A Finite Element Nonoverlapping Domain Decomposition Method with Lagrange Multipliers for the Dual Total Variation Minimizations

Chang-Ock Lee · Jongho Park

Abstract In this paper, we consider a primal-dual domain decomposition method for total variation regularized problems appearing in mathematical image processing. The model problem is transformed into an equivalent constrained minimization problem by tearing-and-interconnecting domain decomposition. Then, the continuity constraints on the subdomain interfaces are treated by introducing Lagrange multipliers. The resulting saddle point problem is solved by the first order primal-dual algorithm. We apply the proposed method to image denoising, inpainting, and segmentation problems with either L^2-fidelity or L^1-fidelity. Numerical results show that the proposed method outperforms the existing state-of-the-art methods.

1 Introduction

After a pioneering work of Rudin et al. [24], total variation minimization has been widely used in image processing. In their work, authors proposed an image

Chang-Ock Lee
Department of Mathematical Sciences, KAIST, Daejeon 34141, Korea
E-mail: colee@kaist.edu

Jongho Park
Department of Mathematical Sciences, KAIST, Daejeon 34141, Korea
Tel.: +82-42-350-2790
E-mail: jongho.park@kaist.ac.kr

denoising model with the total variation regularizer, which is called Rudin–Osher–Fatemi (ROF) model, as follows:

$$\min_{u \in BV(\Omega)} \left\{ \frac{\alpha}{2} \int_{\Omega} (u - f)^2 \, dx + TV(u) \right\}, \tag{1.1}$$

where Ω is an image domain, f is a corrupted image, $\alpha > 0$ is a positive denoising parameter, and $TV(u)$ is the total variation of u defined as

$$TV(u) = \sup \left\{ \int_{\Omega} u \operatorname{div} \mathbf{q} \, dx : \mathbf{q} \in (C_0^1(\Omega))^2, |\mathbf{q}(x)| \leq 1 \;\; \forall x \in \Omega \right\},$$

where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^2$. The solution space $BV(\Omega)$ is the collection of L^1 functions with finite total variation. Thanks to the anisotropic diffusion property of the total variation term, the model (1.1) effectively removes Gaussian noise and preserves edges and discontinuities of the image [26].

Total variation minimization can be applied to not only image denoising, but also other interesting problems in image processing. In [25], an image inpainting model with the total variation regularizer was proposed. The model in [25] is a simple modification of (1.1), which excludes the inpainting domain from the domain of integration of the fidelity term in (1.1). In addition, (1.1) can be generalized to the image deconvolution problem by replacing u in the fidelity term by Au, where A is either a specific convolution operator [7, 29]. In order to treat impulse noise, the TV-L^1 model, which uses L^1 fidelity instead of L^2 fidelity, was introduced in [5, 22]. It is well-known that the TV-L^1 model preserves contrast of the image, while the conventional ROF model does not. We also note that variational image segmentation problem can be represented as the total variation minimization by appropriate change of variables [6]. One can refer [4] for further results of total variation minimization.

This paper is concerned with domain decomposition methods (DDMs) for such total variation regularized minimization problems. DDMs are suitable for parallel computation since they solve a large scale problem by dividing it into smaller problems and treating them in parallel. While DDMs for elliptic partial differential equations have been successfully developed over past decades, there have been relatively modest achievements in total variation minimization problems due to their own difficulties. At first, the total variation term is nonsmooth and nonseparable. Thus, the energy functional cannot be expressed as the sum of the local energy functionals in the subdomains in general. Even more, the solution space $BV(\Omega)$ allows discontinuities of a solution on the subdomain interfaces, so that it is difficult to impose appropriate boundary conditions to the local problems in the subdomains.

There have been several researches to overcome such difficulties and develop efficient DDMs for the total variation minimization [8, 10, 12, 13, 14, 16, 17, 18, 19, 20, 21]. The Gauss-Seidel and Jacobi type subspace correction methods for the ROF model were proposed in [12, 13, 14, 18], and they were extended to the case of mixed L^1/L^2 fidelity in [16]. However, in [19], a counterexample was provided for convergence of subspace correction methods. In [10], a convergent overlapping DDM for the convex Chan–Vese model [6] was proposed. Recently, a domain decomposition framework for the case of L^1 fidelity was introduced in [20]. One of

the effective ideas in DDMs for the total variation minimization is to consider the Fenchel–Rockafellar dual formulation of the model:

$$\min_{\mathbf{p} \in H_0(\mathrm{div};\Omega)} \frac{1}{2\alpha} \int_\Omega (\mathrm{div}\,\mathbf{p} + \alpha f)^2\,dx \quad\text{subject to } |\mathbf{p}(x)| \le 1 \ \ \forall x \in \Omega$$

instead of the original one. Here, $H_0(\mathrm{div};\Omega)$ denotes the space of vector fields $\mathbf{q}\colon \Omega \to \mathbb{R}^2$ such that $\mathrm{div}\,\mathbf{q} \in L^2(\Omega)$ and $\mathbf{q} \cdot \mathbf{n} = 0$ on $\partial\Omega$. Consideration of the dual formulation resolves some difficulties mentioned above; the dual energy functional is separable, and the solution space $H_0(\mathrm{div};\Omega)$ requires some regularity on the subdomain interfaces. Even if there arises another difficulty of treating the inequality constraint $|\mathbf{p}(x)| \le 1 \ \ \forall x \in \Omega$, several successful researches have been done [8, 17, 19, 21].

In this paper, we generalize the primal-dual DDM proposed in [21] to a wider class of the total variation minimizations. We consider a general model problem

$$\min_{u \in BV(\Omega)} \left\{ \alpha F(u) + TV(u) \right\}, \tag{1.2}$$

where $F\colon BV(\Omega) \to \bar{\mathbb{R}}$ is a proper, convex, and lower semicontinuous functional satisfying additional properties such that F is separable and simple in the sense that (1.2) can be efficiently solved by the first order primal-dual algorithm [3]. Such a class of total variation minimizations contains the ROF model, the TV-L^1 model, their inpainting variants, and the convex Chan–Vese model for image segmentation. As we noted above, we treat the Fenchel–Rockafellar dual problem of (1.2):

$$\min_{\mathbf{p} \in H_0(\mathrm{div};\Omega)} \frac{1}{\alpha} F^*(\mathrm{div}\,\mathbf{p}) \quad\text{subject to } |\mathbf{p}(x)| \le 1 \ \ \forall x \in \Omega, \tag{1.3}$$

where F^* is the Legendre–Fenchel conjugate of F. The image domain is decomposed into a number of disjoint subdomains $\{\Omega_s\}_{s=1}^{\mathcal{N}}$, and the continuity along the subdomain interfaces is enforced. By treating the continuity constraints on the subdomain interfaces by the method of Lagrange multipliers, we obtain an equivalent saddle point problem. Application of the Chambolle–Pock primal-dual algorithm [3] to the resulting saddle point problem yields our proposed method, that shows good performance for various total variation minimization problems such as image denoising, inpainting, and segmentation.

The rest of the paper is organized as follows. We present the basic settings for design of DDM in Sect. 2. In Sect. 3, we state the abstract model problem which generalizes various problems in image processing such as image denoising, inpainting, and segmentation. A convergent nonoverlapping DDM for the model problem is proposed in Sect. 4. We apply the proposed method to several image processing problems and compare with the existing state-of-the-art methods in Sect. 5. We conclude the paper with remarks in Sect. 6.

2 Preliminaries

In this section, we present the basic setting for design of DDM. At first, we introduce notations that will be used throughout the paper. Then, the discrete setting for the dual total variation minimization based on the finite element framework is provided.

2.1 Notations

Let $\mathcal{H}$ be the generic n-dimensional Hilbert space. For $v = [v_1, ..., v_n]^T, w = [w_1, ..., w_n]^T \in \mathcal{H}$ and $1 \leq p < \infty$, the p-norm of v is denoted by

$$\|v\|_{p,\mathcal{H}} = \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}},$$

and the Euclidean inner product of v and w is denoted by

$$\langle v, w \rangle_{\mathcal{H}} = \sum_{i=1}^{n} v_i w_i.$$

We may drop the subscript $\mathcal{H}$ if there is no ambiguity.

In this paper, we use the symbol * as superscripts with two different meanings. At first, for a convex functional $F \colon \mathcal{H} \to \bar{\mathbb{R}}$, $F^* \colon \mathcal{H} \to \bar{\mathbb{R}}$ denotes the Legendre–Fenchel conjugate of F, which is defined as

$$F^*(v^*) = \sup_{v \in \mathcal{H}} \left\{ \langle v, v^* \rangle - F(v) \right\}$$

for every $v^* \in \mathcal{H}$. On the other hand, when $A \colon \mathcal{H} \to \mathcal{H}$ is a linear operator on $\mathcal{H}$, A^* denotes the adjoint of A, that is,

$$\langle v, Aw \rangle = \langle A^* v, w \rangle$$

for every $v, w \in \mathcal{H}$. It is well-known that the matrix representation of A^* is the transpose of the matrix representation of A.

For a convex functional $F \colon \mathcal{H} \to \bar{\mathbb{R}}$, the effective domain of F, denoted by $\mathrm{dom}\, F$, is defined as

$$\mathrm{dom}\, F = \{ u \in \mathcal{H} : F(u) < \infty \}.$$

Also, for a convex subset C of $\mathcal{H}$, we denote the relative interior of C by $\mathrm{ri}\, C$, which is defined as the interior of C when C is regarded as a subset of its affine hull. One may refer Sect. 6 of [23] for detailed topological properties of $\mathrm{ri}\, C$.

For a subset C of $\mathcal{H}$, we define the characteristic functional $\chi_C \colon \mathcal{H} \to \bar{\mathbb{R}}$ of C by

$$\chi_C(v) = \begin{cases} 0 & \text{if } v \in C, \\ \infty & \text{if } v \notin C. \end{cases}$$

It is clear that the functional χ_C is convex if and only if C is a convex subset of $\mathcal{H}$.

2.2 Fenchel–Rockafellar duality

Throughout this paper, we use the notion of Fenchel–Rockafellar duality frequently. For the sake of completeness, we provide its key features. For more rigorous texts, readers may refer [4,23], for instance.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be finite-dimensional Hilbert spaces. Consider the minimization problem of the form

$$\min_{u \in \mathcal{H}_1} \left\{ F(u) + G(Ku) \right\}, \tag{2.1}$$

where $K\colon \mathcal{H}_1 \to \mathcal{H}_2$ is a linear operator and $F\colon \mathcal{H}_1 \to \bar{\mathbb{R}}$, $G\colon \mathcal{H}_2 \to \bar{\mathbb{R}}$ are proper, convex, lower semicontinuous functionals. We use the assumptions on F and G presented in Corollary 31.2.1 of [23].

