



# Total roto-translational variation

Antonin Chambolle<sup>1</sup> · Thomas Pock<sup>2,3</sup>

Received: 10 October 2017 / Revised: 7 January 2019 / Published online: 16 March 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

We consider curvature depending variational models for image regularization, such as Euler’s elastica. These models are known to provide strong priors for the continuity of edges and hence have important applications in shape- and image processing. We consider a lifted convex representation of these models in the roto-translation space: in this space, curvature depending variational energies are represented by means of a convex functional defined on divergence free vector fields. The line energies are then easily extended to any scalar function. It yields a natural generalization of the total variation to curvature-dependent energies. As our main result, we show that the proposed convex representation is tight for characteristic functions of smooth shapes. We also discuss cases where this representation fails. For numerical solution, we propose a staggered grid discretization based on an averaged Raviart–Thomas finite elements approximation. This discretization is consistent, up to minor details, with the underlying continuous model. The resulting non-smooth convex optimization problem is solved using a first-order primal-dual algorithm. We illustrate the results of our numerical algorithm on various problems from shape- and image processing.

**Mathematics Subject Classification** 53A04 · 49Q20 · 26A45 · 35J35 · 53A40 · 65K10

## 1 Introduction

It was observed at least since [58] that line energies such as Euler’s “Elastica” (here for a smooth curve  $\gamma$ ):

---

✉ Antonin Chambolle  
antonin.chambolle@cmap.polytechnique.fr  
Thomas Pock  
pock@icg.tugraz.at

<sup>1</sup> CMAP, Ecole Polytechnique, CNRS, 91128 Palaiseau, France

<sup>2</sup> Institute for Computer Graphics and Vision, Graz University of Technology, 8010 Graz, Austria

<sup>3</sup> Center for Vision, Automation and Control, AIT Austrian Institute of Technology GmbH, 1220 Vienna, Austria

$$\int_{\gamma} \kappa^2 ds \quad \text{or} \quad \int_{\gamma} 1 + \kappa^2 ds$$

could be natural regularizers for the completion of missing contours in images. This idea was based in particular on observations of Kanizsa [46,47] about the way our perception can “invent” apparent contours. The model in [58] (see also [39,40] for interesting attempts to solve it with phase-field methods) was variational in nature and since then, many attempts have been proposed to study and address the minimization of such line energies, both theoretically and numerically.

From the theoretical point of view, the lower-semicontinuity of these energies already is a challenge, which has been studied in many papers (in connexion to the applications to computer vision) since at least the 90s [12–17,30]. It is shown already in [12] that the boundary of many sets with cusps can be approximated by sets with smooth boundaries of bounded energy, showing that the relaxation of the Elastica for boundaries of sets is already far from trivial. The study of this lower semi-continuity of course enters the long history of the study of general curvature dependent energies of manifolds and in particular the Willmore energy [77].

Quite early, it has been suggested to lift the manifold in a larger space where a variable represents its direction or orientation, by means in particular (for co-dimension one manifolds) of the Gauss map  $(x, \nu(x))$ ,  $\nu(x)$  being the normal to the manifold at  $x$  [5,6]. Such approach allows to study very general curvature energies and has been successfully used for establishing lower-semicontinuity and existence results [6,7,31,32], and in particular to lines energies such as ours in higher codimension [1,2] or in dimension 2 [31]. We must mention also in this class an older approach based on “curvature varifolds” (which is very natural since varifolds are defined on the cross product of spatial and directional variables) which has allowed to show existence results since the 80s [45], see also [50].

Interestingly, it was understood much earlier [44] that (cats’) vision was functioning in a similar way (eg., using a sort of “Gauss map”), thanks to neurons sensitive to particular directions which were found to be stacked inside the visual cortex into ordered columns, making us sensitive to changes of orientations (and thus curvature intensity). These findings (which might explain some of Kanizsa’s experiments) inspired some mathematical models quite early [48], however they were formalized into a consistent geometric interpretation later on [29,33,57,59,65,66]. The idea of these authors is to lift 2D curves in the “Roto-translation” space (which is the group of rotations and translations, however what is more important here is less its group structure than the associated *sub-Riemannian* metrics) which can be identified, for our purposes, with  $\Omega \times \mathbb{S}^1$  where  $\Omega \subset \mathbb{R}^2$  is the spatial domain where the image is defined and  $\mathbb{S}^1$  parameterizes the local orientation.

The natural metric in that space prevents from moving spatially in a direction other than the local orientation, which makes it singular (hence “sub”-Riemannian). Diffusion and mean curvature flow in this particular metrics were successfully used as efficient methods for image “inpainting”, which is the task of filling in a gap in an image [19,41,61,72]. Indeed, such diffusion naturally extends missing level sets into smooth curves, and even allows for crossing, since curves with different directions “live” in different locations in the Roto-translation space. This seems to improve in difficult

situations (e.g., crossings) upon more classical diffusion models for inpainting [18,27] (or Weickert's "EED" [67,75,76] which can propagate directional information quite smoothly across inpainted regions).

In computer science, similar ideas were successfully implemented in discrete graphs (with nodes representing a spatial point and orientation) in order to minimize curvature-dependent contour energies [36–38,49,68–71]. Our current work is closer to these approaches, although set up in the continuous setting, as we want to represent and solve variational models involving curvature terms, where the unknowns are scalar (grey-level) functions. A continuous approach similar to these discrete ones proposes to compute minimal paths in the Roto-translational metric by solving the corresponding eikonal equations [28,55,56] (the goal being more here to find paths on an image than to complete boundaries). This is strongly connected to geometric control problems (such as parking a car), and these connections have led to the study of many interesting metrics in these settings, with impressive applications to imaging problems such as the extraction of networks of vessels and fibers in 2D and 3D medical images via fast, globally optimal sub-Riemannian and sub-Finslerian geodesic tracking in the roto-translation group [11,34].

As said, we wish to extend these variational methods to level sets representations. The family of problems we are interested in is introduced in a paper by Masnou and Morel [52] (see also [53]), which addresses the problem of image inpainting. Their initial idea is to minimize, in the inpainting domain  $D$ , an energy of the form

$$\int_D \left( 1 + \left| \operatorname{div} \frac{Du}{|Du|} \right|^p \right) |Du|$$

for some  $p \geq 1$ , where  $u$  is a bounded variation function and appropriate boundary conditions are given. A theoretical study of this energy, which in general is not lower-semicontinuous (lsc), is found in [4], in particular it is shown that if  $p > 1$  (in dimension 2), it is lsc on  $C^2$  functions. Interestingly, also this study relies on a space/direction representation (and more precisely on varifolds). An interesting co-area formula for the relaxed envelope is also shown in [51].

There have been many attempts to numerically solve Masnou and Morel's model, which is very difficult to tackle, being highly non convex. Most of the authors introduce auxiliary variables [9,10,26], for instance representing the orientation, which is already close to the idea of the Gauss map or Roto-translational representation. Recent techniques based on Augmented Lagrangian methods (for coupling the auxiliary variables) [8,42,43,74,78–80] have shown to be quite efficient, despite the lack of convexity and hence convergence guarantees.

What we propose here is to rather introduce a functional which may be seen as a convex relaxation of Masnou and Morel's. The method we propose is based on two ingredients: A lifting in the Roto-translation space of curves which allows to write the Elastica energy (or any convex function of the curvature) as a convex function, as in classical Gauss map based approaches, and (formally) a decomposition of functions as sum of characteristic functions of sets with finite energy. This allows to define a

convex functional which is defined on grey-level functions and penalises the curvature of the level lines. It is however easy to check that it is in general strictly below Masnou and Morel's functional, in particular a function of finite energy needs not have its level sets regular (in practice, they could be intersections of regular sets, in addition to having possible cusps, for the same reasons as in [12]), see for example Figs. 5 and 9e. The only theoretical result which we can show is that  $C^2$  curves are tightly represented by our convex relaxation.

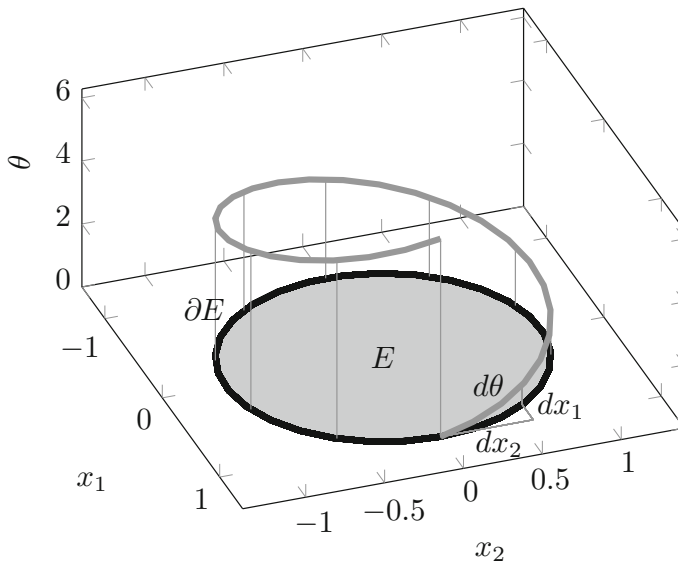
There is a close relationship between our approach and the functionals in [22] ("TVX") and [23], based on similar representations (but [22] omits to preserve the boundary of the lifted current). In fact, the functional we build is, as we show further on, a new expression of the previous convex relaxation of the Elastica energy (and variants) proposed in [23]. This stems from the identity of the corresponding dual problems. Although it boils down to the same energy, it is introduced in a much simpler way, as the primal expression in [23] requires to work in a space where the point, the tangent and the curvature are lifted as independent variables. As a consequence, we also can provide a simpler discretization, and the tightness of the relaxation for  $C^2$  sets was unnoticed in [23]. Eventually, we should mention that part of what we propose here could be generalized to arbitrary dimension or co-dimension, following the techniques in [6]. However, computationally, our construction for 1D curves in the plane already needs to work in a 3D space, and more complex models seem at this moment intractable. Our future research will rather focus on improving the discretization and the optimization of the bidimensional case.

The paper is organised as follows: In Sect. 2 we describe the lifting and introduce our functional. We state our main result, which shows that characteristic functions of  $C^2$  sets are well represented by our convex relaxation (Theorem 1). In the next Sect. 3 we show an approximation result with smooth functions, and compute a dual representation of the functional (which shows it is identical to [23]). Then, in Sect. 4, we describe how the functional is discretized and show a few experimental results. The last Sect. 5 is devoted to the proof of Theorem 1. Some technical tools are found in the Appendix, in particular, "Appendix A" shows, up to (hopefully minor) transformations, the consistency of our implementation in Sect. 4.

## 2 The functional

We consider  $\Omega$  a bounded open domain in the plane, and  $E \subset \Omega$  a set with  $C^2$  (and to simplify, connected) boundary. Assume that this boundary  $\partial E$  is given by a parameterized curve  $(x_1(t), x_2(t))$  with parameter  $t \in [0, 1]$ . The (extrinsic) curvature  $\kappa_E$  of  $\partial E$  is classically defined as the ratio between the variation of the tangential angle  $\theta$  and the variation of its arc length  $s$ , that is

$$\kappa_E = \frac{d\theta}{ds} = \frac{\frac{d\theta}{dt}}{\frac{ds}{dt}} = \frac{\frac{d\theta}{dt}}{\sqrt{\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2}}.$$



**Fig. 1** The gray line is the lifting of the boundary  $\partial E$  of the disk  $E$  to the roto-translational space  $\Omega \times \mathbb{S}^1$

The main idea of the lifting is now to consider a higher dimensional representation of the parametric curve in the 3D roto-translation (RT) space, which is obtained by adding the tangential angle  $\theta$  as an additional dimension to  $\Omega$ . In this space, we now consider a parameterized 3D curve  $(x_1(t), x_2(t), \theta(t))$  which lifts the boundary  $\partial E$  to the roto-translational space. Figure 1 shows an example where we lift the boundary of a disk (which is a circle) to the RT space. Observe that the lifted boundary is represented by a 3D helix.

Now, we define for all  $t \in [0, 1]$  the tangential vector  $p(t) = (p^x(t), p^\theta(t))$  with

$$p^x(t) = \left( \frac{dx_1(t)}{dt}, \frac{dx_2(t)}{dt} \right), \quad p^\theta(t) = \frac{d\theta(t)}{dt},$$

$$|p^x(t)| = \sqrt{\left( \frac{dx_1(t)}{dt} \right)^2 + \left( \frac{dx_2(t)}{dt} \right)^2}.$$

The curvature is therefore given by

$$\kappa_E(t) = \frac{p^\theta(t)}{|p^x(t)|},$$

In this work we consider  $f : \mathbb{R} \rightarrow [0, +\infty]$  a convex, lsc function and want to define a convex lsc extension (to grey-level valued functions) of energies of the type

$$E \mapsto \int_{\partial E} f(\kappa_E) d\mathcal{H}^1$$

where  $E \subset \Omega$  is a set with  $C^2$  boundary, and  $\kappa_E$  is the curvature of the set. Using our tangential vector  $p(t)$ , it is easy to see that the energy can be (formally) written as

$$\int_{\partial E} f(\kappa_E) d\mathcal{H}^1 = \int_0^1 f(p^\theta/|p^x|)|p^x| dt = \int_{\Omega \times \mathbb{S}^1} f(\tau^\theta/|\tau^x|)|\tau^x| d\mathcal{H}^1 \llcorner \Gamma_E$$

where  $\Gamma_E = (x, \theta)([0, 1])$  is the lifted curve and  $\tau(x, \theta)$  its normalized tangential vector, given by  $\tau(x(t), \theta(t)) = p(t)/|p(t)|$  for all  $t \in [0, 1]$ . A precise definition of the energy will be given below.

Let us briefly discuss three instances of energies that will typically appear in applications. In all cases  $\alpha > 0$  will be a tuning parameter that can be used to balance the influence of the curvature term with respect to the length.

1.  $f_1(t) = 1 + \alpha|t|$ . This energy penalizes the arclength plus the absolute curvature, hence we might expect that this type of energy will allow also for corners. This type of energy has been studied with a different approach in [22]. The energy is given by:

$$\int_{\Omega \times \mathbb{S}^1} f_1(\tau^\theta/|\tau^x|)|\tau^x| d\mathcal{H}^1 \llcorner \Gamma_E = \int_{\Omega \times \mathbb{S}^1} |\tau^x| + \alpha|\tau^\theta| d\mathcal{H}^1 \llcorner \Gamma_E.$$

2.  $f_2(t) = \sqrt{1 + \alpha^2|t|^2}$ . This energy penalizes the arclength of the lifted curve in the RT space which, in some sense, corresponds to the “Total Roto-translational Variation”. It yields the energy:

$$\int_{\Omega \times \mathbb{S}^1} f_2(\tau^\theta/|\tau^x|)|\tau^x| d\mathcal{H}^1 \llcorner \Gamma_E = \int_{\Omega \times \mathbb{S}^1} \sqrt{|\tau^x|^2 + \alpha^2|\tau^\theta|^2} d\mathcal{H}^1 \llcorner \Gamma_E$$

which is nothing but the length of the lifted curve  $\Gamma_E$  in a Riemannian metric. This functional was considered for perceptual completion problems probably first in [65]. It coincides with sub-Riemannian [11, 20, 33, 65] and is related to the sub-Finslerian models in [34] (with the constraint of positive direction on the velocity).

3.  $f_3(t) = 1 + \alpha^2|t|^2$ . This is the classical Euler’s Elastica energy studied for example in [12, 23, 53, 58, 68] and many other works already mentioned in our introduction.

$$\int_{\Omega \times \mathbb{S}^1} f_3(\tau^\theta/|\tau^x|)|\tau^x| d\mathcal{H}^1 \llcorner \Gamma_E = \int_{\Omega \times \mathbb{S}^1} |\tau^x| + \alpha^2 \frac{|\tau^\theta|^2}{|\tau^x|} d\mathcal{H}^1 \llcorner \Gamma_E.$$

Observe that the second term in the energy is a “quadratic over linear” function which is still convex, provided  $\tau^x$  is constrained in a half space or a half line (which of course will depend on the  $\theta$  variable).

In all these examples, we find that the energy which we are interested in ends up represented possibly as a *convex* function of the measure  $\sigma = \tau\mathcal{H}^1 \llcorner \Gamma_E$ . Moreover, this  $\sigma$  is not an arbitrary measure. In particular, it satisfies two important constraints:

first, by construction, it is a circulation and has zero divergence in  $\Omega \times \mathbb{S}^1$  (it can have source terms on  $\partial\Omega \times \mathbb{S}^1$ ): indeed, given any smooth function  $\psi$ ,  $\int \nabla \psi \cdot \sigma = \int_{\Gamma_E} \partial_\tau \psi d\mathcal{H}^1$  vanishes if  $\Gamma_E$  is a closed curve, or if it ends on the boundary and  $\psi$  has compact support. Second, its marginals in  $\Omega \times \mathbb{S}^1$ , which we can formally denote  $\int_{\mathbb{S}^1} \sigma$  (and which are also divergence free), coincide with a  $90^\circ$ -rotation of the measure  $D\chi_E$ . We will show now how to generalize this construction to arbitrary sets or functions (with bounded variation).

In the whole paper, we assume that there exists  $\gamma > 0$  such that

$$f(t) \geq \gamma \sqrt{1+t^2}, \quad \forall t \in \mathbb{R}. \quad (1)$$

This is a technical assumption which is not necessary for most of the formal derivation which will follow, but seems crucial to ensure some important properties of the functional we will introduce later on, such as lower semi-continuity, and in the subsequent proofs. We also introduce the *recession function*

$$f^\infty(t) = \lim_{s \rightarrow +\infty} \frac{1}{s} f(st)$$

which is a convex, one-homogeneous function possibly infinite on  $(-\infty, 0)$  or/and  $(0, +\infty)$ . It is easy to check that it is the support function of  $\text{dom } f^*$ , the domain of the convex conjugate  $f^*$  of  $f$ :

$$f^\infty(t) = \sup_{y \in \text{dom } f^*} ty. \quad (2)$$

Indeed, for  $y \in \text{dom } f^*$ ,  $s > 0$ ,

$$ty = \frac{1}{s}(st)y \leq \frac{1}{s}f(st) + \frac{1}{s}f^*(y) \xrightarrow{s \rightarrow \infty} f^\infty(t)$$

as  $s \rightarrow \infty$ ; on the other hand, as  $f^* \geq -\gamma$  thanks to (1), for  $s > 0$  one has

$$\frac{1}{s}f(st) = \sup_{y \in \text{dom } f^*} ty - \frac{1}{s}f^*(y) \leq \sup_{y \in \text{dom } f^*} ty + \frac{\gamma}{s}$$

showing that  $f^\infty(t) \leq \sup_{y \in \text{dom } f^*} ty$ .

We let then, for  $p = (p^x, p^\theta) \in \mathbb{R}^3$  with  $p^x \neq 0$ ,

$$h(\theta, p) = \begin{cases} |p^x| f(p^\theta/|p^x|) & \text{if } p^x \in \mathbb{R}_+ \underline{\theta}, p^x \neq 0, \\ f^\infty(p^\theta) & \text{if } p^x = 0, \\ +\infty & \text{else.} \end{cases} \quad (3)$$

Here,  $\underline{\theta}$  denotes the unit planar vector  $(\cos \theta, \sin \theta)^T$  (by a slight abuse of notation, we will also denote in this way the vector  $(\cos \theta, \sin \theta, 0)^T \in \mathbb{R}^3$ ). It is then classical [64, § 13] and easily follows from (2) that one has

$$h(\theta, p) = \sup \{ \xi \cdot p : \xi^x \cdot \underline{\theta} \leq -f^*(\xi^\theta) \}, \quad (4)$$

that is, the one-homogeneous function  $h$  is the support function of the convex set in the right-hand side of (4). Indeed, first of all, the sup in (4) is  $+\infty$  if  $p^x$  is not  $\lambda \underline{\theta}$ ,  $\lambda \geq 0$ . While if it is of this form, then (taking the supremum over the  $\xi$ 's which satisfy  $\xi^x \cdot \underline{\theta} \leq -f^*(\xi^\theta)$ )

$$\sup_{\xi : \xi^x \cdot \underline{\theta} \leq -f^*(\xi^\theta)} \xi^x \cdot \underline{\theta} \lambda + \xi^\theta p^\theta = \sup_{\xi^\theta} \xi^\theta p^\theta - \lambda f^*(\xi^\theta) = \begin{cases} \lambda f(p^\theta / \lambda) & \text{if } \lambda > 0 \\ f^\infty(p^\theta) & \text{else.} \end{cases}$$

Observe that (1) yields that for all  $p \in \mathbb{R}^3$ ,

$$h(\theta, p) \geq \gamma |p| \quad (5)$$

We now introduce the functional

$$F(u) = \inf \left\{ \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) dx d\theta : \operatorname{div} \sigma = 0, \int_{\mathbb{S}^1} \sigma^x d\theta = Du^\perp \right\}. \quad (6)$$

Here  $x^\perp = (x_2, -x_1)$  is a  $90^\circ$  rotation in the plane. The last condition is understood as follows: for all  $\varphi \in C_c^1(\Omega; \mathbb{R}^2) \subset C_c^1(\Omega \times \mathbb{S}^1; \mathbb{R}^2)$ , one has

$$\int_{\Omega \times \mathbb{S}^1} \varphi^\perp \cdot \sigma^x = \int_{\Omega} \varphi \cdot Du = - \int_{\Omega} u \operatorname{div} \varphi dx. \quad (7)$$

In general, the fields  $\sigma$  appearing in (6) are merely free-divergence bounded Radon measures in  $\Omega \times \mathbb{S}^1$  with values in  $\mathbb{R}^3$  (we denote  $\mathcal{M}^1(\Omega \times \mathbb{S}^1; \mathbb{R}^3)$  the space of such measures) and the proper way to write the integral is rather

$$\int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) = \int_{\Omega \times \mathbb{S}^1} h\left(\theta, \frac{\sigma}{|\sigma|}\right) d|\sigma|$$

where  $\sigma/|\sigma|$  is the Radon-Besicovitch derivative of  $\sigma$  w.r. its total variation  $|\sigma|$ . Notice that for any  $u$  and any admissible  $\sigma$ , thanks to (5),

$$F(u) \geq \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) \geq \gamma \int_{\Omega \times \mathbb{S}^1} |\sigma| \geq \gamma \int_{\Omega} |Du|$$

because of (7), taking the supremum with respect to the test function  $\varphi$  with  $|\varphi(x)| \leq 1$  everywhere. This shows that  $F(u)$  bounds the  $BV$  seminorm. On the other hand, there exist functions  $u \in BV(\Omega)$  with  $F(u) = +\infty$ . An example is as follows: we consider  $w^0 : (0, 1) \rightarrow [-1, 1]$  a smooth function with compact support, and let, for  $x \in (0, 1)$ ,  $w(x) = \sum_{i=1}^{\infty} 4^{-i} w^0(2^i x - 1)$ . Then we let  $\Omega = (0, 1) \times (-1, 1)$ ,  $E := \{(x, y) \in \Omega : y < w(x)\}$  and  $u(x, y) = \chi_E$ . We claim that  $w \in BV(0, 1)$  while  $\int_{-i}^1 |w''| dx = c2^i \rightarrow \infty$  as  $i \rightarrow \infty$ . Hence,  $u$  is easily seen to have also bounded



variation, while Theorem 1, which we can apply in any subdomain  $(2^{-i}, 1) \times (-1, 1)$ , shows that  $F(u) = +\infty$ .

We denote by  $K(\theta)$  the closed convex set of  $\mathbb{R}^3$  whose support function is  $h(\theta, \cdot)$ . As we have seen [cf. (4)], it is given by

$$K(\theta) = \left\{ \xi \in \mathbb{R}^3 : \xi^x \cdot \underline{\theta} \leq -f^*(\xi^\theta) \right\}.$$

Remark that (1) implies that  $f^*(s) \leq -\sqrt{\gamma^2 - s^2}$  for  $|s| \leq \gamma$ , implying in particular that 0 is in the interior of  $K(\theta)$ .

