

CONVERGENCE RATE ANALYSIS OF OVERLAPPING DDMS FOR THE ROF MODEL BASED ON A DUAL FORMULATION

HUIBIN CHANG¹, XUE-CHENG TAI², LI-LIAN WANG³ AND DANPING YANG⁴

ABSTRACT. This paper is concerned with overlapping domain decomposition methods (DDMs), based on successive subspace correction (SSC) and parallel subspace correction (PSC), for the Rudin, Osher and Fatemi (ROF) model in image restoration. In contrast to recent attempts, we work with a dual formulation of the ROF model, where one significant difficulty lies in the decomposition of the global constraint of the dual problem. We propose a stable "Unit Decomposition" and this leads us to come natural overlapping domain decomposition schemes for the dual problem. We further analyze the convergence of the proposed algorithms, and obtain the rate $O(n^{-1/2})$ where n is the number of iterations. Move, the dependence of the convergence rate on the overlapping size, regularization parameter and relaxation parameter is clearly given. To the best of our knowledge, such a convergence has not been claimed so far for domain decomposition related algorithms the ROF model.

1. INTRODUCTION

The Rudin-Osher-Fatemi model given in [27] is one of the fundamental variational models for image processing. It is known that the ROF model restores a noise image g on a domain Ω (e.g. in \mathbb{R}^2) by solving the minimization problem:

$$\min_{u \in BV(\Omega)} \left\{ \lambda TV(u) + \frac{1}{2} \|u - g\|_{L^2(\Omega)}^2 \right\}, \quad (1.1)$$

where $\lambda > 0$ and $BV(\Omega)$ is the space of functions of bounded variation, and the total variation of u is defined as in [1] by

$$TV(u) := \sup_{\mathbf{p} \in \mathcal{K}} \int_{\Omega} u \operatorname{div} \mathbf{p} \, d\mathbf{x} \quad \text{with} \quad (1.2)$$

$$\mathcal{K} := \{ \mathbf{p} = (p_1, p_2) \in (C_0^1(\Omega))^2 : |\mathbf{p}| := (p_1^2 + p_2^2)^{1/2} \leq 1 \}. \quad (1.3)$$

As usual, \mathbf{p} is known as the dual variable, while u is the primal variable.

1991 *Mathematics Subject Classification.* 68U10, 65M55, 74S20 .

Key words and phrases. Dual model, Successive Subspace Correction, Parallel Subspace Correction, Convergence rate .

The first three authors were supported by MOE IDM project NRF2007IDM-IDM002-010, Singapore. The first author was also partially supported by PHD Programme 52XB1304 of Tianjin Normal University, PHD Program Scholarship Fund of ECNU with Grant No.2010026 and Overseas Research Fund of East China Normal University, China. The last author was supported by Natural Science Foundation of China, Grant 11071080.

¹School of Mathematical Sciences, Tianjin Normal University, Tianjin, 300387, P.R. China.

²Department of Mathematics, University of Bergen, Bergen, Norway.

³Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore.

⁴Department of Mathematics, East China Normal University, Shanghai, 200241, PR China.

Over the last two decades, many methods have been proposed to solve the ROF model. In general, they can be classified into three categories based on the nature of manipulating the primal and/or dual variables and one can refer to [29].

- Primal approaches. The gradient descent method (cf. [25, 27]) via the evolution of a parabolic equation or lagged diffusivity fixed-point iteration (cf. [1, 37, 38, 39, 40]) via linearized the curvature term is reliable and restores the image of high quality, but slowly convergent. The most popular algorithms based on Bregman iteration (cf. [16, 41, 42]) and augmented Lagrangian methods (cf. [18, 43]) are quite efficient and fast because the reduced problems can be solved by variable-splitting technique and FFT for subproblems. Usually after few iterations degraded images are well restored. For other algorithms like graph cuts method, additive operator splitting (AOS), multigrid method, one can refer to the [44] which gives the concise review and detailed references.
- Dual approaches. A typical and efficient approach (see, e.g., [5]) is to apply the KKT condition to (1.1) and (1.2), which allows to solve the dual variable:

$$\inf_{\mathbf{p} \in \mathcal{K}} \left\{ D(\mathbf{p}) := \int_{\Omega} (\lambda \operatorname{div} \mathbf{p} - g)^2 d\mathbf{x} \right\}, \quad (1.4)$$

and finally update the primal variable by

$$u = g - \lambda \operatorname{div} \mathbf{p}. \quad (1.5)$$

- Primal-dual approaches. This type of approach was first introduced in [2, 3]. Then the applications to image processing were done in [4, 11, 47]. By virtue of Legendre-Fenchel's duality, they formed the saddle point problem. Finally iterations including alternating dual and primal steps were constructed. In [11] they generalized the primal-dual hybrid gradient (PDHG) algorithm [47] and a general framework was set up. Further in [4] they proved the convergence of a primal-dual gap to zero whose convergence was of first order. Some accelerated techniques were adopted to improve the algorithm with second order as $O(\frac{1}{n^2})$ with n is the iteration number.

There has been much recent interest in developing domain decomposition methods for image processing. As an important numerical means for PDEs, the DDM enjoys two apparent merits: (i) it can break down the problem into a sequence of subproblems of much smaller scale, so better-conditioned solvers can be constructed over each subdomain; and (ii) it allows for parallel computations.

For domain decomposition methods [6, 10, 17, 28, 46] they built up the framework for general applications. It is well known that there exist two different approaches to apply domain decomposition methods to the variational problems. One is to use domain decomposition methods to the Euler-Lagrangian equations of the variational problems which are usually linear partial differential equations. Then lots of techniques for linear partial differential equations, such as [10, 22, 23, 24, 28, 46], consisting of two level methods, multigrid methods and precondition technique can be adopted to solve the problem in parallel. But if the Euler-Lagrangian equation is not linear, it is difficult to give the convergence analysis and the energy is not guaranteed to decrease monotonely. The other is to use parallel

or successive subspace correction methods [30, 31, 32, 33, 34] to decompose the variational problems into sub-variational problems over subdomains. The energy decreases monotonely. Besides, this approach can be applied to more general, or complex variational problems.

In the areas for applying DDM in image processing, DDM as mentioned above was adopted. The classical domain decomposition methods with Dirichlet boundary condition transmission (which is equivalent to PSC or SSC) was adopted for in image denoising using Gaussian curvature [15], PSC or SSC was applied to graph cuts for image restoration and segmentation in [9, 35, 44]. During these applications they successfully decomposed the original problem into more small subproblems which were solved parallel. With fewer outer iterations they got the satisfying processed images. Especially in [44], a coarse mesh correction has been added to the proposed domain decomposition scheme which is often not an easy task for nonlinear problems. Xu and Chang [45] also applied the DDMs to the image deblurring problems, which owned the special properties of the blur operator. Chang et al.[7] considered the DDMs for the nonlocal total variation based image restoration problems. They also pointed out their proposed algorithms can be adopted to solve the total variation problems directly.

Meanwhile, some variant of classical DDM was proposed in [12, 13, 14, 20, 21]. Instead of solving the subproblem deduced by SSC or PSC directly, Fornasier, Langer, and Schönlieb introduced the surrogate functional to form an approximation(or iterative proximity-map) of the subproblem. Then the new subproblem was solved by via oblique thresholding. The algorithm was tested to be efficient in image restoration and compress sensing. The algorithm proposed in Hintermuller-Lamnger [20] is a nonlinear subspace correction algorithm for the ROF model. In [21], subspace correction technique is combined with the Bregmann-splitting idea to get domain decomposition type of algorithms. Convergence of the algorithm was proved in [20, 14, 21]. However, convergence rate was unknown.

In contrast with most of existing work, we are interested in the DDM for the dual formulation (1.4). Indeed, the dual approach, e.g., the Chambolle's algorithm, has proven to be efficient for solving ROF model. However, for image of large size, it is advantageous to employ domain decomposition technique for the vector-valued dual variable. However, one difficulty lies in the decomposition of the convex set \mathcal{K} with the constraint $|\mathbf{p}| \leq 1$.

The original ROF model is highly nonlinear and the owns some singularity. The dual model overcomes these difficulties. So we try to apply the domain decomposition methods to dual model. However, the new difficulties exist in two aspects. On one hand, during the realization or construction of the algorithm, the global constraint of the dual variable is difficult to be decomposed over the subdomains. On the other hand, for the convergence analysis, the objective functional of the dual model is not strictly convex or strongly convex. This property is causing problem to use the theory of [30] to obtain the convergence rate.

In the following section, we will supply some techniques to overcome these difficulties and deduce the convergence rate for the proposed domain decomposition methods. We focus on dealing with model (1.4) using subspace correction methods Successive Subspace Correction (SSC) and Parallel Subspace Correction (PSC) methods in our paper.

The paper is organized as follows. We give the overlapping domain decomposition method in Section 2. The convergence analysis is presented in Section 3. Detailed algorithms and numerical examples for the proposed algorithms are given in Section 4. At last we conclude this paper in Section 5.

2. THE OVERLAPPING DOMAIN DECOMPOSITION ALGORITHMS

In this section, we formulate the overlapping domain decomposition algorithms for (1.4). For this purpose, we introduce the space of functions with square-integrable divergence:

$$H(\operatorname{div}; \Omega) = \{\mathbf{p} \in (L^2(\Omega))^2 : \operatorname{div} \mathbf{p} \in L^2(\Omega)\}, \quad (2.1)$$

equipped with the graph norm

$$\|\mathbf{p}\|_{H(\operatorname{div}; \Omega)} = \left(\|\mathbf{p}\|_{(L^2(\Omega))^2}^2 + \|\operatorname{div} \mathbf{p}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}.$$

Define the subspace by density:

$$H_0(\operatorname{div}; \Omega) = \text{closure of } (C_0^\infty(\Omega))^2 \text{ in the } H(\operatorname{div}; \Omega) \text{ graph norm.} \quad (2.2)$$

It is known that (see, e.g., [26, Theorem 3.25]) if Ω is a bounded Lipschitz domain in \mathbb{R}^2 , then

$$H_0(\operatorname{div}; \Omega) = \{\mathbf{p} \in H(\operatorname{div}; \Omega) : \mathbf{p} \cdot \mathbf{n}|_{\partial\Omega} = 0\}, \quad (2.3)$$

where \mathbf{n} is the outer unit vector normal to $\partial\Omega$. In view of this, we define

$$K := \{\mathbf{p} \in H_0(\operatorname{div}; \Omega) : |\mathbf{p}| \leq 1\}, \quad (2.4)$$

and consider the alternative formulation of (1.4):

$$\min_{\mathbf{p} \in K} \left\{ D(\mathbf{p}) := \int_{\Omega} (\lambda \operatorname{div} \mathbf{p} - g)^2 dx \right\}, \quad (2.5)$$

Notice that the functional $D(\mathbf{p})$ is convex but not strictly convex, so the problem (2.5) admits at least one minimizer $\mathbf{p}^* \in K$. To ensure the uniqueness, one may modify the energy functional (see, e.g., [19, 8]) and consider

$$D_\varepsilon(\mathbf{p}) = D(\mathbf{p}) + \varepsilon \|\mathbf{p}\|_{L^2(\Omega)}^2, \quad \text{where } 0 < \varepsilon \ll 1, \quad (2.6)$$

but this requires a delicate convergence analysis for $\varepsilon \rightarrow 0$. Hence, it is more desirable to directly work with the problem (2.5). Indeed, we just need one optimum \mathbf{p}^* to resolve the primal variable via $u^* = g - \lambda \operatorname{div} \mathbf{p}^*$.

