

# SPLIT BREGMAN METHODS AND FRAME BASED IMAGE RESTORATION

JIAN-FENG CAI\*, STANLEY OSHER†, AND ZUOWEI SHEN‡

**Abstract.** Split Bregman methods introduced in [T. Goldstein and S. Osher, *SIAM J. Imaging Sci.*, 2(2009), pp. 323–343] have been demonstrated to be efficient tools for solving total variation norm minimization problems, which arise from partial differential equation based image restoration such as image denoising and magnetic resonance imaging reconstruction from sparse samples. In this paper, we prove the convergence of the split Bregman iterations, where the number of inner iterations is fixed to be one. Furthermore, we show that these split Bregman iterations can be used to solve minimization problems arising from the analysis based approach for image restoration in the literature. We apply these split Bregman iterations to the analysis based image restoration approach whose analysis operator is derived from tight framelets constructed in [A. Ron and Z. Shen, *J. Funct. Anal.*, 148(1997), pp. 408–447]. This gives a set of new frame based image restoration algorithms that cover several topics in image restorations, such as image denoising, deblurring, inpainting, and cartoon-texture image decomposition. Several numerical simulation results are provided.

**Key words.** split Bregman, wavelet frames, image restorations

**AMS subject classifications.** 94A08, 65T60, 90C90

**1. Introduction.** Image restoration is often formulated as an inverse problem. Since only discrete forms will be used eventually in numerical implementations, we focus on discrete settings in this paper. For simplicity, we denote images as vectors in  $\mathbb{R}^n$  by concatenating their columns. The objective of image restoration is to find the unknown true image  $u \in \mathbb{R}^n$  from an observed image (or measurements)  $f \in \mathbb{R}^\ell$  defined by

$$f = Au + \epsilon, \tag{1.1}$$

where  $\epsilon$  is a white Gaussian noise with variance  $\sigma$ , and  $A \in \mathbb{R}^{\ell \times n}$  is a linear operator, typically a convolution operator in image deconvolution, a projection in image inpainting, or the identity in image denoising.

Images usually have sparse representations (or sparse approximations) in some transformed domains. Such transforms can be, e.g., Fourier or windowed Fourier transforms, local cosine transforms, wavelet or framelet transforms, or discrete gradient operators. In order to make use of the sparsity, one solves (1.1) in the corresponding transformed domain and finds a sparse solution of (1.1) in the transformed domain. In many cases, the sparse solution can be approximated by solving an  $\ell_1$  norm regularized minimization problem. By different sparsity a priori assumptions, it is common to divide these  $\ell_1$  regularized minimization problems into two subcategories, namely, the *analysis based* and the *synthesis based* sparsity priors associated with the transform under which images have sparse representations or approximations.

Let  $\mathcal{D} \in \mathbb{R}^{m \times n}$  be a linear transform, called an analysis operator, acting from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ . The matrix  $\mathcal{D}$  can be generated by discrete Fourier transforms, local cosine

---

\*Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, CA 90095, USA. Email: cai@math.ucla.edu.

†Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, CA 90095, USA. Email: sjo@math.ucla.edu. Research partially supported by ONR grant N000140710810 and by the US Department of Defense.

‡Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543, Singapore. Email: matzuows@nus.edu.sg. Research supported in part by Grant R-146-000-113-112 at the National University of Singapore.

transforms, wavelet or framelet transforms, or discrete gradient operators. We will use redundant transforms, so we have that  $m \geq n$  in most cases.

Although the focus of this paper is the analysis based approach, we start with the synthesis based approach, which is well studied in the literature and motivates us to study the analysis based approach. The synthesis based approach is applicable only when  $m \geq n$  and  $\mathcal{D}$  has a left inverse  $\mathcal{R}$ , i.e.,  $\mathcal{R}\mathcal{D} = I$ . In this case, it is necessary that the column vectors of  $\mathcal{R}$  span  $\mathbb{R}^n$ . The matrix  $\mathcal{R} \in \mathbb{R}^{n \times m}$  is also called a synthesis operator, acting from  $\mathbb{R}^m$  to  $\mathbb{R}^n$ . For a given image  $u \in \mathbb{R}^n$ , there are infinitely many ways to represent  $u$  by the columns of  $\mathcal{R}$ , and for some of them the coefficient sequence may not be in the range of  $\mathcal{D}$ . The synthesis based approach explores full sparsity of all possible representations of  $u$  in terms of the columns of  $\mathcal{R}$ . For this, let  $d \in \mathbb{R}^m$  be the coefficient of the given image  $u$  in terms of the columns of  $\mathcal{R}$ . We first find the sparsest possible coefficient  $d$  by solving an  $\ell_1$  norm minimization problem. Then the solution of (1.1) is derived by synthesizing  $u$  via the synthesis operator  $\mathcal{R}$ , i.e.,  $u = \mathcal{R}d$ . More precisely, the synthesis based approach is to solve either the unconstrained minimization

$$\min_d |d| + H(\mathcal{R}d), \quad (1.2)$$

where  $H(\cdot)$  is a smooth convex function as the data fidelity (e.g.,  $H(\mathcal{R}d) = \frac{\lambda}{2} \|\mathcal{A}\mathcal{R}d - f\|^2$ ), or the constrained one

$$\min_d |d| \quad \text{subject to} \quad \mathcal{A}\mathcal{R}d = f. \quad (1.3)$$

Here  $|\cdot|$  stands for the  $\ell_1$  norm or its variants. These formulations cover a lot of problems in the literature by properly choosing  $\mathcal{A}$ ,  $\mathcal{R}$ , and  $H$ , e.g., frame based methods solving inverse problems proposed by [31, 40] and basis pursuit problems [32] in compressed sensing [19, 41].

Problems (1.2) and (1.3) arising from the synthesis based approach are relatively easy to solve, as the nonsmooth term  $|d|$  involved is separable. Hence, it is now well developed so far; see, e.g., [12–14, 31, 38, 40, 68]. Many of these algorithms are based on scalar soft thresholdings. The basic idea is to use the gradient descent to deal with smooth terms or constraints to obtain a latent vector and then find a vector that has a small  $\ell_1$  norm and is in the proximity of the latent vector. Hence, the values of both  $|d|$  and the smooth term are decreased.

More precisely, an iterative thresholding algorithm was proposed in [38] for (1.2), which can be also formulated as a proximal forward-backward splitting algorithm by [35]. See also [24, 46, 49]. The iteration is given as

$$\begin{cases} c^{k+1} = d^k - \delta \mathcal{R}^T \nabla H(\mathcal{R}d^k), \\ d^{k+1} = \arg \min_d \frac{1}{2} \|d - c^{k+1}\|^2 + \delta |d|. \end{cases} \quad (1.4)$$

The key factor to make this algorithm work is that the second step is equivalent to a scalar shrinkage. For example, when  $|d| := \|d\|_1$ , the second step is the soft thresholding

$$d^{k+1} = \mathcal{T}_\delta(c^{k+1}),$$

where

$$\mathcal{T}_\delta(w) = [t_\delta(w_1), t_\delta(w_2), \dots, \dots]^T, \quad \text{with} \quad t_\delta(w_i) = \text{sgn}(w_i) \max\{0, |w_i| - \delta\}. \quad (1.5)$$

Therefore, this iteration is simple and makes use of the sparsity of the underlying solution. The convergence of the above algorithm in various forms was analyzed in the above-mentioned papers.

For the solution of (1.3), the linearized Bregman iteration was first proposed in [58, 68] and was used as an efficient tool for compressed sensing. The iteration is given as

$$\begin{cases} c^{k+1} = c^k - \delta \mathcal{R}^T A^T (A \mathcal{R} d^k - f), \\ d^{k+1} = \arg \min_d \frac{1}{2} \|d - c^{k+1}\|^2 + \mu |d|. \end{cases} \quad (1.6)$$

Again, the second step is equivalent to a scalar shrinkage, which uses the sparsity of the solution. The convergence analysis of this algorithm was given in [12, 13], and its applications to frame based deblurring was given in [14]. With algorithms (1.4) and (1.6), the synthesis based approach can be considered well developed.

The analysis based sparsity prior is to assume that, for the image  $u \in \mathbb{R}^n$  under consideration, its analysis coefficient  $\mathcal{D}u$  is sparse in  $\mathbb{R}^m$ ; i.e., the number of nonzero entries of the coefficient  $\mathcal{D}u$  is small. Finding a solution of (1.1) with a sparse  $\mathcal{D}u$  usually yields a minimization problem involving a term  $|\mathcal{D}u|$ , where  $|\cdot|$  denotes the  $\ell_1$  norm or its variants. In particular, it can be formulated to find a solution of the unconstrained minimization:

$$\min_u |\mathcal{D}u| + H(u), \quad (1.7)$$

where  $H(\cdot)$  is a smooth convex function as the data fidelity. For example,  $H(\cdot)$  can be chosen  $H(u) = \frac{\lambda}{2} \|Au - f\|^2$ . Or, it can be formulated to find a solution of the constrained minimization:

$$\min_u |\mathcal{D}u| \quad \text{subject to} \quad Au = f. \quad (1.8)$$

Formulations (1.7) and (1.8) can be applied to many important problems in imaging science (and other computational areas), including discrete partial differential equation (PDE) based approaches [30, 60] and wavelet frame based ones.

The  $\ell_1$  terms involved in (1.7) and (1.8) are nonsmooth and nonseparable. This prevents us from using optimization methods for smooth functions. A natural idea to solve them is to use a smoothed  $\ell_1$  norm to approximate the  $\ell_1$  norm and then apply optimization methods for smooth functions. This approach is commonly used in some early algorithms for PDE based, particularly for total variation (TV) based, image processing minimization (1.7); see, e.g., [28, 65]. However, in these smoothing methods, the better the approximations (to the  $\ell_1$  norm) are, the more slowly the algorithms converge. Therefore, in order to make these algorithms converge faster, the smooth approximation to the  $\ell_1$  norm cannot be too accurate. Consequently, one cannot obtain a sparse solution by solving the minimization problem with a smoothed norm. However, sparsity is important in many cases of  $\ell_1$  regularization problems. Moreover, some nice properties induced by the  $\ell_1$  term in these minimization problems will be lost. For example, the  $\ell_1$  norm term yields a solution with large strongly homogeneous zones, as shown in [55, 56].

Another difficulty for the analysis based approach of (1.7) and (1.8) is that the term  $|\mathcal{D}u|$  is not separable. Therefore, one cannot simply use the soft thresholding as one normally does in the synthesis based approach, since it is impossible to keep any sequence staying in the range of  $\mathcal{D}$  after applying a thresholding operator. An

iterative method to find a solution in the range of  $\mathcal{D}$  for a special case of (1.7), where  $H(u) = \frac{1}{2}\|u - f\|^2$ , was proposed by [23] by exploring duality.

A different approach was proposed in [47]. In fact, two algorithms, called split Bregman algorithms, were given in [47] to solve (1.7) and (1.8) respectively. Furthermore, they were shown to be powerful in [47, 69] when these two algorithms are applied to various PDE based image restoration approaches, e.g., the ROF model and nonlocal PDE models. One of the objectives of this paper is to give a convergence analysis of these two algorithms when the number of inner iterations is chosen to be one. Another objective is to apply models (1.7) and (1.8) to frame based image restoration, especially, frame based image deblurring and image inpainting.

We now discuss all these in more detail. As mentioned above, compared with the synthesis based approach, the  $\ell_1$  norm of  $|\mathcal{D}u|$  involved in (1.7) and (1.8) is neither smooth nor separable. To overcome this, one transfers (1.7) and (1.8) to problems involving only separable nonsmooth terms. In particular, one first replaces the term  $|\mathcal{D}u|$  in (1.7) and (1.8) by a separable one  $|d|$  and then adds a new constraint  $d = \mathcal{D}u$  into (1.7) and (1.8). Hence, (1.7) becomes

$$\min_u |d| + H(u) \quad \text{subject to} \quad d = \mathcal{D}u, \quad (1.9)$$

and (1.8) becomes

$$\min_u |d| \quad \text{subject to} \quad Au = f; \quad d = \mathcal{D}u. \quad (1.10)$$

Then either the constraint  $d = \mathcal{D}u$  is approximated (cf. [66]), or a method for constrained minimizations with a separable nonsmooth objective function is employed (cf. [47]).

In order to solve (1.9), an iterative algorithm based on the Bregman distance with an inexact solver was proposed in [47]. This leads to the alternating split Bregman iteration for (1.9). When the number of inner iterations is chosen to be one in this alternating split Bregman iteration, it becomes

$$\begin{cases} u^{k+1} = \arg \min_u H(u) + \frac{\lambda}{2} \|\mathcal{D}u - d^k + b^k\|_2^2, \\ d^{k+1} = \arg \min_d |d| + \frac{\lambda}{2} \|d - \mathcal{D}u^{k+1} - b^k\|_2^2, \\ b^{k+1} = b^k + \delta(\mathcal{D}u^{k+1} - d^{k+1}). \end{cases} \quad (1.11)$$

Since this iteration is for the unconstrained minimization problem (1.9), we call it the *unconstrained split Bregman method*. The split Bregman iteration for (1.9) was demonstrated in [47] to be an efficient tool for solving problems arising from TV norm minimization problems of PDE based models for image restoration, such as denoising. Since  $H(u)$  is convex and differentiable, the subproblem in the first line is easy to solve. Further, noting that the first term of the subproblem in the second line is the  $\ell_1$  norm, the subproblem in the second line can be solved by a simple soft shrinkage. Both of these make the iteration efficient and fast for many problems that are difficult to solve by other means. Besides its speed, the split Bregman method has several advantages. It has a relatively small memory footprint compared to second order methods that require explicit representations of the Hessian matrix. Also, the method is easy to code. Both of these characteristics make this split Bregman method a practical algorithm for large scale problems.

As for the solution of (1.10), a split Bregman method for (1.10) was also introduced in [47]. When we fix the number of inner iterations to be one, that iteration

becomes

$$\begin{cases} u^{k+1} = \arg \min_u \frac{\mu}{2} \|Au - f + c^k\|^2 + \frac{\lambda}{2} \|\mathcal{D}u - d^k + b^k\|_2^2, \\ d^{k+1} = \arg \min_d |d| + \frac{\lambda}{2} \|d - \mathcal{D}u^{k+1} - b^k\|_2^2, \\ b^{k+1} = b^k + \delta_b(\mathcal{D}u^{k+1} - d^{k+1}), \\ c^{k+1} = c^k + \delta_c(Au^{k+1} - f). \end{cases} \quad (1.12)$$

Again, as shown in [47], the split Bregman iteration for (1.10) performs extremely well in the problems arising from, e.g., magnetic resonance imaging reconstruction from sparse samples. Since this iteration is for the constrained minimization problem (1.10), we call it the *constrained split Bregman method*. Again, the subproblems involved in this method are very easy to solve: the first step is equivalent to solving a system of linear equations, and the second step is a simple shrinkage. Therefore, the constrained split Bregman method has the same advantages as the unconstrained one.

