Applied Mathematics & Optimization (2024) 89:66
https://doi.org/10.1007/s00245-024-10134-5

Segmentation in Measure Spaces

Salvador Moll[1] · **Vicent Pallardó-Julià**[1,2] · **Marcos Solera**[1,3]

Accepted: 29 March 2024

Abstract

We consider an abstract concept of *perimeter measure space* as a very general framework in which one can properly consider two of the most well-studied variational models in image processing: the Rudin–Osher–Fatemi model for image denoising (ROF) and the Mumford–Shah model for image segmentation (MS). We show the linkage between the ROF model and the two phases piecewise constant case of MS in perimeter measure spaces. We show applications of our results to nonlocal image segmentation, via discrete weighted graphs, and to multiclass classification on high dimensional spaces.

Keywords Perimeter measure space · Nonlocal · Segmentation · Denoising · Labeling · Graphs · Metric graphs · Fractional perimeter

Mathematics Subject Classification 94A08 · 49J10 · 05C21 · 68P99 · 94A13 · 68R10 · 90C35

✉ Salvador Moll
j.salvador.moll@uv.es

Vicent Pallardó-Julià
vicent.pallardo@uv.es

Marcos Solera
marcos.solera@uv.es

1 Department d'Anàlisi Matemàtica, Universitat de València, C/ Dr. Moliner, 50, 46100 Burjassot, Spain

2 Kimera Technologies, Lanzadera, La Marina de València, C/ del Moll de la Duana, s/n, Valencia, Spain

3 Departmento de Matemáticas, Universidad Autónoma de Madrid, C/ Francisco Tomás y Valiente, 7, Facultad de Ciencias, módulo 17, 28049 Madrid, Spain

Published online: 20 April 2024

1 Introduction

Segmentation is a fundamental process in a variety of fields in Computer Science. It arises in several areas related to image or video processing and computer vision. Among the most successful mathematical approaches to partitioning an image into its constituent parts, we find the variational approach introduced by Mumford and Shah (MS) in [1].

In the particular case where one is interested in segmenting an image into two different phases, the background and the foreground, the MS-model reduces to the following optimization problem, which is known as Chan–Vese's model [2] and which we denote by CV:

Given an initial image $f : \Omega \subset \mathbb{R}^2 \to [0, 1]$, find a set of finite perimeter $\Sigma \subset \Omega$ (representing the foreground region) and two constants $0 \leq m_0^*, m_1^* \leq 1$ (representing the mean of f in the background and in the foreground, respectively) that minimize

$$\mathrm{Per}(\Lambda) + \mu \left(\int_\Lambda (m_1 - f(x))^2 \, \mathrm{d}x + \int_{\Omega \setminus \Lambda} (m_0 - f(x))^2 \, \mathrm{d}x \right)$$

among all finite perimeter sets $\Lambda \subset \Omega$ and constants $0 \leq m_0, m_1 \leq 1$. Here $\mu > 0$ is a parameter acting on the fidelity term.

Obtaining a minimizer triplet of the above energy functional is not an easy task due to the lack of convexity. However, as it was observed in [3], a partial minimizer can be obtained with the help of the minimizer of the Rudin–Osher–Fatemi functional (ROF) [4]:

$$\int_\Omega |Dw| + \frac{\lambda}{2} \int_\Omega (w(x) - f(x))^2 \, \mathrm{d}x,$$

with parameter $\lambda > 0$. To be more precise (see [3, Theorem 3.4]), for $0 < m_0 < m_1 \leq 1$ fixed, if u is a minimizer of the ROF functional with parameter $\lambda = 2\mu(m_1 - m_0)$, then a suitable threshold of u provides the only minimizer of the CV functional (with m_0, m_1 fixed). In this case, $\Sigma := \{x \in \Omega : u(x) \geq \frac{m_0+m_1}{2}\}$ is the set of finite perimeter which minimizes the energy. Moreover, the authors give a convergent algorithm to find a triplet (Σ, m_0, m_1) which minimizes the CV-energy in each of its three components.

This linkage between the ROF problem for denoising and the CV model for segmentation has been further studied in [5]. The authors showed that in the ℓ_1 anisotropic case (i.e., changing the perimeter by the rectilinear perimeter in the CV functional and the total variation term $\int_\Omega |Du|$ by the corresponding anisotropic one: $\int_\Omega |Du|_1$) the same linkage holds. Moreover, in the case that f is piecewise constant on rectangles, then a true minimizer triplet of the anisotropic CV model can be obtained thanks to this relation.

The purpose of this paper is to generalise these results to many different settings. Thus, we provide a unified framework in which the linkage between both problems holds. A deep look at the results in [3] shows that the only mathematical tools that are needed are suitable definitions of the perimeter of sets and of the total variation of integrable functions, and a coarea formula relating these two concepts.

In Sect. 3 we give the abstract setting in which we work and which we call *perimeter measure space*. This framework is so vast that we can gather together a big variety of situations to which our results apply. In particular, we can generalise the results to the case of a non-Euclidean setting. We give an up-to-date list of the possibilities; including the cases of discrete weighted graphs or fractional perimeters.

Section 4 is devoted to stating and proving the linkage between both problems in perimeter measure spaces, see Theorem 2. As we will observe in Sect. 3, the ROF-model for denoising has been much more studied than the CV-model in a number of non-Euclidean settings. Therefore, the results in this paper can be used to obtain algorithms and properties of CV-minimizers or partial minimizers.

In Sect. 5, we use the previous results to build a two step algorithm (analogous to the classical Chan–Vese algorithm, see [2]) to find an approximate minimizer triplet to the CV problem in the case of locally finite weighted discrete graphs. Then, we show two possible applications of segmentation in weighted graphs: nonlocal image segmentation and labeling. In the case of image segmentation, we show how our approach improves the results obtained with standard Euclidean segmentation. In the case of labeling, a postprocess using the CV model to the results obtained with standard linear diffusion (i.e., with harmonic extension on graphs) yields an accuracy in the predictions of labeling comparable to Poisson learning.

2 Related Work

As we have already mentioned, the linkage between the ROF model for denoising and the CV model for segmentation has already been observed in the particular cases of the isotropic perimeter (and total variation) and of the ℓ_1-anisotropy; both in a domain in $\mathbb{R}^2$.

On the other hand, segmentation in the non-Euclidean case and, in particular, in the case of weighted undirected graphs, has been studied in [6]. The authors consider the Mumford–Shah functional defined in a weighted undirected graph with n points and a parameter ε representing the scale in which differences of the values of the target minimizer function are considered large, and therefore can be interpreted as jump points. The main result of this work is the Γ-convergence, when the graph is considered in some domain in the Euclidean space $\mathbb{R}^N$, to the classical Euclidean MS functional as the point cloud becomes denser in the domain and a suitable scaling in ε is considered.

Concerning perimeter measure spaces, our definition of perimeter functional is very close to the one of generalised perimeter given in [7]. However, their concept of perimeter is restricted to the case of Lebesgue measurable sets. Moreover, a similar definition of perimeter measure space is considered in [8]. There, the authors analyze a perimeter functional alongside a range of potential assumptions and proceed to derive multiple results by considering various combinations of these assumptions. Here, we do not seek a profound study of the lower semicontinuity properties of the total variation functional or the perimeter functional. Instead, we consider the necessary assumptions for the existence and uniqueness of minimizers for the ROF functional, as well as for the mentioned linkage to be valid. Concretely, weak lower semicontinuity of the functional with respect to the L^2-convergence and finiteness of the measure are assumed (see Sect. 3 for details).

We finally mention [9], in which the authors consider an energy functional very similar to the CV one in the case of undirected weighted graphs. In their functional, the total variation on graphs appears. However, their minimization problem only depends on one variable. To be more precise, they minimize the total variation minus the variance for functions taking values in the canonical basis of $\mathbb{R}^N$.

3 Perimeter Measure Spaces

3.1 Definitions

Definition 1 Let $(X, \mathcal{B}, \nu)$ be a measure space. We say that $\mathrm{Per} : \mathcal{B} \to [0, \infty]$ is a perimeter functional if the following conditions are satisfied:

(i) $\mathrm{Per}(\emptyset) = 0$;
(ii) $\mathrm{Per}(A) = \mathrm{Per}(B)$ for every $A, B \in \mathcal{B}$ such that $\nu(A \triangle B) = 0$;
(iii) $\mathrm{Per}(A) = \mathrm{Per}(X \setminus A)$ for every $A \in \mathcal{B}$;
(iv) Per is sub-modular, i.e., $\mathrm{Per}(A \cup B) + \mathrm{Per}(A \cap B) \leq \mathrm{Per}(A) + \mathrm{Per}(B)$ for every $A, B \in \mathcal{B}$.

Definition 2 Let $(X, \mathcal{B}, \nu)$ be a measure space and $\mathrm{Per} : \mathcal{B} \to [0, \infty]$ a perimeter functional. For a function $u \in L^1(X, \nu)$, and letting $E_t(u) := \{x \in X : u(x) > t\}$ for every $t \in \mathbb{R}$, we define its total variation by

$$\mathrm{TV}(u) := \int_{-\infty}^{\infty} \mathrm{Per}(E_t(u))\mathrm{d}t,$$

with the convention that $\mathrm{TV}(u) = \infty$ if the map $t \mapsto \mathrm{Per}(E_t(u))$ is not measurable. Note that $\mathrm{TV}(\chi_E) = \mathrm{Per}(E)$ for every $E \in \mathcal{B}$.

A perimeter measure space $(X, \mathcal{B}, \nu, \mathrm{Per})$ (PMS in short) is a measure space $(X, \mathcal{B}, \nu)$ such that $\nu(X) < +\infty$, together with a perimeter functional Per such that the total variation functional TV is lower semi-continuous with respect to the weak convergence in $L^2(X, \nu)$. We will denote it by (X, ν, Per) for short.

Observe that the standard Euclidean perimeter in $\mathbb{R}^N$ endowed with the Lebesgue measure is a PMS. Additional examples of PMS, including this one, are given in Sect. 3.3.

Proposition 1 *Let (X, ν, Per) be a PMS. Then, the total variation is a convex functional in $L^2(X, \nu)$.*

Proof The proof of this result essentially follows the lines of [10, Proposition 3.4]. We give the details for the sake of completeness.

Observe that, since $\nu(X) < +\infty$, $L^2(X, \nu) \subset L^1(X, \nu)$ so the total variation is well defined for $L^2(X, \nu)$ functions. First of all, we show that TV is a positively one-homogeneous functional. Let $u \in L^1(X, \nu)$ and $\lambda > 0$. Then,

$$\begin{aligned}\mathrm{TV}(\lambda u) &= \int_{-\infty}^{\infty} \mathrm{Per}(E_t(\lambda u))\,\mathrm{d}t = \int_{-\infty}^{\infty} \mathrm{Per}\left(E_{\frac{t}{\lambda}}(u)\right)\,\mathrm{d}t \\ &= \lambda \int_{-\infty}^{\infty} \mathrm{Per}(E_t(u))\mathrm{d}t = \lambda \mathrm{TV}(u).\end{aligned}$$

Then, to show the convexity, it suffices to prove the following inequality:

$$\mathrm{TV}(u_1 + u_2) \leq \mathrm{TV}(u_1) + \mathrm{TV}(u_2) \quad \text{for every } u_1, u_2 \in L^2(X, \nu). \tag{3.1}$$

We first prove the following representation formula for nonnegative, bounded and integer valued measurable functions u in X:

$$\mathrm{TV}(u) = \min\left\{\sum_{i=1}^{m} \mathrm{Per}(A_i) \,:\, u = \sum_{i=1}^{m} \chi_{A_i},\ A_i \in \mathcal{B},\ m \in \mathbb{N}\right\}. \tag{3.2}$$

Note that, if u satisfies these assumptions, then $u = \sum_{i=1}^{M} \chi_{E_{i-1}(u)}$ for some $M \in \mathbb{N}$. Take $A_i \in \mathcal{B}$, $i = 1, \ldots, m$, such that $u = \sum_{i=1}^{m} \chi_{A_i}$. Using the sub-modularity of the perimeter, for any $j \neq k \in \{1, \ldots, m\}$,

$$\mathrm{Per}(A_j \cap A_k) + \mathrm{Per}(A_j \cup A_k) + \sum_{i \notin \{j,k\}} \mathrm{Per}(A_i) \leq \sum_{i=1}^{m} \mathrm{Per}(A_i).$$

