A Penalty Relaxation Method for Image Processing Using Euler's Elastica Model*

Fang He†, Xiao Wang‡, and Xiaojun Chen†

Abstract. Euler's elastica model has been widely used in image processing. Since it is a challenging nonconvex and nonsmooth optimization model, most existing algorithms do not have convergence theory for it. In this paper, we propose a penalty relaxation algorithm with mathematical guarantee to find a stationary point of Euler's elastica model. To deal with the nonsmoothness of Euler's elastica model, we first introduce a smoothing relaxation problem, and then propose an exact penalty method to solve it. We establish the relationships between Euler's elastica model, the smoothing relaxation problem and the penalty problem in theory regarding optimal solutions and stationary points. Moreover, we propose an efficient block coordinate descent algorithm to solve the penalty problem by taking advantages of convexity of its subproblems. We prove global convergence of the algorithm to a stationary point of the penalty problem. Finally we apply the proposed algorithm to denoise the optical coherence tomography images with real data from an optometry clinic and show the efficiency of the method for image processing using Euler's elastica model.

Key words. Euler's elastica model, smoothing relaxation, exact penalty, block coordinate descent, convergence, OCT images

AMS subject classifications. 68U10, 90C26, 94A08

1. Introduction.

In this paper, we consider the following Euler's elastica model:

$$\min_u \int_\Omega \left(a + b\left|\nabla \cdot \frac{\nabla u}{|\nabla u|}\right|^2\right)|\nabla u| dxdy + \frac{\lambda}{2}\int_\Omega (\mathcal{K}u - u_0)^2 dxdy, \tag{1.1}$$

where a, b and λ are three positive constant parameters, $|\cdot|$ denotes the l_2 norm of a vector, Ω is a bounded domain in $\mathbb{R}^2$ with Lipschitz continuous boundary. Here, $u : \Omega \to \mathbb{R}$ is assumed to be smooth and denotes the reconstructed output image with $u(x, y)$ being the intensity value of the grey level of u at point $(x, y) \in \Omega$. Moreover, $u \in BV(\Omega) = \{u \in \mathcal{L}^1(\Omega) : |\nabla u|(\Omega) < \infty\}$ where $|\nabla u|$ is the total-variation measure of the weak gradient ∇u of u and u_0 denotes the observed input image. And $\nabla\cdot$ denotes the divergence operator. The operator $\mathcal{K} : BV(\Omega) \to \mathcal{L}^2(\Omega)$ is linear and bounded. The first term in (1.1) is the regularizer which captures the geometrical features of the image while the second term is the fidelity term which guarantees that $\mathcal{K}u$ will be close to the input image u_0.

Variational models have a wide range of applications in image processing [1, 6, 22]. Euler's elastica model (1.1), as a kind of variational model, can be used for illusory contour [25], de-

---

*Submitted to the editors DATE.

Funding: The work of the authors was partially supported by Hong Kong Research Grants Council, grant PolyU15300219, the Chinese NSF Grants 11871453 and 11731013, Young Elite Scientists Sponsorship Program by CAST (2018QNRC001), Youth Innovation Promotion Association, CAS, and the University Research Facility in Big Data Analytics of the Hong Kong Polytechnic University.

†Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong (fanghf.he@connect.polyu.hk, maxjchen@polyu.edu.hk).

‡School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, China (wangxiao@ucas.ac.cn).

noising [26], segmentation [31], inpainting [5], etc. However, it is challenging to solve (1.1) due to nonlinearity, nonconvexity and nonsmoothness of the energy functional. Many numerical methods have been studied for solving (1.1) in the literature. In [23], Chan et al. studied the mathematical foundation and properties of Euler's elastica model. Moreover, a computational scheme based on numerical PDEs was proposed to solve the inpainting problem. They derived the Euler-Lagrange equation for (1.1) and applied a weighted steepest descent method [20] to solve the equation. Actually, PDE based methods are used widely in image processing [7, 27]. Later, to find numerical solutions of the equation in [23], the authors of [5] studied two unconditionally stable time marching methods and a fixed point method. Furthermore, a nonlinear multigrid method was proposed by taking the fixed point method as a smoother. In [26], Tai et al. reformulated the minimization of the Euler's elastica energy to a constrained optimization problem and then proposed an augmented Lagrangian method to solve it. In [30], Yashtini and Kang presented two numerical algorithms to solve Euler's elastica inpainting model (1.1). By relaxing the normal vector $\frac{\nabla u}{|\nabla u|}$ in the curvature term and introducing a new vector to replace ∇u, they proposed a RN2Split algorithm based on operator splitting techniques. They also proposed a κTV algorithm to solve Euler's elastica model in the form of a weighted TV model [30], in which $\nabla \cdot \frac{\nabla u}{|\nabla u|}$ is regarded as an independent term. Recently, Deng et al. [13] proposed a new operator splitting method for solving (1.1). Compared with works on the alternating direction method of multipliers, the time discretization step size is the only free parameter to choose, which leads to the robustness and stableness of the proposed algorithm. To overcome the difficulty of nonconvexity of (1.1), some researchers have studied its convex relaxation. For example, [3] studied a convex relaxation of a class of vertex penalizing functionals, which captures the curvature of level lines of images. Bredies et al. [4] proposed a convex, lower semi-continuous, coercive approximation of Euler's elastica energy by functional lifting of the gradient of the image. Some other works on convex relaxations have also been reviewed in [4]. However, although algorithms for solving Euler's elastica model have been studied comprehensively, convergence analysis of these algorithms is rarely provided.

In this paper, we propose a penalty relaxation method for image processing using Euler's elastica model and have the following new contributions.

- As the discrete reformulation of Euler's elastica model is used in practical computation, we propose a smoothing relaxation problem with a smoothing parameter ϵ and inequality constraints for the discrete Euler's elastica model. We show that any accumulation point of stationary points and any accumulation point of optimal solutions of the smoothing relaxation problems are a stationary point and a global minimizer of the discrete Euler's elastica model, respectively, as the parameter ϵ decreases to zero.
- To solve the smoothing relaxation problem, we represent nonconvex inequality constraints by a penalty term added to the objective. We show that a strict local minimizer of the smoothing relaxation problem is a local minimizer of the penalty problem associated with all sufficiently large penalty parameters. Moreover, if a point is a stationary point of the penalty problem for all sufficiently large penalty parameters, then it is a stationary point of the smoothing relaxation problem. These properties ensure that the penalty problem is a promising approach to find a stationary point of

the discrete Euler's elastica model.

- By taking advantage of the bi-convexity of the penalty problem with respect to two groups of variables, denoted as $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{w} \in \mathbb{R}^{2m}$, we propose a smoothing block coordinate descent (BCD) algorithm. This algorithm executes BCD iteration while updating the smoothing parameter simultaneously. More specifically, at each iteration we first solve an unconstrained strictly convex $\mathbf{u}$-problem through a modified fixed-point method. Then by partitioning $\mathbf{w}$ to $(Q_1^T\mathbf{w}, \ldots, Q_m^T\mathbf{w})$, we sequentially solve m subproblems with respect to $Q_i^T\mathbf{w}$, $i = 1, \ldots, m$ and each subproblem is a two-dimensional convex ball-constrained problem whose solution can be easily calculated. We prove that any accumulation point of the sequence generated by the smoothing BCD algorithm is a stationary point of the penalty problem.

This paper is organized as follows. In section 2, we give the discrete Euler's elastica model through discretization and relaxation. Moreover, we propose a smoothing relaxation model and define a penalty problem by representing the nonconvex constraint by a penalty term in the objective. In section 3, we explore the relationships among solutions and stationary points of the discrete model, the smoothing relaxation and the penalty problem. In section 4, we propose the smoothing BCD algorithm to solve the penalty problem and present the convergence results. In section 5, we present some numerical results by applying the proposed method to some image processing problems.

2. Discretization and relaxation.

It is worth noting that the curvature term in (1.1) makes no sense at those pixels of image with $|\nabla u| = 0$. To deal with this, relaxation is normally used in related works (see, e.g. [2, 13, 26]). Following the relaxation approach in [13], we replace $\frac{\nabla u}{|\nabla u|}$ by a function p satisfying

$$\langle p, \nabla u\rangle = |\nabla u| \text{ and } |p| \leq 1. \tag{2.1}$$

By the well-known Hölder's inequality, (2.1) is equivalent to $p = \frac{\nabla u}{|\nabla u|}$ only for u with $|\nabla u| \neq 0$. When $|\nabla u|$ vanishes, (2.1) ensures the boundedness of p. Then Euler's elastica model (1.1) can be relaxed to the following constrained optimization problem [13]

$$\min_{(u,p)\in\mathcal{W}} \int_\Omega \left(a + b|\nabla\cdot p|^2\right)|\nabla u|dxdy + \frac{\lambda}{2}\int_\Omega (\mathcal{K}u - u_0)^2 dxdy, \tag{2.2}$$

where

$$\mathcal{W} = \{(u,p) \in BV(\Omega) \times \mathcal{H}(\Omega, \mathrm{div}), \langle p, \nabla u\rangle = |\nabla u|, |p| \leq 1\}$$

with

$$\mathcal{H}(\Omega, \mathrm{div}) = \{p \in (\mathcal{L}^2(\Omega))^2, \nabla\cdot p \in \mathcal{L}^2(\Omega)\}.$$

We now introduce the discrete form of model (2.2) with a rectangle $\Omega = [x_0, x_1] \times [y_0, y_1]$. Let the mesh size be $\triangle x = (x_1 - x_0)/(n_1 - 1)$ and $\triangle y = (y_1 - y_0)/(n_2 - 1)$. We consider the discrete image domain $\bar\Omega = \{(x_i, y_j);\ x_i = x_0 + (i-1)\triangle x, y_j = y_0 + (j-1)\triangle y, i = 1, \cdots, n_1,\ j = 1, \cdots, n_2\}$ as an $n_1 \times n_2$ grid and rearrange the intensity value of each pixel in the discrete image into a vector $\mathbf{u} \in \mathbb{R}^m$ ($m = n_1 n_2$). In a similar way we can obtain $\mathbf{u}^0 \in \mathbb{R}^n$ from the input image for some n. We denote a discrete operator by $K \in \mathbb{R}^{n\times m}$.

We define the variable $\mathbf{w} = \binom{\mathbf{p}_1}{\mathbf{p}_2}$ with $\mathbf{p}_1, \mathbf{p}_2 \in \mathbb{R}^m$ obtained by rearranging the discrete forms of p. Let $D^{(1)}, D^{(2)} \in \mathbb{R}^{m\times m}$ be the first order forward finite difference matrices with periodic boundary condition in the horizontal and vertical direction, respectively. We define $D_i \in \mathbb{R}^{2\times m}$ to represent the difference matrix at i-th pixel which consists of the i-th rows of $D^{(1)}$ and $D^{(2)}$. Let $\mathbf{d}_i \in \mathbb{R}^{2m}$ be a vector whose first m elements are the i-th row of the matrix $D^{(1)}$ and last m elements are the i-th row of $D^{(2)}$. Then $\mathbf{d}_i^T\mathbf{w}$ is a discrete divergence of $(\mathbf{p}_1, \mathbf{p}_2)$ at i-th pixel. Let $\mathbf{e}_i \in \mathbb{R}^{2m}$ be the vector with the i-th element being one and others being zero and define $Q_i = (\mathbf{e}_i, \mathbf{e}_{m+i})$, $i = 1, \ldots, m$. Then we obtain the discrete form of (2.2) as

$$
\begin{aligned}
&\min_{\substack{\mathbf{u}\in\mathbb{R}^m\\ \mathbf{w}\in\mathbb{R}^{2m}}} \Phi(\mathbf{u},\mathbf{w}),\ \text{where } \Phi(\mathbf{u},\mathbf{w}) = \sum_{i=1}^m \left(a + b(\mathbf{d}_i^T\mathbf{w})^2\right)\|D_i\mathbf{u}\| + \frac{\lambda}{2}\|K\mathbf{u}-\mathbf{u}_0\|^2 \\
&\quad\text{s.t. } \|D_i\mathbf{u}\| - \mathbf{w}^T Q_i(D_i\mathbf{u}) = 0, \\
&\qquad\quad \|Q_i^T\mathbf{w}\|^2 \le 1,\ i = 1,\cdots,m,
\end{aligned}
\tag{2.3}
$$

where $\mathbf{d}_i^T\mathbf{w} \in \mathbb{R}$ and $D_i\mathbf{u} \in \mathbb{R}^2$ denote the curvature of the level line and discrete gradient of image u at i-th pixel respectively and $\|\cdot\|$ denotes the l_2 norm. In (2.3), the second term is the discrete form of $\frac{\lambda}{2}\int_\Omega(\mathcal{K}u - u_0)^2 dxdy$. We omit the constant $\triangle x\triangle y$ for simplicity in our theoretical study.

For various applications, the matrix K may be different. For example, K is the identity matrix for some image denoising problems, while for image inpainting, K is a diagonal matrix with

$$
K_{i,i} = \begin{cases} 1, & i \in \bar\Omega \setminus \Gamma, \\ 0, & i \in \Gamma, \end{cases}
$$

where i corresponds to the i-th pixel in the domain $\bar\Omega$, and Γ represents the inpainting domain.

To overcome computational difficulties caused by the nonsmooth term $\|D_i\mathbf{u}\|$ in (2.3), we propose a smoothing relaxation scheme which replaces the nonsmooth term by a smooth one and relaxes the equality constraints by inequalities. It results in the following smooth optimization problem:

$$
\begin{aligned}
&\min_{\substack{\mathbf{u}\in\mathbb{R}^m\\ \mathbf{w}\in\mathbb{R}^{2m}}} \Phi_\epsilon(\mathbf{u},\mathbf{w}),\ \text{where } \Phi_\epsilon(\mathbf{u},\mathbf{w}) = \sum_{i=1}^m \left(a + b(\mathbf{d}_i^T\mathbf{w})^2\right)\|D_i\mathbf{u}\|_\epsilon + \frac{\lambda}{2}\|K\mathbf{u}-\mathbf{u}_0\|^2 \\
&\quad\text{s.t. } \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) \le 0, \\
&\qquad\quad \|Q_i^T\mathbf{w}\|^2 \le 1,\ i = 1,\cdots,m,
\end{aligned}
\tag{2.4}
$$

where ϵ is a positive parameter,

$$
\|D_i\mathbf{u}\|_\epsilon = \sqrt{\|D_i\mathbf{u}\|^2 + \epsilon^2} \quad\text{and}\quad \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) = \|D_i\mathbf{u}\|_\epsilon - \mathbf{w}^T Q_i(D_i\mathbf{u}) - 2\epsilon,\ i = 1,\ldots,m.
$$

In problem (2.4), we relax the equality constraint of (2.3) into $\varphi_i^\epsilon(\mathbf{u},\mathbf{w}) \le 0$, which can guarantee that the feasible set of (2.3) is contained in the feasible set of problem (2.4). Moreover, for any fixed $\mathbf{w}$, the set $\{\mathbf{u} : \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) \le 0\}$ is convex, which helps us to develop a block coordinate descent algorithm. Note that the term -2ϵ in $\varphi_i^\epsilon(\mathbf{u},\mathbf{w})$ can be replaced by $-c\epsilon$ for any $c > 1$.