We then let (following for instance [21, 63])

$$K = \left\{ \varphi \in C_c^0(\Omega \times \mathbb{S}^1; \mathbb{R}^3) : \varphi(x, \theta) \in K(\theta) \ \forall (x, \theta) \in \Omega \times \mathbb{S}^1 \right\}.$$

Then, we claim that for any measure  $\sigma \in \mathcal{M}^1(\Omega \times \mathbb{S}^1; \mathbb{R}^3)$

$$\int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) = \sup_{\varphi \in K} \int_{\Omega \times \mathbb{S}^1} \varphi \cdot \sigma, \quad (8)$$

showing in particular that this integral is a lower-semicontinuous function of  $\sigma$  (for the weak-\* convergence). To prove this claim (which is standard), first observe that if  $\varphi \in K$ , then for all  $\psi^- \in C_c^0(\Omega \times \mathbb{S}^1; \mathbb{R}_+)$  and  $\psi \in C_c^0(\Omega \times \mathbb{S}^1)$ , also  $\varphi - \psi^- \underline{\theta} + \psi \underline{\theta}^\perp \in K$ . One deduces easily that the supremum in (8) is infinite if one of the measures  $(\underline{\theta} \cdot \sigma^x)^-$  or  $\underline{\theta}^\perp \cdot \sigma^x$  does not vanish. If both vanish, it means that  $\sigma = (\lambda \underline{\theta}, \sigma^\theta)$  for some nonnegative measure  $\lambda$ . Hence the supremum becomes

$$\begin{aligned} \sup_{\varphi \in K} \int_{\Omega \times \mathbb{S}^1} (\varphi^x \cdot \underline{\theta}) \lambda + \varphi^\theta \sigma^\theta &= \sup_{\psi + f^*(\varphi^\theta) \leq 0} \int_{\Omega \times \mathbb{S}^1} \psi \lambda + \varphi^\theta \sigma^\theta \\ &= \sup_{\varphi^\theta} \int_{\Omega \times \mathbb{S}^1} \varphi^\theta \sigma^\theta - f^*(\varphi^\theta) \lambda, \end{aligned}$$

and (8) is then deduced in a standard way: in practice, one needs to build a Besicovitch covering of  $\Omega \times \mathbb{S}^1$  with respect to the measure  $|\sigma|$  by balls where  $\lambda, \sigma^\theta$  are “almost constant”, and then choose for  $\varphi$  the “right function” in each ball of the covering. These are then multiplied by a smooth cut-off in each ball to vanish on the boundary; we use here the fact that  $0 \in K(\theta)$ , so that the product of a function  $\varphi \in K$  with a smooth cut-off (with values in  $[0, 1]$ ) still belongs to  $K$ .

We can deduce that also  $F$  is lower semicontinuous: consider a sequence  $(u_n)$  with  $F(u_n) \leq c < \infty$ . Then  $u_n$  is bounded in  $BV$  and converges (up to a constant and a subsequence) to some  $u$  in  $L^1(\Omega)$ .

If  $\sigma_n$  reaches the value of  $F(u_n)$  up to  $1/n$ , one has that

$$\int_{\Omega \times \mathbb{S}^1} |\sigma_n| \leq c < \infty$$

so that  $\sigma_n \xrightarrow{*} \sigma$  (up to a subsequence), as measures, and obviously  $\operatorname{div} \sigma = 0$ . Clearly, also (7) passes to the limit. Hence, by lower semicontinuity,

$$F(u) \leq \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) \leq \liminf_n \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma_n) = \liminf_n F(u_n).$$

This shows that  $F$  defines a convex, lower semicontinuous functional on  $BV(\Omega)$ . Our main theoretical result is the following theorem, which shows that if the argument  $u$  is the characteristic function of a smooth enough set, then  $F(u)$  coincides, as expected, with a curvature-dependent energy of the boundary of the set  $\{u = 1\}$ . The proof of this result is postponed to Sect. 5.

**Theorem 1** *Let  $E \subset \Omega$  be a set with  $C^2$  boundary. Then*

$$F(\chi_E) = \int_{\partial E \cap \Omega} f(\kappa_E(x)) d\mathcal{H}^1(x). \quad (9)$$

**Remark 2.1** One could hope that  $F$  coincides with the lower semicontinuous envelope of its restriction to  $C^2$  sets, with respect to the  $L^1$  convergence. However, simple examples show that it is not the case. Many examples where it fails are found in [23]. In general, we expect the relaxation  $F$  to be strictly below the relaxations found in the literature since [12], based on the  $(L^1)$  approximation of sets with smooth sets. In particular, in the case of [12, Fig. 4.2], our relaxation will certainly be below twice the energy reached by the pattern in [12, Fig. 1.2], while the energy reached in [12, Fig. 4.2] is strictly larger.

**Remark 2.2** If  $f$  has growth one and is even, one expects that the same result holds for piecewise  $C^2$  sets.

**Remark 2.3** We believe that the proof below could be extended to show that if  $u \in BV(\Omega)$  is a function with  $C^2$  level sets, and such that there exists a continuous function  $\kappa(x)$  which coincides with the curvature of  $\partial\{u > s\}$  for all level  $s \in \mathbb{R}$ , then one should have

$$F(u) = \int_{\mathbb{R}} \int_{\partial\{u>s\}} f(\kappa_{\partial\{u>s\}}(x)) d\mathcal{H}^1(x) = \int_{\Omega} f(\kappa(x)) d|Du|(x). \quad (10)$$

Hence in that case our functional would coincide with Masnou and Morel's [53]. A typical example of a function  $u$  for which such a function  $\kappa(x)$  exists is (in a convex domain  $\Omega$ , see [54]) a solution of the so-called ‘‘Rudin–Osher–Fatemi’’ functional

$$\min_u \int_{\Omega} |Du| + \frac{\lambda}{2} \int_{\Omega} (u(x) - f(x))^2 dx \quad (11)$$

for some  $\lambda > 0$ , in case  $f$  is continuous. Then, also  $u$  is continuous and  $\kappa(x)$  is given by  $\lambda(u - f)$ . In that case, in addition,  $\kappa(x)$  is also the curvature of the level sets of  $u' = h(u)$  for any nondecreasing function  $h$  such that  $h(u)$  is still in  $BV(\Omega)$ .

### 3 Some properties of the functional

#### 3.1 Approximation by smooth functions

**Proposition 3.1** *Assume  $\Omega$  is a bounded convex set. Then for any  $u \in L^1(\Omega)$  with  $F(u) < \infty$ , there exists  $(u_n)_n$  a sequence of functions with  $u_n \in C^\infty(\overline{\Omega})$ , which converge to  $u$  in  $L^1(\Omega)$ , and such that*

$$\lim_{n \rightarrow \infty} F(u_n) = F(u).$$

**Remark 3.2** We believe that, upon replacing  $C^\infty(\overline{\Omega})$  with  $C^\infty(\Omega)$ , the result should be true in any domain.

**Proof** We assume, without loss of generality, that 0 is in the interior of  $\Omega$ . It follows that for any  $t < 1$ ,  $t\Omega \subset \subset \Omega$ .

Consider  $\rho$  a rotationally symmetric mollifier in  $\mathbb{R}^2$  ( $\rho \in C_c^\infty(B_1^2; \mathbb{R}_+)$ ,  $\int \rho = 1$ , and we let  $\rho_\varepsilon = (1/\varepsilon^2)\rho(x/\varepsilon)$ , where here  $B_1^d$  is the unit  $d$ -dimensional ball). For  $\varepsilon > 0$ , we consider the largest  $t_\varepsilon < 1$  such that  $\{x \in \Omega : \text{dist}(x, \partial\Omega) > \varepsilon\} \supseteq t_\varepsilon\Omega$ . We then define  $\sigma_\varepsilon$  as the measure on  $\Omega \times \mathbb{S}^1$  given by

$$\int_{\Omega \times \mathbb{S}^1} \psi \cdot \sigma_\varepsilon = \int_{\Omega \times \mathbb{S}^1} \psi * \rho_\varepsilon(t_\varepsilon x, \theta) d\sigma(x),$$

where the convolution is only in the  $x$ -variable:

$$\psi * \rho_\varepsilon(x, \theta) = \int_{B_1^2} \psi(x - \varepsilon z, \theta) \rho(z) dz$$

and is well defined for  $x \in t_\varepsilon\Omega$ . It is clear that  $\sigma_\varepsilon$  still has free divergence, moreover, if  $\varphi \in C_c(\Omega; \mathbb{R}^2)$ ,

$$\begin{aligned} \int_{\Omega \times \mathbb{S}^1} \varphi^\perp \cdot \sigma_\varepsilon^x &= \int_{\Omega \times \mathbb{S}^1} \rho_\varepsilon * \varphi^\perp(t_\varepsilon x) \cdot \sigma_\varepsilon^x(x) \\ &= \int_{\Omega} \rho_\varepsilon * \varphi(t_\varepsilon x) \cdot Du \\ &= - \int_{\Omega} \int_{B_1^2} t_\varepsilon u(x) \rho(z) \text{div} \varphi(t_\varepsilon x - \varepsilon z) dz dx \\ &= - \frac{1}{t_\varepsilon} \int_{\Omega} \int_{B_1^2} u((y + \varepsilon z)/t_\varepsilon) \rho(z) \text{div} \varphi(y) dz dy \\ &= - \frac{1}{t_\varepsilon} \int_{\Omega} (\rho_{\varepsilon/t_\varepsilon} * u)(y/t_\varepsilon) \text{div} \varphi(y) dy. \end{aligned}$$

Hence,  $\sigma_\varepsilon$  is admissible for the function  $u_\varepsilon : y \mapsto \rho_{\varepsilon/t_\varepsilon} * u(y/t_\varepsilon)/t_\varepsilon$ , which clearly goes to  $u$  in  $L^1(\Omega)$  (or  $BV(\Omega)$  weakly-\*) as  $\varepsilon \rightarrow 0$ . Moreover,  $u_\varepsilon \in C^\infty(\Omega)$  for all  $\varepsilon > 0$ . Eventually, one has that for any  $\varphi \in K$ ,

$$\int_{\Omega \times \mathbb{S}^1} \varphi \cdot \sigma_\varepsilon = \int_{\Omega \times \mathbb{S}^1} \varphi * \rho_\varepsilon(t_\varepsilon x, \theta) \cdot \sigma \leq \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) = F(u)$$

by observing that since the convolution is only in the  $x$  variable,  $\varphi * \rho_\varepsilon(t_\varepsilon x, \theta) \in K(\theta)$  for all  $(x, \theta)$  when  $\varphi \in K$ . Thanks to (8), it follows that

$$\int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma_\varepsilon) \leq F(u).$$

Hence  $F(u_\varepsilon) \leq F(u)$ , and the proposition is proved.  $\square$

### 3.2 Dual representation

We can show the following representation for  $F$ , which in particular establishes that  $F$  coincides exactly with the (more complicated) relaxation  $R_{\text{el}}^{**}$  defined in [23].

**Proposition 3.3** *Assume  $\partial\Omega$  is connected. Then the functional (6) can also be represented in the following form*

$$F(u) = \sup \left\{ \int_{\Omega} \psi \cdot Du^\perp : \psi \in C_c^0(\Omega; \mathbb{R}^2), \right. \\ \left. \exists \varphi \in C_c^1(\Omega \times \mathbb{S}^1), \underline{\theta} \cdot (\nabla_x \varphi + \psi) + f^*(\partial_\theta \varphi) \leq 0 \right\}. \quad (12)$$

*If  $\partial\Omega$  is not connected, then same formula holds, however the functions  $\varphi$  should be in  $C^1(\Omega \times \mathbb{S}^1)$ , with  $\nabla \varphi \in C_c^0(\Omega \times \mathbb{S}^1)$ .*

**Corollary 3.4** *The functional  $F(u)$  coincides with the relaxation of the Elastica energy  $R_{\text{el}}^{**}(u)$  defined in [23].*

Indeed, formula (12) is the same as (35) in [23, Theorem 13]. The primal formulation (6) we introduce here is simpler than the original formulation in [23], though, as the “ $\kappa$ ” variable is implicit in our formulation (it comes naturally as the derivative of the orientation) and does not need to be lifted.

**Proof of Proposition 3.3** We use the standard perturbation approach to duality [35], in the duality  $(C_0^0(\Omega \times \mathbb{S}^1; \mathbb{R}^3), \mathcal{M}^1(\Omega \times \mathbb{S}^1; \mathbb{R}^3))$ , where  $C_0^0$  denotes the  $L^\infty$ -closure of the functions with compact support (hence, the functions which vanish on the boundary) and  $\mathcal{M}^1$  the totally bounded vector-valued Radon measures. Defining, for  $p \in \mathcal{M}^1(\Omega \times \mathbb{S}^1; \mathbb{R}^3)$ , the function

$$G(p) = \inf \left\{ \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma + p) : \operatorname{div} \sigma = 0, \int_{\mathbb{S}^1} \sigma^x = Du^\perp \right\} \quad (13)$$

so that  $F(u) = G(0)$ , we observe that  $G$  is lower semicontinuous. Indeed if  $p_n \xrightarrow{*} p$  and  $\sigma_n$  is a minimizer in the definition (13) of  $G(p_n)$ , as before thanks to (5),  $\sigma_n + p_n$  is

bounded and up to a subsequence, converges to a measure  $\sigma + p$  where  $\sigma$  is admissible in (13). It follows that

$$G(p) \leq \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma + p) \leq \liminf_n \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma_n + p_n) = \liminf_n G(p_n).$$

In particular, we deduce that  $G = G^{**}$  so that

$$F(u) = G(0) = \sup_{\eta \in C_0^0(\Omega \times \mathbb{S}^1; \mathbb{R}^3)} -G^*(\eta).$$

It remains to compute  $G^*(\eta)$ , to show that (12) follows:

$$\begin{aligned} G^*(\eta) &= \sup_{p, \sigma} \int_{\Omega \times \mathbb{S}^1} \eta \cdot p - h(\theta, \sigma + p) \\ &= \sup_{\sigma} - \int_{\Omega \times \mathbb{S}^1} \eta \cdot \sigma + \sup_p \eta \cdot (\sigma + p) - h(\theta, \sigma + p) \end{aligned}$$

and we find that  $G^*(\eta)$  is infinite unless  $\underline{\theta} \cdot \eta^x + f^*(\eta^\theta) \leq 0$  everywhere in  $\Omega \times \mathbb{S}^1$ . In this case, we let now  $\psi(x) = \frac{1}{2\pi} \int_{\mathbb{S}^1} \eta^x(x, \theta) d\theta$ . Recalling that the sup is over the  $\sigma$  admissible, we find that

$$G^*(\eta) = \sup_{\sigma} - \int_{\Omega \times \mathbb{S}^1} (\eta - \psi) \cdot \sigma - \int_{\Omega} \psi \cdot Du^\perp. \quad (14)$$

Now, given any  $\sigma'$  with  $\operatorname{div} \sigma' = 0$ , we can define  $u'$  such that  $\int_{\mathbb{S}^1} \sigma' = Du'^\perp$  and let  $\sigma = \sigma' + \frac{1}{2\pi} D(u - u')^\perp \otimes d\theta$ , in which case

$$\begin{aligned} &\int_{\Omega \times \mathbb{S}^1} (\eta - \psi) \cdot \sigma \\ &= \int_{\Omega \times \mathbb{S}^1} (\eta - \psi) \cdot \sigma' + \frac{1}{2\pi} \int_{\mathbb{S}^1} \int_{\Omega} (\eta(x, \theta) - \psi(x)) \cdot D(u - u')^\perp d\theta \\ &= \int_{\Omega \times \mathbb{S}^1} (\eta - \psi) \cdot \sigma', \end{aligned}$$

hence the first supremum in (14) is the same as the sup over all  $\sigma'$  with vanishing divergence. Hence, it is zero or  $+\infty$ , depending on whether  $\eta - \psi$  is a gradient or not. We find eventually  $G^*(\eta)$  is finite only when there exists  $\varphi \in C^1(\Omega \times \mathbb{S}^1)$  with  $\nabla \varphi = 0$  on  $\partial\Omega \times \mathbb{S}^1$ , such that  $\eta = \psi + \nabla \varphi$ . The proposition follows.  $\square$

**Remark 3.5** Consider  $(\psi, \varphi)$  compactly supported continuous functions as in (12). If  $t \in (0, 1)$ , observe that  $(\psi_t, \varphi_t) := (t\psi, t\varphi)$  satisfies

$$\begin{aligned} &\underline{\theta} \cdot (\nabla_x \varphi_t + \psi_t) + f^*(\partial_\theta \varphi_t) \\ &\leq t (\underline{\theta} \cdot (\nabla_x \varphi + \psi) + f^*(\partial_\theta \varphi)) + (1 - t)f^*(0) \leq -(1 - t)\gamma < 0. \end{aligned}$$

Let  $\rho \in C_c^\infty(B_1^3)$  be a smooth mollifier,  $\rho_\varepsilon(x, \theta) = \varepsilon^{-3} \rho(x/\varepsilon, \theta/\varepsilon)$  and

$$\psi_t^\varepsilon = \psi_t * \rho_\varepsilon = \int_{B_1^3} \psi_t(x - \varepsilon z) \rho(z, \theta) dz d\theta, \quad \varphi_t^\varepsilon = \varphi_t * \rho_\varepsilon.$$

This is well defined if  $\varepsilon$  is small enough, as the functions have compact support. We have that, for  $(x, \theta) \in \Omega \times \mathbb{S}^1$ , thanks to the convexity of  $f^*$ , and denoting as before  $\eta = \psi + \nabla \varphi$  and  $\eta_t = t\eta$ ,

$$\begin{aligned} & \underline{\theta} \cdot (\nabla_x \varphi_t^\varepsilon + \psi_t^\varepsilon) + f^*(\partial_\theta \varphi_t^\varepsilon) \\ & \leq \rho_\varepsilon * (\underline{\theta} \cdot \eta_t + f^*(\partial_\theta \varphi_t)) + \underline{\theta} \cdot \rho_\varepsilon * \eta_t - \rho_\varepsilon * (\underline{\theta} \cdot \eta_t) \\ & \leq -(1-t)\gamma + \int_{B_1^3} \rho(z, \theta') ((\underline{\theta} - (\underline{\theta} - \varepsilon \theta')) \cdot \eta_t(x - \varepsilon z, \theta - \varepsilon \theta')). \end{aligned}$$

Observing that  $\|\underline{\theta} - (\underline{\theta} - \varepsilon \theta')\| \leq \varepsilon |\theta'|$ , we deduce that the above is less than  $-(1-t)\gamma + t\varepsilon \|\eta\|_\infty < 0$  as soon as  $\varepsilon$  is small enough. It follows that the functions  $\psi, \varphi$  in (12) can be assumed to be in  $C_c^\infty$ .

## 4 Numerical experiments

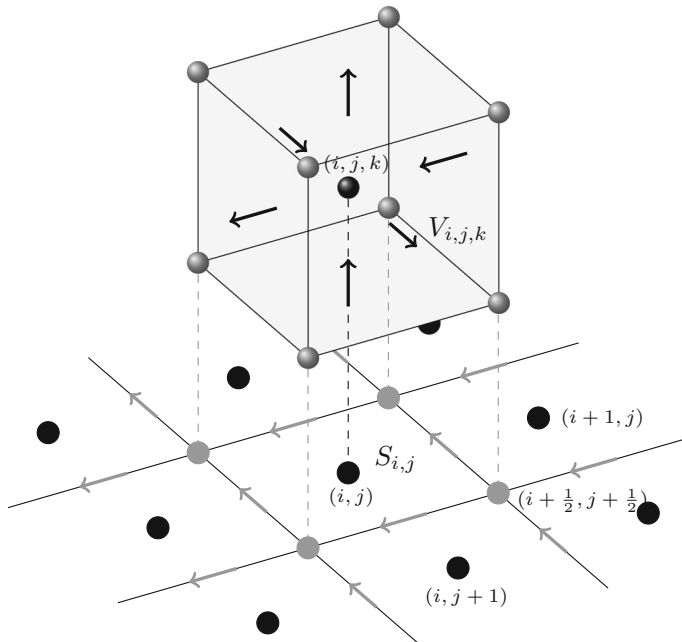
For numerical solution of the proposed model, we need to discretize both the 2D image domain as well as the 3D domain of the roto-translation space. Due to the high anisotropy of the energy, one has to be extremely careful in the choice of the discretization scheme in order to preserve a maximal degree of rotational invariance while keeping the numerical diffusion as low as possible. It turns out that a 2D–3D staggered grid version of an averaged first-order Raviart–Thomas divergence conforming discretization [62] yields the best results. More elaborate discretizations using for example adaptive grids or higher-order approximations will be subject for future study.

### 4.1 Staggered averaged Raviart–Thomas discretization

Figure 2 shows a visualization of our combined 2D–3D staggered grid approach. The first grid is given by a 2D grid of pixels discretizing the image domain  $\Omega$ . The second grid is given by a 3D grid of volumes which discretize the roto-translation space  $\Omega \times \mathbb{S}^1$ . In the visualization we show a few image pixels of the 2D grid (gray nodes) and one cube of the corresponding 3D grid.

First, we start by describing the discretization of the image domain. We assume here that  $\Omega$  is a square or rectangle, or in general a convex domain, which we will discretize at a scale  $\delta_x > 0$ . We will restrict ourselves to the case of a rectangle to simplify the notation. Then, we let  $\Omega_{\delta_x}$  denote the discretized image domain which is given by

$$\Omega_{\delta_x} = \text{int} \left( \bigcup \{ S_{i,j} : (i, j) \in \mathbb{Z}^2, S_{i,j} \subset \Omega \} \right),$$



**Fig. 2** Spatial discretization of the image domain  $\Omega$  and the domain of the RT space  $\Omega \times \mathbb{S}^1$ . The bottom grid represents the 2D image, and the image intensities  $u$  are stored at the vertices of the squares  $S_{i,j}$ . The gray arrows indicate differences between the pixels. The volume  $V_{i,j,k}$  represents one element of the corresponding discretization of the RT space. The black arrows through the faces of the volume represent the 3D vector field  $\sigma$

where

$$S_{i,j} = \left[ \left( i - \frac{1}{2} \right) \delta_x, \left( i + \frac{1}{2} \right) \delta_x \right) \times \left[ \left( j - \frac{1}{2} \right) \delta_x, \left( j + \frac{1}{2} \right) \delta_x \right), \quad (15)$$

denotes the square centered around the index  $\mathbf{i} = (i, j)$ . The intensity values of the discrete image are stored at the vertices half-grid points (such as the four vertices of  $S_{i,j}$ ), and for notational simplicity we introduce the following index set:

$$\mathcal{I} = \left\{ \left( i - \frac{1}{2}, j - \frac{1}{2} \right) : 1 \leq i \leq N_1, 1 \leq j \leq N_2 \right\},$$

with  $N_1, N_2 > 0$ . Hence, the discrete image  $u \in \mathbb{R}^{\mathcal{I}}$  consists of  $N_1 \times N_2$  discrete pixels with values  $(u_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}}$ . We will later also need the following index sets:

$$\begin{aligned} \mathcal{I}^1 &= \left\{ \left( i - \frac{1}{2}, j \right) : 1 \leq i \leq N_1, 1 \leq j < N_2 \right\}, \\ \mathcal{I}^2 &= \left\{ \left( i, j - \frac{1}{2} \right) : 1 < i \leq N_1, 1 \leq j \leq N_2 \right\}, \end{aligned}$$

which index the middle points of the edges between two adjacent pixels along the first and second spatial dimensions.

