

THE LINEARIZED SPLIT BREGMAN ITERATIVE ALGORITHM AND ITS CONVERGENCE ANALYSIS FOR ROBUST TOMOGRAPHIC IMAGE RECONSTRUCTION [†]

CHONG CHEN [‡] AND GUOLIANG XU [§]

Abstract. The split Bregman iteration method has been successfully applied to solve a variety of specially ℓ_1 -regularized problems, such as image denoising, image segmentation, magnetic resonance imaging reconstruction, and so forth. There exists, however, a large-scale and unstructured linear system demanding to be resolved in each iteration for the robust image reconstruction problem in sparse-view X-ray computed tomography. Motivated by this, we propose a new linearized split Bregman iterative algorithm, which is constructed by means of the thoughts of split and linearized Bregman methods. Remarkably, our method can be generalized to efficiently resolve the robust compressed sensing problem, as well as the total variation- ℓ_1 and ℓ_1 - ℓ_1 minimization problems. We also give a rigorous proof for the convergence of the proposed method under appropriate condition, in which the convergence does not depend on the selection of the regularization parameter. Along with the proof idea, we further prove the convergence of the gradient-descent-based split Bregman iteration, yet which relies on the value of the regularization parameter. Hence, our proposed method is more flexible and stable than the split Bregman method. Experimental results demonstrate that our algorithm has better performance in terms of reconstruction quality, effectiveness and robustness, compared with some other frequently used methods for the robust image reconstruction in sparse-view X-ray computed tomography.

Key words. Linearized split Bregman iteration, split Bregman iteration, linearized Bregman iteration, ℓ_1 regularization, robust image reconstruction, sparse-view X-ray computed tomography.

AMS subject classifications. 65K05, 65R32, 92C55

1. Introduction

In modern commercial X-ray computed tomography (CT) imaging, the projections are generally detected at a large number (more than 300) of views from the patient for high-quality image reconstruction by currently analytic-based algorithms, for instance, filtered back-projection (FBP) method [17, 39, 37, 24]. However, data collection from so large number of views, meaning high doses of radiation, likely result in radiation-induced cancer with high probability [16]. It is of practical interest and meaningful, therefore, to develop imaging techniques that are capable of producing high-quality reconstructions from sparse-view projection data. In recent years, the optimization- or variation-based image reconstruction models and the algorithms of interest have been investigated heavily owing to their ability to allow reducing scanning views while maintaining or improving reconstructed image quality for robust image reconstruction problem arising in sparse-view CT [13, 35, 31, 10, 24, 34]. The rationale behind these studies is considerably similar to compressed sensing (CS), though the non-random projection matrix of CT scanner may well be not satisfying

*The work of C. Chen was supported in part by National Natural Science Foundation of China (NSFC) under the grant 11301520. The work of G. Xu was supported in part by NSFC key project under the grant 10990013, NSFC under the grants 11101401, 81173663 and NSFC Funds for Creative Research Groups of China under the grant 11021101.

[†]Received date / Revised version date

[‡]State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100190, P.R. China, (chench@lsec.cc.ac.cn).

[§]State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100190, P.R. China, (xuguo@lsec.cc.ac.cn).

the restricted isometry property which is required to accurately restore the original signal via the CS theory [14, 4, 5, 3].

In a general formulation of the CT reconstruction problem, the reconstructed image can be denoted as a weighted sum of the shifted basis functions (usually pixel/voxel basis) [12]. The discrete projection data $g \in \mathbb{R}^M$ is related to these weights via a linear equations as follows

$$Pf \approx g \quad \text{or} \quad g = Pf + \varepsilon, \quad (1.1)$$

where $f \in \mathbb{R}^N$ is represented as the vector of weights, $P \in \mathbb{R}^{M \times N}$ is the projection matrix for modeling the X-ray transform and ε is the interference component. Due to the data inconsistencies, introduced by the approximation to the coefficient elements and the interference of the noise, error, and some other factors, (1.1) is often used to simulate the CT-imaging process. In addition, we should note that both f and P depend on the selection of the shifted basis functions to expand the image function and the way to discrete the the X-ray transform. In this article, we choose the pixel/voxel functions as the basis, and use the intersection length of the X ray with the image pixel/voxel as the elements of the projection matrix.

Considering the image reconstruction in the sparse-view X-ray CT, we definitely suffer from the severely under-sampling for $M < N$, in many cases, $M \ll N$. In such a case, the study on the optimization- or variation-based image reconstruction models and the induced algorithms has begun to be in fashion. It is well-known that the iterative tomographic reconstruction algorithms, for solving the optimizing or variational models, are able to yield high-quality images from the limited data. That is because they can take into account the *prior* knowledge about the solution to regularize the under-determined problem. Often in medical and other applications, the tomographic images are relatively piecewise-constant over the extended volumes, for instance within an organ. Rapid variation in the image may only occur at boundaries of internal structures. Thus an image itself might not be sparse, yet the image formed by taking the magnitude of its gradient can be approximately sparse [33, 35, 31]. Hence, the total variation (TV) based optimizing or variational image reconstruction models were introduced (See [12, 9, 13, 35, 42, 10] and the references therein).

In this paper, we mainly concern the following model for such a problem

$$\min_f J(f), \quad \text{subject to} \quad \|Pf - g\|_2 \leq \sigma,$$

where $J(f) = \|\nabla f\|_1$ denotes the TV semi-norm. The above constrained optimization is equivalent to the following unconstrained optimization problem

$$\min_f J(f) + \frac{\lambda}{2} \|Pf - g\|_2^2.$$

A brief proof on their equivalency can be found in Section 3. To solve above reconstruction models efficiently, a variety of algorithms have been proposed, such as gradient descend method [33, 29], improved gradient descend method [38], adaptive-steepest-descent projection onto convex sets method [35, 36], split Bregman iteration [22], gradient-flow-based finite element methods [42, 10, 11], etc.

Although the split Bregman iteration is extremely efficient, it just can fast solve the problems that have certain special structure in matrix P , for instance, TV denoising and segmentation (P is an identity matrix), magnetic resonance imaging reconstruction (P comprises a subset of the rows of the Fourier transform matrix), and

so forth [22, 21]. For CT image reconstruction, the projection matrix P has not been figured out the particular structure so as to fast deal with the inversion matrix of $P^T P$ in each iteration. The details about this problem is given in Section 2.

In this work, we present a novel algorithm, called linearized split Bregman (LSB) iteration, that can be well suited to efficiently solving the sparse-view CT image reconstruction problem. Notably, the idea we provide can be generalized to solve a variety of optimization or variation models, which are considerably significant in practice and are hard to resolve by the existing methods. For instance, the robust CS problem

$$\min_f \|f\|_1 + \frac{\lambda}{2} \|Pf - g\|_2^2,$$

where P is a sensing matrix [6]. Recently, there exist many efficiently algorithms to solve this model, for instance, fixed-point continuation [23], alternative direction method (ADM) (See [19, 20, 44] and the references therein). Furthermore, our proposed method is able to solve the models that use the ℓ_1 -norm as the measure of fidelity [27, 8, 18, 41], such as the TV- ℓ_1 minimization problem

$$\min_f \|\nabla f\|_1 + \frac{\lambda}{2} \|Pf - g\|_1,$$

and the ℓ_1 - ℓ_1 minimization problem

$$\min_f \|f\|_1 + \frac{\lambda}{2} \|Pf - g\|_1.$$

We also give a rigorous proof on the convergence of the proposed method under appropriate condition for each extension, in which the convergence does not depend on the selection of the regularization parameter.

As a previous work, Cai *et al.* proved the convergence of the unconstrained/constrained split Bregman iterations in [2], while their proofs are merely concerning the iterative schemes based on solving each subproblem exactly. Admittedly, it is almost impossible to precisely or directly resolve the general large-scale linear system in related subproblem. It means that the practically iterative algorithms of interest must be applied, for instance, the gradient descent method, conjugate gradient method, Gauss-Siedel method, etc. Along with our proof idea, we further prove the convergence of the gradient-decent-based split Bregman (GDSB) method.

The rest of the paper is organized as follows. In Section 2, we first briefly review the classical Bregman methods and point out the existing difficulties of them in solving the robust CT image reconstruction problem. We then propose an LSB iterative algorithm and generalize it to solve a variety of ℓ_1 -optimization problems in Section 3. In Section 4, we give the proof for the convergence of the proposed method, and further analyze the convergence of the GDSB iteration. In Section 5, we apply our algorithm to the robust image reconstruction in sparse-view X-ray CT imaging and make several numerical comparisons with some other methods. Finally, we conclude this paper in Section 6.

2. Classical Bregman iteration algorithms

In this section, we first consider the general optimization problem of a convex function on the n -dimensional (nD) Euclidean space \mathbb{R}^n . Let $J: \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function. A vector $p \in \mathbb{R}^n$ is called a subgradient of J at a point $v \in \mathbb{R}^n$ if

$$J(u) - J(v) - \langle p, u - v \rangle \geq 0, \quad \forall u \in \mathbb{R}^n. \quad (2.1)$$

The subdifferential $\partial J(v)$ is the set of all subgradients of J at v [32]. The Bregman distance associated with J at v is defined as

$$D_J^q(u, v) = J(u) - J(v) - \langle q, u - v \rangle, \quad (2.2)$$

where p is a subgradient of J at v , namely, $p \in \partial J(v)$. Obviously, $D_J^q(u, v) \geq 0$, and $D_J^q(u, v) \geq D_J^q(w, v)$ for any $w \in \mathbb{R}^n$ on the line segment between u and v . Note that $D_J^q(u, v) \neq D_J^q(v, u)$ in general, namely, the symmetry is not met so that $D_J^q(u, v)$ is not a distance in the usual sense [1, 28].

2.1. Bregman iteration algorithm

Considering the following optimization problem

$$\min_f J(f), \quad \text{subject to } Pf = g, \quad (2.3)$$

where if $J(f) = \|\nabla f\|_1$, this model is often applied to image deblurring and tomographic image reconstruction; if $J(f) = \|f\|_1$, that is frequently solved in basis pursuit problem or compressed sensing [28, 40, 22, 46]. Notice that the observational data g are not corrupted by noise in (2.3).

The Bregman iteration was first applied in image processing in [28], and then used to solve the ℓ_1 -minimization problem with application to compressed sensing, which is equivalent to the well-known augmented Lagrangian method [46, 20]. We summarize the Bregman iteration for solving the minimization problem (2.3) as follows

$$f^{k+1} = \arg \min_f D_J^{q^k}(f, f^k) + \frac{\lambda}{2} \|Pf - g\|_2^2, \quad (2.4)$$

$$q^{k+1} = q^k - \lambda P^T (Pf^{k+1} - g). \quad (2.5)$$

As shown in [46, 25], the above seemingly complicated Bregman iteration is equivalent to the following simplified iterative scheme

$$f^{k+1} = \arg \min_f J(f) + \frac{\lambda}{2} \|Pf - b^k\|_2^2, \quad (2.6)$$

$$b^{k+1} = b^k + g - Pf^{k+1}. \quad (2.7)$$

The main advantages of the Bregman iteration include that the convergent rate is very fast when it is employed to certain types of optimization problems and the value of the penalty factor λ is not required to tend to infinity during computation.

2.2. Split Bregman iteration algorithm

When $J(f) = \|\nabla f\|_1$, the unconstrained optimization subproblem (2.6) has no closed-form solution and is difficult to solve since the cost function contains the ℓ_1 and ℓ_2 portions simultaneously. To overcome this problem, Goldstein and Osher presented an algorithm called split Bregman method to solve this optimization problem effectively [22]. The key idea of their algorithm is that they decoupled the ℓ_1 and ℓ_2 terms of the cost function in (2.6). The splitting thought is also proposed in [40] with application to ℓ_1 -regularized deconvolution. Hence, they considered the following minimization problem

$$\min_{f, d} \|d\|_1 + \frac{\lambda}{2} \|Pf - g\|_2^2, \quad \text{subject to } d = \nabla f. \quad (2.8)$$

In what follows the split Bregman iteration algorithm to solve above problem is presented.

$$(f^{k+1}, d^{k+1}) = \arg \min_{f, d} \|d\|_1 + \frac{\lambda}{2} \|Pf - g\|_2^2 + \frac{\mu}{2} \|d - \nabla f - s^k\|_2^2, \quad (2.9)$$

$$s^{k+1} = s^k + \nabla f^{k+1} - d^{k+1}. \quad (2.10)$$

Note that the subproblem (2.9) should be precisely resolved. Hence a lot of inner iterations are needed. The concrete implementation, with the inner iteration number fixed to be 1, is given as the following Algorithm 2.1.