To solve problem (2.4) effectively, we represent the nonconvex inequality constraints $\varphi_i^\epsilon \le 0$ for $i = 1,\dots,m$ by adding a penalty term in the objective, which yields the following penalty problem

$$\begin{aligned} &\min_{\substack{\mathbf{u}\in\mathbb{R}^m\\ \mathbf{w}\in\mathbb{R}^{2m}}} \Psi_{\epsilon,\sigma}(\mathbf{u},\mathbf{w}), \text{ where } \Psi_{\epsilon,\sigma}(\mathbf{u},\mathbf{w}) = \Phi_\epsilon(\mathbf{u},\mathbf{w}) + \sigma\sum_{i=1}^m (\varphi_i^\epsilon(\mathbf{u},\mathbf{w}))_+ \\ &\quad \text{s.t. } \|Q_i^T\mathbf{w}\|^2 \le 1,\ i = 1,\cdots,m, \end{aligned} \tag{2.5}$$

where $\sigma > 0$ is a penalty parameter and $(z)_+ := \max\{z, 0\}$. It is worth noting that the constraints in problem (2.5) are convex and only related to the variable $\mathbf{w}$, and the objective is bi-convex, that is, it is convex with respect to $\mathbf{u}$ and $\mathbf{w}$ for fixed $\mathbf{w}$ and $\mathbf{u}$ respectively. This special structure inspires us to propose an efficient block coordinate descent algorithm in section 4. At the end of this section, we observe that problems (2.3), (2.4) and (2.5) have bounded solution sets. It is easy to see that the feasible sets of the three problems are nonempty and bounded in the components $\mathbf{w}$, since the zero vector in $\mathbb{R}^{3m}$ is their feasible point and $\|Q_i^T\mathbf{w}\|^2 = w_i^2 + w_{m+i}^2$. The three objective functions are continuous and coercive in the components $\mathbf{u}$ for any fixed $\mathbf{w}$. Moreover, the objective function values are nonnegative.

3. Relationships between problems (2.3), (2.4) and (2.5).

In this section, we first give some necessary constraint qualifications and optimality conditions of problems (2.3), (2.4) and (2.5). Next, the theoretical relationships between these three problems are established regarding their optimal solutions and stationary points.

3.1. Constraint qualification and optimality conditions.

In this subsection, we will study constraint qualification and optimality conditions for nonconvex optimization problems (2.3), (2.4) and (2.5). We use notations $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ to represent the feasible sets of problems (2.3), (2.4) and (2.5), respectively, i.e.,

$$\mathcal{A}_1 = \{(\mathbf{u},\mathbf{w}) : \mathbf{u}\in\mathbb{R}^m,\ \mathbf{w}\in\mathbb{R}^{2m},\ \|D_i\mathbf{u}\| - \mathbf{w}^TQ_i(D_i\mathbf{u}) = 0,\ \|Q_i^T\mathbf{w}\|^2 \le 1,\ i = 1,\cdots,m\},$$

$$\mathcal{A}_2 = \{(\mathbf{u},\mathbf{w}) : \mathbf{u}\in\mathbb{R}^m,\ \mathbf{w}\in\mathbb{R}^{2m},\ \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) \le 0,\ \|Q_i^T\mathbf{w}\|^2 \le 1,\ i = 1,\cdots,m\},$$

$$\mathcal{A}_3 = \{(\mathbf{u},\mathbf{w}) : \mathbf{u}\in\mathbb{R}^m,\ \mathbf{w}\in\mathbb{R}^{2m},\ \|Q_i^T\mathbf{w}\|^2 \le 1,\ i = 1,\cdots,m\}.$$

Moreover, let $g_i(\mathbf{w}) = \|Q_i^T\mathbf{w}\|^2 - 1$, $I(\mathbf{u},\mathbf{w}) = \{i : \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) = 0,\ i = 1,\cdots,m\}$ and $J(\mathbf{w}) = \{i : g_i(\mathbf{w}) = 0,\ i = 1,\cdots,m\}$.

Problem (2.3) is nonsmooth at $(\mathbf{u},\mathbf{w})$ if there is $i \in \{1,\dots,m\}$ such that $D_i\mathbf{u} = 0$. The following lemma gives first order optimality conditions for problem (2.3).

Lemma 3.1. *Let $(\mathbf{u}^*,\mathbf{w}^*)$ be a solution of problem (2.3), $L = \{i : D_i\mathbf{u}^* = 0, i = 1,\dots,m\}$, and $\mathcal{U} = \bigcap_{i\in L}\ker(D_i)$. Then there exist multipliers $\xi_i \in \mathbb{R}, i \notin L$, $\zeta_i \in \mathbb{R}^2, i \in L$ and $\eta \in \mathbb{R}_+^m$,* *such that the following conditions hold,*

$$\sum_{i=1,i\notin L}^{m} \left(a+b(\mathbf{d}_i^T\mathbf{w}^*)^2+\xi_i\right) D_i^T\frac{D_i\mathbf{u}^*}{\|D_i\mathbf{u}^*\|}+\lambda K^T(K\mathbf{u}^*-\mathbf{u}_0)-\sum_{i=1,i\notin L}^{m}\xi_i D_i^TQ_i^T\mathbf{w}^*+\sum_{i\in L}D_i^T\zeta_i=0, \tag{3.1}$$

$$\sum_{i=1,i\notin L}^{m} 2b\|D_i\mathbf{u}^*\|\cdot\mathbf{d}_i(\mathbf{d}_i^T\mathbf{w}^*)-\sum_{i=1,i\notin L}^{m}\xi_iQ_iD_i\mathbf{u}^*+2\sum_{i=1}^{m}\eta_iQ_iQ_i^T\mathbf{w}^*=0, \tag{3.2}$$

$$\|D_i\mathbf{u}^*\|-\mathbf{w}^TQ_iD_i\mathbf{u}^*=0,\ i=1,\dots,m, \tag{3.3}$$

$$\min\{\eta_i,1-\|Q_i^T\mathbf{w}^*\|^2\}=0,\ i=1,\dots,m. \tag{3.4}$$

Proof. Let $\hat\Phi(\mathbf{u},\mathbf{w})=\sum_{i=1,i\notin L}^{m}\left(a+b(\mathbf{d}_i^T\mathbf{w})^2\right)\|D_i\mathbf{u}\|+\frac{\lambda}{2}\|K\mathbf{u}-\mathbf{u}_0\|^2$. By the optimality of $(\mathbf{u}^*,\mathbf{w}^*)$ for problem (2.3) and the definition of L, we have

$$\begin{aligned}\hat\Phi(\mathbf{u}^*,\mathbf{w}^*)=\Phi(\mathbf{u}^*,\mathbf{w}^*)&=\min\{\Phi(\mathbf{u},\mathbf{w}):(\mathbf{u},\mathbf{w})\in\mathcal{A}_1\}\\&=\min\{\Phi(\mathbf{u}^*+\mathbf{h},\mathbf{w}):(\mathbf{u}^*+\mathbf{h},\mathbf{w})\in\mathcal{A}_1\}\\&\le\min\{\Phi(\mathbf{u}^*+\mathbf{h},\mathbf{w}):(\mathbf{u}^*+\mathbf{h},\mathbf{w})\in\mathcal{A}_1,\ \mathbf{h}\in\mathcal{U}\}\\&=\min\{\hat\Phi(\mathbf{u}^*+\mathbf{h},\mathbf{w}):(\mathbf{u}^*+\mathbf{h},\mathbf{w})\in\mathcal{A}_1,\ \mathbf{h}\in\mathcal{U}\}.\end{aligned}$$

Hence $(\mathbf{0},\mathbf{w}^*)$ is a solution of the optimization problem

$$\begin{aligned}\min_{\substack{\mathbf{h}\in\mathbb{R}^m\\ \mathbf{w}\in\mathbb{R}^{2m}}}\ &\hat\Phi(\mathbf{u}^*+\mathbf{h},\mathbf{w})\\ \text{s.t. }&\|D_i(\mathbf{u}^*+\mathbf{h})\|-\mathbf{w}^TQ_iD_i(\mathbf{u}^*+\mathbf{h})=0,\ i\notin L,\\ &D_i\mathbf{h}=0,\ i\in L,\\ &\|Q_i^T\mathbf{w}\|^2\le 1, i=1,\dots,m.\end{aligned} \tag{3.5}$$

Notice that there exists a neighborhood $B(\mathbf{0},\mathbf{w}^*)$ such that for any feasible point $(\mathbf{h},\mathbf{w})\in B(\mathbf{0},\mathbf{w}^*)$, $D_i(\mathbf{u}^*+\mathbf{h})\neq 0$, $i\notin L$ and $Q_i^T\mathbf{w}\neq 0$, $i\in J(\mathbf{w}^*)$. Hence problem (3.5) is smooth at any feasible point in $B(\mathbf{0},\mathbf{w}^*)$.

Denote $\mathcal{S}$ by the set consisting of the gradients of equality constraints and active inequality constraints for problem (3.5), namely

$$\mathcal{S}=\mathcal{S}_1\cup\mathcal{S}_2\cup\mathcal{S}_3$$

where

$$\mathcal{S}_1=\left\{\begin{pmatrix}D_i^T\frac{D_i(\mathbf{u}^*+\mathbf{h})}{\|D_i(\mathbf{u}^*+\mathbf{h})\|}-D_i^TQ_i^T\mathbf{w}\\ -Q_iD_i(\mathbf{u}^*+\mathbf{h})\end{pmatrix}, i\notin L\right\},\qquad \mathcal{S}_2=\left\{\begin{pmatrix}D_i^T\\0\end{pmatrix}, i\in L\right\},$$

$$\mathcal{S}_3=\left\{\begin{pmatrix}0\\2Q_iQ_i^T\mathbf{w}\end{pmatrix}, i\in J(\mathbf{w}^*)\right\}.$$

Since $\|D_i(\mathbf{u}^*+\mathbf{h})\|-\mathbf{w}^T Q_i D_i(\mathbf{u}^*+\mathbf{h})=0$, $D_i(\mathbf{u}^*+\mathbf{h})\neq 0$ and $\|Q_i^T\mathbf{w}\|^2\leq 1$ imply $\frac{D_i(\mathbf{u}^*+\mathbf{h})}{\|D_i(\mathbf{u}^*+\mathbf{h})\|}=Q_i^T\mathbf{w}$, we have $D_i^T\frac{D_i(\mathbf{u}^*+\mathbf{h})}{\|D_i(\mathbf{u}^*+\mathbf{h})\|}-D_i^TQ_i^T\mathbf{w}=0$ for $i\notin L$. Moreover, $\|D_i\mathbf{u}^*\|-(\mathbf{w}^*)^TQ_iD_i\mathbf{u}^*=0$ and $\|D_i\mathbf{u}^*\|\neq 0$ imply that $i\in J(\mathbf{w}^*)$. Hence we obtain

$$\mathcal{S}_1=\left\{\begin{pmatrix}0\\ -\|D_i(\mathbf{u}^*+\mathbf{h})\|Q_iQ_i^T\mathbf{w}\end{pmatrix}, i\notin L\right\}\subseteq\left\{\begin{pmatrix}0\\ -\|D_i(\mathbf{u}^*+\mathbf{h})\|Q_iQ_i^T\mathbf{w}\end{pmatrix}, i\in J(\mathbf{w}^*)\right\}.$$

Noticing $D_i(\mathbf{u}^*+\mathbf{h})\neq 0$, $\|Q_i^T\mathbf{w}\|=1$, $i\notin L$ and $Q_i^T\mathbf{w}\neq 0$, $i\in J(\mathbf{w}^*)$, it yields that the rank of $\mathcal{S}_1\cup\mathcal{S}_3$ equals that of $\mathcal{S}_3$ for any $(\mathbf{h},\mathbf{w})\in B(\mathbf{0},\mathbf{w}^*)$. Therefore, as the rank of $\mathcal{S}_2$ is fixed, for any subset $\bar{\mathcal{S}}\subseteq\mathcal{S}$, $\bar{\mathcal{S}}$ has the same rank in the neighborhood $B(\mathbf{0},\mathbf{w}^*)$, which means that the constant rank constraint qualification (CRCQ)[1] holds at point $(\mathbf{0},\mathbf{w}^*)$. Since $\hat{\Phi}$ and all constraint functions are continuously differentiable, the KKT conditions hold at the solution $(\mathbf{0},\mathbf{w}^*)$ of problem (3.5) [17]. Therefore, there exist multipliers $\xi_i, i\notin L$, $\zeta_i, i\in L$ and $\eta\in\mathbb{R}^m_+$, such that conditions (3.1)-(3.4) for problem (2.3) hold at $(\mathbf{u}^*,\mathbf{w}^*)$. ■

We call $(\mathbf{u}^*,\mathbf{w}^*)$ a stationary point of (2.3), if it satisfies conditions (3.1)-(3.4).

We say that the Mangasarian-Fromovitz constraint qualification (MFCQ) holds at a feasible point $(\mathbf{u},\mathbf{w})$ of problem (2.4) if there exist vectors $\mathbf{z}_1\in\mathbb{R}^m, \mathbf{z_2}\in\mathbb{R}^{2m}$ such that for any $i\in I(\mathbf{u},\mathbf{w})$ and $j\in J(\mathbf{w})$, it holds

$$\begin{pmatrix}\nabla_{\mathbf{u}}\varphi_i^\epsilon(\mathbf{u},\mathbf{w})\\ \nabla_{\mathbf{w}}\varphi_i^\epsilon(\mathbf{u},\mathbf{w})\end{pmatrix}^T\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z_2}\end{pmatrix}<0,\quad (\nabla_{\mathbf{w}}g_j(\mathbf{w}))^T\mathbf{z_2}<0.$$

Lemma 3.2. *The MFCQ holds at any feasible point of problem (2.4). Let $(\mathbf{u}^*,\mathbf{w}^*)$ be a local minimizer of problem (2.4). Then there exist multipliers $\xi,\eta\in\mathbb{R}^m_+$, such that the following KKT conditions hold,*

$$\sum_{i=1}^m(a+b(\mathbf{d}_i^T\mathbf{w}^*)^2+\xi_i)D_i^T\frac{D_i\mathbf{u}^*}{\|D_i\mathbf{u}^*\|_\epsilon}+\lambda K^T(K\mathbf{u}^*-\mathbf{u}_0)-\sum_{i=1}^m\xi_iD_i^TQ_i^T\mathbf{w}^*=0, \tag{3.6}$$

$$\sum_{i=1}^m 2b\|D_i\mathbf{u}^*\|_\epsilon\cdot\mathbf{d}_i(\mathbf{d}_i^T\mathbf{w}^*)-\sum_{i=1}^m\xi_iQ_iD_i\mathbf{u}^*+2\sum_{i=1}^m\eta_iQ_iQ_i^T\mathbf{w}^*=0, \tag{3.7}$$

$$\min\{\xi_i,-\varphi_i^\epsilon(\mathbf{u}^*,\mathbf{w}^*)\}=0,\ i=1,\dots,m, \tag{3.8}$$

$$\min\{\eta_i, 1-\|Q_i^T\mathbf{w}^*\|^2\}=0,\ i=1,\dots,m. \tag{3.9}$$

Proof. Let $(\mathbf{u},\mathbf{w})$ be a feasible point of problem (2.4). For simplicity, we abbreviate $I(\mathbf{u},\mathbf{w})$ and $J(\mathbf{w})$ to I and J, respectively. To show MFCQ holds at $(\mathbf{u},\mathbf{w})$, we first introduce a vector $\binom{t\mathbf{u}}{-\mathbf{w}}$, where t is a constant satisfying

$$t<\min_{i\in I}\left\{\frac{(2\epsilon-\|D_i\mathbf{u}\|_\epsilon)\|D_i\mathbf{u}\|_\epsilon}{2\epsilon\|D_i\mathbf{u}\|_\epsilon-\epsilon^2}\right\}, \tag{3.10}$$

[1]We say that the CRCQ holds at a feasible point of a smooth constrained optimization problem, if for each subset of the gradients of equality constraints and active inequality constraints, the rank at a neighborhood of this feasible point is constant [17].

where $2\epsilon\|D_i\mathbf{u}\|_\epsilon - \epsilon^2 \geq \epsilon^2 > 0$. Thus, for any $i \in I$, we have

$$\begin{aligned}
\begin{pmatrix}\nabla_{\mathbf{u}}\varphi_i^\epsilon(\mathbf{u},\mathbf{w})\\ \nabla_{\mathbf{w}}\varphi_i^\epsilon(\mathbf{u},\mathbf{w})\end{pmatrix}^T \begin{pmatrix} t\mathbf{u}\\ -\mathbf{w}\end{pmatrix} &= \begin{pmatrix} D_i^T\frac{D_i\mathbf{u}}{\|D_i\mathbf{u}\|_\epsilon} - D_i^T(Q_i^T\mathbf{w})\\ -Q_i(D_i\mathbf{u})\end{pmatrix}^T \begin{pmatrix} t\mathbf{u}\\ -\mathbf{w}\end{pmatrix}\\
&= (\frac{D_i\mathbf{u}}{\|D_i\mathbf{u}\|_\epsilon} - Q_i^T\mathbf{w})^T D_i(t\mathbf{u}) + \mathbf{w}^T Q_i D_i\mathbf{u}\\
&= t(\frac{\|D_i\mathbf{u}\|^2}{\|D_i\mathbf{u}\|_\epsilon} - \mathbf{w}^T Q_i D_i\mathbf{u}) + \mathbf{w}^T Q_i D_i\mathbf{u}\\
&= t(\frac{\|D_i\mathbf{u}\|^2}{\|D_i\mathbf{u}\|_\epsilon} - \|D_i\mathbf{u}\|_\epsilon + 2\epsilon) + \|D_i\mathbf{u}\|_\epsilon - 2\epsilon\\
&= t\frac{2\epsilon\|D_i\mathbf{u}\|_\epsilon - \epsilon^2}{\|D_i\mathbf{u}\|_\epsilon} + (\|D_i\mathbf{u}\|_\epsilon - 2\epsilon)\\
&< 0,
\end{aligned}$$

where the forth equality is from $\varphi_i^\epsilon(\mathbf{u},\mathbf{w}) = \|D_i\mathbf{u}\|_\epsilon - \mathbf{w}^T Q_i D_i\mathbf{u} - 2\epsilon = 0$ for $i \in I$, and the last inequality is from (3.10) and $2\epsilon\|D_i\mathbf{u}\|_\epsilon - \epsilon^2 > 0$. Moreover, for any $j \in J$, we have $\|Q_j^T\mathbf{w}\|^2 = 1$ and

$$(\nabla_{\mathbf{w}} g_j(\mathbf{w}))^T(-\mathbf{w}) = (2Q_j(Q_j^T\mathbf{w}))^T(-\mathbf{w}) = -2\|Q_j^T\mathbf{w}\|^2 < 0.$$

Therefore, MFCQ holds at $(\mathbf{u},\mathbf{w})$.