2.1. General setup. To formulate the DDM algorithms, we start with the decomposition of the domain Ω and the convex constraint set K . For clarity of presentation, we assume that Ω is a rectangular domain. We partition Ω into M_c classes of overlapping subdomains, and suppose that each class is colored with a different color, that is,

$$\Omega = \bigcup_{j=1}^{M_c} \Omega_j, \quad \text{where } \Omega_j \text{ is a union of } m_j \text{ disjoint subdomains with the same color.} \quad (2.7)$$

We assume that the overlapping size of subdomains is δ . Hence, the total number of subdomains that covers Ω is $m_T = \sum_{j=1}^{M_c} m_j$. We illustrate in Figure 2.1 a typical decomposition of a rectangular domains with subdomains of four colors.

In what follows, let N_0 be the maximum number of subdomains, among all m_T subdomains of the partition, yacht an over at a point $x \in \Omega$. For example, $N_0 = 4$ for the partition depicted in Figure 2.1.

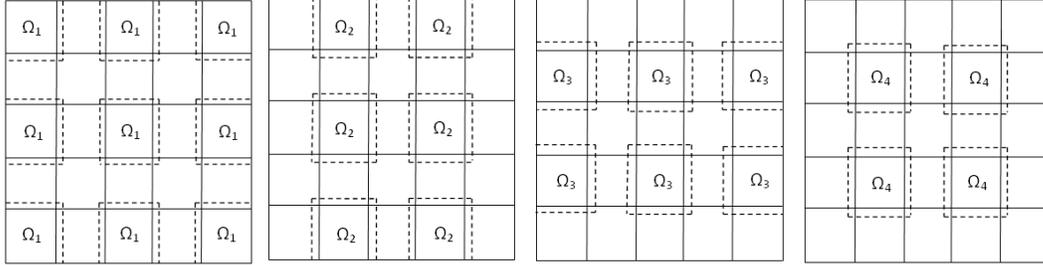


FIGURE 2.1. Domain decomposition with coloring technique. Here, $M_c = 4$, $m_1 = 9$, $m_2 = 6$, $m_3 = 6$, $m_4 = 4$ and $m_T = 25$.

An essential technique for our method is to decompose the convex constraint set K into a sum of convex constraint sets over the subdomains. For this purpose, we need to use an "Unit Partition Function" (UPF), denoted by $\{\theta_j\}_{j=1}^{M_c}$, which are one-to-one associated with $\{\Omega_j\}_{j=1}^{M_c}$, and satisfy

$$(i) \sum_{j=1}^{M_c} \theta_j \equiv 1, \quad \theta_j \geq 0, \quad \text{on } \Omega; \quad (2.8)$$

$$(ii) \theta_j \in H^1(\Omega), \quad \text{supp}(\theta_j) \subset \bar{\Omega}_j, \quad 1 \leq j \leq M_c; \quad (2.9)$$

$$(iii) \|\nabla \theta_j\|_\infty \leq \frac{C_0}{\delta}, \quad 1 \leq j \leq M_c, \quad (2.10)$$

where C_0 is a positive constant independent of δ , and $\|\cdot\|_\infty$ as usual is the L^∞ -norm.

Correspondingly, we define the convex subsets as:

$$K_j = \{\mathbf{p} \in H_0(\text{div}; \Omega) : |\mathbf{p}| \leq \theta_j\}, \quad 1 \leq j \leq M_c. \quad (2.11)$$

It follows from (2.9) that for any $\mathbf{p} \in K_j$, we have $\text{supp}(\mathbf{p}) \subset \bar{\Omega}_j$.

Proposition 2.1. *We have*

$$K = K_1 + K_2 + \cdots + K_{M_c}. \quad (2.12)$$

Proof. On one hand, for any $\mathbf{p} \in K$, we define $\mathbf{p}_j = \theta_j \mathbf{p}$, for $1 \leq j \leq M_c$. By the definition (2.8)-(2.10), we find that $\mathbf{p}_j \in K_j$, and $\mathbf{p} = \sum_{j=1}^{M_c} \mathbf{p}_j$. On the other hand, given $\mathbf{p} = \sum_{j=1}^{M_c} \mathbf{p}_j$ with $\mathbf{p}_j \in K_j$, we have $\mathbf{p} \in H_0(\text{div}; \Omega)$, and

$$|\mathbf{p}| = \left| \sum_{j=1}^{M_c} \mathbf{p}_j \right| \leq \sum_{j=1}^{M_c} |\mathbf{p}_j| \leq \sum_{j=1}^{M_c} \theta_j = 1,$$

so $\mathbf{p} \in K$. □

2.2. Formulation of DDMs. With the above setup, we are now ready to formulate the DDMs based on the formulation (2.5). Hereafter, let $\alpha > 0$ be the relaxation parameter.

Algorithm I: Parallel Subspace Correction (PSC)

1. Initialization: choose \mathbf{p}^0 and select a relaxation parameter¹

$$\alpha \in (0, M_c^{-1}]. \quad (2.13)$$

2. For $n = 0, 1, \dots$,
compute \mathbf{p}^{n+1} by

$$\mathbf{p}^{n+1} = (1 - \alpha)\mathbf{p}^n + \alpha \sum_{j=1}^{M_c} \hat{\mathbf{q}}_j^n, \quad (2.14)$$

where

$$\hat{\mathbf{q}}_j^n = \arg \min_{\mathbf{v} \in K_j} D\left(\mathbf{v} + \sum_{i \neq j} \theta_i \mathbf{p}^n\right), \quad 1 \leq j \leq M_c. \quad (2.15)$$

3. Endfor till some stopping criterion meets.
-

We see that the minimization of (2.5) breaks down to solving a sequence of subdomain problems (2.15) with much smaller scale, which can be resolved in parallel. Notice that when we update \mathbf{p}^{n+1} by (2.14), we could use the available qualities $\{\hat{\mathbf{q}}_j^n\}$ to replace the corresponding values $\{\mathbf{p}^n|_{K_j}\}$. This leads to the successive subspace correction algorithm. Indeed, the situation is reminiscent of the Jacobi and Gauss-Seidel methods for solving linear systems.

Algorithm II: Successive Subspace Correction (SSC)

1. Initialization: choose \mathbf{p}^0 and select a relaxation parameter $\alpha \in (0, 1]$.
2. For $n = 0, 1, \dots$, find \mathbf{p}^{n+1} in the following two steps:

- (i) Find $\{\hat{\mathbf{q}}_j^n\}_{j=1}^{M_c}$ sequentially for $1 \leq j \leq M_c$ such that

$$\hat{\mathbf{q}}_j^n = \arg \min_{\mathbf{v} \in K_j} D\left(\mathbf{v} + \sum_{i < j} \mathbf{q}_i^n + \sum_{i > j} \theta_i \mathbf{p}^n\right), \quad (2.16)$$

and then define

$$\mathbf{q}_j^n = (1 - \alpha)\theta_j \mathbf{p}^n + \alpha \hat{\mathbf{q}}_j^n.$$

- (ii) Update

$$\mathbf{p}^{n+1} = (1 - \alpha)\mathbf{p}^n + \alpha \sum_{j=1}^{M_c} \hat{\mathbf{q}}_j^n. \quad (2.17)$$

3. Endfor till some stopping rule meets.
-

¹The section of α in this specific interval will be justified by the analysis in the forthcoming section. This also applies to **Algorithm II**.

Remark 2.1. The difference between the above algorithms with those in [30] mainly lies in two aspects. First, the functional is strongly and strictly convexity. Second, there exists the convex set constraints over vector \mathbf{p} to the variational problems. These bring about significant difficulties for analysis and lead to different convergence behaviors as well.

Remark 2.2. The coloring of the subdomains is needed for the analysis and implementation for the SSC algorithms. For the PSC algorithms, the coloring is only needed for the analysis, but not needed in the implementations.

3. CONVERGENCE RATE ANALYSIS

This section is devoted to the convergence analysis of the previously proposed overlapping DDMS. We first state the main result, and then present the some necessary lemmas of relevance to the proof, followed by the proof of the main result. Here, we restrict our attentions to analyze **Algorithm I**, since the convergence of **Algorithm II** can be estimated in a similar fashion (see [30]).

3.1. Main result.

Theorem 3.1. *Let \mathbf{p}^* be a minimizer of (2.5), i.e.,*

$$\mathbf{p}^* = \arg \min_{\mathbf{p} \in K} D(\mathbf{p}), \quad (3.1)$$

and let $\{\mathbf{p}^n\}$ be the sequence generated by **Algorithm I**. Define

$$u^n := g - \lambda \operatorname{div} \mathbf{p}^n, \quad u^* := g - \lambda \operatorname{div} \mathbf{p}^*. \quad (3.2)$$

Then we have

$$\|u^n - u^*\|_{L^2(\Omega)} \leq \frac{C}{\sqrt{n}}, \quad (3.3)$$

where

$$C = \sqrt{\zeta^0} \left\{ \frac{2}{\alpha} (1 + 8M_c^2 + 4\sqrt{2}M_c) + \left(16C_0\lambda|\Omega|^{\frac{1}{2}}(\zeta^0)^{-\frac{1}{2}} \right) \frac{M_c\sqrt{N_0}}{\delta\sqrt{\alpha}} + \sqrt{2} - 1 \right\}, \quad (3.4)$$

and $\zeta^0 = D(\mathbf{p}^0) - D(\mathbf{p}^*)$ is the initial error of the energy functional. Here, the constant C_0 is used in (2.10) which implies the upper bound of the unit decomposition ($N_0 = 4$ if $\Omega \in \mathbb{R}^2$), and M_c is the number of subdomains. The parameter α is the relaxation parameter in **Algorithm I**, and δ is the overlapping size.

The proof of Theorem 3.1 is postponed to the Subsection 3.3 and 3.4. Here we give some remarks about the relationship between the convergence rates and the overlapping size δ , and the number of blocks m_T .

Remark 3.1. The theorem indicates the convergence of the algorithm as $n \rightarrow 0$. Further, the convergence rate is of half order $O(n^{-1/2})$, where n is the iteration number of outer loop.

