

Strictly convex realization in two-phase image segmentation

Yunho Kim*

March 28, 2013

Abstract

In this paper, we borrow an idea from already well-known in some geometric problems (e.g., [1], [10]) to find a global minimizer related to the two-phase Mumford-Shah functional via strictly convex formulations. We will briefly review how it applies to find the true minimizers of the original non-convex problem. More importantly, we will clearly address the possibility of a solution's non-uniqueness for non-strict convex problems mentioned but unresolved in [13], which is any non-strict convex problem's issue, using our formulations. The same answer applies to [14], [3] as well. Hence, one should consider this strictly convex formulation where the well-known convex formulations [14], [3] come into play.

1 Introduction

Image segmentation has been an important research topic in computer vision that aims at segmenting a given image into meaningful smaller pieces. The two of the most successful models are a variational model called the Mumford-Shah model and a PDE model called the Perona-Malik equation. There is an extensive collection of works in the literature investigating their analytical aspects as well as the numerical aspect. Moreover, there are many variants of those models fitted into specific problems. In this paper, we will take a variational approach related to the Mumford-Shah model. The model is to minimize the following functional

$$MS(u, K) = \int_{\Omega \setminus K} |\nabla u(x)|^2 dx + \alpha \int_{\Omega} |f(x) - u(x)|^2 dx + \beta \mathcal{H}^1(K) \quad (1)$$

over a piecewise smooth function u and a rectifiable curve K , which is apparently nonlinear and non-convex. Despite its nonlinearity and non-convexity, there were many successful attempts in analyzing the problem both analytically and numerically using various mathematical techniques such as the Γ -convergence, etc. Among the possible applications of the model, we pay attention to the two-phase, piecewise constant Mumford-Shah functional given a function f on Ω . The functional (1) becomes

$$\begin{aligned} MS(\Sigma, c_1, c_2) &= P(\Sigma; \Omega) + \lambda \int_{\Sigma} (c_1 - f(x))^2 dx + \lambda \int_{\Omega \setminus \Sigma} (c_2 - f(x))^2 dx \\ &= P(\Sigma; \Omega) + \lambda \int_{\Sigma} \left[(c_1 - f(x))^2 - (c_2 - f(x))^2 \right] dx + \lambda \int_{\Omega} (c_2 - f(x))^2 dx, \end{aligned}$$

*Department of Diagnostic Radiology, School of Medicine, Yale University. Email: yunho.kim@yale.edu

where u and K from (1) are identified with $c_1 1_\Sigma + c_2 1_{\Omega \setminus \Sigma}$ and $\partial \Sigma$, respectively. All the notations being used will be explained in the next section. Fixing Σ , the functional $MS(\Sigma, c_1, c_2)$ as a function of (c_1, c_2) is strictly convex and is easy to deal with. Indeed, the pair

$$\left(c_1 = \frac{1}{|\Sigma|} \int_{\Sigma} f(x) dx, c_2 = \frac{1}{|\Omega \setminus \Sigma|} \int_{\Omega \setminus \Sigma} f(x) dx \right)$$

minimizes the functional given Σ . Hence, one idea of solving the nonlinear and non-convex problem

$$\min_{\substack{c_1, c_2 \in \mathbb{R} \\ \Sigma \subset \Omega}} MS(\Sigma, c_1, c_2) \quad (2)$$

is to alternate between

$$\min_{\Sigma \subset \Omega} MS(\Sigma, c_1, c_2) \quad (3)$$

and

$$\min_{c_1, c_2 \in \mathbb{R}} MS(\Sigma, c_1, c_2), \quad (4)$$

which, by the way, does not guarantee the convergence of the alternating minimization to (2). Even if it does, we need to solve the non-convex problem (3). A good source of such an approach for our presentation is “active contours without edges” [15] by T. Chan and L. Vese, where the authors proposed to solve (3) by a level set method that made it possible to evolve a curve that is a boundary of a region without curve parametrization unlike its predecessors [5], [9], [22], [23]. However, the formulation in [15] is still non-convex, which suffers from local minima. Then, a work by Chan et al., [14], provided a different way to solve this non-convex problem (3) via a convex problem that resulted in finding a global minimum. The observation made in [14] was that once c_1, c_2 are fixed, a solution of

$$\min_{0 \leq \phi \leq 1} \left\{ F(\phi) = \mathcal{J}(\phi) + \lambda \int_{\Omega} \left[(c_1 - f(x))^2 - (c_2 - f(x))^2 \right] \phi(x) dx \right\}, \quad (5)$$

can provide a solution of (3), which is summarized in the following theorem.

Theorem 1. (Theorem 2, [14]) *For any given fixed $c_1, c_2 \in \mathbb{R}$, a global minimizer for $MS(\cdot, c_1, c_2)$ can be found by carrying out the convex minimization (5) and then setting*

$$\Sigma = \{x \in \Omega : \phi^*(x) \geq \mu\}$$

for a.e. $\mu \in [0, 1]$ with a minimizer ϕ^* of (5).

And the algorithm in [14] to solve this constrained problem (5) was derived from the following proposition.

Proposition 1. (Claim 1, [14]) *Let $f \in L^\infty(\Omega)$. Then, (5) is equivalent to the following unconstrained convex minimization problem:*

$$\min_{\phi} \left\{ \mathcal{J}(\phi) + \int_{\Omega} \left[\alpha \nu(\phi(x)) + \lambda s(x) \phi(x) \right] dx \right\}$$

where

$$\nu(\xi) = \max\{0, 2|\xi - 0.5| - 1\} \quad \text{and} \quad s(x) = (c_1 - f(x))^2 - (c_2 - f(x))^2$$

provided that $\alpha > \frac{\lambda}{2} \|s\|_{L^\infty(\Omega)}$.

The gradient descent method with a regularized version ν_{ϵ_2} of ν was used in [14] and a fast alternating minimization algorithm was reported in [3]. In fact, [3] proposed a more general framework than [14] by considering a more general regularizer \mathcal{J}_g than \mathcal{J} , where the function g called an edge indicator is a positive continuous function. Nevertheless, it is still not known if there is a unique minimizer of (5) simply because it is not strictly convex, which makes it difficult to characterize the minimizers that algorithms to solve (5) can provide.

Our strictly-convex problems in various forms will turn out to be an application to image segmentation of the ROF model [25] that was originally proposed for the image denoising task: given a noisy image $f \in L^2(\Omega)$, recover a clean image u^* that is a solution of

$$\min_u \left\{ \mathcal{J}(u) + \lambda \int_{\Omega} (f(x) - u(x))^2 dx \right\}. \quad (6)$$

The existence and the uniqueness of a solution can be easily obtained and more interesting theoretical and computational results have been found. For instance, the characterization of the solution using the subdifferential of a convex functional was provided in [28] and the regularity of the solution was reported in [7], [8]. In the numerical point of view, there are many efficient algorithms to solve (6) such as a relatively old one in [11] by Chambolle and a relatively new one in [12]. Thanks to the vast amount of research works related to the ROF model, our proposed strictly-convex problems will benefit from the same properties as the ROF model.

The outline of the paper is as follows. In Section 2, we will provide all the the notations and the mathematical background necessary to understand the idea. In Section 3, we will present a review on the idea that we borrowed from some geometric problems and apply the strictly convex formulation to solve (3) and to provide an example where the uniqueness of a solution of (3) fails. In Section 4, we will present our main discussion on resolving an issue of the non-uniqueness of a solution to non-strict convex problems appeared in [13] and also in [14], [3]. It will prove that our strict convex formulations indeed are suited better for image segmentation. Finally, in Section 5, we will show some numerical experiments that prove fast computation and provide visual confirmation with a meaningful stopping criterion proposed for the sake of image segmentation task.

2 Mathematical background

Throughout the paper, we will consider a bounded open subset Ω of \mathbb{R}^N , $N \geq 1$. In addition, we will follow the convention that two functions are the same if they differ only on a set of Lebesgue measure 0 and two sets are the same if their set difference is a set of Lebesgue measure 0.

Definition 1. A function $f \in L^1(\Omega)$ is of bounded variation if

$$\sup \left\{ \int_{\Omega} f(x) \operatorname{div}(\varphi(x)) dx : \varphi \in C_c^1(\Omega; \mathbb{R}^N) \text{ and } \|\varphi\|_{\infty} \leq 1 \right\} < \infty$$

and we denote the set of functions of bounded variation by $\operatorname{BV}(\Omega)$.

One of the most important properties of a function $f \in \operatorname{BV}(\Omega)$ is that it gives rise to a finite Radon measure Df , the distributional derivative of f , on Ω and its total variation is

$$|Df|(\Omega) = \sup \left\{ \int_{\Omega} f(x) \operatorname{div}(\varphi(x)) dx : \varphi \in C_c^1(\Omega; \mathbb{R}^N) \text{ and } \|\varphi\|_{\infty} \leq 1 \right\}. \quad (7)$$

We have an equivalent definition in the case $N = 1$, which will be used later.

Definition 2. Let $a, b \in \mathbb{R}$ be such that $a < b$. A function $f \in L^1(a, b)$ is of bounded variation if

$$\sup \left\{ \sum_{i=1}^m |f(t_{i+1}) - f(t_i)| : a < t_1 < \dots < t_{m+1} < b, t_i \in I \right\} < \infty,$$

where $I \subset (a, b)$ is the set of points of approximate continuity of f and we denote this supremum by $\text{ess } V_a^b f$.

When $N = 1$, we will use any of the three notations $\text{ess } V_a^b f$, $|f'| (a, b)$, $|Df| (a, b)$ for the total variation of f . For every $f \in \text{BV}(\Omega)$, we can define a bounded linear map $J_f : C(\bar{\Omega}) \rightarrow \mathbb{R}$ defined by

$$J_f(g) = \langle g, |Df| \rangle = \int_{\Omega} h(x) d|Df|(x),$$

which gives rise to a convex functional $\mathcal{J}_g : \text{BV}(\Omega) \rightarrow \mathbb{R}$, given a positive function g in $C(\bar{\Omega})$, by

$$\mathcal{J}_g(f) = J_f(g).$$

When $g \equiv 1$, we will just use

$$\mathcal{J}(f) := \mathcal{J}_1(f) = |Df|(\Omega).$$

Note that $\text{BV}(\Omega)$ is a Banach space with norm

$$\|f\|_{\text{BV}(\Omega)} = \|f\|_{L^1(\Omega)} + \mathcal{J}(f)$$

and that we will deal with the functional \mathcal{J}_g for a positive function $g \in C(\bar{\Omega})$ as well as \mathcal{J} on $\text{BV}(\Omega)$. To sum up, \mathcal{J}_g possesses the same properties as \mathcal{J} , which will enable us to propose one-step methods for two non-convex minimization problems involving \mathcal{J} and \mathcal{J}_g , respectively.

We will now discuss some of the properties of \mathcal{J} transferable to the functional \mathcal{J}_g that we need for our presentation. We refer the reader to [18] and [20] for more detailed description of $f \in \text{BV}(\Omega)$.

Lemma 1. Let $g \in C(\bar{\Omega})$ be positive. Given $f \in \text{BV}(\Omega)$,

$$\mathcal{J}_g(f) = \sup \left\{ \int_{\Omega} f(x) \text{div}(\varphi(x)) dx : \varphi \in C_c^1(\Omega; \mathbb{R}^N) \text{ and } |\varphi| \leq g \right\}. \quad (8)$$

Proof. If $g \in C^1(\bar{\Omega})$ is positive, then this follows from Lemma 1, [4]. Now let $g \in C(\bar{\Omega})$ be positive. Then, whenever a sequence $\{g_n\}$ in $C(\bar{\Omega})$ converges uniformly to g , we get

$$\mathcal{J}_g(f) = J_f(g) = \lim_{n \rightarrow \infty} J_f(g_n) = \lim_{n \rightarrow \infty} \mathcal{J}_{g_n}(f).$$

Consider two sequences $\{l_n\}$ and $\{u_n\}$ in $C^1(\bar{\Omega})$ such that

$$l_n \leq g \text{ for all } n, \text{ and } \{l_n\} \text{ converges uniformly to } g,$$

and

$$u_n \geq g \text{ for all } n, \text{ and } \{u_n\} \text{ converges uniformly to } g.$$

It is obvious that for every n ,

$$\mathcal{J}_{l_n}(f) \leq \sup \left\{ \int_{\Omega} f(x) \text{div}(\varphi(x)) dx : \varphi \in C_c^1(\Omega; \mathbb{R}^N) \text{ and } |\varphi| \leq g \right\} \leq \mathcal{J}_{u_n}(f).$$

Therefore, we obtain (8).