Assumption 2.1 There exists $u \in \mathcal{H}_1$ such that $u \in \mathrm{ri}(\mathrm{dom}\,F)$ and $Ku \in \mathrm{ri}(\mathrm{dom}\,G)$.

Under Assumption 2.1, the following relations hold [23]:

$$\min_{u \in \mathcal{H}_1} \{F(u) + G(Ku)\} = \min_{u \in \mathcal{H}_1} \max_{q \in \mathcal{H}_2} \left\{ \langle Ku, q \rangle_{\mathcal{H}_2} + F(u) - G^*(q) \right\} \qquad (2.2a)$$

$$= - \left[\min_{q \in \mathcal{H}_2} \left\{ F^*(-K^*q) + G^*(q) \right\} \right]. \qquad (2.2b)$$

The saddle point problem in the right hand side of (2.2a) is called the primal-dual formulation of (2.1), and the minimization problem in (2.2b) is called the dual formulation of (2.1). Hence, it is enough to solve (2.2a) or (2.2b) instead of (2.1) in many cases.

2.3 Discrete setting

Let $\Omega \subset \mathbb{R}^2$ be a rectangular image domain consisting of a number of rows and columns of pixels. Since each pixel holds a value representing the intensity at a point, we may regard the image as a pixelwise constant function on Ω. We write the collection of all pixels in Ω as $\mathcal{T}$ and define the space X for an image as

$$X = \left\{ u \in L^2(\Omega) : u|_T \text{ is constant } \forall T \in \mathcal{T} \right\}.$$

We note that $X \subset BV(\Omega)$. We regard each pixel in $\mathcal{T}$ as a square element so that X becomes a piecewise constant square finite element space whose side length equals to 1. It is clear that the functions

$$\phi_T(x) = \begin{cases} 1 & \text{if } x \in T, \\ 0 & \text{if } x \notin T, \end{cases} \qquad T \in \mathcal{T}$$

form a basis for X. For $u \in X$ and $T \in \mathcal{T}$, let $(u)_T \in \mathbb{R}$ denote the degree of freedom (dof) of u associated with the basis function ϕ_T. Then, u is represented as

$$u = \sum_{T \in \mathcal{T}} (u)_T \phi_T.$$

To obtain a finite element discretization of the dual problem, the space Y for the dual variables is defined by the lowest order Raviart–Thomas elements:

$$Y = \{ \mathbf{q} \in H_0(\mathrm{div}; \Omega) : \mathbf{q}|_T \in \mathcal{RT}_0(T) \ \forall T \in \mathcal{T} \},$$

where $\mathcal{RT}_0(T)$ is the collection of the vector functions $\mathbf{q}\colon T \to \mathbb{R}^2$ of the form

$$\mathbf{q}(x_1, x_2) = \begin{bmatrix} a_1 + b_1 x_1 \\ a_2 + b_2 x_2 \end{bmatrix}.$$

We notice that the divergence operator in the continuous setting is well-defined on Y and $\mathrm{div}\, Y \subset X$. Each dof of Y is the value of the normal component on a

pixel edge. Let $\mathcal{I}$ be the set of indices of the basis functions for Y and $\{\boldsymbol{\psi}_i\}_{i\in\mathcal{I}}$ be the basis. For $\mathbf{p} \in Y$ and $i \in \mathcal{I}$, we denote the dof of $\mathbf{p}$ associated with the basis function $\boldsymbol{\psi}_i$ as $(\mathbf{p})_i \in \mathbb{R}$; we have

$$\mathbf{p} = \sum_{i\in\mathcal{I}} (\mathbf{p})_i \boldsymbol{\psi}_i.$$

Throughout the paper, the p-norm of $\mathbf{p} \in Y$ is defined as the p-norm of the vector of dofs of $\mathbf{p}$ for $1 \leq p < \infty$. In this case, one may regard $\|\cdot\|_{2,Y}$ as a lumped $(L^2(\Omega))^2$-norm with proper scaling; see Remark 2.2 of [21].

In order to treat the inequality constraints appearing in the dual formulation (1.3), we define the convex subset C of Y by

$$C = \{\mathbf{p} \in Y : |(\mathbf{p})_i| \leq 1 \ \ \forall i \in \mathcal{I}\}. \tag{2.3}$$

The projection of $\mathbf{p} \in Y$ onto C can be easily computed by

$$(\mathrm{proj}_C \, \mathbf{p})_i = \frac{(\mathbf{p})_i}{\max\{1, |(\mathbf{p})_i|\}} \qquad \forall i \in \mathcal{I}. \tag{2.4}$$

2.4 Domain decomposition setting

We decompose the image domain Ω into $\mathcal{N} = N \times N$ disjoint rectangular subdomains $\{\Omega_s\}_{s=1}^{\mathcal{N}}$. For two subdomains Ω_s and Ω_t ($s < t$) sharing a subdomain edge (interface), let $\Gamma_{st} = \partial\Omega_s \cap \partial\Omega_t$ be the shared subdomain edge between them. Also, we define the union of the subdomain interfaces Γ by $\Gamma = \bigcup_{s<t} \Gamma_{st}$. We assume that Γ does not cut through any elements in $\mathcal{T}$.

Now, we define the local function spaces in the subdomains. For $s = 1, ..., \mathcal{N}$, let $\mathcal{T}_s$ be the collection of all pixels in Ω_s. The local primal function space X_s is defined by

$$X_s = \left\{u \in L^2(\Omega_s) : u|_T \text{ is constant } \forall T \in \mathcal{T}_s\right\}.$$

Obviously, we have $X = \bigoplus_{s=1}^{\mathcal{N}} X_s$. Furthermore, we define the natural restriction operator $R_s \colon X \to X_s$ by

$$R_s u = u|_{\Omega_s}. \tag{2.5}$$

Then, the adjoint $R_s^* \colon X_s \to X$ of R_s becomes the extension-by-zero operator

$$(R_s^* u_s)_T = \begin{cases} (u_s)_T & \text{if } T \subset \Omega_s, \\ 0 & \text{if } T \not\subset \Omega_s, \end{cases} \quad T \in \mathcal{T}.$$

The local dual function spaces are defined in the tearing-and-interconnecting sense [11]. More precisely, we define the local dual function space $\tilde{Y}_s$ as

$$\tilde{Y}_s = \{\tilde{\mathbf{q}}_s \in H(\mathrm{div}; \Omega_s) : \tilde{\mathbf{q}}_s \cdot \mathbf{n}_s = 0 \text{ on } \partial\Omega_s \setminus \Gamma, \ \tilde{\mathbf{q}}_s|_T \in \mathcal{RT}_0(T) \ \forall T \in \mathcal{T}_s\}.$$

Note that the boundary condition $\tilde{\mathbf{q}}_s \cdot \mathbf{n}_s = 0$ is not imposed on $\Gamma \cap \partial\Omega_s$ for $\tilde{Y}_s$. Thus, $\tilde{Y}_s$ has dofs on the pixel edges contained in $\partial\Omega_s \cap \Gamma$. Let $\tilde{\mathcal{I}}_s$ be the set of indices of the basis functions for $\tilde{Y}_s$. Similarly to (2.3), the inequality-constrained subset $\tilde{C}_s$ of $\tilde{Y}_s$ is defined by

$$\tilde{C}_s = \left\{\tilde{\mathbf{p}}_s \in \tilde{Y}_s : |(\tilde{\mathbf{p}}_s)_i| \leq 1 \ \ \forall i \in \tilde{\mathcal{I}}_s\right\}.$$

The orthogonal projection onto $\tilde{C}_s$ is computed as

$$(\text{proj}_{\tilde{C}_s} \tilde{\mathbf{p}}_s)_i = \frac{(\tilde{\mathbf{p}}_s)_i}{\max\{1, |(\tilde{\mathbf{p}}_s)_i|\}} \qquad \forall i \in \tilde{\mathcal{I}}_s.$$

Let $\tilde{Y}$ and $\tilde{C}$ be the direct sums of the sets $\tilde{Y}_s$'s and $\tilde{C}_s$'s, respectively. By definition, functions in $\tilde{Y}$ may have discontinuities on Γ. Let $\mathcal{I}_\Gamma$ be the collection of dofs of Y on Γ. The jump operator $B \colon \tilde{Y} \to \mathbb{R}^{|\mathcal{I}_\Gamma|}$ measures the magnitude of such discontinuities, that is, B is defined as

$$B\tilde{\mathbf{p}}|_{\Gamma_{st}} = \tilde{\mathbf{p}}_s \cdot \mathbf{n}_{st} - \tilde{\mathbf{p}}_t \cdot \mathbf{n}_{st}, \qquad s < t,$$

where $\tilde{\mathbf{p}}_s = \tilde{\mathbf{p}}|_{\Omega_s}$. Clearly, there is a natural isomorphism between Y and $\ker B \subset \tilde{Y}$. For later use, we provide an upper bound of the operator norm of B [21].

Proposition 2.2 *The operator norm of $B \colon \tilde{Y} \to \mathbb{R}^{|\mathcal{I}_\Gamma|}$ has a bound $\|B\|_2^2 \leq 2$.*

Proposition 2.2 will be used for the estimation of the range of the parameters in the proposed method.

3 The model problem

The model problem we consider in this paper is the total variation regularized convex minimization problem defined on the image domain Ω:

$$\min_{u \in X} \{\mathcal{J}(u) := \alpha F(u) + TV(u)\}, \tag{3.1}$$

where $F \colon X \to \bar{\mathbb{R}}$ is a proper, convex, lower semicontinuous functional, $TV(u)$ is the discrete total variation of u given by

$$TV(u) = \max_{\mathbf{p} \in C} \langle u, \text{div}\,\mathbf{p} \rangle_X, \tag{3.2}$$

and $\alpha > 0$ is a positive parameter. In addition, we assume that F satisfies the following three assumptions.

Assumption 3.1 F is separable in the sense that there exist proper, convex, lower semicontinuous local energy functionals $F_s \colon X_s \to \bar{\mathbb{R}}$ such that

$$F(u) = \sum_{s=1}^{\mathcal{N}} F_s(R_s u),$$

where R_s is the restriction operator defined in (2.5).

