Institut für Informatik III

der Rheinischen Friedrich-Wilhelms-Universität Bonn

A convex approach for computing minimal partitions

Antonin Chambolle, Daniel Cremers,
Thomas Pock

IAI-TR-2008-5 November 2008

universitätbonn

Institut für Informatik III

der Rheinischen Friedrich-Wilhelms-Universität Bonn

A convex approach for computing minimal partitions

Antonin Chambolle, Daniel Cremers,
Thomas Pock

IAI-TR-2008-5 November 2008

Forschungsbericht
Technical Report

ISSN 0944-8535

Forschungsberichte sind erhältlich durch:

Sekretariat für Forschungsberichte
Institut für Informatik III
Universität Bonn
Römerstraße 164
D-53117 Bonn
Fax: (0228) 734-382
E-mail: techrep@cs.uni-bonn.de

Technical reports can be requested from:

Technical Reports Secretary
Institute of Computer Science III
University of Bonn
Römerstraße 164
D-53117 Bonn
Germany
Fax: +49 228 734-382

Anonymous ftp: ftp.cs.uni-bonn.de /pub/paper/infIII

A convex approach for computing minimal partitions

Antonin Chambolle* Daniel Cremers† Thomas Pock‡

November 26, 2008

Abstract

We describe a convex relaxation for a family of problems of minimal perimeter partitions. The minimization of the relaxed problem can be tackled numerically, we describe an algorithm and show some results. In most cases, our relaxed problem finds a correct (approximate) solution: we give some arguments to explain why it should be so, and also discuss some situation where it fails.

1 Introduction

1.1 Contribution

In this paper, we describe a possible approach for numerically computing minimal partitions, and a few related problems. This problems arises in many fields, in particular, motivated by image analysis applications (segmentation), it has been the subject of extensive study in the beginning of the 90s [35, 5, 31, 41], in connection to the celebrated Mumford-Shah [36] segmentation problem.

The problem of (numerically) finding a partition of a set $\Omega \subset \mathbb{R}^d$, which minimizes the $(d-1)$-dimensional measure of the total interface, plus either boundary conditions, or some external field particular to each set, is a hard task. Its discrete version, known as the "Potts' model" (an extension of Ising's model), is described by an energy whose minimization is NP-hard. What we propose here is to derive some convexification of the problem, which is close enough to the convex envelope (in fact, it is the closest in a quite general class), but, also, seems to be numerically tractable (at least when the number of labels

*CMAP, CNRS UMR 7641, Ecole Polytechnique, 91128 Palaiseau Cedex, France. e-mail: antonin.chambolle@polytechnique.fr

†Department of Computer Science, University of Bonn, Römerstraße 164, 53117 Bonn, Germany. e-mail: dcremers@cs.uni-bonn.de

‡Technische Unversität Graz, Inffeldgasse 16, 8010 Graz, Austria. e-mail: pock@icg.tugraz.at

is not too high). Then, minimizing numerically this envelope, we experience that in general, at least when $d = 2$, the minimizer is actually a minimal partition. See Figure 1 for an example: here, our method is applied to solve a piecewise constant Mumford-Shah model with 10 regions.

Figure 1: Minimal segmentation into 10 regions

Let us point out that this was a surprise for us and remains in part a mystery. We also show a numerical counterexample (a situation where it does not work), and try to explain as best as we can what we expect to be the general situation. The goal of this paper is to describe our setting, and explain how we perform the minimization. Then, we give some arguments to explain why it seems to work in many cases, and why it shouldn't in some other.

In some sense, our approach is related to similar relaxations in discrete optimization (and in particular optimization of Markov Random Fields or MRFs), such as LP-relaxation or roof duality relaxation. The point of view closest to ours in the discrete literature seems to be roof duality [27], indeed, we also look for a convex relaxation of our problem which is *local* in some sense, and is obtained as the supremum of affine functions which satisfy local constraints. Let us mention however that we also use a non-obvious representation of our problem in which our relaxation is, in a precise sense, optimal. Then, we base our analysis on the concept of "calibrations", which may be seen as the continuous counterpart of roof duality, and was developed independently for the study of minimal surfaces in the 70's and 80's (see Appendix B for more recent applications in the calculus of variations).

Our extension to the continuous setting and the representation follow the spirit of [37] where two of the authors of this paper were extending recent approaches in MRFs optimization to the continuous setting. This is crucial, in particular, if one wants to reproduce precisely specific surface energies (and in particular isotropic interfacial energies). On the other hand, the price to pay is that algorithms for minimizing discretization of quite arbitrary convex functionals are less developed than discrete optimization of linear problems. In

particular, our optimization scheme (chosen for its relative simplicity, which is important, since we aim to work on high dimensional problems) is approximate, and not even guaranteed to converge, while purely discrete relaxations can usually be solved exactly by quite efficient algorithms [13, 40].

Another point of view similar to ours is found in a recent paper of C. Zach *et al.* [44], which is also written in a continuous setting although, there again, the interaction potentials are eventually chosen anisotropic. See also [30] for a variant. In fact, with the anisotropy chosen in [44], both approaches boil down to the same representation, while in general, we claim ours is better (meaning, a higher relaxation, closer to the original problem), as shown in Appendix A, both theoretically and experimentally (Fig. 12 and Proposition A.1). Moreover, it is not clear at all how to extend the point of view of [44] to problems with more general interaction energies, except, of course, in the way we propose.

Eventually, we must point out that what we propose is quite different from approaches based on phase-field approximations or level-sets (see for instance [17]) which do not aim at computing a global minimizer of the problem (and, for the latter mentioned, cannot approximate very general interfacial energies, or even uniform ones).

1.2 A convex formulation for multi-label problems

In a recent paper [37], which is inspired by the seminal works of Ishikawa and Geiger [29, 28], the second and third authors of the present paper propose (with co-authors) an approach for minimizing the energy (defined for $u \in L^1(\Omega)$)

$$\mathcal{E}(u) \;=\; \int_\Omega |Du| \;+\; \int_\Omega W(x,u(x))\,dx \tag{1}$$

where here, the first term is the total variation of u, and W some potential (say, measurable in x and continuous in u, but not necessarily convex).

Let us recall that a function $u \in L^1(\Omega)$ has bounded variation (and henceforth is in $BV(\Omega)$) if and only if its total variation, defined by

$$\sup\Big\{ -\int_\Omega u \operatorname{div}\xi \;:\; \xi \in C^1_c(\Omega;\mathbb{R}^d)\,, |\xi(x)| \le 1 \;\forall x \in \Omega\Big\}, \tag{2}$$

is finite. It is equivalent to say that its distributional derivative Du is a bounded Radon measure, and the finite value of (2) is then the total mass $\int_\Omega |Du|$, called the *total variation* of u. It is shown, then, that the measure Du is decomposed as follows:

$$Du \;=\; \nabla u\,dx \;+\; Cu \;+\; (u_+ - u_-)\nu_u\,\mathcal{H}^{d-1}\llcorner J_u\,. \tag{3}$$

Here, $\nabla u\,dx$ is the absolutely continuous part of Du with respect the Lebesgue measure dx, Cu is the "Cantor part," which has the characteristic to be singular

with respect to the Lebesgue measure, but vanish also on $(d-1)$-dimensional sets of finite measure, and the last term is the "jump part," carried on the jump set J_u of u. This set, which is shown to be a $(d-1)$-dimensional, rectifiable set (in the sense of Federer, see [2]), may be defined as the set of points $x \in \Omega$ such that the blowup functions $u(x+\varepsilon y)$, defined for $|y| \leq 1$ and $\varepsilon > 0$ small enough, converge in L^1, as $\varepsilon \to 0$, to a function which takes exactly two values $u_-(x) < u_+(x)$, the value $u_-(x)$ in $\{y : \nu_u(x) \cdot y < 0\}$ and $u_+(x)$ in $\{y : \nu_u(x) \cdot y > 0\}$. It turns out that ν_u defines $\mathcal{H}^{d-1}$-a.e. in J_u a unit normal vector. See [26, 2, 22] and section 2.1 for more details.

The kind of energy $\mathcal{E}$ in (1) appears in many practical problems: in [37], the application was to image processing, and more precisely the computation of the disparity in stereo pairs (see Fig. 10 and 11). The approach suggested in [37] consists in discretizing the variable u on $k+1$ levels (to simplify, the levels $0, \ldots, k$) and trying to solve

$$\min_{u \in BV(\Omega;\{0,\ldots,k\})} \mathcal{E}(u). \tag{4}$$

using the following representation: we introduce k "labels" $u_i \in BV(\Omega;\{0,1\})$ defined by:

$$u_i \;= \begin{cases} 1 & \text{if } u \geq i \\ 0 & \text{else,} \end{cases}$$

so that in particular $0 \leq u_k \leq u_{k-1} \leq \cdots \leq u_1 \leq 1$ and $u = \sum_{i=1}^k u_i$ a.e. in Ω. Then, the co-area for BV functions

$$\int_\Omega |Du| \;=\; \int_{-\infty}^{+\infty} \mathrm{Per}(\{u > s\};\Omega)\, ds \tag{5}$$

(where $\mathrm{Per}(E,\Omega) = \int_\Omega |D\chi_E|$ is the perimeter in Ω of the finite-perimeter set E, see [26]) becomes, if $u \in \{0, \ldots, k\}$,

$$\int_\Omega |Du| \;=\; \sum_{i=1}^k \int_{i-1}^i \mathrm{Per}(\{u \geq i\};\Omega)\, ds \;=\; \sum_{i=1}^k \int_\Omega |Du_i|\,,$$

so that problem (4) can be rewritten

$$\min_{0 \leq u_k \leq u_{k-1} \leq \cdots \leq u_1 \leq 1} \sum_{i=1}^k \int_\Omega |Du_i| \;+\; \int_\Omega (1-u_1(x))W(x,0)\, dx \\ +\; \sum_{i=1}^{k-1} \int_\Omega (u_i(x) - u_{i+1}(x))W(x,i)\, dx \;+\; \int_\Omega u_k(x)W(x,k)\, dx\,. \tag{6}$$

The idea in [37] is to relax (6) by considering now functions which can take all values in $[0,1]$, that is, $u_i \in BV(\Omega;[0,1])$. It is easy to check that

this convexification will work: indeed, the energy in (6) can be written (up to constant factors which play no role in the minimization)

$$\sum_{i=1}^{k} \left(\int_\Omega |Du_i| \, + \, \int_\Omega (W(x,i) - W(x,i-1)) u_i(x) \, dx \right)$$

which, using (5) and the fact that $\int_\Omega ug \, dx = \int_0^1 (\int_{\{u>s\}} g \, dx) \, ds$ for any function u with values in $[0,1]$ and any integrable g, becomes

$$\int_0^1 \sum_{i=1}^{k} \left(\int_\Omega |D\chi_{\{u_i>s\}}| \, + \, \int_\Omega (W(x,i) - W(x,i-1)) \chi_{\{u_i>s\}}(x) \, dx \right) \, ds.$$

We deduce that if the functions $u_i \in BV(\Omega;[0,1])$, $i = 1, \dots, k$, minimize (6) with $0 \le u_k \le \cdots \le u_1 \le 1$, then for any $s \in (0,1]$, the functions $u_i^s = \chi_{\{u_i>s\}}$, which take values in $\{0,1\}$ only, still minimize (6), and $u^s = \sum_{i=1}^k u_i^s$ is then a global minimizer of (4). (One can show even that, "in general", this minimizer will be unique and not depend on s, but this needs not be always the case.)

Then, practically, the problem has been turned into a convex problem. Various approaches can be proposed to minimize (6), in [28], a fully discrete approach is represented on a graph and the optimization is based on max-flow algorithms, while [37] propose a more faithful discretization of (6) (written in a slightly different way) and an optimization scheme based on a primal-dual gradient flow, the same which we will be using in this paper.

It is remarkable that this will work for very general potentials W, on the other hand, it is less surprising if one considers that this is done at the expense of quantizing the values and introducing as many functions as labels (in other words, adding a dimension to the problem[1]), so that the computational cost can become very high. As a matter of fact, the popularity of this kind of methods has grown with the development of computers and microprocessors, and the most recent programs are implemented on graphic cards processors (GPUs) which can perform highly parallel (simple) tasks at a very high speed, and whose cost is quite low.

1.3 Extension to nonconvex neighborhood potentials and the partition problem

For the applications in view in [37] (as well as many other), energy (1) is not optimal: indeed, one would like to penalize the amplitude of the jump between two values whenever these are close, but, after some threshold, to penalize

[1]Think of the (not so simple) problem of minimizing a possibly "arbitrary" nonconvex function $f(x)$ over $x \in \mathbb{R}$, and of how simple it becomes if replaced with $\min_{x \in \{0,\dots,k\}} f(x)$ for reasonable values of k...

just the fact that the values are different. In other words, one would rather replace in (4), for $u \in \{0, \dots, k\}$, the interaction potential (which, by (3), is $\int_\Omega |Du| = \int_{J_u} |u_+ - u_-| \, d\mathcal{H}^{d-1}$), with something like

$$\int_{J_u} \min(|u_+(x) - u_-(x)|, T) \, d\mathcal{H}^{d-1}(x) \tag{7}$$

for some threshold $T > 0$ ("truncated TV"). This is the kind of issue we address in this paper: we propose an approach to extend this framework to the minimization of problems where also the interaction potential is not convex.

Another important typical example is the following variant of (4):

$$\min_{u \in BV(\Omega; \{0, \dots, k\})} \mathcal{H}^{d-1}(J_u) \; + \; \int_\Omega W(x, u(x)) \, dx \; : \tag{8}$$

that is, we just penalize s the total surface $\mathcal{H}^{d-1}(J_u)$ of the discontinuity. This is very important, for instance, in the situations where there is no natural ordering between the sets $\{u = i\}$ that we want to recover, and can be seen as the natural extension in the continuous setting of the Potts' model in statistical physics. Minimizing of (8), here, is equivalent to finding a minimal partition of Ω into $(k+1)$ sets $E_i = \{u = i\}$, $i = 0, \dots, k$ (with the external field $W(x, i)$ weighing each set E_i).

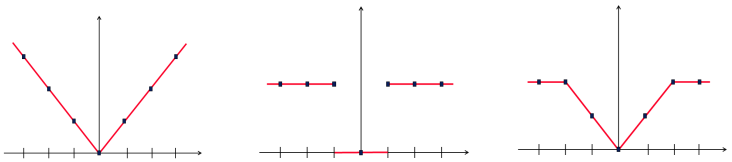

Figure 2: Various popular interaction potentials

The idea to address (8), or the "truncated TV," is to convexify it in the same way as before. We introduce, again, the k ordered functions $u_i \in \{0, 1\}$ such that $u = \sum_{i=1}^k u_i$. We denote by $\mathbf{u}$ the vector-valued function (in $BV(\Omega; \mathbb{R}^k)$) with components $(u_i)_{i=1}^k$. (Observe in particular that $J_u = J_{\mathbf{u}}$.) The definition (2), by duality, of the total variation states that

$$\int_\Omega |Du| \; = \; \sum_{i=1}^k \int_\Omega |Du_i| \; = \; \sup \Big\{ - \sum_{i=1}^k \int_\Omega u_i \operatorname{div} \xi_i \; : \\ (\xi_i)_{i=1}^k \in C_c^1(\Omega; \mathbb{R}^d)^k \, , |\xi_i(x)| \leq 1 \; \forall x \in \Omega, i = 1, \dots, k \Big\}. \tag{9}$$

Using (3), we have, if $u_i \in \{0, 1\}$ a.e.,

$$- \sum_{i=1}^k \int_\Omega u_i \operatorname{div} \xi_i \; = \; \sum_{i=1}^k \int_{J_{u_i}} \xi_i(x) \cdot \nu_{u_i}(x) \, d\mathcal{H}^{d-1}(x) \, . \tag{10}$$

The difference between (8) and (4) is that now, when $x \in J_u$ is a point where u jumps from $u_-(x) = i$ to $u_+(x) = j$ ($0 \leq i < j \leq k$), so that the labels $u_{i+1}, u_{i+2}, \ldots, u_j$ jump simultaneously (while the labels u_l, $l \leq i$ or $l > j$, stay equal respectively to 1 and 0), we want to count only once the jump, and not $j - i$ times. This is the case if, in addition to the constraints $|\xi_i(x)| \leq 1$ for all i and x, one considers the constraint $|\sum_{i<l\leq j} \xi_l| \leq 1$. Indeed, in this case, since all vectors $\nu_{u_l}(x)$ are equal at this particular point x, we will have $\sum_{i<l\leq j} \xi_l \cdot \nu_{u_l}(x) = (\sum_{i<l\leq j} \xi_l) \cdot \nu_{u_{i+1}} \leq 1$, and the supremum 1 will be reached provided one can take $\sum_{i<l\leq j} \xi_l$ arbitrarily close to $\nu_{u_{i+1}}(x)$. Defining the convex

$$K = \Big\{(\xi_1, \ldots, \xi_k) : \xi_i \in \mathbb{R}^d \text{ for all } i, \\ \Big|\sum_{i_1\leq i\leq i_2} \xi_i\Big| \leq 1 \ \forall\ (i_1, i_2) \text{ with } 1 \leq i_1 \leq i_2 \leq k\Big\} \quad (11)$$

we will have that, provided $u_i \in \{0, 1\}$ for all i and a.e. x, and $u = \sum_{i=1}^k u_i$,

$$J(\mathbf{u}) := \sup\Big\{ -\sum_{i=1}^k \int_\Omega u_i(x) \mathrm{div}\, \xi_i(x)\, dx : (\xi_i)_{i=1}^k \in C_c^1(\Omega; K)\Big\} = \mathcal{H}^{d-1}(J_u)\,. \quad (12)$$

Then, our convexification of (8) consists simply in replacing $\mathcal{H}^{d-1}(J_u)$ by the left-hand side expression of (12), and then, as before, taking the functions $u_i \in BV(\Omega; [0, 1])$ instead of binary. We can then minimize the energy (after an appropriate discretization), which is now convex (and defined on the convex set of the constraints $0 \leq u_k \leq u_{k-1} \leq \cdots \leq u_1 \leq 1$). Let us observe that:

- if the minimizer we compute is actually binary, then it is clearly a minimizer of the original problem (8);
- we do not have, however, a simple representation of $J(\mathbf{u})$ as a convex combination, of the form $\int_0^1 J(\mathbf{u}^s)ds$ with binary $\mathbf{u}^s$ (such that $\mathbf{u} = \int_0^1 \mathbf{u}^s ds$), so that there is no guarantee that the minimizer of our relaxed problem is binary-valued (that is $\mathbf{u} \in \{0, 1\}^k$, with moreover $u_{i+1} \leq u_i$ for each $i < k$). In fact, we can provide examples where it does not seem to be true;
- we also have arguments to show that in many useful situations, the solution *should be* binary.

Actually, we were surprised to observe that in almost all our numerical experiments, the minimizers we computed were functions u_i taking mostly the values 0 and 1 (up to quite sharp transitions, which are expected as a byproduct of the discretization).

1.4 Outline

The paper is organized as follows: a preliminary section gives the definition and some basic facts about functions with bounded variation and finite-perimeter partitions. then, we recall how standard relaxation deals with the minimal interface problem (finding a partition in just two sets), and describe our approach for three labels. We propose a numerical approach to solve the problem and show some experiments. Then we detail, in Section 4, the approach in a very general setting. To simplify, we only deal with isotropic interfacial energies, although most of what we do could be easily extended to non-isotropic energies, or nonuniform (spatially). On the other hand, we consider quite general interaction potentials (between the labels).

In Appendix A, we compare our approach to the convexification proposed in the recent paper [44] by C. Zach *et al.* We claim ours is, in general, strictly higher, nevertheless, the latter is simpler and boils down to ours in some situations. Appendix B relates our approach with standard convexification techniques for studying minimization problems in the calculus of variation. A few technical appendices follow.

Acknowledgements

We thank François Alouges, Bastian Goldlücke, Robert V. Kohn, Olivier Pantz and Margherita Solci for fruitful discussions, as well as the Hausdorff Institute for Mathematics in Bonn for financial support. The first author acknowledges the hospitality of the University of Bonn, and in particular the Computer Science Department and the Hausdorff Institute for Mathematics. He is partially supported by the Agence Nationale de la Recherche, project ANR-06-BLAN0082.

2 Preliminaries

2.1 Functions with bounded variation

We recall here some results on functions with bounded variations, see [2, 26, 22, 46] for more details.

Let $d, k \geq 1$, and Ω be an open subset of $\mathbb{R}^d$. Given a function $\mathbf{u} = (u_1, \ldots, u_k) \in L^1(\Omega; \mathbb{R}^k)$, we define its total variation as

$$\sup \Big\{ - \int_\Omega \sum_{i=1}^k u_i \operatorname{div} \xi_i \ : \ \xi \in C_c^1(\Omega; \mathbb{R}^d)^k \, , |\xi(x)|^2 = \sum_{i=1}^k |\xi_i|^2 \leq 1 \ \forall x \in \Omega \Big\}. \quad (13)$$

We say that $\mathbf{u}$ has bounded variation whenever the value of (13) is finite. In this case, one shows (using Riesz' representation theorem) that the distributional derivative $D\mathbf{u}$ is a bounded ($\mathbb{R}^{d\times k}$-valued) Radon measure, whose total mass $\int_\Omega |D\mathbf{u}|$ is precisely the value of (13).

We denote by $BV(\Omega;\mathbb{R}^k)$ the space of (k-dimensional vector valued) functions with bounded variation. Endowed with the norm $\|u\|_{BV} = \|u\|_{L^1(\Omega;\mathbb{R}^k)} + \int_\Omega |D\mathbf{u}|$, it is a Banach space. We use the notation $BV(\Omega)$ whenever $k=1$, and $BV(\Omega;K)$ for functions $\mathbf{u}$ such that $\mathbf{u}(x)\in K$ a.e., where K is a closed, convex subset of $\mathbb{R}^k$.