All functions in problem (2.4) are continuously differentiable. Hence, at any local minimizer of (2.4), under MFCQ there exist multipliers $\xi, \eta \in \mathbb{R}_+^m$ such that the KKT conditions (3.6)-(3.9) hold. ∎

We call $(\mathbf{u}^*, \mathbf{w}^*)$ a stationary point of (2.4), if it satisfies conditions (3.6)-(3.9).

The objective function of problem (2.5) is Lipschitz continuous, but not differentiable. To derive the first order optimality condition of problem (2.5), we use the Clarke subdifferential. For any $\epsilon, \sigma > 0$, it follows from Proposition 2.3.3 and Corollary 1 in [12] that

$$\partial\Psi_{\epsilon,\sigma}(\mathbf{u},\mathbf{w}) = \nabla\Phi_\epsilon(\mathbf{u},\mathbf{w}) + \sigma\partial\sum_{i=1}^m (\varphi_i^\epsilon(\mathbf{u},\mathbf{w}))_+ .$$

Since $\varphi_i^\epsilon(\mathbf{u},\mathbf{w})$, $i = 1,\ldots,m$, are continuously differentiable, and the plus function $(\cdot)_+$ is convex, by Proposition 2.3.6(b) and Theorem 2.3.9(iii) in [12], we have that $(\varphi_i^\epsilon)_+$ is regular [12, Definition 2.3.4] and

$$\partial(\varphi_i^\epsilon(\mathbf{u},\mathbf{w})_+) = r_i(\mathbf{u},\mathbf{w})\nabla\varphi_i^\epsilon(\mathbf{u},\mathbf{w}),$$

where

$$r_i(\mathbf{u},\mathbf{w}) := \begin{cases} 1, & \text{if } \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) > 0,\\ [0,1], & \text{if } \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) = 0,\\ 0, & \text{if } \varphi_i^\epsilon(\mathbf{u},\mathbf{w}) < 0,\end{cases} \tag{3.11}$$

for $i = 1, \cdots, m$. From the regularity of $(\varphi_i^\epsilon)_+$ and Corollary 3 of Proposition 2.3.3 in [12], it indicates

$$\partial \sum_{i=1}^{m} (\varphi_i^\epsilon(\mathbf{u}, \mathbf{w}))_+ = \sum_{i=1}^{m} \partial\left(\varphi_i^\epsilon(\mathbf{u}, \mathbf{w})_+\right) = \sum_{i=1}^{m} r_i(\mathbf{u}, \mathbf{w}) \nabla \varphi_i^\epsilon(\mathbf{u}, \mathbf{w}). \tag{3.12}$$

Lemma 3.3. *Assume that $(\bar{\mathbf{u}}, \bar{\mathbf{w}})$ is a local minimizer of problem (2.5). Then there exist Lagrangian multipliers $\rho_i \geq 0$ and coefficients $\kappa_i \in r_i(\bar{\mathbf{u}}, \bar{\mathbf{w}})$, $i = 1, \ldots, m$, with r_i defined in (3.11), such that*

$$\sum_{i=1}^{m} \left(a + b(\mathbf{d}_i^T \bar{\mathbf{w}})^2\right) D_i^T \frac{D_i \bar{\mathbf{u}}}{\|D_i \bar{\mathbf{u}}\|_\epsilon} + \lambda K^T (K\bar{\mathbf{u}} - \mathbf{u}_0) + \sigma \sum_{i=1}^{m} \kappa_i \left(D_i^T \frac{D_i \bar{\mathbf{u}}}{\|D_i \bar{\mathbf{u}}\|_\epsilon} - D_i^T Q_i^T \bar{\mathbf{w}}\right) = 0, \tag{3.13}$$

$$\sum_{i=1}^{m} 2b \|D_i \bar{\mathbf{u}}\|_\epsilon \cdot \mathbf{d}_i (\mathbf{d}_i^T \bar{\mathbf{w}}) - \sigma \sum_{i=1}^{m} \kappa_i Q_i D_i \bar{\mathbf{u}} + 2 \sum_{i=1}^{m} \rho_i Q_i Q_i^T \bar{\mathbf{w}} = 0, \tag{3.14}$$

$$\min\left(\rho_i, 1 - \|Q_i^T \bar{\mathbf{w}}\|^2\right) = 0, \;\; i = 1, \ldots, m. \tag{3.15}$$

Proof. By the definition of Q_i, $\nabla g_i(w) = 2Q_i Q_i^T \mathbf{w}, i \in J(\mathbf{w})$ are linearly independent at any feasible point of problem (2.5). Hence the linear independence constraint qualification (LICQ) holds at a local minimizer $(\bar{\mathbf{u}}, \bar{\mathbf{w}})$ of problem (2.5). Therefore, there exist $\rho_i \geq 0$ for $i = 1, \ldots, m$ such that

$$0 \in \partial(\Psi_{\epsilon,\sigma}(\bar{\mathbf{u}}, \bar{\mathbf{w}}) + \sum_{i=1}^{m} \rho_i (\|Q_i^T \bar{\mathbf{w}}\|^2 - 1))$$

and (3.15) hold. Since Φ_ϵ and $\|Q_i^T \mathbf{w}\|^2$ are continuously differentiable, and $(\varphi_i^\epsilon)_+$ is regular, by (3.12), we find that the first order necessary conditions (3.13)-(3.15) hold at $(\bar{\mathbf{u}}, \bar{\mathbf{w}})$. ■

We call $(\bar{\mathbf{u}}, \bar{\mathbf{w}})$ a stationary point of (2.5) if conditions (3.13)-(3.15) hold at $(\bar{\mathbf{u}}, \bar{\mathbf{w}})$.

3.2. Relationships between (2.3), (2.4) and (2.5).

In this subsection, we will focus on relationships between problems (2.3), (2.4) and (2.5). First, we consider problems (2.3) and (2.4) regarding global minimizers.

Theorem 3.4. *For any given $\epsilon > 0$, assume that $(\mathbf{u}_\epsilon, \mathbf{w}_\epsilon)$ is an optimal solution of problem (2.4). Let $(\mathbf{u}^*, \mathbf{w}^*)$ be an arbitrary accumulation point of $\{(\mathbf{u}_\epsilon, \mathbf{w}_\epsilon)\}$ as $\epsilon \downarrow 0$. Then $(\mathbf{u}^*, \mathbf{w}^*)$ is an optimal solution of problem (2.3).*

Proof. From

$$\|D_i \mathbf{u}\|_\epsilon - \mathbf{w}^T Q_i D_i \mathbf{u} \leq \|D_i \mathbf{u}\| - \mathbf{w}^T Q_i D_i \mathbf{u} + \epsilon = \epsilon,$$

at any feasible point $(\mathbf{u}, \mathbf{w})$ of problem (2.3), we have that the feasible set $\mathcal{A}_1$ of problem (2.3) is contained in the feasible set $\mathcal{A}_2$ of problem (2.4). Hence from the optimality of $(\mathbf{u}_\epsilon, \mathbf{w}_\epsilon)$, it yields that

$$\Phi_\epsilon(\mathbf{u}_\epsilon, \mathbf{w}_\epsilon) \leq \Phi_\epsilon(\mathbf{u}, \mathbf{w}), \quad \text{for any } (\mathbf{u}, \mathbf{w}) \in \mathcal{A}_1.$$

Then taking limitation on both sides of the above inequality as $\epsilon \downarrow 0$, we have

$$\Phi(\mathbf{u}^*, \mathbf{w}^*) \leq \Phi(\mathbf{u}, \mathbf{w}), \ \text{ for any } (\mathbf{u}, \mathbf{w}) \in \mathcal{A}_1.$$

Moreover, from

$$\begin{aligned} 2\epsilon &\geq \|D_i \mathbf{u}_\epsilon\|_\epsilon - \mathbf{w}_\epsilon Q_i(D_i \mathbf{u}_\epsilon) \\ &\geq \|D_i \mathbf{u}_\epsilon\|_\epsilon - \|Q_i^T \mathbf{w}_\epsilon\| \|D_i \mathbf{u}_\epsilon\| \\ &\geq \|D_i \mathbf{u}_\epsilon\|_\epsilon - \|D_i \mathbf{u}_\epsilon\| \geq 0, \end{aligned} \tag{3.16}$$

we know $(\mathbf{u}^*, \mathbf{w}^*)$ is feasible for problem (2.3), which completes the proof. ∎

The following theorem considers problems (2.3) and (2.4) regarding stationary points.

Theorem 3.5. *Let $(\mathbf{u}_\epsilon, \mathbf{w}_\epsilon)$ be a stationary point of (2.4) with multipliers $\xi^\epsilon, \eta^\epsilon \in \mathbb{R}^m_+$. If $\{\mathbf{u}_\epsilon, \xi^\epsilon\}$ is bounded for all sufficiently small ϵ. Then as ϵ decreases to zero, there exists an accumulation point $(\mathbf{u}^*, \mathbf{w}^*)$ that is a stationary point of (2.3).*

Proof. Following from the feasibility of $\mathbf{w}_\epsilon$ and (3.9), it implies $\eta_i^\epsilon = 0$ for $i \notin J(\mathbf{w}_\epsilon)$. Then by (3.7) we obtain

$$\sum_{i \in J(\mathbf{w}_\epsilon)} \eta_i^\epsilon Q_i Q_i^T \mathbf{w}_\epsilon = -\sum_{i=1}^m b \|D_i \mathbf{u}_\epsilon\|_\epsilon \cdot \mathbf{d}_i(\mathbf{d}_i^T \mathbf{w}_\epsilon) + \frac{1}{2} \sum_{i=1}^m \xi_i^\epsilon Q_i D_i \mathbf{u}_\epsilon$$

which is bounded from the boundedness of $\{\mathbf{u}_\epsilon, \xi^\epsilon\}$ for all sufficiently small ϵ. Furthermore, due to the structure of $Q_i = (\mathbf{e}_i, \mathbf{e}_{m+i})$, $\|Q_i Q_i^T \mathbf{w}_\epsilon\| = \|Q_i^T \mathbf{w}_\epsilon\| = 1$ and $\eta_i^\epsilon \geq 0$ for $i \in J(\mathbf{w}_\epsilon)$, we know

$$\Big\| \sum_{i \in J(\mathbf{w}_\epsilon)} \eta_i^\epsilon Q_i Q_i^T \mathbf{w}_\epsilon \Big\| = \sum_{i \in J(\mathbf{w}_\epsilon)} \eta_i^\epsilon \|Q_i Q_i^T \mathbf{w}_\epsilon\| = \sum_{i \in J(\mathbf{w}_\epsilon)} \eta_i^\epsilon. \tag{3.17}$$

Hence, $\eta_i^\epsilon, i \in J(\mathbf{w}_\epsilon)$ is bounded. Therefore, $\{\eta^\epsilon\}$ is bounded for all sufficiently small ϵ.

By the boundedness of $\{\mathbf{u}_\epsilon, \mathbf{w}_\epsilon, \xi^\epsilon, \eta^\epsilon\}$, there are a subsequence ϵ_k of $\{\epsilon : \epsilon \to 0\}$, $\mathbf{u}^*, \mathbf{w}^*, \xi \geq 0, \eta \geq 0$ and $\pi_i, i \in L := \{i : D_i \mathbf{u}^* = 0\}$ such that

$$\mathbf{u}_{\epsilon_k} \to \mathbf{u}^*, \quad \mathbf{w}_{\epsilon_k} \to \mathbf{w}^*, \quad \xi^{\epsilon_k} \to \xi, \quad \eta^{\epsilon_k} \to \eta$$

and $\frac{D_i \mathbf{u}_{\epsilon_k}}{\|D_i \mathbf{u}_{\epsilon_k}\|_{\epsilon_k}} \to \pi_i, i \in L$. Then taking limit on both sides of (3.6) as $\epsilon_k \to 0$, we have

$$\begin{aligned} 0 = &\sum_{i=1, i \notin L}^m (a + b(d_i^T \mathbf{w}^*)^2 + \xi_i) D_i^T \frac{D_i \mathbf{u}^*}{\|D_i \mathbf{u}^*\|} + \lambda K^T (K\mathbf{u}^* - \mathbf{u}_0) - \sum_{i=1}^m \xi_i D_i^T Q_i^T \mathbf{w}^* \\ &+ \sum_{i \in L} (a + b(d_i^T \mathbf{w}^*)^2 + \xi_i) D_i^T \pi_i. \end{aligned}$$

Then by denoting $\zeta_i = (a + b(d_i^T \mathbf{w}^*)^2 + \xi_i)\pi_i$ for $i \in L$, we obtain that (3.1) holds at $(\mathbf{u}^*, \mathbf{w}^*)$. Similarly, we can derive (3.2) from (3.7). Moreover the feasibility of $(\mathbf{u}_\epsilon, \mathbf{w}_\epsilon)$ for (2.4) together with (3.16) yields that $(\mathbf{u}^*, \mathbf{w}^*)$ is a feasible point of (2.3), which leads to (3.3)-(3.4). This proof is completed. ∎

The following theorem considers problems (2.4) and (2.5) regarding local minimizers, which is a direct application of Lemma 3.2 in this paper and Theorems 4.4 and 4.6 in [16].

Theorem 3.6. *If* $(\mathbf{u}^*, \mathbf{w}^*)$ *is a strict local minimizer of problem (2.4), then* $(\mathbf{u}^*, \mathbf{w}^*)$ *is a local minimizer of problem (2.5) for all* $\sigma > \max_{1\le i\le m} \xi_i$*, where* $\xi \in \mathbb{R}^m_+$ *is the corresponding multiplier vector in (3.7).*

Proof. By Lemma 3.2, MFCQ holds for problem (2.4) at $(\mathbf{u}^*, \mathbf{w}^*)$. Then there exists a σ^* such for all $\sigma > \sigma^*$, $(\mathbf{u}^*, \mathbf{w}^*)$ is a local minimizer of problem (2.5) according to Theorem 4.4 in [16]. In addition, we have $\sigma^* = \max_{1\le i\le m} \xi_i$ by Theorem 4.6 in [16]. ■

Now we consider problems (2.4) and (2.5) regarding their stationary points and optimal solutions.

