Nonstationary Iterated Thresholding Algorithms for Image Deblurring

J. Huang[*], M. Donatelli[†], and R. Chan[‡]

Abstract

We propose iterative thresholding algorithms based on the iterated Tikhonov method for image deblurring problems. Our method is similar in idea to the modified linearized Bregman algorithm (MLBA) so is easy to implement. In order to obtain good restorations, MLBA requires an accurate estimate of the regularization parameter α which is hard to get in real applications. Based on previous results in iterated Tikhonov method, we design two nonstationary iterative thresholding algorithms which give near optimal results without estimating α. One of them is based on the iterative soft thresholding algorithm and the other is based on MLBA. We show that the nonstationary methods, if converge, will converge to the same minimizers of the stationary variants. Numerical results show that the accuracy and convergence of our nonstationary methods are very robust with respect to the changes in the parameters and the restoration results are comparable to those of MLBA with optimal α.

1 Introduction

We consider the problem of finding $f \in \mathbb{R}^M$ in the system

$$g = Kf + e, \tag{1}$$

where $g \in \mathbb{R}^N$ is given, $e \in \mathbb{R}^N$ is the noise and $K \in \mathbb{R}^{N\times M}$, $N \leq M$, is an ill-conditioned matrix. Such problem arises in many practical problems in science and engineering [3, 43, 49, 29, 36]. For image deblurring problems, K in (1) can simply be a blurring operator or be written as $K = A\mathcal{D}^T$, where A is a blurring matrix and $\mathcal{D}^T$ is a wavelet or tightframe synthesis operator [13, 7]. In the latter case, f will be the wavelets or tightframe coefficients of the original image.

Due to the ill-conditioning of K, to find an estimation of the solution f, it is necessary to resort to a regularization method [48, 18]. The following iterative soft thresholding algorithm

[*]School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, PR China. This research is supported in part by 973 Program (2013CB329404), NSFC (61170311), Sichuan Province Sci. & Tech. Research Project (2012GZX0080). E-mail: jiehuangjh86@gmail.com

[†]Dipartimento di Scienza e Alta Tecnologia, Università dell'Insubria, Como 22100, Italy. E-mail: marco.donatelli@uninsubria.it

[‡]Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong. Research is supported in part by HKRGC Grant No. CUHK400412, CUHK DAG 4053007, and CUHK FIS Grant 1902036. Email: rchan@math.cuhk.edu.hk.

(ISTA) was proposed independently from several authors [14, 32, 33, 20, 21]. We consider here the formulation in [21]:

$$f^n = \mathbf{S}_\mu(f^{n-1} + K^*(g - Kf^{n-1})), \tag{2}$$

where the nonlinear operator $\mathbf{S}_\mu$ is defined component-wise by

$$[\mathbf{S}_\mu(g)]_i = S_\mu(g_i), \tag{3}$$

with S_μ the soft-thresholding function

$$S_\mu(x) = \operatorname{sgn}(x)\,(|x| - \mu)_+\,.$$

In [21] it was proven that, under the assumption that $\|K\| < 1$, the iteration (2) converges to a minimizer of

$$\Phi_\mu(f) = \|Kf - g\|^2 + 2\mu\|f\|_1, \tag{4}$$

where $\|\cdot\|$ denotes the ℓ^2-norm. Since the convergence in general is slow, many improvements are proposed in literature to accelerate the convergence, see [1, 22, 40] and references therein.

A new fast algorithm is the modified linearized Bregman algorithm introduced in [8]. It is derived in a different way, by considering the constraint minimization

$$\min_f \{\|f\|_1 : Kf = g\}. \tag{5}$$

The method we propose in this paper is also derived by considering the same minimization (5). If f is a solution of (5), it is the only intersection point of the ℓ_1-ball $\mathcal{B} = \{x : \|x\|_1 \le \|f\|_1\}$ and the hyperplane $\mathcal{P} = \{x : Kx = g\}$. Since both $\mathcal{B}$ and $\mathcal{P}$ are convex, the intersection can be determined by alternate projections onto convex sets as proposed in [12]. The projection onto $\mathcal{P}$ is obtained by $\hat{x} = x + K^*(KK^*)^{-1}(g - Kx)$, while the projection onto $\mathcal{B}$ requires the soft-thresholding. Since K is ill-conditioned, one can replace $(KK^*)^{-1}$ by the regularized inverse $(KK^* + \alpha I)^{-1}$, where $\alpha > 0$ is the regularization parameter and I is the identity operator. This results in our method:

$$f^n = \mathbf{S}_\mu(f^{n-1} + K^*(KK^* + \alpha I)^{-1}(g - Kf^{n-1})). \tag{6}$$

We remark that the iteration inside $\mathbf{S}_\mu$ is what is called *iterated Tikhonov iteration*, see [29]. Thus we name our method *iterated Tikhonov thresholding algorithm* (ITTA). Following the analysis in [21], we can prove that it converges to a minimizer of

$$\|Kf - g\|^2_{(KK^*+\alpha I)^{-1}} + 2\mu\|f\|_1,$$

where $\|x\|_P^2 = \langle x, Px\rangle$ for any positive definite matrix P.

We will see that our method (6) is very similar to the modified linearized Bregman algorithm (MLBA), and requires the same computational cost at each iteration. The quality of the restorations of the two methods are comparable. We will see that both methods are very sensitive to the choice of α. If α is slightly off the optimal one, the quality of the restorations can deteriorate quickly, see Figures 1 and 2 in Section 5. For practical problems, it may be difficult to obtain the optimal α. Since our method (6) involves an iterated Tikhonov step, well-known results for choosing optimal α in iterated Tikhonov method, such as those in [34], can be applied here. We apply the strategy in [34] to both ITTA in (6) and MLBA to obtain two nonstationary algorithms. Numerical results show that the resulting nonstationary algorithms provide the solution with near optimal α. Hence there is no need to estimate the optimal α in our algorithms.

The paper is organized as follows. In Section 2 we recall the classical iterated Tikhonov method and some convergence results. In Section 3 we derive our method ITTA, provide a convergence analysis and investigate its relationship with MLBA. In Section 4 we propose nonstationary version for both ITTA and MLBA and prove that the resulting methods, if converge, will converge to the same minimizers of the stationary variants. Numerical examples are given in Section 5 to show that our nonstationary methods are robust against α. Concluding remarks are in Section 6.

2 Iterated Tikhonov Method

A classical approach for computing a regularized solution of (1) is to solve the minimization problem $\min_f \|Kf - g\|^2$ by iterative methods but stop before it converges using, for instance, the discrepancy principle [42, 36]. One example is the Landweber method:

$$f^n = f^{n-1} + \tau K^*(g - Kf^{n-1}), \qquad 0 < \tau < 2/\|K^*K\|. \tag{7}$$

It has the nice property that a precise estimation of the stopping iteration is not crucial [39, 47, 29, 36]. However, it converges slowly. Its convergence can be accelerated by a proper choice of τ or by adding a regularization preconditioner [46]. Alternatively, one can use other iterative methods to get fast convergence. For example, conjugate gradient for normal equations [29] can be used with similar regularization feature. In this paper we consider the iterated Tikhonov method [29].

Iterated Tikhonov method can be defined as an iterative refinement of the Tikhonov method [38, 37, 31]. The classical Tikhonov method computes the solution $\hat{f} = (K^*K + \alpha I)^{-1}K^*g$ of the following minimization problem

$$\min_f \|Kf - g\|^2 + \alpha\|f\|^2, \qquad \alpha > 0.$$

Refining the previous approximation by solving the error equation with the Tikhonov method, we obtain the iterated Tikhonov method:

$$f^n = f^{n-1} + K^*(KK^* + \alpha I)^{-1}(g - Kf^{n-1}). \tag{8}$$

The previous iteration (8) can be interpreted in many different ways: (i) it is a preconditioned Landweber method with regularization preconditioner $(K^*K + \alpha I)^{-1}$ and $\tau = 1$ (see [46]), (ii) it is equivalent to the Levenberg-Marquardt method applied to $\min_f \frac{1}{2}\|Kf - g\|^2$ (see [35]), and (iii) it is a gradient descent method for $\min_f \frac{1}{2}\|Kf - g\|^2_{(KK^*+\alpha I)^{-1}}$. The method is characterized by the *semiconvergence property*: the iteration starts to converge to the true solution for small values of n, but then when n becomes large, it diverges. If α is chosen properly, it can converge to the true solution in few iterations. If α is too large the convergence slows down, while if α is too small the noise is already amplified at the early iterations and the computed solution will not be accurate.