When  $\mathcal{D}$  is an invertible matrix, the analysis based and the synthesis based approaches can be transferred from one to the other. For example, by letting  $\mathcal{R} = \mathcal{D}^{-1}$ , (1.7) and (1.8) are equivalent to (1.2) and (1.3) respectively. However, for noninvertible matrices  $\mathcal{D}$  and/or  $\mathcal{R}$ , the analysis based and the synthesis based approaches cannot be transferred from one to the other. In fact, it was observed in, for examples, [32, 42] that there is a gap between the analysis based and the synthesis based approaches. Both of them have their own favorable data sets and applications. In general, it is hard to draw the conclusion as to which approach is better, without specifying the applications and the data sets. However, numerical simulation results show that the analysis based approach tends to generate smoother images in application. It is natural, since the coefficient  $\mathcal{D}u$  is quite often able to be linked with the smoothness of the underlying image. Meanwhile, the synthesis approach tends to explore more sparsity. In fact, a model that bridges the analysis based and the synthesis based approaches in image inpainting and deconvolution was proposed in [8, 10, 11, 15]. We forgo the detailed discussion here, and interested readers should consult [8, 10, 11, 15] or the last section of this paper for details.

In this paper, we will give a convergence analysis for iterations (1.11) and (1.12). Furthermore, both algorithms will be applied to deriving frame based image restoration algorithms. The rest of the paper is organized as follows. In Section 2, we will derive split Bregman methods (1.11) and (1.12) from the Bregman iteration with inexact inner solvers. In Section 3, we prove the convergence of split Bregman methods. The new applications of split Bregman methods to frame based image restoration are illustrated in Section 4.

**2. Formulation.** In this section, we derive split Bregman methods (1.11) and (1.12) from the Bregman iteration with inexact inner solvers, where the number of inner iterations is chosen to be one. We start with an introduction to the Bregman iteration in Section 2.1, and then split Bregman methods are described in Section 2.2.

Iterative algorithms involving the Bregman distance were introduced to image and signal processing by, e.g., [21] and by many other authors. See [57] for an overview. In [57], a Bregman iteration was proposed for the nondifferentiable TV energy for image restoration. Then, in [68], it was shown to be remarkably successful for  $\ell_1$  norm minimization problems in compressive sensing. To further improve the performance of the Bregman iteration, a linearized Bregman iteration was invented in [36]; see also [68]. More details and an improvement called “kicking” of the linearized

Bregman iteration were described in [58], a rigorous theory was given in [12, 13], and applications to frame based image deblurring can be found in [14]. Moreover, in [7], the linearized Bregman iteration was successfully applied to matrix completion (see e.g. [18]) to obtain an efficient singular value thresholding algorithm. Split Bregman iterations were introduced in [47], which extended the utility of the Bregman iteration and the linearized Bregman iteration to minimizations of more general  $\ell_1$  based regularizations including TV, Besov norms, and sums of such things. Other variants of Bregman iterations include Bregmanized operator splitting (BOS) and preconditioned BOS (PBOS); see [69]. Wavelet based denoising using the Bregman iteration was introduced in [67], and it was further extended by using translation invariant wavelets in [53].

**2.1. Bregman iteration.** Since split Bregman methods are motivated by the Bregman method, we start with the Bregman iteration. It is based on the concept of the Bregman distance [6]. For any convex function  $J(u)$ , the Bregman distance is defined by

$$B_J^p(u, v) = J(u) - J(v) - \langle u - v, p \rangle, \quad \text{where } p \in \partial J(v). \quad (2.1)$$

In general  $B_J^p(u, v) \neq B_J^p(v, u)$  and the triangle inequality is not satisfied, so  $B_J^p(u, v)$  is not a distance in the usual sense. However it does measure the closeness between  $u$  and  $v$  in the sense that  $B_J^p(u, v) \geq 0$  and  $B_J^p(u, v) \geq B_J^p(w, v)$  for all points  $w$  on the line segment connecting  $u$  and  $v$ .

The goal of the Bregman iteration is to solve the general constrained convex minimization problem

$$\min_u J(u) \quad \text{subject to} \quad Lu = f, \quad (2.2)$$

where  $J(u)$  is a convex energy, and  $L$  is a linear operator. Traditionally, this problem may be solved by continuation methods, where we solve sequentially the unconstrained problems  $\min_u J(u) + \frac{\lambda^k}{2} \|Lu - f\|^2$ . By choosing a predefined sequence  $\{\lambda^k\}_{k \in \mathbb{Z}}$  that tends to infinity, we get a solution of the constrained minimization problem (2.2).

The idea of the Bregman iteration is also to transfer the constrained problem (2.2) into a series of unconstrained ones. Instead of varying the regularization parameter  $\lambda^k$  as in continuation methods, the Bregman iteration fixes this parameter and varies the data. More explicitly, given  $u^0 = 0$  and  $p^0 = 0$ , we iteratively solve

$$\begin{aligned} u^{k+1} &= \arg \min_u B_J^{p^k}(u, u^k) + \frac{\lambda}{2} \|Lu - f\|^2 \\ &= \arg \min_u J(u) - J(u^k) - \langle u - u^k, p^k \rangle + \frac{\lambda}{2} \|Lu - f\|^2. \end{aligned} \quad (2.3)$$

In order that (2.3) is well defined for  $k+1$ , it must hold that  $p^{k+1} \in \partial J(u^{k+1})$ . Define

$$p^{k+1} = p^k - \lambda L^T (Lu^{k+1} - f). \quad (2.4)$$

By differentiating the energy in (2.3), we obtain that  $p^{k+1} \in \partial J(u^{k+1})$ . Iteration (2.3) together with (2.4) is the Bregman iteration. In [57], the authors analyzed the convergence of the Bregman iterative scheme. In particular, it is shown that, under fairly weak assumptions on  $J(\cdot)$ ,  $\|Lu^k - f\| \rightarrow 0$  as  $k \rightarrow \infty$ . The Bregman iteration has several nice denoising properties which were discussed and proved in [57, 68].

The Bregman iteration has several advantages over traditional penalty function and continuation methods. First, the Bregman iteration converges very quickly when applied to certain types of objective functions, especially for problems involving an  $\ell_1$  regularization term. A nice explanation of why this is true was given in [47]. When the Bregman iteration converges quickly, we need only solve a small number of unconstrained problems. The second (and perhaps most significant) advantage of the Bregman iteration over continuation methods is that the value of  $\lambda$  in (2.3) remains constant. We can therefore choose a value for  $\lambda$  that minimizes the condition number of the subproblems, resulting in fast convergence for iterative optimization methods, such as Newton or Gauss–Seidel. The Bregman iteration also avoids the problem of numerical instabilities that occur as  $\lambda^k \rightarrow \infty$  that arise when using continuation methods.

The Bregman iteration (2.3) and (2.4) can be reformulated into a compact form. By  $p^0 = 0$  and (2.4), we obtain  $p^{k+1} = -\lambda L^T \sum_{i=1}^{k+1} (Lu^i - f)$ . Substitute this into (2.3), and it yields

$$u^{k+1} = \arg \min_u J(u) + \frac{\lambda}{2} \left\| Lu - f + \sum_{i=1}^k (Lu^i - f) \right\|^2.$$

Let  $b^k = \sum_{i=1}^k (Lu^i - f)$  and  $u^0 = 0$ . Then, we get a compact form of the Bregman iteration (2.3) and (2.4) as follows: Given  $b^0 = 0$  and  $u^0 = 0$ ,

$$\begin{cases} u^{k+1} = \arg \min_u J(u) + \frac{\lambda}{2} \|Lu - f + b^k\|^2, \\ b^{k+1} = b^k + (Lu^{k+1} - f). \end{cases} \quad (2.5)$$

It is this form that will be used in the remaining part of this paper.

Since there is generally no explicit expression for the solution of the subminimization problem involved in (2.5), one has to employ iterative algorithms to solve it. Different approximations to the solution of the sub minimization problem in (2.5) lead to different new algorithms. Here, we briefly recall two of them before we get to split Bregman iterations. The first one is the Bregmanized operator splitting (BOS) algorithm introduced in [69], which solves the sub minimization problems by using the proximal forward-backward splitting iteration (1.4) with one executing step in each iteration. More precisely, the BOS algorithm is

$$\begin{cases} v^{k+1} = u^k - \delta \lambda L^T (Lu^k - f + b^k), \\ u^{k+1} = \arg \min_u \delta J(u) + \frac{1}{2} \|u - v^{k+1}\|^2, \\ b^{k+1} = b^k + (Lu^{k+1} - f). \end{cases} \quad (2.6)$$

A preconditioned version of BOS is also introduced in [69], and the convergence analysis is also given there.

If the term  $\frac{1}{2} \|Lu - f\|^2$  in (2.3) is approximated by  $\frac{1}{2\delta} \|u - (u^k - \delta L^T (Lu^k - f))\|^2$ , we obtain

$$\begin{cases} u^{k+1} = \arg \min_u J(u) - J(u^k) - \langle u - u^k, p^k \rangle + \frac{\lambda}{2\delta} \|u - (u^k - \delta L^T (Lu^k - f))\|^2, \\ p^{k+1} = p^k - \lambda L^T (Lu^{k+1} - f). \end{cases} \quad (2.7)$$

This can be shown to be BOS. Note that  $p^{k+1}$  does not belong to  $\partial J(u^{k+1})$ . By adjusting the second equation in (2.7) to keep  $p^{k+1} \in \partial J(u^{k+1})$ , one gets

$$\begin{cases} u^{k+1} = \arg \min_u J(u) - J(u^k) - \langle u - u^k, p^k \rangle + \frac{\lambda}{2\delta} \|u - (u^k - \delta L^T(Lu^k - f))\|^2, \\ p^{k+1} = p^k - \frac{\lambda}{\delta}(u^{k+1} - u^k) - \lambda L^T(Lu^{k+1} - f). \end{cases} \quad (2.8)$$

Denote  $v^k = \frac{\delta}{\lambda} p^k + u^k$ . Then (2.8) becomes the second algorithm, the linearized Bregman iteration for  $J(u)$ :

$$\begin{cases} v^{k+1} = v^k - \delta L^T(Lu^k - f), \\ u^{k+1} = \arg \min_u \frac{\lambda}{2} \|u - v^{k+1}\|^2 + \delta J(u). \end{cases} \quad (2.9)$$

When  $J(u)$  is the  $\ell_1$  norm as in the synthesis based approach, the subminimization problems involved in the BOS algorithm (2.6) and the linearized Bregman algorithm (2.9) can be solved by scalar shrinkages. Hence, these algorithms are easy to implement and are efficient. However, when  $J(u)$  is  $|\mathcal{D}u|$  as in the analysis based approach, there is no such simple solver for the subminimization problems involved. Hence, the BOS algorithm (2.6) and the linearized Bregman iteration (2.9) are not simple anymore. One has to use inner iterative solvers to solve subminimizations involved there, so the efficiency of (2.6) and (2.9) depends heavily on the inner solvers. This is one of the reasons why we are using the split Bregman iterations for the analysis based approach.

**2.2. Split Bregman methods.** The goal of split Bregman methods [47] is to extend the utility of the Bregman iteration and the linearized Bregman iteration to minimizations of more general  $\ell_1$  regularization in the form of (1.7) and (1.8). The basic idea is to introduce an intermediate variable  $d$  such that  $\mathcal{D}u = d$ , and the term  $|\mathcal{D}u|$  in (1.7) and (1.8) is separable and easy to minimize. Then the Bregman iteration with an inexact solver is applied.

**2.2.1. Unconstrained minimization problem.** The unconstrained minimization problem (1.7) is transferred into a constrained one (1.9). Then, the Bregman iteration (2.5) for solving (1.9) can be written as follows: let  $u^0 = 0$ ,  $d^0 = 0$ , and  $b^0 = 0$ , and define the sequence  $(u^k, d^k, b^k)$  by

$$\begin{cases} (u^{k+1}, d^{k+1}) = \arg \min_{u,d} |d| + H(u) + \frac{\lambda}{2} \|\mathcal{D}u - d + b^k\|^2, \\ b^{k+1} = b^k + (\mathcal{D}u^{k+1} - d^{k+1}). \end{cases} \quad (2.10)$$

The convergence of (2.10) was shown in [47, 52] under the assumption that the subproblem in the first step is solved exactly. However, in practice, it is not so easy to find the exact solution of the subproblem

$$(u^{k+1}, d^{k+1}) = \arg \min_{u,d} |d| + H(u) + \frac{\lambda}{2} \|\mathcal{D}u - d + b^k\|^2.$$

One commonly used method to solve it is the alternative minimization, or a block nonlinear Gauss-Seidel algorithm. Therefore, we get an algorithm as follows:

$$\begin{cases} \text{for } n = 1 \text{ to } N \\ \quad u^{k+1} \leftarrow \arg \min_u H(u) + \frac{\lambda}{2} \|\mathcal{D}u - d^{new} + b^k\|_2^2, \\ \quad d^{k+1} \leftarrow \arg \min_d |d| + \frac{\lambda}{2} \|d - \mathcal{D}u^{new} - b^k\|_2^2, \\ \text{end} \\ b^{k+1} = b^k + (\mathcal{D}u^{k+1} - d^{k+1}), \end{cases} \quad (2.11)$$



where  $d^{new}$  (respectively  $u^{new}$ ) is either  $d^{k+1}$  (respectively  $u^{k+1}$ ) if it is available or  $d^k$  (respectively  $u^k$ ) otherwise.

As pointed out in [47], it is not desirable to solve the first subproblem to full convergence; i.e.,  $N$  is not infinity in (2.11). Intuitively, the reason for this is that if the error in our solution for this subproblem is small compared to  $\|b^k - b^*\|^2$ , where  $b^*$  is the “true  $b$ ,” then this extra precision will be “wasted” when the Bregman parameter is updated. In fact, it was found empirically in [47] that for many applications optimal efficiency is obtained when only one iteration of the inner loop is performed (i.e.  $N = 1$  in (2.11)).

In a special case when  $\mathcal{D}^T \mathcal{D} = I$  and  $H(u) = \frac{\mu}{2} \|u - f\|^2$  (this corresponds to frame based image denoising), the convergence of (2.11) with  $N = 1$  was analyzed in [62]. In this case, one simply apply the idea of forward-backward splitting in [35]. Recently, we became aware that a proof of the convergence of (2.11) with  $N = 1$  was given in [61]. The idea is to reformulate (1.11) as a Douglas–Rachford splitting on the dual problem.

We will give the convergence proof of the constrained split Bregman method (1.12) in Section 3.1. With a slight modification, the proof leads to a proof of the convergence of (1.11) in Section 3.2. More precisely, it yields a convergence proof of the iteration

$$\begin{cases} u^{k+1} = \arg \min_u H(u) + \frac{\lambda}{2} \|\mathcal{D}u - d^k + b^k\|_2^2, \\ d^{k+1} = \arg \min_d |d| + \frac{\lambda}{2} \|d - \mathcal{D}u^{k+1} - b^k\|_2^2, \\ b^{k+1} = b^k + \delta(\mathcal{D}u^{k+1} - d^{k+1}) \end{cases}$$

whenever  $0 < \delta \leq 1$ . When  $\delta = 1$ , it becomes the unconstrained split Bregman iteration (2.11) with  $N = 1$ .

