that the free energy e is strongly quasiconvex, see Definition 3.1, and hence, if in addition it satisfies the assumptions (H_1), (H_2) and (H_3), we show that smooth solutions of the system (1.4) are unique within a suitable class of dissipative measure-valued solutions, Theorem 4.3. As we mentioned in the Introduction, for polyconvex energy e, stability of classical solutions in the class of entropy weak and dissipative measure-valued solutions of the system (1.4) has been established in [22, 23, 24].

Dissipative measure-valued solutions:

For $p \geqslant 0$, let $C_p(\mathbb{R}^d)$ denote the space of continuous functions such that

$$C_p(\mathbb{R}^d) := \left\{ g \in C(\mathbb{R}^d) : \lim_{|z| \to \infty} \frac{g(z)}{|z|^p} = 0 \right\}$$

while the space $C_0(\mathbb{R}^d)$ is defined as

$$C_0(\mathbb{R}^d) := \left\{ g \in C(\mathbb{R}^d) : \lim_{|z| \to \infty} g(z) = 0 \right\}.$$

Identifying the space of signed Radon measures $\mathcal{M}(\mathbb{R}^d)$ equipped with the total variation norm as isometrically isomorphic to the dual of $C_0(\mathbb{R}^d)$, a (parametrised) Young measure $\nu = (\nu_x)_{x \in Q}$ is an element of the space $L^\infty_{w*}(Q, \mathcal{M}(\mathbb{R}^d))$ taking values in the space of probability measures. The space $L^\infty_{w*}(Q, \mathcal{M}(\mathbb{R}^d))$ consists of all weak-$*$ measurable, essentially bounded maps $\nu : Q \ni x \mapsto \nu_x \in \mathcal{M}(\mathbb{R}^d)$, i.e. all maps such that

$$x \mapsto \langle \nu, g \rangle := \int g(z) \, d\nu(z)$$

is measurable for all $g \in C_0(\mathbb{R}^d)$ and

$$\sup_{x \in Q} \|\nu_x\|_{\mathcal{M}(\mathbb{R}^d)} < \infty.$$

Since $C_0(\mathbb{R}^d)$ is separable, we have

$$L^\infty_{w*}(Q,\mathcal{M}(\mathbb{R}^d)) = L^1(Q, C_0(\mathbb{R}^d))^*$$

and this defines the weak-$*$ limits of Young measures. The Fundamental Theorem of Young measures in L^p states that given a bounded sequence (U_n) in $L^p(Q)$ $(1 \leqslant p < \infty)$, there exists a subsequence and a parametrized family of Young measures $\nu = (\nu_x)_{x\in Q}$ such that

$$g(U_n) \rightharpoonup \langle \nu_x, g\rangle \quad \text{in } L^1(Q),\ \forall\, g \in C_p(\mathbb{R}^d), \tag{4.28}$$

and we say that the sequence (U_n) generates the Young measure ν. We call ν a $p-$*Young measure* since it is generated by a bounded sequence in L^p. We note that the space of Young measures as defined here coincides with the space $Y(Q;\mathbb{R}^d)$ defined in Section 2.3, see [81] for more details. In fact, as we already mentioned in section 2.3, the sequence $(g(U_n))$ converges as in (4.28) whenever it is equiintegrable and the barycentre $\langle \nu_x, id\rangle$ of the generated Young measure gives the weak limit of the sequence U_n, i.e.

$$U_n \rightharpoonup \langle \nu_x, id\rangle \quad \text{in } L^p(Q).$$

If $U_n = \nabla u_n$ for $u_n \in W^{1,p}(Q)$, then we call ν a *gradient* $p-$*Young measure*. Below, we wish to consider generating sequences (U_n) bounded in the Bochner space $L^\infty(0,T;L^p(Q))$, defined both in time and space, for some $T > 0$. Then, (U_n) is also bounded in $L^p(Q_T)$ and generates a p-Young measure $\nu_{t,x}$, with $(t,x) \in Q_T := (0,T) \times Q$, which satisfies

$$\int_{Q_T} \langle \boldsymbol{\nu}_{t,x}, |\cdot|^p\rangle dx < \infty.$$

However, the above integrability can be improved to obtain L^∞ bounds in the time variable, see [16] for the proof. In particular, it holds that

$$\sup_{0\leqslant t\leqslant T} \int \langle \nu_{t,x}, |\cdot|^p\rangle < \infty.$$

In our context, we naturally consider measure-valued solutions as limits of approximations that satisfy the uniform bound

$$\sup_{0\leqslant t\leqslant T} \int_{\mathbb{T}^d} e(F^\varepsilon, \eta^\varepsilon) + \frac{1}{2}|v^\varepsilon|^2\, dx \leqslant C, \tag{4.29}$$

coming by integrating in Q_T the energy conservation equation $(1.4)_3$, given that the radiative heat supply r is bounded in $L^1(Q_T)$ and that the initial data have bounded energy. Since the energy satisfies the growth-coercivity condition (H_2) i.e.

$$c(|F|^p + |\eta|^q - 1) \leqslant e(F,\eta) \leqslant c(|F|^p + |\eta|^q + 1), \tag{4.30}$$

together with (4.29), implies the following uniform in ε bounds:

$$F^\varepsilon \in L^\infty(0,T;L^p(Q)), \quad v^\varepsilon \in L^\infty(0,T;L^2(Q)), \quad \eta^\varepsilon \in L^\infty(0,T;L^q(Q)). \tag{4.31}$$

Then, the sequence $\{(F^\varepsilon, v^\varepsilon, \eta^\varepsilon)\}$ generates a family of probability measures

$$\nu_{t,x} \in \mathcal{M}_1^+(\mathbb{M}^{3\times 3} \times \mathbb{R}^3 \times \mathbb{R})$$

given by the mapping $(\nu_{t,x}) : Q_T \ni (t,x) \mapsto \nu_{t,x}$. The Young measure $(\nu_{t,x})$ is an element of the space $L^\infty_{w*}(Q_T, \mathcal{M}(\mathbb{R}^{13}))$ representing weak limits of the form

$$\text{wk-}*\text{-}\lim_{\varepsilon\to 0} \psi(F^\varepsilon, v^\varepsilon, \eta^\varepsilon) = \langle \nu_{t,x}, \psi(\lambda_F, \lambda_v, \lambda_\eta)\rangle, \tag{4.32}$$

for all continuous functions $\psi = \psi(\lambda_F, \lambda_v, \lambda_\eta)$ such that

$$\lim_{|\lambda_F|^p + |\lambda_v|^2 + |\lambda_\eta|^q \to \infty} \frac{|\psi(\lambda_F, \lambda_v, \lambda_\eta)|}{|\lambda_F|^p + |\lambda_v|^2 + |\lambda_\eta|^q} = 0, \tag{4.33}$$

where in (4.32) the notation $\langle \nu_{t,x}, \cdot\rangle$ stands for the average

$$\langle \nu_{t,x}, \psi(\lambda_F, \lambda_v, \lambda_\eta)\rangle = \int \psi(\lambda_F, \lambda_v, \lambda_\eta)\, \nu_{t,x}(d\lambda_F, d\lambda_v, d\lambda_\eta)$$

and $\lambda_F \in \mathbb{M}^{3\times 3}$, $\lambda_v \in \mathbb{R}^3$, $\lambda_\eta \in \mathbb{R}$. The marginal of $\nu_{t,x}$ generated by $(F^\varepsilon)_\varepsilon = (\nabla y^\varepsilon)_\varepsilon$ is a gradient $p-$Young measure, while the marginals generated by (v^ε) and (η^ε) are a 2- and a q-Young measure respectively. In particular,

$$\begin{aligned} F^\varepsilon &\overset{*}{\rightharpoonup} \langle \nu_{t,x}, \lambda_F\rangle =: F \quad \text{weak-}* \text{ in } L^\infty(0,T;L^p(Q))\,, \\ v^\varepsilon &\overset{*}{\rightharpoonup} \langle \nu_{t,x}, \lambda_v\rangle =: v \quad \text{weak-}* \text{ in } L^\infty(0,T;L^2(Q))\,, \\ \eta^\varepsilon &\overset{*}{\rightharpoonup} \langle \nu_{t,x}, \lambda_\eta\rangle =: \eta \quad \text{weak-}* \text{ in } L^\infty(0,T;L^q(Q))\,. \end{aligned}$$

We note that the space $C_p(\mathbb{R}^d)$ is separable equipped with the norm $\|g(\cdot)/(1+|\cdot|^p)\|_{L^\infty}$, and so is the space $C_{p,q}(\mathbb{R}^d\times\mathbb{R})$ defined as

$$C_{p,q}(\mathbb{R}^d \times \mathbb{R}) := \left\{ g \in C(\mathbb{R}^d\times\mathbb{R}) : \lim_{|z_1|^p + |z_2|^q \to\infty} \frac{g(z_1,z_2)}{|z_1|^p + |z_2|^q} = 0 \right\},$$

equipped with the norm $\|g(\cdot)/(1+|\cdot|^p+|\cdot|^q)\|_{L^\infty}$. As a result, the internal energy function $e(\lambda_F, \lambda_\eta)$ belongs to the separable space $C_{p,q}(\mathbb{R}^9\times\mathbb{R})$ ($\mathbb{M}^{3\times 3} \simeq \mathbb{R}^9$) under the aformentioned norm.

To take into account the formation of concentration effects, we introduce the concentration measure γ, that depends on the total energy. This is a well-defined nonnegative Radon measure for a subsequence of

$$\frac{1}{2}|v^\varepsilon|^2 + e(F^\varepsilon, \eta^\varepsilon).$$

Since we know that the functions $(F^{\varepsilon}, v^{\varepsilon}, \eta^{\varepsilon})$ are all bounded in some L^p space -because of (4.31)- we may apply the Theorem 2.7, in order to pass to the limit. Indeed, letting Ω be an open subset of $\mathbb{R}^d$, the theorem asserts that given a sequence of functions (u_n), $u_n : \Omega \to \mathbb{R}^m$, bounded in $L^p(\Omega)$, $(p \geqslant 1)$ there exists a subsequence (which we will not relabel), a parametrized family of probability measures $\nu \in L^{\infty}_{w*}(\Omega; (\mathbb{R}^m))$, a nonnegative measure $\lambda \in \mathcal{M}^+(\Omega)$ and a parametrized probability measure on a sphere $\nu^{\infty} \in L^{\infty}_{w*}((\Omega, \mu); \mathcal{M}^+_1(S^{m-1}))$ such that

$$\psi(x, u_n)\, dx \overset{*}{\rightharpoonup} \int_{\mathbb{R}^m} \psi(x, z)\, d\nu dx + \int_{S^{m-1}} \psi^{\infty}(x, z)\, d\nu^{\infty}(z) d\lambda(x), \tag{4.34}$$

for all ψ continuous with well-defined recession function

$$\psi^{\infty}(x, z) := \lim_{\substack{s\to\infty \\ z'\to z}} \frac{\psi(x, sz')}{s^p}.$$

The sequences $(F^{\varepsilon}, v^{\varepsilon}, \eta^{\varepsilon})$ are bounded in different spaces and have different growth, and as a result, we need to apply a refinement of the aforementioned theorem as, for instance, in [64]: consider a sequence of maps $u_n = (u_n^1, u_n^2)$ where (u_n^1) is bounded in some $L^p(\Omega; \mathbb{R}^b)$ and (u_n^2) is bounded in $L^q(\Omega; \mathbb{R}^l)$ and define the non-homogeneous unit sphere

$$S^{b+l-1}_{pq} := \left\{(\beta_1, \beta_2) \in \mathbb{R}^{b+l} : |\beta_1|^p + |\beta_2|^q = 1\right\},$$

for exponents $p, q > 1$. Then one can pass to the limit as in (4.34) where

$$\psi^{\infty}(x, z) := \lim_{\substack{x'\to x \\ s\to\infty \\ (\beta_1', \beta_2')\to(\beta_1,\beta_2)}} \frac{\psi(x', s^q\beta_1', s^p\beta_2')}{s^{pq}} = \lim_{\substack{x'\to x \\ \tau\to\infty \\ (\beta_1', \beta_2')\to(\beta_1,\beta_2)}} \frac{\psi(x', \tau^{\frac{1}{p}}\beta_1', \tau^{\frac{1}{q}}\beta_2')}{\tau}.$$

We define the generalized sphere

$$S^{12} = \left\{(F, v, \eta) \in \mathbb{R}^{13} : |F|^p + |v|^2 + |\eta|^q = 1\right\}.$$

The form of the recession function for the energy follows from [2, Thm 2.5] and reads

$$\left(\frac{1}{2}|v|^2 + e(F, \eta)\right)^{\infty} = \lim_{\tau\to\infty} \left(\frac{1}{2}|v|^2 + \frac{e(\tau^{\frac{1}{p}}F, \tau^{\frac{1}{q}}\eta)}{\tau}\right),$$

and we require it to be continuous on S^{12}. Then, along a subsequence in ε,

$$\frac{1}{2}|v^{\varepsilon}|^2 + e(F^{\varepsilon}, \eta^{\varepsilon}) \overset{*}{\rightharpoonup} \left\langle \nu_{t,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta)\right\rangle dx + \left\langle \nu^{\infty}, \left(\frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta)\right)^{\infty}\right\rangle \lambda$$

weak-$*$ in the sense of measures, where $\nu \in \mathcal{M}^+_1(Q_T; \mathbb{R}^{13})$, $\nu^{\infty} \in \mathcal{M}^+_1((Q_T, \lambda); S^{12})$ and $\lambda \in \mathcal{M}^+(Q_T)$. Then (4.30) implies

$$\left(\frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta)\right)^{\infty} > 0$$

so that the concentration measure $\gamma \in \mathcal{M}^+(Q_T)$ is nonnegative, i.e.

$$\gamma := \left\langle \nu^\infty, \left(\frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta)\right)^\infty \right\rangle \lambda \geq 0 \,. \tag{4.35}$$

The following definition of a measure-valued solution for system (1.4) thus arises:

Definition 4.1. A dissipative measure-valued solution to adiabatic thermoelasticity (1.4), (1.6) consists of a thermomechanical process $(y(t,x), \eta(t,x)) : [0,T] \times Q \to \mathbb{R}^3 \times \mathbb{R}$ for any $T > 0$,

$$y \in W^{1,\infty}(0,T; L^2(Q)) \cap L^\infty(0,T; W^{1,p}(Q))\,, \quad \eta \in L^\infty(0,T; L^q(Q))\,, \tag{4.36}$$

a parametrized family of probability measures $\nu = (\nu_{t,x})_{(t,x)\in Q_T}$ and a nonnegative Radon measure $\gamma \in \mathcal{M}^+(Q_T)$. The measure ν is generated by a sequence $(v^\varepsilon, \nabla y^\varepsilon, \eta^\varepsilon)$ such that

$$\begin{aligned}
&(y^\varepsilon) \text{ is bounded in } L^\infty(0,T; W^{1,p}(Q)) \\
&(\partial_t \nabla y^\varepsilon) \text{ is bounded in } L^\infty(0,T; H^{-1}(Q)) \\
&(\eta^\varepsilon) \text{ is bounded in } L^\infty(0,T; L^q(Q)).
\end{aligned} \tag{4.37}$$

If (v, F, η) denote the averages

$$F = \langle \nu_{t,x}, \lambda_F \rangle\,, \quad v = \langle \nu_{t,x}, \lambda_v \rangle\,, \quad \eta = \langle \nu_{t,x}, \lambda_\eta \rangle\,,$$

then $\nu_{t,x}$ and γ satisfy

$$F = \nabla y \in L^\infty(L^p), \quad v = \partial_t y \in L^\infty(L^2)\,, \tag{4.38}$$

and the relations

$$\begin{aligned}
\partial_t F &= \partial_\alpha v_i \\
\partial_t \langle \nu_{t,x}, \lambda_{v_i} \rangle &= \partial_\alpha \left\langle \nu_{t,x}, \frac{\partial e}{\partial F_{i\alpha}}(\lambda_F, \lambda_\eta) \right\rangle \\
\partial_t \langle \nu_{t,x}, \lambda_\eta \rangle &\geqslant \left\langle \nu_{t,x}, \frac{r}{\theta(\lambda_F, \lambda_\eta)} \right\rangle
\end{aligned} \tag{4.39}$$

in the sense of distributions. Moreover, they satisfy the integrated form of the averaged energy identity,

$$\begin{aligned}
&\int \varphi(0) \left\langle \nu_{0,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta) \right\rangle dx \\
&\int_0^T \int \varphi'(t) \left(\left\langle \nu_{t,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta) \right\rangle (t,x)\, dx\, dt + \gamma(dx\, dt) \right) \\
&= -\int_0^T \int \langle \nu_{t,x}, r \rangle \varphi(t)\, dx\, dt,
\end{aligned} \tag{4.40}$$

holding for all $\varphi \in C_c^\infty([0,T))$, $\varphi \geq 0$.

Remark 4.4. On the definition of the dissipative measure-valued solution:

(1) We remark that in addition to the uniform estimate (4.29), natural approximations of (1.4), (1.6) produce a uniform bound on the time derivatives of (F^ε) and (v^ε) in a negative Sobolev space. We take all this into account by assuming (4.37).

(2) The first equation holds in a classical weak form, due to its linearity.

(3) Henceforth, we assume the measure $\gamma_0 = 0$, meaning that we consider initial data with no concentrations at time $t = 0$.

(4) We choose to work with dissipative measure-valued solutions, namely solutions that satisfy the integrated form of the averaged energy equation (4.40). This approach has the technical advantage that one does not need to place any integrability condition on the right hand-side of the energy equation $(1.4)_3$, namely on the term $\Sigma_{i\alpha} v_i$, since it appears as a divergence and its contribution integrates to zero.

The averaged relative entropy inequality

Consider a strong solution $(\bar{F}, \bar{v}, \bar{\eta})^T \in W^{1,\infty}(Q_T)$ to (1.4) that satisfies the entropy identity (1.6) and a dissipative measure-valued solution to (1.4), (1.6) according to Definition 4.1. Similarly with the first part of this section, subsection 4.1.1, we use the relative entropy method to estimate the distance between the above solutions. To this end, we first write the difference of the weak form of equations (1.4) and (4.39), to obtain the following three integral identities

$$\begin{aligned}\int (F_{i\alpha} - \bar{F}_{i\alpha})(0,x)\phi_1(0,x)\, dx + \int_0^T \int (F_{i\alpha} - \bar{F}_{i\alpha})\partial_t \phi_1(t,x)\, dx\, dt& \\ = \int_0^T \int (v_i - \bar{v}_i)\partial_\alpha \phi_1(t,x)\, dx\, dt,&\end{aligned} \tag{4.41}$$

$$\begin{aligned}\int (\langle \nu_{0,x}, \lambda_{v_i}\rangle - \bar{v}_i(0,x))\phi_2(0,x)\, dx + \int_0^T \int (\langle \nu_{t,x}, \lambda_{v_i}\rangle - \bar{v}_i)\partial_t \phi_2(t,x)\, dx\, dt& \\ = \int_0^T \int \big(\langle \nu_{t,x}, \Sigma_{i\alpha}(\lambda_F, \lambda_\eta)\rangle - \Sigma_{i\alpha}(\bar{F}, \bar{\eta})\big)\, \partial_\alpha \phi_2(t,x)\, dx\, dt\ ,&\end{aligned} \tag{4.42}$$

and

$$\begin{aligned}&\int \left(\left\langle \nu_{0,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta) \right\rangle - \left(\frac{1}{2}|\bar{v}|^2 + e(\bar{F}, \bar{\eta}) \right)(0,x) \right) \phi_3(0)\, dx \\ &+ \int_0^T \int \left\{ \left(\left\langle \nu_{t,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F, \lambda_\eta) \right\rangle - \frac{1}{2}|\bar{v}|^2 - e(\bar{F}, \bar{\eta}) \right) + \gamma \right\} \partial_t \phi_3(t)\, dx\, dt \\ &= - \int_0^T \int (\langle \nu_{t,x}, r\rangle - \bar{r})\phi_3(t)\, dx\, dt,\end{aligned} \tag{4.43}$$

for any $\phi_i \in C_c^1([0,T)\times Q)$, $i=1,2$ and $\phi_3 \in C_c^1([0,T))$. Similarly, testing the difference of (1.6) and (4.39)$_3$ against $\phi_4 \in C_c^1([0,T)\times Q)$, with $\phi_4 \geq 0$, we have

$$
\begin{aligned}
-\int (\langle \nu_{0,x}, \lambda_\eta\rangle - \bar\eta(0,x))\phi_4(0,x)\,dx - \int_0^T \int (\langle \nu_{t,x}, \lambda_\eta\rangle - \bar\eta)\partial_t\phi_4(t,x)\,dx\,dt \\
\geqslant \int_0^T \int \left(\left\langle \nu_{t,x}, \frac{r}{\theta(\lambda_F,\lambda_\eta)}\right\rangle - \frac{\bar r}{\theta(\bar F,\bar\eta)}\right)\phi_4(t,x)\,dx\,dt.
\end{aligned}
\tag{4.44}
$$

We then choose

$$
(\phi_1,\phi_2,\phi_3) = -\theta(\bar F,\bar\eta)\,G(\bar U)\varphi(t) = (-\Sigma(\bar F,\bar\eta), -\bar v, 1)^T\varphi(t),
$$

for some $\varphi \in C_c^1([0,T])$. Thus, by virtue of (1.7), equations (4.41), (4.42) and (4.43) become

$$
\begin{aligned}
&\int \left(-\frac{\partial e}{\partial F_{i\alpha}}(\bar F,\bar\eta)(F_{i\alpha}-\bar F_{i\alpha})\right)(0,x)\varphi(0)\,dx \\
&\quad + \int_0^T \int \left(-\frac{\partial e}{\partial F_{i\alpha}}(\bar F,\bar\theta)(F_{i\alpha}-\bar F_{i\alpha})\right)\varphi'(t)\,dx\,dt \\
&\quad = \int_0^T \int \left[\partial_t\Big(\frac{\partial e}{\partial F_{i\alpha}}(\bar F,\bar\eta)\Big)(F_{i\alpha}-\bar F_{i\alpha}) - \partial_\alpha\Big(\frac{\partial e}{\partial F_{i\alpha}}(\bar F,\bar\eta)\Big)(v_i-\bar v_i)\right]\varphi(t)\,dx\,dt\,,
\end{aligned}
\tag{4.45}
$$

$$
\begin{aligned}
&\int -\bar v_i(\langle \nu_{0,x},\lambda_{v_i}\rangle - \bar v_i(0,x))\varphi(0)\,dx + \int_0^T \int -\bar v_i(\langle \nu_{t,x},\lambda_{v_i}\rangle - \bar v_i)\varphi'(t)\,dx\,dt \\
&\quad = -\int_0^T \int \Bigg[-\partial_\alpha\left(\frac{\partial e}{\partial F_{i\alpha}}(\bar F,\bar\eta)\right)(\langle \nu_{t,x},\lambda_{v_i}\rangle - \bar v_i) \\
&\qquad\qquad + \partial_\alpha\bar v_i\left(\left\langle \nu_{t,x}, \frac{\partial e}{\partial F_{i\alpha}}(\lambda_F,\lambda_\eta)\right\rangle - \frac{\partial e}{\partial F_{i\alpha}}(\bar F,\bar\eta)\right)\Bigg]\varphi(t)\,dx\,dt\ ,
\end{aligned}
\tag{4.46}
$$

and

$$
\begin{aligned}
&\int \left(\left\langle \nu_{0,x}, \frac12|\lambda_v|^2 + e(\lambda_F,\lambda_\eta)\right\rangle - \left(\frac12|\bar v|^2 - e(\bar F,\bar\eta)\right)(0,x)\right)\ \varphi(0)\,dx \\
&\quad + \int_0^T \int \left\{\left(\left\langle \nu_{t,x}, \frac12|\lambda_v|^2 + e(\lambda_F,\lambda_\eta)\right\rangle - \frac12|\bar v|^2 - e(\bar F,\bar\eta)\right) + \gamma\right\}\varphi'(t)\,dx\,dt \\
&\quad = -\int_0^T \int (\langle \nu_{t,x}, r\rangle - \bar r)\varphi(t)\,dx\,dt.
\end{aligned}
\tag{4.47}
$$

For inequality (4.44), we choose accordingly $\phi_4 := \theta(\bar F,\bar\eta)\varphi(t) \geqslant 0$, $\varphi \geqslant 0$ so that

$$
\begin{aligned}
&-\int \theta(\bar F,\bar\eta)(\langle \nu_{0,x},\lambda_\eta\rangle - \bar\eta(0,x))\varphi(0)\,dx - \int_0^T \int \theta(\bar F,\bar\eta)(\langle \nu_{t,x},\lambda_\eta\rangle - \bar\eta)\varphi'(t)\,dx\,dt \\
&\quad \geqslant \int_0^T \int \Bigg[\partial_t\theta(\bar F,\bar\eta)(\langle \nu_{t,x},\lambda_\eta\rangle - \bar\eta) \\
&\qquad\qquad + \theta(\bar F,\bar\eta)\left(\left\langle \nu_{t,x}, \frac{r}{\theta(\lambda_F,\lambda_\eta)}\right\rangle - \frac{\bar r}{\theta(\bar F,\bar\eta)}\right)\Bigg]\varphi(t)\,dx\,dt.
\end{aligned}
\tag{4.48}
$$

Adding together (4.45), (4.46), (4.47) and (4.48), we obtain the integral inequality

$$
\int \varphi(0)\Big[-\frac{\partial e}{\partial F_{i\alpha}}(\bar F,\bar\eta)(F_{i\alpha}-\bar F_{i\alpha})(0,x) - \langle \nu_{0,x}, \bar v_i(\lambda_{v_i}-\bar v_i)\rangle(0,x)
$$

$$
\begin{aligned}
&+ \left\langle \nu_{0,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F,\lambda_\eta) - \frac{1}{2}|\bar{v}|^2 - e(\bar{F},\bar{\eta}) \right\rangle (0,x) \\
&\qquad\qquad\qquad - \theta(\bar{F},\bar{\eta})\langle \nu_{0,x}, \lambda_\eta - \bar{\eta}\rangle(0,x) \Big] \, dx \\
&+ \int_0^T\!\!\int \varphi'(t) \Big[- \frac{\partial e}{\partial F_{i\alpha}}(\bar{F},\bar{\eta})(F_{i\alpha} - \bar{F}_{i\alpha}) - \langle \nu_{t,x}, \bar{v}_i(\lambda_{v_i} - \bar{v}_i)\rangle \\
&\qquad + \left\langle \nu_{t,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F,\lambda_\eta) - \frac{1}{2}|\bar{v}|^2 - e(\bar{F},\bar{\eta}) \right\rangle \\
&\qquad\qquad\qquad - \theta(\bar{F},\bar{\eta})\langle \nu_{t,x}, \lambda_\eta - \bar{\eta}\rangle + \gamma \Big] \, dx\, dt \\
&\geqslant - \int_0^T\!\!\int \varphi(t) \Big[- \partial_t \Big(\frac{\partial e}{\partial F_{i\alpha}}(\bar{F},\bar{\eta}) \Big)(F_{i\alpha} - \bar{F}_{i\alpha}) \\
&\qquad + \partial_\alpha \bar{v}_i \left(\left\langle \nu_{t,x}, \frac{\partial e}{\partial F_{i\alpha}}(\lambda_F,\lambda_\eta) \right\rangle - \frac{\partial e}{\partial F_{i\alpha}}(\bar{F},\bar{\eta}) \right) \\
&\qquad - \partial_t \theta(\bar{F},\bar{\eta})(\langle \nu_{t,x}, \lambda_\eta\rangle - \bar{\eta}) \\
&\qquad - \theta(\bar{F},\bar{\eta}) \left(\left\langle \nu_{t,x}, \frac{r}{\theta(\lambda_F,\lambda_\eta)} \right\rangle - \frac{\bar{r}}{\theta(\bar{F},\bar{\eta})} \right) + \langle \nu_{t,x}, r - \bar{r}\rangle \Big] \, dx\, dt \\
&=: - \int_0^T \int \varphi(t) \mathcal{R}(t,x) \, dx\, dt \, .
\end{aligned} \tag{4.49}
$$