The distributional derivative $D\mathbf{u}$ has the decomposition [23, 43]:

$$D\mathbf{u} \;=\; \nabla\mathbf{u}\,dx \;+\; C\mathbf{u} \;+\; (\mathbf{u}_+ - \mathbf{u}_-)\otimes\nu_{\mathbf{u}}\,\mathcal{H}^{d-1}\llcorner J_{\mathbf{u}}\,. \tag{14}$$

Here, $\nabla\mathbf{u}\,dx$ is the part of $D\mathbf{u}$ which is absolutely continuous with respect to the Lebesgue measure. The approximate gradient $\nabla\mathbf{u}(x)$ is a L^1 vector field, which corresponds to the weak gradient whenever $\mathbf{u}\in W^{1,1}$. The jump set $J_{\mathbf{u}}$ is the sets of points $x\in\Omega$ where there exist $\mathbf{u}_-(x)\neq\mathbf{u}_+(x)\in\mathbb{R}^k$ and $\nu_{\mathbf{u}}(x)\in\mathbb{S}^{d-1}$ such that

$$(y\mapsto\mathbf{u}(x+\varepsilon y)) \overset{L^1(B(0,1);\mathbb{R}^k)}{\longrightarrow} \big(\, y\mapsto \chi_{\{y\cdot\nu_{\mathbf{u}}(x)>0\}}\mathbf{u}_+(x) \;+\; \chi_{\{y\cdot\nu_{\mathbf{u}}(x)>0\}}\mathbf{u}_-(x)\big)$$

as $\varepsilon\to 0$. Of course, here, we could replace the triplet $(\mathbf{u}_+,\mathbf{u}_-,\nu_{\mathbf{u}})$ with $(\mathbf{u}_-,\mathbf{u}_+,-\nu_{\mathbf{u}})$: when $k=1$, the convention is to choose $u_+>u_-$, if $k>1$, we can for instance choose that the component with the lowest index i such that $(\mathbf{u}_+)_i\neq(\mathbf{u}_-)_i$ satisfies $(\mathbf{u}_+)_i>(\mathbf{u}_-)_i$. Then, the tensor product $(\mathbf{u}_+-\mathbf{u}_-)\otimes\nu_{\mathbf{u}}$ is the $k\times d$ matrix $((\mathbf{u}_+-\mathbf{u}_-)_i(\nu_{\mathbf{u}})_j)_{i=1,\dots,k,j=1,\dots,d}$. The set $J_{\mathbf{u}}$ is shown to be a $(d-1)$-dimensional set which is rectifiable in the sense of Federer (that is, can be covered by countably many C^1 hypersurfaces, up to a set which is negligible for the $(d-1)$-dimensional Hausdorff measure, $\mathcal{H}^{d-1}$). In particular, $\nu_{\mathbf{u}}$ coincides $\mathcal{H}^{d-1}$-a.e. with a normal to $J_{\mathbf{u}}$. Eventually, the cantor part $C\mathbf{u}$ is the part of the measure which, essentially, has dimension between d and $d-1$: it is singular with respect to the Lebesgue measure, but also satisfies $|Cu|(A)=0$ for any set A with $\mathcal{H}^{d-1}(A)<+\infty$.

Given a convex, continuous function $\Psi:\mathbb{R}^{d\times k}\to[0,+\infty)$, one can define the integral $\int_\Omega\Psi(D\mathbf{u})$ as follows: one introduces the recession function

$$\overline{\Psi}(\mathbf{p}) \;=\; \lim_{t\to 0} t\Psi(\mathbf{p}/t)$$

which is a convex, one-homogeneous function (possibly taking the value $+\infty$).

Then, [12, 38]

$$\int_\Omega \Psi(D\mathbf{u}) \;=\; \int_\Omega \Psi(\nabla \mathbf{u}(x))\,dx \\ + \int_\Omega \overline{\Psi}\left(\tfrac{C\mathbf{u}}{|C\mathbf{u}|}\right) d|Cu| \;+\; \int_{J_\mathbf{u}} \overline{\Psi}((\mathbf{u}_+(x)-\mathbf{u}_-(x))\otimes \nu_\mathbf{u}(x))\,d\mathcal{H}^{d-1}(x). \quad (15)$$

Moreover, if Ψ^* is the Legendre-Fenchel conjugate [21, 39] of Ψ, one also have the dual representation

$$\int_\Omega \Psi(D\mathbf{u}) \;=\; \sup\left\{ \int_\Omega \mathbf{u}(x)\cdot \operatorname{div}\xi(x) \;-\; \Psi^*(\xi(x))\,dx \;:\; \xi\in C_c^\infty(\Omega;\mathbb{R}^{d\times k})\right\} \quad (16)$$

(here $\operatorname{div}\xi$ is the vector $(\sum_{j=1}^d (\partial(\xi_i)_j/\partial x_j))_{i=1}^k$).

We have the following approximation result, whose proof follows Meyers-Serrin's classical proof for Sobolev functions and is found for instance in [26]:

Theorem 2.1. *For any $\mathbf{u}\in BV(\Omega;\mathbb{R}^k)$, there exists a sequence $(\mathbf{u}_n)_{n\geq 1}$ of functions in $C^\infty(\Omega;\mathbb{R}^k)$ such that $\mathbf{u}_n\to u$ in $L^1(\Omega;\mathbb{R}^k)$ and*

$$\lim_{n\to\infty}\int_\Omega |\nabla \mathbf{u}_n(x)|\,dx \;=\; |D\mathbf{u}|(\Omega). \quad (17)$$

Combining this result with a celebrated theorem of Reshetnyak [2, Theorem 2.39], we get in addition that, if Ψ is a convex, continuous and one-homogeneous function over $\mathbb{R}^{d\times k}$,

$$\lim_{n\to\infty}\int_\Omega \Psi(\nabla \mathbf{u}_n(x))\,dx \;=\; \int_\Omega \Psi(D\mathbf{u})\,. \quad (18)$$

2.2 Caccioppoli sets

If $E\subset\Omega$ is a measurable set, then its *perimeter* in Ω is defined as the total variation (2) of χ_E. A Caccioppoli set, or set with finite perimeter, is a set such that $\chi_E\in BV(\Omega)$: then we let $\mathrm{Per}(E,\Omega):=\int_\Omega |D\chi_E|$. In this case, the jump set J_{χ_E} is also called the "reduced boundary" of E, denoted by $\partial_* E$, and is equal, up to a set of $\mathcal{H}^{d-1}$ measure zero, to $\Omega\setminus(E^1\cup E^0)$, where E^1, resp., E^0, are the points where E has Lebesgue density 1, resp., 0.

A Caccioppoli partition of Ω is a (finite or countable) sequence of subsets of Ω, $(E_i)_{i\geq 1}$ such that $|E_i\cap E_j|=0$ if $i\neq j$, $|\Omega\setminus\bigcup_i E_i|=0$ (equivalently, $\sum_i \chi_{E_i}=1$ a.e. in Ω), and $\sum_i \mathrm{Per}(E_i,\Omega)<+\infty$. The total perimeter of the partition is half the sum of the perimeters of the sets, since in the latter sum each interface between two sets E_i and E_j is counted twice.

We recall eventually that if $u\in BV(\Omega)$ is a scalar-valued BV function, it enjoys the co-area formula (5) which we have mentioned in the introduction.

In the next section, we introduce a class of variational problem whose unknown is a Caccioppoli partition, with a maximal number of sets $k+1$.

3 Minimal partitions

3.1 Representation in the framework of [37]

Let us first, to simplify, focus on the problem of computing a minimal partition of Ω, bounded open subset of $\mathbb{R}^d$. We will generalize the problem afterwards. We assume we want to find a partition of a set Ω into $(k+1)$ (at most) sets $E_0, \dots, E_k$, which solves the following problem:

$$\min \left\{ \frac{1}{2} \sum_{i=0}^{k} \mathrm{Per}(E_i, \Omega) \, + \, \sum_{i=0}^{k} \int_{E_i} g_i(x)\, dx \;\; : \;\; \sum_{i=0}^{k} \chi_{E_i}(x) = 1 \text{ a.e. in } \Omega \right\} \quad (19)$$

where here $\mathrm{Per}(E_i, \Omega)$ is the perimeter of E_i in Ω (defined below) and g_i are nonnegative weight functions.

The first term in energy (19) is the total length of the boundaries of the partition. The weight $1/2$ is there to take into account that each interface between to sets E_i and E_j is contained twice in the sum, as the boundary of each set. If we introduce the function $u : \Omega \to \{0, \dots, k\}$ defined by $u(x) = i$ if and only if $x \in E_i$, then $u \in BV(\Omega)$, $\sum_{i=0}^{k} \mathrm{Per}(E_i, \Omega) = 2\mathcal{H}^{d-1}(J_u)$, and (19) can be rewritten as the following variant of (8):

$$\min_{u \in BV(\Omega; \{0, \dots, k\})} \mathcal{H}^{d-1}(J_u) \, + \, \sum_{i=0}^{k} \int_{\{u=i\}} g_i(x)\, dx$$

As in the introduction, or in [37], we then introduce the functions $(u_i)_{i=1}^{k}$ defined by $u_i = \chi_{\{u \geq i\}}$, so that:

- $u_i \in \{0, 1\}$ a.e. in Ω, for each $i = 1, \dots, k$;
- $0 \leq u_k \leq u_{k-1} \leq \cdots \leq u_1 \leq 1$;
- $u = \sum_{i=1}^{k} u_i$, and: $\chi_{E_k} = u_k$, $\chi_{E_i} = u_i - u_{i+1}$ for $i = 1, \dots, k-1$, $\chi_{E_0} = 1 - u_1$.

The problem, then, can be written in terms of the vector $\mathbf{u} = (u_i)_{i=1}^{k} \in BV(\Omega; \{0,1\}^k)$ as the minimization of:

$$\begin{aligned} &\mathcal{H}^{d-1}(J_{\mathbf{u}}) \\ &+ \int_\Omega (1 - u_0(x)) g_0(x)\, dx \, + \, \sum_{i=1}^{k-1} (u_i(x) - u_{i+1}(x)) g_i(x)\, dx \, + \, \int_\Omega u_k(x) g_k(x)\, dx \\ &\qquad = \; \mathcal{H}^{d-1}(J_{\mathbf{u}}) \, + \, \sum_{i=1}^{k} u_i(x)(g_i(x) - g_{i-1}(x))\, dx \, + \text{ a constant}\,, \end{aligned} \quad (20)$$

over all $\mathbf{u}$, with the constraint $u_i \geq u_{i+1}$ for $i = 1, \dots, k-1$.

3.2 The case $k = 1$

3.2.1 Convex formulation of the problem

If $k = 1$, that is, we want to partition Ω into two regions, it is clear that we just need one set $E = E_1$, while E_0 will simply be the complement of E. This case is standard and quite straightforward. It is written

$$\min_E \mathrm{Per}(E, \Omega) \; + \; \int_\Omega g_1(x)\chi_E(x) + g_0(x)(1 - \chi_E(x))\, dx. \tag{21}$$

In this case, the representation above in terms of levels u_i is just $u = u_1 = \chi_E$, and since $\mathrm{Per}(E, \Omega) = \mathcal{H}^{d-1}(J_u) = \int_\Omega |Du|$, the problem boils down to

$$\min_{u \in BV(\Omega)} \int_\Omega |Du| \; + \; \int_\Omega (g_1(x) - g_0(x))u(x)\, dx \; + \; \int_\Omega g_0(x)\, dx \tag{22}$$

with the constraint $u \in \{0, 1\}$ a.e. in Ω. (the last term here is constant and can be dropped). This problem admits an obvious convexification: we replace the constraint $u \in \{0, 1\}$ with $u \in [0, 1]$.

Then, using the co-area formula for BV functions (5), we have that the energy in (22) is also

$$\int_0^1 \left(\mathrm{Per}(\{u > s\}, \Omega) \; + \; \int_{\{u>s\}} (g_1(x) - g_0(x))\, dx \right) ds$$

from which we deduce that if u is a minimizer of (22), each set $E_s = \{u > s\}$, $s > 0$, is a minimizer of (21).

This type of convexification is classical in optimization (and known as *LP-relaxation*), as well as in the calculus of variations. It is a standard tool for the study of minimal surfaces which was developed in the end of the 50's [23, 26]. In numerical analysis, it was probably first used in [6] for minimal perimeter problems.

3.2.2 Numerical Analysis

There are many ways to numerically tackle problem (22). We present a simple approach which is quite efficient, in the framework of finite differences.

Approximation To simplify we consider only the 2D case ($d = 2$), and we assume that $\Omega = (0, 1)^2$. Let $N > 1$ and $h = 1/N > 0$ be a discretization step. If $u^h = (u^h_{i,j})_{0 \le i,j \le N-1} \in X^h \sim \mathbb{R}^{N \times N}$ is a discrete function, which is identified with the function in $L^1(\Omega)$

$$u^h(x) \; = \; \sum_{0 \le i,j < N} u^h_{i,j} \chi_{(ih,jh)+[0,h)^2}(x)\,, \tag{23}$$

we define its discrete total variation by

$$J^h(u^h) \;=\; h^2 \sum_{0\le i,j<N} \Big|(\nabla^h u^h)_{i,j}\Big| \tag{24}$$

where $|\cdot|$ is the standard Euclidean norm and $\nabla^h : X^h \to X^h \times X^h$ is defined by

$$(\nabla^h u^h)_{i,j} \;= \frac{1}{h}\begin{cases} (u^h_{i+1,j}-u^h_{i,j},\, u^h_{i,j+1}-u^h_{i,j})^T & \text{if } 0\le i,j<N-1\,,\\ (u^h_{i+1,j}-u^h_{i,j},\, 0)^T & \text{if } 0\le i<N-1, j=N-1\,,\\ (0,\, u^h_{i,j+1}-u^h_{i,j})^T & \text{if } i=N-1, 0\le j<N-1\,,\\ (0,0)^T & \text{if } (i,j)=(N-1,N-1)\end{cases}$$

Then, assuming that J^h is extended to $L^1(\Omega)$ by letting $J^h(u)=+\infty$ whenever u is not of the form (23), we have the following elementary result. For the definition and properties of Γ-convergence, see [15, 18], and Remark 3.2.

Proposition 3.1. *J^h Γ-converges to the total variation $\int_\Omega |Du|$ as $h\to 0$.*

Proof. Let just quickly sketch a proof of this result. By definition of the Γ-convergence we need to show that [15] for any $u\in BV(\Omega)$:

(i) If $u^h\to u$, then $\int_\Omega |Du| \le \liminf_{h\to 0} J^h(u^h)$;

(ii) There exists $u^h\to u$ with $\int_\Omega |Du| \ge \limsup_{h\to 0} J^h(u^h)$.

The proof of (ii) is standard and based on the fact that any $u\in BV(\Omega)$ can be approximated with smooth functions $u^n\in C^\infty(\Omega)$ with $\int_\Omega |\nabla u_n|\,dx \to \int_\Omega |Du|$, see Theorem 2.1. Then, it is easy to show that if $u^n\in C^\infty(\Omega)$, by simply defining $u^{n,h}_{i,j}=u^n(ih,jh)$ we have $\lim_{h\to 0} J^h(u^{n,h}) = \int_\Omega |\nabla u^n|\,dx$ (with, in this case, an error bound depending on D^2u_n). Point (ii) follows from a standard diagonal argument.

The proof of (i) is based on the dual formulation of the total variation. We consider $(u^h)_h$, a sequence of discrete functions at scale h, $h=1/N\to 0$, which converges to u in $L^1(\Omega)$. To show (i), it is enough to show that for any $\xi\in C^\infty_c(\Omega;\mathbb{R}^2)$ with $|\xi|\le 1$ in Ω, we have

$$-\int_\Omega u(x)\mathrm{div}\,\xi(x)\,dx \;\le\; \liminf_{h\to 0} J^h(u^h). \tag{25}$$

Then, taking the supremum over all such ξ's will yield (i). We have

$$\begin{aligned} -\int_\Omega u(x)\mathrm{div}\,\xi(x)\,dx \;&=\; \lim_{h\to 0} -\int_\Omega u^h(x)\mathrm{div}\,\xi(x)\,dx\\ &=\; \lim_{h\to 0}\sum_{i,j}\int_0^h \xi_1((i+1)h, jh+s)(u^h_{i+1,j}-u^h_{i,j})\,ds\\ &\qquad + \int_0^h \xi_2(ih+s,(j+1)h)(u^h_{i,j+1}-u^h_{i,j})\,ds\,. \end{aligned} \tag{26}$$

Now, if $s \in (0,h)$, $|((i+1)h, jh+s) - (ih+s, (j+1)h)| = |(h-s, s-h)| \leq \sqrt{2}h$, so that

$$\sqrt{\xi_1((i+1)h, jh+s)^2 + \xi_2(ih+s, (j+1)h)^2} \\ \leq \ |\xi((i+1)h, jh+s)| \ + \ \sqrt{2}\|\nabla\xi\|_{L^\infty(\Omega)}h \ \leq \ 1 + Ch$$

and it follows that

$$\xi_1((i+1)h, jh+s)(u^h_{i+1,j} - u^h_{i,j}) \ + \ \xi_2(ih+s, (j+1)h)(u^h_{i,j+1} - u^h_{i,j}) \\ \leq \ (1+Ch)h\left|(\nabla^h u^h)_{i,j}\right|$$

We deduce from (26) that (25) holds, hence (i). □

Remark 3.2. By standard properties of the Γ-convergence [15], it follows that if u^h is a minimizer of

$$J^h(u^h) \ + \ \int_\Omega (g_1(x) - g_0(x))u^h(x)\,dx\,, \tag{27}$$

then as $h \to 0$, u^h converges to a solution of (22) (up to subsequences if this solution is non unique). Hence, to solve our problem, we will minimize (27).

Numerical Scheme We propose here a primal-dual Arrow-Hurwicz [4, 21] type algorithm for minimizing (27). The advantage of this algorithm is that it extends easily to the problems we will address in the next sections. It has been recently suggested in this framework (more precisely, total variation minimization for image denoising and reconstruction), first by H. Talbot and B. Appleton [3] (who, in fact, suggested a primal-dual flow in a continuous setting, which was then discretized using classical methods for hyperbolic schemes) and more recently, in a setting closer to ours, by M. Zhu and T. Chan [45].

We fix a scale $h = 1/N > 0$, and first discretize the external field by letting

$$W^h_{i,j} \ = \ \frac{1}{h^2}\int_{(ih,jh)+[0,h)^2} (g_1(x) - g_0(x))\,dx$$

(if not already given in a discrete form). Then, we introduce a primal variable $U = (U_{i,j}) \in X^h$ and a dual $\Xi = (\Xi_{i,j}) \in X^h \times X^h$. The problem we want to optimize may be written as

$$\min_{0 \leq U_{i,j} \leq 1} \ \max_{\Xi_{i,j} \in B(0,1)} \ h^2 \sum_{i,j} \Xi_{i,j} \cdot (\nabla^h U)_{i,j} \ + \ h^2 \sum_{i,j} U_{i,j} W^h_{i,j}. \tag{28}$$

We implement a Arrow-Hurwicz type algorithm, which corresponds to follow simultaneously a gradient descent in U and ascent in Ξ, until (hopefully) convergence to the saddle-point.

First we initialize with $U^0 = 0$, $\Xi^0 = 0$. Then, we fix two "time steps" $\tau, \tau' > 0$ and we update U^n, Ξ^n by letting

$$\begin{array}{rcll} U^{n+1}_{i,j} & = & \Pi_{[0,1]}\left(U^n_{i,j} + \tau\left((\mathrm{div}^{\,h}\Xi^n)_{i,j} - W^h_{i,j}\right)\right) & \text{for all } i,j \\ \Xi^{n+1}_{i,j} & = & \Pi_{B(0,1)}\left(\Xi^n_{i,j} + \tau'(\nabla^h U^{n+1})_{i,j}\right) & \text{for all } i,j \end{array} \tag{29}$$

where here, the discrete divergence $\mathrm{div}^{\,h} = -(\nabla^h)^*$ is the opposite of the adjoint of the discrete gradient. The projections $\Pi_{[0,1]}$ and $\Pi_{B(0,1)}$ are respectively the projections on the segment $[0,1]$ (hence the truncation $x \mapsto \max\{0, \min\{x, 1\}\}$), and on the unit ball in $\mathbb{R}^2$. Experimentally, the scheme is stable and seems to converge if $\tau\tau' \leq 1/2$ (for $h = 1$), although we did not find any convincing proof of this fact. Nevertheless, even if this approach might not be the best for this problem, it turns out that for the extensions we present later on, it seems to be one of the most convenient since it does not require too much space for storing the variables (most variables need be stored just once).

An advantage of this approach is that the primal-dual gap, given by

$$\begin{aligned} \max_{\Xi_{i,j}\in B(0,1)} \; & h^2 \sum_{i,j} \Xi_{i,j}\cdot(\nabla^h U^n)_{i,j} \;+\; h^2\sum_{i,j} U^n_{i,j} W^h_{i,j} \\ & - \min_{0\leq U_{i,j}\leq 1} h^2 \sum_{i,j} (-\mathrm{div}^{\,h}\Xi^n_{i,j}) U_{i,j} \;+\; h^2 \sum_{i,j} U_{i,j} W^h_{i,j} \\ = \; & J^h(U^n) + h^2 \sum_{i,j} W^h_{i,j} U^n_{i,j} \;+\; h^2 \sum_{i,j} ((\mathrm{div}^{\,h}\Xi^n)_{i,j} - W^h_{i,j})^+ \;\geq\; 0 \end{aligned} \tag{30}$$

(where $x^+ := \max\{x, 0\}$), may be computed at each step and goes to zero at convergence, hence it may be taken as a criterion for convergence.

3.3 The case $k = 2$: minimal partitions into three subsets

Assume now we want to find the minimizer of (19) with three phases, that is, for $k = 2$. Hence, we are given three sets E_0, E_1, E_2, forming a partition of Ω, and we represent these as above with now two functions, $u_1 = 1 - \chi_{E_0}$, and $u_2 = 1 - \chi_{E_1\cup E_0} = \chi_{E_2}$. The functions satisfy $0 \leq u_2 \leq u_1 \leq 1$, and we recover the sets by letting $\chi_{E_0} = 1 - u_1$, $\chi_{E_1} = u_1 - u_2$, $\chi_{E_2} = u_2$. The vector $\mathbf{u} = (u_1, u_2)$ is in $BV(\Omega; \{0,1\}^2)$ and the energy to be minimized is

$$\mathcal{E}(\mathbf{u}) \;=\; \mathcal{H}^{d-1}(J_{\mathbf{u}}) \;+\; \int_\Omega (g_1(x) - g_0(x))u_1(x) + (g_2(x) - g_1(x))u_2(x)\,dx\,.$$

Define $\mathcal{E}(\mathbf{u}) = +\infty$ whenever $\mathbf{u} \notin BV(\Omega; \{0,1\}^2)$ with $u_2 \leq u_1$. Now, in view of the numerical minimization of $\mathcal{E}$, we look for a reasonable convexification. The best would be to use the convex lower semicontinous envelope of $\mathcal{E}$

in $L^1(\Omega;\mathbb{R}^2)$, which is

$$\overline{\mathcal{E}}(\mathbf{u}) \;=\; \inf\Big\{\liminf_j \textstyle\sum_{l=1}^{n_j}\theta^{j,l}\mathcal{E}(\mathbf{u}^{j,l}) \;:\; n_j\geq 1, \theta^{j,l}\in[0,1]\,, \\ \textstyle\sum_{l=1}^{n_j}\theta^{j,l}=1\,, \lim_{j\to\infty}\big\|\mathbf{u}-\sum_{l=1}^{n_j}\theta^{j,l}\mathbf{u}^{j,l}\big\|_{L^1(\Omega)}=0\Big\}. \quad (31)$$

Its domain (that is, the set where it is finite) are the functions $\mathbf{u}\in BV(\Omega;\mathbb{R}^2)$ such that $0\leq u_2\leq u_1\leq 1$ a.e. in Ω. Then, one can show that for a binary $\mathbf{u}\in\{0,1\}$ a.e., $\overline{\mathcal{E}}(\mathbf{u})=\mathcal{E}(\mathbf{u})$. This is true because binary vectors $\mathbf{u}$ are extremal points in the domain of $\overline{\mathcal{E}}$, and hence cannot be approximated by convex combinations of elements which do not all converge to $\mathbf{u}$ (and $\mathcal{E}$ is l.s.c.) See [16, Prop. 1] for a proof in a different, but similar context.

Hence, any minimizer of $\mathcal{E}$ is a minimizer of $\overline{\mathcal{E}}$ while, in some sense, minimizers of $\overline{\mathcal{E}}$ are convex combinations of minimizers of $\mathcal{E}$ (and in case $\mathcal{E}$ has a unique minimizer, which cannot be asserted in general but is likely to be often true, it is also the unique minimizer of $\overline{\mathcal{E}}$).