**2.2.2. Constrained minimization problem.** The Bregman iteration (2.5) for the constrained minimization problem (1.8) can be formulated as follows: Given  $c^0 = 0$  and  $u^0 = 0$ , define

$$\begin{cases} u^{k+1} = \arg \min_u |\mathcal{D}u| + \frac{\mu}{2} \|Au - f + c^k\|^2, \\ c^{k+1} = c^k + (Au^{k+1} - f). \end{cases} \quad (2.12)$$

In the first step, we have to solve subproblems

$$u^{k+1} = \arg \min_u |\mathcal{D}u| + \frac{\mu}{2} \|Au - f + c^k\|^2,$$

which is a special case of (1.7). These subproblems are solved by the unconstrained alternating split Bregman method (2.11) with  $N = 1$ , i.e., by the iteration

$$\begin{cases} u^{k+1} = \arg \min_u \frac{\mu}{2} \|Au - f + c^k\|^2 + \frac{\lambda}{2} \|\mathcal{D}u - d^k + b^k\|_2^2, \\ d^{k+1} = \arg \min_d |d| + \frac{\lambda}{2} \|d - \mathcal{D}u^{k+1} - b^k\|_2^2, \\ b^{k+1} = b^k + (\mathcal{D}u^{k+1} - d^{k+1}). \end{cases}$$

Combining this inner solver with the outer iteration (2.12), we obtain the constrained split Bregman method for (1.8) as follows:

$$\left\{ \begin{array}{l} \text{for } m = 1 \text{ to } M \\ \quad u^{k+1} \leftarrow \arg \min_u \frac{\mu}{2} \|Au - f - c^k\|^2 + \frac{\lambda}{2} \|d^{new} - \mathcal{D}u - b^{new}\|_2^2, \\ \quad d^{k+1} \leftarrow \arg \min_d |d| + \frac{\lambda}{2} \|d - \mathcal{D}u^{new} - b^{new}\|_2^2, \\ \quad b^{k+1} \leftarrow b^{new} + (\mathcal{D}u^{new} - d^{new}), \\ \text{end} \\ c^{k+1} = c^k + (Au^{k+1} - f). \end{array} \right. \quad (2.13)$$

Here  $b^{new}$  (respectively  $u^{new}$  and  $d^{new}$ ) is either  $b^{k+1}$  (respectively  $u^{k+1}$  and  $d^{k+1}$ ) if it is available or  $b^k$  (respectively  $u^k$  and  $d^k$ ) otherwise.

For applications in image restoration, it is not necessary to solve each unconstrained subproblem entirely to numerical precision. Therefore,  $M$  is a finite integer. Furthermore, when parameter values are properly chosen, it has been found in [47] that the outer loop of this algorithm need only be executed a small number of times. As a result, this algorithm is very fast when properly chosen parameter values are used.

In Section 3, we will prove that one just needs to pick  $M = 1$  in (2.13) to ensure the convergence of (2.13). In fact, we will prove that iteration (1.12), which is

$$\left\{ \begin{array}{l} u^{k+1} = \arg \min_u \frac{\mu}{2} \|Au - f + c^k\|^2 + \frac{\lambda}{2} \|\mathcal{D}u - d^k + b^k\|_2^2, \\ d^{k+1} = \arg \min_d |d| + \frac{\lambda}{2} \|d - \mathcal{D}u^{k+1} - b^k\|_2^2, \\ b^{k+1} = b^k + \delta_b (\mathcal{D}u^{k+1} - d^{k+1}), \\ c^{k+1} = c^k + \delta_c (Au^{k+1} - f), \end{array} \right.$$

converges when  $\mu > 0$ ,  $\lambda > 0$ ,  $0 < \delta_b \leq 1$ , and  $0 < \delta_c < 2$ . In particular, when  $\delta_b = \delta_c = 1$ , it becomes the constrained split Bregman iteration (2.13) with  $M = 1$ .

**3. Convergence.** In this section, we prove the convergence of (1.11) and (1.12). Similar to the arguments in [7] for the linearized Bregman and in [69] for the BOS, split Bregman methods (1.11) and (1.12) can be recast as inexact Uzawa algorithms. Inexact Uzawa algorithms have been analyzed in, for examples, [1, 4, 5, 20, 33, 50, 51]. The available analyses seem to be either for linear saddle point problems (cf. [1, 4, 5, 20, 50, 51]) or for strongly convex function minimizations (cf. [33, 51]). Models (1.7) and (1.8) are neither linear saddle point problems nor strongly convex function minimization problems. Therefore, we cannot directly use methods in the literature for proving (1.11) and (1.12). However, some ideas in the literature can still be used. For example, our analysis here is motivated by [33] for inexact Uzawa algorithms for linear saddle point problems, which is also closely related to the approach of [69] for the convergence analysis of the (preconditioned) BOS algorithm.

While we were preparing the final version of this paper, we were referred to [61], where an independent, but different, proof of the convergence of (1.11) had been given recently by reformulating (1.11) as a Douglas–Rachford splitting on the dual problem. This led us to reorganizing the paper as presented now, i.e. giving a detailed proof of the convergence of (1.12) and an outline of the proof of the convergence of (1.11) by using an approach similar to but different from that in [61].

**3.1. Convergence of iteration (1.12).** In this subsection, we prove the convergence of (1.12) for the constrained minimization problem (1.8). For this, we note

that, since all the subproblems involved in (1.12) are convex, the first order optimality condition gives the following fact:

$$\begin{cases} 0 = \mu A^T(Au^{k+1} - f + c^k) + \lambda \mathcal{D}^T(\mathcal{D}u^{k+1} - d^k + b^k), \\ 0 = p^{k+1} + \lambda(d^{k+1} - \mathcal{D}u^{k+1} - b^k), \quad \text{with } p^{k+1} \in \partial|d^{k+1}|, \\ b^{k+1} = b^k + \delta_b(\mathcal{D}u^{k+1} - d^{k+1}), \\ c^{k+1} = c^k + \delta_c(Au^{k+1} - f), \end{cases} \quad (3.1)$$

which will be used in the proof of the next result.

**THEOREM 3.1.** *Assume that there exists at least one solution  $u^*$  of (1.8). Assume that  $\mu > 0$ ,  $\lambda > 0$ ,  $0 < \delta_b \leq 1$ , and  $0 < \delta_c < 2$ . Then, the following properties for the constrained split Bregman iteration (1.12) hold:*

$$\lim_{k \rightarrow +\infty} \|Au^k - f\| = 0, \quad \lim_{k \rightarrow +\infty} |\mathcal{D}u^k| = |\mathcal{D}u^*|. \quad (3.2)$$

Furthermore,

$$\lim_{k \rightarrow +\infty} \|u^k - u^*\| = 0 \quad (3.3)$$

whenever (1.8) has a unique solution.

*Proof.* Let

$$L(u, w) = |\mathcal{D}u| + \langle w, Au - f \rangle$$

be the Lagrangian of (1.8) with  $w$  being the Lagrangian multiplier. Since there exists at least one solution  $u^*$  of (1.8) by the assumption, the KKT condition asserts that there must exist a vector  $w^*$  such that

$$\begin{cases} 0 = \mathcal{D}^T p^* + A^T w^*, \quad \text{with } p^* \in \partial|d^*|, \quad \text{where } d^* = \mathcal{D}u^*, \\ Au^* = f. \end{cases} \quad (3.4)$$

Let  $b^* = p^*/\lambda$  and  $c^* = w^*/\mu$ . Then (3.4) leads to

$$\begin{cases} 0 = \mu A^T(Au^* - f + c^*) + \lambda \mathcal{D}^T(\mathcal{D}u^* - d^* + b^*), \\ 0 = p^* + \lambda(d^* - \mathcal{D}u^* - b^*), \quad \text{with } p^* \in \partial|d^*|, \\ b^* = b^* + \delta_b(\mathcal{D}u^* - d^*), \\ c^* = c^* + \delta_c(Au^* - f). \end{cases} \quad (3.5)$$

This means that  $u^*, d^*, b^*, c^*$  is a fixed point of (1.12).

Denote the errors by

$$u_e^k = u^k - u^*, \quad d_e^k = d^k - d^*, \quad c_e^k = c^k - c^*, \quad b_e^k = b^k - b^*.$$

Subtracting the first equation of (3.1) by the first equation of (3.5), we obtain

$$0 = \mu A^T(Au_e^{k+1} + c_e^k) + \lambda \mathcal{D}^T(\mathcal{D}u_e^{k+1} - d_e^k + b_e^k).$$

Taking the inner product of the left- and right- hand sides with respect to  $u_e^{k+1}$ , we have

$$0 = \mu \|Au_e^{k+1}\|^2 + \lambda \|\mathcal{D}u_e^{k+1}\|^2 - \lambda \langle \mathcal{D}^T d_e^k, u_e^{k+1} \rangle + \mu \langle A^T c_e^k, u_e^{k+1} \rangle + \lambda \langle \mathcal{D}^T b_e^k, u_e^{k+1} \rangle. \quad (3.6)$$

The same manipulations applied to the second equation of (3.1) and the second equation of (3.5) lead to

$$0 = \langle p^{k+1} - p^*, d^{k+1} - d^* \rangle + \lambda \|d_e^{k+1}\|^2 - \lambda \langle \mathcal{D}u_e^{k+1}, d_e^{k+1} \rangle - \lambda \langle b_e^k, d_e^{k+1} \rangle, \quad (3.7)$$

where  $p^{k+1} \in \partial|d^{k+1}|$  and  $p^* = \lambda b^* \in \partial|d^*|$ . By summing (3.6) and (3.7), we get

$$0 = \mu \|Au_e^{k+1}\|^2 + \langle p^{k+1} - p^*, d^{k+1} - d^* \rangle + \mu \langle A^T c_e^k, u_e^{k+1} \rangle + \lambda (\|\mathcal{D}u_e^{k+1}\|^2 + \|d_e^{k+1}\|^2 - \langle \mathcal{D}u_e^{k+1}, d_e^k + d_e^{k+1} \rangle + \langle b_e^k, \mathcal{D}u_e^{k+1} - d_e^{k+1} \rangle). \quad (3.8)$$

Furthermore, by subtracting the fourth equation of (3.1) by the fourth equation of (3.5), we have

$$c_e^{k+1} = c_e^k + \delta_c Au_e^{k+1},$$

which leads to

$$\|c_e^{k+1}\|^2 = \|c_e^k\|^2 + \delta_c^2 \|Au_e^{k+1}\|^2 + 2\delta_c \langle c_e^k, Au_e^{k+1} \rangle,$$

and further

$$\langle c_e^k, Au_e^{k+1} \rangle = \frac{1}{2\delta_c} (\|c_e^{k+1}\|^2 - \|c_e^k\|^2) - \frac{\delta_c}{2} \|Au_e^{k+1}\|^2. \quad (3.9)$$

Similarly, the same manipulations applied to the third equation of (3.1) and the third equation of (3.5) yield

$$\langle b_e^k, \mathcal{D}u_e^{k+1} - d_e^{k+1} \rangle = \frac{1}{2\delta_b} (\|b_e^{k+1}\|^2 - \|b_e^k\|^2) - \frac{\delta_b}{2} \|\mathcal{D}u_e^{k+1} - d_e^{k+1}\|^2. \quad (3.10)$$

Substituting (3.9) and (3.10) into (3.8), we have

$$\begin{aligned} & \frac{\mu}{2\delta_c} (\|c_e^k\|^2 - \|c_e^{k+1}\|^2) + \frac{\lambda}{2\delta_b} (\|b_e^k\|^2 - \|b_e^{k+1}\|^2) \\ &= \mu \left(1 - \frac{\delta_c}{2}\right) \|Au_e^{k+1}\|^2 + \langle p^{k+1} - p^*, d^{k+1} - d^* \rangle \\ & \quad + \lambda \left( \|\mathcal{D}u_e^{k+1}\|^2 + \|d_e^{k+1}\|^2 - \langle \mathcal{D}u_e^{k+1}, d_e^k + d_e^{k+1} \rangle - \frac{\delta_b}{2} \|\mathcal{D}u_e^{k+1} - d_e^{k+1}\|^2 \right). \end{aligned} \quad (3.11)$$

By summing the above equation from  $k = 0$  to  $k = K$ , we get

$$\begin{aligned} & \frac{\mu}{2\delta_c} (\|c_e^0\|^2 - \|c_e^{K+1}\|^2) + \frac{\lambda}{2\delta_b} (\|b_e^0\|^2 - \|b_e^{K+1}\|^2) \\ &= \mu \left(1 - \frac{\delta_c}{2}\right) \sum_{k=0}^K \|Au_e^{k+1}\|^2 + \sum_{k=0}^K \langle p^{k+1} - p^*, d^{k+1} - d^* \rangle \\ & \quad + \lambda \left( \frac{1 - \delta_b}{2} \sum_{k=0}^K \|\mathcal{D}u_e^{k+1} - d_e^{k+1}\|^2 + \frac{1}{2} \sum_{k=0}^K \|\mathcal{D}u_e^{k+1} - d_e^k\|^2 + \frac{1}{2} \|d_e^{K+1}\|^2 \right) - \frac{\lambda}{2} \|d_e^0\|^2. \end{aligned} \quad (3.12)$$

Since  $p^{k+1} \in \partial|d^{k+1}|$  and  $p^* \in \partial|d^*|$ , and  $|\cdot|$  is convex, we have

$$\langle p^{k+1} - p^*, d^{k+1} - d^* \rangle \geq 0 \quad \forall k. \quad (3.13)$$

Therefore, all terms involved in (3.12) are nonnegative. This fact leads to the following inequality:

$$\begin{aligned} & \frac{\mu}{2\delta_c} \|c_e^0\|^2 + \frac{\lambda}{2\delta_b} \|b_e^0\|^2 + \frac{\lambda}{2} \|d_e^0\|^2 \\ & \geq \mu \left(1 - \frac{\delta_c}{2}\right) \sum_{k=0}^K \|Au_e^{k+1}\|^2 + \sum_{k=0}^K \langle p^{k+1} - p^*, d^{k+1} - d^* \rangle + \frac{\lambda}{2} \sum_{k=0}^K \|\mathcal{D}u_e^{k+1} - d_e^k\|^2. \end{aligned} \quad (3.14)$$

This leads to the following conclusions: it follows from (3.14) and the assumptions of  $\mu > 0$  and  $0 < \delta_c < 2$  that

$$\sum_{k=0}^{+\infty} \|Au_e^{k+1}\|^2 < +\infty.$$

This, together with  $Au_e^{k+1} = Au^{k+1} - f$ , leads to the first equation in (3.2) holding, i.e.,

$$\lim_{k \rightarrow +\infty} \|Au^{k+1} - f\| = 0.$$

Second, (3.14) also leads to

$$\sum_{k=0}^{+\infty} \langle p^{k+1} - p^*, d^{k+1} - d^* \rangle < +\infty,$$

and hence

$$\lim_{k \rightarrow +\infty} \langle p^k - p^*, d^k - d^* \rangle = 0. \quad (3.15)$$