Theorem 3.7. *Assume that* $(\mathbf{u}^*, \mathbf{w}^*)$ *is a stationary point of problem (2.5) for all* σ *greater than a certain threshold* $\hat{\sigma} > 0$*. Then* $(\mathbf{u}^*, \mathbf{w}^*)$ *is a stationary point of problem (2.4).*

Proof. For simplicity, we introduce the following notations:

$$\varphi_i^{\epsilon *} := \varphi_i^{\epsilon}(\mathbf{u}^*, \mathbf{w}^*), \quad g_i^* := g_i(\mathbf{w}^*),\ i = 1, \dots, m,$$
$$\mathbf{v}^* := -\sum_{i=1}^m 2b\|D_i\mathbf{u}^*\|_\epsilon \cdot \mathbf{d}_i(\mathbf{d}_i^T\mathbf{w}^*).$$

Suppose to the contrary that $(\mathbf{u}^*, \mathbf{w}^*)$ is not a feasible point of problem (2.4). Then there exists some i such that $\varphi_i^{\epsilon}(\mathbf{u}^*, \mathbf{w}^*) > 0$. By the definition (3.13)-(3.15) of a stationary point of (2.5), there are $\rho, \kappa \in \mathbb{R}^m$, which depend on σ, such that

$$\mathbf{v}^* = 2\sum_{i=1}^m \rho_i Q_i Q_i^T \mathbf{w}^* - \sigma \sum_{i:\varphi_i^{\epsilon *}>0} Q_i D_i \mathbf{u}^* - \sigma \sum_{i:\varphi_i^{\epsilon *}\le 0} \kappa_i Q_i D_i \mathbf{u}^*,$$

for any $\sigma > \hat{\sigma}$. Therefore, for any $\beta > \alpha > 1$ and $\sigma_1 > \hat{\sigma}$, by letting $\sigma_2 = \alpha\sigma_1$ and $\sigma_3 = \beta\sigma_1$, there exist corresponding coefficients $\{\kappa_i^j\}, \{\rho_i^j\}$, $j = 1, 2, 3$ with $\kappa_i^j \in [0, 1]$ and $\rho_i^j \ge 0$ such that

$$\mathbf{v}^* = 2\sum_{i=1}^m \rho_i^j Q_i Q_i^T \mathbf{w}^* - \sigma_j \sum_{i:\varphi_i^{\epsilon *}>0} Q_i D_i \mathbf{u}^* - \sigma_j \sum_{i:\varphi_i^{\epsilon *}\le 0} \kappa_i^j Q_i D_i \mathbf{u}^*,\ j = 1, 2, 3. \tag{3.18}$$

Then performing the subtraction of equality (3.18) with different j and division by $\sigma_2 - \sigma_1$ and $\sigma_3 - \sigma_1$, respectively, we have

$$0 = 2\sum_{i:g_i^*=0} \frac{\rho_i^2 - \rho_i^1}{\sigma_1(\alpha - 1)} Q_i Q_i^T \mathbf{w}^* - \sum_{i:\varphi_i^{\epsilon *}>0} Q_i D_i \mathbf{u}^* - \sum_{i:\varphi_i^{\epsilon *}=0} \frac{\alpha\kappa_i^2 - \kappa_i^1}{\alpha - 1} Q_i D_i \mathbf{u}^* \tag{3.19}$$

and

$$0 = 2\sum_{i:g_i^*=0} \frac{\rho_i^3 - \rho_i^1}{\sigma_1(\beta - 1)} Q_i Q_i^T \mathbf{w}^* - \sum_{i:\varphi_i^{\epsilon *}>0} Q_i D_i \mathbf{u}^* - \sum_{i:\varphi_i^{\epsilon *}=0} \frac{\beta\kappa_i^3 - \kappa_i^1}{\beta - 1} Q_i D_i \mathbf{u}^*, \tag{3.20}$$

where we use $\rho_i^j = 0$ if $g_i^* > 0$ and $\kappa_i = 0$ if $\varphi_i^{\epsilon *} < 0$.

Since $Q_i = (\mathbf{e}_i, \mathbf{e}_{m+i})$, (3.19) indicates that $\{i : \varphi_i^{\epsilon *} > 0\} \subseteq \{i : g_i^* = 0\}$. Thus for $i \in \{i : \varphi_i^{\epsilon *} > 0\}$, we have $Q_i(\tilde{\alpha} Q_i^T \mathbf{w}^* - D_i \mathbf{u}^*) = 0$, which implies $\tilde{\alpha} Q_i^T \mathbf{w}^* = D_i \mathbf{u}^*$ by $Q_i = (\mathbf{e}_i, \mathbf{e}_{m+i})$, where $\tilde{\alpha} = \frac{2(\rho_i^2 - \rho_i^1)}{\sigma_1(\alpha - 1)}$. Moreover, from $g_i^* = \|Q_i \mathbf{w}^*\|^2 - 1 = 0$, we have $\|D_i \mathbf{u}^*\| = |\tilde{\alpha}|$ and $\varphi_i^{\epsilon *} = \sqrt{\tilde{\alpha}^2 + \epsilon^2} - \tilde{\alpha} - 2\epsilon > 0$, which further imply $\tilde{\alpha} < 0$ by contradiction. It follows that $\rho_i^2 < \rho_i^1$. Subtracting (3.19) from (3.20), we obtain

$$0 = 2 \sum_{i : g_i^* = 0} \left(\frac{\rho_i^2 - \rho_i^1}{\sigma_1(\alpha - 1)} - \frac{\rho_i^3 - \rho_i^1}{\sigma_1(\beta - 1)}\right) Q_i Q_i^T \mathbf{w}^* + \sum_{i : \varphi_i^{\epsilon *} = 0} \left(\frac{\alpha \kappa_i^2 - \kappa_i^1}{\alpha - 1} - \frac{\beta \kappa_i^3 - \kappa_i^1}{\beta - 1}\right) Q_i D_i \mathbf{u}^*. \tag{3.21}$$

As the vector group $\{Q_i Q_i^T \mathbf{w}^*,\ i \in J(\mathbf{w}^*)\}$ is linearly independent, the coefficients of $Q_i Q_i^T \mathbf{w}^*$ in (3.21) for $i \in \{i : \varphi_i^{\epsilon *} > 0\}$ must be zero, i.e.,

$$\frac{\rho_i^2 - \rho_i^1}{\sigma_1(\alpha - 1)} = \frac{\rho_i^3 - \rho_i^1}{\sigma_1(\beta - 1)},$$

which derives

$$\rho_i^3 = \frac{\beta - 1}{\alpha - 1}(\rho_i^2 - \rho_i^1) + \rho_i^1.$$

As $\rho_i^3 \geq 0$, for $i \in \{i : \varphi_i^{\epsilon *} > 0\}$, which is non-empty by assumption, β is upper bounded by

$$\beta \leq \frac{\rho_i^1}{\rho_i^1 - \rho_i^2}(\alpha - 1) + 1,$$

which contradicts the arbitrariness of β. Therefore, $(\mathbf{u}^*, \mathbf{w}^*)$ is a feasible point of problem (2.4), i.e., $\varphi_i^{\epsilon *} \leq 0$ and $g_i^* \leq 0$ for $i = 1, \cdots, m$.

As $(\mathbf{u}^*, \mathbf{w}^*)$ is a stationary point of problem (2.5), conditions (3.13)-(3.15) hold at $(\mathbf{u}^*, \mathbf{w}^*)$. Then it is equivalent to that there exist coefficients $\kappa_i^* \in [0, 1], i = 1, \ldots, m$ such that

$$\begin{gathered}
\sum_{i=1}^m \big(a + b(\mathbf{d}_i^T \mathbf{w}^*)^2\big) D_i^T \frac{D_i \mathbf{u}^*}{\|D_i \mathbf{u}^*\|_\epsilon} + \lambda K^T(K\mathbf{u}^* - \mathbf{u}_0) + \sigma \sum_{i : \varphi_i^{\epsilon *} = 0} \kappa_i^* \big(D_i^T \frac{D_i \mathbf{u}^*}{\|D_i \mathbf{u}^*\|_\epsilon} - D_i^T Q_i^T \mathbf{w}^*\big) = 0, \\
\sum_{i=1}^m 2b \|D_i \mathbf{u}^*\|_\epsilon \cdot \mathbf{d}_i(\mathbf{d}_i^T \mathbf{w}^*) - \sigma \sum_{i : \varphi_i^{\epsilon *} = 0} \kappa_i^* Q_i D_i \mathbf{u}^* + 2 \sum_{i=1}^m \rho_i Q_i Q_i^T \mathbf{w}^* = 0, \\
\min\{\rho_i, 1 - \|Q_i^T \mathbf{w}^*\|^2\} = 0,\ i = 1, \ldots, m.
\end{gathered}$$

Taking $\eta_i = \rho_i$ for $i = 1, \ldots, m$, $\xi_i = \sigma \kappa_i^*$ if $\varphi_i^{\epsilon *} = 0$ and $\xi_i = 0$ if $\varphi_i^{\epsilon *} < 0$ ensures that conditions (3.6)-(3.9) hold at $(\mathbf{u}^*, \mathbf{w}^*)$. Therefore, $(\mathbf{u}^*, \mathbf{w}^*)$ satisfies the KKT conditions for problem (2.4) which completes the proof. ■

Corollary 3.8. *Assume that $(\mathbf{u}^*, \mathbf{w}^*)$ is an optimal point of problem (2.5) for all σ greater than a certain threshold $\hat{\sigma} > 0$. Then it is also an optimal solution of problem (2.4).*

Proof. By Theorem 3.7, $(\mathbf{u}^*, \mathbf{w}^*)$ is feasible for problem (2.4). Then for any feasible point $(\mathbf{u}, \mathbf{w})$ of problem (2.4), we have

$$\Phi_\epsilon(\mathbf{u}^*, \mathbf{w}^*) = \Psi_{\epsilon,\sigma}(\mathbf{u}^*, \mathbf{w}^*) \le \Psi_{\epsilon,\sigma}(\mathbf{u}, \mathbf{w}) = \Phi_\epsilon(\mathbf{u}, \mathbf{w}),$$

which implies $(\mathbf{u}^*, \mathbf{w}^*)$ is an optimal solution of problem (2.4). ∎

Remark 3.9. In this section, we establish the theoretical relationships between the discrete Euler's elastica model (2.3), the smoothing relaxation problem (2.4) and the penalty problem (2.5) regarding their optimal solutions and stationary points. From these theoretical results, optimal solutions and stationary points of problem (2.5) with a large penalty parameter σ and a sufficiently small smoothing parameter ϵ are good approximate optimal solutions and stationary points of the discrete Euler's elastica model (2.3), respectively. Moreover, when ϵ decreases to zero, any accumulation point of optimal solutions or stationary points of problem (2.5) is that of (2.3) under certain conditions. According to the relationships, the discrete Euler's elastica model (2.3) can be solved via problem (2.5). Compared with (2.3), problem (2.5) is easier to solve, since the feasible set of (2.5) is convex and the functions Φ_ϵ and φ_i^ϵ in the objective are continuously differentiable. Moreover, problem (2.5) is convex in the $\mathbf{u}$-subspace and $\mathbf{w}$-subspace. Inspired by this special structure, efficient optimization algorithms can be developed.

4. Smoothing BCD method and convergence analysis.

In this section, we will present a smoothing block coordinate descent (BCD) method for solving problem (2.5). Considering that $\Psi_{\epsilon,\sigma}(\mathbf{u}, \mathbf{w})$ is locally Lipschitz continuous but nondifferentiable, we first give a definition of its smoothing function.

Definition 4.1. *We call $\tilde{\Psi}_{\epsilon,\sigma} : \mathbb{R}^m \times \mathbb{R}^{2m} \times \mathbb{R}^+ \to \mathbb{R}$ a smoothing function of $\Psi_{\epsilon,\sigma}$, if $\tilde{\Psi}_{\epsilon,\sigma}(\cdot, \cdot, \mu)$ is continuously differentiable in $\mathbb{R}^m \times \mathbb{R}^{2m}$ for any fixed $\mu > 0$, and satisfies the following two conditions for any $(\mathbf{u}, \mathbf{w}) \in \mathbb{R}^m \times \mathbb{R}^{2m}$,*

- *function value consistency*

$$\lim_{\substack{\mu \downarrow 0 \\ (\tilde{\mathbf{u}}, \tilde{\mathbf{w}}) \to (\mathbf{u}, \mathbf{w})}} \tilde{\Psi}_{\epsilon,\sigma}(\tilde{\mathbf{u}}, \tilde{\mathbf{w}}, \mu) = \Psi_{\epsilon,\sigma}(\mathbf{u}, \mathbf{w}),$$

- *gradient consistency*

$$\mathrm{co}\{\lim \nabla \tilde{\Psi}_{\epsilon,\sigma}(\tilde{\mathbf{u}}, \tilde{\mathbf{w}}, \mu) : (\tilde{\mathbf{u}}, \tilde{\mathbf{w}}) \to (\mathbf{u}, \mathbf{w}), \mu \downarrow 0\} = \partial \Psi_{\epsilon,\sigma}(\mathbf{u}, \mathbf{w}),$$

where "co" *denotes the convex hull.*

To obtain a smoothing function of $\Psi_{\epsilon,\sigma}$, we use the smoothing function

$$\phi(z, \mu) = \begin{cases} (z)_+, & \text{if } |z| > \frac{\mu}{2}, \\ \frac{1}{2\mu}(z + \frac{\mu}{2})^2, & \text{if } |z| \le \frac{\mu}{2} \end{cases} \tag{4.1}$$

to approximate the plus function $(z)_+$. For other smoothing functions in image processing, interested readers are referred to [10, 11]. Let

$$\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}, \mathbf{w}, \mu) = \Phi_\epsilon(\mathbf{u}, \mathbf{w}) + \sigma \sum_{i=1}^{m} \tilde{\varphi}_i^\epsilon(\mathbf{u}, \mathbf{w}, \mu),$$

where $\tilde{\varphi}_i^\epsilon(\mathbf{u},\mathbf{w},\mu) := \phi(\varphi_i^\epsilon(\mathbf{u},\mathbf{w}),\mu)$ for $i = 1,\dots,m$. Since Φ_ϵ is continuously differentiable, and $(\varphi_i^\epsilon)_+$ is regular, $\tilde{\Psi}_{\epsilon,\sigma}$ is a smoothing function of $\Psi_{\epsilon,\sigma}$ by Theorem 1 in [9].

4.1. Algorithms. By replacing the objective function $\Psi_{\epsilon,\sigma}$ in (2.5) with its smoothing approximation $\tilde{\Psi}_{\epsilon,\sigma}$, we obtain the following problem

$$\begin{aligned} &\min_{\mathbf{u},\mathbf{w}} \ \tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u},\mathbf{w},\mu), \text{ where } \tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u},\mathbf{w},\mu) = \Phi_\epsilon(\mathbf{u},\mathbf{w}) + \sigma\sum_{i=1}^m \tilde{\varphi}_i^\epsilon(\mathbf{u},\mathbf{w},\mu) \\ &\text{s.t. } \ \|Q_i^T\mathbf{w}\|^2 \le 1, \ i = 1,\cdots,m. \end{aligned} \tag{4.2}$$

Recall that for $i = 1,\dots,m$, $Q_i = (\mathbf{e}_i, \mathbf{e}_{m+i})$, so $Q_i^T\mathbf{w} = (w_i, w_{m+i})^T \in \mathbb{R}^2$ where w_i and w_{m+i} are the i-th and $(m+i)$-th element of $\mathbf{w}$, respectively. By defining $\mathbf{w}_i = Q_i^T\mathbf{w}$, we obtain a partition of $\mathbf{w}$: $(\mathbf{w}_1,\dots,\mathbf{w}_m)$. By taking advantage of the problem structure, we now present a smoothing block coordinate descent algorithm for solving (4.2).

Algorithm 1

Choose positive parameters c, $\theta \in (0,1)$ and initial variables $\mathbf{u}^0 = \mathbf{u}_0 \in \mathbb{R}^m$, $\mathbf{w}^0 = \mathbf{0} \in \mathbb{R}^{2m}$, $\mu^0 > 0$.