Assumption 3.2 For any $u \in X$ and $\sigma > 0$, the proximity operator $\text{prox}_{\sigma F}(u)$ defined by

$$\text{prox}_{\sigma F}(u) = \arg\min_{v \in X} \left\{ F(v) + \frac{1}{2\sigma}\|v - u\|_{2,X}^2 \right\}$$

has a closed-form formula.

Assumption 3.3 $\text{ri}(\text{dom}\,F) \neq \emptyset$.

Assumption 3.1 makes (3.1) more suitable for designing DDMs. It will be explained in Sect. 4. Also, by Assumption 3.2, we are able to adopt the primal-dual algorithm [3] to solve (3.1). Indeed, if Assumption 3.2 holds, we can solve an equivalent primal-dual form of (3.1),

$$\min_{u \in X} \max_{\mathbf{p} \in Y} \left\{ - \langle u, \operatorname{div} \mathbf{p} \rangle_X + \alpha F(u) - \chi_C(\mathbf{p}) \right\}, \tag{3.3}$$

by the primal-dual algorithm. For the model problem (3.1), Assumption 3.3 is equivalent to Assumption 2.1 when $G(Ku) = TV(u)$ since $\operatorname{ri}(\operatorname{dom}(\| \cdot \|_{1,Y})) = Y$. That is, Assumption 3.3 is essential for the primal-dual equivalence. It is trivial that all the problems we will deal with in this paper satisfy Assumption 3.3.

Next, we consider the dual formulation of (3.1):

$$\min_{\mathbf{p} \in Y} \left\{ \frac{1}{\alpha} F^*(\operatorname{div} \mathbf{p}) + \chi_C(\mathbf{p}) \right\}. \tag{3.4}$$

From (3.3), it is possible to deduce a relationship between solutions of the primal problem (3.1) and the dual problem (3.4):

$$\begin{aligned} 0 &\in -\operatorname{div} \mathbf{p} + \alpha \partial F(u), \\ 0 &\in -\operatorname{div}^* u - \partial \chi_C(\mathbf{p}). \end{aligned} \tag{3.5}$$

As the standard discretization for the total variation minimization is the finite difference method, we provide a relation between the finite difference discretization and our finite element discretization. The following proposition means that a solution of the finite difference discretization of (1.2) can be recovered from a solution of (3.4).

Proposition 3.4 *Assume that the image domain Ω consists of $M \times N$ pixels. Let $\mathbf{p}^* \in Y$ be a solution of (3.4). If $u^* \in X$ and $\mathbf{p}^*$ satisfy the primal-dual relation (3.5), then u^* is a solution of the minimization problem*

$$\min_{u \in X} \left\{ \alpha F(u) + \||Du|\|_1 \right\},$$

where Du is the forward finite difference operator

$$\begin{aligned} (Du)^1_{ij} &= \begin{cases} u_{i+1,j} - u_{ij} & \text{if } i = 1, ..., M-1, \\ 0 & \text{if } i = M, \end{cases} \\ (Du)^2_{ij} &= \begin{cases} u_{i,j+1} - u_{ij} & \text{if } j = 1, ..., N-1, \\ 0 & \text{if } j = N \end{cases} \end{aligned} \tag{3.6}$$

and $(|Du|)_{ij} = |(Du)^1_{ij}| + |(Du)^2_{ij}|$.

Proof It is straightforward by the same argument as Theorem 2.3 of [21]. □

Remark 3.5 In Proposition 3.4, $TV(u)$ equals to the finite difference anisotropic total variation $\||Du|\|_1$ due to the structure of the constraint set C in (2.3). On the other hand, one can make $TV(u)$ equal to the finite difference isotropic total variation by replacing C by an appropriate convex subset of Y; see [21] for details.

Remark 3.6 We give a remark on the relation between the continuous problem (1.2) and the discrete problem (3.1). For simplicity, we assume that the resolution of the image is $N \times N$. We fix the side length of Ω by 1 and introduce a side length parameter $h = 1/N$. The space X^h is defined in the same manner as in Sect. 2.3. Then the question we have is whether a solution $u^{h*} \in X^h$ of the minimization problem

$$\min_{u^h \in X^h} \left\{ \alpha F(u^h) + TV(u^h) \right\}$$

accumulates at a solution $u^* \in BV(\Omega)$ of (1.2) in $BV(\Omega)$ as $h \to 0$. However, since $TV(u^h)$ does not Γ-converge to the BV-seminorm in $BV(\Omega)$ as $h \to 0$ (see Example 4.1 of [1]), this is not true in general. It means that the space X^h may not be considered as a correct space for solving (1.2). To avoid this uncomfortable situation, one may use higher order Raviart–Thomas elements to discretize the dual formulation [1, 15].

Nevertheless, thanks to Proposition 3.4, one can construct a sequence accumulating at u^* in the $L^1(\Omega)$-topology from $\{u^{h*}\}_{h>0}$. For any $u^h \in X^h$, we have

$$TV(u^h) = h^2 \left\| |D^h \text{dof}^h(u^h)| \right\|_1, \tag{3.7}$$

where $\text{dof}^h(u^h) \in \mathbb{R}^{N \times N}$ is the dofs of u^h and $D^h \colon \mathbb{R}^{N \times N} \to \mathbb{R}^{N \times N} \times \mathbb{R}^{N \times N}$ is the forward finite difference operator defined similarly to (3.6). As it can be shown without much difficulty the right hand side of (3.7) Γ-converges to the BV-seminorm in $L^1(\Omega)$ as $h \to 0$ [2, 27], there exists an interpolation operator $I^h \colon \mathbb{R}^{N \times N} \to L^1(\Omega)$ such that $I^h \text{dof}^h(u^{h*})$ accumulates at u^* in $L^1(\Omega)$ as $h \to 0$. We note that for $u^h \in X^h$, $I^h \text{dof}^h(u^h) \notin X^h$ in general.

The model problem (3.1) occurs in various areas of mathematical image processing. One of the typical examples is the image denoising problem. In [24], authors proposed the well-known ROF model which consists of the L^2-fidelity term and the total variation regularizer. In the ROF model, F is given by

$$F(u) = \frac{1}{2} \|u - f\|_{2,X}^2$$

and Assumptions 3.1 and 3.2 are satisfied with

$$F_s(u_s) = \frac{1}{2} \|u_s - f\|_{2,X_s}^2,$$
$$\text{prox}_{\sigma F}(u) = \frac{u + \sigma f}{1 + \sigma}.$$

In order to preserve the contrast of an image, the $TV\text{-}L^1$ model which uses the L^1-fidelity term was introduced in [5, 22]:

$$F(u) = \|u - f\|_{1,X}.$$

The local functionals F_s and the proximity operator $\text{prox}_{\sigma F}$ are readily obtained as follows:

$$F_s(u_s) = \|u_s - f\|_{1,X_s},$$
$$\text{prox}_{\sigma F}(u) = f + \text{shr}_\Omega(u - f, \sigma),$$

where the elementwise shrinkage operator $\mathrm{shr}_S \colon X \times \mathbb{R}_{>0} \to X$ on $S \subset \Omega$ is defined as follows [28]:

$$(\mathrm{shr}_S(v,\sigma))_T = \begin{cases} \max\{|(v)_T| - \sigma, 0\} \frac{(v)_T}{|(v)_T|} & \text{if } T \subset S, \\ (v)_T & \text{if } T \not\subset S, \end{cases}, \quad T \in \mathcal{T}, \qquad (3.8)$$

with the convention $\frac{0}{|0|} = 0$. For simplicity, S is assumed to be a union of elements of $\mathcal{T}$ in (3.8).

The models for the image denoising problem are easily extended to the image inpainting problem [25]. Let $D \subset \Omega$ be the inpainting domain and $f \in X$ be the known part of an image. We assume that D does not cut through any elements in $\mathcal{T}$. We set $f = 0$ on D for simplicity. Also, let $A \colon X \to X$ be the restriction operator onto $\Omega \setminus D$, that is, $Au = 0$ on D for all $u \in X$. Then, F is given by

$$F(u) = \frac{1}{2}\|Au - f\|_{2,X}^2.$$

Note that Assumptions 3.1 and 3.2 are ensured with

$$F_s(u_s) = \frac{1}{2}\|A_s u_s - f\|_{2,X_s}^2,$$

$$(\mathrm{prox}_{\sigma F}(u))_T = \begin{cases} \frac{(u)_T + \sigma(f)_T}{1+\sigma} & \text{if } T \subset \Omega \setminus D, \\ (u)_T & \text{if } T \subset D, \end{cases}, \quad T \in T,$$

where $A_s u_s = R_s A R_s^* u_s$ for $s = 1, ..., \mathcal{N}$. One can easily check that $A = \bigoplus_{s=1}^{\mathcal{N}} A_s$. Similarly to the denoising model, we may use the L^1 fidelity term instead of the L^2 fidelity term as follows:

$$F(u) = \|Au - f\|_{1,X}.$$

In this case, F_s and $\mathrm{prox}_{\sigma F}$ are given by

$$F_s(u_s) = \|A_s u_s - f\|_{1,X_s},$$
$$\mathrm{prox}_{\sigma F}(u) = \mathrm{shr}_{\Omega \setminus D}(u - f, \sigma).$$

Another typical example is the image segmentation problem. In [6], authors proposed a convex image segmentation model with the total variation regularizer as follows:

$$\min_{u \in X} \left\{ \alpha \left(\left\langle u, (f - c_1)^2 \right\rangle_X + \left\langle 1 - u, (f - c_2)^2 \right\rangle_X \right) + \chi_{\{0 \le u \le 1\}}(u) + TV(u) \right\},$$
$$(3.9)$$

where f is a given image, c_1 and c_2 are predetermined intensity values. Writing $g = (f - c_1)^2 - (f - c_2)^2$ in (3.9) yields the following simpler form:

$$\min_{u \in X} \left\{ \alpha \left\langle u, g \right\rangle_X + \chi_{\{0 \le u \le 1\}}(u) + TV(u) \right\}. \qquad (3.10)$$

Then, (3.10) is of the form (3.1) with

$$F(u) = \left\langle u, g \right\rangle_X + \chi_{\{0 \le u \le 1\}}(u),$$

and satisfies Assumptions 3.1 and 3.2 with

$$F_s(u_s) = \langle u_s, g \rangle_{X_s} + \chi_{\{0 \leq u_s \leq 1\}}(u_s)$$
$$\mathrm{prox}_{\sigma F}(u) = \mathrm{proj}_{\{0 \leq \cdot \leq 1\}}(u - \sigma g).$$

Here, $\mathrm{proj}_{\{0 \leq \cdot \leq 1\}}$ can be computed pointwise like (2.4).