It can be difficult to get a good α in practice, so in [6], the authors proposed the *nonstationary* iterated Tikhonov method:

$$f^n = f^{n-1} + K^*(KK^* + \alpha_n I)^{-1}(g - Kf^{n-1}), \qquad f^0 = 0. \tag{9}$$

Here α_n is changing at each iteration n. Below we give some results on the convergence of (9) for the noise free case, i.e., $e = 0$ in (1). Define

$$\gamma_n := \sum_{j=1}^{n} \alpha_j^{-1}. \tag{10}$$

Theorem 1 ([6]) *The method* (9) *converges to the solution* $f^\star$ *of* $K^*Kf = K^*g$ *if and only if*

$$\lim_{n\to\infty} \gamma_n = \infty. \tag{11}$$

Theorem 2 ([34]) *If* $f^\star = (K^*K)^\nu w$ *for some* $\nu > 0$ *with some* w *in the domain of* $(K^*K)^\nu$, *and if there exists a constant* $c_\nu > 0$ *such that*

$$\alpha_n^{-1} \le c_\nu \gamma_{n-1}, \tag{12}$$

then $\|f^n - f^\star\| \le c_\nu \gamma_n^{-\nu} \|w\|$.

A classical choice for α_n is the successive geometric sequence

$$\alpha_n = \alpha_0 q^{n-1}, \qquad 0 < q < 1. \tag{13}$$

In this case γ_n in (10) satisfies $\gamma_n \ge q^{1-n}/\alpha_0$ and (12) holds with $c_\nu = 1/q$ as shown in [34]. Therefore, from Theorem 2 we have $\|f^n - f^\star\| = \mathcal{O}(q^{\nu n})$ and hence we have a linear rate of convergence. The sequence (13) provides a convergence faster than the stationary case where $\alpha_n = \alpha$ for all $n \in \mathbb{N}$. In the latter case, $\gamma_n = n/\alpha$ and (12) holds with $c_\nu = 1/(n-1)$. Therefore, from Theorem 2 $\|f^n - f^\star\| = \mathcal{O}(n^{-(\nu+1)})$ and hence we only have a sublinear rate of convergence.

For perturbed data, i.e., $e \neq 0$ in (1), it was proven in [34] that the nonstationary iterated Tikhonov (9) with the α_n defined in (13) converges again faster than the stationary method (8). Moreover, numerical experiments in [26] show that the nonstationary method avoids an accurate estimation of optimal α which may not be readily available in real applications. Indeed, α_0 in (13) has only to be chosen large enough (an over-estimation of the optimal α), while q controls how fast the sequence decreases.

3 Iterated Tikhonov with Thresholding

The ISTA iteration in (2) can be viewed as a Landweber method (7) combined with a soft-thresholding, hence it inherits the slow convergence of Landweber method. As mentioned in the last section, the Landweber part can be accelerated in several ways, e.g.,

$$f^n = \mathbf{S}_\mu(f^{n-1} + \tau^n K^*(g - Kf^{n-1})),$$

with a proper choice of τ^n [1, 22, 40]. For instance, the choice

$$\tau^n = \frac{\|K^*(g - Kf^{n-1}\|^2}{\|KK^*(g - Kf^{n-1})\|^2},$$

is equivalent to replacing the inner Landweber iteration with one step of conjugate gradient for normal equation. Following this idea, another possible strategy to speed up the ISTA iteration (2) is to replace the inner Landweber step with a faster convergent iterative regularization method. In this paper we propose to replace it by the iterated Tikhonov method in (8). We thus arrive at our iterated Tikhonov thresholding algorithm (ITTA):

$$f^n = \mathbf{S}_\mu(f^{n-1} + K^*(KK^* + \alpha I)^{-1}(g - Kf^{n-1})). \tag{14}$$

3.1 Convergence analysis of ITTA

To obtain the convergence of ITTA in (14), we first review the convergence of ISTA which was given in [21] for general Hilbert spaces. Here for simplicity, we rewrite the lemma for finite dimensional spaces.

Lemma 3 ([21], Theorem 3.1) *Let $K \in \mathbb{R}^{N\times M}$, $N \leq M$ with $\|K\|_2 < 1$. Then the sequence of iterates* (2)*, with $\mathbf{S}_\mu$ defined in* (3) *and f^0 arbitrarily chosen in $\mathbb{R}^M$, converges to a minimizer of the functional*

$$\Phi_\mu(f) = \|Kf - g\|^2 + 2\mu\|f\|_1 \,. \tag{15}$$

If $\mathrm{Null}(K) = \{0\}$*, then the minimizer $f^\star$ of Φ_μ is unique, and every sequence of iterates f^n converges to $f^\star$ (i.e., $\|f^n - f^\star\| \to 0$), regardless of the choice of f^0.*

Defining $P = (KK^* + \alpha I)^{-1}$ and using $P^{1/2}K$ and $P^{1/2}g$ to replace K and g in Lemma 3, respectively, we have the following convergence theorem for ITTA.

Theorem 4 *Let $K \in \mathbb{R}^{N\times M}$, $N \leq M$ and assume $\mathbf{S}_\mu$ as defined in* (3)*. Then for any given $\alpha > 0$, the sequence of iterates in* (14) *with f^0 arbitrarily chosen in $\mathbb{R}^M$ converges to a minimizer of the functional*

$$\Phi_{\mu,\alpha}(f) = \|Kf - g\|^2_{(KK^*+\alpha I)^{-1}} + 2\mu\|f\|_1 \,. \tag{16}$$

If $\mathrm{Null}(K) = \{0\}$*, then the minimizer $f^\star$ of $\Phi_{\mu,\alpha}$ is unique, and every sequence of iterates f^n converges to $f^\star$ (i.e., $\|f^n - f^\star\| \to 0$), regardless of the choice of f^0.*

Proof: For any given $\alpha > 0$, denote $P = (KK^* + \alpha I)^{-1}$ for convenience. To employ Lemma 3 for the matrix $P^{1/2}K$, we first show that $\|P^{1/2}K\|_2 < 1$. Clearly, we have

$$K^*(KK^* + \alpha I)^{-1} = (K^*K + \alpha I)^{-1}K^*.$$

This equation, combining with the definition of P, gives

$$\begin{aligned}\|P^{1/2}K\|^2 &= \rho\left((P^{1/2}K)^*(P^{1/2}K)\right) = \rho(K^*PK) \leq \|K^*PK\| \\ &= \left\|K^*(KK^* + \alpha I)^{-1}K\right\| = \left\|(K^*K + \alpha I)^{-1}K^*K\right\| < 1,\end{aligned} \tag{17}$$

where $\rho(\cdot)$ is the spectral radius. Thus replacing K and g by $P^{1/2}K$ and $P^{1/2}g$ in Lemma 3, respectively, we find that the sequence of iterates (2) becomes

$$\begin{aligned} f^n &= \mathbf{S}_\mu\left(f^{n-1} + (P^{1/2}K)^*\left(P^{1/2}g - P^{1/2}Kf^{n-1}\right)\right) \\ &= \mathbf{S}_\mu\left(f^{n-1} + K^*P\left(g - Kf^{n-1}\right)\right) \\ &= \mathbf{S}_\mu\left(f^{n-1} + K^*(KK^* + \alpha I)^{-1}\left(g - Kf^{n-1}\right)\right), \quad n = 1, 2, \ldots. \end{aligned}$$

This is exactly the iteration (14). Moreover, it converges to a minimizer of the functional

$$\Phi_{\mu,\alpha}(f) = \left\|P^{1/2}Kf - P^{1/2}g\right\|^2 + 2\mu\|f\|_1 = \|Kf - g\|^2_{(KK^*+\alpha I)^{-1}} + 2\mu\|f\|_1 \,.$$

$\square$

According to Remark 2.4 in [21], if $\{\phi_\gamma\}$ is an orthonormal basis for $\mathbb{R}^M$ and P is diagonalizable in $\{\phi_\gamma\}$, then the surrogate function in (16) with P-norm can be reformulated as the surrogate function in (15) with ℓ_2 norm, but at a different scale μ. Unfortunately, for the image deblurring problems considered in this paper, P is not diagonalizable by $\{\phi_\gamma\}$ which are piecewise linear B-spline framelets.

3.2 Relationship with modified linearized Bregman iteration

Our proposed ITTA in (14) is similar to the modified linearized Bregman algorithm in [8]. To see this clearly, we first recall that linearized Bregman algorithm is a method designed to solve (5) by using the iteration

$$\begin{cases} z^n = z^{n-1} + K^* \left(g - K f^{n-1} \right), \\ f^n = \delta \mathbf{S}_\mu (z^n), \end{cases} \tag{18}$$

where $z^0 = f^0 = 0$; see [45, 50]. The convergence of linearized Bregman algorithm is given in [10, 11] and we restate it as follows.

Lemma 5 ([10], Theorem 2.4) *Assume $K \in \mathbb{R}^{N \times M}$, $N \leq M$, be any surjective matrix and $0 < \delta < \frac{1}{\|KK^*\|}$. Then the sequence $\{f^n\}$ generated by* (18) *converges to the unique solution of*

$$\min_{f \in \mathbb{R}^M} \{\mu \|f\|_1 + \frac{1}{2\delta} \|f\|^2 : Kf = g\}, \tag{19}$$

i.e., $\lim_{n\to\infty} \|f^n - f_\mu^\star\| = 0$, where $f_\mu^\star$ is the unique solution of (19)*. Furthermore,*

$$\lim_{\mu \to \infty} \|f_\mu^\star - f_1\| = 0,$$

where f_1 is the solution of (5) *that has the minimal ℓ_2 norm among all the solutions of* (5)*.*

A modified linearized Bregman algorithm (MLBA) for frame-based image deblurring was proposed and analyzed in [8]. Its corresponding convergence, given in the next lemma, is followed from similar analysis in [10, 11].