Using relations (1.7), the entropy production identity (1.6) that holds for strong solutions and equation $(1.4)_1$, since

$$
r - \bar{r} = \frac{r}{\theta}\,\theta - \frac{\bar{r}}{\bar{\theta}}\,\bar{\theta},
$$

the quantity $\mathcal{R}(t,x)$ in the integrand on the right hand-side of (4.49) becomes

$$
\begin{aligned}
\mathcal{R} &= -\partial_t \bar{F}_{j\beta} \frac{\partial^2 e}{\partial F_{i\alpha}\partial F_{j\beta}}(\bar{F},\bar{\eta})(F_{i\alpha} - \bar{F}_{i\alpha}) - \partial_t \bar{\eta} \frac{\partial^2 e}{\partial F_{i\alpha}\partial \eta}(\bar{F},\bar{\eta})(F_{i\alpha} - \bar{F}_{i\alpha}) \\
&\quad + \partial_t \bar{F}_{i\alpha} \left\langle \nu_{t,x}, \frac{\partial e}{\partial F_{i\alpha}}(\lambda_F,\lambda_\eta) - \frac{\partial e}{\partial F_{i\alpha}}(\bar{F},\bar{\eta}) \right\rangle - \partial_t \bar{F}_{i\alpha} \frac{\partial^2 e}{\partial F_{i\alpha}\partial \eta}(\bar{F},\bar{\eta})\langle \nu_{t,x}, \lambda_\eta - \bar{\eta}\rangle \\
&\quad - \partial_t \bar{\eta} \frac{\partial^2 e}{\partial \eta^2}(\bar{F},\bar{\eta})\langle \nu_{t,x}, \lambda_\eta - \bar{\eta}\rangle \\
&\quad + \partial_t \bar{\eta} \langle \nu_{t,x}, \theta(\lambda_F,\lambda_\eta) - \theta(\bar{F},\bar{\eta})\rangle - \partial_t \bar{\eta} \langle \nu_{t,x}, \theta(\lambda_F,\lambda_\eta) - \theta(\bar{F},\bar{\eta})\rangle \\
&\quad - \theta(\bar{F},\bar{\eta}) \left(\left\langle \nu_{t,x}, \frac{r}{\theta(\lambda_F,\lambda_\eta)} \right\rangle - \frac{\bar{r}}{\theta(\bar{F},\bar{\eta})} \right) + \langle \nu_{t,x}, r - \bar{r}\rangle \\
&= \partial_t \bar{\eta} \Big[\big\langle \nu_{t,x}, \theta(\lambda_F,\lambda_\eta) - \theta(\bar{F},\bar{\eta}) \big\rangle \\
&\qquad\qquad - \frac{\partial \theta}{\partial F_{i\alpha}}(\bar{F},\bar{\eta})(F_{i\alpha} - \bar{F}_{i\alpha}) - \frac{\partial \theta}{\partial \eta}(\bar{F},\bar{\eta})\langle \nu_{t,x}, \lambda_\eta - \bar{\eta}\rangle \Big] \\
&\quad + \partial_t \bar{F}_{j\beta} \Big[\big\langle \nu_{t,x}, \Sigma_{j\beta}(\lambda_F,\lambda_\eta) - \Sigma_{j\beta}(\bar{F},\bar{\eta}) \big\rangle \\
&\qquad\qquad - \frac{\partial^2 e}{\partial F_{i\alpha}\partial F_{j\beta}}(\bar{F},\bar{\eta})(F_{i\alpha} - \bar{F}_{i\alpha}) - \frac{\partial^2 e}{\partial \eta\, \partial F_{j\beta}}(\bar{F},\bar{\eta})\langle \nu_{t,x}, \lambda_\eta - \bar{\eta}\rangle \Big] \\
&\quad - \frac{\bar{r}}{\theta(\bar{F},\bar{\eta})} \langle \nu_{t,x}, \theta(\lambda_F,\lambda_\eta) - \theta(\bar{F},\bar{\eta})\rangle
\end{aligned}
$$

$$
\begin{aligned}
&- \theta(\bar{F},\bar{\eta}) \left(\left\langle \nu_{t,x}, \frac{r}{\theta(\lambda_F,\lambda_\eta)} \right\rangle - \frac{\bar{r}}{\theta(\bar{F},\bar{\eta})} \right) + \langle \nu_{t,x}, r - \bar{r} \rangle \\
&= \partial_t \bar{\eta} \left\langle \nu_{t,x}, \theta(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) \right\rangle + \partial_t \bar{F}_{j\beta} \left\langle \nu_{t,x}, \Sigma_{j\beta}(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) \right\rangle \\
&\quad + \left\langle \nu_{t,x}, (\theta(\lambda_F,\lambda_\eta) - \theta(\bar{F},\bar{\eta})) \left(\frac{r}{\theta(\lambda_F,\lambda_\eta)} - \frac{\bar{r}}{\theta(\bar{F},\bar{\eta})} \right) \right\rangle .
\end{aligned}
\tag{4.50}
$$

Above, we have used the following notation:

$$
\begin{aligned}
\left\langle \nu_{t,x}, \theta(\lambda_F, \lambda_\eta | \bar{F}, \bar{\theta}) \right\rangle := \Big\langle & \nu_{t,x}, \theta(\lambda_F,\lambda_\eta) - \theta(\bar{F},\bar{\theta}) \\
& - \frac{\partial \theta}{\partial F_{i\alpha}}(\bar{F},\bar{\eta})(F_{i\alpha} - \bar{F}_{i\alpha}) - \frac{\partial \theta}{\partial \eta}(\bar{F},\bar{\eta})(\lambda_\eta - \bar{\eta}) \Big\rangle ,
\end{aligned}
\tag{4.51}
$$

and

$$
\begin{aligned}
\left\langle \nu_{t,x}, \Sigma_{i\alpha}(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) \right\rangle := \Big\langle & \nu_{t,x}, \Sigma_{i\alpha}(\lambda_F,\lambda_\eta) - \Sigma_{i\alpha}(\bar{F},\bar{\eta}) \\
& - \frac{\partial^2 e}{\partial F_{i\alpha}\partial F_{j\beta}}(\bar{F},\bar{\eta})(\lambda_F - \bar{F}) - \frac{\partial^2 e}{\partial F_{i\alpha}\partial \eta}(\bar{F},\bar{\eta})(\lambda_\eta - \bar{\eta}) \Big\rangle .
\end{aligned}
\tag{4.52}
$$

If we define the averaged quantity

$$
I(\lambda_U | \bar{U}) = I(\lambda_F, \lambda_v, \lambda_\eta | \bar{F}, \bar{v}, \bar{\eta}) := \frac{1}{2} |\lambda_v - \bar{v}|^2 + e(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) \tag{4.53}
$$

for

$$
e(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) := e(\lambda_F,\lambda_\eta) - e(\bar{F},\bar{\eta}) - \frac{\partial e}{\partial F_{i\alpha}}(\bar{F},\bar{\eta})(\lambda_F - \bar{F})_{i\alpha} - \frac{\partial e}{\partial \eta}(\bar{F},\bar{\eta})(\lambda_\eta - \bar{\eta}),
$$

we observe that the term on the left hand-side of (4.49) becomes

$$
\begin{aligned}
&- \frac{\partial e}{\partial F_{i\alpha}}(\bar{F},\bar{\eta})(F_{i\alpha} - \bar{F}_{i\alpha}) - \langle \nu_{t,x}, \bar{v}_i(\lambda_{v_i} - \bar{v}_i) \rangle \\
&\quad + \left\langle \nu_{t,x}, \frac{1}{2}|\lambda_v|^2 + e(\lambda_F,\lambda_\eta) - \frac{1}{2}|\bar{v}|^2 - e(\bar{F},\bar{\eta}) \right\rangle - \theta(\bar{F},\bar{\eta}) \langle \nu_{t,x}, \lambda_\eta - \bar{\eta} \rangle \\
&\quad = \left\langle \nu_{t,x}, \frac{1}{2}|\lambda_v - \bar{v}|^2 \right\rangle + \left\langle \nu_{t,x}, e(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) \right\rangle \\
&\quad = \left\langle \nu_{t,x}, I(\lambda_F, \lambda_v, \lambda_\eta | \bar{F}, \bar{v}, \bar{\eta}) \right\rangle .
\end{aligned}
\tag{4.54}
$$

We then combine (4.49),(4.50), and (4.54) to arrive at the relative entropy inequality

$$
\begin{aligned}
&\int \varphi(0) [\left\langle \nu_{0,x}, I(\lambda_{U_0} | \bar{U}_0) \right\rangle \, dx] \\
&\quad + \int_0^T \int \varphi'(t) \left[\left\langle \nu_{t,x}, I(\lambda_U | \bar{U}) \right\rangle \, dx \, dt + \gamma(dx\, dt) \right] \\
&\quad \geqslant - \int_0^T \int \varphi(t) \Big[\partial_t \bar{\eta} \left\langle \nu_{t,x}, \theta(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) \right\rangle + \partial_t \bar{F}_{j\beta} \left\langle \nu_{t,x}, \Sigma_{j\beta}(\lambda_F, \lambda_\eta | \bar{F}, \bar{\eta}) \right\rangle \\
&\qquad\qquad + \left\langle \nu_{t,x}, (\theta(\lambda_F,\lambda_\eta) - \theta(\bar{F},\bar{\eta})) \left(\frac{r}{\theta(\lambda_F,\lambda_\eta)} - \frac{\bar{r}}{\theta(\bar{F},\bar{\eta})} \right) \right\rangle \Big] \, dx \, dt.
\end{aligned}
\tag{4.55}
$$

Measure-valued versus strong uniqueness

Due to the relative entropy inequality (4.55), we may now show that classical solutions are unique in the class of dissipative measure-valued solutions. Before we proceed, we summarise the assumptions on the internal energy:

(H_1) $e \in C^3(\mathbb{R}^{d\times d} \times \mathbb{R})$

(H_2) $c_2(|F|^p + |\eta|^q - 1) \leqslant e(F,\eta) \leqslant c_1(|F|^p + |\eta|^q + 1)$

(H_3) $|e_F(F,\eta)| \lesssim 1 + |F|^{p-1} + |\eta|^{q\frac{p-1}{p}}$, and $|e_\eta(F,\eta)| \lesssim 1 + |F|^{p\frac{q-1}{q}} + |\eta|^{q-1}$.

To establish the measure-valued vs strong uniqueness result, we first assert that the following bounds on the relative entropy and the terms on the right hand side of (4.55) can be obtained given the above hypotheses and the quasiconvexity assumption, Definition 3.3.

Lemma 4.1. *Given hypotheses $(H_1)-(H_3)$, for p, $q \geqslant 2$, assume that the smooth solution $(\bar{F},\bar{v},\bar{\eta})$ lies in the compact set*

$$\Gamma_K := \left\{(\bar{F},\bar{v},\bar{\eta}) : |\bar{F}(t,\cdot)| \leqslant K, |\bar{v}(t,\cdot)| \leqslant K, |\bar{\eta}(t,\cdot)| \leqslant K\right\}$$

for a positive constant K. Then there exist constants $C_1, C_2, C_3 > 0$ such that

$$|I(F,v,\eta|\bar{F},\bar{v},\bar{\eta})| \leqslant C_1\left(|v-\bar{v}|^2 + |V_p(F-\bar{F})|^2 + |V_q(\eta-\bar{\eta})|^2\right) \tag{4.56}$$

$$|\theta(F,\eta|\bar{F},\bar{\eta})| \leqslant C_2\left(|V_p(F-\bar{F})|^2 + |V_q(\eta-\bar{\eta})|^2\right) \tag{4.57}$$

and

$$|\Sigma(F,\eta|\bar{F},\bar{\eta})| \leqslant C_3\left(|V_p(F-\bar{F})|^2 + |V_q(\eta-\bar{\eta})|^2\right). \tag{4.58}$$

Under the additional hypothesis:

$$\theta(F,\eta) = \frac{\partial e}{\partial \eta}(F,\eta) \geq \delta > 0\,, \tag{4.59}$$

and given that $r(t,x) = \bar{r}(t,x) \in L^\infty(Q_T)$, there exist a constant $C_4 > 0$ such that

$$\left|(\theta(F,\eta) - \theta(\bar{F},\bar{\eta}))\left(\frac{r}{\theta(F,\eta)} - \frac{\bar{r}}{\theta(\bar{F},\bar{\eta})}\right)\right| \leqslant C_4\left(|V_p(F-\bar{F})|^2 + |V_q(\eta-\bar{\eta})|^2\right) \tag{4.60}$$

for all $(\bar{F},\bar{v},\bar{\eta}) \in \Gamma_K$.

Proof. For the proof of (4.56), observe that

$$e(F,\eta|\bar{F},\bar{\eta}) = \int_0^1 (1-s)D^2 e\left(\bar{F} + s(F-\bar{F}), \bar{\eta} + s(\eta-\bar{\eta})\right)(F-\bar{F},\eta-\bar{\eta}) : (F-\bar{F},\eta-\bar{\eta}).$$

Given the set of hypotheses $(H_1)-(H_3)$, (4.56) follows from Lemma 3.3 (a) by setting $\xi_1 = F-\bar{F}$, $\xi_2=\eta-\bar{\eta}$, $\lambda_1=\bar{F}$, $\lambda_2=\bar{\eta}$ and $z_1=z_2=0$.

Moving to bound (4.57), we cannot use the proof in the Appendix directly, as θ does not satisfy the same growth conditions as e. We start by expressing $\theta(F,\eta|\bar{F},\bar{\eta})$ as follows:

$$\theta(F,\eta|\bar{F},\bar{\eta}) = \int_0^1 (1-s)D^2\theta(\bar{F}+s(F-\bar{F}),\bar{\eta}+s(\eta-\bar{\eta}))(F-\bar{F},\eta-\bar{\eta}) : (F-\bar{F},\eta-\bar{\eta})$$

so that

$$|\theta(F,\eta|\bar{F},\bar{\eta})| \leqslant C\left(|F-\bar{F}|^2+|\eta-\bar{\eta}|^2\right),$$

where $C=C(d,\max D^2\theta)$ in the region $|F-\bar{F}|+|\eta-\bar{\eta}|\leqslant 1$ and $(\bar{F},\bar{\eta})\in\Gamma_K$. If $|F-\bar{F}|+|\eta-\bar{\eta}|>1$ and $(\bar{F},\bar{\eta})\in\Gamma_K$ we have

$$\begin{aligned}
|\theta(F,\eta|\bar{F},\bar{\eta})| &\leqslant |\theta(F,\eta)-\theta(\bar{F},\bar{\eta})| + \left|\frac{\partial\theta}{\partial F_{i\alpha}}(\bar{F},\bar{\eta})\right| |F_{i\alpha}-\bar{F}_{i\alpha}| + \left|\frac{\partial\theta}{\partial\eta}(\bar{F},\bar{\eta})\right| |\eta-\bar{\eta}| \\
&\lesssim |\theta(F,\eta)| + |F-\bar{F}| + |\eta-\bar{\eta}| + 1 \\
&\lesssim |F|^{p\frac{q-1}{q}} + |\eta|^{q-1} + |F-\bar{F}| + |\eta-\bar{\eta}| + 1 \\
&\lesssim 2^{p\frac{q-1}{q}-1}(|F-\bar{F}|^{p\frac{q-1}{q}} + |\bar{F}|^{p\frac{q-1}{q}}) + 2^{q-2}(|\eta-\bar{\eta}|^{q-1}+|\bar{\eta}|^{q-1}) \\
&\qquad\qquad\qquad\qquad + |F-\bar{F}| + |\eta-\bar{\eta}| + 1 \\
&\lesssim |F-\bar{F}|^p + |\eta-\bar{\eta}|^q + |F-\bar{F}| + |\eta-\bar{\eta}| + 1 \\
&\lesssim |F-\bar{F}|^p + |\eta-\bar{\eta}|^q + |F-\bar{F}| + |\eta-\bar{\eta}| \\
&\lesssim |F-\bar{F}|^p + |\eta-\bar{\eta}|^q + (|F-\bar{F}| + |\eta-\bar{\eta}|)^2 \\
&\lesssim |F-\bar{F}|^p + |\eta-\bar{\eta}|^q + |F-\bar{F}|^2 + |\eta-\bar{\eta}|^2 \\
&\leqslant C\left(|V_p(F-\bar{F})|^2 + |V_q(\eta-\bar{\eta})|^2\right),
\end{aligned}$$

because of $(H_3)_2$, Minkowski's inequality and Young's inequality.

Bound (4.58) can be obtained similarly by employing $(H_3)_1$, as now the function $\Sigma(F,\eta)$ is given by $(1.7)_2$ as a partial derivative of $e(F,\eta)$.

Finally for (4.60) in the region where $|F-\bar{F}|+|\eta-\bar{\eta}|\leqslant 1$ and $(\bar{F},\bar{\eta})\in\Gamma_K$ we have

$$\begin{aligned}
\left|(\theta(F,\eta)-\theta(\bar{F},\bar{\eta}))\left(\frac{r}{\theta(F,\eta)}-\frac{\bar{r}}{\theta(\bar{F},\bar{\eta})}\right)\right| &\lesssim \frac{|\theta(F,\eta)-\theta(\bar{F},\bar{\eta})|^2}{\theta(F,\eta)\theta(\bar{F},\bar{\eta})} \\
&\lesssim |\theta(F,\eta)-\theta(\bar{F},\bar{\eta})|^2 \\
&\lesssim |F-\bar{F}|^2 + |\eta-\bar{\eta}|^2,
\end{aligned}$$

where the constants involved depend on $(||r||_{L^\infty},K)$. If $|F-\bar{F}|+|\eta-\bar{\eta}|>1$, given (4.59) it holds that

$$\left|(\theta(F,\eta)-\theta(\bar{F},\bar{\eta}))\left(\frac{r}{\theta(F,\eta)}-\frac{\bar{r}}{\theta(\bar{F},\bar{\eta})}\right)\right| \lesssim |\theta(F,\eta)-\theta(\bar{F},\bar{\eta})| \lesssim |\theta(F,\eta)|+1,$$

for all $(\bar{F},\bar{\eta}) \in \Gamma_K$ and for a constant depending on $(||r||_{L^\infty},\delta,K)$. Then we can proceed as above to bound the term $|\theta(F,\eta)|$. This concludes the proof. $\square$

As in the proof of Theorem 4.2, in order to establish weak-strong uniqueness for the system of adiabatic thermoelasticity (1.4), we use our Gårding-type inequality, Theorem 3.2. In the current setting, due to the fact that the involutions correspond to the differential operator curl, inequality (3.15) takes the form

$$\int |V_p(\nabla\phi)|^2 + |V_q(\psi)|^2 \leqslant C_0 \int e(\bar{F}+\nabla\phi,\bar{\eta}+\psi|\bar{F},\bar{\eta}) + C_1 \int |V_p(\phi)|^2, \tag{4.61}$$

where $C_0 = C_0(e,K) > 0$, $C_1 = C_1(e,K) > 0$ and $\psi \in L^q(Q)$ with $\int \psi = 0$ and $\phi \in W_0^{1,p}(Q)$. We note here that in the RHS we used the equivalent norm for the penalty term, expressed with respect to the primitive functions. In the sequel, we use (4.61) for $(\bar{F},\bar{\eta})$ a classical (Lipschitz) solution of (1.4), to prove the measure-valued versus strong uniqueness result. Hence, the proof of Theorem 4.3 relies on Theorems 3.1 and 3.2, and we refer the reader to those statements for the proof of (4.61).

Lemma 4.2. *Suppose that (ν,γ,F,v,η) is a dissipative measure-valued solution to adiabatic thermoelasticity according to Definition 4.1 and that $(\bar{F},\bar{v},\bar{\eta})$ is a classical solution to* (1.4) *with initial data (F^0,v^0,η^0) and $(\bar{F}^0,\bar{v}^0,\bar{\eta}^0)$ respectively. Under the assumptions of Theorem* 3 *and by denoting $\nu_0 = \nu_{t_0,x}$, it holds that*

$$\begin{aligned}\int &\langle \nu_0, |V_p(\lambda_F - \bar{F}(t_0,x))|^2 + |V_q(\lambda_\eta - \bar{\eta}(t_0,x))|^2\rangle\, dx \\ &\leqslant \tilde{C}_0 \int \langle \nu_0, e(\lambda_F,\lambda_\eta|\bar{F}(t_0,x),\bar{\eta}(t_0,x))\rangle\, dx + \tilde{C}_1 \int |V_p(y(t_0,x)-\bar{y}(t_0,x))|^2\, dx,\end{aligned} \tag{4.62}$$

and

$$\begin{aligned}\int &\langle \nu_0, |V_p(\lambda_F - \bar{F}(t_0,x))|^2 + |V_q(\lambda_\eta - \bar{\eta}(t_0,x))|^2 + |\lambda_v - \bar{v}(t_0,x)|^2\rangle\, dx \\ &\leqslant \tilde{C}_0 \int \langle \nu_0, I(F,v,\eta|\bar{F},\bar{v},\bar{\eta};(t_0,x))\rangle\, dx + \tilde{C}_1 \int |V_p(y(t_0,x)-\bar{y}(t_0,x))|^2\, dx,\end{aligned} \tag{4.63}$$

for almost all $t_0 \in (0,T)$. In addition, at $t=0$

$$\int I(F^0,v^0,\eta^0|\bar{F}^0,\bar{v}^0,\bar{\eta}^0) \leqslant \tilde{c} \int |v^0-\bar{v}^0|^2 + |V_p(F^0-\bar{F}^0)|^2 + |V_q(\eta^0-\bar{\eta}^0)|^2\,. \tag{4.64}$$

Proof. We prove (4.62) as a result of Theorems 4.4 and 3.2. In (4.61) take $\bar{F} = \bar{F}(t_0,\cdot)$ and $\bar{\eta} = \bar{\eta}(t_0,\cdot)$ and then fix $t \in (0,T)$ and choose $\phi = z_k(t,\cdot) - \bar{y}(t_0,\cdot)$ and $\psi = w_k(t,\cdot) - \bar{\eta}(t_0,\cdot)$, see Theorem 4.4. Observe in this case, that the relative quantity $e(\bar{F}+\nabla\phi, \bar{\eta}+\psi|\bar{F},\bar{\eta})$ becomes $e(\nabla z_k, w_k|\bar{F}(t_0,x), \bar{\eta}(t_0,x))$ and Theorem 3.2 gives that

$$\int_Q |V_p(\nabla z_k(t,x)-\bar{F}(t_0,x))|^2 + |V_q(w_k(t,x)-\bar{\eta}(t_0,x))|^2\,dx$$
$$\leqslant \tilde{C}_0 \int_Q e\left(z_k(t,x),\eta_k(t,x)|\bar{F}(t_0,x),\bar{\eta}(t_0,x)\right)\,dx + \tilde{C}_1 \int_Q |V_p(z_k(t,x)) - \bar{y}(t_0,x))|^2\,dx.$$

Integrating the resulting inequality in time and since, from Theorem 4.4 $(\nabla z_k, w_k)$ generates the measure $(\nu_{t_0,x})_{x\in Q}$, $(|\nabla z_k|^p + |w_k|^q)$ is weakly relatively compact in $L^1(Q_T)$ and $z_k \to y(t_0,\cdot)$ strongly in $L^p(Q)$, by taking the limit $k\to\infty$, we obtain inequality (4.62).

For (4.63), we exploit inequality (4.62), together with the fact that the relative entropy I is given as a sum in (4.53). Indeed, let (F^k, v^k, η^k) be a generating sequense satisfying

$$F^k \in L^\infty(0,T;L^p(Q)), \quad v^k \in L^\infty(0,T;L^2(Q)), \quad \eta^k \in L^\infty(0,T;L^q(Q)).$$

Note that whenever $g : Q_T \times \mathbb{R}^{d\times d}\times\mathbb{R}^d\times\mathbb{R}$ is a function that admits an additive decomposition

$$g(t,x,F,v,\eta) = g_v(t,x,v) + g_{F,\eta}(t,x,F,\eta),$$

where

$$|g_v| \leqslant c(1+|v|^2) \quad \text{and} \quad |g_{F,\eta}| \leqslant c(1+|F|^p+|\eta|^q),$$

the action of the generated measure $\nu_{t,x}$ is equivalent with the action of $\nu^v\otimes\nu^{F,\eta}$ where ν^v and $\nu^{F,\eta}$ are generated by the sequenses (v^k) and (F^k,η^k) respectively. Therefore, it suffices to add the term

$$\int \langle \nu_0^v, |v-\bar{v}(t_0,x))|^2\rangle\,dx$$

to the inequality (4.61). Finally, (4.64) follows directly from Lemma 4.1 and in particular bound (4.56). □

Combining the above lemma with the averaged relative entropy inequality (4.55), we are now in a position to prove that in the presence of a classical solution, given that the associated Young measure is initially a Dirac mass, the dissipative measure-valued solution must coincide with the classical one.