Algorithm 2.1. *Split Bregman Iteration Algorithm*

Step 1. Given initial points $f^0 = 0$, $d^0 = 0$, $s^0 = 0$, $0 < \epsilon \ll 1$ and an integer $K > 0$. Set $k := 0$.

Step 2. Update d :

$$d^{k+1} = \arg \min_d \|d\|_1 + \frac{\mu}{2} \|d - \nabla f^k - s^k\|_2^2. \quad (2.11)$$

Step 3. Update f :

$$f^{k+1} = \arg \min_f \frac{\lambda}{2} \|Pf - g\|_2^2 + \frac{\mu}{2} \|d^{k+1} - \nabla f - s^k\|_2^2. \quad (2.12)$$

Compute $r_k = \|f^{k+1} - f^k\|_2$. If $r_k < \epsilon$ or $k + 1 > K$, stop the iteration, otherwise go to the next step.

Step 4. Update s :

$$s^{k+1} = s^k + \nabla f^{k+1} - d^{k+1}.$$

Step 5. Set $k := k + 1$, return to Step 2.

Notably, the Step 2 and Step 3 should be precisely resolved to guarantee the convergence of the algorithm (See [2]). Fortunately, one can explicitly figure out the minimum d^{k+1} in Step 2 by using the generalized shrinkage formula as follows

$$d^{k+1} = \max\left(h^k - \frac{1}{\mu}, 0\right) \frac{\nabla f^k + s^k}{h^k}, \quad (2.13)$$

where

$$h^k = \|\nabla f^k + s^k\|_2. \quad (2.14)$$

For the purpose of solving the minimization problem in Step 3 of Algorithm 2.1, one can easily obtain the first-order optimality condition

$$(\lambda P^T P + \mu \nabla^T \nabla) f^{k+1} = \lambda P^T g + \mu \nabla^T (d^{k+1} - s^k). \quad (2.15)$$

Obviously, to solve f^{k+1} , the inversion of $\lambda P^T P + \mu \nabla^T \nabla$ needs to be computed, or a number of iterations require to be implemented for solving the large-scale linear system. Thus this step is time-consuming and one can hardly obtain the exact solution. In practice, we are able to apply the gradient descend, conjugate gradient or Newton-type methods, etc. to solve (2.12). For instance, by gradient descend method, we have

$$f^{k+1} = f^k - \alpha (\lambda P^T (Pf^k - g) + \mu \nabla^T (\nabla f^k - d^{k+1} + s^k)). \quad (2.16)$$

In addition, (2.11) is equivalent to

$$q_d^{k+1} + \mu(d^{k+1} - \nabla f^k - s^k) = 0, \quad q_d^{k+1} \in \partial J(d^{k+1}). \quad (2.17)$$

Hence, summarizing the GDSB iterative algorithm leads to

$$\begin{cases} q_d^{k+1} + \mu(d^{k+1} - \nabla f^k - s^k) = 0, \\ f^{k+1} = f^k - \alpha(\lambda P^T(Pf^k - g) + \mu \nabla^T(\nabla f^k - d^{k+1} + s^k)), \\ s^{k+1} = s^k + \nabla f^{k+1} - d^{k+1}. \end{cases} \quad (2.18)$$

2.3. Linearized Bregman iteration algorithm

The linearized Bregman iteration algorithm is used for solving the image deblurring or CS problem

$$\min_f J(f), \quad \text{subject to } Pf = g, \quad (2.19)$$

where $J(f) = \|\nabla f\|_1$ or $\|f\|_1$, the matrix P is a projection operator or sensing matrix, and g is a known measurements without noise, respectively [46, 30]. The procedure of linearized Bregman iteration consists of

$$f^{k+1} = \arg \min_f D_J^{q^k}(f, f^k) + \frac{\lambda}{2\alpha} \|f - (f^k - \alpha P^T(Pf^k - g))\|_2^2, \quad (2.20)$$

$$q^{k+1} = q^k + \frac{\lambda}{\alpha} (f^{k+1} - (f^k - \alpha P^T(Pf^k - g))), \quad (2.21)$$

where α is positive and can be thought as the step size in each iteration.

In order to solve (2.20), we give its equivalent form as follows

$$f^{k+1} = \arg \min_f J(f) + \frac{\lambda}{2\alpha} \|f - (f^k + \alpha(\frac{1}{\lambda} q^k - P^T(Pf^k - g)))\|_2^2. \quad (2.22)$$

Supposing $J(f) = \|f\|_1$, we can easily use the shrinkage operator to obtain the explicit solution of (2.22) as (2.11). However, assuming $J(f) = \|\nabla f\|_1$, as a result of the special formulation and nondifferential of $\|\nabla f\|_1$, we cannot obtain the closed formula of f^{k+1} from (2.22). Hence, the linearized Bregman iteration has limitations to solve the problems appeared in tomographic image reconstruction.

3. Linearized split Bregman iteration algorithm

In practice, the observational data are often contaminated by noise. Therefore, the general model for treating problems appeared in robust tomographic image reconstruction is presented as the following constrained optimization problem

$$\min_f J(f), \quad \text{subject to } \|Pf - g\|_2 \leq \sigma, \quad (3.1)$$

or the equivalently unconstrained optimization problem

$$\min_f J(f) + \frac{\lambda}{2} \|Pf - g\|_2^2, \quad (3.2)$$

where $J(f) = \|\nabla f\|_1$, σ is the standard deviation of the noise in observational data g which is blind in usual situation, the positive value λ stands for the regularization parameter which is utilized to trade off the fitting term and the regularized term.

Actually, the aforementioned two optimization problems are equivalent under appropriate condition. This result can be immediately given by the following lemma.

LEMMA 3.1. *Let $\lambda \geq 0$ and f_λ the minimum of (3.2). Then f_λ is also the minimum of (3.1) with $\sigma = \|Pf_\lambda - g\|_2$.*

Proof. Since f_λ is the minimum of model (3.2), for any f satisfying the constraint in model (3.1) we have

$$J(f_\lambda) + \frac{\lambda}{2} \|Pf_\lambda - g\|_2^2 \leq J(f) + \frac{\lambda}{2} \|Pf - g\|_2^2. \quad (3.3)$$

Then by rearranging the terms in (3.3), we obtain

$$J(f_\lambda) + \frac{\lambda}{2} (\|Pf_\lambda - g\|_2^2 - \|Pf - g\|_2^2) \leq J(f). \quad (3.4)$$

Using

$$\sigma = \|Pf_\lambda - g\|_2, \quad (3.5)$$

and $\sigma \geq \|Pf - g\|_2$, we obtain the following formula immediately

$$J(f_\lambda) \leq J(f). \quad (3.6)$$

Hence, f_λ is also the minimum of optimization problem (3.1). \square

Assume that the optimization problems (3.1) and (3.2) both have unique solutions, by Lemma 3.1, we can know that if $\sigma = \|Pf_\lambda - g\|_2$ is satisfied, the optimization problems (3.1) and (3.2) are equivalent. Moreover, because the algorithm solving for the constrained optimization problem (3.1) is more difficult than that for the unconstrained optimization problem (3.2), so we can solve (3.2) in stead of (3.1). However, the parameter λ is impossible to figure out since the inverse function of (3.5) is in calculable. An important problem is how to choose the proper regularization parameter λ in an effective way. Here the methods of how to choose regularization parameter is out of the scope of our discussion. The attention of the remaining part in this section is concentrated on developing efficient algorithm for solving (3.2).

Motivated by the limitations of split Bregman method to solve the ℓ_1 -regularized problem arising in tomographic image reconstruction, we present an innovative computational method called LSB algorithm.

Rather than considering the minimization problem (3.2) with $J(f) = \|\nabla f\|_1$, we treat the other optimization problem

$$\min_{f,d,b} \|d\|_1 + \frac{\lambda}{2} \|b\|_2^2, \quad \text{subject to } d = \nabla f, \quad b = Pf - g. \quad (3.7)$$

Problem (3.7) is clearly equivalent to problem (3.2). Here is the distinction between our method and the classical split Bregman method. In addition, the natural idea of splitting two terms has been found in [15, 45].

For simplicity, let

$$E(f, d, b) = \|d\|_1 + \frac{\lambda}{2} \|b\|_2^2. \quad (3.8)$$

Clearly, $E(f, d, b)$ is separable with respect to d , b and f . Let

$$E(f, d, b) = E_1(f) + E_2(d) + E_3(b), \quad (3.9)$$

$$D_E^{\mathbf{q}}(f, \tilde{f}, d, \tilde{d}, b, \tilde{b}) = D_{E_1}^{\mathbf{q}_f}(f, \tilde{f}) + D_{E_2}^{\mathbf{q}_d}(d, \tilde{d}) + D_{E_3}^{\mathbf{q}_b}(b, \tilde{b}), \quad (3.10)$$

where $\mathbf{q} = (q_f, q_d, q_b)$ and

$$E_1(f) \equiv 0, \quad E_2(d) = \|d\|_1, \quad E_3(b) = \frac{\lambda}{2} \|b\|_2^2. \quad (3.11)$$

Let $F^k = (f^k, d^k, b^k)$ and $\mathbf{q}^k = (q_f^k, q_d^k, q_b^k)$. In order to penalize the equality constraints, we solve the optimization problem (3.2) by the Bregman iterative algorithm as follows.

$$F^{k+1} = \arg \min_{f, d, b} D_E^{\mathbf{q}^k}(f, f^k, d, d^k, b, b^k) + \frac{\beta_1}{2} \|d - \nabla f\|_2^2 + \frac{\beta_2}{2} \|b - (Pf - g)\|_2^2, \quad (3.12)$$

$$q_f^{k+1} = q_f^k - \beta_1 \nabla^T (\nabla f^{k+1} - d^{k+1}) - \beta_2 P^T (Pf^{k+1} - g - b^{k+1}), \quad (3.13)$$

$$q_d^{k+1} = q_d^k - \beta_1 (d^{k+1} - \nabla f^{k+1}), \quad (3.14)$$

$$q_b^{k+1} = q_b^k - \beta_2 (b^{k+1} - Pf^{k+1} + g). \quad (3.15)$$

To simplify the formulation (3.12), we have to consider f^{k+1} , d^{k+1} and b^{k+1} respectively. By (3.7)–(3.12), we compute

$$\begin{aligned} b^{k+1} &= \arg \min_b D_{E_3}^{q_b^k}(b, b^k) + \frac{\beta_2}{2} \|b - Pf^k + g\|_2^2 \\ &= \frac{q_b^k + \beta_2 Pf^k - \beta_2 g}{\lambda + \beta_2}. \end{aligned} \quad (3.16)$$

Also according to (3.7)–(3.12), we obtain

$$\begin{aligned} d^{k+1} &= \arg \min_d D_{E_2}^{q_d^k}(d, d^k) + \frac{\beta_1}{2} \|d - \nabla f^k\|_2^2 \\ &= \max\left(h^k - \frac{1}{\beta_1}, 0\right) \frac{\nabla f^k + \frac{q_d^k}{\beta_1}}{h^k}, \end{aligned} \quad (3.17)$$

where $h^k = \|\nabla f^k + \frac{q_d^k}{\beta_1}\|_2$.