Lemma 6 ([8], Theorem 3.2) *Assume P is a symmetric positive definite matrix and let $0 < \delta < \frac{1}{\|KPK^*\|}$. Then the sequence $\{f^n\}$ generated by the modified linearized Bregman algorithm:*

$$\begin{cases} z^n = z^{n-1} + K^* P \left(g - K f^{n-1} \right), \\ f^n = \delta \mathbf{S}_\mu (z^n), \end{cases} \tag{20}$$

converges to the unique solution of

$$\min_{f \in \mathbb{R}^M} \{\mu \|f\|_1 + \frac{1}{2\delta} \|f\|^2 : f = \operatorname{argmin} \|Kf - g\|_P^2\}. \tag{21}$$

Furthermore, as $\mu \to \infty$, the limit of this iteration is the solution of

$$\min_{f \in \mathbb{R}^M} \{\|f\|_1 : f = \operatorname{argmin} \|Kf - g\|_P^2\} \tag{22}$$

that has a minimum ℓ_2 norm among all the solutions of (22)*.*

We note that the matrix P in Lemma 6 serves as a preconditioner to accelerate the convergence. Often P is chosen as $(KK^* + \alpha I)^{-1}$ and thus $\|KPK^*\| < 1$ (see (17)); so that we can fix $\delta = 1$. In this case the modified linearized Bregman iteration (20) becomes

$$\begin{cases} z^n = z^{n-1} + K^*(KK^* + \alpha I)^{-1}(g - K f^{n-1}), \\ f^n = \mathbf{S}_\mu (z^n). \end{cases} \tag{23}$$

Clearly, MLBA in (23) is similar to ITTA in (14). The only difference is in the first part of the first equation in (23) where we replace f^{n-1} in (14) by z^{n-1} in (23). As a result, both methods have the same computational cost. The only small overhead of algorithm (23) is that it requires two vectors for z^{n-1} and f^{n-1}, while ITTA in (14) can be implemented with only one vector.

Finally, we observe that ITTA (14) is modified from ISTA (2) exactly like MLBA (23) is modified from the linearized Bregman iteration (18).

4 Nonstationary Methods

Numerical results in Section 5 will show that the regularization parameter α affects not only the speed of convergence but also the accuracy of the restoration in both ITTA in (14) and MLBA in (23), see Figures 1 and 2 in Section 5. In practice, it could be difficult to have a good estimation of α. Recall in Section 2 one idea to get the optimal result in iterated Tikhonov method is to use a geometric decreasing sequence of α_n, such as those defined in [34]. Applying the same idea to both ITTA in (14) and MLBA in (23), we get two nonstationary methods:

1. Nonstationary iterated Tihkonov thresholding algorithm (NITTA)

$$f^n = \mathbf{S}_\mu \left(f^{n-1} + K^*(KK^* + \alpha_n I)^{-1} \left(g - Kf^{n-1} \right) \right), \quad \alpha_n > 0. \tag{24}$$

2. Nonstationary modified linearized Bregman algorithm (NMLBA):

$$\begin{cases} z^n = z^{n-1} + K^* \left(KK^* + \alpha_n I \right)^{-1} \left(g - Kf^{n-1} \right), \\ f^n = \mathbf{S}_\mu(z^n), \end{cases} \quad \alpha_n > 0. \tag{25}$$

We will see in the numerical results that the nonstationary methods are much more robust against the parameter α than their stationary variants.

In the following, we derive convergence limits of both nonstationary methods (24) and (25) under the following assumption.

Assumption 1 *The sequence of regularization parameters $\alpha_n > 0$ satisfies*

$$\lim_{n\to\infty} \alpha_n = \bar{\alpha}, \qquad 0 < \bar{\alpha} < \infty.$$

Notice that Assumption 1 implies (11) and hence the nonstationary iterated Tikhonov method (9) with such $\{\alpha_n\}$ will converge according to Theorem 1.

4.1 Convergence limit of nonstationary iterated Tikhonov thresholding algorithm (NITTA)

In this subsection, we derive the convergent limit of NITTA in (24) if it converges.

Theorem 7 *Let $K \in \mathbb{R}^{N\times M}$, $N \leq M$, be any matrix and assume $\mathbf{S}_\mu$ as defined in* (3)*. With Assumption 1, if the sequence $\{f^n\}$ of iterates* (24) *in NITTA converges, then its limit is a minimizer of the ITTA functional in* (16) *with $\alpha = \bar{\alpha}$:*

$$\Phi_{\mu,\bar{\alpha}}(f) = \|Kf - g\|^2_{(KK^* + \bar{\alpha} I)^{-1}} + 2\mu \|f\|_1 .$$

If $\mathrm{Null}(K) = \{0\}$*, then the minimizer $f^\star$ of $\Phi_{\mu,\bar{\alpha}}$ is unique, and every convergent sequence $\{f^n\}$ converges to $f^\star$ (i.e., $\|f^n - f^\star\| \to 0$).*

Proof: Assume f^n converges to $f^\star$. From (24) and Assumption 1, as $n \to \infty$, we have

$$f^\star = \mathbf{S}_\mu \left(f^\star + K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star) \right). \tag{26}$$

We now prove that $f^\star$ is a minimizer of $\Phi_{\mu,\bar{\alpha}}$. Since $\Phi_{\mu,\bar{\alpha}}$ is convex, using the first-order optimality condition of $\Phi_{\mu,\bar{\alpha}}$ we equivalently show that $0 \in \partial\Phi_{\mu,\bar{\alpha}}(f^\star)$, where ∂ denotes the subdifferential. It is clear that

$$\partial\Phi_{\mu,\bar{\alpha}}(f^\star) = 2K^*(KK^* + \bar{\alpha} I)^{-1}(Kf^\star - g) + 2\mu\partial\|f^\star\|_1.$$

In the following we will prove $0 \in K^*(KK^* + \bar{\alpha} I)^{-1}(Kf^\star - g) + \mu\partial\|f^\star\|_1$. It means that for each index i, we need to show that

$$0 \in \left(K^*(KK^* + \bar{\alpha} I)^{-1}(Kf^\star - g)\right)_i + \mu\left(\partial\|f^\star\|_1\right)_i = \left(K^*(KK^* + \bar{\alpha} I)^{-1}(Kf^\star - g)\right)_i + \mu\partial|f_i^\star|.$$

This relation is established by considering three cases.

(i) If $f_i^\star + \left(K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star)\right)_i < -\mu$, then the definition of soft-thresholding and (26) give

$$\begin{aligned} f_i^\star &= S_\mu \left(f_i^\star + \left(K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star)\right)_i \right) \\ &= f_i^\star + \left(K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star)\right)_i + \mu < 0. \end{aligned}$$

It follows that $0 = \left(K^*(KK^* + \bar{\alpha} I)^{-1} (Kf^\star - g)\right)_i + \mu \cdot (-1)$. Since $f_i^\star < 0$, we have $\partial|f_i^\star| = \{-1\}$, and hence

$$0 \in \left(K^*(KK^* + \bar{\alpha} I)^{-1} (Kf^\star - g)\right)_i + \mu\partial|f_i^\star|.$$

(ii) If $\left|f_i^\star + \left(K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star)\right)_i\right| \le \mu$, then soft-thresholding and (26) give

$$f_i^\star = S_\mu \left(f_i^\star + \left(K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star)\right)_i \right) = 0.$$

Thus $\left|\left(K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star)\right)_i\right| \le \mu$. Let $p_0 = -\frac{1}{\mu} \left(K^*(KK^* + \bar{\alpha} I)^{-1} (Kf^\star - g)\right)_i$, then $|p_0| \le 1$ and $0 = \left(K^*(KK^* + \bar{\alpha} I)^{-1} (Kf^\star - g)\right)_i + \mu \cdot p_0$. It says that

$$0 \in \left\{ \left(K^*(KK^* + \bar{\alpha} I)^{-1} (Kf^\star - g)\right)_i + \mu \cdot p, \quad p \in [-1, 1] \right\}.$$

Since $f_i^\star = 0$, we have $\partial|f_i^\star| = [-1, 1]$, and hence

$$0 \in \left(K^*(KK^* + \bar{\alpha} I)^{-1} (Kf^\star - g)\right)_i + \mu\partial|f_i^\star|.$$

(iii) The case where $f_i^\star + \left(K^*(KK^* + \bar{\alpha} I)^{-1} (g - Kf^\star)\right)_i > \mu$ can be established analogously as case (i), we therefore will omit it.

Combining cases (i), (ii) and (iii), we have established the result.

Moreover, if $\mathrm{Null}(K) = \{0\}$, then $\|Kf - g\|^2_{(KK^* + \bar{\alpha} I)^{-1}}$ is strictly convex in f. Thus $\Phi_{\mu,\bar{\alpha}}$ is strictly convex so that it has a unique minimizer. □

We note that under Assumption 1, NITTA in (24), if converges, will converge to the same minimization functional as ITTA with $\alpha = \bar{\alpha}$.