From this inequality, by induction on m, one easily gets

$$\sum_{i=1}^{m} \mathrm{Per}(A_i) \geq \sum_{i=1}^{M} \mathrm{Per}(E_{i-1}(u)).$$

Indeed, if $m = 1$ then $A_1 = E_0(u)$ up to a ν-null set thus $\mathrm{Per}(A_1) = \mathrm{Per}(E_0(u))$. Now, suppose that the inequality holds for m and let $u = \sum_{i=1}^{m+1} \chi_{A_i}$. Let $v := \sum_{i=1}^{m} \chi_{A_i}$. By induction hypothesis we have that

$$\sum_{i=1}^{m+1} \mathrm{Per}(A_i) \geq \sum_{i=1}^{m} \mathrm{Per}(E_{i-1}(v)) + \mathrm{Per}(A_{m+1}).$$

Now,

$$E_j(u) = E_j(v) \cup (A_{m+1} \cap E_{j-1}(v))$$

so, by submodularity:

$$\mathrm{Per}(E_j(v)) \geq \mathrm{Per}(E_j(u)) + \mathrm{Per}(A_{m+1} \cap E_j(v)) - \mathrm{Per}(A_{m+1} \cap E_{j-1}(v)).$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{m}\operatorname{Per}(E_{i-1}(v)) &\geq \sum_{i=1}^{m}\operatorname{Per}(E_{i-1}(u))+\operatorname{Per}(A_{m+1}\cap E_{i-1}(v))-\operatorname{Per}(A_{m+1}\cap E_{i-2}(v))\\
&=\sum_{i=1}^{m}\operatorname{Per}(E_{i-1}(u))+\operatorname{Per}(A_{m+1}\cap E_{m-1}(v))-\operatorname{Per}(A_{m+1}\cap E_{-1}(v))\\
&=\sum_{i=1}^{m}\operatorname{Per}(E_{i-1}(u))+\operatorname{Per}(A_{m+1}\cap E_{m-1}(v))-\operatorname{Per}(A_{m+1})\\
&=\sum_{i=1}^{m}\operatorname{Per}(E_{i-1}(u))+\operatorname{Per}(E_{m}(v)\cup(A_{m+1}\cap E_{m-1}(v)))-\operatorname{Per}(A_{m+1})\\
&=\sum_{i=1}^{m}\operatorname{Per}(E_{i-1}(u))+\operatorname{Per}(E_{m}(u))-\operatorname{Per}(A_{m+1})\\
&=\sum_{i=1}^{m+1}\operatorname{Per}(E_{i-1}(u))-\operatorname{Per}(A_{m+1}).\end{aligned}$$

Thus, by taking the infimum, we conclude the proof of (3.2):

$$\begin{aligned}&\inf\left\{\sum_{i=1}^{m}\operatorname{Per}(A_i)\ :\ u=\sum_{i=1}^{m}\chi_{A_i}\ ,\ A_i\in\mathcal{B}\right\}\\
&\quad=\sum_{i=1}^{M}\operatorname{Per}(E_{i-1}(u))=\int_0^{M-1}\operatorname{Per}(E_t(u))\,\mathrm{d}t=\operatorname{TV}(u).\end{aligned}$$

We now prove (3.1): Observe that we can suppose that $\operatorname{TV}(u_i)$ is finite for $i=1,2$. As a first step, suppose that $0\leq u_i\leq 1$, $i=1,2$, and consider the following approximations:

$$u_{i,n}:=\frac{1}{n}\sum_{k=-1}^{n}\chi_{E_{\frac{k+t}{n}}(u_i)}\quad\text{for } n\in\mathbb{N},\ t\in]0,1[\text{ and } i=1,2.$$

By construction, $0\leq u_{i,n}\leq 3$ and $u_{i,n}\to u_i$ in $L^1(X,\nu)$, $i=1,2$. Then, we can extract a subsequence (not renamed) converging a.e. to u_i. Since $\nu(X)<+\infty$, we obtain that $u_{i,n}\to u_i$ in $L^2(X,\nu)$. Therefore, from the lower semicontinuity and the one-homogeneity of the total variation we get

$$\begin{aligned}\operatorname{TV}(u_1+u_2)&\leq\liminf_{n\to\infty}\operatorname{TV}(u_{1,n}+u_{2,n})\\
&=\liminf_{n\to\infty}\frac{1}{n}\operatorname{TV}\left(\sum_{k=-1}^{n}\chi_{E_{\frac{k+t}{n}}(u_1)}+\chi_{E_{\frac{k+t}{n}}(u_2)}\right)\end{aligned}$$

$$\stackrel{(3.2)}{\leq} \liminf_{n\to\infty} \frac{1}{n} \sum_{k=-1}^{n} \operatorname{Per}\left(E_{\frac{k+t}{n}}(u_1)\right) + \operatorname{Per}\left(E_{\frac{k+t}{n}}(u_2)\right).$$

Finally, applying [10, Lemma 3.3], we can choose $t \in [0, 1[$ such that, up to a subsequence,

$$\frac{1}{n} \sum_{k=-1}^{n} \operatorname{Per}\left(E_{\frac{k+t}{n}}(u_i)\right) \to \int_0^1 \operatorname{Per}(E_t(u_i))\, \mathrm{d}t = \operatorname{TV}(u_i) \quad \text{as } n \to \infty,\ i = 1, 2,$$

which finishes the proof in the case $0 \leq u_i \leq 1$.

The case where the u_i are bounded, $i = 1, 2$, easily follows by considering $\tilde{u}_i := \frac{u_i - m}{M - m}$ with $m \leq u_i \leq M$ and applying the previous step.

Finally, if $u_i \in L^2(X, \nu)$, $i = 1, 2$, we choose as approximations the truncations at level $\pm K$; i.e., $u_i^K := \max\{\min\{u_i, K\}, -K\}$ (observe that $u_i^K \to u_i$ in $L^2(X, \nu)$ as $K \to +\infty$) and use the lower semicontinuity of TV:

$$\begin{aligned} \operatorname{TV}(u_1 + u_2) &\leq \liminf_{K\to\infty} \operatorname{TV}(u_1^K + u_2^K) \\ &\leq \liminf_{K\to\infty} \operatorname{TV}(u_1^K) + \operatorname{TV}(u_2^K) \leq \limsup_{K\to\infty} \operatorname{TV}(u_1^K) + \operatorname{TV}(u_2^K) \\ &= \limsup_{K\to\infty} \int_{-K}^{K} (\operatorname{Per}(E_t(u_1)) + \operatorname{Per}(E_t(u_2)))\, \mathrm{d}t \\ &= \operatorname{TV}(u_1) + \operatorname{TV}(u_2). \end{aligned}$$

□

3.2 CV and ROF Models in PMS

We continue with the generalization of the CV and ROF models to the PMS setting:

From now on, we assume that $(X, \nu, \operatorname{Per})$ is a PMS. Let $f : X \to [0, 1]$ be a measurable function not ν-a.e. equal to a constant. The **CV model** on $(X, \nu, \operatorname{Per})$ aims to minimize the following energy functional over $\Sigma \in \mathcal{B}$ and $m_0, m_1 \in [0, 1]$:

$$\mathcal{E}_\mu^{CV}(\Sigma, m_0, m_1) := \operatorname{Per}(\Sigma) + \mu \left(\int_\Sigma (m_1 - f)^2\, d\nu + \int_{X\setminus\Sigma} (m_0 - f)^2\, \mathrm{d}\nu \right)$$

where $\mu > 0$ is a scale parameter.

The **ROF model** on the perimeter measure space $(X, \nu, \operatorname{Per})$ takes the following form, for $f \in L^2(X, \nu)$:

$$\min \left\{ \mathcal{E}_\lambda^{ROF}(u) := \operatorname{TV}(u) + \frac{\lambda}{2} \int_X |u(x) - f(x)|^2 d\nu(x) :\ u \in L^2(X, \nu) \right\}. \tag{3.3}$$

We now prove that (3.3) has a unique solution.

Theorem 1 *The $\mathcal{E}_\lambda^{ROF}$ functional has a unique minimizer $u \in L^2(X,\nu)$. Moreover, it is the only function in $L^2(X,\nu)$ satisfying $0 \in \partial(\mathcal{E}_\lambda^{ROF})(u)$. If $0 \le f \le 1$ ν-a.e., then $0 \le u \le 1$ ν-a.e.*

Here, the subdifferential of the $\mathcal{E}_\lambda^{ROF}$ functional in $L^2(X,\nu)$ is defined as follows: $v \in \partial(\mathcal{E}_\lambda^{ROF})(u)$ if, and only if, $u, v \in L^2(X,\nu)$ and, for all $w \in L^2(X,\nu)$,

$$\int_X v(w-u)\,\mathrm{d}\nu \le \mathcal{E}_\lambda^{ROF}(w) - \mathcal{E}_\lambda^{ROF}(u).$$

Proof Since the TV functional is lower semicontinuous with respect to the weak convergence in $L^2(X,\nu)$, existence of a minimizer follows from the direct method of the calculus of variations. Indeed, let $\{u_k\}_{k\in\mathbb{N}} \subset L^2(X,\nu)$ be a minimizing sequence. Then,

$$\begin{aligned}\int_X |u_k(x)|^2\,d\nu(x) &\le 2\int_X (u_k(x)-f(x))^2\,d\nu(x) + 2\int_X |f(x)|^2\,d\nu(x)\\ &\le \frac{4}{\lambda}\sup_{k\in\mathbb{N}}\{\mathcal{E}_\lambda^{ROF}(u_k)\} + 2\int_X |f(x)|^2\,d\nu(x).\end{aligned}$$

Therefore, up to a subsequence, u_k converges to a function $u \in L^2(X,\nu)$ weakly in $L^2(X,\nu)$. Finally, applying the lower semicontinuity of the total variation we obtain that u is a minimizer of $\mathcal{E}_\lambda^{ROF}$.

Moreover, by Proposition 1, it is a convex functional. Therefore, the $\mathcal{E}_\lambda^{ROF}$ functional is strictly convex. This yields uniqueness of the minimizer.

Note that $0 \in \partial(\mathcal{E}_\lambda^{ROF})(u)$ is equivalent to u being a minimizer.

Finally, if $0 \le f \le 1$, the bound $0 \le u \le 1$ is a direct consequence of the fact that the truncation of any function $v \in L^2(X,\nu)$ by $T_0^1(s) := \max\{\min\{1,s\},0\}$ satisfies

$$\mathcal{E}_\lambda^{ROF}(T_0^1(v)) = \int_0^1 \mathrm{Per}(E_t(v))\,\mathrm{d}t + \frac{\lambda}{2}\int_X |T_0^1(v(x)) - f(x)|^2 d\nu(x) \le \mathcal{E}_\lambda^{ROF}(v).$$

□

3.3 Examples of PMS

We now give an extensive list of examples of PMS:

3.3.1 Isotropic Euclidean Case

Let $\Omega \subset \mathbb{R}^N$ be an open bounded set and $\nu = \mathcal{L}^N$ be the Lebesgue measure in $\mathbb{R}^N$. Recall that, for an integrable function $u \in L^1(\Omega)$, the total variation of u is defined as

$$\mathrm{TV}(u) := \sup\left\{\int_\Omega u \,\mathrm{div}\,\mathbf{z}\,\mathrm{d}x : \mathbf{z} \in C_0^\infty(\Omega;\mathbb{R}^N),\ \|\mathbf{z}(x)\| \le 1 \text{ a.e. in } \Omega\right\}, \quad (3.4)$$

where $\|\cdot\|$ is the Euclidean norm. The perimeter of a Lebesgue measurable subset $E \subset \Omega$ is defined as the total variation of the characteristic function of E; i.e.,

$$\mathrm{Per}(E) := \mathrm{TV}(\chi_E).$$

By the coarea formula (see [11, Theorem 3.40]), we get that TV defined as above coincides with the one in Definition 2. The submodularity of the perimeter can be found in [11, Proposition 3.38]. On the other hand, the lower semicontinuity with respect to weak convergence in $L^2(X,\nu)$ follows easily from the expression in (3.4) as we next show:

Suppose that $u_n \rightharpoonup u$ weakly in $L^2(X,\nu)$. Then, for any $\mathbf{z} \in C_0^\infty(\Omega;\mathbb{R}^N)$ such that $\|\mathbf{z}(x)\| \le 1$,

$$\int_X u(x)\mathrm{div}\,\mathbf{z}\,(x)\,dx = \liminf_{n\to\infty} \int_X u_n(x)\mathrm{div}\,\mathbf{z}\,(x)\,dx \le \liminf_{n\to\infty} \mathrm{TV}(u_n).$$

Taking the supremum in $\mathbf{z}$ we get that

$$\mathrm{TV}(u) \le \liminf_{n\to\infty} \mathrm{TV}(u_n).$$

Therefore, $(\Omega,\ \mathcal{L}^N\big|_\Omega, \mathrm{Per})$ is a PMS.