Remark 3.2. We see that the constant C depends on the parameters such as overlapping size δ and the number of colors M_c . Usually we can use fixed number of colors to divide the partitions. From (3.4), one readily obtains C increases as overlapping size δ decreases. That is to say, the convergence becomes slow as the overlapping size decreases. This constant also relies on the regularized parameter λ , which implies that our proposed DDMs are sensitive to λ . One can observe that from Figure 4.12 in the numerical tests.

3.2. Some lemmas. For clarity of presentation, we denote the norm of $L^2(\Omega)$ by $\|\cdot\|$, and define $|\mathbf{p}|_{1,*} = \|\lambda \operatorname{div} \mathbf{p}\|$. Let $D(\mathbf{p})$ be the cost functional defined in (1.4).

$$D(\mathbf{p}) - D(\mathbf{q}) - D'(\mathbf{q})(\mathbf{p} - \mathbf{q}) = |\mathbf{p} - \mathbf{q}|_{1,*}^2, \quad (3.5)$$

where $D'(\mathbf{q})$ defines the Gâteaux derivative. One verifies readily that

$$(D'(\mathbf{p}) - D'(\mathbf{q}), \mathbf{p} - \mathbf{q}) = 2|\mathbf{p} - \mathbf{q}|_{1,*}^2. \quad (3.6)$$

Lemma 3.1. *There hold the following two inequalities:*

$$\left(\sum_{i=1}^{M_c} |\theta_i \mathbf{p}|_{1,*}^2 \right)^{\frac{1}{2}} \leq \sqrt{2} |\mathbf{p}|_{1,*} + \frac{C_0 \sqrt{2N_0} \lambda}{\delta} \|\mathbf{p}\|, \quad \forall \mathbf{p} \in K \quad (3.7)$$

and

$$\sum_{i,j=1}^{M_c} |\langle D'(\mathbf{q}_{ij} + \mathbf{p}_i) - D'(\mathbf{q}_{ij}), \hat{\mathbf{p}}_j \rangle| \leq 2M_c \left(\sum_{i=1}^{M_c} |\mathbf{p}_i|_{1,*}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{M_c} |\hat{\mathbf{p}}_i|_{1,*}^2 \right)^{\frac{1}{2}} \quad (3.8)$$

for any $\mathbf{q}_{ij} \in K$, $\mathbf{p}_i \in K_i$ and $\hat{\mathbf{p}}_j \in K_j$ for $1 \leq i, j \leq M_c$, where C_0 already exists in (2.10).

Proof. Firstly, we prove (3.7). It is clear that

$$|\theta_i \mathbf{p}|_{1,*}^2 = \lambda^2 \|\operatorname{div}(\theta_i \mathbf{p})\|^2 = \lambda^2 \|\nabla \theta_i \cdot \mathbf{p} + \theta_i \operatorname{div} \mathbf{p}\|^2 \leq 2\lambda^2 \|\nabla \theta_i \cdot \mathbf{p}\|^2 + 2\lambda^2 \|\theta_i \operatorname{div} \mathbf{p}\|^2.$$

Summing up the above equations over i from 1 to M_c and using (2.10) yields

$$\begin{aligned} \sum_{i=1}^{M_c} |\theta_i \mathbf{p}|_{1,*}^2 &\leq 2\lambda^2 \int_{\Omega} \left(\sum_{i=1}^{M_c} |\nabla \theta_i|^2 \right) \mathbf{p}^2 + 2\lambda^2 \int_{\Omega} \left(\sum_{i=1}^{M_c} \theta_i^2 \right) (\operatorname{div} \mathbf{p})^2 \\ &\leq 2N_0 \lambda^2 \left(\frac{C_0}{\delta} \right)^2 \|\mathbf{p}\|^2 + 2\lambda^2 \int_{\Omega} \left(\sum_{i=1}^{M_c} \theta_i \right) (\operatorname{div} \mathbf{p})^2 \leq 2N_0 \lambda^2 \left(\frac{C_0}{\delta} \right)^2 \|\mathbf{p}\|^2 + 2|\mathbf{p}|_{1,*}^2. \end{aligned}$$

Thus we obtain

$$\left(\sum_{i=1}^{M_c} |\theta_i \mathbf{p}|_{1,*}^2 \right)^{\frac{1}{2}} \leq \sqrt{2} |\mathbf{p}|_{1,*} + \frac{\sqrt{2N_0} C_0 \lambda}{\delta} \|\mathbf{p}\|.$$

Now, we prove the second inequality (3.8). One readily verifies that

$$\langle D'(\mathbf{q}_{ij} + \mathbf{p}_i) - D'(\mathbf{q}_{ij}), \hat{\mathbf{p}}_j \rangle = 2\lambda^2 (\operatorname{div} \mathbf{p}_i, \operatorname{div} \hat{\mathbf{p}}_j).$$

Summing up the above over i, j from 1 to M_c , leads to

$$\begin{aligned} \sum_{i,j=1}^{M_c} |\langle D'(\mathbf{q}_{ij} + \mathbf{p}_i) - D'(\mathbf{q}_{ij}), \hat{\mathbf{p}}_j \rangle| &\leq 2 \sum_{i,j=1}^{M_c} \lambda^2 |(\operatorname{div} \mathbf{p}_i, \operatorname{div} \hat{\mathbf{p}}_j)| \\ &\leq 2M_c \left(\sum_{i=1}^{M_c} |\mathbf{p}_i|_{1,*}^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{M_c} |\hat{\mathbf{p}}_j|_{1,*}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

This ends the proof. \square

Define

$$\begin{aligned} \mathbf{e}_i^n &:= \hat{\mathbf{q}}_i^n - \theta_i \mathbf{p}^n, & \hat{\mathbf{q}}^n &:= \sum_{i=1}^{M_c} \hat{\mathbf{q}}_i^n = \mathbf{p}^n + \sum_{i=1}^{M_c} \mathbf{e}_i^n, \\ \mathbf{q}_{\frac{i}{M_c}}^n &:= \left(\sum_{j \neq i} \theta_j \mathbf{p}^n \right) + \hat{\mathbf{q}}_i^n = \mathbf{p}^n + \mathbf{e}_i^n. \end{aligned} \quad (3.9)$$

From (2.15) we derive the first order optimality condition

$$\langle D'(\mathbf{q}_{\frac{i}{M_c}}^n), \tilde{\mathbf{q}} - \hat{\mathbf{q}}_i^n \rangle \geq 0, \quad \forall \tilde{\mathbf{q}} \in K_i, \quad 1 \leq i \leq M_c,$$

which is equivalent to

$$\int_{\Omega} (\lambda \operatorname{div} \mathbf{q}_{\frac{i}{M_c}}^n - g) \operatorname{div}(\tilde{\mathbf{q}} - \hat{\mathbf{q}}_i^n) \geq 0, \quad \forall \tilde{\mathbf{q}} \in K_i. \quad (3.10)$$

The following Lemmas in this subsection are important, which give the analysis for two successive iteration solutions.

Lemma 3.2. *There holds*

$$D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}) \geq \alpha \sum_{i=1}^{M_c} |\mathbf{e}_i^n|_{1,*}^2. \quad (3.11)$$

Proof. It follows from (2.17) and (3.9) that

$$\mathbf{p}^{n+1} = \mathbf{p}^n + \alpha \sum_{i=1}^{M_c} (\hat{\mathbf{q}}_i^n - \theta_i \mathbf{p}^n) = \mathbf{p}^n + \alpha \sum_{i=1}^{M_c} (\mathbf{q}_{\frac{i}{M_c}}^n - \mathbf{p}^n) = (1 - M_c \alpha) \mathbf{p}^n + \alpha \sum_{i=1}^{M_c} \mathbf{q}_{\frac{i}{M_c}}^n. \quad (3.12)$$

Thus by using (3.5), (3.12) and $1 - M_c \alpha \geq 0$ due to (2.13), we obtain

$$\begin{aligned} D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}) &= D(\mathbf{p}^n) - D\left((1 - M_c \alpha) \mathbf{p}^n + \alpha \sum_{i=1}^{M_c} \mathbf{q}_{\frac{i}{M_c}}^n\right) \\ &\geq D(\mathbf{p}^n) - (1 - M_c \alpha) D(\mathbf{p}^n) - \alpha \sum_{i=1}^{M_c} D(\mathbf{q}_{\frac{i}{M_c}}^n) \\ &= \alpha \sum_{i=1}^{M_c} (D(\mathbf{p}^n) - D(\mathbf{q}_{\frac{i}{M_c}}^n)) \\ &= -\alpha \sum_{i=1}^{M_c} \langle D'(\mathbf{q}_{\frac{i}{M_c}}^n), \mathbf{e}_i^n \rangle + \alpha \sum_{i=1}^{M_c} |\mathbf{e}_i^n|_{1,*}^2. \end{aligned} \quad (3.13)$$

Setting $\tilde{\mathbf{q}} := \theta_i \mathbf{p}^n \in K_i$ in (3.10) yields

$$\langle D'(\mathbf{q}_{\frac{i}{M_c}}^n), \mathbf{e}_i^n \rangle \leq 0. \quad (3.14)$$

Therefore we deduce (3.11) from (3.13) and (3.14). \square

Lemma 3.2 guarantees the decreasing of the energy. The following lemma provides more precise estimate of the energy decay.