We also note that the condition $\lim_{n\to\infty}\alpha_n = \bar{\alpha} > 0$ is only sufficient and it is stronger than the necessary and sufficient condition in (11). Nevertheless, similarly to the decreasing geometric sequence (13), we choose

$$\alpha_n = \alpha_0 q^{n-1} + \bar{\alpha}, \qquad 0 < \bar{\alpha} < \alpha_0,\ 0 < q < 1, \tag{27}$$

in our numerical tests such that $\lim_{n\to\infty}\alpha_n = \bar{\alpha} > 0$. The parameters α_0, q and $\bar{\alpha}$ in (27) do not need to be estimated accurately to give the best results, see Section 5. They can be fixed easily by keeping in mind that α_0 should be large enough, $\bar{\alpha}$ should be small (a good regularization parameter for the method (14) has to be in the interval $[\bar{\alpha}, \alpha_0]$), and q controls how fast α_n approximates the limiting value $\bar{\alpha}$.

4.2 Convergence limit of nonstationary modified linearized Bregman algorithm (NMLBA)

We now derive the convergence limit of NMLBA defined in (25). Define the *Bregman distance* as

$$D_J^p(u,v) = J(u) - J(v) - \langle u - v, p\rangle,$$

where J is a convex function, $p \in \partial J(v)$ is a subgradient in the subdifferential of J at the point v, see, e.g., [5] for details. The linearized Bregman algorithm is: given $f^0 = p^0 = 0$,

$$\begin{cases} f^n = \arg\min_{f\in\mathbb{R}^M}\left\{\frac{1}{2\delta}\left\|f - \left(f^{n-1} - \delta K^*\left(Kf^{n-1} - g\right)\right)\right\|^2 + \mu D_J^{p^{n-1}}(f, f^{n-1})\right\}, \\ p^n = p^{n-1} - \frac{1}{\mu\delta}\left(f^n - f^{n-1}\right) - \frac{1}{\mu}K^*\left(Kf^{n-1} - g\right). \end{cases}$$

If $J(f) = \|f\|_1$, with a change of variable for p, the iteration can be rewritten as the compact form (18); see [11]. We note that its convergence analysis has been stated in Lemma 5.

For the nonstationary case, define

$$\tilde{K}_n = (KK^* + \alpha_n I)^{-\frac{1}{2}}K, \quad \tilde{g}_n = (KK^* + \alpha_n I)^{-\frac{1}{2}}g, \quad \alpha_n > 0. \tag{28}$$

Given $f^0 = p^0 = 0$, we iterate

$$\begin{cases} f^n = \arg\min_{f\in\mathbb{R}^M}\left\{\frac{1}{2\delta}\left\|f - \left(f^{n-1} - \delta\tilde{K}_n^*\left(\tilde{K}_n f^{n-1} - \tilde{g}_n\right)\right)\right\|^2 + \mu D_J^{p^{n-1}}(f, f^{n-1})\right\}, \\ p^n = p^{n-1} - \frac{1}{\mu\delta}(f^n - f^{n-1}) - \frac{1}{\mu}\tilde{K}_n^*(\tilde{K}_n f^{n-1} - \tilde{g}_n). \end{cases} \tag{29}$$

Then similarly we obtain the compact form of the iterations:

$$\begin{cases} z^n = z^{n-1} + \tilde{K}_n^*\left(\tilde{g}_n - \tilde{K}_n f^{n-1}\right), \\ f^n = \delta\mathbf{S}_\mu(z^n). \end{cases} \tag{30}$$

We note from (28) that this iteration with $\delta = 1$ is exactly NMLBA in (25). Moreover, using a similar analysis as in [10, 11] for linearized Bregman iteration and Assumption 1, we will derive the convergence limit of NMLBA in the following.

First, we give two lemmas. For the first equation of (29), we have the following result from [50], see also Lemma 3.1 in [10].

Lemma 8 ([50]) *Assume that* $\|\tilde{K}_n^*\tilde{K}_n\| < 1/\delta$. *Then*

$$\|\tilde{K}_n f^n - \tilde{g}_n\|^2 + \left(\frac{1}{\delta} - \|\tilde{K}_n^*\tilde{K}_n\|\right)\|f^n - f^{n-1}\|^2 \le \|\tilde{K}_n f^{n-1} - \tilde{g}_n\|^2.$$

Lemma 9 *Assume Assumption 1 holds. Let* $P = (KK^* + \bar{\alpha} I)^{-1}$*. If* $\{f^n\}$ *generated by* (30) *converges, then*

$$\lim_{n\to\infty} K^* P(g - Kf^n) = 0. \tag{31}$$

Consequently, we have $\lim_{n\to\infty} K^*(g - Kf^n) = 0$.

Proof: From the assumption that $\{f^n\}$ converges, we obtain that $\{K^*P(g - Kf^n)\}$ also converges. Let $\lim_{n\to\infty} K^*P(g - Kf^n) = d$, we now prove that $d = 0$ by contradiction.

Assume $d \neq 0$. Let $P_n = (KK^* + \alpha_n I)^{-1}$, then $\tilde{K}_n$ and $\tilde{g}_n$ defined in (28) can be rewritten as $\tilde{K}_n = P_n^{\frac{1}{2}} K$ and $\tilde{g}_n = P_n^{\frac{1}{2}} g$. From the first equation of (30), for any positive integer l, we have

$$\begin{aligned} z^{n+l} - z^n &= \sum_{j=1}^{l} \tilde{K}^*_{n+j}(\tilde{g}_{n+j} - \tilde{K}_{n+j} f^{n+j-1}) \\ &= \sum_{j=1}^{l} \tilde{K}^*_{n+j}(\tilde{g}_{n+j} - \tilde{K}_{n+j} f^{n}) + \sum_{j=1}^{l} \tilde{K}^*_{n+j}\tilde{K}_{n+j}(f^n - f^{n+j-1}) \\ &= \sum_{j=1}^{l} K^* P_{n+j}(g - Kf^n) + \sum_{j=1}^{l} K^* P_{n+j} K(f^n - f^{n+j-1}). \end{aligned}$$

By the assumption that $\{f^n\}$ converges and Assumption 1, we have

$$\lim_{n\to\infty} P_n = P,\ \lim_{n\to\infty} \|f^{n+1} - f^n\| = 0,\ \lim_{n\to\infty} K^* P_{n+j}(g - Kf^n) = \lim_{n\to\infty} K^* P(g - Kf^n).$$

Thus,

$$\lim_{n\to\infty}(z^{n+l} - z^n) = l \cdot \lim_{n\to\infty} K^* P(g - Kf^n) = ld.$$

Therefore, there exists an n_0 such that for all $n > n_0$, $\|z^{n+l} - z^n - ld\| \leq 1$. Hence

$$\|z^{n+l}\| \geq \|z^n + ld\| - 1 \geq l\|d\| - \|z^n\| - 1. \tag{32}$$

Notice that $f^n = \mathbf{S}_\mu(z^n)$ and $\{f^n\}$ converges, hence $\{z^n\}$ is bounded, i.e., there exists a $c > 0$, $\|z^n\| \leq c, \forall n$. However, if we choose $\lceil l = (2c+2)/\|d\| \rceil$ which is finite because $d \neq 0$, then from (32) we have $\|z^{n+l}\| \geq c + 1$. We arrive at a contradiction.

Notice that $K^*P = (K^*K + \bar{\alpha} I)^{-1} K^*$ and $(K^*K + \bar{\alpha} I)^{-1}$ is symmetric positive definite, then from $\lim_{n\to\infty} K^*P(g - Kf^n) = 0$ we get $\lim_{n\to\infty} K^*(g - Kf^n) = 0$. □

Following the analysis in [8], we derive the convergence limit of NMLBA. We note that the uniqueness of the solution of (21) is guaranteed by the strict convexity of the object function of (21).

Theorem 10 *With Assumption 1, if the sequence* $\{f^n\}$ *generated by* (25) *converges, then its limit is the unique solution of* (21) *with* $\delta = 1$ *and* $P = (KK^* + \bar{\alpha} I)^{-1}$*. More precisely, let* $f^\star_\mu$ *be the unique solution of*

$$\min_{f\in\mathbb{R}^M} \{\mu \|f\|_1 + \frac{1}{2}\|f\|^2 : f = \operatorname{argmin} \|Kf - g\|^2_{(KK^*+\bar{\alpha} I)^{-1}}\}, \tag{33}$$

then $\lim_{n\to\infty} \|f^n - f^\star_\mu\| = 0$*. Furthermore,* $\lim_{\mu\to\infty} \|f^\star_\mu - f_1\| = 0$*, where* f_1 *is the solution of* (22) *that has the minimal* ℓ_2 *norm among all the solutions of* (22).