3.3.2 Anisotropic Euclidean Case

Let $\Omega \subset \mathbb{R}^N$ be an open bounded set, $\nu = \mathcal{L}^N$ be the Lebesgue measure in $\mathbb{R}^N$ and $\Phi : \Omega \times \mathbb{R}^N \to [0,+\infty[$ be a coercive anisotropy; i.e.:

- $\Phi(x,\cdot)$ is convex and lower semicontinuous for any $x \in \Omega$;
- $\Phi(x,t\xi) = |t|\Phi(x,\xi)$ for any $(x,\xi) \in \Omega \times \mathbb{R}^N$ and any $t \in \mathbb{R}$ (positive 1-homogeneity);
- There exists $C \ge 1$ such that $\frac{1}{C}\|\xi\| \le \Phi(x,\xi) \le C\|\xi\|$ for all $(x,\xi) \in \Omega \times \mathbb{R}^N$ (coercivity and sublinear growth).

We denote by Φ^0 the dual anisotropy of Φ; i.e.,

$$\Phi^0(x,\xi) := \sup\{\xi\cdot\eta\ :\ \eta \in \mathbb{R}^N,\ \Phi(x,\eta) \le 1\}.$$

For any $u \in L^1(\Omega)$, one can define

$$\mathrm{TV}_\Phi(u) := \sup\left\{\int_\Omega u\,\mathrm{div}\,\mathbf{z}\,\mathrm{dx}\ :\ \mathbf{z} \in \mathcal{H}_\phi(\Omega)\right\},$$

with

$$\mathcal{H}_\phi(\Omega) := \{\mathbf{z} \in L^\infty(\Omega;\mathbb{R}^N)\ :\mathrm{div}\,\mathbf{z} \in L^N(\Omega),\ \Phi^0(x,\mathbf{z}(x)) \le 1 \text{ a.e. in } \Omega\}.$$

We define the anisotropic perimeter as

$$\mathrm{Per}_\Phi(E) := \mathrm{TV}_\Phi(\chi_E) \quad \text{for every Lebesgue measurable set } E \subset \Omega.$$

The coarea formula and the submodularity of the perimeter can be found in [12] and [13, Remark 2.4], respectively. Then, as in the previous example, we can easily show that $(\Omega, \mathcal{L}^N|_\Omega, \mathrm{Per}_\Phi)$ is a PMS.

The anisotropic ROF model (for anisotropies depending only on the gradient) was proposed by Esedoglu and Osher in [14] as a tool for image denoising for images with some particular geometric features. The subdifferential of the functional TV_Φ was completely characterised (under Dirichlet constraints on the boundary or if $\Omega = \mathbb{R}^N$) in [15].

3.3.3 Fractional Perimeter

Let $\Omega \subset \mathbb{R}^N$ be an open bounded set. Given $s \in (0, 1)$, the fractional Sobolev space $W^{s,1}(\Omega)$ is defined as the set of functions

$$W^{s,1}(\Omega) := \left\{ u \in L^1(\Omega) \,:\, \mathrm{TV}_s(u) := \int_\Omega \int_\Omega \frac{|u(x) - u(y)|}{|x - y|^{N+s}} \, \mathrm{d}x\mathrm{d}y < +\infty \right\}.$$

Moreover, defining the nonlocal perimeter as

$$\mathrm{Per}_s(E) := \int_{E\cap\Omega} \int_{(\mathbb{R}^N \setminus E)\cap\Omega} \frac{1}{|x - y|^{N+s}} \, \mathrm{d}x\mathrm{d}y,$$

for every Lebesgue measurable set $E \subset \mathbb{R}^N$, one can show as above that $\left(\mathbb{R}^N, \mathcal{L}^N|_\Omega, \mathrm{Per}_s\right)$ is a PMS. In fact, the coarea formula for TV_s can be found in [16, Proposition 3.1] or [17, Lemma 10], the submodularity of the perimeter can be proved as in [18, Sect. 2.1.2] and one also has [19, Theorem 3.4] for a representation formula of TV_s similar to that in (3.4). Let us mention that this nonlocal concept of perimeter was first introduced in [20].

The ROF model for the fractional perimeter in the case $\Omega = \mathbb{R}^N$ has been recently studied in [21] (and in [22] or [23] with L^1 fidelity term).

3.3.4 Random Walk Spaces

Let $(X, \mathcal{B})$ be a measurable space such that the σ-field $\mathcal{B}$ is countably generated. A random walk on $(X, \mathcal{B})$ is a family of probability measures $(m_x)_{x\in X}$ on $\mathcal{B}$ such that $x \mapsto m_x(B)$ is a measurable function on X for each fixed $B \in \mathcal{B}$. Moreover, a σ-finite measure ν on $\mathcal{B}$ is reversible with respect to the random walk m if the following balance condition holds:

$$\mathrm{d}m_x(y)\mathrm{d}\nu(x) = \mathrm{d}m_y(x)\mathrm{d}\nu(y) \quad \text{for every } x, y \in X.$$

$(X, \mathcal{B})$ together with a random walk m and a σ-finite measure ν which is reversible with respect to m is called a reversible random walk space and denoted by $[X, \mathcal{B}, m, \nu]$.

Given $E \in \mathcal{B}$, the m-perimeter of E is defined as

$$\mathrm{Per}_m(E) := \int_E \int_{X \setminus E} \mathrm{d}m_x(y)\mathrm{d}\nu(x)$$

and the m-total variation of a measurable function $u : X \to \mathbb{R}$ is defined by

$$\mathrm{TV}_m(u) := \frac{1}{2} \int_X \int_X |u(y) - u(x)| \, \mathrm{d}m_x(y)\mathrm{d}\nu(x).$$

Then, if $\nu(X) < +\infty$, (X, ν, Per_m) is a PMS (see [24] or [25] for the coarea formula, the submodularity of the perimeter and the lower semicontinuity result). The ROF model in reversible random walk spaces has been studied in [26] and the subdifferential of the m-total variation functional was completely characterized in [24].

We point out that many different examples can be included in this category of reversible random walk spaces. We give a non-exhaustive list of them, including only the more relevant to us and without entering into details, for which we refer to [24] or [25]:

1. Nonlocal perimeter with an integrable Kernel: Let $J : \mathbb{R}^N \to [0, +\infty[$ be a Lebesgue measurable, nonnegative and radially symmetric function such that $\int_{\mathbb{R}^N} J(x)\mathrm{d}x = 1$. Let $\Omega \subset \mathbb{R}^N$ be a closed set of finite Lebesgue measure and

$$m_x(A) := \int_A J(x-y)dy + \left(\int_{\mathbb{R}^n \setminus \Omega} J(x-z)dz \right) \delta_x(A)$$

   for any $x \in \Omega$ and any Lebesgue measurable set $A \subset \Omega$. Then, $m = (m_x)$ is a random walk on Ω with respect to which the Lebesgue measure on Ω is reversible. This is the usual random walk arising from a nonsingular kernel J, but modified so that it does not jump outside of Ω (immediately reflected back to the starting position).
2. Markov chains on a countable space X with a reversible probability measure π: Given a Markov Kernel $K : X \times X \to \mathbb{R}$, one defines

$$m_x(A) := \sum_{y \in A} K(x, y) \quad \text{for any } x \in X \text{ and any } A \subset X.$$

   Then, if π is a reversible probability measure with respect to K, $[X, \mathcal{B}, m, \pi]$ is a reversible random walk space (here $\mathcal{B}$ is the σ-algebra of all subsets of X).
3. Locally finite undirected weighted discrete graphs. Let $G = (V(G), E(G))$ be a locally finite weighted discrete graph with vertex set $V(G)$ and suppose that each edge $(x, y) \in E(G)$ has a positive weight $w_{xy} = w_{yx}$ assigned. For each

$x \in V(G)$, the random walk m is defined as follows:

$$m_x := \frac{\sum_{y\sim x} w_{xy}\delta_y}{\sum_{y\sim x} w_{xy}},$$

with $y \sim x$ denoting that $(x, y) \in E(G)$ (alternatively, we can suppose that $w_{xy} = 0$ if $(x, y) \notin E(G)$ and consider the sum over all the vertices $y \in V(G)$). Then, for $A \subset V(G)$, one considers

$$\nu(A) := \sum_{x\in A}\sum_{y\sim x} w_{xy}.$$

It follows that ν is a reversible measure with respect to the random walk m.

3.3.5 Carnot-Carathéodory Spaces

Given an open set $\Omega \subset \mathbb{R}^N$, $m < N$ and m locally Lipschitz vector fields $X = (X_1, \ldots, X_m)$, a distance in Ω is defined as follows:

$$d(x, y) = \inf\left\{\int_0^1 |\dot\gamma(t)|\,dt \ :\ \gamma(0) = x,\ \gamma(1) = y,\ \dot\gamma(t) = \sum_{i=1}^m a_i(t)X_i\ ,\ 0 < t < 1\right\}.$$

Assuming that this distance is everywhere finite, the space of bounded variation $BV_X(\Omega)$ is defined as the space of functions $u \in L^1(\Omega)$ such that

$$\mathrm{TV}_X(u) := \sup\left\{\int_\Omega u\,\mathrm{div}_{X^*}\mathbf{z}\ :\ \mathbf{z} \in C_0^1(\Omega; \mathbb{R}^N),\ \|\mathbf{z}\|_\infty \le 1\right\},$$

where $\mathrm{div}_{X^*}\mathbf{z} = \sum_{i=1}^m X_i^* z^i$, with X^* the adjoint vector field of X. For any Lebesgue measurable $E \subset \Omega$, its perimeter is defined as

$$\mathrm{Per}_X(E) := \mathrm{TV}_X(\chi_E).$$

The coarea formula and the submodularity of the perimeter can be found in [27, Propositions 4.2 and 4.7] and the lower semicontinuity can be proved as in the Euclidean case. Then, $(\Omega, \mathcal{L}^N, \mathrm{Per}_X)$ is a PMS (see [27]). The subdifferential of the total variation functional coupled with different Neumann or Dirichlet conditions in general metric measure spaces (which includes the case of Carnot-Carathéodory spaces) has been recently studied in [28].

3.3.6 Metric Graphs

We follow [29] for the definition of a metric graph, which we now briefly recall. Let E be a finite set. Given $\{\ell_e\}_{e\in E} \subset]0, \infty[$, consider the family $[0, \ell_e]_{e\in E}$ of metric measure subspaces of $\mathbb{R}$ (with the Euclidean metric d_e and Lebesgue measure λ_e) and their disjoint union

$$\mathcal{E} := \bigsqcup_{e\in E} [0, \ell_e].$$

We adopt the notation (x, e) for the element of $\mathcal{E}$ with $x \in [0, \ell_e]$ and $e \in E$. We endow $\mathcal{E}$ with the disjoint union topology. For this, let $\varphi_e : [0, \ell_e] \ni x \mapsto (x, e) \in \mathcal{E}$ be the canonical injection. Then, a set $U \subset \mathcal{E}$ is open if and only if each $\varphi_e^{-1}(U)$ is a union of sets of the form $[0, \varepsilon_1[$, $]\varepsilon_2, \ell_e]$ or $]\varepsilon_3, \varepsilon_4[$, for $\varepsilon_i \in]0, \ell_e[$.