Lemma 3.3. *Given $\mu \in (0, 1)$, we have*

$$D(\mathbf{p}^{n+1}) - D(\mathbf{p}^*) \leq \gamma(D(\mathbf{p}^n) - D(\mathbf{p}^*)) + C_3(D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}))^{\frac{1}{2}}, \quad (3.15)$$

where

$$C_3 = \frac{2C_1\sqrt{\alpha\mu}|\Omega|^{\frac{1}{2}}}{\mu + C_2}, \quad \gamma = 1 - \frac{\alpha(1-\mu)\mu}{\mu + C_2}, \quad (3.16)$$

with

$$C_1 = \frac{4C_0\lambda M_c\sqrt{N_0}}{\delta}, \quad C_2 = 4M_c^2 + 2\sqrt{2}M_c \quad (3.17)$$

Proof. By the convexity of $D(\cdot)$, we have

$$\begin{aligned} D(\mathbf{p}^{n+1}) - D(\mathbf{p}^*) &\leq (1-\alpha)D(\mathbf{p}^n) + \alpha D(\hat{\mathbf{q}}^n) - D(\mathbf{p}^*) \\ &= (1-\alpha)(D(\mathbf{p}^n) - D(\mathbf{p}^*)) + \alpha(D(\hat{\mathbf{q}}^n) - D(\mathbf{p}^*)). \end{aligned} \quad (3.18)$$

Let us estimate $D(\hat{\mathbf{q}}^n) - D(\mathbf{p}^*)$. Introduce functions as

$$\phi_j^i = \begin{cases} \mathbf{p}^n + \sum_{k=i}^{j+i-1} \mathbf{e}_k^n, & j \in [1, M_c - i + 1], \\ \mathbf{p}^n + \sum_{k=i}^{M_c} \mathbf{e}_k^n + \sum_{k=1}^{j-M_c+i-1} \mathbf{e}_k^n, & j \in [M_c - i + 2, M_c]. \end{cases} \quad (3.19a)$$

$$\quad (3.19b)$$

We therefore obtain

$$D'(\mathbf{p}^n + \sum_{i=1}^{M_c} \mathbf{e}_i^n) - D'(\mathbf{p}^n + \mathbf{e}_i^n) = \sum_{j=2}^{M_c} (D'(\phi_j^i) - D'(\phi_{j-1}^i)). \quad (3.20)$$

Using (3.7), (3.8), (3.10) and (3.20), we derive

$$\begin{aligned} \langle D'(\hat{\mathbf{q}}^n), \hat{\mathbf{q}}^n - \mathbf{p}^* \rangle &= \sum_{i=1}^{M_c} \langle D'(\hat{\mathbf{q}}^n), \hat{\mathbf{q}}_i^n - \theta_i \mathbf{p}^* \rangle \leq \sum_{i=1}^{M_c} \langle D'(\hat{\mathbf{q}}^n) - D'(\mathbf{p}^n + \mathbf{e}_i^n), \hat{\mathbf{q}}_i^n - \theta_i \mathbf{p}^* \rangle \\ &= \sum_{i=1, j=2}^{M_c} \langle D'(\phi_j^i) - D'(\phi_{j-1}^i), \hat{\mathbf{q}}_i^n - \theta_i \mathbf{p}^* \rangle \leq 2M_c \left(\sum_{j=1}^{M_c} |\mathbf{e}_j^n|_{1,*}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{M_c} |\hat{\mathbf{q}}_i^n - \theta_i \mathbf{p}^*|_{1,*}^2 \right)^{\frac{1}{2}} \\ &= 2M_c \left(\sum_{j=1}^{M_c} |\mathbf{e}_j^n|_{1,*}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{M_c} |\mathbf{e}_i^n + \theta_i \mathbf{p}^n - \theta_i \mathbf{p}^*|_{1,*}^2 \right)^{\frac{1}{2}}, \end{aligned}$$

so

$$\begin{aligned} \langle D'(\hat{\mathbf{q}}^n), \hat{\mathbf{q}}^n - \mathbf{p}^* \rangle &\leq 2\sqrt{2}M_c \left(\sum_{j=1}^{M_c} |\mathbf{e}_j^n|_{1,*}^2 \right)^{\frac{1}{2}} \left(\left(\sum_{i=1}^{M_c} |\mathbf{e}_i^n|_{1,*}^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^{M_c} |\theta_i \mathbf{p}^n - \theta_i \mathbf{p}^*|_{1,*}^2 \right)^{\frac{1}{2}} \right) \\ &\leq 2\sqrt{2}M_c \left(\sum_{j=1}^{M_c} |\mathbf{e}_j^n|_{1,*}^2 \right)^{\frac{1}{2}} \left(\left(\sum_{i=1}^{M_c} |\mathbf{e}_i^n|_{1,*}^2 \right)^{\frac{1}{2}} + \sqrt{2}|\mathbf{p}^n - \mathbf{p}^*|_{1,*} + \frac{C\sqrt{2N_0}\lambda}{\delta} \|\mathbf{p}^n - \mathbf{p}^*\| \right) \\ &= 2\sqrt{2}M_c \left(\sum_{j=1}^{M_c} |\mathbf{e}_j^n|_{1,*}^2 \right)^{\frac{1}{2}} + 4M_c |\mathbf{p}^n - \mathbf{p}^*|_{1,*} \left(\sum_{j=1}^{M_c} |\mathbf{e}_j^n|_{1,*}^2 \right)^{\frac{1}{2}} \\ &\quad + \frac{4C\lambda M_c \sqrt{N_0}}{\delta} \|\mathbf{p}^n - \mathbf{p}^*\| \left(\sum_{j=1}^{M_c} |\mathbf{e}_j^n|_{1,*}^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (3.21)$$

From the first order optimal condition of (2.5) that $\langle D'(\mathbf{p}^*), \tilde{\mathbf{q}} - \mathbf{p}^* \rangle \geq 0$ for any $\tilde{\mathbf{q}}$ in K and (3.5), we obtain

$$D(\mathbf{p}^n) - D(\mathbf{p}^*) \geq |\mathbf{p}^n - \mathbf{p}^*|_{1,*}^2. \quad (3.22)$$

Using inequality $ab \leq \frac{a^2}{4\mu} + \mu b^2$ with $0 < \mu < 1$ and (3.11), (3.21) and (3.22), we obtain

$$\begin{aligned} D(\hat{\mathbf{q}}^n) - D(\mathbf{p}^*) &\leq \langle D'(\hat{\mathbf{q}}^n), \hat{\mathbf{q}}^n - \mathbf{p}^* \rangle \\ &\leq \frac{2\sqrt{2}M_c}{\alpha} (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1})) + \frac{4M_c}{\sqrt{\alpha}} (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}))^{\frac{1}{2}} (D(\mathbf{p}^n) - D(\mathbf{p}^*))^{\frac{1}{2}} \\ &\quad + \frac{4C\lambda M_c \sqrt{N_0}}{\delta \sqrt{\alpha}} (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}))^{\frac{1}{2}} \|\mathbf{p}^n - \mathbf{p}^*\| \\ &\leq \frac{1}{\alpha\mu} (4M_c^2 + 2\sqrt{2}M_c) (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1})) + \mu (D(\mathbf{p}^n) - D(\mathbf{p}^*)) \\ &\quad + \frac{4C\lambda M_c \sqrt{N_0}}{\delta \sqrt{\alpha}} (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}))^{\frac{1}{2}} \|\mathbf{p}^n - \mathbf{p}^*\| \end{aligned} \quad (3.23)$$

such that

$$\begin{aligned} D(\hat{\mathbf{q}}^n) - D(\mathbf{p}^*) &\leq \frac{C_2}{\alpha\mu} (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1})) + \mu (D(\mathbf{p}^n) - D(\mathbf{p}^*)) \\ &\quad + \frac{C_1}{\sqrt{\alpha}} \|\mathbf{p}^n - \mathbf{p}^*\| (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}))^{\frac{1}{2}}. \end{aligned} \quad (3.24)$$

Applying (3.24) to (3.18) leads to

$$\begin{aligned} D(\mathbf{p}^{n+1}) - D(\mathbf{p}^*) &\leq \left(1 - \alpha + \alpha\mu + \frac{C_2}{\mu}\right) (D(\mathbf{p}^n) - D(\mathbf{p}^*)) - \frac{C_2}{\mu} (D(\mathbf{p}^{n+1}) - D(\mathbf{p}^*)) \\ &\quad + C_1 \sqrt{\alpha} \|\mathbf{p}^n - \mathbf{p}^*\| (D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}))^{\frac{1}{2}}. \end{aligned}$$

Since $\mathbf{p}^n, \mathbf{p}^* \in K$ such that $\|\mathbf{p}^n - \mathbf{p}^*\| \leq 2|\Omega|^{\frac{1}{2}}$, we derive (3.15) from the above estimates. \square

Lemma 3.4. *If a positive monotone decreasing $\{a^n\}_{n=0}^\infty$ satisfies the following inequality*

$$a^{n+1} \leq \gamma a^n + C_3 (a^n - a^{n+1})^{\frac{1}{2}} \quad (3.25)$$

with $0 \leq \gamma < 1$ and $0 \leq C_3 < \infty$, then

$$a^n \leq \frac{a^0}{\tilde{C} a^{0n} + 1}, \quad (3.26)$$

where \tilde{C} is given by

$$\tilde{C} = \frac{(1 - \gamma)^2}{2a^0(1 - \gamma)^2 + (\gamma\sqrt{a^0} + C_3)^2}.$$

Proof. We can also easily prove this lemma following [34]. It follows from (3.25) that

$$\begin{aligned} (1 - \gamma)a^{n+1} &\leq \gamma(a^n - a^{n+1}) + C_3(a^n - a^{n+1})^{\frac{1}{2}} \\ &= (\gamma(a^n - a^{n+1})^{\frac{1}{2}} + C_3)(a^n - a^{n+1})^{\frac{1}{2}} \leq (\gamma\sqrt{a^0} + C_3)(a^n - a^{n+1})^{\frac{1}{2}}. \end{aligned}$$

Thus we obtain

$$(a^{n+1})^2 \leq \left(\frac{\gamma\sqrt{a^0} + C_3}{1 - \gamma}\right)^2 (a^n - a^{n+1}).$$

By Lemma 3.2 [34, Page 110] and proofs [34, Page 113], one readily obtains (3.26) \square

With the above preparations, we are now ready to prove the main results.

3.3. Proof of Theorem 3.1 for Algorithm I. We split the proof into two steps.

Step I. Let $\zeta^n = D(\mathbf{p}^n) - D(\mathbf{p}^*)$. We prove $\zeta^n \rightarrow 0$ as $n \rightarrow \infty$. By Lemma 3.2 and (3.22), one can readily obtain the properties of the sequence as:

$$D(\mathbf{p}^{n+1}) \leq D(\mathbf{p}^n), \quad 0 \leq \zeta^{n+1} \leq \zeta^n. \quad (3.27)$$

Thus by Lemma 3.3, we have

$$D(\mathbf{p}^{n+1}) - D(\mathbf{p}^*) \leq \gamma(D(\mathbf{p}^n) - D(\mathbf{p}^*)) + C_3(D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}))^{\frac{1}{2}}, \quad (3.28)$$

that is

$$\zeta^{n+1} \leq \gamma\zeta^n + C_3(\zeta^n - \zeta^{n+1})^{\frac{1}{2}} \quad (3.29)$$

with $0 \leq \gamma < 1$. The sequence $\{\zeta^n\}_{n=1}^{\infty}$ is convergent due to the decreasing property (3.27). Thus, by taking limit in (3.29), one readily obtains that $\zeta^n \rightarrow 0$ as $n \rightarrow \infty$.

Step II. It follows from (3.29) and Lemma 3.4 that

$$\zeta^n \leq \frac{\zeta^0}{(C^*)^{-1}n + 1} \leq \frac{C^*}{n}, \quad (3.30)$$

with $C^* = \zeta^0 \frac{2\zeta^0(1-\gamma)^2 + (\gamma\sqrt{\zeta^0} + C_3)^2}{\zeta^0(1-\gamma)^2}$. Then by (3.22) and (3.2), (3.3) is derived.