Proof: We first prove that $\{f^n\}$ converges to $f^\star_\mu$. Let $\lim_{n\to\infty} f^n = \tilde{f}$, we now show that $\tilde{f} = f^\star_\mu$. Let $P_n = (KK^* + \alpha_n I)^{-1}$. From (30) with $z^0 = 0$, we obtain

$$z^n = z^{n-1} + K^* P_n(g - Kf^{n-1}) = \sum_{j=1}^{n} K^* P_j(g - Kf^{j-1}) = K^* \sum_{j=1}^{n} P_j(g - Kf^{j-1}).$$

Define $w^n = \sum_{j=1}^n P_j(g - Kf^{j-1})$. Then $z^n = K^* w^n$. Decompose $w^n = w^n_{\mathrm{Ran(K)}} + w^n_{\mathrm{Ker(K^*)}}$, where $w^n_{\mathrm{Ran(K)}}$ is in the range of K and $w^n_{\mathrm{Ker(K^*)}}$ is in the kernel of K^*. It is clear that

$$z^n = K^* w^n_{\mathrm{Ran(K)}}. \tag{34}$$

It is easy to check that K^* is one-to-one from Ran(K) to $\mathbb{R}^M$ and z^n is bounded. Hence $w^n_{\mathrm{Ran(K)}}$ is bounded, i.e., $\|w^n_{\mathrm{Ran(K)}}\| \le C$ for all n.

Notice that $f^n = \mathbf{S}_\mu(z^n) = \operatorname{argmin} \frac{1}{2}\|z^n - f\|^2 + \mu\|f\|_1$, thus $z^n - f^n \in \partial\mu\|f^n\|_1$, where ∂ denotes the subdifferential. Rewrite $z^n = f^n + (z^n - f^n)$, and define

$$H(f) = \mu\|f\|_1 + \frac{1}{2}\|f\|^2.$$

We have $z^n \in \partial H(f^n)$. Thus, by the definition of subdifferential and (34), we have

$$H(f^n) \le H(f^\star_\mu) - \langle f^\star_\mu - f^n, z^n\rangle = H(f^\star_\mu) - \langle K(f^\star_\mu - f^n), w^n_{\mathrm{Ran(K)}}\rangle.$$

For the second term of the above equation and using the Cauchy-Schwarz inequality, we get

$$\left|\langle K(f^\star_\mu - f^n), w^n_{\mathrm{Ran(K)}}\rangle\right| \le \|K(f^\star_\mu - f^n)\| \cdot \|w^n_{\mathrm{Ran(K)}}\| \le C\|K(f^\star_\mu - f^n)\|.$$

From Lemma 9, we have $\lim_{n\to\infty} K^*(g - Kf^n) = 0$. Thus, $\lim_{n\to\infty} Kf^n = g_{\mathrm{Ran(K)}}$. Recall that $f^\star_\mu$ is the unique solution of (33), hence it satisfies $K^*PKf^\star_\mu = K^*Pg$ and thus $(K^*K + \bar{\alpha}I)^{-1}K^*(g - Kf^\star_\mu) = 0$. Consequently, we have $Kf^\star_\mu = g_{\mathrm{Ran(K)}}$. Therefore, $\lim_{n\to\infty}\|K(f^\star_\mu - f^n)\| = 0$ and $\lim_{n\to\infty}\left|\langle K(f^\star_\mu - f^n), w^n_{\mathrm{Ran(K)}}\rangle\right| = 0$. It follows that $H(\tilde{f}) = \lim_{n\to\infty} H(f^n) \le H(f^\star_\mu)$. The uniqueness of $f^\star_\mu$ forces $f^\star_\mu = \tilde{f}$.

The proof of the remaining part of the theorem, i.e., as $\mu \to \infty$, $\{f^\star_\mu\}$ converges to the solution of (22) that has the minimal ℓ_2 norm among all the solutions of (22), is exactly the same as in Theorem 4.4 of [11]. We therefore omit it here. □

We also note that with $\delta = 1$ and Assumption 1, if NMLBA in (25) converges, then both nonstationary and stationary modified linearized Bregman iteration converge to a minimizer of (33).

5 Numerical Results for Image Deblurring

In this section, we apply the algorithms we derived in the previous section to deblur images corrupted by Gaussian noise. We use the tight-frame synthesis approach [8, 21] and consider $K = AD^T$ where A is a blurring operator and D^T is a tight-frame synthesis operator. 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 [23, 16, 17]. The tight-frame we used in our tests is the piecewise linear B-spline framelets given in [8, 9]. Replacing K by AD^T in our ITTA in (14), we obtain

$$f^n = \mathbf{S}_\mu\left(f^{n-1} + DA^*(AA^* + \alpha I)^{-1}(g - AD^T f^{n-1})\right). \tag{35}$$

Similarly, the MLBA in (23) becomes

$$\begin{cases} z^n = z^{n-1} + DA^*(AA^* + \alpha I)^{-1}(g - AD^T f^{n-1}), \\ f^n = \mathbf{S}_\mu(z^n). \end{cases} \tag{36}$$

The nonstationary variants of the two methods are obtained by replacing α with α_n in (35) and (36). Therefore the methods (35), (36) and their nonstationary variants all have the same computational cost. In the tests, we compare all four methods. We refer to [8, 9] for a large experimentation that compares MLBA (i.e. iteration (36)) with other recent methods, including the iterative soft thresholding algorithm.

In the experiment, we assume periodic boundary conditions on all images, and hence A can be diagonalized by 2D Fast Fourier Transforms (FFTs) and $(AA^* + \alpha I)^{-1}$ is easily computable. In fact, the matrix vector product with the matrix $A^*(AA^* + \alpha I)^{-1}$ requires only 2 FFTs. For other boundary conditions associated to fast trigonometric transforms that could be used as well, see [44, 25] and the references therein.

Our tests were done by using MATLAB 7.10.0 (R2010a) on an HP laptop with Intel(R) Core(TM) i5 CPU 2.27 GHz and 2 GB memory. The floating-point precision is 10^{-16}. The initial guess of each algorithm is set to be the zero vector. According to [8], we stop all methods using the discrepancy principle, i.e., stop at the first iteration n such that $\|AD^T f^n - g\| \leq (1 + 10^{-15})\|e\|$, or when the number of iterations reached 300. Here e is the error defined in (1). Note that for an image with N pixels, $\|e\|^2 = N\sigma^2$, where σ^2 is the noise variance which can be estimated by using the median rule [41] in practice. The accuracy of the solution is measured by the PSNR value which is defined by $20\log_{10}\frac{255\times N}{\|f-\tilde{f}\|_2}$ with f and $\tilde{f}$ being the original and restored images.

For the stationary methods (35) and (36), the optimal parameters $\mu^\dagger$ and $\alpha^\dagger$ are chosen by trial and error. For their nonstationary variants, we will use the same $\mu^\dagger$ but with the following decreasing sequences of α_n replacing α in (35) and (36):

1. nonstationary iterated Tikhonov thresholding (NITTA): $\alpha_n = \alpha_0 \times 0.95^{n-1} + 10^{-15}$,
2. nonstationary modified linearized Bregman (NMLBA): $\alpha_n = \alpha_0 \times 0.9^{n-1} + 10^{-15}$.

Note that these choices of α_n satisfy Assumption 1. For each test, we tried four different choice of α_0: $\alpha_0 = 2\alpha^\dagger$, $10\alpha^\dagger$, $100\alpha^\dagger$ and 0.5 to show the robustness of our methods with respect to α_0. We will see that the value of α_0 does not affect the accuracy of the restoration but only the number of iterations.

We have performed five tests.

Example 5.1. The true image is the 256×256 *Cameraman* image. It is distorted by a 15×15 Gaussian blur with standard deviation 2 and a white Gaussian noise with standard deviation $\sigma = 2, 5, 10$.

Example 5.2. Same *Cameraman* image is blurred by a 7×7 disk kernel generated by MATLAB command `fspecial('disk',3)` and then corrupted by a Gaussian white noise with standard deviation $\sigma = 2, 5, 10$.

Example 5.3. The 256×256 *Bridge* image is corrupted by a 9×15 motion kernel generated by `fspecial('motion',15,30)` and a Gaussian white noise with standard deviation $\sigma = 2, 5, 10$.

Example 5.4. The 256×256 *Boat* image is blurred by `fspecial('average',9)` and degraded by a Gaussian white noise with standard deviation $\sigma = 3$.

Example 5.5. The 256×256 *Peppers* is corrupted by a 15×15 Gaussian kernel with standard deviation 2 and a Gaussian white noise with standard deviation $\sigma = 7$.

