METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 20, No. 4, pp. 425–438, December 2013

007

SIMULTANEOUS DATA RECOVERY IN IMAGE AND TRANSFORM DOMAINS*

SAY SONG GOH†, ZUOWEI SHEN†, AND JUNQI ZHOU†

Abstract. This paper addresses the problem of image recovery from partially given data in both the image and tight frame transform domains. Motivated by an uncertainty principle, a sufficient condition that ensures the exact recovery of an image is derived. Then an analysis-based model is proposed to handle situations in which exact recovery is impossible, such as when insufficient or only inaccurate data is available. An efficient iterative algorithm is obtained for the model by applying the split Bregman method. Several numerical examples are presented to demonstrate the potential of the algorithm.

Key words. Image restoration, uncertainty principle, analysis-based approach, split Bregman method.

AMS subject classifications. 49N45, 42C40.

1. Introduction. Image restoration is one of the most important areas in image processing and its wide-ranging applications include image inpainting, deblurring and denoising, and super-resolution image reconstruction. The tight frame transform gives rise to a recent and fast emerging approach for image restoration. Interested readers can refer to the survey article [24] and the more detailed lecture notes [12] for the development of tight frame based image restoration. Furthermore, this approach for image restoration, based on wavelet frames, is closely related to the methods based on partial differential equation. We refer readers to [5] for its connection to the well-established total variational approaches, and the references therein for various applications and aspects of image restoration. For such transform-based approach, the transform matrix maps an image from the image domain into the transform domain. We are interested in the problem of recovering the original image from partially available information in both the image and transform domains. A solution of this problem would have important ramifications in many practical applications.

For simplicity of notations, by concatenating the columns of an image, let $\mathbf{f} \in \mathbb{R}^N$ be the original image, $\mathbb{R}^N$ the image domain and $\mathbb{R}^M$ the transform domain, where the transform $\mathbf{W}$ is given by an $M \times N$ matrix. The transform considered in this paper is a *tight frame transform*, i.e.

$$\mathbf{W}^T\mathbf{W} = \mathbf{I}. \tag{1.1}$$

Recall that a vector system $\mathcal{W}$ in $\mathbb{R}^N$ is a *tight frame* if for every $\mathbf{f} \in \mathbb{R}^N$, there holds

$$\mathbf{f} = \sum_{\mathbf{w}\in\mathcal{W}} \langle \mathbf{f}, \mathbf{w}\rangle \mathbf{w}.$$

Then the identity $\mathbf{W}^T\mathbf{W} = \mathbf{I}$ is equivalent to (the transpose of) the rows of $\mathbf{W}$ forming a tight frame for $\mathbb{R}^N$.

In the image domain, suppose that the data available is $\mathbf{f} \in \mathbb{R}^N$ restricted to an index set $\Lambda \subset \mathcal{N} := \{1, \ldots, N\}$. This can be written as $\mathbf{P}_\Lambda \mathbf{f}$, where $\mathbf{P}_\Lambda$ is the $N \times N$

*Received January 5, 2013; accepted for publication July 11, 2013. S. S. Goh and Z. Shen were supported in part by MOE Academic Research Fund Tier 1 Grant R-146-000-165-112.

†Department of Mathematics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Republic of Singapore ({matgohss; matzuows; zhoujunqi}@nus.edu.sg).

projection matrix defined by

$$\mathbf{P}_\Lambda(i,j) = \begin{cases} 1, & \text{if } i = j \in \Lambda, \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, in the transform domain $\mathbb{R}^M$, suppose that the information available is $\mathbf{Wf}$ restricted to an index set $\Gamma \subset \mathcal{M} := \{1, \ldots, M\}$, i.e. $\mathbf{P}_\Gamma \mathbf{Wf}$, where the $M \times M$ projection matrix $\mathbf{P}_\Gamma$ is defined similarly.

We are interested to know when the original $\mathbf{f}$ can be recovered from the partially given data $\mathbf{P}_\Lambda \mathbf{f}$ and $\mathbf{P}_\Gamma \mathbf{Wf}$, where Λ and Γ are proper subsets of $\mathcal{N}$ and $\mathcal{M}$ respectively. We will show that if the transform $\mathbf{W}$ is a tight frame transform, and the index sets Λ and Γ satisfy $\sum_{i \notin \Gamma} \sum_{j \notin \Lambda} \mathbf{W}(i,j)^2 < 1$, where $\mathbf{W}(i,j)$ is the (i,j)-entry of $\mathbf{W}$, then the original data $\mathbf{f}$ can be reconstructed exactly by applying the iterative algorithm

$$\mathbf{f}_{k+1} = \mathbf{P}_\Lambda \mathbf{f} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{W}^T(\mathbf{P}_\Gamma \mathbf{Wf} + (\mathbf{I} - \mathbf{P}_\Gamma)\mathbf{Wf}_k). \tag{1.2}$$

This algorithm essentially performs interpolation in both the image and transform domains. Indeed, in (1.2), we first enforce the available data on Γ in the transform domain to the output of the previous iteration, then transform the result back to the image domain via $\mathbf{W}^T$, followed by enforcing the known information on Λ in the image domain to it.

A related, but different, problem was considered in [2]. The formulation of the problem is as follows. Let $\mathbf{x}$ be given data in $\mathbb{R}^N$ and $\mathbf{y}$ given data in $\mathbb{R}^M$. The objective is to find $\mathbf{f} \in \mathbb{R}^N$ such that

$$\begin{cases} \mathbf{P}_\Lambda \mathbf{f} = \mathbf{x}, \\ \mathbf{P}_\Gamma \mathbf{Wf} = \mathbf{y}. \end{cases} \tag{1.3}$$

This is a very ill-posed problem. First of all, it may not have a solution, which occurs, for example, when the data set $\mathbf{y}$ lies outside the range of $\mathbf{P}_\Gamma \mathbf{W}$. The range of $\mathbf{W}$ is the orthogonal complement of the kernel of $\mathbf{W}^T$ which is nonempty when the rows of $\mathbf{W}$ form a redundant system. Even when $\mathbf{y}$ does lie within the range of $\mathbf{P}_\Gamma \mathbf{W}$, the given data $\mathbf{x}$ could be incompatible with $\mathbf{y}$, in the sense that among all the $\mathbf{f}$'s satisfying $\mathbf{P}_\Gamma \mathbf{Wf} = \mathbf{y}$, there may not be any that gives $\mathbf{P}_\Lambda \mathbf{f} = \mathbf{x}$. This would again result in (1.3) having no solution. The problem can also have infinitely many solutions. For instance, when $\Lambda \subset \mathcal{N}$ and $\Gamma = \emptyset$, the values of $\mathbf{f}$ on the set $\mathcal{N} \backslash \Lambda$ can be arbitrarily chosen. Furthermore, in practical applications, since the data is always contaminated by noise, it may not be desirable to interpolate the data exactly even if it is possible to do so.