For $k \ge 0$:

- **STEP 1** Solve

$$\min_{\mathbf{u}} \tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u},\mathbf{w}^k,\mu^k) + \frac{c}{2}\|\mathbf{u}-\mathbf{u}^k\|^2 \tag{4.3}$$

  to obtain $\mathbf{u}^{k+1}$.
- **STEP 2** For $i = 1,\dots,m$, solve

$$\min_{\|\mathbf{w}_i\|^2\le 1} \tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k+1}, \mathbf{w}^{k+1}_{1\le j<i}, \mathbf{w}_i, \mathbf{w}^k_{i<j\le m}, \mu^k) \tag{4.4}$$

  to obtain $\mathbf{w}^{k+1} = \sum_{i=1}^m Q_i\mathbf{w}_i$.
- **STEP 3** If the stopping criteria are satisfied, terminate the algorithm. Otherwise, let $\mu^{k+1} = \theta\mu^k$, $k = k+1$ and go to Step 1.

The stopping criteria in Step 3 depend on optimality conditions for (4.3) and (4.4). More details will be given in section 5. In the following, we will first study $\mathbf{u}$-subproblem in (4.3) and then propose Algorithm 2 to solve it. Secondly, we will consider solutions to $\mathbf{w}_i$-subproblem (4.4). We will show that each $\mathbf{w}_i$-subproblem has a unique solution which is easy to obtain.

- **u-subproblem.** The objective function in (4.3) has the form

$$\Phi_\epsilon(\mathbf{u},\mathbf{w}^k) + \frac{c}{2}\|\mathbf{u}-\mathbf{u}^k\|^2 + \sigma\sum_{i=1}^m \tilde{\varphi}_i^\epsilon(\mathbf{u},\mathbf{w}^k,\mu^k). \tag{4.5}$$

  By the definition of the smoothing function in (4.1), $\tilde{\varphi}_i^\epsilon(\mathbf{u},\mathbf{w}^k,\mu) = \frac{1}{2\mu}(\varphi_i^\epsilon(\mathbf{u},\mathbf{w}^k) + \frac{\mu}{2})^2$ if $|\varphi_i^\epsilon(\mathbf{u},\mathbf{w}^k)| \le \frac{\mu}{2}$. To solve (4.3) efficiently by a fixed point method, we split $\tilde{\varphi}_i^\epsilon(\mathbf{u},\mathbf{w}^k,\mu^k)$ into two parts for $|\varphi_i^\epsilon(\mathbf{u},\mathbf{w}^k)| \le \frac{\mu}{2}$, namely, we write

$$\Phi_\epsilon(\mathbf{u},\mathbf{w}^k) + \frac{c}{2}\|\mathbf{u}-\mathbf{u}^k\|^2 + \sigma\sum_{i=1}^m \tilde{\varphi}_i^\epsilon(\mathbf{u},\mathbf{w}^k,\mu^k) = h^k(\mathbf{u}) + \frac{\sigma}{\mu}\sum_{i:|\varphi_i^\epsilon|\le\frac{\mu}{2}} f_i^k(\mathbf{u}) \tag{4.6}$$

where

$$h^k(\mathbf{u}) = \Phi_\epsilon(\mathbf{u}, \mathbf{w}^k) + \frac{c}{2}\|\mathbf{u}-\mathbf{u}^k\|^2 + \sigma \sum_{i:|\varphi_i^\epsilon|>\frac{\mu}{2}} \varphi_i^\epsilon(\mathbf{u}, \mathbf{w}^k)$$
$$+ \frac{\sigma}{2\mu} \sum_{i:|\varphi_i^\epsilon|\le\frac{\mu}{2}} \left(\|D_i\mathbf{u}\|_\epsilon^2 - (4\epsilon-\mu)\|D_i\mathbf{u}\|_\epsilon + ((\mathbf{w}^k)^T Q_i D_i \mathbf{u} + 2\epsilon - \frac{\mu}{2})^2 \right)$$

and $f_i^k(\mathbf{u}) = -((\mathbf{w}^k)^T Q_i D_i \mathbf{u})\|D_i\mathbf{u}\|_\epsilon$. Here we use φ_i^ϵ to denote $\varphi_i^\epsilon(\mathbf{u}, \mathbf{w}^k)$ for simplicity.

As (4.6) is strictly convex, its optimal solution also solves the following nonlinear equations:

(4.7)

$$0 = \nabla h^k(\mathbf{u}) + \frac{\sigma}{\mu} \sum_{i:|\varphi_i^\epsilon|\le\frac{\mu}{2}} \nabla f_i^k(\mathbf{u})$$
$$= \sum_{i=1}^m (a + b(\mathbf{d}_i^T\mathbf{w}^k)^2) D_i^T \frac{D_i\mathbf{u}}{\|D_i\mathbf{u}\|_\epsilon} + \lambda K^T(K\mathbf{u}-\mathbf{u}_0) + c(\mathbf{u}-\mathbf{u}^k)$$
$$+ \sigma \sum_{|\varphi_i^\epsilon|>\frac{\mu}{2}} \left(D_i^T \frac{D_i\mathbf{u}}{\|D_i\mathbf{u}\|_\epsilon} - D_i^T Q_i^T \mathbf{w}^k \right)$$
$$+ \frac{\sigma}{2\mu} \sum_{|\varphi_i^\epsilon|\le\frac{\mu}{2}} \left(2D_i^T D_i\mathbf{u} - (4\epsilon-\mu) D_i^T \frac{D_i\mathbf{u}}{\|D_i\mathbf{u}\|_\epsilon} + (2(\mathbf{w}^k)^T Q_i D_i \mathbf{u} + 4\epsilon - \mu)(D_i^T Q_i^T \mathbf{w}^k) \right)$$
$$+ \frac{\sigma}{\mu} \sum_{|\varphi_i^\epsilon|\le\frac{\mu}{2}} \left(-(\mathbf{w}^k)^T Q_i D_i \mathbf{u} D_i^T \frac{D_i\mathbf{u}}{\|D_i\mathbf{u}\|_\epsilon} - \|D_i\mathbf{u}\|_\epsilon D_i^T Q_i^T \mathbf{w}^k \right).$$

∎

We next present an iterative algorithm to solve (4.7).

Algorithm 2

Set positive parameters: $d^k, \varepsilon_{tol} > 0$. Initialize variables: $\mathbf{u}^t = \mathbf{u}^k$ with $t = 0$.

For $t \ge 0$:

- **STEP 1** Solve the following equation with respect to $\mathbf{u}$

$$0 = \nabla h^k(\mathbf{u}) + \frac{\sigma}{\mu} \sum_{i:|\varphi_i^\epsilon|\le\frac{\mu}{2}} \nabla f_i^k(\mathbf{u}^t) + d^k(\mathbf{u}-\mathbf{u}^t) \quad (4.8)$$

  obtaining $\mathbf{u}^{t+1}$.
- **STEP 2** If the residual of (4.7) at $\mathbf{u}^{t+1}$ is no more than ε_{tol}, terminate the algorithm and return $\mathbf{u}^{k+1} = \mathbf{u}^{t+1}$. Otherwise, let $t = t+1$ and go to Step 1.

In Algorithm 2, we choose

$$d^k = 2\tilde{\epsilon} + \sum_{i:|\varphi_i^\epsilon|\le\frac{\mu}{2}} \frac{\sigma d_i^k}{\mu},$$

where $\tilde{\epsilon}$ is a positive constant and d_i^k is an upper bound of Lipschitz constant of ∇f_i^k for i such that $|\varphi_i^\epsilon| \le \frac{\mu}{2}$. Thus, for each $t \ge 0$, problem (4.8) is an approximation to (4.7) at $\mathbf{u}^t$. Moreover, it is easy to check that (4.8) is equivalent to minimize $F_t^k(\mathbf{u})$, where

$$F_t^k(\mathbf{u}) = h^k(\mathbf{u}) + \frac{\sigma}{\mu} \sum_{i:|\varphi_i^\epsilon| \le \frac{\mu}{2}} (f_i^k(\mathbf{u}^t) + (\nabla f_i^k(\mathbf{u}^t))^T(\mathbf{u} - \mathbf{u}^t)) + \frac{d^k}{2}\|\mathbf{u} - \mathbf{u}^t\|^2.$$

Thus $\nabla F_t^k(\mathbf{u}^{t+1}) = 0$ at the minimizer $\mathbf{u}^{t+1}$. Function F_t^k will play a key role in proving the convergence of Algorithm 2, as shown in Lemma 4.2.

In the following numerical experiments, we apply the lagged diffusivity fixed point method to solve (4.8). The lagged diffusivity approach was firstly proposed for solving TV functional minimization in [28] with theoretical convergence analyzed in [8].

- **$\mathbf{w}_i$-subproblem.** Notice that for each $\mathbf{w}_i$-subproblem, $i = 1, \dots, m$, (4.4) can be reformulated as a two-dimensional ball constrained optimization problem as below:

$$\min_{\|\mathbf{w}_i\| \le 1} \ \mathbf{w}_i^T P_i^k \mathbf{w}_i + (\mathbf{q}_i^k)^T \mathbf{w}_i, \tag{4.9}$$

where $\mathbf{w}_i = Q_i^T \mathbf{w}$ and P_i^k is a symmetric and positive definite matrix. For illustrations we present here a special case of $\mathbf{w}_i$-subproblem with the i-th pixel point not on the boundary of $\bar{\Omega}$ and $\tilde{\varphi}_i^\epsilon = \frac{1}{2\mu}(\varphi_i^\epsilon(\mathbf{u}, \mathbf{w}) + \frac{\mu}{2})^2$. In this case the $\mathbf{w}_i$-subproblem is of the form:

$$\begin{aligned}
\min_{\|\mathbf{w}_i\|^2 \le 1} \ & b\|D_i \mathbf{u}^{k+1}\|_\epsilon \big((1\ 1) \cdot \mathbf{w}_i - (w_{i+1}^k + w_{m+i+n_1}^k)\big)^2 \\
& + b\|D_{i-1} \mathbf{u}^{k+1}\|_\epsilon \big((1\ 0) \cdot \mathbf{w}_i - (w_{i-1}^{k+1} + w_{m+i-1}^{k+1} - w_{m+n_1+i-1}^k)\big)^2 \\
& + b\|D_{i-n_1} \mathbf{u}^{k+1}\|_\epsilon \big((0\ 1) \cdot \mathbf{w}_i - (w_{i-n_1}^k + w_{m+i-n_1}^k - w_{i-n_1+1}^k)\big)^2 \\
& + \frac{\sigma}{2\mu^k} \left((D_i \mathbf{u}^{k+1})^T \mathbf{w}_i - \|D_i \mathbf{u}^{k+1}\|_\epsilon + 2\epsilon - \frac{\mu^k}{2} \right)^2 .
\end{aligned} \tag{4.10}$$

Then we can reformulate (4.10) into the form of (4.9) where

$$\begin{aligned}
P_i^k = & \, b\|D_i \mathbf{u}^{k+1}\|_\epsilon \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + b \begin{pmatrix} \|D_{i-1}\mathbf{u}^{k+1}\|_\epsilon & 0 \\ 0 & \|D_{i-n_1}\mathbf{u}^{k+1}\|_\epsilon \end{pmatrix} \\
& + \frac{\sigma}{2\mu^k} (D_i \mathbf{u}^{k+1})(D_i \mathbf{u}^{k+1})^T
\end{aligned}$$

which is symmetric and positive definite and $\mathbf{q}_i^k = \binom{q_i^k(1)}{q_i^k(2)}$ with

$$\begin{aligned}
q_i^k(1) = & -2b\|D_i \mathbf{u}^{k+1}\|_\epsilon (w_{i+1}^k + w_{m+i+n_1}^k) - \frac{\sigma}{\mu^k}(\|D_i \mathbf{u}^{k+1}\|_\epsilon - 2\epsilon + \frac{\mu^k}{2}) u_{i1}^{k+1} \\
& - 2b\|D_{i-1}\mathbf{u}^{k+1}\|_\epsilon (w_{i-1}^{k+1} + w_{m+i-1}^{k+1} - w_{m+n_1+i-1}^k), \\
q_i^k(2) = & -2b\|D_i \mathbf{u}^{k+1}\|_\epsilon (w_{i+1}^k + w_{m+i+n_1}^k) - \frac{\sigma}{\mu^k}(\|D_i \mathbf{u}^{k+1}\|_\epsilon - 2\epsilon + \frac{\mu^k}{2}) u_{i2}^{k+1} \\
& - 2b\|D_{i-n_1}\mathbf{u}^{k+1}\|_\epsilon (w_{i-n_1}^{k+1} + w_{m+i-n_1}^{k+1} - w_{i-n_1+1}^{k+1}).
\end{aligned}$$

Here u_{i1}^{k+1} and u_{i2}^{k+1} correspond to the first and second elements of $D_i\mathbf{u}^{k+1}$ respectively. By optimality conditions for (4.9) it has a unique solution

$$\mathbf{w}_i^{k+1} = -(2P_i^k + 2\tau_i^k I_2)^{-1}\mathbf{q}_i^k, \tag{4.11}$$

where $I_2 \in \mathbb{R}^{2\times 2}$ is the identity matrix and τ_i^k is zero if $\|(2P_i^k)^{-1}\mathbf{q}_i^k\|^2 \leq 1$ and a positive number satisfying $\|(2P_i^k + 2\tau_i^k I_2)^{-1}\mathbf{q}_i^k\|^2 = 1$ otherwise, where the equation is quartic with respect to τ_i^k and has a unique positive root.

4.2. Convergence analysis. The following lemma establishes the convergence of Algorithm 2.

Lemma 4.2. *Let the sequence $\{\mathbf{u}^t\}$ be generated by Algorithm 2. Then it converges to the solution of problem (4.5).*

Proof. Recall that problem (4.5) is strictly convex, thus it has a unique solution. As $\mathbf{u}^{t+1}$ minimizes F_t^k, we have $F_t^k(\mathbf{u}^{t+1}) \leq F_t^k(\mathbf{u}^t)$. Recall that d_i^k was introduced in Algorithm 2 to denote the upper bound of Lipschitz constant of ∇f_i^k. Then by Taylor's theorem we have that for every i,

$$f_i^k(\mathbf{u}^{t+1}) \leq f_i^k(\mathbf{u}^t) + (\nabla f_i^k(\mathbf{u}^t))^T(\mathbf{u}^{t+1} - \mathbf{u}^t) + \frac{d_i^k}{2}\|\mathbf{u}^{t+1} - \mathbf{u}^t\|^2.$$

By the definition of F_t^k and above inequality, we can derive that

$$\begin{aligned}
&F_{t+1}^k(\mathbf{u}^{t+1}) + \tilde{\epsilon}\|\mathbf{u}^{t+1} - \mathbf{u}^t\|^2 \\
&= h^k(\mathbf{u}^{t+1}) + \left(\frac{d^k}{2} - \sum_{i:|\varphi_i^\epsilon|\leq\frac{\mu}{2}} \frac{\sigma d_i^k}{2\mu}\right)\|\mathbf{u}^{t+1} - \mathbf{u}^t\|^2 + \frac{\sigma}{\mu}\sum_{i:|\varphi_i^\epsilon|\leq\frac{\mu}{2}} f_i^k(\mathbf{u}^{t+1}) \\
&\leq h^k(\mathbf{u}^{t+1}) + \left(\frac{d^k}{2} - \sum_{i:|\varphi_i^\epsilon|\leq\frac{\mu}{2}} \frac{\sigma d_i^k}{2\mu}\right)\|\mathbf{u}^{t+1} - \mathbf{u}^t\|^2 \\
&\quad + \frac{\sigma}{\mu}\sum_{i:|\varphi_i^\epsilon|\leq\frac{\mu}{2}}\left(f_i^k(\mathbf{u}^t) + (\nabla f_i^k(\mathbf{u}^t))^T(\mathbf{u}^{t+1} - \mathbf{u}^t) + \frac{d_i^k}{2}\|\mathbf{u}^{t+1} - \mathbf{u}^t\|^2\right) \\
&= h^k(\mathbf{u}^{t+1}) + \frac{d^k}{2}\|\mathbf{u}^{t+1} - \mathbf{u}^t\|^2 + \frac{\sigma}{\mu}\sum_{i:|\varphi_i^\epsilon|\leq\frac{\mu}{2}}\left(f_i^k(\mathbf{u}^t) + (\nabla f_i^k(\mathbf{u}^t))^T(\mathbf{u}^{t+1} - \mathbf{u}^t)\right) \\
&= F_t^k(\mathbf{u}^{t+1}) \leq F_t^k(\mathbf{u}^t).
\end{aligned}$$