In Tables 1–4, we report the restoration results by all four methods. We see that the value of α in the stationary methods ITTA and MLBA affects not only the speed of the convergence

Method		$\sigma = 2$		$\sigma = 5$		$\sigma = 10$	
		PSNR($\alpha^\dagger$)	Itr.	PSNR($\alpha^\dagger$)	Itr.	PSNR($\alpha^\dagger$)	Itr.
MLBA	$\alpha^\dagger_{\text{MLBA}}/10$	23.23	21	22.03	10	21.3	8
	$\alpha^\dagger_{\text{MLBA}}$	25.49(0.02)	33	24.73(0.04)	16	24.04(0.08)	14
	$\alpha^\dagger_{\text{MLBA}} \times 10$	25.13	87	24.42	48	23.74	46
NMLBA	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 2$	25.23	28	24.71	16	24.03	14
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 10$	25.48	38	24.65	24	23.93	22
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 100$	25.49	56	24.55	42	23.86	41
	$\alpha_0 = 0.5$	25.49	44	24.62	25	23.96	19
ITTA	$\alpha^\dagger_{\text{ITTA}}/10$	24.42	8	24.15	2	23.31	2
	$\alpha^\dagger_{\text{ITTA}}$	25.47(0.05)	117	24.65(0.2)	89	23.89(0.3)	53
	$\alpha^\dagger_{\text{ITTA}} \times 10$	24.66	300	23.15	300	21.89	300
NITTA	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 2$	25.37	36	24.53	28	23.84	25
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 10$	25.41	62	24.53	57	23.82	54
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 100$	25.38	102	24.48	101	23.78	99
	$\alpha_0 = 0.5$	25.41	62	24.54	32	23.83	22

Table 1: Restoration results for different algorithms for Example 5.1.

but also the accuracy of the restoration. In fact if α is underestimated, the PSNR can drop significantly. If α is overestimated, the PSNR also drops while the number of iterations required for convergence increases rapidly. Thus for these two stationary methods, α should be carefully chosen.

For the nonstationary methods, the PSNR values are very robust with respect to the parameter α_0. For different α_0, the changes in PSNR is always within 0.1dB except for one case (NMLBA in Table 1 for $\sigma = 2$). Moreover, the best PSNR values they achieve are very close to the optimal one obtained by their stationary counterparts. In fact, for all test cases, the differences between the best PSNR's from MLBA and NMLBA are always within 0.1dB. We note that NMLBA performs better than NITTA both in terms of accuracy and speed and hence *NMLBA is the best method among the four if one has no idea what the best α should be.*

To further analyze the difference between the stationary and nonstationary methods, we plot α against PSNR of the solution and iteration number for stationary methods ITTA and MLBA, and α_0 against the same for nonstationary methods NITTA and NMLBA for Example 5.1, see Figures 1 and 2. (The graphs for other examples are similar so we omit them here.) The figures clearly illustrate that one has to get a very good estimate of the optimal α for the stationary methods while α_0 for the nonstationary methods can be chosen to be any reasonably large number. In particular, $\alpha_0 = 0.5$ can be a good choice for nonstationary methods for all the examples.

In Figures 3–4, we give the restored images by all four methods with different choices of α and α_0 for Examples 5.2 and 5.4. (The images for other examples show similar conclusion, so again we omit them here.) Clearly, nonstationary methods provide comparable restorations in visual quality to stationary methods with optimal $\alpha^\dagger$.

Method		$\sigma = 2$		$\sigma = 5$		$\sigma = 10$	
		PSNR($\alpha^\dagger$)	Itr.	PSNR($\alpha^\dagger$)	Itr.	PSNR($\alpha^\dagger$)	Itr.
MLBA	$\alpha^\dagger_{\text{MLBA}}/10$	20.30	16	16.62	8	19.65	8
	$\alpha^\dagger_{\text{MLBA}}$	27.70(0.04)	34	25.61(0.06)	17	24.51(0.2)	20
	$\alpha^\dagger_{\text{MLBA}} \times 10$	26.85	153	25.12	65	24.24	93
NMLBA	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 2$	26.81	26	25.54	16	24.50	18
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 10$	27.70	37	25.53	26	24.47	29
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 100$	27.71	58	25.45	46	24.42	49
	$\alpha_0 = 0.5$	27.70	39	25.56	25	24.50	19
ITTA	$\alpha^\dagger_{\text{ITTA}}/10$	25.76	2	24.13	2	23.11	2
	$\alpha^\dagger_{\text{ITTA}}$	27.50(0.1)	108	25.46(0.3)	69	24.35(0.6)	46
	$\alpha^\dagger_{\text{ITTA}} \times 10$	24.93	300	23.37	300	21.92	300
NITTA	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 2$	27.26	31	25.29	27	24.26	24
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 10$	27.33	60	25.28	57	24.28	55
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 100$	27.27	104	25.24	102	24.24	99
	$\alpha_0 = 0.5$	27.32	48	25.29	24	24.13	10

Table 2: Restoration results for all algorithms for Example 5.2.

Method		$\sigma = 2$		$\sigma = 5$		$\sigma = 10$	
		PSNR($\alpha^\dagger$)	Itr.	PSNR($\alpha^\dagger$)	Itr.	PSNR($\alpha^\dagger$)	Itr.
MLBA	$\alpha^\dagger_{\text{MLBA}}/10$	18.61	7	15.93	7	17.10	5
	$\alpha^\dagger_{\text{MLBA}}$	23.68(0.02)	11	21.75(0.07)	15	20.60(0.2)	13
	$\alpha^\dagger_{\text{MLBA}} \times 10$	23.12	50	21.39	76	20.39	68
NMLBA	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 2$	23.64	12	21.71	15	20.59	13
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 10$	23.58	22	21.64	25	20.54	24
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 100$	23.58	43	21.63	46	20.50	44
	$\alpha_0 = 0.5$	23.56	30	21.69	23	20.57	14
ITTA	$\alpha^\dagger_{\text{ITTA}}/10$	22.87	2	20.67	2	18.38	2
	$\alpha^\dagger_{\text{ITTA}}$	23.63(0.07)	36	21.67(0.2)	32	20.53(0.3)	14
	$\alpha^\dagger_{\text{ITTA}} \times 10$	21.26	300	19.83	300	19.22	300
NITTA	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 2$	23.61	24	21.64	23	20.53	18
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 10$	23.54	53	21.57	53	20.48	48
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 100$	23.47	97	21.54	98	20.46	93
	$\alpha_0 = 0.5$	23.57	47	21.63	27	20.53	15

Table 3: Restoration results for all algorithms for Example 5.3.

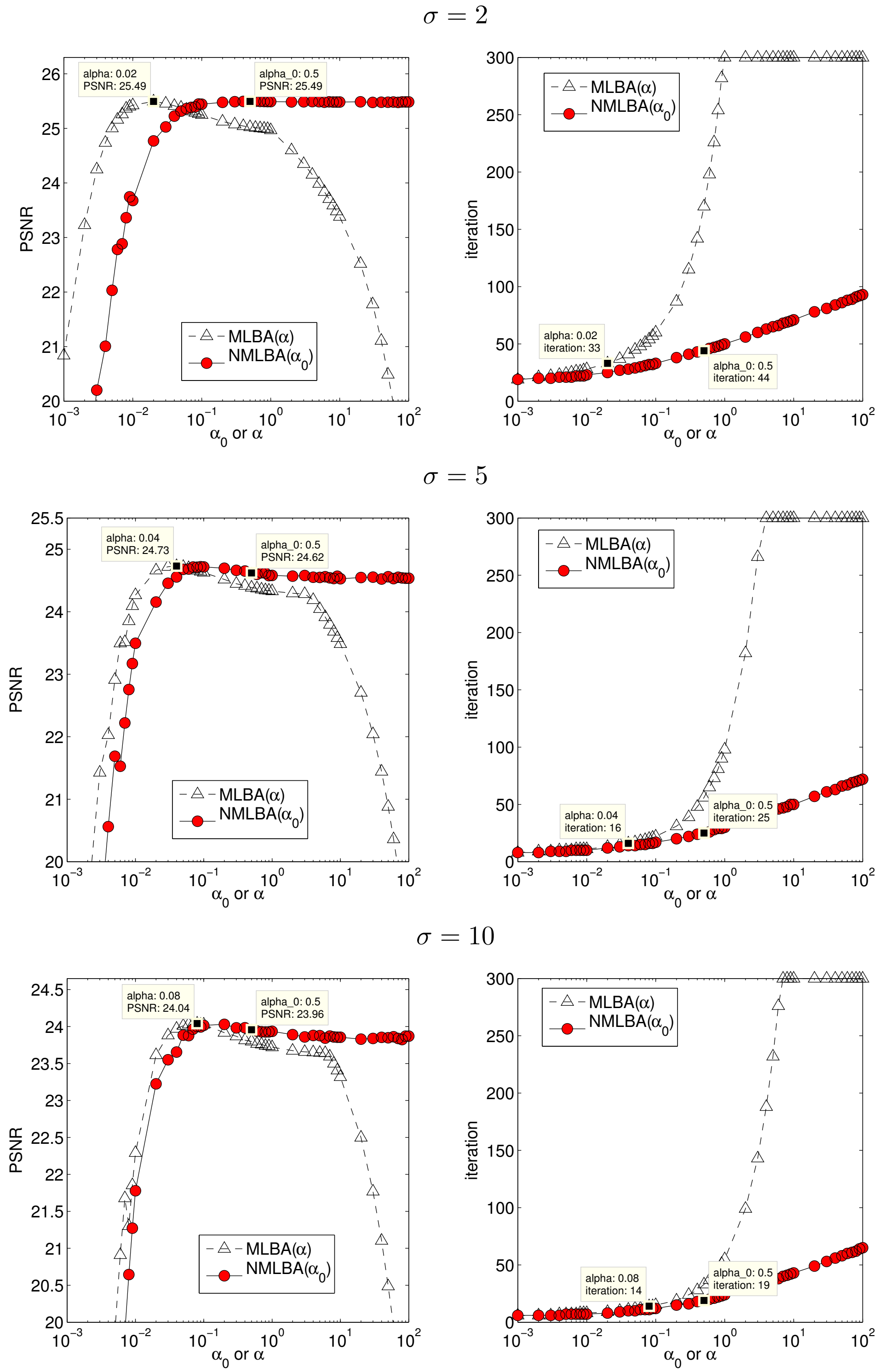


Figure 1: Plots of α and α_0 against PSNR and iteration for Example 5.1.