In view of all these difficulties with the problem (1.3), the following variational model was proposed in [2]. Let $\mathbf{t}^* \in \mathbb{R}^M$ be a minimizer of the minimization problem

$$\min_{\{\mathbf{t} \in \mathbb{R}^M : \mathbf{P}_\Gamma \mathbf{t} = \mathbf{T}_\mathbf{u} \mathbf{y}\}} \Big\{ \frac{1}{2} \|\mathbf{P}_\Lambda \mathbf{W}^T \mathbf{t} - \mathbf{x}\|_2^2 + \frac{1}{2} \|(\mathbf{I} - \mathbf{WW}^T)\mathbf{t}\|_2^2 + \|\mathrm{diag}(\mathbf{u})\mathbf{t}\|_1 \Big\}, \tag{1.4}$$

where $\mathbf{T}_\mathbf{u}$, $\mathbf{u} \in \mathbb{R}^M$, is the soft thresholding operator

$$\mathbf{T}_\mathbf{u}(\mathbf{y}) := \big(t_{\mathbf{u}(1)}(\mathbf{y}(1)), \cdots, t_{\mathbf{u}(i)}(\mathbf{y}(i)), \cdots, t_{\mathbf{u}(M)}(\mathbf{y}(M))\big)^T, \quad \mathbf{y} \in \mathbb{R}^M,$$

defined in [11] with

$$t_{\mathbf{u}(i)}(\mathbf{y}(i)) := \begin{cases} 0, & \text{if } |\mathbf{y}(i)| < \mathbf{u}(i), \\ \mathbf{y}(i) - \operatorname{sgn}(\mathbf{y}(i))\mathbf{u}(i), & \text{if } |\mathbf{y}(i)| \geq \mathbf{u}(i). \end{cases}$$

Then the solution is given by $\mathbf{f} = \mathbf{W}^T\mathbf{t}^*$. This model solves the problem in the transform domain. For arbitrarily given $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^M$, under the assumption that the underlying solution has a good sparse approximation in the transform domain, the model balances the approximation to the data's fidelity and the data's sparsity in the transform domain. An iterative algorithm on its realization was also proposed in [2], which is

$$\mathbf{f}_{k+1} = \mathbf{x} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{W}^T\mathbf{T}_\mathbf{u}(\mathbf{y} + (\mathbf{I} - \mathbf{P}_\Gamma)\mathbf{W}\mathbf{f}_k). \tag{1.5}$$

It is interesting to note that if $\mathbf{P}_\Lambda\mathbf{f}$ and $\mathbf{P}_\Gamma\mathbf{W}\mathbf{f}$ are known and we define $\mathbf{x}$ and $\mathbf{y}$ to be them, then (1.5) becomes

$$\mathbf{f}_{k+1} = \mathbf{P}_\Lambda\mathbf{f} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{W}^T\mathbf{T}_\mathbf{u}(\mathbf{P}_\Gamma\mathbf{W}\mathbf{f} + (\mathbf{I} - \mathbf{P}_\Gamma)\mathbf{W}\mathbf{f}_k).$$

This is very similar to the exact recovery algorithm (1.2). The only difference is the soft thresholding operator $\mathbf{T}_\mathbf{u}$ applied to the second term. This suggests that we may use (1.5) to handle our original problem of recovering missing data when exact recovery is impossible or when the data is contaminated by noise.

While the model (1.4) is efficient to implement for arbitrarily given data $\mathbf{x}$ and $\mathbf{y}$, it may not have fully exploited the possible close resemblance of $\mathbf{y}$ to $\mathbf{W}\mathbf{f}$. When $\mathbf{x}$ and $\mathbf{y}$ are closely related to $\mathbf{f}$ and $\mathbf{W}\mathbf{f}$, it would be appropriate to have a model whose approximation term in the transform domain is reflected by $\mathbf{W}\mathbf{f}$. In the frame literature, the entries of $\mathbf{W}\mathbf{f}$ are called *canonical coefficients* of the frame transform of $\mathbf{f}$. In many cases, the sparsity assumption is also imposed on the canonical coefficients, which should be reflected by the regularity term in the model. Taking into account all these considerations, we propose the following analysis-based model when exact recovery is impossible or unnecessary and when approximation and regularity of $\mathbf{W}\mathbf{f}$ is desirable. The solution we seek is a minimizer of the minimization problem

$$\min_{\mathbf{f}\in\mathbb{R}^N} \left\{ \frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f} - \mathbf{x}\|_2^2 + \frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{W}\mathbf{f} - \mathbf{y}\|_2^2 + \|\operatorname{diag}(\mathbf{u})\mathbf{W}\mathbf{f}\|_1 \right\}, \tag{1.6}$$

where $\mathbf{u} \in \mathbb{R}^M$ has positive entries which together with $\nu > 0$ form weighted parameters. The first two terms in (1.6) penalize the distance of the given data to the solution in both the image and transform domains. The third term guarantees the regularity and sparsity of the underlying solution. In this paper, we will derive an efficient iterative algorithm for the model (1.6) by using the split Bregman method (see [6]).

In general, there are two different assumptions on the sparse approximation of the underlying image, which lead to three different approaches for image restoration, namely the analysis-based approach, the synthesis-based approach and the balanced approach. In the analysis-based approach, we assume that the coefficient vector $\mathbf{W}\mathbf{f}$ can be sparsely approximated. The image restoration problem is usually reduced to a minimization problem involving the term $\mathbf{W}\mathbf{f}$. On the other hand, in the synthesis-based approach, we assume that the underlying image $\mathbf{f}$ can be synthesized by a sparse coefficient vector $\mathbf{t} \in \mathbb{R}^M$. The image restoration is usually formulated as a solution of a minimization problem on $\mathbb{R}^M$ which is of higher dimension than $\mathbb{R}^N$ when

$\mathbf{W}$ is a redundant tight frame transform. The balanced approach balances between the analysis-based and synthesis-based approaches. The model (1.4) is a balanced approach and the method proposed in this paper is an analysis-based approach. Interested readers should refer to [8, 16] for more details on these approaches. The three approaches for image restoration were developed independently in the literature. In [14], the authors gave a comprehensive analysis on the analysis-based and synthesis-based approaches, and pointed out a gap between them. The balanced approach bridges the gap between the analysis-based and synthesis-based approaches (see, e.g., [3, 4, 12, 21, 24]).

The rest of the paper is organized as follows. In Section 2, we give a sufficient condition, motivated by an uncertainty principle, which enables $\mathbf{f}$ to be exactly recovered. The iterative reconstruction algorithm is also derived. In Section 3, we focus on applying the analysis-based model (1.6) to handle the problem (1.3) and derive an iterative algorithm by use of the split Bregman method. Some applications of this algorithm are presented in Section 4 as illustration.

2. Exact recovery. In this section, we give a sufficient condition to exactly recover the underlying data $\mathbf{f}$ when incomplete data in the image and transform domains are $\mathbf{P}_\Lambda\mathbf{f}$ and $\mathbf{P}_\Gamma\mathbf{W}\mathbf{f}$ respectively. We also propose an algorithm to recover the data and compare it with the algorithm in [2].

To simplify notations, we denote $\Lambda^c := \mathcal{N}\backslash\Lambda$ and $\Gamma^c := \mathcal{M}\backslash\Gamma$.