Next, we let  $\delta_\theta > 0$  be the discretization step for the angular variable. We must assume that  $\delta_\theta = 2\pi/N_\theta$  with  $N_\theta > 0$ . Hence  $\mathbb{S}^1$  is discretized by  $N_\theta$  intervals  $[(k - \frac{1}{2})\delta_\theta, (k + \frac{1}{2})\delta_\theta)$  for  $1 \leq k \leq N_\theta$ . We then denote by  $\Gamma_\delta$  with  $\delta = (\delta_x, \delta_\theta)$  the discretized 3D roto-translation space. It is given by

$$\Gamma_\delta = \text{int} \left( \bigcup \left\{ V_{i,j,k} : (i, j, k) \in \mathbb{Z}^3, V_{i,j,k} \subset \Omega \times \mathbb{S}^1 \right\} \right),$$

where

$$V_{i,j,k} = S_{i,j} \times \left[ (k - \frac{1}{2})\delta_\theta, (k + \frac{1}{2})\delta_\theta \right) \quad (16)$$

denotes the volume centered around the 3D index  $\mathbf{j} = (i, j, k)$ . We associate an index set  $\mathcal{J}$  of center locations of the volumes  $V_{i,j,k}$ :

$$\mathcal{J} = \{(i, j, k) : 1 \leq i < N_1, 1 \leq j < N_2, 1 \leq k \leq N_\theta\}. \quad (17)$$

We shall also introduce the following three index sets:

$$\begin{aligned} \mathcal{J}^1 &= \{(i - \frac{1}{2}, j, k) : 1 \leq i \leq N_1, 1 \leq j < N_2, 1 \leq k \leq N_\theta\}, \\ \mathcal{J}^2 &= \{(i, j - \frac{1}{2}, k) : 1 \leq i < N_1, 1 \leq j \leq N_2, 1 \leq k \leq N_\theta\}, \\ \mathcal{J}^\theta &= \{(i, j, k - \frac{1}{2}) : 1 \leq i < N_1, 1 \leq j < N_2, 1 \leq k \leq N_\theta\}, \end{aligned}$$

which we will use to index the facets of the volumes. Observe that the 3D points  $((i - \frac{1}{2})\delta_x, j\delta_x, k\delta_\theta)$  for all  $(i - \frac{1}{2}, j, k) \in \mathcal{J}^1$  correspond to the middle points of the facets

$$\mathcal{F}_{(i-\frac{1}{2},j,k)}^1 = \{(i - \frac{1}{2})\delta_x\} \times [(j - \frac{1}{2})\delta_x, (j + \frac{1}{2})\delta_x) \times [(k - \frac{1}{2})\delta_\theta, (k + \frac{1}{2})\delta_\theta), \quad (18)$$

which are orthogonal to the first spatial dimension. Similarly, the 3D points  $(i\delta_x, (j - \frac{1}{2})\delta_x, k\delta_\theta)$  for all  $(i, j - \frac{1}{2}, k) \in \mathcal{J}^2$  correspond to the middle points of the facets

$$\mathcal{F}_{(i,j-\frac{1}{2},k)}^2 = [(i - \frac{1}{2})\delta_x, (i + \frac{1}{2})\delta_x) \times \{(j - \frac{1}{2})\delta_x\} \times [(k - \frac{1}{2})\delta_\theta, (k + \frac{1}{2})\delta_\theta), \quad (19)$$

which are orthogonal to the second spatial dimension. Finally, the 3D points  $(i\delta_x, j\delta_x, (k - \frac{1}{2})\delta_\theta)$  for all  $(i, j, k - \frac{1}{2}) \in \mathcal{J}^\theta$  correspond to the middle points of the facets

$$\mathcal{F}_{(i,j,k-\frac{1}{2})}^\theta = [(i - \frac{1}{2})\delta_x, (i + \frac{1}{2})\delta_x) \times [(j - \frac{1}{2})\delta_x, (j + \frac{1}{2})\delta_x) \times \{(k - \frac{1}{2})\delta_\theta\}, \quad (20)$$

which are orthogonal to the angular dimension.

Now, we consider a discrete vector field  $\sigma = (\sigma^1, \sigma^2, \sigma^\theta)$  with  $\sigma^1 \in \mathbb{R}^{\mathcal{J}^1}$ ,  $\sigma^2 \in \mathbb{R}^{\mathcal{J}^2}$ , and  $\sigma^\theta \in \mathbb{R}^{\mathcal{J}^\theta}$ . The values  $(\sigma_{\mathbf{j}}^1)_{\mathbf{j} \in \mathcal{J}^1}$ ,  $(\sigma_{\mathbf{j}}^2)_{\mathbf{j} \in \mathcal{J}^2}$ ,  $(\sigma_{\mathbf{j}}^\theta)_{\mathbf{j} \in \mathcal{J}^\theta}$  define the average fluxes through the facets  $(\mathcal{F}_{\mathbf{j}}^1)_{\mathbf{j} \in \mathcal{J}^1}$ ,  $(\mathcal{F}_{\mathbf{j}}^2)_{\mathbf{j} \in \mathcal{J}^2}$  and  $(\mathcal{F}_{\mathbf{j}}^\theta)_{\mathbf{j} \in \mathcal{J}^\theta}$ , respectively.



The discrete vector field  $\sigma$  can also be seen as a Raviart–Thomas vector field [62], defined everywhere in  $\Gamma_{\delta_x, \delta_\theta}$ . It is obtained by an affine extension:

$$\sigma_{RT}(x_1, x_2, \theta) = \left( \sum_{\mathbf{j}=(i-\frac{1}{2}, j, k) \in \mathcal{J}^1} \sigma_{\mathbf{j}}^1 \Delta \left( \frac{x_1 - (i - \frac{1}{2})\delta_x}{\delta_x} \right), \right. \\ \left. \sum_{\mathbf{j}=(i, j-\frac{1}{2}, k) \in \mathcal{J}^2} \sigma_{\mathbf{j}}^2 \Delta \left( \frac{x_2 - (j - \frac{1}{2})\delta_x}{\delta_x} \right), \sum_{\mathbf{j}=(i, j, k-\frac{1}{2}) \in \mathcal{J}^\theta} \sigma_{\mathbf{j}}^\theta \Delta \left( \frac{\theta - (k - \frac{1}{2})\delta_\theta}{\delta_\theta} \right) \right), \quad (21)$$

where  $\Delta(t) = \max(0, 1 - |t|)$  is the usual linear interpolation kernel.

It is well known that such a Raviart–Thomas field has a divergence (defined everywhere in the distributional sense) which is obtained in each volume by summing the finite differences of the average fluxes through each pair of opposite facets. In our model, the discrete field  $\sigma$  is constrained to be divergence free, hence for each volume  $V_{\mathbf{j}}$ , we must impose that:

$$\frac{\sigma_{i+\frac{1}{2}, j, k}^1 - \sigma_{i-\frac{1}{2}, j, k}^1}{\delta_x} + \frac{\sigma_{i, j+\frac{1}{2}, k}^2 - \sigma_{i, j-\frac{1}{2}, k}^2}{\delta_x} + \frac{\sigma_{i, j, k+\frac{1}{2}}^\theta - \sigma_{i, j, k-\frac{1}{2}}^\theta}{\delta_\theta} = 0. \quad (22)$$

The roto-translation space is periodic in the  $\theta$  direction and hence, whenever  $k = N_\theta$ , we identify the index  $N_\theta + \frac{1}{2}$  with the index  $\frac{1}{2}$  in the computation of the last term of the discrete divergence.

To simplify our presentation, we introduce a discrete divergence operator  $\mathcal{D} : \mathbb{R}^{\mathcal{J}^1 \cup \mathcal{J}^2 \cup \mathcal{J}^\theta} \rightarrow \mathbb{R}^{\mathcal{J}}$ , so that the divergence free constraint (22) can be compactly written as

$$\mathcal{D}\sigma = 0.$$

Next, to implement a discrete version of the consistency condition we observe that averaging (in the direction of  $\theta$ ) the first two components of the field  $\sigma$  yields a rotated (by  $\pi/2$ ) gradient which should locally coincide with the finite differences of the discrete image  $u$ . This yields the following discrete consistency condition:

$$\begin{cases} \delta_\theta \sum_k \sigma_{i+\frac{1}{2}, j, k}^1 = \frac{1}{\delta_x} (u_{i+\frac{1}{2}, j+\frac{1}{2}} - u_{i+\frac{1}{2}, j-\frac{1}{2}}) \\ \delta_\theta \sum_k \sigma_{i, j+\frac{1}{2}, k}^2 = -\frac{1}{\delta_x} (u_{i+\frac{1}{2}, j+\frac{1}{2}} - u_{i-\frac{1}{2}, j+\frac{1}{2}}). \end{cases} \quad (23)$$

We introduce a projection operator  $\mathcal{P} : \mathbb{R}^{\mathcal{J}^1 \cup \mathcal{J}^2} \rightarrow \mathbb{R}^{\mathcal{I}^1 \cup \mathcal{I}^2}$  and a discrete (rotated) gradient operator  $\mathcal{G} : \mathbb{R}^{\mathcal{I}} \rightarrow \mathbb{R}^{\mathcal{I}^1 \cup \mathcal{I}^2}$  such that we can write the above discrete compatibility condition as

$$\mathcal{P}\sigma = \mathcal{G}u.$$

In order to approximate the continuous energy (6) with our discrete Raviart–Thomas field  $\sigma_{RT}$  we can use different types of quadrature rules. After various attempts, we found out that simply summing the values at the center points of the volumes  $V_{i,j,k}$  provides the highest flexibility for the discrete field  $\sigma$  to concentrate on thin lines, and, in turn, yields the most faithful numerical results.

Letting thus  $\sigma_{i,j,k} = \sigma_{RT}(i\delta_x, j\delta_x, k\delta_\theta)$ , we have that this discrete field is obtained by the following formula:

$$\begin{aligned} \hat{\sigma}_{i,j,k} &= (\hat{\sigma}_{i,j,k}^1, \hat{\sigma}_{i,j,k}^2, \hat{\sigma}_{i,j,k}^\theta) \\ &= \frac{1}{2} \left( \sigma_{i+\frac{1}{2},j,k}^1 + \sigma_{i-\frac{1}{2},j,k}^1, \sigma_{i,j+\frac{1}{2},k}^2 + \sigma_{i,j-\frac{1}{2},k}^2, \sigma_{i,j,k+\frac{1}{2}}^\theta + \sigma_{i,j,k-\frac{1}{2}}^\theta \right). \end{aligned} \quad (24)$$

We again introduce an operator  $\mathcal{A} : \mathbb{R}^{\mathcal{J}^1 \cup \mathcal{J}^2 \cup \mathcal{J}^\theta} \rightarrow (\mathbb{R}^3)^{\mathcal{J}}$  such that the above averaging operation can be written

$$\hat{\sigma} = \mathcal{A}\sigma.$$

Now, using the volume centered quadrature rule, the discrete energy is

$$F_\delta(u) = \min_{\sigma} \left\{ \delta_x^2 \delta_\theta \sum_{\mathbf{j}=(i,j,k) \in \mathcal{J}} h(k\delta_\theta, (\mathcal{A}\sigma)_{\mathbf{j}}) : \mathcal{D}\sigma = 0, \mathcal{P}\sigma - \mathcal{G}u = 0 \right\}. \quad (25)$$

The consistency of this energy, as  $\delta := (\delta_x, \delta_\theta) \rightarrow 0$ , with the continuous energy  $F$  defined in (6) is studied in “Appendix A”.

The regularization function  $h(\theta, p)$ , where  $\theta \in \mathbb{S}^1$  and  $p = (p^x, p^\theta) \in \mathbb{R}^3$  is defined as in (3), and the function  $f(t)$ , which appears in its definition will be one of the following classical examples of convex functions:

$$f_1(t) = 1 + \alpha|t|, \quad (\text{TAC})$$

$$f_2(t) = \sqrt{1 + \alpha^2|t|^2}, \quad (\text{TRV})$$

$$f_3(t) = 1 + \alpha^2|t|^2. \quad (\text{TSC})$$

In all examples, the regularizing function is a combination of length and curvature regularization and the parameter  $\alpha > 0$  can be used to adjust the influence of the curvature regularization.

The first function,  $f_1$  is the sum of length and absolute curvature, hence we call the corresponding regularizer “total absolute curvature” (TAC). One of its main features is that it allows for sharp corners in the level sets of the image. Interestingly, since integrating the absolute curvature along the boundary of a shape is constantly  $2\pi$  for all convex shapes (and in general, is scale independent) then the main effect of the curvature term in this energy is to penalize non-convex shapes, regardless of their size.

The function  $f_2$  combines length and curvature through an Euclidean metric and hence, it corresponds to the total variation of the lifted curve in the RT space, hence we consequently denote this regularizer “total roto-translational variation” (TRV). For relatively small curvature, it favors smooth shapes, but it also allows sharp discontinuities.

Function  $f_3$  penalizes squared curvature plus length and hence is equivalent to the Elastica energy. We call this regularizer “total squared curvature” (TSC). It is well-known that while length regularization favors smaller shapes, quadratic curvature favors larger shapes. Hence, the interplay between length and curvature regularization removes the shrinkage bias and leads to smooth shapes.

Based on the “basis” functions  $f_l$ ,  $l = 1, 2, 3$ , we obtain the following convex regularization functions in the RT space.

$$\text{(TAC)} \quad h_1(\theta, p) = \begin{cases} |p^x| + \alpha |p^\theta| & \text{if } p^x = \underline{\theta}s, \quad s \geq 0 \\ +\infty & \text{else,} \end{cases} \quad (26)$$

$$\text{(TRV)} \quad h_2(\theta, p) = \begin{cases} \sqrt{|p^x|^2 + \alpha^2 |p^\theta|^2} & \text{if } p^x = \underline{\theta}s, \quad s \geq 0 \\ +\infty & \text{else,} \end{cases} \quad (27)$$

$$\text{(TSC)} \quad h_3(\theta, p) = \begin{cases} |p^x| + \alpha^2 \frac{|p^\theta|^2}{|p^x|} & \text{if } p^x = \underline{\theta}s, \quad s > 0 \\ +\infty & \text{else.} \end{cases} \quad (28)$$

Observe in particular, that the above definitions already properly take into account the correct values of the function  $h_l$ ,  $l = 1, 2, 3$  for  $p^x = 0$ .

In the following sections we will apply the proposed curvature based regularization functions to a variety of image- and shape processing problems. For this, we consider generic optimization problems of the form

$$\min_u F_\delta(u) + G(u), \quad (29)$$

where  $G(u)$  is a convex, lsc. function defining an image-based convex data fidelity term, possibly dependent on the pixel location.

## 4.2 Primal-dual optimization

In this section, we show how to compute a minimizer of the non-smooth convex optimization problem (29). For notational simplicity, we first divide the whole objective function (29) by  $\delta_x^2 \delta_\theta$  and assume that the remaining factors are absorbed by the data fitting term.

In order to solve the constrained optimization problem (29), we consider its Lagrangian (saddle-point) formulation:

$$\min_{u, \sigma} \max_{\phi, \psi, \xi} \sum_{\mathbf{j} \in \mathcal{J}} (\mathcal{A}\sigma)_{\mathbf{j}} \cdot \xi_{\mathbf{j}} - \sum_{\mathbf{j}=(i,j,k) \in \mathcal{J}} h^*(k\delta_\theta, \xi_{\mathbf{j}}) + G(u)$$

$$+ \sum_{\mathbf{j} \in \mathcal{J}} (\mathcal{D}\sigma)_{\mathbf{j}} \phi_{\mathbf{j}} + \sum_{\mathbf{i} \in \mathcal{I}^1 \cup \mathcal{I}^2} ((\mathcal{P}\sigma)_{\mathbf{i}} - (\mathcal{G}u)_{\mathbf{i}}) \psi_{\mathbf{i}}, \quad (30)$$

where  $\phi \in \mathbb{R}^{\mathcal{J}}$ ,  $\xi = (\xi^1, \xi^2, \xi^\theta) \in (\mathbb{R}^3)^{\mathcal{J}}$  and  $\psi = (\psi^1, \psi^2) \in \mathbb{R}^{\mathcal{I}^1 \cup \mathcal{I}^2}$  with  $\psi^1 \in \mathbb{R}^{\mathcal{I}^1}$  and  $\psi^2 \in \mathbb{R}^{\mathcal{I}^2}$  are the dual variables (Lagrange multipliers or discrete test functions). The function  $h^*$  denotes the convex conjugate of the function  $h$ . Recall that  $h$  is the support function of the convex set

$$H(\theta) = \{\xi = (\xi^x, \xi^\theta) \in \mathbb{R}^3 : \xi^x \cdot \underline{\theta} \leq -f^*(\xi^\theta)\}, \quad \underline{\theta} = (\cos \theta, \sin \theta),$$

where  $f^*$  denotes the convex conjugate of  $f$ . Hence, the convex conjugate  $h^*$  is simply the indicator function of the set  $H$ :

$$h^*(\theta, \xi) = \begin{cases} 0 & \text{if } \xi \in H(\theta) \\ \infty & \text{else.} \end{cases}$$

The Problem (30) is a saddle-point problem which is separable in the non-linear terms and hence falls into the class of problems that can be solved by the first-order-primal-dual algorithm with diagonal preconditioning and overrelaxation [24,25,60]. In order to make the algorithm implementable, we need efficient algorithms to compute the projection operators  $\text{proj}_{H(\theta)}$  and proximity operators  $\text{prox}_{\tau G}$ . They are defined as the unique minimizers of the minimization problems:

$$\text{proj}_{H(\theta)}(\eta) = \arg \min_{\xi \in H(\theta)} \frac{1}{2\tau} |\xi - \eta|^2, \quad \forall \theta \in \mathbb{S}^1,$$

and,

$$\text{prox}_{\tau G}(v) = \arg \min_u G(u) + \frac{1}{2\tau} \|u - v\|^2,$$

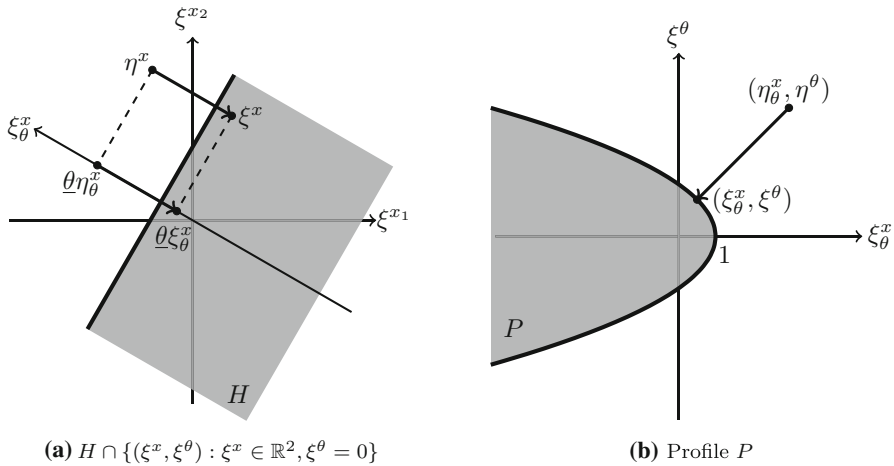
for some  $\tau > 0$ . In what follows, we will detail the projection operator for different instances of the convex sets  $H$  (respectively regularization functions  $h$ ), the proximity maps for  $G$  will be detailed as soon as they are needed in the numerical results.

The general idea for performing the projection  $\xi = \text{proj}_{H(\theta)}(\eta)$  of a point  $\eta = (\eta^x, \eta)$  onto the set  $H(\theta)$  is outlined in Fig. 3. The Figure on the left hand side represents the set  $H$  in the  $\xi^x$  plane, and the Figure on the right hand side shows the “profile” of the set which is obtained by cutting along the  $\xi_\theta^x \times \xi^\theta$  plane.

We let  $\eta_\theta^x = \eta^x \cdot \underline{\theta}$  and define  $(\xi_\theta^x, \xi^\theta) = \text{proj}_P(\eta_\theta^x, \eta^\theta)$  as the projection of the point  $(\eta_\theta^x, \eta^\theta)$  onto the “profile”

$$P = \{(\xi_\theta^x, \xi^\theta) \in \mathbb{R}^2 : \xi_\theta^x \leq -f^*(\xi^\theta)\}. \quad (31)$$

Then, using simple geometric reasoning, we see that the variable  $\xi^x$  can be recovered from  $\xi^\theta$  by



**Fig. 3** Projection of the point  $(\eta^x, \eta^\theta)$  onto the convex sets  $H$ . First, the point  $(\eta_\theta^x, \eta^\theta)$  is projected onto the profile  $P$  which yields  $(\xi_\theta^x, \xi^\theta)$ , then  $\xi^x$  is computed from  $\xi_\theta^x$  via the geometric relationship  $\xi^x = \eta^x - \underline{\theta}(\eta_\theta^x - \xi_\theta^x)$

$$\xi^x = \eta^x - \underline{\theta}(\eta_\theta^x - \xi_\theta^x).$$

It remains to detail the projections onto the three different profiles  $P$ .

- TAC:  $f_1(t) = 1 + \alpha|t|$ : The convex conjugate of the function  $f_1$  is given by

$$f_1^*(s) = \begin{cases} -1 & \text{if } |s| \leq \alpha \\ \infty & \text{else,} \end{cases}$$

and in turn the profile  $P_1$  is given by

$$P_1 = \{(\xi_\theta^x, \xi^\theta) \in \mathbb{R}^2 : \xi_\theta^x \leq 1, |\xi^\theta| \leq \alpha\}$$

It is straightforward that the projection of a point  $(\eta_\theta^x, \eta^\theta)$  onto the profile  $P_1$  can be performed via simple truncation operations

$$\begin{pmatrix} \xi_\theta^x \\ \xi^\theta \end{pmatrix} = \begin{pmatrix} \min(1, \eta_\theta^x) \\ \max(-\alpha, \min(\alpha, \eta^\theta)) \end{pmatrix}.$$

- TRV:  $f_2(t) = \sqrt{1 + \alpha^2 t^2}$ : A simple computation shows that the convex conjugate of  $f_2(t)$  is given

$$f_2^*(s) = \begin{cases} -\sqrt{1 - s^2/\alpha^2} & \text{if } |s| \leq \alpha \\ \infty & \text{else.} \end{cases}$$

Inserting the expression of the convex conjugate into the profile (31) and squaring both sides, we obtain

$$P_2 = \{(\xi_\theta^x, \xi^\theta) \in \mathbb{R}^2 : \max(0, \xi_\theta^x)^2 + (\xi^\theta/\alpha)^2 \leq 1\}.$$

In what follows, we assume that  $\max(0, \eta_\theta^x)^2 + (\eta^\theta/\alpha)^2 > 1$  since otherwise we do not need to project the point. We first treat the case  $\eta_\theta^x \leq 0$ . It is easy to see that in this case the solution of the projected point is given by

$$\begin{pmatrix} \xi_\theta^x \\ \xi^\theta \end{pmatrix} = \begin{pmatrix} \eta_\theta^x \\ \max(-\alpha, \min(\alpha, \eta^\theta)) \end{pmatrix}.$$

In case  $\eta_\theta^x > 0$ , computing the projection of a point  $(\eta_\theta^x, \eta^\theta)$  onto the boundary of  $P_2$  amounts to solve the following equality constrained optimization problem

$$\min_{(\xi_\theta^x)^2 + (\xi^\theta/\alpha)^2 = 1} \frac{1}{2} |\xi_\theta^x - \eta_\theta^x|^2 + \frac{1}{2} |\xi^\theta - \eta^\theta|^2.$$

The Karush–Kuhn–Tucker (KKT) optimality conditions for the above problem are given by

$$\begin{aligned} \xi_\theta^x - \eta_\theta^x + 2\lambda(\xi_\theta^x) &= 0, \\ \xi^\theta - \eta^\theta + 2\lambda\xi^\theta/\alpha^2 &= 0, \\ (\xi_\theta^x)^2 + (\xi^\theta/\alpha)^2 - 1 &= 0, \end{aligned}$$

where  $\lambda > 0$  is the Lagrange multiplier which is positive since the point  $(\eta_\theta^x, \eta^\theta)$  was assumed to be outside  $P_2$ . Combining the first three equations shows that the Lagrange multiplier is computed from the roots of the fourth-order polynomial

$$(\alpha + 2\lambda/\alpha)^2(\eta_\theta^x)^2 + (1 + 2\lambda)^2(\eta^\theta)^2 - (1 + 2\lambda)^2(\alpha + 2\lambda/\alpha)^2 = 0.$$