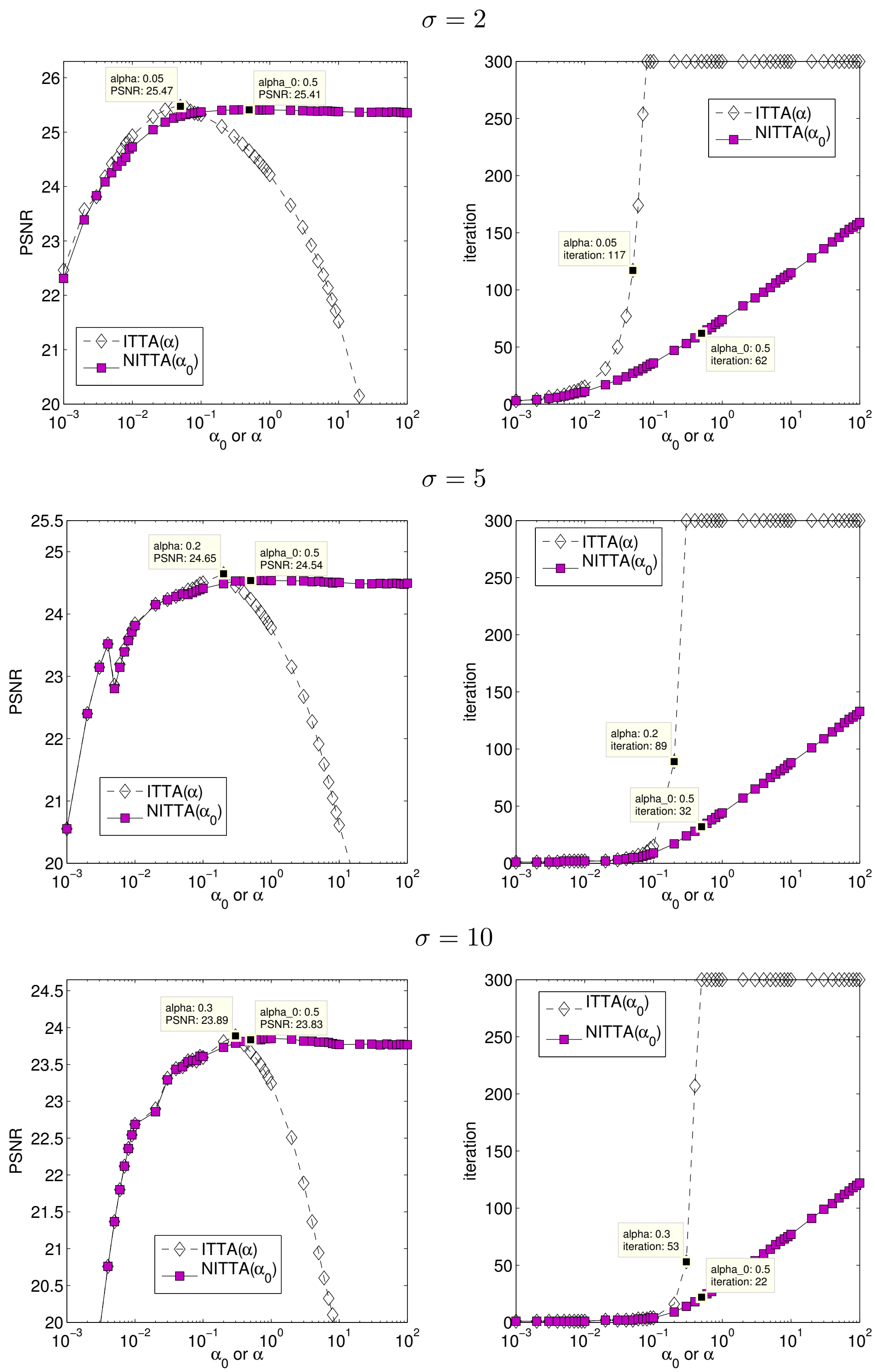


Figure 2: Plots of α and α_0 against PSNR and iteration for Example 5.1.

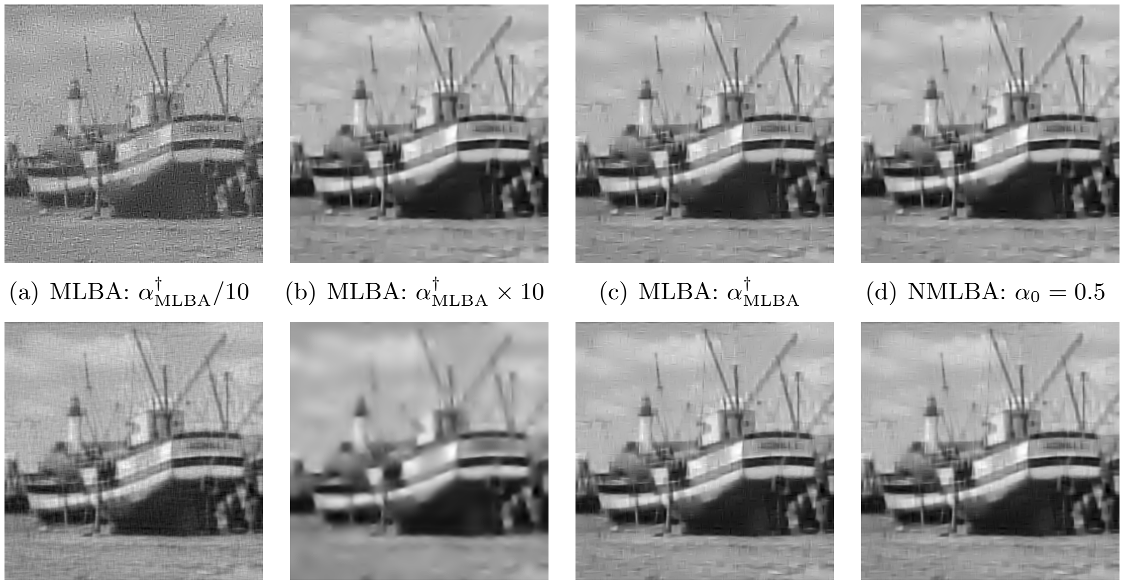
Method		Example 5.4		Example 5.5	
		PSNR($\alpha^\dagger$)	Itr.	PSNR($\alpha^\dagger$)	Itr.
MLBA	$\alpha^\dagger_{\text{MLBA}}/10$	21.10	7	22.39	10
	$\alpha^\dagger_{\text{MLBA}}$	25.96(0.04)	14	24.95(0.07)	17
	$\alpha^\dagger_{\text{MLBA}} \times 10$	25.60	77	24.64	52
NMLBA	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 2$	25.95	14	24.93	16
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 10$	25.90	25	24.86	24
	$\alpha_0 = \alpha^\dagger_{\text{MLBA}} \times 100$	25.81	45	24.79	43
	$\alpha_0 = 0.5$	25.81	26	24.89	22
ITTA	$\alpha^\dagger_{\text{ITTA}}/10$	25.15	2	24.19	2
	$\alpha^\dagger_{\text{ITTA}}$	25.77(0.2)	48	24.84(0.3)	56
	$\alpha^\dagger_{\text{ITTA}} \times 10$	23.29	300	22.97	300
NITTA	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 2$	25.64	26	24.74	25
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 10$	25.65	57	24.74	54
	$\alpha_0 = \alpha^\dagger_{\text{ITTA}} \times 100$	25.57	101	24.64	98
	$\alpha_0 = 0.5$	25.65	30	24.73	22

Table 4: Restoration results for all algorithms for Examples 5.4 and 5.5.

(a) MLBA: $\alpha^\dagger_{\text{MLBA}}/10$ (b) MLBA: $\alpha^\dagger_{\text{MLBA}} \times 10$ (c) MLBA: $\alpha^\dagger_{\text{MLBA}}$ (d) NMLBA: $\alpha_0 = 0.5$

(e) ITTA: $\alpha^\dagger_{\text{ITTA}}/10$ (f) ITTA: $\alpha^\dagger_{\text{ITTA}} \times 10$ (g) ITTA: $\alpha^\dagger_{\text{ITTA}}$ (h) NITTA: $\alpha_0 = 0.5$

Figure 3: Deblurring results for all four algorithms for Example 5.2 with $\sigma = 10$.