Unfortunately, we do not know a general approach to compute (31). The standard way to compute this envelope is through the Legendre-Fenchel conjugate, see [21]. For $\mathbf{v}\in L^\infty(\Omega;\mathbb{R}^2)$, the Legendre-Fenchel transform of $\mathcal{E}$ is:

$$\mathcal{E}^*(\mathbf{v}) \;=\; \sup_{\mathbf{u}\in BV(\Omega;\{0,1\}^2)}\int_\Omega \mathbf{v}(x)\cdot\mathbf{u}(x)\,dx \;-\; \mathcal{E}(\mathbf{u})$$

and, then, again, the transform of $\mathcal{E}^*$ is

$$\mathcal{E}^{**}(\mathbf{u}) \;=\; \sup_{\mathbf{v}\in L^\infty(\Omega;\mathbb{R}^2)}\int_\Omega \mathbf{v}(x)\cdot\mathbf{u}(x)\,dx \;-\; \mathcal{E}^*(\mathbf{v})$$

and it happens to be equal to $\overline{\mathcal{E}}(\mathbf{u})$.

What we propose is to try to guess a reasonable envelope, as close as possible to $\mathcal{E}^{**}=\overline{\mathcal{E}}$ but with an explicit expression. First of all, since the term involving the external fields (g_0,g_1,g_2) is linear in $\mathbf{u}$, it is clearly enough to focus on the convex envelope of the interaction term $\mathcal{H}^{d-1}(J_\mathbf{u})$. Let us therefore define

$$J^0(\mathbf{u}) \;=\; \begin{cases}\mathcal{H}^{d-1}(J_\mathbf{u}) & \text{if } \mathbf{u}\in BV(\Omega;\{0,1\}^2)\,, u_2\leq u_1 \text{ a.e.}\,,\\ +\infty & \text{else.}\end{cases} \quad (32)$$

We have, as before, that any minimizer of $\mathcal{E}$ is also a minimizer of the problem

$$\min_{\mathbf{u}=(u_1,u_2)} (J^0)^{**}(\mathbf{u}) \;+\; \int_\Omega (g_1(x)-g_0(x))u_1(x)+(g_2(x)-g_1(x))u_2(x)\,dx\,.$$

We have the following elementary estimate:

Lemma 3.3. *For $\mathbf{p} = (p_1, p_2) \in \mathbb{R}^{d\times 2}$, let*

$$\Psi^1(\mathbf{p}) \;:=\; |p_1| + |p_2| \;\text{ and }\; \Psi^\infty(\mathbf{p}) \;:=\; \max\{|p_1|, |p_2|\}\,.$$

Then for any $\mathbf{u} \in \mathcal{C}_0 := \{u \in BV(\Omega; [0,1]^2) \,:\, u_2 \le u_1 \text{ a.e.}\}$, we have:

$$\int_\Omega \Psi^\infty(D\mathbf{u}) \;\le\; (J^0)^{**}(\mathbf{u}) \;\le\; \int_\Omega \Psi^1(D\mathbf{u}) \tag{33}$$

whereas if $\mathbf{u} \in L^1(\Omega;\mathbb{R}^2) \setminus \mathcal{C}_0$, $(J^0)^{**}(\mathbf{u}) = +\infty$.

Proof. The fact that the domain of $(J^0)^{**}$ is contained in $\mathcal{C}_0$ is obvious, as one can show that $\mathcal{C}_0$ is the closed convex envelope of $\{\mathbf{u} \in BV(\Omega;\{0,1\}^2) \,:\, u_2 \le u_1 \text{ a.e.}\}$. We therefore just need to show the estimate (33), for $\mathbf{u} \in \mathcal{C}_0$.

We recall that if $\mathbf{u} \in BV(\Omega;\mathbb{R}^2)$, then $D\mathbf{u}$ has the decomposition (14) and for any $\Psi : \mathbb{R}^d \to [0,+\infty)$ which is convex, one-homogeneous, $\int_\Omega \Psi(D\mathbf{u})$ is given by (15), with $\overline{\Psi} = \Psi$.

The first inequality in (33), $\int_\Omega \Psi^\infty(D\mathbf{u}) \le (J^0)^{**}(\mathbf{u})$, simply comes from the fact that $\int_\Omega \Psi^\infty(D\mathbf{u}) \le J^0(\mathbf{u})$ for any $\mathbf{u}$ (and the left-hand side is already convex, l.s.c.).

A sketch of proof of the second inequality is as follows: using the co-area formula (5), we have for any $\mathbf{u} \in BV(\Omega;[0,1]^2)$ with $u_2 \le u_1$ a.e.:

$$\begin{aligned}\int_\Omega \Psi^1(D\mathbf{u}) \;&=\; \int_0^1 \left(\int_\Omega |D\chi_{\{u_1>s\}}| + |D\chi_{\{u_2>s\}}| \right) ds \\ &\ge\; \int_0^1 J^0(\chi_{\{u_1>s\}}, \chi_{\{u_2>s\}})\, ds \;\ge\; (J^0)^{**}(\mathbf{u})\,,\end{aligned}$$

where we have used $\mathbf{u} = (u_1,u_2) = \int_0^1 (\chi_{\{u_1>s\}}, \chi_{\{u_2>s\}})\, ds$. (A rigorous proof would require an extra approximation procedure.) □

Now, we do not believe that $(J^0)^{**}$ has a simple representation. Instead, we propose to try to find a (possibly lower) convex envelope which is local, in the sense that it can be written $J(\mathbf{u}) = \int_\Omega \Psi(x, D\mathbf{u})$ for a function Ψ which is to be determined. Assume $\Psi : \Omega \times \mathbb{R}^{d\times k} \to [0,+\infty)$ is a nonnegative function which is continuous, and convex in its second variable. Then, we have the generalization of (15)

$$\begin{aligned}&\int_\Omega \Psi(x, D\mathbf{u}) \;=\; \int_\Omega \Psi(x, \nabla\mathbf{u}(x))\, dx \\ &+ \int_\Omega \overline{\Psi}\left(x, \tfrac{C\mathbf{u}}{|C\mathbf{u}|}\right) d|Cu| \;+\; \int_{J_\mathbf{u}} \overline{\Psi}(x, (\mathbf{u}_+(x) - \mathbf{u}_-(x)) \otimes \nu_\mathbf{u}(x))\, d\mathcal{H}^{d-1}(x),\end{aligned} \tag{34}$$

where $\overline{\Psi}(x,\mathbf{p}) = \lim_{t\to 0} t\Psi(x, \mathbf{p}/t)$ is the recession function of $\Psi(x,\cdot)$. In fact, a convex local functional may have a more general form, as detailed in [11],

however, we believe that considering such a form would make necessary much more complex notation to get, in the end, essentially the same result.

Let us first restrict our research to convex functions $\Psi : \mathbb{R}^{d\times k} \to [0, +\infty)$ which do not depend on x: we are thus looking for the maximal functional below J^0 of the form (15). We denote $J(\mathbf{u}) = \int_\Omega \Psi(D\mathbf{u})$. We will have $J(\mathbf{u}) \leq J^0(\mathbf{u})$ for any $\mathbf{u}$, if and only if $\Psi(0,0) \leq 0$, while the three values $\overline{\Psi}(\nu, 0)$, $\overline{\Psi}(\nu,\nu)$, $\overline{\Psi}(0,\nu)$ are less than, or equal to 1, for any $\nu \in \mathbb{R}^d$ with norm one. Of course, the "best choice" is to take these values equal to one.

On the other hand, if J is the largest functional of this form below J^0, we must have by (33) that $\Psi \geq \Psi^\infty \geq 0$. It follows that $\Psi \leq \overline{\Psi}$, but then, if $\Psi < \overline{\Psi}$ somewhere, substituting Ψ with $\overline{\Psi}$ in the definition of J results in a functional which remains below J^0 and is higher, in contradiction with the maximality of J. Hence, we must have $\Psi = \overline{\Psi}$, that is, Ψ must be one-homogeneous.

We see that the largest possible Ψ, therefore, is the convex envelope of

$$\Psi^0 : \mathbf{p} \mapsto \begin{cases} |p| & \text{if } \mathbf{p} = (p,p) \text{ or } \mathbf{p} = (p,0) \text{ or } \mathbf{p} = (0,p)\,, p \in \mathbb{R}^d \\ +\infty & \text{else.} \end{cases}$$

Now, we can compute, for $\mathbf{q} \in \mathbb{R}^{d\times 2}$,

$$\begin{aligned} (\Psi^0)^*(\mathbf{q}) &= \sup_{\mathbf{p}} \mathbf{p}\cdot\mathbf{q} - \Psi^0(\mathbf{p}) \\ &= \sup_{p\in\mathbb{R}^d} (\max\{p\cdot(q_1+q_2), p\cdot q_1, p\cdot q_2\} - |p|) \\ &= \begin{cases} 0 & \text{if } |q_1| \leq 1\,, |q_2| \leq 1\,, |q_1+q_2| \leq 1\,, \\ +\infty & \text{else.} \end{cases} \end{aligned}$$

This function is usually known as the "characteristic function" of the convex set

$$K = \left\{ \mathbf{q} = (q_1, q_2) \in \mathbb{R}^{d\times 2} \,:\, |q_1| \leq 1, |q_2| \leq 1, |q_1+q_2| \leq 1 \right\}.$$

We can now compute

$$\Psi(\mathbf{p}) = (\Psi^0)^{**}(\mathbf{p}) = \sup_{\mathbf{q}\in K} \mathbf{q}\cdot\mathbf{p}\,,$$

and we have:

Lemma 3.4. *The function Ψ is given as the inf-convolution:*

$$\Psi(\mathbf{p}) = \min_{q\in\mathbb{R}^d} |q| + |p_1 - q| + |p_2 - q|. \tag{35}$$

Proof. It follows from the fact that given any two convex, l.s.c. functions, $(g+f)^*(x) = \min_{y+z=x} g^*(y) + f^*(z)$. In this case, Ψ is the Legendre-Fenchel conjugate of sum of the three characteristic functions of the convex sets $\{(q_1,q_2) : |q_1| \leq 1\}$, $\{(q_1,q_2) : |q_2| \leq 1\}$, and $\{(q_1,q_2) : |q_1+q_2| \leq 1\}$. □

We can now show the following:

Proposition 3.5. *The largest local convex, l.s.c. functional below J^0, of the form* (34)*, is given by, for* $\mathbf{u} \in L^1(\Omega;\mathbb{R}^2)$*,*

$$J(\mathbf{u}) \;=\; \begin{cases} \int_\Omega \Psi(D\mathbf{u}) & \text{if } \mathbf{u} \in BV(\Omega;\mathbb{R}^2), \\ +\infty & \text{else} \end{cases}$$

where Ψ is given by (35)*. In particular, we have that $J = J^0$ on the domain of J^0 (that is, whenever it is finite).*

Proof. The proposition has been proved for functionals of the form (15), with Ψ not depending on x. We need to show that we cannot to better by considering an integrand $\Psi(x,\mathbf{p})$ (continuous in x) as in (34). Let us denote by J the functional $\int_\Omega \Psi(Du)$ found so far, with Ψ given by (35), and let $J'(\mathbf{u}) = \int_\Omega \Psi'(x,Du)$ a possibly larger convex, local functional below J^0.

Observe that, as before, the best choice is to take $\Psi'(x,\cdot) = \overline{\Psi}'(x,\cdot)$ for all $x \in \Omega$ (hence Ψ' one-homogeneous). Indeed, we always have for any x and $\mathbf{p}$: $\Psi'(x,\mathbf{p}) \le t\Psi'(x,\mathbf{p}/t) + (1-t)\Psi'(x,0)$ if $t<1$, and sending t to zero we deduce $\Psi'(x,\mathbf{p}) \le \overline{\Psi}'(x,\mathbf{p}) + \Psi'(x,0)$. Now, it follows, for any $\mathbf{u}$

$$J'(\mathbf{u}) \;\le\; \int_\Omega \overline{\Psi}'(x,D\mathbf{u}) \,+\, \int_\Omega \Psi'(x,0) \;\le\; \int_\Omega \overline{\Psi}'(x,D\mathbf{u})$$

since $\int_\Omega \Psi'(x,0) = J'(0) \le J^0(0) = 0$. Hence the functional $\int_\Omega \overline{\Psi}'(x,D\mathbf{u})$ (which is obviously below J^0, since it coincides with J' on the domain of J^0), is larger than J', hence must be equal. It follows that Ψ' is one-homogeneous.

Then, since $J(\mathbf{u}) = J^0(\mathbf{u})$ whenever $J^0(\mathbf{u})$ is finite, we also must have $J'(\mathbf{u}) = J^0(\mathbf{u})$ in this case. Let $x_0 \in \Omega$ and $\rho > 0$ such that $B(x_0,\rho) \subset \Omega$, choose $\overline{\nu} \in \mathbb{S}^{d-1}$ a direction and let

$$\begin{aligned} \Gamma^+ &= \partial B(x_0,\rho) \cap \{x : (x-x_0)\cdot\overline{\nu} > 0\}, \\ \Gamma^- &= \partial B(x_0,\rho) \cap \{x : (x-x_0)\cdot\overline{\nu} < 0\}, \\ \Delta &= B(x_0,\rho) \cap \{x : (x-x_0)\cdot\overline{\nu} = 0\} \end{aligned}$$

For $\mathbf{a},\mathbf{b} \in \{(0,0),(1,0),(1,1)\}$ with $\mathbf{a} \neq \mathbf{b}$, we let

$$\begin{aligned} \lambda^{\pm}_{\mathbf{a},\mathbf{b}} &= \int_{\Gamma^\pm} \Psi'(x,(\mathbf{a}-\mathbf{b})\otimes\nu(x))\, d\mathcal{H}^{d-1}(x)\,, \\ \lambda_{\mathbf{a},\mathbf{b}} &= \int_{\Delta} \Psi'(x,(\mathbf{a}-\mathbf{b})\otimes\overline{\nu})\, d\mathcal{H}^{d-1}(x) \end{aligned}$$

where $\nu(x) = (x-x_0)/|x-x_0|$ denotes the outer normal to $B(x_0,\rho)$, and $\overline{\Psi}'(x,\cdot)$ is the recession function of $\Psi'(x,\cdot)$.

Then, considering all possible $\mathbf{u}$ piecewise constant with $J_{\mathbf{u}} \subseteq \Gamma^+ \cup \Gamma^- \cup \Delta$, and using the identity $J'(\mathbf{u}) = J^0(\mathbf{u})$ for such $\mathbf{u}$, we have the following relationships:

$$\begin{array}{rcl}
\lambda^+_{\mathbf{a},\mathbf{b}} + \lambda^-_{\mathbf{a},\mathbf{c}} + \lambda_{\mathbf{b},\mathbf{c}} &=& \mathcal{H}^{d-1}(\Gamma^+) + \mathcal{H}^{d-1}(\Gamma^+) + \mathcal{H}^{d-1}(\Delta) \\
\lambda^+_{\mathbf{a},\mathbf{b}} + \lambda^-_{\mathbf{a},\mathbf{b}} &=& \mathcal{H}^{d-1}(\Gamma^+) + \mathcal{H}^{d-1}(\Gamma^-) \\
\lambda^+_{\mathbf{a},\mathbf{b}} + \lambda_{\mathbf{b},\mathbf{a}} &=& \mathcal{H}^{d-1}(\Gamma^+) + \mathcal{H}^{d-1}(\Delta) \\
\lambda^-_{\mathbf{a},\mathbf{b}} + \lambda_{\mathbf{a},\mathbf{b}} &=& \mathcal{H}^{d-1}(\Gamma^-) + \mathcal{H}^{d-1}(\Delta)
\end{array}$$

for all possible $\{\mathbf{a},\mathbf{b},\mathbf{c}\} = \{(0,0),(1,0),(1,1)\}$. Combining the three last equations, we deduce:

$$\lambda_{\mathbf{a},\mathbf{b}} + \lambda_{\mathbf{b},\mathbf{a}} = 2\mathcal{H}^{d-1}(\Delta)$$

and letting $\rho \to 0$ this implies

$$\Psi'(x_0,(\mathbf{a}-\mathbf{b})\otimes\overline{\nu})) + \Psi'(x_0,(\mathbf{b}-\mathbf{a})\otimes\overline{\nu})) = 2 = 2\Psi^0((\mathbf{a}-\mathbf{b})\otimes\overline{\nu})$$

for any x_0, any $\overline{\nu} \in \mathbb{S}^{d-1}$, and any $\mathbf{a},\mathbf{b} \in \{(0,0),(1,0),(1,1)\}$. It follows (using also the homogeneity) $\Psi'(x,\mathbf{p}) + \Psi'(x,-\mathbf{p}) = 2\Psi^0(\mathbf{p})$ for any x, and $\mathbf{p} \in \{(p,0),(0,p),(p,p) : p \in \mathbb{R}^d\}$. Since $\Psi'(x,\cdot)$ is convex and $\Psi = (\Psi^0)^{**}$, we deduce

$$\frac{1}{2}\left(\Psi'(x,\mathbf{p}) + \Psi'(x,-\mathbf{p})\right) \leq \Psi(\mathbf{p}) \tag{36}$$

for any $\mathbf{p}$.

Let now $\mathbf{u} = (u_1,u_2) \in \mathcal{C}_0$. By definition of J and J' we must have $J(\mathbf{u}) \leq J'(\mathbf{u}) \leq (J^0)^{**}(\mathbf{u}) < +\infty$. We define $\mathbf{u}' = (u_1',u_2') = ((1+u_1)/2, u_2/2)$ and $\mathbf{u}'' = (u_1'',u_2'') = (1-u_1/2,(1-u_2)/2)$. We also have $J(\mathbf{u}'') \leq J'(\mathbf{u}'') < +\infty$ and $J(\mathbf{u}'') \leq J'(\mathbf{u}'') < +\infty$. But as $D\mathbf{u}' = -D\mathbf{u}'' = D\mathbf{u}/2$, it follows, using (36),

$$\begin{aligned}
J'(\mathbf{u}') + J'(\mathbf{u}'') &= \frac{1}{2}\left(\int_\Omega \Psi'(x,D\mathbf{u}) + \int_\Omega \Psi'(x,-D\mathbf{u})\right) \\
&\leq \int_\Omega \Psi(D\mathbf{u}) = J(\mathbf{u}) = 2J(\mathbf{u}') = 2J(\mathbf{u}'')
\end{aligned}$$

and we deduce $J'(\mathbf{u}') = J(\mathbf{u}') = J'(\mathbf{u}'') = J(\mathbf{u}'')$, in other words, $J'(\mathbf{u}) = \int_\Omega \Psi'(x,D\mathbf{u}) = \int_\Omega \Psi'(x,-D\mathbf{u})$ for any $\mathbf{u}$ in the domain $\mathcal{C}_0$ of $(J^0)^{**}$. Hence, for any such $\mathbf{u}$, using (36) again:

$$J'(\mathbf{u}) = \frac{1}{2}\left(\int_\Omega \Psi'(x,D\mathbf{u}) + \int_\Omega \Psi'(x,-D\mathbf{u})\right) \leq \int_\Omega \Psi(D\mathbf{u}) = J(\mathbf{u})$$

which shows the maximality of J. □

It follows from the definition of Ψ that J has the dual representation given in equation (16):

$$J(\mathbf{u}) = \sup\left\{ -\int_\Omega u_1 \mathrm{div}\,\xi_1 + u_2 \mathrm{div}\,\xi_2\, dx \ : \ (\xi_1,\xi_2) \in C_c^\infty(\Omega;K)\right\}.$$

Now, our method for finding a minimizer of $\mathcal{E}$ is simply the following: we solve the convex problem

$$\min_{\mathbf{u}\in\mathcal{C}_0} J(\mathbf{u}) \;+\; \int_\Omega (g_1(x)-g_0(x))u_1(x) + (g_2(x)-g_1(x))u_2(x)\,dx\,. \tag{37}$$

Then, it is not clear that the solution $\mathbf{u}$ will be a binary vector ($u_i \in \{0,1\}$ a.e. for $i=1,2$). However, if it is, then $\mathbf{u}$ will also be a minimizer of $\mathcal{E}$, hence $E_0=\{u_1=0\}$, $E_1=\{u_2-u_1=1\}$, $E_2=\{u_2=1\}$ will be a minimizer of (19) (for $k=2$). In addition, it will follow from the analysis in Section 4 (Proposition 4.3) that it is equivalent to solve

$$\min_{\mathbf{u}} J(\mathbf{u}) \;+\; \int_\Omega g_0(x)|1-u_1(x)| + g_1(x)|u_1(x)-u_2(x)| + g_2(x)|u_2(x)|\,dx \tag{38}$$

without any constraint on $\mathbf{u}$. We discuss the implementation of these problems in the following section.

3.4 Numerical analysis

Let us explain quickly how we solve (37) or (38). Both have a structure very similar to (22), so that we will use the same techniques.

As before, we consider the 2D case, in $\Omega=(0,1)^2$. At scale $h=1/N>0$, we introduce the discrete version of J, defined for $\mathbf{u}^h=((u_1^h)_{i,j},(u_2^h)_{i,j})_{0\le i,j\le N-1}\in X^h\times X^h$:

$$J^h(\mathbf{u}^h) \;=\; h^2 \sum_{0\le i,j<N} \Psi((\nabla^h u_1^h)_{i,j},(\nabla^h u_2^h)_{i,j}) \tag{39}$$

where Ψ is given by (35). Then, J^h is considered as a functional in $L^1(\Omega;\mathbb{R}^2)$, $\mathbf{u}^h$ being identified with the vectorial function

$$\mathbf{u}^h(x) \;=\; \sum_{0\le i,j<N} \mathbf{u}^h_{i,j}\chi_{(ih,jh)+[0,h)^2}(x)\,, \tag{40}$$

as previously in (40). The proof of the following result is then identical to the proof of the similar scalar result in Proposition 3.1, provided we use in addition to the approximation Theorem 2.1, the convergence (18):

Proposition 3.6. *J^h Γ-converges to $J(\mathbf{u})=\int_\Omega \Psi(D\mathbf{u})$ as $h\to 0$.*

This means that, again, we can approximate the minimization of (37) with the discrete minimization of

$$\min_{0\le u_2^h\le u_1^h\le 1} J^h(\mathbf{u}^h) \;+\; \int_\Omega (g_1(x)-g_0(x))u_1^h(x) + (g_2(x)-g_1(x))u_2^h(x)\,dx\,. \tag{41}$$

This has the same form as (27). We can introduce, therefore, a primal variable Our unknown is now $\mathbf{U} = (U_1, U_2) \in (X^h)^2$ and a dual variable $(\Xi_1, \Xi_2) \in (X^h \times X^h)^2$, and rewrite the problem as

$$\min_{0\leq(U_2)_{i,j}\leq(U_1)_{i,j}\leq 1} \max_{((\Xi_1)_{i,j},(\Xi_2)_{i,j})\in K} h^2 \sum_{i,j} (\Xi_1)_{i,j} \cdot (\nabla^h U_1)_{i,j} + (\Xi_2)_{i,j} \cdot (\nabla^h U_2)_{i,j} + h^2 \sum_{i,j} (U_1)_{i,j}(W_1^h)_{i,j} + (U_2)_{i,j}(W_2^h)_{i,j}, \quad (42)$$

where now,

$$(W_l^h)_{i,j} = \frac{1}{h^2} \int_{(ih,jh)+[0,h)^2} (g_l(x) - g_{l-1}(x))\, dx$$

for $l = 1, 2$, and each (i, j). Then, we proceed as in (29), with $\mathbf{W}^h = (W_1^h, W_2^h)$: our Arrow-Hurwicz type iteration is

$$\begin{array}{ll} \mathbf{U}_{i,j}^{n+1} = \Pi_{C_0}\left(\mathbf{U}_{i,j}^n + \tau\left((\mathrm{div}^h(\Xi_1,\Xi_2)^n)_{i,j} - \mathbf{W}_{i,j}^h\right)\right) & \text{for all } i,j \\ (\Xi_1,\Xi_2)_{i,j}^{n+1} = \Pi_K\left((\Xi_1,\Xi_2)_{i,j}^n + \tau'(\nabla^h \mathbf{U}^{n+1})_{i,j}\right) & \text{for all } i,j \end{array} \quad (43)$$

where here as in the forthcoming section 4.3, C_0 is the simplex $\{\mathbf{u} \in \mathbb{R}^2 : 0 \leq u_2 \leq u_1 \leq 1\}$. Projecting onto this simplex is very easy in 2D, see Appendix D, while projecting onto K is a bit more complicated and is done using Dykstra's iterarive algorithm, see Appendix C. The bound on $\tau\tau'$ is the same as in (29).