We close this section by mentioning that the image deconvolution problem is not a case of (3.1) in general. For an operator $A \colon X \to X$ defined by the matrix convolution with some matrix kernel, $A \neq \bigoplus_{s=1}^{\mathcal{N}} R_s A R_s^*$ since the computation of $R_s A u$ needs values of u in Ω_s as well as adjacent subdomains. Consequently, Assumption 3.1 cannot be satisfied for the deconvolution problem.

4 Proposed method

In this section, we extend the primal-dual DDM for the ROF model introduced in [21] to the more general model problem (3.1). The continuity of a solution on the subdomain interfaces is imposed in the *dual* sense, that is, it is imposed by the method of Lagrange multipliers. As a result, we obtain an equivalent saddle point problem, which is solved by the first order primal-dual algorithm [3].

We start the section by stating the following simple proposition, which means that the Legendre–Fenchel conjugate is separable if the original functional is separable.

Proposition 4.1 *Let $F \colon X \to \bar{\mathbb{R}}$ be a proper, convex, lower semicontinuous functional satisfying Assumption 3.1. Then, its Legendre–Fenchel conjugate F^* also satisfies Assumption 3.1.*

Proof For $u^* = \bigoplus_{s=1}^{\mathcal{N}} u_s^* \in X$, we have

$$\begin{aligned}
F^*(u^*) &= \sup_{u \in X} \left\{ \langle u, u^* \rangle_X - F(u) \right\} \\
&= \sum_{s=1}^{\mathcal{N}} \sup_{u_s \in X_s} \left\{ \langle u_s, u_s^* \rangle_{X_s} - F_s(u_s) \right\} \\
&= \sum_{s=1}^{\mathcal{N}} F_s^*(u_s^*).
\end{aligned}$$

Letting $(F^*)_s = F_s^*$ for $s = 1, ..., \mathcal{N}$ completes the proof. $\qquad\square$

Thanks to Proposition 4.1, we can transform the dual model problem (3.4) to an equivalent constrained minimization problem

$$\min_{\tilde{\mathbf{p}} \in \tilde{Y}} \sum_{s=1}^{\mathcal{N}} \left\{ \frac{1}{\alpha} F_s^*(\mathrm{div}\, \tilde{\mathbf{p}}_s) + \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) \right\} \qquad \text{subject to } B\tilde{\mathbf{p}} = 0. \qquad (4.1)$$

In order to treat the continuity constraint $B\tilde{\mathbf{p}} = 0$, the method of Lagrange multipliers for (4.1) yields the saddle point formulation

$$\min_{\tilde{\mathbf{p}} \in \tilde{Y}} \max_{\lambda \in \mathbb{R}^{|\mathcal{I}_\Gamma|}} \mathsf{L}(\tilde{\mathbf{p}}, \lambda), \qquad (4.2)$$

where

$$\mathrm{L}(\tilde{\mathbf{p}}, \lambda) = \sum_{s=1}^{\mathcal{N}} \left\{ \frac{1}{\alpha} F_s^* (\mathrm{div}\, \tilde{\mathbf{p}}_s) + \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) \right\} + \langle B\tilde{\mathbf{p}}, \lambda \rangle_{\mathbb{R}^{|\mathcal{I}_{\Gamma}|}} .$$

The following proposition summarizes the equivalence between the dual model problem (3.4) and the resulting saddle point problem (4.2).

Proposition 4.2 *If $\mathbf{p}^* \in C$ is a solution of (3.4), then $\tilde{\mathbf{p}}^* = \bigoplus_{s=1}^{\mathcal{N}} \mathbf{p}^*|_{\Omega_s} \in \tilde{C}$ is a primal solution of (4.2). Conversely, if $\tilde{\mathbf{p}}^*$ is a primal solution of (4.2), then $\tilde{\mathbf{p}}^* \in \ker B$. Hence $\tilde{\mathbf{p}}^* \in C$ and $\tilde{\mathbf{p}}^*$ is a solution of (3.4).*

Now, we are ready to propose the main algorithm of this paper. In the case of ROF model, recovering a primal solution u^* from the computed dual solution $\mathbf{p}^*$ can be easily done by the primal-dual relation $u^* = f + \mathrm{div}\,\mathbf{p}^*/\alpha$. However, in the general case, the primal solution may not be obtained as in the ROF case since the primal-dual relation (3.5) does not always give an explicit formula for u^*. Instead, we consider an algorithm to obtain a primal solution u^* and a dual solution $\mathbf{p}^*$ simultaneously. We begin with the primal-dual algorithm [3] applied to (4.2). In each iteration of the primal-dual algorithm, we need to solve the local problems of the following form:

$$\min_{\tilde{\mathbf{p}} \in \tilde{Y}} \sum_{s=1}^{\mathcal{N}} \left\{ \frac{1}{\alpha} F_s^*(\mathrm{div}\, \tilde{\mathbf{p}}_s) + \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) + \frac{1}{2\tau} \|\tilde{\mathbf{p}}_s - \hat{\mathbf{p}}_s^{(n+1)}\|_{2,\tilde{Y}_s}^2 \right\} \qquad (4.3)$$

for $\hat{\mathbf{p}}_s^{(n+1)} \in \tilde{Y}_s$. The term $\frac{1}{2\tau} \|\tilde{\mathbf{p}}_s - \hat{\mathbf{p}}_s^{(n+1)}\|_{2,\tilde{Y}_s}^2$ appears because we compute proximal descent/ascent in each step of the primal-dual algorithm. To obtain a primal solution $u^{(n+1)}$ and a dual solution $\tilde{\mathbf{p}}^{(n+1)}$ simultaneously, we replace (4.3) by the following primal-dual formulation of (4.3):

$$\min_{u \in X} \max_{\tilde{\mathbf{p}} \in \tilde{Y}} \sum_{s=1}^{\mathcal{N}} \left\{ -\langle u_s, \mathrm{div}\, \tilde{\mathbf{p}}_s \rangle_{X_s} + \alpha F_s(u_s) - \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) - \frac{1}{2\tau} \|\tilde{\mathbf{p}}_s - \hat{\mathbf{p}}_s^{(n+1)}\|_{2,\tilde{Y}_s}^2 \right\} .$$
$$(4.4)$$

There is another advantage to solve (4.4) instead of (4.3). Differently from the ROF model, it is sometimes cumbersome to get an explicit formula for F_s^*, which makes the design of a local solver difficult. We note that an explicit formula for F^* in the case of $F(u) = \|Au - f\|_{1,X}$ with nonsingular AA^* is given in [9], but it is somewhat complicated. However, considering (4.4) does not require an explicit formula for F^*.

Similarly to the ROF case, the solution pair $(u^{(n+1)}, \tilde{\mathbf{p}}^{(n+1)})$ can be constructed by assembling the local solution pairs $(u_s^{(n+1)}, \tilde{\mathbf{p}}_s^{(n+1)})$ in the subdomains, i.e., $u^{(n+1)} = \bigoplus_{s=1}^{\mathcal{N}} u_s^{(n+1)}$ and $\tilde{\mathbf{p}}^{(n+1)} = \bigoplus_{s=1}^{\mathcal{N}} \tilde{\mathbf{p}}_s^{(n+1)}$. The local solution pair $(u_s^{(n+1)}, \tilde{\mathbf{p}}_s^{(n+1)})$ in the subdomain Ω_s is obtained by solving the local problem

$$\min_{u_s \in X_s} \max_{\tilde{\mathbf{p}}_s \in \tilde{Y}_s} \left\{ -\langle u_s, \mathrm{div}\, \tilde{\mathbf{p}}_s \rangle_{X_s} + \alpha F_s(u_s) - \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) - \frac{1}{2\tau} \|\tilde{\mathbf{p}}_s - \hat{\mathbf{p}}_s^{(n+1)}\|_{2,\tilde{Y}_s}^2 \right\}$$
$$(4.5)$$

and each local problem can be solved in parallel. We will address how to solve (4.5) in Sect. 5 in detail.

Algorithm 1 Primal-dual domain decomposition method for the model problem (3.1)

Choose $L > 2$, $\tau, \sigma > 0$ with $\tau\sigma = \frac{1}{L}$. Let $\tilde{\mathbf{p}}^{(0)} = \tilde{\mathbf{p}}^{(-1)} = \mathbf{0}$ and $\lambda^{(0)} = 0$.
for $n = 0, 1, 2, ...$

$\qquad \lambda^{(n+1)} = \lambda^{(n)} + \sigma B \left(2\tilde{\mathbf{p}}^{(n)} - \tilde{\mathbf{p}}^{(n-1)} \right)$

$\qquad \hat{\mathbf{p}}^{(n+1)} = \tilde{\mathbf{p}}^{(n)} - \tau B^* \lambda^{(n+1)}$

$\qquad$ **for** $s = 1, 2, ..., \mathcal{N}$ **in parallel**

$\qquad\qquad (u_s^{(n+1)}, \tilde{\mathbf{p}}_s^{(n+1)}) \in \arg \min_{u_s \in X_s} \max_{\tilde{\mathbf{p}}_s \in \tilde{Y}_s} \Big\{ -\langle u_s, \operatorname{div} \tilde{\mathbf{p}}_s \rangle_{X_s} + \alpha F_s(u_s)$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad -\chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) - \frac{1}{2\tau} \|\tilde{\mathbf{p}}_s - \hat{\mathbf{p}}^{(n+1)}|_{\Omega_s}\|_{2, \tilde{Y}_s}^2 \Big\}$

$\qquad$ **end**

$\qquad u^{(n+1)} = \bigoplus_{s=1}^{\mathcal{N}} u_s^{(n+1)}$

$\qquad \tilde{\mathbf{p}}^{(n+1)} = \bigoplus_{s=1}^{\mathcal{N}} \tilde{\mathbf{p}}_s^{(n+1)}$

end

In summary, the proposed algorithm is presented in Algorithm 1. As in the ROF case, the range $\tau\sigma < 1/2$ comes from Proposition 2.2.

Next, we analyze convergence of the proposed method. Since the sequence $\{(\tilde{\mathbf{p}}^{(n)}, \lambda^{(n)})\}$ generated by Algorithm 1 agrees with the one generated by the standard primal-dual algorithm, $O(1/n)$ ergodic convergence is guaranteed.