□

Indeed, for a positive function g in $C(\overline{\Omega})$,

$$\|f\|_{\text{BV}(\Omega),g} = \|f\|_{L^1(\Omega)} + \mathcal{J}_g(f)$$

is an equivalent norm to $\|f\|_{\text{BV}(\Omega)}$.

Theorem 2. *Let $g \in C(\overline{\Omega})$ be positive and $1 \leq p < \infty$. Given a sequence $\{f_i\}$ in $\text{BV}(\Omega)$ converging weakly to $f \in \text{BV}(\Omega)$ in $L^p(\Omega)$, we obtain*

$$\mathcal{J}_g(f) \leq \liminf_{i \rightarrow \infty} \mathcal{J}_g(f_i).$$

Proof. For $\varphi \in C_c^1(\Omega; \mathbb{R}^N)$ with $|\varphi| \leq g$, we obtain $\text{div}(\varphi) \in L^q(\Omega)$, $1 \leq q \leq \infty$, and

$$\int_{\Omega} f(x) \text{div}(\varphi(x)) dx = \lim_{i \rightarrow \infty} \int_{\Omega} f_i(x) \text{div}(\varphi(x)) dx \leq \liminf_{i \rightarrow \infty} \mathcal{J}_g(f_i).$$

Then, Lemma 1 finishes the proof. □

Theorem 2 implies that for $1 \leq p < \infty$, the extension of \mathcal{J}_g to $L^p(\Omega)$ by

$$\mathcal{J}_g(f) = \begin{cases} \mathcal{J}_g(f), & \text{if } f \in L^p(\Omega) \cap \text{BV}(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

is lower semi-continuous. Especially, we will use the extension of \mathcal{J}_g to $L^2(\Omega)$.

Theorem 3. *Let $g \in C(\overline{\Omega})$ be positive. Given $f \in \text{BV}(\Omega)$, there exists a sequence $\{f_i\}$ in $C^\infty(\Omega)$ such that*

$$\lim_{i \rightarrow \infty} \|f_i - f\|_{L^1(\Omega)} = 0, \quad \text{and} \quad \mathcal{J}_g(f) = \lim_{i \rightarrow \infty} \mathcal{J}_g(f_i).$$

Proof. We are going to construct such a sequence $\{f_i\}$ in $C^\infty(\Omega)$ in the same way as it was in [18] and [20]. So the proof will essentially be the same and we will emphasize only the difference at the end of the proof. The Given $\epsilon > 0$, there exists a positive integer $m \in \mathbb{N}$ such that

$$|Df|(\Omega \setminus \Omega_0) < \epsilon,$$

where $\Omega_0 = \left\{x \in \Omega : \text{dist}(x, \partial\Omega) > \frac{1}{m}\right\}$. For each $i = 1, 2, \dots$, we let

$$\Omega_i = \left\{x \in \Omega : \text{dist}(x, \partial\Omega) > \frac{1}{m+i}\right\}.$$

Then we define $A_1 = \Omega_2$ and

$$A_i = \Omega_{i+1} \setminus \overline{\Omega}_{i-1}, \quad i = 2, 3, \dots$$

Since $\{A_i\}$ is an open cover of Ω , we can choose a partition of unity $\{\phi_i\}$ subordinate to the cover $\{A_i\}$. Let η be a positive smooth function supported in $B_1 = \{x \in \mathbb{R}^N : |x| < 1\}$ satisfying

$$\eta(x) = \eta(|x|) \quad \text{and} \quad \int_{\mathbb{R}^N} \eta(x) dx = 1.$$

For each $i = 1, 2, \dots$, we may choose $\epsilon_i > 0$ such that

$$\begin{aligned} \text{supp}(\eta_{\epsilon_i} * (f\phi_i)) &\subset \Omega_{i+2} \setminus \overline{\Omega}_{i-2} \\ \text{and } \|\eta_{\epsilon_i} * (f\phi_i) - f\phi_i\|_{L^1(\Omega)} &< \epsilon 2^{-i} \\ \text{and } \|\eta_{\epsilon_i} * (f\nabla\phi_i) - f\nabla\phi_i\|_{L^1(\Omega)} &< \epsilon 2^{-i}. \end{aligned}$$

Then we set

$$f_\epsilon = \sum_{i=1}^{\infty} \eta_{\epsilon_i} * (f \phi_i).$$

Since f_ϵ converges to f in $L^1(\Omega)$, we obtain from Theorem 2

$$\mathcal{J}_g(f) \leq \liminf_{\epsilon \rightarrow 0} \mathcal{J}_g(f_\epsilon).$$

On the other hand, for any smooth and bounded function ϕ on \mathbb{R}^N , we have

$$\int_{\Omega} f(x) \operatorname{div}(\phi_1(x)(\eta_{\epsilon_1} * \phi)(x)) dx \leq \mathcal{J}_{|\eta_{\epsilon_1} * \phi|}(f).$$

We can eventually obtain

$$\int_{\Omega} f_\epsilon(x) \operatorname{div}(\phi(x)) dx \leq \mathcal{J}_{|\eta_{\epsilon_1} * \phi|}(f) + 4\epsilon.$$

If $|\phi| \leq g$, then $|\eta_{\epsilon_1} * \phi| \leq \eta_{\epsilon_1} * g$ and $\mathcal{J}_{|\eta_{\epsilon_1} * \phi|}(f) \leq \mathcal{J}_{\eta_{\epsilon_1} * g}(f)$, which implies, together with Lemma 1,

$$\mathcal{J}_g(f_\epsilon) = \int_{\Omega} h(x) |\nabla f_\epsilon(x)| dx \leq \mathcal{J}_{\eta_{\epsilon_1} * g}(f) + 4\epsilon.$$

Note that $\epsilon_1 \rightarrow 0$ and $\|\eta_{\epsilon_1} * g - g\|_{\infty} \rightarrow 0$ as $\epsilon \rightarrow 0$, which implies

$$\lim_{\epsilon \rightarrow 0} \mathcal{J}_{\eta_{\epsilon_1} * g}(f) = \lim_{\epsilon \rightarrow 0} J_f(\eta_{\epsilon_1} * g) = J_f(g) = \mathcal{J}_g(f).$$

That is,

$$\limsup_{\epsilon \rightarrow 0} \mathcal{J}_g(f_\epsilon) \leq \mathcal{J}_g(f).$$

This finishes the proof. □

We will now define the perimeter of a set E in Ω .

Definition 3. For a set $E \subset \Omega$,

$$P(E; \Omega) = \mathcal{J}(1_E) \quad \text{and} \quad P_g(E; \Omega) = \mathcal{J}_g(1_E).$$

Then, just as it is given in [20] that

$$P(E \cup F; \Omega) + P(E \cap F; \Omega) \leq P(E; \Omega) + P(F; \Omega)$$

we can prove the same inequality for P_g .

Lemma 2. Let $g \in C(\overline{\Omega})$ be positive. Then,

$$P_g(E \cup F; \Omega) + P_g(E \cap F; \Omega) \leq P_g(E; \Omega) + P_g(F; \Omega)$$

Proof. The same proof from [20] can be applied with the use of Theorem 2 and Theorem 3. □

Since a function $f \in \operatorname{BV}(\Omega)$ satisfies the co-area formula:

$$\mathcal{J}(f) = |Df|(\Omega) = \int_{-\infty}^{\infty} \mathcal{J}(1_{\Omega_t^f}) dt = \int_{-\infty}^{\infty} P(\Omega_t^f; \Omega) dt,$$

where $\Omega_t^f = \{x \in \Omega : f(x) > t\}$, the same co-area formula holds for \mathcal{J}_g due to Strang [26]:

Lemma 3. Let $g \in C(\bar{\Omega})$ be positive. Given $f \in \text{BV}(\Omega)$,

$$\mathcal{J}_g(f) = \int_{-\infty}^{\infty} \mathcal{J}_g(1_{\Omega_t^f}) dt = \int_{-\infty}^{\infty} P_g(\Omega_t^f; \Omega) dt,$$

where $\Omega_t^f = \{x \in \Omega : f(x) > t\}$.

Just as we can characterize various properties of the functional \mathcal{J} using convex analysis, we can do the same for \mathcal{J}_g . Among those, we want to mention one lemma whose proof can be easily derived from [17] and [28].

Lemma 4. Let $g \in C(\bar{\Omega})$ be positive. Then for any $u, p \in L^2(\Omega)$,

$$\mathcal{J}_g^{**}(u) = \mathcal{J}_g(u) \quad \text{and} \quad \mathcal{J}_g(u) + \mathcal{J}_g^*(p) \geq \int_{\Omega} p(x)u(x) dx.$$

Moreover, we obtain

$$\mathcal{J}_g^*(p) = \sup_{u \in L^2(\Omega)} \left\{ \int_{\Omega} p(x)u(x) dx - \mathcal{J}_g(u) \right\} = \begin{cases} 0, & p \in K_g, \\ +\infty, & p \notin K_g. \end{cases}$$

and

$$\partial \mathcal{J}_g(u) = \left\{ p \in K_g : \int_{\Omega} p(x)u(x) dx = \mathcal{J}_g(u) \right\},$$

where K_g is the closure in $L^2(\Omega)$ of

$$\left\{ \text{div}(\varphi) : \varphi \in C_c^1(\Omega), |\varphi| \leq g \right\}.$$

3 Proposed method

3.1 The original problem using \mathcal{J}

We now propose a strictly convex minimization problem for (3):

Theorem 4. Let $h \in L^2(\Omega)$ be such that $\{x \in \Omega : h(x) < 0\}$ has positive measure. We consider the following problem:

$$\min_{\omega} \left\{ \mathcal{J}(\omega) + \int_{\Omega} \left(\omega(x) + \frac{\lambda}{2} h(x) \right)^2 dx \right\}. \quad (9)$$

Then, there exists a unique minimizer ω^* of (9). Moreover, $\Sigma_* = \{\omega^* > 0\}$ is a minimizer of

$$\min_{\Sigma \subset \Omega} \left\{ P(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx \right\}. \quad (10)$$

In addition, (10) has a unique minimizer if and only if $\{\omega^* = 0\}$ has measure 0. Otherwise, $\{\omega^* > 0\}$ and $\{\omega^* \geq 0\}$ are the minimal and the maximal solutions of (10).

Remark 1. Note first that $h \geq 0$ a.e. makes the problem (10) obvious. And the problem (9) unlike the ones in [14], [3] is a strictly convex problem that has been intensively studied in the image processing community, which resulted in very efficient algorithms for finding the unique solution and does not require the given data h to be in $L^\infty(\Omega)$ that was a crucial condition in proposing an unconstrained convex problem in [14]. In addition, (9) has the same complexity as one of the subproblems in [3], which leads to faster computation of (10). We will discuss this in the next section.

Remark 2. If h is set to be $(c_1 - f)^2 - (c_2 - f)^2$, then (10) is the same problem as (3). Therefore, Theorem 4 provides a way to solve (3) directly via a strictly convex problem.