But the constant C^* depends on μ . We need to optimal value of the factor C^* by choosing suitable μ . That is to say, we should choose optimal μ to minimize C^* where μ is a relaxed factor in Lemma 3.3.

$$\begin{aligned} \mu^* &= \arg \min_{0 < \mu < 1} C^* \\ &= \arg \max_{\mu} \frac{\zeta^0}{2\zeta^0 + \left(\frac{\gamma\sqrt{\zeta^0} + C_3}{1-\gamma}\right)^2} = \arg \min_{\mu} \frac{\gamma\sqrt{\zeta^0} + C_3}{1-\gamma} \\ &= \arg \min_{\mu} \frac{\left(1 - \frac{\alpha(1-\mu)\mu}{\mu + C_2}\right)\sqrt{\zeta^0} + \frac{2C_1\sqrt{\alpha}\mu}{\mu + C_2}|\Omega|^{\frac{1}{2}}}{\frac{\alpha(1-\mu)\mu}{\mu + C_2}} \\ &= \arg \min_{\mu} \frac{\alpha\mu + \frac{C_2}{\mu} + (1-\alpha + 2\sqrt{\alpha}C_1|\Omega|^{\frac{1}{2}})(\zeta^0)^{-\frac{1}{2}}}{1-\mu} \\ &= \arg \min_{\mu} \frac{\mu + C_4}{\mu(1-\mu)} = \sqrt{(C_4)^2 + C_4} - C_4 \end{aligned} \quad (3.31)$$

with $C_4 = \frac{C_2}{1 + 2\sqrt{\alpha}C_1|\Omega|^{\frac{1}{2}}(\zeta^0)^{-\frac{1}{2}}}$. Thus, we have

$$\begin{aligned} \frac{C^*}{\zeta^0} &= 2 + ((\alpha^{-1} + 2\frac{C_1}{\sqrt{\alpha}}|\Omega|^{\frac{1}{2}}(\zeta^0)^{-\frac{1}{2}})(\sqrt{C_4} + \sqrt{C_4 + 1})^2 - 1)^2 \\ &= 2 + \alpha^{-2} \left(\left(\sqrt{C_2} + \sqrt{1 + 2\sqrt{\alpha}C_1|\Omega|^{\frac{1}{2}}(\zeta^0)^{-\frac{1}{2}} + C_2} \right)^2 - \alpha \right)^2 \\ &\leq 2 + \left(\frac{2}{\alpha} \left(1 + 2C_2 + 2\sqrt{\alpha}C_1|\Omega|^{\frac{1}{2}}(\zeta^0)^{-\frac{1}{2}} \right) - 1 \right)^2 \\ &= 2 + \left(\frac{2}{\alpha} (1 + 2C_2) + \frac{4C_1}{\sqrt{\alpha}}|\Omega|^{\frac{1}{2}}(\zeta^0)^{-\frac{1}{2}} - 1 \right)^2 \\ &\leq \left(\frac{2}{\alpha} (1 + 2C_2) + \frac{4C_1}{\sqrt{\alpha}}|\Omega|^{\frac{1}{2}}(\zeta^0)^{-\frac{1}{2}} + \sqrt{2} - 1 \right)^2. \end{aligned}$$

The proof of Theorem 3.1 for **Algorithm I** is completed.

3.4. Proof of Theorem 3.1 for Algorithm II. The convergence analysis of **Algorithm II** is similar to that of **Algorithm I** (see [30]). The difference lies in the proof of (3.21) in Lemma 3.3. However, the second term in the 3rd row of (3.21) can be enlarged to be the same as the term for PSC. Therefore, the deduced convergence rate is the same as **Algorithm I**, and we just give the sketch of the proof.

Proof. Similarly to [30], define

$$\mathbf{q}_{\frac{i}{M_c}}^n := \sum_{j \leq i} \mathbf{q}_j^n + \sum_{j > i} \theta_j \mathbf{p}^n, \quad \hat{\mathbf{q}}_{\frac{i}{M_c}}^n := \sum_{j \leq i} \mathbf{q}_j^n + \hat{\mathbf{q}}_j^n + \sum_{j > i} \theta_j \mathbf{p}^n$$

One can readily obtain that

$$\langle D'(\hat{\mathbf{q}}_{\frac{i}{M_c}}^n), \bar{\mathbf{q}} - \hat{\mathbf{q}}_i^n \rangle \geq 0, \quad \forall \bar{\mathbf{q}} \in K_i,$$

as $\hat{\mathbf{q}}_{\frac{i}{M_c}}^n$ is the minimizer of the subproblem (2.16). Therefore, we have

$$\begin{aligned} D(\mathbf{p}^n) - D(\mathbf{p}^{n+1}) &= \sum_{i=1}^{M_c} \left(D(\mathbf{q}_{\frac{i-1}{M_c}}^n) - D(\mathbf{q}_{\frac{i}{M_c}}^n) \right) \\ &\geq \alpha \sum_{i=1}^{M_c} \left(D(\mathbf{q}_{\frac{i-1}{M_c}}^n) - D(\hat{\mathbf{q}}_{\frac{i}{M_c}}^n) \right) \geq \alpha \sum_{i=1}^{M_c} |\mathbf{e}_i^n|_{1,*}^2 \end{aligned}$$

By introducing the functions ϕ_j^i

$$\phi_j^i = \begin{cases} \mathbf{p}^n + \sum_{k=1}^{i-j} \mathbf{e}_k^n + \sum_{k=i-j+1}^i \mathbf{e}_k^n, & j \leq i, \\ \mathbf{p}^n + \sum_{k=1}^j \mathbf{e}_k^n, & j > i, \end{cases}$$

We can obtain the same result as Lemma 3.3. Finally by Lemma 3.4, we can finish the proof. \square

4. NUMERICAL EXPERIMENTS

In this section, we first give an algorithm to solve the sub-problems and then perform several numerical tests to verify the theoretical results given in Subsection 3.

4.1. Algorithm for sub-problems of dual model. To solve (2.15) and (2.16) we can apply the gradient projection methods similarly to [5]. One can choose other solvers for the subproblem. We just realize this algorithm for examples. We will display how to solve (2.15), while to solve (2.16) is similar. Assume \mathbf{p} is the given initial value for arbitrary iterations. Denoting $\mathbf{q}_i^0 := \sum_{j \neq i} \theta_j \mathbf{p}$ and $g_i = \frac{g}{\lambda} - \operatorname{div} \mathbf{q}_i^0$. From Karush-Kuhn-Tucker conditions, there exists Lagrangian multiplier $\mu_i \geq 0$ satisfying the following equation:

$$D'(\hat{\mathbf{q}}_i + \mathbf{q}_i^0) + 2\mu_i \hat{\mathbf{q}}_i = 0. \quad (4.1)$$

associated with either $\mu_i > 0$ as $|\hat{\mathbf{q}}_i| = \theta_i$, or $\mu_i = 0$ as $|\hat{\mathbf{q}}_i| < \theta_i$. Thus

$$\theta_i \left(-\lambda^2 \nabla(\operatorname{div} \hat{\mathbf{q}}_i - g_i) \right) + \left| -\lambda^2 \nabla(\operatorname{div} \hat{\mathbf{q}}_i - g_i) \right| \hat{\mathbf{q}}_i = 0. \quad (4.2)$$

The iterative scheme is constructed as follows:

$$\hat{\mathbf{q}}_i^{n+1} = \frac{\theta_i \hat{\mathbf{q}}_i^n + \theta_i \tau (\nabla(\operatorname{div} \hat{\mathbf{q}}_i^n - g_i))}{\theta_i + \tau |\nabla(\operatorname{div} \hat{\mathbf{q}}_i^n - g_i)|} \quad (4.3)$$

with suitable $\tau > 0$.

4.2. Numerical examples. First we introduce difference schemes for **Algorithm I** and **Algorithm II**. We use the classical difference scheme for PDE-based image processing problem. For simplicity, set $\Omega = [0, 1] \times [0, 1]$. Divide the domain as

$$\Omega_h := \{(x_i, y_i) \mid x_i = ih_x, y_j = jh_y, 0 \leq i \leq m, 0 \leq j \leq n\}$$

with the grid size of h_x and h_y . The mesh size is defined by $h = \min\{h_x, h_y\}$. Then define the gradient and divergence of each $u_{i,j}$ as

$$(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2),$$

where

$$(\nabla u)_{i,j}^1 = \begin{cases} \frac{1}{h_x} (u_{i+1,j} - u_{i,j}), & i < m, \\ 0, & i = m, \end{cases} \quad (\nabla u)_{i,j}^2 = \begin{cases} \frac{1}{h_y} (u_{i,j+1} - u_{i,j}), & j < n, \\ 0, & j = n. \end{cases}$$

The divergence of $\mathbf{p} = (p^1, p^2) \in \mathcal{R}^2$ satisfying $\operatorname{div} = -\nabla^*$ in discrete form is defined as

$$(\operatorname{div} \mathbf{p})_{i,j} = \begin{cases} \frac{1}{h_x} (p_{i,j}^1 - p_{i-1,j}^1) & (\text{as } 0 < i < m) \\ \frac{1}{h_x} p_{i,j}^1 & (\text{as } i = 0) \\ -\frac{1}{h_x} p_{i-1,j}^1 & (\text{as } i = m) \end{cases} + \begin{cases} \frac{1}{h_y} (p_{i,j}^2 - p_{i,j-1}^2) & (\text{as } 0 < j < n) \\ \frac{1}{h_y} p_{i,j}^2 & (\text{as } j = 0) \\ -\frac{1}{h_y} p_{i,j-1}^2 & (\text{as } j = n) \end{cases}$$

In the following tests, we set $h_x = h_y = 1$. Then the domain Ω_h is decomposed into small subdomains Ω_h^i . For simple realization we use the color techniques and the subdomains are classified into four ‘‘colors’’ that are shown in Figure 2.1. The unit decomposition functions are shown in Figure 4.1.