As before, the primal-dual gap may be computed at each step and used as a stopping criterion. Its expression is

$$J^h(\mathbf{U}^n) + h^2 \sum_{i,j} \mathbf{W}_{i,j}^h \cdot \mathbf{U}_{i,j}^n + h^2 \sum_{i,j} \max\{0, (\mathcal{P}_1^n)_{i,j}, (\mathcal{P}_1^n + \mathcal{P}_2^n)_{i,j}\} \geq 0 \quad (44)$$

where for $l = 1, 2$, $0 \leq i, j < N$, and $n \geq 0$,

$$(\mathcal{P}_l^n)_{i,j} = (\mathrm{div}^h \Xi_l^n)_{i,j} - (W_l^h)_{i,j}.$$

Here, $\Psi(p_1, p_2)$ in J^h is computed using expression (35): one just needs to compute the minimal distance between three points $0, p_1, p_2$ in $\mathbb{R}^2$, see Appendix E. This step becomes quite impractical in with more labels ($k \geq 3$), as it becomes quite difficult to evaluate (quickly) the primal energy $J^h(\mathbf{U})$.

It is suggested in [37], in order to avoid the projection step onto C_0, to replace problem (37) with (38). This is not so relevant in the problem with three phases, but will be quite important when the number of functions gets larger, in the next Section 4. In this case, the iterations are slightly changed. We let

$$(G_l^h)_{i,j} = \frac{1}{h^2} \int_{(ih,jh)+[0,h)^2} g_l(x)\, dx$$

for $l = 0, 1, 2$, and each (i, j). Then, we introduce an additional dual variable $(V_l)_{l=0,1,2}$, and write the problem as

$$\min_{\mathbf{U}=(U_1,U_2)} \max_{\substack{((\Xi_1)_{i,j},(\Xi_2)_{i,j}\in K \\ |(V_l)_{i,j}|\leq(G_l^h)_{i,j}}} h^2 \sum_{i,j} (\Xi_1)_{i,j} \cdot (\nabla^h U_1)_{i,j} + (\Xi_2)_{i,j} \cdot (\nabla^h U_2)_{i,j} + h^2 \sum_{i,j} (V_0)_{i,j}(1 - (U_1)_{i,j}) + (V_1)_{i,j}(U_1 - U_2)_{i,j} + (V_2)_{i,j}(U_2)_{i,j} \quad (45)$$

and the iterations (43) are replaced with the following (where each step is, of course, computed on all the grid $1 \leq i, j < N$):

$$\begin{aligned} (U_l^{n+1})_{i,j} &= (U_l^n)_{i,j} + \tau \left((\mathrm{div}^h \Xi_l^n)_{i,j} - (V_l - V_{l-1})_{i,j} \right) \qquad \text{for } l = 1, 2 \\ (\Xi_1, \Xi_2)_{i,j}^{n+1} &= \Pi_K \left((\Xi_1, \Xi_2)_{i,j}^n + \tau' (\nabla^h \mathbf{U}^{n+1})_{i,j} \right) \\ (V_l^{n+1})_{i,j} &= \Pi_{[-(G_l^h)_{i,j}, (G_l^h)_{i,j}]} \left((V_l^n)_{i,j} + \tau' (U_l^{n+1} - U_{l+1}^{n+1})_{i,j} \right) \end{aligned} \quad (46)$$

where here, by convention, $U_0^n \equiv 1$ and $U_{k+1}^n \equiv 0$ for all $n \geq 0$.

Then, we have convergence to the saddle point in (45), provided $\tau\tau' \leq 1/3$ (for $h = 1$).

3.5 Examples

To test this approach, we have implemented a very simple program where the input is a color image and the weight on each set E_0, E_1, E_2 is equal to minus the level of red, green and blue respectively. The result is a kind of projection to the closest of these three colors (of course, this gives terrible output for most images). The example on Figure 3, right, is an output for an input image with areas dominantly red, green or blue. The example corresponds to a value of $h = 1$ (the discretization step in (41)), and the weight is $(-2)\times$ the level of each channel, normalized between 0 and 1. The image has almost 120000 points (412×291). The gap (44) at convergence is ≤ 1 (for an energy of -122355).

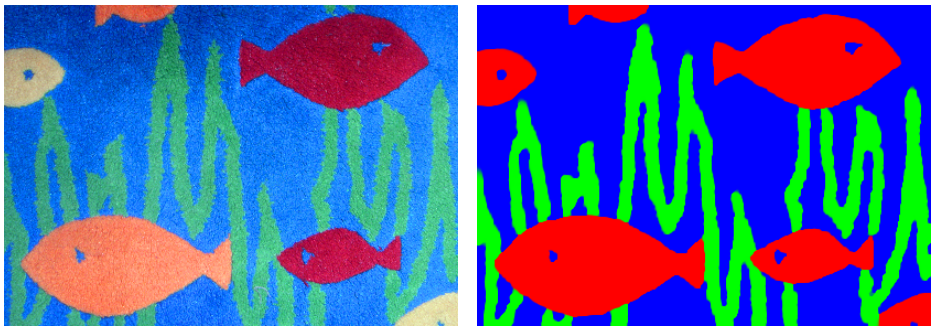

Figure 3: An example. Left: input image - Right: output

The total length energy ($J^h(\mathbf{u})$) of the output is 5897. (Hence, "most" of the energy is the external field term.) The output is "almost" binary (i.e., the fields u_1, u_2 are most of the time 0 and 1). In fact, we can compute a "width" of the interface, defined for instance as the ratio of the number of points where u_1 and u_2 are between 0.1 and 0.9, over the total "length" $J^h(\mathbf{u})$. In this example, we found .75 pixel units, which is quite narrow. This is not expected to be zero, because the discretization (39) of J requires a fuzzy interface to approximate precisely the length (just as (24), this is clear if one thinks of how the Γ-limit superior is established, point (ii) in the proof of Proposition 3.1).

Why it works so well. It is remarkable that in all the examples we tried, the three interfaces (R-B, R-G, G-B) seem really to be treated equally. To test this, we have tried an example where the output is driven to these three values by a very strong external field (pure red, green or blue) except in a grey area where no particular color (i.e., set E_i, $i = 0, 1, 2$) is favored. The result is impressive: what is expected, that is, a sharp discontinuity set with a triple point where all three interfaces meet with and angle of 120°, is actually computed by the program: see Figure 4. In this case, the weights g_i are 1/20

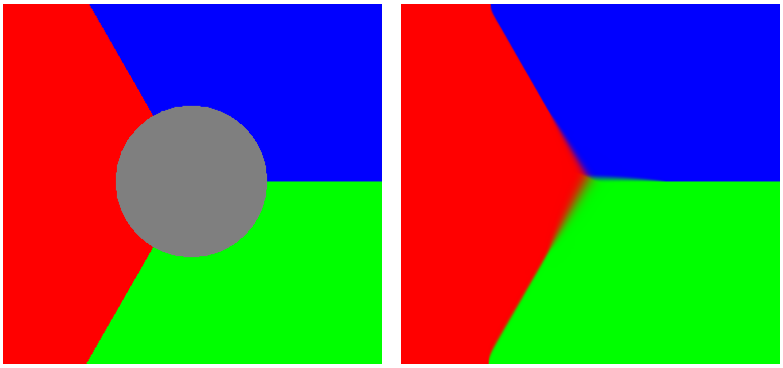

Figure 4: Reconstruction of a triple point (left: input, right: output)

in the red, green, blue area and equal in the grey area. This makes the length term quite important in the total energy, and actually the total length $J^h(\mathbf{u})$ which is computed is about 603 which is not much more of the "true" (isotropic) length of the total interface, estimated between 599 and 600 pixels. The size of the image is 376×357.

In fact, in trying to understand why this seemed to work so well, we could find both a reason which explains Figure 4 and a "counterexample," that is, a situation where we do not expect a binary solution ($u_1, u_2 \in \{0, 1\}$ a.e.), but rather a mixed one.

First of all, we can prove the following very easy proposition:

Proposition 3.7. *Let $\Omega \subset \mathbb{R}^2$ be a Lipschitz, connected set. Let $\bar{E}_1 = \{x = (x_1, x_2) \in \Omega \,:\, x_2 > 0\,, x_1 > -\sqrt{3}x_2\}$, $\bar{E}_2 = \{x = (x_1, x_2) \in \Omega \,:\, x_2 < 0\,, x_1 < \sqrt{3}x_2\}$, $\bar{E}_0 = \Omega \setminus (\bar{E}_1 \cup \bar{E}_2)$. Then*

$$\bar{\mathbf{u}} \;=\; (\bar{u}_1, \bar{u}_2) : x \;\mapsto \begin{cases} (0,0) & \text{if } x \in \bar{E}_0\,, \\ (1,0) & \text{if } x \in \bar{E}_1\,, \\ (1,1) & \text{if } x \in \bar{E}_2\,, \end{cases} \tag{47}$$

is a minimizer of $J(\mathbf{u})$ with prescribed boundary conditions $\mathbf{u} = \bar{\mathbf{u}}$ on $\partial\Omega$.

This is of particular importance, if one considers that in 2D, this is the only possible type of singularity in the boundary of the partition [5, 8], beyond a regular interface between two different labels.

Proof. The proof is straightforward and relies on a "calibration" argument. Let ξ_1, ξ_2 be the constant vector fields given by $\xi_1 \equiv (\sqrt{3}/2, 1/2)$ and $\xi_2 \equiv (0, -1)$ in Ω (see Figure 5). Let $\mathbf{u} : \Omega \to \mathbb{R}^2$ be a BV vector-valued function with $\mathbf{u} = \bar{\mathbf{u}}$

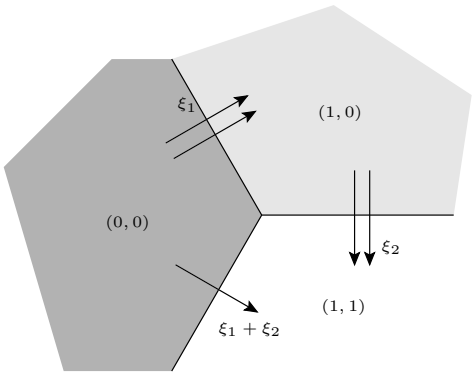

Figure 5: The "calibration" for the triple point

on $\partial\Omega$. Then, since $|\xi_1 + \xi_2| = |(\sqrt{3}/2, -1/2)| = 1$, $(\xi_1, \xi_2) \in K$ everywhere in Ω and we have

$$\begin{aligned} J(\mathbf{u}) \;&=\; \int_\Omega \Psi(D\mathbf{u}) \;\geq\; \int_\Omega \xi_1 \cdot Du_1 + \xi_2 \cdot Du_2 \\ &\geq\; \int_\Omega \xi_1 \cdot D\bar{u}_1 + \xi_2 \cdot D\bar{u}_2 \;+\; \int_\Omega \xi_1 \cdot D(u_1 - \bar{u}_1) + \xi_2 \cdot D(u_2 - \bar{u}_2) \end{aligned}$$

Now, by construction, $\int_\Omega \xi_1 \cdot D\bar{u}_1 + \xi_2 \cdot D\bar{u}_2 = J(\bar{\mathbf{u}})$ (indeed, ξ_1 has norm one and is orthogonal to the interface $\partial\bar{E}_0 \cap \partial\bar{E}_1$, ξ_2 has norm one and is orthogonal to the interface $\partial\bar{E}_1 \cap \partial\bar{E}_2$, and $\xi_1 + \xi_2$ has norm one and is orthogonal to the interface $\partial\bar{E}_0 \cap \partial\bar{E}_2$). On the other hand, since $\mathbf{u} = \bar{\mathbf{u}}$ on $\partial\Omega$,

$$\int_\Omega \xi_1 \cdot D(u_1 - \bar{u}_1) + \xi_2 \cdot D(u_2 - \bar{u}_2) = -\int_\Omega (u_1 - \bar{u}_1)\mathrm{div}\,\xi_1 + (u_2 - \bar{u}_2)\mathrm{div}\,\xi_2\,dx = 0$$

since both fields ξ_1, ξ_2 have vanishing divergence. We deduce $J(\mathbf{u}) \geq J(\bar{\mathbf{u}})$. $\square$

Remark 3.8. Although we still have not performed simulations in three dimensions, it seems that the structure of minimal surfaces in 3D and their singularities, as described by J. Taylor [42], is such that for all three types of singular minimal cones which she describes, there should exist a calibration made of constant vector fields as in the 2D situation. We mention that two of these minimal cones only can appear in the three phases problem (a plane and three planes meeting on one line, that is the 2D triple point extended by translation in the orthogonal direction), while the third one (built upon a regular tetraedron) involves at least four phases, however, it still has the property that the normal to any interface is just the sum of vectors which are normal to two or three other interfaces.

And why, sometimes, it does not. Now, let us consider the same geometry, but instead of a boundary datum, a weight g_0, g_1, g_2 such that on each $\bar{E}_i$, $g_i = 1$ while $g_j = 0$ for $j \neq i$, except on a disk centered at the triple point where we choose $g_0 = g_1 = g_2$. Equivalently, we run our program with as input, $\bar{E}_0$ colored pure cyan $((R, G, B) = (0, 1, 1))$, $\bar{E}_1$ pure yellow $(1, 1, 0)$, $\bar{E}_2$ pure magenta $(1, 0, 1)$, except a grey area in the middle (see Figure 6, left). Then, we find that the grey circle is completed with a perfect triple point, but this time with mixed colors, that is, cyan $((R, G, B) = (0, .5, .5))$ in $\bar{E}_0$, yellow $(.5, .5, 0)$ in $\bar{E}_1$, magenta $(.5, 0, .5)$ in $\bar{E}_2$, or more precisely, $\mathbf{u} = (1, 1/2)$ in $\bar{E}_0$, $(1/2, 1/2)$ in $\bar{E}_1$, $(1/2, 0)$ in $\bar{E}_2$. In other words, the optimal solution is a *mixture* of half E_j and half E_k in $\bar{E}_i$, for $\{i, j, k\} = \{0, 1, 2\}$. The issue is that, now, the opposite of the "calibration" (ξ_1, ξ_2) in Figure 5 can show that this is actually a local minimizer inside the grey circle, as before (with the same proof as Proposition 3.7, replacing $\bar{\mathbf{u}}$ with the solution above and ξ_i with $-\xi_i$). In particular, the length energy $J^h(\mathbf{u})$ which is computed now is 306, about half the true length, thus strictly below the length of any binary solution with the same pattern of discontinuity (and, actually, any reasonable binary solution for this input data). We must add that convergence of our scheme is very slow (and poor) for this particular example.

In fact, it really seems that the example in Figure 6 corresponds to a situation where $J(\mathbf{u}) < (J^0)^{**}(\mathbf{u})$, were we recall that $(J^0)^{**}$ is the convex, l.s.c. envelope of J^0 defined in (32). Consider indeed in $\Omega = B(0, 1)$ and the functions $\mathbf{u}^1, \mathbf{u}^2 : \Omega \to [0, 1]^2$ given by $\mathbf{u}^1 = (0, 0)$ in $\Omega \cap \bar{E}_1$, $\mathbf{u}^1 = (1, 0)$ in $\Omega \setminus \bar{E}_1$, and $\mathbf{u}^2 = (0, 0)$ in $\Omega \cap \bar{E}_2$, $\mathbf{u}^2 = (1, 1)$ in $\Omega \setminus \bar{E}_2$. Then, $J(\mathbf{u}^l) = J^0(\mathbf{u}^l) = 2$, $l = 1, 2$. Letting $\mathbf{u} = (\mathbf{u}^1 + \mathbf{u}^2)/2$, we find the function $\mathbf{u}$ computed above and we have $J(\mathbf{u}) = \int_{J_\mathbf{u}} \Psi((\mathbf{u}_+ - \mathbf{u}_-) \otimes \nu_\mathbf{u})\, d\mathcal{H}^1 = 3/2$ since the total length of the jump is 3,

Figure 6: A nonbinary solution (left: input, right: output)

while the amplitude is $1/2$. We believe that, on the other hand, $(J^0)^{**}(\mathbf{u}) = 2$. Actually, a way to define the convex l.s.c. envelope $(J^0)^{**}$ of J^0 is through formula (31) (with $\mathcal{E}$ replaced with J^0), and we conjecture that in this formula, $\mathbf{u} = (\mathbf{u}^1 + \mathbf{u}^2)/2$ is the optimal decomposition of $\mathbf{u}$ as a convex combination of functions in the domain of J^0, so that $(J^0)^{**}(\mathbf{u}) = (J(\mathbf{u}_1) + J(\mathbf{u}_2))/2 = 2 > J(\mathbf{u})$.

More generally, a standard strategy to prove that $J = (J^0)^{**}$ in $\mathcal{C}_0 = \{\mathbf{u} \in BV(\Omega; \mathbb{R}^2) \, : \, 0 \leq u_2 \leq u_1 \leq 1 \text{ a.e.}\}$ could be, using definition (31) of the convex envelope, as follows: first, using the approximation Theorem 2.1 and (18), and standard finite elements theory, we can show that it is enough to establish the equality for piecewise affine $\mathbf{u}$. Then, one first shows that for affine $\mathbf{u}$ (of the form $x \mapsto \mathbf{a} + \mathbf{p}^T x$, and with $0 \leq u_2 \leq u_1 \leq 1$ a.e.), there actually exist $\mathbf{u}^l, \theta^l \geq 0$ with $\sum_l \theta^l = 1$, $\sum_l \theta^l \mathbf{u}^l$ arbitrarily close to $\mathbf{u}$ in L^1, and $\sum_l \theta^l J^0(\mathbf{u}^l)$ arbitrarily close to $J(\mathbf{u}) = |\Omega| \Psi(\mathbf{p})$. The construction is based on the expression (35) of $\Psi(\mathbf{p})$ which yields the existence of $q \in \mathbb{R}^d$, with $\Psi(\mathbf{p}) = |q| + |p_1 - q| + |p_2 - q|$. Then, the functions $(q \cdot x, q \cdot x)$, $(p_1 \cdot x, p_2 \cdot x)$ are approximated by piecewise constant functions.

The next step would be to "glue" together the approximations of affine pieces of $\mathbf{u}$ to approximate a piecewise affine $\mathbf{u}$. It seems that this can not be done, in general, in dimension $d > 1$.

Hence, a natural question is: why does the numerical method seem to work, most of the time? We have no real answer. First of all, when the data term dominates, then clearly binary solutions are favored. Indeed, $\mathbf{u}$ is constrained to stay in the simplex $\{0 \leq u_2 \leq u_1 \leq 1\}$, whose vertices are $(0,0)$, $(1,0)$ and $(1,1)$ and except in degenerate cases (as in Figure 6), optimal values of the term

$$\int_\Omega (g_1(x) - g_0(x)) u_1^h(x) + (g_2(x) - g_1(x)) u_2^h(x) \, dx \, .$$

in (41) are reached at the vertices.

On the other hand, when the data term gets smaller (in all our computations, a parameter $\lambda > 0$ is introduces and the input $(g_i)_{i=0}^2$ is first divided by λ), then the perimeter term J^h becomes more important and situations such as illustrated in Figure 6 should appear more often. It seems, though, that it is not the case. The reason could be that J, minimized with binary boundary conditions as in Proposition 3.7, produces solutions which take inside the domain only (binary) boundary values? We do not know whether this is true, but it is far from excluded and is the subject of future study.

In the next section, we discuss our approach with more labels and more general interaction energies. Further experiments are then shown in Section 4.5.

4 Extension to more general partition problems

4.1 A convex, local envelope of the general partition problem

Now, we explain how our setting is generalized to higher dimension and an arbitrary number of labels, and possibly to more complex interaction energies (including, for instance, the "truncated TV" (7). Let $d \geq 1$ be the space dimension, and $k+1 \geq 2$ the number of labels. Let Ω be a bounded open subset of $\mathbb{R}^d$. We need k functions $u_1, \dots, u_k$ to represent the $k+1$ labels (as before, $\chi_{E_i} \sim u_i - u_{i+1}$, $1 \leq i < k$, $\chi_{E_0} \sim 1 - u_1$, $\chi_{E_k} \sim u_k$).

For $u \in BV(\Omega; \{0, \dots, k\})$, the interfacial energy we consider is of the following type:

$$\mathcal{J}(u) \;=\; \int_{J_u} \sigma(u_+(x) - u_-(x))\, d\mathcal{H}^{d-1}(x)$$

where $\sigma : (0, +\infty) \to (0, +\infty)$ is a nondecreasing, positive concave function (hence subbaditive: $\sigma(a) + \sigma(b) \leq \sigma(a+b)$ for any $a, b > 0$). We are interested in solving

$$\min_{u \in BV(\Omega; \{0,\dots,k\})} \mathcal{J}(u) \;+\; \int_\Omega W(x, u(x))\, dx \tag{48}$$

for some potential W such that $W(x, i) \in L^\infty(\Omega; \mathbb{R}_+)$ for each $i = 0, \dots, k$. If $\sigma \equiv 1$ and $W(x, i) = g_i(x)$, this boils down to (19). The interaction (7) corresponds to $\sigma(t) = t$ if $t < T$, and T else.

Introducing the same representation as before, that is a vector valued map $\mathbf{u} \in BV(\Omega; \{0,1\}^k)$, with $u_{i+1} \leq u_i$ for $1 \leq i < k$, and $u(x) = \sum_{i=1}^k u_i(x)$ (in other words, $u_i = \chi_{\{u \geq i\}}$), we recast our variational problem as

$$\min J^0(\mathbf{u}) \;+\; \sum_{i=1}^k \int_\Omega (W(x, i) - W(x, i-1)) u_i(x)\, dx \tag{49}$$

where now,

$$J^0(\mathbf{u}) \;=\; \mathcal{J}(u) \;=\; \sum_{1\le i_1\le i_2\le k} \sigma(i_2-(i_1-1))\mathcal{H}^{d-1}\left(\bigcap_{i_1\le i\le i_2} J_{u_i}\setminus (J_{u_{i_1-1}}\cup J_{u_{i_2}})\right),$$

that is: we pay $\sigma(i_2-(i_1-1))$ when all functions u_i between i_1 and i_2 jump simultaneously from 0 to 1, while the other do not jump. (Remember that the constraints $0\le u_k\le u_{k-1}\le\cdots\le 1$, and $u_i\in\{0,1\}$ a.e., yield that at each jump point a maximal group of functions $(u_i)_{i_1\le i\le i_2}$ jump simultaneously.)