Consider now the set

$$\mathcal{V} := \bigsqcup_{e\in E} \{0, \ell_e\}$$

of endpoints of $\mathcal{E}$. Given any equivalence relation $\sim$ on $\mathcal{V}$, we extend it to an equivalence relation on $\mathcal{E}$ by equality: i.e., two elements $(x_1, e_1), (x_2, e_2) \in \mathcal{E}$ belong to the same equivalence class if and only if $(x_1, e_1) = (x_2, e_2)$ or else $(x_1, e_1), (x_2, e_2) \in \mathcal{V}$ and $(x_1, e_1) \sim (x_2, e_2)$. We continue to denote this equivalence relation on $\mathcal{E}$ by $\sim$.

We call $\Gamma := {}^{\mathcal{E}}\!\big/\!\sim$ a *metric graph* and $V := {}^{\mathcal{V}}\!\big/\!\sim$ its set of vertices.

Therefore, a metric graph is uniquely determined by a family $\{\ell_e\}_{e\in E}$ and an equivalence relation on $\mathcal{V}$. Its vertices are the cells of the partition of $\mathcal{V}$ induced by $\sim$. Two vertices $\mathbf{v}, \mathbf{w} \in V$ are said to be adjacent if there exists some $e \in E$ such that $\{x, y\} = \{0, \ell_e\}$ for some representatives (x, e) of $\mathbf{v}$ and (y, e) of $\mathbf{w}$; in this case we write $\mathbf{v} \sim \mathbf{w}$ and, with an abuse of notation, also $\mathbf{v} \sim e$. The cardinality $\deg(\mathbf{v})$ of the set $\{\mathbf{w} \in V : \mathbf{w}$ is adjacent to $\mathbf{v}\}$ is called degree of $\mathbf{v} \in V$. We denote $E_{\mathbf{v}} := \{e \in E : \mathbf{v} \sim e\}$ and $\mathrm{int}(V) := \{\mathbf{v} \in V : \#E_{\mathbf{v}} > 1\}$ (here $\#E_{\mathbf{v}}$ denotes the cardinality of $E_{\mathbf{v}}$). Moreover, Γ is a measure space with respect to the direct sum measure $\mu = \bigoplus_{e\in E} \lambda_e$ (see [29] and the references therein).

A function u on a metric graph Γ is a collection of functions $\{[u]_e\}_{e\in E}$ with $[u]_e :]0, \ell_e[\to \mathbb{R}$. If $[u]_e$ is integrable for all $e \in E$ then

$$\int_\Gamma u(x)\,\mathrm{d}x := \sum_{e\in E} \int_0^{\ell_e} [u]_e(x)\,\mathrm{d}x.$$

The Sobolev space $H^1(\Gamma)$ is defined as the space of functions u such that $[u]_e \in H^1(0, \ell_e)$ for all $e \in E$ and

$$\sum_{e\in E} \|[u]_e\|_{H^1(0,\ell_e)} < +\infty;$$

if $\mathbf{v} \sim e$ and $(0, e)$ (alternatively, (ℓ_e, e)) is a representative of $\mathbf{v}$, the trace of $[u]_e$ at 0 (alternatively, at ℓ_e) is denoted by $[u]_e(\mathbf{v})$.

The total variation of a measurable function u is defined as

$$\mathrm{TV}_\Gamma(u) := \sup\left\{\int_\Gamma u(x)z'(x)\,\mathrm{d}x \ : \ z \in X_k(\Gamma), \ \|z\|_{L^\infty(\Gamma)} \leq 1\right\},$$

where $X_k(\Gamma)$ is the set of vector fields in $H^1(\Gamma)$ satisfying a Kirchhoff condition on all the vertices of the graph; i.e.,

$$X_K(\Gamma) = \left\{z \in H^1(\Gamma) : \sum_{e\in E_\mathbf{v}} [z]_e(\mathbf{v})\nu^e(\mathbf{v}) = 0, \ \forall \mathbf{v} \in \mathrm{int}(V)\right\}.$$

Here, ν^e is the unit outer exterior normal to e; i.e., $\nu^e(\mathbf{v}) = 1$ if (ℓ_e, e) is a representative of $\mathbf{v}$ and $\nu^e(\mathbf{v}) = -1$ if $(0, e)$ is a representative of $\mathbf{v}$.

Note that, as in the previous subsections, weak lower semicontinuity holds by the definition of TV_Γ. In order to prove the coarea formula and the submodularity of the perimeter, we will give a representation formula for TV_Γ. We first introduce some notation. Given $\mathbf{v} \in \mathrm{int}(V)$, we consider the maximum and the minimum of the traces of u at $\mathbf{v}$:

$$u_\mathbf{v}^{\max} := \max\{[u]_e(\mathbf{v}) : e \in E_\mathbf{v}\}, \quad u_\mathbf{v}^{\min} := \min\{[u]_e(\mathbf{v}) : e \in E_\mathbf{v}\}.$$

For a function u on Γ such that each $[u]_e$ is a function of bounded variation, we denote $|Du|(\Gamma) := \sum_{e\in E} |D[u]_e|(0, \ell_e)$. With this notation, we obtain the following representation formula:

Proposition 2

$$\mathrm{TV}_\Gamma(u) = |Du|(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int} V(\Gamma)} (u_\mathbf{v}^{\max} - u_\mathbf{v}^{\min}). \tag{3.5}$$

Proof To prove (3.5) we use the following integration by parts formula [30]:

$$\int_\Gamma uz'\,dx = -(z, Du)(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} \sum_{e\in E_\mathbf{v}} [z]_e(\mathbf{v})\nu^e(\mathbf{v})[u]_e(\mathbf{v}), \tag{3.6}$$

with

$$(z, Du)(\Gamma) := \sum_{e\in E} ([z]_e, D[u]_e)(0, \ell_e);$$

$([z]_e, D[u]_e)$ being the one dimensional Anzellotti Radon measure product (see [31]). We recall that $|(z, Du)| \leq \|z\|_\infty |Du|$ as measures.

From (3.6) we directly have

$$\mathrm{TV}_\Gamma(u) \leq |Du|(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} \max\left\{\sum_{e\in E_\mathbf{v}} w_e[u]_e(\mathbf{v}) : w_e \in [-1, 1], \ \sum_{e\in E_\mathbf{v}} w_e = 0\right\}.$$

The last maximum is easily computed by the simplex method and we obtain

$$\mathrm{TV}_\Gamma(u) \leq |Du|(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} (u_{\mathbf{v}}^{\max} - u_{\mathbf{v}}^{\min}).$$

Let us see the opposite inequality. Since

$$|D[u]_e|(0,\ell_e) = \sup\left\{(z, Du)(0,\ell_e) = -\int_0^{\ell_e} [u]_e z' : z \in C_c^1(0,\ell_e)\right\},$$

we can find $z^n := \{[z^n]_e\}_{e\in E}$ with $[z^n]_e \in C_c^1(0,\ell_e)$ such that $\mathrm{dist}(\mathrm{supp}([z^n]_e), \{0,\ell_e\}) > \frac{1}{n}$ for all $e \in E_{\mathbf{v}}$, and

$$|Du|(\Gamma) \leq (z^n, Du)(\Gamma) + \frac{1}{n}.$$

We now fix $\mathbf{v} \in \mathrm{int}(V)$ and take $e_{\max} \in \mathrm{argmax}\{[u]_e(\mathbf{v}) : e \in E_{\mathbf{v}}\}$ and $e_{\min} \in \mathrm{argmin}\{[u]_e(\mathbf{v}) : e \in E_{\mathbf{v}}\}$ with $e_{\max} \neq e_{\min}$. Suppose for simplicity that $(0, e_{\min})$ and $(\ell_{e_{\max}}, e_{\max})$ are representatives of $\mathbf{v}$ (the other cases follow similarly). Then, we define

$$[\tilde{z}^n]_{e_{\max}} := \begin{cases} \dfrac{nx-1}{n} & \text{if } 0 \leq x \leq \frac{1}{n} \\ [z^n]_{e_{\max}} & \text{otherwise} \end{cases}, \quad [\tilde{z}^n]_{e_{\min}} := \begin{cases} \dfrac{1-nx}{n} & \text{if } 0 \leq x \leq \frac{1}{n} \\ [z^n]_{e_{\min}} & \text{otherwise} \end{cases}.$$

Then, repeating this with each $\mathbf{v} \in \mathrm{int}\,(V)$ (changing $(0, \frac{1}{n})$ with $(\ell_{e_{\min}} - \frac{1}{n}, \ell_{e_{\min}})$ or $(\ell_{e_{\max}} - \frac{1}{n}, \ell_{e_{\max}})$ when necessary) we end up with $\tilde{z}^n$ such that

$$\begin{aligned}
&(\tilde{z}^n, Du)(\Gamma) - \sum_{\mathbf{v}\in \mathrm{int}(V)} \sum_{e\in E_{\mathbf{v}}} [z]_e(\mathbf{v})\nu^e(\mathbf{v})[u]_e(\mathbf{v}) \\
&= (z^n, Du)(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} (u_{\mathbf{v}}^{\max} - u_{\mathbf{v}}^{\min}) \\
&\quad + \sum_{\mathbf{v}\in \mathrm{int}(V)} \left(\left([\tilde{z}^n]_{e_{\max}}, D[u]_{e_{\max}}\right)\left(0, \frac{1}{n}\right) + \left([\tilde{z}^n]_{e_{\min}}, D[u]_{e_{\min}}\right)\left(0, \frac{1}{n}\right)\right) \\
&\geq |Du|(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} (u_{\mathbf{v}}^{\max} - u_{\mathbf{v}}^{\min}) - \sum_{\mathbf{v}\in \mathrm{int}(V)} (|D[u]_{e_{\max}}| + |D[u]_{e_{\min}}|)\left(0, \frac{1}{n}\right) - \frac{1}{n}.
\end{aligned}$$

Letting $n \to +\infty$, with the use of (3.6),we conclude that

$$\mathrm{TV}_\Gamma(u) \geq |Du|(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} (u_{\mathbf{v}}^{\max} - u_{\mathbf{v}}^{\min}).$$

Therefore,

$$
\begin{aligned}
\mathrm{TV}_\Gamma(u) &= |Du|(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} (u_{\mathbf{v}}^{\max} - u_{\mathbf{v}}^{\min}) \\
&= |Du|(\Gamma) + \sum_{\mathbf{v}\in \mathrm{int}(V)} \max\left\{ \sum_{e\in E_{\mathbf{v}}} w_e[u]_e(\mathbf{v}) : w_e \in [-1,1],\ \sum_{e\in E_{\mathbf{v}}} w_e = 0 \right\}.
\end{aligned}
$$

□

We define $\mathrm{Per}_\Gamma(A) := \mathrm{TV}_\Gamma(\chi_A)$ for any μ-measurable $A \subset \Gamma$.
Here, given a μ-measurable set $A \subset \Gamma$ the function $\chi_A = \{[\chi_A]_e\}_{e\in E}$ on Γ is defined by

$$
[\chi_A]_e(x) := \begin{cases} 1 \text{ if } (x,e) \text{ is a representative of an element in } A, \\ 0 \text{ if } (x,e) \text{ is not a representative of any element in } A. \end{cases}
$$

Then, $(\Gamma, \mu, \mathrm{Per}_\Gamma)$ is a PMS. Indeed, observe that, since the classical total variation satisfies the coarea formula and the usual perimeter is submodular, and thanks to the last equality, we only need to show that:

(i) The trace at the endpoints of an interval of a one dimensional BV function satisfies the layer cake formula, in particular,

$$
[u]_e(\mathbf{v}) = \int_0^{+\infty} \left[\chi_{E_t(u)}\right]_e(\mathbf{v})dt - \int_0^{+\infty} \left[\chi_{\Gamma\setminus E_t(u)}\right]_e(\mathbf{v})dt
$$

for every $\mathbf{v} \in \mathrm{int}(V)$ and $e \in E_{\mathbf{v}}$ (here $\chi_{E_t(u)} = \left\{\chi_{E_t([u]_e)}\right\}_{e\in E}$).

(ii) The trace of the characteristic function of a set is submodular.