We now focus on the update of f^{k+1} and q_f^{k+1} . The subproblems (3.12) and (3.13) for solving f^{k+1} and q_f^{k+1} are equivalent to the following Bregman iteration.

$$f^{k+1} = \arg \min_f D_{E_1}^{q_f^k}(f, f^k) + \frac{\beta_1}{2} \|d^{k+1} - \nabla f\|_2^2 + \frac{\beta_2}{2} \|b^{k+1} - (Pf - g)\|_2^2, \quad (3.18)$$

$$q_f^{k+1} = q_f^k - \beta_1 \nabla^T (\nabla f^{k+1} - d^{k+1}) - \beta_2 P^T (Pf^{k+1} - g - b^{k+1}). \quad (3.19)$$

As a result of $E_1(f) \equiv 0$ by (3.11), we obtain $\partial E_1(f) \equiv 0$ so that

$$q_f^k \equiv 0. \quad (3.20)$$

Hence through simple calculations, it is easy to find that the minimization problem (3.18) and (3.19) are equivalent. Namely, they are both introducing

$$(\beta_1 \nabla^T \nabla + \beta_2 P^T P) f^{k+1} = \beta_1 \nabla^T d^{k+1} + \beta_2 P^T b^{k+1} + \beta_2 P^T g. \quad (3.21)$$

In general, the matrix $\beta_1 \nabla^T \nabla + \beta_2 P^T P$ is considerably large-scale and dense in tomographic image reconstruction. It is, hence, very time-consuming and likely to be unstable to compute f^{k+1} by (3.21). In order to solve the subproblems for updating

f^{k+1} and q_f^{k+1} efficiently, we can convert the Bregman iteration (3.18)–(3.19) into the linearized Bregman iteration as follows.

$$f^{k+1} = \operatorname{argmin}_f \left[D_{E_1}^{q_f^k}(f, f^k) + \frac{\beta_1}{2\alpha_1} \|f - (f^k - \alpha_1 \nabla^T(\nabla f^k - d^{k+1}))\|_2^2 + \frac{\beta_2}{2\alpha_2} \|f - (f^k - \alpha_2 P^T(P f^k - g - b^{k+1}))\|_2^2 \right], \quad (3.22)$$

$$q_f^{k+1} = q_f^k - \frac{\beta_1}{\alpha_1} (f^{k+1} - (f^k - \alpha_1 \nabla^T(\nabla f^k - d^{k+1}))) - \frac{\beta_2}{\alpha_2} (f^{k+1} - (f^k - \alpha_2 P^T(P f^k - g - b^{k+1}))). \quad (3.23)$$

Using the properties of $E_1(f)$ and q_f^k in (3.11) and (3.20), we can combine (3.22)–(3.23) into just one step of updating f^{k+1} as

$$f^{k+1} = f^k - \frac{\beta_1 \nabla^T(\nabla f^k - d^{k+1}) + \beta_2 P^T(P f^k - g - b^{k+1})}{\omega_1 + \omega_2}, \quad (3.24)$$

where $\omega_1 = \frac{\beta_1}{\alpha_1}$ and $\omega_2 = \frac{\beta_2}{\alpha_2}$. Note that the parameters ω_1 and ω_2 of interest can be combined into one parameter.

To sum up, in what follows we conclude our LSB iteration for Robust Tomographic Image Reconstruction with $J(f) = \|\nabla f\|_1$.

Algorithm 3.1. *LSB Iterative Algorithm*

Step 1. Given initial points $f^0 = 0$, $d^0 = 0$, $b^0 = 0$, $q^0 = 0$, $0 < \epsilon \ll 1$ and an integer $K > 0$. Set $k := 0$.

Step 2. Update b :

$$b^{k+1} = \frac{q_b^k + \beta_2 P f^k - \beta_2 g}{\lambda + \beta_2}.$$

Step 3. Update d :

$$d^{k+1} = \max\left(h^k - \frac{1}{\beta_1}, 0\right) \frac{\nabla f^k + \frac{q_d^k}{\beta_1}}{h^k},$$

where $h^k = \|\nabla f^k + \frac{q_d^k}{\beta_1}\|_2$.

Step 4. Update f :

$$f^{k+1} = f^k - \frac{\beta_1 \nabla^T(\nabla f^k - d^{k+1}) + \beta_2 P^T(P f^k - g - b^{k+1})}{\omega_1 + \omega_2}.$$

Compute $r_k = \|f^{k+1} - f^k\|_2$. If $r_k < \epsilon$ or $k+1 > K$, stop the iteration, otherwise go to the next step.

Step 5. Update q_d :

$$q_d^{k+1} = q_d^k - \beta_1 (d^{k+1} - \nabla f^{k+1}).$$

Step 6. Update q_b :

$$q_b^{k+1} = q_b^k - \beta_2 (b^{k+1} - P f^{k+1} + g).$$

Step 7. Set $k := k+1$, return to Step 2.

REMARK 3.2. We can also directly apply linearized Bregman iteration to (3.12)–(3.15) with replacing the implementation just to (3.18)–(3.19) and then obtain the following iterative scheme

$$\begin{aligned}
F^{k+1} &= \operatorname{argmin}_{f,d,b} \left[D_E^{q^k}(f, f^k, d, d^k, b, b^k) + \frac{\beta_1}{2} \|d - \nabla f^k\|_2^2 + \frac{\beta_2}{2} \|b - (Pf^k - g)\|_2^2 \right. \\
&\quad + \frac{\beta_1}{2\alpha_1} \|f - (f^k + \alpha_1(\Delta f^k - \operatorname{div} d^k))\|_2^2 \\
&\quad \left. + \frac{\beta_2}{2\alpha_2} \|f - (f^k - \alpha_2 P^T(Pf^k - g - b^k))\|_2^2 \right], \\
q_f^{k+1} &= q_f^k - \frac{\beta_1}{\alpha_1} (f^{k+1} - (f^k + \alpha_1(\Delta f^k - \operatorname{div} d^k))) \\
&\quad - \frac{\beta_2}{\alpha_2} (f^{k+1} - (f^k - \alpha_2 P^T(Pf^k - g - b^k))), \\
q_d^{k+1} &= q_d^k - \beta_1 (d^{k+1} - \nabla f^k), \\
q_b^{k+1} &= q_b^k - \beta_2 (b^{k+1} - Pf^k + g).
\end{aligned}$$

It is easy to find that the updating of b^{k+1} and d^{k+1} is the same as that in Algorithm 3.1, nevertheless, the updating of f^{k+1} , q_d^{k+1} and q_b^{k+1} is deferent. Since the latest b^{k+1} and d^{k+1} are not used in updating f^{k+1} , and the latest f^{k+1} is not used to update q_d^{k+1} and q_b^{k+1} .

REMARK 3.3. We cannot directly apply the similarly linearized idea onto (2.12) in split Bregman iteration, i.e., Algorithm 2.1. Since the term $\|d^{k+1} - \nabla f - s^k\|_2^2$ can be linearized with respect to f . Nevertheless, the term $\|Pf - g\|_2^2$ cannot be linearized, for which is not concerning an equality constraint as the above term. This is why we split two terms in our algorithm.

REMARK 3.4. The linearization and fixed point techniques have been successfully applied to various inverse problems, including robust CS problem, nuclear-norm minimization, matrix rank minimization, constrained linear least-squares problem and so forth [23, 26, 7, 43]. We further generalize these techniques to deal with more complicated problems. The proposed method can simultaneously address two split terms in the robust tomographic image reconstruction, which is the main difference with the above cases.

REMARK 3.5. The computational cost of Algorithm 3.1 comes mainly from Step 3, in particular, the matrix-vector multiplication $P^T P f$. Suppose that the dimension of f is $N (= n^c, c=2 \text{ or } 3)$, the number of projection views is m and each view has $O(n^{c-1})$ bins. Correspondingly, the projection matrix P is sparse and each row of P has $O(n)$ non-zero components at most. Hence the computational complexity of $P f$ is $O(mn^c)$. Furthermore, the computational cost of $P^T P f$ is $O(mn^c)$. To summarize, the computational complexity of each iteration of Algorithm 3.1 is $O(mn^c)$.

3.1. Equivalent form

The LSB iteration can be reformulated into an equivalent form. As in Algorithm

3.1, let $\mathbf{q}^0 = 0$, then by (3.13–3.15) we immediately have

$$q_f^{k+1} = -\sum_{i=1}^{k+1} (\beta_1 \nabla^T (\nabla f^i - d^i) + \beta_2 P^T (P f^i - g - b^i)), \quad (3.25)$$

$$q_d^{k+1} = -\beta_1 \sum_{i=1}^{k+1} (d^i - \nabla f^i), \quad (3.26)$$

$$q_b^{k+1} = -\beta_2 \sum_{i=1}^{k+1} (b^i - P f^i + g). \quad (3.27)$$

For ease of description, let

$$s^k = \sum_{i=1}^k (d^i - \nabla f^i), \quad (3.28)$$

$$t^k = \sum_{i=1}^k (b^i - P f^i + g), \quad (3.29)$$

then we obtain

$$q_f^k = \beta_1 \nabla^T s^k + \beta_2 P^T t^k, \quad (3.30)$$

$$q_d^k = -\beta_1 s^k, \quad (3.31)$$

$$q_b^k = -\beta_2 t^k. \quad (3.32)$$

Because $E_1(f) \equiv 0$, we have $q_f^k \equiv 0$, namely,

$$\beta_1 \nabla^T s^k + \beta_2 P^T t^k \equiv 0. \quad (3.33)$$

Replacing the q_f^k, q_d^k, q_b^k in (3.12) by (3.30)–(3.32), the equivalent form of (3.12)–(3.15) can be written as

$$\begin{aligned} F^{k+1} &= \arg \min_{f,d,b} E_1(f) + E_2(d) + E_3(b) \\ &\quad + \frac{\beta_1}{2} \|d - \nabla f + s^k\|_2^2 + \frac{\beta_2}{2} \|b - (P f - g) + t^k\|_2^2, \end{aligned} \quad (3.34)$$

$$s^{k+1} = s^k + (d^{k+1} - \nabla f^{k+1}), \quad (3.35)$$

$$t^{k+1} = t^k + (b^{k+1} - P f^{k+1} + g). \quad (3.36)$$

To solve the optimization problem (3.34), we can apply alternatively the strategy as follows. Like the procedure of our algorithm, we first update b :

$$\begin{aligned} b^{k+1} &= \arg \min_b E_3(b) + \frac{\beta_2}{2} \|b - (P f^k - g) + t^k\|_2^2 \\ &= \frac{-\beta_2 t^k + \beta_2 P f^k - \beta_2 g}{\lambda + \beta_2}. \end{aligned} \quad (3.37)$$

Then, update d :

$$\begin{aligned} d^{k+1} &= \arg \min_d E_2(d) + \frac{\beta_1}{2} \|d - \nabla f^k + s^k\|_2^2 \\ &= \max(h^k - \frac{1}{\beta_1}, 0) \frac{\nabla f^k - s^k}{h^k}, \end{aligned} \quad (3.38)$$

where $h^k = \|\nabla f^k - s^k\|_2$. Finally, update f by solving the following minimization subproblem

$$f^{k+1} = \arg \min_f \frac{\beta_1}{2} \|d^{k+1} - \nabla f + s^k\|_2^2 + \frac{\beta_2}{2} \|b^{k+1} - (Pf - g) + t^k\|_2^2, \quad (3.39)$$

for $E_1(f) \equiv 0$. We also use the linearization as our algorithm, namely,

$$f^{k+1} = \arg \min_f \left[\frac{\beta_1}{2\alpha_1} \|f - (f^k - \alpha_1 \nabla^T (\nabla f^k - d^{k+1} - s^k))\|_2^2 + \frac{\beta_2}{2\alpha_2} \|f - (f^k - \alpha_2 P^T (Pf^k - g - b^{k+1} - t^k))\|_2^2 \right]. \quad (3.40)$$

Hence by the fact (3.33), in what follows we obtain the updating of f

$$f^{k+1} = f^k - \frac{\beta_1 \nabla^T (\nabla f^k - d^{k+1}) + \beta_2 P^T (Pf^k - g - b^{k+1})}{\omega_1 + \omega_2}. \quad (3.41)$$

Therefore, we convert our algorithm into the following equivalent form.