Therefore, $\{F_t^k(\mathbf{u}^t)\}$ is a monotonically decreasing sequence and lower bounded, then it converges and

$$\lim_{t\to\infty} \tilde{\epsilon}\|\mathbf{u}^{t+1} - \mathbf{u}^t\|^2 \leq \lim_{t\to\infty} F_t^k(\mathbf{u}^t) - F_{t+1}^k(\mathbf{u}^{t+1}) = 0,$$

which implies $\lim_{t\to\infty}\|\mathbf{u}^{t+1} - \mathbf{u}^t\| = 0$. Furthermore, since $F_t^k(\mathbf{u})$ is level bounded, the sequence $\{\mathbf{u}^t\}$ is bounded and there is a subsequence $\{\mathbf{u}^{t_l}\}$ converging to $\mathbf{u}^*$ and satisfying $\nabla F_{t_l}^k(\mathbf{u}^{t_l+1}) = 0$ by the optimality condition for minimizing $F_{t_l}^k$. Then we obtain

$$
\begin{aligned}
0 &= \nabla F_{t_l}^k(\mathbf{u}^{t_l+1}) = \lim_{l\to\infty} \nabla F_{t_l}^k(\mathbf{u}^{t_l+1}) \\
&= \lim_{l\to\infty}\left(\nabla h^k(\mathbf{u}^{t_l+1}) + \frac{\sigma}{\mu}\sum_{i:|\varphi_i^\epsilon|\le\frac{\mu}{2}} \nabla f_i^k(\mathbf{u}^{t_l}) + d^k(\mathbf{u}^{t_l+1}-\mathbf{u}^{t_l})\right) \\
&= \nabla h^k(\mathbf{u}^*) + \frac{\sigma}{\mu}\sum_{i:|\varphi_i^\epsilon|\le\frac{\mu}{2}} \nabla f_i^k(\mathbf{u}^*) \qquad (4.12)
\end{aligned}
$$

where $\mathbf{u}^{t_l+1} \to \mathbf{u}^*$ is from $\lim_{t\to\infty}\|\mathbf{u}^{t+1}-\mathbf{u}^t\| = 0$. Consequently, (4.12) indicates that $\mathbf{u}^*$ is the solution of problem (4.5), which completes the proof. ∎

Lemma 4.3. *Let $\{(\mathbf{u}^k,\mathbf{w}^k)\}$ be generated by Algorithm 1. Then $\{(\mathbf{u}^k,\mathbf{w}^k)\}$ is bounded and*

$$
\lim_{k\to\infty}\|(\mathbf{u}^{k+1},\mathbf{w}^{k+1})-(\mathbf{u}^k,\mathbf{w}^k)\| = 0. \qquad (4.13)
$$

Proof. Following the framework in Algorithm 1, the sequence $\{(\mathbf{u}^k,\mathbf{w}^k)\}$ satisfies

$$
\tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^k,\mu^k) + \frac{c}{2}\|\mathbf{u}^{k+1}-\mathbf{u}^k\|^2 \le \tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^k,\mathbf{w}^k,\mu^k). \qquad (4.14)
$$

For $i=1,\dots,m$, as $\mathbf{w}_i^{k+1}$ is the solution of problem (4.9), by (4.11) it also solves the problem

$$
\min_{\mathbf{w}_i\in\mathbb{R}^2} \quad \mathbf{w}_i^T P_i^k \mathbf{w}_i + (\mathbf{q}_i^k)^T\mathbf{w}_i + \tau_i^k(\|\mathbf{w}_i\|^2-1). \qquad (4.15)
$$

Note that the objective function in (4.15) is strongly convex, $\tau_i^k(\|\mathbf{w}_i^{k+1}\|^2-1)=0$ and $P_i^k+\tau_i^k I_2 \succeq P_i^k \succeq b\epsilon I_2$ for $i=1,\dots,m$, where $A\succeq B$ means that $A-B$ is positive semidefinite for symmetric matrices A and B with the same dimension. Then we obtain

$$
\begin{aligned}
(\mathbf{w}_i^k)^T P_i^k\mathbf{w}_i^k + \mathbf{q}_k^T\mathbf{w}_i^k &\ge (\mathbf{w}_i^k)^T P_i^k\mathbf{w}_i^k + \mathbf{q}_k^T\mathbf{w}_i^k + \tau_i^k(\|\mathbf{w}_i^k\|^2-1) \\
&\ge (\mathbf{w}_i^{k+1})^T P_i^k\mathbf{w}_i^{k+1} + \mathbf{q}_k^T\mathbf{w}_i^{k+1} + \tau_i^k(\|\mathbf{w}_i^{k+1}\|^2-1) + \frac{b\epsilon}{2}\|\mathbf{w}_i^k-\mathbf{w}_i^{k+1}\|^2, \\
&= (\mathbf{w}_i^{k+1})^T P_i^k\mathbf{w}_i^{k+1} + \mathbf{q}_k^T\mathbf{w}_i^{k+1} + \frac{b\epsilon}{2}\|\mathbf{w}_i^k-\mathbf{w}_i^{k+1}\|^2,
\end{aligned}
$$

which implies

$$
\tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^{k+1}_{1\le j<i},\mathbf{w}^k_{i\le j\le m},\mu^k) \ge \tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^{k+1}_{1\le j\le i},\mathbf{w}^k_{i<j\le m},\mu^k) + \frac{b\epsilon}{2}\|\mathbf{w}_i^k-\mathbf{w}_i^{k+1}\|^2.
$$

Thus we have

$$
\tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^{k+1},\mu^k) + \frac{b\epsilon}{2}\|\mathbf{w}^k-\mathbf{w}^{k+1}\|^2 \le \tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^k,\mu^k). \qquad (4.16)
$$

Moreover, as for fixed z the smoothing function $\phi(z,\mu)$ is nondecreasing with respect to $\mu>0$, by $\mu^{k+1}=\theta\mu^k$ with $\theta\in(0,1)$ we obtain

$$
\tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^{k+1},\mu^{k+1}) \le \tilde\Psi_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^{k+1},\mu^k). \qquad (4.17)
$$

By inequalities (4.14), (4.16) and (4.17), we know

$$\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^{k+1},\mu^{k+1})+\frac{b\epsilon}{2}\|\mathbf{w}^k-\mathbf{w}^{k+1}\|^2+\frac{c}{2}\|\mathbf{u}^k-\mathbf{u}^{k+1}\|^2\le\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^k,\mathbf{w}^k,\mu^k),$$

which implies that the sequence $\{\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^k,\mathbf{w}^k,\mu^k)\}$ is monotonically decreasing. Then as $\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u},\mathbf{w},\mu)$ is coercive, it is level bounded, i.e., the sequence set

$$\{(\mathbf{u}^k,\mathbf{w}^k,\mu^k)\}\subseteq\{(\mathbf{u},\mathbf{w},\mu):\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u},\mathbf{w},\mu)\le\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^0,\mathbf{w}^0,\mu^0)\}$$

is bounded, which yields the boundedness of $\{(\mathbf{u}^k,\mathbf{w}^k)\}$. Moreover, $\{\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^k,\mathbf{w}^k,\mu^k)\}$ converges due to its lower boundedness. Therefore, we obtain that

$$\lim_{k\to\infty}\frac{c}{2}\|\mathbf{u}^{k+1}-\mathbf{u}^k\|^2+\frac{b\epsilon}{2}\|\mathbf{w}^k-\mathbf{w}^{k+1}\|^2\le\lim_{k\to\infty}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^k,\mathbf{w}^k,\mu^k)-\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k+1},\mathbf{w}^{k+1},\mu^{k+1})=0,$$

which yields (4.13). This completes the proof. ∎

Theorem 4.4. *Let $\{(\mathbf{u}^k,\mathbf{w}^k)\}$ be generated by Algorithm 1. Then any accumulation point $(\mathbf{u}^*,\mathbf{w}^*)$ is a stationary point of problem (2.5).*

Proof. By Lemma 4.3 there exists a subsequence $\{(\mathbf{u}^{k_l},\mathbf{w}^{k_l})\}$ converging to $(\mathbf{u}^*,\mathbf{w}^*)$. Since $\mathbf{u}^{k_l+1}$ and $\mathbf{w}^{k_l+1}$ are obtained by solving their first order conditions, we have

$$\nabla_{\mathbf{u}}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k_l+1},\mathbf{w}^{k_l},\mu^{k_l})+c(\mathbf{u}^{k_l+1}-\mathbf{u}^{k_l})=0,\tag{4.18}$$

$$\nabla_{\mathbf{w}_i}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k_l+1},\mathbf{w}^{k_l+1}_{1\le j<i},\mathbf{w}^{k_l+1}_i,\mathbf{w}^{k_l}_{i<j\le m},\mu^{k_l})+2\tau_i^{k_l}\mathbf{w}_i^{k_l+1}=0, i=1,\dots,m,\tag{4.19}$$

where $\tau_i^{k_l}$ is the corresponding multiplier defined in (4.11) for $i=1,\dots,m$. And (4.19) implies

$$\|\nabla_{\mathbf{w}_i}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k_l+1},\mathbf{w}^{k_l+1}_{1\le j<i},\mathbf{w}^{k_l+1}_i,\mathbf{w}^{k_l}_{i<j\le m},\mu^{k_l})\|=2\tau_i^{k_l}\|\mathbf{w}_i^{k_l+1}\|, i=1,\dots,m.\tag{4.20}$$

Considering that $\tau_i^{k_l}(\|\mathbf{w}^{k_l+1}\|-1)=0$, we obtain from (4.20) that

$$\frac{1}{2}\|\nabla_{\mathbf{w}_i}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k_l+1},\mathbf{w}^{k_l+1}_{1\le j<i},\mathbf{w}^{k_l+1}_i,\mathbf{w}^{k_l}_{i<j\le m},\mu^{k_l})\|=\tau_i^{k_l}, i=1,\dots,m.\tag{4.21}$$

According to the definition of smoothing function [9], $\nabla_{\mathbf{w}_i}\tilde{\Psi}_{\epsilon,\sigma}$ is bounded, then there exists a convergent subsequence of $\nabla_{\mathbf{w}_i}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k_l+1},\mathbf{w}^{k_l+1}_{1\le j<i},\mathbf{w}^{k_l+1}_i,\mathbf{w}^{k_l}_{i<j\le m},\mu^{k_l})$. Without loss of generality, we still label this subsequence by k_l. Accordingly, there exists a point τ^* such that τ^{k_l} converges to τ^*.

As it follows from $\|\mathbf{u}^{k+1}-\mathbf{u}^k\|\to0$, $\|\mathbf{w}^{k+1}-\mathbf{w}^k\|\to0$ and $\|\mu^{k+1}-\mu^k\|\to0$ that $\{\mathbf{u}^{k_l+1}\}\to\mathbf{u}^*$, $\{\mathbf{w}^{k_l+1}\}\to\mathbf{w}^*$,$\{\mu^{k_l+1}\}\to0$, taking limits on both sides of (4.18) as $k_l\to\infty$ yields

$$\lim_{k_l\to\infty}\nabla_{\mathbf{u}}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^{k_l+1},\mathbf{w}^{k_l},\mu^{k_l})=0.\tag{4.22}$$

Besides, since (4.19) is equivalent to

$$\nabla_{\mathbf{w}_i}\Phi_\epsilon(\mathbf{u}^{k_l+1},\mathbf{w}^{k_l+1}_{1\le j<i},\mathbf{w}^{k_l+1}_i,\mathbf{w}^{k_l}_{i<j\le m})+\sigma\nabla_{\mathbf{w}_i}\tilde{\varphi}_i^\epsilon(\mathbf{u}^{k_l+1},\mathbf{w}_i^{k_l+1},\mu^{k_l})+2\tau_i^{k_l}\mathbf{w}_i^{k_l+1}=0,$$

for $i = 1, \ldots, m$. Taking limits on the above equation, we have

$$-\left(\nabla_{\mathbf{w}_i}\Phi_\epsilon(\mathbf{u}^*, \mathbf{w}^*) + 2\tau_i^*\mathbf{w}_i^*\right) \in \mathrm{co}\left\{\lim_{k_l\to\infty} \sigma\nabla_{\mathbf{w}_i}\tilde{\varphi}_i^\epsilon(\mathbf{u}^{k_l+1}, \mathbf{w}_i^{k_l+1}, \mu^{k_l})\right\},\ i = 1, \ldots, m. \tag{4.23}$$

Since $Q_i(Q_i^T\mathbf{w}^*) = \mathbf{w}_i^*$, then we obtain

$$-\left(\nabla_{\mathbf{w}}\Phi_\epsilon(\mathbf{u}^*, \mathbf{w}^*) + \sum_{i=1}^m 2\tau_i^* Q_i(Q_i^T\mathbf{w}^*)\right) \in \mathrm{co}\left\{\lim_{k_l\to\infty} \sigma\sum_{i=1}^m \nabla_{\mathbf{w}}\tilde{\varphi}_i^\epsilon(\mathbf{u}^{k_l+1}, \mathbf{w}^{k_l+1}, \mu^{k_l})\right\}. \tag{4.24}$$

Hence, from the gradient consistency for the smoothing function in Definition 4.1, (4.22) and (4.24), it follows that $(\mathbf{u}^*, \mathbf{w}^*)$ satisfies conditions (3.13) and (3.14) in Theorem 3.3, and $(\tau_i^*, \mathbf{w}^*)$ satisfies condition (3.15) for $i = 1, \ldots, m$. Consequently, $(\mathbf{u}^*, \mathbf{w}^*)$ is a stationary point of problem (2.5) which completes the proof. ■

5. Numerical Experiments.

In this section, we report numerical experiments using our Algorithm 1 with Algorithm 2 for image inpainting and image denoising by Euler's elastica model. The experiments were performed in MATLAB version R2016b on a laptop of 8GB RAM and Intel Core i5-8350 CPU: @1.70GHz 1.90GHz.

All the stopping criteria in the algorithms are based on optimality conditions. Based on Theorem 4.4 with (4.22) and (4.24), we set the termination criterion in Algorithm 1 as $Res_1 = \max\{r_1, r_2, r_3\} \le 10^{-4}$ where

$$\begin{cases} r_1 = \|\nabla_{\mathbf{u}}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^k, \mathbf{w}^k, \mu^k)\|_\infty, \\ r_2 = \|\nabla_{\mathbf{w}}\tilde{\Psi}_{\epsilon,\sigma}(\mathbf{u}^k, \mathbf{w}^k, \mu^k) + \sum_{i=1}^m \tau_i^k Q_i(Q_i^T\mathbf{w}^k)\|_\infty, \\ r_3 = \max_{1\le i\le m} |\min(\tau_i^k, 1 - \|Q_i^T\mathbf{w}^k\|^2)|. \end{cases}$$

For Algorithm 2 we adopt the stopping criterion $\epsilon_{tol} \le 10^{-4}$ with $\mathbf{w}^k$ at $\mathbf{u}^{t+1}$ outer iteration.