Before proving Theorem 4, we will modify the form of the functional in (9) to emphasize where it originated from. Fix $\beta > 0$ and define

$$\phi(x) = \frac{1}{2} + \frac{1}{\beta}\omega(x) \quad \text{and} \quad H(x) = \frac{1}{2} - \frac{\lambda}{2\beta}h(x). \quad (11)$$

Then, $H \in L^2(\Omega)$ and

$$\mathcal{J}(\phi) = \frac{1}{\beta}\mathcal{J}(\omega) \quad \text{and} \quad \phi(x) - H(x) = \frac{1}{\beta}\left(\omega(x) + \frac{\lambda}{2}h(x)\right).$$

Hence, ω^* is the unique minimizer of (9) if and only if $\phi^* = \frac{1}{2} + \frac{1}{\beta}\omega^*$ is the unique minimizer of

$$\min_{\phi} \left\{ \mathcal{J}(\phi) + \beta \int_{\Omega} \left(\phi(x) - H(x) \right)^2 dx \right\}. \quad (12)$$

And $\{\phi^* > \frac{1}{2} + \frac{t}{\beta}\} = \{\omega^* > t\}$ and $\{\phi^* \geq \frac{1}{2} + \frac{t}{\beta}\} = \{\omega^* \geq t\}$ for all $t \in \mathbb{R}$. Note also that (12) has the same minimizer as

$$\min_{\phi} \left\{ \mathcal{J}(\phi) + \lambda \int_{\Omega} \phi(x)h(x)dx + \beta \int_{\Omega} \left(\phi(x) - \frac{1}{2} \right)^2 dx \right\}. \quad (13)$$

Moreover, if we consider $\phi = 1_{\Sigma}$, then (13) becomes the same problem as (10) since the last term in (13) becomes a constant $\beta|\Omega|$.

Proof. Since there exist a great number of works in the literature about theoretical and computational aspects of the problem (9) including existence and uniqueness of the solution, we will not discuss those further in detail here. Instead, we would like to focus on the relationship between (9) and (10).

By Proposition 2 below, $\Omega_{\frac{1}{2}}^{\phi^*} = \{\phi^* > \frac{1}{2}\}$ is a minimizer of

$$\min_{\Sigma \subset \Omega} \left\{ P(\Sigma; \Omega) + 2\beta \int_{\Sigma} \left(\frac{1}{2} - H(x) \right) dx \right\} = \min_{\Sigma \subset \Omega} \left\{ P(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx \right\}.$$

Especially, we obtain from (11) that $\{\omega^* > 0\}$ is the unique minimizer of (10) if and only if $\{\omega^* = 0\}$ has measure 0, where

$$\omega^* = \beta \left(\phi^* - \frac{1}{2} \right).$$

Otherwise, $\{\omega^* > 0\}$ and $\{\omega^* \geq 0\}$ are the minimal and the maximal solutions of (10). \square

Proposition 2. (Proposition 3.1, [7]) For any $t \in \mathbb{R}$, consider the minimal surface problem

$$\min_{\Sigma \subset \Omega} \left\{ P(\Sigma; \Omega) + 2\beta \int_{\Sigma} (t - H(x)) dx \right\} \quad (14)$$

(whose solution is defined in the class of finite-perimeter sets and hence up to a Lebesgue-negligible set). Then, for the minimizer ϕ^* of (12), $\{\phi^* > t\}$ (resp., $\{\phi^* \geq t\}$) is the minimal (resp., maximal) solution of (14). In particular, for all t but a countable set, the solution of this problem is unique.

We would like to mention that our strictly convex formulation is not new because Proposition 2 has already been known in some geometric problems (e.g. [1], [10]) and we just realized its connection to the two-phase Mumford-Shah minimization model. [1] presents a characterization of convex calibrable sets in \mathbb{R}^N via the mean curvature motion and [10] also discussed an algorithm for mean curvature motion, both of which are related to the ROF model with an input $f = 1_C$. Later, [6] characterized convex calibrable sets in \mathbb{R}^N with respect to anisotropic norms generalizing the result of [1].

Despite this nice characterization and existence results related to (9) based on the works mentioned above, the source of difficulty of the two-phase image segmentation problem lies in the fact that a solution of the underlying problem (10) may not be unique, which can be observed by the following example.

Proposition 3. *Let $\Omega = (0, 1) \times (0, 3) \subset \mathbb{R}^2$ and $\lambda > 0$. We choose a function h on Ω to be*

$$h(x) = \begin{cases} -\frac{2}{\lambda}, & \text{for } x \in (0, 1) \times (0, 1), \\ 0, & \text{for } x \in (0, 1) \times (0, 2), \\ \frac{2}{\lambda}, & \text{for } x \in (0, 1) \times (2, 3). \end{cases}$$

Then, all the solutions of (10) are of the form $\Sigma_\gamma = (0, 1) \times (0, \gamma)$ for $\gamma \in [1, 2]$. That is, Σ_1 and Σ_2 are the minimal and the maximal solutions of (10) and the solution ω^ of (9) satisfies $\{\omega^* = 0\} = (0, 1) \times (1, 2)$.*

Proof. Let $\Sigma \subset \Omega$ be a solution. Then,

$$P(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx \leq -1 = P((0, 1) \times (0, 1); \Omega) + \lambda \int_{(0,1) \times (0,1)} h(x) dx. \quad (15)$$

Note that $P(\Sigma; \Omega) > 0$. We will prove now that $P(\Sigma; \Omega) = 1$. Suppose that $P(\Sigma; \Omega) < 1$. If we choose a ball B of radius r centered at $(\frac{1}{2}, \frac{1}{2})$ with

$$r = \frac{P(\Sigma; \Omega)}{2\pi} < \frac{1}{2\pi},$$

then $\bar{B} \subset (0, 1) \times (0, 1)$ and

$$P(B; \Omega) + \lambda \int_B h(x) dx = P(\Sigma; \Omega) - 2|B| \leq P(\Sigma; \Omega) - 2|\Sigma| \leq P(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx.$$

However, $P(B; \Omega) + \lambda \int_B h(x) dx = 2\pi r - 2\pi r^2 = \pi r(1 - r) > 0$ implying

$$P(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx > 0,$$

which contradicts (15). Therefore, $P(\Sigma; \Omega) \geq 1$ and this implies

$$P(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx \geq P((0, 1) \times (0, 1); \Omega) + \lambda \int_{(0,1) \times (0,1)} h(x) dx. \quad (16)$$

Indeed, by (15) and (16), we know that

$$P(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx = P((0, 1) \times (0, 1); \Omega) + \lambda \int_{(0,1) \times (0,1)} h(x) dx = -1$$

and $P(\Sigma; \Omega) = 1$. Since

$$\int_{\Sigma} h(x) dx = \int_{(0,1) \times (0,1)} h(x) dx = \int_{(0,1) \times (0,2)} h(x) dx = -2,$$

we also know that Σ satisfies

$$(0, 1) \times (0, 1) \subset \Sigma \subset (0, 1) \times (0, 2) \quad \text{a.e.}$$

If we set γ_1, γ_2 such that

$$\begin{aligned} \gamma_1 &= \sup\{\gamma \in [1, 2] : |\Sigma_\gamma \setminus \Sigma| = 0\}, \\ \gamma_2 &= \inf\{\gamma \in [1, 2] : |\Sigma \setminus \Sigma_\gamma| = 0\}, \end{aligned}$$

then $\gamma_1 = \gamma_2$. Otherwise, $P(\Sigma; \Omega) > 1$ that can be easily seen by [2] and [19]. Therefore,

$$\Sigma = \Sigma_\gamma,$$

where $\gamma = \inf\{\gamma \in [1, 2] : |\Sigma \setminus \Sigma_\gamma| = 0\}$. Therefore, all the solutions of (10) are of the form

$$\Sigma_\gamma = (0, 1) \times (0, \gamma) \quad \text{for } \gamma \in [1, 2].$$

Since Σ_1 and Σ_2 are the minimal and the maximal solutions of (10), given the solution ω^* of (9) we know that

$$\{\omega^* > 0\} = \Sigma_1 \quad \text{and} \quad \{\omega^* \geq 0\} = \Sigma_2,$$

implying that $\{\omega^* = 0\} = (0, 1) \times (1, 2)$. \square

The above example is probable when one tries to segment an image taking 3 values into two regions in the following setting. Let f be defined on $\Omega = (0, 1) \times (0, 3)$ by

$$f(x) = \begin{cases} 0, & \text{for } x \in (0, 1) \times (0, 1), \\ 1, & \text{for } x \in (0, 1) \times (0, 2), \\ 2, & \text{for } x \in (0, 1) \times (2, 3) \end{cases}$$

and we choose $c_1 = 1 - \frac{1}{2\lambda}$ and $c_2 = 1 + \frac{1}{2\lambda}$. If we solve (3) with these data c_1, c_2, f , then we end up solving (10) with $h = (c_1 - f)^2 - (c_2 - f)^2$, which is the same problem given in Proposition 3. Hence, we can view this as a problem that finds an image taking only two values c_1 and c_2 which best approximates the image f in the sense of (3) and this explains a limitation of (3).

Despite this uncertainty, Theorem 4 guarantees that the strictly convex problem (9) always provides the minimal and the maximal solutions of (3), which explains whether there is a unique solution to (3).

3.2 A variant problem using \mathcal{J}_g

For the purpose of comprehensiveness, we will now consider a slight modification of the problem investigated in the previous sections: given a positive continuous and bounded function g ,

$$\min_{\Sigma \subset \Omega} \left\{ P_g(\Sigma; \Omega) + \lambda \int_{\Sigma} h(x) dx \right\}. \quad (17)$$

This problem was discussed in [3] by considering the following convex minimization problem:

$$\min_{0 \leq \phi \leq 1} \left\{ \mathcal{J}_g(\phi) + \lambda \int_{\Omega} \phi(x) h(x) dx \right\}. \quad (18)$$

Its unconstrained version was obtained in the same way as in [14] and then an alternating minimization algorithm was proposed to take care of the non-strictly convex term that ensures

that a minimizer is bounded between 0 and 1. To be more precise, the alternating minimization algorithm in [3] to solve (18) is to solve the following two problems iteratively:

$$\min_u \left\{ \mathcal{J}_g(u) + \frac{1}{2\theta} \int_{\Omega} (u(x) - v(x))^2 dx \right\} \quad (19)$$

and

$$\min_v \left\{ \frac{1}{2\theta} \int_{\Omega} (u(x) - v(x))^2 dx + \int_{\Omega} (\lambda v(x)h(x) + \alpha v(x)) dx \right\}. \quad (20)$$

It turns out that both of the two subproblems, (19) and (20), are easy to solve and this method provides more accurate results than (5), where \mathcal{J} is used, due to the edge indicator function g in \mathcal{J}_g . However, there are convergence issues in alternating minimization problems. For instance, we need to make sure that alternating between the two subproblems converges and it can eventually find a true minimizer of (18). There is also an issue of how many times we should alternate between the two subproblems for numerical computations. Alternating between them 10 times was considered in [3].

We now propose the following strictly-convex problem to solve (17), which can resolve those issues mentioned above. Note that our proposed problem has the same computational complexity as (19) guaranteeing that our method is faster in nature than the alternating minimization algorithm. We can simply transfer the analysis related to \mathcal{J} in the previous sections to this variant related to \mathcal{J}_g , which is what we will do now.

Theorem 5. *Let $g \in C(\overline{\Omega})$ be positive. Given $h \in L^2(\Omega)$, consider the following problem:*

$$\min_{\omega} \left\{ \mathcal{J}_g(\omega) + \int_{\Omega} \left(\omega + \frac{\lambda}{2} h(x) \right)^2 dx \right\}. \quad (21)$$

Then, there exists a unique minimizer ω^ of (21). Moreover, $\Sigma_* = \{\omega^* > 0\}$ is a minimizer of (17). In addition, (17) has a unique minimizer if and only if $\{\omega^* = 0\}$ has measure 0. Otherwise, $\{\omega^* > 0\}$ and $\{\omega^* \geq 0\}$ are the minimal and the maximal solutions of (17).*

Note that by defining ϕ and H in the same way as (11) with fixed $\beta > 0$, (21) becomes equivalent to

$$\min_{\phi} \left\{ \mathcal{J}_g(\phi) + \beta \int_{\Omega} \left(\phi(x) - H(x) \right)^2 dx \right\}. \quad (22)$$

Hence, ϕ^* is the unique minimizer of (22) if and only if ω^* is the unique minimizer of (21). And $\{\phi^* > \frac{1}{2} + \frac{t}{\beta}\} = \{\omega^* > t\}$ and $\{\phi^* \geq \frac{1}{2} + \frac{t}{\beta}\} = \{\omega^* \geq t\}$ for all $t \in \mathbb{R}$.

Before proving this theorem, we need to establish the same tools for \mathcal{J}_g as those for \mathcal{J} .