The interaction potential may be rewritten

$$J^0(\mathbf{u}) \;=\; \int_\Omega \Psi^0(D\mathbf{u})$$

where $\Psi^0(0)=0$, and $\Psi(\mathbf{p})=\sigma(i_2-(i_1-1))|p|$ when there exists $p\in\mathbb{R}^d$ such that $\mathbf{p}=(p_1,\dots,p_k)$ is given by $p_i=p$ when $i_1\le i\le i_2$, and $p_i=0$ if $i<i_1$ or $i>i_2$. We let $\Psi^0(\mathbf{p})=+\infty$ for vectors not of these forms. As in the previous section, we introduce the convexification $(J^0)^{**}$, and, the maximal local functional of the form $\int_\Omega\Psi(D\mathbf{u})$ below J^0. Then, the following results, which generalize Lemma 3.3 and Proposition 3.5 in the previous section, are true again, with the same proofs:

Lemma 4.1. *For $\mathbf{p}\in\mathbb{R}^{d\times k}$, let*

$$\Psi^1(\mathbf{p}) \;:=\; \sum_{i=1}^k |p_i| \quad \textit{and} \quad \Psi^\infty(\mathbf{p}) \;:=\; \max_{i=1,\dots,k}|p_i|\,.$$

Then for any $\mathbf{u}\in\mathcal{C}_0:=\{\mathbf{u}\in BV(\Omega;[0,1]^k)\,:\,u_k\le u_{k-1}\le\cdots\le u_1$ a.e.$\}$ we have:

$$\sigma(1)\int_\Omega\Psi^\infty(D\mathbf{u}) \;\le\; (J^0)^{**}(\mathbf{u}) \;\le\; \sigma(k)\int_\Omega\Psi^1(D\mathbf{u})\,, \tag{50}$$

while if $\mathbf{u}\in L^1(\Omega;\mathbb{R}^k)\setminus\mathcal{C}_0$, $(J^0)^{**}(\mathbf{u})=+\infty$.

Proposition 4.2. *Let $\Psi=(\Psi_0)^{**}$: then, the largest convex, l.s.c. and local functional, of the form (34), below J^0 is given, for $\mathbf{u}\in L^1(\Omega;\mathbb{R}^k)$, by*

$$J(\mathbf{u}) \;=\; \begin{cases} \int_\Omega\Psi(D\mathbf{u}) & \textit{if } \mathbf{u}\in BV(\Omega;\mathbb{R}^k)\\ +\infty & \textit{else.}\end{cases}$$

We now compute Ψ. As before,

$$\begin{aligned}(\Psi^0)^*(\mathbf{q}) \;&=\; \sup_{\mathbf{p}}\,\mathbf{p}\cdot\mathbf{q} \;-\; \Psi^0(\mathbf{p})\\ &=\; \sup_{p\in\mathbb{R}^d}\max_{1\le i_1\le i_2\le k}\Big(\big(p\cdot\sum_{i_1\le i\le i_2}q_i\big)-\sigma(i_2-(i_1-1))|p|\Big)\\ &=\; \begin{cases}0 & \text{if } \left|\sum_{i_1\le i\le i_2}q_i\right|\le\sigma(i_2-(i_1-1))\ \forall 1\le i_1\le i_2\le k\,,\\ +\infty & \text{else.}\end{cases}\end{aligned}$$

Hence, it is the characteristic function of the convex set $K \subset \mathbb{R}^{d\times k}$, defined as

$$K \;=\; \Big\{\mathbf{q}=(q_1,\dots,q_k)\in\mathbb{R}^{d\times k} \;:\; \text{for all } i, \\ \Big|\textstyle\sum_{i_1\le i\le i_2} q_i\Big| \le \sigma(i_2-(i_1-1))\;\forall\;(i_1,i_2) \text{ with } 1\le i_1\le i_2\le k\Big\} \quad (51)$$

Thus,

$$\Psi(\mathbf{p}) \;=\; \sup_{\mathbf{q}\in K} \mathbf{q}\cdot\mathbf{p}$$

and the functional $J(\mathbf{u}) \;=\; \int_\Omega \Psi(D\mathbf{u})$ is given, for any $\mathbf{u}\in L^1(\Omega;\mathbb{R}^k)$, by

$$J(\mathbf{u}) \;=\; \sup\Big\{ -\int_\Omega \sum_{i=1}^k u_i \mathrm{div}\,\xi_i\,dx \;:(\xi_1,\dots\xi_k)\in C_c^\infty(\Omega;K)\Big\}.$$

It is clear that this defines a convex, one-homogeneous and lower-semicontinuous functional on $L^1(\Omega;\mathbb{R}^k)$. It clearly also satisfies (50), just as $(J^0)^{**}$.

Again, to solve (49), we will solve numerically

$$\min\Big\{J(\mathbf{u})+\sum_{i=1}^k\int_\Omega (W(x,i)-W(x,i-1))u_i(x)\,dx, \mathbf{u}=(u_1,\dots,u_k)\in\mathcal{C}_0\Big\} \quad (52)$$

and: if the solution is "binary", that is if $u_i\in\{0,1\}$ a.e. for all i, then it will be a solution of the original problem.

It is proposed in [37], where a similar problem is solved with $\sigma(t)=t$, to replace (52) by the following, unconstrained problem:

$$\min_{\mathbf{u}\in BV(\Omega;\mathbb{R}^k)} J(\mathbf{u}) \;+\; \int_\Omega W(x,0)|1-u_1(x)|\,dx \\ +\sum_{i=1}^{k-1}\int_\Omega W(x,i)|u_i(x)-u_{i+1}(x)|\,dx \;+\; \int_\Omega W(x,k)|u_k(x)|\,dx \quad (53)$$

Indeed, we can show the following proposition:

Proposition 4.3. *Assume $W(x,i)>0$ a.e, and for each i. Then, any solution of* (53) *satisfies $0\le u_k\le u_{k-1}\le\cdots\le u_1\le 1$ a.e. in Ω, and hence is a solution of* (52)

Remark 4.4. There is no loss of generality in assuming that the $W(x,i)$ are positive, as soon as they are bounded from below (even, in fact, by a given L^1 function which does not depend on the label i): indeed, given any $V\in L^1(\Omega)$, any solution of (52) is clearly a solution of the same problem with the potentials $W(x,i)$ replaced with $W'(x,i)=W(x,i)-V(x)+1$ which is positive if $W(\cdot,i)\ge V$ a.e., for each i.

The proof of Proposition 4.3 relies on some truncation properties of the functional J and will be a consequence of the Lemmas 4.5 and 4.6 which are proved in the following sections 4.2 and 4.3: we postpone it to section 4.4.

4.2 Ordering of the vector u

Let us denote by $C \subset \mathbb{R}^k$ the convex $C = \{\mathbf{z} \in \mathbb{R}^k \,:\, z_1 \geq z_2 \geq \cdots \geq z_k\}$. For $\mathbf{u} \in \mathbb{R}^k$, we denote by $\Pi_C(\mathbf{u})$ the orthogonal projection of $\mathbf{u}$ onto C. The projection $\mathbf{u}' = \Pi_C(\mathbf{u})$ is a vector of the following form: there exist $1 = k_1 \leq k_2 \leq \cdots \leq k_{l+1} = k+1$ such that for each $n = 1, \ldots, l$ and each i with $k_n \leq i < k_{n+1}$,

$$u'_i \;=\; \frac{1}{k_{n+1} - k_n} \sum_{j=k_n}^{k_{n+1}-1} u_j \tag{54}$$

Indeed, the sets $\{k_n, \ldots, k_{n+1} - 1\}$, when containing more than one index, are just the maximal clusters of indices whose associated coefficients become equal in the projection, and $\mathbf{u}'$ is then simply the projection of $\mathbf{u}$ onto $\{\mathbf{z} \in \mathbb{R}^k \,:\, z_i = z_j \text{ if } \exists n\,, k_n \leq i \leq j < k_{n+1}\}$.

In addition, we claim that the projection $\mathbf{u}'$ must satisfy, for each $n = 1, \ldots, l$,

$$u_{k_n} \;\leq\; u'_{k_n} \;=\; u'_{k_{n+1}-1} \;\leq\; u_{k_{n+1}-1} \tag{55}$$

Indeed, if for instance $u_{k_n} > u'_{k_n}$ for some n, we define a vector $\mathbf{u}'' \in C$ by $u''_i = u'_i$ if $i \neq k_n$, and $u''_{k_n} = \min\{u'_{k_{n-1}}, u_{k_n}\} > u'_{k_n}$: then, since clearly $(u_{k_n} - u''_{k_n})^2 < (u_{k_n} - u'_{k_n})^2$, the vector $\mathbf{u}''$ is closer to $\mathbf{u}$ than $\mathbf{u}'$, a contradiction.

Observe in particular that (55) yields that if $u'_{i-1} > u'_i$ (that is, the index i is one of the k_n, $2 \leq n \leq l$), we must have $u_{i-1} > u_i$, hence, conversely, if $u_{i-1} \leq u_i$ then $u'_{i-1} = u'_i$. An recursive algorithm for computing the projection $\Pi_C(\mathbf{u})$ can be deduced from this remark, see Appendix D.

Now, let us show that the projection onto C decreases the energy J:

Lemma 4.5. *For any* $\mathbf{u} \in BV(\Omega; \mathbb{R}^k)$, $J(\Pi_C(\mathbf{u})) \leq J(\mathbf{u})$.

Here, $\Pi_C(\mathbf{u})$ denotes (clearly) the function $x \mapsto \Pi_C(\mathbf{u}(x))$.

Proof. First, using Theorem 2.1 and the convergence (18), together with the lower-semicontinuity of J, it is enough to show the lemma for a smooth $\mathbf{u} \in C^\infty(\Omega; \mathbb{R}^k)$. In particular, $\Pi_C(\mathbf{u})$ is continuous: we even have (since Π_C is 1-Lipschitz) $|\nabla \Pi_C(\mathbf{u})(x)| \leq |\nabla \mathbf{u}(x)|$ for a.e. $x \in \Omega$. (In particular, the projection decreases the Euclidean total variation.)

As seen above, we can cover Ω by finitely many closed sets A associated each to a particular partition of $\{1, \ldots, k\}$ into l clusters $\{k_n, \ldots, k_{n+1}\}$, with the same notation as above. Letting $\hat{\mathbf{u}}$ be defined by

$$\hat{u}_i \;=\; \frac{1}{k_{n+1} - k_n} \sum_{j=k_n}^{k_{n+1}-1} u_j$$

whenever $k_n \le i < k_{n+1}$, the set A is simply $\{x \in \Omega : \Pi_C(\mathbf{u}) = \hat{\mathbf{u}}\}$. This implies $\nabla \Pi_C(\mathbf{u}) = \nabla \hat{\mathbf{u}}$ a.e. in A, hence

$$\int_A \Psi(\nabla \hat{\mathbf{u}}(x))\, dx \;=\; \int_A \Psi(\nabla \Pi_C(\mathbf{u})(x))\, dx\,. \tag{56}$$

Hence to prove the lemma, it is enough to show that

$$\Psi(\nabla \hat{\mathbf{u}}(x)) \le \Psi(\nabla \mathbf{u}(x)) \tag{57}$$

a.e. in A. But in Ω we have:

$$\nabla \hat{u}_i \;=\; \frac{1}{k_{n+1} - k_n} \sum_{j=k_n}^{k_{n+1}-1} \nabla u_j$$

for $k_n \le i < k_{n+1}$. Now, if $(\xi_i)_{i=1}^k \in K$, we have for any $x \in \Omega$

$$\sum_{i=1}^k \xi_i \cdot \nabla \hat{u}_i(x) \;=\; \sum_{n=1}^l \sum_{i=k_n}^{k_{n+1}-1} \frac{\left(\sum_{j=k_n}^{k_{n+1}-1} \xi_j\right)}{k_{n+1} - k_n} \cdot \nabla u_i(x) \;=\; \sum_{i=1}^k \hat{\xi}_i \cdot \nabla u_i(x) \tag{58}$$

where for each i with $k_n \le i < k_{n+1}$, we have let

$$\hat{\xi}_i \;=\; \frac{1}{k_{n+1} - k_n} \left(\sum_{j=k_n}^{k_{n+1}-1} \xi_j \right).$$

If we show that $(\hat{\xi}_i)_{i=1}^k \in K$, it will follow from (58) that $\sum_i \xi_i \nabla \hat{u}_i(x) \le \Psi(\nabla \mathbf{u}(x))$, from which (57) will follow.

We let $\xi_i^0 = \xi_i$ for each i, and for $n = 1, \ldots, l$ we let $\xi_i^n = \hat{\xi}_i$ if $i < k_{n+1}$, and $\xi_i^n = \xi_i^{n-1} = \xi_i$ if $i \ge k_{n+1}$. In other words, ξ^n is obtained from ξ^{n-1} by averaging all components between k_n and $k_{n+1} - 1$, and leaving the other unchanged. Let us show by induction that $(\xi_i^n)_{i=1}^k \in K$, for each $n \le l$: since $(\hat{\xi}_i)_{i=1}^k = (\xi_i^l)_{i=1}^k$, the thesis will follow.

This is true for $n = 0$, by assumption. Assume $n \ge 1$ and $(\xi_i^{n-1})_{i=1}^k \in K$. The only non-obvious conditions to check to prove that $(\xi_i^n)_{i=1}^k \in K$ are of the kind

$$\left| \textstyle\sum_{i_1 \le i \le i_2} \xi_i^n \right| \;\le\; \sigma(i_2 - (i_1 - 1))\,,$$

with either $i_1 < k_n \le i_2 < k_{n+1}$ or $k_n \le i_1 < k_{n+1} \le i_2$. In the first case, for instance, simple algebra shows that this sum is the convex combination

$$\left(1 - \frac{1 + i_2 - k_n}{k_{n+1} - k_n}\right) \sum_{i_1 \le i < k_n} \xi_i^{n-1} \;+\; \frac{1 + i_2 - k_n}{k_{n+1} - k_n} \sum_{i_1 \le i < k_{n+1}} \xi_i^{n-1}\,,$$

hence its norm is less than, using the concavity of σ:

$$\left(1 - \frac{1+i_2-k_n}{k_{n+1}-k_n}\right)\sigma(k_n - i_1) + \frac{1+i_2-k_n}{k_{n+1}-k_n}\sigma(k_{n+1}-i_1) \\ \leq \sigma\left(\left(1 - \tfrac{1+i_2-k_n}{k_{n+1}-k_n}\right)(k_n - i_1) + \tfrac{1+i_2-k_n}{k_{n+1}-k_n}(k_{n+1}-i_1)\right).$$

Writing $k_{n+1} - i_1 = (k_{n+1} - k_n) + (k_n - i_1)$ in the last term, we find that the argument of σ is in fact equal to $(k_n - i_1) + (1 + i_2 - k_n) = i_2 - (i_1 - 1)$, and the thesis follows. The second case is treated in the same way.

This achieves the proof of the lemma. □

4.3 Truncation of the coordinates

Now, we define $C_0 = C \cap [0,1]^k$. Observe that $\Pi_{C_0} = \Pi_{[0,1]^k} \circ \Pi_C$: indeed, for any $\mathbf{u} \in \mathbb{R}^k$ and $\mathbf{z} \in C_0$,

$$\begin{aligned}
&(\mathbf{u} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}) \cdot (\mathbf{z} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}) \\
&= (\mathbf{u} - \Pi_C\mathbf{u}) \cdot (\mathbf{z} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}) + (\Pi_C\mathbf{u} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}) \cdot (\mathbf{z} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}) \\
&\leq (\mathbf{u} - \Pi_C\mathbf{u}) \cdot (\mathbf{z} - \Pi_C\mathbf{u}) + (\mathbf{u} - \Pi_C\mathbf{u}) \cdot (\Pi_C\mathbf{u} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}) \\
&\leq (\mathbf{u} - \Pi_C\mathbf{u}) \cdot (\Pi_C\mathbf{u} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}) = 0
\end{aligned}$$

since the vector $\Pi_C\mathbf{u} - \Pi_{[0,1]^k}\Pi_C\mathbf{u}$ has constant coefficients on each cluster of constants coefficients of $\Pi_C\mathbf{u}$. Then, we have in addition:

Lemma 4.6. *For any $\mathbf{u} \in BV(\Omega;\mathbb{R}^k)$, $J(\Pi_{C_0}(\mathbf{u})) \leq J(\mathbf{u})$.*

Here again, $\Pi_{C_0}(\mathbf{u}) \in \mathcal{C}_0$ is $x \mapsto \Pi_{C_0}(\mathbf{u}(x))$.

Proof. Since $J(\Pi_C(\mathbf{u})) \leq J(\mathbf{u})$, we just need to show that $J(\Pi_{[0,1]^k}(\mathbf{u})) \leq J(\mathbf{u})$ whenever $\mathbf{u} \in C$ a.e, moreover, as before, we may assume that $\mathbf{u}$ is smooth. Then, the same arguments as in the proof of the previous lemma show that the inequality is true provided we can show that for any i_1, i_2 with $1 \leq i_1 \leq i_1 \leq k$, and any $\mathbf{p} \in \mathbb{R}^{d\times k}$, if $\hat{\mathbf{p}}$ is defined by $\hat{p}_i = p_i$ when $i_1 \leq i \leq i_2$ and $\hat{p}_i = 0$ else, $\Psi(\hat{\mathbf{p}}) \leq \Psi(\mathbf{p})$. But this is an obvious consequence of the fact that if $(\xi_i)_{i=1}^k \in K$, also $(\hat{\xi}_i)_{i=1}^k$, defined by $\hat{\xi}_i = \xi_i$ whenever $i_1 \leq i \leq i_2$ and $\hat{\xi}_i = 0$ else, belongs to K. □

4.4 Proof of proposition 4.3

Proof. Let $\mathbf{u}$ solve (53). Let $\mathbf{u}' = \Pi_{C_0}\mathbf{u}$. By Lemma 4.6, we have $J(\mathbf{u}') \leq J(\mathbf{u})$. Let us now consider the other terms in (53), that is,

$$\int_\Omega \left(W(x,0)|1 - u_1(x)| + \sum_{i=1}^{k-1} W(x,i)|u_i(x) - u_{i+1}(x)| + W(x,k)|u_k(x)|\right) dx$$

We see that when we replace $\mathbf{u}$ with, first, $\mathbf{u}'' = \Pi_C \mathbf{u}$ in this expression, some terms of the form $|u_i(x) - u_{i+1}(x)|$ will vanish (hence decreasing the energy, since $W(\cdot, i) > 0$ a.e.), while, thanks to (55), if $|u_i''(x) - u_{i+1}''(x)| \neq 0$, we have $u_i(x) \geq u_i''(x) > u_{i+1}''(x) \geq u_{i+1}(x)$ so that $|u_i''(x) - u_{i+1}''(x)| \leq |u_i(x) - u_{i+1}(x)|$. On the other hand, the effect of projecting then $\mathbf{u}''$ onto $[0,1]^k$, to find $\mathbf{u}'$, does not alter most terms in the sum from $i = 1$ to $k-1$, but strictly reduces the first and last term ($|1 - u_1(x)|$ and $|u_k(x)|$) if $\mathbf{u}''$ was not already in $[0,1]^k$, while it might also reduce some other terms of the form $|u_i(x) - u_{i+1}(x)|$ whenever $u_i(x) > u_{i+1}(x)$. Hence, the energy is strictly reduced by replacing $\mathbf{u}$ with $\mathbf{u}'$, unless $\mathbf{u} = \mathbf{u}'$, that is, $\mathbf{u} \in C_0$. □

4.5 Numerical analysis and results

We do not detail our implementation here: indeed, it follows the same lines as in the three-labels case, in Section 3.4. The discrete approximation is the same and the Γ-convergence result in Proposition 3.6 still holds. Then, we use the same type of Arrow-Hurwicz' iteration, as described in (46). (Also (43) could be used, but then a reprojection onto C_0 is necessary, as described in Appendix D.)

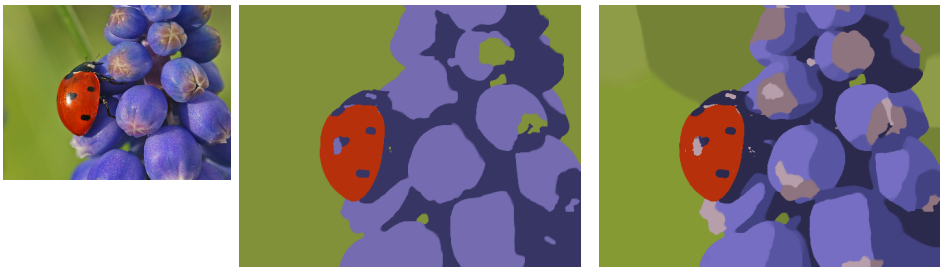

Figure 7: Basic piecewise constant segmentation: Middle: 4 regions – Right: 10 regions

The slowest part in this method is the computation of the projection onto K of the dual variable $(\Xi_1, \ldots, \Xi_k) \in \mathbb{R}^{d \times k}$. This motivates the use of a GPU. This step is done using Dykstra's algorithm as described in Appendix C. The set K is described as the intersection of $k(k+1)/2$ "simple" convex sets and an iterate of Dykstra's algorithm must therefore involve as many projections and dual variables. Also, the primal-dual gap, which has essentially the same expression as in (44), is not as easy to compute than in the 2 or 3-labels cases (since evaluating Ψ requires an iterative algorithm), so that we must choose a more arbitrary stopping criterion for our iterations (like the variation between two successive iterates).

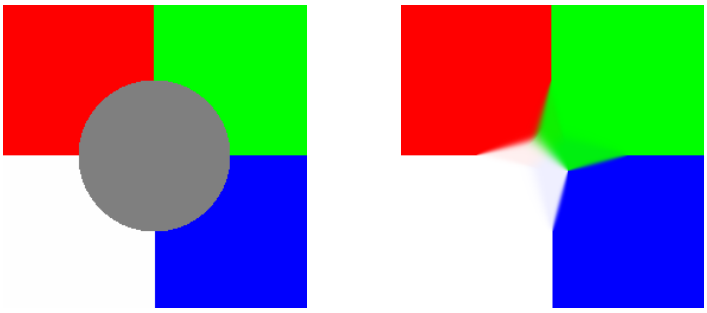

Figure 8: Completion of four regions

We first show an example of "basic" image segmentation (following a piecewise constant Mumford-Shah model, see [36, 17]). The idea is to solve alternatively (for fixed k)

$$\min_{(E_i,a_i)_{i=0}^k} \min \left\{ \frac{\lambda}{2} \sum_{i=0}^{k} \mathrm{Per}(E_i, \Omega) + \sum_{i=0}^{k} \int_{E_i} (I(x) - a_i)^2 \, dx \right\}$$

with respect to $(E_i)_{i=0}^k$ and then to $(a_i)_{i=0}^k$. Here, $I : \Omega \to [0,1]^3$ is the color information (intensity of red, green, blue channels) of the original image. The initial values $(a_i)_{i=0}^k$ (also vector-valued) are initialized using the k-means algorithm, and then, once a new partition (E_i) is found, each a_i is updated by computing the average value of I in E_i. Figure 1 was computed in this way, while Figure 7 shows another example, once with 4 and then with 10 labels. Again, as for three labels, in almost all our experiment the solution was nearly binary (up to a smoothing of the discontinuities due to the discretization).