These two properties follow directly from the continuity of a precise representative of a one dimensional BV function outside of the jump set since, working edge by edge, we have that the precise representative of the extension of $[u]_e$ with constant value $[u]_e(\mathbf{v})$ to the right of ℓ_e (or left of 0) is continuous at the point ℓ_e (or 0) and the same thing happens with the precise representative of the extension of χ_A of any finite perimeter set A.

4 Linkage

In this section we prove the relation between the ROF and CV models in the framework of PMS. We need several previous results concerning the ROF model and a related minimization problem.

We consider the following energy functional defined on elements of the σ-algebra:

$$
\mathcal{E}_\lambda(\Sigma) := \mathrm{Per}(\Sigma) - \lambda \int_\Sigma f \,\mathrm{d}\nu\,, \quad \Sigma \in \mathcal{B} \tag{4.1}
$$

for $f \in L^2(X, \nu)$.

The following three results have been proved in the isotropic Euclidean case (Example 3.3.1) in [32] (Proposition 2.1, Lemma 2.4 and Lemma 2.5, respectively).

Proposition 3 *Given a minimizer $u \in L^2(X, \nu)$ of*

$$\min_{\substack{u \in L^2(X,\nu) \\ 0 \le u \le 1}} \mathrm{TV}(u) - \lambda \int_X uf \, \mathrm{d}\nu \tag{4.2}$$

and $t \in [0, 1)$, we have that $E_t(u) := \{x \in X \ : \ u(x) > t\}$ is a minimizer of (4.1).

Proof First of all, by the lower semi-continuity of TV with respect to the weak convergence in $L^2(X, \nu)$, the direct method of the calculus of variations yields the existence of a minimizer solving (4.2).

Now observe that for any $u \in L^2(X, \nu)$ with $0 \le u \le 1$ ν-a.e.,

$$\begin{aligned} \int_X uf \, \mathrm{d}\nu &= \int_X \left(\int_0^{u(x)} \mathrm{d}s \right) f(x) \, \mathrm{d}\nu(x) \\ &= \int_X \int_0^1 \chi_{E_s(u)}(x) f(x) \, \mathrm{d}s \, \mathrm{d}\nu(x) = \int_0^1 \int_{E_s(u)} f(x) \, \mathrm{d}\nu(x) \, ds. \end{aligned}$$

Therefore, it follows that

$$\mathrm{TV}(u) - \lambda \int_X uf \, \mathrm{d}\nu = \int_0^1 \mathcal{E}_\lambda \left(E_s(u) \right) \mathrm{d}s \ge \min_{\Sigma \in \mathcal{B}} \mathcal{E}_\lambda(\Sigma). \tag{4.3}$$

Consequently, the minimum in (4.2) is greater than or equal to $\min_{\Sigma \in \mathcal{B}} \mathcal{E}_\lambda(\Sigma)$. However, (4.2) is smaller than $\min_{\Sigma \in \mathcal{B}} \mathcal{E}_\lambda(\Sigma)$ since $\min_{\Sigma \in \mathcal{B}} \mathcal{E}_\lambda(\Sigma)$ is just (4.2) over characteristic functions of Borel sets. Therefore, they coincide.

Moreover, from (4.3) it follows that, if u is a minimizer of (4.2) then, for a.e. $t \in (0, 1)$, $E_t(u)$ is a minimizer of (4.1). Let us see that, in fact, this is true for every $t \in [0, 1)$.

Let $t \in [0, 1)$ and let $(t_n)_{n \ge 1}$ be a decreasing sequence such that $E_{t_n}(u)$ is a minimizer of (4.1) for every $n \ge 1$ and $t_n \downarrow t$ as $n \to \infty$. Then, since $\chi_{E_{t_n}(u)} \xrightarrow{n} \chi_{E_t(u)}$ in $L^2(X, \nu)$, from the lower semi-continuity of TV with respect to the weak convergence in $L^2(X, \nu)$, we infer that

$$\mathcal{E}_\lambda(E_t(u)) \le \liminf_{n \to \infty} \mathcal{E}_\lambda(E_{t_n}(u)) = \min_{\Sigma \in \mathcal{B}} \mathcal{E}_\lambda(\Sigma)$$

as desired. □

As a consequence of Proposition 3, for a fixed $\tau \in \mathbb{R}$, the superlevel sets $E_t(u)$, $t \in [0, 1)$, of a minimizer $u \in L^2(X, \nu)$ of

$$\min_{\substack{u \in L^2(X,\nu) \\ 0 \le u \le 1}} \mathrm{TV}(u) + \lambda \int_X u(\tau - f)\,\mathrm{d}\nu \tag{4.4}$$

are minimizers of

$$\mathcal{E}_{\lambda,\tau}(\Sigma) := \mathrm{Per}(\Sigma) + \lambda \int_\Sigma (\tau - f)\,\mathrm{d}\nu\,, \quad \Sigma \in \mathcal{B}.$$

Lemma 1 *Let $\tau_0 < \tau_1$. If Σ_{λ,τ_0} and $\Sigma_{\lambda,\tau_1} \in \mathcal{B}$ are minimizers of $\mathcal{E}_{\lambda,\tau_0}$ and $\mathcal{E}_{\lambda,\tau_1}$, respectively, then $\Sigma_{\lambda,\tau_1} \subseteq \Sigma_{\lambda,\tau_0}$ up to a ν-null set.*

Proof Since Σ_{λ,τ_0} and Σ_{λ,τ_1} are minimizers we have that

$$\mathrm{Per}(\Sigma_{\lambda,\tau_0}) + \lambda \int_{\Sigma_{\lambda,\tau_0}} (\tau_0 - f)\,\mathrm{d}\nu \le \mathrm{Per}(\Sigma_{\lambda,\tau_0} \cup \Sigma_{\lambda,\tau_1}) + \lambda \int_{\Sigma_{\lambda,\tau_0} \cup \Sigma_{\lambda,\tau_1}} (\tau_0 - f)\,\mathrm{d}\nu,$$
$$\mathrm{Per}(\Sigma_{\lambda,\tau_1}) + \lambda \int_{\Sigma_{\lambda,\tau_1}} (\tau_1 - f)\,\mathrm{d}\nu \le \mathrm{Per}(\Sigma_{\lambda,\tau_0} \cap \Sigma_{\lambda,\tau_1}) + \lambda \int_{\Sigma_{\lambda,\tau_0} \cap \Sigma_{\lambda,\tau_1}} (\tau_1 - f)\,\mathrm{d}\nu.$$

Adding these equations we get

$$\begin{aligned}
&\mathrm{Per}(\Sigma_{\lambda,\tau_0}) + \mathrm{Per}(\Sigma_{\lambda,\tau_1}) + \lambda \int_{\Sigma_{\lambda,\tau_0}} (\tau_0 - f)\,\mathrm{d}\nu + \lambda \int_{\Sigma_{\lambda,\tau_1}} (\tau_1 - f)\,\mathrm{d}\nu \\
&\le \mathrm{Per}(\Sigma_{\lambda,\tau_0} \cup \Sigma_{\lambda,\tau_1}) + \mathrm{Per}(\Sigma_{\lambda,\tau_0} \cap \Sigma_{\lambda,\tau_1}) \\
&\quad + \lambda \int_{\Sigma_{\lambda,\tau_0} \cup \Sigma_{\lambda,\tau_1}} (\tau_0 - f)\,\mathrm{d}\nu + \lambda \int_{\Sigma_{\lambda,\tau_0} \cap \Sigma_{\lambda,\tau_1}} (\tau_1 - f)\,\mathrm{d}\nu
\end{aligned}$$

which, by the sub-modularity of Per, yields

$$\begin{aligned}
&\int_{\Sigma_{\lambda,\tau_0}} (\tau_0 - f)\,\mathrm{d}\nu + \int_{\Sigma_{\lambda,\tau_1}} (\tau_1 - f)\,\mathrm{d}\nu \le \int_{\Sigma_{\lambda,\tau_0} \cup \Sigma_{\lambda,\tau_1}} (\tau_0 - f)\,\mathrm{d}\nu \\
&+ \int_{\Sigma_{\lambda,\tau_0} \cap \Sigma_{\lambda,\tau_1}} (\tau_1 - f)\,\mathrm{d}\nu,
\end{aligned}$$

i.e.,

$$\int_{\Sigma_{\lambda,\tau_1} \setminus \Sigma_{\lambda,\tau_0}} (\tau_1 - f)\,\mathrm{d}\nu \le \int_{\Sigma_{\lambda,\tau_1} \setminus \Sigma_{\lambda,\tau_0}} (\tau_0 - f)\,\mathrm{d}\nu,$$

thus $(\tau_1 - \tau_0)\nu\left(\Sigma_{\lambda,\tau_1} \setminus \Sigma_{\lambda,\tau_0}\right) \le 0$. □

Remark 1 It follows that the minimizer $\Sigma_{\lambda,\tau}$ of $\mathcal{E}_{\lambda,\tau}$ is unique up to a ν-null set, except for at most countably many values of τ. Indeed, for each $\tau \in \mathbb{R}$, fix a minimizer $\Sigma_{\lambda,\tau}$ of $\mathcal{E}_{\lambda,\tau}$. Then, $\tau \mapsto \nu(\Sigma_{\lambda,\tau})$ is monotone so it is continuous except for at most countably many values. Let us see that if τ is a continuity point then $\Sigma_{\lambda,\tau}$ is unique up to a ν-null set. Suppose otherwise that there exist two minimizers $\Sigma^1_{\lambda,\tau}$ and $\Sigma^2_{\lambda,\tau}$ such that $\nu\left(\Sigma^1_{\lambda,\tau} \triangle \Sigma^2_{\lambda,\tau}\right) > 0$. Then, $\Sigma^1_{\lambda,\tau} \cup \Sigma^2_{\lambda,\tau} \subseteq \Sigma_{\tau'}$ up to a ν-null set for any $\tau' < \tau$; and $\Sigma^1_{\lambda,\tau} \cap \Sigma^2_{\lambda,\tau} \supseteq \Sigma_{\tau^\star}$ up to a ν-null set for any $\tau^\star > \tau$. However, this implies that

$$\nu\left(\Sigma_{\tau'}\right) \geq \nu\left(\Sigma_{\tau^\star}\right) + \nu\left(\Sigma^1_{\lambda,\tau} \triangle \Sigma^2_{\lambda,\tau}\right)$$

so τ cannot be a continuity point.

Let $\mathfrak{Q}$ be a countable dense subset of $\mathbb{R}$ such that the minimizer $\Sigma_{\lambda,\tau}$ of $\mathcal{E}_{\lambda,\tau}$ is unique for every $\tau \in \mathfrak{Q}$.

Lemma 2 *Let $\Sigma_{\lambda,\tau}$ be a minimizer of $\mathcal{E}_{\lambda,\tau}$. Then, the function*

$$u_\lambda(x) := \sup\{\tau \in \mathfrak{Q} \,:\, x \in \Sigma_{\lambda,\tau}\}$$

is the unique minimizer of $\mathcal{E}^{ROF}_\lambda$ in $L^2(X,\nu)$.