Lemma 5. *Let $f_1, f_2 \in L^1(\Omega)$ and E and F , respectively, minimizers of*

$$\min_{\Sigma} \left\{ P_g(\Sigma, \Omega) - \int_{\Sigma} f_1(x) dx \right\} \quad \text{and} \quad \min_{\Sigma} \left\{ P_g(\Sigma, \Omega) - \int_{\Sigma} f_2(x) dx \right\}.$$

Then if $f_1 < f_2$ a.e., then $|E \setminus F| = 0$.

Proof. We only need the following inequality

$$P_g(A \cap B, \Omega) + P_g(A \cup B, \Omega) \leq P_g(A, \Omega) + P_g(B, \Omega)$$

from Lemma 2. The rest of the proof is the same as that of Lemma 4 in [1]. \square

Proposition 4. For any $t \in \mathbb{R}$, consider the minimal surface problem

$$\min_{\Sigma \subset \Omega} \left\{ \mathcal{J}_g(1_\Sigma) + 2\beta \int_{\Sigma} (t - H(x)) dx \right\} \quad (23)$$

(whose solution is defined in the class of finite-perimeter sets and hence up to a Lebesgue-negligible set). Then, for the minimizer ϕ^* of (22), $\{\phi^* > t\}$ (resp., $\{\phi^* \geq t\}$) is the minimal (resp., maximal) solution of (23). In particular, for all t but a countable set, the solution of this problem is unique.

Proof. The proof is fundamentally the same as that of Proposition 2.2 in [10]. There are just minor modifications. Hence, we will omit the details. \square

Proof. (Proof of Theorem 5) The same method to prove Theorem 4 using Proposition 2 applies with the help of Proposition 4. \square

4 Main discussion on an ambiguity for an L^1 problem

We would like to describe how we can solve a non-strict convex minimization L^1 problem by solving a strictly convex L^2 problem to answer an open question arisen in [13], which is related to how certain one can be that a solution from the non-strict convex problem is meaningful. Let us first recall the following theorem:

Theorem 6. *If the observed image $f(x)$ is the characteristic function of a bounded domain $\Omega_1 \subset \Omega$, then for any $\lambda \geq 0$, there exists a minimizer of $E_1(\cdot, \lambda)$ that is also the characteristic function of a (possibly different) domain. In other words, when the observed image is binary, then for each $\lambda \geq 0$, there is at least one $u(x) \in M(\lambda)$ which is also binary.*

In fact, if $u_\lambda(x) \in M(\lambda)$ is any minimizer of $E_1(\cdot, \lambda)$, then for almost every $\gamma \in [0, 1]$ we have that the binary function

$$1_{\{x \in \Omega: u_\lambda > \gamma\}}(x)$$

is also a minimizer of $E_1(\cdot, \lambda)$.

This theorem for the case of $\Omega = \mathbb{R}^N$ was provided in [13], where the functional $E_1(\cdot, \lambda)$ was defined by

$$E_1(u, \lambda) = \int_{\Omega} |\nabla u| + \lambda \int_{\Omega} |f - u| dx = \mathcal{J}(u) + \lambda \int_{\Omega} |f - u| dx$$

and $M(\lambda)$ is the set of minimizers of the functional $E_1(\cdot, \lambda)$. The same proof works for an open set $\Omega \subset \mathbb{R}^N$ in a more general framework, [3]. The main contribution of [13] was to solve the non-convex L^2 problem

$$\min_{\substack{\Sigma \subset \Omega, \\ u(x) = 1_\Sigma(x)}} \mathcal{J}(u) + \lambda \int_{\Omega} (f(x) - u(x))^2 dx \quad (24)$$

by the convex L^1 problem

$$\min_{u \in \text{BV}(\Omega)} \mathcal{J}(u) + \lambda \int_{\Omega} |f(x) - u(x)| dx, \quad (25)$$

when $f(x) = 1_{\Omega_1}(x)$. Note that Theorem 6 guarantees that

$$1_{\{x \in \Omega: u_\lambda > \gamma\}}$$

is a minimizer of (24) for a.e. $\gamma \in [0, 1]$ whenever u_λ is a minimizer of (25). However, the convex problem (25) is not completely satisfactory because of the following reasons. Firstly, Theorem 6 guarantees a minimizer of (10) only when the observed image f is binary. Secondly, it is not clear for which $\gamma \in [0, 1]$ values the solution $1_{\{x \in \Omega: u_\lambda > \gamma\}}$ of (10) is meaningful when u_λ is a minimizer of (25). Thirdly, the L^1 part in (25) is neither differentiable nor strictly convex. In this sense, the form of (24) is preferred as long as an exact solution can be found with at most the same complexity as (25). With the same process as what was observed in the previous section, we can modify (24) to have a form of (9) that can find a minimizer of (10) even if the observed image f is not binary. Note that the L^2 part of (9) is not only differentiable, but also strictly convex.

For $f = 1_{\Omega_1}$, (24) is equivalent to

$$\min_{\substack{\Sigma \subset \Omega, \\ u(x)=1_\Sigma(x)}} \left\{ \mathcal{J}(u) + \lambda \int_{\Omega} (f(x) - u(x))^2 dx - \lambda \int_{\Omega} \left(u(x) - \frac{1}{2} \right)^2 dx \right\},$$

which becomes the following convex problem:

$$\min_{\substack{\Sigma \subset \Omega, \\ u(x)=1_\Sigma(x)}} \left\{ \mathcal{J}(u) + 2\lambda \int_{\Omega} \left(\frac{1}{2} - f(x) \right) u(x) dx \right\}. \quad (26)$$

Theorem 7 below states the connection between (25) and (24) in the reverse order in the sense that (25) can be solved using (26) efficiently.

Theorem 7. *If the observed image $f(x)$ is the characteristic function of a bounded domain $\Omega_1 \subset \Omega$, then for any $\lambda \geq 0$, the unique minimizer ω^* of*

$$\min_{\omega} \left\{ \mathcal{J}(\omega) + \int_{\Omega} \left(\omega(x) + \lambda h(x) \right)^2 dx \right\}, \quad (27)$$

where

$$h(x) = \frac{1}{2} - f(x) = \frac{1}{2} 1_{\Omega \setminus \Omega_1}(x) - \frac{1}{2} 1_{\Omega_1}(x),$$

provides $1_{\{\omega^* > 0\}}$ and $1_{\{\omega^* \geq 0\}}$ as the minimizers of the convex L^1 minimization problem (25). Moreover, (25) has a unique minimizer if and only if $\{\omega^* = 0\}$ has measure 0.

Proof. This is a corollary of Theorem 4. □

We will now discuss an open question arisen in [13], i.e., if it is possible to find a solution of (25) that has more than two values when the given input data f is a characteristic function. This is equivalent to if there exists a unique solution of (24) since it can be easily seen from Theorem 6 that there exists a solution of (25) taking more than two values if and only if there exist two distinct sets $\Sigma_1, \Sigma_2 \subset \Omega$ such that $\Sigma_1 \subset \Sigma_2 \neq \Omega$ and 1_{Σ_1} and 1_{Σ_2} are two distinct solutions of (24). The difficulty of this question in the non-strict convex setting lies in characterizing the set of minimizers with an arbitrary binary input and one can ask the same question in [14], [3]. The answer to the open question is that it is indeed possible to find such a solution, which can be seen as a corollary of the following proposition. Below, Proposition 5 solves a particular case of the strictly convex formulation (27) that plays a very important role of confirming that a chosen function is indeed a minimizer of (25) in Corollary 1 and Corollary 2 and eventually constructing solutions with more than two values.

Proposition 5. *Let $\Omega = (0, 3)$. Let $\Sigma = (0, 1) \cup (\frac{3}{2}, 2)$ and $f = 1_\Sigma$. Then, for $\lambda = 2$ and*

$$h = \frac{1}{2} 1_{\Omega \setminus \Sigma} - \frac{1}{2} 1_\Sigma \quad \text{in } \Omega,$$

the solution s of (27) is

$$s = \frac{1}{2}1_{(0,1)} - \frac{1}{2}1_{(2,3)}.$$

Proof. Let s be the solution of (27) with $\lambda = 2$ and $h = \frac{1}{2}1_{\Omega \setminus \Sigma} - \frac{1}{2}1_{\Sigma}$. Note that s is unique and $-1 \leq s \leq 1$ in Ω .

On $(1, \frac{3}{2})$, we define

$$s_1(r) = \sup_{\substack{q \text{ is affine,} \\ q \leq s \text{ a.e.}}} q(r).$$

Then, s_1 is convex and $s \geq s_1$ a.e. in $(1, \frac{3}{2})$. We set $\gamma = \min_{r \in (1, \frac{3}{2})} s_1(r)$ and

$$\beta_1 = \lim_{r \rightarrow 1^+} s_1(r), \quad \beta_2 = \lim_{r \rightarrow \frac{3}{2}^-} s_1(r).$$

Let

$$s^* = s1_{\Omega \setminus (1, \frac{3}{2})} + s_11_{(1, \frac{3}{2})}.$$

Note that

$$|s'| \left(1, \frac{3}{2}\right) = \text{ess } V_1^{\frac{3}{2}} s \geq (\beta_1 - \gamma) + (\beta_2 - \gamma) = |s'_1| \left(1, \frac{3}{2}\right).$$

and

$$\int_1^{\frac{3}{2}} (s(r) + 2h(r))^2 dr = \int_1^{\frac{3}{2}} (s(r) + 1)^2 dr \geq \int_1^{\frac{3}{2}} (s_1(r) + 1)^2 dr = \int_1^{\frac{3}{2}} (s_1(r) + 2h(r))^2 dr.$$

Let $I \subset \Omega$ be the set of points of approximate continuity of s and

$$l_0 \in I \cap (0, 1] \quad \text{and} \quad l_{00} \in I \cap \left[\frac{3}{2}, 3\right).$$

For any $\epsilon > 0$, we may choose $l_1 < l_2 < l_3 < l_4$ in $I \cap (1, \frac{3}{2})$ such that

$$s(l_2) < \beta_1 < s(l_1) + \epsilon, \quad |s(l_2) - \gamma| < \epsilon, \quad |s(l_3) - \gamma| < \epsilon, \quad s(l_3) < \beta_2 < s(l_4) + \epsilon.$$

Then,

$$|\beta_1 - s(l_1)| + |s(l_1) - s(l_2)| \geq \beta_1 - s(l_2) > \beta_1 - \gamma - \epsilon,$$

and

$$|\beta_2 - s(l_4)| + |s(l_4) - s(l_3)| \geq \beta_2 - s(l_3) > \beta_2 - \gamma - \epsilon,$$

which implies that

$$\begin{aligned} & |s(l_0) - \beta_1| + (\beta_1 - \gamma) + (\beta_2 - \gamma) + |s(l_{00}) - \beta_2| \\ & \leq |s(l_0) - (s(l_1) + \epsilon)| + |(s(l_1) + \epsilon) - \beta_1| + (\beta_1 - \gamma) \\ & \quad + (\beta_2 - \gamma) + |(s(l_4) + \epsilon) - \beta_2| + |s(l_{00}) - (s(l_4) + \epsilon)| \\ & \leq |s(l_0) - s(l_1)| + s(l_1) - \gamma + s(l_4) - \gamma + |s(l_4) - s(l_{00})| + 4\epsilon \\ & \leq |s(l_0) - s(l_1)| + \sum_{i=1}^3 |s(l_{i+1}) - s(l_i)| + |s(l_4) - s(l_{00})| + 6\epsilon. \end{aligned} \tag{28}$$