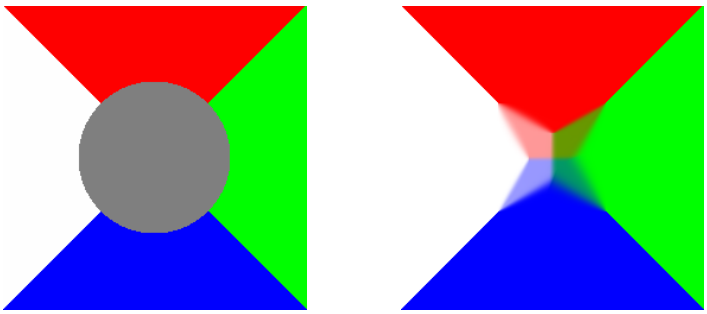

Figure 9: Completion of four regions: in case of non uniqueness, the method may find a combination of the solutions

It is interesting to check how, as in the experiments in Figure 4 for three-labels segmentation, the method completes a boundary datum inside a given area. Figure 8 shows the completion of four regions with equidistant labels Here the values of $(a_i)_{i=0}^3$ are kept fixed, and span the four color intensity values

$(1,0,0),(0,1,0),(0,0,1),(1,1,1)$. The result is what is theoretically expected (two triple junctions, with angles of 120°). However, we must point out that the solution which is found here is only one out of two. It seems that the program selects this particular solution because the discretization makes it minimal. If one rotates the original image by 45°, then both solutions have the same energy even for the discrete problem, and our program produces an output which is not binary, but a convex combination of the two minimal binary solutions, as shown in Figure 9. We insist that it does not correspond to a situation where

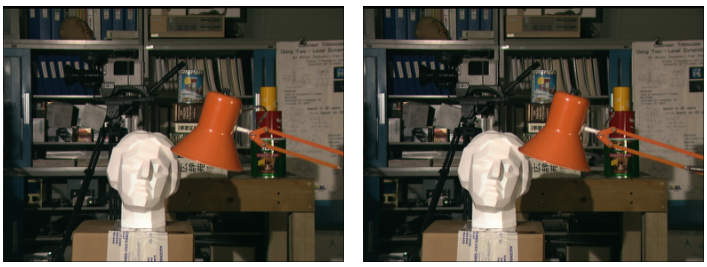

Figure 10: A stereo pair

the convexification is too low (as in Fig. 6), but just a case of non uniqueness, where any convex combination of the binary solutions is also minimizing.

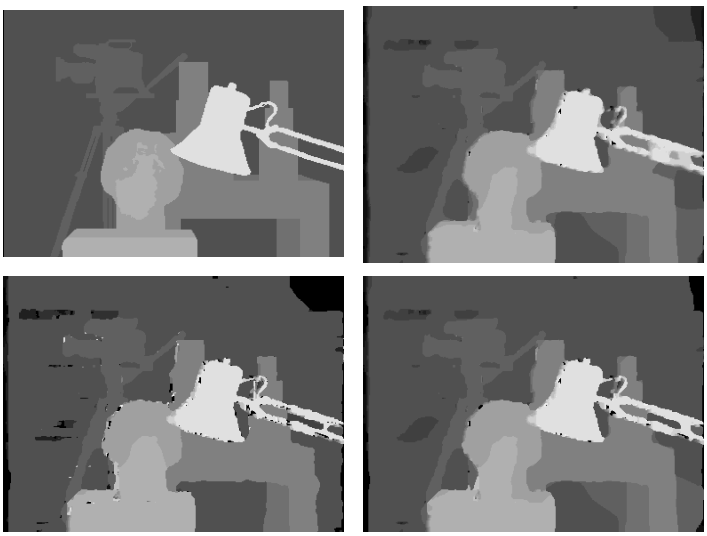

Figure 11: Stereo reconstruction. Top left, disparity (ground truth) for the pair in Fig. 10 – Top right, reconstructed with TV regularization as in [37] – Bottom left: with "Potts" energy – Bottom right: with Truncated TV

Let us eventually show an example of stereo reconstruction, as described in [37]. We have implemented with 17 labels ($k = 16$) a disparity reconstruction algorithm, where the data term is made of absolute differences of the color

channels between the left and right image, with a linear interpolation as described in [7]. We show here the reconstruction with three different interaction potentials: the total variation as in [37], a "Potts" energy (the length of the discontinuity), and the "truncated TV" (7). Results are shown on Figure 11. The best, and more precise results are obtained with the truncated TV. More precisely, we can count the number of "wrong" pixels where the error in disparity is more than one: in this case, we have found 2.55% of wrong pixels in the result with the TV interaction, 4.22% with the Potts model, and 1.88% with the truncated TV.

A An alternate approach

A much more natural approach to the optimal partition problem (19), presented in [44], is the following: nonnegative labels $v_0, \dots, v_k \in BV(\Omega; [0,1])$ are introduced, with a constraint $\sum_{i=0}^k v_i = 1$. Then, the problem is simply relaxed as

$$\min_{\mathbf{v}} \frac{1}{2} \sum_{i=0}^{k} \int_\Omega |Dv_i| \, + \, \sum_{i=0}^{k} \int_\Omega g_i(x) v_i(x) \, dx \, . \tag{59}$$

Clearly, if $v_i \in \{0,1\}$ a.e., this is just another formulation of the problem. Now, we claim that this relaxation is below ours, and that it will therefore fail more often. Let us also mention that, instead of the 1-norm of the vector $(Dv_1, \dots, Dv_k)$ used in (59), a 2-norm could be used as well (that is, a "standard" classical TV, this time normalized with $1/\sqrt{2}$), see for instance [30]. Since the results we computed with this approach are even worse than with the 1-norm (see Figure 12), we limited our discussion to the study of (59).

The relationship between this representation using the labels $(v_i)_{i=0}^k$, and ours is very simple. The dual formulation of the energy in (59) is

$$\sup \Big\{ - \int_\Omega \sum_{i=0}^{k} v_i(x) \mathrm{div}\, \eta_i(x) \, dx \, + \, \sum_{i=0}^{k} \int_\Omega g_i(x) v_i(x) \, dx \, : \\ \eta_i \in C_c^\infty(\Omega; \mathbb{R}^d), \|\eta_i\| \leq 1/2 \;\; \forall x \in \Omega, i = 0, \dots, k \Big\}$$

By making the change of variable $v_0 = 1 - u_1$, $v_i = u_i - u_{i+1}$ for $i = 1, \dots, k-1$, $v_k = u_k$ (that is, since $\sum_j v_j = 1$, $u_1 = 1 - v_0 = \sum_{j=1}^k v_j$, $u_i = \sum_{j=i}^k v_k$ for $i = 1, \dots, k$), it becomes

$$\sup \Big\{ - \int_\Omega \sum_{i=1}^{k} u_i(x) \mathrm{div}\, (\eta_i(x) - \eta_{i-1}(x)) \, dx + \sum_{i=1}^{k} \int_\Omega (g_i(x) - g_{i-1}(x)) u_i(x) \, dx \, : \\ \eta_i \in C_c^\infty(\Omega; \mathbb{R}^d), \|\eta_i\| \leq 1/2 \;\; \forall x \in \Omega, i = 0, \dots, k \Big\} . \tag{60}$$

We see that minimizing (60) with the constraints $0 \leq u_k \leq u_{k-1} \leq \cdot \leq u_1 \leq 1$ is equivalent to solving (52) (with $g_i = W(\cdot, i)$), with the interaction potential J replaced with

$$I(\mathbf{u}) \;=\; \int_\Omega \Phi(D\mathbf{u})$$

where $\Phi(\mathbf{p}) \;=\; \sup_{\mathbf{q}\in L} \mathbf{q}\cdot\mathbf{p}$ is the support function of the convex set L:

$$L \;=\; \Big\{ \mathbf{q} = (q_1, \ldots, q_k) \in \mathbb{R}^{d\times k} \;: \\ \exists (\eta_i)_{i=0}^k \in (\mathbb{R}^d)^{k+1}\,, |\eta_i| \leq 1/2\,, q_i = \eta_i - \eta_{i-1} \;\forall\; i = 1, \ldots, k \Big\}$$

We clearly have $|\sum_{i_1\leq i\leq i_2} q_i| = |\eta_{i_2} - \eta_{i_1-1}| \leq 1$ so that $L \subset K$, where K is defined in (51) (with $\sigma = 1$), so that $I \leq J$. In dimension $d = 1$, $L = K$: indeed, in this case, given any $\mathbf{q} \in K$, we let $\hat\eta_0 = 0$ and $\hat\eta_i = \sum_{j=1}^i q_j$. Then, we have

$$|\hat\eta_{i_2} - \hat\eta_{i_1-1}| \;=\; \left| \sum_{i_1\leq i\leq i2} q_i \right| \;\leq\; 1$$

for any $1 \leq i_1 \leq i_2 \leq k$. Hence, letting $M = \max_{0\leq i\leq k} \hat\eta_i$ and $m = \min_{0\leq i\leq k} \hat\eta_i$, we deduce $M - m \leq 1$. Letting now $\eta_i = \hat\eta_i - (M-m)/2$, we check easily that $|\eta_i| \leq 1/2$ for each i, and $q_i = \eta_i - \eta_{i-1}$, so that $\mathbf{q} \in L$. Hence: in dimension one, this approach and ours boil down to the same convexification of the problem. Let us mention that in this case, it can be shown (see [16] for a similar problem and a proof) that it produces the actual convex envelope of the initial k-labels problem (which is easily solved using dynamic programming anyway).

If $d \geq 2$, this is not true anymore. A simple counterexample is as follows, in dimension 2: let $\mathbf{q}$ be given by $q_1 = (1/2, \sqrt{3}/2)$, $q_2 = (1/2, -\sqrt{3}/2)$, $q_i = 0$ if $i \geq 3$. Then, $q_1 + q_2 = (1, 0)$ and we check that $\mathbf{q} \in K$. On the other hand, if there existed $(\eta_i)_{i=0}^k \in B(0, 1/2)^{k+1}$ with $q_i = \eta_i - \eta_{i-1}$, we should have $|\eta_0| \leq 1/2$ and $|\eta_1| = |\eta_0 + q_1| \leq 1/2$, which is possible only if $\eta_0 = -q_1/2$. But, then, $|\eta_2| = |\eta_0 + q_1 + q_2| = |q_2 + q_1/2| = \sqrt{3}/2 > 1/2$. Hence $(\eta_i)_{i=1}^k \notin L$. In fact, this explains why the triple point in Figure 12 is not well recovered by this technique, while it is with our approach (the input image and fields are described in Section 3.5, see Figure 4 and comments). Indeed, it is shown that the solution in the figure has an energy which is below the energy of a "binary" solution with three regions and sharp interfaces. More precisely, there holds the following result:

Proposition A.1. *Let Ω, $\bar E_i$, $i = 0, 1, 2$ be as in Proposition 3.7. Define also $\bar{\mathbf{u}}$ in the same way. Then $\bar{\mathbf{u}}$ is* not *a minimizer of $I(\mathbf{u}) = \int_\Omega \Phi(D\mathbf{u})$ with prescribed boundary condition $\mathbf{u} = \bar{\mathbf{u}}$ on $\partial\Omega$.*

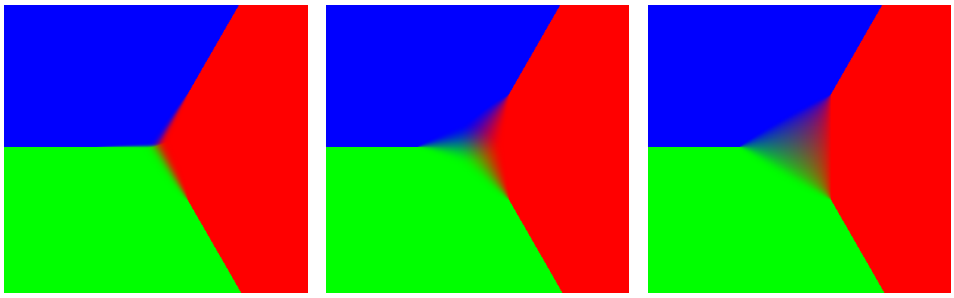

Figure 12: Left: our approach – Middle: with energy (59) – Right: using the vectorial TV as in [30]

In other words, letting for each i, $\bar{v}_i = \chi_{\bar{E}_i}$, we can find v_i such that $\sum_{i=0}^2 v_i = 1$ a.e., $v_i = \bar{v}_i$ on $\partial\Omega$ and $(1/2)\sum_{i=0}^2 \int_\Omega |Dv_i| < (1/2)\sum_{i=0}^2 \int_\Omega |D\bar{v}_i|$. The proof is postponed to the end of this section.

On the other hand, the interfacial energy which is actually chosen in [44] is anisotropic and, in fact, a sum of 1D energies in each direction. In this case, one can show again that their approach is equivalent to (the anisotropic version of) ours. We claim that the correct isotropic generalization of the approach in [44] is what we presented in this paper, or, in terms of the variable $\mathbf{v}$, what we propose below.

In fact, we see that the "correct" constraint on the vector $(\eta_i)_{i=0}^k$, should be that not on its norm but not its *oscillation*: the vector $\mathbf{q}$ given by $q_i = \eta_i - \eta_{i-1}$ will belong to K if and only if $|\eta_{i_2} - \eta_{i_1-1}| \le 1$ for each i_1, i_2 with $1 \le i_1 \le i_2 \le k$. If we reintroduce the interfacial energy σ as in (51), we get a constraint $|\eta_i - \eta_j| \le \sigma(|i-j|)$ for all $i, j \in \{0, \dots, k\}$. It is then easy to show that letting, for $\mathbf{v} \in BV(\Omega; \mathbb{R}^{k+1})$,

$$\hat{J}(\mathbf{v}) \;=\; \int_\Omega \hat{\Psi}(D\mathbf{v})$$

where now, for $\mathbf{p} \in \mathbb{R}^{d\times(k+1)}$, $\hat{\Psi}(\mathbf{p}) \;=\; \sup_{\mathbf{q}\in\hat{K}} \mathbf{q}\cdot\mathbf{p}$ is the support function of

$$\hat{K} = \Big\{ \mathbf{q} = (q_0, \dots, q_k) \in \mathbb{R}^{d\times(k+1)} : |q_i - q_j| \le \sigma(|i-j|) \;\forall\; i, j = 0, \dots, k \Big\} \quad (61)$$

then, problem (52) is equivalent (up to the change of variable $u_i = \sum_{j=i}^k v_j$) to

$$\min_{\mathbf{v}} \hat{J}(\mathbf{v}) \;+\; \sum_{i=0}^k \int_\Omega W(x,i) v_i(x)\, dx\,, \quad (62)$$

with the constraint $0 \le v_i$ for each i, and $\sum_{i=0}^k v_k = 1$. Moreover, the first constraint can be relaxed, as in (53), provided the problem is replaced with

$$\min_{\sum_{i=0}^k v_k = 1} \hat{J}(\mathbf{v}) \;+\; \sum_{i=0}^k \int_\Omega W(x,i)|v_i(x)|\, dx\,. \quad (63)$$

This alternate formulation can also be tackled numerically: it seems to be about as complex as (53), but some details in the implementation might be a bit simpler. Of course, it should produce essentially the same results.

Our choice to write the problem in terms of the variable $\mathbf{u}$ rather than $\mathbf{v}$ ($\mathbf{u}$ is some kind of "primitive" of $\mathbf{v}$) comes from related convexification approaches for more general variational problems, that have been studied in the past decade to address other problems in the calculus of variations. We mention this in the next Appendix B. Before, let us prove Proposition A.1.

Proof of Proposition A.1. To simplify the notation, in this proof, let the sets $\bar{E}_i$ be defined in all $\mathbb{R}^2$ (and $\bar{v}_i = \chi_{\bar{E}_i \cap \Omega}$ for each i).

Step 1 First, let us assume that Ω is a convex set and let us not worry about the boundary condition. We choose a rotationally symmetric, smooth mollifier $\rho \in C_c^\infty(B(0,1); \mathbb{R}_+)$, with $\int_{B(0,1)} \rho \, dx = 1$. As usual, for $\varepsilon > 0$, we let $\rho_\varepsilon(x) = \rho(x/\varepsilon)/\varepsilon^2$. We then define, for $i = 0, 1, 2$, a smooth function v_i^ε as the restriction to Ω of $\rho_\varepsilon * \chi_{\bar{E}_i}$. By linearity, clearly, we still have $\sum_{i=0}^2 v_i^\varepsilon = 1$. By the co-area formula (5), the variation $\int_\Omega |Dv_i^\varepsilon|$ is the average of the lengths of the level lines $\{v_i^\varepsilon = s\}$ for $s \in (0,1)$.

Let us choose a coordinate system (y_1, y_2) such that $\bar{E}_0 = \{y_2 < -|y_1|/\sqrt{3}\}$, and the jump set $\partial \bar{E}_0$ is the graph $y_2 = w(y_1) := -|y_1|/\sqrt{3}$. Assume for a while, to simplify, that Ω, near this graph, coincides with the set $|y_1| < 2$. First, assume also $\varepsilon = 1$ and consider the function $v_0^1 = \rho * \chi_{\bar{E}_0}$. It is nondecreasing with respect to y_2, and even strictly nondecreasing at points (y_1, y_2) such that the support of $\rho(\cdot - (y_1, y_2))$ meets $\partial \bar{E}_0$, that is, as soon as $0 < v_0^1 < 1$. Invoking the implicit functions theorem, we see that for any $0 < s < 1$, the level line $\{v_0^1 = s\}$ is the graph of a smooth, even function w_s, defined by $v_0^1(y_1, w_s(y_1)) = s$. In particular, $w_s'(y_1) = -(\partial_1 v_0^1)/(\partial_2 v_0^1)(y_1, w_s(y_1))$.

Outside of $B(0,1)$ (hence, in particular, for $|y_1| > 1$), we clearly have $w_s = w + c_s$ for a constant $c_s \in (-2/\sqrt{3}, 2/\sqrt{3})$, and with $c_s = -c_{1-s}$ for each $s \in (0,1)$. For each s, hence, $w_s'(y_1) = w'(y_1) = \pm 1/\sqrt{3}$ if $|y_1| > 1$. We check that $1/\sqrt{3}$ is an upper bound for the derivative of w_s': indeed, since any translate of $\bar{E}_0$ in a direction $(t, -1)$ with $|t| \le \sqrt{3}$ is included in $\bar{E}_0$, we also have $t\partial_1 v_0^1 - \partial_2 v_0^1 \ge 0$: hence $tw_s' + 1 \ge 0$ for all t with $|t| \le \sqrt{3}$, that is, $|w_s'| \le 1/\sqrt{3} = |w'|$. Since, moreover, $w_s'(0) = 0$ (w_s is even), we deduce

$$\mathcal{H}^1(\{v_0^1 = s\}) \;=\; \int_{-2}^{2} \sqrt{1 + |w_s'|^2}\, dy_1 \;<\; 4 \times \frac{2}{\sqrt{3}} \;=\; \mathcal{H}^1(\partial \bar{E}_0 \cap \Omega)\,.$$

Hence, from the co-area formula we deduce $\int_\Omega |Dv_0^1| < \mathcal{H}^1(\partial \bar{E}_0 \cap \Omega)$. A simple scaling argument will then show that there exists $c > 0$ such that

$\int_\Omega |Dv_0^\varepsilon| \le \mathcal{H}^1(\partial\bar{E}_0 \cap \Omega) - c\varepsilon$ for $\varepsilon \le 1$. Now, if the boundary of $\partial\Omega$ is made of two straight lines (not necessarily vertical) in the neighborhood of its intersection with $\partial\bar{E}_0$, the same is still true, because the possible increase or loss of length of $\{v_0^\varepsilon = s\}$ at the boundary (with respect to the previous case) is compensated exactly by the loss or increase of the symmetric line $\{v_0^\varepsilon = 1-s\}$, so that the average length remains the same as when $\partial\Omega$ is vertical. If Ω is a generic convex set, then the length of the lines $\{v_0^\varepsilon = s\}$ in Ω are even shorter than if $\partial\Omega$ were replaced with two tangent straight lines at its intersection with $\partial\bar{E}_0$, so that the variation of v_0^ε is even lower.

Step 2 Now, we consider any open set Ω and $\eta > 0$ such that $\overline{B(0,2\eta)} \subset \Omega$. We choose $\varepsilon << \eta$ and let as above, for each $i \in \{0,1,2\}$, $v_i^\varepsilon = \rho_\varepsilon * \chi_{\bar{E}_i}$ in $B(0,\eta)$, while $v_i^\varepsilon = \bar{v}_i = \chi_{\bar{E}_i}$ in $\Omega \setminus B(0,2\eta)$. In $B(0,2\eta) \setminus B(0,\eta)$, we build v_i^ε by joining with straight lines the level lines of v_i^ε inside $B(0,\eta)$ and outside $B(0,2\eta)$: one can check that the total variation of each v_i^ε in the crown is then of order at most $2\eta + \varepsilon^2/\eta$ as $\varepsilon \to 0$, while, using the first step, it is less than $2\eta - c\varepsilon$ inside the ball $B(0,\eta)$. Hence, if ε is small enough, we find that $\int_\Omega |Dv_i^\varepsilon| < \int_\Omega |D\bar{v}_i|$, and this shows the proposition. □

B Related representations in the calculus of variations

Let us explain some connections between the representation we have introduced in this paper, and recent similar approaches in the calculus of variations. In fact, our formulation belongs to a wide class of convexifications for possibly non-convex scalar minimization problems, introduced in the past ten or twenty years. The very general idea is as follows: assume you want to minimize a function $F : L^1(\Omega) \to [0,+\infty]$ which depends on a scalar function u. Then, letting, for $(x,t) \in \Omega \times \mathbb{R}$, $\mathbf{1}_u(x,t) = 1$ if $u(x) > t$ and 0 else, that is, $\mathbf{1}_u$ is the characteristic function of the subgraph of u, you can define an energy on measurable functions $v : \Omega \times \mathbb{R} \to \mathbb{R}$ as

$$\mathcal{F}(v) \;=\; \begin{cases} F(u) & \text{if } v = \mathbf{1}_u \text{ for some } u \in L^1(\Omega)\,, \\ +\infty & \text{else.} \end{cases}$$

Then, the idea is to introduce the convex envelope $\bar{\mathcal{F}}$ of $\mathcal{F}$, defined on the functions $v : \Omega \times \mathbb{R} \to [0,1]$ which are non-increasing in the last variable (and still having value $+\infty$ elsewhere).

It is shown for instance in [16] that, provided F is l.s.c. in L^1, and under a few technical assumptions, $\bar{\mathcal{F}}$ is equal to $\mathcal{F}$ on subgraphs. Hence, the problem

of minimizing F is essentially equivalent to the problem of minimizing a convex functional.