$$\begin{cases} b^{k+1} = \frac{-\beta_2 t^k + \beta_2 Pf^k - \beta_2 g}{\lambda + \beta_2}, \\ d^{k+1} = \max(h^k - \frac{1}{\beta_1}, 0) \frac{\nabla f^k - s^k}{h^k}, \\ f^{k+1} = f^k - \frac{\nabla^T (\nabla f^k - d^{k+1}) + \beta_2 P^T (Pf^k - g - b^{k+1})}{\omega_1 + \omega_2}, \\ s^{k+1} = s^k + (d^{k+1} - \nabla f^{k+1}), \\ t^{k+1} = t^k + (b^{k+1} - Pf^{k+1} + g). \end{cases} \quad (3.42)$$

3.2. Extensions

On the other hand, we will consider the issue of robust CS with $J(f) = \|f\|_1$, in the constrained minimization problem (3.1). By Lemma 3.1, we can treat the following equivalent problem

$$\min_{f,b} \|f\|_1 + \frac{\lambda}{2} \|b\|_2^2, \quad \text{subject to } b = Pf - g. \quad (3.43)$$

For simplicity, let

$$\hat{E}(f, b) = \|f\|_1 + \frac{\lambda}{2} \|b\|_2^2. \quad (3.44)$$

Also, the $\hat{E}(f, b)$ is separable with respect to f and b . Let

$$\hat{E}(f, b) = \hat{E}_1(f) + \hat{E}_2(b), \quad (3.45)$$

$$D_{\hat{E}}^{\hat{\mathbf{q}}^k}(f, f^k, b, b^k) = D_{\hat{E}_1}^{\hat{q}_f^k}(f, f^k) + D_{\hat{E}_2}^{\hat{q}_b^k}(b, b^k), \quad (3.46)$$

where $\hat{\mathbf{q}}^k = (\hat{q}_f^k, \hat{q}_b^k)$, $\hat{q}_f^k \in \partial \hat{E}_1(f^k)$, $\hat{q}_b^k \in \partial \hat{E}_2(b^k)$ and

$$\hat{E}_1(f) = \|f\|_1, \quad \hat{E}_2(b) = \frac{\lambda}{2} \|b\|_2^2. \quad (3.47)$$

By Bregman iteration, (3.43) can be resolved by

$$\min_{f,b} D_{\hat{E}}^{\hat{q}^k}(f, f^k, b, b^k) + \frac{\beta}{2} \|b - (Pf - g)\|_2^2, \quad (3.48)$$

$$\hat{q}_f^{k+1} = \hat{q}_f^k - \beta P^T (Pf^{k+1} - g - b^{k+1}), \quad (3.49)$$

$$\hat{q}_b^{k+1} = \hat{q}_b^k - \beta (b^{k+1} - Pf^{k+1} + g). \quad (3.50)$$

In order to efficiently resolve the subproblem involving f , we solve the optimization problem by our LSB iterative algorithm as follows.

Algorithm 3.2. *LSB Iterative Algorithm for Robust Compressed Sensing*

Step 1. Given initial points $f^0 = 0$, $b^0 = 0$, $\hat{q}^0 = 0$, $0 < \epsilon \ll 1$ and an integer $K > 0$. Set $k := 0$.

Step 2. Update b :

$$b^{k+1} = \frac{\hat{q}_b^k + \beta Pf^k - \beta g}{\lambda + \beta}. \quad (3.51)$$

Step 3. Update f :

$$f^{k+1} = \arg \min_f D_{\hat{E}_1}^{\hat{q}_f^k}(f, f^k) + \frac{\beta}{2\alpha} \|f - (f^k - \alpha P^T (Pf^k - g - b^{k+1}))\|_2^2. \quad (3.52)$$

Compute $r_k = \|f^{k+1} - f^k\|_2$. If $r_k < \epsilon$ or $k+1 > K$, stop the iteration, otherwise go to the next step.

Step 4. Update \hat{q}_f :

$$\hat{q}_f^{k+1} = \hat{q}_f^k - \frac{\beta}{\alpha} (f^{k+1} - (f^k - \alpha P^T (Pf^k - g - b^{k+1}))). \quad (3.53)$$

Step 5. Update \hat{q}_b :

$$\hat{q}_b^{k+1} = \hat{q}_b^k - \beta (b^{k+1} - Pf^{k+1} + g). \quad (3.54)$$

Step 6. Set $k := k+1$, return to Step 2.

Note that the minimization problem (3.52) in Step 3 of Algorithm 3.2 is easy to solve by the shrinkage or soft-threshold operator. As in Subsection 3.1, we can obtain the equivalent form of Algorithm 3.2 as follows.

$$\begin{cases} \hat{q}_b^{k+1} + \beta (b^{k+1} - Pf^k + g - t^k) = 0, \\ \hat{q}_f^{k+1} + \frac{\beta}{\alpha} (f^{k+1} - f^k + \alpha P^T (Pf^k - g - b^{k+1} + t^k)) = 0, \\ t^{k+1} = t^k + (Pf^{k+1} - g - b^{k+1}), \end{cases} \quad (3.55)$$

where $t^k = -\sum_{i=1}^k (b^i - Pf^i + g) = \hat{q}_b^k / \beta$ and $\hat{q}_f^k = -\beta P^T t^k$.

We can further extend our proposed method to effectively solve the following problem

$$\min_f J(f), \quad \text{subject to } \|Pf - g\|_1 \leq \sigma, \quad (3.56)$$

where $J(f) = \|\nabla f\|_1$ or $\|f\|_1$. Model (3.56) can be used to solve the problem in robust tomographic image reconstruction or robust compressed sensing, respectively. The corresponding unconstrained optimization problem is

$$\min_f J(f) + \lambda \|Pf - g\|_1, \quad (3.57)$$

which can be asserted by the following lemma.

LEMMA 3.6. *Let $\lambda \geq 0$ and f_λ the minimum of (3.57). Thus f_λ is also the minimum of (3.56) with $\sigma = \|Pf_\lambda - g\|_1$.*

Proof. The proof is the same as that of Lemma 3.1. \square

If $J(f) = \|f\|_1$, (3.56) is an ℓ_1 - ℓ_1 problem and equivalent to

$$\min_{f,b} \|f\|_1 + \lambda \|b\|_1, \quad \text{subject to } b = Pf - g. \quad (3.58)$$

Redefining $\hat{E}_2(b) = \lambda \|b\|_1$, we can resolve (3.58) along with the thought of Algorithm 3.2, where the difference is just the updating of b . Here we need to use the shrinkage operator to update b . The equivalent iterative scheme is the same as (3.55).

On the other hand, if $J(f) = \|\nabla f\|_1$, (3.56) is a TV- ℓ_1 problem, and equivalent to

$$\min_{f,b,d} \|d\|_1 + \lambda \|b\|_1, \quad \text{subject to } d = \nabla f, \quad b = Pf - g. \quad (3.59)$$

Assume $E_3(b) = \lambda \|b\|_1$. Besides the first step of updating b by the shrinkage operator, the following steps are the same as those in Algorithm 3.1. The equivalent iterative scheme is the same as (4.1) in the next section.

4. Convergence analysis

In this section, we study the convergence of the proposed algorithms as applied to the robust tomographic image reconstruction and the other extensional cases. In addition, the convergence analysis of the GDSB method will be involved.

4.1. Convergence of the linearized split Bregman iterative algorithm

In this subsection, we give the convergence analysis of the LSB iterative algorithm for the robust tomographic image reconstruction and the other extensional cases. It is well-known that if the convergence of the equivalent form is proved, then we obtain the convergence of the originally iterative algorithm. Hence we first consider the convergence of the iterative scheme in (3.42). The proof idea is inspired by [2].

Admittedly, the first and second formulas of (3.42) can be rewritten as

$$\begin{aligned} q_b^{k+1} + \beta_2(b^{k+1} - (Pf^k - g) + t^k) &= 0, & q_b^{k+1} &\in \partial E_3(b^{k+1}), \\ q_d^{k+1} + \beta_1(d^{k+1} - \nabla f^k + s^k) &= 0, & q_d^{k+1} &\in \partial E_2(d^{k+1}). \end{aligned}$$

Then (3.42) can be translated into

$$\begin{cases} q_b^{k+1} + \beta_2(b^{k+1} - (Pf^k - g) + t^k) = 0, \\ q_d^{k+1} + \beta_1(d^{k+1} - \nabla f^k + s^k) = 0, \\ f^{k+1} = f^k - \frac{\beta_1 \nabla^T (\nabla f^k - d^{k+1}) + \beta_2 P^T (Pf^k - g - b^{k+1})}{\omega_1 + \omega_2}, \\ s^{k+1} = s^k + (d^{k+1} - \nabla f^{k+1}), \\ t^{k+1} = t^k + (b^{k+1} - Pf^{k+1} + g). \end{cases} \quad (4.1)$$

The following theorem gives the convergence of the LSB iteration under the appropriate condition.

THEOREM 4.1. *Let $J(f) = \|\nabla f\|_1$. Suppose that there exists a unique solution f^* of (3.2), and $\beta_1 > 0$, $\beta_2 > 0$, $\alpha > 0$, $I - (\alpha\beta_1 \nabla^T \nabla + \alpha\beta_2 P^T P)$ is positive semi-definite*

where I is an identity operator. Then we have the convergence results for the iterative scheme (4.1), namely,

$$\lim_{k \rightarrow +\infty} (\|\nabla f^k\|_1 + \frac{\lambda}{2} \|Pf^k - g\|_2^2) = \|\nabla f^*\|_1 + \frac{\lambda}{2} \|Pf^* - g\|_2^2, \quad (4.2)$$

and

$$\lim_{k \rightarrow +\infty} \|f^k - f^*\|_2 = 0, \quad (4.3)$$

where $\alpha = 1/(\omega_1 + \omega_2)$.

Proof. By the assumption that f^* is the unique solution of (3.2), we have the following first-order optimality condition

$$\nabla^T q_d^* + P^T q_b^* = 0, \quad (4.4)$$

where $q_d^* \in \partial E_2(d^*)$ with $d^* = \nabla f^*$ and $q_b^* \in \partial E_3(b^*)$ with $b^* = Pf^* - g$.

Then we introduce $s^* = -\frac{q_d^*}{\beta_1}$ and $t^* = -\frac{q_b^*}{\beta_2}$. Obviously,

$$\beta_1 \nabla^T s^* + \beta_2 P^T t^* = 0. \quad (4.5)$$

Next it is also easy to obtain the following equations

$$\begin{cases} q_b^* + \beta_2(b^* - (Pf^* - g) + t^*) = 0, \\ q_d^* + \beta_1(d^* - \nabla f^* + s^*) = 0, \\ f^* = f^* - \frac{\beta_1 \nabla^T (\nabla f^* - d^*) + \beta_2 P^T (Pf^* - g - b^*)}{\omega_1 + \omega_2}, \\ s^* = s^* + (d^* - \nabla f^*), \\ t^* = t^* + (b^* - Pf^* + g). \end{cases} \quad (4.6)$$

Let

$$\begin{aligned} f_e^k &= f^k - f^*, & t_e^k &= t^k - t^*, & s_e^k &= s^k - s^*, \\ d_e^k &= d^k - d^*, & q_{d_e}^k &= q_d^k - q_d^*, \\ b_e^k &= b^k - b^*, & q_{b_e}^k &= q_b^k - q_b^*. \end{aligned}$$