(a) MLBA: $\alpha^{\dagger}_{\text{MLBA}}/10$ (b) MLBA: $\alpha^{\dagger}_{\text{MLBA}} \times 10$ (c) MLBA: $\alpha^{\dagger}_{\text{MLBA}}$ (d) NMLBA: $\alpha_0 = 0.5$

(e) ITTA: $\alpha^{\dagger}_{\text{ITTA}}/10$ (f) ITTA: $\alpha^{\dagger}_{\text{ITTA}} \times 10$ (g) ITTA: $\alpha^{\dagger}_{\text{ITTA}}$ (h) NITTA: $\alpha_0 = 0.5$

Figure 4: Deblurring results for all four algorithms for Example 5.4.

6 Conclusions

Combining iterated Tikhonov method with soft-thresholding we obtain our iterated Tikhonov thresholding algorithm (ITTA) that has a very similar form as the modified linearized Bregman iteration (MLBA). Inspired by the results on nonstationary iterated Tikhonov we propose to vary the regularization parameter α in each iteration, and we arrive at the nonstationary version of these two algorithms. We proved the convergence of ITTA and gave some convergence properties of the nonstationary version of ITTA and MLBA under Assumption 1. Numerical tests show that ITTA and MLBA are very sensible to the optimal choice of α, while our proposed nonstationary versions do not require any accurate parameter estimation (e.g., $\alpha_0 = 0.5$ is a good choice for image deblurring). In particular, the nonstationary MLBA (NMLBA) is the best method among the four: it is robust both in terms of accuracy and speed with respect to the choice of α_0.

In this paper we have investigated only the synthesis approach while a nonstationary sequence of regularization parameters could be useful also for the analysis and the balanced approaches, cf [9]. Moreover, the combination of iterated Tikhonov with thresholding could be useful also for other kinds of convex constraints (e.g. total variation) [24]. Concerning the applications, we point out that our proposal does not depend on the particular problem of image deblurring considered in this paper, so it could be applied also to multiresolution [4, 15, 16, 17, 28], inpainting [2, 7, 19, 30], etc. Finally, all methods investigated in this paper require solution of a linear system with the coefficient matrix $KK^* + \alpha I$ which could be hard in some applications. In such case, a preconditioning strategy should be considered where K is replaced with an easily invertible approximation as in [27].

References

[1] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci., 2–1 (2009), pp. 183–202.

[2] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester, *Image inpainting*, SIGGRAPH, 34 (2000), pp. 417–424.

[3] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, Bristol, UK, 1998.

[4] N. Bose and K. Boo, *High-resolution image reconstruction with multisensors*, International Journal of Imaging Systems and Technology, 9 (1998), pp. 294–304.

[5] L. M. Bregman, *A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming*, Z . Vycisl. Mat. i Mat. Fiz., 7 (1967), pp. 620–631.

[6] M. Brill and E. Schock, *Iterative solution of ill-posed problems - a survey*, 1987.

[7] J. F. Cai, R. H. Chan, and Z. Shen, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal., 24 (2008), pp. 131–149.

[8] J. F. Cai, S. Osher, and Z. Shen, *Linearized Bregman iterations for frame-based image deblurring*, SIAM J. Imaging Sci., 2–1 (2009), pp. 226–252.

[9] J. F. Cai, S. Osher, and Z. Shen, *Split Bregman methods and frame based image restoration*, Multiscale Model. Simul., 8–2 (2009), pp. 337–369.

[10] J. F. Cai, S. Osher, and Z. Shen, *Convergence of the linearized Bregman iteration for ℓ_1-norm minimization*, Math. Comput., 78–268 (2009), pp. 2127–2136.

[11] J. F. Cai, S. Osher, and Z. Shen, *Linearized Bregman iterations for compressed sensing*, Math. Comput., 78–267 (2009), pp. 1515–1536.

[12] E. J. Candés and J. Romberg, *Practical signal recovery from random projections*, Wavelet Applications in Signal and Image Processing XI Proc. SPIE Conf. Vol. 5914 (2004).

[13] A. Chai and Z. Shen, *Deconvolution: A wavelet frame approach*, Numer. Math., 106 (2007), pp. 529–587.

[14] A. Chambolle, R. A. De Vore, 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.

[15] 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.

[16] 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.

[17] R. H. Chan, Z. Shen, and T. Xia, *A framelet algorithm for enhancing video stills*, Appl. Comput. Harmon. Anal., 23 (2007), pp. 153–170.

[18] T. Chan and J. H. Shen, *Image Processing and Analysis–Variational, PDE, wavelet, and stochastic methods*, SIAM, Philadelphia, 2005.

[19] T. Chan, J. H. Shen, and H. M. Zhou, *Total variation wavelet inpainting*, J. Math. Imaging Vision, 25 (2006), pp. 107–125.

[20] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., 4–4 (2005), pp. 1168–1200.

[21] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math., 57–11 (2004), pp. 1413–1457.

[22] I. Daubechies, M. Fornasier, and I. Loris, *Accelerated projected gradient method for linear inverse problems with sparsity constraints*, J. Fourier Anal. Appl. 14 (2008), pp. 764–792.

[23] 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.

[24] I. Daubechies, G. Teschke, and L. Vese, *Iteratively solving linear inverse problems under general convex constraints*, Inverse Problems Imaging, 1–1 (2007), pp. 29–46.

[25] M. Donatelli, *Fast transforms for high order boundary conditions in deconvolution problems*, BIT, 50–3 (2010), pp. 559–576.

[26] M. Donatelli, *On nondecreasing sequences of regularization parameters for nonstationary iterated Tikhonov*, Numer. Algor., 60–4 (2012), pp. 651–668.

[27] M. Donatelli and M. Hanke, *Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring*, submitted manuscript.

[28] M. Elad and A. Feuer, *Restoration of a single superresolution image from several blurred, noisy and undersampled measured images*, IEEE Trans. Image Process., 6 (1997), pp. 1646–1658.

[29] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht (1996).

[30] M. J. Fadili and J. L. Starck, *Sparse representations and Bayesian image inpainting*, Proc. SPARS'05, Vol. I, Rennes, France, 2005.

[31] A. G. Fakeev, *A class of iterative processes for solving degenerate systems of linear algebraic equations*, U.S.S.R. Comput. Math. Math. Phys., 21 (1981), pp. 15–22.

[32] M. Figueiredo and R. Nowak, *An EM algorithm for wavelet-based image restoration*, IEEE Trans. Image Process., 12–8 (2003), pp. 906–916.

[33] E. Hale, W. Yin, and Y. Zhang, *Fixed-point continuation for ℓ_1-minimization: methodology and convergence*, SIAM J. Optim., 19–3 (2008), pp. 1107–1130.

[34] M. Hanke and C. W. Groetsh, *Nonstationary iterated Tikhonov regularization*, J. Optim. Theory Appl., 98–1 (1998), pp. 37–53.

[35] M. Hanke and P. C. Hansen, *Regularization methods for large-scale problems*, Surveys Math. Indust., 3 (1993), pp. 253–315.

[36] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1997.

[37] J. T. King and D. Chillingworth, *Approximation of generalized inverses by iterated regularization*, Numer. Func. Anal. Opt., 1 (1979), pp. 499–513.

[38] A. V. Kryanev, *An iterative method for solving incorrectly posed problems*, U.S.S.R. Comput. Math. Math. Phys., 14 (1974), pp. 24–33.

[39] L. Landweber, *An iteration formula for Fredholm integral equations of the first kind*, Am. J. Math., 73 (1951), pp. 615–624.

[40] I. Loris, M. Bertero, C. De Mol, R. Zanella, and L. Zanni, *Accelerating gradient projection methods for ℓ_1-constrained signal recovery by steplength selection rules*, Appl. Comput. Harmon. Anal., 27–2 (2009), pp. 247–254.

[41] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd edition. Academic Press: San Diego (1999).

[42] V. A. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl., 7 (1966), pp. 414–417.

[43] F. Natterer, *The Mathematics of Computerized Tomography*, SIAM, Philadelphia, 2001.

[44] M. K. Ng, R. H. Chan, and W. C. Tang, *A fast algorithm for deblurring models with Neumann boundary conditions*, SIAM J. Sci. Comput., 21 (1999), pp. 851–866.

[45] S. Osher, Y. Mao, B. Dong, and W. Yin, *Fast linearized Bregman iteration for compressed sensing and sparse denoising*, Commun. Math. Sci., 8–1 (2010), pp. 93–111.

[46] M. Piana and M. Bertero, *Projected Landweber method and preconditioning*, Inverse Problems, 13 (1997), pp. 441–464.

[47] O. N. Strand, *Theory and methods related to the singular-function expansion and Landweber's iteration for integral equations of the first kind*, SIAM J. Numer. Anal, 11 (1974), pp. 798–825.

[48] A. N. Tikhonov, *Solution of incorrectly formulated problems and the regularization method*, Soviet Math. Dokl., 4 (1963), pp. 1035–1038.

[49] C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.

[50] 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 (2008), pp. 143–168.