Proof Let us first see that $u_\lambda \in L^2(X,\nu)$. Since

$$\mathrm{Per}(\Sigma_{\lambda,\tau}) + \lambda \int_{\Sigma_{\lambda,\tau}} (\tau - f)\,\mathrm{d}\nu \leq 0 = \mathcal{E}_{\lambda,\tau}(\emptyset)$$

we have that

$$\tau\nu(\Sigma_{\lambda,\tau}) \leq \int_{\Sigma_{\lambda,\tau}} f\,\mathrm{d}\nu.$$

Thus

$$\int_0^M \tau\nu(\Sigma_{\lambda,\tau})\mathrm{d}\tau \leq \int_0^M \int_{\Sigma_{\lambda,\tau}} f\,\mathrm{d}\nu\mathrm{d}\tau. \tag{4.5}$$

Now, by Fubini's theorem,

$$\begin{aligned}\int_0^M \tau\nu(\Sigma_{\lambda,\tau})\mathrm{d}\tau &= \int_0^M \int_{\Sigma_{\lambda,0}} \tau\chi_{\Sigma_{\lambda,\tau}}(x)\,\mathrm{d}\nu(x)\,\mathrm{d}\tau \\ &= \int_{\Sigma_{\lambda,0}} \int_0^M \tau\chi_{[0,u_\lambda(x)]}(\tau)\,\mathrm{d}\tau\,\mathrm{d}\nu(x)\end{aligned}$$

$$= \int_{\Sigma_{\lambda,0}} \int_0^{\min\{u_\lambda(x), M\}} \tau \, \mathrm{d}\tau \, \mathrm{d}\nu(x)$$
$$= \int_{\Sigma_{\lambda,0}} \frac{1}{2} \left(\min\{u_\lambda(x), M\}\right)^2 \, \mathrm{d}\nu(x)$$

and, similarly,

$$\int_0^M \int_{\Sigma_{\lambda,\tau}} f(x) \, \mathrm{d}\nu(x) \, \mathrm{d}\tau = \int_{\Sigma_{\lambda,0}} \min\{u_\lambda(x), M\} f(x) \, \mathrm{d}\nu(x).$$

Consequently, by (4.5),

$$\frac{1}{2} \int_{\Sigma_{\lambda,0}} (\min\{u_\lambda, M\})^2 \, \mathrm{d}\nu \leq \int_{\Sigma_{\lambda,0}} \min\{u_\lambda, M\} f \, \mathrm{d}\nu$$
$$\leq \left(\int_{\Sigma_{\lambda,0}} (\min\{u_\lambda, M\})^2 \mathrm{d}\nu \int_{\Sigma_{\lambda,0}} f^2 \mathrm{d}\nu \right)^{\frac{1}{2}}$$

thus

$$\int_{\Sigma_{\lambda,0}} (\min\{u_\lambda, M\})^2 \, \mathrm{d}\nu \leq 4 \int_{\Sigma_{\lambda,0}} f^2 \, \mathrm{d}\nu.$$

From this, letting $M \to \infty$, it follows that

$$\int_{\{u_\lambda > 0\}} u_\lambda^2 \, \mathrm{d}\nu \leq 4 \int_{\{u_\lambda \geq 0\}} f^2 \, \mathrm{d}\nu. \tag{4.6}$$

Similarly, we get that

$$\int_{\{u_\lambda < 0\}} u_\lambda^2 \, \mathrm{d}\nu \leq 4 \int_{\{u_\lambda \leq 0\}} f^2 \, \mathrm{d}\nu. \tag{4.7}$$

Indeed, since $\mathrm{Per}(E) = \mathrm{Per}(X \setminus E)$ for every $E \in \mathcal{B}$, we have that $\{-u_\lambda > -\tau\} = \{u_\lambda < \tau\}$ is a minimizer of

$$\min_{\Sigma} \mathrm{Per}(\Sigma) + \lambda \int_{\Sigma} (f - \tau) \, \mathrm{d}\nu = \min_{\Sigma} \mathrm{Per}(\Sigma) + \lambda \int_{\Sigma} (-\tau - (-f)) \, \mathrm{d}\nu.$$

Therefore, it follows that if we replace f with $-f$ in $\mathcal{E}_{\lambda,\tau}$ then we obtain $-u_\lambda$ instead of u_λ.

It follows from (4.6) and (4.7) that $u_\lambda \in L^2(X, \nu)$.

We now consider

$$\Sigma_{\lambda,\tau}^{-} := \bigcup_{\tau' > \tau,\, \tau' \in \mathfrak{Q}} \Sigma_{\lambda,\tau'}, \quad \tau \in \mathbb{R}.$$

Observe that $\Sigma^-_{\lambda,\tau} = E_\tau(u_\lambda)$ up to a ν-null set.

Let us prove that $\Sigma^-_{\lambda,\tau}$ is a minimizer of $\mathcal{E}_{\lambda,\tau}$. Let $\{\tau_n\}_{n\in\mathbb{N}} \subset \mathfrak{Q}$ be a decreasing sequence such that $\tau_n \downarrow \tau$ as $n \to \infty$. Since for any $\Sigma \in \mathcal{B}$ we have

$$\mathrm{Per}(\Sigma) + \lambda \int_\Sigma (\tau_n - f)\,\mathrm{d}\nu \geq \mathrm{Per}(\Sigma_{\lambda,\tau_n}) + \lambda \int_{\Sigma_{\lambda,\tau_n}} (\tau_n - f)\,\mathrm{d}\nu$$

and $\chi_{\Sigma_{\lambda,\tau_n}}$ is nonincreasing (thus $\chi_{\Sigma_{\lambda,\tau_n}} \xrightarrow{n} \chi_{\Sigma^-_{\lambda,\tau}}$ in $L^2(X,\nu)$), the lower semi-continuity of Per with respect to the weak convergence in $L^2(X,\nu)$ yields

$$\mathrm{Per}(\Sigma) + \lambda \int_\Sigma (\tau - f)\,\mathrm{d}\nu \geq \mathrm{Per}(\Sigma^-_{\lambda,\tau}) + \lambda \int_{\Sigma^-_{\lambda,\tau}} (\tau - f)\,\mathrm{d}\nu.$$

Therefore, $\Sigma^-_{\lambda,\tau}$ is a minimizer of $\mathcal{E}_{\lambda,\tau}$.

Now let $v \in L^2(X,\nu)$. Since $\Sigma^-_{\lambda,\tau}$ is a minimizer of $\mathcal{E}_{\lambda,\tau}$ we have that, for every $M > 0$,

$$\begin{aligned}
&\int_{-M}^{M} \left(\mathrm{Per}(\Sigma^-_{\lambda,\tau}) + \lambda \int_{\Sigma^-_{\lambda,\tau}} (\tau - f)\mathrm{d}\nu\right) \mathrm{d}\tau \\
&\quad \leq \int_{-M}^{M} \left(\mathrm{Per}(E_\tau(v)) + \lambda \int_{E_\tau(v)} (\tau - f)\mathrm{d}\nu\right) \mathrm{d}\tau. \qquad (4.8)
\end{aligned}$$

Now, for any $w \in L^2(X,\nu)$, by Fubini's theorem,

$$\begin{aligned}
&\int_{-M}^{M} \int_{E_\tau(w)} (\tau - f(x))\,\mathrm{d}\nu(x)\,\mathrm{d}\tau = \int_X \int_{-M}^{M} \chi_{E_\tau(w)}(x)(\tau - f(x))\,\mathrm{d}\tau\,\mathrm{d}\nu(x) \\
&= \int_X \int_{-M}^{\min\{w(x),M\}} (\tau - f(x))\,\mathrm{d}\tau\,\mathrm{d}\nu(x) \\
&= \frac{1}{2}\int_X ((\min\{w,M\} - f)^2 - (-M - f)^2)\,\mathrm{d}\nu.
\end{aligned}$$

Thus

$$\begin{aligned}
&\int_{-M}^{M} \int_{E_\tau(w)} (\tau - f(x))\,\mathrm{d}\nu(x)\mathrm{d}\tau + \frac{1}{2}\int_X (M + f)^2\,\mathrm{d}\nu \\
&= \frac{1}{2}\int_X (w - f)^2\,\mathrm{d}\nu + \Phi(w,M) \qquad (4.9)
\end{aligned}$$

where

$$\Phi(w,M) = \frac{1}{2}\int_X \left((\min\{w,M\} - f)^2 - (w - f)^2\right)\,\mathrm{d}\nu,$$

satisfies

$$\lim_{M\to\infty} \Phi(w, M) = 0$$

for any $w \in L^2(X, \nu)$. Therefore, from (4.8) and (4.9), it follows that

$$\int_{-M}^{M} \mathrm{Per}(E_\tau(u_\lambda))\, \mathrm{d}\tau + \frac{\lambda}{2}\int_X (u_\lambda - f)^2\, \mathrm{d}\nu + \Phi(u_\lambda, M)$$
$$\leq \int_{-M}^{M} \mathrm{Per}(E_\tau(v))\, \mathrm{d}\tau + \frac{\lambda}{2}\int_X (v - f)^2\, \mathrm{d}\nu + \Phi(v, M).$$

Then, letting $M \to \infty$, we get

$$\mathrm{TV}(u_\lambda) + \frac{\lambda}{2}\int_X (u_\lambda - f)^2\, \mathrm{d}\nu \leq \mathrm{TV}(v) + \frac{\lambda}{2}\int_X (v - f)^2\, \mathrm{d}\nu.$$

Consequently, since $v \in L^2(X, \nu)$ is arbitrary, u_λ is a minimizer for the ROF model. $\square$

Remark 2 Note that, in the case that $0 \leq f \leq 1$:

- If $\tau > 1$, then $\emptyset$ is the unique (up to ν-null sets) minimizer of $\mathcal{E}_{\lambda,\tau}$.
- If $\tau < 0$, then X is the unique (up to ν-null sets) minimizer of $\mathcal{E}_{\lambda,\tau}$.

Therefore, $0 \leq u_\lambda \leq 1$ and the proof of Lemma 2 becomes much simpler.

We have obtained the following result.

Proposition 4 *A function u solves the ROF model (3.3) if, and only if, $E_\tau(u)$ is a minimizer of $\mathcal{E}_{\lambda,\tau}$ for every $\tau \in \mathbb{R}$.*

As in [32, Proposition 2.6], we observe that the change of τ from almost any value to all value in $\mathbb{R}$ is achieved by approximation.

Next, we will show the relation between the ROF and CV models. We recall that an n-tuple is a partial minimizer of some functional defined on a product space $X_1 \times \cdots \times X_n$ if it minimizes the functional over each space X_i, $i = 1, \ldots, n$, when the other components are fixed.

Theorem 2 *Given a measurable function $f : X \to [0, 1]$, let u_λ be the unique minimizer of the ROF functional $\mathcal{E}_\lambda^{ROF}$. Let $0 < m_0 < m_1 \leq 1$ be such that $\widetilde{\Sigma} := E_{(m_0+m_1)/2}(u_\lambda)$ satisfies $0 < \nu(\widetilde{\Sigma}) < \nu(X)$. Then $\mathcal{E}_\mu^{CV}(\widetilde{\Sigma}, m_0, m_1) \leq \mathcal{E}_\mu^{CV}(\Sigma, m_0, m_1)$ for all $\Sigma \in \mathcal{B}$ and $\mu := \frac{\lambda}{2(m_1 - m_0)}$. In particular, $\big(\widetilde{\Sigma}, m_0, m_1\big)$ is a partial minimizer of the Chan–Vese model if*

$$m_0 := \frac{1}{\nu(X \setminus \tilde{\Sigma})}\int_{X\setminus\tilde{\Sigma}} f\, \mathrm{d}\nu \quad \textit{and} \quad m_1 := \frac{1}{\nu(\tilde{\Sigma})}\int_{\tilde{\Sigma}} f\, \mathrm{d}\nu.$$

Proof First of all, observe that, thanks to Proposition 4, $\widetilde{\Sigma}$ minimizes $\mathcal{E}_{\lambda, \frac{m_0+m_1}{2}}$. Then, it is obvious that it is also a minimizer of $\mathcal{E}_{\lambda, \frac{m_0+m_1}{2}} + C$ for any $C \in \mathbb{R}$; in particular for $C := \frac{\lambda}{2(m_1-m_0)} \int_X (m_0 - f)^2 \mathrm{d}\nu$. On the other hand, for any $\Sigma \in \mathcal{B}$,

$$\begin{aligned}
\mathcal{E}_{\lambda, \frac{m_0+m_1}{2}}(\Sigma) + C &= \mathrm{Per}(\Sigma) + \lambda \int_\Sigma \left(\frac{m_0 + m_1}{2} - f \right) \mathrm{d}\nu + C \\
&= \mathrm{Per}(\Sigma) + \frac{\lambda}{2(m_1 - m_0)} \int_\Sigma \left((m_1 - f)^2 - (m_0 - f)^2 \right) \mathrm{d}\nu + C \\
&= \mathrm{Per}(\Sigma) + \frac{\lambda}{2(m_1 - m_0)} \left(\int_\Sigma (m_1 - f)^2 \, \mathrm{d}\nu + \int_{X \setminus \Sigma} (m_0 - f)^2 \, \mathrm{d}\nu \right) \\
&= \mathcal{E}_\mu^{CV}(\Sigma, m_0, m_1).
\end{aligned}$$

Thus $\mathcal{E}_\mu^{CV}(\widetilde{\Sigma}, m_0, m_1) \leq \mathcal{E}_\mu^{CV}(\Sigma, m_0, m_1)$ for any $\Sigma \in \mathcal{B}$.