Recall that, for any convex function  $J$ , the Bregman distance (2.1) satisfies

$$B_J^p(u, v) + B_J^q(v, u) = \langle q - p, u - v \rangle \quad \forall p \in \partial J(v), q \in \partial J(u). \quad (3.16)$$

This, together with (3.15) and the nonnegativity of the Bregman distance, implies that  $\lim_{k \rightarrow +\infty} B_{|\cdot|}^{p^*}(d^k, d^*) = 0$ , i.e.,

$$\lim_{k \rightarrow +\infty} |d^k| - |d^*| - \langle d^k - d^*, p^* \rangle = 0. \quad (3.17)$$

Moreover, since  $\lambda > 0$ , (3.12) also asserts that  $\sum_{k=0}^K \|\mathcal{D}u_e^{k+1} - d_e^k\|^2 < +\infty$ , which implies that  $\lim_{k \rightarrow +\infty} \|\mathcal{D}u_e^{k+1} - d_e^k\| = 0$ . By  $\mathcal{D}u^* = d^*$ , we conclude that

$$\lim_{k \rightarrow +\infty} \|\mathcal{D}u^{k+1} - d^k\| = 0. \quad (3.18)$$

Since  $|\cdot|$  is continuous, by (3.17) and (3.18), we obtain

$$\lim_{k \rightarrow +\infty} |\mathcal{D}u^k| - |\mathcal{D}u^*| - \langle \mathcal{D}u^k - \mathcal{D}u^*, p^* \rangle = 0. \quad (3.19)$$

Furthermore, since  $Au^k \rightarrow f$  and  $Au^* = f$ ,

$$\lim_{k \rightarrow +\infty} \langle Au^k - Au^*, w^* \rangle = 0.$$

Summing this with (3.19) yields

$$\begin{aligned}
0 &= \lim_{k \rightarrow +\infty} |\mathcal{D}u^k| - |\mathcal{D}u^*| - \langle \mathcal{D}u^k - \mathcal{D}u^*, p^* \rangle - \langle Au^k - Au^*, w^* \rangle \\
&= \lim_{k \rightarrow +\infty} |\mathcal{D}u^k| - |\mathcal{D}u^*| - \langle u^k - u^*, \mathcal{D}^T p^* + A^T w^* \rangle \\
&= \lim_{k \rightarrow +\infty} |\mathcal{D}u^k| - |\mathcal{D}u^*|,
\end{aligned} \tag{3.20}$$

where the last equality comes from  $0 = \mathcal{D}^T p^* + A^T w^*$  in (3.4). This gives the second equation in (3.2).

Next, we prove (3.3) by assuming that (1.8) has a unique solution. It is proved by contradiction. Let  $w^*$  be the vector in (3.4). Define

$$E(u) := L(u, w^*) + \|Au - f\|^2 = |\mathcal{D}u| + \langle w^*, Au - f \rangle + \|Au - f\|^2$$

Then  $E(u)$  is a convex and continuous function. Also, since  $(u^*, w^*)$  is a saddle point of  $L(u, w)$  and  $Au^* = f$ , we have  $E(u) \geq E(u^*)$ . In case of  $u \neq u^*$ , it holds that  $E(u) > E(u^*)$ . This can be seen from the following. When  $u \neq u^*$ , if  $Au = f$ , then  $E(u) > E(u^*)$  follows immediately from the uniqueness of the solution of (1.8); otherwise,  $\|Au - f\|^2 > 0 = \|Au^* - f\|^2$  and therefore

$$E(u) = L(u, w^*) + \|Au - f\|^2 \geq L(u^*, w^*) + \|Au - f\|^2 > L(u^*, w^*) + \|Au^* - f\|^2 = E(u^*).$$

Now we suppose that (3.3) does not hold, so there exists a subsequence  $u^{k_i}$  such that  $\|u^{k_i} - u^*\| > \epsilon$  for some  $\epsilon > 0$  and for all  $i$ . Then,  $E(u^{k_i}) > \min\{E(u) : \|u - u^*\| = \epsilon\}$ . Indeed, let  $v$  be the intersection of the sphere  $\{u : \|u - u^*\| = \epsilon\}$  and the line segment from  $u^*$  to  $u^{k_i}$ ; then there exists a positive number  $t \in (0, 1)$  such that  $v = tu^* + (1 - t)u^{k_i}$ . By the convexity of  $E$  and the definition of  $u^*$ , we have

$$\begin{aligned}
E(u^{k_i}) &> tE(u^*) + (1 - t)E(u^{k_i}) \geq E(tu^* + (1 - t)u^{k_i}) = E(v) \\
&\geq \min\{E(u) : \|u - u^*\| = \epsilon\}.
\end{aligned}$$

Denote  $\tilde{u} = \arg \min\{E(u) : \|u - u^*\| = \epsilon\}$ . By applying (3.2), we have

$$E(u^*) = \lim_{i \rightarrow +\infty} E(u^{k_i}) \geq E(\tilde{u}) > E(u^*),$$

which is a contradiction.  $\square$

Here are some remarks on Theorem 3.1. Assume that  $|\cdot|$  is the  $\ell_1$  norm and  $\mathcal{D}$  has a left inverse (e.g.,  $\mathcal{D}$  is a tight frame). Since  $|\mathcal{D}u|$  is equivalent to its  $\ell_2$  norm in finite-dimensional spaces,  $|\mathcal{D}u|$  is equivalent to the  $\ell_2$  norm of  $u$ . This leads to the fact that the convex energy functional  $|\mathcal{D}u|$  is coercive. Hence (1.8) has at least one solution. Consequently, (3.2) in Theorem 3.1 implies that, while  $Au^k$  converges to  $f$ , the energy  $|\mathcal{D}u^k|$  of  $u^k$  converges to the minimum value of  $|\mathcal{D}u|$  subject to  $Au = f$ . In other words, the constrained split Bregman algorithm (1.12) reaches a solution that is arbitrarily close to the constraint and has an energy arbitrarily close to the minimum energy as long as sufficiently many iterations are executed. Since the key for model (1.8) is to find a solution that attains the minimum of the energy functional  $|\mathcal{D}u|$  subject to  $Au = f$ , (3.2) implies that algorithm (1.12) is the right one for this. Furthermore, it is easy to show that  $\{u^k\}_{k \in \mathbb{Z}}$  has a convergent subsequence. Indeed, since  $|\mathcal{D}u|$  is equivalent to its  $\ell_2$  norm, (3.2) implies that  $u^k$  is bounded. Hence, there exists a convergent subsequence of  $\{u^k\}_{k \in \mathbb{Z}}$ . Finally, when  $A$  is invertible, the sequence  $u^k$  converges. This follows from Theorem 3.1, because (1.8) has a unique solution.

**3.2. Convergence of iteration (1.11).** Our proof presented in Section 3.1 can be modified to prove the convergence of (1.11).

Similar to Section 3.1, we note that, since all the subproblems involved in (1.11) are convex, the first order optimality condition gives

$$\begin{cases} 0 = \nabla H(u^{k+1}) + \lambda \mathcal{D}^T(\mathcal{D}u^{k+1} - d^k + b^k), \\ 0 = p^{k+1} + \lambda(d^{k+1} - \mathcal{D}u^{k+1} - b^k), \quad \text{with } p^{k+1} \in \partial|d^{k+1}|, \\ b^{k+1} = b^k + \delta(\mathcal{D}u^{k+1} - d^{k+1}). \end{cases} \quad (3.21)$$

This simple observation will be used in our proof of the convergence of the unconstrained split Bregman method (1.11). Since the proof is similar to the proof of (1.12), we will omit some computational details. However, we will give enough details to understand the idea of the proof.

**THEOREM 3.2.** *Assume that there exists at least one solution  $u^*$  of (1.7). Assume that  $0 < \delta \leq 1$  and  $\lambda > 0$ . Then, we have the following properties for the unconstrained split Bregman iteration (1.11):*

$$\lim_{k \rightarrow +\infty} |\mathcal{D}u^k| + H(u^k) = |\mathcal{D}u^*| + H(u^*). \quad (3.22)$$

Furthermore,

$$\lim_{k \rightarrow +\infty} \|u^k - u^*\| = 0 \quad (3.23)$$

whenever (1.7) has a unique solution.

*Proof.* Let  $u^*$  be an arbitrary solution of (1.7). By the first order optimality condition,  $u^*$  must satisfy

$$0 = \mathcal{D}^T p^* + \nabla H(u^*), \quad (3.24)$$

where  $p^* \in \partial|d^*|$  with  $d^* = \mathcal{D}u^*$ . Let

$$b^* = \frac{1}{\lambda} p^*.$$

We obtain

$$\begin{cases} 0 = \nabla H(u^*) + \lambda \mathcal{D}^T(\mathcal{D}u^* - d^* + b^*), \\ 0 = p^* + \lambda(d^* - \mathcal{D}u^* - b^*), \quad \text{with } p^* \in \partial|d^*|, \\ b^* = b^* + \delta(\mathcal{D}u^* - d^*). \end{cases} \quad (3.25)$$

Therefore,  $u^*, d^*, b^*$  is a fixed point of (3.21) of the unconstrained split Bregman iteration (1.11). Consequently, if the unconstrained split Bregman iteration converges, it converges to a solution of (1.7), as stated in [47].

Denote the errors by

$$u_e^k = u^k - u^*, \quad d_e^k = d^k - d^*, \quad b_e^k = b^k - b^*.$$

An argument similar to the corresponding part of the proof of Theorem 3.1 allows us

obtain

$$\begin{aligned}
& \frac{\lambda}{2\delta} (\|b_e^0\|^2 - \|b_e^{K+1}\|^2) \\
&= \sum_{k=0}^K \langle \nabla H(u^{k+1}) - \nabla H(u^*), u^{k+1} - u^* \rangle + \sum_{k=0}^K \langle p^{k+1} - p^*, d^{k+1} - d^* \rangle \\
&+ \lambda \left( \frac{1-\delta}{2} \sum_{k=0}^K \|\mathcal{D}u_e^{k+1} - d_e^{k+1}\|^2 + \frac{1}{2} \sum_{k=0}^K \|\mathcal{D}u_e^{k+1} - d_e^k\|^2 + \frac{1}{2} \|d_e^{K+1}\|^2 \right) - \frac{\lambda}{2} \|d_e^0\|^2.
\end{aligned} \tag{3.26}$$

Note that all terms involved in the above equation are nonnegative. This observation leads to the following inequality:

$$\frac{\lambda}{2\delta} \|b_e^0\|^2 + \frac{\lambda}{2} \|d_e^0\|^2 \geq \sum_{k=0}^K \langle \nabla H(u^{k+1}) - \nabla H(u^*), u^{k+1} - u^* \rangle. \tag{3.27}$$

By assumption  $\lambda > 0$ , we have  $\sum_{k=0}^{+\infty} \langle \nabla H(u^{k+1}) - \nabla H(u^*), u^{k+1} - u^* \rangle < +\infty$ , which leads to

$$\lim_{k \rightarrow +\infty} \langle \nabla H(u^k) - \nabla H(u^*), u^k - u^* \rangle = 0. \tag{3.28}$$

This, together with (3.16) and the nonnegativity of the Bregman distance, implies that  $\lim_{k \rightarrow +\infty} B_H^{\nabla H(u^*)}(u^k, u^*) = 0$ , i.e.,

$$\lim_{k \rightarrow +\infty} H(u^k) - H(u^*) - \langle u^k - u^*, \nabla H(u^*) \rangle = 0. \tag{3.29}$$

Similarly, we can prove that  $\lim_{k \rightarrow +\infty} B_{|\cdot|}^{p^*}(d^k, d^*) = 0$ , i.e.,

$$\lim_{k \rightarrow +\infty} |d^k| - |d^*| - \langle d^k - d^*, p^* \rangle = 0, \tag{3.30}$$

and

$$\lim_{k \rightarrow +\infty} \|\mathcal{D}u^{k+1} - d^k\| = 0. \tag{3.31}$$

Since  $|\cdot|$  is continuous, by (3.30) and (3.31), we obtain

$$\lim_{k \rightarrow +\infty} |\mathcal{D}u^k| - |\mathcal{D}u^*| - \langle \mathcal{D}u^k - \mathcal{D}u^*, p^* \rangle = 0. \tag{3.32}$$

Summing this and (3.29), it yields

$$\lim_{k \rightarrow +\infty} (|\mathcal{D}u^k| + H(u^k)) - (|\mathcal{D}u^*| + H(u^*)) - \langle u^k - u^*, \nabla H(u^*) + \mathcal{D}^T p^* \rangle = 0. \tag{3.33}$$

This, together with (3.24), proves (3.22).

Next, we prove (3.23) by assuming that (1.7) has the unique solution  $u^*$ . It is proved by contradiction. Let  $E(u) = |\mathcal{D}u| + H(u)$ . Then  $E(u)$  is a convex, lower semicontinuous function. The remainder of the proof follows the same argument as the proof of (3.3) in the proof of Theorem 3.1.  $\square$



Assume that  $|\cdot|$  is the  $\ell_1$  norm and  $\mathcal{D}$  has a left inverse (e.g.,  $\mathcal{D}$  is a tight frame). Similar to the corresponding discussion in Section 3.1, this theorem implies that the energy  $|\mathcal{D}u^k| + H(u^k)$  of  $u^k$  converges to the minimum value of  $|\mathcal{D}u| + H(u)$ . In other words, the unconstrained split Bregman algorithm (1.11) reaches a solution that has an energy arbitrarily close to the minimum energy with a sufficient number of iterations. This shows that one can get the right solution numerically by (1.11). Furthermore,  $u^k$  has a convergent subsequence, as  $|\mathcal{D}u^k|$  and  $\|u^k\|$  are bounded. Finally, when  $H(u) = \|Au - f\|^2$ , which is commonly used in many applications, and when  $A$  is invertible,  $u^k$  converges to the unique solution of (1.7).

**4. Frame based image restoration.** In this section, we will apply iterations (1.11) and (1.12) to derive algorithms for frame based image restoration, especially image deblurring and image inpainting. Furthermore, image denoising and image decomposition are also discussed as special cases of image inpainting. Frame based image restoration has been studied in [7, 8, 10, 14, 15, 26, 31, 40, 43, 45]. While algorithms in [14, 31, 40, 43, 45] are the model of the synthesis based approach, algorithms in [7, 8, 10, 15, 26] balance the synthesis based and the analysis based approaches.