Subtracting the third equation of (4.6) from the third equation of (4.1), and then multiplying by $(f_e^{k+1} - f_e^k)$ on the two sides of the result, we obtain

$$\begin{aligned} & -\|f_e^{k+1} - f_e^k\|_2^2 + \alpha \beta_1 \|\nabla f_e^{k+1} - \nabla f_e^k\|_2^2 + \alpha \beta_2 \|Pf_e^{k+1} - Pf_e^k\|_2^2 \\ & = \alpha \beta_1 \langle \nabla f_e^{k+1} - d_e^{k+1}, \nabla f_e^{k+1} - \nabla f_e^k \rangle + \alpha \beta_2 \langle Pf_e^{k+1} - b_e^{k+1}, Pf_e^{k+1} - Pf_e^k \rangle, \end{aligned} \quad (4.7)$$

and then multiplying by $(f_e^{k+1} + f_e^k)$ on the two sides of the same result, we have

$$\begin{aligned} & -\|f_e^{k+1}\|_2^2 + \|f_e^k\|_2^2 + \alpha \beta_1 (\|\nabla f_e^{k+1}\|_2^2 - \|\nabla f_e^k\|_2^2) + \alpha \beta_2 (\|Pf_e^{k+1}\|_2^2 - \|Pf_e^k\|_2^2) \\ & = \alpha \beta_1 \langle \nabla f_e^{k+1} - d_e^{k+1}, \nabla f_e^{k+1} + \nabla f_e^k \rangle + \alpha \beta_2 \langle Pf_e^{k+1} - b_e^{k+1}, Pf_e^{k+1} + Pf_e^k \rangle. \end{aligned} \quad (4.8)$$

Adding (4.7) to (4.8) leads to

$$\begin{aligned} & \frac{1}{2\alpha} (\|Af_e^k\|_2^2 - \|Af_e^{k+1}\|_2^2 - \|A(f_e^{k+1} - f_e^k)\|_2^2) \\ & = \beta_1 \|\nabla f_e^{k+1}\|_2^2 + \beta_2 \|Pf_e^{k+1}\|_2^2 - \beta_1 \langle d_e^{k+1}, \nabla f_e^{k+1} \rangle - \beta_2 \langle b_e^{k+1}, Pf_e^{k+1} \rangle \end{aligned} \quad (4.9)$$

where $A = \sqrt{I - (\alpha\beta_1\nabla^T\nabla + \alpha\beta_2P^TP)}$. By the positive semi-definite assumption of $I - (\alpha\beta_1\nabla^T\nabla + \alpha\beta_2P^TP)$, we know that A is well defined and also positive semi-definite.

Then, subtracting the first equation of (4.6) from the first equation of (4.1), and then multiplying by b_e^{k+1} on the two sides of the result, we obtain

$$\langle q_{b_e}^{k+1}, b_e^{k+1} \rangle + \beta_2 \|b_e^{k+1}\|_2^2 - \beta_2 \langle Pf_e^k, b_e^{k+1} \rangle + \beta_2 \langle t_e^k, b_e^{k+1} \rangle = 0. \quad (4.10)$$

Similarly,

$$\langle q_{d_e}^{k+1}, d_e^{k+1} \rangle + \beta_1 \|d_e^{k+1}\|_2^2 - \beta_1 \langle \nabla f_e^k, d_e^{k+1} \rangle + \beta_1 \langle s_e^k, d_e^{k+1} \rangle = 0. \quad (4.11)$$

From the left hands of (4.10) and (4.11) plus the right hand of (4.9), we have

$$\begin{aligned} & \frac{1}{2\alpha} (\|Af_e^k\|_2^2 - \|Af_e^{k+1}\|_2^2 - \|A(f_e^{k+1} - f_e^k)\|_2^2) \\ &= \langle q_{b_e}^{k+1}, b_e^{k+1} \rangle + \langle q_{d_e}^{k+1}, d_e^{k+1} \rangle \\ & \quad + \beta_1 (\|\nabla f_e^{k+1}\|_2^2 + \|d_e^{k+1}\|_2^2 - \langle \nabla f_e^{k+1} + \nabla f_e^k, d_e^{k+1} \rangle + \langle s_e^k, d_e^{k+1} \rangle) \\ & \quad + \beta_2 (\|Pf_e^{k+1}\|_2^2 + \|b_e^{k+1}\|_2^2 - \langle Pf_e^{k+1} + Pf_e^k, b_e^{k+1} \rangle + \langle t_e^k, b_e^{k+1} \rangle). \end{aligned} \quad (4.12)$$

Subtracting the fourth equation of (4.6) from the fourth equation of (4.1), we have

$$s_e^{k+1} = s_e^k + d_e^{k+1} - \nabla f_e^{k+1}, \quad (4.13)$$

then squaring the both sides of (4.13) and rearranging the terms, we get

$$\langle s_e^k, d_e^{k+1} \rangle = \frac{1}{2} (\|s_e^{k+1}\|_2^2 - \|s_e^k\|_2^2 - \|d_e^{k+1} - \nabla f_e^{k+1}\|_2^2) + \langle s_e^k, \nabla f_e^{k+1} \rangle. \quad (4.14)$$

Similarly, for the fifth equations of (4.1) and (4.6), we have

$$\langle t_e^k, b_e^{k+1} \rangle = \frac{1}{2} (\|t_e^{k+1}\|_2^2 - \|t_e^k\|_2^2 - \|b_e^{k+1} - Pf_e^{k+1}\|_2^2) + \langle t_e^k, Pf_e^{k+1} \rangle. \quad (4.15)$$

For the facts of (3.33) and (4.5), we have

$$\beta_1 \nabla^T s_e^k + \beta_2 P^T t_e^k = 0. \quad (4.16)$$

In (4.12), we can rewrite

$$\begin{aligned} & \|\nabla f_e^{k+1}\|_2^2 + \|d_e^{k+1}\|_2^2 - \langle \nabla f_e^{k+1} + \nabla f_e^k, d_e^{k+1} \rangle \\ &= \frac{1}{2} (\|\nabla f_e^{k+1} - d_e^{k+1}\|_2^2 + \|\nabla f_e^k - d_e^{k+1}\|_2^2 + \|\nabla f_e^{k+1}\|_2^2 - \|\nabla f_e^k\|_2^2), \end{aligned} \quad (4.17)$$

$$\begin{aligned} & \|Pf_e^{k+1}\|_2^2 + \|b_e^{k+1}\|_2^2 - \langle Pf_e^{k+1} + Pf_e^k, b_e^{k+1} \rangle \\ &= \frac{1}{2} (\|Pf_e^{k+1} - b_e^{k+1}\|_2^2 + \|Pf_e^k - b_e^{k+1}\|_2^2 + \|Pf_e^{k+1}\|_2^2 - \|Pf_e^k\|_2^2). \end{aligned} \quad (4.18)$$

Substituting (4.14)–(4.18) into (4.12), we obtain the following equation

$$\begin{aligned} & \frac{1}{2\alpha} (\|Af_e^k\|_2^2 - \|Af_e^{k+1}\|_2^2) + \frac{\beta_1}{2} (\|s_e^k\|_2^2 - \|s_e^{k+1}\|_2^2) + \frac{\beta_2}{2} (\|t_e^k\|_2^2 - \|t_e^{k+1}\|_2^2) \\ &= \frac{1}{2\alpha} \|A(f_e^{k+1} - f_e^k)\|_2^2 + \langle q_{b_e}^{k+1}, b_e^{k+1} \rangle + \langle q_{d_e}^{k+1}, d_e^{k+1} \rangle \\ & \quad + \frac{\beta_1}{2} (\|\nabla f_e^k - d_e^{k+1}\|_2^2 + \|\nabla f_e^{k+1}\|_2^2 - \|\nabla f_e^k\|_2^2) \\ & \quad + \frac{\beta_2}{2} (\|Pf_e^k - b_e^{k+1}\|_2^2 + \|Pf_e^{k+1}\|_2^2 - \|Pf_e^k\|_2^2). \end{aligned} \quad (4.19)$$

Summing (4.19) from $k=0$ to K , we have

$$\begin{aligned}
& \frac{1}{2\alpha} \sum_{k=0}^K \|A(f_e^{k+1} - f_e^k)\|_2^2 + \sum_{k=0}^K \langle q_{b_e}^{k+1}, b_e^{k+1} \rangle + \sum_{k=0}^K \langle q_{d_e}^{k+1}, d_e^{k+1} \rangle \\
& + \frac{\beta_1}{2} \sum_{k=0}^K \|\nabla f_e^k - d_e^{k+1}\|_2^2 + \frac{\beta_2}{2} \sum_{k=0}^K \|P f_e^k - b_e^{k+1}\|_2^2 + \frac{\beta_1}{2} \|\nabla f_e^{K+1}\|_2^2 \\
& + \frac{\beta_2}{2} \|P f_e^{K+1}\|_2^2 + \frac{1}{2\alpha} \|A f_e^{K+1}\|_2^2 + \frac{\beta_1}{2} \|s_e^{K+1}\|_2^2 + \frac{\beta_2}{2} \|t_e^{K+1}\|_2^2 \\
& = \frac{1}{2\alpha} \|A f_e^0\|_2^2 + \frac{\beta_1}{2} \|s_e^0\|_2^2 + \frac{\beta_2}{2} \|t_e^0\|_2^2 + \frac{\beta_1}{2} \|\nabla f_e^0\|_2^2 + \frac{\beta_2}{2} \|P f_e^0\|_2^2 < C, \quad (4.20)
\end{aligned}$$

where K is an arbitrary positive integer and C is a positive constant.

Since for any convex function J , the Bregman distance (2.2) is nonnegative. Then the following formula is satisfied

$$D_J^p(u, v) + D_J^q(v, u) = \langle q - p, u - v \rangle \geq 0, \quad \forall p \in \partial J(v), \forall q \in \partial J(u). \quad (4.21)$$

Hence, by (4.21), we get

$$D_{E_2}^{q_d^*}(d^k, d^*) + D_{E_2}^{q_d^k}(d^*, d^k) = \langle q_d^k - q_d^*, d^k - d^* \rangle = \langle q_{d_e}^k, d_e^k \rangle \geq 0, \quad (4.22)$$

for $\forall q_d^* \in \partial E_2(d^*), \forall q_d^k \in \partial E_2(d^k)$. Similarly,

$$D_{E_3}^{q_b^*}(b^k, b^*) + D_{E_3}^{q_b^k}(b^*, b^k) = \langle q_b^k - q_b^*, b^k - b^* \rangle = \langle q_{b_e}^k, b_e^k \rangle \geq 0, \quad (4.23)$$

for $\forall q_b^* \in \partial E_3(b^*), \forall q_b^k \in \partial E_3(b^k)$.

From (4.20), (4.22) and (4.23), we have

$$\sum_{k=0}^K \langle q_{b_e}^{k+1}, b_e^{k+1} \rangle < C, \quad \forall K > 0, \quad (4.24)$$

$$\sum_{k=0}^K \langle q_{d_e}^{k+1}, d_e^{k+1} \rangle < C, \quad \forall K > 0. \quad (4.25)$$

According to (4.22)–(4.25), we obtain

$$\lim_{k \rightarrow +\infty} D_{E_2}^{q_d^*}(d^k, d^*) = \lim_{k \rightarrow +\infty} (E_2(d^k) - J_2(d^*) - \langle q_d^*, d^k - d^* \rangle) = 0, \quad (4.26)$$

$$\lim_{k \rightarrow +\infty} D_{E_3}^{q_b^*}(b^k, b^*) = \lim_{k \rightarrow +\infty} (E_3(b^k) - J_3(b^*) - \langle q_b^*, b^k - b^* \rangle) = 0. \quad (4.27)$$

Applying (4.20), and $\beta_1 > 0, \beta_2 > 0$, we easily get

$$\sum_{k=0}^K \|\nabla f_e^k - d_e^{k+1}\|_2^2 < C, \quad \forall K > 0, \quad (4.28)$$

$$\sum_{k=0}^K \|P f_e^k - b_e^{k+1}\|_2^2 < C, \quad \forall K > 0. \quad (4.29)$$

Therefore,

$$\lim_{k \rightarrow +\infty} \|\nabla f_e^k - d_e^{k+1}\|_2^2 = 0, \quad (4.30)$$

$$\lim_{k \rightarrow +\infty} \|Pf_e^k - b_e^{k+1}\|_2^2 = 0. \quad (4.31)$$

By the known relations of $d^* = \nabla f^*$ and $b^* = Pf^* - g$, using (4.30) and (4.31), the following formulas are satisfied

$$\lim_{k \rightarrow +\infty} \|\nabla f^k - d^{k+1}\|_2^2 = 0, \quad (4.32)$$

$$\lim_{k \rightarrow +\infty} \|Pf^k - g - b^{k+1}\|_2^2 = 0. \quad (4.33)$$

As a result of the continuity of E_2 and E_3 , by (4.26), (4.27), (4.32) and (4.33), we have

$$\lim_{k \rightarrow +\infty} (E_2(\nabla f^k) - E_2(\nabla f^*) - \langle q_d^*, \nabla(f^k - f^*) \rangle) = 0, \quad (4.34)$$

$$\lim_{k \rightarrow +\infty} (E_3(Pf^k - g) - E_3(Pf^* - g) - \langle q_b^*, P(f^k - f^*) \rangle) = 0. \quad (4.35)$$

Applying (4.4), then (4.34) plus (4.35) leads to

$$\lim_{k \rightarrow +\infty} (E_2(\nabla f^k) + E_3(Pf^k - g) - E_2(\nabla f^*) - E_3(Pf^* - g)) = 0,$$

in other words,

$$\lim_{k \rightarrow +\infty} (\|\nabla f^k\|_1 + \frac{\lambda}{2} \|Pf^k - g\|_2^2) = \|\nabla f^*\|_1 + \frac{\lambda}{2} \|Pf^* - g\|_2^2.$$

Hence, (4.2) is proved.