Now let

$$\widetilde{m}_0 = \frac{1}{\nu(X \setminus \widetilde{\Sigma})} \int_{X \setminus \widetilde{\Sigma}} f \mathrm{d}\nu \quad \text{and} \quad \widetilde{m}_1 := \frac{1}{\nu(\widetilde{\Sigma})} \int_{\widetilde{\Sigma}} f \mathrm{d}\nu.$$

It is straightforward to check that they minimize $\mathcal{E}_\mu^{CV}(\widetilde{\Sigma}, \cdot, \cdot)$. To finish, we must prove that $\widetilde{m}_0 < \widetilde{m}_1$.

Since $\mathcal{E}_{\lambda, \frac{\widetilde{m}_0+\widetilde{m}_1}{2}}(\widetilde{\Sigma}) \leq \mathcal{E}_{\lambda, \frac{\widetilde{m}_0+\widetilde{m}_1}{2}}(\emptyset) = 0$, we have that $\int_{\widetilde{\Sigma}} \left(\frac{\widetilde{m}_0+\widetilde{m}_1}{2} - f \right) \mathrm{d}\nu \leq 0$. Thus $\frac{\widetilde{m}_0+\widetilde{m}_1}{2} \leq \widetilde{m}_1$. If $\widetilde{m}_0 = \widetilde{m}_1$, then $\emptyset$ is a minimizer of $\mathcal{E}_{\lambda, \frac{\widetilde{m}_0+\widetilde{m}_1}{2}}(\cdot)$ thus, by uniqueness (up to a ν-null set) of the minimizer, we get $\nu(\widetilde{\Sigma}) = 0$, a contradiction. Therefore, $\widetilde{m}_0 < \widetilde{m}_1$. □

5 Applications

In this section we see how our previous results allow us to use the classical CV model for image segmentation in settings different than those for which it was originally designed. We consider a locally finite weighted discrete graph $G = (V(G), E(G))$, where the edge weights $\{w_{xy}\}_{x,y \in V(G)}$ are given by the Gaussian kernel $\eta_\sigma(x, y) = \exp(-d(x, y)/\sigma^2)$ for some distance d defined on $V(G)$. We recall the definition of the CV and ROF functionals in a discrete graph (see Example 3.3.4):

$$\mathcal{E}_\mu^{\mathrm{CV}}(\Sigma, m_0, m_1) = \mathrm{Per}_G(\Sigma) + \mu \sum_{x \in \Sigma} (m_0 - f(x))^2 d_x + \mu \sum_{x \in V(G) \setminus \Sigma} (m_1 - f(x))^2 d_x,$$

$$\mathcal{E}_\lambda^{\mathrm{ROF}}(u) = \frac{1}{2} \sum_{x,y \in V(G)} |u(x) - u(y)| w_{xy} + \frac{\lambda}{2} \sum_{x \in V(G)} (u(x) - f(x))^2 d_x$$

where $f : V(G) \to [0, 1]$, $\Sigma \subseteq V(G)$, $u \in L^2(V(G), \nu_G)$, $m_0, m_1 > 0$,

$$d_x := \sum_{y \sim x} w_{xy} \quad \text{and} \quad \mathrm{Per}_G(A) = \sum_{x \in A} \sum_{y \in V(G) \setminus A} w_{xy} \quad \text{for all} A \subseteq V(G).$$

Thanks to the results in Sect. 4, we know that it suffices to minimize the ROF problem to obtain a partial minimizer of $\mathcal{E}_\mu^{\mathrm{CV}}$. To compute the approximate minimizer of $\mathcal{E}_\lambda^{\mathrm{ROF}}$, we will use the source code provided in [6], based on the IRLS method [33]. In our experiments, we will specify the parameters ε, σ and K used in this method, which we will denote hereafter by $\mathrm{IRLS}_{(\varepsilon,\sigma,K)}$. We recall that σ is the parameter of the kernel η_σ and that K is the number of active edges per vertex (i.e., with $w_{xy} \neq 0$), a choice that will be determined by the K nearest neighbours to each node. That said, we present the following strategy based on Theorem 2 to estimate a $\mathcal{E}_\mu^{\mathrm{CV}}$ minimizer:

Algorithm 1 $\mathcal{E}_\mu^{\mathrm{CV}}$ approximate minimizer

initiation: $k = 1,\ \mu > 0,\ (m_0, m_1) \in [0, 1]^2$ s.t. $m_1 > m_0$.

$w \leftarrow \mathcal{E}^{\mathrm{ROF}}_{2\mu(m_1 - m_0)}$ minimizer using $\mathrm{IRLS}_{(\varepsilon,\sigma,K)}$.

$\Sigma_0 \leftarrow \emptyset$.

$\Sigma_1 \leftarrow \{x \in V \ : \ w(x) > \frac{1}{2}(m_1 - m_0)\}$.

while $|\Sigma_k \triangle \Sigma_{k-1}|^2 > \varepsilon_{tol} \ \wedge \ k < n_{max}$ **do**

$m_0 \leftarrow |\Sigma_k|^{-1} \sum_{x \in \Sigma_k} f(x),\ \ m_1 \leftarrow |V \setminus \Sigma_k|^{-1} \sum_{x \in V \setminus \Sigma_k} f(x)$.

$w \leftarrow \mathcal{E}^{\mathrm{ROF}}_{2\mu(m_1 - m_0)}$ minimizer using $\mathrm{IRLS}_{(\varepsilon,\sigma,K)}$.

$\Sigma_{k+1} \leftarrow \{x \in V \ : \ w(x) > \frac{1}{2}(m_1 - m_0)\},\ \ k \leftarrow k + 1$.

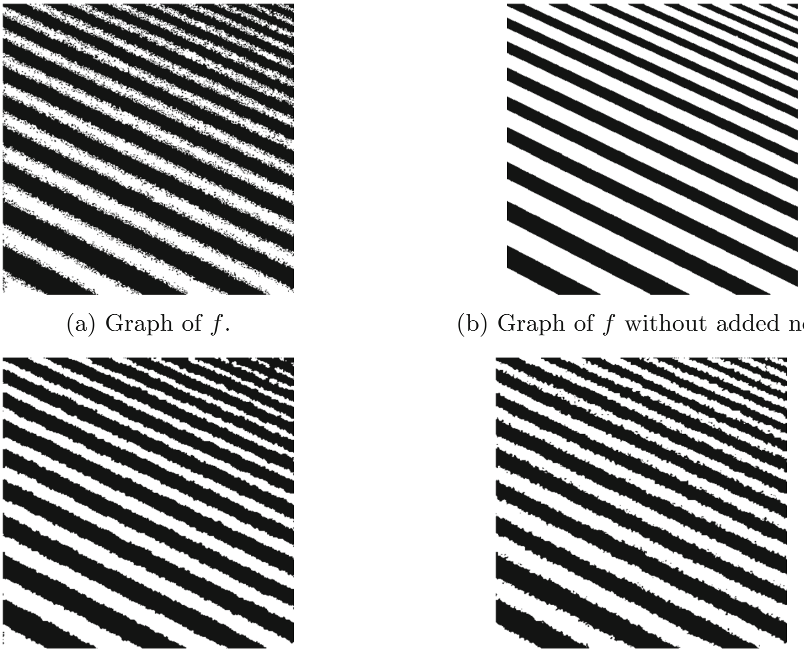
end while

return $\boldsymbol{\Sigma}_k, w$

where $n_{\max}$ and ε_{tol} denote the maximum number of iterations and the tolerance of the algorithm, respectively. In our experiments, we set $n_{\max} = \varepsilon_{tol} = 10$ and we take d as the Euclidean distance. We note that the above scheme is equivalent to those presented in [3, 5], proposed for the classical CV model in its isotropic and anisotropic forms, respectively.

5.1 Nonlocal Image Segmentation

As a first example, we define $f : \Omega \to \{0, 1\}$ such that the image represented by f consists of diagonal lines with added noise. The noise fulfills the condition that the closer it is to a line, the denser it is (cf. Fig. 1a, b). Suppose that $\Omega := \{1, ..., n\} \times \{1, ..., m\}$, where (n, m) is the image size. Taking into account the particular features



(a) Graph of f.

(b) Graph of f without added noise.

(c) [3]'s approach with $\mu = 1.2$

(d) Our approach with $\mu = 3$ on each V_i

Fig. 1 Comparison of approximate solutions to the CV model

of this image, we will divide Ω into disjoint subsets $\{V_i\}_{i=1}^{10}$ as follows:

$$V_i := \begin{cases} \left\{x \in \Omega : \bar{f}(x) \le D_i\left(\bar{f}(\Omega)\right)\right\} & \text{if } i = 1, \\ \left\{x \in \Omega \setminus V_{i-1} : \bar{f}(x) \le D_i\left(\bar{f}(\Omega)\right)\right\} & \text{if } i > 1, \end{cases}$$
$$\text{s.t.} \quad \bar{f}(x) := \sum_{y \in B_1(x)} \frac{f(y)}{|B_1(x)|},$$

where $B_1(x)$ is the ℓ^∞ unit ball in Ω centered at x, and $D_i(\bar{f}(\Omega))$ is the i-th decile of the finite set $\bar{f}(\Omega) \subset \mathbb{R}$. Then, we will minimize the CV functional on each $\left(V_i, \nu_{G_i}, \mathrm{Per}_{G_i}\right)$ (here $G_i = (V_i(G_i), E(G_i))$). With this procedure, the segmentation process is carried out independently on each element of the partition, which is constructed taking into account the density of 1's in the neighborhood of each node of Ω.

We apply the Algorithm 1 with $\mathrm{IRLS}_{(3,0.3,10)}$ and compare our approach with the classical Euclidean CV segmentation. We computed the latter by making use of the algorithm proposed in [3], conceptually equivalent to ours. In Fig. 1c and d, we present the segmentations that differ the least, in quadratic error terms, from the original denoised image shown in Fig. 1b. In these examples, we can see that these approaches produce noticeably different segmentations, especially where the density of the noise is highest. In fact, if we compute the Frobenius norm of the differences between the denoising results and Fig. 1b, we see that our approach, in this case, is 13.27% more accurate than the classical one. This is due to the way in which the segmentation process is carried out. The classical approach handles the image as a whole, solving both problems from a global viewpoint. Our approach, however, solves both problems in specific regions of the image, which have been previously identified through a particular criterion (the noise density around each pixel). Therefore, our proposal provides higher flexibility in the segmentation or denoising process.

5.2 Labeling

Additionally, extending the ROF and CV models to perimeter measure spaces allows us to use these models in other fields beyond image processing. In the next example, we show an application of these models in a multiclass classification problem in higher-dimensional spaces: the *labeling problem*.

We consider again the PMS $(V(G), \nu_G, \mathrm{Per}_G)$ associated to a weighted graph. The labeling problem consists in assigning a label from the label set $\{y_i\}_{i=1}^k$ to each vertex in $V(G)$. As the initial condition, we assume that labels have already been assigned to each of the vertices in $\{x_i\}_{i=1}^m$ for a given $m \le k$, i.e., $y_{j_i} \in \{y_j\}_{j=1}^k$ is already assigned to x_i for $1 \le i \le m$. One of the methods most commonly used to solve this problem is the so-called Laplacian learning algorithm. It finds an approximate solution as follows: First, an approximate solution $u = (u_i)_{i=1}^k : V(G) \to [0,1]^k$ of the following PDE on the graph is found:

$$\begin{cases} \mathcal{L}u(x_i) = 0 & \text{if } i > m, \\ u(x_i) = e_{j_i} & \text{otherwise}, \end{cases} \tag{5.1}$$

where $\{e_i\}_{i=1}^k$ is the canonical basis of $\mathbb{R}^k$ and $\mathcal{L}$ is the graph Laplacian functional defined as

$$\mathcal{L}u(x) := \sum_{y \in V} w_{xy}(u(x) - u(y)) \quad \text{for } x \in V.$$

Then, a label is assigned to each $x \in V(G)\backslash\{x_i\}_{i=1}^m$ corresponding to the largest component of $u(x)$. Unfortunately, this labeling method can be inconsistent and is largely influenced by the number of vertices per label initially known [34, 35].