Theorem 4.3 *Let $\{(\tilde{\mathbf{p}}^{(n)}, \lambda^{(n)})\}$ be the sequence generated by Algorithm 1. Then, $(\tilde{\mathbf{p}}^{(n)}, \lambda^{(n)})$ converges to a saddle point $(\tilde{\mathbf{p}}^*, \lambda^*)$ of (4.2) and it satisfies*

$$L(\tilde{\mathbf{p}}_n, \lambda) - L(\tilde{\mathbf{p}}, \lambda_n) \leq \frac{1}{n} \left(\frac{1}{\tau} \|\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^{(0)}\|_{2,\tilde{Y}}^2 + \frac{1}{\sigma} \|\lambda - \lambda^{(0)}\|_{2, \mathbb{R}^{|\mathcal{I}_\Gamma|}}^2 \right)$$

for any $\tilde{\mathbf{p}} \in \tilde{Y}$ and $\lambda \in \mathbb{R}^{|\mathcal{I}_\Gamma|}$, where $\tilde{\mathbf{p}}_n = \frac{1}{n} \sum_{i=1}^{n} \tilde{\mathbf{p}}^{(i)}$ and $\lambda_n = \frac{1}{n} \sum_{i=1}^{n} \lambda^{(i)}$.

Proof Since the term

$$\sum_{s=1}^{\mathcal{N}} \left\{ \frac{1}{\alpha} F_s^*(\operatorname{div} \tilde{\mathbf{p}}_s) + \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) \right\}$$

in (4.2) is convex, by Theorem 5.1 in [4], we get the desired result. $\qquad \square$

By Theorem 4.3, we ensure the convergence of the sequences $\{\tilde{\mathbf{p}}^{(n)}\}$ and $\{\lambda^{(n)}\}$. On the other hand, since the sequence $\{u^{(n)}\}$ is a kind of byproduct of the primal-dual algorithm, the convergence theory for the primal-dual algorithm developed in [3, 4] does not ensure global convergence of $\{u^{(n)}\}$. Thus, we will prove that it tends to a solution of (3.1). For the sake of convenience, we rewrite (4.4) as the following more compact form:

$$\min_{u \in X} \max_{\tilde{\mathbf{p}} \in \tilde{Y}} \left\{ -\left\langle u, \widetilde{\operatorname{div}} \tilde{\mathbf{p}} \right\rangle_X + \alpha F(u) - \chi_{\tilde{C}}(\tilde{\mathbf{p}}) - \frac{1}{2\tau} \|\tilde{\mathbf{p}} - \hat{\mathbf{p}}^{(n+1)}\|_{2,\tilde{Y}}^2 \right\}, \qquad (4.6)$$

where $\widetilde{\operatorname{div}} : \tilde{Y} \to X$ is defined as $\widetilde{\operatorname{div}} \tilde{\mathbf{p}} = \bigoplus_{s=1}^{\mathcal{N}} \operatorname{div} \tilde{\mathbf{p}}_s$. Note that if $\tilde{\mathbf{p}} \in \ker B$, then $\tilde{\mathbf{p}} \in Y$ so that $\widetilde{\operatorname{div}} \tilde{\mathbf{p}} = \operatorname{div} \tilde{\mathbf{p}}$. Furthermore, since $\tilde{Y}$ has dofs on Γ, we have

$\ker \widetilde{\mathrm{div}}^* = \{0\}$. Then, we readily see that $u^{(n+1)}$ is a solution of the minimization problem

$$\min_{u \in X} \left\{ \alpha F(u) + G_{n+1}(-\widetilde{\mathrm{div}}^* u) \right\},$$

where G_n is defined as

$$G_n(\tilde{\mathbf{q}}) = \max_{\tilde{\mathbf{p}} \in \tilde{Y}} \left\{ \langle \tilde{\mathbf{p}}, \tilde{\mathbf{q}} \rangle_{\tilde{Y}} - \chi_{\tilde{C}}(\tilde{\mathbf{p}}) - \frac{1}{2\tau} \| \tilde{\mathbf{p}} - \hat{\mathbf{p}}^{(n)} \|_{2,\tilde{Y}}^2 \right\},$$

the Legendre–Fenchel conjugate of

$$\chi_{\tilde{C}}(\tilde{\mathbf{p}}) + \frac{1}{2\tau} \| \tilde{\mathbf{p}} - \hat{\mathbf{p}}^{(n)} \|_{2,\tilde{Y}}^2.$$

At first, we verify the boundedness of $\{G_n\}$.

Lemma 4.4 *There exist a finite functional* $\overline{G} \colon \tilde{Y} \to \mathbb{R}$ *and a coercive functional* $\underline{G} \colon \tilde{Y} \to \mathbb{R}$ *such that*

$$\underline{G} \le G_n \le \overline{G} \qquad \forall n \in \mathbb{Z}_{>0}.$$

Proof We first see that for any $\tilde{\mathbf{q}} \in \tilde{Y}$, we have $\frac{\tilde{\mathbf{q}}}{|\tilde{\mathbf{q}}|} \in \tilde{C}$, where $|\tilde{\mathbf{q}}|$ is the dof-wise absolute value of $\tilde{\mathbf{q}}$ and the division is done dof-wise with convention $\frac{0}{|0|} = 0$. An upper bound of $\{G_n\}$ is obtained as follows:

$$\begin{aligned}
G_n(\tilde{\mathbf{q}}) &= \max_{\tilde{\mathbf{p}} \in \tilde{C}} \left\{ \langle \tilde{\mathbf{p}}, \tilde{\mathbf{q}} \rangle_{\tilde{Y}} - \frac{1}{2\tau} \| \tilde{\mathbf{p}} - \hat{\mathbf{p}}^{(n)} \|_{2,\tilde{Y}}^2 \right\} \\
&\le \max_{\tilde{\mathbf{p}} \in \tilde{C}} \langle \tilde{\mathbf{p}}, \tilde{\mathbf{q}} \rangle_{\tilde{Y}} \\
&= \left\langle \frac{\tilde{\mathbf{q}}}{|\tilde{\mathbf{q}}|}, \tilde{\mathbf{q}} \right\rangle_{\tilde{Y}} = \| \tilde{\mathbf{q}} \|_{1,\tilde{Y}}.
\end{aligned}$$

Take $\overline{G}(\tilde{\mathbf{q}}) = \| \tilde{\mathbf{q}} \|_{1,\tilde{Y}}$, which is clearly finite.

Now, we find a lower bound of $\{G_n\}$. Note that $\hat{\mathbf{p}}^{(n)} = \tilde{\mathbf{p}}^{(n-1)} - \tau B^* \lambda^{(n)}$ in Algorithm 1. Since $\{\tilde{\mathbf{p}}^{(n)}\}$ and $\{\lambda^{(n)}\}$ are convergent by Theorem 4.3, $\{\hat{\mathbf{p}}^{(n)}\}$ is also convergent. Hence, $\{\hat{\mathbf{p}}^{(n)}\}$ is bounded, i.e. there exists a constant $L > 0$ such that

$$\| \hat{\mathbf{p}}^{(n)} \|_{2,\tilde{Y}}^2 \le L \qquad \forall n \in \mathbb{Z}_{>0}.$$

Then, we have

$$\begin{aligned}
G_n(\tilde{\mathbf{q}}) &= \max_{\tilde{\mathbf{p}} \in \tilde{C}} \left\{ \langle \tilde{\mathbf{p}}, \tilde{\mathbf{q}} \rangle_{\tilde{Y}} - \frac{1}{2\tau} \| \tilde{\mathbf{p}} - \hat{\mathbf{p}}^{(n)} \|_{2,\tilde{Y}}^2 \right\} \\
&\ge \max_{\tilde{\mathbf{p}} \in \tilde{C}} \left\{ \langle \tilde{\mathbf{p}}, \tilde{\mathbf{q}} \rangle_{\tilde{Y}} - \frac{1}{\tau} \left(\| \tilde{\mathbf{p}} \|_{2,\tilde{Y}}^2 + \| \hat{\mathbf{p}}^{(n)} \|_{2,\tilde{Y}}^2 \right) \right\} \\
&\ge \max_{\tilde{\mathbf{p}} \in \tilde{C}} \left\{ \langle \tilde{\mathbf{p}}, \tilde{\mathbf{q}} \rangle_{\tilde{Y}} - \frac{1}{\tau} \left(\| \tilde{\mathbf{p}} \|_{2,\tilde{Y}}^2 + L \right) \right\} \\
&\ge \left\langle \frac{\tilde{\mathbf{q}}}{|\tilde{\mathbf{q}}|}, \tilde{\mathbf{q}} \right\rangle_{\tilde{Y}} - \frac{1}{\tau} \left(\left\| \frac{\tilde{\mathbf{q}}}{|\tilde{\mathbf{q}}|} \right\|_{2,\tilde{Y}}^2 + L \right) \\
&\ge \| \tilde{\mathbf{q}} \|_{1,\tilde{Y}} - \frac{1}{\tau} \left(\sum_{s=1}^{\mathcal{N}} |\tilde{\mathcal{I}}_s| + L \right).
\end{aligned}$$

Note that $\tilde{\mathcal{I}}_s$ denotes the set of indices of the basis functions for $\tilde{Y}_s$ and the value of $\sum_{s=1}^{\mathcal{N}} |\tilde{\mathcal{I}}_s|$ depends on the image size and the number of subdomains $\mathcal{N}$. Take

$$\underline{G}(\tilde{\mathbf{q}}) = \|\tilde{\mathbf{q}}\|_{1,\tilde{Y}} - \frac{1}{\tau}\left(\sum_{s=1}^{\mathcal{N}} |\tilde{\mathcal{I}}_s| + L\right),$$

which is coercive due to the term $\|\tilde{\mathbf{q}}\|_{1,\tilde{Y}}$. $\qquad\square$

Next lemma provides a criterion for equi-coercivity of a collection of functionals.