Theorem 4.3. *Let $\bar{U}$ be a Lipschitz bounded solution of (1.4),(1.6) with initial data $\bar{U}^0$ and $(\nu_{t,x},\gamma,U)$ be a dissipative measure-valued solution satisfying (4.39),(4.40), with initial data also $\bar{U}^0$, both under the constitutive assumptions (1.7) and such that $r(t,x) = \bar{r}(t,x)\in L^\infty(Q_T)$. Suppose that e is strongly quasiconvex according to Definition 3.1 and the hypotheses $(H_1)-(H_3)$ hold for p, $q\geqslant 2$, together with* (4.59). *If $\nu_{0,x}=\delta_{\bar{U}^0(x)}$ and $\gamma_0 = 0$, we have that $\nu_{t,x}=\delta_{\bar{U}}$ and $U=\bar{U}$ a.e. on Q_T.*

Proof. Let $\{\varphi_n\}$ be a sequence of monotone decreasing functions such that $\varphi_n \geqslant 0$, for all $n \in \mathbb{N}$, converging as $n \to \infty$ to the Lipschitz function

$$\varphi(\tau) = \begin{cases} 1 & 0 \leqslant \tau \leqslant t \\ \dfrac{t-\tau}{\varepsilon} + 1 & t \leqslant \tau \leqslant t+\varepsilon \\ 0 & \tau \geqslant t+\varepsilon \end{cases}$$

for some $\varepsilon > 0$. Writing the relative entropy inequality (4.55) for $r(t,x) = \bar{r}(t,x)$, tested against the functions φ_n we have

$$\begin{aligned} &\int \varphi_n(0)[\langle \nu_{0,x}, I(\lambda_{U_0}|\bar{U_0})\rangle \, dx] \\ &\quad + \int_0^T \int \varphi_n'(t) \left[\langle \nu_{t,x}, I(\lambda_U|\bar{U})\rangle \, dx\, dt + \gamma(dx\, dt)\right] \\ &\quad \geqslant - \int_0^T \int \varphi_n(t) \Big[\partial_t \bar{\eta} \left\langle \nu_{t,x}, \theta(\lambda_F, \lambda_\eta|\bar{F}, \bar{\eta})\right\rangle + \partial_t \bar{F}_{j\beta} \left\langle \nu_{t,x}, \Sigma_{j\beta}(\lambda_F, \lambda_\eta|\bar{F}, \bar{\eta})\right\rangle \\ &\qquad\qquad + \left\langle \nu_{t,x}, (\theta(\lambda_F, \lambda_\eta) - \theta(\bar{F}, \bar{\eta})) \left(\frac{\bar{r}}{\theta(\lambda_F, \lambda_\eta)} - \frac{\bar{r}}{\theta(\bar{F}, \bar{\eta})} \right) \right\rangle \Big] \, dx\, dt. \end{aligned}$$

Passing to the limit as $n \to \infty$ we get

$$\begin{aligned} &\int \left\langle \nu_{0,x}, I(\lambda_F, \lambda_v, \lambda_\eta|\bar{F}, \bar{v}, \bar{\eta})\right\rangle (0,x) \, dx \\ &\quad - \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \int \left[\left\langle \nu_{\tau,x}, I(\lambda_F, \lambda_v, \lambda_\eta|\bar{F}, \bar{v}, \bar{\eta})\right\rangle \, dx\, d\tau + \gamma(dxd\tau)\right] \\ &\quad \geqslant - \int_0^{t+\varepsilon} \int \Big[\partial_t \bar{\eta} \left\langle \nu_{\tau,x}, \theta(\lambda_F, \lambda_\eta|\bar{F}, \bar{\eta})\right\rangle + \partial_t \bar{F}_{j\beta} \left\langle \nu_{\tau,x}, \Sigma_{j\beta}(\lambda_F, \lambda_\eta|\bar{F}, \bar{\eta})\right\rangle \\ &\qquad\qquad + \left\langle \nu_{\tau,x}, (\theta(\lambda_F, \lambda_\eta) - \theta(\bar{F}, \bar{\eta})) \left(\frac{\bar{r}}{\theta(\lambda_F, \lambda_\eta)} - \frac{\bar{r}}{\theta(\bar{F}, \bar{\eta})} \right) \right\rangle \Big] \, dx\, d\tau, \end{aligned}$$

and using the estimates (4.57), (4.58), and (4.60) we arrive at

$$\begin{aligned} &\int \left\langle \nu_{0,x}, I(\lambda_F, \lambda_v, \lambda_\eta|\bar{F}, \bar{v}, \bar{\eta})\right\rangle (0,x) \, dx \\ &\quad - \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \int \left[\left\langle \nu_{\tau,x}, I(\lambda_F, \lambda_v, \lambda_\eta|\bar{F}, \bar{v}, \bar{\eta})\right\rangle \, dx\, d\tau + \gamma(dxd\tau)\right] \\ &\quad \geqslant -C \int_0^{t+\varepsilon} \int \left\langle \nu_{\tau,x}, |V_p(\lambda_F - \bar{F})|^2 + |V_q(\lambda_\eta - \bar{\eta})|^2 \right\rangle \, dx\, d\tau. \end{aligned}$$

Passing now to the limit as $\varepsilon \to 0^+$ and using the fact that $\gamma \geqslant 0$ we get

$$\begin{aligned} \int [\langle \nu_{t,x}, I(\lambda_F, \lambda_v, \lambda_\eta|\bar{F}, \bar{v}, \bar{\eta})\rangle \, dx &\leqslant \int \left\langle \nu_{0,x}, I(\lambda_F, \lambda_v, \lambda_\eta|\bar{F}, \bar{v}, \bar{\eta})\right\rangle (0,x) \, dx \\ &\quad + C \int_0^t \int \left\langle \nu_{\tau,x}, |V_p(\lambda_F - \bar{F})|^2 + |V_q(\lambda_\eta - \bar{\eta})|^2 \right\rangle \, dx\, d\tau, \end{aligned} \tag{4.65}$$

which together with (4.56), (4.63), and (4.64) yields

$$\int \langle \nu_{t,x}, |\lambda_v - \bar{v}|^2 + |V_p(\lambda_F - \bar{F})|^2 + |V_q(\lambda_\eta - \bar{\eta})|^2 \rangle \, dx$$
$$\leqslant C \int_0^t \int \left\langle \nu_{\tau,x}, |V_p(\lambda_F - \bar{F})|^2 + |V_q(\lambda_\eta - \bar{\eta})|^2 \right\rangle dx \, d\tau + C \int |V_p(y - \bar{y})|^2 \, dx, \tag{4.66}$$

for a.e. $t \in (0, T)$. Here we used the assumptions that the two solutions have the same initial data and that $\gamma_0 = 0$. Note that the constant C depends only on the smooth bounded solution $\bar{U}$.

To apply Grönwall's inequality and close our argument it remains to estimate the last term on the right hand-side of (4.66). This was done in [73], using elliptic estimates and equation $(1.4)_1$, together with (4.20). Here, for the sake of completeness, we briefly sketch their arguments. Note that, due to the fact that F and $\bar{F}$ are deformation gradients, involution (4.20), we have that

$$\int_0^T \int_Q (\nabla y - \nabla \bar{y}) \, \phi_t - (u - \bar{u}) \mathrm{div} \phi \, dx dt = 0, \tag{4.67}$$

for all $\phi \in C_c^\infty([0, T); C^\infty(Q))$. Then, consider the unique solution of the system

$$-\Delta g(t, x) = \psi(t, x)$$
$$\int_Q g(t, x) = 0,$$

where $\psi \in C_c^\infty((0, T); C^\infty(Q))$ such that $\int_Q \psi(t, x) = 0$. Taking $\phi = \nabla g$ in (4.67), we infer that

$$\int_0^T \int_Q (y - \bar{y})_t \, \psi - (u - \bar{u}) \psi \, dx dt = 0. \tag{4.68}$$

Note that in the above equation we integrate the time derivative of $y - \bar{y}$. This is indeed well defined, since we have that $\partial_t \nabla y$, $\partial_t \nabla \bar{y} \in L^\infty\left(0, T; H^{-1}(Q)\right)$. We now test (4.68) with the function $(y - \bar{y})\left(1 + |y - \bar{y}|^{p-2}\right) - \int_Q (y - \bar{y})\left(1 + |y - \bar{y}|^{p-2}\right)$, and since $\int_Q (u - \bar{u}) = \int_Q (y - \bar{y}) = 0$, we infer that

$$\frac{d}{dt} \int_Q \frac{|y(t, x) - \bar{y}(t, x)|^2}{2} + \frac{|y(t, x) - \bar{y}(t, x)|^p}{p} dx$$
$$\leqslant \int_Q |u(t, x) - \bar{u}(t, x)| \, |y(t, x) - \bar{y}(t, x)| dx + \int_Q |u(t, x) - \bar{u}(t, x)| \, |y(t, x) - \bar{y}(t, x)|^{p-1} dx.$$

Now, by integrating in time and applying Young's inequality, we get that for almost all $t \in (0, T)$,

$$\int_Q \frac{|y(t, x) - \bar{y}(t, x)|^2}{2} + \frac{|y(t, x) - \bar{y}(t, x)|^p}{p} dx \lesssim \int_0^t \int_Q |u(\tau, x) - \bar{u}(\tau, x)|^2 \, dx d\tau +$$

$$+\int_0^t\int_Q |y(\tau,x)-\bar{y}(\tau,x)|^2\,dxd\tau+\int_0^t\int_Q |y(\tau,x)-\bar{y}(\tau,x)|^{2p-2}\,dxd\tau. \tag{4.69}$$

Then, since $2p-2\geqslant p\geqslant 2$, we use the fact that $W^{1,p}(Q)\hookrightarrow L^{2p-2}(Q)$ to deduce that

$$\begin{aligned}\int_0^t\int_Q |y(\tau,x)-\bar{y}(\tau,x)|^{2p-2} &\lesssim \int_0^t\left(\|y(\tau,\cdot)-\bar{y}(\tau,\cdot)\|^{2p-2}_{L^p(Q)}+\|\nabla y(\tau,\cdot)-\nabla\bar{y}(\tau,\cdot)\|^{2p-2}_{L^p(Q)}\right)\\ &\lesssim \int_0^t\left(\|y(\tau,\cdot)-\bar{y}(\tau,\cdot)\|^{p}_{L^p(Q)}+\|\nabla y(\tau,\cdot)-\nabla\bar{y}(\tau,\cdot)\|^{p}_{L^p(Q)}\right),\end{aligned}$$

where in the last inequality we used the fact that

$$\sup_{\tau\in(0,t)}\left\{\|y(\tau,\cdot)-\bar{y}(\tau,\cdot)\|^{p-2}_{L^p(Q)}+\|\nabla y(\tau,\cdot)-\nabla\bar{y}(\tau,\cdot)\|^{p-2}_{L^p(Q)}\right\}<\infty.$$

Going back to (4.69), we finally conclude that,

$$\begin{aligned}\int |V_p(y-\bar{y})|^2\,dx \leqslant C\int_0^t\int \left\langle \nu_{t,x}, |V_p(\lambda_F-\bar{F})|^2+|\lambda_v-\bar{v}|^2\right\rangle\,dx\,d\tau\\ +C\int_0^t\int V_p(y-\bar{y})|^2\,dx\,d\tau.\end{aligned}$$

Adding the term $\int |V_p(y-\bar{y})|^2\,dx$ on both sides of (4.66) we arrive at

$$\begin{aligned}\int \big(\langle\nu_{t,x}, &|\lambda_v-\bar{v}|^2+|V_p(\lambda_F-\bar{F})|^2+|V_q(\lambda_\eta-\bar{\eta})|^2\rangle+|V_p(y-\bar{y})|^2\big)\,dx\\ &\leqslant C\int_0^t\int \big(\left\langle \nu_{t,x}, |V_p(\lambda_F-\bar{F})|^2+|V_q(\lambda_\eta-\bar{\eta})|^2\right\rangle+|V_p(y-\bar{y})|^2\big)\,dx\,d\tau.\end{aligned}$$

Grönwall's inequality completes the proof. □

Remark 4.5. The radiative heat supply $r(t,x)$ is a field that can be regulated externally. Therefore, one could think instead the theory of thermoelasticity with zero radiative heat supply and prove Theorem 4.3 in this less general setting. In the case $r(t,x)=\bar{r}(t,x)=0$, the result of Theorem 4.3 holds without the assumption on the temperature (4.59), and bound (4.60).

Localisation in time

In Theorem 4.3 we are required to localise our measure-valued solution in time and the generating sequences for these localised measures must be given by a proper time modification of the generating sequence for ν. However, due to the lack of equiintegrability of the assumed generating sequence, we need to construct a new sequence which lies on the desired wave cone, has suitable equiintegrability and convergence properties and generates the localised measure $(\nu_{t_0,x})_{x\in Q}$. To this end, we first prove a technical result which is used for our time-dependent decomposition lemma, Lemma 4.4. The latter is not needed if, instead of measure-valued solutions, weak solutions of (1.4) are considered.

Lemma 4.3. *([Lemma 15,[73]]) Let $v \in L^\infty(0,T;L^p(Q))$ for any $p \in [1,\infty)$. Then, up to a subsequence which is not relabelled and for almost all $t_0 \in (0,T)$, it holds that*

$$\lim_{\varepsilon\to 0}\int_{Q_T} |v(t_0+\varepsilon t/T,x)-v(t_0,x)|^p\,dxdt = 0.$$

Proof. By the continuity of the translations, for almost all t,

$$\|v(\cdot+\varepsilon t/T,\cdot)-v(\cdot,\cdot)\|^p_{L^p(Q_T)} \to 0, \text{ as } \varepsilon \to 0.$$

Also, due to the fact that $v \in L^\infty(0,T;L^p(Q))$, the above quantity is uniformly bounded in ε and hence, by dominated convergence,

$$\lim_{\varepsilon\to 0}\int_0^T\int_{Q_T} |v(t_0+\varepsilon t/T,x)-v(t_0,x)|^p\,dxdtdt_0 = 0,$$

where t_0 is considered as a variable. Indeed, the above limit holds since

$$\begin{aligned}\int_0^T\int_{Q_T} |v(t_0+\varepsilon t/T,x)-v(t_0,x)|^p\,dxdtdt_0 &= \int_0^T\int_{Q_T} |v(t_0+\varepsilon t/T,x)-v(t_0,x)|^p\,dxdt_0dt\\ &= \int_0^T \|v(\cdot+\varepsilon t/T,\cdot)-v(\cdot,\cdot)\|^p_{L^p(Q_T)}\,dt \overset{D.C.T.}{\to} 0.\end{aligned}$$

Hence, up to a subsequence, for almost t_0

$$\lim_{\varepsilon\to 0}\int_{Q_T} |v(t_0+\varepsilon t/T,x)-v(t_0,x)|^p\,dxdt = 0.$$

□

Theorem 4.4. *Let $\nu = (\nu_{t,x})_{(t,x)\in Q_T}$ be a family of probability measures generated by a sequence $(\nabla y_k, \eta_k)$ such that*

$$\begin{aligned}&(y_k) \text{ is bounded in } L^\infty(0,T;W^{1,p}(Q))\\ &(\partial_t \nabla y_k) \text{ is bounded in } L^\infty(0,T;H^{-1}(Q))\\ &(\eta_k) \text{ is bounded in } L^\infty(0,T;L^q(Q)),\end{aligned}$$

and write $(\nabla y,\eta) = \langle \nu, \mathrm{id}\rangle$ for its centre of mass. Then, for almost all $t_0 \in (0,T)$, there exists a sequence $(\nabla z_k, w_k)$ also bounded in $L^\infty(0,T;L^p(Q)) \times L^\infty(0,T;L^q(Q))$ with the following properties

(1) $(\nabla z_k, w_k)$ generates the measure $(\nu_{t_0,x})_{x\in Q}$ as a p-q-Young measure;

(2) $(|\nabla z_k|^p + |w_k|^q)$ is weakly relatively compact in $L^1(Q_T)$;

(3) $z_k \to y(t_0,\cdot)$ strongly in $L^p(Q_T)$.

Proof. For $t_0 \in (0,T)$ define

$$y^{k,\varepsilon}(t,x) := y_k(t_0 + \varepsilon t/T, x), \ \ \eta^{k,\varepsilon}(t,x) := \eta_k(t_0 + \varepsilon t/T, x).$$

We claim that for a.e. t_0 an appropriate subsequence of (ε_k) can be chosen such that $(\nabla y^{k,\varepsilon}, \eta^{k,\varepsilon})$ generates the measure $(\nu_{t_0,x})_{x\in Q}$ and that $y^{k,\varepsilon_k} \to y(t_0,\cdot)$ in $L^p(Q_T)$. To this end, note that, up to a subsequence which is not relabelled, for any $g \in C_{p,q}(\mathbb{R}^{d\times d}\times\mathbb{R})$ and any Borel set $E \subseteq Q_T$ for a.e. $t_0 \in (0,T)$ it holds that

$$\lim_{\varepsilon\to 0}\int_E |\langle \nu_{t_0+\varepsilon t/T,x}, g(\lambda_F,\lambda_\eta)\rangle - \langle \nu_{t_0,x}, g(\lambda_F,\lambda_\eta)\rangle| = 0. \tag{4.70}$$

This is a consequence of Lemma 4.3 noting that the function $v(t,x) = \langle \nu_{t,x}, g\rangle$ is an element of $L^\infty(0,T;L^1(Q))$ since, due to the growth behaviour of g,

$$\sup_t \int_Q |\langle \nu_{t,x}, g\rangle| \lesssim \sup_t \int_Q \langle \nu_{t,x}, |\lambda_F|^p + |\lambda_\eta|^q\rangle < \infty.$$

Hence, it follows that for any such g and E, denoting by $\mathcal{X}_E$ the characteristic function of E and t_0 fixed a.e. in $(0,T)$ using (4.70), we infer that

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\lim_{k\to\infty}\int_E g(\nabla y^{k,\varepsilon}(t,x), \eta^{k,\varepsilon}(t,x)) \\
&= \lim_{\varepsilon\to 0}\lim_{k\to\infty}\frac{T}{\varepsilon}\int_{t_0}^{t_0+\varepsilon}\int_Q \mathcal{X}_E\left((t-t_0)T/\varepsilon, x\right) g(\nabla y^k(t,x), \eta^k(t,x)) \\
&= \lim_{\varepsilon\to 0}\int_{Q_T} \mathcal{X}_E(t,x)\langle \nu_{t_0+\varepsilon t/T,x}, g(\lambda_F,\lambda_\eta)\rangle \\
&= \int_E \langle \nu_{t_0,x}, g(\lambda_F,\lambda_\eta)\rangle.
\end{aligned} \tag{4.71}$$

In addition, similarly with [[73],Lemma 16] we have that

$$\lim_{\varepsilon\to 0}\lim_{k\to\infty}\int_{Q_T} |y_k(t_0+\varepsilon t/T,x) - y(t_0,x)|^p = 0. \tag{4.72}$$

Indeed,

$$\begin{aligned}
\int_{Q_T} |y^k(t_0+\varepsilon t/T,x) - y(t_0,x)|^p \leqslant{} & C\int_{Q_T} |y^k(t_0+\varepsilon t/T,x) - y(t_0+\varepsilon t/T,x)|^p \\
&+ C\int_{Q_T} |y(t_0+\varepsilon t/T,x) - y(t_0,x)|^p =: I + II.
\end{aligned}$$

Then , since $y \in L^\infty(0,T;L^p(Q))$, Lemma 4.3 says that (up to a further subsequence), for a.e. $t_0 \in (0,T)$

$$\lim_{\varepsilon\to 0} II = 0.$$

Regarding the first term, noting that the sequences

$$(y^k) \subset L^\infty(0,T;W^{1,p}(Q)) \ \text{ and } \ (\partial_t y^k) \subset L^2(0,T;L^2(Q))$$

are bounded in the respective spaces, we apply the Aubin-Lions lemma to infer that

$$y^k \to y \text{ in } C(0,T;L^p(Q)),$$

which is enough to conclude that

$$\lim_{k\to\infty} I = 0, \ .$$

Now, for g and E in a countable dense subset of $C_{p,q}(\mathbb{R}^{d\times d}\times\mathbb{R})$ and of the collection of Borel subsets of Q_T, respectively, we may choose a subsequence (ε_k) such that (4.70) and (4.72) hold. In particular, for t_0 fixed almost everywhere in $(0,T)$,

$$\lim_{k\to\infty}\int_E g\left(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k}\right) = \int_E \langle \nu_{t_0,x}, g(\lambda_F,\lambda_\eta)\rangle,$$

for all the elements of the countable subsets where g and E belong and, by density, for all $g\in C_{p,q}(\mathbb{R}^{d\times d}\times\mathbb{R})$ and all $E\subseteq Q_T$, i.e.

$$g\left(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k}\right) \rightharpoonup \langle \nu_{t_0,x}, g(\lambda_F,\lambda_\eta)\rangle \text{ in } L^1(Q_T),$$

and $\left(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k}\right)$ generates the measure $(\nu_{t_0,x})_x$. Note also that

$$(\nabla y^{k,\varepsilon_k})\subseteq L^\infty(0,T;L^p(Q)) \text{ and } (\eta^{k,\varepsilon_k})\subseteq L^\infty(0,T;L^q(Q)).$$

For $n\in\mathbb{N}$ and $(z_1,z_2)\in\mathbb{R}^{d\times d}\times\mathbb{R}$ consider the truncation operator

$$\tau_n(z_1,z_2) := \begin{cases} (z_1,z_2), & |z_1|^2+|z_2|^2\leqslant n^2,\\ n\,(z_1,z_2)/|(z_1,z_2)|, & |z_1|^2+|z_2|^2 > n^2.\end{cases}$$

We observe that $\tau_n(z_1,z_2) = \left(\tau_n^F(z_1,z_2),\tau_n^\eta(z_1,z_2)\right)$ where

$$\tau_n^F(z_1,z_2) := \begin{cases} z_1, & |z_1|^2+|z_2|^2\leqslant n^2,\\ n\,z_1/|(z_1,z_2)|, & |z_1|^2+|z_2|^2 > n^2,\end{cases}$$

and $\tau_n^\eta(z_1,z_2)$ is defined respectively. It is straightforward to see that for fixed $n\in\mathbb{N}$ the sequence $\left(|\tau_n^F(z_1,z_2)|^p+|\tau_n^\eta(z_1,z_2)|^q\right)$ is equiintegrable and so

$$\begin{aligned}&\lim_{n\to\infty}\lim_{k\to\infty}\int_{Q_T}|\tau_n^F(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k})|^p+|\tau_n^\eta(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k})|^q\\ &=\lim_{n\to\infty}\int_{Q_T}\langle\nu_{t_0,x},|\tau_n^F(\lambda_F\lambda_\eta)|^p+|\tau_n^\eta(\lambda_F,\lambda_\eta)|^q\rangle = \int_{Q_T}\langle\nu_{t_0,x},|\lambda_F|^p+|\lambda_\eta|^q\rangle,\end{aligned} \tag{4.73}$$

where the second equality uses monotone convergence. Moreover,

$$\lim_{n\to\infty}\lim_{k\to\infty}\int_{Q_T}|\tau_n(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k})-(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k})| = 0 \tag{4.74}$$

due to the L^1-equiintegrability of $(\nabla y^{k,\varepsilon_k},\eta^{k,\varepsilon_k})$. Then, from (4.73) and (4.74), there exists a subsequence k_n, such that

$$V_n := \left(\tau_n^F(\nabla y^{k_n,\varepsilon_{k_n}},\eta^{k_n,\varepsilon_{k_n}}),\tau_n^\eta(\nabla y^{k_n,\varepsilon_{k_n}},\eta^{k_n,\varepsilon_{k_n}})\right)\xrightarrow{Y}(\nu_{t_0,x})_{x\in Q};$$
$$\left(|\tau_n^F(\nabla y^{k_n,\varepsilon_{k_n}},\eta^{k_n,\varepsilon_{k_n}})|^p+|\tau_n^\eta(\nabla y^{k_n,\varepsilon_{k_n}},\eta^{k_n,\varepsilon_{k_n}})|^q\right)\quad\text{is equiinegrable.}$$

Next, for almost all t, consider the decomposition

$$\tilde V_n := \left(\mathcal{P}_{\text{curl}}\left(V_n^{(1)}-\int_Q V_n^{(1)}\right),\, V_n^{(2)}-\int_Q V_n^{(2)}\right), \tag{4.75}$$

where

$$\tilde V_n^{(1)} := \mathcal{P}_{\text{curl}}\left(V_n^{(1)}-\int_Q V_n^{(1)}\right),\qquad V_n^{(1)} := \tau_n^F(\nabla y^{k_n,\varepsilon_{k_n}},\eta^{k_n,\varepsilon_{k_n}}),$$
$$\tilde V_n^{(2)} := V_n^{(2)}-\int_Q V_n^{(2)},\qquad V_n^{(2)} := \tau_n^\eta(\nabla y^{k_n,\varepsilon_{k_n}},\eta^{k_n,\varepsilon_{k_n}})$$

and $\mathcal{P}_{\text{curl}}$ denotes the projection operator onto curl-free vector fields. For convenience, let us write $y^n := y^{k_n,\varepsilon_{k_n}}$ and $\eta^n := \eta^{k_n,\varepsilon_{k_n}}$ and recall that $\mathcal{P}_{\text{curl}}$ is a strong (r,r) operator, $1<r<\infty$. Then, for a.e. $t\in(0,T)$,

$$\|\tilde V_n^{(1)}(t,\cdot)\|_{L^p(Q)}\leqslant C\|\nabla y^n(t,\cdot)\|_{L^p(Q)}\leqslant C\sup_t\|\nabla y^n(t,\cdot)\|_{L^p(Q)},$$
$$\|\tilde V_n^{(2)}(t,\cdot)\|_{L^q(Q)}\leqslant C\|\eta^n(t,\cdot)\|_{L^q(Q)}\leqslant C\sup_t\|\eta^n(t,\cdot)\|_{L^q(Q)},$$

which shows that the sequences $(\tilde V_n^{(1)})$ and $(\tilde V_n^{(2)})$ are bounded in $L^\infty(0,T;L^p(Q))$ and $L^\infty(0,T;L^q(Q))$ respectively. To see that $(\tilde V_n)$ generates the measure $(\nu_{t_0,x})_{x\in Q}$, note that, denoting by $\mathcal{P}_{\text{div}}$ the projection onto divergence-free vector fields,

$$\begin{aligned}
&|(\nabla y^n(t,\cdot),\eta^n(t,\cdot))-\tilde V_n(t,\cdot)|\leqslant|\nabla y^n(t,\cdot)-\tilde V_n^{(1)}(t,\cdot)|+|\eta^n(t,\cdot)-\tilde V_n^{(2)}(t,\cdot)|\\
&=|\nabla y^n(t,\cdot)-\tau_n^F(\nabla y^n(t,\cdot),\eta^n(t,\cdot))+\mathcal{P}_{\text{div}}\left(\tau_n^F(\nabla y^n(t,\cdot),\eta^n(t,\cdot))-\nabla y^n(t,\cdot)\right)|\\
&\qquad+\left|\eta^n(t,\cdot)-\tau_n^\eta(\nabla y^n(t,\cdot),\eta^n(t,\cdot))+\int_Q\tau_n^\eta(\nabla y^n(t,\cdot),\eta^n(t,\cdot))\right|\\
&\leqslant|\nabla y^n(t,\cdot)-\tau_n^F(\nabla y^n(t,\cdot),\eta^n(t,\cdot))|+|\mathcal{P}_{\text{div}}\left(\tau_n^F(\nabla y^n(t,\cdot),\eta^n(t,\cdot))-\nabla y^n(t,\cdot)\right)|\\
&\qquad+|\eta^n(t,\cdot)-\tau_n^\eta(\nabla y^n(t,\cdot),\eta^n(t,\cdot))|+\left|\int_Q\tau_n^\eta(\nabla y^n(t,\cdot),\eta^n(t,\cdot))\right|=:\sum_{i=1}^4 I_i^n.
\end{aligned}\tag{4.76}$$

However, for any $\varepsilon>0$ and almost all t

$$\begin{aligned}
\mathcal{L}^d\big(\{|(\nabla y^n(t,\cdot),\eta^n(t,\cdot))-\tilde V_n(t,\cdot)|>\varepsilon\}\big)&\leqslant\frac{1}{\varepsilon}\int_{\{|(\nabla y^n(t,\cdot),\eta^n(t,\cdot))-\tilde V_n(t,\cdot)|>\varepsilon\}}\varepsilon dx\\
&\leqslant\frac{1}{\varepsilon}\int_Q|(\nabla y^n(t,\cdot),\eta^n(t,\cdot))-\tilde V_n(t,\cdot)|\overset{(4.76)}{\leqslant}\frac{1}{\varepsilon}\int_Q I_1^n+I_2^n+I_3^n+I_4^n.
\end{aligned}$$

Then, we claim that

$$\mathcal{L}^{d+1}\big(\{\,|\,(\nabla y^n,\eta^n)-\tilde V_n|>\varepsilon\}\big)=\int_0^T \mathcal{L}^d\big(\{|\,(\nabla y^n(t,\cdot),\eta^n(t,\cdot))-\tilde V_n(t,\cdot)|>\varepsilon\}\big)$$
$$\leqslant \frac{C}{\varepsilon}\int_{Q_T} I_1^n+I_2^n+I_3^n+I_4^n\to 0.$$

Indeed, for the first term

$$\begin{aligned}\int_{Q_T} I_1^n &= \|\nabla y^n-\tau_n^F(\nabla y^n,\eta^n)\|_{L^1(Q_T)}\leqslant 2\int_{\{|\nabla y^n|^2+|\eta^n|^2>n^2\}}|\nabla y^n|\\ &\leqslant 2\int_{\{|\nabla y^n|^2+|\eta^n|^2>n^2\}}\frac{(|\nabla y^n|+|\eta^n|)^2}{|\nabla y^n|+|\eta^n|}\leqslant 2\int_{\{|\nabla y^n|+|\eta^n|>n\}}\frac{(|\nabla y^n|+|\eta^n|)^2}{|\nabla y^n|+|\eta^n|}\\ &\leqslant \frac{2}{n}\int_{Q_T}|\nabla y^n|^2+|\eta^n|^2\leqslant\frac{2T}{n}\sup_t\int_Q|\nabla y^n|^2+|\eta^n|^2\to 0,\end{aligned}$$

whenever $n\to\infty$. Similarly we may prove that $\int_{Q_T} I_3^n\to 0$ as $n\to\infty$. Since $\mathcal{P}_{\mathrm{div}}$ is a weak $(1,1)$ operator, the term $\int_{Q_T} I_2^n$ behaves like $\int_{Q_T} I_1^n$ and thus

$$\int_{Q_T} I_2^n\to 0,\quad\text{whenever}\quad n\to\infty.$$

Concerning the last term, since $\|\eta^n-\tau_n^\eta(\nabla y^n,\eta^n)\|_{L^1(Q_T)}\to 0$ and $\int_Q\eta^n(t,\cdot)dx=0$ for a.e. $t\in(0,T)$, we infer that

$$\lim_n\int_0^T\Big|\int_Q\tau_n^\eta(\nabla y^n(t,\cdot),\eta^n(t,\cdot))\Big|dt=0,$$

which concludes the proof of the claim. From the above estimates we infer that the sequence $\tilde V_n:=(\tilde V_n^{(1)},\tilde V_n^{(2)})$ generates the Young measure $(\nu_{t_0,x})_{x\in Q}$.