THEOREM 2.1. *Let $\mathbf{W} \in \mathbb{R}^{M\times N}$ be a tight frame transform. For $\mathbf{f} \in \mathbb{R}^N$, suppose that the data available in the image domain is $\{\mathbf{f}(j)\}_{j\in\Lambda}$ and that in the transform domain is $\{\mathbf{W}\mathbf{f}(i)\}_{i\in\Gamma}$. If the inequality*

$$\sum_{i\in\Gamma^c}\sum_{j\in\Lambda^c}\mathbf{W}(i,j)^2 < 1 \tag{2.1}$$

holds, then $\mathbf{f}$ can be exactly recovered by

$$\mathbf{f} = (\mathbf{I} - \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c})^{-1}(\mathbf{W}^T\mathbf{P}_\Gamma\mathbf{W}\mathbf{f} + \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_\Lambda\mathbf{f}) \tag{2.2}$$

or

$$\mathbf{f} = (\mathbf{I} - \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W})^{-1}(\mathbf{P}_\Lambda\mathbf{f} + \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_\Gamma\mathbf{W}\mathbf{f}). \tag{2.3}$$

The key issue in proving the theorem is the invertibility of the operators $\mathbf{I} - \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}$ in (2.2) and $\mathbf{I} - \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}$ in (2.3), which we shall establish in due course.

As an example to illustrate how the sufficient condition (2.1) can be satisfied, consider the $M \times M$ matrix $\mathbf{C}$ of the type-II discrete cosine transform. Let $\mathbf{W}$ be an $M \times N$ matrix formed by any N distinct columns of $\mathbf{C}$. Since $\mathbf{C}$ is an orthogonal matrix, its columns are orthonormal and so $\mathbf{W}$ is a matrix with orthonormal columns, i.e. (1.1) holds. The entries of $\mathbf{W}$ are given by

$$\mathbf{W}(i,j) = \sigma_{i-1}\cos\left(\frac{(2\ell_j+1)(i-1)\pi}{2M}\right), \quad i = 1,\dots,M,\ j = 1,\dots,N,$$

where $\{\ell_1,\dots,\ell_N\} \subset \{0,\dots,M-1\}$, and

$$\sigma_\ell := \begin{cases} \frac{1}{\sqrt{M}}, & \text{if } \ell = 0, \\ \frac{\sqrt{2}}{\sqrt{M}}, & \text{if } \ell = 1,\dots,M-1. \end{cases}$$

Then $\mathbf{W}(i,j)^2 \leq \frac{2}{M}$ for all $i = 1, \ldots, M$, $j = 1, \ldots, N$. Thus

$$\sum_{i \in \Gamma^c} \sum_{j \in \Lambda^c} \mathbf{W}(i,j)^2 \leq \frac{2}{M} \left|\Gamma^c\right| \left|\Lambda^c\right| = 2\left(1 - \frac{|\Gamma|}{M}\right)\left(N - |\Lambda|\right), \tag{2.4}$$

where $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$. If $\Gamma = \mathcal{M}$ or $\Lambda = \mathcal{N}$, then (2.1) holds trivially. On the other hand, for any $\Gamma \subset \mathcal{M}$ such that $\frac{M}{2} < |\Gamma| < M$, since $\frac{1}{2\left(1 - \frac{|\Gamma|}{M}\right)} > 1$, there exists $\Lambda \subset \mathcal{N}$ which satisfies

$$N - \frac{1}{2\left(1 - \frac{|\Gamma|}{M}\right)} < |\Lambda| < N.$$

By (2.4), this gives (2.1).

Returning to the recovery result of Theorem 2.1, in order to fully recover the original data, we should have enough information from the available data. The sufficient condition (2.1) guarantees that a significant proportion of data in the image or transform domain is known. This fact is suggested by the proposition below which is a form of uncertainty principle in the same spirit as Theorem 2 in [13]. Various uncertainty principles of this type were derived in [18, 19]. While the general result Theorem 2.1 in [19] is applicable to the concrete setting on hand, our direct derivation here, utilizing the specifics of the case, gives a stronger result.

PROPOSITION 2.1. *Let $\mathbf{W} \in \mathbb{R}^{M \times N}$ be a tight frame transform, and $\Lambda \subset \mathcal{N}$ and $\Gamma \subset \mathcal{M}$ be given sets. If there exist a nonzero $\mathbf{f} \in \mathbb{R}^N$ and constants $\varepsilon, \eta \geq 0$ with $\varepsilon + \eta \leq 1$ such that*

$$\|\mathbf{P}_\Lambda \mathbf{f}\|_2 \leq \varepsilon \|\mathbf{f}\|_2 \tag{2.5}$$

and

$$\|\mathbf{P}_\Gamma \mathbf{W}\mathbf{f}\|_2 \leq \eta \|\mathbf{W}\mathbf{f}\|_2, \tag{2.6}$$

then

$$\sum_{i \in \Gamma^c} \sum_{j \in \Lambda^c} \mathbf{W}(i,j)^2 \geq (1 - \varepsilon - \eta)^2.$$

Proof. We first observe that

$$\begin{aligned} \|\mathbf{P}_{\Gamma^c} \mathbf{W} \mathbf{P}_{\Lambda^c} \mathbf{f}\|_2^2 = \sum_{i \in \Gamma^c} (\mathbf{W} \mathbf{P}_{\Lambda^c} \mathbf{f}(i))^2 &= \sum_{i \in \Gamma^c} \left(\sum_{j=1}^{N} \mathbf{W}(i,j) \mathbf{P}_{\Lambda^c} \mathbf{f}(j) \right)^2 \\ &= \sum_{i \in \Gamma^c} \left(\sum_{j \in \Lambda^c} \mathbf{W}(i,j) \mathbf{f}(j) \right)^2. \end{aligned}$$

By the Cauchy-Schwarz inequality, for every $i \in \Gamma^c$, we have

$$\left(\sum_{j \in \Lambda^c} \mathbf{W}(i,j) \mathbf{f}(j) \right)^2 \leq \left(\sum_{j \in \Lambda^c} \mathbf{W}(i,j)^2 \right) \left(\sum_{j \in \Lambda^c} \mathbf{f}(j)^2 \right) \leq \left(\sum_{j \in \Lambda^c} \mathbf{W}(i,j)^2 \right) \|\mathbf{f}\|_2^2.$$

Hence,

$$\|\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\mathbf{f}\|_2 \le \Big(\sum_{i\in\Gamma^c}\sum_{j\in\Lambda^c}\mathbf{W}(i,j)^2\Big)^{1/2}\|\mathbf{f}\|_2. \tag{2.7}$$

On the other hand, by the triangle inequality, we have

$$\begin{aligned}\|\mathbf{W}\mathbf{f}\|_2 - \|\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\mathbf{f}\|_2 &\le \|\mathbf{W}\mathbf{f} - \mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\mathbf{f}\|_2\\ &\le \|\mathbf{W}\mathbf{f} - \mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{f}\|_2 + \|\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{f} - \mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\mathbf{f}\|_2\\ &\le \eta\|\mathbf{W}\mathbf{f}\|_2 + \|\mathbf{P}_{\Gamma^c}\|\|\mathbf{W}\|\|\mathbf{f} - \mathbf{P}_{\Lambda^c}\mathbf{f}\|_2\\ &\le \eta\|\mathbf{f}\|_2 + \varepsilon\|\mathbf{f}\|_2,\end{aligned}$$

where the operator norms $\|\mathbf{P}_{\Gamma^c}\|$ and $\|\mathbf{W}\|$ are both equal to 1, and (2.6), (1.1) and (2.5) are employed in the last two inequalities. Combining this with (2.7) and again using (1.1), we obtain

$$\begin{aligned}\Big(\sum_{i\in\Gamma^c}\sum_{j\in\Lambda^c}\mathbf{W}(i,j)^2\Big)^{1/2}\|\mathbf{f}\|_2 &\ge \|\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\mathbf{f}\|_2\\ &\ge \|\mathbf{W}\mathbf{f}\|_2 - \eta\|\mathbf{f}\|_2 - \varepsilon\|\mathbf{f}\|_2\\ &= (1-\eta-\varepsilon)\|\mathbf{f}\|_2.\end{aligned}$$

Then the result follows from dividing both sides by $\|\mathbf{f}\|_2$. □

Proposition 2.1 shows that if (2.5) and (2.6) hold with small values of ε and η, i.e. there is little information given in both the image and transform domains, then the chance for (2.1) to hold is much lower, which means that exact recovery via Theorem 2.1 becomes more difficult.