Lemma 4.5 *Let $\{F_i\}$, $i = 1, 2, \ldots$, be a collection of functionals from X into $\bar{\mathbb{R}}$. If there exists a coercive functional $F_0 \colon X \to \bar{\mathbb{R}}$ such that $F_i \geq F_0$ for all i, then $\{F_i\}$ is equi-coercive, that is, for every $t \geq 0$, there exists a compact subset K_t of X such that*

$$\{u \in X : F_i(u) \leq t\} \subset K_t \quad \text{for all } i = 1, 2, \ldots.$$

Proof Let $t \geq 0$. Since F_0 is coercive, there exists a compact subset K_t of X such that

$$\{u \in X : F_0(u) \leq t\} \subset K_t.$$

On the other hand, for every i, $F_i \geq F_0$ implies that

$$\{u \in X : F_i(u) \leq t\} \subset \{u \in X : F_0(u) \leq t\}.$$

Therefore, $\{F_i\}$ is equi-coercive. $\qquad\square$

Now, with the help of the lemmas above, we state the main theorem, which ensures that the sequence $\{u^{(n)}\}$ generated by Algorithm 1 approaches to solutions of (3.1).

Theorem 4.6 *Let $\{u^{(n)}\}$ be the sequence generated by Algorithm 1. Then, $\{u^{(n)}\}$ is bounded and every limit point of $\{u^{(n)}\}$ is a solution of (3.1).*

Proof Recall that $u^{(n)}$ is a solution of the minimization problem

$$\min_{u \in X} \left\{ \alpha F(u) + G_n(-\widetilde{\text{div}}^* u) \right\}.$$

Since F is proper, we may choose $u_0 \in X$ with $F(u_0) < \infty$. By Lemma 4.4, we have

$$t = \alpha F(u_0) + \overline{G}(-\widetilde{\text{div}}^* u_0) < \infty.$$

Thanks to the minimization property of $u^{(n)}$, we get

$$\alpha F(u^{(n)}) + G_n(-\widetilde{\text{div}}^* u^{(n)}) \leq \alpha F(u_0) + G_n(-\widetilde{\text{div}}^* u_0)$$
$$\leq \alpha F(u_0) + \overline{G}(-\widetilde{\text{div}}^* u_0) = t.$$

By Lemmas 4.4 and 4.5, $\{G_n\}$ is equi-coercive, that is, there exists a compact subset K_t of $\tilde{Y}$ independent of n such that

$$\left\{ \tilde{\mathbf{q}} \in \tilde{Y} : G_n(\mathbf{q}) \leq t \right\} \subset K_t.$$

Thus, $G_n(-\widetilde{\operatorname{div}}^* u^{(n)}) \leq t$ implies that $-\widetilde{\operatorname{div}}^* u^{(n)} \in K_t \cap \operatorname{ran} \widetilde{\operatorname{div}}^*$. Since $\ker \widetilde{\operatorname{div}}^* = \{0\}$, the map $-\widetilde{\operatorname{div}}^*$ is a continuous isomorphism between X and $\operatorname{ran} \widetilde{\operatorname{div}}^*$. Therefore, we can deduce that

$$u^{(n)} \in (-\widetilde{\operatorname{div}}^*)^{-1}(K_t \cap \operatorname{ran} \widetilde{\operatorname{div}}^*),$$

which is a compact subset of X independent of n. This implies that $\{u^{(n)}\}$ is bounded.

Now, we may refine $\{u^{(n)}\}$ so that it converges to its limit point u^*. Since $(u^{(n)}, \tilde{\mathbf{p}}^{(n)})$ is a solution of a saddle point problem (4.6), it satisfies

$$\frac{\widetilde{\operatorname{div}} \tilde{\mathbf{p}}^{(n)}}{\alpha} \in \partial F(u^{(n)}).$$

By Theorem 4.3, $\{\tilde{\mathbf{p}}^{(n)}\}$ converges to $\tilde{\mathbf{p}}^*$. As $u^{(n)} \to u^*$ and $\frac{\widetilde{\operatorname{div}} \tilde{\mathbf{p}}^{(n)}}{\alpha} \to \frac{\widetilde{\operatorname{div}} \tilde{\mathbf{p}}^*}{\alpha}$, by closedness of the graph of ∂F (See Theorem 24.4 in [23]), we get

$$\frac{\widetilde{\operatorname{div}} \tilde{\mathbf{p}}^*}{\alpha} \in \partial F(u^*).$$

By Proposition 4.2, $\tilde{\mathbf{p}}^* \in \ker B$, and hence $\tilde{\mathbf{p}}^* \in Y$. We obtain the relation

$$0 \in -\operatorname{div} \tilde{\mathbf{p}}^* + \alpha \partial F(u^*).$$

From the facts that $\tilde{\mathbf{p}}^*$ is a solution of (3.4) (See Proposition 4.2) and $(u^*, \tilde{\mathbf{p}}^*)$ satisfies the relation (3.5), we conclude that u^* is a solution of (3.1). $\square$

As a direct consequence of of Theorem 4.6, we get the following result.

Corollary 4.7 *Let $\{u^{(n)}\}$ be the sequence generated by Algorithm 1. Then, $\{\mathcal{J}(u^{(n)})\}$ converges to $\mathcal{J}(u^*)$, where u^* is a solution of* (3.1).

5 Applications

In this section, we apply our proposed DDM to various total variation based image processing problems mentioned in Sect. 3. The proposed method was implemented in ANSI C with OpenMPI and compiled by Intel Parallel Studio XE. All the computations were done on a cluster composed of seven machines, where each machine has two Intel Xeon SP-6148 CPUs (2.4GHz, 20C), 192GB memory, and the operating system CentOS 7.4 64bit.

To emphasize efficiency of the proposed method as a parallel algorithm, we compare the wall-clock time of the proposed method with the primal-dual algorithm [3] for the full dimension problem (3.4). The wall-clock time measures the total elapsed time including the communication time.

At each iteration of Algorithm 1, we solve the local saddle point problems of the form

$$\min_{u_s \in X_s} \max_{\tilde{\mathbf{p}}_s \in \tilde{Y}_s} \left\{ -\langle u_s, \operatorname{div} \tilde{\mathbf{p}}_s \rangle_{X_s} + \alpha F_s(u_s) - G_s^*(\tilde{\mathbf{p}}_s) \right\}, \tag{5.1}$$

where F_s is given in Assumption 3.1 and

$$G_s^*(\tilde{\mathbf{p}}_s) = \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) + \frac{1}{2\tau} \|\tilde{\mathbf{p}}_s - \hat{\mathbf{p}}_s^{(n+1)}\|_{2, \tilde{Y}_s}^2.$$

Test image	$\mathcal{N}$	ALG1				LNP			
		PSNR	iter	max inner iter	wall-clock time (sec)	PSNR	iter	max inner iter	wall-clock time (sec)
Cameraman 2048 × 2048	1	47.68	-	546	86.30				
	2 × 2	48.26	39	266	155.85	46.75	33	1761	587.86
	4 × 4	48.26	45	277	50.40	46.75	61	1679	141.31
	8 × 8	48.25	49	274	14.79	46.78	134	1796	42.37
	16 × 16	48.24	54	279	2.19	46.79	175	1200	7.82
Boat 2048 × 3072	1	33.82	-	2093	476.02				
	2 × 2	33.92	61	283	531.97	33.60	43	1815	970.12
	4 × 4	33.92	63	276	165.45	33.62	71	1823	255.70
	8 × 8	33.92	66	258	53.68	33.62	114	1413	87.23
	16 × 16	33.92	71	246	8.18	33.63	161	844	14.07

Table 1 Performance of the proposed method for the image denoising problem with the L^1 fidelity term (5.2)

The primal-dual algorithm for (5.1) consists of computation of the proximity operators of $\alpha F_s(u_s)$ and $G_s^*(\tilde{\mathbf{p}}_s)$, that is,

$$u_s = \text{prox}_{\sigma_0 \alpha F_s}(\bar{u}_s),$$
$$\tilde{\mathbf{p}}_s = \text{prox}_{\tau_0 G_s^*}(\bar{\mathbf{p}}_s),$$

for some $\sigma_0, \tau_0 > 0$. For more details, we refer readers to see [3]. The proximity operator of $G_s^*(\tilde{\mathbf{p}}_s)$ is computed easily as follows:

$$\text{prox}_{\tau_0 G_s^*}(\bar{\mathbf{p}}_s) = \underset{\tilde{\mathbf{p}}_s \in \tilde{Y}_s}{\arg\min} \left\{ \chi_{\tilde{C}_s}(\tilde{\mathbf{p}}_s) + \frac{1}{2\tau}\|\tilde{\mathbf{p}}_s - \hat{\mathbf{p}}_s^{(n+1)}\|_{2,\tilde{Y}_s}^2 + \frac{1}{2\tau_0}\|\tilde{\mathbf{p}}_s - \bar{\mathbf{p}}_s\|_{2,\tilde{Y}_s}^2 \right\}$$
$$= \text{proj}_{\tilde{C}_s}\left(\frac{\tau \bar{\mathbf{p}}_s + \tau_0 \hat{\mathbf{p}}_s^{(n+1)}}{\tau + \tau_0} \right).$$

Computation of the proximity operator of $\alpha F_s(u_s)$ depends on the problem to solve. Thus, we will give details in each subsection.

Since G_s^* is uniformly convex with parameter $1/\tau$, we are able to adopt the $O(1/n^2)$ convergent primal-dual algorithm [3, Algorithm 2] for the local problems. Such acceleration of the local solvers for DDMs was discussed in [20, 21].

Next, we provide the setting of the used parameters. We set the parameters for the outer iterations by $\tau = 50$ and $\sigma\tau = 1/2$ for the image denoising and inpainting problems, and $\tau = 1, \sigma\tau = 1/2$ for the image segmentation problem. For the $O(1/n^2)$ convergent algorithms for the local problems, we set $\gamma = 1/8\tau$, $\tau_0 = 10$, and $\sigma_0\tau_0 = 1/8$. (The same notations for the parameters are used as in [3]). The full dimension problems ($\mathcal{N} = 1$) are solved by the primal-dual algorithm with the same parameters as the corresponding local problems.

Finally, we note that the local solutions $(u_s^{(n-1)}, \tilde{\mathbf{p}}_s^{(n-1)})$ from the previous outer iteration were chosen as initial guesses for the local problems to reduce the number of inner iterations.