The equiintegrability of the sequences $\tilde V_n^{(1)}$ and $\tilde V_n^{(2)}$ comes directly from the equiintegrability of $V_n^{(1)}$ and $V_n^{(2)}$. We note that since $\tilde V_n^{(1)}(t,\cdot)$ is curl-free for a.e. $t\in(0,T)$, there exists $z_k\in L^\infty(0,T;W^{1,p}(Q))$ s.t. $\tilde V_n^{(1)}(t,\cdot)=\nabla z^n(t,\cdot)$. In addition, we set $w_n:=\tilde V_n^{(2)}(t,\cdot)$ to serve the requirements of the theorem.

Finally, for the strong convergence of the primitives of the sequence $\tilde V_n^{(1)}$ we follow the proof of [[73],Lemma 16]. We note that by (4.74) and the Lebesgue interpolation theorem, it holds that

$$\lim_n\int_0^T\int_Q|V_n^{(1)}-\nabla y^n|=0\Rightarrow\lim_n\|V_n^{(1)}-\nabla y^n\|_{L^r(L^m)}=0,\tag{4.77}$$

for all $r<\infty$, $m<p$. Since $V_n^{(1)}:=\nabla z^n+\mathcal{P}_{\mathrm{div}}(V_n^{(1)})$, by adding ∇y^n to both sides and taking the divergence we get that

$$-\Delta(z^n-y^n)=\mathrm{div}(\nabla y^n-V_n^{(1)}).\tag{4.78}$$

Then, by standard elliptic estimates, for all $1<m<\infty$ it holds that

$$\|\nabla z^n(t,\cdot)-\nabla y^n(t,\cdot)\|_{L^m(Q)}\lesssim\|\nabla y^n(t,\cdot)-V_n^{(1)}(t,\cdot)\|_{L^m(Q)}.\tag{4.79}$$

In our setting we treat the case $d = 3$ and so letting $m = 3p/(p+3)$, by Sobolev embedding and (4.79) we have that

$$\|z^n(t,\cdot) - y^n(t,\cdot)\|_{L^p(Q)} \lesssim \|\nabla y^n(t,\cdot) - V_n^{(1)}(t,\cdot)\|_{L^m(Q)},$$

and by integrating in time,

$$\int_0^T \|z^n(t,\cdot) - y^n(t,\cdot)\|_{L^p(Q)}^p dt \lesssim \int_0^T \|\nabla y^n(t,\cdot) - V_n^{(1)}(t,\cdot)\|_{L^m(Q)}^p \to 0,$$

as $n \to \infty$. The last convergence comes from (4.77) and concludes the proof of the theorem since, from (4.72), $y^n \to y(t_0,\cdot)$ in $L^p(Q_T)$. □

4.2 Statics: sufficient conditions for local minimisers

The problem of finding sufficient conditions for a smooth extremal to be a strong local minimiser is an old problem. It has been first solved by Weierstrass for the case of one single independent variable, while for scalar variational problems see [65, 21].

Motivated by applications arising from continuum mechanics and material sciences, the question of finding necessary and sufficient conditions for strong local minima in the vectorial case gained great interest. It turned out by the work of Meyers [85] that the notion of quasiconvexity is a necessary condition for strong local minima, while in [10] Ball and Marsden, by introducing the notion of quasiconvexity at the boundary, proved a similar result for minimisers that take free values on part of the boundary of their domain. In addition to this, regarding the sufficiency questions, Ball in [9] conjectured that if the solutions of the weak Euler-Lagrange is sufficiently smooth, then the combination of the strict positivity of the second variation with suitable quasiconvexity-based assumptions should imply that the extremals are strong local minimisers. Later, the work of Taheri [101] in L^p-local minimisers enriched the existing theory around the Weierstrass problem, however it resulted in a convexity-based sufficiency theorem, and hence it left Ball's conjecture open.

After a few attempts which partially answered Ball's conjecture, see [110, 59], it was Grabovsky and Mengesha in [61] who, by following the strategy of [59] enriched with a proper decomposition result, settled a sufficiency theorem for C^1-extremals in the quasiconvex setting. The importance of the C^1-smoothness requirement on the extremals is highlighted by the example of Kristensen and Taheri in [83]. More recently, Campos Cordero in [31] presented an alternative strategy for the sufficiency theorem for C^1-extremals, which also formed the basis of the proof of our Gårding inequality, Theorem 3.2. At this point, we want to emphasise that the strategy of Zhang in [110], which shows that smooth extremals of quasiconvex integrands are minimisers with respect to localised variations, constitutes one of the most crucial tools in this alternative approach by Campos Cordero. The latter can be also revealed from our arguments in Section 3.3, see Lemma 3.6.

In this section, we study functionals of the form

$$\mathcal{W}[U] := \int_Q W(U(x))dx, \tag{4.1}$$

for $U \in L^p_{\mathcal{A}}(Q)$ where we remind that

$$L^p_{\mathcal{A}}(Q) := \left\{ U \in L^p(Q) \,:\, \mathcal{A}U = 0, \int_Q U = 0 \right\}.$$

Motivated by recent developments in the vectorial Weierstrass problem [31, 29, 60], we provide an appropriate generalisation for functionals of the form (4.1) and differential operators other than curl, that is we establish sufficient conditions for local minimisers in the strong $W^{-1,p}$ topology based on $\mathcal{A}$-quasiconvexity assumptions. We remark that the presented result entails a quantitative version of uniqueness for these minimisers, see also Corollary 4.1, which had not been previously observed. The proof comes as a direct consequence of Theorem 3.1 which formed the basis for the Gårding inequality, and its proof has been largely motivated by these recent developments on the Weierstrass problem.

In particular, we prove the following theorem. We note that the natural space of variations for $\mathcal{W}$ is given by

$$\left\{ \varphi \in C(Q) \,:\, \mathcal{A}\varphi = 0, \int_Q \varphi = 0 \right\}.$$

However, under the growth assumptions (h3), see (4.1.1), one may equivalently consider the closure of variations in L^p given by the space $L^p_{\mathcal{A}}(Q)$.

In the sequel, similarly with subsection 4.1.1, we remove the z_2-dependence from the analysis of the Gårding inequality part of the thesis, section 3.3. Having this in mind, we are able to apply Theorem 3.1 for our integrand W, whenever the latter satisfies the growth assumptions $(h1)$-$(h4)$.

Theorem 4.5. *Assume that $W \in C^3(\mathbb{R}^N)$ satisfies (h2), (h3) (see* (4.1.1)*) and let $\bar{U} \in L^p_{\mathcal{A}}(Q) \cap C(\overline{Q})$ such that the following conditions hold:*

- *$\bar{U}$ is a weak solution of the Euler-Lagrange equations, $\mathcal{B}^* DW(\bar{U}) = 0$, i.e.*

$$\int_Q DW(\bar{U}(x))\varphi(x)dx = 0,$$

for all $\varphi \in L^p_{\mathcal{A}}(Q)$;

- *the second variation is strongly positive at $\bar{U}$, i.e. there exists $c > 0$ such that*

$$\int_Q D^2W(\bar{U}(x))\varphi(x) \cdot \varphi(x)dx \geqslant c \int_Q |\varphi(x)|^2 dx,$$

for all $\varphi \in L^p_{\mathcal{A}}(Q)$;

- *W is strongly $\mathcal{A}$-quasiconvex at $\bar{U}(x_0)$ for all $x_0 \in Q$, i.e. there exists $c_0 > 0$ such that*

$$\int_Q \left[W(\bar{U}(x_0) + \varphi(x)) - W(\bar{U}(x_0))\right] dx \geqslant c_0 \int_Q |V(\varphi(x))|^2 dx,$$

for all $\varphi \in L^p_{\mathcal{A}}(Q)$.

Then, there exists $\varepsilon_0 > 0$ and $C > 0$ such that

$$\mathcal{W}[U] - \mathcal{W}[\bar{U}] \geqslant C \int_Q |V(U(x) - \bar{U}(x))|^2 dx,$$

for all $U \in L^p_{\mathcal{A}}(Q)$ with $\|U - \bar{U}\|_{W^{-1,p}(Q)} \leqslant \varepsilon_0$.

Proof. The main ingredient in the proof is Theorem 3.1 combined with the simple observation that if $\bar{U}$ solves the Euler-Lagrange system, then

$$\int_Q W(\bar{U} + \varphi|\bar{U}) = \int_Q \left[W(\bar{U} + \varphi) - W(\bar{U})\right] = \mathcal{W}(\bar{U} + \varphi) - \mathcal{W}(\bar{U}),$$

for any $\varphi \in L^p_{\mathcal{A}}(Q)$. Note that the relative energy $W(\cdot|\cdot)$ is precisely the so-called Weierstrass excess or E-function for the functional $\mathcal{W}$. Thus, given $\bar{U}$ as in the statement, let $U \in L^p_{\mathcal{A}}(Q)$ and set $\varphi = U - \bar{U} \in L^p_{\mathcal{A}}(Q)$. We prove that there exists $\varepsilon_0 > 0$ and $C > 0$ such that

$$\int_Q W(\bar{U} + \varphi|\bar{U}) \geqslant C \int_Q |V(\psi)|^2$$

whenever $\|\varphi\|_{W^{-1,p}(Q)} \leqslant \varepsilon_0$. Note that this is precisely the statement of Theorem 3.1 without the penalty term $\|\varphi\|^2_{W^{-1,(2,p)}}$. One may indeed proceed in the exact same way as in the proof of Theorem 3.1 setting $c_1 = 0$. The only difference lies in Step 5 where Proposition 3.1 is replaced by the stronger assertion that

$$\int_Q D^2\tilde{W}(\bar{U})\varphi \cdot \varphi \gtrsim \int_Q |\varphi|^2 \tag{4.2}$$

which is a consequence of the strong positivity of the second variation of W at $\bar{U}$. Indeed, $\tilde{W}$ is defined in (3.1) as $\tilde{W}(z) = W(z) - C_1|V(z)|^2$ where $C_1 = C_1(W, K)$ can be chosen even smaller if necessary. For $|\lambda| \leqslant K$ and $z \in \mathbb{R}^N$, we compute that

$$|D^2\left(|V(\lambda)|^2\right) z \cdot z| \leqslant 2|z|^2 + p(p-2)|\lambda|^{p-4}\lambda^2|z|^2 + p|\lambda|^{p-2}|z|^2 \leqslant \left(2 + p(p-1)K^{p-2}\right)|z|^2,$$

and hence, due to the positivity of the second variation of W at $\bar{U}$, setting $C := 2 + p(p-1)K^{p-2}$,

$$\int_Q D^2\tilde{W}(\bar{U})\varphi\cdot\varphi = \int_Q D^2W(\bar{U})\varphi\cdot\varphi - C_1\int_Q D^2\left(|V(\bar{U})|^2\right)\varphi\cdot\varphi$$
$$\geqslant c\int_Q |\varphi|^2 - C_1\, C\int_Q |\varphi|^2$$

and we may thus choose $C_1 = C_1(p,K) \leqslant c/(2C)$ so that for $\|\bar{U}\|_{L^\infty} \leqslant K$ and $\varphi \in L^p_{\mathcal{A}}(Q)$, (4.2) holds. This completes the proof. □

Remark 4.6. Note that in the case $\mathcal{A} = \operatorname{curl}$, Theorem 4.5 reduces to a statement about L^p local minimisers, thus recovering partially the result in [31]. In fact this is a statement about L^p local minimisers for any operator $\mathcal{A}$ that admits an elliptic, first-order potential $\mathcal{B}$. Indeed, ellipticity is required to control the L^p norm of the primitive by the $W^{-1,p}$ norm of the function without reverting to properties of the potential operator as in Lemma 2.1.

We also remark that extending the presented result to the case of a bounded domain Ω is nontrivial as, working on the torus, allows for Fourier Analysis tools that are otherwise not available. However, for $\mathcal{A} = \operatorname{curl}$, the above result can be extended in a straightforward way for pure displacement boundary conditions. In fact, with slight modifications one may treat problems with mixed boundary conditions, whereby a part of the boundary remains free. Then, one needs to append the sufficient conditions of Theorem 4.5 with quasiconvexity at the boundary, see [31] as well as [60, 29] for L^∞ local minimisers. Below we show that this is indeed true in the form of a corollary that extends existing results to include a quantitative estimate of uniqueness. The case of functionals depending on lower order terms and L^∞ local minimisers lies outside the scope of the present thesis. We refer the reader to [60, 31, 29] for discussions on quasiconvexity at the boundary. Note that a notion of $\mathcal{A}$-quasiconvexity at the boundary for p-homogeneous functions was defined in [76] in the context of lower semicontinuity for signed integrands.

For the following corollary, let $\Omega \subset \mathbb{R}^d$ a bounded domain with C^1 boundary $\partial\Omega$ such that

$$\partial\Omega = \Gamma_D \cap \Gamma_N$$

where Γ_D is a relatively open subset of $\partial\Omega$ and $\Gamma_N = \partial\Omega\backslash\overline{\Gamma_D}$. We consider the minimisation problem

$$\mathcal{W}(y) = \int_\Omega W(\nabla y(x))\,dx$$

for $y \in W^{1,p}_{y_0,D}(\Omega)$ where for a generic function g we write

$$W^{1,p}_{g,D}(\Omega) = \left\{y \in W^{1,p}(\Omega)\,:\, y = g \text{ on } \Gamma_D\right\},$$

in the sense of trace. We thus interpret Γ_D as the Dirichlet part of the boundary, and Γ_N as the Neumann boundary. Moreover, for a unit vector n, we define the half ball

$$B_n^- := \left\{ x \in \mathbb{R}^d \, : \, |x| < 1, \, x \cdot n < 0 \right\}.$$

Corollary 4.1. *Assume that $W \in C^2(\mathbb{R}^{n \times d})$ satisfies (h2), (h3) (see (4.1.1)) and let $\bar{y} \in C^1(\overline{\Omega}) \cap W^{1,p}_{y_0,D}(\Omega)$ such that the following conditions hold:*

- *$\bar{y}$ is a weak solution of the Euler-Lagrange equations, $\mathrm{div} DW(\nabla \bar{y}) = 0$, i.e.*

$$\int_\Omega DW(\nabla \bar{y}(x)) \nabla \phi(x) dx = 0,$$

for all $\phi \in C^1(\Omega) \cap W^{1,p}_{0,D}(\Omega)$;

- *the second variation is strongly positive at $\bar{y}$, i.e.*

$$\int_\Omega D^2 W(\nabla \bar{y}(x)) \nabla \phi(x) \cdot \nabla \phi(x) dx \geqslant c \int_\Omega |\nabla \phi(x)|^2 dx,$$

for all $\phi \in C^1(\Omega) \cap W^{1,p}_{0,D}(\Omega)$;

- *W is strongly quasiconvex at $\nabla \bar{y}(x_0)$ for all $x_0 \in \overline{\Omega}$, i.e.*

$$\int_B \left[W(\nabla \bar{y}(x_0) + \nabla \phi(x)) - W(\nabla \bar{y}(x_0)) \right] dx \geqslant c_0 \int_B |V(\nabla \phi(x))|^2 dx,$$

for all $\phi \in W^{1,p}_0(B)$, where B denotes the unit ball in $\mathbb{R}^d$;

- *W is strongly quasiconvex at $\nabla \bar{y}(x_0)$ for all $x_0 \in \Gamma_N$, i.e. denoting by $n(x_0)$ the outward pointing unit normal at $x_0 \in \Gamma_N$,*

$$\int_{B^-_{n(x_0)}} W(\nabla \bar{y}(x_0) + \nabla \phi(x) | \nabla \bar{y}(x_0)) dx \geqslant c_0 \int_{B^-_{n(x_0)}} |V(\nabla \phi(x))|^2 dx,$$

for all $\phi \in W^{1,p}(B^-_{n(x_0)})$ such that $\phi = 0$ on $\partial B \cap \overline{B^-_{n(x_0)}}$.

Then, there exists $\varepsilon_0 > 0$ and $C > 0$ such that

$$\mathcal{W}[y] - \mathcal{W}[\bar{y}] \geqslant C \int_\Omega |V(\nabla y(x) - \nabla \bar{y}(x))|^2 dx,$$

for all $y \in W^{1,p}_{y_0,D}(\Omega)$ with $\|y - \bar{y}\|_{L^p(\Omega)} \leqslant \varepsilon_0$.

Proof. The proof that (h2), (h3), the strong positivity of the second variation and the quasiconvexity conditions imply that

$$\int_\Omega W(\nabla \bar{y} + \nabla \phi | \nabla \bar{y}) \geqslant 0, \tag{4.3}$$

is given in [31, 29]. Note that the proof relies on proving Proposition 3.2 also for points on Γ_N and appropriate test functions, using the quasiconvexity at the boundary. This is the content of [29, Proposition 4.6] where, due to the presence of lower order terms, L^∞ assumptions are needed which are not required here. Proposition 3.2 replaces the quasiconvexity conditions for the rest of the proof which thus remains the same. Then, the satisfaction of the Euler-Lagrange equations implies that (4.3) gives the minimality of $\bar{y}$.

Thus, in order to obtain the lower bound and the quantitative estimate of uniqueness, we prove (4.3) for the function $\tilde{W}$, in place of W. In particular, we need to find a constant $C_1 = C_1(W, \|\bar{y}\|_{C^1})$ such that $\tilde{W}$ satisfies (h2), (h3), the strong positivity of the second variation, as well as the quasiconvexity conditions. That C_1 can be chosen so that $(h2)$ and $(h3)$ holds is straightforward, while quasiconvexity is the content of Lemma 3.4. That the second variation is strongly positive is part of the proof of Theorem 4.5 and we are thus left to infer the quasiconvexity at the boundary. Denoting by $f(\lambda) = |V(\lambda)|^2$, we compute

$$\begin{aligned}\int_{B^-_{n(x_0)}} \tilde{W}(\nabla\bar{y}(x_0) + \nabla\phi|\nabla\bar{y}(x_0)) &= \int_{B^-_{n(x_0)}} W(\nabla\bar{y}(x_0) + \nabla\phi|\nabla\bar{y}(x_0)) \\ - C_1 \int_{B^-_{n(x_0)}} f(\nabla\bar{y}(x_0) + \nabla\phi(x)|\nabla\bar{y}(x_0)) &\geqslant (c_0 - C_1 C) \int_{B^-_{n(x_0)}} |V(\nabla\phi)|^2,\end{aligned}$$

by the strong quasiconvexity at the boundary of W and Lemma 3.3 (a). We may thus choose $C_1 \leqslant c_0/(2C)$ to complete the proof. □

Chapter 5

Bilevel training schemes in imaging: total-variation-type functionals with convex integrands

5.1 Mathematical setup

We collect in this section all the assumptions and the notation to be used in the sequel. We also include some heuristic motivation for the definition of the class of admissible weights.

5.1.1 Functional setting: $\mathrm{BV}_p^{\mathscr{B}}$ spaces.

We work in the d-dimensional Euclidean space $\mathbb{R}^d$, $d \geqslant 2$, that we endow with the Lebesgue measure $\mathscr{L}^d$. We let $\Omega \subset \mathbb{R}^d$ be a fixed open and bounded set with Lipschitz boundary, which stands as the image domain. In typical applications $d = 2$ and Ω is a rectangle. We suppose that the image functions take values in a finite dimensional inner product space $\mathbb{U}$, which, for instance, is $\mathbb{R}$ for grayscale images, $\mathbb{R}^3$ for RGB images, or it can be even more structured like e.g. $\mathcal{S}^{d\times d}$ for diffusion tensor imaging [106]. In order to describe further the functional setting in which our analysis is carried out, we need to introduce the class of differential operators that we consider.

Let $\mathbb{V}$ be another finite dimensional inner product space and let $\mathrm{Lin}(\mathbb{U},\mathbb{V})$ be the space of linear maps from $\mathbb{U}$ to $\mathbb{V}$. Hereafter, for $l \in \mathbb{N}\setminus\{0\}$, $\mathscr{B}$ denotes a l-th order, homogeneous and linear differential operator with constant coefficients. Explicitly, as we discussed in section 2.1, given $B_i \in \mathrm{Lin}(\mathbb{U},\mathbb{V})$ for any d-dimensional multi-index j, we define for a smooth function $u\colon \mathbb{R}^d \to \mathbb{U}$

$$\mathscr{B}u \coloneqq \sum_{|j|=l} B_j \partial^j u.$$

When u is less regular, we interpret $\mathcal{B}u$ in the distributional sense. In particular, we are interested in the case in which $\mathcal{B}u$ is a finite Radon measure.

Given a generic open set $O \subset \mathbb{R}^d$, we recall that a finite ($\mathbb{U}$-valued) Radon measure on O is a measure on the σ-algebra of the Borel sets of O. We denote the space of such measures by $\mathcal{M}(O, \mathbb{U})$, and, by means of the classical Riesz's representation theorem, we can identify it as the dual of the space

$$\mathrm{C}_0(O, \mathbb{U}) := \big\{u \colon O \to \mathbb{U} : \{|u| > \delta\} \text{ is relatively compact for all } \delta > 0\big\},$$

equipped with the uniform norm. The dual norm induced on $\mathcal{M}(O, \mathbb{U})$ turns out to be the one associated with the total variation, which we denote by $|\bullet|$. We refer to [5, Chapter 1] for further reading on measure theory.

In our case, given $\mu \in \mathcal{M}(\Omega, \mathbb{U})$, we have that $\mathcal{B}\mu \in \mathcal{M}(\Omega, \mathbb{V})$ if and only if there exists $\nu \in \mathcal{M}(\Omega, \mathbb{V})$ such that

$$\langle \nu, \phi \rangle = \int_\Omega \mathcal{B}^* \phi \, \mathrm{d}\mu \qquad \text{for all } \phi \in \mathrm{C}_c^\infty(\Omega, \mathbb{V}),$$

where $\langle \bullet, \bullet \rangle$ denotes the duality pairing and $\mathcal{B}^*$ is the formal adjoint of $\mathcal{B}$, i.e.

$$\mathcal{B}^* \phi := -\sum_{|j|=l} B_j^* \partial^j \phi \qquad \text{for all } \phi \in \mathrm{C}_c^\infty(\mathbb{R}^d, \mathbb{V}),$$

B_j^* being the transpose of B_j.

It is convenient to have at our disposal a specific notation for the spaces that we are going to work with. For Ω, $\mathbb{U}$ and $\mathbb{V}$ as above, and for $p \in (1, +\infty)$, we set

$$\mathrm{BV}_p^{\mathcal{B}}(\Omega) := \{u \in L^p(\Omega, \mathbb{U}) \colon \mathcal{B}u \in \mathcal{M}(\Omega, \mathbb{V})\},$$

and we abbreviate $\mathrm{BV}^{\mathcal{B}}(\Omega) := \mathrm{BV}_p^{\mathcal{B}}(\Omega)$ when $p = 1$. The spaces above are naturally endowed with weak-$*$ notions of convergence, namely

$$u_j \overset{*}{\rightharpoonup} u \text{ in } \mathrm{BV}_p^{\mathcal{B}}(\Omega) \quad \text{if and only if} \quad u_j \rightharpoonup u \text{ in } L^p(\Omega) \text{ and } \mathcal{B}u_j \overset{*}{\rightharpoonup} \mathcal{B}u \text{ in } \mathcal{M}(\Omega, \mathbb{V}).$$

Stronger convergence may be retrieved if the class of differential operators is restricted. We give a brief account on this point in the following lines.

Due to the interaction between Fourier transform and linear PDE, often analytic properties of $\mathrm{BV}^{\mathcal{B}}$ spaces (and of the equation $\mathcal{B}u = v$ in general) can be expressed in terms of algebraic properties of the characteristic polynomial. We recall that the characteristic polynomial, or symbol, of $\mathcal{B}$ is

$$\mathbb{B}(\xi) := \sum_{|j|=l} B_j \xi^j \in \mathrm{Lin}(\mathbb{U}, \mathbb{V}), \qquad \xi \in \mathbb{C}^d,$$

where $\xi^j := \xi_1^{j_1} \cdots \xi_d^{j_d}$. In our study, the following property will be particularly relevant:

Definition 5.1 ([99, 15, 58])**.** An operator $\mathcal{B}$ is said to be $\mathbb{C}$-elliptic if

$$\ker_{\mathbb{C}} \mathbb{B}(\xi) = \{0\} \quad \text{for all } \xi \in \mathbb{C}^d \setminus \{0\}.$$

It was shown in [99] that $\mathbb{C}$-ellipticity is equivalent with full Sobolev regularity for the equation $\mathcal{B}u = v$ on Lipschitz domains, provided that $v \in L^p(\Omega, \mathbb{V})$, $p \in (1, +\infty)$. For $p = 1$ we have the counterpart:

Theorem 5.2 ([58])**.** *Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain. An operator $\mathcal{B}$ is $\mathbb{C}$-elliptic if and only if*

$$\|u\|_{\mathrm{W}^{k-1,d/(d-1)}(\Omega)} \leqslant c \left(|\mathcal{B}u|(\Omega) + \|u\|_{L^1(\Omega)} \right) \quad \textit{for } u \in \mathrm{BV}^{\mathcal{B}}(\Omega),$$

where $|\mathcal{B}u|$ is the total variation measure associated with $\mathcal{B}u$.