Proof of Theorem 2.1. By (1.1), (2.7) and (2.1),

$$\|\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\| \le \|\mathbf{W}^T\|\|\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\| \le \Big(\sum_{i\in\Gamma^c}\sum_{j\in\Lambda^c}\mathbf{W}(i,j)^2\Big)^{1/2} < 1, \tag{2.8}$$

which ensures the invertibility of $\mathbf{I}-\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}$. Furthermore, a simple calculation shows that

$$\begin{aligned}(\mathbf{I} - \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c})\mathbf{f} &= \mathbf{W}^T(\mathbf{I}-\mathbf{P}_{\Gamma^c})\mathbf{W}\mathbf{f} + \mathbf{W}^T(\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{f} - \mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\mathbf{f})\\ &= \mathbf{W}^T\mathbf{P}_{\Gamma}\mathbf{W}\mathbf{f} + \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}(\mathbf{I}-\mathbf{P}_{\Lambda^c})\mathbf{f}\\ &= \mathbf{W}^T\mathbf{P}_{\Gamma}\mathbf{W}\mathbf{f} + \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda}\mathbf{f}.\end{aligned}$$

We note that the data on the right-hand side is known since both $\mathbf{P}_\Lambda\mathbf{f}$ and $\mathbf{P}_\Gamma\mathbf{W}\mathbf{f}$ are available. Then by the invertibility of $\mathbf{I} - \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}$, we can restore the original $\mathbf{f}$ exactly from the known information $\mathbf{P}_\Lambda\mathbf{f}$ and $\mathbf{P}_\Gamma\mathbf{W}\mathbf{f}$ via (2.2).

As for the second recovery formula (2.3), (2.8) implies that

$$\|\mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\| = \|(\mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W})^T\| = \|\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}\| < 1,$$

which ensures the invertibility of $\mathbf{I} - \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}$. Then we observe that

$$\begin{aligned}(\mathbf{I} - \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W})\mathbf{f} &= (\mathbf{I}-\mathbf{P}_{\Lambda^c})\mathbf{f} + \mathbf{P}_{\Lambda^c}(\mathbf{f} - \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{f})\\ &= \mathbf{P}_\Lambda\mathbf{f} + \mathbf{P}_{\Lambda^c}\mathbf{W}^T(\mathbf{I}-\mathbf{P}_{\Gamma^c})\mathbf{W}\mathbf{f}\\ &= \mathbf{P}_\Lambda\mathbf{f} + \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma}\mathbf{W}\mathbf{f}.\end{aligned}$$

The right-hand side of this inequality contains the known data $\mathbf{P}_\Lambda \mathbf{f}$ and $\mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$. Thus the invertibility of $\mathbf{I} - \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}$ enables the exact recovery of $\mathbf{f}$ via (2.3). □

The exact recovery provided by Theorem 2.1 necessarily implies that the recovery is unique. Indeed, under the assumption (2.1), suppose that two different $\widetilde{\mathbf{f}}_1$ and $\widetilde{\mathbf{f}}_2$ in $\mathbb{R}^N$ satisfy $\mathbf{P}_\Lambda \widetilde{\mathbf{f}}_1 = \mathbf{P}_\Lambda \widetilde{\mathbf{f}}_2$ and $\mathbf{P}_\Gamma \mathbf{W}\widetilde{\mathbf{f}}_1 = \mathbf{P}_\Gamma \mathbf{W}\widetilde{\mathbf{f}}_2$. Then $\mathbf{h} = \widetilde{\mathbf{f}}_1 - \widetilde{\mathbf{f}}_2 \neq 0$ satisfies $\mathbf{P}_\Lambda \mathbf{h} = \mathbf{P}_\Gamma \mathbf{W}\mathbf{h} = 0$, which contradicts Proposition 2.1 with $\varepsilon = \eta = 0$. We record the implication of this observation in the corollary below.

COROLLARY 2.1. *For an image $\mathbf{f} \in \mathbb{R}^N$, suppose that partial data $\mathbf{P}_\Lambda \mathbf{f}$ and $\mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$ are given, where $\mathbf{W} \in \mathbb{R}^{M\times N}$ is a tight frame transform, and $\Lambda \subset \mathcal{N}$ and $\Gamma \subset \mathcal{M}$ are index sets in the image and transform domains respectively. Then the condition* (2.1) *guarantees that any reconstructed image by interpolation is the original* $\mathbf{f}$.

Theorem 2.1 ensures the exact recovery of $\mathbf{f}$ under the condition (2.1), and (2.2) and (2.3) provide the recovery formulas. Since it is complicated to compute the inverse of a matrix, we now describe iterative algorithms for the realization of these two formulas.

The algorithms are based on the following well-known approach (see, e.g., [13, 18, 19]). Let $\mathbf{L}$ be an $N \times N$ matrix that satisfies $\|\mathbf{L}\| < 1$. Then the matrix $\mathbf{I} - \mathbf{L}$ is invertible. Starting from any given vector $\mathbf{v}_0 \in \mathbb{R}^N$, for $k \geq 0$, we set

$$\mathbf{v}_{k+1} = \mathbf{v}_0 + \mathbf{L}\mathbf{v}_k.$$

The contraction mapping principle ensures that this iterative procedure converges in a geometric rate to the unique fixed point $\mathbf{v}^*$ given by

$$\mathbf{v}^* = \mathbf{v}_0 + \mathbf{L}\mathbf{v}^*.$$

In other words, $\mathbf{v}^* = (\mathbf{I} - \mathbf{L})^{-1}\mathbf{v}_0$.

Thus, for (2.2), by substituting $\mathbf{L} = \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_{\Lambda^c}$ and $\mathbf{v}_0 = \mathbf{W}^T\mathbf{P}_\Gamma \mathbf{W}\mathbf{f} + \mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}\mathbf{P}_\Lambda \mathbf{f}$, we arrive at the iterative algorithm

$$\mathbf{f}_{k+1} = \mathbf{W}^T(\mathbf{P}_\Gamma \mathbf{W}\mathbf{f} + (\mathbf{I} - \mathbf{P}_\Gamma)\mathbf{W}(\mathbf{P}_\Lambda \mathbf{f} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{f}_k)). \tag{2.9}$$

In (2.9), the data is updated according to the following procedure. In the image domain, we update the coefficients on Λ by the known data $\mathbf{P}_\Lambda \mathbf{f}$, and then transform the resulting image to the transform domain and use the known data $\mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$ in the transform domain to replace the coefficients on Γ. The output is transformed back to the image domain for the next iteration.