However, this is useless if one does not have an explicit (and relatively simple) expression of $\bar{\mathcal{F}}$. In some cases, such as (1), such an expression exists and is relatively easy to produce. A general framework is introduced in [25, 10, 1], as follows: consider

$$F(u) \;=\; \int_\Omega f(x, u(x), \nabla u(x))\, dx$$

with a Lagrangian $f(x,t,p)$ convex in the last variable $p \in \mathbb{R}^d$ (and with some coercivity). Then, let us define a convex K of vector fields $\phi = (\phi^x, \phi^t)$ defined on $\Omega \times \mathbb{R}$, satisfying pointwise $\phi^t(x,t) \geq f^*(x,t,\phi^x(x,t))$ (f^* denotes here the Legendre-Fenchel conjugate of f with respect to p, see [21, 39]). One can show that:

$$F(u) \;=\; \sup_{\phi \in K} \int_{\Omega\times\mathbb{R}} \phi \cdot D\mathbf{1}_u. \tag{64}$$

The reason for (64) is the following: the integrand in (64) represents the flux of ϕ across the graph of u. After a change of variable, it can be written

$$\int_{\Omega\times\mathbb{R}} \phi \cdot D\mathbf{1}_u \;=\; \int_\Omega \phi^x(x,u(x)) \cdot \nabla u(x) \;-\; \phi^t(x,u(x))\, dx\,.$$

Then, when the supremum is taken, the best choice for $\phi^t(x,u(x))$ is of course the minimal value $f^*(x,u(x),\phi^x(x,t))$, but then the integrand becomes

$$\sup_{\phi^x} \phi^x \cdot \nabla u(x) - f^*(x,u(x),\phi^x) \;=\; f(x,u(x),\nabla u(x))\,.$$

Now, it turns out that if one replaces in (64) the function $\mathbf{1}_u$ with a generic function $v \in BV(\Omega\times\mathbb{R})$, with $0 \leq v \leq 1$, one obtains a function

$$\bar{\mathbf{F}}(v) \;=\; \sup_{\phi\in K} \int_{\Omega\times\mathbb{R}} \phi \cdot Dv \tag{65}$$

which is nothing but the convex envelope $\bar{\mathcal{F}}$ introduced above. This can be checked by showing that

- it coincides with $\mathcal{F}$ when v is binary (i.e., takes only the values 0 and 1): this follows from (64);
- it satisfies a "generalized co-area formula:"

$$\bar{\mathbf{F}}(v) \;=\; \int_0^1 \bar{\mathbf{F}}(\chi_{\{v>s\}})\, ds \;=\; \int_0^1 \mathcal{F}(\chi_{\{v>s\}})\, ds$$

  which shows that it is larger than any convex function below $\mathcal{F}$.

It is interesting to point out that F needs not be convex in the variable u. In the discrete setting, this approach corresponds exactly to the representations which have been introduced by Ishikawa and Geiger [29, 28] about ten years ago, and widely used in image processing (in particular in stereo reconstruction) by many authors after these seminal papers. We refer also to [20] for a general study, in the discrete setting, of the kind of energies which can be minimized in this manner.

The convex set K above is defined by a local constraint: then, it was observed by Alberti, Bouchitté and Dal Maso [1] that with a nonlocal constraint, as we propose here in (11), more general (and more "nonconvex") functionals could be studied in the same way. They convex set K of vector fields (ϕ^x, ϕ^t) that they introduce is defined by the relationships:

$$\begin{array}{lll} \textit{(i)} & \phi^t(x,t) \geq \frac{1}{4}\phi^x(x,t)^2 & \text{for any } (x,t) \in \Omega \times \mathbb{R} \\ \textit{(ii)} & \left| \int_{t_1}^{t_2} \phi^x(x,s)\, ds \right| \leq 1 & \text{for any } x \in \Omega \text{ and } t_1 < t_2\,. \end{array} \tag{66}$$

Condition *(i)* corresponds to the previous condition $\phi^t \geq f^*(x,t,\phi^x)$, in the particular case $f(x,t,p) = p^2$ (hence $f^*(x,t,q) = q^2/4$). To understand the second condition, assume u is, away of a jump set J_u, a smooth function. If $x \in J_u$, we assume u jumps from a lower value $u_-(x)$ to a higher value $u_+(x)$. In particular, u is in the space "$SBV(\Omega)$" of special functions with bounded variation (which are BV functions such that the Cantor part Cu in (3) vanishes, [2]), and its distributional derivative is the measure

$$Du \;=\; \nabla u(x)\, dx \;+\; (u_+(x) - u_-(x))\nu_u(x)\mathcal{H}^{d-1} \llcorner J_u\,,$$

where $\nu_u(x)$ is the normal vector to J_u at x, pointing towards the value u_+. Then, the boundary of the subgraph of u, $\partial\{t < u(x)\}$, has a "horizontal part" $(x, u(x))$, $x \notin J_u$, with (inner) normal vector $\nu = (\nabla u, -1)/\sqrt{1+|\nabla u|^2}$, and a "vertical part" $\cup_{x\in J_u}\{x\} \times [u_-(x), u_+(x)]$, with normal vector ν_u. Then, we have

$$\begin{aligned} \int_{\Omega\times\mathbb{R}} \phi \cdot D\mathbf{1}_u \;=\; & \int_\Omega \phi^x(x,u(x)) \cdot \nabla u(x) - \phi^t(x,u(x))\, dx \\ & + \int_{J_u} \left(\int_{u_-(x)}^{u_+(x)} \phi^x(x,s) \cdot \nu_u(x)\, ds \right) d\mathcal{H}^{d-1}(x) \end{aligned}$$

and we check easily that the supremum of this expression for $\phi \in K$ is the celebrated Mumford-Shah functional

$$F(u) \;=\; \int_\Omega |\nabla u(x)|^2\, dx \;+\; \mathcal{H}^{d-1}(J_u)$$

The idea of by Alberti, Bouchitté and Dal Maso in [1] was to provide, through this framework, necessary conditions of optimality for minimizers of the Mumford-Shah functional. The idea is simple: if you can find a free divergence vector field ϕ such that

$$F(u) \;=\; \int_\Omega |\nabla u(x)|^2\,dx \;+\; \mathcal{H}^{d-1}(J_u) \;=\; \int_{\Omega\times\mathbb{R}} \phi\cdot D\mathbf{1}_u\,,$$

then for any u' with $u'=u$ on $\partial\Omega$, one must have $F(u')\geq F(u)$: indeed, one always have

$$F(u') \;\geq\; \int_{\Omega\times\mathbb{R}} \phi\cdot D\mathbf{1}_{u'} \;=\; \int_{\Omega\times\mathbb{R}} \phi\cdot D\mathbf{1}_u \;-\; \int_{\Omega\times\mathbb{R}} (\mathbf{1}_{u'}-\mathbf{1}_u)\mathrm{div}\,\phi\,dxdt \;=\; F(u).$$

This approached proved very useful in a number of cases [33, 32, 34, 19]. It is still unknown, though, whether for any minimizer of the Mumford Shah functional, there exists such a field ϕ (called a "calibration"). Such existence is related to the minimality of $\mathbf{1}_u$ for the convex functional $\bar{\mathbf{F}}$ defined in (65). What remains unknown, in this case, is whether $\bar{\mathbf{F}}=\bar{\mathcal{F}}$, the convex envelope of $\mathcal{F}$ (in which case it would have the same minimizers as F). A positive answer would for instance (up to regularity issues) imply the existence of a "calibration," for showing the minimality of the "crack-tip," and provide a much shorter proof than Bonnet and David's [9], but the experiment illustrated in Figure 6, although on a very simplified variant, seems to show that the answer is rather negative.

C Algorithm for projecting onto K

Let us explain here how we implemented, in our practical computation, the projection of a vector $(\xi_i)_{i=1}^k$ in $\mathbb{R}^{d\times k}$ onto the convex K. This convex is described as a finite intersection of simple sets, on which a projection is easily computed. In this case, a converging algorithm was first proposed by Dykstra in the 1980's [14].

C.1 Dykstra's algorithm

The idea is essentially as follows: we are given k convex sets in $\mathbb{R}^N$, $K_1,\ldots,K_k$, each on which has a simple algorithm for projecting onto it, and you want to compute the projection of a vector x onto $\bigcap_{i=1}^k K_i$. Letting $\psi_i(z)=0$ if $z\in K_i$, and $+\infty$ else, it means you want to solve

$$\min_{z\in\mathbb{R}^N} \frac{|x-z|}{2}^2 \;+\; \sum_{i=1}^k \psi_i(z)\,. \tag{67}$$

Then, a good idea is to consider the dual problem, which in this case can be written [21, 39]

$$\min_{y\in\mathbb{R}^N} \frac{|x-y|^2}{2} \;+\; \left(\sum_{i=1}^k \psi_i\right)^*(y)\,, \tag{68}$$

moreover, $\bar{z}$ solves (67) if and only if $\bar{y} = x - \bar{z}$ solves (68). Here, $(\sum_{i=1}^k \psi_i)^*$ denotes the Legendre-Fenchel conjugate [21, 39] of $\sum_{i=1}^k \psi_i$, that is,

$$\left(\sum_{i=1}^k \psi_i\right)^*(y) \;=\; \sup_{z\in\mathbb{R}^N} y\cdot z \;-\; \sum_{i=1}^k \psi_i(z)$$

Now, it is known (and easy to check) that the Legendre-Fenchel conjugate of a sum of convex, lower semicontinuous functions is given by the inf-convolution of the Legendre-Fenchel conjugate of each function. In this case, that means:

$$\left(\sum_{i=1}^k \psi_i\right)^*(y) \;=\; \inf\left\{\sum_{i=1}^k \psi_i^*(y_i) \;:\; \sum_{i=1}^k y_i \;=\; y\right\}, \tag{69}$$

so that another way to write (68) is:

$$\inf_{(y_i)_{i=1}^k\in(\mathbb{R}^N)^k} \frac{1}{2}\Big|x - \textstyle\sum_{i=1}^k y_i\Big|^2 \;+\; \displaystyle\sum_{i=1}^k \psi_i^*(y_i)\,, \tag{70}$$

The idea behind Dykstra's algorithm is to minimize now (70) alternatively in each variable (y_i), for $i = 1,\dots,k$, until convergence (this point of view was is found for instance in Gaffke and Mathar [24]).

Hence, we start with $y_i^0 = 0$ for each i, and $(y_i^{n+1})_{i=1}^k$ is found from $(y_i^n)_{i=1}^k$ by letting, for $i = 1,\dots,k$:

$$y_i^{n+1} \;=\; \arg\min_{y\in\mathbb{R}^N} \frac{1}{2}\Big|\Big(x - \textstyle\sum_{j<i} y_j^{n+1} - \sum_{j>i} y_j^n\Big) - y\Big|^2 \;+\; \psi_i^*(y)\,. \tag{71}$$

Now, let us set for each $n \geq 1$, $y^n = \sum_{i=1}^k y_i^n$, $x^n = x - y^n$ (and $x^0 = x$). We also set $x_0^n = x^{n-1}$, $x_i^n = x - (\sum_{j\leq i} y_j^n + \sum_{j>i} y_j^{n-1})$ (in particular $x_k^n = x^n = x_0^{n+1}$).

The primal version of (71) is, for $n \geq 1$ and $i = 1,\dots,k$:

$$x_i^{n+1} \;=\; \arg\min_{z\in\mathbb{R}^N} \frac{1}{2}\left|x_{i-1}^n + y_i^n - z\right|^2 \;+\; \psi_i(z) \tag{72}$$

while y_i^{n+1} is then given by $y_i^{n+1} = x_{i-1}^n + y_i^n - x_i^{n+1}$. But, as ψ_i is defined, it means nothing else than

$$\left\{\begin{array}{rcl} x_i^{n+1} & = & \Pi_{K_i}(x_{i-1}^n + y_i^n) \\ y_i^{n+1} & = & x_{i-1}^n + y_i^n - x_i^{n+1} \end{array}\right. \tag{73}$$

This is the main iteration of Dykstra's algorithm, as described in the original work. Moreover, it is shown there that as $n\to\infty$, $x^n \to \Pi_K(x)$ (strongly, if we replace $\mathbb{R}^N$ with a Hilbert space).

C.2 The projection onto K

For solving problem (53), we need to project onto the convex set $K \subset \mathbb{R}^{d\times k}$ defined in (51), which may be written as the intersection of convex sets

$$\bigcap_{1\leq i_1\leq i_2\leq k} K_{i_1,i_2}, \text{ with } K_{i_1,i_2} \;=\; \left\{(q_i)_{i=1}^k \in \mathbb{R}^{d\times k} \;:\; \left|\sum_{i=i_1}^{i_2} q_i\right| \leq \sigma(i_2-i_1+1))\right\}.$$

Given $\mathbf{q} = (q_i)_{i=1}^k \in \mathbb{R}^{d\times k}$, the projection onto K_{i_1,i_2} is given by

$$(\Pi_{K_{i_1,i_2}}(\mathbf{q}))_i \;=\; \begin{cases} q_i & \text{if } i < i_1 \text{ or } i > i_2\,; \\ q_i - q^{i_1,i_2} & \text{if } i_1 \leq i \leq i_2 \end{cases}$$

where, letting $\hat{q} = \sum_{i=i_1}^{i_2} q_i$, the vector $q^{i_1,i_2} \in \mathbb{R}^d$ is given by $q^{i_1,i_2} = (|\hat{q}| - \sigma(i_2 - i_1 + 1))^+/(i_2 - i_1 + 1)(\hat{q}/|\hat{q}|)$. Dykstra's algorithm is now simple to implement: we are given a vector $\mathbf{q}$ we want to project onto K. We fist let $q^{i_1,i_2} = 0$, and recursively update as follows:

1. For $1 \leq i_1 \leq i_2 \leq k$:
   (a) Compute $\hat{q} = \sum_{i=i_1}^{i_2}(q_i + q^{i_1,i_2})$;
   (b) Let then $\hat{\hat{q}} = (|\hat{q}| - \sigma(i_2 - i_1 + 1))^+/(i_2 - i_1 + 1)(\hat{q}/|\hat{q}|)$;
   (c) For $i_1 \leq i \leq i_2$, replace q_i with $q_i + q^{i_1,i_2} - \hat{\hat{q}}$;
   (d) Replace q^{i_1,i_2} with $\hat{\hat{q}}$
2. Test how much $\mathbf{q} = (q_i)_{i=1}^k$ has changed during the loop: if change is less than some tolerance, end. Else: go back to step 1.

D Algorithm for projecting onto C

Let $\mathbf{u} \in \mathbb{R}^k$: we give here a simple algorithm for computing the projection $\Pi_C\mathbf{u}$ of $\mathbf{u}$ onto $C = \{\mathbf{z} \in R^k \,:\, z_1 \geq z_2 \geq \cdots z_k\}$. Of course, as in the previous section, we could still use Dykstra's algorithm: however, this will usually converge in infinitely many iterations (although the error decreases very fast), whereas the alternative solution we propose here needs at most $k-1$ iterates to produce the exact solution.

Let us first modify slightly the problem, and consider the minimization:

$$\min_{\mathbf{z}\in C} \sum_{i=1}^k \omega_i (z_i - u_i)^2 \tag{74}$$

where ω_i are positive weight. Then, the same analysis as in section 4.2 holds true, and in particular (54): we hence know that if for some index $i \in \{1,\ldots,k-$

1}, $u_i \leq u_{i-1}$, the solution $\mathbf{z}$ of (74) satisfies $z_i = z_{i+1}$ (while if there is no such index, $\mathbf{u} \in C$ and $\mathbf{z} = \mathbf{u}$ is the solution).

In this case, (74) is obviously equivalent to the minimization of

$$\begin{aligned}\sum_{j\neq i,i+1} \omega_j (z_j - u_j)^2 \;+\; \big(\omega_i (z_i - u_i)^2 + \omega_{i+1}(z_i - u_{i+1})^2\big) \\ = \sum_{j\neq i,i+1} \omega_j (z_j - u_j)^2 \;+\; (\omega_i + \omega_{i+1}) \left(z_i - \frac{\omega_i u_i + \omega_{i+1} u_{i+1}}{\omega_i + \omega_{i+1}} \right)^2 \\ +\; \omega_i u_i^2 + \omega_{i+1} u_{i+1}^2 - \frac{(\omega_i u_i + \omega_{i+1} u_{i+1})^2}{\omega_i + \omega_{i+1}} \qquad (75)\end{aligned}$$

over all $\mathbf{z} \in C$ with $z_i = z_{i+1}$, which (since the last line in (75) does not depend on $\mathbf{z}$) is a new problem of the form (74), but now in dimension $(k-1)$, and with the coordinates u_i, u_{i+1} replaced with their average with respective weights ω_i and ω_{i+1}, and their common weight in the new distance replaced with the sum of the weights $\omega_i + \omega_{i+1}$.

Hence, a straightforward recursive algorithm for computing $\Pi_C u$ is as follows:

1. Let first $\mathbf{z} = \mathbf{u}$, and define k "clusters" $\mathcal{C}_i = \{i\}$, $i = 1, \ldots, k$, of only one element.

2. Identify two coordinates z_i, z_{i+1} with $z_i < z_{i+1}$. If there are no such coordinates, $\mathbf{z} \in C$: the procedure ends and $\mathbf{z} = \Pi_C \mathbf{u}$.

3. Replace the clusters $\mathcal{C}_j$, $j \in \mathcal{C}_i \cup \mathcal{C}_{i+1}$, with the union $\mathcal{C}_i \cup \mathcal{C}_{i+1}$. Replace then z_i and z_{i+1} with the average $\left(\sum_{j\in\mathcal{C}_i} z_j\right) / (\sharp\, \mathcal{C}_i)$.

4. go back to 2.

E Evaluation of Ψ for three labels

Let us quickly mention how the value of $\Psi(\mathbf{p})$, defined in (35), can be computed in order to estimate the primal-dual gap (44). This is probably standard. The formula is

$$\Psi(\mathbf{p}) \;=\; \min_{q\in\mathbb{R}^d} |q| \;+\; |p_1 - q| \;+\; |p_2 - q|\,,$$

that is, q is the point whose total distance to 0, p_1 and p_2 in $\mathbb{R}^2$ is minimal.

Hence, the problem boils down to finding a simple way to compute, given $a, b, c \in \mathbb{R}^2$, the point $x \in \mathbb{R}^2$ which minimizes $|x-a| + |x-b| + |x-c|$.

If a, b, c are aligned (in our case, if $\det(p_1, p_2) = 0$), then x is optimal when it is equal to a point which lies in between the two other (or is equal to one

of the other). For instance, if $b \in [a, c]$, $x = b$ is optimal and the solution is $|a - b| + |b - c|$.

If the points are not on the same line, then they form a non-degenerate triangle. It is shown, then, that if x is optimal, it must lie inside the triangle formed by the three points. Then, from the optimality conditions, one checks that either x is one of the points a, b, c (say, for instance, a), in which case the angle $\widehat{bac}$ must be greater than, or equal to 120°, or $x \notin \{a, b, c\}$, in which case the angles $\widehat{abc}$, $\widehat{bac}$, $\widehat{acb}$ are all three smaller than 120°. In this last non-degenerate case, the three angles $\widehat{axb}, \widehat{bxc}, \widehat{cxa}$ must be exactly 120°, so that x lies at the intersection of three circles, which are the circles made of the points which "see" two vertices of the triangle under an angle of 120°. This is very easy to compute. The centers of these circles are A, B, C: A is at distance $|bc|/(2\sqrt{3})$ of the segment $[b, c]$ and its projection on this segment is the middle point $(b + c)/2$ (and a and A must be separated by (bc)). B is the same with respect to a and c, and C with respect to a and b. Then, the circles of center A and radius $|bc|/\sqrt{3}$ and of center B and radius $|ac|/\sqrt{3}$ have the two intersection points c and x: x is then the symmetric of c with respect to the line (AB). See Fig. 13.

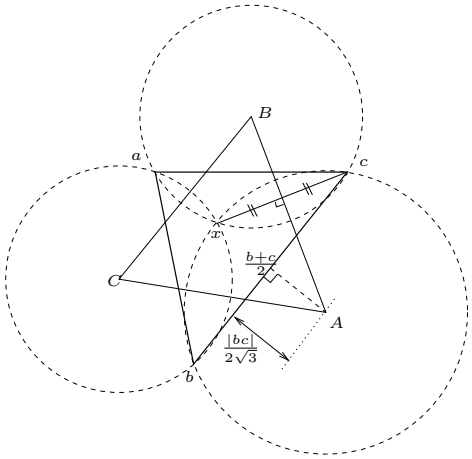

Figure 13: Construction of the optimal point x

References

[1] G. Alberti, G. Bouchitté, and G. Dal Maso. The calibration method for the Mumford-Shah functional and free-discontinuity problems. *Calc. Var. Partial Differential Equations*, 16(3):299–333, 2003.

[2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[3] B. Appleton and H. Talbot. Globally optimal geodesic active contours. *J. Math. Imaging Vision*, 23(1):67–86, 2005.

[4] K. J. Arrow, L. Hurwicz, and H. Uzawa. *Studies in linear and non-linear programming*. With contributions by H. B. Chenery, S. M. Johnson, S. Karlin, T. Marschak, R. M. Solow. Stanford Mathematical Studies in the Social Sciences, vol. II. Stanford University Press, Stanford, Calif., 1958.

[5] G. Bellettini. An almost everywhere regularity result for minimal partitions. *Boll. Unione Mat. Ital., VII. Ser., A*, 4(1):57–63, 1990.

[6] G. Bellettini, M. Paolini, and C. Verdi. Convex approximations of functionals with curvature. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 2(4):297–306, 1991.

[7] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):401–406, 1998.

[8] A. Bonnet. On the regularity of edges in image segmentation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 13(4):485–528, 1996.

[9] A. Bonnet and G. David. Cracktip is a global Mumford-Shah minimizer. *Astérisque*, (274):vi+259, 2001.

[10] G. Bouchitté. Recent convexity arguments in the calculus of variations. (lecture notes from the 3rd Int. Summer School on the Calculus of Variations, Pisa), 1998.

[11] G. Bouchitté and G. Dal Maso. Integral representation and relaxation of convex local functionals on BV(Ω). *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 20(4):483–533, 1993.

[12] G. Bouchitté and M. Valadier. Integral representation of convex functionals on a space of measures. *J. Funct. Anal.*, 80(2):398–420, 1988.

[13] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.

[14] J. P. Boyle and R. L. Dykstra. A method for finding projections onto the intersection of convex sets in Hilbert spaces. In *Advances in order restricted statistical inference (Iowa City, Iowa, 1985)*, volume 37 of *Lecture Notes in Statist.*, pages 28–47. Springer, Berlin, 1986.

[15] A. Braides. Γ-*convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[16] A. Chambolle. Convex representation for lower semicontinuous envelopes of functionals in L^1. *J. Convex Anal.*, 8(1):149–170, 2001.

[17] T. F. Chan and L. A. Vese. Active contour and segmentation models using geometric PDE's for medical imaging. In *Geometric methods in bio-medical image processing*, Math. Vis., pages 63–75. Springer, Berlin, 2002.

[18] G. Dal Maso. *An introduction to* Γ-*convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[19] G. Dal Maso, M. G. Mora, and M. Morini. Local calibrations for minimizers of the Mumford-Shah functional with rectilinear discontinuity sets. *J. Math. Pures Appl. (9)*, 79(2):141–162, 2000.

[20] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation. II. Levelable functions, convex priors and non-convex cases. *J. Math. Imaging Vision*, 26(3):277–291, 2006.

[21] I. Ekeland and R. Témam. *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.

[22] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. CRC Press, Boca Raton, FL, 1992.

[23] H. Federer. *Geometric measure theory*. Springer-Verlag New York Inc., New York, 1969.

[24] N. Gaffke and R. Mathar. A cyclic projection algorithm via duality. *Metrika*, 36(1):29–54, 1989.

[25] M. Giaquinta, G. Modica, and J. Souček. *Cartesian currents in the calculus of variations I. Cartesian currents.* Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, Berlin, 1998.

[26] E. Giusti. *Minimal surfaces and functions of bounded variation*, volume 80 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1984.