In particular, the above result constitutes a generalisation of the Gagliardo-Niremberg-Sobolev inequality on domains Ω and of the Korn-Sobolev inequality in [104]. The proof relies on the extension of maps in $C^\infty(\bar{\Omega}, \mathbb{U})$ to the full-space boundedly and the application of an estimate of the form

$$\|D^{l-1}u\|_{L^{1^*}(\mathbb{R}^d)} \lesssim \|\mathcal{B}u\|_{L^1(\mathbb{R}^d)}.$$

The latter inequality, which characterise a particular class of operators, see [107], is guaranteed by a novel result of [58], where the authors proved that the $\mathbb{C}$-ellipticity of $\mathcal{B}$ implies that

$$\bigcap_{\xi \in \mathbb{R}^d \setminus \{0\}} \operatorname{im} \mathbb{B}(\xi) = \{0\},$$

the so-called cancelling property. In other words, Theorem 5.2 says that an operator $\mathcal{B}$, which is defined as above, is $\mathbb{C}$-elliptic if and only if

$$\mathrm{BV}^{\mathcal{B}}(\Omega) \subset W^{l-1,d/(d-1)}(\Omega, \mathbb{U}).$$

5.1.2 The bilevel scheme

We are now in a position to formulate our problem rigorously. Let us fix $p \in (1, +\infty)$. As we have touched upon in the introduction, our goal is to provide an existence result for solutions to the following training scheme: given $g \in L^p(\Omega, \mathbb{U})$,

$$\text{find } \alpha^* \in \operatorname{argmin} \{F(u_\alpha) : \alpha \in \mathrm{Adm}\} \tag{L1}$$

$$\text{such that } u_\alpha \in \operatorname{argmin} \left\{ I[u; \alpha] \colon u \in \mathrm{BV}_p^{\mathcal{B}}(\Omega) \right\}, \tag{L2}$$

where

$$I[u;\alpha] := \Phi_g(u) + \int_\Omega \alpha(x) f(x, \mathrm{d}\mathcal{B}u). \tag{5.2}$$

All the due definitions and assumptions are collected below.

– *Cost functional:* As for the upper level problem (L1), $F\colon L^p(\Omega,\mathbb{U}) \to \mathbb{R}$ is a proper, convex and weakly lower semicontinuous functional. Typical choices for this functional are the Peak Signal-to-Noise Ratio (PSNR) maximising F_{PSNR} in (1.9), which makes use of the ground truth u_{gt}, and the statistics-based, ground truth-free F_{stat} in (1.10), in the spirit of supervised and unsupervised learning respectively.

– *Fidelity term:* The assumptions on the functional $\Phi_g\colon L^p(\Omega,\mathbb{U}) \to \mathbb{R}$ in (5.2) are similar to the ones on F, namely Φ_g is a proper, convex and weakly lower semicontinuous functional that is also coercive. This means that

$$\lim_{j\to+\infty} \|u_j - g\|_{L^p(\Omega,\mathbb{U})} = +\infty \quad \text{implies} \quad \lim_{j\to+\infty} \Phi_g(u_j) = +\infty.$$

In particular,

$$\Phi_g(u) = \|u_j - g\|^p_{L^p(\Omega,\mathbb{U})}$$

is a simple instance of a fidelity term.

– *Weights:* Given $\underline{\alpha}, \overline{\alpha} \geqslant 0$ with $\underline{\alpha} < \overline{\alpha}$, the scalar fields $\alpha \in \mathrm{C}(\bar{\Omega}, [\underline{\alpha}, \overline{\alpha}])$ are supposed to share the same uniform modulus of continuity ω, that is, an increasing function $\omega\colon [0,+\infty) \to [0,+\infty)$ such that $\omega(0) = 0$. As a consequence, the class of admissible weights

$$\mathrm{Adm} := \left\{\alpha \in \mathrm{C}(\bar{\Omega}, [\underline{\alpha}, \overline{\alpha}]) : |\alpha(x) - \alpha(y)| \leqslant \omega(|x-y|) \text{ for every } x, y \in \overline{\Omega}\right\}, \tag{5.3}$$

is compact with respect to the uniform norm by Arzelà–Ascoli theorem. We will motivate the definition of the set Adm below, see Subsection 5.1.3.

– *Integrand:* The function $f\colon \Omega \times \mathbb{V} \to [0,+\infty)$ is a *Carathéodory* integrand such that $z \mapsto f(x,z)$ is convex for $\mathcal{L}^d$-a.e. $x \in \Omega$. We remind that, Carathéodory integrand means that $f(\cdot, z)$ is Borel measurable for all $z \in \mathbb{V}$ and $f(x,\cdot)$ is continuous for a.e. $x \in \Omega$. We also suppose that the integrand satisfies the linear coercivity and growth bounds

$$c(|\bullet| - 1) \leqslant f(x,\bullet) \leqslant C(1 + |\bullet|) \quad \text{for } \mathcal{L}^d\text{-a.e. } x \in \Omega, \tag{5.4}$$

for some $c, C \geqslant 0$.

We note that (5.2) makes sense via the discussion we made in section 2.3. In particular, for an integrand $f\colon \Omega \times \mathbb{V} \to \mathbb{R}$ satisfying (5.4), we recall that the *recession function* is defined as

$$f^\infty(x,z) := \lim_{(x',z',t)\to(x,z,+\infty)} \frac{f(x',tz')}{t} \quad \text{for } (x,z) \in \bar{\Omega} \times \mathbb{V}, \tag{5.5}$$

which we assume exists and is jointly continuous. We also remind that, for $\mu \in \mathcal{M}(\Omega, \mathbb{V})$,

$$\int_\Omega f(x, \mathrm{d}\mu) := \int_\Omega f\left(x, \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^n}(x)\right) \mathrm{d}x + \int_\Omega f^\infty\left(x, \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu|}(x)\right) \mathrm{d}|\mu|(x), \tag{5.6}$$

where μ^s denotes the singular part of μ with respect to Lebesgue measure and $\mathrm{d}\mu/\mathrm{d}\nu$ is the Radon-Nikodým derivative of μ with respect to the measure ν. To summarise, in this chapter we consider integrands

$$f \in \mathbb{L}^+(\Omega, \mathbb{U}) \quad \text{such that} \quad z \mapsto f(x,z) \text{ is convex, and } c|\cdot| \leqslant f(x,\cdot) \text{ for } \mathcal{L}^d\text{-a.e. } x \in \Omega,$$

where the space $\mathbb{L}^+(\Omega, \mathbb{U})$ contains all the non-negative valued functions of $\mathbb{L}(\Omega, \mathbb{U})$, see Section 2.3 for the respective definition.

5.1.3 Rationale for the definition of the set of admissible weights

In order to highlight the main technical obstacles that are encountered in the analysis of bilevel training schemes with space-dependent weights, we start with an example involving the total variation with spatially varying weight, which, in spite of its simplicity, exhibits the typical features of such class of problems. The model we address has been already studied in [68], using a slightly different approach from the one we outline.

Let $\Omega \subset \mathbb{R}^d$ be a bounded open set with Lipschitz boundary. For $p \in [1, \frac{d}{d-1})$, we suppose that a training pair $(u_{\mathrm{gt}}, g) \in L^2(\Omega, \mathbb{R}) \times L^p(\Omega, \mathbb{R})$ is assigned, where u_{gt} and g encode the ground truth and the corrupted datum respectively. We also fix two positive parameters $\underline{\alpha}$ and $\overline{\alpha}$, and we provisionally allow the regularising weights to vary in $\mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])$, the space of lower semicontinuous functions on Ω with range in $[\underline{\alpha}, \overline{\alpha}]$.

For $u \in \mathrm{BV}(\Omega)$ and $\alpha \in \mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])$, we introduce the first order functional

$$J[u;\alpha] := \int_\Omega |u-g|^p \mathrm{d}x + \int_\Omega \alpha(x) \mathrm{d}|\mathrm{D}u|(x), \tag{5.7}$$

and the ensuing corresponding training scheme:

$$\text{find } \alpha^* \in \operatorname{argmin}\left\{F_{\mathrm{PSNR}}(u_\alpha) : \alpha \in \mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])\right\} \tag{5.8}$$

$$\text{such that } u_\alpha \in \operatorname{argmin}\left\{J[u;\alpha] \colon u \in \mathrm{BV}(\Omega)\right\}. \tag{5.9}$$

The functional in (5.7) is reminiscent of the one considered in [8], where, motivated by vortex density models, the authors studied the property of minimisers, i.e., of solutions to (5.9).

Before discussing the existence of solutions to the scheme (5.8)–(5.9) as a whole, let us justify the choice of the class of weights in (5.8). Note that the definition of J itself calls for some degree of regularity for α. Indeed, if in (5.7) $\alpha\colon \Omega \to [\underline{\alpha}, \overline{\alpha}]$ is a given function and u is allowed to vary in $\mathrm{BV}(\Omega)$ (as it is the case of (5.9)), there might be choices of u for which the coupling

$$\int_\Omega \alpha(x) \mathrm{d}|\mathrm{D}u|(x)$$

is not well-defined. Prescribing lower semicontinuity for the admissible weights α allows to circumvent the issue, because lower semicontinuous functions are Borel measurable and $\mathrm{D}u \in \mathcal{M}(\Omega, \mathbb{R}^d)$ is a Borel measure. Besides, for any $\alpha \in \mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])$ the existence of a solution u_α to (5.9) follows by the direct method of the calculus of variations. Indeed, we firstly observe that the coercivity of J in L^1 is deduced by the following standard result (see e.g. [5, Theorem 3.23]):

Theorem 5.3 (Compactness in BV)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let $(u_j)_j$ be a bounded sequence in $\mathrm{BV}(\Omega)$. Then, there exist $u \in \mathrm{BV}(\Omega)$ and a subsequence of (u_j) (which is not relabelled), such that (u_j) weakly-$*$ converges to u, that is, $u_j \to u$ in $L^1(\Omega)$ and*

$$\lim_{j\to+\infty} \int_\Omega \phi \, \mathrm{d}\mathrm{D}u_j = \int_\Omega \phi \, \mathrm{d}\mathrm{D}u \qquad \textit{for all } \phi \in \mathrm{C}_0(\Omega).$$

Additionally, we notice that $J[\,\bullet\,;\alpha]$ is lower semicontinuous with respect to the L^1-convergence, because, when $\alpha \in \mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])$, general lower semicontinuity results in BV may be invoked (see e.g. [50]; and also [4] for lower semicontinuity and relaxation results with BV integrands).

Once we know that, due to the lower semicontinuity of the weights, solutions to (5.9) exist, we can handle the complete scheme. So, let $(\alpha_j)_j \subset \mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])$ be a minimising sequence for (5.8). Then, by definition, the integrals

$$F_{\mathrm{PSNR}}(u_j) = \|u_j - u_{\mathrm{gt}}\|^2_{L^2(\Omega)}, \qquad \text{with } u_j := u_{\alpha_j},$$

converge, and we deduce that $(u_j)_j$ is a bounded sequence in $L^2(\Omega)$. Denote by $u \in L^2(\Omega)$ the weak L^2-limit of (a subsequence of) $(u_j)_j$. By lower semicontinuity of the L^2-norm, we obtain

$$F_{\mathrm{PSNR}}(u) \leqslant \liminf_j F_{\mathrm{PSNR}}(u_j) = \inf\left\{F_{\mathrm{PSNR}}(u_j) : \alpha \in \mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])\right\}.$$

If we manage to show that $u = u_{\alpha^*}$ for some admissible α^*, then the latter is a solution to (5.8). The natural choice for α^* would be the weak-$*$ limit of $(\alpha_j)_j$ in $L^\infty(\Omega)$, which can fail in general to have any lower semicontinuous representative. On the positive side, $J[u;\bullet]$ is continuous with respect to a suitable weak-$*$ convergence. Indeed, if $(\alpha_j)_j \subset \mathrm{LSC}(\Omega, [\underline{\alpha}, \overline{\alpha}])$ is bounded and $u \in \mathrm{BV}(\Omega)$, then there exist a subsequence, which we do not relabel, and $\alpha^* \in L^\infty(\Omega, [\underline{\alpha}, \overline{\alpha}]; |\mathrm{D}u|)$ such that

$$\lim_{j\to+\infty} \int_\Omega \alpha_j(x)\phi(x)\mathrm{d}|\mathrm{D}u|(x) = \int_\Omega \alpha^*(x)\phi(x)\mathrm{d}|\mathrm{D}u|(x) \qquad \text{for all } \phi \in L^1(\Omega; |\mathrm{D}u|).$$

In particular,

$$\lim_{j\to+\infty} J[u;\alpha_j] = J[u;\alpha^*]. \tag{5.10}$$

The previous lines suggest that what is missing to solve the scheme (5.8)-(5.9) is a compactness property for the class of admissible weights. This leads us to reduce ourselves to the problem

$$\begin{aligned} &\text{find } \alpha^* \in \operatorname{argmin}\{F_{\mathrm{PSNR}}(u_\alpha) : \alpha \in \mathrm{Adm}\} \\ &\text{with } u_\alpha \in \operatorname{argmin}\{J[u;\alpha] \colon u \in \mathrm{BV}(\Omega)\}, \end{aligned} \tag{5.11}$$

where we assume *a priori* that $\mathrm{Adm} \subset \mathrm{C}(\bar{\Omega}, [\underline{\alpha}, \overline{\alpha}])$ is compact with respect to the uniform convergence. Under the compactness assumptions on the class of admissible weights, if $(\alpha_j)_j \subset \mathrm{Adm}$ is a minimising sequence for (5.11), and if $(u_j)_j$ and u are constructed as above, we are actually able to prove that $u = u_{\alpha^*}$, where $\alpha^* \in \mathrm{Adm}$ is the uniform limit of (α_j). In other words, the couple (α^*, u) is a solution to the scheme consisting of (5.11)–(5.9).

To prove the claim, we need to show that

$$J[u;\alpha^*] \leqslant J[v;\alpha^*] \qquad \text{for any } v \in \mathrm{BV}(\Omega). \tag{5.12}$$

We start from observing that from the definition of u_j we get

$$J[u_j;\alpha_j] \leqslant J[v;\alpha_j] \qquad \text{for any } v \in \mathrm{BV}(\Omega),$$

and hence, for any $v \in \mathrm{BV}(\Omega)$,

$$\liminf_{j\to+\infty} J[u_j;\alpha_j] \leqslant \liminf_{j\to+\infty} J[v;\alpha_j] = J[v;\alpha^*], \tag{5.13}$$

where the equality follows by (5.10). In particular,

$$\liminf_{j\to+\infty} J[u_j;\alpha_j] < +\infty.$$

Then, the uniform lower bound $\alpha_j \geqslant \underline{\alpha}$ and Theorem 5.3 yield that (u_j) converges weakly-$*$ in $\mathrm{BV}(\Omega)$ (again upon extraction of subsequences) to a limit function which is necessarily u. We thereby infer

$$u \in \mathrm{BV}(\Omega) \cap L^2(\Omega).$$

Finally, from the weak-$*$ lower semicontinuity of $J[\cdot, \alpha^*]$ and the uniform convergence of (α_j) we obtain that

$$J[u; \alpha^*] \leqslant \liminf_{j \to +\infty} J[u_j; \alpha^*] = \liminf_{j \to +\infty} J[u_j; \alpha_j] \overset{(5.13)}{\leqslant} J[v; \alpha^*], \tag{5.14}$$

which is the desired inequality.

Remark 5.1. In the absence of compactness for the set Adm under uniform convergence, the analysis becomes more delicate. We outline here some of the issues.

Keeping in force the notation above, let $u \in \mathrm{BV}(\Omega)$ be the weak-$*$ limit of (u_j) and let $\alpha^* \in L^\infty(\Omega, [\underline{\alpha}, \overline{\alpha}]; |\mathrm{D}u|)$ be the weak-$*$ limit of (α_j). Proving the optimality of u, i.e. $u = u_{\alpha^*}$, means

$$J[u; \alpha^*] \leqslant J[v; \alpha^*] \qquad \text{for any } v \in \mathrm{BV}(\Omega).$$

However, the right-hand side might not be well-defined. Intuitively, an ideal class of weights should be *a priori* "sufficiently compact", and at the same time it should give rise to "well-behaved" weighted BV functions.

Another passage that is needed in the proof of existence (cf. (5.6), (5.7)) is the following semicontinuity inequality:

$$J[u; \alpha^*] \leqslant \liminf_{j \to +\infty} J[u_j; \alpha^{(j)}].$$

Knowing that (u_j) weakly-$*$ converges to u, its validity is undermined if only weak-$*$ convergence is available for the weights.

5.2 Existence theorems for the lower level problems

We begin with a general lower semicontinuity result for convex integrands with rough x-dependence, Proposition 5.1. Before we present this result, we will prove a technical lemma which is needed in the latter's proof.

Lemma 5.1. *Let $g : \mathbb{V} \to \mathbb{R}^+$ be a convex function with linear growth at infinity. Then,*

$$g(z) + tg^\infty(w) \geqslant g(z + tw),$$

for all z, $w \in \mathbb{V}$ and $t \geqslant 0$.

Proof. We observe that

$$z + tw = \frac{t}{s}(z + sw) + (1 - \frac{t}{s})z,$$

and hence, due to the convexity of g and for all $s \geqslant t$, it holds that

$$g(z + tw) \leqslant \frac{t}{s} g(z + sw) + (1 - \frac{t}{s})g(z) \Rightarrow \frac{g(z + tw) - g(z)}{t} \leqslant \frac{g(z + sw) - g(z)}{s}.$$

Taking the limit $s \to \infty$ in the above inequality, we infer that

$$\frac{g(z+tw)-g(z)}{t} \leqslant \lim_{s\to\infty} \frac{g(z+sw)-g(z)}{s} = \lim_{s\to\infty} \frac{g(z+sw)}{s}.$$

However, since g is convex and it has linear growth, it is also Lipshitz and thus

$$\lim_{s\to\infty} \left[\frac{g(z+sw)}{s} - \frac{g(sw)}{s} \right] \lesssim \lim_{s\to\infty} \frac{|z|}{s} = 0.$$

Combining the above we finally deduce the required inequality

$$\frac{g(z+tw)-g(z)}{t} \leqslant \lim_{s\to\infty} \frac{g(sw)}{s} = g^{\infty}(w).$$

□

Proposition 5.1. *Let $f \in \mathbb{L}^{+}(\Omega, \mathbb{V})$ such that $f(x, \bullet)$ is convex for almost every $x \in \Omega$. Then, it holds that*

$$\mu_j \overset{*}{\rightharpoonup} \mu \text{ in } \mathcal{M}(\Omega, \mathbb{V}) \implies \liminf_{j\to\infty} \int_{\Omega} f(x, \mathrm{d}\mu_j(x)) \geqslant \int_{\Omega} f(x, \mathrm{d}\mu(x)).$$

Proof. From the weak-* convergence, due to Theorem 2.7, we consider the associated Young measure $\boldsymbol{\nu}$ which is generated by the sequence $(\mu_j)_j$. By Proposition 2.2 and Jensen's inequality Theorem 2.8, we have that

$$\begin{aligned} \liminf_{j\to\infty} \int_{\Omega} f(x, \mathrm{d}\mu_j(x)) &= \int_{\Omega} \langle \nu_x, f(x, \bullet) \rangle \mathrm{d}x + \int_{\bar{\Omega}} \langle \nu_x^{\infty}, f^{\infty}(x, \bullet) \rangle \mathrm{d}\lambda \\ &\geqslant \int_{\Omega} f(x, \bar{\nu}_x) \mathrm{d}x + \int_{\bar{\Omega}} f^{\infty}(x, \bar{\nu}_x^{\infty}) \mathrm{d}\lambda, \\ &= \int_{\Omega} \left[f(x, \bar{\nu}_x) + \lambda^{a}(x) f^{\infty}(x, \bar{\nu}_x^{\infty}) \right] \mathrm{d}x + \int_{\bar{\Omega}} f^{\infty}(x, \bar{\nu}_x^{\infty}) \mathrm{d}\lambda^{s} \end{aligned}$$

where $\lambda = \lambda^{\alpha} \mathcal{L}^{d} + \lambda^{s}$ is the Radon-Nikodým decomposition of λ. We note that in the last inequality we used that the recession function of a convex integrand is itself convex. The latter constitutes a straightforward consequence of the definition of the recession function.

Now, due to the convexity of $f(x, \bullet)$, we apply Lemma 5.1 to get that

$$\int_{\Omega} \left[f(x, \bar{\nu}_x) + \lambda^{a}(x) f^{\infty}(x, \bar{\nu}_x^{\infty}) \right] \mathrm{d}x \geqslant \int_{\Omega} f(x, \bar{\nu}_x + \lambda^{a}(x) \bar{\nu}_x^{\infty}) dx,$$

and hence, combining the above, we conclude

$$\begin{aligned} \liminf_{j\to\infty} \int_{\Omega} f(x, \mathrm{d}\mu_j(x)) &\geqslant \int_{\Omega} f(x, \bar{\nu}_x + \lambda^{a}(x) \bar{\nu}_x^{\infty}) \mathrm{d}x + \int_{\bar{\Omega}} f^{\infty}(x, \bar{\nu}_x^{\infty}) \mathrm{d}\lambda^{s} \\ &\geqslant \int_{\Omega} f(x, \mu^{a}(x)) \mathrm{d}x + \int_{\Omega} f^{\infty}(x, \mathrm{d}\mu^{s}) = \int_{\Omega} f(x, \mathrm{d}\mu), \end{aligned}$$

where $\mu = \mu^{\alpha} \mathcal{L}^{d} + \mu^{s}$ is the Radon-Nikodým decomposition of μ. In the last inequality above we applied Lemma 2.4 for the sequence $(\mu_j)_j$ together with the fact that f^{∞} is non-negative, while in the last equation we used the 1-homogeneity of f^{∞}. □

We next prove two general existence results for convex integrals defined on $\mathrm{BV}_p^{\mathcal{B}}$ spaces. The first one holds for *arbitrary* operators $\mathcal{B}$.

Theorem 5.4. *Let us fix* $p \in (1, +\infty)$, $g \in L^p(\Omega, \mathbb{V})$, $\underline{\alpha} > 0$ *and* $\alpha \in \mathrm{C}(\bar{\Omega}, [\underline{\alpha}, \overline{\alpha}])$. *Then, if* $f\colon \Omega \times \mathbb{V} \to [0, +\infty)$ *satisfies the assumptions outlined in Subsection 5.1.2, the functional* I, *defined in* (5.2), *is weakly-$*$ lower semicontinous and admits a minimiser* $u \in \mathrm{BV}_p^{\mathcal{B}}(\Omega)$. *If in addition the fidelity term is strictly convex, then the minimiser is unique.*

Proof. There is no loss of generality in assuming that $\alpha \equiv 1$, and hence denote the functional $I[\cdot, 1]$ just by $I[\cdot]$. We will employ the direct method of the calculus of variations. From the non-negativity of f we have that the functional I is bounded from below, and hence, there exists a minimising sequence $(u_j)_j \subset \mathrm{BV}_p^{\mathcal{B}}(\Omega)$ such that the limit of $I[u_j]$ as $j \to +\infty$ is finite. Since Φ_g is coercive on $L^p(\Omega, \mathbb{U})$ and f satisfies the growth condition in (5.4), we have that

$$\sup_j \left(\int_\Omega |u_j(x) - g(x)|^p dx + \int_\Omega |\mathcal{B}u_j(x)| \right) \lesssim \sup_j \left(\Phi_g(u_j) + \int_\Omega f(x, \mathrm{d}\mathcal{B}u_j) \right) \leqslant \infty,$$

and so $(u_j)_j$ must be bounded in $\mathrm{BV}_p^{\mathcal{B}}(\Omega)$. Thus, on a subsequence that we do not relabel, we have $u_j \rightharpoonup u$ in $L^p(\Omega, \mathbb{U})$ and $\mathcal{B}u_j \overset{*}{\rightharpoonup} \mathcal{B}u$ in $\mathcal{M}(\Omega, \mathbb{V})$. By the weak lower semicontinuity of the fidelity term we have that

$$\Phi_g(u) \leqslant \liminf_{j \to +\infty} \Phi_g(u_j), \tag{5.1}$$

whereas by Proposition 5.1 we obtain

$$\int_\Omega f(x, \mathrm{d}\mathcal{B}u) \leqslant \liminf_{j \to \infty} \int_\Omega f(x, \mathrm{d}\mathcal{B}u_j). \tag{5.2}$$

On the whole, we deduce

$$I[u] \leqslant \liminf_{j \to +\infty} I[u_j],$$

and we conclude that $u \in \mathrm{BV}_p^{\mathcal{B}}(\Omega)$ is a minimiser of I.

Uniqueness follows easily when Φ_g is strictly convex. Indeed, let u^1, u^2 be distinct minimisers and $u^0 := (u^1 + u^2)/2$. Then, $\Phi_g(u^0) < \frac{1}{2}\Phi_g(u^1) + \frac{1}{2}\Phi_g(u^2)$, while the convexity of the second term gives

$$\int_\Omega f(x, \mathrm{d}\mathcal{B}u^0) \leqslant \frac{1}{2} \int_\Omega f(x, \mathrm{d}\mathcal{B}u^1) + \frac{1}{2} \int_\Omega f(x, \mathrm{d}\mathcal{B}u^2).$$

By adding the last two inequalities we infer $I[u^0; \alpha] < \min I$, a contradiction. ◻

Remark 5.2. Notably, in the previous theorem uniqueness holds for $\Phi_g(u) = \|u - g\|^p_{L^p(\Omega)}$. Indeed, for instance by the uniform convexity of the L^p spaces, we have for $u^0 := (u^1 + u^2)/2$

$$2 \int_\Omega |u^0 - g|^p \mathrm{d}x < \int_\Omega |u^1 - g|^p \mathrm{d}x + \int_\Omega |u^2 - g|^p \mathrm{d}x.$$

Remark 5.3. There is no immediate counterpart of Theorem 5.4 when $p = 1$, due to the fact that bounded sequences in $\mathrm{BV}^{\mathcal{B}}$ are not weakly-$*$ precompact. One possibility would be to embed $\mathrm{BV}^{\mathcal{B}}$ in the larger space of measures $\{\mu \in \mathcal{M}(\Omega, \mathbb{U}) : \mathcal{B}\mu \in \mathcal{M}(\Omega, \mathbb{V})\}$. A second option is to assume $\mathcal{B}$ to be $\mathbb{C}$-elliptic, as we do below.