Our focus here is on the analysis based approach by using split Bregman algorithms. As pointed out in [11, 42] and mentioned before, it is hard to determine which one is better between the analysis based and the synthesis based approaches. It depends on the applications at hand. In general, the synthesis based approach emphasizes the full sparsity of the frame coefficient, and the analysis based approach emphasizes the smoothness of the image since it penalizes the  $\ell_1$  norm of the canonical frame coefficients that normally is linked with the smoothness of solutions. There is also a choice that balances these two approaches as given in [7, 8, 10, 15, 26]. Nevertheless, some simulations for various algorithms based on different models are given. We will see that split Bregman iterations (1.11) and (1.12) provide convenient tools for the analysis based approach of image restoration. Hence, it enriches the discussions in the literature on the analysis based approach for image restorations.

**4.1. Frames and framelets.** Real images usually have sparse approximations under some tight frame systems. Examples of tight frames that can sparsely approximate images are curvelets [16, 17], orthonormal wavelets [37], translation invariant wavelets [34], and framelets [39, 59]. Tight frames are redundant orthogonal bases in  $\mathbb{R}^n$ . The row vectors of  $\mathcal{D} \in \mathbb{R}^{m \times n}$  form a tight frame in an Euclidean space if and only if  $\mathcal{D}^T \mathcal{D} = I$ . The matrix  $\mathcal{D}$  and  $\mathcal{D}^T$  are the analysis operator and the synthesis operator respectively. In general,  $\mathcal{D} \mathcal{D}^T \neq I$ . The redundancy of the tight frame leads to robust signal representations in which partial loss of the data can be tolerated without adverse effects; see, for example, [8, 10, 26, 34, 37, 39]. Since the system  $\mathcal{D}$  is redundant, for a given image  $u$ , there are infinitely many representations  $d$  such that  $u = \mathcal{D}^T d$ . Among them, the one  $\mathcal{D}u$  is called the canonical frame coefficient of  $u$ .

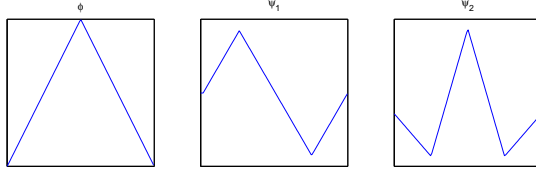
We will use the tight frame system that is derived from the filter banks of framelets (wavelet tight frames) constructed in [39, 59]. In the rest of this subsection, we give a brief introduction to the concept of framelet. To save notations, we use the same lowercase letter to denote a function in  $L^2(\mathbb{R}^2)$  and its discrete version in  $\mathbb{R}^n$ .

A countable set  $X \subset L^2(\mathbb{R})$  is called a tight frame of  $L^2(\mathbb{R})$  if

$$f = \sum_{h \in X} \langle f, h \rangle h \quad \forall f \in L^2(\mathbb{R}), \quad (4.1)$$

where  $\langle \cdot, \cdot \rangle$  is the inner product of  $L^2(\mathbb{R})$ . For given  $\Psi := \{\psi_1, \dots, \psi_r\} \subset L^2(\mathbb{R})$ , the affine (or wavelet) system is defined by the collection of the dilations and the shifts

FIG. 4.1. Piecewise linear framelets.



of  $\Psi$  as

$$X(\Psi) := \{\psi_{\ell,j,k} : 1 \leq \ell \leq r; j, k \in \mathbb{Z}\} \quad \text{with} \quad \psi_{\ell,j,k} := 2^{j/2} \psi_{\ell}(2^j \cdot -k). \quad (4.2)$$

When  $X(\Psi)$  forms a tight frame of  $L^2(\mathbb{R})$ , it is called a tight wavelet frame, and  $\psi_{\ell}$ ,  $\ell = 1, \dots, r$ , are called the (tight) framelets.

To construct a set of framelets, usually, one starts from a compactly supported refinable function  $\phi \in L^2(\mathbb{R})$  (a scaling function) with a refinement mask  $\tau_{\phi}$  satisfying

$$\widehat{\phi}(2\cdot) = \tau_{\phi} \widehat{\phi}.$$

Here  $\widehat{\phi}$  is the Fourier transform of  $\phi$ , and  $\tau_{\phi}$  is a trigonometric polynomial with  $\tau_{\phi}(0) = 1$ ; i.e., a refinement mask of a refinable function must be a lowpass filter. For a given compactly supported refinable function, the construction of a tight framelet system is to find a finite set  $\Psi$  that can be represented in the Fourier domain as

$$\widehat{\psi}(2\cdot) = \tau_{\psi} \widehat{\phi}$$

for some  $2\pi$ -periodic  $\tau_{\psi}$ . The unitary extension principle (UEP) of [59] says that  $X(\Psi)$  in (4.2) generated by  $\Psi$  forms a tight frame in  $L^2(\mathbb{R})$  provided that the masks  $\tau_{\phi}$  and  $\{\tau_{\psi}\}_{\psi \in \Psi}$  satisfy

$$\tau_{\phi}(\omega) \overline{\tau_{\phi}(\omega + \gamma\pi)} + \sum_{\psi \in \Psi} \tau_{\psi}(\omega) \overline{\tau_{\psi}(\omega + \gamma\pi)} = \delta_{\gamma,0}, \quad \gamma = 0, 1, \quad (4.3)$$

for almost all  $\omega$  in  $\mathbb{R}$ . While  $\tau_{\phi}$  corresponds to a lowpass filter,  $\{\tau_{\psi}\}_{\psi \in \Psi}$  must correspond to highpass filters by the UEP. The sequences of Fourier coefficients of  $\tau_{\psi}$ , as well as  $\tau_{\psi}$  itself, are called *framelet masks*. In our implementation, we adopt the piecewise linear B-spline framelet constructed in [59]. The refinement mask is  $\tau_{\phi}(\omega) = \cos^2(\frac{\omega}{2})$ , whose corresponding lowpass filter is  $h_0 = \frac{1}{4}[1, 2, 1]$ . Two framelets are  $\tau_{\psi_1} = -\frac{\sqrt{2}i}{2} \sin(\omega)$  and  $\tau_{\psi_2} = \sin^2(\frac{\omega}{2})$ , whose corresponding highpass filters are

$$h_1 = \frac{\sqrt{2}}{4}[1, 0, -1], \quad h_2 = \frac{1}{4}[-1, 2, -1].$$

The associated refinable function and framelets are given in Figure 4.1. With a one-dimensional framelet system for  $L(\mathbb{R})$ , the two-dimensional framelet system for  $L^2(\mathbb{R}^2)$  can be easily constructed by tensor products of one-dimensional framelets.

In the discrete setting, a discrete image  $f$  is considered as the coefficients  $\{f_i = \langle f, \phi(\cdot - i) \rangle\}$  up to a dilation, where  $\phi$  is the refinable function associated with the framelet system, and  $\langle \cdot, \cdot \rangle$  is the inner product in  $L^2(\mathbb{R}^2)$ . The  $L$ -level discrete framelet

decomposition of  $f$  is then the coefficients  $\{\langle f, 2^{-L/2}\phi(2^{-L} \cdot -j) \rangle\}$  at a prescribed coarsest level  $L$ , and the framelet coefficients are

$$\{\langle f, 2^{-l/2}\psi_i(2^{-l} \cdot -j) \rangle, 1 \leq i \leq r^2 - 1\}$$

for  $0 \leq l \leq L$ . This decomposition can be written into a linear operator applied to the discrete image  $f$ , i.e.,  $\mathcal{D}f$  with  $\mathcal{D} \in \mathbb{R}^{k \times n}$ . By the UEP (4.3),  $\mathcal{D}^T \mathcal{D} = I$ ; thus the row vectors of  $\mathcal{D}$  form a tight frame system in  $\mathbb{R}^n$ . In our implementations, we use a multilevel tight framelet decomposition without downsampling under the Neumann (symmetric) boundary condition. The detailed description can be found in [10, 22].

**4.2. Image deblurring.** In this section, we give some new applications of split Bregman methods to frame based image deblurring by the analysis based approach. In other words, we want to find the underlying image  $u$  from its noisy blurred observation  $f$  in (1.1), where  $A$  is a convolution operator, by solving problems in the form of (1.7) or (1.8).

We assume that we are given a tight frame  $\mathcal{D}$  under which real images have sparse representations. Since the tight frame system  $\mathcal{D}$  is redundant and has a left inverse  $\mathcal{D}^T$ , we can have different approaches: the analysis based and the synthesis based approaches. Although our focus here is on the computational side of the analysis based approach, we will give some simulation results for the synthesis based approach for reference.

**4.2.1. Algorithms for the analysis based approach.** In the analysis based approach, we solve one of the following minimization problems:

$$\min_u \|\mathcal{D}u\|_1 + \frac{\mu}{2} \|Au - f\|^2 \quad (4.4)$$

and

$$\min_u \|\mathcal{D}u\|_1 \quad \text{subject to} \quad \|Au - f\|^2 \leq \sigma^2. \quad (4.5)$$

In other words, we minimize the  $\ell_1$  norm of the canonical coefficient  $\|\mathcal{D}u\|_1$  with reasonable constraints.

Since (4.4) is a special case of (1.7), one can use the unconstrained split Bregman method to solve it. Recall that  $\mathcal{D}^T \mathcal{D} = I$ . Then, the unconstrained alternating split Bregman method (1.11) for solving (4.4) becomes

$$\begin{cases} u^{k+1} = (\mu A^T A + \lambda I)^{-1} (\mu A^T f + \lambda \mathcal{D}^T (d^k - b^k)), \\ d^{k+1} = \mathcal{T}_{\frac{1}{\lambda}}(\mathcal{D}u^{k+1} + b^k), \\ b^{k+1} = b^k + \delta(\mathcal{D}u^{k+1} - d^{k+1}). \end{cases} \quad (4.6)$$

Because  $A$  is a convolution operator, the first step in (4.6) is solved by the fast Fourier transform if the circular boundary condition is used or by the discrete cosine transform if the Neumann boundary condition is used. Also, the second step is an entrywise soft thresholding. Therefore, at each step, the computational cost of (4.6) is  $O(n \log n)$ , where  $n$  is the number of pixels of the image. A preliminary result for algorithm (4.6) was shown in [14].

Since  $|\cdot| = \|\cdot\|_1$  and  $\mathcal{D}$  is a tight frame, (4.4) has a solution. Hence, it follows from Theorem 3.2 that the energy  $\|\mathcal{D}u^k\|_1 + \frac{\mu}{2} \|Au^k - f\|^2$  converges to the corresponding minimum value of the energy functional  $\|\mathcal{D}u\|_1 + \frac{\mu}{2} \|Au - f\|^2$ . Furthermore, the

boundedness of the sequence  $\{u^k\}_{k \in \mathbb{Z}}$  asserts that there exists at least one convergent subsequence of  $\{u^k\}_{k \in \mathbb{Z}}$ . Finally, when  $A$  is invertible, the sequence  $u^k$  converges to the unique solution of (4.5). We further remark that, in many image deblurring applications,  $A$  is usually invertible, although the condition number is large.

To solve (4.5), one can use the Bregman iteration (2.12) for the equality constrained minimization problem (1.8) with an early stopping criterion

$$\|Au^k - f\|^2 \leq \sigma^2 \quad (4.7)$$

to find a good approximate solution of (4.5). This approach has already been used and discussed in, for example, [14, 47, 57, 68], when Bregman or linearized Bregman iterations are used. Since the constrained split Bregman method (1.11) is essentially the Bregman iteration (2.12) with the split Bregman method as an inner solver, we apply (1.11) with the stopping criterion (4.7) to solve (4.5). This leads to the following iteration:

$$\begin{cases} u^{k+1} = (\mu A^T A + \lambda I)^{-1} (\mu A^T (f - c^k) + \lambda \mathcal{D}^T (d^k - b^k)), \\ d^{k+1} = \mathcal{T}_{\frac{1}{\lambda}} (\mathcal{D} u^{k+1} + b^k), \\ b^{k+1} = b^k + \delta_b (\mathcal{D} u^{k+1} - d^{k+1}), \\ c^{k+1} = c^k + \delta_c (A u^{k+1} - f). \end{cases} \quad (4.8)$$

We stop it whenever (4.7) is satisfied. Similar to the discussion for (4.6), the cost of algorithm (4.8) at each step is also  $O(n \log n)$ .

A similar discussion as before leads to the following: (4.5) has a solution; for  $\{u^k\}_{k \in \mathbb{Z}}$  defined in (4.8),  $Au^k$  converges to  $f$ , and the energy  $\|\mathcal{D} u^k\|_1$  converges to the minimum value of the energy functional  $\|\mathcal{D} u\|_1$ ; there is at least one convergent subsequence of  $\{u^k\}_{k \in \mathbb{Z}}$ ; finally, when  $A$  is invertible, the sequence  $u^k$  converges to the unique solution of (4.5).

**4.2.2. Algorithms for the synthesis based approach.** We review some algorithms for the synthesis based approach that are used here. Since the tight frame satisfies  $\mathcal{D}^T \mathcal{D} = I$ ,  $\mathcal{D}^T$  is a synthesis operator. Therefore, the synthesis based approach is to solve one of the following minimization problems:

$$\min_d \|d\|_1 + \frac{\mu}{2} \|A \mathcal{D}^T d - f\|^2 \quad (4.9)$$

and

$$\min_d \|d\|_1 \quad \text{subject to} \quad \|A \mathcal{D}^T d - f\|^2 \leq \sigma^2. \quad (4.10)$$

The solution of (1.1) is taken as  $u = \mathcal{D}^T d$ .

Since (4.9) is a special case of (1.2), one can use (1.4) to solve it. This is essentially the approach in [31, 40]; see also [24, 35, 46, 49]. This leads to the following iterative algorithm:

$$d^{k+1} = \mathcal{T}_{\delta} (d^k - \delta \mu \mathcal{D} A^T (A \mathcal{D}^T d^k - f)), \quad (4.11)$$

where  $\mathcal{T}$  is the soft-thresholding operator defined in (1.5).

For the model (4.10), we use a modified linearized Bregman iteration (1.6) with an early stopping criterion as suggested by [14]. More precisely, given  $d^0 = c^0 = 0$ , we iterate as

$$\begin{cases} c^{k+1} = c^k - \delta \mathcal{D} A^T (A A^T + \lambda I)^{-1} (A \mathcal{D}^T d^k - f), \\ d^{k+1} = \mathcal{T}_{\mu} (c^{k+1}) \end{cases} \quad (4.12)$$

until the error is within the noise level, e.g.,  $\|A\mathcal{D}^T d^k - f\|^2 \leq \sigma^2$ . Here  $(AA^T + \lambda I)^{-1}$  serves as a preconditioner to accelerate the convergence.