Similarly, we formulate the iterative algorithm for (2.3) by letting $\mathbf{L} = \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_{\Gamma^c}\mathbf{W}$ and $\mathbf{v}_0 = \mathbf{P}_\Lambda \mathbf{f} + \mathbf{P}_{\Lambda^c}\mathbf{W}^T\mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$:

$$\mathbf{f}_{k+1} = \mathbf{P}_\Lambda \mathbf{f} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{W}^T(\mathbf{P}_\Gamma \mathbf{W}\mathbf{f} + (\mathbf{I} - \mathbf{P}_\Gamma)\mathbf{W}\mathbf{f}_k). \tag{2.10}$$

In (2.10), from $\mathbf{f}_k$ to $\mathbf{f}_{k+1}$, we first transform the image $\mathbf{f}_k$ to get the transform coefficients $\mathbf{W}\mathbf{f}_k$. Then we replace the coefficients on Γ by the known data $\mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$. After that, we transform the output back to the image domain and update the data on Λ by the known data $\mathbf{P}_\Lambda \mathbf{f}$ in the image domain.

While we have provided two algorithms for the exact recovery case, the computations below show that they are essentially the same. To this end, set the known

data in the image and transform domains as $\mathbf{x} = \mathbf{P}_\Lambda \mathbf{f}$ and $\mathbf{y} = \mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$. Then the algorithm (2.9) can be written as the iteration

$$\begin{cases} \mathbf{t}_k = \mathbf{x} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{f}_k, \\ \mathbf{r}_k = \mathbf{y} + (\mathbf{I} - \mathbf{P}_\Gamma)\mathbf{W}\mathbf{t}_k, \\ \mathbf{f}_{k+1} = \mathbf{W}^T \mathbf{r}_k. \end{cases} \tag{2.11}$$

The algorithm (2.10) is equivalent to the iteration

$$\begin{cases} \mathbf{r}_k = \mathbf{y} + (\mathbf{I} - \mathbf{P}_\Gamma)\mathbf{W}\mathbf{f}_k, \\ \mathbf{t}_k = \mathbf{W}^T \mathbf{r}_k, \\ \mathbf{f}_{k+1} = \mathbf{x} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{t}_k. \end{cases} \tag{2.12}$$

Comparing the two algorithms, they are both essentially alternating interpolation in the image and transform domains, since it is difficult to interpolate data in two domains simultaneously. When the iteration converges, the limit eventually interpolates the data in both domains. The only difference is that we choose $\mathbf{W}^T\mathbf{r}^*$ as the result in (2.9) and $\mathbf{P}_\Lambda \mathbf{f} + (\mathbf{I} - \mathbf{P}_\Lambda)\mathbf{t}^*$ in (2.10), where $\mathbf{r}^*$ and $\mathbf{t}^*$ are the limits of $\mathbf{r}_k$ in (2.11) and $\mathbf{t}_k$ in (2.12) respectively. Therefore, we may only consider (2.10) as our exact recovery algorithm. Note that the exact recovery algorithm (2.10) is slightly different from the balanced approach (1.5) with $\mathbf{x} = \mathbf{P}_\Lambda \mathbf{f}$ and $\mathbf{y} = \mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$ by a plunging denoising, i.e. soft thresholding, operator $\mathbf{T}_\mathbf{u}$. This is reasonable since the data $\mathbf{f}$ in practice is mostly contaminated by noise.

3. Analysis-based approach. As noted in Section 1, for given data $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^M$, the problem of finding $\mathbf{f} \in \mathbb{R}^N$ such that

$$\begin{cases} \mathbf{P}_\Lambda \mathbf{f} = \mathbf{x}, \\ \mathbf{P}_\Gamma \mathbf{W}\mathbf{f} = \mathbf{y}, \end{cases}$$

is generally very ill posed. It may have no solution in many cases; and since the data is normally contaminated by noise, it is unnecessary and impossible to have exact recovery in such situations.

We consider the following analysis-based model under the assumption that the data given, $\mathbf{x}$ and $\mathbf{y}$, are somewhat close to $\mathbf{P}_\Lambda \mathbf{f}$ and $\mathbf{P}_\Gamma \mathbf{W}\mathbf{f}$ respectively, where $\mathbf{f}$ is the underlying solution which has a sparse approximation in terms of the canonical coefficients $\mathbf{W}\mathbf{f}$. The solution is a minimizer of the minimization problem formulated in (1.6), i.e.

$$\min_{\mathbf{f} \in \mathbb{R}^N} \Big\{ \frac{1}{2} \|\mathbf{P}_\Lambda \mathbf{f} - \mathbf{x}\|_2^2 + \frac{\nu}{2} \|\mathbf{P}_\Gamma \mathbf{W}\mathbf{f} - \mathbf{y}\|_2^2 + \|\mathrm{diag}(\mathbf{u})\mathbf{W}\mathbf{f}\|_1 \Big\}.$$

Let us examine each of the terms in (1.6). The first term penalizes the distance between $\mathbf{f}$ and the given data $\mathbf{x}$ on Λ in the image domain. The second term penalizes the distance from $\mathbf{W}\mathbf{f}$ to the known data $\mathbf{y}$ on Γ in the transform domain. Thus the first two terms guarantee the fidelity of the solution. The third term employs the assumption that the underlying solution has a good sparse approximation provided by its canonical coefficients.

Since the minimization problem (1.6) is not separable, it cannot be solved directly by thresholding in contrast to the balanced and synthesis-based approaches (see, e.g.,

[2, 3]). We shall derive an iterative algorithm for (1.6) by using the split Bregman method.

The derivation of the split Bregman algorithm in [6, 20] is based on the Bregman distance. Furthermore, the split Bregman method can be understood as the augmented Lagrangian method (see [17]) applied to (1.6) (see, e.g., [15, 25]). It is clear that (1.6) is equivalent to the minimization problem

$$\min_{\{\mathbf{f}\in\mathbb{R}^N,\,\mathbf{d}\in\mathbb{R}^M:\,\mathbf{W}\mathbf{f}=\mathbf{d}\}}\Big\{\frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f}-\mathbf{x}\|_2^2+\frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{d}-\mathbf{y}\|_2^2+\|\mathrm{diag}(\mathbf{u})\mathbf{d}\|_1+\frac{\lambda}{2}\|\mathbf{d}-\mathbf{W}\mathbf{f}\|_2^2\Big\} \tag{3.1}$$

for any $\lambda>0$. The Lagrangian for the problem (3.1) is given by

$$\mathcal{L}(\mathbf{f},\mathbf{d},\mathbf{p})=\frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f}-\mathbf{x}\|_2^2+\frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{d}-\mathbf{y}\|_2^2+\|\mathrm{diag}(\mathbf{u})\mathbf{d}\|_1+\frac{\lambda}{2}\|\mathbf{d}-\mathbf{W}\mathbf{f}\|_2^2+\langle\mathbf{p},\mathbf{d}-\mathbf{W}\mathbf{f}\rangle.$$