We can also choose $0 < t_1 < \dots < t_{m+1} < 3$, $t_i \in I$ such that

$$\begin{aligned} |(s^*)'|(\Omega) &= \text{ess } V_0^3 s^* \leq \sum_{i=1}^m |s^*(t_{i+1}) - s^*(t_i)| + \epsilon \\ &= \sum_{i=1}^{k_1-1} |s^*(t_{i+1}) - s^*(t_i)| + \sum_{i=k_1}^{k_2} |s^*(t_{i+1}) - s^*(t_i)| + \sum_{i=k_2+1}^m |s^*(t_{i+1}) - s^*(t_i)| + \epsilon, \end{aligned}$$

where $t_{k_1} \leq 1 < t_{k_1+1} \leq t_{k_2} < \frac{3}{2} \leq t_{k_2+1}$. For $1 < p_1 < t_{k_1+1} \leq t_{k_2} < p_2 < \frac{3}{2}$ and $p_1, p_2 \in I$ such that

$$|s_1(p_1) - \beta_1| < \epsilon \quad \text{and} \quad |s_1(p_2) - \beta_2| < \epsilon,$$

we have

$$\begin{aligned} |(s^*)'|(\Omega) &\leq \sum_{i=1}^{k_1-1} |s(t_{i+1}) - s(t_i)| + |s(t_{k_1}) - s_1(p_1)| + (\beta_1 - \gamma) + (\beta_2 - \gamma) \\ &\quad + |s_1(p_2) - s(t_{k_2+1})| + \sum_{i=k_2+1}^m |s(t_{i+1}) - s(t_i)| + \epsilon \\ &\leq \sum_{i=1}^{k_1-1} |s(t_{i+1}) - s(t_i)| + |s(t_{k_1}) - \beta_1| + (\beta_1 - \gamma) + (\beta_2 - \gamma) \\ &\quad + |s(t_{k_2+1}) - \beta_2| + \sum_{i=k_2+1}^m |s(t_{i+1}) - s(t_i)| + 3\epsilon. \end{aligned}$$

By (28), we obtain

$$|(s^*)'|(\Omega) \leq |s'|(\Omega) + 9\epsilon.$$

Since $\epsilon > 0$ is arbitrary, $|(s^*)'|(\Omega) \leq |s'|(\Omega)$ and

$$\int_{\Omega} (s^*(r) + 2h(r))^2 dr \leq \int_{\Omega} (s(r) + 2h(r))^2 dr,$$

that is, $s^* = s$. Hence, s is convex in $(1, \frac{3}{2})$. Likewise, s can be proved to be convex in $(2, 3)$. If we change min to max and inf to sup, then it can be easily seen that s is concave in $(0, 1)$ and in $(\frac{3}{2}, 2)$. Therefore, s can have jump discontinuities only at $r = 1, \frac{3}{2}, 2$. We will now prove that s is constant on each of the intervals, $(0, 1)$, $(1, \frac{3}{2})$, $(\frac{3}{2}, 2)$, $(2, 3)$. We will only show that s is constant on $(1, \frac{3}{2})$ since the same proof applies to the other intervals. First of all, let

$$\gamma = \min_{r \in (1, \frac{3}{2})} s(r) \quad \text{and} \quad \beta_1 = \lim_{r \rightarrow 1^+} s(r), \quad \beta_2 = \lim_{r \rightarrow \frac{3}{2}^-} s(r).$$

If we define $s_2(r) = \gamma$ for $r \in (1, \frac{3}{2})$, then

$$|s'| \left(1, \frac{3}{2}\right) = \int_1^{\frac{3}{2}} |s'(r)| dr = \beta_1 + \beta_2 - 2\gamma = \beta_1 + \beta_2 - 2\gamma + |(s_2)'| \left(1, \frac{3}{2}\right)$$

and

$$\int_1^{\frac{3}{2}} (s(r) + 2h(r))^2 dr = \int_1^{\frac{3}{2}} (s(r) + 1)^2 dr \geq \int_1^{\frac{3}{2}} (\gamma + 1)^2 dr = \int_1^{\frac{3}{2}} (s_2(r) + 2h(r))^2 dr.$$

Now if we set $s^* = s_1 \mathbf{1}_{\Omega \setminus (1, \frac{3}{2})} + s_2 \mathbf{1}_{(1, \frac{3}{2})}$, then

$$\begin{aligned} |s'|(\Omega) &= \int_0^1 |s'(r)| dr + \int_1^{\frac{3}{2}} |s'(r)| dr + \int_{\frac{3}{2}}^2 |s'(r)| dr + \int_2^3 |s'(r)| dr + A_1 + A_{\frac{3}{2}} + A_2 \\ &= \int_0^1 |s'(r)| dr + \int_1^{\frac{3}{2}} |(s_2)'(r)| dr + \int_{\frac{3}{2}}^2 |s'(r)| dr + \int_2^3 |s'(r)| dr \\ &\quad + A_1 + (\beta_1 - \gamma) + A_{\frac{3}{2}} + (\beta_2 - \gamma) + A_2 \\ &\geq |(s^*)'|(\Omega), \end{aligned}$$

where A_t is the size of the jump of s at t , i.e.,

$$A_t = \left| \lim_{r \rightarrow t^-} s(r) - \lim_{r \rightarrow t^+} s(r) \right|,$$

since

$$A_1 + (\beta_1 - \gamma) = \left| \lim_{r \rightarrow 1^-} s(r) - \lim_{r \rightarrow 1^+} s(r) \right| + \left| \lim_{r \rightarrow 1^+} s(r) - \lim_{r \rightarrow 1^+} s_2(r) \right| \geq \left| \lim_{r \rightarrow 1^-} s^*(r) - \lim_{r \rightarrow 1^+} s^*(r) \right|$$

and, likewise,

$$A_{\frac{3}{2}} + (\beta_2 - \gamma) \geq \left| \lim_{r \rightarrow \frac{3}{2}^-} s^*(r) - \lim_{r \rightarrow \frac{3}{2}^+} s^*(r) \right|.$$

Hence, we obtain

$$|(s^*)'(\Omega) \leq |s'|(\Omega) \text{ and } \int_{\Omega} (s^*(r) + 2h(r))^2 dr \leq \int_{\Omega} (s(r) + 2h(r))^2 dr,$$

that is, $s^* = s$ implying that s is constant on $(1, \frac{3}{2})$. We can now represent the solution s as

$$c_1 \mathbf{1}_{(0,1)} + c_2 \mathbf{1}_{(1,\frac{3}{2})} + c_3 \mathbf{1}_{(\frac{3}{2},2)} + c_4 \mathbf{1}_{(2,3)}$$

for some constants $-1 \leq c_1, c_2, c_3, c_4 \leq 1$ and

$$\begin{aligned} |s'|(\Omega) + \int_{\Omega} (s(r) + 2h(r))^2 dr &= |c_1 - c_2| + |c_2 - c_3| + |c_3 - c_4| \\ &\quad + (c_1 - 1)^2 + \frac{1}{2}(c_2 + 1)^2 + \frac{1}{2}(c_3 - 1)^2 + (c_4 + 1)^2 \\ &\leq \min_{-1 \leq a_1, a_2, a_3, a_4 \leq 1} \left\{ |a_1 - a_2| + |a_2 - a_3| + |a_3 - a_4| \right. \\ &\quad \left. + (a_1 - 1)^2 + \frac{1}{2}(a_2 + 1)^2 + \frac{1}{2}(a_3 - 1)^2 + (a_4 + 1)^2 \right\}. \end{aligned}$$

Let us define

$$H(a_1, a_2, a_3, a_4) = |a_1 - a_2| + |a_2 - a_3| + |a_3 - a_4| + (a_1 - 1)^2 + \frac{1}{2}(a_2 + 1)^2 + \frac{1}{2}(a_3 - 1)^2 + (a_4 + 1)^2$$

and set

$$\mathcal{M} = H(c_1, c_2, c_3, c_4) = \min_{-1 \leq a_1, a_2, a_3, a_4 \leq 1} H(a_1, a_2, a_3, a_4).$$

If $a_1 < a_2 \leq 1$, then

$$\begin{aligned} H(a_1, a_2, a_3, a_4) &= \left(a_1 - \frac{3}{2}\right)^2 - \frac{5}{4} + \frac{1}{2}(a_2^2 + 4a_2 + 1) + |a_2 - a_3| + |a_3 - a_4| + \frac{1}{2}(a_3 - 1)^2 + (a_4 + 1)^2 \\ &> H(a_2, a_2, a_3, a_4) \end{aligned}$$

since $\left(a_1 - \frac{3}{2}\right)^2 > \left(a_2 - \frac{3}{2}\right)^2$. This implies $c_1 \geq c_2$. In the same way, $c_3 \geq c_4$ can be easily observed. We will now show that $c_3 \geq c_2$. Suppose that $c_3 < c_2$. Together with $c_1 \geq c_2$ and $c_3 \geq c_4$, we have

$$c_1 \geq c_2 > c_3 \geq c_4.$$

Note that

$$\begin{aligned} H(c_1, c_2, c_3, c_4) &= c_1 - c_4 + (c_1 - 1)^2 + \frac{1}{2}(c_2 + 1)^2 + \frac{1}{2}(c_3 - 1)^2 + (c_4 + 1)^2 \\ &> c_1 - c_4 + (c_1 - 1)^2 + \frac{1}{2}(c_2 + 1)^2 + \frac{1}{2}(c_2 - 1)^2 + (c_4 + 1)^2 \\ &= H(c_1, c_2, c_2, c_4). \end{aligned}$$

This is a contradiction since

$$H(c_1, c_2, c_3, c_4) = \mathcal{M} = \min_{-1 \leq a_1, a_2, a_3, a_4 \leq 1} H(a_1, a_2, a_3, a_4).$$

Therefore, we have obtained $c_1 \geq c_2$ and $c_3 \geq c_2$ and $c_3 \geq c_4$ and

$$\begin{aligned} \mathcal{M} &= \left(c_1 - \frac{1}{2}\right)^2 + \frac{1}{2}(c_2 - 1)^2 + \frac{1}{2}(c_3 + 1)^2 + \left(c_4 + \frac{1}{2}\right)^2 + \frac{3}{2} \\ &= \left(c_1 - \frac{1}{2}\right)^2 + \frac{1}{2}(c_2^2 + c_3^2) + (c_3 - c_2) + \left(c_4 + \frac{1}{2}\right)^2 + \frac{5}{2} \geq \frac{5}{2}. \end{aligned}$$

It is easy to see that $c_1 = \frac{1}{2}, c_2 = c_3 = 0, c_4 = -\frac{1}{2}$ and $\mathcal{M} = \frac{5}{2}$ and the solution s of (27) with $\lambda = 2$ and $g = \frac{1}{2}1_{\Omega \setminus \Sigma} - \frac{1}{2}1_{\Sigma}$ is

$$s = \frac{1}{2}1_{(0,1)} - \frac{1}{2}1_{(2,3)}.$$

□

We can now explain the possibility of non-uniqueness of a solution of (25).

Corollary 1. For $N \geq 1$, given $\Omega = (0, 3)^N$ and $f = 1_{\Sigma}$ in Ω , where

$$\Sigma = \left((0, 1) \cup \left(\frac{3}{2}, 2\right) \right) \times (0, 3)^{N-1},$$

the solution ω^* of (27) with $\lambda = 2$ and $g = \frac{1}{2}1_{\Omega \setminus \Sigma} - \frac{1}{2}1_{\Sigma}$ is

$$\omega^* = \frac{1}{2}1_{(0,1) \times (0,3)^{N-1}} - \frac{1}{2}1_{(2,3) \times (0,3)^{N-1}}.$$

Moreover, $\{\omega^* = 0\} = (1, 2) \times (0, 3)^{N-1}$ implies that for $t \in (0, 1)$,

$$t1_{\Sigma_1} + (1-t)1_{\Sigma_2}$$

is a solution of (25) taking more than two values, where

$$\Sigma_1 = (0, 1) \times (0, 3)^{N-1}, \quad \Sigma_2 = (0, 2) \times (0, 3)^{N-1}.$$

Proof. The case $N = 1$ was proved in Proposition 5. Let ω^* be the solution of (27) with the given data with $N \geq 2$. We can choose a sequence of C^∞ functions ϕ_n such that

$$\lim_{n \rightarrow \infty} \left\{ \mathcal{J}(\phi_n) + \int_{\Omega} (\phi_n(x) + 2h(x))^2 dx \right\} = \mathcal{J}(\omega^*) + \int_{\Omega} (\omega^*(x) + 2h(x))^2 dx.$$