**4.2.3. Balanced approaches.** While it depends on applications to determine whether to use the synthesis based or the analysis based approach, there is a third choice that balances these two approaches, and it can be helpful. This approach is initiated in [25–27] and developed in [8–10, 15, 22, 25–27] for various applications. A typical deblurring algorithm in the above references is given as follows:

$$u^{k+1} = \mathcal{D}^T \mathcal{T}_\mu \mathcal{D} (u^k - \delta A^T (Au^k - f)). \quad (4.13)$$

It was proved that the coefficient  $d^k := \mathcal{T}_\delta \mathcal{D} (u^k - \delta \mu A^T (Au^k - f))$  converges to a solution of the minimization problem

$$\min_d \mu \|d\|_1 + \frac{1}{2} \|(I - \mathcal{D}\mathcal{D}^T)d\|^2 + \frac{\delta}{2} \|A\mathcal{D}^T d - f\|^2. \quad (4.14)$$

The second term penalizes the distance of the frame coefficient  $d$  to the range of  $\mathcal{D}$ . This term makes (4.14) balance the synthesis based and the analysis based approaches. To see this, we consider a more general approach, which has been proposed in [7, 15], by introducing a relative weight  $\gamma$  before the second term as follows:

$$\min_d \mu \|d\|_1 + \frac{\gamma}{2} \|(I - \mathcal{D}\mathcal{D}^T)d\|^2 + \frac{\delta}{2} \|A\mathcal{D}^T d - f\|^2. \quad (4.15)$$

Corresponding iterative algorithms that converge to the solution of (4.15) are also developed in [7, 15].

In this model, when  $\gamma = 0$ , (4.15) becomes the synthesis based approach (4.9), while when  $\gamma = \infty$ , the tight frame coefficient  $d$  must be in the range of  $\mathcal{D}$ ; hence (4.15) becomes the analysis based approach. Therefore, (4.15) is between the analysis based and the synthesis based approaches, and so is (4.14) as a special case of (4.15). See [11, 15] for more details.

**4.2.4. Numerical simulations.** Simulation results for the algorithms discussed in previous sections are shown in Figures 4.2 and 4.3. Here the tight frame  $\mathcal{D}$  is generated by the masks of the piecewise linear framelet of [59]. In all the numerical experiments here and hereafter, the iterations are stopped when (4.7) is satisfied or the relative error is smaller than  $5 \times 10^{-4}$  for algorithms solving constrained minimization problems, and they are stopped when  $\frac{\|u^{k+1} - u^k\|}{\|f\|} \leq 10^{-4}$  is met for the other algorithms.

Split Bregman methods for the analysis based approach take only a few steps of iterations to give out good results. Furthermore, since the computational cost of one iteration of the split Bregman methods (4.6) and (4.8) is only  $O(n \log n)$ , split Bregman methods are very efficient algorithms for models (4.4) and (4.5) in the analysis based approach.

It can be seen from Figures 4.2 and 4.3 that the images restored by the analysis based approach are smoother and have fewer artifacts. This is mainly because the minimization is taken in the range of the analysis operator (i.e. among canonical coefficients of the frame system used) in the analysis based approach. The norm of a canonical coefficient is usually linked with the smoothness of the corresponding solution; hence, implicitly, the analysis based approach penalizes the regularity of the underlying image. This is also shown in the restored image based on the balanced

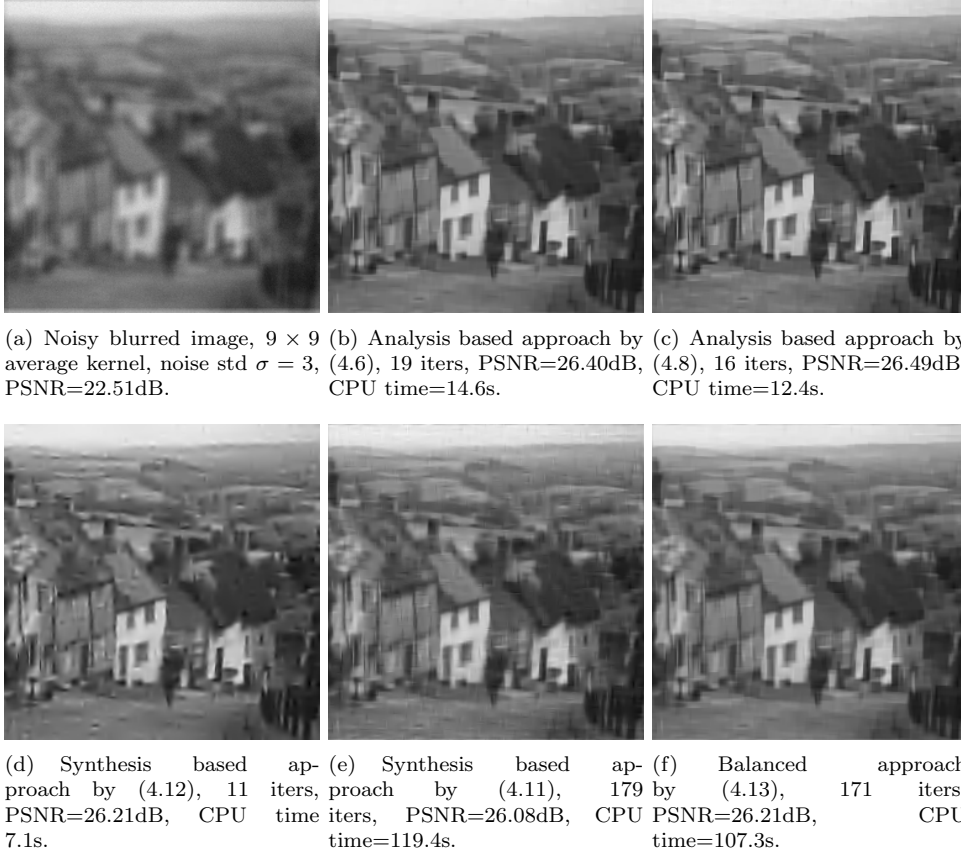


FIG. 4.2. Deblurring results for  $256 \times 256$  Goldhill image.

approach, as there is a penalty term that penalizes the distance to the canonical coefficient. The fact that the balanced methods given smoother restored images has already been observed in, e.g., [8–10, 15, 22, 25–27]. As mentioned before, the algorithm in [14] (i.e. iteration (4.12)) is the synthesis based approach. Since it penalizes the sparsity of the frame coefficient (normally not the canonical coefficient), it lacks the penalty of the smoothness. The artifacts in the restored image were removed by a bilateral filter in [14].

Finally, we note that, among all the algorithms, those involving the Bregman distance (cf. (b)–(d) in Figures 4.2 and 4.3) take many fewer steps than other algorithms. This again shows that iterations based on the Bregman distance (e.g. linearized and splitting Bregman iterations) are particularly efficient for various types of  $\ell_1$  minimization as observed in [47].

**4.3. Image inpainting.** In this section, we apply split Bregman methods to frame based image inpainting. Inpainting refers to problems of filling in the missing part in images. Let  $\Lambda$  be the region of known pixels. Then, we want to recover  $u$  from

$$P_{\Lambda}u = P_{\Lambda}f,$$

where  $P_{\Lambda}$  is a projection, or more precisely, a diagonal matrix with diagonals 1 if the corresponding pixel is known or 0 otherwise. It arises, for example, from removing

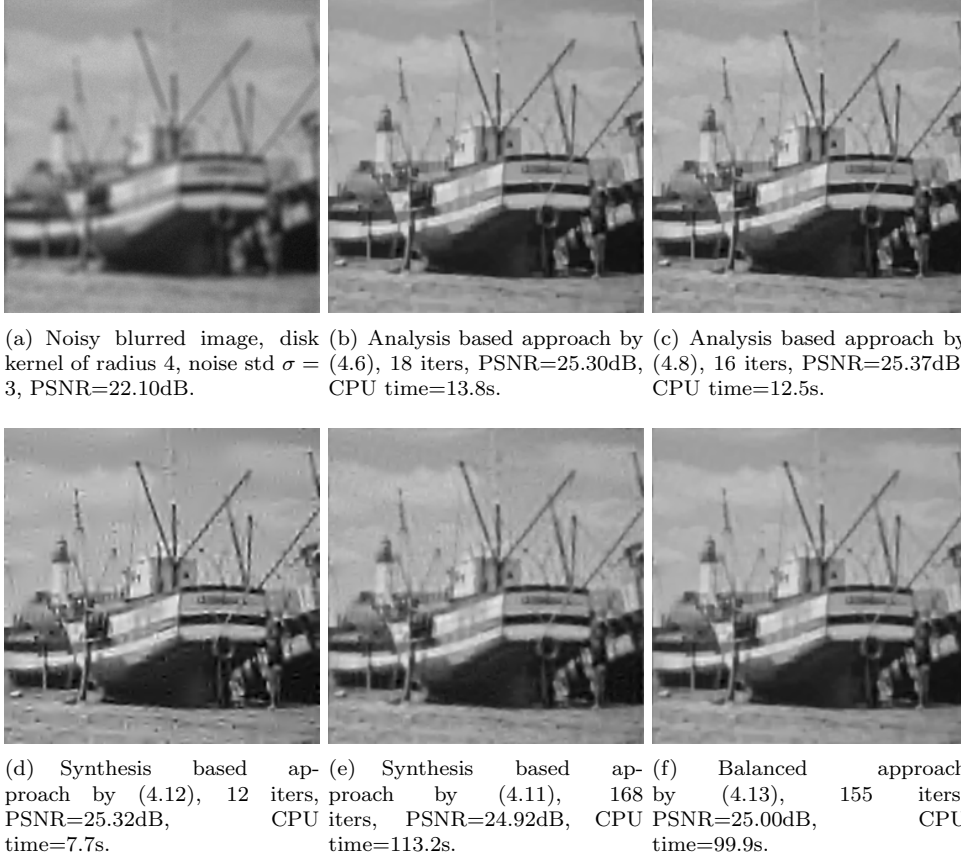


FIG. 4.3. Deblurring results for  $256 \times 256$  Boat image.

scratches in photos, in restoring ancient drawings, and in filling in the missing pixels of images transmitted through a noisy channel. We need to extract information such as edges and textures from the observed data to fill in the missing part such that shapes and patterns are consistent in the human vision. There are basically two kinds of methods for image inpainting. The first one is the variational or PDE based method [2, 29, 30]. The other one is the wavelet based method [10, 43, 48]. Our tight frame falls into the second category. We also remark that many of the variational or PDE based methods can also be understood as analysis based approaches under more general transforms.

Our focus here is on the analysis based approach. Other formulations such as the synthesis based approach or the balanced approach of image inpainting can be found in, e.g., [11, 43]. For the analysis based approach, we solve a constrained minimization problem

$$\min_u \|\mathcal{D}u\|_1 \quad \text{subject to} \quad P_\Lambda u = P_\Lambda f. \quad (4.16)$$

By letting  $A = P_\Lambda$  and  $|\cdot| = \|\cdot\|_1$ , this problem is transferred into a problem in the form of (1.8). By applying (1.12) to image inpainting, we obtain an algorithm for

image inpainting as follows:

$$\begin{cases} u^{k+1} = (\mu P_\Lambda + \lambda I)^{-1} (\mu P_\Lambda (f - c^k) + \lambda \mathcal{D}^T (d^k - b^k)), \\ d^{k+1} = T_{\frac{1}{\lambda}}(\mathcal{D} u^{k+1} + b^k), \\ b^{k+1} = b^k + \delta_b (\mathcal{D} u^{k+1} - d^{k+1}), \\ c^{k+1} = c^k + \delta_c P_\Lambda (u^{k+1} - f). \end{cases} \quad (4.17)$$

Since  $\mu P_\Lambda + \lambda I$  is a diagonal matrix, it is easy to be inverted; hence (4.17) can be easily implemented.

Since  $|\cdot| = \|\cdot\|_1$  and  $\mathcal{D}$  is a tight frame, the previous discussions assert that (4.16) has a solution. By applying Theorem 3.1 and the discussions following it, we conclude that  $P_\Lambda u^k$  converges to  $P_\Lambda f$ . Moreover, the energy  $\|\mathcal{D} u^k\|_1$  converges to the minimum value of the energy functional  $\|\mathcal{D} u\|_1$ , and there exists at least one convergent subsequence of  $\{u^k\}_{k \in \mathbb{Z}}$ .

The simulation result of (4.17) for the image inpainting is shown in Figure 4.4, where  $\mathcal{D}$  is generated by the piecewise linear framelet masks. It takes only 51 iterations to give a very good inpainted image. Notice that, in each step of (4.19), the computational cost is only  $O(n \log n)$ , where  $n$  is the number of pixels. Therefore, (4.17) is a very efficient and effective algorithm for image inpainting.

A simulation result by the algorithm from [44, 45] for the synthesis based approach, which is essentially (4.11) with  $A = P_\Lambda$  and the parameters  $\mu, \delta$  changing step by step, is also given in Figure 4.4. We also show in Figure 4.4 a result by the algorithm for the balanced approach, which is (4.13) with  $A = P_\Lambda$  as proposed in [10]. Here again, the analysis based approach generates a visually smoother restored image.

**4.4. Cartoon-texture image inpainting.** Real images usually have two layers, namely, the cartoon part and the texture part. The former one refers to the piecewise smooth part of images, while the latter one models the oscillating part. These two layers have very different characteristics; see [3, 54, 64]. Hence, we need to use more than one system to represent the image with both cartoon and texture, and adapt the inpainting algorithms to more than one tight frame systems. More precisely, we use two tight frames, say  $\mathcal{D}_1$  and  $\mathcal{D}_2$ , that can represent sparsely the cartoon and texture parts of the image respectively. While it is known that framelets of [39, 59], curvelets of [17], and translation invariant wavelets [34] can represent the cartoon part sparsely, it is hard to say which system can represent the texture part sparsely. However, it is believed that overlapped local discrete cosine transforms (LDCTs) can sparsely approximate the texture part; see, e.g., [43, 63]. Here again, we focus on the analysis based approach, and the other approaches can be found in [11, 43, 63].

The analysis based approach is to solve the following constrained minimization problem:

$$\min_{u_1, u_2} \mu_1 \|\mathcal{D}_1 u_1\|_1 + \mu_2 \|\mathcal{D}_2 u_2\|_1 \quad \text{subject to} \quad P_\Lambda (u_1 + u_2) = P_\Lambda f, \quad (4.18)$$

where  $u_1$  and  $u_2$  are the cartoon and the texture parts of the image respectively. By letting

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad A = \begin{bmatrix} P_\Lambda & P_\Lambda \end{bmatrix}, \quad \mathcal{D} = \begin{bmatrix} \mathcal{D}_1 & 0 \\ 0 & \mathcal{D}_2 \end{bmatrix}, \quad |u| = \mu_1 \|u_1\|_1 + \mu_2 \|u_2\|_1,$$

(4.18) is transferred into a problem in the form of (1.8). By applying (1.12), we obtain





(a) Observed  $256 \times 256$  image.

(b) Analysis based approach by (4.17), 51 iters, PSNR=33.86dB, CPU time=10.1s.



(c) Synthesis based approach by (4.11), 639 iters, PSNR=32.43dB, CPU time=219.6s.

(d) Balanced approach by (4.13), 329 iters, PSNR=33.82dB, CPU time=46.0s.