Let us observe that in case  $\alpha = 1$ , the solution for  $\lambda$  is particularly simple, indeed

$$\lambda = \frac{1}{2} \left( \sqrt{(\eta_\theta^x)^2 + (\eta^\theta)^2} - 1 \right),$$

such that from the first two equations we obtain for the projected point

$$\begin{pmatrix} \xi_\theta^x \\ \xi^\theta \end{pmatrix} = \frac{1}{\sqrt{(\eta_\theta^x)^2 + (\eta^\theta)^2}} \begin{pmatrix} \eta_\theta^x \\ \eta^\theta \end{pmatrix}.$$

In the general case  $\alpha \neq 1$  we compute  $\lambda$  by applying Newton's algorithm to find the (correct) root from the fourth order polynomial. In our experiments it turns out that a large enough initial value, e.g.  $\lambda = 10^3$  provides a good initialization

for Newton's algorithm. Usually, we need less than 5–10 iterations of Newton's algorithm to converge to a solution with feasibility error less than  $10^{-9}$ . From the first two equations of the KKT conditions the expression of the projected point is given by:

$$\begin{pmatrix} \xi_\theta^x \\ \xi^\theta \end{pmatrix} = \begin{pmatrix} \eta_\theta^x / (1 + 2\lambda) \\ \eta^\theta / (1 + 2\lambda / \alpha^2) \end{pmatrix}.$$

- TSC:  $f_3(t) = 1 + \alpha^2 t^2$ : The convex conjugate of  $f_3$  is computed as

$$f_3^*(s) = (s / (2\alpha))^2 - 1,$$

and hence the profile  $P_3$  is given by

$$P_3 = \{(\xi_\theta^x, \xi^\theta) \in \mathbb{R}^2 : \xi_\theta^x + (\xi^\theta / (2\alpha))^2 \leq 1\}.$$

Following the same approach as before a point  $(\eta_\theta^x, \eta^\theta)$  with  $\eta_\theta^x + (\eta^\theta / (2\alpha))^2 > 1$  is projected onto  $P_3$  by solving the KKT optimality conditions

$$\begin{aligned} \xi_\theta^x - \eta_\theta^x + \lambda &= 0, \\ \xi^\theta - \eta^\theta + \lambda \xi^\theta / (2\alpha^2) &= 0, \\ \xi_\theta^x + (\xi^\theta / (2\alpha))^2 - 1 &= 0. \end{aligned}$$

Combining the above equations, the Lagrange multiplier is obtained by finding the correct root from the third-order polynomial

$$(2\alpha^2 + \lambda)^2 (\eta_\theta^x - 1 - \lambda) + (\alpha \eta^\theta)^2 = 0.$$

We again apply Newton's method and observe that for a large enough initial value, e.g.  $\lambda = 10^3$ , Newton's algorithm rapidly converges to the correct root of the polynomial. The projected point is finally given by

$$\begin{pmatrix} \xi_\theta^x \\ \xi^\theta \end{pmatrix} = \begin{pmatrix} \eta_\theta^x - \lambda \\ \eta^\theta / (1 + \lambda / (2\alpha^2)) \end{pmatrix}.$$

### 4.3 Computing a disk

In the first example, we consider the most basic numerical experiment, which is using TSC energy to compute a disk of a given radius. The aim of this experiment is to investigate the quality of our proposed discretization scheme, in particular when using a different number of discrete orientations. Consider a 2D disk  $B(0, r) = \{x \in \mathbb{R}^2 : |x| \leq r\}$  of radius  $r > 0$  centered around the origin. The TSC energy of the boundary of the disk  $B(0, r)$  is given by

$$\int_{\partial B(0,r)} (1 + \alpha^2 \kappa^2) d\mathcal{H}^1 = 2\pi(r + \alpha^2/r).$$

Here our disk will be represented by its characteristic function  $\chi_{B(0,r)}(x)$ . We can force the minimizer of the TSC energy to yield a disk using suitable boundary conditions. For example, we can force at least one point inside the disk to be one and at least one point outside the disk to be zero. The following simple computation shows that the minimizer of the TSC energy will be a disk of radius  $r = \alpha$ , indeed:

$$\begin{aligned} & \int_{\partial B(0,r)} (1 + \alpha^2 \kappa^2) d\mathcal{H}^1 \\ &= \int_{\partial B(0,r)} \left( \alpha^2 \left( \kappa - \frac{1}{\alpha} \right)^2 + 2\alpha\kappa \right) d\mathcal{H}^1 = \int_{\partial B(0,r)} \alpha^2 \left( \kappa - \frac{1}{\alpha} \right)^2 d\mathcal{H}^1 + 4\pi\alpha, \end{aligned}$$

which is minimized if  $\kappa = 1/r = 1/\alpha$  everywhere.

We set up our optimization problem using a 2D grid of  $N_1 \times N_2 = 40 \times 40$  pixels. We will use a different number of discrete orientations  $N_\theta \in \{4, 8, 16, 32, 64\}$  in order to investigate the quality of the approximation depending on  $N_\theta$ . In this and the subsequent experiments, the discretization width of the spatial grid is set to  $\delta_x = 1$  and the discretization width of the angular dimension is set to  $\delta_\theta = 2\pi/N_\theta$ . The aim of this experiment is to compute a disk of radius  $r = 10$  pixels. We therefore define an inpainting domain  $\mathcal{D} \subseteq \mathcal{I}$  forming a band of 10 pixels width around a disk of radius  $r = 10$ . In this domain we minimize the TSC energy and we use the remaining part of the image as boundary condition. The corresponding data term in our optimization problem (29) is given by

$$G(u) = \sum_{\mathbf{i} \in \mathcal{D}} \iota_{u_{\mathbf{i}}^0}(u_{\mathbf{i}}), \quad (32)$$

where  $\iota_C$  denotes the indicator function of the convex set  $C$ . The proximal map for  $G(u)$  is given by

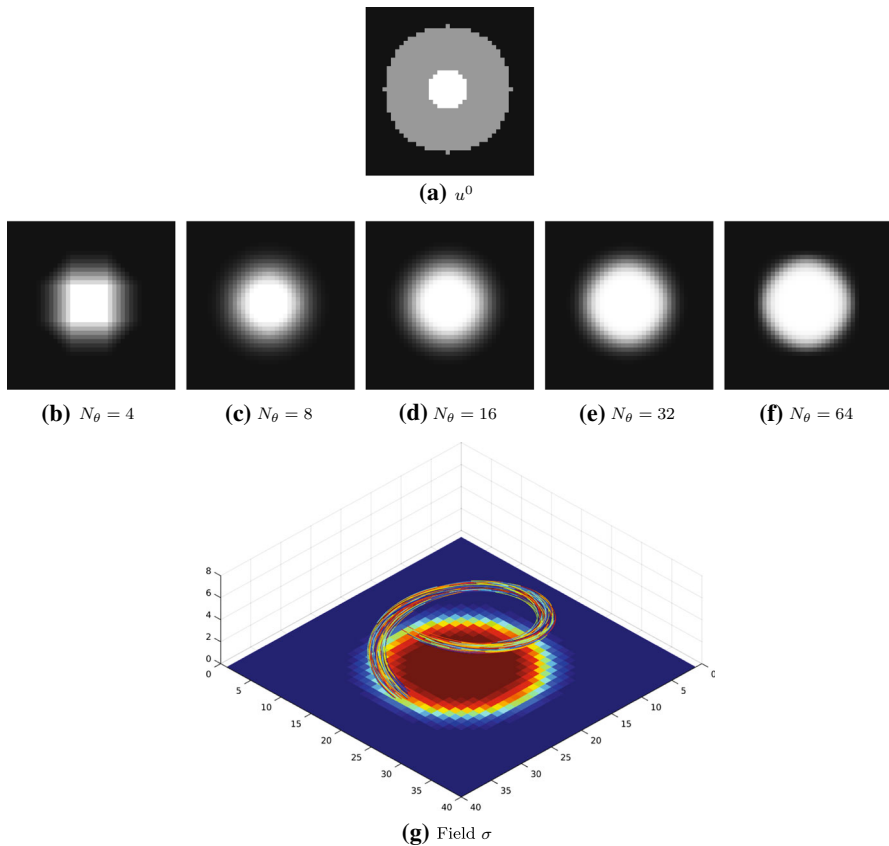
$$u = \text{prox}_{\tau G}(v) \iff u_{\mathbf{i}} = \begin{cases} v_{\mathbf{i}} & \text{if } \mathbf{i} \in \mathcal{D} \\ u_{\mathbf{i}}^0 & \text{else} \end{cases}, \quad \forall \mathbf{i} \in \mathcal{I}. \quad (33)$$

Figure 4a shows the input image  $(u_{\mathbf{i}}^0)_{\mathbf{i} \in \mathcal{I}}$  and the inpainting domain  $\mathcal{D}$  is indicated by the gray area.

In order to quantify the approximation quality of our discretization for a varying number of discrete orientations, we report the total variation (TV), the absolute curvature (AC) and the squared curvature (SC). All three quantities are computed from the averaged field  $\hat{\sigma} = \mathcal{A}\sigma$  using the volume-centered based discrete energy

$$H(\hat{\sigma}) = \delta_x^2 \delta_\theta \sum_{\mathbf{j} \in \mathcal{J}} h(\hat{\sigma}_{\mathbf{j}}),$$





**Fig. 4** Computing a disk by minimizing the TSC energy. We set  $\alpha = 10$  and hence the optimal solution of the inpainting problem is given by a disk of radius  $r = 10$ . **a** The input image  $u^0$  of size  $40 \times 40$  pixels together with its inpainting domain indicated by gray pixels and **b–f** show the computed solution images  $u$  for different numbers of discrete orientations  $N_\theta$ . **g** A “stream-line” plot of the averaged Raviart–Thomas vector field  $\hat{\sigma}$  on top of a false-color version of the corresponding image  $u$

where the function  $h$  is one of the three instances:

$$h_{\text{TV}}(\hat{\sigma}_{\mathbf{j}}) = \sqrt{(\hat{\sigma}_{\mathbf{j}}^1)^2 + (\hat{\sigma}_{\mathbf{j}}^2)^2}, \quad h_{\text{AC}}(\hat{\sigma}_{\mathbf{j}}) = |\hat{\sigma}_{\mathbf{j}}^\theta|, \quad h_{\text{SC}}(\hat{\sigma}_{\mathbf{j}}) = \frac{|\hat{\sigma}_{\mathbf{j}}^\theta|^2}{\sqrt{(\hat{\sigma}_{\mathbf{j}}^1)^2 + (\hat{\sigma}_{\mathbf{j}}^2)^2}}.$$

Table 1 details the values of the discrete energies we obtained for different numbers of discrete orientations  $N_\theta$ . From the results, one can see that a higher number of discrete energies generally leads to a better approximation of the true energy. This is particularly true for the value of the squared curvature which seems to be well approximated only when using a quite high number of discrete orientations. The absolute curvature, however seems to be well approximated even when using only a small number of discrete orientations. The reason for this is that the absolute curvature of a smooth curve is easy to approximate by means of a piecewise linear curve.

**Table 1** Approximating a disk of radius  $r = 10$ 

$N_\theta$	$H_{TV}$ ( $2\pi r \approx 62.8319$ )	$H_{AC}$ ( $2\pi \approx 6.2832$ )	$H_{SC}$ ( $2\pi/r \approx 0.6283$ )
4	60.1063	6.3456	1.7504
8	54.8043	6.2847	0.8930
16	58.5041	6.2874	0.7041
32	61.5257	6.2835	0.6448
64	62.9336	6.2835	0.6277

The table shows the values of the computed total variation (TV), absolute curvature (AC), and squared curvature (SC) for a varying number of discrete orientations  $N_\theta$ . In parentheses we also give the values of the respective energies for the true solution. Observe that the squared curvature is approximated well only when using a quite large number of discrete orientations

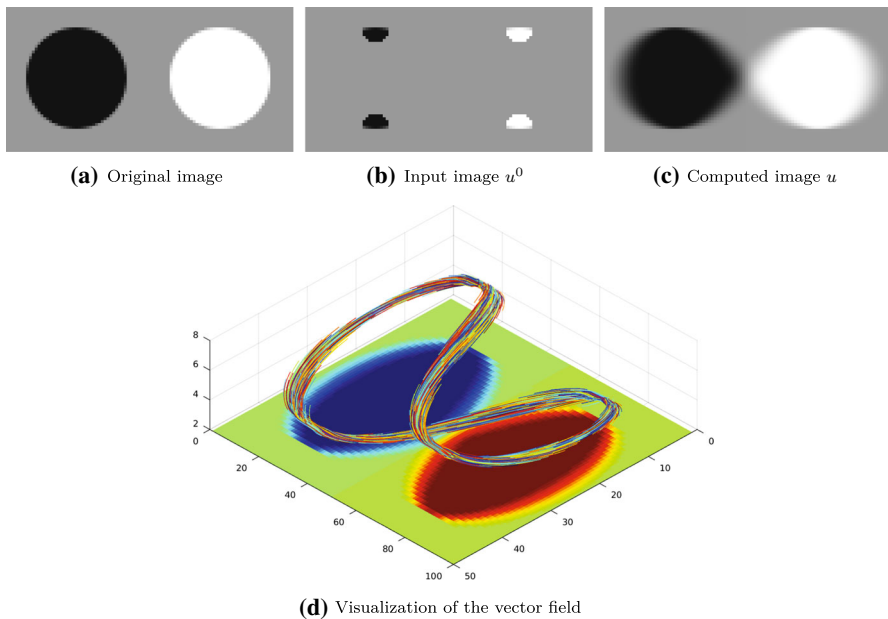
Figure 4g visualizes the averaged Raviart–Thomas vector field  $\hat{\sigma}$  of the disk example in case of  $N_\theta = 64$  discrete orientations. Observe that the vector field nicely corresponds to the expected shape of a helix, shown in Fig. 1. However, due to diffusive effects of our numerical scheme, the vector field does not perfectly concentrate on a one-dimensional structure.

#### 4.4 Non-smooth level sets

Here we demonstrate a typical effect of our convexification which finds a low-energy solution for an image with non-smooth level sets which should have infinite energy in more standard relaxations [53] of the Elastica energy. Figure 5a shows an image of size  $N_1 \times N_2 = 100 \times 50$  pixels of a black and a white disk in front of a gray background. Similarly to the inpainting problem of the previous example, we fix only the four small parts of the original image (see Fig. 5b) and minimize the TSC energy in the inpainting domain  $\mathcal{D}$ , indicated by the gray area. In this experiment we used  $N_\theta = 64$  discrete orientations and we set  $\alpha = 17$  to match the radius of the disks. Figure 5c shows the computed minimizer of the TSC energy. Observe that the solution  $u$  does not yield the expected two disks of the original image but rather drop-like shapes that form a sharp cusp in the middle of the image. Figure 5d shows a stream-line representation of the minimizing vector field  $\hat{\sigma}$  in the roto-translation space. Inspecting the field, one can immediately see the reason for this behavior. The vector field  $\hat{\sigma}$  forms a twisted  $\infty$ -shape curve which “skips” the strong curvature of the cusp.

#### 4.5 Shape completion

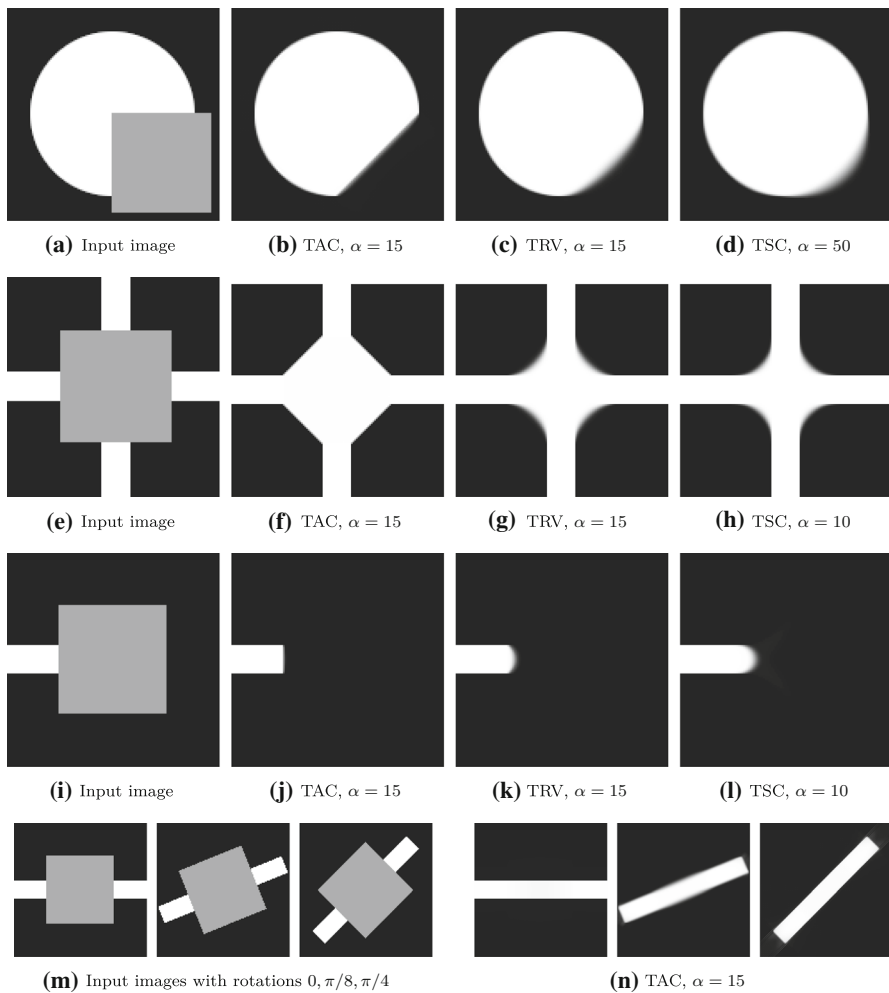
In this section we provide some qualitative results on a number of different shape completion problems. Similar to the previous two examples, we define an inpainting domain  $\mathcal{D}$  which is indicated by the gray area and keep the remaining image as boundary condition. Figure 6 shows various input shapes with their inpainting domains and the solutions of minimizing different curvature energies using different settings of the parameter  $\alpha$ . One can see that while minimizing the TAC energy usually leads



**Fig. 5** Effect of convexifying curvature depending energies in the roto-translation space. **a** The original image, **b** the input image  $u^0$  of size  $100 \times 50$  pixels, where gray pixels indicate the inpainting domain. **c** The computed solution of minimizing the TSC energy. **d** Visualizes the vector field in the roto-translation space. Observe that the twisted  $\infty$ -shape curve skips the strong curvature at the cusp

to straight connections and sharp corners, TSC leads to a smooth continuation of the boundaries. The TRV energy leads to results which are somewhere in between the results of TAC and TSC. In the last row of Fig. 6, we additionally demonstrate the behavior for completing straight lines at different rotations by minimizing the TAC energy. It turns out that our discretization scheme performs quite well for rotations of  $0, \pi/4, \pi/2, \dots$  but leads to more diffusive results for rotations of  $\pi/8, 3\pi/8, \dots$ . The development of a more isotropic scheme will be subject of future research.

In Fig. 7, we show the application to shape completion from dipoles using the original data of “Weickert’s cat” [76]. The input image  $u^0$  is of size  $N_1 \times N_2 = 400 \times 380$  pixels and the inpainting domain  $\mathcal{D}$  is again indicated by the gray pixels. In this example we used  $N_\theta = 64$  discrete orientations and the curvature parameter was set to  $\alpha = 50$ . In order to avoid relaxation artifacts (self-intersections) inside the dipoles, we additionally use zero-boundary conditions for the field  $\sigma^{1,2}$  within the constant areas of the dipoles. The results show that our numerical scheme can successfully reconstruct the curvilinear shape of the cat from only very little information given by the dipoles. However, we can also observe diffusion artifacts of our numerical scheme, especially if the reconstructed boundaries are relatively long. Sharper results for this kind of problems are usually obtained using sophisticated anisotropic diffusion schemes such as the edge enhancing diffusion (EED) [67,75]. It would be interesting to understand whether these schemes also minimize an underlying variational energy.

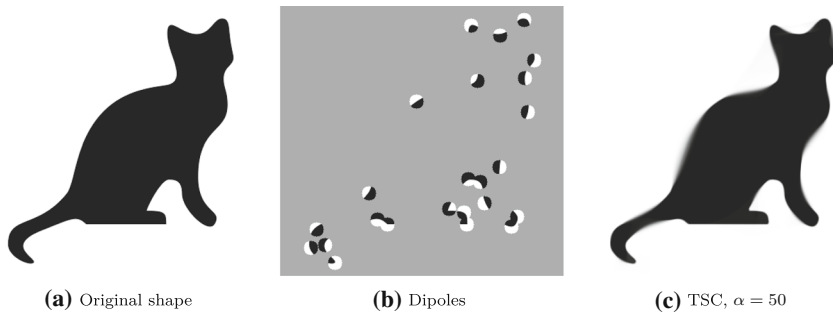


**Fig. 6** Shape completion for various shapes using total absolute curvature (TAC), total roto-translational variation (TRV) and total squared curvature (TSC). The inpainting domain is indicated by the gray pixels. In the last row we provide results for completing straight lines at different angles

#### 4.6 Shape regularization

In our next experiment, we apply our curvature based energies for shape regularization. Given an input image  $(u_i^0)_{i \in \mathcal{I}}$ , which is the characteristic function of a given shape, our aim is to compute a simplified (or regularized) shape which is represented by means of a (relaxed) binary image  $(u_i)_{i \in \mathcal{I}}$ . We make use of a simple linear fidelity term which is frequently used in image segmentation:

$$G(u) = \sum_{i \in \mathcal{I}} g_i, \quad g_i(u_i) = u_i w_i + \iota_{[0,1]}(u_i),$$



**Fig. 7** Computation of “Weickert’s cat” [76]: **a** the original shape, **b** depicts the given dipole data and **c** is the result of minimizing the TSC energy in the gray inpainting domain

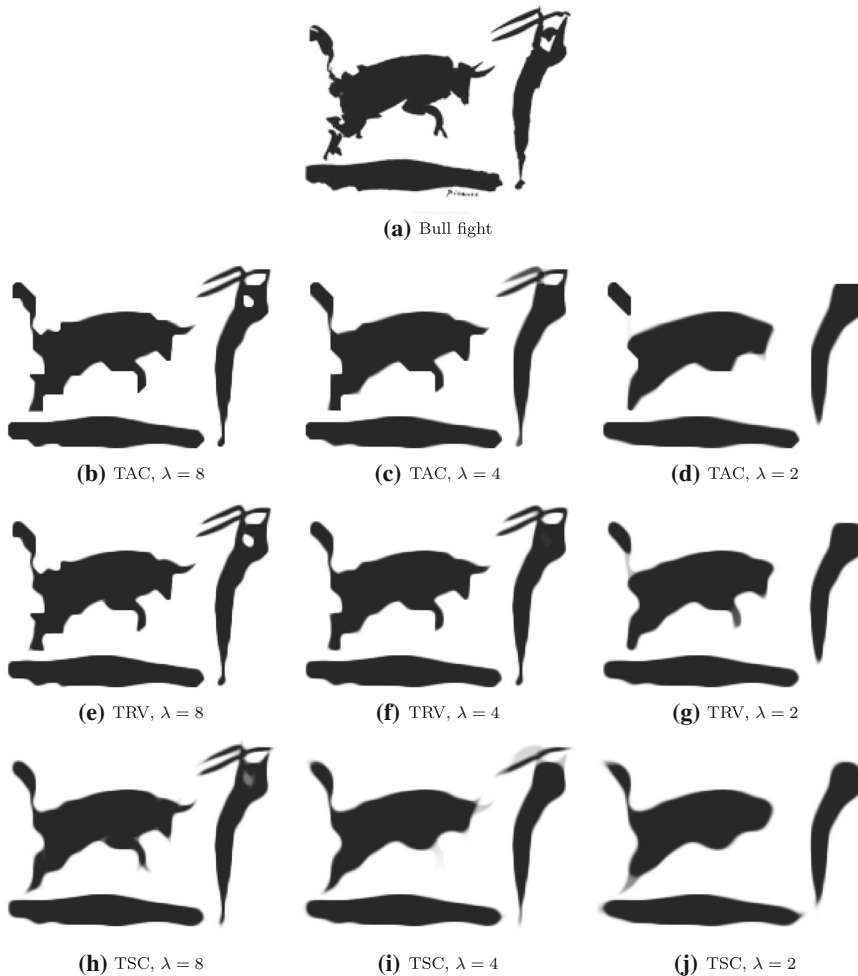
where  $w \in \mathbb{R}^{\mathcal{I}}$  is a force field. For the application to shape regularization we use  $w = \lambda(\frac{1}{2} - u^0)$ , where  $\lambda > 0$  defines the strength of the force field. The proximal map for this data term is easily computed:

$$u = \text{prox}_{\tau G}(v) \iff u_i = \max(0, \min(1, v_i - \tau w_i)), \quad \forall i \in \mathcal{I}.$$

In Fig. 8, we apply TAC, TRV and TSC regularization to regularize the shape of Picasso’s “Bull fight” image. In all three cases, we set  $\alpha = 10$  and we use different settings of the parameter  $\lambda$  to obtain gradually simplified shapes. From the results one can see that TAC regularization yields shapes with relatively straight boundaries and sharp corners. TRV regularization yields smooth shapes but also allows for sharp corners. TSC regularization yields smooth shapes. Observe that whenever it seems energetically preferable, the solution of TSC produces sharp cusps with “hidden” edges to bypass locations of strong curvature, as predicted by the theory [12]. This effect is usually less visible when minimizing the TAC or TRV energies.