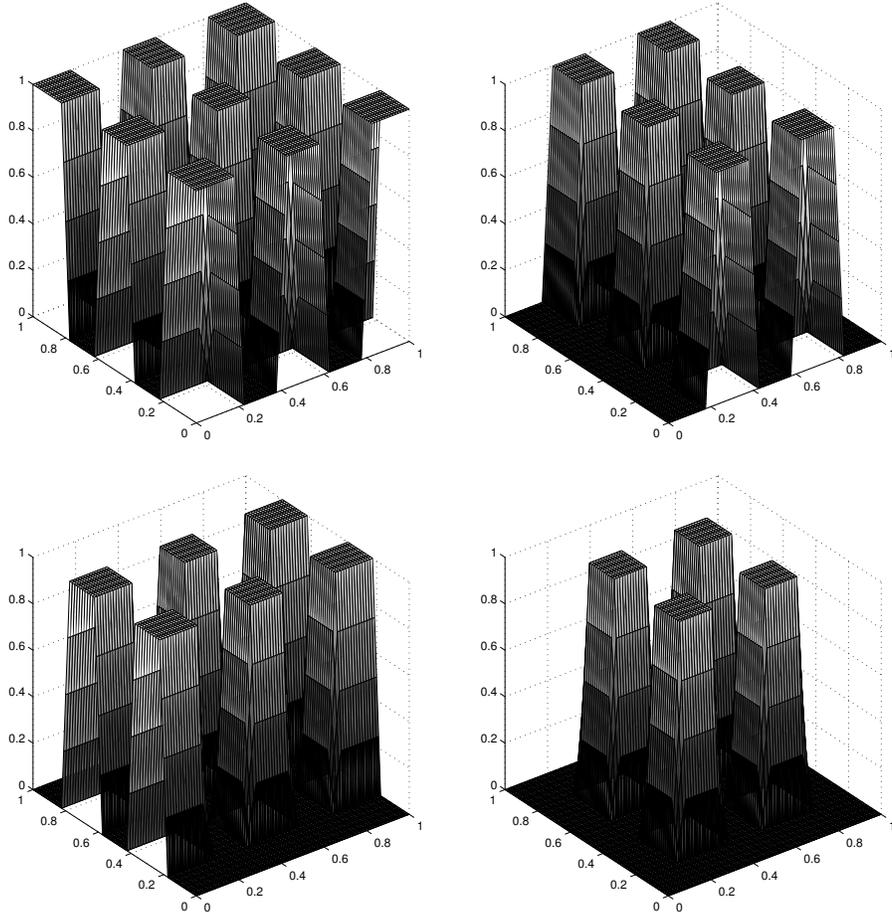


FIGURE 4.1. Unit decomposition functions $\{\theta_i\}_{i=1}^4$: the subdomains have four colors, overlapping size $\delta=2h$.

Remark 4.1. In the real computation that is the discrete forms, we will employ a modified unit partition as in [36]. Assume there exists a mesh on the domain Ω . Then interpolate the partition function θ_j on the mesh as follows:

$$\tilde{\theta}_j := I^h(\theta_j),$$

where I^h is the piecewise linear operator interpolating in the nodes on the mesh. One readily knows that the modified functions $\{\tilde{\theta}_j\}$ satisfy the above properties in (2.8), (2.9) and (2.10). For 1-dimension case, assume that $\Omega = [0, 1]$. The domain is decomposed into two overlapping subdomains $\Omega_1 = [0, a + \delta]$, and $\Omega_2 = [a - \delta, 1]$ with $0 < a < \frac{1}{2}$, δ is overlapping size. Then

$$\tilde{\theta}_1 = \begin{cases} 1, & x \in [0, a - \delta], \\ 1 - \frac{x-a+\delta}{2\delta}, & x \in [a - \delta, a + \delta], \\ 0, & x \in [a + \delta, 1], \end{cases} \quad \text{and} \quad \tilde{\theta}_2 = \begin{cases} 1, & x \in [a + \delta, 1], \\ \frac{x-a+\delta}{2\delta}, & x \in [a - \delta, a + \delta], \\ 0, & x \in [0, a - \delta]. \end{cases}$$

One readily verifies that $\tilde{\theta}_j$ satisfies (2.8) and (2.9). we need to require $\nabla\tilde{\theta}_j \in L^\infty(\Omega)$ instead of $\tilde{\theta}_j \in C^1(\Omega)$. It is easily known that $\tilde{\theta}_j$ satisfies that

$$|\tilde{\theta}_j(x) - \tilde{\theta}_j(y)| \leq \frac{C_0}{\delta} |x - y|$$

for $y \in \Omega_j$, sufficiently close to x . Thus $\tilde{\theta}_j$ satisfies (2.10) (see [36, Page57]). See Figure 4.1 for 2-dimension case for example as well. These partition functions are the tensors of that in 1-dimension.

Next we give the numerical examples by **Algorithm II** by using coloring technique. Two images in Figure 4.2 are tested. Define energy

$$E(\mathbf{p}) := \sum_{i=1}^m \sum_{j=1}^n (\lambda \operatorname{div} \mathbf{p} - g)_{i,j}^2, \quad (4.4)$$

where the image resolution is $m \times n$, and \mathbf{p}_{DDM} is computed by DDM. The error is defined as

$$e := \frac{\left(\sum_{i=1}^m \sum_{j=1}^n (\lambda \operatorname{div}(\mathbf{p}_{DDM} - \mathbf{p}^*))_{i,j}^2 \right)^{\frac{1}{2}}}{\left(\sum_{i=1}^m \sum_{j=1}^n (g - \lambda \operatorname{div} \mathbf{p}^*)_{i,j}^2 \right)^{\frac{1}{2}}}, \quad (4.5)$$

with \mathbf{p}^* is approximated by dual methods after 10^7 iterations without DDM.

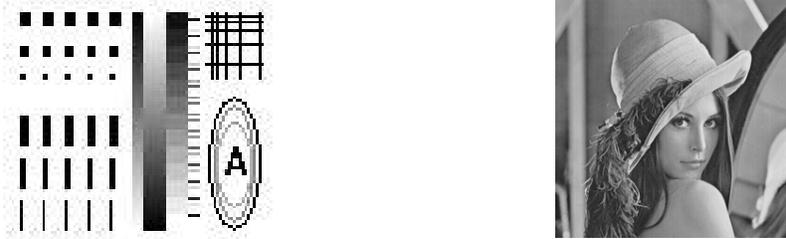


FIGURE 4.2. Left: resolution 128×128 . Right: resolution 256×256

Some notations: *subsize* is the size of subdomain. The following tests are qualified by the the decreasing of the energy defined by $Energy := E(\mathbf{p}_{DDM}) - E(\mathbf{p}^*)$ and $Error := e$.

4.2.1. *Test 0: Verifying Convergence.* We first show the performance of the proposed DDMs in Figure 4.3. The denoising images are present in Fig. 4.3. The restored images, and the differences between the solutions by proposed DDMs and the exact minimizer are shown in Figure 4.4 within 4 iterations, and the convergence curve is shown in Figure 4.5. The proposed DDMs work as good as the gradient projection method in Figure 4.3. Inferred from Figure 4.4, denoising results are satisfactory within 4 iterations of the proposed DDMs. Indeed, the proposed DDMs are convergent by observing Figure 4.5.

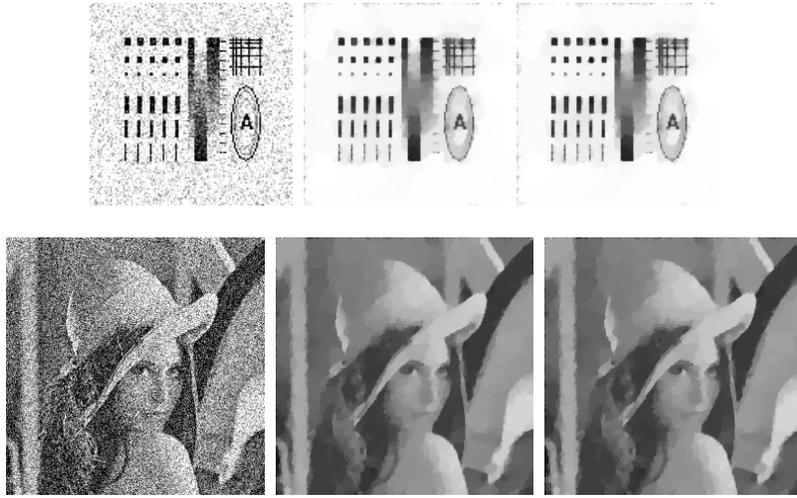


FIGURE 4.3. $\alpha = 1, \tau = \frac{1}{4}, \text{subsize} = 64, \delta = 4, \sigma = 50, \lambda = 60, N_{in} = 1000$. From left to right: Noised Image, Denoised Image without DDM and Denoised by DDM; from up to down: 128×128 , and 256×256

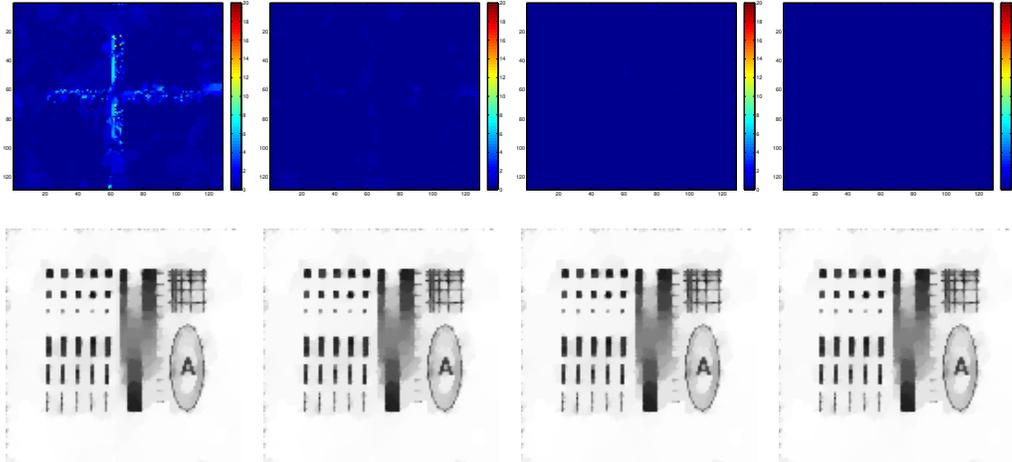


FIGURE 4.4. First row: difference between the DDMs ($\delta=4, \text{subsize}=64, \sigma = 50, \alpha = 1, \text{and } \lambda = 60$) and the exact minimizer within the first 4 iterations from the left to right; Second row: restored images by proposed DDMs within first 4 iterations

4.2.2. *Test 1: Verifying convergence rates.* Convergence rates are tested in Fig. 4.6. From the tests we see that the convergence rate of energy decay is between $O(N^{-1})$ and $O(N^{-2})$, and that of the dual variables error is a bit larger than $O(N^{-1})$, where N is the iterative number. But we can just prove the rate of the convergence rate of average energy decay is about $O(N^{-1})$, and that of the dual variables average error is $O(N^{-\frac{1}{2}})$. We just give a

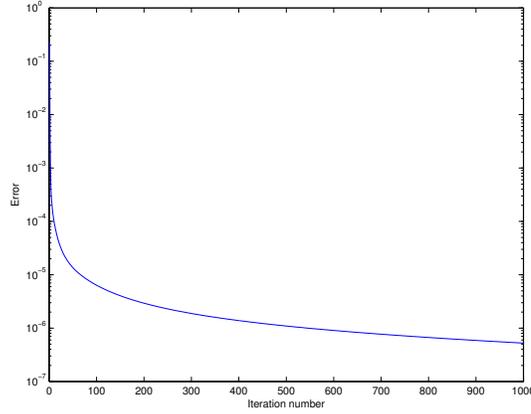


FIGURE 4.5. Convergence performance of the proposed DDMs after 1000 iterations. The error is computed as $\frac{\|\text{div}(\mathbf{p}^k - \mathbf{p}^*)\|}{\|\text{div}\mathbf{p}^*\|}$, where \mathbf{p}^k is the iterated solution by proposed DDMs, and \mathbf{p}^* is computed by gradient projection method after 10^7 iterations.

general bound estimate for the DDMs, and cannot give proof of the optimal order. Most importantly, the convergence rate behavior is of pow function as $O(N^{-\sigma})$, with $\sigma > 0$. The following tests are based on the image with resolution 128×128 .