FIG. 4.4. Image inpainting results.

the following algorithm for simultaneous cartoon and texture image inpainting

$$\begin{cases}
 u_1^{k+1} = \frac{1}{\lambda(\lambda + 2\mu)} P_\Lambda \left( (\lambda\mu P_\Lambda f - \mu(\mu + \lambda)c_1^k + \mu^2 c_2^k) + \lambda(\mu + \lambda) \mathcal{D}_1^T (d_1^k - b_1^k) - \lambda\mu \mathcal{D}_2^T (d_2^k - b_2^k) \right) \\
 \quad + (I - P_\Lambda) \mathcal{D}_1^T (d_1^k - b_1^k), \\
 u_2^{k+1} = \frac{1}{\lambda(\lambda + 2\mu)} P_\Lambda \left( (\lambda\mu P_\Lambda f - \mu(\mu + \lambda)c_2^k + \mu^2 c_1^k) + \lambda(\mu + \lambda) \mathcal{D}_2^T (d_2^k - b_2^k) - \lambda\mu \mathcal{D}_1^T (d_1^k - b_1^k) \right) \\
 \quad + (I - P_\Lambda) \mathcal{D}_2^T (d_2^k - b_2^k), \\
 d_1^{k+1} = \mathcal{T}_{\frac{\mu_1}{\lambda}} (\mathcal{D}_1 u_1^{k+1} + b_1^k), \\
 d_2^{k+1} = \mathcal{T}_{\frac{\mu_2}{\lambda}} (\mathcal{D}_2 u_2^{k+1} + b_2^k), \\
 b_1^{k+1} = b_1^k + \delta_b (\mathcal{D}_1 u_1^{k+1} - d_1^{k+1}), \\
 b_2^{k+1} = b_2^k + \delta_b (\mathcal{D}_2 u_2^{k+1} - d_2^{k+1}), \\
 c_1^{k+1} = c_1^k + \delta_c (P_\Lambda u_1^{k+1} - P_\Lambda f), \\
 c_2^{k+1} = c_2^k + \delta_c (P_\Lambda u_2^{k+1} - P_\Lambda f).
 \end{cases}$$

Similar to the inpainting algorithm with one frame system, we have that  $P_\Lambda(u_1^k + u_2^k)$  converges to  $P_\Lambda f$ . Furthermore, the energy  $\mu_1\|\mathcal{D}_1 u_1^k\|_1 + \mu_2\|\mathcal{D}_2 u_2^k\|_1$  converges to the minimum value of the energy functional  $\mu_1\|\mathcal{D}_1 u_1\|_1 + \mu_2\|\mathcal{D}_2 u_2\|_1$ , and there exists at least one convergent subsequence of  $\{u_1^k, u_2^k\}_{k \in \mathbb{Z}}$ .

In the simulations, the frame system  $\mathcal{D}_1$  is generated by the piecewise linear framelet filter masks and  $\mathcal{D}_2$  is an LDCT. Two simulation results of (4.19) are given in Figures 4.5 and 4.6. It takes fewer than 100 iterations to give a very good inpainted result, where both the cartoon part (e.g. the legs of the table in Figure 4.5) and the texture part (e.g. the clothes of the girl in Figure 4.5) are well preserved. As a by-product, the cartoon part and the texture part of the inpainted image are also shown. We see that the two layers are very well separated. Note that, in each step of (4.19), the computational cost is only  $O(n \log n)$ , where  $n$  is the number of pixels. Therefore, (4.19) is a very efficient and effective algorithm for simultaneous cartoon and texture image inpainting.

In the following, we give two special cases of the simultaneous cartoon and texture inpainting, namely, the cartoon-texture image decomposition and the cartoon-texture denoising. When  $\Lambda$  is the whole domain, i.e.,  $P_\Lambda = I$ , the simultaneous cartoon-texture inpainting algorithm (4.19) becomes a cartoon-texture image decomposition algorithm. The result is shown in Figure 4.7, which again shows that the split Bregman method is very efficient for decomposing images into their cartoon and texture parts.

When  $\Lambda$  is the whole domain, i.e.,  $P_\Lambda = I$ , and there is noise, the model of image inpainting becomes denoising. Correspondingly, the constrained split Bregman algorithm (4.19) with an early stopping criterion becomes a cartoon-texture denoising algorithm.

Finally, we can also use the unconstrained split Bregman algorithm for the analysis based approach for image inpainting, cartoon-texture image decomposition, and image denoising. For simplicity, we next employ the unconstrained split Bregman algorithm to get a denoising scheme. More precisely, we consider to solve the unconstrained minimization algorithm

$$\min_{u_1, u_2} \mu_1\|\mathcal{D}_1 u_1\|_1 + \mu_2\|\mathcal{D}_2 u_2\|_1 + \frac{1}{2}\|u_1 + u_2 - f\|_2^2.$$

Then by choosing

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad A = \begin{bmatrix} I & I \end{bmatrix}, \quad H(u) = \frac{1}{2}\|Au - f\|_2^2, \quad |u| = \mu_1\|u_1\|_1 + \mu_2\|u_2\|_1, \quad \mathcal{D} = \begin{bmatrix} \mathcal{D}_1 & \\ & \mathcal{D}_2 \end{bmatrix}$$

and applying (1.11), we obtain another cartoon-texture denoising algorithm as follows:

$$\begin{cases} u_1^{k+1} = \frac{1}{\lambda(\lambda+2)} ((1+\lambda)\mathcal{D}_1^T(d_1^k - b_1^k) - \lambda\mathcal{D}_2^T(d_2^k - b_2^k) + f), \\ u_2^{k+1} = \frac{1}{\lambda(\lambda+2)} ((1+\lambda)\mathcal{D}_2^T(d_2^k - b_2^k) - \lambda\mathcal{D}_1^T(d_1^k - b_1^k) + f), \\ d_1^{k+1} = \mathcal{T}_{\frac{\mu_1}{\lambda}}(\mathcal{D}_1 u_1^{k+1} + b_1^k), \\ d_2^{k+1} = \mathcal{T}_{\frac{\mu_2}{\lambda}}(\mathcal{D}_2 u_2^{k+1} + b_2^k), \\ b_1^{k+1} = b_1^k + \delta_b(\mathcal{D}_1 u_1^{k+1} - d_1^{k+1}), \\ b_2^{k+1} = b_2^k + \delta_b(\mathcal{D}_2 u_2^{k+1} - d_2^{k+1}). \end{cases} \quad (4.20)$$

Simulation results of the split Bregman methods (4.19) and (4.20) are given in Figure 4.8. Again, this shows that they are very efficient for denoising images while preserving textures.



FIG. 4.5. *Simultaneous cartoon and texture inpainting by the split Bregman method.*

**Acknowledgement.** We would like to thank Dr. Xiaoqun Zhang, from the Department of Mathematics at UCLA, for her comments which improved the presentation of this paper.

#### REFERENCES

- [1] M. BENZI, G. H. GOLUB, AND J. LIESEN, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.
- [2] M. BERTALMÍO, G. SAPIRO, V. CASELLES, AND C. BALLESTER, *Image inpainting*, in Proceedings of the 27th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley, New York, 2000, pp. 417–424.
- [3] M. BERTALMÍO, L. VESE, G. SAPIRO, AND S. OSHER, *Simultaneous structure and texture image inpainting*, IEEE Trans. Image Process., 12 (2003), pp. 882–889.

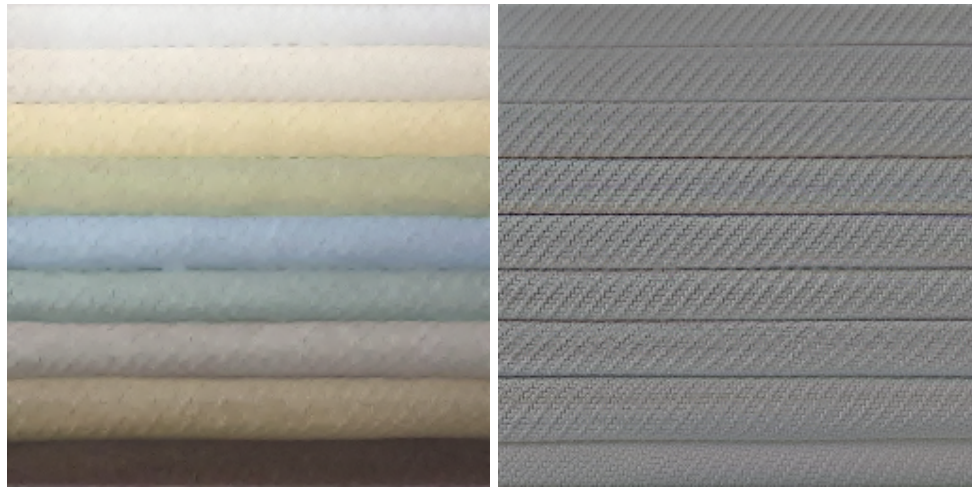
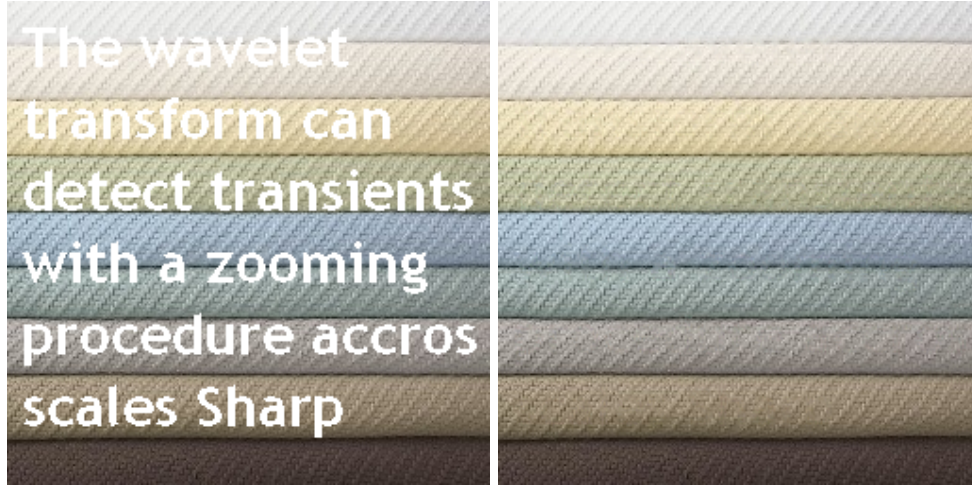


FIG. 4.6. Simultaneous cartoon and texture inpainting by the split Bregman method.

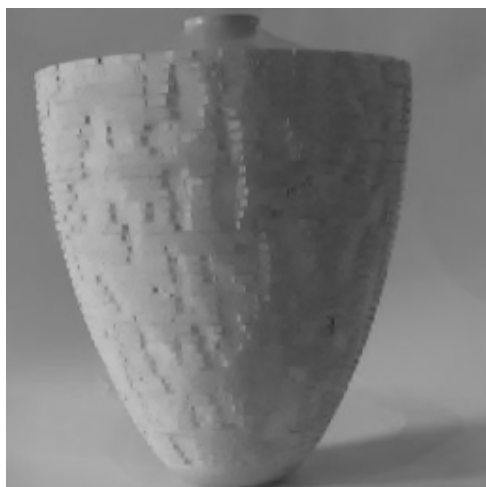
- [4] J. H. BRAMBLE, J. E. PASCIAK, AND A. T. VASSILEV, *Analysis of the inexact Uzawa algorithm for saddle point problems*, SIAM J. Numer. Anal., 34 (1997), pp. 1072–1092.
- [5] J. H. BRAMBLE, J. E. PASCIAK, AND A. T. VASSILEV, *Uzawa type algorithms for nonsymmetric saddle point problems*, Math. Comp., 69 (2000), pp. 667–689.
- [6] L. M. BRÈGMAN, *A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming*, Ž. Vychisl. Mat. i Mat. Fiz., 7 (1967), pp. 620–631.
- [7] J.-F. CAI, E. J. CANDÈS, AND Z. SHEN, *A Singular Value Thresholding Algorithm for Matrix Completion*, arXiv:0810.3286, 2008.
- [8] J.-F. CAI, R. CHAN, L. SHEN, AND Z. SHEN, *Restoration of chopped and noded images by framelets*, SIAM J. Sci. Comput., 30 (2008), pp. 1205–1227.
- [9] J.-F. CAI, R. H. CHAN, L. SHEN, AND Z. SHEN, *Convergence analysis of tight framelet approach for missing data recovery*, Adv. Comput. Math., 31 (2009), pp. 87–113.



(a) Original  $256 \times 256$  image



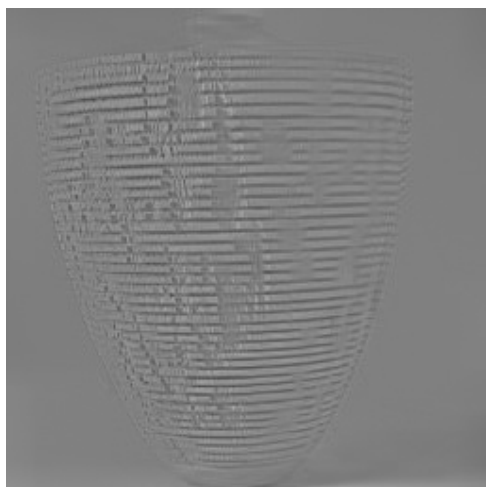
(b) Original  $512 \times 512$  image



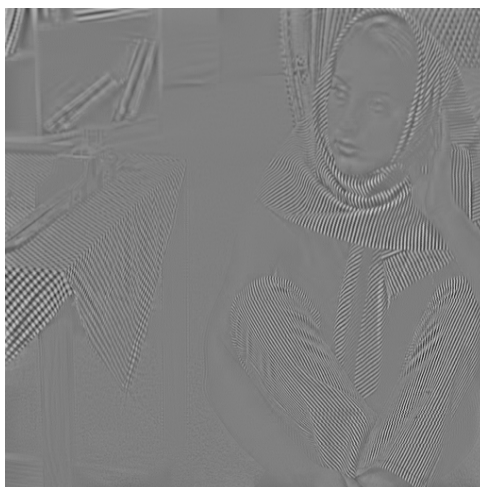
(c) Cartoon part



(d) Cartoon part

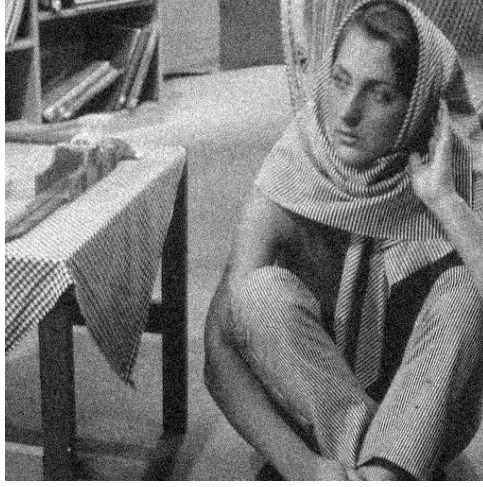


(e) Texture part

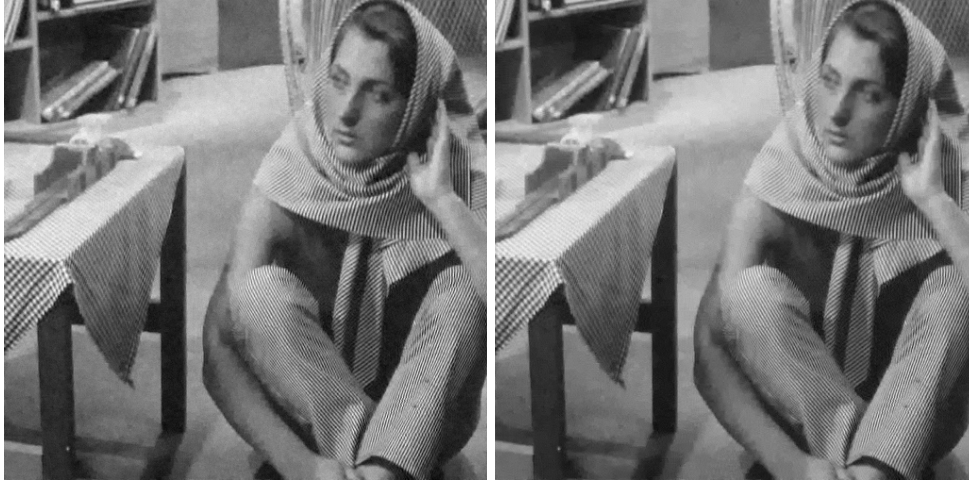


(f) Texture part

FIG. 4.7. Simultaneous cartoon and texture inpainting by the split Bregman method. The number of iterations is 44 for the images in the left column and 64 for those in the right column, and the CPU times are 12.3s, and 69.1s, respectively.