#### 4.7 Image inpainting

In this section, we apply our proposed curvature energies to the classical problem of image inpainting. Similar to the previous examples, we use the data term (32) with the only difference that the input image  $u^0$  is now a gray level image. In both examples we used  $N_\theta = 32$  discrete orientations and the curvature parameter was set to  $\alpha = 15$ . Figure 9 shows the results for two different inpainting problems. In the “Einstein” image, we randomly remove 80% lines and in the “Picasso” image we randomly remove 90% of the pixels. In case of the “Einstein” image, we use TSC regularization with  $\alpha = 10$  and for the “Picasso” we set  $\alpha = 5$ . In case of the “Einstein” image, the gaps are much larger and hence we use a larger value of  $\alpha$ . For comparison, we also provide results of standard TV regularization, which is equivalent to using  $\alpha = 0$  in one of the three curvature energies. From the results one can clearly see that curvature regularization leads to significantly better inpainting results. On the downside, we can also observe some artifacts which are caused by our convex representation in the roto-translation space.



**Fig. 8** Shape regularization using curvature regularization. **a** The original “Bull fight” image. **b–d** The result of TAC regularization, **e–g** the result of TRV regularization and **h–j** show the result of TSC regularization

In Fig. 10 we also provide a comparison between the three proposed curvature models TAC, TRV, and TSC. At first glance, all three models give comparable results. However, a more detailed inspection of high-contrast details such as the bright buttons on Picasso’s shirt unveils that TSC (which is based on the squared curvature) has the tendency to produce larger and rounder objects. In contrast TAC (which is based on the scale invariant total absolute curvature) gives smaller and less rounder objects. TRV produces results somewhere in between the results of TAC and TSC.



**Fig. 9** Image inpainting using the proposed curvature-based energies. **a** A blend of the original “Einstein” image with its degraded version where we have removed 80% of the lines. **d** A blend of the “Picasso” image with a version where we have removed 90% of the pixels. **b** Is the inpainting result of the “Einstein” and **e** is the inpainting result of the “Picasso” image. For comparison, **c, f** the result of standard TV regularization. Observe that our proposed curvature based regularization leads to significant better inpainting results

#### 4.8 Image denoising

In our last experiments we investigate our curvature energies for the classical problem of image denoising. We investigate two different types of noise: Zero-mean Gaussian noise and impulse noise such as “salt and pepper” noise. For Gaussian noise, it is well-known data a quadratic  $\ell_2$  data term

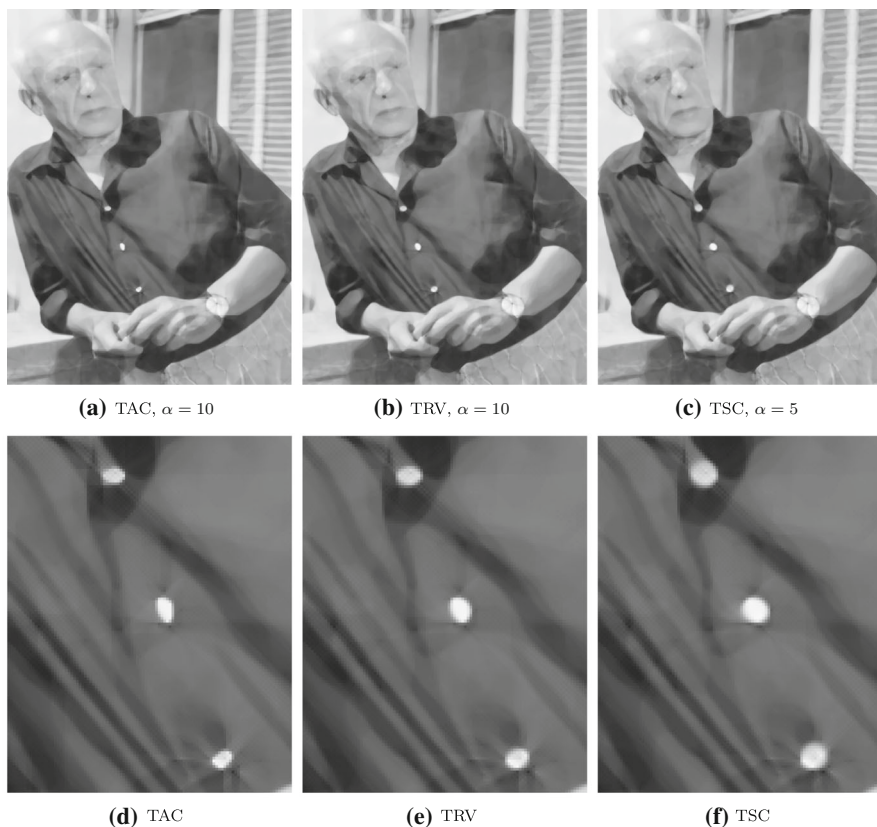
$$G(u) = \frac{\lambda}{2} \|u - f\|^2,$$

where  $\lambda > 0$  is a good data fidelity parameter. The proximal map is given by

$$u = \text{prox}_{\tau G}(v) \iff u_i = \frac{v_i + \tau \lambda f_i}{1 + \tau \lambda}, \quad \forall i \in \mathcal{I}.$$

In case of impulse noise, a  $\ell_1$  data term is more suitable:

$$G(u) = \lambda \|u - f\|_1,$$



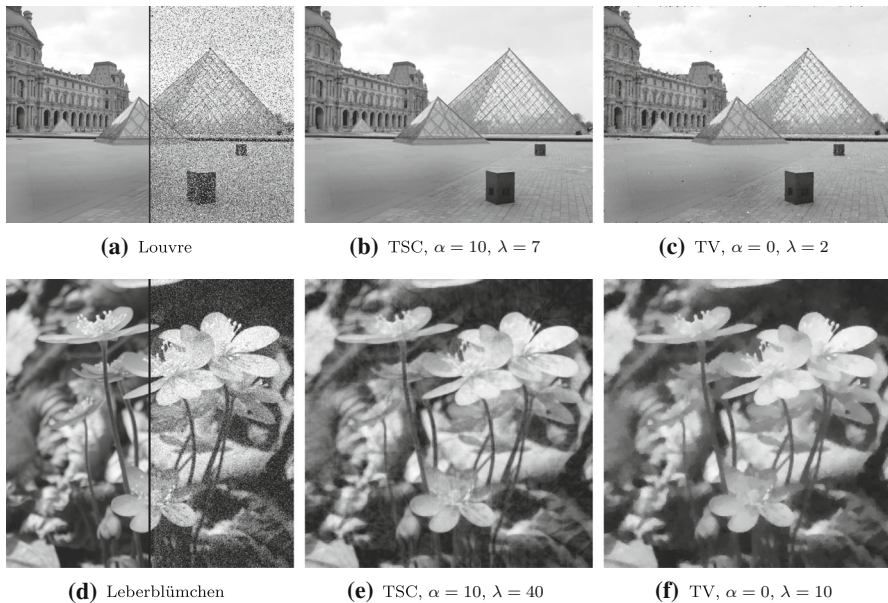
**Fig. 10** Comparison between TAC, TRV and TSC on the Picasso inpainting problem of Fig. 9. The inpainting quality of all three models is comparable but TSC has the tendency to give larger, rounder objects (see the detail views)

since it is more robust with respect to outliers. The proximal map for the  $\ell_1$  data term is given by the classical soft-shrinkage formula

$$u = \text{prox}_{\tau G}(v) \iff u_i = f_i + \max(0, |v_i - f_i| - \tau\lambda) \cdot \text{sgn}(v_i - f_i), \quad \forall i \in \mathcal{I}.$$

In Fig. 11 we show the results of our curvature-based regularization energies for image denoising. In case of the “Louvre” image we generated the noisy image by adding 25% “salt and pepper” noise. The noisy “Leberblümchen” image was generated by adding zero-mean Gaussian noise with standard deviation 0.1. In case of Gaussian noise we use a  $\ell_2$  data term and in case of “salt and pepper” noise we used the  $\ell_1$  fidelity term. In both cases, we use TSC regularization and the curvature parameter was set to  $\alpha = 10$ . For comparison, we also show the results of standard total variation (TV) denoising, which is obtained from our models by setting  $\alpha = 0$ . One can clearly see that TSC regularization leads to a better preservation of image edges than TV regularization, in particular at small and elongated structures. This effect is more visible in case of





**Fig. 11** Image denoising using curvature-based regularization. **a** The original “Louvre” image blended with a noisy version, where we have added 25% “salt and pepper” noise. **d** The clean “Leberblümchen” image together with its noisy variant where we have added zero-mean Gaussian noise with standard deviation 0.1 **b**, **e** the result of image denoising using TSC regularization. For comparison we show in **(c)** and **(f)** the result when using classical total variation (TV) regularization

“salt and pepper” noise since the  $\ell_1$  data term leads to a behavior similar to image inpainting, once an outlier is detected.

## 5 Proof of Theorem 1

This section and the following are devoted to the proof of Theorem 1. We first show a preliminary result which shows that admissible curves (of finite energy) in  $\Omega \times \mathbb{S}^1$  project onto curves in  $\Omega$  with bounded energy as well.

### 5.1 Control of curves

Let  $\gamma : [0, L] \rightarrow \Omega \times \mathbb{S}^1$  a rectifiable curve of length  $L$  in  $\Omega \times \mathbb{S}^1$ , parameterized by its length ( $|\dot{\gamma}| = 1$  a.e.). We define its energy as

$$\mathcal{E}(\gamma) := \int_0^1 h(\gamma(s), \dot{\gamma}(s)) ds$$

and assume that it is finite, which in particular yields that for a.e.  $s$ ,

$$\dot{\gamma}^x(s) = |\dot{\gamma}^x(s)|(\cos \gamma^\theta(s), \sin \gamma^\theta(s))^T.$$

We consider the projected curve  $\gamma^x : [0, L] \rightarrow \Omega$  (possibly overlapping even if  $\gamma$  is simple) and wish to show that the energy  $\mathcal{E}$  controls an energy on this curve. First we reparameterize (as usual) the rectifiable curve  $\gamma^x$  by its length. For this we introduce the length, for  $s \in [0, L]$

$$\ell(s) := \int_0^s |\dot{\gamma}^x(t)| dt \in [0, L],$$

which is a 1-Lipschitz, nondecreasing function. Since clearly for  $s' > s$  one has  $\ell(s) = \ell(s')$  if and only if  $\dot{\gamma}^x(t) = 0$  a.e. in  $[s, s']$ , we also find that  $\gamma^x(t) = \gamma^x(s)$  for all  $t \in [s, s']$ . This means that (obviously) one can reparameterize  $\gamma^x$  by defining a curve  $\tilde{\gamma} : [0, \ell(L)] \rightarrow \Omega$  such that  $\tilde{\gamma}(\ell(s)) = \gamma^x(s)$  for all  $s \in [0, L]$ . By definition it is clear that  $\tilde{\gamma}$  is 1-Lipschitz, moreover if  $l = \ell(s)$  and  $l' = \ell(s') > l$ ,

$$\frac{\tilde{\gamma}(l') - \tilde{\gamma}(l)}{l' - l} = \frac{\int_s^{s'} \dot{\gamma}^x(t) dt}{\int_s^{s'} |\dot{\gamma}^x(t)| dt} \xrightarrow{l' \rightarrow l} \frac{\dot{\gamma}^x(s)}{|\dot{\gamma}^x(s)|} = \gamma^\theta(s)$$

if  $s$  is a Lebesgue point of  $\dot{\gamma}^x$  where  $\dot{\gamma}^x(s) \neq 0$ . One deduces that for a.e.  $l \in [0, \ell(L)]$ ,  $\tilde{\gamma}$  has the tangent vector  $\tau(l) = (\cos \theta(l), \sin \theta(l))^T$  given by

$$\tau(l) = \dot{\tilde{\gamma}}(l) = \begin{pmatrix} \cos \gamma^\theta(s) \\ \sin \gamma^\theta(s) \end{pmatrix}$$

where  $l = \ell(s)$ . Observe that for any smooth  $\psi : (0, \ell(L)) \rightarrow \mathbb{R}$  with compact support one has

$$\begin{aligned} \int_0^{\ell(L)} \dot{\psi}(l) \theta(l) dl &= \int_0^L \dot{\psi}(\ell(s)) \theta(\ell(s)) \dot{\ell}(s) ds \\ &= \int_0^L \gamma^\theta(s) (\psi \circ \ell)'(s) ds \\ &= - \int_0^L \dot{\gamma}^\theta(s) \psi(\ell(s)) ds \leq L \|\psi\|_{C^0} \end{aligned}$$

which shows that  $\theta(l)$  is a  $BV$  function, or equivalently that  $\tilde{\gamma}$  has a curvature  $\kappa = \dot{\theta}$  which is a bounded measure; in particular taking the supremum over  $\psi$  with  $|\psi| \leq 1$  one finds

$$\int_0^{\ell(L)} d|\kappa|(l) \leq \int_0^L |\dot{\gamma}^\theta(s)| ds \leq L.$$

In addition, it follows, given  $\psi \in C_c^\infty(\ell(a), \ell(b))$  with  $0 \leq a < b \leq L$  that:

$$\begin{aligned} & \int_{\ell(a)}^{\ell(b)} (-\dot{\psi}(l)\theta(l) - f^*(\psi(l)))dl \\ &= \int_a^b (-(\psi \circ \ell)'(s)\gamma^\theta(s) - f^*(\psi(\ell(s)))|\dot{\gamma}^x(s)|)ds \\ &= \int_a^b (\psi(\ell(s))\dot{\gamma}^\theta(s) - f^*(\psi(\ell(s)))|\dot{\gamma}^x(s)|)ds \leq \int_a^b h(\gamma(s), \dot{\gamma}(s))ds \end{aligned}$$

and taking the supremum one deduces that

$$\int_{\ell(a)}^{\ell(b)} f(\kappa) \leq \int_a^b h(\gamma(s), \dot{\gamma}(s))ds$$

(where the left-hand side integral is on the open interval and denotes a convex function of a measure). Now, if  $\varphi \in C^0(\Omega; \mathbb{R}_+)$  is a bounded, continuous and nonnegative function, one deduces that

$$\begin{aligned} \int_0^{\ell(L)} \varphi(\tilde{\gamma}(l))f(\kappa) &= \int_0^\infty \left( \int_{\{l: \varphi(\tilde{\gamma}(l)) > t\}} f(\kappa) \right) dt \\ &\leq \int_0^\infty \left( \int_{\{s: \varphi(\gamma^x(s)) > t\}} h(\gamma(s), \dot{\gamma}(s)) \right) dt \end{aligned}$$

and it follows

$$\int_0^{\ell(L)} \varphi(\tilde{\gamma}(l))f(\kappa) \leq \int_0^L \varphi(\gamma^x(s))h(\gamma(s), \dot{\gamma}(s))ds. \quad (34)$$

We now show the following lemma:

**Lemma 5.1** *Let  $\Gamma \subset \Omega$  be a  $C^2$  (oriented) curve with tangent  $\tau_\Gamma$  and curvature  $\kappa_\Gamma$ , and let  $\gamma : [0, L] \rightarrow \Omega \times \mathbb{S}^1$  be a rectifiable curve. Define  $\Gamma^+ = \{x \in \Gamma : \exists s \in (0, L), \gamma^x(s) = x \text{ and } \dot{\gamma}^x(s) \cdot \tau_\Gamma \geq 0\}$ . Then for any bounded, nonnegative, continuous function  $\varphi \in C^0(\Omega; \mathbb{R}_+)$ ,*

$$\int_{\Gamma^+} \varphi(x)f(\kappa_\Gamma)d\mathcal{H}^1 \leq \int_0^L \varphi(\gamma^x(s))h(\gamma(s), \dot{\gamma}(s))ds. \quad (35)$$

**Proof** We just need to show that, defining the measure  $\kappa$  as before,

$$\int_{\Gamma^+} \varphi(x)f(\kappa_\Gamma)d\mathcal{H}^1 \leq \int_0^{\ell(L)} \varphi(\tilde{\gamma})f(\kappa),$$

and the conclusion will follow from (34).

A first observation is that as  $\tau(l)$  defined above is  $BV$ , it has at most a countable number of jumps and one can cover (up to the jump points)  $[0, \ell(L)]$  with an at most countable union of intervals  $(a_i, b_i)_{i \in I}$  on which  $\tau(l)$  is continuous (hence  $\tilde{\gamma}$  is a  $C^1$  curve on  $(a_i, b_i)$ ). Moreover, possibly dropping further a finite number of points, one may assume that  $|\kappa|(a_i, b_i) < \pi$  so that the image  $\gamma^x(a_i, b_i)$  is a simple curve (without crossing).

Assume  $|\{l \in (a_i, b_i) : \tilde{\gamma}(l) \in \Gamma, \dot{\tilde{\gamma}}(l) \cdot \tau_\Gamma \geq 0\}| > 0$ . Choose such an  $l \in (a_i, b_i)$ , with  $x = \tilde{\gamma}(l) \in \Gamma^+$ , and such that  $\tilde{\gamma}$  is differentiable in  $l$ . We assume in addition that  $x$  is a point of  $(\mathcal{H}^1)$ -density one in  $\tilde{\gamma}(a_i, b_i) \cap \Gamma^+$ , that  $l$  is a Lebesgue point of  $\kappa^a$ , the absolutely continuous part (w.r. the Lebesgue measure) of  $\kappa$ , and that  $\theta(l)$  is approximately differentiable at  $l$ : for any  $\eta > 0$ ,

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \left| \left\{ s \in (l - \varepsilon, l + \varepsilon) : \frac{|\theta(s) - \theta(l) - \kappa^a(l)(s - l)|}{|s - l|} ds > \eta \right\} \right| = 0.$$

All this is true  $\mathcal{H}^1$ -a.e. in  $\tilde{\gamma}(a_i, b_i) \cap \Gamma^+$ , so there is no loss of generality, see for instance [3].

If  $\dot{\tilde{\gamma}}(l) \neq \tau_\Gamma(x)$ , then  $x$  must be an isolated point in  $\tilde{\gamma}(a_i, b_i) \cap \Gamma^+$ , a contradiction. Hence  $\dot{\tilde{\gamma}}(l) = \tau_\Gamma(x)$ . In the same way, one has that for a.e.  $s$  near  $l$  such that  $\tilde{\gamma}(s) \in \Gamma$ ,  $\dot{\tilde{\gamma}}(s) = \tau_\Gamma(\tilde{\gamma}(s))$ . For such  $s$ , one therefore has that

$$\frac{|\theta(s) - \theta(l) - \kappa^a(l)(s - l)|}{|s - l|} = \frac{|\theta_\Gamma(\tilde{\gamma}(s)) - \theta_\Gamma(x) - \kappa^a(l)(s - l)|}{|s - l|},$$

where  $\theta_\Gamma$  is the angle between  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $\tau_\Gamma$ . Using that for such  $s$ ,

$$\begin{aligned} \theta_\Gamma(\tilde{\gamma}(s)) - \theta_\Gamma(x) &= \kappa_\Gamma(x)(\tilde{\gamma}(s) - x) \cdot \tau_\Gamma(x) + o(|\tilde{\gamma}(s) - x|) \\ &= \kappa_\Gamma(x)\dot{\tilde{\gamma}}(l) \cdot \tau_\Gamma(x)(s - l) + o(|\dot{\tilde{\gamma}}(s)||s - l|) \end{aligned}$$

we obtain that for such  $s$ ,

$$\frac{|\theta(s) - \theta(l) - \kappa^a(l)(s - l)|}{|s - l|} = |\kappa_\Gamma(x) - \kappa^a(l)| + o(1)$$

(where  $o(1)$  goes to 0 as  $\varepsilon \rightarrow 0$ ). Hence, for  $\varepsilon > 0$  small and  $\eta > 0$ ,

$$\begin{aligned} &\left\{ s \in (l - \varepsilon, l + \varepsilon) : \frac{|\theta(s) - \theta(l) - \kappa^a(l)(s - l)|}{|s - l|} ds > \eta \right\} \\ &\supseteq \{s \in (l - \varepsilon, l + \varepsilon) : \tilde{\gamma}(s) \in \Gamma \text{ and } |\kappa_\Gamma(x) - \kappa^a(l)| + o(1) > \eta\} \end{aligned}$$

In the limit, using that  $l$  is a point of density one in the set  $\{\tilde{\gamma}(\cdot) \in \Gamma\}$ , we obtain that  $|\kappa_\Gamma(x) - \kappa^a(l)| \leq \eta$  and since  $\eta$  is arbitrary,  $\kappa_\Gamma(x) = \kappa^a(l)$ . Since the intervals  $(a_i, b_i)$  cover  $\mathcal{H}^1$ -almost all of  $[0, \ell(L)]$ , we find that this equality holds  $\mathcal{H}^1$ -a.e. in  $\Gamma^+$ .

The thesis of the Lemma easily follows, since

$$\int_0^{\ell(L)} \varphi(\tilde{\gamma}) f(\kappa) \geq \int_0^{\ell(L)} \varphi(\tilde{\gamma}) f(\kappa^a) dx$$

□

## 5.2 Decomposition of $\sigma$

Let  $u \in BV(\Omega)$  with  $F(u) < \infty$ . We observe that we can decompose an admissible  $\sigma$  (with in particular  $\operatorname{div} \sigma = 0$ ) as follows:

$$\sigma = \int_{\mathcal{M}} \lambda d\mu(\lambda), \quad |\sigma| = \int_{\mathcal{M}} |\lambda| d\mu(\lambda), \quad (36)$$

where  $\mu$  is a nonnegative measure on the set of charges  $\mathcal{M}$  which is here a shorthand notation for  $\mathcal{M}(\Omega \times \mathbb{S}^1; \mathbb{R}^3)$ : see “Appendix B” which discusses the results of Smirnov in [73].