By Proposition 5, we note that for all n ,

$$\begin{aligned} &\mathcal{J}(\phi_n) + \int_{\Omega} (\phi_n(x) + 2h(x))^2 dx = \int_{\Omega} \left[|\nabla \phi_n(x)| + (\phi_n(x) + 2h(x))^2 \right] dx \\ &= \int_{(0,3)^{N-1}} \left\{ \int_0^3 \left[|\nabla \phi_n(x_1, X_{N-1})| + (\phi_n(x_1, X_{N-1}) + 2\tilde{h}(x_1))^2 \right] dx_1 \right\} dX_{N-1} \\ &\geq \int_{(0,3)^{N-1}} \left\{ \int_0^3 \left[|\partial_{x_1} \phi_n(x_1, X_{N-1})| + (\phi_n(x_1, X_{N-1}) + 2\tilde{h}(x_1))^2 \right] dx_1 \right\} dX_{N-1} \\ &\geq \int_{(0,3)^{N-1}} \left\{ |s'(0, 3)| + \int_0^3 (s(x_1) + 2\tilde{h}(x_1))^2 dx_1 \right\} dX_{N-1}, \end{aligned}$$

where s is the solution in Proposition 5 with

$$\tilde{h}(x_1) = \frac{1}{2}1_{(1, \frac{3}{2}) \cup (2, 3)}(x_1) - \frac{1}{2}1_{(0, 1) \cup (\frac{3}{2}, 2)}(x_1) \quad \text{in } (0, 3)$$

and $X_{N-1} = (x_2, \dots, x_N)$. Therefore, we obtain that

$$\omega^*(x_1, X_{N-1}) = s(x_1), \quad \text{i.e., } \omega^* = \frac{1}{2}1_{(0, 1) \times (0, 3)^{N-1}} - \frac{1}{2}1_{(2, 3) \times (0, 3)^{N-1}}.$$

Since $\{\omega^* = 0\} = (1, 2) \times (0, 3)^{N-1}$ has positive measure, Theorem 7 implies that (25) with $\lambda = 2$ and $f = 1_\Sigma$ has two distinct solutions 1_{Σ_1} and 1_{Σ_2} , where

$$\Sigma_1 = (0, 1) \times (0, 3)^{N-1}, \quad \Sigma_2 = (0, 2) \times (0, 3)^{N-1}.$$

Since (25) is a convex problem,

$$t1_{\Sigma_1} + (1-t)1_{\Sigma_2}$$

is a solution of (25) taking more than two values for any $t \in (0, 1)$. \square

Corollary 2. *Let $\Omega = \mathbb{R}$ and $\Sigma = (0, 1) \cup (\frac{3}{2}, 2)$ and $f = 1_\Sigma$. Then, (25) with $\lambda = 2$ has a solution $t1_{(0, 1)} + (1-t)1_{(0, 2)}$ taking more than two values for $t \in (0, 1)$.*

Proof. In this proof, for an open set $A = (a, b)$, we will use the notation $\mathcal{J}(\cdot; A)$ for

$$\mathcal{J}(u; A) = \text{ess } V_a^b u$$

to emphasize the domain A . Let $\tilde{\Omega} = (-\frac{1}{2}, \frac{5}{2})$ and $h = \frac{1}{2}1_{\tilde{\Omega} \setminus \Sigma} - \frac{1}{2}1_\Sigma$ and $\lambda = 2$. Then, the unique solution s of (27) with $\Omega = \tilde{\Omega}$ and $h = \frac{1}{2}1_{\tilde{\Omega} \setminus \Sigma} - \frac{1}{2}1_\Sigma$ is constant on each of the intervals, $(-\frac{1}{2}, 0)$, $(0, 1)$, $(1, \frac{3}{2})$, $(\frac{3}{2}, 2)$, $(2, \frac{5}{2})$, as was seen in Proposition 5. Now, we consider

$$\begin{aligned} H(a_1, a_2, a_3, a_4, a_5) &= |a_1 - a_2| + |a_2 - a_3| + |a_3 - a_4| + |a_4 - a_5| \\ &\quad + \frac{1}{2}(a_1 + 1)^2 + (a_2 - 1)^2 + \frac{1}{2}(a_3 + 1)^2 + \frac{1}{2}(a_4 - 1)^2 + \frac{1}{2}(a_5 + 1)^2. \end{aligned}$$

For some $-1 \leq c_1, c_2, c_3, c_4, c_5 \leq 1$, the solution s is

$$c_1 1_{(-\frac{1}{2}, 0)} + c_2 1_{(0, 1)} + c_3 1_{(1, \frac{3}{2})} + c_4 1_{(\frac{3}{2}, 2)} + c_5 1_{(2, \frac{5}{2})}$$

where $H(c_1, c_2, c_3, c_4, c_5) = \min_{-1 \leq a_1, a_2, a_3, a_4, a_5 \leq 1} H(a_1, a_2, a_3, a_4, a_5)$. Note that if $1 \geq a_1 > a_2$, then

$$H(a_1, a_2, a_3, a_4, a_5) > H(a_2, a_2, a_3, a_4, a_5).$$

Likewise, if $a_4 < a_5 \leq 1$, then

$$H(a_1, a_2, a_3, a_4, a_5) > H(a_1, a_2, a_3, a_4, a_4).$$

Hence, we know that $c_1 \leq c_2$ and $c_4 \geq c_5$. It is also easy to see that $c_3 \leq \min(c_2, c_4)$. Then,

$$\begin{aligned} H(c_1, c_2, c_3, c_4, c_5) &= 2c_2 - c_1 - c_3 + 2c_4 - c_3 - c_5 \\ &\quad + \frac{1}{2}(c_1 + 1)^2 + (c_2 - 1)^2 + \frac{1}{2}(c_3 + 1)^2 + \frac{1}{2}(c_4 - 1)^2 + \frac{1}{2}(c_5 + 1)^2 \\ &= \frac{1}{2}c_1^2 + c_2^2 + \frac{1}{2}(c_3 - 1)^2 + \frac{1}{2}(c_4 + 1)^2 + \frac{1}{2}c_5^2 + 2 \\ &= \frac{1}{2}c_1^2 + c_2^2 + \frac{1}{2}(c_3^2 + c_4^2) + (c_4 - c_3) + \frac{1}{2}c_5^2 + 3 \geq 3 \end{aligned}$$

implies

$$c_1 = c_2 = c_3 = c_4 = c_5 = 0 \quad \text{and} \quad H(c_1, c_2, c_3, c_4, c_5) = 3.$$

Then, $s = 0$ is the solution of (27) with $\Omega = \tilde{\Omega}$ and $h = \frac{1}{2}1_{\tilde{\Omega} \setminus \Sigma} - \frac{1}{2}1_{\Sigma}$. In other words, $s = 0$ and $s = 1_{\tilde{\Omega}}$ are solutions of (25) with $\Omega = \tilde{\Omega}$ with $f = 1_{\Sigma}$ by Theorem 7. Note that

$$\mathcal{J}(0; \tilde{\Omega}) + \lambda \int_{\tilde{\Omega}} |1_{\Sigma}(x) - 0| dx = 2|\tilde{\Omega} \cap \Sigma| = 3,$$

and

$$\mathcal{J}(1_{\tilde{\Omega}}; \tilde{\Omega}) + \lambda \int_{\tilde{\Omega}} |1_{\Sigma}(x) - 1_{\tilde{\Omega}}(x)| dx = 2|\tilde{\Omega} \setminus \Sigma| = 3.$$

If \tilde{s} is a minimizer of

$$\min_u \left\{ \mathcal{J}(u; \mathbb{R}) + \lambda \int_{\mathbb{R}} |1_{\Sigma}(x) - u(x)| dx \right\},$$

then

$$\begin{aligned} 3 &= \mathcal{J}(0; \tilde{\Omega}) + \lambda \int_{\tilde{\Omega}} |1_{\Sigma}(x) - 0| dx \leq \mathcal{J}(\tilde{s}; \tilde{\Omega}) + \lambda \int_{\tilde{\Omega}} |1_{\Sigma}(x) - \tilde{s}(x)| dx \\ &\leq \mathcal{J}(\tilde{s}; \mathbb{R}) + \lambda \int_{\mathbb{R}} |1_{\Sigma}(x) - \tilde{s}(x)| dx \leq \mathcal{J}(0; \mathbb{R}) + \lambda \int_{\mathbb{R}} |1_{\Sigma}(x) - 0| dx = 3, \end{aligned}$$

which implies that $s = 0$ is also a minimizer of (25) with $\Omega = \mathbb{R}$ and

$$\mathcal{J}(0; \mathbb{R}) + \lambda \int_{\mathbb{R}} |1_{\Sigma}(x) - 0| dx = 3.$$

Since

$$\mathcal{J}(1_{(0,1)}; \mathbb{R}) + \lambda \int_{\mathbb{R}} |1_{\Sigma}(x) - 1_{(0,1)}(x)| dx = 2 + 2|\Sigma \setminus (0, 1)| = 3$$

and

$$\mathcal{J}(1_{(0,2)}; \mathbb{R}) + \lambda \int_{\mathbb{R}} |1_{\Sigma}(x) - 1_{(0,2)}(x)| dx = 2 + 2|(0, 2) \setminus \Sigma| = 3,$$

we know that $1_{(0,1)}$ and $1_{(0,2)}$ are two distinct solutions of (25) with $\Omega = \mathbb{R}$, $\lambda = 2$ and that

$$t1_{(0,1)} + (1 - t)1_{(0,2)}$$

is also a solution of (25) with $\Omega = \mathbb{R}$, $\lambda = 2$ taking more than two values for any $t \in (0, 1)$. \square

5 Numerical experiments

We will present some numerical computations in this section solving (9) via the algorithm by Chambolle [11]. We will not pursue an exhaustive numerical investigation on various algorithms because there have already been many important works in the literature dedicated to the numerical aspects and their applications. Our experiments are simply to confirm visually that we can solve image segmentation problems fast by using the proposed strictly-convex minimization problems. Nevertheless, we would like to draw attention to [12], one of the most recent algorithms, that guarantees the optimal convergence $O(1/N^2)$ for the ROF model. We noticed that our proposed stopping criterion (30) below produces satisfying results with the algorithm by Chambolle [11] as fast as with the one in [12] despite the fact that the convergence of the algorithm in [11] is slower than that of [12], which might indicate that any decent convergent algorithms should work. This comparison will follow after the presentation with the algorithm by Chambolle [11]. All the experiments in this section were obtained by repeating the following steps 3 times with $\lambda = 5$ fixed:

1. Set the two constants c_1 and c_2 by

$$c_1 = \frac{1}{|\Sigma|} \int_{\Sigma} f(x) dx \quad \text{and} \quad c_2 = \frac{1}{|\Omega \setminus \Sigma|} \int_{\Omega \setminus \Sigma} f(x) dx.$$

2. Having c_1, c_2 defined as above, we compute the solution ω^* of (9) with h replaced by

$$(c_1 - f(x))^2 - (c_2 - f(x))^2,$$

and set $\Sigma = \{\omega^* > 0\}$.

When setting c_1 and c_2 for the first time, we simply chose $\Sigma \subset \Omega$ to be a circular disk in the center of the domain Ω whose area is about half of the area of Ω . And we run the algorithm by Chambolle [11] to solve (9), which is described by

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(\nabla(\operatorname{div}(p^n) + \lambda h))_{i,j}}{1 + \tau|\nabla(\operatorname{div}(p^n) + \lambda h))_{i,j}|}$$

with $\tau = 0.1$. The usual stopping criterion for this algorithm is

$$\max_{i,j} |p_{i,j}^{n+1} - p_{i,j}^n| < \epsilon \quad (29)$$

for a chosen $\epsilon > 0$. Then,

$$\omega_{i,j}^* = -\frac{\lambda}{2} h_{i,j} - \frac{1}{2} (\operatorname{div}(p^{n+1}))_{i,j}$$

is considered to be the unique solution of (9) and $\Sigma = \{(i, j) : \omega_{i,j}^* > 0\}$ is a solution of (10). Note, however, that what we really need to find is the region where ω^* is positive. This leads to another stopping criterion,

$$|\Sigma_{n+1} \Delta \Sigma_n| < \epsilon, \quad (30)$$

where

$$\Sigma_n = \left\{ (i, j) : \omega_{i,j}^n = -\frac{\lambda}{2} h_{i,j} - \frac{1}{2} (\operatorname{div}(p^n))_{i,j} > 0 \right\}$$

and

$$\Sigma_{n+1} \Delta \Sigma_n = (\Sigma_{n+1} \setminus \Sigma_n) \cup (\Sigma_n \setminus \Sigma_{n+1}).$$

In this discrete setting, we choose $\epsilon = 1$ which we would like to consider as exact segmentation since $|\Sigma_{n+1} \Delta \Sigma_n|$ is an integer for each n and $|\Sigma_{n+1} \Delta \Sigma_n| < 1$ means $\Sigma_{n+1} = \Sigma_n$. The results are shown in Figure 1. We want to point out that there should be a discussion about how often we want to check (30) and we found that checking (30) at every 10^{th} iteration helped compute the results almost the fastest for all the experiments.