The saddle points of $\mathcal{L}(\mathbf{f},\mathbf{d},\mathbf{p})$ are obtained by the iteration

$$\begin{cases}(\mathbf{f}_{k+1},\mathbf{d}_{k+1})=\underset{\mathbf{f}\in\mathbb{R}^N,\,\mathbf{d}\in\mathbb{R}^M}{\arg\min}\big\{\frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f}-\mathbf{x}\|_2^2+\frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{d}-\mathbf{y}\|_2^2+\|\mathrm{diag}(\mathbf{u})\mathbf{d}\|_1\\ \qquad\qquad\qquad\qquad+\frac{\lambda}{2}\|\mathbf{d}-\mathbf{W}\mathbf{f}\|_2^2+\langle\mathbf{p}_k,\mathbf{d}-\mathbf{W}\mathbf{f}\rangle\big\},\\ \mathbf{p}_{k+1}=\mathbf{p}_k+\lambda(\mathbf{d}_{k+1}-\mathbf{W}\mathbf{f}_{k+1}).\end{cases}$$

By letting $\mathbf{b}_k=-\mathbf{p}_k/\lambda$, this iteration becomes

$$\begin{cases}(\mathbf{f}_{k+1},\mathbf{d}_{k+1})=\underset{\mathbf{f}\in\mathbb{R}^N,\,\mathbf{d}\in\mathbb{R}^M}{\arg\min}\big\{\frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f}-\mathbf{x}\|_2^2+\frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{d}-\mathbf{y}\|_2^2+\|\mathrm{diag}(\mathbf{u})\mathbf{d}\|_1\\ \qquad\qquad\qquad\qquad+\frac{\lambda}{2}\|\mathbf{W}\mathbf{f}-\mathbf{d}+\mathbf{b}_k\|_2^2\big\},\\ \mathbf{b}_{k+1}=\mathbf{b}_k+(\mathbf{W}\mathbf{f}_{k+1}-\mathbf{d}_{k+1}).\end{cases}$$

For the first subproblem, we usually use an alternative minimization method to solve. Thus we obtain the following iteration for (1.6):

$$\begin{cases}\mathbf{f}_{k+1}=\underset{\mathbf{f}\in\mathbb{R}^N}{\arg\min}\big\{\frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f}-\mathbf{x}\|_2^2+\frac{\lambda}{2}\|\mathbf{W}\mathbf{f}-\mathbf{d}_k+\mathbf{b}_k\|_2^2\big\},\\ \mathbf{d}_{k+1}=\underset{\mathbf{d}\in\mathbb{R}^M}{\arg\min}\big\{\frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{d}-\mathbf{y}\|_2^2+\frac{\lambda}{2}\|\mathbf{W}\mathbf{f}_{k+1}-\mathbf{d}+\mathbf{b}_k\|_2^2+\|\mathrm{diag}(\mathbf{u})\mathbf{d}\|_1\big\},\\ \mathbf{b}_{k+1}=\mathbf{b}_k+(\mathbf{W}\mathbf{f}_{k+1}-\mathbf{d}_{k+1}).\end{cases} \tag{3.2}$$

The first subproblem in (3.2) is easy to solve and implement. Its solution is given by

$$\mathbf{f}_{k+1}=(\mathbf{P}_\Lambda+\lambda\mathbf{I})^{-1}(\mathbf{x}+\lambda\mathbf{W}^T(\mathbf{d}_k-\mathbf{b}_k)).$$

Since $\mathbf{P}_\Lambda+\lambda\mathbf{I}$ is a diagonal matrix, it can be readily inverted and computed.

For the second subproblem in (3.2), as it is separable with respect to $\mathbf{d}$, we can handle it by considering the cases $\mathbf{d}(i)$, $i\in\Gamma^c$, and $\mathbf{d}(i)$, $i\in\Gamma$, separately, where $\mathbf{d}(i)$ is the ith element of the vector $\mathbf{d}$. It turns out that for both cases, the solution is provided by a simple threshold. Indeed, for $i\in\Gamma^c$, the subproblem reduces to

$$\mathbf{d}_{k+1}(i)=\underset{\mathbf{d}(i)\in\mathbb{R}}{\arg\min}\Big\{\frac{\lambda}{2}(\mathbf{W}\mathbf{f}_{k+1}(i)-\mathbf{d}(i)+\mathbf{b}_k(i))^2+|\mathbf{u}(i)\mathbf{d}(i)|\Big\},$$

which has the well-known soft-thresholding solution $\mathbf{d}_{k+1}(i) = t_{\frac{\mathbf{u}(i)}{\lambda}}(\mathbf{W}\mathbf{f}_{k+1}(i)+\mathbf{b}_k(i))$ (see, e.g., [9]). On the other hand, for $i \in \Gamma$, the subproblem is

$$\mathbf{d}_{k+1}(i) = \underset{\mathbf{d}(i)\in\mathbb{R}}{\arg\min}\Big\{\frac{\nu}{2}(\mathbf{d}(i)-\mathbf{y}(i))^2 + \frac{\lambda}{2}(\mathbf{W}\mathbf{f}_{k+1}(i)-\mathbf{d}(i)+\mathbf{b}_k(i))^2 + |\mathbf{u}(i)\mathbf{d}(i)|\Big\},$$

which again is solved by a threshold. This follows from the observation that

$$\begin{aligned}
&\underset{\mathbf{d}(i)\in\mathbb{R}}{\arg\min}\Big\{\frac{\nu}{2}(\mathbf{d}(i)-\mathbf{y}(i))^2 + \frac{\lambda}{2}(\mathbf{W}\mathbf{f}_{k+1}(i)-\mathbf{d}(i)+\mathbf{b}_k(i))^2 + |\mathbf{u}(i)\mathbf{d}(i)|\Big\}\\
&= \underset{\mathbf{d}(i)\in\mathbb{R}}{\arg\min}\Big\{|\mathbf{u}(i)\mathbf{d}(i)| + \frac{\nu+\lambda}{2}\Big[\mathbf{d}(i) - \frac{\nu\mathbf{y}(i)+\lambda(\mathbf{W}\mathbf{f}_{k+1}(i)+\mathbf{b}_k(i))}{\nu+\lambda}\Big]^2\Big\}\\
&= t_{\frac{\mathbf{u}(i)}{\nu+\lambda}}\Big(\frac{\nu\mathbf{y}(i)+\lambda(\mathbf{W}\mathbf{f}_{k+1}(i)+\mathbf{b}_k(i))}{\nu+\lambda}\Big).
\end{aligned}$$

Combining these two cases, we obtain the solution for the second subproblem of (3.2) as

$$\mathbf{d}_{k+1} = (\mathbf{I}-\mathbf{P}_\Gamma)\mathbf{T}_{\frac{\mathbf{u}}{\lambda}}(\mathbf{W}\mathbf{f}_{k+1}+\mathbf{b}_k) + \mathbf{P}_\Gamma\mathbf{T}_{\frac{\mathbf{u}}{\nu+\lambda}}\Big(\frac{\nu\mathbf{y}+\lambda(\mathbf{W}\mathbf{f}_{k+1}+\mathbf{b}_k)}{\nu+\lambda}\Big).$$

Hence, the algorithm (3.2) can be written in a more explicit form, and its steps are as follows:

$$\text{(3.3)}\qquad \begin{cases}
\mathbf{f}_{k+1} = (\mathbf{P}_\Lambda + \lambda\mathbf{I})^{-1}(\mathbf{x}+\lambda\mathbf{W}^T(\mathbf{d}_k-\mathbf{b}_k)),\\
\mathbf{d}_{k+1} = (\mathbf{I}-\mathbf{P}_\Gamma)\mathbf{T}_{\frac{\mathbf{u}}{\lambda}}(\mathbf{W}\mathbf{f}_{k+1}+\mathbf{b}_k) + \mathbf{P}_\Gamma\mathbf{T}_{\frac{\mathbf{u}}{\nu+\lambda}}\big(\frac{\nu\mathbf{y}+\lambda(\mathbf{W}\mathbf{f}_{k+1}+\mathbf{b}_k)}{\nu+\lambda}\big),\\
\mathbf{b}_{k+1} = \mathbf{b}_k + (\mathbf{W}\mathbf{f}_{k+1}-\mathbf{d}_{k+1}),
\end{cases}$$

where $\mathbf{T}_\mathbf{u}$ is the soft thresholding operator as defined before. The steps in (3.3) are easy to implement and can also be efficiently executed.