**Lemma 5.2** *Assume (36) holds. Then for any  $g(\xi, p)$  nonnegative, continuous, convex and one-homogeneous in its second argument,*

$$\int_{\Omega \times \mathbb{S}^1} g(\xi, \sigma) = \int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} g(\xi, \lambda) \right) d\mu(\lambda). \quad (37)$$

**Proof** For  $\rho > 0$  we define  $\Omega_\rho = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \rho\}$ . We let for  $\rho > 0$ ,  $\varepsilon \in (0, \rho)$  and  $\xi \in \Omega_{2\rho} \times \mathbb{S}^1$

$$f_{\rho, \varepsilon}(\xi) = \begin{cases} \frac{\sigma(\Omega_{3\rho} \times \mathbb{S}^1 \cap \overline{B}(\xi, \rho))}{|\sigma|(B(\xi, \rho + \varepsilon))} & \text{if } |\sigma|(B(\xi, \rho + \varepsilon)) > 0, \\ 0 & \text{else.} \end{cases}$$

We also let  $f_{\rho, \varepsilon}(\xi) = 0$  if  $\xi \in \Omega \times \mathbb{S}^1 \setminus \Omega_{2\rho} \times \mathbb{S}^1$ : remark then that  $f_{\rho, \varepsilon}$  is the pointwise limit of functions in  $C_c^0(\Omega \times \mathbb{S}^1; B(0, 1))$  and that

$$f_\rho(\xi) := \lim_{\varepsilon \rightarrow 0} f_{\rho, \varepsilon}(\xi) = \begin{cases} \frac{\sigma(\Omega_{3\rho} \times \mathbb{S}^1 \cap \overline{B}(\xi, \rho))}{|\sigma|(B(\xi, \rho))} & \text{if } \xi \in \Omega_{2\rho} \times \mathbb{S}^1, |\sigma|(B(\xi, \rho)) > 0, \\ 0 & \text{else.} \end{cases}$$

By (36) we have for any  $\varphi \in C_c^0(\Omega \times \mathbb{S}^1; \mathbb{R}_+)$ :

$$\int_{\Omega \times \mathbb{S}^1} \varphi(|\sigma| - f_{\rho, \varepsilon} \cdot \sigma) = \int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} \varphi(|\lambda| - f_{\rho, \varepsilon} \cdot \lambda) \right) d\mu.$$

Observe that

$$\begin{aligned} & \int_{\Omega \times \mathbb{S}^1} \varphi(|\lambda| - f_{\rho, \varepsilon} \cdot \lambda) \\ &= \int_{\Omega \times \mathbb{S}^1} \varphi(|\lambda| - f_{\rho} \cdot \lambda) \frac{|\sigma|(\overline{B}(\xi, \rho))}{|\sigma|(B(\xi, \rho + \varepsilon))} \\ &+ \int_{\Omega \times \mathbb{S}^1} \varphi \frac{|\sigma|(B(\xi, \rho + \varepsilon) \setminus \overline{B}(\xi, \rho))}{|\sigma|(B(\xi, \rho + \varepsilon))} |\lambda| \end{aligned}$$

and using the monotone convergence theorem, we can send  $\varepsilon \rightarrow 0$  and deduce that

$$\int_{\Omega \times \mathbb{S}^1} \varphi(|\sigma| - f_{\rho} \cdot \sigma) = \int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} \varphi(|\lambda| - f_{\rho} \cdot \lambda) \right) d\mu. \quad (38)$$

Then, thanks to Radon-Nikodym's derivation theorem (and Lebesgue's convergence theorem),

$$\lim_{\rho \rightarrow 0} \int_{\Omega \times \mathbb{S}^1} \varphi(|\sigma| - f_{\rho} \cdot \sigma) = 0.$$

Observe that all the integrands in (38) are nonnegative. Let  $\varphi_n$  be a nondecreasing sequence of compactly supported nonnegative smooth functions which converges to 1. Then one can build a subsequence  $\rho_n \downarrow 0$  such that

$$\sum_n \int_{\Omega \times \mathbb{S}^1} \varphi_n(|\sigma| - f_{\rho_n} \cdot \sigma) < +\infty,$$

so that

$$\int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} \sum_n \varphi_n(|\lambda| - f_{\rho_n} \cdot \lambda) \right) d\mu < +\infty.$$

It follows that there exists a set  $E \subset \mathcal{M}$  with  $\mu(E) = 0$  such that if  $\lambda \notin E$ ,

$$\lim_n \int_{\Omega \times \mathbb{S}^1} \varphi_n(|\lambda| - f_{\rho_n} \cdot \lambda) = 0$$

and  $f_{\rho_n} \cdot \frac{\lambda}{|\lambda|} \rightarrow 1$   $|\lambda|$ -a.e. in  $\Omega \times \mathbb{S}^1$  (using  $|\lambda| - f_{\rho_n} \cdot \lambda = |\lambda|(1 - f_{\rho_n} \cdot \frac{\lambda}{|\lambda|}) \geq 0$ ). Using  $|f_{\rho_n} \cdot \frac{\lambda}{|\lambda|}|^2 = |f_{\rho_n}^2| + 1 - 2f_{\rho_n} \cdot \frac{\lambda}{|\lambda|} \leq 2(1 - f_{\rho_n} \cdot \frac{\lambda}{|\lambda|})$  it follows that  $f_{\rho_n} \rightarrow \frac{\lambda}{|\lambda|}$ ,  $|\lambda|$ -a.e.

Let  $g : (\Omega \times \mathbb{S}^1) \times \mathbb{R}^3 \rightarrow \mathbb{R}$  be a continuous, nonnegative, and convex one-homogeneous function in its second argument. Then as before for  $\varphi \in C_c^0(\Omega \times \mathbb{S}^1; \mathbb{R}_+)$

$$\int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, f_{\rho, \varepsilon}(\xi)) |\sigma| = \int_{\mathcal{M}} \int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, f_{\rho, \varepsilon}) |\lambda| d\mu$$

and since  $g(\xi, f_{\rho, \varepsilon}(\xi)) = (|\sigma|(\overline{B}(\xi, \rho)) / |\sigma|(B(\xi, \rho + \varepsilon)))g(\xi, f_\rho(\xi)) \uparrow g(\xi, f_\rho(\xi))$  as  $\varepsilon \downarrow 0$ , the monotone convergence theorem yields that

$$\int_{\Omega \times \mathbb{S}^1} \varphi g(\xi, f_\rho) |\sigma| = \int_{\mathcal{M}} \int_{\Omega \times \mathbb{S}^1} \varphi g(\xi, f_\rho) |\lambda| d\mu.$$

Using Lebesgue's theorem, we have

$$\lim_{\rho \rightarrow 0} \int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, f_\rho(\xi)) |\sigma| = \int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, \frac{\sigma}{|\sigma|}(\xi)) |\sigma| = \int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, \sigma).$$

On the other hand, we know that if  $\lambda \notin E$ , still thanks to Lebesgue's theorem,

$$\lim_n \int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, f_{\rho_n}(\xi)) |\lambda| = \int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, \frac{\lambda}{|\lambda|}(\xi)) |\lambda| = \int_{\Omega \times \mathbb{S}^1} \varphi g(\xi, \lambda).$$

We invoke one last time Lebesgue's theorem to deduce that (for all nonnegative test function  $\varphi$ )

$$\lim_n \int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} \varphi(\xi) g(\xi, f_{\rho_n}(\xi)) |\lambda| \right) d\mu = \int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} \varphi g(\xi, \lambda) \right) d\mu.$$

We deduce (37).  $\square$

**Corollary 5.3** *Let  $\sigma$  be admissible for problem (6), and assume it is decomposed as in (36). Then  $\mu$ -a.e. measure  $\lambda$  satisfies*

$$\frac{\lambda}{|\lambda|} \cdot \theta \geq 0, \quad \frac{\lambda}{|\lambda|} \cdot \theta^\perp = 0 \quad \text{a.e. in } \Omega \times \mathbb{S}^1$$

**Proof** Apply the Lemma with  $g(\xi, p) = (\theta \cdot p)^-$  and then with  $g(\xi, p) = |\theta^\perp \cdot p|$ .  $\square$

**Corollary 5.4** *Let  $\sigma$  be admissible for problem (6), and assume it is decomposed as in (36). Let  $h$  be defined as in (3), and  $\varphi \in C_c^0(\Omega)$  a nonnegative test function. Then*

$$\int_{\Omega \times \mathbb{S}^1} \varphi(x) h(\theta, \sigma) = \int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} \varphi(x) h(\theta, \lambda) \right) d\mu(\lambda). \quad (39)$$

**Proof** This is a consequence of Lemma 5.2, however now  $h$  can take the value  $+\infty$ . We simply observe that if we let, for  $\varepsilon > 0$ ,

$$h_\varepsilon(\theta, p) := \min_{q \in \mathbb{R}^3} h(\theta, q) + \frac{1}{\varepsilon} |p - q|$$

then  $h_\varepsilon$  is  $(1/\varepsilon)$ -Lipschitz, convex, one-homogeneous in the second variable, and continuous (and hence also  $(\xi, p) = ((x, \theta), p) \mapsto \varphi(x) h_\varepsilon(\theta, p)$ ). Hence (37) yields

$$\int_{\Omega \times \mathbb{S}^1} \varphi(x) h_\varepsilon(\theta, \sigma) = \int_{\mathcal{M}} \left( \int_{\Omega \times \mathbb{S}^1} \varphi(x) h_\varepsilon(\theta, \lambda) \right) d\mu(\lambda).$$

The results follows from the monotone convergence theorem.  $\square$

### 5.3 Conclusion: Proof of Theorem 1

We can now complete the proof of Theorem 1. We consider  $E$  a  $C^2$  set. First, if  $\gamma(t) : [0, 1] \rightarrow \Omega$  is an oriented parameterization of (the closure of)  $\partial E \cap \Omega$ , (assuming it is connected, otherwise one needs to introduce a curve for each connected component), one can define in  $\Omega \times \mathbb{S}^1$  the curve  $\lambda(t) = (\gamma(t), \theta(t))$  where  $\theta(t)$  is defined by  $\lambda(t) = |\gamma'(t)|(\cos \theta(t), \sin \theta(t))^T$ . We then let  $\Gamma = \lambda([0, 1]) \cap \Omega \times \mathbb{S}^1$  and  $\sigma = \tau_\Gamma \mathcal{H}^1 \llcorner \Gamma$ . Then,  $\sigma$  is admissible for (6) with  $u = \chi_E$ , and one finds that

$$F(\chi_E) \leq \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) = \int_{\partial E \cap \Omega} f(\kappa_E) d\mathcal{H}^1.$$

We need therefore to prove the reverse inequality. Let  $\sigma$  be admissible for (6). Thanks to [73, Theorem A] [cf “Appendix B”, Eq. (59)], one can decompose  $\sigma$  as

$$\sigma = \int_{\mathfrak{C}_1} \lambda d\mu(\lambda), \quad |\sigma| = \int_{\mathfrak{C}_1} |\lambda| d\mu(\lambda),$$

where  $\lambda$  are of the form

$$\lambda_\gamma = \tau_\gamma \mathcal{H}^1 \llcorner \gamma$$

for rectifiable (possibly closed) curves  $\gamma \subset \Omega \times \mathbb{S}^1$  of length at most one.

Equation (39) is valid for this decomposition and shows that for any  $\varphi \in C_c^0(\Omega; [0, 1])$ ,

$$\int_{\Omega \times \mathbb{S}^1} \varphi(x) h(\theta, \sigma) = \int_{\mathfrak{C}_1} \left( \int_\gamma \varphi(x) h(\theta, \tau_\gamma) d\mathcal{H}^1 \right) d\mu(\lambda_\gamma). \quad (40)$$

Now, for any  $\psi \in C_c^0(\Omega; \mathbb{R}^2)$ , one has

$$\int_{\Omega \times \mathbb{S}^1} (\psi, 0)^T \cdot \sigma = \int_\Omega \psi \cdot Du^\perp = \int_{\partial\Omega} \psi \cdot \tau_E d\mathcal{H}^1 \quad (41)$$

where  $\tau_E$  is a tangent vector to  $\partial E$  (oriented with  $E$  on the left-hand side and  $E^c$  on the right-hand side<sup>1</sup>).

Let us introduce the signed distance function  $d_E(x) = \text{dist}(x, \Omega \setminus E) - \text{dist}(x, E)$  which is  $C^2$  in a neighborhood of  $\partial E$  and is such that  $\nabla d_E = \nu_E$  (the inner normal)

<sup>1</sup> This just depends on the choice of the  $90^\circ$  rotation  $x \mapsto x^\perp$ .



on  $\partial\Omega$  and  $\Delta d_E = -\kappa_E$  (we assume the curvature is nonnegative where the set is convex). If we consider, for  $\varepsilon > 0$ , a test function of the form

$$\psi(x) = (\nabla d_E(x))^\perp \left(1 - \frac{|d_E(x)|}{\varepsilon}\right)^+ \varphi(x),$$

with  $\varphi \in C_c^0(\Omega)$ , (41) yields

$$\begin{aligned} \int_{\partial E} \varphi(x) d\mathcal{H}^1 &= \int_{\Omega \times \mathbb{S}^1} (\psi, 0)^T \cdot \sigma \\ &= \int_{\mathfrak{C}_1} \left( \int_{\gamma} \varphi(x) \left(1 - \frac{|d_E(x)|}{\varepsilon}\right)^+ ((\nabla d_E(x))^\perp, 0)^T \cdot \tau_\gamma(x, \theta) d\mathcal{H}^1 \right) d\mu(\lambda_\gamma). \end{aligned}$$

Sending  $\varepsilon \rightarrow 0$  we find that

$$\int_{\partial E} \varphi(x) d\mathcal{H}^1 = \int_{\mathfrak{C}_1} \left( \int_{\gamma \cap (\partial E \times \mathbb{S}^1)} \varphi(\tau_E, 0)^T \cdot \tau_\gamma d\mathcal{H}^1 \right) d\mu(\lambda_\gamma).$$

Now, we can choose  $\varphi$  of the form  $\varphi(x) f(\kappa_E(x))$ , with  $\varphi \in C_c^0(\Omega; [0, 1])$ , assuming  $f$  is finite-valued. We obtain that

$$\int_{\partial E} \varphi f(\kappa_E) d\mathcal{H}^1 = \int_{\mathfrak{C}_1} \left( \int_{\gamma \cap (\partial E \times \mathbb{S}^1)} \varphi f(\kappa_E) \tau_E \cdot \tau_\gamma^x d\mathcal{H}^1 \right) d\mu(\lambda_\gamma). \quad (42)$$

Given the rectifiable curve  $\gamma$ , one has that

$$\begin{aligned} \int_{\gamma \cap (\partial E \times \mathbb{S}^1)} \varphi f(\kappa_E) \tau_E \cdot \tau_\gamma^x d\mathcal{H}^1 &\leq \int_{\Pi(\gamma) \cap \partial E} \varphi f(\kappa_E) (\tau_E \cdot \tau_{\Pi(\gamma)})^+ d\mathcal{H}^1 \\ &\leq \int_{\Omega \times \mathbb{S}^1} \varphi(x) h(\theta, \lambda_\gamma) \end{aligned} \quad (43)$$

thanks to (35) in Lemma 5.1. Then, (40), (42) and (43) yield

$$\int_{\partial E} \varphi(x) f(\kappa_E) d\mathcal{H}^1 \leq \int_{\Omega \times \mathbb{S}^1} \varphi(x) h(\theta, \sigma)$$

and Theorem 1 is easily deduced. In case  $f$  takes the value  $+\infty$ , then we first approximates  $f$  with a finite-valued function from below (replacing  $f$  with  $\min_{t'} f(t') + |t - t'|/\varepsilon$  for  $\varepsilon > 0$  small), and once the inequality is established for this function we send  $\varepsilon \rightarrow 0$ , so that (9) also holds.

**Acknowledgements** The authors would like to thank Prof. R. Duits (T.U. Eindhoven), as well as the referees for their accurate reading of the paper and their helpful and positive comments. They also wish to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme “Variational methods, new optimisation techniques and new fast numerical algorithms” (Sept.–Oct., 2017), when this paper was completed. This work was supported by: EPSRC Grant No. EP/K032208/1.

The work of A.C. was also partially supported by a Grant of the Simons Foundation. T.P. acknowledges support by the Austrian science fund (FWF) under the project EANOI, No. I1148 and the ERC starting Grant HOMOVIS, No. 640156.

## Appendix A: Consistency of the discretization

In this appendix, we study the consistency of the discrete approximation of the problem which is used in Sect. 4.

### A.1 Preliminary results

Let us introduce, for  $(s, t) \in \mathbb{R}^2$ ,

$$\bar{h}(s, t) = \begin{cases} sf(t/s) & \text{if } s > 0, \\ f^\infty(t) & \text{if } s = 0, \\ +\infty & \text{else,} \end{cases}$$

which is such that

$$\bar{h}(s, t) = \sup_{a+f^*(b) \leq 0} as + bt.$$

Observe that if the convex function  $f$  is differentiable, then for  $s > 0$ ,  $\partial_s \bar{h}(s, t) = f(t/s) - (t/s)f'(t/s) \leq f(0)$ .

Consider a vector-valued measurable function  $\sigma(\theta) = (\lambda(\theta)\underline{\theta}, \mu(\theta))$  where  $\lambda \geq 0$ . Let then  $\bar{\theta} \in \mathbb{S}^1$ ,  $\delta_\theta > 0$  and

$$\bar{\sigma} = \frac{1}{\delta_\theta} \int_{\bar{\theta} - \frac{\delta_\theta}{2}}^{\bar{\theta} + \frac{\delta_\theta}{2}} \sigma(\theta) d\theta.$$

We first observe that

$$\bar{\sigma} \cdot \bar{\theta} = \frac{1}{\delta_\theta} \int_{\bar{\theta} - \frac{\delta_\theta}{2}}^{\bar{\theta} + \frac{\delta_\theta}{2}} \lambda(\theta) \underline{\theta} \cdot \bar{\theta} d\theta \in \left[ \bar{\lambda} \cos \frac{\delta_\theta}{2}, \bar{\lambda} \right]$$

where

$$\bar{\lambda} = \frac{1}{\delta_\theta} \int_{\bar{\theta} - \frac{\delta_\theta}{2}}^{\bar{\theta} + \frac{\delta_\theta}{2}} \lambda(\theta) d\theta.$$

Assuming  $f$  is smooth, it follows that

$$\begin{aligned}\bar{h}\left(\frac{\bar{\sigma} \cdot \bar{\theta}}{\cos \frac{\delta_{\theta}}{2}}, \bar{\sigma}^{\theta}\right) &= \bar{h}(\bar{\lambda}, \bar{\sigma}^{\theta}) + \int_0^1 \partial_s \bar{h}\left(\bar{\lambda} + s\left(\frac{\bar{\sigma} \cdot \bar{\theta}}{\cos \frac{\delta_{\theta}}{2}} - \bar{\lambda}\right), \bar{\sigma}^{\theta}\right) \left(\frac{\bar{\sigma} \cdot \bar{\theta}}{\cos \frac{\delta_{\theta}}{2}} - \bar{\lambda}\right) ds \\ &\leq \bar{h}(\bar{\lambda}, \bar{\sigma}^{\theta}) + f(0) \bar{\lambda} \frac{1 - \cos \frac{\delta_{\theta}}{2}}{\cos \frac{\delta_{\theta}}{2}}.\end{aligned}$$

Using that  $\bar{h}$  is 1-homogeneous and that  $\bar{h}(s, t) \geq \gamma \sqrt{s^2 + t^2}$ , it follows

$$\begin{aligned}\bar{h}(\bar{\sigma} \cdot \bar{\theta}, \cos \frac{\delta_{\theta}}{2} \bar{\sigma}^{\theta}) &\leq \cos \frac{\delta_{\theta}}{2} \bar{h}(\bar{\lambda}, \bar{\sigma}^{\theta}) + (1 - \cos \frac{\delta_{\theta}}{2}) \frac{f(0)}{\gamma} \bar{h}(\bar{\lambda}, \bar{\sigma}^{\theta}) \\ &\leq \left(1 + \frac{\delta_{\theta}^2}{4} (f(0)/\gamma - 1)\right) \bar{h}(\bar{\lambda}, \bar{\sigma}^{\theta})\end{aligned}$$

if  $\delta_{\theta}$  is small enough. Observe that if  $f$  is not smooth, this still holds by approximation. It follows, thanks to Jensen's inequality and the fact that  $h(\theta, \sigma) = \bar{h}(\lambda, \sigma^{\theta})$  for all  $\theta$ , that provided  $\delta_{\theta}$  is small enough,

$$\bar{h}(\bar{\sigma} \cdot \bar{\theta}, \cos \frac{\delta_{\theta}}{2} \bar{\sigma}^{\theta}) \leq (1 + C \delta_{\theta}^2) \frac{1}{\delta_{\theta}} \int_{\bar{\theta} - \frac{\delta_{\theta}}{2}}^{\bar{\theta} + \frac{\delta_{\theta}}{2}} h(\theta, \sigma(\theta)) d\theta, \quad (44)$$

where the constant  $C = (f(0)/\gamma - 1)/4$  only depends on the function  $f$ .

**Remark A.1** If  $\sigma$  is a bounded measure such that  $\int h(\theta, \sigma) < +\infty$ , then (44) still holds with now all integrals and averages replaced with integrals over  $[\bar{\theta} - \delta_{\theta}/2, \bar{\theta} + \delta_{\theta}/2]$ . Indeed, in this case, approximating  $\sigma$  with smooth measures by convolution one easily deduces that it holds for almost all  $\bar{\theta}$  (whenever  $|\sigma|(\{\bar{\theta} - \delta_{\theta}/2, \bar{\theta} + \delta_{\theta}/2\}) = 0$ ). Then, for the other values, it is enough to find a sequence  $\varepsilon_n > 0$  with  $\varepsilon_n \downarrow 0$  such that the result holds for  $\bar{\theta} - \varepsilon_n$  and use the fact that for any bounded measure  $\mu$ ,

$$\lim_{n \rightarrow \infty} \mu([\bar{\theta} - \varepsilon_n - \frac{\delta_{\theta}}{2}, \bar{\theta} - \varepsilon_n + \frac{\delta_{\theta}}{2}]) = \mu([\bar{\theta} - \frac{\delta_{\theta}}{2}, \bar{\theta} + \frac{\delta_{\theta}}{2}])$$

as  $\mu([\bar{\theta} - \varepsilon_n, \bar{\theta})) \rightarrow 0$  as  $n \rightarrow \infty$  for any  $\bar{\theta} \in \mathbb{S}^1$ .

We can now show the following lemma:

**Lemma A.2** Let  $\delta_x, \delta_{\theta} > 0$  small enough,  $\sigma \in \mathcal{M}^1(\Omega \times \mathbb{S}^1; \mathbb{R}^3)$  with  $\operatorname{div} \sigma = 0$  and such that  $\int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) < \infty$ . Define, for  $(\bar{x}, \bar{\theta}) \in \Omega \times \mathbb{S}^1$ ,

$$S = \left[\bar{x}_1 - \frac{\delta_x}{2}, \bar{x}_1 + \frac{\delta_x}{2}\right) \times \left[\bar{x}_2 - \frac{\delta_x}{2}, \bar{x}_2 + \frac{\delta_x}{2}\right)$$

and

$$\bar{\sigma} = \frac{1}{\delta_x^2 \delta_{\theta}} \sigma(S \times [\bar{\theta} - \frac{\delta_{\theta}}{2}, \bar{\theta} + \frac{\delta_{\theta}}{2})).$$

Then,

$$\bar{h}(\bar{\sigma} \cdot \bar{\theta}, \cos \frac{\delta_\theta}{2} \bar{\sigma}^\theta) \leq \frac{1 + C\delta_\theta^2}{\delta_x^2 \delta_\theta} \int_{S \times [\bar{\theta} - \frac{\delta_\theta}{2}, \bar{\theta} + \frac{\delta_\theta}{2}]} h(\theta, \sigma) \quad (45)$$

where  $C$  depends only on  $f$ . Moreover,  $\bar{\sigma}^x/|\bar{\sigma}|$  lies in the cone  $\mathbb{R}_+(\theta - \delta_\theta/2) + \mathbb{R}_+(\theta + \delta_\theta/2)$ .

**Proof** If we introduce the (averaged) marginal  $\sigma' \in \mathcal{M}^1(\mathbb{S}^1; \theta)$  defined by  $\int_{\mathbb{S}^1} \psi \sigma' = \frac{1}{\delta_x^2} \int_{S \times \mathbb{S}^1} \psi \sigma$  for all  $\psi \in C^0(\mathbb{S}^1)$ , then one observes that

$$\bar{\sigma} = \frac{1}{\delta_\theta} \sigma'([\bar{\theta} - \frac{\delta_\theta}{2}, \bar{\theta} + \frac{\delta_\theta}{2}])$$

and

$$\int_{[\bar{\theta} - \frac{\delta_\theta}{2}, \bar{\theta} + \frac{\delta_\theta}{2}]} h(\theta, \sigma') \leq \int_{S \times [\bar{\theta} - \frac{\delta_\theta}{2}, \bar{\theta} + \frac{\delta_\theta}{2}]} h(\theta, \sigma).$$

This follows by a disintegration argument and using Jensen's inequality in each "slice" corresponding to a fixed value of  $\theta$ . The result then follows from (44), together with Remark A.1.