5.1 Image denoising

We present the results of numerical experiments for the $TV\text{-}L^1$ model for image denoising:

$$\min_{u \in X} \left\{ \alpha\|u - f\|_{1,X} + TV(u) \right\}. \tag{5.2}$$

(a) Noisy "Cameraman 2048 × 2048" (PSNR: 12.07)

(b) $\mathcal{N} = 1$ (PSNR: 47.68)

(c) $\mathcal{N} = 16 \times 16$ (PSNR: 48.24)

(d) Noisy "Boat 2048 × 3072" (PSNR: 12.34)

(e) $\mathcal{N} = 1$ (PSNR: 33.82)

(f) $\mathcal{N} = 16 \times 16$ (PSNR: 33.92)

Fig. 1 Results of the proposed method for the image denoising problem with the L^1 fidelity term (5.2)

(a) Cameraman 2048 × 2048

(b) Boat 2048 × 3072

Fig. 2 Decay of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{\mathcal{J}(u^*)}$ with respect to the number of iterations n for various algorithms, applied to the image denoising problem with the L^1 fidelity term (5.2)

We note that numerical results of the proposed method for ROF model were given in [21]. In (5.2), $\alpha F_s(u_s)$ is given by

$$\alpha F_s(u_s) = \alpha \|u_s - f\|_{1, X_s}$$

and its proximity operator can be computed as

$$\mathrm{prox}_{\sigma_0 \alpha F_s}(\bar{u}_s) = f + \mathrm{shr}_{s, \Omega_s}(\bar{u}_s - f, \sigma_0 \alpha),$$

where the shrinkage operator $\text{shr}_{s,S}$ on $S \subset \Omega_s$ is defined by replacing Ω and $\mathcal{T}$ by Ω_s and $\mathcal{T}_s$ in (3.8), respectively.

We use two test images "Cameraman 2048×2048" and "Boat 2048×3072" with 20% salt-and-pepper noise (See Fig. 1). PSNR denotes peak signal-to-noise ratio. We set $\alpha = 1$.

Thanks to Proposition 3.4, it is able to compare the proposed method with the existing methods for (5.2) based on the finite difference discretization. The following algorithms are used for our performance evaluation:

- ALG1: Proposed method described in Algorithm 1, $\mathcal{N} = 8 \times 8$.
- LNP: DDM proposed by Lee, Nam, and Park [20] with the anisotropic total variation, $\mathcal{N} = 8 \times 8$, $\tau = 0.05$, $\sigma\tau = 1/8$.
- CP: Primal-dual algorithm proposed by Chambolle and Pock [3] with the anisotropic total variation, $\tau = 10$, $\sigma\tau = 1/8$.

In the numerical experiments, the following stopping criterion is used for outer iterations:

$$\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{\mathcal{J}(u^*)} < 10^{-5}, \tag{5.3}$$

and the following ones are used for inner iterations of ALG1 and LNP, respectively:

$$\frac{\|\mathbf{p}_s^{(n+1)} - \mathbf{p}_s^{(n)}\|_2}{\|\mathbf{p}_s^{(n+1)}\|_2} < 10^{-6}, \tag{5.4a}$$

$$\frac{\|u_s^{(n+1)} - u_s^{(n)}\|_2}{\|u_s^{(n+1)}\|_2} < 10^{-6}. \tag{5.4b}$$

Fig. 1 shows the denoised images obtained by ALG1 with $\mathcal{N} = 16 \times 16$ and CP ($\mathcal{N} = 1$). In order to highlight the efficiency of the proposed method as a parallel solver, Table 1 shows the performance of the methods with the varying number of subdomains. In addition, to compare the convergence rate of the proposed method with existing algorithms, we present Fig. 2 which shows decay of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{\mathcal{J}(u^*)}$ for 1000 iterations of three algorithms ALG1, LNP, and CP, where the minimum primal energy $\mathcal{J}(u^*)$ is computed approximately by 10^6 iterations of the primal-dual algorithm.

5.2 Image inpainting

We first consider the inpainting model with the L^2 fidelity term:

$$\min_{u \in X} \left\{ \frac{\alpha}{2} \|Au - f\|_{2,X}^2 + TV(u) \right\}. \tag{5.5}$$

Here, $A\colon X \to X$ is the restriction operator onto $\Omega \setminus D$, so that its matrix representation is a diagonal matrix whose diagonal entries are either 0 or 1. Thus, computation of the proximity operator of $\alpha F_s(u_s) = \frac{\alpha}{2}\|A_s u_s - f\|_{2,X_s}^2$ is as easy as the case of the denoising problem. Indeed, we have

$$\left(\text{prox}_{\sigma_0 \alpha F_s}(\bar{u}_s)\right)_T = \begin{cases} \frac{(\bar{u}_s)_T + \sigma_0 \alpha (f)_T}{1 + \sigma_0 \alpha} & \text{if } T \subset \Omega_s \setminus D, \\ (\bar{u}_s)_T & \text{if } T \subset D, \end{cases} \quad T \in \mathcal{T}_s.$$

(a) Corrupted "Cameraman 2048 × 2048" (PSNR: 15.83)

(b) $\mathcal{N} = 1$ (PSNR: 24.96)

(c) $\mathcal{N} = 16 \times 16$ (PSNR: 25.00)

(d) Corrupted "Boat 2048 × 3072" (PSNR: 16.66)

(e) $\mathcal{N} = 1$ (PSNR: 24.32)

(f) $\mathcal{N} = 16 \times 16$ (PSNR: 24.33)

Fig. 3 Results of the proposed method for the image inpainting problem with the L^2 fidelity term (5.5)

(a) Cameraman 2048 × 2048

(b) Boat 2048 × 3072

Fig. 4 Decay of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{\mathcal{J}(u^*)}$ with respect to the number of iterations n for various algorithms, applied to the image inpainting problem with the L^2 fidelity term (5.5)

Two test images "Cameraman 2048 × 2048" and "Boat 2048 × 3072" corrupted by additive Gaussian noise with mean 0 and variance 0.05 and a text mask are used for numerical experiments (See Fig. 3). The model parameter α is set as $\alpha = 10$. Table 2 shows the computation results of the proposed method for the varying number of subdomains. Stop conditions (5.3) and (5.4a) are used for outer iterations and inner iterations of ALG1, respectively. Fig. 3 shows the resulting images of ALG1 with $\mathcal{N} = 16 \times 16$ and CP. Fig. 4 shows decay of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{\mathcal{J}(u^*)}$

Test image	$\mathcal{N}$	PSNR	iter	max inner iter	wall-clock time (sec)
Cameraman 2048×2048	1	24.96	-	2324	389.88
	2×2	25.00	438	613	805.88
	4×4	25.00	440	713	260.06
	8×8	25.00	445	99	88.96
	16×16	25.00	460	99	13.40
Boat 2048×3072	1	24.32	-	1861	436.32
	2×2	24.33	285	817	1085.96
	4×4	24.33	286	588	361.04
	8×8	24.33	289	111	132.14
	16×16	24.33	298	117	25.63

Table 2 Performance of the proposed method for the image inpainting problem with the L^2 fidelity term (5.5)

(a) Corrupted "Cameraman 2048×2048" (PSNR: 11.41)

(b) $\mathcal{N} = 1$ (PSNR: 34.70)

(c) $\mathcal{N} = 16 \times 16$ (PSNR: 35.53)

(d) Corrupted "Boat 2048×3072" (PSNR: 11.92)

(e) $\mathcal{N} = 1$ (PSNR: 30.83)

(f) $\mathcal{N} = 16 \times 16$ (PSNR: 31.07)

Fig. 5 Results of the proposed method for the image inpainting problem with the L^1 fidelity term (5.6)

for two algorithms ALG1 and CP. Here, $\mathcal{J}(u^*)$ is computed by 10^6 iterations of the primal-dual algorithm.

Now, we consider the following L^1 inpainting model:

$$\min_{u \in X} \left\{ \alpha \|Au - f\|_{1,X} + TV(u) \right\}. \tag{5.6}$$

Similarly to the L^2 inpainting problem, the proximity operator of $\alpha F_s(u_s) = \alpha \|A_s u_s - f\|_{1,X}$ is given by

$$\text{prox}_{\sigma_0 \alpha F_s}(\bar{u}_s) = f + \text{shr}_{s, \Omega_s \setminus D}(\bar{u}_s - f, \sigma_0 \alpha).$$

(a) Cameraman 2048 × 2048   (b) Boat 2048 × 3072

Fig. 6 Decay of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{\mathcal{J}(u^*)}$ with respect to the number of iterations n for various algorithms, applied to the image inpainting problem with the L^1 fidelity term (5.6)

Test image	$\mathcal{N}$	ALG1				LNP			
		PSNR	iter	max inner iter	wall-clock time (sec)	PSNR	iter	max inner iter	wall-clock time (sec)
Cameraman 2048 × 2048	1	34.70	-	1693	304.05				
	2 × 2	35.56	78	921	615.5	34.18	173	1530	1132.71
	4 × 4	35.53	86	958	244.21	34.24	190	1672	323.98
	8 × 8	35.53	90	274	68.44	34.26	205	1796	97.65
	16 × 16	35.53	101	279	16.10	34.37	239	1200	24.67
Boat 2048 × 3072	1	30.83	-	2262	604.34				
	2 × 2	31.08	80	860	1283.93	30.53	368	1703	2134.45
	4 × 4	31.08	84	765	432.68	30.55	374	1888	635.88
	8 × 8	31.08	88	258	128.09	30.56	384	1413	217.33
	16 × 16	31.07	97	246	31.95	30.58	402	844	60.68

Table 3 Performance of the proposed method for the image inpainting problem with the L^1 fidelity term (5.6)

Test images are corrupted by 20% salt-and-pepper noise and the same text mask as the L^2 inpainting problem (See Fig. 5). We use $\alpha = 1$ as the model parameter. Numerical results of ALG1, LNP, and CP for (5.6) are given in Table 3. Both ALG1 and LNP uses (5.3) as the stopping criterion for outer iterations. Stopping criteria for local problems are given in (5.4a) and (5.4b) for ALG1 and LNP, respectively. The recovered images obtained by ALG1 and CP are given in Fig. 5. Fig. 6 shows decay of the value of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{\mathcal{J}(u^*)}$ for three algorithms ALG1, LNP, and CP.