If we compare our algorithm with the alternating minimization algorithm in [3], then it is clear that our algorithm computes results faster because for each pair (c_1, c_2) , our algorithm solves a problem of the form (6) once, however, the one in [3] solves the same kind of a problem many times as an internal step of the alternating minimization algorithm. Moreover, our stopping criterion (30) helps us find numerical solutions even faster.

For all the experiments, we used either noisy images or non-binary images simply because it takes about a few milliseconds to obtain the exact segmentation result ($\epsilon = 1$ for (30)) with a binary image, which is not of much interest. We normalized the input data f to have values in $[0, 1]$. So the solution ω^* of (9) is obtained from this normalized input data. For the visualization purpose, we presented the normalized view of the solution ω^* .

Other than that we can compute the solution faster and more efficiently, another advantage of using (9) for segmentation is that since it is guaranteed to find a true minimizer of (10) by solving (9) once, it can recognize an object disguised in heavy noise, which we show in Figure 2.

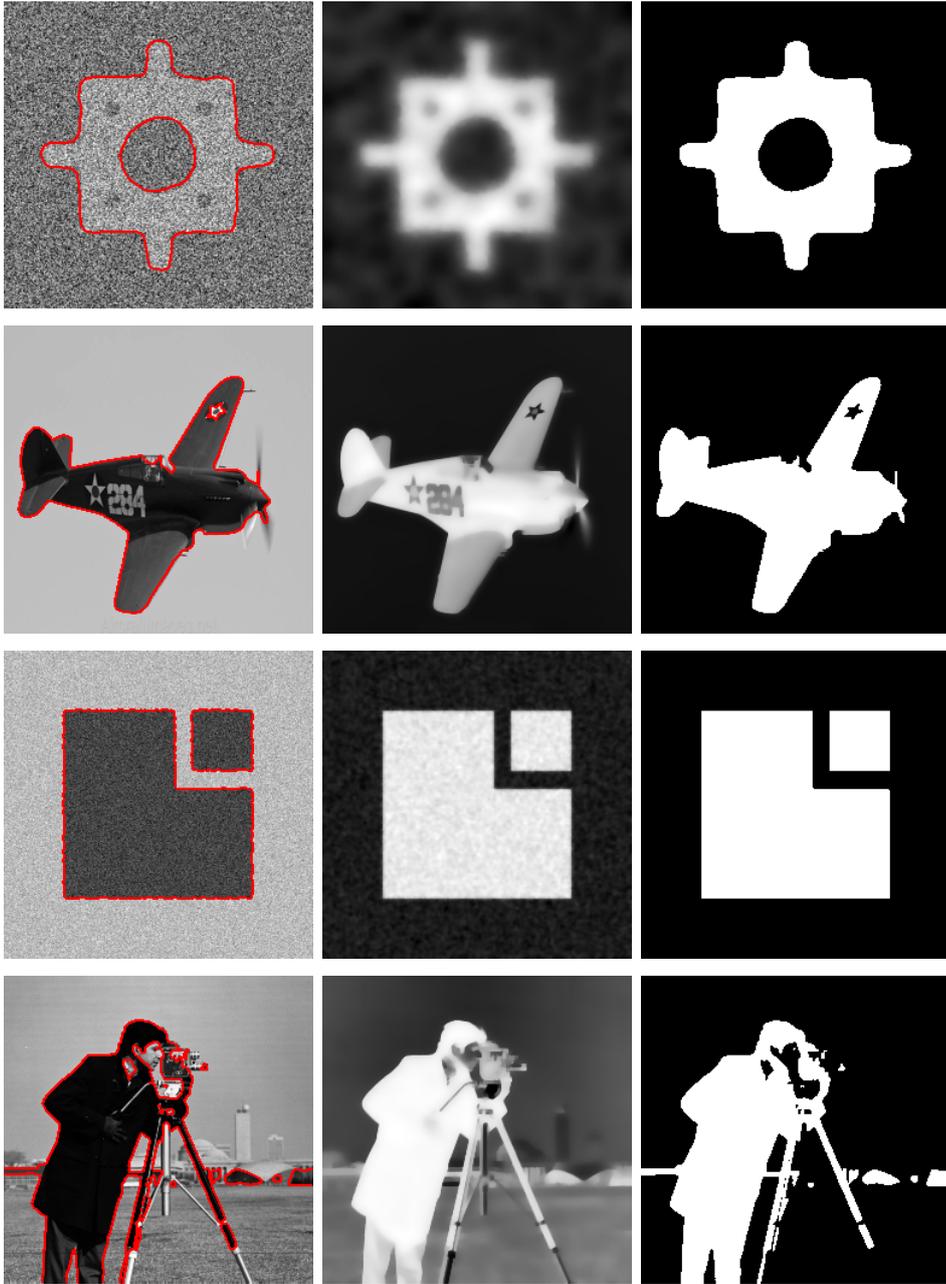


Figure 1: **Left column** : Input images with the segmentation results indicated in red, **Middle column** : Computed solutions ω^* , **Right column** : The set $\{\omega^* > 0\}$ shown in white. The elapsed time for each experiment from top to bottom : 1.54 sec, 0.31 sec, 0.06 sec, 1.02 sec. The stopping criterion was $|\Sigma_{n+1}\Delta\Sigma_n| = 0$. Images of size 256×256 were used. When visualizing the images in the middle column, we normalized them for better recognition.

In this figure, we also presented two results by considering the alternating minimization method to solve (19) and (20) iteratively. We note that Theorem 6 and its generalization with \mathcal{J}_g in [3]

are valid only for binary input data, resulting in uncertainty in the final results to some extent such as how to set the segmentation result from the final solution u . The level curve $u = 0.5$ is usually used to denote the segmentation. However, this is not supported by the theory. On the other hand, our solution $\{\omega^* > 0\}$ always serves as a correct segmentation result that is a solution of either (10) or (17). For both results, we tried to compute the results as fast as we can while trying to obtain the results of the best quality by considering various factors. The most important issues were the choice of ϵ in (29) and the number of internal alternation between (19) and (20) for each pair (c_1, c_2) and the level of u for the final segmentation, which do not exist in our models (9) and (21). We used three colors to indicate three different level curves of the solution u . The yellow, blue, green contours represent the level curves $u = 0.4, 0.5, 0.6$, respectively. On the left, $u = 0.6$ seems to be a reasonable segmentation, whereas $u = 0.5$ seems to be a reasonable one.

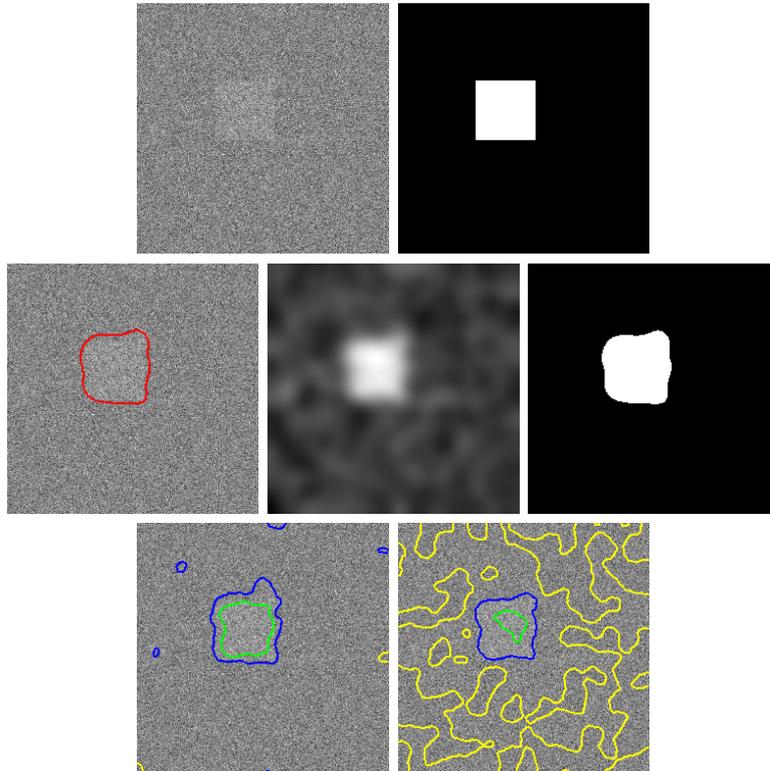


Figure 2: **Top Row** : Extremely noisy input image and the original noiseless image. **Middle Row** : Input image with the segmentation result in red, Normalized view of the computed solution ω^* using the stopping criterion $\epsilon = 1$ for (30), the set $\{\omega^* > 0\}$. The elapsed time was 4.01 seconds. **Bottom Row** : The best results that we obtained by (19) and (20) with $\theta = 1, \lambda = 1, \tau = 0.1$ and $\epsilon = 0.001$ for (29). The internal alternation between (19) and (20) was repeated 3 times. Yellow contour $u = 0.4$, Blue contour $u = 0.5$, Green contour $u = 0.6$. The elapsed time for the left and the right : 4.38 seconds, 7.85 seconds.

In Figure 3, we compared two stopping criteria: a) our stopping criterion (30), b) the combination of (29) and (30). We chose $\epsilon = 1$ for a). As for b), we chose $\epsilon = 0.00001$ for (29) and $\epsilon = 1$ for (30). The purpose of this comparison is not to show which one is better, but to confirm that our proposed stopping criterion (30) alone works very well and can be used for

exact segmentation when $\epsilon = 1$.

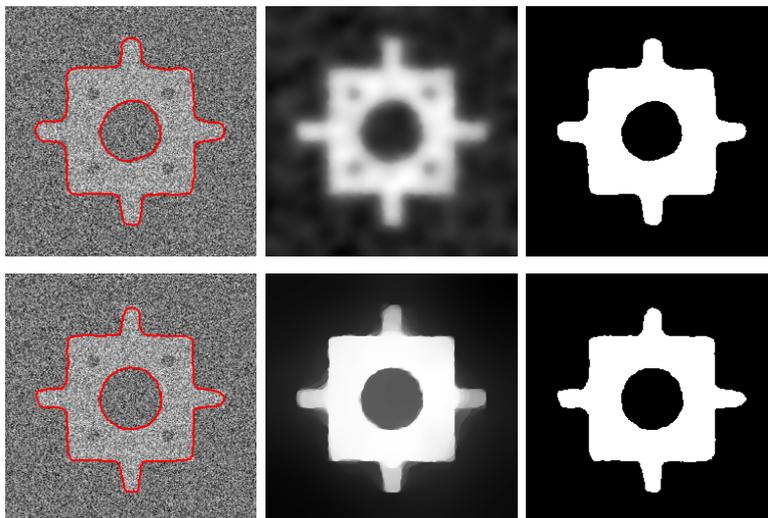


Figure 3: **Top Row** : Result obtained by the stopping criterion a), **Bottom Row** : Result obtained by the stopping criterion b). In the middle column, we show the computed solutions ω^* corresponding to the two different criteria: a) and b), where the bottom image looks brighter than the image above it, which is because we normalized each image for better recognition. Indeed, the actual function values of the bottom image are less than those of the top image.

In Figure 4, we present an experiment showing that $\epsilon = 1$ for (30) is necessary when the input image is noisy. In the absence of noise, it has been observed that we can choose $\epsilon > 1$ for (30), which results in a faster computation. Even in the absence of noise, if the input image to be segmented is complicated, then it is expected that we use $\epsilon = 1$.