The second existence result involves the smaller class of $\mathbb{C}$-elliptic operators, which was introduced in Definition 5.1. In this case, we are able to treat regularisers that also involve lower order terms, see (1.11), and we can obtain much more precise information on the minimisers. We make the unconventional convention that $\frac{d}{d-l} = +\infty$ if $l \geqslant d$, and we denote by $\mathrm{sym}^i(\mathbb{R}^d, \mathbb{U})$ the space of symmetric $\mathbb{U}$-valued i-linear maps on $\mathbb{R}^d$.

Theorem 5.5. *Let us fix $p \in [1, \frac{d}{d-l})$, $g \in L^p(\Omega, \mathbb{U})$, $\overline{\alpha} > \underline{\alpha} > 0$, $\alpha_i \in \mathrm{C}(\bar{\Omega}, [0, \overline{\alpha}])$ for $i = 1, \ldots, l-1$ and $\alpha_l \in \mathrm{C}(\bar{\Omega}, [\underline{\alpha}, \overline{\alpha}])$. Let $f_i \colon \Omega \times \mathrm{sym}^i(\mathbb{R}^d, \mathbb{U}) \to \mathbb{R}$ be Carathéodory integrands such that $f_i(x, \cdot)$ is convex with linear gowth at infinity i.e. $f_i(x, \cdot) \leqslant C(1+|\cdot|)$ for all $i = 1, \ldots, l-1$ and almost every $x \in \Omega$, and let $f_l \colon \Omega \times \mathbb{V} \to [0, +\infty)$ satisfying the assumptions outlined in Subsection 5.1.2. Then, if $\mathcal{B}$ is $\mathbb{C}$-elliptic, the functional*

$$\tilde{I}[u; \alpha] := \Phi_g(u) + \sum_{i=1}^{l-1} \int_\Omega \alpha_i(x) f_i(x, \nabla^i u(x)) \mathrm{d}x + \int_\Omega \alpha_l(x) f_l(x, \mathrm{d}\mathcal{B}u). \tag{5.3}$$

is weakly-$*$ lower semicontinuous in $\mathrm{BV}^{\mathcal{B}}$ and admits a minimiser

$$u \in \mathrm{BV}^{\mathcal{B}}(\Omega) \cap \mathrm{W}^{l-1, d/(d-1)}(\Omega, \mathbb{U}).$$

If the fidelity term is strictly convex, then the minimiser is unique.

Proof. If $l = 1$ the statement collapses to Theorem 5.4. The $\mathbb{C}$-ellipticity of $\mathcal{B}$ is still needed to make use of Theorem 5.2, which grants that the minimiser $u \in L^{d/(d-1)}(\Omega, \mathbb{U})$. We now turn to the case $l \geqslant 2$.

If $(u_j)_j \subset \mathrm{BV}^{\mathcal{B}}(\Omega)$ is a minimising sequence, as in the proof of Theorem 5.4, due to the coercivity conditions on the fidelity term Φ_g, the integrand f_l and the lower bound of the higher order weight α_l, we have that

$$\sup_j \left(\int_\Omega |u_j(x) - g(x)|^p dx + \int_\Omega \underline{\alpha} |\mathcal{B}u_j(x)| \right) \lesssim \sup_j \tilde{I}[u_j, \alpha] \leqslant \infty.$$

The above inequality guarantees that the sequence $(u_j)_j$ is bounded in $\mathrm{BV}^{\mathcal{B}}(\Omega)$, and hence also in $\mathrm{W}^{l-1, d/(d-1)}(\Omega, \mathbb{U})$ thanks to the $\mathbb{C}$-ellipticity of $\mathcal{B}$, Theorem 5.2. Let $u \in \mathrm{BV}^{\mathcal{B}}$ be a weak-$*$ limit point of $(u_j)_j$. Following the same arguments with the proof of Theorem 5.4, we have that (5.1) and (5.2) with $f = \alpha_l f_l$ hold. We now fix $1 \leqslant i \leqslant l-1$, and we look at the Young measure $\boldsymbol{\nu}$ generated by $(\nabla^i u_j)_j$, which is bounded in $L^{d/(d-1)}(\Omega)$. Thus, due to the linear growth of f_i, we have that $(\alpha_i f_i(\,\boldsymbol{\cdot}\,, \nabla^i u_j))_j$ is uniformly bounded in $L^{d/(d-1)}(\Omega)$ and hence uniformly integrable. We can thus employ Proposition 2.1 for $f = \alpha_i f_i$ to obtain

$$\begin{aligned}\liminf_{j\to\infty} \int_\Omega \alpha_i(x) f_i(x, \nabla^i u_j(x))\mathrm{d}x = \int_\Omega \alpha_i(x)\langle \nu_x, f_i(x, \boldsymbol{\cdot})\rangle \mathrm{d}x \geqslant \int_\Omega \alpha_i(x) f_i(x, \bar{\nu}_x)\mathrm{d}x \\ = \int_\Omega \alpha_i(x) f_i(x, \nabla^i u(x))\mathrm{d}x,\end{aligned}$$

where, due to the convexity of $f_i(x, \boldsymbol{\cdot})$, we used Jensen's inequality and Lemma 2.4. Combining the above, we infer that

$$\liminf_{j\to+\infty} \tilde{I}[u_j;\alpha] \geqslant \tilde{I}[u;\alpha],$$

which translates to $u \in \mathrm{BV}^{\mathcal{B}}(\Omega) \subset \mathrm{W}^{l-1,d/(d-1)}(\Omega, \mathbb{U})$ is a minimiser of $\tilde{I}[\boldsymbol{\cdot}, \alpha]$.

The uniqueness follows exactly by the same argument as in the proof of Theorem 5.4, so the conclusion is achieved. □

Remark 5.4. If $\Phi_g(u) = \|u - g\|_{L^1(\Omega)}$, uniqueness might fail in Theorem 5.5.

5.3 The bilevel training scheme in the space $\mathrm{BV}_p^{\mathcal{B}}$

We devote this section to the proof of our main theoretical result, that is, the existence of solutions to the bilevel scheme (L1)–(L2). The study of the lower level problem will be addressed by Theorem 5.4. A variant involving functionals as in Theorem 5.5 will also be presented, see Remark 5.5.

Theorem 5.6. *Let us fix $p \in (1, +\infty)$, $g \in L^p(\Omega, \mathbb{U})$, $\overline{\alpha} > \underline{\alpha} > 0$ and $\alpha \in \mathrm{C}(\bar{\Omega}, [\underline{\alpha}, \overline{\alpha}])$. Let $f\colon \Omega\times\mathbb{V} \to [0, +\infty)$ be an integrand satisfying the assumptions outlined in Subsection 5.1.2. Then, the training scheme* (L1)–(L2) *in Subsection 5.1.2 admits a solution $\alpha^* \in \mathrm{Adm}$ and it provides an associated optimally reconstructed image $u_{\alpha^*} \in \mathrm{BV}_p^{\mathcal{B}}(\Omega)$.*

Proof. Let $(\alpha_j)_j \subset \mathrm{Adm}$ be a minimising sequence for the upper level objective F i.e.,

$$\lim_j F(u_j) = \inf_{\alpha\in\mathrm{Adm}} F(u_\alpha), \tag{5.1}$$

where we abbreviated $u_j \coloneqq u_{\alpha_j} \in \mathrm{BV}_p^{\mathcal{B}}(\Omega)$ for a minimiser of (L2) associated to the weight α_j, which, together with $u_\alpha \in \mathrm{BV}_p^{\mathcal{B}}(\Omega)$ above, exists in light of Theorem 5.4. Since $\alpha_j \in \mathrm{Adm}$ which is a compact space (see the respective discussion in the subsection 5.1.2), we find $\alpha^* \in \mathrm{Adm}$ such that $\alpha_j \to \alpha^*$ uniformly in $\bar{\Omega}$. Thus, to prove our theorem, it suffices to show that

$$F(u_{\alpha^*}) \leqslant \lim_{j\to+\infty} F(u_j), \tag{5.2}$$

where $u_{\alpha^*} \in \mathrm{BV}_p^{\mathcal{B}}(\Omega)$ is a minimiser of (L2) with respect to the weight α^*.

We firstly show that (u_j) is weakly-$*$ precompact in $\mathrm{BV}_p^{\mathcal{B}}(\Omega)$. To see this, we observe that by the definition of u_j we have

$$I[u_j;\alpha_j] \leqslant I[v;\alpha_j], \quad \text{for any } v \in \mathrm{BV}_p^{\mathcal{B}}(\Omega). \tag{5.3}$$

In particular, by selecting $v = 0$ and recalling that $\|\alpha_j\|_{L^\infty} \leqslant \overline{\alpha}$, we find that $I[u_j;\alpha_j] \leqslant C$ for some $C \geqslant 0$ independent of j. Then, owing to the coercivity of I and similarly with Theorem 5.4, we infer that $(u_j)_j$ is bounded in $\mathrm{BV}_p^{\mathcal{B}}(\Omega)$. Hence, there exists $u \in \mathrm{BV}_p^{\mathcal{B}}(\Omega)$ such that, upon extraction of subsequences, $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}_p^{\mathcal{B}}(\Omega)$.

We now claim that $u = u_{\alpha^*}$, which is enough to conclude due to the lower semicontinuity of F. In other words, we need to show that

$$I[u;\alpha^*] \leqslant I[v;\alpha^*], \quad \text{for any } v \in \mathrm{BV}_p^{\mathcal{B}}(\Omega). \tag{5.4}$$

The uniform convergence of $(\alpha_j)_j$ along with (5.3) yields

$$\liminf_{j\to+\infty} I[u_j;\alpha_j] \leqslant \liminf_{j\to+\infty} I[v;\alpha_j] = I[v;\alpha^*], \quad \text{for any } v \in \mathrm{BV}_p^{\mathcal{B}}(\Omega). \tag{5.5}$$

Further, in view of the growth condition (5.4) and the fact that the sequence $(u_j)_j$ is bounded in $\mathrm{BV}_p^{\mathcal{B}}(\Omega)$, we obtain the estimate

$$\begin{aligned} |I[u_j;\alpha_j] - I[u_j;\alpha^*]| &\leqslant \int_\Omega |\alpha_j - \alpha^*| f(x, \mathrm{d}\mathcal{B}u_j) \leqslant C(1 + |\mathcal{B}u_j|(\Omega))\|\alpha_j - \alpha^*\|_{L^\infty} \\ &\leqslant C\|\alpha_j - \alpha^*\|_{L^\infty} \to 0, \end{aligned} \tag{5.6}$$

for $j \to +\infty$. Hence, by the lower semicontinuity result in Theorem 5.4, we obtain that

$$I[u;\alpha^*] \leqslant \liminf_{j\to\infty} I[u_j;\alpha^*] \overset{(5.6)}{\leqslant} \liminf_{j\to\infty} I[u_j;\alpha_j] \overset{(5.6)}{\leqslant} I[v,\alpha^*], \quad \text{for any } v \in \mathrm{BV}_p^{\mathcal{B}}(\Omega), \tag{5.7}$$

which proves our claim. Finally, due to the weak lower semicontinuity of the upper level objective F, since $u_j \rightharpoonup u \equiv u_{\alpha^*}$ in $L^p(\Omega)$, we conclude that

$$\inf_{\alpha\in\mathrm{Adm}} F(u_\alpha) \leqslant F(u_{\alpha^*}) = F(u) \leqslant \liminf_{j\to\infty} F(u_j) \overset{(5.1)}{=} \inf_{\alpha\in\mathrm{Adm}} F(u_\alpha),$$

which completes the proof. □

If in the lower level problem (L2) the functional I is replaced by $\tilde{I}$ as in Theorem 5.5, a result in the same spirit of the above holds. We only sketch it in the next remark, since it parallels closely Theorem 5.6.

Remark 5.5. Within the general framework of subsection 5.1.2, we introduce a variant of the scheme (L1)–(L2). For $l \in \mathbb{N}$, $l \geqslant 2$, we define the sets

$$\mathrm{Adm}_{\mathrm{low}} := \left\{ \alpha \in \mathrm{C}(\bar{\Omega}, [0, \overline{\alpha}]^{l-1}) : |\alpha_i(x) - \alpha_i(y)| \leqslant \omega(|x-y|) \text{ for every } i \text{ and } x, y \in \overline{\Omega} \right\},$$
$$\widetilde{\mathrm{Adm}} := \mathrm{Adm}_{\mathrm{low}} \times \mathrm{Adm},$$

where ω is the same modulus of uniform continuity as in (5.3). We consider the bilevel problem

$$\text{find } \alpha^* \in \operatorname{argmin} \left\{ F(u_\alpha) : \alpha \in \widetilde{\mathrm{Adm}} \right\} \tag{5.8}$$
$$\text{such that } u_\alpha \in \operatorname{argmin} \left\{ \tilde{I}[u;\alpha] \colon u \in \mathrm{BV}^{\mathcal{B}}(\Omega) \right\}, \tag{5.9}$$

where $\tilde{I}$ is as in (5.3). Under the assumptions of Theorem 5.5, notably $\mathbb{C}$-ellipticity for $\mathcal{B}$, we are able to prove the existence of a solution, that is, an optimal regulariser $\alpha^* \in \widetilde{\mathrm{Adm}}$ for (5.8). Let us outline the argument.

If $(\alpha^j)_j \subset \widetilde{\mathrm{Adm}}$ is a minimising sequence, as in the proof of Theorem 5.6, we may assume that $\alpha^j \to \alpha^* \in \widetilde{\mathrm{Adm}}$ uniformly. By Theorem 5.5, we can pick a sequence $(u_j)_j \subset \mathrm{BV}^{\mathcal{B}}(\Omega)$ made of minimisers for (5.9) associated with $(\alpha^j)_j$. As a consequence of the coercivity of $\tilde{I}$, $(u_j)_j$ is bounded in $\mathrm{BV}^{\mathcal{B}}(\Omega)$, and thus, owing to Theorem 5.2, also in $\mathrm{W}^{l-1,d/(d-1)}(\Omega)$. Denoting by $u \in \mathrm{BV}^{\mathcal{B}}(\Omega)$ the weak-$*$ limit (up to subsequences) of $(u_j)_j$, the remainder of the proof follows the one of Theorem 5.6, the most significant difference being the use of Theorem 5.5 instead of Theorem 5.4 to obtain the analogue of (5.7).

5.4 Numerical examples

In this section, due to the cotribution of my collaborator Kostas Papafitsoros, we are able to provide some numerical results for image reconstruction by focusing on some specific instances of the differential operators considered above. These numerical examples show the applicability and versatility of our approach, which, as we will see, is able to yield results that are comparable, and in certain cases even better, than the ones obtained by using some standard high quality regularisers, such as the Total Generalized Variation (TGV) [13] and its version with spatially varying weights [67]. Since our main target here is to evaluate the performance of the types of regularisers that we introduced, we restrict ourselves to two particular cases of image denoising. Firstly, in the class of first-order functionals, we consider a Huber-type TV regularisation, with both the regularisation parameter α and the Huber parameter γ being spatially dependent. This can be considered as a functional that incorporates a local choice between TV and Tikhonov regularisation. The second example is a spatially varying TV^2 regularisation, which is a second-order functional and has the capability to improve the reconstructions by eliminating the undesirable staircasing effect of TV [91]. Even though in theory the TV^2 regularisation is not able to preserve sharp edges, we will see that its spatially varying version produces high quality results and can even outperform both the scalar and the spatially varying versions of TGV.

5.4.1 Spatially varying Huber versions of TV and TV^2

Let $\gamma \in L^\infty(\Omega)$, with $\gamma \geq 0$, be fixed. We define the spatially varying Huber function $f_\gamma \colon \Omega \times \mathbb{R}^N \to [0, +\infty)$ as follows:

$$f_\gamma(x,z) = \begin{cases} |z| - \dfrac{1}{2}\gamma(x), & \text{if } |z| \geq \gamma(x), \\ \dfrac{1}{2\gamma(x)}|z|^2, & \text{if } |z| < \gamma(x). \end{cases} \tag{5.1}$$

Obviously, for all $z \in \mathbb{R}^N$ and for almost all $x \in \Omega$, f_γ satisfies the coercivity and growth conditions in (5.4), namely

$$|z| - \|\gamma\|_{L^\infty(\Omega)} \leq f_\gamma(x,z) \leq |z|. \tag{5.2}$$

Indeed, for $|z| \geqslant \gamma(x)$ and since $\|\gamma\|_\infty \geqslant \gamma(x)$ for a.e. $x \in \Omega$,

$$\begin{aligned} f_\gamma(x,z) &= |z| - \frac{1}{2}\gamma(x) \leqslant |z|, \\ f_\gamma(x,z) &= |z| - \frac{1}{2}\gamma(x) \geqslant |z| - \frac{1}{2}\|\gamma\|_\infty \geqslant |z| - \|\gamma\|_\infty. \end{aligned}$$

On the other hand, if $|z| < \gamma(x)$,

$$\begin{aligned} f_\gamma(x,z) &= \frac{1}{2\gamma(x)}|z|^2 = \frac{|z|}{2\gamma(x)}|z| < \frac{1}{2}|z| \leqslant |z|, \\ f_\gamma(x,z) &= \frac{1}{2\gamma(x)}|z|^2 \geqslant 0 \geqslant |z| - \|\gamma\|_\infty, \end{aligned}$$

since, in this case, $|z| \leqslant \|\gamma\|_\infty$.

Then, if $u \in \mathrm{BV}(\Omega)$, we define the ensuing convex function of the measure $\mathrm{D}u$ with the alternative notations

$$\mathrm{TV}_\gamma(u) := |f_\gamma(\mathrm{D}u)|(\Omega) := \int_\Omega f_\gamma(x, d\mathrm{D}u).$$

A straightforward check shows that the recession function of f_γ (cf. (5.5)) is

$$f_\gamma^\infty(x,z) := \lim_{(x',z',t)\to(x,z,+\infty)} \frac{f_\gamma(x',tz')}{t} = \lim_{(x',z',t)\to(x,z,+\infty)} \frac{t|z'| - \frac{1}{2}\gamma(x')}{t} = |z|.$$

Thus, all the assumptions of Theorem 5.6 are trivially satisfied. Consequently, TV_γ is indeed well-defined as

$$\mathrm{TV}_\gamma(u) = \int_\Omega f_\gamma(x, \nabla u)\mathrm{d}x + \int_\Omega d|\mathrm{D}^s u|,$$

and for $\alpha \in C(\overline{\Omega})$ with $\overline{\alpha} \geqslant \alpha(x) \geq \underline{\alpha} > 0$ we can define its spatially varying version

$$\mathrm{TV}_{\alpha,\gamma}(u) = \int_\Omega \alpha f_\gamma(x, \nabla u)\mathrm{d}x + \int_\Omega \alpha\, d|D^s u|.$$

Similarly, for a function $u \in \mathrm{BV}^2(\Omega) := \{u \in W^{1,1}(\Omega) : D^2 u \in \mathcal{M}(\Omega, \mathcal{S}^{N\times N})\}$ and $\alpha \in \mathrm{C}(\overline{\Omega})$ with $\overline{\alpha} \geqslant \alpha(x) \geqslant \underline{\alpha} > 0$, we define the spatially varying Huber TV^2 functional as

$$\mathrm{TV}^2_{\alpha,\gamma}(u) = \int_\Omega \alpha f_\gamma(x, \nabla^2 u)\,\mathrm{d}x + \int_\Omega \alpha\,\mathrm{d}|(\mathrm{D}^2 u)^s|,$$

where f_γ now is a function defined on $\Omega \times \mathbb{R}^{N\times N}$ defined by the natural analogue of (5.1).

Our examples concern the following lower level image denoising problems:

$$I^1_\gamma[u;\alpha] := \int_\Omega |u-g|^2 \mathrm{d}x + \mathrm{TV}_{\alpha,\gamma}(u), \tag{5.3}$$

$$I^2_\gamma[u;\alpha] := \int_\Omega |u-g|^2 \mathrm{d}x + \mathrm{TV}^2_{\alpha,\gamma}(u). \tag{5.4}$$

5.4.2 The bilevel problems

The family of bilevel problems for the automatic computation of the spatial regularisation parameter α associated with the functional I^i_γ for $i = 1, 2$ is:

$$\text{find } \alpha^* \in \operatorname{argmin}\{F(u_\alpha) :\ \alpha \in \mathrm{Adm}\} \tag{5.5}$$

$$\text{such that } u_\alpha = \operatorname{argmin}\left\{I^i_\gamma[u;\alpha] \colon u \in \mathrm{BV}^i(\Omega)\right\}. \tag{5.6}$$

In view of (5.2), the lower level problems (5.6) are well-defined, while from Theorem 5.6 we know that the overall schemes (5.5)–(5.6) admit a solution for the two alternative upper level objectives F considered next. As we discussed in the introduction and repeat here for the sake of reading flow, we take into account two alternatives for the upper level objective functional F:

$$F_{\mathrm{PSNR}}(u) = \int_\Omega |u - u_{\mathrm{gt}}|^2 \mathrm{d}x, \tag{5.7}$$

$$F_{\mathrm{stat}}(u) = \frac{1}{2}\int_\Omega \max(Ru - \overline{\sigma}^2, 0)^2 \mathrm{d}x + \frac{1}{2}\int_\Omega \min(Ru - \underline{\sigma}^2, 0)^2 \mathrm{d}x, \tag{5.8}$$

$$\text{where } Ru(x) := \int_\Omega w(x,y)(u-g)^2(y)\mathrm{d}y \quad \text{for } w \in L^\infty(\Omega\times\Omega),\ \int_\Omega\int_\Omega w(x,y)\mathrm{d}x\mathrm{d}y = 1.$$

The first cost functional corresponds to a maximisation of the PSNR of the reconstruction and requires the knowledge of the ground truth u_{gt} [19, 44, 43, 42, 84], while the second enforces the localised residuals Ru to belong in a certain tight corridor $[\underline{\sigma}^2, \overline{\sigma}^2] := [\sigma^2 - \epsilon, \sigma^2 + \epsilon]$, σ^2 being the variance of the noise η, which is assumed here to be Gaussian, see also [68, 69, 66, 67]. The latter option has the advantage of being ground truth free, but knowledge or a good estimate for the noise variance σ^2 is needed. For the discrete version of the averaging filter w in the definition of the localised residuals (5.8) we use a filter of size $n_w \times n_w$, with entries of equal value that sum to one.

Since a numerical projection to the admissible set Adm is not practical, here we also follow [68, 69, 66, 67] and add instead a small H^1 term of the weight function α in the upper level objective, together with a supplementary box constraint $\mathcal{C} := \{\alpha \in H^1(\Omega) : \underline{\alpha} \le \alpha \le \overline{\alpha}\}$ for some $\underline{\alpha}, \overline{\alpha} \in \mathbb{R}$ with $0 < \underline{\alpha} < \overline{\alpha}$. On the whole, we will use the following upper level objectives:

$$\begin{aligned}
\mathcal{F}_{\mathrm{PSNR}}(\alpha, u) &= \int_\Omega |u - u_{\mathrm{gt}}|^2 \mathrm{d}x + \frac{\lambda}{2}\|\alpha\|^2_{H^1(\Omega)},\\
\mathcal{F}_{\mathrm{stat}}(\alpha, u) &= \frac{1}{2}\int_\Omega \max(Ru - \overline{\sigma}^2, 0)^2 \mathrm{d}x + \frac{1}{2}\int_\Omega \min(Ru - \underline{\sigma}^2, 0)^2 \mathrm{d}x + \frac{\lambda}{2}\|\alpha\|^2_{H^1(\Omega)},
\end{aligned}$$

for some small $\lambda > 0$. We will denote by $\hat{\mathcal{F}}$ the corresponding reduced objective functionals, that is $\hat{\mathcal{F}}_{\mathrm{PSNR/stat}}(\alpha) := \mathcal{F}_{\mathrm{PSNR/stat}}(\alpha, u_\alpha)$. That leads us to the bilevel minimisation problems that we tackle numerically: for $i = 1, 2$

$$\text{find } \alpha^* \in \underset{\alpha}{\operatorname{argmin}}\, \mathcal{F}_{\mathrm{PSNR/stat}}(\alpha, u_\alpha) \tag{5.9}$$

$$\text{such that } \begin{cases} u_\alpha = \underset{u}{\operatorname{argmin}}\, I^i_\gamma[u; \alpha], \\ \alpha \in \mathcal{C}. \end{cases} \tag{5.10}$$

Note that in this setting it is not guaranteed that $\alpha \in C(\overline{\Omega})$, since $H^1(\Omega)$ does not embed in that space for dimensions higher than 1. However, one can take advantage of a regularity result of the H^1-projection onto $\mathcal{C}$, denoted by $P_\mathcal{C}$ see [69, Corollary 2.3]. This projection is applied to every iteration of the projected gradient algorithm, which is to be used for the numerical solution of (5.9)–(5.10) and is described in [89]. In that case, it is ensured that the computed weight α^l at the l-th projected gradient iteration belongs to $H^2(\Omega)$, which for $d = 2$ embeds compactly into any Hölder space $C^\beta(\overline{\Omega})$, $\beta \in (0, 1)$.

5.4.3 Strategy for fixing γ

Since in our set up the function γ is not part of the minimising variables, it has to be fixed from the start. Our rationale for fixing γ is that we would like to regularise high detailed areas with a Tikhonov term with a spatially varying weight $\frac{1}{2}\int_\Omega \tilde{\alpha}|\nabla u|^2 \mathrm{d}x$, with $\tilde{\alpha}$ having as low regularity as possible, e.g. $L^\infty(\Omega)$, in order to increase flexibility in the regularisation. In the other areas we would like to regularise using a TV or TV^2 term with a spatially varying weight α. This will happen if γ is large in such detailed areas in order to allow for the second case in (5.1) and small otherwise. We thus adopt the following strategy: We first solve an auxiliary bilevel problem with a weighted Tikhonov regulariser using the upper level objective $\mathcal{F}_{\text{stat}}$. The output is a spatially varying $\tilde{\alpha}$ that essentially acts as an edge detector, since it is small on the edges and on the detailed areas of the image. We then invert this weight and set

$$\gamma = s\frac{1}{\tilde{\alpha}} \tag{5.11}$$

for some constant $s > 0$. By choosing the function γ as in (5.11), we have that when $\tilde{\alpha}$ is small (fine scale details), γ will be large and thus the second case in (5.1) will be selected with a weight $\frac{1}{2\gamma(x)} = \frac{s}{2}\tilde{\alpha}$ in front of the term $|\nabla u|^2$. On the other hand, when $\tilde{\alpha}$ is large, then γ will be small and thus a TV or TV^2 term will be preferred, i.e., first case in (5.1). In the third images of the top rows of Figures 5.1 and 5.3, we see how the resulting γ function looks like for the example images. Details for the computation of $\tilde{\alpha}$ via the auxiliary bilevel Tikhonov problem are discussed in [89].