The convergence analysis of the split Bregman method is similar to the discussions in [6, 12], so we only state the results without proof here. Assume that there is at least one solution $\mathbf{f}^*$ for (1.6). Then the sequence $\{\mathbf{f}_k\}_{k=0}^\infty$ generated by (3.2) satisfies

$$\begin{aligned}
&\lim_{k\to\infty}\Big(\frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f}_k-\mathbf{x}\|_2^2 + \frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{W}\mathbf{f}_k-\mathbf{y}\|_2^2 + \|\mathrm{diag}(\mathbf{u})\mathbf{W}\mathbf{f}_k\|_1\Big)\\
&= \frac{1}{2}\|\mathbf{P}_\Lambda\mathbf{f}^*-\mathbf{x}\|_2^2 + \frac{\nu}{2}\|\mathbf{P}_\Gamma\mathbf{W}\mathbf{f}^*-\mathbf{y}\|_2^2 + \|\mathrm{diag}(\mathbf{u})\mathbf{W}\mathbf{f}^*\|_1.
\end{aligned}$$

Furthermore, when (1.6) has a unique solution, we have $\lim_{k\to\infty}\|\mathbf{f}_k-\mathbf{f}^*\|_2 = 0$.

4. Numerical implementations. We apply our proposed analysis-based algorithm (3.3) of the model (1.6) to the following three image restoration problems:

(1) Image inpainting;
(2) Super-resolution image reconstruction with multiple sensors;
(3) Super-resolution image reconstruction with different zooms.

These three applications correspond to three different types of missing data recovery. The quality of the reconstructed images is evaluated by the peak signal-to-noise ratio (PSNR) defined as

$$\mathrm{PSNR} = 20\log_{10}\frac{255\sqrt{N}}{\|\mathbf{f}-\mathbf{f}^\infty\|_2},$$

where $\mathbf{f}$ and $\mathbf{f}^\infty$ are the original image and the reconstructed image respectively, and N is the number of pixels of $\mathbf{f}$. We also present the results obtained by the balanced approach (1.5) from [2] to compare with those by our analysis-based algorithm (3.3). In all our implementations, we set the initial data $\mathbf{f}_0$ to be the zero vector and the iteration result is recorded when the reconstructed image achieves the highest PSNR value. In practice, as the original image is not known, other criteria for stopping the iterations need to be employed. One possibility is to terminate the iterations when the corresponding values of the cost function in (1.6) stabilize.

The tight frame transform we use in the three applications is based on the piecewise linear tight wavelet frame constructed by the unitary extension principle in [23]. The construction begins with the linear B-spline as the refinable function and produces two piecewise linear mother wavelets. The corresponding filters of these three functions are

$$h_0 = (1/4, 2/4, 1/4),\ h_1 = (-1/4, 2/4, -1/4),\ h_2 = (\sqrt{2}/4, 0, -\sqrt{2}/4).$$

Readers can refer to [3] on how the transform matrix $\mathbf{W}$ is obtained from the given filters. In our implementations, we apply the fast decomposition and reconstruction algorithms derived in [10] (see also [12]).

EXAMPLE 4.1. (Image Inpainting) The first application of our algorithm is image inpainting, where only part of the information in the image domain is given and no information in the transform domain is available. In the context of the paper, the setting means that $\Lambda \neq \emptyset$ and $\Gamma = \emptyset$. This is an interesting and important inverse problem. Examples of practical situations in which the problem arises include removing scratches in photographs, restoring ancient drawings, and filling in missing pixels of digital images transmitted through a noisy channel. See, e.g., [3, 12] for more details on image inpainting.

FIG. 1. *Inpainting in the image domain for the "cameraman" image. Columns (from left to right) are the observed corrupted image, the recovered image by the analysis-based approach* (3.3)*, and the recovered image by the balanced approach* (1.5)*. The PSNR values of the recovered images are* 35.7742 *and* 34.3899 *respectively, with corresponding number of iterations being* 9 *and* 100.

Figure 1 presents the inpainting results of our algorithm (3.3) on the analysis-based approach and the algorithm (1.5) on the balanced approach. The missing data in the test image are the white words. In this experiment, (3.3) returns a slightly better result than (1.5) and in much fewer iterations. This is due to the common fact that the Bregman iteration recovers edges quickly.

EXAMPLE 4.2. (Super-Resolution with Multiple Sensors) This application involves reconstructing an image $\mathbf{f}$, as a super-resolution image, from its low-resolution images using K multiple sensors of the same resolution but with different subpixel displacements. This translates into knowing part of the information in the transform

domain but not the image domain. Therefore the application is for the case $\Lambda = \emptyset$ and $\Gamma \neq \emptyset$ in our model, and we take $K = 2$ in our implementation. Readers can refer to [2] on how the index set Γ is obtained from the low-resolution images and [1, 7] for more details about super-resolution image reconstruction with multiple sensors.

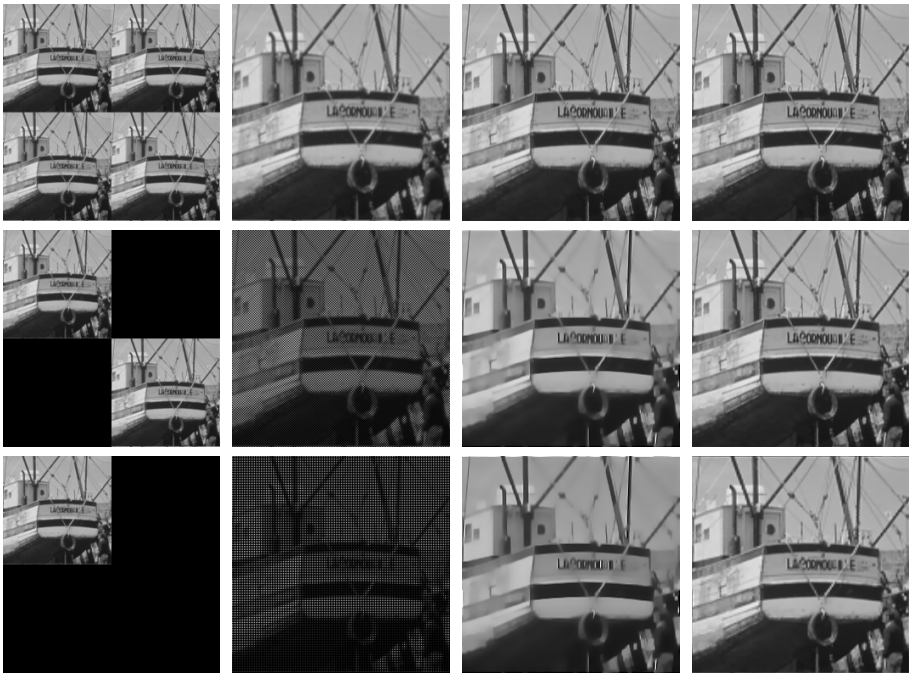

FIG. 2. *Super-resolution image reconstruction using* 2×2 *sensors for the "boat" image. Columns (from left to right) are the available low-resolution images, the observed high-resolution images, the reconstructed high-resolution images by the analysis-based approach* (3.3)*, and the reconstructed high-resolution images by the balanced approach* (1.5)*. The PSNR values of the reconstructed images in the third column (analysis-based approach) are* 31.7281*,* 28.0557 *and* 22.4243 *respectively, and those in the fourth column (balanced approach) are* 29.2638*,* 29.1752 *and* 24.5309 *respectively.*