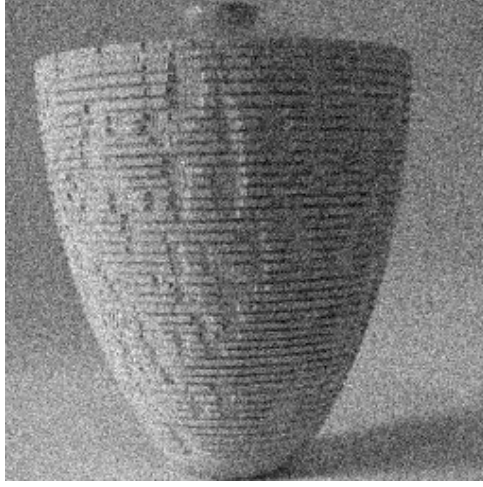
(a) Noisy image,  $\sigma = 20$ , PSNR=22.12dB.



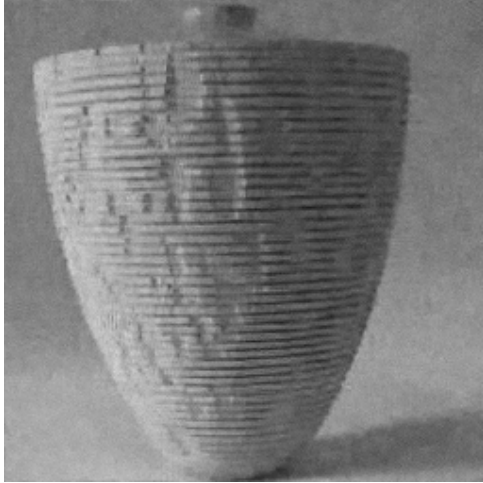
(b) Denoised by (4.19) (67 iters), PSNR=29.25dB, CPU time=72.9s. (c) Denoised by (4.20) (55 iters), PSNR=29.01dB, CPU time=71.6s.

FIG. 4.8. Simultaneous cartoon and texture denoising by the constrained and the unconstrained split Bregman methods (4.19) and (4.20).

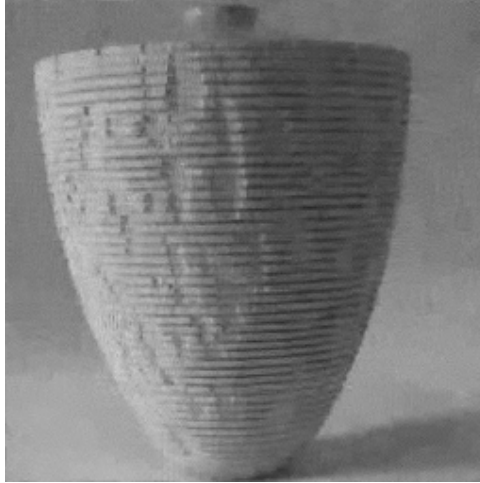
- [10] J.-F. CAI, R. H. CHAN, AND Z. SHEN, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal., 24 (2008), pp. 131–149.
- [11] J.-F. CAI, R. H. CHAN, AND Z. SHEN, *Simultaneous Cartoon and Texture Inpainting*, 2008. preprint.
- [12] J.-F. CAI, S. OSHER, AND Z. SHEN, *Convergence of the Linearized Bregman Iteration for  $\ell_1$ -norm Minimization*, Math. Comput., 78 (2009), pp. 2127–2136.
- [13] J.-F. CAI, S. OSHER, AND Z. SHEN, *Linearized Bregman iterations for compressed sensing*, Math. Comput., 78 (2009), pp. 1515–1536.
- [14] J.-F. CAI, S. OSHER, AND Z. SHEN, *Linearized Bregman Iterations for Frame-Based Image Deblurring*, SIAM J. Imaging Sci., 2 (2009), pp. 226–252.
- [15] J.-F. CAI AND Z. SHEN, *Framelet Based Deconvolution*, J. Comput. Math., to appear.
- [16] E. J. CANDÈS AND D. L. DONOHO, *Curvelets – a surprisingly effective nonadaptive representation for objects with edges*, in Curves and Surfaces, Vanderbilt University Press, Nashville,



(a) Noisy image,  $\sigma = 20$ , PSNR=22.11dB.



(b) Denoised by (4.19) (74 iters), PSNR=30.11dB, CPU time=21.1s.



(c) Denoised by (4.20) (53 iters), PSNR=30.22dB, CPU time=18.1s.

FIG. 4.9. Simultaneous cartoon and texture denoising by the constrained and the unconstrained split Bregman methods (4.19) and (4.20).

- TN, 2000, pp. 105–120.
- [17] E. J. CANDÈS AND D. L. DONOHO, *New tight frames of curvelets and optimal representations of objects with piecewise  $C^2$  singularities*, Comm. Pure Appl. Math., 57 (2004), pp. 219–266.
  - [18] E. J. CANDÈS AND B. RECHT, *Exact Matrix Completion via Convex Optimization*, 2008.
  - [19] E. J. CANDÈS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, 52 (2006), pp. 489–509.
  - [20] Z.-H. CAO, *Fast Uzawa algorithm for generalized saddle point problems*, Appl. Numer. Math., 46 (2003), pp. 157–171.
  - [21] A. E. ÇETIN, *Reconstruction of signals from Fourier transform samples*, Signal Process., 16 (1989), pp. 129–148.
  - [22] A. CHAI AND Z. SHEN, *Deconvolution: A wavelet frame approach*, Numer. Math., 106 (2007), pp. 529–587.



- [23] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97.
- [24] A. CHAMBOLLE, R. A. DEVORE, N.-Y. LEE, AND B. J. LUCIER, *Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Trans. Image Process., 7 (1998), pp. 319–335.
- [25] 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.
- [26] R. H. CHAN, S. D. RIEMENSCHNEIDER, L. SHEN, AND Z. SHEN, *Tight frame: an efficient way for high-resolution image reconstruction*, Appl. Comput. Harmon. Anal., 17 (2004), pp. 91–115.
- [27] R. H. CHAN, Z. SHEN, AND T. XIA, *A framelet algorithm for enhancing video stills*, Appl. Comput. Harmon. Anal., 23 (2007), pp. 153–170.
- [28] T. F. CHAN, G. H. GOLUB, AND P. MULET, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput., 20 (1999), pp. 1964–1977 (electronic).
- [29] T. F. CHAN AND J. SHEN, *Mathematical models for local nontexture inpaintings*, SIAM J. Appl. Math., 62 (2001/02), pp. 1019–1043 (electronic).
- [30] T. F. CHAN AND J. SHEN, *Image processing and analysis*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. Variational, PDE, wavelet, and stochastic methods.
- [31] C. CHAUX, P. L. COMBETTES, J.-C. PESQUET, AND V. R. WAJS, *A variational formulation for frame-based inverse problems*, Inverse Problems, 23 (2007), pp. 1495–1518.
- [32] S. S. CHEN, D. L. DONOHO, AND M. A. SAUNDERS, *Atomic decomposition by basis pursuit*, SIAM Rev., 43 (2001), pp. 129–159 (electronic). Reprinted from SIAM J. Sci. Comput. **20** (1998), no. 1, 33–61 (electronic).
- [33] X. CHEN, *Global and superlinear convergence of inexact Uzawa methods for saddle point problems with nondifferentiable mappings*, SIAM J. Numer. Anal., 35 (1998), pp. 1130–1148 (electronic).
- [34] R. R. COIFMAN AND D. L. DONOHO, *Translation-invariant de-noising*, in Wavelets and Statistics, A. Antoniadis and G. Oppenheim, eds., vol. 103 of Lecture Notes in Statistics, New York, 1995, Springer-Verlag, pp. 125–150.
- [35] P. L. COMBETTES AND V. R. WAJS, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., 4 (2005), pp. 1168–1200 (electronic).
- [36] J. DARBON AND S. OSHER, *Fast discrete optimization for sparse approximations and deconvolutions*, 2007. preprint.
- [37] I. DAUBECHIES, *Ten lectures on wavelets*, vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- [38] I. DAUBECHIES, M. DEFRISE, AND C. DE MOL, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math., 57 (2004), pp. 1413–1457.
- [39] 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.
- [40] I. DAUBECHIES, G. TESCHKE, AND L. VESE, *Iteratively solving linear inverse problems under general convex constraints*, Inverse Probl. Imaging, 1 (2007), pp. 29–46.
- [41] D. L. DONOHO, *Compressed sensing*, IEEE Trans. Inform. Theory, 52 (2006), pp. 1289–1306.
- [42] M. ELAD, P. MILANFAR, AND R. RUBINSTEIN, *Analysis versus synthesis in signal priors*, Inverse Problems, 23 (2007), pp. 947–968.
- [43] M. ELAD, J.-L. STARCK, P. QUERRE, AND D. L. DONOHO, *Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)*, Appl. Comput. Harmon. Anal., 19 (2005), pp. 340–358.
- [44] M. FADILI AND J.-L. STARCK, *Sparse representations and bayesian image inpainting*, in Proc. SPARS’05, Vol. I, Rennes, France, 2005.
- [45] M. FADILI, J.-L. STARCK, AND F. MURTAGH, *Inpainting and zooming using sparse representations*, The Computer Journal, 52 (2009), pp. 64–79.
- [46] M. A. T. FIGUEIREDO AND R. D. NOWAK, *An EM algorithm for wavelet-based image restoration*, IEEE Trans. Image Process., 12 (2003), pp. 906–916.
- [47] T. GOLDSTEIN AND S. OSHER, *The Split Bregman Algorithm for L1 Regularized Problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.
- [48] 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.
- [49] E. HALE, W. YIN, AND Y. ZHANG, *Fixed-point continuation for l1-minimization: methodology*



- and convergence, SIAM J. Optim., 19 (2008), pp. 1107–1130.
- [50] Q. HU AND J. ZOU, *Two new variants of nonlinear inexact Uzawa algorithms for saddle-point problems*, Numer. Math., 93 (2002), pp. 333–359.
  - [51] Q. HU AND J. ZOU, *Nonlinear inexact Uzawa algorithms for linear and nonlinear saddle-point problems*, SIAM J. Optim., 16 (2006), pp. 798–825 (electronic).
  - [52] R.-Q. JIA, H. ZHAO, AND W. ZHAO, *Convergence analysis of the Bregman method for the variational model of image denoising*, Appl. Comput. Harmon. Anal., (27) 2009, pp. 367–379.
  - [53] M. LI, B. HAO, AND X. FENG, *Iterative regularization and nonlinear inverse scale space based on translation invariant wavelet shrinkage*, Int. J. Wavelets Multiresolut. Inf. Process., 6 (2008), pp. 83–95.
  - [54] Y. MEYER, *Oscillating patterns in image processing and nonlinear evolution equations*, vol. 22 of University Lecture Series, American Mathematical Society, Providence, RI, 2001. The fifteenth Dean Jacqueline B. Lewis memorial lectures.
  - [55] M. NIKOLOVA, *Local strong homogeneity of a regularized estimator*, SIAM J. Appl. Math., 61 (2000), pp. 633–658 (electronic).
  - [56] M. NIKOLOVA, *Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers*, SIAM J. Numer. Anal., 40 (2002), pp. 965–994 (electronic).
  - [57] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, AND W. YIN, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489 (electronic).
  - [58] S. OSHER, Y. MAO, B. DONG, AND W. YIN, *Fast linearized Bregman iteration for compressed sensing and sparse denoising*, Commu. Math. Sci., to appear.
  - [59] A. RON AND Z. SHEN, *Affine systems in  $L_2(\mathbf{R}^d)$ : the analysis of the analysis operator*, J. Funct. Anal., 148 (1997), pp. 408–447.
  - [60] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
  - [61] S. SETZER, *Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage*, in Proceedings of SSVM '09, Lecture Notes in Comput. Sci. 5567, Springer, Berlin, 2009, pp. 464–476.
  - [62] S. SETZER AND G. STEIDL, *Split Bregman method, gradient descent reprojction method and Parseval frames*, 2008. preprint.
  - [63] J.-L. STARCK, M. ELAD, AND D. L. DONOHO, *Image decomposition via the combination of sparse representations and a variational approach*, IEEE Trans. Image Process., 14 (2005), pp. 1570–1582.
  - [64] L. A. VESE AND S. J. OSHER, *Modeling textures with total variation minimization and oscillating patterns in image processing*, J. Sci. Comput., 19 (2003), pp. 553–572. Special issue in honor of the sixtieth birthday of Stanley Osher.
  - [65] C. VOGEL AND M. OMAN, *Fast, robust total variation-based reconstruction of noisy, blurred images*, IEEE Trans. Image Process., 7 (1998), pp. 813–824.
  - [66] Y. WANG, J. YANG, W. YIN, AND Y. ZHANG, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM J. Imaging Sci., 1 (2008), pp. 248–272.
  - [67] J. XU AND S. J. OSHER, *Iterative regularization and nonlinear inverse scale space applied to wavelet-based denoising*, IEEE Trans. Image Process., 16 (2007), pp. 534–544.
  - [68] W. YIN, S. OSHER, D. GOLDFARB, AND J. DARBON, *Bregman iterative algorithms for  $\ell_1$ -minimization with applications to compressed sensing*, SIAM J. Imaging Sci., 1 (2008), pp. 143–168.
  - [69] X. ZHANG, M. BURGER, X. BRESSON, AND S. OSHER, *Bregmanized Nonlocal Regularization for Deconvolution and Sparse Reconstruction*, 2009. UCLA CAM Report (09-03).