5.1. Choice of parameters.

For model parameters, a and b are positive constant weights associated with the TV term and the curvature term, respectively. The parameter λ balances the weight of the fidelity term relative to the regularization term, ϵ makes the objective function well defined everywhere in the domain Ω and σ is the penalty parameter for the constraint $\varphi_i^\epsilon(\mathbf{u}, \mathbf{w}) \le 0$, $i = 1, \ldots, m$ in (2.4). Besides, there are three parameters μ^0, c and θ for Algorithm 1. Although a big penalty penalty parameter σ can ensure the feasibility of $(\mathbf{u}, \mathbf{w})$ for the constraint $\varphi_i^\epsilon(\mathbf{u}, \mathbf{w}) \le 0$, we experimentally found that a large σ may lead to unsatisfactory numerical results and $\sigma \in [1, 10]$ can have good numerical performance in image inpainting. Parameter $\theta \in (0, 1)$ is the reduction factor of μ^k in the iteration as $\mu^{k+1} = \theta\mu^k$ and c is the proximal coefficient in the $\mathbf{u}$-subproblem. We set $c = 0.01$ in all experiments. In image inpainting, we set $a = 5$, $b = 10$, $\sigma = 1$, $\lambda = 1000$, $\epsilon = 0.1$, $\theta = 0.9$ and $\mu^0 = 0.7$ in Figures 1-2, and $a = 1$, $b = 5$, $\sigma = 1$, $\lambda = 1000$, $\epsilon = 0.001$, $\theta = 0.999$ and $\mu^0 = 0.5$ in Figure 3. Here we use a large λ to ensure that the known information in input data is properly preserved. In image denoising, parameters are set as $a = 1$, $b = 5$, $\sigma = 5$, $\lambda = 2.4$, $\epsilon = 0.0001$, $\theta = 0.9$ and $\mu^0 = 0.1$.

(a) (b) (c)

Figure 1. *(a) Input Image. Red region is inpainting domain. (b) Result by TV model. (c) Result by Euler's elastica model.*

5.2. Image inpainting. In this subsection, we apply the proposed algorithms to image inpainting problems. The matrix K is an diagonal matrix as mentioned in section 2.

We first report in Figure 1 the comparison between Algorithm 1 for Euler's elastica model and the lagged diffusion fixed method method [28] for Total Variation (TV) model ($b = 0$ in (1.1)). The red region in Figure 1(a) is the inpainting domain. Figure 1(b) is the inpainting result by using the lagged diffusion fixed point method in [28] for TV model and while Figure 1(c) is the result by using Algorithm 1 for Euler's elastica model. Although the edge obtained by our proposed method is blurry, we can see that Algorithm 1 for Euler's elastica model performs superior in connecting the disconnected edges.

We also report some results for synthetic images in Figure 2. The red regions in Figures 2(a)-2(c) all represent unknown regions while Figures 2(d)-2(f) are inpainting results. These examples display that our proposed method is effective as the obtained results are visually reasonable and correct. Moreover, we can observe that although the edges are a little bit blurry, Algorithm 1 for Euler's elastica model shows a quality of extending connectivity and connecting the missing region smoothly along the curves of images in inpainting domains.

In Figure 3, we test Algorithm 1 with some real images downloaded from world wide web[2]. The missing pixels shown in red region of Figure 3(a) are chosen randomly and the portion is 50% of image size. The type of inpainting region in Figure 3(a) often appears in archaeological artifacts which cannot be recovered manually. The red lines in Figure 3(c) are made randomly and can simulate the distorted area in some old photos. We can see from Figures 3(b) and 3(d) that features of recovery results are restored well, which can demonstrate that Algorithm 1 for Euler's elastica model can be used in some practical image inpainting problems.

In Table 1, we present the comparison results between Algorithm 1, THC method [26], vanilla Stochastic Gradient Descent (SGD) method [18, 21] and Adaptive Moment Estimation (ADAM) method [19] for Euler's elastica model. THC method is a fast and efficient numerical algorithm by using an augmented Lagrangian approach to solve problem (2.2). However, as discussed in [13], one drawback of THC method is that it is sensitive to some parameters. In recent years, SGD method and ADAM method have been widely recognized as efficient methods for finite-sum nonconvex optimization and used in Matlab, Python Optimization Toolbox for solving unconstrained general finite-sum nonconvex optimization problems [18].

[2]Datasets were downloaded at: http://www.robots.ox.ac.uk/~vgg/data.

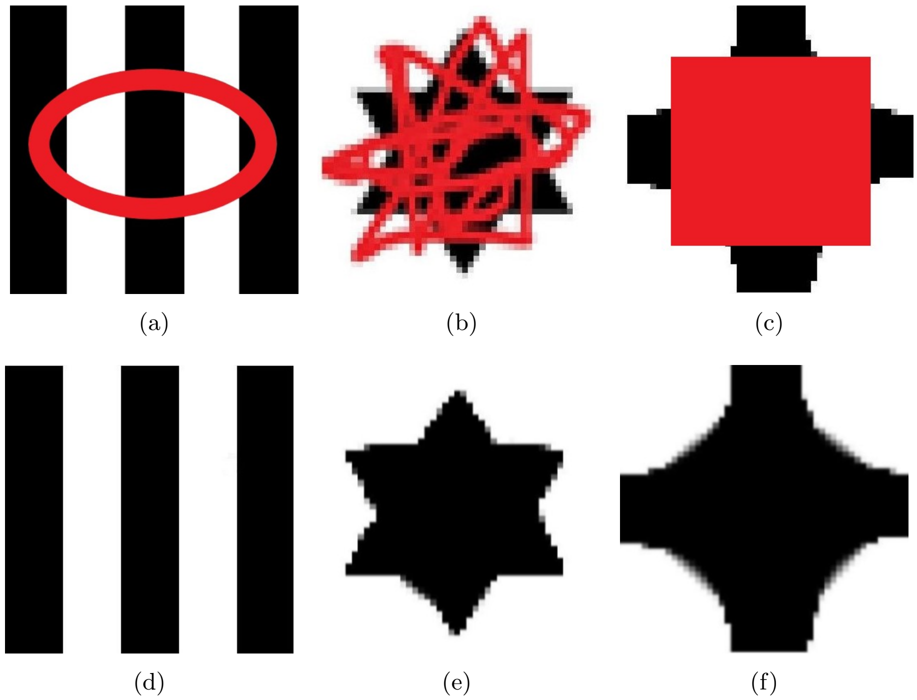
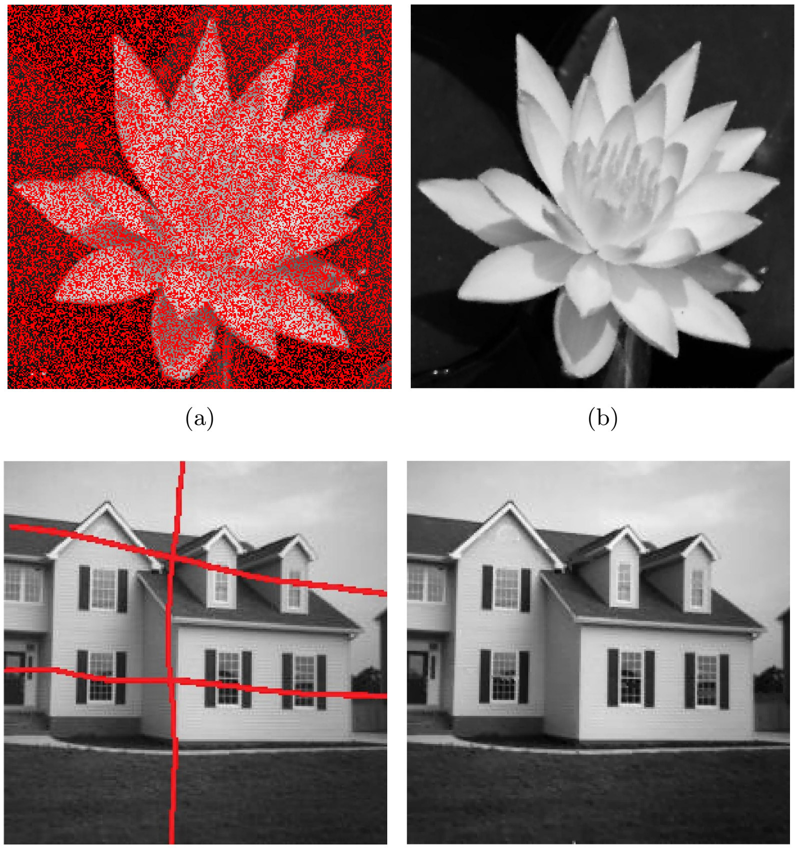


Figure 2. *Input images with red inpainting regions in the first row, results by Algorithm 1 in the second row.*

Table 1
Relative error comparison on images.

RelErr	**Fig. 2(a)**	**Fig. 2(b)**	**Fig. 2(c)**	**Fig. 3(a)**	**Fig. 3(c)**
Algorithm 1	**0.0524**	**0.1316**	**0.2270**	**0.0732**	**0.0360**
THC	0.0947	0.1799	0.2539	0.0918	0.0503
SGD	0.2093	0.1637	0.2956	0.0936	0.0450
ADAM	0.2086	0.1368	0.2563	0.0927	0.0528

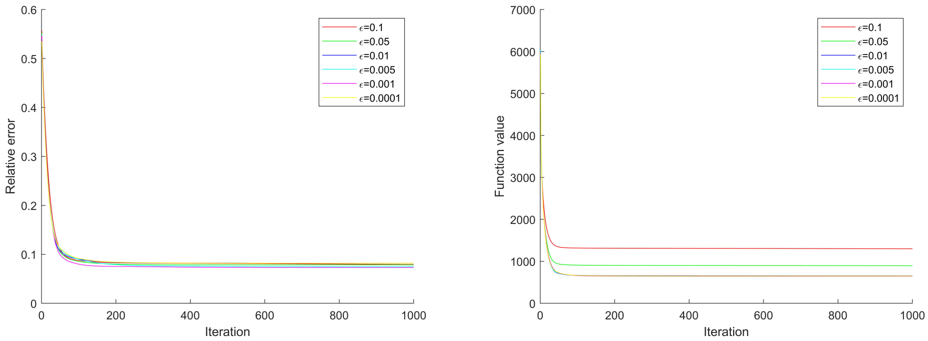
We apply these two methods to solve an unconstrained problem that is obtained by penalizing all constraints in problem (4.2) to the objective function in a quadratic term $\sigma_2 \sum_{i=1}^m \|Q_i^T \mathbf{w}\|^2$. For THC method, the parameters are set same as in [26] and [30]. For SGD method and ADAM method, the sample size is 100, the stepsize is 0.001, the penalty parameter for penalizing constraints in problem (4.2) is $\sigma_2 = 100$, and other parameters are same as ours. Moreover, for ADAM method, the exponential rates for first- and second-moment estimates are set as 0.9 and 0.999, respectively, and the constant for numerical stability is set as 10^{-8}. We implement SGD method and ADAM method by calling solvers **sgd( )** and **adam( )** in a SGDLibrary



(a) (b)

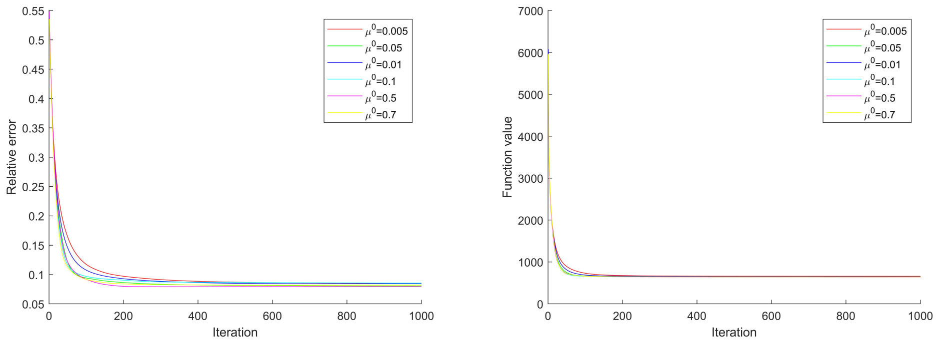
(c) (d)

Figure 3. *The input images with red missing region are shown in left column, while the corresponding results are displayed in right column.*



(a) (b)

Figure 4. *Performances of Algorithm 1 with different ϵ for Figure 3(a).*



(a) (b)

Figure 5. *Performances of Algorithm 1 with* $\theta = 0.999$ *and different* μ^0 *for Figure 3(a).*

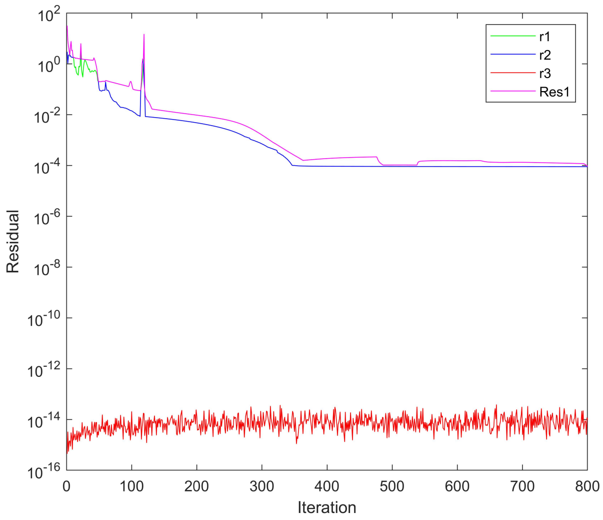
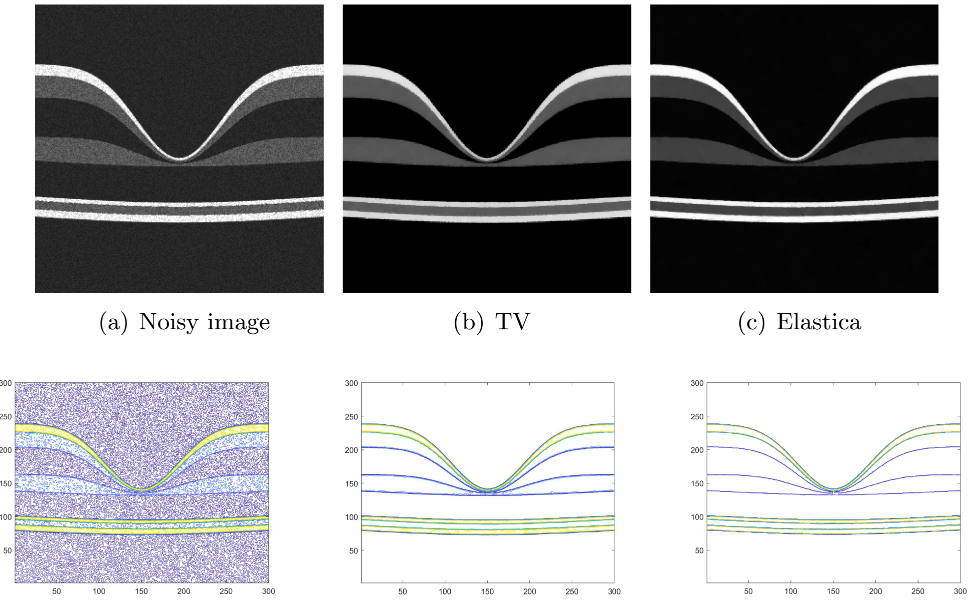


Figure 6. *Values of* r_1*,* r_2*,* r_3 *and* Res_1 *for Figure 3(a).*

[18] in MATLAB. Relative error is defined by

$$\text{RelErr} := \frac{\|\mathbf{u}^k - \mathbf{u}_{org}\|}{\|\mathbf{u}_{org}\|},$$

where $\mathbf{u}_{org}$ is the original image without any inpainting domain and $\mathbf{u}^k$ is the output image. In order to achieve a small relative error, our proposed method is more stable than THC method, SGD method and ADAM method, which are sensitive with some parameters in numerical computation. In Figures 4 and 5, we report performances of Algorithm 1 for Figure 3(a) on relative error as well as function values with respect to different smoothing parameters ϵ and the reduction factor θ and initial smoothing parameter μ^0 for $\mu^{k+1} = \theta\mu^k$.