5.3 Image segmentation

As we mentioned in Sect. 3, the convex Chan–Vese model for the image segmentation model is represented as

$$\min_{u \in X} \left\{ \alpha \langle u, g \rangle_X + \chi_{\{0 \leq u \leq 1\}}(u) + TV(u) \right\}, \tag{5.7}$$

where $g = (f - c_1)^2 - (f - c_2)^2$. The proximity operator of $\alpha F_s(u_s) = \langle u, g \rangle_X + \chi_{\{0 \leq u \leq 1\}}(u)$ is computed as

$$\mathrm{prox}_{\sigma_0 \alpha F_s}(\bar{u}_s) = \mathrm{proj}_{\{0 \leq \cdot \leq 1\}}(\bar{u}_s - \sigma_0 \alpha g).$$

(a) Cameraman 2048×2048 (b) $\mathcal{N} = 1$ (c) $\mathcal{N} = 16 \times 16$

Fig. 7 Results of the proposed method for the image segmentation problem (5.7)

Fig. 8 Decay of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{|\mathcal{J}(u^*)|}$ with respect to the number of iterations n for various algorithms, applied to the image segmentation problem (5.7)

Test image	$\mathcal{N}$	ALG1			DCT		
		iter	max inner iter	wall-clock time (sec)	iter	max inner iter	wall-clock time (sec)
Cameraman 2048×2048	1	-	1070	172.47			
	2×2	16	258	153.72	14	290	85.67
	4×4	16	63	37.58	15	52	18.21
	8×8	16	55	6.65	15	52	5.31
	16×16	16	56	1.10	16	55	1.43

Table 4 Performance of the proposed method for the image segmentation problem (5.7)

As shown in Fig. 7, we use a test image "Cameraman 2048×2048." We set the model parameters $\alpha = 10$, $c_1 = 0.6$, and $c_2 = 0.1$ heuristically. Also, to convert the results to binary functions, we use a threshold parameter $1/2$. Numerical results of ALG1, DCT, and CP for (5.7) are presented in Table 4, where DCT denotes the following algorithm:

– DCT: DDM proposed by Duan, Chang, and Tai [10] with the anisotropic total variation, $\mathcal{N} = 8 \times 8$, $\tau = 1$, $\sigma\tau = 1/8$.

Both ALG1 and DCT uses (5.3) as a stop condition for outer iterations. For stop conditions for local problems, ALG1 uses (5.4a) while DCT uses (5.4b). The segmentation results obtained by ALG1 and CP are provided in Fig. 7. Fig. 8

presents the energy decay of several algorithms containing the proposed method for (5.7). Since the minimum primal energy $\mathcal{J}(u^*)$ is negative in this case, we plot the value of $\frac{\mathcal{J}(u^{(n)}) - \mathcal{J}(u^*)}{|\mathcal{J}(u^*)|}$, where $\mathcal{J}(u^*)$ is computed by 10^6 iterations of the primal-dual algorithm.

5.4 Discussion

As shown in Figs. 1, 3, 5, and 7, the results of the full dimension problem and the proposed method are not visually distinguishable. Note that the resulting images of the proposed method show no trace of the subdomain interfaces. In Tables 1–4, one can observe that solutions for different values of $\mathcal{N}$ have different PSNRs. It means that the algorithms converge to different solutions. This is because (1.2) admits nonunique solutions in general. However, as the stopping criterion (5.3) indicates, their primal energies tend to the minimum.

In Tables 1–4, one can see that the number of outer iterations monotonically increases as the number of subdomains $\mathcal{N}$ increases. We also observe that the numbers of maximum inner iterations of the proposed method are much smaller than the numbers of iterations of the full dimension problem. The reason is that we utilize more accelerated solvers for the local problems than standard ones. Reduction of the numbers of inner iterations makes the proposed method faster.

For every problem, we see that the wall-clock time is decreasing as the number of subdomains $\mathcal{N}$ grows. In particular, large scale images such as "Cameraman 2048×2048" and "Boat 2048×3072" can be processed in a minute with sufficiently many subdomains, while it takes quite a long time with a single domain. This shows efficiency of the proposed method as a parallel solver for image processing.

Finally, as shown in Figs. 2, 4, 6, and 8, the proposed method shows good performance in terms of the decay rate of the primal energy $\mathcal{J}(u^{(n)})$. To be more precise, the primal energy of the proposed method decreases much faster than the one of CP for the full-dimension problem. The proposed method also outperforms LNP which is a recently developed DDM for $TV\text{-}L^1$ problems. Even more, for the segmentation problem, the convergence rate of the proposed method seems to be faster than $O(1/n)$. However, a theoretical evidence for such fast convergence is still missing. It is observed in Figs. 2, 4, 6, and 8 that the primal energy of the proposed method becomes stagnant when the relative error is sufficiently small. This is due to that local problems are solved inexactly by iterative methods in each iteration. Such stagnation of the primal energy is not problematic in practice because the quality of the recovered image becomes acceptable enough before the stagnation starts. Meanwhile, in terms of wall-clock time, the proposed method outperforms LNP for both denoising and inpainting problems while it shows the similar performance to DCT for the segmentation problem.

6 Conclusion

In this paper, we generalized the primal-dual DDM for the ROF model proposed in [21] to more general total variation minimization problem. The Fenchel–Rockafellar dual of the model problem was considered. We constructed the constrained minimization problem which has domain decomposition structure and is

equivalent to the dual model problem. The constrained minimization problem was converted to the equivalent saddle point problem by the method of Lagrange multipliers. The resulting saddle point problem was solved by the first order primal-dual algorithm and convergence of the dual solution was guaranteed. We also proved convergence of the primal solution. Numerical results showed that the proposed method is superior to the existing methods in the sense of convergence rate, and is much faster with sufficiently many subdomains than the full dimension problem.

Even though the proposed DDM is applicable for various total variation regularized problems, we point out that the proposed method is not appropriate for the image deconvolution problem. We will develop a DDM with similar strategy which is applicable for the image deconvolution problem later.

References

1. Bartels, S.: Total variation minimization with finite elements: convergence and iterative solution. SIAM J. Numer. Anal. **50**(3), 1162–1180 (2012)
2. Chambolle, A., Levine, S.E., Lucier, B.J.: An upwind finite-difference method for total variation-based image smoothing. SIAM J. Imaging Sci. **4**(1), 277–299 (2011)
3. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**(1), 120–145 (2011)
4. Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. Acta Numer. **25**, 161–319 (2016)
5. Chan, T.F., Esedoglu, S.: Aspects of total variation regularized L^1 function approximation. SIAM J. Appl. Math. **65**(5), 1817–1837 (2005)
6. Chan, T.F., Esedoglu, S., Nikolova, M.:Algorithms for finding global minimizers of image segmentation and denoising models. SIAM J. Appl. Math. **66**(5), 1632–1648 (2006)
7. Chan, T.F., Wong, C.-K.: Total variation blind deconvolution. IEEE Trans. Image Process. **7**(3), 370–375 (1998)
8. Chang, H., Tai, X.-C., Wang, L.-L., Yang, D.: Convergence rate of overlapping domain decomposition methods for the Rudin-Osher-Fatemi model based on a dual formulation. SIAM J. Imaging Sci. **8**(1), 564–591 (2015)
9. Dong, Y., Hintermüller, M., Neri, M.: An efficient primal-dual method for L^1TV image restoration. SIAM J. Imaging Sci. **2**(4), 1168–1189 (2009)
10. Duan, Y., Chang, H., Tai, X.-C.: Convergent non-overlapping domain decomposition methods for variational image segmentation. J. Sci. Comput. **69**(2), 532–555 (2016)
11. Farhat, C., Roux, F.-X.: A method of finite element tearing and interconnecting and its parallel solution algorithm. Int. J. Numer. Methods Eng. **32**(6), 1205–1227 (1991)
12. Fornasier, M.: Domain decomposition methods for linear inverse problems with sparsity constraints. Inverse Probl. **23**(6), 2505–2526 (2007)
13. Fornasier, M., Langer, A., Schönlieb, C.-B.: A convergent overlapping domain decomposition method for total variation minimization. Numer. Math. **116**(4), 645–685 (2010)
14. Fornasier, M., Schönlieb, C.-B.: Subspace correction methods for total variation and $l1$-minimization. SIAM J. Numer. Anal. **47**(5), 3397–3428 (2009)
15. Hermann, M., Herzog, R., Schmidt, S., Vidal-Núñez, J., Wachsmuth, G.: Discrete total variation with finite elements and applications to imaging. J. Math. Imaging Vis. **61**(4), 411–431 (2019)
16. Hintermüller, M., Langer, A.: Subspace correction methods for a class of nonsmooth and nonadditive convex variational problems with mixed L^1/L^2 data-fidelity in image processing. SIAM J. Imaging Sci. **6**(4), 2134–2173 (2013)
17. Hintermüller, M., Langer, A.: Non-overlapping domain decomposition methods for dual total variation based image denoising. J. Sci. Comput. **62**(2), 456–481 (2015)
18. Lee, C.-O., Lee, J.H., Woo, H., Yun, S.: Block decomposition methods for total variation by primal–dual stitching. J. Sci. Comput. **68**(1), 273–302 (2016)
19. Lee, C.-O., Nam, C.: Primal domain decomposition methods for the total variation minimization, based on dual decomposition. SIAM J. Sci. Comput. **39**(2), B403–B423 (2017)
20. Lee, C.-O., Nam, C., Park, J.: Domain decomposition methods using dual conversion for the total variation minimization with L^1 fidelity term. J. Sci. Comput. **78**(2), 951–970 (2019)

21. Lee, C.-O., Park, E.-H., Park, J.: A finite element approach for the dual Rudin–Osher–Fatemi model and its nonoverlapping domain decomposition methods. SIAM J. Sci. Comput. **41**(2), B205–B228 (2019)
22. Nikolova, M.: A variational approach to remove outliers and impulse noise. J. Math. Imaging Vis. **20**(1), 99–120 (2004)
23. Rockafellar, R.T.: Convex Analysis. Princeton University Press, New Jersey (2015)
24. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D. **60**(1-4), 259–268 (1992)
25. Shen, J., Chan, T.F.: Mathematical models for local nontexture inpaintings. SIAM J. Appl. Math. **62**(3), 1019–1043 (2002)
26. Strong, D., Chan, T.F.: Edge-preserving and scale-dependent properties of total variation regularization. Inverse Probl. **19**(6), S165–S187 (2003)
27. Wang, J., Lucier, B.J.: Error bounds for finite-difference methods for Rudin–Osher–Fatemi image smoothing. SIAM J. Numer. Anal. **49**(2), 845–868 (2011)
28. Wang, Y., Yang, J., Yin, W., Zhang, Y.: A new alternating minimization algorithm for total variation image reconstruction. SIAM J. Imaging Sci. **1**(3), 248–272 (2008)
29. You, Y.-L., Kaveh, M.: A regularization approach to joint blur identification and image restoration. IEEE Trans. Image Process. **5**(3), 416–428 (1996)