5.4.4 Numerical results

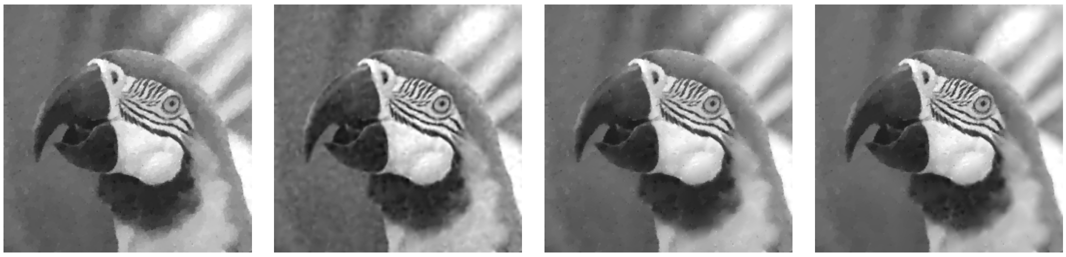
Our main tools for the quality of reconstructions are the peak signal-to-noise ratio PSNR and the structural similarity index SSIM. Those two metrics has been widely used in the literature to assess the quality of the reconstructed image, and more precisely, even though PSNR is a measurement tool that is more popular and more widely used than SSIM, the later (SSIM) is designed based on three factors i.e. luminance, contrast, and structure to better suit the workings of the human visual system. Note that a perfect reconstruction has SSIM$=1$.

In Figure 5.1 we report our numerical results on the *Parrot* image, see also Figure 5.2 for zoom-in details. Here the spatially varying regularisation weights α are produced with the ground truth-free bilevel approach, i.e., using $\mathcal{F}_{\text{stat}}$ as an upper lever objective. Among the regularisers with scalar parameters, second row, first three images, the best reconstruction both in terms of PSNR and SSIM is achieved by the scalar TGV. The bilevel Huber TV reconstruction with spatially varying α and scalar γ is able to better preserve the details around the eye of the parrot, third row first image. When we use the spatially varying γ, the details in that area become even more pronounced, compare the first two images in the third row of Figure 5.2. This is also accompanied with a slight increase of the SSIM index but also with a decrease in PSNR. Observe that the weights

Ground truth — PSNR=∞, SSIM=1.000

Gaussian noise, $\sigma^2 = 0.01$ — PSNR=20.04, SSIM=0.2773

Spatially varying γ



scalar Huber TV — PSNR=29.25, SSIM=0.8354

scalar Huber TV^2 — PSNR=29.28, SSIM=0.8305

scalar TGV — PSNR=29.50, SSIM=0.8509

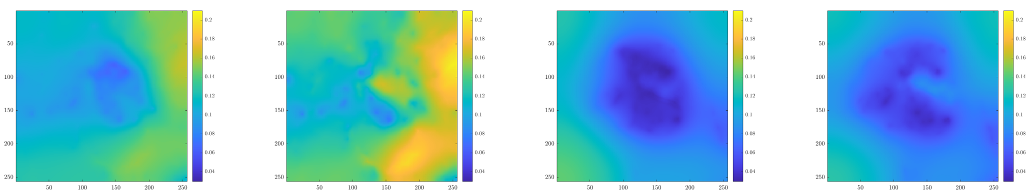
Bilevel TGV with spatially varying α and β — PSNR=**29.84**, SSIM=0.8606

Bilevel Huber TV with spatially varying α and scalar γ — PSNR=29.20, SSIM=0.8549

Bilevel Huber TV with spatially varying α and spatially varying γ — PSNR=28.92, SSIM=0.8571

Bilevel Huber TV^2 with spatially varying α and scalar γ — PSNR=29.81, SSIM=**0.8705**

Bilevel Huber TV^2 with spatially varying α and spatially varying γ — PSNR=**29.84**, SSIM=0.8700



Weight α of Huber TV

Weight α of Huber TV

Weight α of Huber TV^2

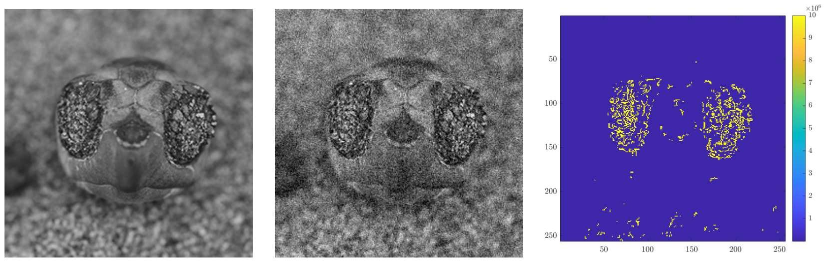
Weight α of Huber TV^2

Figure 5.1: *Parrot* image: Huber TV and TV^2 denoising with spatially varying Huber parameter γ and regularisation parameter α. The weights α are produced with the ground truth-free bilevel approach using $\mathcal{F}_{\text{stat}}$. The highest PSNR and SSIM values are highlighted in bold font.

Figure 5.2: Details of images shown in Figure 5.1

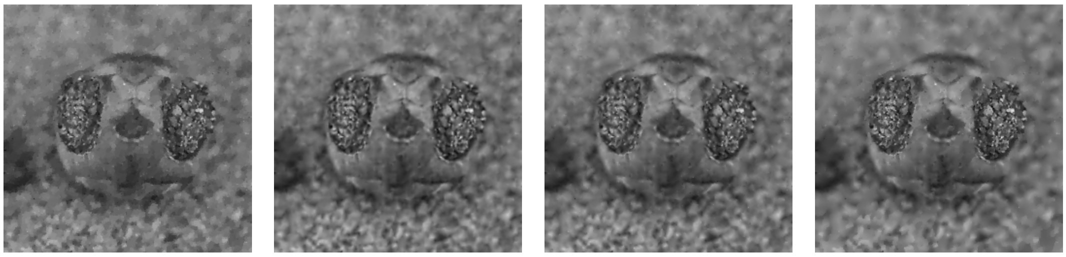
α that are computed in these two cases are quite different, see first two images of the last row of Figure 5.1. The bilevel Huber TV^2 approach with spatially varying α, produces similar reconstructions for both the scalar (slighly higher SSIM) and the spatially varying γ case (slightly higher PSNR). These reconstructions are of very good quality and even outperform the spatially varying TGV in terms of SSIM, having also the same PSNR. This is due to the fact that the combination of the statistics-based upper level objective and the second order TV is forcing the weight α to drop significantly in the detailed areas of the image, see the last two images of the last row of Figure 5.1. It is characteristic that while the PSNR of scalar TV^2 is only 0.03 dB higher than the one of scalar TV, the PSNR of bilevel Huber TV^2 with spatially varying α and scalar γ is 0.61 dB higher compared to the corresponding Huber TV result.

The superiority of the bilevel Huber TV^2 with spatially varying weights, is even more



Ground truth | Gaussian noise, $\sigma^2 = 0.01$ | Spatially varying γ

PSNR=∞, SSIM=1.000 | PSNR=20.00, SSIM=0.3349

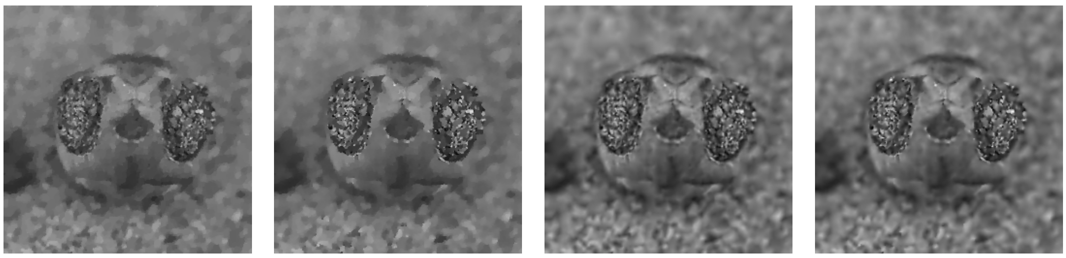


scalar Huber TV — PSNR=27.75, SSIM=0.7701

scalar Huber TV^2 — PSNR=28.22, SSIM=0.8142

scalar TGV — PSNR=28.20, SSIM=0.8132

Bilevel TGV with spatially varying α and β — PSNR=28.33, SSIM=0.8145

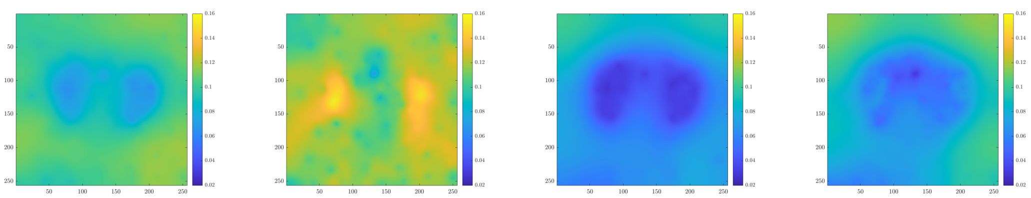


Bilevel Huber TV with spatially varying α and scalar γ — PSNR=27.50, SSIM=0.7702

Bilevel Huber TV with spatially varying α and spatially varying γ — PSNR=27.15, SSIM=0.7688

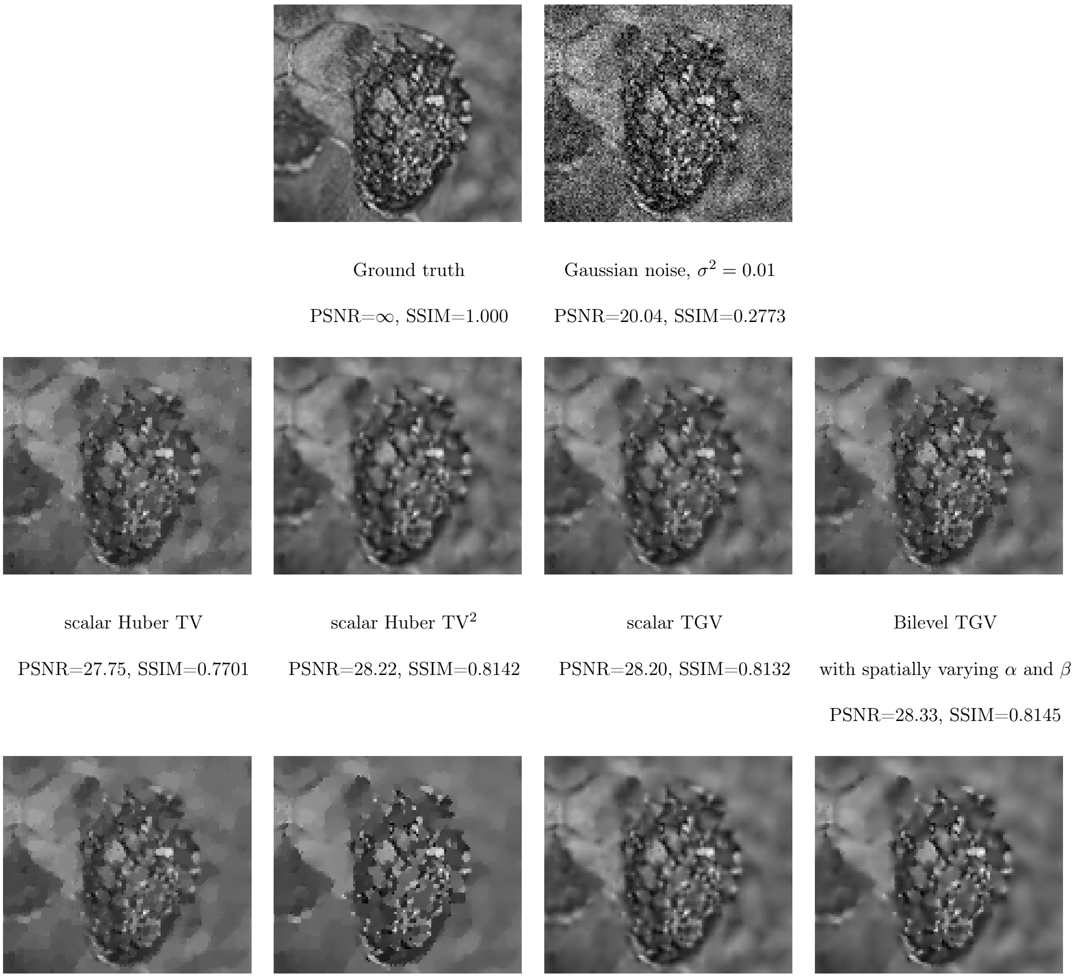
Bilevel Huber TV^2 with spatially varying α and scalar γ — PSNR=**28.66**, SSIM=**0.8367**

Bilevel Huber TV^2 with spatially varying α and spatially varying γ — PSNR=28.44, SSIM=0.8285



Weight α of Huber TV | Weight α of Huber TV | Weight α of Huber TV^2 | Weight α of Huber TV^2

Figure 5.3: *Hatchling* image: Huber TV and TV^2 denoising with spatially varying Huber parameter γ and regularisation parameter α. The weights α are produced with the ground truth-free bilevel approach using $\mathcal{F}_{\text{stat}}$. The highest PSNR and SSIM values are highlighted in bold font.



Ground truth
PSNR=∞, SSIM=1.000

Gaussian noise, $\sigma^2 = 0.01$
PSNR=20.04, SSIM=0.2773

scalar Huber TV
PSNR=27.75, SSIM=0.7701

scalar Huber TV^2
PSNR=28.22, SSIM=0.8142

scalar TGV
PSNR=28.20, SSIM=0.8132

Bilevel TGV
with spatially varying α and β
PSNR=28.33, SSIM=0.8145

Bilevel Huber TV
with spatially varying α
and scalar γ
PSNR=27.50, SSIM=0.7702

Bilevel Huber TV
with spatially varying α
and spatially varying γ
PSNR=27.15, SSIM=0.7688

Bilevel Huber TV^2
with spatially varying α
and scalar γ
PSNR=**28.66**, SSIM=**0.8367**

Bilevel Huber TV^2
with spatially varying α
and spatially varying γ
PSNR=28.44, SSIM=0.8285

Figure 5.4: Details of images shown in Figure 5.3

evident in the second image example *Hatchling*, Figures 5.3 and 5.4. Here the reconstruction is more challenging due to the oscillatory nature of the ground truth image. The bilevel Huber TV^2 with scalar γ gives by far the best result with respect to both PSNR and SSIM. Again, the automatically computed regularisation weights α have much lower values in bilevel TV^2 than in bilevel TV, compare the first two versus the last two figures of the last row of Figure 5.3. In this example, the spatially varying γ leads to a reduction of PSNR and SSIM in all cases, but nevertheless also to more highlighted details in the eye area, see second and fourth images of the last row of Figure 5.4.

In order to verify further the regularisation capabilities of these regularisers, we make another series of experiments with these two example images, using the ground truth-based upper level objective $\mathcal{F}_{\text{PSNR}}$, see Figure 5.5. In both images, the highest PSNR and SSIM

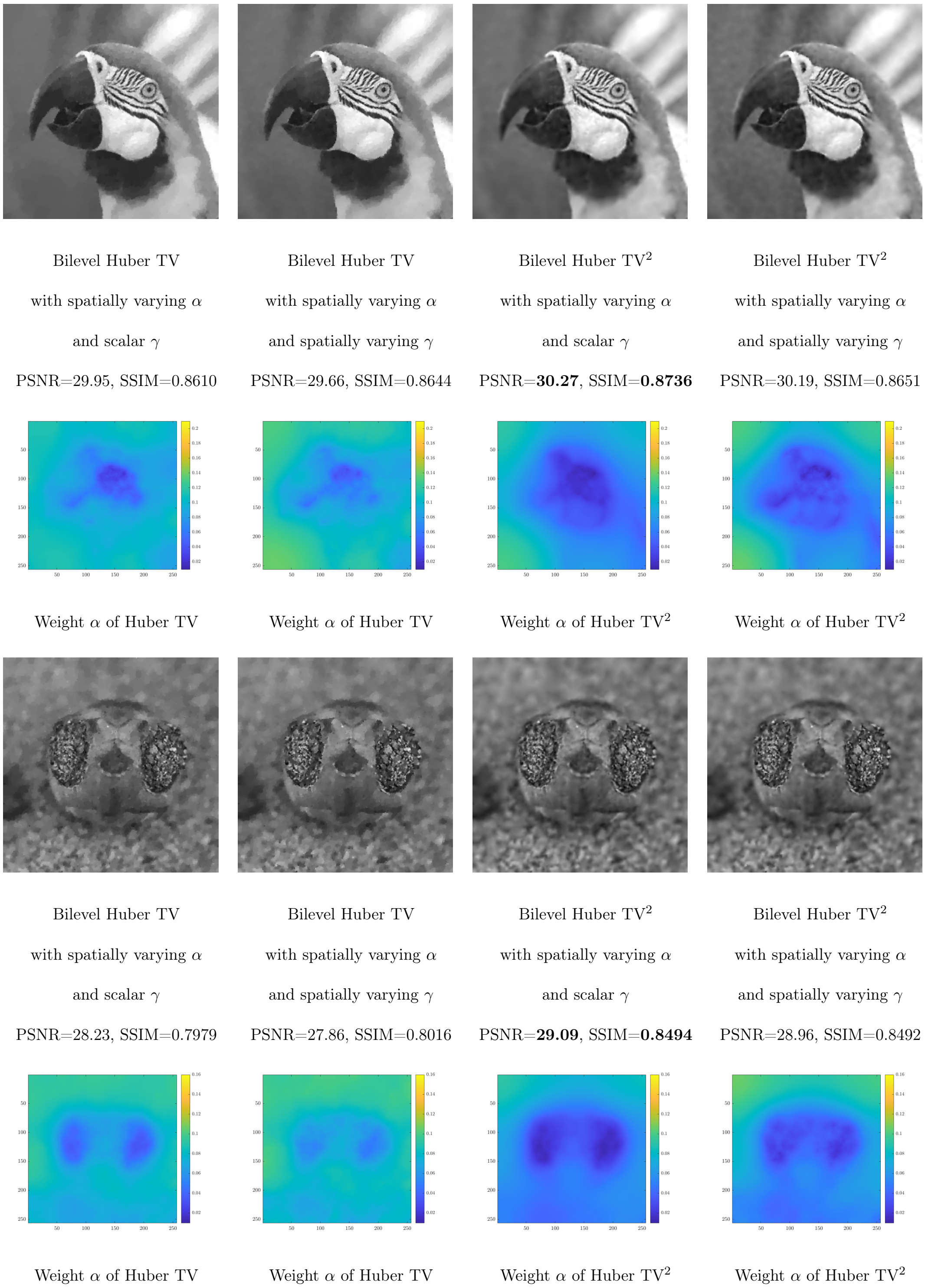
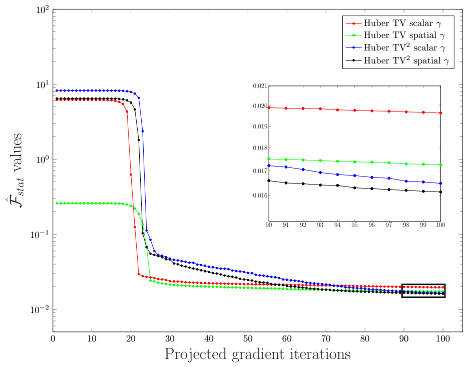
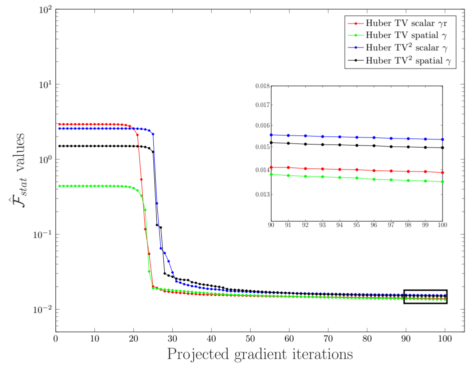


Figure 5.5: Huber TV and TV^2 denoising with spatially varying Huber parameter γ and regularisation parameter α. The weights α are produced with the ground truth-based bilevel approach using $\mathcal{F}_{\text{PSNR}}$. The highest PSNR and SSIM values are highlighted in bold font.

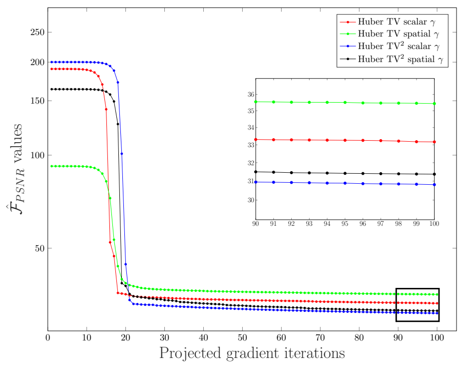
is achieved by the bilevel TV^2 with spatially varying α and scalar Huber parameter γ, third images of first and third row, with the corresponding regularisation weight α having



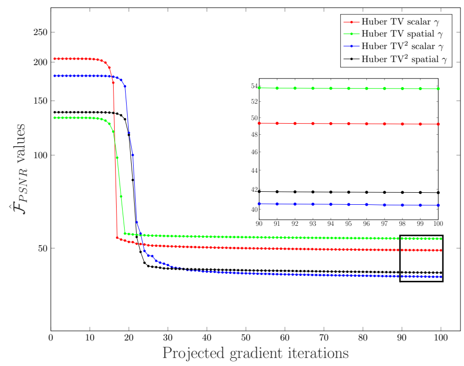
$\hat{\mathcal{F}}_{\text{stat}}$ for *parrot*



$\hat{\mathcal{F}}_{\text{stat}}$ for *hatchling*



$\hat{\mathcal{F}}_{\text{PSNR}}$ for *parrot*



$\hat{\mathcal{F}}_{\text{PSNR}}$ for *hatchling*

Figure 5.6: Values of the reduced objective $\hat{\mathcal{F}}(u^k)$ along the projected gradient iterations. The inner boxes show zoom-in plots of the last 10 iterations.

again smaller values compared to the TV one. Nevertheless, we observe that the spatially varying γ results in higher SSIM in the Huber TV examples in both images, again with more pronounced features around the eye.

Finally in Figure 5.6, we have plotted the values of the reduced objective $\hat{\mathcal{F}}(u^k)$ along the projected gradient iterations, for all bilevel Huber TV and TV^2 examples. The top row shows these plots for the reduced statistics-based upper level objective $\hat{\mathcal{F}}_{\text{stat}}$. We observe that in both images, the introduction of the spatially varying γ in both Huber TV and Huber TV^2 functionals, helps towards a further reduction of this objective, compare red versus green and blue versus black plots. We observed already that in some cases this is accompanied with a larger SSIM index and more pronounced details in the images, but in most cases the PSNR in decreased. This is in accordance with the plots of the second row, where we see that the reduced PSNR-maximising upper level objective $\hat{\mathcal{F}}_{\text{PSNR}}$ is not further decreased by the introduction of the spatially varying γ, compare again the red versus green and blue versus black plots.

We conclude that the bilevel Huber TV^2 is able to produce remarkably good results.

This is perhaps even surprising as the use of its scalar version is not that popular due to its inability to preserve sharp edges. We showed that the use of a spatially varying Huber parameter γ can result in improved results both quantitatively and qualitatively, thus justifying our rigorous analytical study on spatially inhomogeneous integrands acting on TV-type regularisers. We also stress that by no means our strategy for setting γ is necessarily the optimal one. In fact, future work will involve setting up a bilevel framework where also this parameter is included in the upper level minimisation variables along with the parameter α, adding further flexibility to the regularisation process.

Appendix A. Proof of Lemma 3.3, pg. 34

Proof. We define the associated Hessian L_f as follows

$$L_f(\lambda_1,\lambda_2)[(\xi_1,\xi_2),(\xi_1,\xi_2)] := f_{FF}(\lambda_1,\lambda_2)\,\xi_1 : \xi_1 + f_{\eta\eta}(\lambda_1,\lambda_2)\,\xi_2\,\xi_2 + 2f_{F\eta}(\lambda_1,\lambda_2)\,\xi_1\,\xi_2.$$

In the sequel we write simply L instead of L_f.

For (a), note that if $|\xi_1|+|\xi_2|+|z_1|+|z_2|\leqslant 1$, it holds that

$$\begin{aligned}
&f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2) - f(\lambda_1+z_1,\lambda_2+z_2|\lambda_1,\lambda_2)\\
&= \int_0^1 (1-t)\Big(L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\,(\xi_1,\xi_2),(\xi_1,\xi_2)\,]\\
&\qquad\qquad - L(\lambda_1+tz_1,\lambda_2+tz_2)[\,(z_1,z_2),(z_1,z_2)\,]\Big)dt\\
&= \int_0^1 (1-t)\Big(L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\,(\xi_1,\xi_2),(\xi_1,\xi_2)\,]\\
&\qquad\qquad - L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\,(\xi_1,\xi_2),(z_1,z_2)\,]\Big)dt\\
&+ \int_0^1 (1-t)\Big(L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\,(\xi_1,\xi_2),(z_1,z_2)\,]\\
&\qquad\qquad - L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\,(z_1,z_2),(z_1,z_2)\Big)dt\\
&+ \int_0^1 (1-t)\Big(L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\,(z_1,z_2),(z_1,z_2)\,]\\
&\qquad\qquad - L(\lambda_1+tz_1,\lambda_2+tz_2)[\,(z_1,z_2),(z_1,z_2)\,]\Big)dt\\
&=: \int_0^1 (1-t)(I_1+I_2+I_3)dt.
\end{aligned}$$

Concerning the term I we observe that

$$\begin{aligned}
I_1 &\leqslant |I_1| \leqslant |f_{FF}(\lambda_1+t\xi_1,\lambda_2+t\xi_2)||\xi_1||\xi_1-z_1|\\
&+|f_{\eta\eta}(\lambda_1+t\xi_1,\lambda_2+t\xi_2)|\xi_2||\xi_2-z_2| + 2|f_{F\eta}(\lambda_1+t\xi_1,\lambda_2+t\xi_2)||\xi_1\xi_2-z_1z_2|,
\end{aligned}$$

and since $\xi_1\xi_2 - z_1z_2 = \xi_1\xi_2 - z_1\xi_2 + z_1\xi_2 - z_1z_2$ we infer that

$$\begin{aligned}
I_1 &\leqslant C\big(|\xi_1||\xi_1-z_1| + |\xi_2||\xi_1-z_1| + |z_1||\xi_2-z_2| + |\xi_2||\xi_2-z_2|\big)\\
&= C\big(|\xi_1|+|\xi_2|\big)|\xi_1-z_1| + C\big(|\xi_1|+|\xi_2|\big)|\xi_2-z_2|.
\end{aligned}$$

Similarly for I_2 we find that

$$\begin{aligned}
I_2 &\leqslant C\big(|\xi_1||\xi_1-z_1| + |\xi_2||\xi_1-z_1| + |z_1||\xi_2-z_2| + |\xi_2||\xi_2-z_2|\big)\\
&= C\big(|z_1|+|z_2|\big)|\xi_1-z_1| + C\big(|\xi_1|+|z_2|\big)|\xi_2-z_2|.
\end{aligned}$$

Considering the third term since for $|z_1|+|z_2|\leqslant 1$ we have that $\big(|z_1|+|z_2|\big)^2\leqslant |z_1|+|z_2|$ it holds that

$$I_3\leqslant |I_3|\leqslant C\big(|\xi_1-z_1|+|\xi_2-z_2|\big)\big(|z_1|+|z_2|\big)^2\leqslant C\big(|\xi_1-z_1|+|\xi_2-z_2|\big)\big(|z_1|+|z_2|\big).$$

Combining the above we deduce that

$$\begin{aligned}&|f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)-f(\lambda_1+z_1,\lambda_2+z_2|\lambda_1,\lambda_2)|\\&\lesssim\big(|\xi_1|+|\xi_2|+|z_1|+|z_2|\big)|\xi_1-z_1|+\big(|\xi_1|+|\xi_2|+|z_1|+|z_2|\big)|\xi_2-z_2|.\end{aligned}$$

Now, for $|\xi_1|+|\xi_2|+|z_1|+|z_2|>1$

$$\begin{aligned}&|f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)-f(\lambda_1+z_1,\lambda_2+z_2|\lambda_1,\lambda_2)|\\&=|f(\lambda_1+\xi_1,\lambda_2+\xi_2)-f(\lambda_1+z_1,\lambda_2+z_2)+f_F(\lambda_1,\lambda_2)(\xi_1-z_1)+f_\eta(\lambda_1,\lambda_2)(\xi_2-z_2)|\\&\leqslant|f(\lambda_1+\xi_1,\lambda_2+\xi_2)-f(\lambda_1+z_1,\lambda_2+\xi_2)|+|f(\lambda_1+z_1,\lambda_2+\xi_2)-f(\lambda_1+z_1,\lambda_2+z_2)|\\&\quad+C|\xi_1-z_1|+C|\xi_2-z_2|=:J_1+J_2+J_3\end{aligned}$$

So, from the growth assumptions on f and on its partial derivatives we infer that

$$\begin{aligned}J_1&\leqslant\int_0^1|f_F\big(\lambda_1+z_1+t(\xi_1-z_1),\lambda_2+\xi_2\big)|dt\cdot|\xi_1-z_1|\\&\lesssim(1+|z_1|^{p-1}+|\xi_1|^{p-1}+|\xi_2|^{q\frac{p-1}{p}})|\xi_1-z_1|\\&\lesssim(|\xi_1|+|\xi_2|+|z_1|+|z_2|+|z_1|^{p-1}+|\xi_1|^{p-1}+|\xi_2|^{q\frac{p-1}{p}})|\xi_1-z_1|.\end{aligned}$$

Similarly for the second term

$$J_2\lesssim\big(|\xi_1|+|\xi_2|+|z_1|+|z_2|+|\xi_2|^{q-1}+|z_2|^{q-1}+|z_1|^{p\frac{q-1}{q}}\big)|\xi_2-z_2|.$$

For J_3, since $|\xi_1|+|\xi_2|+|z_1|+|z_2|>1$ we have that

$$J_3\lesssim\big(|\xi_1|+|\xi_2|+|z_1|+|z_2|\big)\,|\xi_1-z_1|+\big(|\xi_1|+|\xi_2|+|z_1|+|z_2|\big)\,|\xi_2-z_2|,$$

and together with the first two terms we deduce that

$$\begin{aligned}&|f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)-f(\lambda_1+z_1,\lambda_2+z_2|\lambda_1,\lambda_2)|\\&\leqslant C(|\xi_1|+|\xi_2|+|z_1|+|z_2|+|\xi_1|^{p-1}+|z_1|^{p-1}+|\xi_2|^{q\frac{p-1}{p}})|\xi_1-z_1|\\&\quad+C(|\xi_1|+|\xi_2|+|z_1|+|z_2|+|\xi_2|^{q-1}+|z_2|^{q-1}+|z_1|^{p\frac{q-1}{q}})|\xi_2-z_2|.\end{aligned}$$

For the second part we just set $(z_1,z_2)=(0,0)$ in the above inequality and apply Young's inequality to conclude the proof of (a).

