Regularizing preconditioners by non-stationary iterated Tikhonov with general penalty term

Alessandro Buccini[a]

[a]*University of Insubria, Department of Science and High Technology, Como, Italy*

Abstract

The nonstationary preconditioned iteration proposed in a recent work by Donatelli and Hanke appeared on IP can be seen as an approximated iterated Tikhonov method. Starting from this observation we extend the previous iteration in two directions: the introduction of a regularization operator different from the identity (e.g., a differential operator) and the projection into a convex set (e.g., the nonnegative cone). Depending on the application both generalizations can lead to an improvement in the quality of the computed approximations. Convergence results and regularization properties of the proposed iterations are proved. Finally, the new methods are applied to image deblurring problems and compared with the iteration in the original work and other methods with similar properties recently proposed in the literature.

Keywords: Iterative regularization, Regularizing preconditioners, Image deblurring
2000 MSC: 65F22, 65R32, 15A29

1. Introduction

Many applications in physics and engineering lead to a linear problem of the form

$$T\mathbf{x} = \mathbf{y}, \tag{1}$$

where $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator between the Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$, see [1, 2, 3]. In this paper we consider numerical examples concerning the image deblurring problem, but other applications could be investigated as well.

The situation we want to consider is when T is severely ill-conditioned and the noise free right-hand side term $\mathbf{y}$ is not available, as we can only have a noise contaminated data $\mathbf{y}^\delta$, such that

$$\left\| \mathbf{y}^\delta - \mathbf{y} \right\| \le \delta, \tag{2}$$

with $\delta > 0$ assumed to be known and $\|\cdot\|$ denoting the ℓ_2–norm. Because of the ill conditioning of T in order to find a good approximation of $\mathbf{x}^\dagger := T^\dagger \mathbf{y}$, where $T^\dagger$ denotes the Moore-Penrose pseudo inverse, we have to use regularization methods. An example of these is the Tikhonov regularization method, that in its standard form is defined as

$$\min_{\mathbf{x}} \left\| T\mathbf{x} - \mathbf{y}^\delta \right\|^2 + \alpha \left\| \mathbf{x} \right\|^2, \tag{3}$$

where $\alpha > 0$ is the regularization parameter that balances the data fidelity and the regularization term, see [4].

In order to improve the quality of the computed approximation of $\mathbf{x}^\dagger$, a regularization matrix L can be introduced:

$$\min_{\mathbf{x}} \left\| T\mathbf{x} - \mathbf{y}^\delta \right\|^2 + \alpha \left\| L\mathbf{x} \right\|^2. \tag{4}$$

Denoting by $\mathcal{N}(L)$ the null space of L, the regularization matrix L has to be chosen such that important components of the solution to restore belong to $\mathcal{N}(L)$ and

$$\mathcal{N}(L) \cap \mathcal{N}(T) = \{0\}, \tag{5}$$

see [5] and references therein.

Email address: `alessandro.buccini@uninsubria.it` (Alessandro Buccini)

Let $\mathbf{x}_0$ be an approximation of the true solution $\mathbf{x}^\dagger$, it can be included in the Tikhonov method (3) computing

$$\min_{\mathbf{x}} \left\| T\mathbf{x} - \mathbf{y}^\delta \right\|^2 + \alpha \left\| \mathbf{x} - \mathbf{x_0} \right\|^2 \qquad \Longleftrightarrow \qquad \min_{\mathbf{h}} \left\| T\mathbf{h} - \mathbf{r}_0 \right\|^2 + \alpha \left\| \mathbf{h} \right\|^2,$$

where $\mathbf{r}_0 = \mathbf{y}^\delta - T\mathbf{x}_0$ and $\mathbf{h} = \mathbf{x} - \mathbf{x_0}$. Therefore $\mathbf{h}$ provides an approximation of the error $\mathbf{e}_0 = \mathbf{x}^\dagger - \mathbf{x}_0$ and an improved restoration is

$$\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{h}.$$

Applying iteratively the same refinement strategy, we obtain the Iterated Tikhonov (IT) method [1]. Given $\mathbf{x}_0 \in \mathcal{X}$, for $n = 0, 1, \ldots$

1. compute $\mathbf{r}_n = \mathbf{y}^\delta - T\mathbf{x_n}$,

2. solve $\min_{\mathbf{h}} \left\| T\mathbf{h} - \mathbf{r}_n \right\|^2 + \alpha_n \left\| \mathbf{h} \right\|^2$,

3. update $\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{h}_n$.

The previous IT algorithm can be formulated as the following iteration

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{x}_n + (T^*T + \alpha_n I)^{-1} T^* (\mathbf{y}^\delta - T\mathbf{x}_n), \\ &= \mathbf{x}_n + T^* (TT^* + \alpha_n I)^{-1} (\mathbf{y}^\delta - T\mathbf{x}_n), \qquad n = 0, 1, \ldots. \end{aligned}$$

As a stopping rule for this iterative process is often chosen the *discrepancy principle*, so that the method stops after $n = n_\delta \geq 0$ iterations with

$$\left\| \mathbf{r}_{n_\delta} \right\| \leq \tau\delta < \left\| \mathbf{r}_n \right\|, \qquad n = 0, 1, \ldots, n_\delta - 1.$$

where $\tau > 1$.

In this algorithm the choice of α has a crucial role and many strategies have been proposed. If the same α is used for each iteration we call the method *stationary*, whereas if α depends on n then we call the method *nonstationary*. In many applications the latter type has proved to obtain better results than the first type and it has been thoroughly investigated. For instance, a common choice is the geometric sequence $\alpha_n = \alpha_0 q^{n-1}$, $n = 1, 2, \ldots$ with $0 < q < 1$ and $\alpha_0 > 0$, studied in [6]. Another possible choice is that of nondecreasing sequences proposed in [7].