Finally, it was observed that the algorithm by Chambolle [11] and the one in [12], despite the fact that the latter provides a better convergence rate than the former meaning that the latter possibly gets us a better result during the same amount of time, presented comparable results in terms of quality and cost. At this point, we would like to stress again that we do not pursue the unique solution of (9) when solving it, but are interested in the region where the unique solution is positive and the stopping criterion (30) does the work. The exact algorithm in [12] to solve (9) is as follows: we set $\tau_0 = 1/8 = \sigma_0$, $\gamma = 1.4$ and $\bar{x}^0 = 0 = y^0 = x^0$ and for each $n = 1, 2, \dots$ we compute

$$\begin{aligned} y_{i,j}^{n+1} &= \frac{y_{i,j}^n + \sigma_n (\nabla \bar{x}^n)_{i,j}}{\max\{1, |y_{i,j}^n + \sigma_n (\nabla \bar{x}^n)_{i,j}|\}}, \\ x_{i,j}^{n+1} &= \frac{x_{i,j}^n + \tau_n (\operatorname{div}(y^{n+1}))_{i,j} - \tau_n \lambda h_{i,j}}{1 + 2\tau_n}, \\ \bar{x}^{n+1} &= x^{n+1} + \theta_n (x^{n+1} - x^n), \end{aligned}$$

where

$$\theta_n = \frac{1}{1 + 2\gamma\tau_n}, \quad \tau_{n+1} + \theta_n \tau_n, \quad \sigma_{n+1} = \frac{\sigma_n}{\theta_n}.$$

We show comparison results in Figure 5 between the algorithm by Chambolle [11] and the algorithm in [12], where we can notice that they perform almost equally well in terms of quality and cost. However, the algorithm in [12] did not work for the heavy noisy case shown in Figure 2 in the exactly same setting.

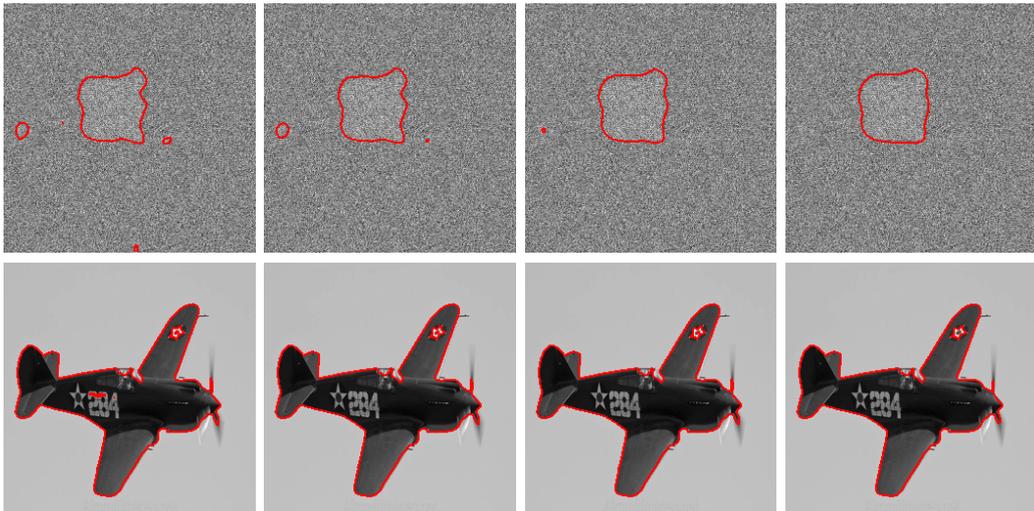


Figure 4: Each column from the left to the right corresponds to the stopping criterion (30) with $\epsilon = 10, 5, 2, 1$. The comparison is to see how the value of ϵ in (30) affects the final segmentation with or without noise. We used the same noisy input image from Figure 2 and one of the clean input images from Figure 1. **Top Row from left to right** : $\epsilon = 10$ and TE (Time elapsed) = 1.55 seconds, $\epsilon = 5$ and TE = 2.11 seconds, $\epsilon = 2$ and TE = 2.80 seconds, $\epsilon = 1$ (exact segmentation) and TE = 4.01 seconds. **Bottom Row from left to right** : $\epsilon = 10$ and TE (Time elapsed) = 0.12 seconds, $\epsilon = 5$ and TE = 0.16 seconds, $\epsilon = 2$ and TE = 0.28 seconds, $\epsilon = 1$ (exact segmentation) and TE = 0.31 seconds. The segmentation result with $\epsilon = 5$ is already satisfactory in the absence of noise.

6 Conclusion

We presented strictly-convex minimization formulations for image segmentation and showed that this idea, well-known in some geometric problems, turned out to be useful indeed for segmentation as well. Strict convexity helped us find meaningful solutions, i.e., minimal and maximal solutions, and provided us with an easier and faster way to find global minimizers related to the two-phase image segmentation. Moreover, we realized that the strictly convex formulations could do more than this as was discussed in Section 4. In other words, non-uniqueness of solutions to non-strict convex problems usually hinders us from characterizing the set of minimizers, which was also discussed in [13] and the strictly convex formulations confirmed this ambiguity in such a rare case as the one in [13] by writing an equivalent strictly convex problem to the non-strict convex problem. The same issue can be addressed in [14] and [3] and the same answer of ours applies to them as well. These concrete examples confirm the fundamental belief that strict convexity is always preferred. As numerically dealing with the strictly convex formulations, we found that any decent algorithms as well as the ones with optimal convergent rate could do the work with our proposed stopping criterion. This provides another advantage because we do not need to compute the exact solution of a strictly convex problem, resulting in faster computation than the strictly convex problem itself and, not to mention, than the related non-strict convex problems. More importantly, whenever 2-level or multi-level image segmentation algorithms solving (10) or (17) are considered, the strictly convex method can be implemented easily for more general classes of image segmentation problems.

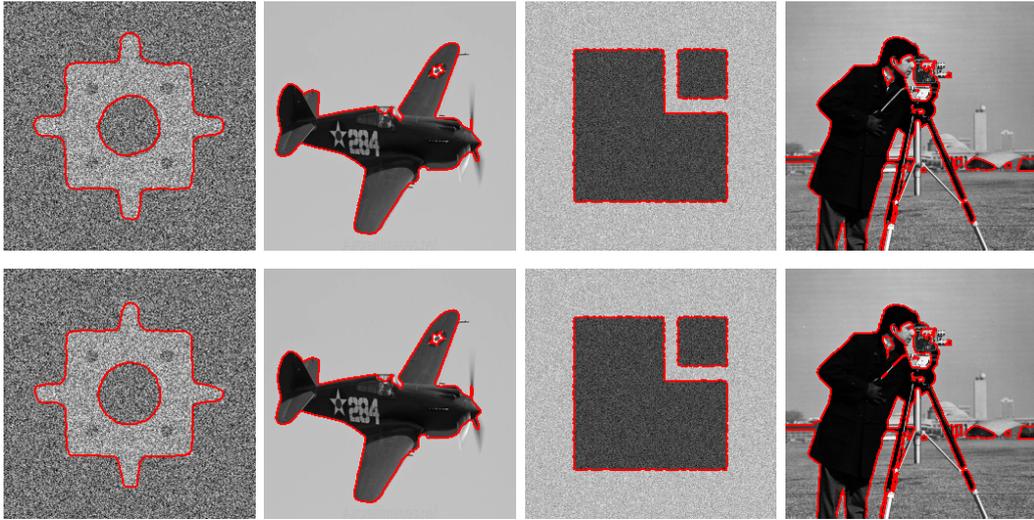


Figure 5: **Top Row** : The four results shown in Figure 1 with the algorithm by Chambolle [11]. The elapsed time from left to right : 1.54 sec, 0.31 sec, 0.06 sec, 1.02 sec. **Bottom Row** : The corresponding results with the algorithm in [12] to the ones above them. The elapsed time from left to right : 2.10 sec, 0.41 sec, 0.06 sec, 0.80 sec.

References

- [1] F. Alter, V. Caselles and A. Chambolle, *A chracterization of convex calibrable sets in \mathbb{R}^N* , Math. Ann., 332(2), 329–366 (2005)
- [2] L. Ambrosio, V. CVaselles, S. Masnou and J.-M. Morel, *Connected components of sets of finite perimeter and applications to image processing*, J. Eur. Math. Soc., 3, 39–92 (2001)
- [3] X. Bresson, S. Esedoglu, P. Vanderghyest, J.-P. Thiran and S. Osher, *Fast Global Minimization of the Active Contour/Snake Model*, J. Math. Imaging Vis., 28(2), 151–167 (2007)
- [4] G. Carlier and M. Comte, *On a weighted total variation minimization problem*, Journal of Functional Analysis, 250, 214–226 (2007)
- [5] V. Caselles, F. Catté, T. Coll and F. Dibos, *A geometric model for active contours in image processing*, Numer. Math., 66, 1–31 (1993)
- [6] V. Caselles, A. Chambolle, S. Moll, and M. Novaga, *A chracterization of convex calibrable sets in \mathbb{R}^N with respect to anisotropic norms*, Ann. I. H. Poincaré AN 25, 803832 (2008)
- [7] V. Caselles, A. Chambolle and M. Novaga, *The discontinuity set of solutions of the TV denoising problem and some extensions*, SIAM Multiscale Model. Simul. 6(3), 879–894 (2007)
- [8] V. Caselles, A. Chambolle and M. Novaga, *Regularity for solutions of the total variation denoising problem*, Rev. Mat. Iberoamericana Volume 27, Number 1 (2011), 233-252.
- [9] V. Caselles, R. Kimmel and G. Sapiro, *On geodesic active contours*, Int. J. Comput. Vis., 22(1), 61–79 (1997)
- [10] A. Chambolle, *An algorithm for mean curvature motion*, Interfaces and Free Boundaries, 6, 195–218 (2004)
- [11] A. Chambolle, *An Algorithm for Total Variation Minimization and Applications*, J. Math. Imaging Vis., 20, 89–97 (2004)

- [12] A. Chambolle and T. Pock, *A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging*, J. Math. Imaging Vis., 40, 120–145 (2011)
- [13] T. Chan and S. Esedoglu, *Aspects of total variation regularized L^1 function approximation*, SIAM J. Appl. Math., 65(5), 1817–1837 (2005)
- [14] T. Chan, S. Esedoglu and M. Nikolova, *Algorithms for finding global minimizers of image segmentation and denoising models*, SIAM J. Appl. Math., 66(5), 1632–1648
- [15] T. Chan and L. Vese, *Active contours without edges*, IEEE Transactions on image processing, 10(2), 266–277 (2001)
- [16] T. Chan and L. Vese, *A multiphase level set framework for image segmentation using the Mumford and Shah model*, Inter. J. Comput. Vis., 50(3), 271–293 (2002)
- [17] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, SIAM
- [18] C. Evance and R. Gariepy, *Measure Theory and Fine Properties of Functions*, Stud. Adv. Math., CRC Press, Inc., 1992
- [19] A. Ferriero and N. Fusco, *A note on the convex hull of sets of finite perimeter in the plane*, DCDS-B, 11(1), 103–108 (2009)
- [20] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, 1984
- [21] T. Goldstein and S. Osher, *The Split Bregman Method for $L1$ -Regularized Problems*, SIAM J. Ing. Sci., 2(2), 323–343 (2009)
- [22] M. Kass, A. Witkin and D. Terzopoulos, *Snakes: Active contour models*, Int. J. Comput. Vis., 1, 321–331 (1988)
- [23] R. Malladi, J.A. Sethian and B.C. Vemuri, *A topology independent shape modeling scheme*, in Proc. SPIE Conf. Geometric Methods Computer Vision II, vol. 2031, San Diego, CA, 246–258 (1993)
- [24] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Win, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4(2), 460 – 489 (2005)
- [25] L. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physics D, 60, 259 – 268 (1992)
- [26] G. Strang, *Maximal flow through a domain*, Mathematical Programming, 26, 123–143 (1983)
- [27] R. Temam, *Dual variational principles in mechanics and physics. In: Semi-Infinite Programming and Applications* (AV Fiacco and KO Kortanec, eds). Lecture Notes in Economics and Mathematical Systems 215. Springer-Verlag, Berlin
- [28] L. Vese, *A Study in the BV Space of a Denoising-Deblurring Variational Problem*, Appl. Math. Optim. 44:131–161 (2001)