Concerning (b) we see that for $|\xi_1|+|\xi_2|\leqslant 1$ and $\xi=(\xi_1,\xi_2)$, it holds

$$
\begin{aligned}
&|f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)-f(\mu_1+\xi_1,\mu_2+\xi_2|\mu_1,\mu_2)|\\
&\leqslant \int_0^1 \Big| L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\xi,\xi]-L(\mu_1+t\xi_1,\mu_2+t\xi_2)[\xi,\xi]\Big| dt\\
&\leqslant C\big(|\lambda_1-\mu_1|+|\lambda_2-\mu_2|\big)\big(|\xi_1|^2+|\xi_2|^2\big).
\end{aligned}
$$

When $|\xi_1|+|\xi_2|>1$, from the growth of $Df=(f_F,f_\eta)$, we have that

$$
\begin{aligned}
&|f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)-f(\mu_1+\xi_1,\mu_2+\xi_2|\mu_1,\mu_2)|\\
&\leqslant \big|f(\lambda_1+\xi_1,\lambda_2+\xi_2)-f(\mu_1+\xi_1,\mu_2+\xi_2)\big|+\big|f(\lambda_1,\lambda_2)-f(\mu_1,\mu_2)\big|\\
&+\big|Df(\lambda_1,\lambda_2)-Df(\mu_1,\mu_2)\big|\cdot|\xi|\\
&\leqslant \big|f(\lambda_1+\xi_1,\lambda_2+\xi_2)-f(\mu_1+\xi_1,\lambda_2+\xi_2)\big|\\
&+\big|f(\mu_1+\xi_1,\lambda_2+\xi_2)-f(\mu_1+\xi_1,\mu_2+\xi_2)\big|+\big(|\xi_1|+|\xi_2|\big)\big(|\lambda_1-\mu_1|+|\lambda_2-\mu_2|\big)\\
&\lesssim \int_0^1 |f_F\big(\xi_1+\mu_1+t(\lambda_1-\mu_1),\lambda_2+\xi_2\big)|dt\cdot|\lambda_1-\mu_1|\\
&+\int_0^1 |f_\eta\big(\mu_1+\xi_1,\xi_2+\mu_2+t(\lambda_2-\mu_2)\big)|dt\cdot|\lambda_2-\mu_2|\\
&+\big(|\xi_1|+|\xi_2|\big)\big(|\lambda_1-\mu_1|+|\lambda_2-\mu_2|\big)\\
&\lesssim \Big(1+|\xi_1|^{p-1}+|\xi_2|^{q\frac{p-1}{p}}\Big)|\lambda_1-\mu_1|+\Big(1+|\xi_1|^{p\frac{q-1}{q}}+|\xi_2|^{q-1}\Big)|\lambda_2-\mu_2|\\
&+\big(|\xi_1|+|\xi_2|\big)\big(|\lambda_1-\mu_1|+|\lambda_2-\mu_2|\big).
\end{aligned}
$$

But since $|\xi_1|+|\xi_2|>1$ we have that $|\xi_1|^{p-1}+|\xi_2|^{q\frac{p-1}{p}}$, $|\xi_1|^{p\frac{q-1}{q}}+|\xi_2|^{q-1}\leqslant 2+|\xi_1|^p+|\xi_2|^q\lesssim |\xi_1|+|\xi_2|+|\xi_1|^p+|\xi_2|^q\lesssim |\xi_1|^2+|\xi_2|^2+|\xi_1|^p+|\xi_2|^q$. Combining both cases we infer that

$$
\begin{aligned}
&|f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)-f(\mu_1+\xi_1,\mu_2+\xi_2|\mu_1,\mu_2)|\\
&\qquad\qquad\leqslant C\big(|\lambda_1-\mu_1|+|\lambda_2-\mu_2|\big)\big(|\xi_1|^2+|\xi_2|^2+|\xi_1|^p+|\xi_2|^q\big),
\end{aligned}
$$

and by choosing $R<\delta/C$ we conclude the proof of part (b).

Regarding the proof of (c), again for $|\xi_1|+|\xi_2|\leqslant 1$, since $p,q\geqslant 2$

$$
\begin{aligned}
f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)&=\int_0^1(1-t)L(\lambda_1+t\xi_1,\lambda_2+t\xi_2)[\xi,\xi]dt\geqslant -\tilde d_2|\xi|^2\\
&\geqslant \tilde d_1\big[(|\xi_1|^p-|\xi_1|^2)+(|\xi_2|^q-|\xi_2|^2)\big]-\tilde d_2(|\xi_1|^2+|\xi_2|^2).
\end{aligned}
$$

For the case $|\xi_1|+|\xi_2|>1$, from the coercivity of f we have that

$$
\begin{aligned}
f(\lambda_1+\xi_1,\lambda_2+\xi_2|\lambda_1,\lambda_2)&=f(\lambda_1+\xi_1,\lambda_2+\xi_2)-f(\lambda_1,\lambda_2)-f_F(\lambda_1,\lambda_2)\xi_1-f_\eta(\lambda_1,\lambda_2)\xi_2\\
&\geqslant \tilde d_3\big(-1+|\xi_1|^p+|\xi_2|^q\big)-\tilde d_4\big(|\xi_1|+|\xi_2|\big)\\
&\geqslant \tilde d_3\big(|\xi_1|^p+|\xi_2|^q\big)-\tilde d_4\big(|\xi_1|^2+|\xi_2|^2\big).
\end{aligned}
$$

So, combining the two cases, we may choose $d_1, d_2 > 0$ to conclude the proof of the lemma. $\square$

Bibliography

[1] M.S. Agranovich and M.I. Vishik. Elliptic problems with a parameter and parabolic problems of general type. *Uspekhi Matematicheskikh Nauk*, 19(3):53–161, 1964.

[2] J. J. Alibert and G. Bouchitté. Non-uniform integrability and generalized Young measures. *J. Convex Anal.*, 4(1):129–147, 1997.

[3] J.J. Alibert and G. Bouchitté. Non-uniform integrability and generalized young measure. *Journal of Convex Analysis*, 4:129–148, 1997.

[4] M. Amar, V. De Cicco, and N. Fusco. Lower semicontinuity and relaxation results in bv for integral functionals with bv integrands. *ESAIM: Control, Optimisation and Calculus of Variations*, 14(3):456–477, 2008.

[5] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Courier Corporation, 2000.

[6] A. Arroyo-Rabasa. Characterization of generalized young measures generated by $\mathcal{A}$-free measures. *arXiv preprint arXiv:1908.03186*, 2019.

[7] A. Arroyo-Rabasa and A. Skorobogatova. On the fine properties of elliptic operators. *arXiv preprint arXiv:1911.08474*, 2019.

[8] P. Athavale, R.L. Jerrard, M. Novaga, and G. Orlandi. Weighted tv minimization and applications to vortex density models. *arXiv preprint arXiv:1509.03713*, 2015.

[9] J.M. Ball. The calculus of variations and materials science. *Quarterly of Applied Mathematics*, 56(4):719–740, 1998.

[10] J.M. Ball and J.E. Marsden. Quasiconvexity at the boundary, positivity of the second variation and elastic stability. *Archive for rational mechanics and analysis*, 86(3):251–277, 1984.

[11] M. Bergounioux and L. Piffet. A second-order model for image denoising. *Set-Valued and Variational Analysis*, 18(3):277–306, 2010.

[12] K. Bredies and M. Holler. Regularization of linear inverse problems with total generalized variation. *Journal of inverse and ill-posed problems*, 22(6):871–913, 2014.

[13] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[14] K. Bredies and T. Valkonen. Inverse problems with second-order total generalized variation constraints. *arXiv preprint arXiv:2005.09725*, 2020.

[15] D. Breit, L. Diening, and F. Gmeineder. On the trace operator for functions of bounded $\mathbb{A}$-variation. *Analysis & PDE*, 13(2):559–594, 2020.

[16] Y. Brenier, C. De Lellis, and L. Székelyhidi. Weak-strong uniqueness for measure-valued solutions. *Communications in mathematical physics*, 305(2):351–361, 2011.

[17] Yann Brenier. Hydrodynamic structure of the augmented born-infeld equations. *Archive for Rational Mechanics and Analysis*, 172(1):65–91, 2004.

[18] M. Burger, K. Papafitsoros, E. Papoutsellis, and C.B. Schönlieb. Infimal convolution regularisation functionals of bv and l^p spaces. *Journal of Mathematical Imaging and Vision*, 55(3):343–369, 2016.

[19] L. Calatroni, C. Cao, J.C. De Los Reyes, C.B. Schönlieb, and T. Valkonen. Bilevel approaches for learning of variational imaging models. In *Variational methods*, pages 252–290. De Gruyter, 2017.

[20] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, 1997.

[21] B. Charles and Jr. Morrey. Multiple integrals in the calculus of variations. *Springer-Verlag New York, Inc., New York, 1966, Die Grundlehren der mathematischen Wissenschaften, Band 130. MR0202511 (34:2380.*

[22] C. Christoforou, M. Galanopoulou, and A. E. Tzavaras. A symmetrizable extension of polyconvex thermoelasticity and applications to zero-viscosity limits and weak-strong uniqueness. *Comm. Partial Differential Equations*, 43(7):1019–1050, 2018.

[23] C. Christoforou, M. Galanopoulou, and A. E. Tzavaras. Measure-valued solutions for the equations of polyconvex adiabatic thermoelasticity. *Discrete Contin. Dyn. Syst.*, 39(11):6175–6206, 2019.

[24] C. Christoforou, M. Galanopoulou, and A. E. Tzavaras. A discrete variational scheme for isentropic processes in polyconvex thermoelasticity. *Discrete Contin. Dyn. Syst.*, 59(122), 2020.

[25] C. Christoforou and A. E. Tzavaras. Relative entropy for hyperbolic-parabolic systems and application to the constitutive theory of thermoviscoelasticity. *Arch. Ration. Mech. Anal.*, 229(1):1–52, 2018.

[26] C.V. Chung, J.C. De los Reyes, and C.B. Schönlieb. Learning optimal spatially-dependent regularization parameters in total variation image denoising. *Inverse Problems*, 33(7):074005, 2017.

[27] B. D Coleman and Ellis H Dill. Thermodynamic restrictions on the constitutive equations of electromagnetic theory. *Zeitschrift für angewandte Mathematik und Physik*, 22(4):691–702, 1971.

[28] B. D. Coleman and W. Noll. The thermodynamics of elastic materials with heat conduction and viscosity. In *The Foundations of Mechanics and Thermodynamics*, pages 145–156. Springer, 1974.

[29] J. C. Cordero and K. Koumatos. Necessary and sufficient conditions for the strong local minimality of c^1 extremals on a class of non-smooth domains. *ESAIM: Control, Optimisation and Calculus of Variations*, 26:49, 2020.

[30] J.C. Cordero. *Regularity and uniqueness in the Calculus of Variations*. PhD thesis, Oxford University, UK, 2014.

[31] J.C. Cordero. Boundary regularity and sufficient conditions for strong local minimizers. *Journal of Functional Analysis*, 272(11):4513–4587, 2017.

[32] B. Dacorogna. Quasi-convexité et semi-continuité inférieure faible des fonctionnelles non linéaires. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 9(4):627–644, 1982.

[33] B. Dacorogna. *Weak continuity and weak lower semicontinuity of non-linear functionals*, volume 922. Springer, 2006.

[34] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78. Springer Science & Business Media, 2007.

[35] C. M. Dafermos. The second law of thermodynamics and stability. *Arch. Rational Mech. Anal.*, 70(2):167–179, 1979.

[36] C. M. Dafermos. Stability of motions of thermoelastic fluids. *Journal of Thermal Stresses*, 2(1):127–134, 1979.

[37] C. M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, fourth edition, 2016.

[38] C.M. Dafermos. Quasilinear hyperbolic systems with involutions. *Archive for Rational Mechanics and Analysis*, 94(4):373–389, 1986.

[39] C.M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 3. Springer, 2005.

[40] C.M. Dafermos. Non-convex entropies for conservation laws with involutions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(2005):20120344, 2013.

[41] E. Davoli, I. Fonseca, and P. Liu. Adaptive image processing: first order pde constraint regularizers and a bilevel training scheme. *arXiv preprint arXiv:1902.01122*, 2019.

[42] J.C. De los Reyes and C.B. Schönlieb. Image denoising: learning the noise model via nonsmooth pde-constrained optimization. *Inverse Problems & Imaging*, 7(4):1183, 2013.

[43] J.C. De Los Reyes, C.B. Schönlieb, and T. Valkonen. The structure of optimal parameters for image restoration problems. *Journal of Mathematical Analysis and Applications*, 434(1):464–500, 2016.

[44] J.C. De Los Reyes, C.B. Schönlieb, and T. Valkonen. Bilevel parameter learning for higher-order total variation regularisation models. *Journal of Mathematical Imaging and Vision*, 57(1):1–25, 2017.

[45] J.C. De los Reyes and D. Villacís. Optimality conditions for bilevel imaging learning problems with total variation regularization. *arXiv preprint arXiv:2107.08100*, 2021.

[46] Guido De Philippis and Filip Rindler. On the structure of $\mathcal{A}$-free measures and applications. *Annals of Mathematics*, pages 1017–1039, 2016.

[47] H.P. Decell Jr. An application of the Cayley-Hamilton theorem to generalized matrix inversion. *SIAM review*, 7(4):526–528, 1965.

[48] F. Demengel and R. Temam. Convex functions of a measure and applications. *Indiana University Mathematics Journal*, 33(5):673–709, 1984.

[49] R. J. DiPerna. Uniqueness of solutions to hyperbolic conservation laws. *Indiana University Mathematics Journal*, 28(1):137–188, 1979.

[50] I. Fonseca and G. Leoni. On lower semicontinuity and relaxation. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 131(3):519–565, 2001.

[51] I. Fonseca and G. Leoni. *Modern Methods in the Calculus of Variations: L^p Spaces*. Springer Science & Business Media, 2007.

[52] I. Fonseca and S. Müller. $\mathcal{A}$-Quasiconvexity, Lower Semicontinuity, and Young Measures. *SIAM Journal on Mathematical Analysis*, 30(6):1355–1390, 1999.

[53] I. Fonseca and S. Müller. $\mathcal{A}$-quasiconvexity, Lower Semicontinuity, and Young Measures. *SIAM Journal on Mathematical Analysis*, 30(6):1355–1390, 1999.

[54] K. O. Friedrichs and P. D. Lax. Systems of conservation equations with a convex extension. *Proc. Nat. Acad. Sci. U.S.A.*, 68:1686–1688, 1971.

[55] M. Galanopoulou, K. Koumatos, and A. Vikelis. Weak-strong uniqueness for measure-valued solutions to the equations of quasiconvex adiabatic thermoelasticity. *arXiv preprint arXiv:2109.15151*, 2021.

[56] Lars Gårding. Dirichlet's problem for linear elliptic partial differential equations. *Mathematica Scandinavica*, pages 55–72, 1953.

[57] F. Gmeineder and J. Kristensen. Partial regularity for BV minimizers. *Archive for Rational Mechanics and Analysis*, 232(3):1429–1473, 2019.

[58] F. Gmeineder and B. Raiţă. Embeddings for $\mathbb{A}$-weakly differentiable functions on domains. *Journal of Functional Analysis*, 277(12):108278, 2019.

[59] Y. Grabovsky and T. Mengesha. Direct approach to the problem of strong local minima in calculus of variations. *Calculus of Variations and Partial Differential Equations*, 29(1):59–83, 2007.

[60] Y. Grabovsky and T. Mengesha. Sufficient conditions for strong local minima: The case of C^1 extremals. *Transactions of the American Mathematical Society*, 361(3):1495–1541, 2009.

[61] Y. Grabovsky and T. Mengesha. Sufficient conditions for strong local minima: The case of C^1 extremals. *Transactions of the American Mathematical Society*, 361(3):1495–1541, 2009.

[62] A. Guerra, J. Kristensen, and B. Raiţă. Oscillation and concentration under constant rank constraints, 2020. OxPDE Technical Report.

[63] A. Guerra and B. Raiţă. Quasiconvexity, null lagrangians, and Hardy space integrability under constant rank constraints. *arXiv preprint arXiv:1909.03923*, 2019.

[64] P. Gwiazda, A. Świerczewska Gwiazda, and E. Wiedemann. Weak-strong uniqueness for measure-valued solutions of some compressible fluid models. *Nonlinearity*, 28(11):3873–3890, 2015.

[65] M.R. Hestenes. Sufficient conditions for multiple integral problems in the calculus of variations. *American Journal of Mathematics*, 70(2):239–276, 1948.

[66] M. Hintermüller and K. Papafitsoros. Generating structured nonsmooth priors and associated primal-dual methods. In *Handbook of numerical analysis*, volume 20, pages 437–502. Elsevier, 2019.

[67] M. Hintermüller, K. Papafitsoros, C.N. Rautenberg, and H. Sun. Dualization and automatic distributed parameter selection of total generalized variation via bilevel optimization. *arXiv preprint arXiv:2002.05614*, 2020.

[68] M. Hintermüller and C.N. Rautenberg. Optimal selection of the regularization function in a weighted total variation model. part i: Modelling and theory. *Journal of Mathematical Imaging and Vision*, 59(3):498–514, 2017.

[69] M. Hintermüller, C.N. Rautenberg, T. Wu, and A. Langer. Optimal selection of the regularization function in a weighted total variation model. part ii: Algorithm, its analysis and numerical tests. *Journal of Mathematical Imaging and Vision*, 59(3):515–533, 2017.

[70] M. Hintermüller and G. Stadler. An infeasible primal-dual algorithm for total bounded variation–based inf-convolution-type image restoration. *SIAM Journal on Scientific Computing*, 28(1):1–23, 2006.

[71] D. Kinderlehrer and P. Pedregal. Characterizations of young measures generated by gradients. *Archive for rational mechanics and analysis*, 115(4):329–365, 1991.

[72] D. Kinderlehrer and P. Pedregal. Gradient young measures generated by sequences in sobolev spaces. *The Journal of Geometric Analysis*, 4(1):59–90, 1994.

[73] K. Koumatos and S. Spirito. Quasiconvex elastodynamics: Weak-strong uniqueness for measure-valued solutions. *Communications on Pure and Applied Mathematics*, 07 2017.

[74] K. Koumatos and S. Spirito. Quasiconvex elastodynamics: Weak-strong uniqueness for measure-valued solutions. *Communications on Pure and Applied Mathematics*, 72(6):1288–1320, 2019. [Corrigendum: in press].

[75] K. Koumatos and A. Vikelis. $\mathcal{A}$-quasiconvexity, gårding inequalities, and applications in pde constrained problems in dynamics and statics. *SIAM Journal on Mathematical Analysis*, 53(4):4178–4211, 2021.

[76] J. Krämer, S. Krömer, M. Kružík, and G. Pathó. $\mathcal{A}$-quasiconvexity at the boundary and weak lower semicontinuity of integral functionals. *Advances in Calculus of Variations*, 10(1):49–67, 2017.

[77] J. Kristensen. *Finite functionals and Young measures generated by gradients of Sobolev functions*. Technical University of Denmark. Department of Mathematics, 1994.

[78] J. Kristensen. Lower semicontinuity in spaces of weakly differentiable functions. *Mathematische Annalen*, 313(4):653–710, 1999.

[79] J. Kristensen and J. Campos Cordero. Personal communication, 2012.

[80] J. Kristensen and B. Raiţă. Oscillation and concentration in sequences of PDE constrained measures. *arXiv preprint arXiv:1912.09190*, 2019.

[81] J. Kristensen and B. Raiţă. An introduction to generalized young measures. *Lecture notes 45/2020, Max Planck Institute for Mathematics in the Sciences*, 2020.

[82] J. Kristensen and F. Rindler. Characterization of generalized gradient young measures generated by sequences in $w^{1,1}$ and bv. *Archive for rational mechanics and analysis*, 197(2):539–598, 2010.

[83] J. Kristensen and A. Taheri. Partial regularity of strong local minimizers in the multi-dimensional calculus of variations. *Archive for rational mechanics and analysis*, 170(1):63–89, 2003.

[84] K. Kunisch and T. Pock. A bilevel optimization approach for parameter learning in variational models. *SIAM Journal on Imaging Sciences*, 6(2):938–983, 2013.

[85] N.G. Meyers. Quasi-convexity and lower semi-continuity of multiple variational integrals of any order. *Transactions of the American Mathematical Society*, 119(1):125–149, 1965.

[86] S. Müller. Variational models for microstructure and phase transitions. In *Calculus of Variations and geometric evolution problems*, pages 85–210. Springer, 1999.

[87] F. Murat. Compacité par compensation. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 5(3):489–507, 1978.

[88] J. Necas. *Les méthodes directes en théorie des équations elliptiques.* Masson, 1967.

[89] V. Pagliari, K. Papafitsoros, B. Raiţă, and A. Vikelis. Bilevel training schemes in imaging for total-variation-type functionals with convex integrands. *arXiv preprint arXiv:2112.10682*, 2021.

[90] K. Papafitsoros. *Novel higher order regularisation methods for image reconstruction.* PhD thesis, University of Cambridge, 2015.

[91] K. Papafitsoros and C.B. Schönlieb. A combined first and second order variational approach for image reconstruction. *Journal of mathematical imaging and vision*, 48(2):308–338, 2014.

[92] B. Raiţă. Potentials for $\mathcal{A}$-quasiconvexity. *Calculus of Variations and Partial Differential Equations*, 58(3), 2019.

[93] B. Raiţă and A. Skorobogatova. Continuity and canceling operators of order n on $\mathbb{R}^n$. *arXiv preprint arXiv:1903.03574*, 2019.

[94] F. Rindler. *Calculus of Variations.* Springer, 2018.

[95] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.

[96] D. Serre. Hyperbolicity of the nonlinear models of maxwell?s equations. *Archive for Rational Mechanics and Analysis*, 172(3):309–331, 2004.

[97] Swarnendu Sil. Regularity for elliptic systems of differential forms and applications. *Calculus of Variations and Partial Differential Equations*, 56(6):172, 2017.

[98] Jack WD Skipper and Emil Wiedemann. Lower semi-continuity for $\mathcal{A}$-quasiconvex functionals under convex restrictions. *arXiv preprint arXiv:1909.11543*, 2019.

[99] K.T. Smith. Formulas to represent functions by their derivatives. *Mathematische Annalen*, 188(1):53–77, 1970.

[100] MA.. Sychev. A new approach to young measure theory, relaxation and convergence in energy. In *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, volume 16, pages 773–812. Elsevier, 1999.

[101] A. Taheri. Sufficiency theorems for local minimizers of the multiple integrals of the calculus of variations. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 131(1):155–184, 2001.

[102] L Tartar. Une nouvelle méthode de résolution d?équations aux dérivées partielles non linéaires. In *Journées d?Analyse non linéaire*, pages 228–241. Springer, 1978.

[103] L. Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Watt symposium*, volume 4, pages 136–212, 1979.

[104] R. Temam and G. Strang. Functions of bounded deformation. *Archive for Rational Mechanics and Analysis*, 75(1):7–21, 1980.

[105] C. Truesdell and W. Noll. *The non-linear field theories of mechanics*. Springer-Verlag, Berlin, third edition, 2004. Edited and with a preface by Stuart S. Antman.

[106] T. Valkonen, K. Bredies, and F. Knoll. Total generalized variation in diffusion tensor imaging. *SIAM Journal on Imaging Sciences*, 6(1):487–525, 2013.

[107] J. Van Schaftingen. Limiting sobolev inequalities for vector fields and canceling linear differential operators. *Journal of the European Mathematical Society*, 15(3):877–921, 2013.

[108] L. Vese. A study in the *bv* space of a denoising—deblurring variational problem. *Applied Mathematics and optimization*, 44(2):131–161, 2001.

[109] K. Zhang. Remarks on quasiconvexity and stability of equilibria for variational integrals. *Proceedings of the American Mathematical Society*, 114(4):927–930, 1992.

[110] K. Zhang. Remarks on quasiconvexity and stability of equilibria for variational integrals. *Proceedings of the American Mathematical Society*, 114(4):927–930, 1992.