Alternatively, we will use the CV model to improve the results. First, we obtain an approximate solution u of (5.1) by using the well-known algorithm proposed in [36]. However, we then follow a different approach to assign the labels, first using the CV model on a modified version of each u_i, $1 \le i \le k$. To begin with, let

$$\tilde{u} = (\tilde{u}_i)_{i=1}^k \; : \; \tilde{u}_i(x) := u_i(x) - w_i, \quad x \in V(G),$$

where w_i is the mean of $\{u_i(x)\}_{x \in V(G)}$; and then normalize between 0 and 1 (without renaming) the values of $\{\tilde{u}_i(x)\}_{1 \le i \le k}$ for each $x \in V(G)$. This first step balances out the prevalence of the different labels. Then, for each $1 \le i \le k$, we use Algorithm 1

Table 1 Average accuracy for the MNIST dataset

n_k	1	2	3	4	5
PoissonMBO learning	96.1 (2.5)	97.0 (0.2)	97.2 (0.1)	97.2 (0,1)	97.2 (0.1)
Laplacian learning	16.8 (6.1)	27.5 (7.8)	41.1 (12.4)	57.9 (11.5)	69.1 (12.3)
CV post-processing	83.3 (4.7)	94.2 (2.6)	95.5 (0.3)	96.1 (0.3)	96.3 (0.3)

Table 2 Average accuracy for the FashionMNIST dataset

n_k	1	2	4	8	16
PoissonMBO learning	61.9 (5.8)	67.0 (4.9)	72.2 (2.5)	74.2 (2.1)	76.9 (1.8)
Laplacian learning	17.4 (7.2)	32.0 (8.1)	51.6 (5.9)	69.7 (3.3)	74.6 (1.3)
CV post-processing	64.3 (6.7)	65.4 (5.8)	69.5 (3.3)	72.1 (1.7)	76.3 (1.3)

Table 3 Average accuracy for the CIFAR-10 dataset

n_k	10	20	40	80	160
PoissonMBO learning	61.6 (2.3)	64.6 (1.7)	66.7 (0.7)	68.5 (0.6)	70.4 (0.5)
Laplacian learning	21.5 (7.3)	38.2 (7.9)	55.0 (4.7)	62.8 (1.8)	66.5 (1.2)
CV post-processing	57.8 (3.3)	62.0 (2.1)	66.2 (1.5)	68.3 (0.9)	70.5 (0.5)

with $f = \tilde{u}_i$ and denote the output by $(\Sigma_i, \tilde{w}_i)$. Finally, we define $v := (\tilde{w}_i \chi_{\Sigma_i})_{i=1}^k : V \to \mathbb{R}^k$ and assign a label to $x \in V(G)\backslash\{x_i\}_{i=1}^m$ according to the largest component of $v(x)$.

In order to evaluate the effectiveness of this post-processing we will compare it with the PoissonMBO learning [37], which, to the best of our knowledge, is one of the best ways to solve the labeling problem (see, for example, [37, Sect. 4]). To compare the accuracy, we will use three databases composed by $k = 10$ categories: MNIST [38], FashionMNIST [39] (in both cases, $n = 60{,}000$ and $V \subset \{0.1\}^{784}$) and CIFAR-10 [40] ($n = 70000$ and $V \subset \{0, ..., 255\}^{3072}$). We perform experiments with varying number n_k of initially labeled elements per label. To compute the approximate solutions of the Laplacian (u according to the previous notation) and PoissonMBO learnings, we use the source code provided in [37]; while $v = (\tilde{w}_i \chi_{\Sigma_i})_{i=1}^{10}$ is computed by minimizing $\mathcal{E}_5^{\text{cv}}$ (via Algorithm 1 with IRLS$_{(3,1,4)}$). Note that, to reduce the computational cost of obtaining v, we have split the set of vertices $V(G)$ into 30 disjoint subsets $\{V_i\}_{i=1}^{30}$ of equal size and used the CV model on each of the V_i with $f = \tilde{u}|_{V_i}$.

The results are shown in Tables 1, 2 and 3; where the same scheme as the one in [37] is used. To be more precise, we show the average accuracy for each of the approaches over 100 trials together with the standard deviation inside the brackets.

According to the above tables, the application of the CV post-processing to the Laplacian approximation provides a considerable increase in accuracy in the labeling process. This increase is significant in all three datasets, specially in the cases where the number n_k is small. Moreover, the accuracy obtained is comparable to that of the

PoissonMBO learning. Consequently, we see that our CV model proposal can improve existing processes in settings different from image segmentation, such as the classical Laplacian learning algorithm for labeling.

Acknowledgements The authors would like to thank the anonymous referee for the thorough review of the manuscript, which contributed to its improvement.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The authors have been partially supported by Conselleria d'Innovació, Universitats, Ciència i Societat Digital, project AICO/2021/223, by Grant PID2022-136589NB-I00 funded by MCIN/AEI/10.13039/501100011033 and FEDER "A way to make Europe" and by Grant RED2022-134784-T funded by MCIN/AEI/10.13039/501100011033. The second author has also been supported by the Agència Valenciana de la Innovació, programme "Promoció del talent" (Innodocto, Ref. INNTA3/2022/16). The third author has also been supported by the Ministerio de Universidades (Spain) and NextGenerationUE, programme "Recualificación del sistema universitario español" (Margarita Salas), Ref. UP2021-044; the "Conselleria d'Innovació, Universitats, Ciència i Societat Digital", programme "Subvenciones para la contratación de personal investigador en fase postdoctoral" (APOSTD 2022), Ref. CIAPOS/2021/28; the MICINN (Spain) and FEDER, project PID2021-124195NB-C32; and by the ERC Advanced Grant 834728.

Declarations

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. **42**(5), 577–685 (1989)
2. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001)
3. Cai, X., Chan, R., Schönlieb, C., Steidl, G., Zeng, T.: Linkage between piecewise constant Mumford-Shah Model and Rudin-Osher-Fatemi model and its virtue in image segmentation. SIAM J. Sci. Comput. **41**(6), 1310–1340 (2019)
4. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)
5. Moll, S., Pallardó-Julià, V.: Anisotropic Chan-Vese segmentation. Nonlinear Anal. Real World Appl. **73**, 103908 (2023)
6. Caroccia, M., Chambolle, A., Slepčev, D.: Mumford-Shah functionals on graphs and their asymptotics. Nonlinearity **33**(8), 3846–3888 (2020)
7. Chambolle, A., Morini, M., Ponsiglione, M.: Nonlocal curvature flows. Arch. Ration. Mech. Anal. **218**(3), 1263–1329 (2015)
8. Franceschi, V., Pinamonti, A., Saracco, G., Stefani, G.: The Cheeger problem in abstract measure spaces. J. Lond. Math. Soc. **109**(1), 12840 (2024)
9. Hu, H., Laurent, T., Porter, M.A., Bertozzi, A.L.: A method based on total variation for network modularity optimization using the MBO scheme. SIAM J. Appl. Math. **73**(6), 2224–2246 (2013)

10. Chambolle, A., Giacomini, A., Lussardi, L.: Continuous limits of discrete perimeters. ESAIM Math. Model. Numer. Anal. **44**(2), 207–230 (2010)
11. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York (2000)
12. Amar, M., Bellettini, G.: A notion of total variation depending on a metric with discontinuous coefficients. Ann. Inst. H. Poincaré C Anal. Non Linéaire **11**(1), 91–133 (1994)
13. De Philippis, G., Maggi, F.: Regularity of free boundaries in anisotropic capillarity problems and the validity of Young's law. Arch. Ration. Mech. Anal. **216**(2), 473–568 (2015)
14. Esedoglu, S., Osher, S.J.: Decomposition of images by the anisotropic Rudin-Osher-Fatemi model. Commun. Pure Appl. Math. **57**(12), 1609–1626 (2004)
15. Moll, J.S.: The anisotropic total variation flow. Math. Ann. **332**(1), 177–218 (2005)
16. Lombardini, L.: Approximation of sets of finite fractional perimeter by smooth sets and comparison of local and global s-minimal surfaces. Interfaces Free Bound. **20**(2), 261–296 (2018)
17. Ambrosio, L., De Philippis, G., Martinazzi, L.: Gamma-convergence of nonlocal perimeter functionals. Manuscripta Math. **134**(3–4), 377–403 (2011)
18. Cesaroni, A., Novaga, M.: Periodic partitions with minimal perimeter. Nonlinear Anal. **243**, 113522 (2024)
19. Antil, H., Díaz, H., Jing, T., Schikorra, A.: Nonlocal bounded variations with applications. SIAM J. Math. Anal. **56**(2), 1903–1935 (2024)
20. Caffarelli, L., Roquejoffre, J.-M., Savin, O.: Nonlocal minimal surfaces. Commun. Pure Appl. Math. **63**(9), 1111–1144 (2010)
21. Novaga, M., Onoue, F.: Local Hölder regularity of minimizers for nonlocal variational problems. Commun. Contemp. Math. **25**(10), 2250058 (2023)
22. Bessas, K.: Fractional total variation denoising model with L^1 fidelity. Nonlinear Anal. **222**, 112926-20 (2022)
23. Bessas, K., Stefani, G.: Non-local BV functions and a denoising model with L^1 fidelity. Adv. Calc. Var. (2024). https://doi.org/10.48550/arXiv.2210.11958
24. Mazón, J.M., Solera, M., Toledo, J.: The total variation flow in metric random walk spaces. Calc. Var. Partial Differ. Equ. **59**(1), 29–64 (2020)
25. Mazón, J.M., Solera-Diana, M., Toledo-Melero, J.: Variational and Diffusion Problems in Random Walk Spaces. Progress in Nonlinear Differential Equations and Their Applications, vol. 103. Birkhäuser, Basel (2023)
26. Mazón, J.M., Solera, M., Toledo, J.: (BV, L^p)-decomposition, $p = 1, 2$, of functions in metric random walk spaces. Adv. Calc. Var. **15**(3), 515–550 (2022)
27. Miranda, M.: Functions of bounded variation on "good" metric spaces. J. Math. Pures Appl. **82**(8), 975–1004 (2003)
28. Górny, W., Mazón, J.M.: The Neumann and Dirichlet problems for the total variation flow in metric measure spaces. Adv. Calc. Var. (2022). https://doi.org/10.48550/arXiv.2105.11424
29. Mugnolo, D.: What is actually a metric graph? (2021) arXiv:1912.07549
30. Mazón, J.M.: The total variation flow in metric graphs. Math. Eng. **5**(1), 1–38 (2023)
31. Anzellotti, G.: Pairings between measures and bounded functions and compensated compactness. Ann. Mat. Pura Appl. **4**(135), 293-3181984 (1983)
32. Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: An Introduction to Total Variation for Image Analysis, pp. 263–340. De Gruyter, Berlin (2010)
33. Geman, D., Reynolds, G.: Constrained restoration and the recovery of discontinuities. IEEE Trans. Pattern Anal. Mach. Intell. **14**(3), 367–383 (1992)
34. Calder, J., Slepčev, D., Thorpe, M.: Rates of convergence for Laplacian semi-supervised learning with low labeling rates. Res. Math. Sci. (2023). https://doi.org/10.1007/s40687-022-00371-x
35. Nadler, B., Srebro, N., Zhou, X.: Semi-supervised learning with the graph Laplacian: the limit of infinite unlabelled data. Adv. Neural Inf. Process. Syst. **22**, 1330–1338 (2009)
36. Zhu, X., Ghahramani, Z., Lafferty, J.D.: Semi-supervised learning using gaussian fields and harmonic functions. In: Proceedings of the 20th International Conference on Machine Learning (ICML-03), pp. 912–919 (2003)
37. Calder, J., Cook, B., Thorpe, M., Slepčev, D.: Poisson learning: graph based semi-supervised learning at very low label rates. In: International Conference on Machine Learning, pp. 1306–1316 (2020)
38. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)

39. Xiao, H., Rasul, K., Vollgraf, R.: Fashion–MNIST: a novel image dataset for benchmarking machine learning algorithms. arXiv preprint arXiv:1708.07747 (2017)
40. Krizhevsky, A., Hinton, G.: Learning multiple layers of features from tiny images. Technical Report, University of Toronto, Toronto, Ontario (2009)