[27] P. L. Hammer, P. Hansen, and B. Simeone. Roof duality, complementation and persistency in quadratic 0-1 optimization. *Math. Programming*, 28(2):121–155, 1984.

[28] H. Ishikawa. Exact optimization for Markov random fields with convex prio rs. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(10):1333–1336, 2003.

[29] H. Ishikawa and D. Geiger. Segmentation by grouping junctions. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 125–131, 1998.

[30] J. Lellman, J. Kappes, J. Yuan, F. Becker, and C. Schnörr. Convex multiclass image labeling by simplex-constrained total variation. Technical report, IPA, HCI, Dept. Of Mathematics and Computer Science, Univ. Heidelberg, October 2008.

[31] U. Massari and I. Tamanini. Regularity properties of optimal segmentations. *J. Reine Angew. Math.*, 420:61–84, 1991.

[32] M. G. Mora. The calibration method for free-discontinuity problems on vector-valued maps. *J. Convex Anal.*, 9(1):1–29, 2002.

[33] M. G. Mora. Local calibrations for minimizers of the Mumford-Shah functional with a triple junction. *Commun. Contemp. Math.*, 4(2):297–326, 2002.

[34] M. G. Mora and M. Morini. Local calibrations for minimizers of the Mumford-Shah functional with a regular discontinuity set. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 18(4):403–436, 2001.

[35] J.-M. Morel and S. Solimini. Segmentation of images by variational methods: a constructive approach. *Rev. Mat. Univ. Complutense Madr.*, 1(1-3):169–182, 1988.

[36] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.*, 42(5):577–685, 1989.

[37] T. Pock, T. Schoenemann, G. Graber, H. Bischof, and D. Cremers. A convex formulation of continuous multi-label problems. In *European Conference on Computer Vision (ECCV)*, Marseille, France, October 2008.

[38] R. T. Rockafellar. Integrals which are convex functionals. II. *Pacific J. Math.*, 39:439–469, 1971.

[39] R. T. Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.

[40] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields. In *Computer Vision - ECCV 2006*, volume 3952 of *Lecture Notes in Computer Science*, pages 16–29, 2006.

[41] I. Tamanini and G. Congedo. Optimal segmentation of unbounded functions. *Rend. Semin. Mat. Univ. Padova*, 95:153–174, 1996.

[42] J. E. Taylor. The structure of singularities in soap-bubble-like and soap-film-like minimal surfaces. *Ann. of Math. (2)*, 103(3):489–539, 1976.

[43] A. I. Vol′pert. Spaces BV and quasilinear equations. *Mat. Sb. (N.S.)*, 73 (115):255–302, 1967.

[44] C. Zach, D. Gallup, J.-M Frahm, and M. Niethammer. Fast global labeling for real-time stereo using multiple plane sweeps. In *Vision, Modeling and Visualization Workshop VMV 2008*, October 2008. to appear.

[45] M. Zhu and T. Chan. An efficient primal-dual hybrid gradient algorithm for total variation image restoration. Technical Report 08-34, UCLA Cam Reports, 2008.

[46] W. P. Ziemer. *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.

Liste der erschienenen Berichte / Publication list

IAI-TR-91-1 Axioms for Concurrent Processes
Armin B. Cremers and Thomas N. Hibbard
Juni 1991.

IAI-TR-92-1 The Categorical Abstract Machine:
Basics and Enhancements
Ralf Hinze
Dezember 1992.

IAI-TR-92-2 On Perfect Introspection with Quantifying-in
Gerhard Lakemeyer
Dezember 1992.

IAI-TR-92-3 All You Ever Wanted to Know about Tweety
Gerhard Lakemeyer
Dezember 1992.

IAI-TR-93-1 Exploration and Model Building in Mobile
Robot Domains
Sebastian B. Thrun
Januar 1993.

IAI-TR-93-2 Best Approximate General Ellipses on Integer
Grids
Dieter W. Fellner and Christoph Helmberg
Januar 1993.

IAI-TR-93-3 Programming in Propositional Logic
or Reductions: Back to the Roots (Satisfiability)
Hermann Stamm-Wilbrandt
März 1993

IAI-TR-93-4 Extensible Image Synthesis
Dieter W. Fellner
April 1993.

IAI-TR-93-5 Extracting Provably Correct Rules
from Artificial Neural Networks
Sebastian B. Thrun
Mai 1993.

IAI-TR-93-6 Adaptive Steuerung von Strategieparametern
bei genetischen Algorithmen
Andreas Bergmann
Juni 1993.

IAI-TR-93-7 Lifelong Robot Learning
Sebastian B. Thrun and Tom M. Mitchell
Juli 1993.

IAI-TR-93-8 Situation Schemata
as Representation Format and Query Language
for a Non-Standard Database System
Armin B. Cremers, Thomas Lemke,
Volker Paul and Marion Schulz
August 1993.

IAI-TR-93-9 CGI++
A 2D Graphics Interface
Based on CGI
Dieter W. Fellner and Martin Fischer
August 1993.

IAI-TR-93-10 A simple linear time algorithm
for embedding maximal planar graphs
Hermann Stamm-Wilbrandt
Oktober 1993.

IAI-TR-94-1 Logic Program Modules for
Interoperable Information
Systems
Armin B. Cremers, Thomas Lemke and
Mechthild Rohen
Mai 1994.

IAI-TR-94-2 Planar embedding of hamiltonian graphs
via efficient bipartation of circle graphs
Christoph Hundack and
Hermann Stamm-Wilbrandt
Juni 1994.

IAI-TR-94-3 Implementation of Dynamic Delegation
in Strongly Typed Inheritance-Based Systems
Günter Kniesel
Oktober 1994.

IAI-TR-94-4 Update Propagation in Chimera,
an Active DOOD Language
Ulrike Griefahn and Rainer Manthey
Oktober 1994.

IAI-TR-94-5 Reflections on Some Fundamental Issues of
Rule-Based Incremental Update Propagation
Rainer Manthey
Oktober 1994.

IAI-TR-95-1 Modelle für die Photogrammetrische Gebäudekonstruktion
C. Braun, T.H. Kolbe, F. Lang, W. Schickler, V. Steinhage, A.B. Cremers, W. Förstner and L. Plümer
Januar 1995.

IAI-TR-95-2 The Mobile Robot Rhino
J. Buhmann, W. Burgard, A.B. Cremers, D. Fox, T. Hofmann, F. Schneider, J. Strikos and S. Thrun
Januar 1995.

IAI-TR-95-3 Levesque's Axiomatization of Only Knowing is Incomplete
Joseph Y. Halpern and Gerhard Lakemeyer
Januar 1995.

IAI-TR-95-4 First Specification of Phoenix, an Active Database Programming Language
Ulrike Griefahn
Januar 1995.

IAI-TR-95-5 Extended circle graphs I
Christoph Hundack and Hermann Stamm-Wilbrandt
März 1995.

IAI-TR-95-6 Knowledge-enhanced CO-monitoring in Coal Mines
Wolfram Burgard and Armin B. Cremers and Dieter Fox and Martin Heidelbach and Angelica M. Kappel and Stefan Lüttringhaus-Kappel
März 1995.

IAI-TR-95-7 Pairwise Data Clustering by Deterministic Annealing
Joachim M. Buhmann and Thomas Hofmann
Juni 1995.

IAI-TR-95-8 Penguin - A Portable Environment for a Graphical User Interface
Dieter W. Fellner and Martin Fischer
Juni 1995.

IAI-TR-95-9 A Portable Algorithm for Dynamic Color Management
Martin Fischer
Juni 1995.

IAI-TR-95-10 Distributed CGI^{++} - eine Client/Server-Lösung zur Entwicklung leistungsfähiger Computergraphik-Anwendungen
Dieter W. Fellner and Martin Fischer
Juli 1995.

IAI-TR-95-11 Runtime Type for Class Hierarchies
Martin Fischer
Juli 1995.

IAI-TR-95-12 Reference Pointer for Class Hierarchies
Martin Fischer
August 1995.

IAI-TR-95-13 The Dynamic Window Approach to Collision Avoidance
Dieter Fox, Wolfram Burgard and Sebastian Thrun
August 1995.

IAI-TR-96-1 Estimating the Absolute Position of a Mobil Robot Using Position Probability Grids
Wolfram Burgard, Dieter Fox Daniel Hennig and Timo Schmidt
Februar 1996.

IAI-TR-96-2 A Deterministic Annealing Framework for Unsupervised Texture Segmentation
Thomas Hofmann, Jan Puzicha and Joachim M. Buhmann
Februar 1996.

IAI-TR-96-3 Map Learning and High-Speed Navigation in RHINO
S. Thrun, A. Bücken, W. Burgard, D. Fox, T. Fröhlinghaus, D. Hennig, T. Hofmann, M. Krell and T. Schmidt
Juli 1996.

IAI-TR-96-4 Belief Revision in a Nonclassical Logic
Gerhard Lakemeyer and Wilfried Lang
Juli 1996.

IAI-TR-96-5 Bildverarbeitung und Mustererkennung
W. Förstner und A.B. Cremers
Juli 1996.

IAI-TR-96-6 Implementing Chimera on Top of an Active Relational Database System
Ulrike Griefahn and Thomas Lemke
Juli 1996.

IAI-TR-96-7 Propagation Rule Compiler:
Technical Documentation
Ulrike Griefahn and Thomas Rath
Juli 1996.

IAI-TR-96-8 Data Visualization by Multimensional Scaling:
A Deterministic Annealing Approach
Joachim Buhmann and Hansjörg Klock
Oktober 1996.

IAI-TR-96-9 Monadic-style Backtracking
Ralf Hinze
Oktober 1996.

IAI-TR-96-10 The system Cobrep: Automatic Conversion of
CSG Building Data to Boundary Representation
Roman Englert
Oktober 1996.

IAI-TR-96-11 Objects don't migrate!
Günter Kniesel
April 1996.

IAI-TR-96-12 Encapsulation - Visability and Access Rights
Günter Kniesel
November 1996.

IAI-TR-96-13 Maintaining Library Catalogues with an RDBMS
- A Performance Study -
O. Balownew, T. Bode, A.B. Cremers,
J. Kalinski, J.E. Wolff, H. Rottmann
November 1996.

IAI-TR-96-14 Hierarchical Radiosity on Topological Data Structures
Heinzgerd Bendels, Dieter W. Fellner,
Stephan Schäfer
November 1996.

IAI-TR-97-1 Closing the Gap between Modeling and Radiosity
Dieter W. Fellner, Stephan Schäfer
Januar 1997.

IAI-TR-97-2 P-Subgraph Isomorphism Computation
and Upper Bound Complexity Estimation
Roman Englert, Jörg Seelmann-Eggebert
Januar 1997.

IAI-TR-97-3 Active Mobile Robot Localization
Wolfram Burgard, Dieter Fox and Sebastain Thrun
Februar 1997.

IAI-TR-97-4 Multiscale Annealing for Real-Time Unsupervised Texture Segmentation
Jan Puzicha and Joachim M. Buhmann
April 1997.

IAI-TR-97-5 The MYVIEW System: Tackling the Interface Problem
J.E. Wolff and J. Kalinski
Dezember 1997.

IAI-TR-98-1 On Spatial Quantization of Color Images
Jan Puzicha, Marcus Held, Jens Ketterer,
Joachim M. Buhmann, Dieter W. Fellner
Januar 1998.

IAI-TR-98-2 A Reconfigurable Image Capture and Image Processing System for Autonomous Robots - A Proposal -
Tom Arbuckle, Michael Beetz,
and Boris Trouvain
Januar 1998.

IAI-TR-98-3 Empirical Risk Approximation: An Induction Principle for Unsupervised Learning
Joachim M. Buhmann
April 1998.

IAI-TR-98-4 Type-safe Delegation for Dynamic Component Adaptation
Günter Kniesel
Mai 1998.

IAI-TR-98-5 Delegation for Java: API or Language Extension?
Günter Kniesel
Mai 1998.

IAI-TR-98-6 Experiences with an Interactive Museum Tour-Guide Robot
Wolfram Burgard, Armin B. Cremers, Dieter Fox,
Dirk Hähnel, Gerhard Lakemeyer, Dirk Schulz,
Walter Steiner, Sebastian Thrun
Juni 1998.

IAI-TR-98-7 Support Vector Machines for Land Usage Classification in Landsat TM Imagery
Lothar Hermes, Jan Puzicha, Dieter Frieauff,
and Joachim M. Buhmann
Juni 1998.

IAI-TR98-8 Aufnahme und Intelligente Auswertung Grossflächiger Szenen
A. Brunn, R. Englert, Th. Hau,
A.B. Cremers und W. Förstner
September 1998.

IAI-TR-98-9 Komponentenbasierte Anpassbarkeit von Groupware
Oliver Stiemerling
September 1998.

IAI-TR-98-10 Polytypic Functions Over Nested Datatypes
Ralf Hinze
Oktober 1998.

IAI-TR-98-11 Generalizing Generalized Tries
Ralf Hinze
November 1998.

IAI-TR-98-12 Numerical Representations as Higher-Order Nested Datatypes
Ralf Hinze
Dezember 1998.

IAI-TR-98-13 The MyVIEW Project: from Heterogeneous Bibliographic Information Repositories to Personalized Digital Libraries
Jens E. Wolff and Armin B. Cremers
Dezember 1998.

IAI-TR-98-14 The Use of Cooperation Scenarios in the Design and Evaluation of a CSCW System
Oliver Stiemerling and Armin B. Cremers
Dezember 1998.

IAI-TR-98-15 Portfoliooptimierung mit Verfahren aus der Informatik
Armin B. Cremers, Christoph Hundack, Jens Lüssem
Dezember 1998.

IAI-TR-99-1 Deriving Monad Transformers
Ralf Hinze
Januar 1999.

IAI-TR-99-2 Polytypic Programming With Ease
Ralf Hinze
Februar 1999.

IAI-TR-99-3 Experience- and Model-based Transformational Learning of Symbolic Behavior Specifications
Michael Beetz and Thorsten Belker
März 1999.

IAI-TR-99-4 Perfect Trees and Bit-reversal Permutations
Ralf Hinze
März 1999.

IAI-TR-99-5 Manufacturing Datatypes
Ralf Hinze
April 1999.

IAI-TR-99-6 Constructing Red-Black Trees
Ralf Hinze
Mai 1999.

IAI-TR-99-7 Nonlinear Discriminant Analysis Using Kernel Functions
Volker Roth and Volker Steinhage
Juni 1999.

IAI-TR-99-8 Efficient Generalized Folds
Ralf Hinze
Juni 1999.

IAI-TR-99-9 A New Approach to generic Functional Programming
Ralf Hinze
Juli 1999.

IAI-TR-99-10 Distributed Component-Based Tailorability for CSCW Applications
Oliver Stiemerling, Ralph Hinken, and Armin B. Cremers
März 1999.

IAI-TR-99-11 Flexible Aliasing with Protection
Günter Kniesel and Dirk Theisen
Juni 1999.

IAI-TR-99-12 *XPRES*: a Ranking Approach to Retrieval on Structured Documents
J.E. Wolff, H. Flörke, A.B. Cremers
Juli 1999.

IAI-TR-99-13 The EVOLVE Tailoring Platform: Supporting the Evolution of Component -based Groupware
Oliver Stiemerling, Ralph Hinken, Armin B. Cremers
September 1999.

IAI-TR-99-14 Mutiple SIP Strategies -
A Dynamic Approach to
Deductive Query Processing
Andreas Behrend
Dezember 1999.

IAI-TR-99-15 Polytypic values possess polykind types
Ralf Hinze
Dezember 1999.

IAI-TR-2000-1 Probabilistic Hybrid Action Models for Predicting Concurrent Percept-driven Robot Behavior
Michael Beetz and Henrik Grosskreutz
Juni 2000.

IAI-TR-2001-1 NewsSIEVE
Ein selbstadaptiver Filter für textuelle Informationen
Elmar Haneke
Januar 2001.

IAI-TR-2001-2 Perspektiven: Persistente Objekte mit anwendungsspezifischer Struktur und Funktionalität
Wolfgang Reddig
November 2001.

IAI-TR-2001-3 A New Adaptive Algorithm for the Polygonization of Noisy Imagery
Lothar Hermes and Joachim M. Buhmann
Dezember 2001.

IAI-TR-2002-1 Second Workshop on aspect-Oriented Software development
P. Costanza, G. Kniesel, K. Mehner, E. Pulvermüller and A. Speck
Februar 2002.

IAI-TR-2002-2 Constructing Tournament Representations: An Exercise in Pointwise Relational Programming
Ralf Hinze
Februar 2002.

IAI-TR-2002-3 Church Numerals, Twice!
Ralf Hinze
Februar 2002.

IAI-TR-2002-4 Schnelle Numerische Verfahren zur Bewertung von Europäischen Optionen in Erweiterten Black-Scholes Marktmodellen
Stefan Alex Popovici
Mai 2002.

IAI-TR-2002-5 Optimal Cluster Preserving Embedding of Non-Metric Proximity Data
Volker Roth, Julian Laub, Motoaki Kawanabe, Joachim M. Buhmann
Juni 2002.

IAI-TR-2002-6 Internet-Based Robotic Tele-Presence
Dirk Schulz
Mai 2002.

IAI-TR-2002-7 Path Based Clustering
Bernd Fischer and Joachim M. Buhmann
August 2002.

IAI-TR-2002-8 The Generalized Lasso: a Wrapper approach to gene selection for microarray data
Volker Roth
August 2002.

IAI-TR-2002-9 Planning to Optimize the UMTS Call Set-up for the Execution of Mobile Applications
Roman Englert
Oktober 2002.

IAI-TR-2003-1 Program-independent Composition of Conditional Transformations
Günter Kniesel and Helge Koch
Juni 2003.

IAI-TR-2003-2 The JMangler Tutorial
Günter Kniesel
Juli 2003.

IAI-TR-2003-3 Static Dependency Analysis for Conditional Program Transformations
Günter Kniesel
Juli 2003.

IAI-TR-2003-4 Objektorientierte Datenmodell-Integration für ein offenes, Paläoökologisches Informationssystem
A.B. Cremers, R. Dikau (Hrsg.)
Dezember 2003.

IAI-TR-2003-5 Concepts for the Representation, Storage, and Retrieval of Spatio-Temporal Objects in 3D/4D Geo-Information-Systems
Jörg Siebeck
Dezember 2003.

IAI-TR-2004-1 Modell- und diagnosebasiertes Transformationslernen von symbolischen Plänen für mobile Roboter
Michael Beetz, Thorsten Belker, Armin B. Cremers
Februar 2004.

IAI-TR-2004-2 Unantizipierte Software-Evolution - Abschlussbericht des TAILOR Projekts (DFG-Projekt CR 65/13)
Pascal Costanza, Günter Kniesel, Armin B. Cremers
Oktober 2004.

IAI-TR-2004-3 A Meta Model for AspectJ
YAN Han, Günter Kniesel, Armin B. Cremers
Oktober 2004.

IAI-TR-2004-4 Uniform Genericity for Aspect Languages
Tobias Rho, Günter Kniesel
Dezember 2004.

IAI-TR-2006-1 A Logic Foundation for Program Transformations
Günter Kniesel
Januar 2006.

IAI-TR-2006-2 "Scrap Your Boilerplate" Reloaded
Ralf Hinze, Andres Löh, and Bruno C. d. S. Oliveira
Januar 2006.

IAI-TR-2006-3 'AOSD.06 Industry Track Proceedings'
Fifth International Conference on Aspect-Oriented Software Development, Bonn, Germany, March 20-24
Matt Chapman, Alexandre Vasseur, Günter Kniesel (Eds.)
März 2006.

IAI-TR-2006-4 Independent Evolution of Design Patterns and Application Logic with Generic Aspects - A Case Study
Tobias Rho, Günter Kniesel
April 2006.

IAI-TR-2006-5 Nonparametric Density Estimation for Human Pose Tracking
Thomas Brox, Bodo Rosenhahn, Uwe Kersting, and Daniel Cremers
April 2006.

IAI-TR-2006-6 EWAS 2006
3rd European Workshop on Aspects in Software
Günter Kniesel (Ed.)
August 2006.

IAI-TR-2006-7 A Fixpoint Approach to State Generation for Stratifiable Disjunctive Deductive Databases
Andreas Behrend
Dezember 2006.

IAI-TR-2006-8 Interaktive, plan-basierte Softwareagenten im verteilten Supply Chain Management
Abschlussbericht im Rahmen des DFG-Schwerpunktprogramms Intelligente Softwareagenten und betriebswirtschaftliche Anwendungsszenarien
Michael Beetz, Armin B. Cremers, Axel Kuhn, Bernd Hellingrath, Jürgen Schumacher, Andre Alberti
Dezember 2006.

IAI-TR-2006-9 "Zeitsynchronisation in TmoteSky Sensornetzen"
Serge Shumilov, Boris Iven, Oliver Mali, Roman Saul
Dezember 2006.

IAI-TR-2007-1 Introducing Curvature into Globally Optimal Image Segmentation: Minimum Ratio Cycles on Product Graphs
Thomas Schoenemann, Daniel Cremers
April 2007.

IAI-TR-2007-2 Globally Optimal Image Segmentation with an Elastic Shape Prior
Thomas Schoenemann, Daniel Cremers
April 2007.

IAI-TR-2007-3 Dual-space Graph Cuts for Motion Layer Decomposition
Thomas Schoenemann and Daniel Cremers
April 2007.

IAI-TR-2007-4 Iterated and Efficient Nonlocal Means for Denoising of Textural Patterns
Thomas Brox, Oliver Kleinschmidt, and Daniel Cremers
August 2007.

IAI-TR-2007-5 Design and Implementation of a Workflow Engine
Sebastian Bergmann
September 2007.

IAI-TR-2007-6 On Local Region Models and the Statistical Interpretation of the Piecewise Smooth Mumford-Shah Functional
Thomas Brox and Daniel Cremers
September 2007.

IAI-TR-2007-7 Time-based Triggers for SQL: A Lingering Issue Revisited
Andreas Behrend, Christian Dorau, and Rainer Manthey
November 2007.

IAI-TR-2007-8 High Resolution Motion Layer Decomposition using Dual-space Graph Cuts
Thomas Schoenemann and Daniel Cremers
Dezember 2007.

IAI-TR-2007-9 Matching Non-rigidly Deformable Shapes Across Images: A Globally Optimal Solution
(DFG-ProjekT CR 250/1-1)
Thomas Schoenemann and Daniel Cremers
Dezember 2007.

IAI-TR-2007-10 Globally Optimal Shape-based Tracking in Real-time
(DFG-ProjekT CR 250/1-1)
Thomas Schoenemann and Daniel Cremers
Dezember 2007.

IAI-TR-2008-1 Software Security Metrics for Malware Resilience
Hanno Langweg
Februar 2008.

IAI-TR-2008-2 Rule-Based Programming
Günter Kniesel, Jorge Sousa Pinto
Juni 2008.

IAI-TR-2008-3 Efficient Planar Graph Cuts with Applications in Computer Vision
Frank R. Schmidt, Eno Töppe and Daniel Cremers
Juni 2008.

IAI-TR-2008-4 The Elastic Ratio: Introducing Curvature into Ratio-based Globally Optimal Image Segmentation
Thomas Schoenemann, Simon Masnou and Daniel Cremers
Dezember 2008.

IAI-TR-2008-5 A convex approach for computing minimal partitions
Antonin Chambolle, Daniel Cremers, Thomas Pock
November 2008.