4.2.3. *Test 2: Convergence rate V.S. overlapping size δ and number of blocks.*

- 1) Fixing *subsize* = 64, we test how the convergence rate relies on the overlapping size. $\delta = 2, 4, 8$, and 16. See Figure 4.7. From the tests the convergence become fast when overlapping size becomes large. But the rates vary not so much that implies the convergence rate is not sensitive to the overlapping size δ which just affects the coefficient of the convergence order. From the results we guess the coarse grid correction may do not help to improve the convergence order which is also implied in [44], but that can improve the coefficient.
- 2) Fixing $\delta = 4$, we test how the convergence rate relies on the number of subdomain. *subsize* = 8, 16, 32 and 64. See Figure 4.8. The convergence becomes fast when the size of subdomain becomes small. We use color techniques that fixes $M_c = 4$. Then the smaller size of the subdomain, the relative larger is the overlapping size.
- 3) Fixing *subsize* = 128 when the testing image is Lena(256×256), we test how the convergence rate relies on the overlapping size. $\delta = 2, 4, 8, 16$, and 24. See Figure 4.9. The convergence varies more obviously as the δ varies than that of Figure 4.7.

4.2.4. *Test 3: Convergence rate V.S. relaxation parameter α .* We test four different values. $\alpha = 1/8, 1/4, 1/2$ and 1. See Figure 4.10. The convergence is fast as α is close to 1, that is consistent with (3.4) in Theorem 3.1.

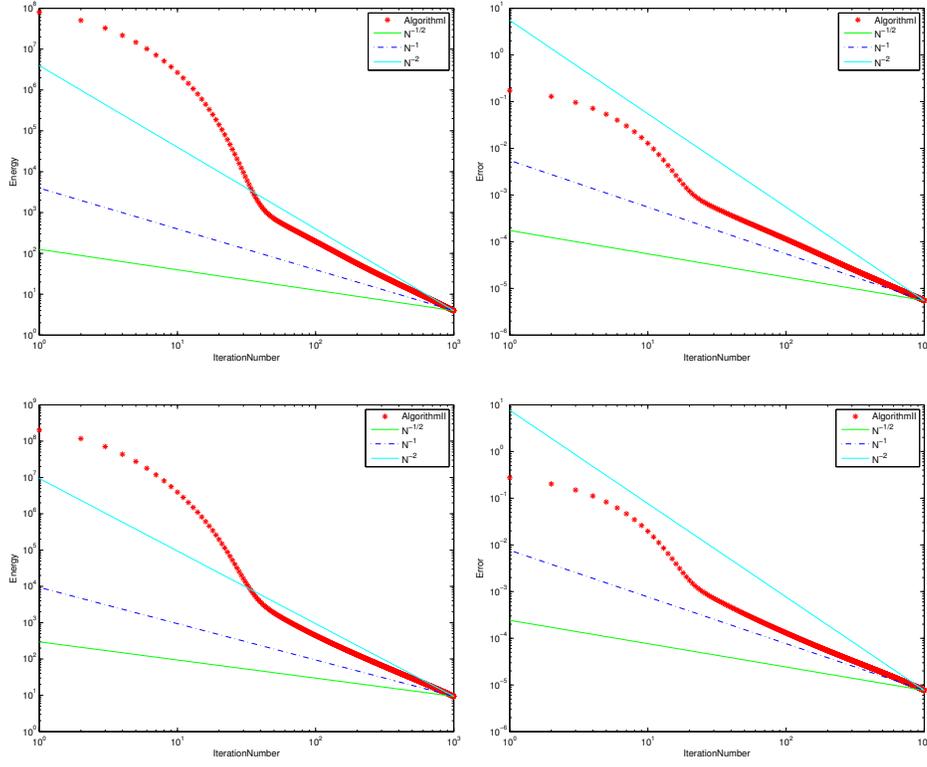


FIGURE 4.6. $\alpha = \frac{1}{4}, \tau = \frac{1}{4}, \text{subsize} = 64, \delta = 4, \sigma = 50, \lambda = 60, N_{in} = 500$. From left to right: Energy, and Error; from up to down: 128×128 , and 256×256

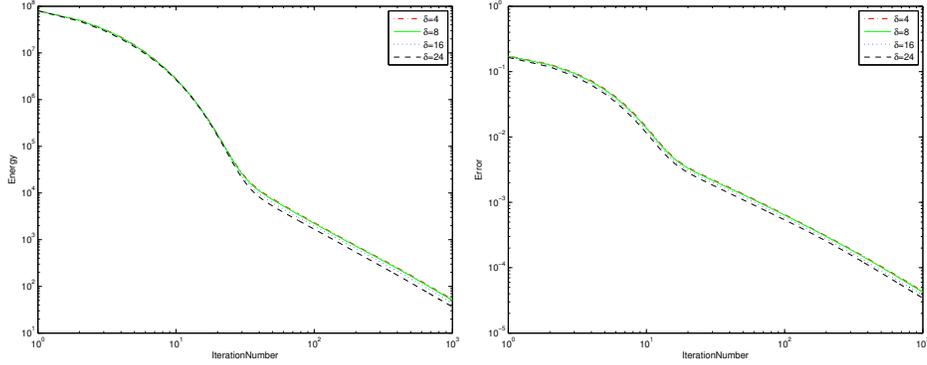


FIGURE 4.7. $\alpha = \frac{1}{4}, \tau = \frac{1}{4}, \text{subsize} = 64, \sigma = 50, \lambda = 60, N_{in} = 500$. From left to right: Energy, and Error.

4.2.5. *Test 4: comparing with the gradient projection algorithm* [5]. At last we show the convergence behavior comparing with the gradient projection (GP) algorithm by Chambolle [5]. One can readily infer that our proposed DDMS convergent much faster.

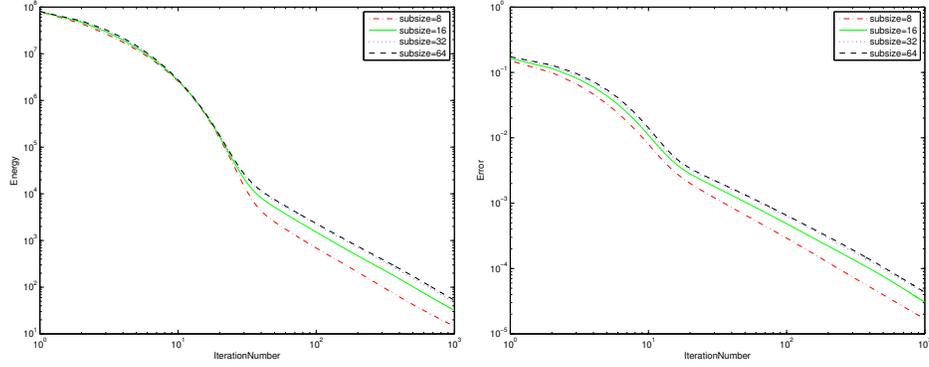


FIGURE 4.8. $\alpha = \frac{1}{4}, \tau = \frac{1}{4}, \sigma = 50, \lambda = 60, N_{in} = 500$. From left to right: Energy, and Error.

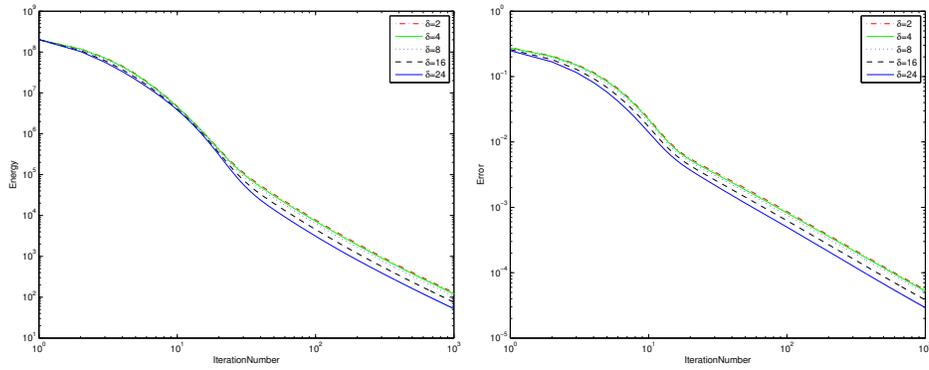


FIGURE 4.9. $\alpha = \frac{1}{4}, \tau = \frac{1}{4}, subsize = 128, \sigma = 50, \lambda = 60, N_{in} = 500$. From left to right: Energy, and Error.

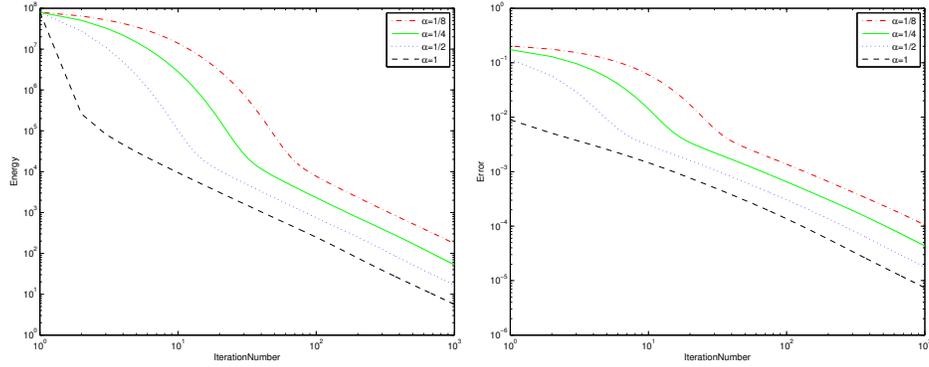


FIGURE 4.10. $\tau = \frac{1}{4}, \sigma = 50, subsize = 64, \delta = 4, \lambda = 60, N_{in} = 500$. From left to right: Energy, and Error.