(a) Noisy image (b) TV (c) Elastica

(d) Noisy image (e) TV (f) Elastica

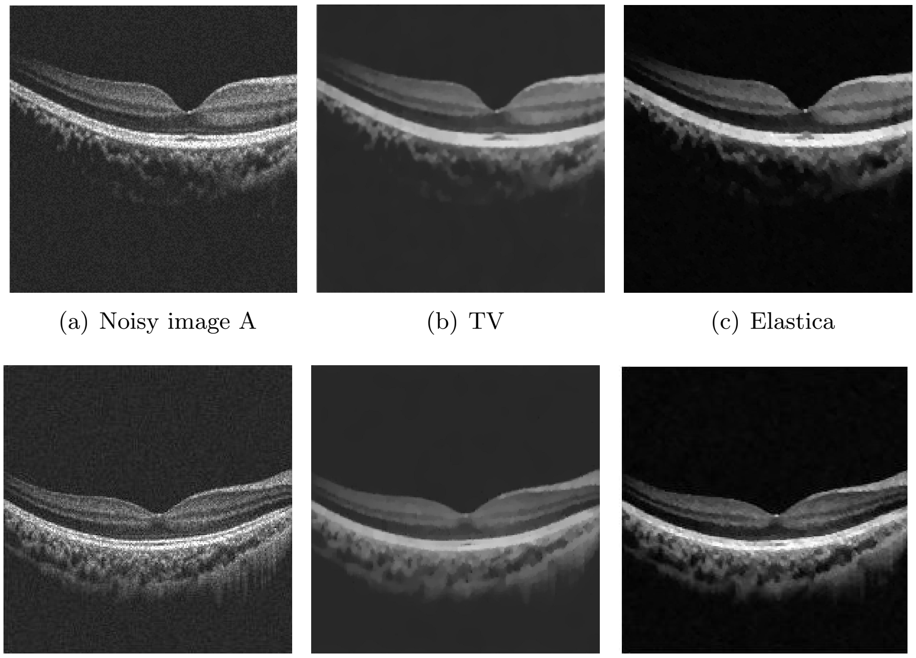
Figure 7. *The images in first row are noisy image, result of TV model by FTVd method and result of Euler's elastica model by Algorithm 1, respectively. The corresponding contour maps are shown in the second row.*

Results show that the numerical performance is slightly different when varying parameters. But overall speaking, Algorithm 1 is stable and insensitive to smoothing parameters ϵ, μ^0 and θ. Convergence behavior of residuals r_1, r_2, r_3 and $Res_1 = \max(r_1, r_2, r_3)$ are presented in Figure 6. We can see that the optimality residual is reduced to a small number eventually which verifies the theoretical analysis in previous sections.

5.3. Image denoising. In this subsection, the matrix K is an identity matrix. We use Algorithm 1 to denoise the optical coherence tomography (OCT) images. OCT is a high-resolution imaging technology mainly used in clinical medicine and can yield good effects in diagnosis of retinal diseases especially. However, the high-resolution of OCT means high demanding for environment and collection process of signal data. OCT images are always damaged by speckle noise, which has an adverse impact on observation and estimation of OCT images [14]. Thus pre-processing is necessary and often the first step in OCT image analysis. According to statistical optics, speckle noise in OCT images is multiplicative noise and can be converted into additive noise by logarithmic compression [24]. Therefore we can apply Algorithm 1 for Euler's elastica model to denoise the corrupted OCT images. The variance of speckle noise is 0.02. The peak signal-to-noise ratio (PSNR) is defined by

$$\mathrm{PSNR} = 10 \times \log_{10} \frac{m \cdot (\max \mathbf{u}^k)^2}{\|\mathbf{u}^k - \mathbf{u}_{org}\|^2}.$$

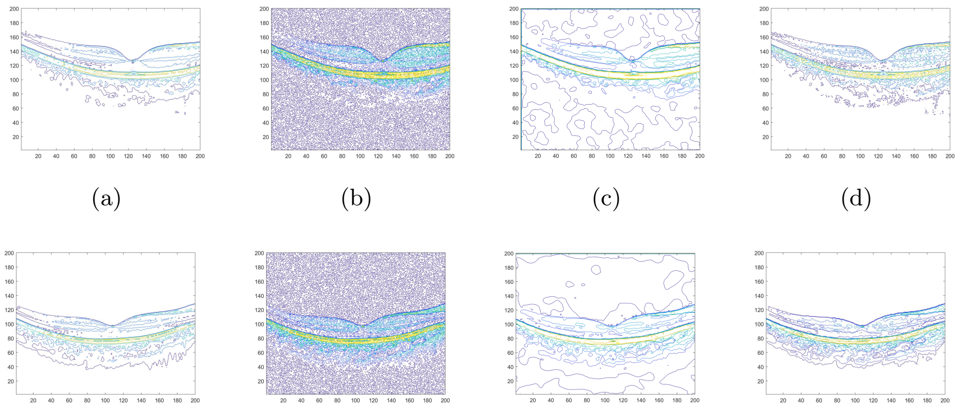
In Figure 7, we use a Gaussian function [15] to generate a synthetic OCT image to compare



(a) Noisy image A (b) TV (c) Elastica

(d) Noisy image B (e) TV (f) Elastica

Figure 8. *Two experiments in real OCT images. The parameters associated with the fidelity term in FTVd for (b) and (e) are set as $\lambda = 8$ and 5, respectively.*



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 9. *The first row are contour maps of original OCT image A without noise and Figures 8(a)-8(c). The second row are contour maps of original OCT image B without noise and Figures 8(d)-8(f).*

Algorithm 1 for Euler's elastica model and fast total variation deconvolution (FTVd) method for the TV model in [29] to denoise OCT images. Figure 7(a) is the input noisy image, Figures 7(b) and 7(c) are results by FTVd method for the TV model and Algorithm 1 for

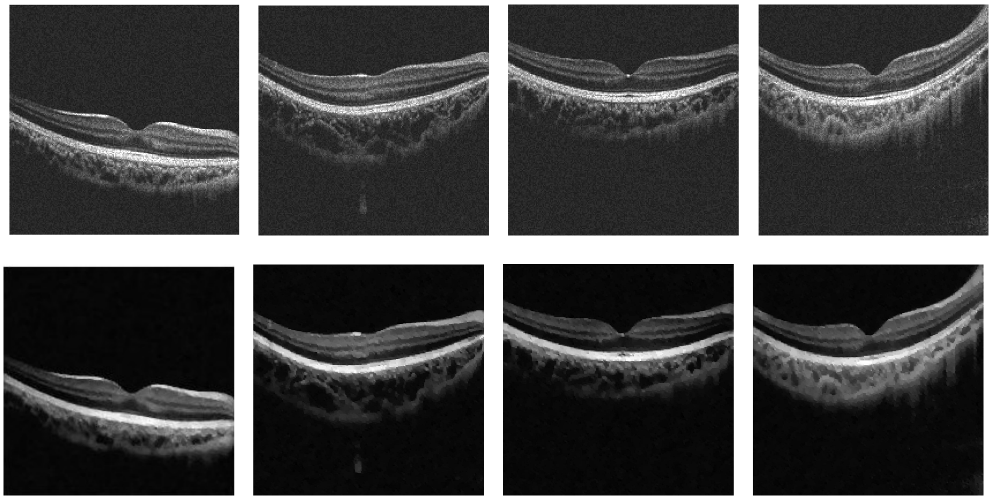


Figure 10. *The first row are noisy images. The second row are corresponding results of Euler's elastica model by Algorithm 1, the values of PSNR from left to right are: 27.9292, 26.7305,26.6660, 26.5482.*

Euler's elastica model, respectively. Figures 7(d)-7(f) are corresponding contour maps given for visual comparison. It can be found that Algorithm 1 for Euler's elastica model yields a more pleasant and smoothing restoration in layers which is important for subsequent processing such as layer segmentation.

We report two more experiments on real OCT images in Figure 8. Noisy OCT images are shown in the left column and the denoising results by FTVd method for TV model and Algorithm 1 for Euler's elastica model are shown in middle and right columns, respectively. To give a more vivid description, we draw the corresponding contour maps in Figure 9. It can be observed visually that Algorithm 1 for Euler's elastica model significantly eliminates the noise in OCT images and preserves the continuity and integrity of choroid layer. Although it looks more smoothing in the results of FTVd method for the TV model, the tiny features easily confused with noise are vanished which may seriously affect the segmentation of choroid layer and detection of disease. Moreover, staircase effect appears in the results by FTVd method for the TV model. For the results of Algorithm 1 for Euler's elastica model, details of physiological tissue in images are kept without over-smoothing and staircase effect, which means Figures 8(c) and 8(f) are better than Figures 8(b) and 8(e) in OCT image analysis. Four more OCT denoising results using Algorithm 1 for Euler's elastica model are given in Figure 10.

Remark 5.1 The operator splitting method in [13] is a new and efficient method for the Euler elastica model for image smoothing. However, the operator splitting method needs a unique solution of a strongly convex problem at each step (see section 3.6 in [13]) and does not have convergence guarantees. Note that the objective function in the Euler elastica model (1.1) for inpainting problem is not strongly convex due to the singularity of the linear operator $\mathcal{K}$. It is worth noting that Algorithm 1 has convergence guarantees for Euler's elastica model with $\mathcal{K}$ being an identity operator for image smoothing.

6. Conclusion. In this paper, we propose a penalty relaxation method to solve the discrete Euler's elastica model (2.3), which has wide applications in image processing. To deal with the nonsmoothness of problem (2.3), we introduce a smoothing relaxation problem (2.4) and establish the relationship between solutions and stationary points of problem (2.3) and problem (2.4) in Theorems 3.4-3.5. Moreover, we propose the penalty problem (2.5) to overcome the difficulties caused by the nonconvex constraints in problem (2.4). We derive the relationship between problem (2.4) and problem (2.5) regarding their local minimizers, stationary points and optimal solutions in Theorems 3.6-3.7 and Corollary 3.8. Using the special structure of problem (2.5), we propose a smoothing block coordinate descent algorithm (Algorithm 1). In the algorithm, we split problem (2.5) into an unconstrained strictly convex subproblem in variable $\mathbf{u}$ and m two-dimensional ball constrained subproblems with a unique solution in variable $\mathbf{w}_i$. We prove that any accumulation point of the sequence generated by Algorithm 1 is a stationary point of problem (2.5). Finally, we present some numerical results in image inpainting and OCT image denoising to show the effectiveness of the proposed method.

Acknowledgment. The authors would like to thank the PolyU Optometry Clinic for providing real OCT images for numerical experiments in this paper. The authors also thank Professor Xuecheng Tai for providing references and codes for Euler's elastica model. The authors are grateful to the referees for their helpful comments.

REFERENCES

[1] M. Bacák, R. Bergmann, G. Steidl, and A. Weinmann, *A second order nonsmooth variational model for restoring manifold-valued images*, SIAM J. Sci. Comput., 38 (2016), pp. A567–A597.

[2] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, *Filling-in by joint interpolation of vector fields and gray levels*, IEEE Trans. Image Process., 10 (2001), pp. 1200–1211.

[3] K. Bredies, T. Pock, and B. Wirth, *Convex relaxation of a class of vertex penalizing functionals*, J. Math. Imaging Vision, 47 (2013), pp. 278–302.

[4] K. Bredies, T. Pock, and B. Wirth, *A convex, lower semicontinuous approximation of Euler's elastica energy*, SIAM J. Math. Anal., 47 (2015), pp. 566–613.

[5] C. Brito-Loeza and K. Chen, *Fast numerical algorithms for Euler's elastica inpainting model*, Int. J. Mod. Math., 5 (2010), pp. 157–182.

[6] M. Burger, A. Sawatzky, and G. Steidl, *First order algorithms in variational image processing*, in Splitting Methods in Communication, Imaging, Science, and Engineering, Springer, 2016, pp. 345–407.

[7] R. H. Chan, S. Setzer, and G. Steidl, *Inpainting by flexible haar-wavelet shrinkage*, SIAM J. Imaging Sci., 1 (2008), pp. 273–293.

[8] T. F. Chan and P. Mulet, *On the convergence of the lagged diffusivity fixed point method in total variation image restoration*, SIAM J. Numer. Anal., 36 (1999), pp. 354–367.

[9] X. Chen, *Smoothing methods for nonsmooth, nonconvex minimization*, Math. Program., 134 (2012), pp. 71–99.

[10] X. Chen, M. K. Ng, and C. Zhang, *Non-Lipschitz l_p-regularization and box constrained model for image restoration*, IEEE Trans. Image Process., 21 (2012), pp. 4709–4721.

[11] X. Chen and W. Zhou, *Smoothing nonlinear conjugate gradient method for image restoration using nonsmooth nonconvex minimization*, SIAM J. Imaging Sci., 3 (2010), pp. 765–790.

[12] F. H. Clarke, *Optimization and Nonsmooth Analysis*, SIAM Philadelphia, PA, 1990.

[13] L.-J. Deng, R. Glowinski, and X.-C. Tai, *A new operator splitting method for the Euler elastica model for image smoothing*, SIAM J. Imaging Sci., 12 (2019), pp. 1190–1230.

[14] W. Drexler and J. G. Fujimoto, *Optical coherence tomography: technology and applications*, Springer

Science & Business Media, 2008.

[15] J. Duan, W. Lu, C. Tench, I. Gottlob, F. Proudlock, N. N. Samani, and L. Bai, *Denoising optical coherence tomography using second order total generalized variation decomposition*, Biomed. Signal Proc. Control, 24 (2016), pp. 120–127.

[16] S.-P. Han and O. L. Mangasarian, *Exact penalty functions in nonlinear programming*, Math. Program., 17 (1979), pp. 251–269.

[17] R. Janin, *Directional derivative of the marginal function in nonlinear programming*, in Sensitivity, Stability and Parametric Analysis, Springer, 1984, pp. 110–126.

[18] H. Kasai, *Sgdlibrary: A matlab library for stochastic optimization algorithms*, J. Mach. Learn. Res., 18 (2017), pp. 7942–7946.

[19] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, ICLR, (2015).

[20] A. Marquina and S. Osher, *A new time dependent model based on level set motion for nonlinear deblurring and noise removal*, in Int. Conf. Scale-Space Theories Comp. Vision, Springer, 1999, pp. 429–434.

[21] H. Robbins and S. Monro, *A stochastic approximation method*, Ann. Math.Statist., 22 (1951), pp. 400–407.

[22] S. Setzer and G. Steidl, *Variational methods with higher order derivatives in image processing*, Approximation XII, (San Antonio, USA), M.Neamtu and L. L. Schumaker, Eds., Brentwood: Nashboro Press, 12 (2008), pp. 360–386.

[23] J. Shen, S. H. Kang, and T. F. Chan, *Euler's elastica and curvature-based inpainting*, SIAM J. Appl. Math., 63 (2003), pp. 564–592.

[24] C. Sheng, Y. Xin, Y. Liping, and S. Kun, *Total variation-based speckle reduction using multi-grid algorithm for ultrasound images*, in Int. Conf. Image Anal. Processing, Springer, 2005, pp. 245–252.

[25] X.-C. Tai and J. Duan, *A simple fast algorithm for minimization of the elastica energy combining binary and level set representations*, Int. J. Numer. Anal. Model, 14 (2017), pp. 809–821.

[26] X.-C. Tai, J. Hahn, and G. J. Chung, *A fast algorithm for Euler's elastica model using augmented Lagrangian method*, SIAM J. Imaging Sci., 4 (2011), pp. 313–344.

[27] X.-C. Tai, K.-A. Lie, T. F. Chan, and S. Osher, *Image Processing Based on Partial Differential Equations: Proceedings of the International Conference on PDE-Based Image Processing and Related Inverse Problems, CMA, Oslo, August 8-12, 2005*, Springer Science & Business Media, 2006.

[28] C. R. Vogel and M. E. Oman, *Iterative methods for total variation denoising*, SIAM J. Sci. Comput., 17 (1996), pp. 227–238.

[29] Y. Wang, J. Yang, W. Yin, and Y. Zhang, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM J. Imaging Sci., 1 (2008), pp. 248–272.

[30] M. Yashtini and S. H. Kang, *A fast relaxed normal two split method and an effective weighted TV approach for Euler's elastica image inpainting*, SIAM J. Imaging Sci., 9 (2016), pp. 1552–1581.

[31] W. Zhu, X.-C. Tai, and T. Chan, *Image segmentation using Euler's elastica as the regularization*, J. Sci. Comput., 57 (2013), pp. 414–438.