Now, we prove (4.3) by contradiction. If (4.3) is false, there exists a subsequence $\{f^{k_i}\}$ such that $\|f^{k_i} - f^*\|_2 > \varepsilon$, $\forall k_i > 0$, for certain small positive number ε . Let \hat{f}^{k_i} be the intersection of the sphere $\{f : \|f - f^*\|_2 = \varepsilon\}$ and the line segment from f^* to f^{k_i} . Then there is only a $t \in (0, 1)$ such that $\hat{f}^{k_i} = tf^* + (1-t)f^{k_i}$ on the above sphere. Let $J_0(f) = \|\nabla f\|_1 + \frac{\lambda}{2} \|Pf - g\|_2^2$, and $\hat{f} = \arg \min_f \{J_0(f) : \|f - f^*\|_2 = \varepsilon\}$. As J_0 is convex and f^* is a unique solution of (3.2), hence,

$$J_0(f^*) < J_0(\hat{f}) \leq J_0(\hat{f}^{k_i}) = J_0(tf^* + (1-t)f^{k_i}) \leq tJ_0(f^*) + (1-t)J_0(f^{k_i}) < J_0(f^{k_i}). \quad (4.36)$$

Taking the limit of (4.36) leads to

$$J_0(f^*) < J_0(\hat{f}) \leq \lim_{k \rightarrow +\infty} J_0(f^{k_i}) = J_0(f^*). \quad (4.37)$$

Hence, the contradiction is yielded, that is to say, (4.3) is proved. \square

REMARK 4.2. *It is easy to verify that the updating order of b^{k+1} , d^{k+1} and f^{k+1} of Algorithm 3.1 can be changed into any other combinational order, which has no bearing on the convergence of the algorithm.*

THEOREM 4.3. *Let $J(f) = \|f\|_1$. Suppose that there exists a unique solution f^* of (3.2), and $\alpha > 0$, $\beta > 0$, $I - \alpha P^T P$ is positive semi-definite. Then we have the following convergence results for the iterative scheme (3.55)*

$$\lim_{k \rightarrow +\infty} (\|f^k\|_1 + \frac{\lambda}{2} \|Pf^k - g\|_2^2) = \|f^*\|_1 + \frac{\lambda}{2} \|Pf^* - g\|_2^2, \quad (4.38)$$

and

$$\lim_{k \rightarrow +\infty} \|f^k - f^*\|_2 = 0. \quad (4.39)$$

Proof. Let

$$b^* = Pf^* - g, \quad \beta t^* = \hat{q}_b^*, \quad \beta P^T t^* = -\hat{q}_f^*.$$

Hence,

$$P^T \hat{q}_b^* + \hat{q}_f^* = 0. \quad (4.40)$$

Then there exist the following equations

$$\begin{cases} \hat{q}_b^* + \beta(b^* - (Pf^* - g) - t^*) = 0, \\ \hat{q}_f^* + \frac{\beta}{\alpha}(f^* - f^* + \alpha P^T(Pf^* - g - b^* + t^*)) = 0, \\ t^* = t^* + (Pf^* - g - b^*). \end{cases} \quad (4.41)$$

Taking the same operations for the second equations of (3.55) and (4.41) as those in Theorem 4.1, we obtain

$$\begin{aligned} & \frac{\beta}{2\alpha} (\|\hat{A}f_e^k\|_2^2 - \|\hat{A}f_e^{k+1}\|_2^2 - \|\hat{A}(f_e^{k+1} - f_e^k)\|_2^2) \\ & = \langle \hat{q}_{f_e}^{k+1}, f_e^{k+1} \rangle + \beta \|Pf_e^{k+1}\|_2^2 - \beta \langle b_e^{k+1} - t_e^k, Pf_e^{k+1} \rangle, \end{aligned} \quad (4.42)$$

where $\hat{A} = \sqrt{I - \alpha P^T P}$ and \hat{A} is positive semi-definite.

Subtracting the first equation of (4.41) from the first equation of (3.55), and then multiplying by b_e^{k+1} on the two sides of the above result, we have

$$\langle \hat{q}_{b_e}^{k+1}, b_e^{k+1} \rangle + \beta \|b_e^{k+1}\|_2^2 - \beta \langle Pf_e^k, b_e^{k+1} \rangle - \beta \langle t_e^k, b_e^{k+1} \rangle = 0. \quad (4.43)$$

The third equation of (3.55) minus the third equation of (4.41) is

$$t_e^{k+1} = t_e^k + Pf_e^{k+1} - b_e^{k+1}, \quad (4.44)$$

then squaring the both sides of (4.44) and rearranging the terms, we obtain that

$$\langle t_e^k, Pf_e^{k+1} - b_e^{k+1} \rangle = \frac{1}{2} (\|t_e^{k+1}\|_2^2 - \|t_e^k\|_2^2 - \|Pf_e^{k+1} - b_e^{k+1}\|_2^2). \quad (4.45)$$

By (4.42), (4.43) and (4.45), we have

$$\begin{aligned} & \sum_{k=1}^K \langle \hat{q}_{f_e}^{k+1}, f_e^{k+1} \rangle + \sum_{k=1}^K \langle \hat{q}_{b_e}^{k+1}, b_e^{k+1} \rangle + \frac{\beta}{2} \sum_{k=1}^K \|Pf_e^k - b_e^{k+1}\|_2^2 + \frac{\beta}{2} \|Pf_e^{K+1}\|_2^2 \\ & + \frac{\beta}{2} \|t_e^{K+1}\|_2^2 + \frac{\beta}{2\alpha} \sum_{k=1}^K \|\hat{A}(f_e^{k+1} - f_e^k)\|_2^2 + \frac{\beta}{2\alpha} \|\hat{A}f_e^{K+1}\|_2^2 \\ & = \frac{\beta}{2\alpha} \|\hat{A}f_e^0\|_2^2 + \frac{\beta}{2} \|Pf_e^0\|_2^2 + \frac{\beta}{2} \|t_e^0\|_2^2 < C. \end{aligned} \quad (4.46)$$

As the derivation of Theorem 4.1, we obtain

$$\lim_{k \rightarrow +\infty} D_{\hat{E}_1}^{\hat{q}_f^*}(f^k, f^*) = \lim_{k \rightarrow +\infty} (\hat{E}_1(f^k) - \hat{E}_1(f^*) - \langle \hat{q}_f^*, f^k - f^* \rangle) = 0, \quad (4.47)$$

$$\lim_{k \rightarrow +\infty} D_{\hat{E}_2}^{\hat{q}_b^*}(b^k, b^*) = \lim_{k \rightarrow +\infty} (\hat{E}_2(b^k) - \hat{E}_2(b^*) - \langle \hat{q}_b^*, b^k - b^* \rangle) = 0. \quad (4.48)$$

By (4.46) and $b^* = Pf^* - g$, we get

$$\lim_{k \rightarrow +\infty} \|Pf^k - g - b^{k+1}\|_2 = 0. \quad (4.49)$$

Then combining (4.48) and (4.49) gives

$$\lim_{k \rightarrow +\infty} (\hat{E}_2(Pf^k - g) - \hat{E}_2(Pf^* - g) - \langle \hat{q}_b^*, Pf^k - Pf^* \rangle) = 0. \quad (4.50)$$

Utilizing (4.40), (4.47) plus (4.50) leads to (4.38).

Relation (4.39) can be proved by the method in the proof of Theorem 4.1. \square

THEOREM 4.4. *Let $J(f) = \|\nabla f\|_1$. Suppose that there exists a unique solution f^* of (3.57), and $\alpha_1 > 0$, $\alpha_2 > 0$, $\beta_1 > 0$, $\beta_2 > 0$, $I - (\alpha\beta_1 \nabla^T \nabla + \alpha\beta_2 P^T P)$ is positive semi-definite. Then in what follows we have the convergence results for the iterative scheme (4.1), namely,*

$$\lim_{k \rightarrow +\infty} (\|\nabla f^k\|_1 + \lambda \|Pf^k - g\|_1) = \|\nabla f^*\|_1 + \lambda \|Pf^* - g\|_1, \quad (4.51)$$

and

$$\lim_{k \rightarrow +\infty} \|f^k - f^*\|_2 = 0, \quad (4.52)$$

where $\alpha = 1/(\omega_1 + \omega_2)$.

Proof. The proof procedure is the same as that of Theorem 4.1. \square

THEOREM 4.5. *Let $J(f) = \|f\|_1$. Suppose that there exists a unique solution f^* of (3.57), and $\alpha > 0$, $\beta > 0$, $I - \alpha P^T P$ is positive semi-definite. Then we have the following results for the iterative scheme (3.55)*

$$\lim_{k \rightarrow +\infty} (\|f^k\|_1 + \lambda \|Pf^k - g\|_1) = \|f^*\|_1 + \lambda \|Pf^* - g\|_1, \quad (4.53)$$

and

$$\lim_{k \rightarrow +\infty} \|f^k - f^*\|_2 = 0. \quad (4.54)$$

Proof. The proof of Theorem 4.3 can be used directly here. \square

4.2. Convergence of the gradient-descent-based split Bregman iteration

In this subsection, we provide the convergence analysis of the GDSB algorithm for robust tomographic image reconstruction. In [2], the authors presented a convergence on ideal unconstrained/constrained split Bregman methods with updating f^{k+1} accurately. In practice, the accurate update is impossible to attain since one cannot solve a general large-scale linear system exactly and frequently turns to certain

iterative method. Here we recall the practical unconstrained split Bregman method in Subsection 2.2 as follows.

$$\begin{cases} q_d^{k+1} + \mu(d^{k+1} - \nabla f^k - s^k) = 0, \\ f^{k+1} = f^k - \alpha(\lambda P^T(Pf^k - g) + \mu \nabla^T(\nabla f^k - d^{k+1} + s^k)), \\ s^{k+1} = s^k + \nabla f^{k+1} - d^{k+1}. \end{cases} \quad (4.55)$$

THEOREM 4.6. *Suppose that there exists a unique solution f^* of (3.2), and $\alpha > 0$, $\lambda > 0$, $\mu > 0$, $I - (\alpha\mu\nabla^T\nabla + \alpha\lambda P^T P)$ is positive semi-definite. Then we have the following convergence results for the iterative scheme (4.55)*

$$\lim_{k \rightarrow +\infty} (\|\nabla f^k\|_1 + \frac{\lambda}{2}\|Pf^k - g\|_2^2) = \|\nabla f^*\|_1 + \frac{\lambda}{2}\|Pf^* - g\|_2^2. \quad (4.56)$$

In particular,

$$\lim_{k \rightarrow +\infty} \|\nabla f^k\|_1 = \|\nabla f^*\|_1, \quad (4.57)$$

$$\lim_{k \rightarrow +\infty} \|Pf^k - g\|_2 = \|Pf^* - g\|_2. \quad (4.58)$$

Moreover,

$$\lim_{k \rightarrow +\infty} \|f^k - f^*\|_2 = 0. \quad (4.59)$$

Proof. By the assumption that f^* is the unique solution of (3.2), we have the following first-order optimality condition

$$\nabla^T q_d^* + \lambda P^T(Pf^* - g) = 0, \quad (4.60)$$

where $q_d^* \in \partial E_2(d^*)$ with $d^* = \nabla f^*$. Let $s^* = q_d^*/\mu$. Hence, we have

$$\begin{cases} q_d^* + \mu(d^* - \nabla f^* - s^*) = 0, \\ f^* = f^* - \alpha(\lambda P^T(Pf^* - g) + \mu \nabla^T(\nabla f^* - d^* + s^*)), \\ s^* = s^* + \nabla f^* - d^*. \end{cases} \quad (4.61)$$

Taking the same operations for the second equations of (4.55) and (4.61) as those in the proof of Theorem 4.1, we obtain

$$\begin{aligned} & \frac{1}{2\alpha} (\|Bf_e^k\|_2^2 - \|Bf_e^{k+1}\|_2^2 - \|B(f_e^{k+1} - f_e^k)\|_2^2) \\ &= \mu \|\nabla f_e^{k+1}\|_2^2 + \lambda \|Pf_e^{k+1}\|_2^2 - \mu \langle d_e^{k+1} - s_e^k, \nabla f_e^{k+1} \rangle, \end{aligned} \quad (4.62)$$

where $B = \sqrt{I - (\alpha\mu\nabla^T\nabla + \alpha\lambda P^T P)}$ and B is positive semi-definite by assumption.