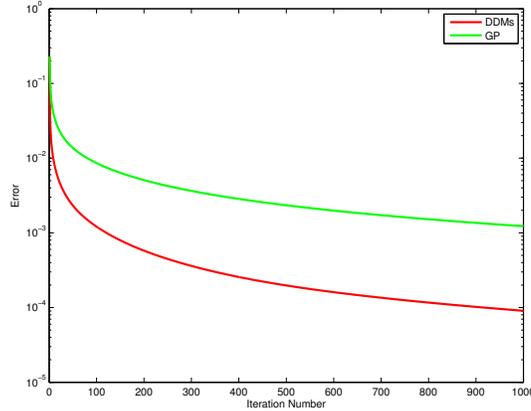


FIGURE 4.11. Error comparison for gradient projection method(GP) and the proposed DDMS($N_{in} = 10$)

Remark 4.2. From the tests we see that this algorithm converges fast in the beginning but stagnates after some iterations with the order $O(N^{-\sigma})$. That maybe deduce by (3.29). For the first iterations the errors are governed by

$$\zeta^{n+1} \leq \gamma \zeta^n,$$

that implies the convergence rate is linear. Then after many iterations the errors are governed by

$$\zeta^{n+1} \leq C_3 \left(\zeta^n - \zeta^{n+1} \right)^{\frac{1}{2}},$$

which implies the convergence rate is $O(N^{-1})$.

Finally we test the performance of the proposed DDMS with respect to the regularized parameter λ . The energy decay is shown in Figure 4.12. Our proposed DDMS are sensitive to the parameter λ , that is consistent with the estimate of the convergence rate.

5. CONCLUSION

We propose the domain decomposition methods for the dual model of ROF. The convergence rates are deduced as well. Big value of λ is needed if one wants to get the “scales” images from the TV model. Then the coarse mesh correction shall be needed in order to increase the robustness of the proposed DDMS. Thus for the future work, the results should be extended to the case with a coarse mesh. The dual models with more applications in the image processing ares should be considered as well.

REFERENCES

- [1] R. ACAR AND C. R. VOGEL, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Probl., 10(6):1217-1230, 1994.
- [2] B. APPLETON AND H. TALBOT, *Globally optimal geodesic active contours*, J. Math. Imaging Vision, 23(1):67-86, 2005
- [3] K. J. ARROW, L. HURWICZ, AND H. UZAWA, *Studies in linear and non-linear programming*, With contributions by H. B. Chenery, S. M. Johnson, S. Karlin, T. Marschak, R. M. Solow. Stanford Mathematical Studies in the Social Sciences, vol. II. Stanford University Press, Stanford, Calif., 1958.

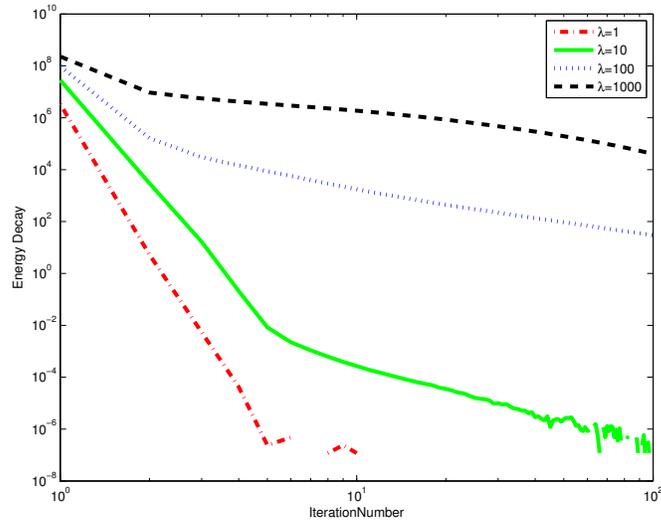


FIGURE 4.12. Energy decay with respect to the regularized parameter λ

- [4] A. CHAMBOLLE AND T. POCK, *A First-order Primal-dual Algorithm for Convex Problems with Applications to Imaging*, Ecole Polytechnique/T.U. Graz, Report, 2010.
- [5] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, Math. Imaging Vis., 20:89-97, 2004.
- [6] T. F. . CHAN AND T. P. MATHEW, *Domain decomposition algorithms*, Acta numerica 3, 61-143, 1994.
- [7] H. CHANG, X. ZHANG, X.C. TAI, AND D. YANG, *Domain Decomposition Methods for Nonlocal Total Variation Image Restoration*, J. Sci. Comput., DOI 10.1007/s10915-013-9786-9, online published Oct 2013.
- [8] Y. DONG, M. HINTERMÜLLER, AND M. NERI, *An Efficient Primal-Dual Method for L^1 -TV Image Restoration*, SIAM J. Imaging Sci. 2(4):1168-1189, 2009
- [9] Y. DUAN AND X. C. TAI, *Domain decomposition methods with Graph cuts algorithms for total variation minimization*, Adv. Comput. Math., 36(2):175-199, 2012
- [10] M. DRYJA AND O. B. WIDLUND, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, Houston, Texas, T. Chan et. al., eds, 1989.
- [11] E. ESSER, X. ZHANG AND T. CHAN, *A General Framework for a Class of First Order Primal-Dual Algorithms for TV Minimization*, UCLA, Center for Applied Math., CAM Reports no. 09-67, 2009.
- [12] M. FORNASIER, A. LANGER, AND C. B. SCHÖNLIEB, *Domain decomposition methods for compressed sensing*, In print, 2009.
- [13] M. FORNASIER, A. LANGER, AND C. B. SCHÖNLIEB, *A convergent overlapping domain decomposition method for total variation minimization*, Numer. Math., 116(4):645-685, 2010.
- [14] M. FORNASIER, AND C. B. SCHÖNLIEB, *subspace correction methods for total variation and l_1 -minimization*, SIAM J. NUMER. ANAL., 47(5):3397-3428, 2009
- [15] D. FIRSOV, S. H. LUI, *Domain decomposition methods in image denoising using Gaussian curvature*, J. Comput. Appl. Math., 193:460-473, 2006
- [16] T. GOLDSTEIN AND S. OSHER, *The split Bregman method for L^1 -regularized problems*, SIAM J. Imaging Sci., 2(2):323-343, 2009.
- [17] M. GRIEBEL AND P. OSWALD, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math., 70(2):163-180, 1995.
- [18] R. GLOWINSKI AND P. LE TALLEC, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia, 1989.
- [19] M. HINTERMÜLLER, AND K. KUNISCH, *Total bounded variation regularization as a bilaterally constrained optimization problem*, SIAM. J. Appl. Math., 64(4):1311-1333, 2004

- [20] M. HINTERMÜLLER, AND A. LANGER, *Subspace Correction Methods for a Class of Nonsmooth and Non-additive Convex Variational Problems with Mixed L^1/L^2 Data-Fidelity in Image Processing*, SIAM Journal on Imaging Sciences, 6(4):2134-2173, 2013.
- [21] A. LANGER, S. OSHER AND C. SCHONLIEB, *Bregmanized domain decomposition for image restoration*, Journal of Scientific Computing 54(2-3):549-576, 2013.
- [22] P. L. LIONS, *On Schwarz alternating method I*, Proceedings of the second international symposium on domain decomposition methods for partial differential equations, SIAM. Philadelphia, Chan, T., Glowinski, R., Périaux, J. and Widlund, O. B., (eds), 401-430, 1989.
- [23] P. L. LIONS, *On the Schwarz alternating method II*, SIAM, Philadelphia, PA, 47-70, 1989.
- [24] P. L. LIONS, *On the Schwarz alternating method III*. SIAM, Philadelphia, PA, 202-223, 1990.
- [25] A. MARQUINA AND S. OSHER. EXPLICIT ALGORITHMS FOR A NEW TIME DEPENDENT MODEL BASED ON LEVEL SET MOTION FOR NONLINEAR DEBLURRING AND NOISE REMOVAL, SIAM J. Sci. Comput., 22(4):387-405, 2000.
- [26] P. MONK, *Finite Element Methods for Maxwell's Equations*, Oxford University Press, New York, 2003.
- [27] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60(14):259-268, 1992.
- [28] B. F. SMITH, P. E. BJØRSTAD AND W. D. GROPP, *Domain decomposition, Parallel multilevel methods for elliptic partial differential equations*, Cambridge University Press, Cambridge, 1996.
- [29] O. SCHERZER, EDITOR, *Handbook of Mathematical Methods in Imaging*, Springer, New York, 2011
- [30] X. C. TAI, *Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities*, Numer. Math., 93(4), Springer, 2003.
- [31] X. C. TAI AND M. ESPEDAL, *Applications of a space decomposition method to linear and nonlinear elliptic problems*, Numer. Meth. Part. D. E., 14(6):717-737, 1998.
- [32] X. C. TAI AND M. ESPEDAL, *Rate of convergence of some space decomposition methods for linear and nonlinear problems*, SIAM J. Numer. Anal., 35(14):1558-1570, 1998.
- [33] X. C. TAI AND P. TSENG, *Convergence rate analysis of an asynchronous space decomposition method for convex minimization*, Math. Comput., 71(239):1105-1136, 2002.
- [34] X. C. TAI AND J. XU, *Global and uniform convergence of subspace correction methods for some convex optimization problems*, Math. Comput., 71(237):105-124, 2002.
- [35] X. C. TAI, AND Y. P. DUAN, *domain decomposition methods with graph cuts algorithms for image segmentation*, Int. J. Numer. Anal. Model., 8(1):137-155, 2011.
- [36] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods-Algorithm and Theory*, Springer-Verlag Berlin Heidelberg, 2005
- [37] C. R. VOGEL, *A multigrid method for total variation-based image denoising*, Computation and Control IV, Progress in Systems and Control Theory, 20:323-323, Birkhäuser Boston,1995.
- [38] C. R. VOGEL, *Computational methods for inverse problems*, Society for Industrial Mathematics, 2002.
- [39] C. R. VOGEL AND M. E. OMAN, *Iterative methods for total variation denoising*, SIAM J. Sci. Comput., 17(1):227-238, 1996.
- [40] C. R. VOGEL, M. E. OMAN, ET AL, *Fast, robust total variation-based reconstruction of noisy, blurred images*, IEEE Trans. Image Process., 7(6):813-824, 1998.
- [41] Y. WANG, J. YANG, W. YIN AND Y. ZHANG, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM J. Imaging Sci., 1(3):248-272, 2008.
- [42] Y. WANG, W. YIN AND Y. ZHANG, *A fast algorithm for image deblurring with total variation regularization*, Rice University CAAM Technical Report TR07-10, 2007.
- [43] C. L. WU AND X. C. TAI, *Augmented lagrangian method, dual methods and split-bregman iterations for rof, vectorial TV and higher order models*, SIAM J. Imaging Sci., 3(3):300-339, 2010.
- [44] J. XU, X. C. TAI AND L. L. WANG. *A two-level domain decomposition method for image restoration*, Inverse Probl. Image, 4(3):523-545, 2010.
- [45] J. XU, AND H. CHANG *Domain decomposition method for image deblurring*, UCLA, Center for Applied Math., CAM Reports no. 13-63.
- [46] J. C. XU, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34(4):581-613, 1992.
- [47] M. ZHU AND T. CHAN, *An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration*, UCLA, Center for Applied Math., CAM Reports No. 08-34, 2008.