The last statement comes from the fact that  $\sigma_x$  is the average of measures all contained in the cone, which is convex.  $\square$

Consider now a measure  $\sigma$  admissible for some function  $u$ , and assume that  $\sigma$  is "smooth" in  $x$ : we assume for instance that it is the result of a convolution  $\rho_\varepsilon * \sigma'$  for some  $\sigma'$  admissible (possibly extended in a larger domain), with  $\rho_\varepsilon(x)$  a rotationally symmetric mollifier, as in the Proof of Proposition 3.1. In this case,  $x \mapsto \sigma$  can be seen as a  $\mathcal{M}^1(\mathbb{S}^1; \mathbb{R}^3)$ -valued smooth function.

Consider  $(\bar{x}, \bar{\theta}) \in \Omega \times \mathbb{S}^1$ ,  $\delta_x, \delta_\theta$  small, and define in the volume  $V = S \times [\bar{\theta} - \delta_\theta/2, \bar{\theta} + \delta_\theta/2]$  the average  $\bar{\sigma}$  as before and the "Raviart–Thomas" approximation of  $\sigma$  defined by the average fluxes through the 6 facets of  $V$  (linearly extended inside the volume, as in Eq. 21). There are several ways to define this properly, at least for all  $\bar{\theta}$  but a countable number. In our case, one can disintegrate the measure in  $\Omega \times \mathbb{S}^1$  as  $\sigma = \sigma_\theta d\mu$  where  $\mu$  is a bounded positive measure in  $\mathbb{S}^1$  and for all  $\theta, \sigma_\theta \in C^\infty(\Omega)$ . Then

$$\begin{aligned} \tau_\pm^1 &= \frac{1}{\delta_x \delta_\theta} \int_{[\bar{\theta} - \delta_\theta/2, \bar{\theta} + \delta_\theta/2]} \left( \int_{\bar{x}_2 - \frac{\delta_x}{2}}^{\bar{x}_2 + \frac{\delta_x}{2}} \sigma_\theta^1(\bar{x}_1 \pm \frac{\delta_x}{2}, x_2, \theta) dx_2 \right) \mu, \\ \tau_\pm^2 &= \frac{1}{\delta_x \delta_\theta} \int_{[\bar{\theta} - \delta_\theta/2, \bar{\theta} + \delta_\theta/2]} \left( \int_{\bar{x}_1 - \frac{\delta_x}{2}}^{\bar{x}_1 + \frac{\delta_x}{2}} \sigma_\theta^1(x_1, \bar{x}_2 \pm \frac{\delta_x}{2}, \theta) dx_1 \right) \mu. \end{aligned}$$

To define the vertical fluxes  $\tau_\pm^\theta$  we assume in addition that  $\mu(\{\bar{\theta} \pm \delta_\theta/2\}) = 0$  (which is true for all values but a countable number). In this case, observe that if  $\phi \in C_c^1(\mathbb{S} \times$

$\{\theta = -\delta_\theta/2\}$ ), it can be extended into a  $C^1$  function in  $V$  vanishing near the 5 other boundaries, and then

$$\int_V \nabla \phi \cdot \sigma$$

defines a measure  $\tilde{\tau}_-^\theta$  on  $\hat{S} \times \{\theta = -\delta_\theta/2\}$ . Then one simply let  $\tau_-^\theta = \tilde{\tau}_-^\theta(\hat{S} \times \{\theta = -\delta_\theta/2\})/\delta_x^2$ . The value  $\tau_+^\theta$  is defined in the same way. The assumption that  $\mu(\{\bar{\theta} - \delta_\theta/2\}) = 0$  guarantees that the same construction from below will build the same measure and the same value, and that one actually has  $(\tau_+^\theta - \tau_-^\theta + \tau_+^2 - \tau_-^2)/\delta_x + (\tau_+^\theta - \tau_-^\theta)/\delta_\theta = 0$ .

We can show the following lemma.

**Lemma A.3** *Let  $\tau = (\tau_a^1, \tau_b^2, \tau_c^\theta)^T$  for any  $(a, b, c) \in \{-, +\}^3$ . Then, for all  $\bar{\theta}$  but a countable number,*

$$\delta_x^2 \delta_\theta |\tau - \bar{\sigma}| \leq \sqrt{\delta_x^2 + \delta_\theta^2} \int_V |\partial_1 \sigma^1| + |\partial_2 \sigma^2|. \quad (46)$$

**Proof** We prove the result for  $(a, b, c) = (-, -, -)$ , the proof in the other cases being identical. We first assume that  $\sigma$  is also  $C^1$  in  $\theta$  (which can be achieved by convolution). In this case, one has for all  $(x_1, x_2, \theta) \in V$ ,

$$\sigma^1(\bar{x}_1 - \frac{\delta_x}{2}, x_2, \theta) = \sigma^1(x_1, x_2, \theta) - \int_{\bar{x}_1 - \frac{\delta_x}{2}}^{x_1} \partial_1 \sigma^1(s, x_2, \theta) ds$$

so that

$$\begin{aligned} & \delta_x \sigma^1(\bar{x}_1 - \frac{\delta_x}{2}, x_2, \theta) \\ &= \int_{\bar{x}_1 - \frac{\delta_x}{2}}^{\bar{x}_1 + \frac{\delta_x}{2}} \sigma^1(x_1, x_2, \theta) - \int_{\bar{x}_1 - \frac{\delta_x}{2}}^{\bar{x}_1 + \frac{\delta_x}{2}} (\bar{x}_1 + \frac{\delta_x}{2} - s) \partial_1 \sigma^1(s, x_2, \theta) ds \end{aligned}$$

Averaging over  $x_2, \theta$ , we deduce that

$$\tau_-^1 = \bar{\sigma}^1 - \frac{1}{\delta_x^2 \delta_\theta} \int_V (\bar{x}_1 + \frac{\delta_x}{2} - x_1) \partial_1 \sigma^1 dx_1 dx_2 d\theta.$$

In the same way,

$$\begin{aligned} \tau_-^2 &= \bar{\sigma}^2 - \frac{1}{\delta_x^2 \delta_\theta} \int_V (\bar{x}_2 + \frac{\delta_x}{2} - x_2) \partial_2 \sigma^2 dx_1 dx_2 d\theta, \\ \tau_-^\theta &= \bar{\sigma}^\theta - \frac{1}{\delta_x^2 \delta_\theta} \int_V (\bar{\theta} + \frac{\delta_\theta}{2} - \theta) \partial_\theta \sigma^\theta dx_1 dx_2 d\theta. \end{aligned}$$

Using that  $\operatorname{div} \sigma = 0$ , the latter can be rewritten

$$\tau_-^\theta = \bar{\sigma}^\theta + \frac{1}{\delta_x^2 \delta_\theta} \int_V (\bar{\theta} + \frac{\delta_\theta}{2} - \theta) \operatorname{div}_x \sigma^x dx_1 dx_2 d\theta.$$

The estimate (46) follows. If  $\sigma$  is not  $C^1$  in  $\theta$ , as before we can smooth  $\sigma$ , then in the limit we will obtain (46) for all  $\bar{\theta}$  such that  $\mu(\{\bar{\theta} \pm \delta_\theta/2\}) = 0$ .  $\square$

**Corollary A.4** *Let  $\tau(x, \theta)$  be the Raviart–Thomas extension of the fluxes  $\tau_\pm^\bullet$  in  $V$ : then it holds*

$$\int_V |\tau - \bar{\sigma}| dx d\theta \leq \sqrt{\delta_x^2 + \delta_\theta^2} \int_V |\partial_1 \sigma^1| + |\partial_2 \sigma^2| \quad (47)$$

(for the same values of  $\bar{\theta}$ ).

This is proven in the same way, as inside  $V$ ,  $\tau^\bullet(x, \theta)$  is a convex combination of the two fluxes  $\tau_\pm^\bullet$ . Moreover, by construction since  $\operatorname{div} \sigma = 0$ , it is easy to check that one also has  $\operatorname{div} \tau = 0$ . The following is also immediate:

**Corollary A.5** *Let  $\tau(x, \theta)$  be the Raviart–Thomas extension of the fluxes  $\tau_\pm^\bullet$  in  $V$  and  $\bar{\tau}$  the value in the middle of the cell (in other words,*

$$\bar{\tau} = \begin{pmatrix} \frac{\tau_-^1 + \tau_+^1}{2} \\ \frac{\tau_-^2 + \tau_+^2}{2} \\ \frac{\tau_-^\theta + \tau_+^\theta}{2} \end{pmatrix}$$

*is given by the average of the fluxes through the facets of  $V$ ). Then*

$$|V| |\bar{\tau} - \bar{\sigma}| \leq \sqrt{\delta_x^2 + \delta_\theta^2} \int_V |\partial_1 \sigma^1| + |\partial_2 \sigma^2|. \quad (48)$$

## A.2 Consistent discretization of the energy $F$

We now are in a position to define almost consistent approximations of  $F$ . We will build a discrete approximation which enjoys a sort of discrete-to-continuum  $\Gamma$ -convergence property to the limiting functional  $F$ .

Assume to simplify  $\Omega$  is a convex set,<sup>2</sup> and even a rectangle  $[0, a] \times [0, b]$ ,  $a, b > 0$ , which further simplifies our notation.

Let  $u \in BV(\Omega)$  and  $\sigma_u$  be admissible for  $u$ , such that  $F(u) = \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma_u) < \infty$  and first, for  $\varepsilon > 0$  fixed,  $\sigma_\varepsilon$  by convolution as in the Proof of Proposition 3.1. In particular,

$$\int_{\Omega \times \mathbb{S}^1} |\partial_1 \sigma_\varepsilon^1| + |\partial_2 \sigma_\varepsilon^2| \leq \frac{c}{\varepsilon} \int_{\Omega \times \mathbb{S}^1} |\sigma_u^1| + |\sigma_u^2|$$

<sup>2</sup> This is not really important, as one could approximate  $\sigma$  by smooth function only inside  $\Omega$  and then let the corresponding set invade  $\Omega$  in the limit.

where  $c = 2\pi \int_{B_1} |\nabla \rho| dx$  depends only on the convolution kernel  $\rho$ . Fix  $\delta_x, \delta_\theta$  small enough, assume  $\delta_\theta = 2\pi/N_\theta$  for some integer  $N_\theta$ , and consider all the volumes  $V_{i,j,k}$ , defined in (16) [with  $S_{i,j}$  defined by (15)], and which are inside  $\Omega \times \mathbb{S}^1$ , for  $(i, j, k) \in \mathcal{J}$  (17). We define a Raviart–Thomas vector field from the (averaged) fluxes of  $\sigma_\varepsilon$  through the facets of the volumes:  $\sigma_{i-\frac{1}{2},j,k}^1$  through the facets  $\mathcal{F}_{i-\frac{1}{2},j,k}^1$ ,  $\sigma_{i,j-\frac{1}{2},k}^2$  through the facets  $\mathcal{F}_{i,j-\frac{1}{2},k}^2$ , and  $\sigma_{i,j,k-\frac{1}{2}}^\theta$  through  $\mathcal{F}_{i,j,k-\frac{1}{2}}^\theta$ , see (18), (19), (20). The latter flux is well-defined up to an infinitesimal vertical translation of the origin of the discretization in  $\theta$  (without loss of generality we thus assume it is well defined). The Raviart–Thomas field inside the cube is defined then as in (21). We also define  $\bar{\sigma}_{i,j,k} = (\delta_x^{-2} \delta_\theta^{-1}) \int_{V_{i,j,k}} \sigma_\varepsilon$  as the average of  $\sigma_\varepsilon$  in  $V_{i,j,k}$ , and let  $\hat{\sigma}_{i,j,k}$  be defined by (24), which corresponds to averaging the fluxes of the facets, or equivalently to consider the value of the Raviart–Thomas extension in the middle of the volume  $V_{i,j,k}$ .

From Corollary A.5, letting

$$e_{i,j,k} = \hat{\sigma}_{i,j,k} - \bar{\sigma}_{i,j,k},$$

one has

$$\delta_x^2 \delta_\theta \sum_{i,j,k} |e_{i,j,k}| \leq c \frac{\sqrt{\delta_x^2 + \delta_\theta^2}}{\varepsilon} \int_{\Omega \times \mathbb{S}^1} |\sigma_u^x|.$$

Moreover by Lemma A.2, one has (we denote, for every  $k$ ,  $\theta_k = k\delta_\theta$  and  $\theta_{k+\frac{1}{2}} = (k + \frac{1}{2})\delta_\theta$ )

$$\delta_x^2 \delta_\theta \sum_{i,j,k} \bar{h}(\bar{\sigma}_{i,j,k} \cdot \underline{\theta}_k, \cos \frac{\delta_\theta}{2} \bar{\sigma}_{i,j,k}^\theta) \leq (1 + C\delta_\theta^2) \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma_u)$$

(using that  $\int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma_\varepsilon) \leq \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma_u)$ , cf the Proof of Proposition 3.1).

Eventually, letting now

$$e'_{i,j,k} = e_{i,j,k} + (1 - \cos \frac{\delta_\theta}{2}) \bar{\sigma}_{i,j,k}^\theta,$$

which is such that

$$\begin{aligned} \delta_x^2 \delta_\theta \sum_{i,j,k} |e'_{i,j,k}| &\leq c \frac{\sqrt{\delta_x^2 + \delta_\theta^2}}{\varepsilon} \int_{\Omega \times \mathbb{S}^1} |\sigma_u^x| + \delta_\theta^2 \int_{\Omega \times \mathbb{S}^1} |\sigma_u^\theta| \\ &\leq c \frac{\sqrt{\delta_x^2 + \delta_\theta^2}}{\varepsilon} \int_{\Omega \times \mathbb{S}^1} |\sigma_u| \leq c \frac{\sqrt{\delta_x^2 + \delta_\theta^2}}{\gamma \varepsilon} F(u). \end{aligned} \quad (49)$$

We deduce that we can find an center-averaged Raviart–Thomas field  $\hat{\sigma}$  and an error term  $e'$  such that (49) holds and

$$\sum_{i,j,k} \bar{h}((\hat{\sigma}_{i,j,k} - e'_{i,j,k}) \cdot \underline{\theta}_k, (\hat{\sigma}_{i,j,k} - e'_{i,j,k})^\theta) \leq (1 + C\delta_\theta^2)F(u)$$

where  $\theta^{\delta_\theta} = \sum_k \theta_k \chi_{\{k\delta_\theta, (k+1)\delta_\theta\}}$ . In particular if we introduce, for  $\delta = (\delta_x, \delta_t)$  small, the inf-convolution

$$\bar{h}_\delta(s, t) = \min_{s', t'} \bar{h}(s - s', t - t') + \frac{\gamma}{(\delta_x^2 + \delta_\theta^2)^{1/4}} \sqrt{s'^2 + t'^2} \quad (50)$$

we find that

$$\delta_x^2 \delta_\theta \sum_{i,j,k} \bar{h}_\delta(\hat{\sigma}_{i,j,k} \cdot \underline{\theta}_k, \hat{\sigma}_{i,j,k}^\theta) \leq (1 + C\delta_\theta^2 + \frac{c}{\varepsilon}(\delta_x^2 + \delta_\theta^2)^{1/4})F(u).$$

Moreover, one easily sees that

$$\hat{\sigma}_{i,j,k}^x \in \mathbb{R} + \underline{\theta}_{k-\frac{1}{2}} + \mathbb{R} + \underline{\theta}_{k+\frac{1}{2}}. \quad (51)$$

**Remark A.6** If  $\bar{h}$  is  $L$ -Lipschitz (as it is the case when  $f$  as growth one, for instance if  $f(t) = \gamma\sqrt{1+t^2}$ ), then the inf-convolution step is not necessary. One directly obtains

$$\delta_x^2 \delta_\theta \sum_{i,j,k} \bar{h}(\hat{\sigma}_{i,j,k} \cdot \underline{\theta}_k, \hat{\sigma}_{i,j,k}^\theta) \leq \left(1 + C\delta_\theta^2 + \frac{cL}{\gamma\varepsilon} \sqrt{\delta_x^2 + \delta_\theta^2}\right)F(u).$$

Now, we check the consistency between  $\hat{\sigma}$  and  $u$ . By construction,  $\delta_x \delta_\theta \sum_k \sigma_{i+\frac{1}{2},j,k}^1$  is the flux of  $Du^\perp$  through the edge  $\{i + \frac{1}{2}\delta_x\} \times [(j - \frac{1}{2})\delta_x, (j + \frac{1}{2})\delta_x]$  in  $\Omega$ , hence it is equal to the value  $u((i + \frac{1}{2})\delta_x, (j + \frac{1}{2})\delta_x) - u((i + \frac{1}{2})\delta_x, (j - \frac{1}{2})\delta_x)$ . Accordingly, if we let, for all  $i, j$ ,  $u_{i+\frac{1}{2},j+\frac{1}{2}}^\delta := u((i + \frac{1}{2})\delta_x, (j + \frac{1}{2})\delta_x)$ , we obtain that (23) holds (with  $u$  replaced with  $u^\delta$ ).

Eventually we observe that the free divergence condition simply translates as (22) for all admissible  $i, j, k$ , as this is the global flux of  $\sigma_\varepsilon$  across the boundaries of  $V_{i,j,k}$ .

It is now easy to deduce the following upper approximation result:

**Proposition A.7** *Let  $u \in BV(\Omega)$ ,  $\sigma$  be admissible for  $u$  and such that*

$$F(u) = \int_{\Omega \times \mathbb{S}^1} h(\theta, \sigma) < \infty.$$

*Then for  $\delta = (\delta_x, \delta_\theta) \rightarrow 0$  one can find a discrete field  $(\sigma_{i+\frac{1}{2},j,k}^1, \sigma_{i,j+\frac{1}{2},k}^2, \sigma_{i,j,k+\frac{1}{2}}^\theta)$  and a discrete image  $u_{i+\frac{1}{2},j+\frac{1}{2}}^\delta$  with*

$$\sum_{i,j} u_{i+\frac{1}{2},j+\frac{1}{2}}^\delta \chi_{[i\delta_x, (i+1)\delta_x] \times [j\delta_x, (j+1)\delta_x]} \rightarrow u \quad (52)$$



(strongly in  $L^2(\Omega)$ ) and such that for all  $i, j$ , (23) holds, for all  $i, j, k$ , (51) and (22) hold, and:

$$\limsup_{\delta \rightarrow 0} \delta_x^2 \delta_\theta \sum_{i,j,k} \bar{h}_\delta(\hat{\sigma}_{i,j,k} \cdot \underline{\theta}_k, \hat{\sigma}_{i,j,k}^\theta) \leq F(u) \quad (53)$$

where  $\hat{\sigma}$  is defined by (24), and where  $\bar{h}_\delta$  is defined in (50) (or is  $\bar{h}$  in case it is Lipschitz).

To show that the discretization is consistent, we must now show a similar lower bound: namely that given any  $u$  and  $u^\delta, \sigma, \hat{\sigma}$  which satisfy (24), (51), (23), and (52) (weakly, for instance as distributions), then one has

$$\liminf_{\delta \rightarrow 0} \delta_x^2 \delta_\theta \sum_{i,j,k} \bar{h}_\delta(\hat{\sigma}_{i,j,k} \cdot \underline{\theta}_k, \hat{\sigma}_{i,j,k}^\theta) \geq F(u). \quad (54)$$

A first obvious remark is that the field

$$\hat{\sigma}^\delta := \sum_{i,j,k} \hat{\sigma}_{i,j,k} \chi_{V_{i,j,k}}$$

is bounded in measure, and hence, up to subsequences, converges (weakly-\*) to a measure  $\sigma$ . It is then easy to deduce from (51) and the convexity of  $\bar{h}$  that

$$\int_{\Omega \times \mathbb{S}^1} h(\theta, u) \leq \liminf_{\delta \rightarrow 0} \delta_x^2 \delta_\theta \sum_{i,j,k} \bar{h}_\delta(\hat{\sigma}_{i,j,k} \cdot \underline{\theta}_{k+\frac{1}{2}}, \hat{\sigma}_{i,j,k}^\theta).$$

One can also check that  $\text{div } \sigma = 0$  by passing to the limit in (22) [after a suitable integration against a smooth test function, exactly as in (55) below]. Hence it is enough to show that the limiting  $\sigma$  is compatible with  $u$ .

But this is quite obvious from (23), which one can integrate against a smooth test function, then “integrate by part” before passing to the limit. More precisely, for  $\varphi \in C_c^\infty(\Omega)$ , one has (dropping the superscripts  $\delta$  and denoting  $\varphi_{i,j} = (1/\delta_x^2) \int_{S_{i,j}} \varphi(x) dx$ ):

$$\begin{aligned} & \int_{\Omega \times \mathbb{S}^1} \varphi(x) \hat{\sigma}^1 dx d\theta \\ &= \sum_{i,j,k} \hat{\sigma}_{i,j,k}^1 \int_{V_{i,j,k}} \varphi(x) dx d\theta \\ &= \delta_x^2 \delta_\theta \sum_{i,j} \varphi_{i,j} \sum_k \delta_\theta \frac{\sigma_{i+\frac{1}{2},j,k}^1 + \sigma_{i-\frac{1}{2},j,k}^1}{2} \\ &= \frac{\delta_x^2 \delta_\theta}{2} \sum_{i,j} \varphi_{i,j} \left( \frac{u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{1}{2},j-\frac{1}{2}}}{\delta_x} + \frac{u_{i-\frac{1}{2},j+\frac{1}{2}} - u_{i-\frac{1}{2},j-\frac{1}{2}}}{\delta_x} \right) \end{aligned}$$

$$\begin{aligned}
&= -\delta_x^2 \delta_\theta \sum_{i,j} \frac{u_{i+\frac{1}{2},j+\frac{1}{2}} + u_{i-\frac{1}{2},j+\frac{1}{2}}}{2} \frac{\varphi_{i,j+1} - \varphi_{i,j}}{\delta_x} \\
&\rightarrow - \int_{\Omega} u(x) \partial_2 \varphi(x) dx
\end{aligned} \tag{55}$$

as  $\delta \rightarrow 0$ .

Eventually, we need to show a compactness property, which is that if

$$\sup_{\delta} \delta_x^2 \delta_\theta \sum_{i,j,k} \bar{h}(\hat{\sigma}_{i,j,k} \cdot \underline{\theta}_k, \hat{\sigma}_{i,j,k}^\theta) < +\infty \tag{56}$$

the discrete image  $u^\delta$  which is recovered from (23) (up to a constant) converges to a  $u(x)$ ,  $x \in \Omega$  (in a weak sense which will be made clear). The point here is that a priori, from (56) and (1), one has only

$$\begin{aligned}
&\sup_{\delta} \delta_x^2 \sum_{i,j} \left( \left( \frac{u_{i+\frac{1}{2},j+\frac{1}{2}}^\delta - u_{i+\frac{1}{2},j-\frac{1}{2}}^\delta}{\delta_x} + \frac{u_{i-\frac{1}{2},j+\frac{1}{2}}^\delta - u_{i-\frac{1}{2},j-\frac{1}{2}}^\delta}{\delta_x} \right)^2 \right. \\
&\quad \left. + \left( \frac{u_{i+\frac{1}{2},j-\frac{1}{2}}^\delta - u_{i-\frac{1}{2},j-\frac{1}{2}}^\delta}{\delta_x} + \frac{u_{i+\frac{1}{2},j+\frac{1}{2}}^\delta - u_{i-\frac{1}{2},j+\frac{1}{2}}^\delta}{\delta_x} \right)^2 \right)^{\frac{1}{2}} < +\infty
\end{aligned} \tag{57}$$

so that the discrete total variation of  $u^\delta$  is a priori not well controlled. However, one can easily check that the kernel of the operator which appears in the energy (57) is two-dimensional, and made of the oscillating discrete images

$$v_{i-\frac{1}{2},j-\frac{1}{2}}^\delta = \alpha + \beta(-1)^{i+j}, \tag{58}$$

$\alpha, \beta \in \mathbb{R}^2$ . Hence it is possible to show that one can decompose  $u^\delta$  as a sum of a non-oscillating function with zero average  $\bar{u}^\delta$  and an oscillation  $v^\delta$ , and obtain a strong control on the discrete total variation of  $\bar{u}^\delta$ . Therefore one easily deduce that any suitably built continuous extension of  $\bar{u}^\delta$  will converge to some  $u$  strongly in  $L^p(\Omega)$ , for any  $p < 2$  (as  $BV(\Omega)$  is compactly embedded in such spaces), and weakly in  $L^2(\Omega)$ .