Subtracting the first equation of (4.61) from the first equation of (4.55), and then multiplying by d_e^{k+1} on the two sides of the above result, we have

$$\langle q_{d_e}^{k+1}, d_e^{k+1} \rangle + \mu \|d_e^{k+1}\|_2^2 - \mu \langle \nabla f_e^k, d_e^{k+1} \rangle - \mu \langle s_e^k, d_e^{k+1} \rangle = 0. \quad (4.63)$$

Adding the left-hand side of (4.63) onto the right-hand side of (4.62), we get

$$\begin{aligned} & \frac{1}{2\alpha} (\|Bf_e^k\|_2^2 - \|Bf_e^{k+1}\|_2^2 - \|B(f_e^{k+1} - f_e^k)\|_2^2) \\ &= \mu \|\nabla f_e^{k+1}\|_2^2 + \lambda \|Pf_e^{k+1}\|_2^2 + \langle q_{d_e}^{k+1}, d_e^{k+1} \rangle + \mu \|d_e^{k+1}\|_2^2 \\ & \quad - \mu \langle \nabla f_e^{k+1} + \nabla f_e^k, d_e^{k+1} \rangle + \mu \langle s_e^k, \nabla f_e^{k+1} - d_e^{k+1} \rangle. \end{aligned} \quad (4.64)$$

Subtracting the third equation of (4.61) from the third equation of (4.55), we have

$$s_e^{k+1} = s_e^k + \nabla f_e^{k+1} - d_e^{k+1}. \quad (4.65)$$

Then squaring the both sides of (4.65) and rearranging the terms, we obtain

$$\langle s_e^k, \nabla f_e^{k+1} - d_e^{k+1} \rangle = \frac{1}{2} (\|s_e^{k+1}\|_2^2 - \|s_e^k\|_2^2 - \|\nabla f_e^{k+1} - d_e^{k+1}\|_2^2). \quad (4.66)$$

The terms in (4.64)

$$\begin{aligned} & \|\nabla f_e^{k+1}\|_2^2 + \|d_e^{k+1}\|_2^2 - \langle \nabla f_e^{k+1} + \nabla f_e^k, d_e^{k+1} \rangle \\ &= \frac{1}{2} (\|\nabla f_e^{k+1} - d_e^{k+1}\|_2^2 + \|\nabla f_e^k - d_e^{k+1}\|_2^2 + \|\nabla f_e^{k+1}\|_2^2 - \|\nabla f_e^k\|_2^2). \end{aligned} \quad (4.67)$$

Substituting (4.66) and (4.67) into (4.64) and then summing from $k=0$ to K leads to

$$\begin{aligned} & \sum_{k=1}^K \langle q_{d_e}^{k+1}, d_e^{k+1} \rangle + \frac{\mu}{2} \sum_{k=1}^K \|\nabla f_e^k - d_e^{k+1}\|_2^2 + \lambda \sum_{k=1}^K \|P f_e^{k+1}\|_2 \\ &+ \frac{\mu}{2} \|\nabla f_e^{K+1}\|_2^2 + \frac{\mu}{2} \|s_e^{K+1}\|_2^2 + \frac{1}{2\alpha} \sum_{k=1}^K \|B(f_e^{k+1} - f_e^k)\|_2^2 + \frac{1}{2\alpha} \|B f_e^{K+1}\|_2^2 \\ &= \frac{1}{2\alpha} \|B f_e^0\|_2^2 + \frac{\mu}{2} \|\nabla f_e^0\|_2^2 + \frac{\mu}{2} \|s_e^0\|_2^2 < C. \end{aligned}$$

Hence,

$$\lim_{k \rightarrow +\infty} \|P f_e^k\|_2 = 0. \quad (4.68)$$

Then

$$\lim_{k \rightarrow +\infty} \|P f^k - g\|_2 = \|P f^* - g\|_2. \quad (4.69)$$

Namely, (4.58) is proved. By (4.68), we can also get

$$\lim_{k \rightarrow +\infty} \langle P^T (P f^* - g), f^k - f^* \rangle = 0. \quad (4.70)$$

As the proof of Theorem 4.1, we have

$$\lim_{k \rightarrow +\infty} (\|\nabla f^k\|_1 - \|\nabla f^*\| - \langle \nabla^T q_d^*, f^k - f^* \rangle) = 0. \quad (4.71)$$

By (4.60), combining (4.70) and (4.71) concludes (4.57).

Next, the proof of (4.59) is similar as that in the proof of Theorem 4.1. So here we omit it. \square

REMARK 4.7. *The constrained split Bregman method in [2] can be treated as the practical unconstrained case in this article. Also, the convergence of the practical constrained split Bregman method can be easily obtained via the technique in the proofs of Theorems 4.1 and 4.6.*

REMARK 4.8. *Comparing Theorem 4.1 with Theorem 4.6, we conclude that the convergence of the proposed method is independent of the selection of the regularization parameter λ , whereas the convergence of the GDSB iteration depends. That is to say, the parameter λ in our method has merely effect on the reconstructed quality; In Theorem 4.6, however, not only does the parameter λ have bearings on the reconstructed quality, but also on the convergence of the iterative scheme. Therefore, the proposed algorithm should be more flexible and robust than the practical split Bregman iteration.*

5. Numerical experiments

In this section, we present several numerical experiments using synthetic and real data to illustrate that our LSB method can produce desirable image reconstruction results for the data detected from uniformly and sparsely distributed views and contaminated by the additive Gaussian white noise. Suppose that the projections from one view are uniformly spacing. Actually, this hypothesis is reasonable for the real tomography. All the involved algorithms other than FBP are programmed in C/C++ language and all the implementations ran on a desktop with Intel Xeon X5550 2.67GHz CPU, Fedora 11 OS and GCC 4.3.2 compiler, and no any parallel computing was conducted.

5.1. Comparisons with the FBP and the gradient-flow-based semi-implicit finite element method

We first investigate the performance of our algorithm compared with the classical FBP and the gradient-flow-based semi-implicit finite element method (SFEM). For this numerical simulation, the test image is taken to be a synthetic Shepp-Logan phantom as shown in Fig. 5.1, which is discretized on a 257×257 pixel grid within the gray interval $[0,1]$. This phantom is often utilized in evaluating tomographic reconstruction algorithms. A set of uniformly spacing 257 parallel projections from each sampled view is obtained. The totally used angles is 60, namely, projecting once every 3° . Obviously, the sampling is severely insufficient. Then the additive Gaussian white noise is added onto the projection data from each view resulting in a set of corrupted data with signal-to-noise (SNR) ($=24.7dB$), where the SNR in decibels is defined to be

$$\text{SNR} = 10 \log_{10} \frac{\sum_{n_v=1}^{N_v} \sum_{n_p=1}^{M_p} |Pf_{n_v, n_p} - \overline{Pf}|^2}{\sum_{n_v=1}^{N_v} \sum_{n_p=1}^{M_p} |\varepsilon_{n_v, n_p} - \bar{\varepsilon}|^2}, \quad (5.1)$$

where N_v is the number of total angles, M_p is the number of total projections from each angle, \overline{Pf} is the average of totally noiseless measured data, ε is the added noise and $\bar{\varepsilon}$ is the average of ε .

By Theorem 4.1, we recall that just the parameters β_1 , β_2 and α have effect on the convergence of our algorithm, and λ has impact on the reconstructed quality. Under this guidance, selecting these parameters is, hence, relatively easy. Here we choose the parameters $\lambda=0.1$, $\beta_1=0.06$, $\beta_2=3 \times 10^{-5}$, $\alpha=10/3$ in our algorithm. The iteration number is 1500. In addition, different filters (such as Ramp, Shepp-Logan, Hamming and Hann) and different interpolation schemes (such as nearest, linear, cubic spline) can be used in FBP. However, the performance of FBP has no evident improvement when we choose different combinations in terms of filters and interpolations. Without loss of generality, we only display in Fig. 5.1 the image reconstructed by FBP using a ramp filter and linear interpolating technique. In SFEM, we choose the modified parameter $\epsilon=0.001$, the temporal step size $\tau=0.01$ and regularization parameter $\lambda=2$. The iteration number is 50 for the SFEM.

As shown in Fig. 5.1, in terms of reconstruction quality, the LSB and the SFEM are much better than FBP; yet the SFEM is a little more blurry on edges than the LSB, which can be clearly figured out by Fig. 5.2. From the view of computational time, the new method is totally 73.8s for 1500 iterations, and the SFEM is about 1831s for 50 iterations, where the iterations are almost the least demanded steps for the two methods.



FIG. 5.1. The SNR = 24.7dB. From left to right, reconstructed images by FBP, the proposed method, SFEM, and the original image with gray scale over $[0.0, 1.0]$.

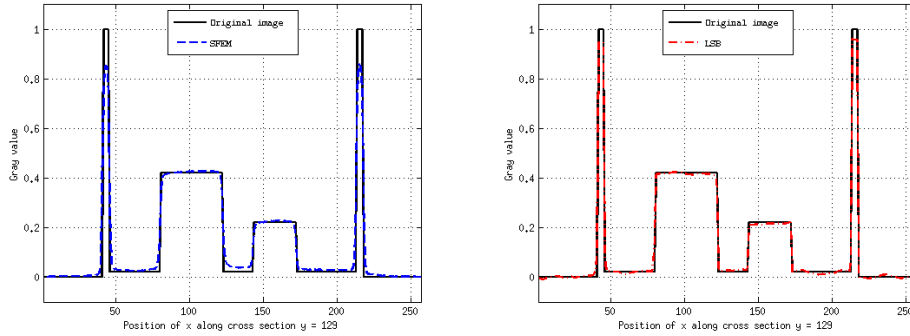


FIG. 5.2. The cross sections of the images, as shown in Fig. 5.1, along the centers are plotted in the horizontal direction. The solid lines denote the cross sections of the original image. The dashed line and the dash-dotted line denote the cross sections of the reconstructed images by the SFEM (left) and the LSB (right), respectively.

5.2. Comparison with the gradient-descent-based split Bregman iteration

As described in Subsection 2.2, we can directly apply the GDSB method to resolve the concerned optimization model (3.2). What is more, in this paper we have proved that the GDSB method is convergent under the proper condition. So why we do not use it but turn to propose the LSB method? The theoretical reason can be refer to Remark 4.8. The numerical aspect will be presented in this subsection.