The first column of the images in Figure 2 contains the given data. The second column comprises the observed high-resolution images obtained by directly applying the inverse transform $\mathbf{W}^T$ without any further processing. The third and fourth columns show the reconstructed high-resolution images by our analysis-based approach (3.3) and the balanced approach (1.5) respectively. The three rows in Figure 2 depict the results of three sets of experiments. All the low-resolution images are given in the first experiment and only some of the low-resolution images are available for the second and third experiments.

EXAMPLE 4.3. (Super-Resolution with Different Zooms) A super-resolution image may be constructed from images of different zooms, which could be part of the original image and low-resolution images of different resolutions. Figure 3 compares the results obtained by the analysis-based approach (3.3) and the balanced approach (1.5).

The setup of our experiment is as follows. We are given part of the data in the image domain, which is the first image in Figure 3. This amounts to $\Lambda \neq \emptyset$ in accordance to our formulation. In the transform domain, we use sensors of two

different resolutions to acquire two different data sets. By using 2×2 sensors, we first obtain a low-resolution image and then choose an appropriate index set to extract part of this low-resolution image. The output is the second image in Figure 3. By using 4×4 sensors (which corresponds to the second level of the tight frame transform associated with 2×2 sensors), we generate the low-resolution image displayed as the third image in Figure 3. The second and third images are the given data in the transform domain, which implies that $\Gamma \neq \emptyset$. See [2] for details on how the index set Γ is obtained from the given data in the transform domain and [2, 22] for more information on super-resolution image reconstruction with different zooms.

FIG. 3. *Super-resolution image reconstruction with different zooms for the "boat" image. Columns (from left to right) are part of the original image, part of a low-resolution image from* 2×2 *sensors, a low-resolution image from* 4×4 *sensors, the reconstructed high-resolution image by the analysis-based approach* (3.3)*, and the reconstructed high-resolution image by the balanced approach* (1.5)*. The PSNR values for the analysis-based approach and the balanced approach are* 25.7972 *and* 24.9855 *respectively.*

REFERENCES

[1] N. BOSE AND K. BOO, *High-resolution image reconstruction with multisensors*, Int. J. Imaging Syst. Tech., 9 (1998), pp. 294–304.

[2] J.-F. CAI, R. H. CHAN, L. SHEN, AND Z. SHEN, *Simultaneously inpainting in image and transformed domains*, Numer. Math., 112 (2009), pp. 509–533.

[3] J.-F. CAI, R. H. CHAN, AND Z. SHEN, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal., 24 (2008), pp. 131–149.

[4] J.-F. CAI, R. H. CHAN, AND Z. SHEN, *Simultaneous cartoon and texture inpainting*, Inverse Probl. Imaging, 4 (2010), pp. 379–395.

[5] J.-F. CAI, B. DONG, S. OSHER, AND Z. SHEN, *Image restoration: total variation, wavelet frames, and beyond*, J. Amer. Math. Soc., 25 (2012), pp. 1033–1089.

[6] J.-F. CAI, S. OSHER, AND Z. SHEN, *Split Bregman methods and frame based image restoration*, Multiscale Model. Simul., 8 (2009), pp. 337–369.

[7] R. H. CHAN, T. F. CHAN, L. SHEN, AND Z. SHEN, *Wavelet algorithms for high-resolution image reconstruction*, SIAM J. Sci. Comput., 24 (2003), pp. 1408–1432.

[8] S. S. CHEN, D. L. DONOHO, AND M. A. SAUNDERS, *Atomic decomposition by basis pursuit*, SIAM J. Sci. Comput., 20 (1998), pp. 33–61.

[9] P. L. COMBETTES AND J.-C. PESQUET, *Proximal thresholding algorithm for minimization over orthonormal bases*, SIAM J. Optim., 18 (2007), pp. 1351–1376.

[10] I. DAUBECHIES, B. HAN, A. RON, AND Z. SHEN, *Framelets: MRA-based constructions of wavelet frames*, Appl. Comput. Harmon. Anal., 14 (2003), pp. 1–46.

[11] M. N. DO AND M. VETTERLI, *The contourlet transform: an efficient directional multiresolution image representation*, IEEE Trans. Image Process., 14 (2005), pp. 2091–2106.

[12] B. DONG AND Z. SHEN, *MRA-based wavelet frames and applications*, in: IAS Lecture Notes Series 19, Summer Program on "The Mathematics of Image Processing", Park City Mathematics Institute, 2010, pp. 7–158.

[13] D. L. DONOHO AND P. B. STARK, *Uncertainty principles and signal recovery*, SIAM J. Appl. Math., 49 (1989), pp. 906–931.

[14] M. ELAD, P. MILANFAR, AND R. RUBINSTEIN, *Analysis versus synthesis in signal priors*, Inverse Probl., 23 (2007), pp. 947–968.

[15] E. Esser, *Applications of Lagrangian-based alternating direction methods and connections to split Bregman*, UCLA CAM Report 09-31, 2009.
[16] M. Fadili and J. L. Starck, *Sparse representations and Bayesian image inpainting*, in: Proceedings of the Workshop on Signal Processing with Adaptative Sparse Structured Representations, vol. 1, Rennes, France, 2005.
[17] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, 1987.
[18] S. S. Goh and T. N. T. Goodman, *Uncertainty principles in Banach spaces and signal recovery*, J. Approx. Theory, 143 (2006), pp. 26–35.
[19] S. S. Goh and J. Li, *Uncertainty principles of concentration type for time-frequency representations*, preprint.
[20] T. Goldstein and S. Osher, *The split Bregman method for l1-regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.
[21] O. G. Guleryuz, *Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising part II: adaptive algorithms*, IEEE Trans. Image Process., 15 (2006), pp. 555–571.
[22] M. V. Joshi, S. Chaudhuri, and R. Panuganti, *Super-resolution imaging: use of zoom as a cue*, Image Vision Comput., 22 (2004), pp. 1185–1196.
[23] A. Ron and Z. Shen, *Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator*, J. Funct. Anal., 148 (1997), pp. 408–447.
[24] Z. Shen, *Wavelet frames and image restorations*, in: R. Bhatia (Ed.), Proceedings of the International Congress of Mathematicians, vol. IV, Hindustan Book Agency, Hyderabad, India, 2010, pp. 2834–2863.
[25] X. C. Tai and C. Wu, *Augmented Lagrangian method, dual methods and split Bregman iteration for ROF, vectorial TV, and high order models*, SIAM J. Imaging Sci., 3 (2010), pp. 300–339.