In addition, any control on the average of  $u^\delta$  and on its oscillation (which cannot be given by (57) and has to come from other terms in the energy, such as a boundary condition or a penalization: note that it is enough to control two adjacent pixels) will ensure in addition that  $v^\delta$  remains bounded and converges (only weakly in  $L^p(\Omega)$ , if the control is only on the  $L^p$  norm, however in this case it is obvious that the oscillating term in (58) goes to zero and  $v^\delta$  can only go to a constant).

To sum up, we have shown the following.

**Proposition A.8** For  $\delta \rightarrow 0$ , assume we are given  $\sigma^\delta$ ,  $u^\delta$  and  $\hat{\sigma}^\delta$  with (24), (23), which in addition satisfy (51) and (22), and (56). Then, up to an oscillating function  $v^\delta$  of the form (58), there is  $u \in BV(\Omega)$  such that  $u^\delta \rightarrow u$ , and (54) holds.

**Remark A.9** In practice, we did not use the inf-convolutions  $\bar{h}_\delta$  (only  $\bar{h}$ ) in our discrete scheme. Also, we replaced the constraint (51) with the stronger constraint  $\hat{\sigma}_{i,j,k}^x \in \mathbb{R}_{+\theta_k}$ , after having experimentally observed that there was no qualitative difference in the output. It seems the results we compute are still consistent with what is expected from the energy.

## B. Smirnov's theorem in $\Omega \times \mathbb{S}^1$

In this whole paper  $\Omega \subset \mathbb{R}^2$  is assumed to be a Lipschitz set. In particular, locally its boundary can be represented as the subgraph  $\{(x, y) : y < h(x)\}$  of a Lipschitz function  $h$ . Consider a ball  $B$  where this representation holds and assume first  $h$  is  $C^1$ , then one can extend in  $B$  a bounded Radon measure  $\sigma$  with  $\text{div } \sigma = 0$  into  $\tilde{\sigma}$  defined (for  $\psi \in C_c^0(B; \mathbb{R}^2)$ )

$$\int_B \tilde{\sigma} \cdot \psi := \int_{B \cap \Omega} \sigma \cdot \left( \psi(x, y) - \begin{pmatrix} 1 & 2h'(x) \\ 0 & -1 \end{pmatrix} \psi(x, 2h(x) - y) \right).$$

Then, it is standard that  $\text{div } \tilde{\sigma} = 0$  in  $B$ , indeed, if  $\varphi \in C_c^1(B)$ , one has that

$$\int_B \tilde{\sigma} \cdot \nabla \varphi = \int_{B \cap \Omega} \sigma \cdot \nabla [\varphi(x, y) - \varphi(x, 2h(x) - y)].$$

The function  $\varphi^s(x, y) := \varphi(x, y) - \varphi(x, 2h(x) - y)$  is  $C^1$  and vanishes on  $\partial\Omega$ , hence this expression is zero: Indeed if for  $\tau > 0$  one lets  $\varphi_\tau^s(x, y) = S_\tau(\varphi^s(x, y))$  where  $S_\tau \in C^\infty(\mathbb{R})$  is a smooth approximation of a “shrinkage operator”:

$$S_\tau(t) = \begin{cases} t - \tau & \text{if } t \geq \frac{3}{2}\tau, \\ 0 & \text{if } |t| < \frac{1}{2}\tau, \\ t + \tau & \text{if } t \leq -\frac{3}{2}\tau, \end{cases}$$

with smooth and 1-Lipschitz interpolation in  $\pm[\tau/2, 3\tau/2]$ , then  $\varphi_\tau^s \in C_c^1(B \cap \Omega)$  so that

$$\int_{B \cap \Omega} \sigma \cdot \nabla \varphi_\tau^s = 0$$

and, using  $\nabla \varphi_\tau^s = S'_\tau(\varphi^s) \nabla \varphi^s$

$$\int_{B \cap \Omega} \sigma \cdot (\nabla \varphi^s - \nabla \varphi_\tau^s) \leq C|\sigma|(B \cap \Omega \cap \{0 < |\varphi^s| < 3\tau/2\}) \rightarrow 0$$

as  $\tau \rightarrow 0$ , showing our claim. Hence  $\operatorname{div} \tilde{\sigma} = 0$ . If  $h$  is not  $C^1$  but just Lipschitz, one can approximate it from below by smooth functions  $h_n$ , build in such a way a sequence  $\sigma_n$  of extensions of  $\sigma|_{\{y < h_n(x)\}}$  and pass to the limit to deduce that the extension still exists.

Using cut-off functions, one can therefore assume that  $\sigma$  can be extended into a field  $\tilde{\sigma}$  which is a measure in  $\mathbb{R}^2$  with free divergence in a neighborhood of  $\Omega$ .

A similar construction would allow to extend a field  $\sigma \in \mathcal{M}(\Omega \times \mathbb{R}^2; \mathbb{R}^4)$  to  $\mathcal{M}(\mathbb{R}^4; \mathbb{R}^4)$  with free divergence (either in a neighborhood of  $\Omega \times \mathbb{R}^4$ , or even everywhere). This remark allows to localize Smirnov's theorems in [73].

Consider indeed now a free divergence field  $\sigma \in \mathcal{M}(\Omega \times \mathbb{S}^1; \mathbb{R}^3)$ . It can be seen, after extension, as a field in  $\mathcal{M}(\mathbb{R}^2 \times \mathbb{R}^2; \mathbb{R}^4)$  with  $\operatorname{spt} \sigma \subseteq \mathbb{R}^2 \times \mathbb{S}^1$ .

As in Smirnov's paper [73], for  $l > 0$  we introduce  $\mathcal{C}_l$  the set of oriented curves  $\gamma$  in  $\mathbb{R}^4$  with length  $l$ , with the topology corresponding to the weak convergence of the measures  $\tau_\gamma \mathcal{H}^1 \llcorner \gamma$ . Then, thanks to [73, Theorem A],  $\sigma$  can be decomposed as

$$\sigma = \int_{\mathcal{C}_l} \lambda d\mu(\lambda), \quad |\sigma| = \int_{\mathcal{C}_l} |\lambda| d\mu(\lambda),$$

for some measure  $\mu$  on  $\mathcal{C}_l$ . Moreover thanks to Remark 5 in [73],  $\mu$ -a.e. curve in the decomposition lies in  $\mathbb{R}^2 \times \mathbb{S}^1$ .

For this work, it is enough to consider  $l = 1$ . Moreover, if we restrict then all these measures to  $\Omega \times \mathbb{S}^1$  (and take for  $\mu$  the corresponding marginal), we get a decomposition on curves of length less or equal to 1 (possibly entering/exiting the domain). By a slight abuse of notation we still denote  $\mathcal{C}_1$  such a set of curves. One finds that

$$\sigma = \int_{\mathcal{C}_1} \lambda d\mu(\lambda), \quad |\sigma| = \int_{\mathcal{C}_1} |\lambda| d\mu(\lambda), \quad (59)$$

with now  $\lambda \in \mathcal{C}_1$ , the curves of length at most one in  $\Omega \times \mathbb{S}^1$ .

**Remark B.1** Theorem B in [73] is a more precise statement. It shows that one can obtain a similar decomposition with now curves  $\lambda$  with  $\operatorname{div} \lambda = 0$  a.e.: being either finite curves entering and exiting the domain, or “elementary solenoids”, which are objects of the form

$$\begin{aligned} \lambda &= \mathcal{M} - \lim_{k \rightarrow \infty} \frac{1}{2k} f_{\#} \overrightarrow{[-k, k]} \llcorner \Omega \times \mathbb{S}^1 \\ \operatorname{Lip}(f) &\leq 1 \\ \operatorname{var}(\lambda) &= 1 \\ f(\mathbb{R}) &\subset \operatorname{spt}(\lambda), \end{aligned}$$

(in particular one should have  $|f'(t)| = 1$  a.e.), meaning that for any  $\varphi \in C_c^1(\Omega \times \mathbb{S}^1)$ ,

$$\lambda(\varphi) = \lim_{s \rightarrow \infty} \frac{1}{2s} \int_{-s}^s \langle f'(t), \varphi(f(t)) \rangle dt. \quad (60)$$

This expresses that either  $\lambda$  is defined by the closed curve  $f(\mathbb{R})$  (if  $f$  is periodic), or  $\lambda$  is a limit of curves which densify and do not loose mass in the limit. We do not need such a precise result for our construction.

## References

1. Acerbi, E., Mucci, D.: Curvature-dependent energies. *Milan J. Math.* **85**(1), 41–69 (2017)
2. Acerbi, E., Mucci, D.: Curvature-dependent energies: the elastic case. *Nonlinear Anal.* **153**, 7–34 (2017)
3. Ambrosio, L., Fusco, N., Pallara, D.: *Functions of Bounded Variation and Free Discontinuity Problems*. The Clarendon Press/Oxford University Press, New York (2000)
4. Ambrosio, L., Masnou, S.: A direct variational approach to a problem arising in image reconstruction. *Interfaces Free Bound.* **5**(1), 63–81 (2003)
5. Anzellotti, G.: Functionals depending on curvatures. *Rend. Sem. Mat. Univ. Politec. Torino (Special Issue)*:47–62 (1990), 1989. Conference on Partial Differential Equations and Geometry (Torino, 1988)
6. Anzellotti, G., Serapioni, R., Tamanini, I.: Curvatures, functionals, currents. *Indiana Univ. Math. J.* **39**(3), 617–669 (1990)
7. Anzellotti, G., Delladio, S.: Minimization of functionals of curvatures and the Willmore problem. In: *Advances in Geometric Analysis and Continuum Mechanics* (Stanford, CA, 1993), pp. 33–43. International Press Press, Cambridge (1995)
8. Bae, E., Tai, X.-C., Zhu, W.: Augmented Lagrangian method for an Euler's elastica based segmentation model that promotes convex contours. *Inverse Probl. Imaging* **11**(1), 1–23 (2017)
9. Ballester, C., Bertalmio, M., Caselles, V., Sapiro, G., Verdera, J.: Filling-in by joint interpolation of vector fields and gray levels. *IEEE Trans. Image Process.* **10**(8), 1200–1211 (2001)
10. Ballester, C., Caselles, V., Verdera, J.: Disocclusion by joint interpolation of vector fields and gray levels. *Multiscale Model. Simul.* **2**(1), 80–123 (2003)
11. Bekkers, E.J., Duits, R., Mashtakov, A., Sanguinetti, G.R.: A PDE approach to data-driven sub-Riemannian geodesics in  $SE(2)$ . *SIAM J. Imaging Sci.* **8**(4), 2740–2770 (2015)
12. Bellettini, G., Dal Maso, G., Paolini, M.: Semicontinuity and relaxation properties of a curvature depending functional in 2D. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **20**(2), 247–297 (1993)
13. Bellettini, G., Mugnai, L.: Characterization and representation of the lower semicontinuous envelope of the elastica functional. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21**(6), 839–880 (2004)
14. Bellettini, G., Mugnai, L.: On the approximation of the elastica functional in radial symmetry. *Calc. Var. Partial Differ. Equ.* **24**(1), 1–20 (2005)
15. Bellettini, G., Mugnai, L.: A varifolds representation of the relaxed elastica functional. *J. Convex Anal.* **14**(3), 543–564 (2007)
16. Bellettini, G., March, R.: An image segmentation variational model with free discontinuities and contour curvature. *Math. Models Methods Appl. Sci.* **14**(1), 1–45 (2004)
17. Bellettini, G., March, R.: Asymptotic properties of the Nitzberg–Mumford variational model for segmentation with depth. In: *Free Boundary Problems. Volume 154 of International Series of Numerical Mathematics*, pp. 75–84. Birkhäuser, Basel (2007)
18. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 417–424. ACM Press/Addison-Wesley Publishing Co. (2000)
19. Boscain, U., Chertovskih, R.A., Gauthier, J.P., Remizov, A.O.: Hypoelliptic diffusion and human vision: a semidiscrete new twist. *SIAM J. Imaging Sci.* **7**(2), 669–695 (2014)
20. Boscain, U., Duits, R., Rossi, F., Sachkov, Y.: Curve cusplless reconstruction via sub-Riemannian geometry. *ESAIM Control Optim. Calc. Var.* **20**(3), 748–770 (2014)
21. Bouchitté, G., Valadier, M.: Integral representation of convex functionals on a space of measures. *J. Funct. Anal.* **80**(2), 398–420 (1988)
22. Bredies, K., Pock, T., Wirth, B.: Convex relaxation of a class of vertex penalizing functionals. *J. Math. Imaging Vis.* **47**(3), 278–302 (2013)
23. Bredies, K., Pock, T., Wirth, B.: A convex, lower semicontinuous approximation of Euler's elastica energy. *SIAM J. Math. Anal.* **47**(1), 566–613 (2015)

24. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145 (2011)
25. Chambolle, A., Pock, T.: On the ergodic convergence rates of a first-order primal-dual algorithm. *Math. Program.* **159**(1–2, Ser. A), 253–287 (2016)
26. Chan, T.F., Kang, S.H., Shen, J.: Euler's elastica and curvature-based inpainting. *SIAM J. Appl. Math.* **63**(2), 564–592 (2002)
27. Chan, T.F., Shen, J.: Mathematical models for local nontexture inpaintings. *SIAM J. Appl. Math.* **62**(3), 1019–1043 (2001/2002)
28. Chen, D., Mirebeau, J.-M., Cohen, L.D.: Global minimum for a Finsler elastica minimal path approach. *Int. J. Comput. Vis.* **122**(3), 458–483 (2017)
29. Citti, G., Franceschiello, B., Sanguinetti, G., Sarti, A.: Sub-Riemannian mean curvature flow for image processing. *SIAM J. Imaging Sci.* **9**(1), 212–237 (2016)
30. Dayrens, F., Masnou, S., Novaga, M.: Existence, regularity and structure of confined elasticae. In: *ESAIM:COCV, 2017 (to appear)*
31. Delladio, S.: Minimizing functionals depending on surfaces and their curvatures: a class of variational problems in the setting of generalized Gauss graphs. *Pac. J. Math.* **179**(2), 301–323 (1997)
32. Delladio, S.: Special generalized Gauss graphs and their application to minimization of functionals involving curvatures. *J. Reine Angew. Math.* **486**, 17–43 (1997)
33. Duits, R., Boscaïn, U., Rossi, F., Sachkov, Y.: Association fields via cusplless sub-Riemannian geodesics in SE(2). *J. Math. Imaging Vis.* **49**(2), 384–417 (2014)
34. Duits, R., Meesters, S.P.L., Mirebeau, J.-M., Portegies, J. M.: Optimal Paths for Variants of the 2D and 3D Reeds-Shepp Car with Applications in Image Analysis. *ArXiv e-prints* (2016)
35. Ekeland, I., Témam, R.: *Convex Analysis and Variational Problems*, Volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition (1999). Translated from the French
36. El-Zehiry, N., Grady, L.: Optimization of weighted curvature for image segmentation. *arXiv preprint arXiv:1006.4175* (2010)
37. El-Zehiry, N.Y., Grady, L.: Fast global optimization of curvature. In: *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3257–3264. IEEE (2010)
38. El-Zehiry, N.Y., Grady, L.: Contrast driven elastica for image segmentation. *IEEE Trans. Image Process.* **25**(6), 2508–2518 (2016)
39. Esedoglu, S., March, R.: Segmentation with depth but without detecting junctions. *J. Math. Imaging Vis.* **18**(1), 7–15 (2003). Special issue on imaging science (Boston, MA, 2002)
40. Esedoglu, S., Shen, J.: Digital inpainting based on the Mumford–Shah–Euler image model. *Eur. J. Appl. Math.* **13**(4), 353–370 (2002)
41. Franken, E., Duits, R.: Crossing-preserving coherence-enhancing diffusion on invertible orientation scores. *Int. J. Comput. Vis.* **85**(3), 253 (2009)
42. Glowinski, R.: ADMM and non-convex variational problems. In: *Splitting Methods in Communication, Imaging, Science, and Engineering*, Science, and Engineering, pp. 251–299. Springer, Cham (2016)
43. Glowinski, R., Pan, T.-W., Tai, X.-C.: Some facts about operator-splitting and alternating direction methods. In: *Splitting Methods in Communication, Imaging, Science, and Engineering*. Sci. Comput., pp. 19–94. Springer, Cham (2016)
44. Hubel, D.H., Wiesel, T.N.: Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**(3), 574–591 (1959)
45. Hutchinson, J.E.: Second fundamental form for varifolds and the existence of surfaces minimising curvature. *Indiana Univ. Math. J.* **35**(1), 45–71 (1986)
46. Kanizsa, G.: *Grammatica del vedere: saggi su percezione e gestalt*. Biblioteca: Mulino. Il Mulino (1980), 2nd edition 1997
47. Kanizsa, G.: *Organization in Vision : Essays on Gestalt Perception*. Praeger, New York (1979). Foreword by Paolo Legrenzi and Paolo Bozzi
48. Koenderink, J.J., van Doorn, A.J.: Representation of local geometry in the visual system. *Biol. Cybernet.* **55**(6), 367–375 (1987)
49. Krueger, M., Delmas, P., Gimel'farb, G.: Efficient Image Segmentation Using Weighted Pseudo-Elastica, pp. 59–67. Springer, Berlin (2011)
50. Mantegazza, C.: Curvature varifolds with boundary. *J. Differ. Geom.* **43**(4), 807–843 (1996)
51. Masnou, S., Nardi, G.: A coarea-type formula for the relaxation of a generalized elastica functional. *J. Convex Anal.* **20**(3), 617–653 (2013)

52. Masnou, S., Morel, J.-M.: Level lines based disocclusion. In: International Conference on Image Processing, 1998. ICIP 98. Proceedings, pp. 259–263. IEEE (1998)
53. Masnou, S., Morel, J.-M.: On a variational theory of image amodal completion. *Rend. Sem. Mat. Univ. Padova* **116**, 211–252 (2006)
54. Mercier, G.: Continuity results for TV-minimizers. *Indiana Univ. Math. J.* **67**(4), 1499–1545 (2018)
55. Mirebeau, J.-M.: Anisotropic fast-marching on Cartesian grids using lattice basis reduction. *SIAM J. Numer. Anal.* **52**(4), 1573–1599 (2014)
56. Mirebeau, J.-M.: Fast Marching Methods for Curvature Penalized Shortest Paths, preprint (2017)
57. Moiseev, I., Sachkov, Y.L.: Maxwell strata in sub-Riemannian problem on the group of motions of a plane. *ESAIM Control Optim. Calc. Var.* **16**(2), 380–399 (2010)
58. Nitzberg, M., Mumford, D., Shiota, T.: Filtering, segmentation and depth. *Lecture Notes in Computer Science*, vol. 662. Springer, Berlin (1993)
59. Petitot, J., Tondut, Y.: Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux. *Math. Inform. Sci. Hum.* **145**, 5–101 (1999)
60. Pock, T., Chambolle, A.: Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In: Proceedings of the 2011 International Conference on Computer Vision, ICCV’11, pp. 1762–1769, Washington, DC, USA, IEEE Computer Society (2011)
61. Prandi, D., Boscain, U., Gauthier, J.-P.: Image processing in the semidiscrete group of rototranslations. In: Geometric science of information. Volume 9389 of Lecture Notes in Computer Science, pp. 627–634. Springer, Cham (2015)
62. Raviart, P.-A., Thomas, J.M.: A mixed finite element method for 2nd order elliptic problems. In: Mathematical Aspects of Finite Element Methods (Proceedings of the International Conference on Consiglio Nazionale delle Ricerche (C.N.R.), Rome, 1975). *Lecture Notes in Mathematics*, Vol. 606, pp. 292–315. Springer, Berlin (1977)
63. Rockafellar, R.T.: Integrals which are convex functionals. II. *Pac. J. Math.* **39**, 439–469 (1971)
64. Rockafellar R.T.: Convex analysis. Princeton Landmarks in Mathematics. Princeton University Press, Princeton (1997). Reprint of the 1970 original, Princeton paperbacks
65. Sarti, A., Citti, G.: Subjective surfaces and Riemannian mean curvature flow of graphs. *Acta Math. Univ. Comen. (N.S.)* **70**(1), 85–103 (2000)
66. Sarti, A., Citti, G., Petitot, J.: The symplectic structure of the primary visual cortex. *Biol. Cybernet.* **98**(1), 33–48 (2008)
67. Schmaltz, C., Peter, P., Mainberger, M., Ebel, F., Weickert, J., Bruhn, A.: Understanding, optimising, and extending data compression with anisotropic diffusion. *Int. J. Comput. Vis.* **108**(3), 222–240 (2014)
68. Schoenemann, T., Cremers, D.: Introducing curvature into globally optimal image segmentation: minimum ratio cycles on product graphs. In: IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007, pp. 1–6. IEEE (2007)
69. Schoenemann, T., Kahl, F., Masnou, S., Cremers, D.: A linear framework for region-based image segmentation and inpainting involving curvature penalization. *Int. J. Comput. Vis.* **99**(1), 53–68 (2012)
70. Schoenemann, T., Masnou, S., Cremers, D.: The elastic ratio: introducing curvature into ratio-based image segmentation. *IEEE Trans. Image Process.* **20**(9), 2565–2581 (2011)
71. Schoenemann, T., Masnou, S., Cremers, D.: On a linear programming approach to the discrete Willmore boundary value problem and generalizations. In: Curves and Surfaces. Volume 6920 of Lecture Notes in Computer Science, pp. 629–646. Springer, Heidelberg (2012)
72. Sharma, U., Duits, R.: Left-invariant evolutions of wavelet transforms on the similitude group. *Appl. Comput. Harmon. Anal.* **39**(1), 110–137 (2015)
73. Smirnov, S.K.: Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows. *Algebra i Analiz* **5**(4), 206–238 (1993)
74. Tai, X.-C., Hahn, J., Chung, G.J.: A fast algorithm for Euler’s elastica model using augmented Lagrangian method. *SIAM J. Imaging Sci.* **4**(1), 313–344 (2011)
75. Weickert, J.: Theoretical Foundations of Anisotropic Diffusion in Image Processing, pp. 221–236. Springer, Vienna (1996)
76. Weickert, J.: Mathematische Bildverarbeitung mit Ideen aus der Natur. *Mitt. Dtsch. Math.-Ver.* **20**(2), 82–90 (2012)
77. Willmore, T.J.: A survey on Willmore immersions. In: Geometry and Topology of Submanifolds, IV (Leuven, 1991), pp. 11–16. World Science Publishing, River Edge (1992)
78. Yashtini, M., Kang, S.H.: Alternating direction method of multiplier for Euler’s elastica-based denoising. In: Aujol, J.-F., Nikolova, M., Papadakis, N. (eds.) Scale Space and Variational Methods in

- Computer Vision: 5th International Conference, SSVM 2015, Lège-Cap Ferret, France, May 31–June 4, 2015, Proceedings, pp. 690–701. Springer International Publishing, Cham (2015)
79. Yashtini, M., Kang, S.H.: A fast relaxed normal two split method and an effective weighted TV approach for Euler's elastica image inpainting. *SIAM J. Imaging Sci.* **9**(4), 1552–1581 (2016)
80. Zhu, W., Tai, X.-C., Chan, T.: Augmented Lagrangian method for a mean curvature based image denoising model. *Inverse Probl. Imaging* **7**(4), 1409–1432 (2013)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.