Here we consider a real CT image within the gray interval $[0,1]$ as show in Fig. 5.3. The size of this image is also 257×257 . We obtain uniformly spacing 257 parallel projections along each sampled view. The totally used views is 60. Then the additive Gaussian white noise is added onto the projection data along each view, which leads to a set of corrupted data with SNR=29.6dB.

As the size of the original image and the projection views are the same as the above experiment, so the parameters β_1 , β_2 and α can be selected as before to guarantee the LSB iterative convergence, namely, 0.06, 3×10^{-5} and $10/3$, respectively. The regularization parameter λ is chosen as 1.0. In addition, we select the same λ as the LSB method to assure good reconstructed quality in the GDSB method. According to Theorems 4.1 and 4.6, to assure convergence of the GDSB method, we should sufficiently reduce the step size α for the relatively large λ replacing the small β_2 in

the convergence condition. Here the $\alpha = 2 \times 10^{-6}$, $\mu = 0.06$. The rate of convergence of the GDSB method is, therefore, slower than that of the LSB method, which can be indicated by the decreasing amount of the objective energy as the iteration progressing as shown in Fig. 5.3.

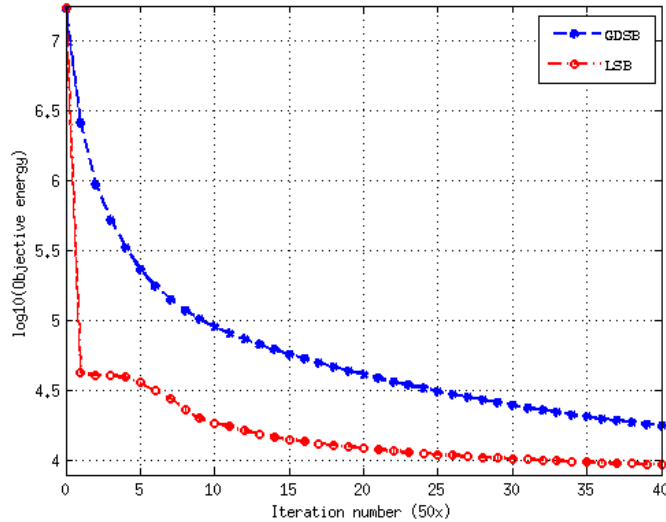


FIG. 5.3. The star dashed and circle dash-dotted lines denote the values of the objective energy of the LSB and GDSB methods as the increase of the iteration number, respectively.

After 2000 iterations, the reconstructed images are shown in Fig. 5.4. It is easy to figure out that the performance of the LSB method is much better than that of the GDSB method. Remarkably, the latter introduces a lot of artifacts in the background, and leads to the blurry result. We further select another $\lambda = 20$ in both methods. Then, the LSB method produces the pretty near image as $\lambda = 1$, whereas the GDSB method is not convergent.

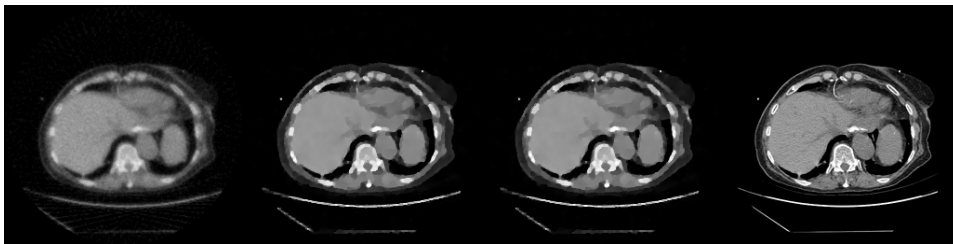


FIG. 5.4. The $SNR = 29.6dB$. From left to right, reconstructed images by the GDSB method, the LSB method with $\lambda = 1$, the LSB method with $\lambda = 20$, and the original image with gray scale over $[0.0, 1.0]$.

6. Conclusions

In this work, we have presented a novel linearized split Bregman iterative algorithm which is constructed by means of the thoughts of split and linearized Bregman

methods. We also have given the rigorous proof for the convergence of the proposed method under appropriate condition, where the convergence does not depend on the selection of the regularization parameter. Using the idea of the proof, we further proved the convergence of the GDSB method, yet which relies on the value of regularization parameter. In other words, our proposed method is more flexible and robust than the GDSB method. Notably, our method can be generalized to efficiently resolve the robust CS problem, as well as the TV- ℓ_1 and ℓ_1 - ℓ_1 minimization problems.

Finally, we have evaluated the performance of our algorithm in dealing with the sparse-view X-ray CT reconstruction problem using the synthetic and real data. As the results in these numerical comparisons, our algorithm yields high quality and almost accurate tomographic reconstructions in the above difficult problem. The numerical results presented in this paper also suggest that our algorithm greatly outperforms the classical FBP method in the reconstructed quality. Also, our algorithm possesses a bit better reconstructed performance and more effective than the earlier proposed SFEM method. Moreover, based on the numerical experiments, we can figure out that our method has much faster convergent rate and better reconstructed quality compared with the popular split Bregman method. Besides the advantages as the split Bregman method, such as efficient to solve several special ℓ_1 problems, easy to program, simple to parallelize, saving memory, etc., our algorithm is more suitable for dealing with the general ℓ_1 -regularized problems, for instance, the tomographic image reconstruction, and so forth.

Acknowledgement. We would like to thank Prof. Chunlin Wu, from the School of Mathematical Sciences at Nankai University, for his helpful discussions, comments and corrections. We also want to thank Dr. Ming Li, from Institute of Computational Mathematics at Chinese Academy of Sciences, for his helpful suggestions on the part of numerical experiments.

REFERENCES

- [1] L. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR computational mathematics and mathematical physics*, 7(3):200–217, 1967.
- [2] J. Cai, S. Osher, and Z. Shen. Split Bregman methods and frame based image restoration. *Multiscale Model. Simul.*, 8(2):337–369, 2009.
- [3] E. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathématique*, 346(9):589–592, 2008.
- [4] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2):489–509, 2006.
- [5] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59(8):1207–1223, 2006.
- [6] E. Candès and M. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.
- [7] R. Chan, M Tao, and X. Yuan. Linearized Alternating Direction Method for Constrained Linear Least-Squares Problem. *East Asian J. Appl. Math.*, 2(4):326–341, 2012.
- [8] T. Chan and S. Esedoglu. Aspects of total variation regularized l^1 function approximation. *SIAM J. Appl. Math.*, 65(5):1817–1837, 2005.
- [9] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.*, 6(2):298–311, 1997.
- [10] C. Chen and G. Xu. Gradient-flow-based semi-implicit finite-element method and its convergence analysis for image reconstruction. *Inverse Problems*, 28(3):035006, 2012.

- [11] C. Chen and G. Xu. Computational inversion of electron micrographs using L2-gradient flows—convergence analysis. *Math. Methods Appl. Sci.*, DOI: 10.1002/mma.2770, 2013.
- [12] A. Delaney and Y. Bresler. A fast and accurate fourier algorithm for iterative parallel-beam tomography. *IEEE Trans. Image Process.*, 5(5):740–753, 1996.
- [13] A. Delaney and Y. Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Trans. Image Process.*, 7(2):204–221, 1998.
- [14] D. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.
- [15] E. Esser. Applications of Lagrangian-based alternating direction methods and connections to split Bregman. *CAM report*, 9:31, 2009.
- [16] B. Fahimian, Y. Zhao, R. Fung, Y. Mao, C. Zhu, Z. Huang, M. Khatonabadi, J. DeMarco, S. Osher, M. McNitt-Gray, and Miao J. Radiation dose reduction in medical CT through equally sloped tomography. *Institute for Mathematics and Its Applications*, 90095, 1770.
- [17] L. Feldkamp, L. Davis, and J. Kress. Practical cone-beam algorithm. *JOSA A*, 1(6):612–619, 1984.
- [18] H. Fu, M. Ng, M. Nikolova, and J. Barlow. Efficient minimization methods of mixed l2-l1 and l1-l1 norms for image restoration. *SIAM J. Sci. Comput.*, 27(6):1881–1902, 2006.
- [19] R. Glowinski. *Lectures on Numerical Methods for Non-Linear Variational Problems*. Springer, 2008.
- [20] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and operator-splitting methods in non-linear mechanics*, volume 9. SIAM, 1989.
- [21] T. Goldstein, X. Bresson, and S. Osher. Geometric applications of the split Bregman method: segmentation and surface reconstruction. *J. Sci. Comput.*, 45:272–293, 2010.
- [22] T. Goldstein and S. Osher. The split Bregman method for L1-regularized problems. *SIAM J. Imaging Sci.*, 2(2):323–343, 2009.
- [23] E. Hale, W. Yin, and Y. Zhang. Fixed-point continuation for ℓ_1 -minimization: Methodology and convergence. *SIAM J. Optim.*, 19(3):1107–1130, 2008.
- [24] X. Han, J. Bian, E. Ritman, E. Sidky, and X. Pan. Optimization-based reconstruction of sparse images from few-view projections. *Physics in Medicine and Biology*, 57(16):5245, 2012.
- [25] R. Jia, H. Zhao, and W. Zhao. Convergence analysis of the Bregman method for the variational model of image denoising. *Appl. Comput. Harmon. Anal.*, 27(3):367–379, 2009.
- [26] S. Ma, D. Goldfarb, and L. Chen. Fixed point and Bregman iterative methods for matrix rank minimization. *Math. Program.*, 128(1-2):321–353, 2011.
- [27] M. Nikolova. Minimizers of cost-functions involving nonsmooth data-fidelity terms. application to the processing of outliers. *SIAM J. Numer. Anal.*, 40(3):965–994, 2002.
- [28] 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(2):460–489, 2005.
- [29] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153. Springer, 2003.
- [30] S. Osher, Y. Mao, B. Dong, and W. Yin. Fast linearized Bregman iteration for compressive sensing and sparse denoising. *arXiv preprint arXiv:1104.0262*, 2011.
- [31] X. Pan, E. Sidky, and M. Vannier. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse problems*, 25(12):123009, 2009.
- [32] R. Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1997.
- [33] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992.
- [34] E. Sidky, J. Jørgensen, and X. Pan. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm. *Physics in medicine and biology*, 57(10):3065, 2012.
- [35] E. Sidky, C. Kao, and X. Pan. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *Journal of X-ray Science and Technology*, 14(2):119–139, 2006.
- [36] E. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in medicine and biology*, 53(17):4777, 2008.
- [37] X. Tang, J. Hsieh, R. Nilsen, S. Dutta, D. Samsonov, and A. Hagiwara. A three-dimensional-weighted cone beam filtered backprojection (CB-FBP) algorithm for image reconstruction in volumetric CT–helical scanning. *Physics in Medicine and Biology*, 51(4):855, 2006.
- [38] C. Vogel and M. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.*, 17(1):227–238, 1996.

- [39] G. Wang, T. Lin, P. Cheng, and D. Shinozaki. A general cone-beam reconstruction algorithm. *IEEE Transactions on Medical Imaging*, 12(3):486–496, 1993.
- [40] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.*, 1(3):248–272, 2008.
- [41] C. Wu, J. Zhang, and X. Tai. Augmented Lagrangian method for total variation restoration with non-quadratic fidelity. *Inverse Probl. Imaging*, 5(1):237–261, 2011.
- [42] G. Xu, M. Li, A. Gopinath, and C. Bajaj. INVERSION OF ELECTRON TOMOGRAPHY IMAGES USING L2-GRADIENT FLOW—COMPUTATIONAL METHODS. *J. Comput. Math.*, 29:501–25, 2011.
- [43] J. Yang and X. Yuan. Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization. *Math. Comp.*, 82(281):301–329, 2013.
- [44] J. Yang and Y. Zhang. Alternating direction algorithms for ℓ_1 -problems in compressive sensing. *SIAM J. Sci. Comput.*, 33(1):250–278, 2011.
- [45] J. Yang, Y. Zhang, and W. Yin. An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise. *SIAM J. Sci. Comput.*, 31(4):2842–2865, 2009.
- [46] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for ℓ_1 -minimization with applications to compressed sensing. *SIAM J. Imaging Sci.*, 1(1):143–168, 2008.