In [8] the authors developed an iterative method with a nonstationary preconditioner, that can be seen as an approximated iterated Tikhonov regularization. In particular they consider an operator C which is spectrally equivalent to T (see Assumption 1 in the next section) and form the preconditioner at step n as

$$C^* (CC^* + \alpha_n I)^{-1} \approx T^* (TT^* + \alpha_n I)^{-1},$$

where α_n is determined by a damped version of the well-known discrepancy principle. In this way they are able to both achieve fast computation, by wisely choosing the structure of C, and have a parameter free method, which is very important. The estimation of the parameter α_n can be difficult. For example in the above mentioned geometric sequence there are two parameters to be estimated: α_0 and q. Even though small changes in either α_0 and q have only a limited effect on the quality of the reconstruction, an imprudent choice can still lead to poor results. Because, roughly speaking, they are approximating the operator T with C we will refer to this method as *Approximated Iterated Tikhonov (AIT)*.

Another extension of the AIT method has been proposed in [9]. In the this work the authors consider the case of image deblurring and use the eigenvalue of the preconditioner generated by AIT as a generating function for a structured preconditioner inside an iterative refinement technique.

In this work we want to add some features to the algorithm in [8] and test the resulting methods on image deblurring. Since we know that the introduction of a regularization operator L in (3) can improve the quality of the obtained reconstructions, at first we study the introduction of L in place of I in AIT. We call the resulting method *AIT-GP (Approximated Iterated Tikhonov with General Penalty term)*. If we know that $\mathbf{x}^\dagger$ lies in some closed and convex $\Omega \subset \mathcal{X}$, we constrain the algorithm in order to get $\mathbf{x}_n \in \Omega$, $\forall n$. Hence, we modify AIT introducing the metric projection into Ω. We refer to this method as *APIT (Approximated Projected Iterated Tikhonov)*. For both the previous generalizations of the iterative method proposed in [8], namely AIT-GP and APIT, we prove that the new iterations are convergent in the noise free case and that are regularization methods

in the noisy case. Finally, we combine the regularization term and the projection into Ω developing a third algorithm called *APIT-GP (Approximated Projected Iterated Tikhonov with General Penalty term)*.

This work is structured as follows. In section 2 we summarized some results in [10] and we give an iterated Tikhonov algorithm which includes the regularization term. Section 3 describes the AIT method proposed in [8]. In Sections 4 and 5 we define and study the theoretical properties of our new three iterative regularization methods. Finally, in Section 6 the proposed methods are applied to the image deblurring problem and compared with other methods proposed in the literature.

2. Iterated Tikhonov in general form

In order to theoretically analyze (4), it can be useful to reduce the problem in the standard form (3).

If L is invertible, then the minimization problem (4) becomes

$$\min_{\overline{\mathbf{x}} = L\mathbf{x}} \left\| TL^{-1}\overline{\mathbf{x}} - \mathbf{y}^\delta \right\|^2 + \alpha \left\| \overline{\mathbf{x}} \right\|^2. \tag{6}$$

Solving (6) leads to $\overline{\mathbf{x}}_\alpha$ from which we can retrieve the solution $\mathbf{x}_\alpha$ of (4) by multiplying times L^{-1}:

$$\mathbf{x}_\alpha = L^{-1}\overline{\mathbf{x}}_\alpha.$$

When L is not invertible, we follow Eldén [10]. Let $T : \mathcal{X} \to \mathcal{Y}$ and $L : \mathcal{X} \to \mathcal{X}$ be two linear operator between Hilbert spaces, the $T-$weighted pseudoinverse of L is

$$L_T^\dagger = (I - (T(I - L^\dagger L))^\dagger T)L^\dagger. \tag{7}$$

We define the vectors

$$\begin{cases} \overline{\mathbf{x}} = L\mathbf{x} \\ \mathbf{x}^{(0)} = (T(I - L^\dagger L))^\dagger \mathbf{y}^\delta \\ \overline{\mathbf{y}}^\delta = \mathbf{y}^\delta - T\mathbf{x}^{(0)} \end{cases}$$

and consider the problem

$$\min_{\overline{\mathbf{x}}} \left\| TL_T^\dagger \overline{\mathbf{x}} - \mathbf{y}^\delta \right\|^2 + \alpha \left\| \overline{\mathbf{x}} \right\|^2. \tag{8}$$

The solution $\mathbf{x}_\alpha$ of (4) is obtained from the solution $\overline{\mathbf{x}}_\alpha$ of (8) by

$$\mathbf{x}_\alpha = L_T^\dagger \overline{\mathbf{x}}_\alpha + \mathbf{x}^{(0)}.$$

Since a given approximation $\mathbf{x}_0$ of $\mathbf{x}^\dagger$ can be included also in the generalized Tikhonov method (4) computing

$$\min_{\mathbf{x}} \left\| T\mathbf{x} - \mathbf{y}^\delta \right\|^2 + \alpha \left\| L(\mathbf{x} - \mathbf{x_0}) \right\|^2 \qquad \Longleftrightarrow \qquad \min_{\mathbf{h}} \left\| T\mathbf{h} - \mathbf{r}_0 \right\|^2 + \alpha \left\| L\mathbf{h} \right\|^2,$$

the iterated Tikhonov method can be defined also for a general regularization operator L. However, to the best of our knowledge, a complete theory about this method is currently not available in the literature. After completing this work we discovered that a generalization similar to the one here proposed has been introduced in [11] in relation with the generalized Krilov spaces.

Applying iteratively the same refinement strategy, given $\mathbf{x}_0 \in \mathcal{X}$, for $n = 0, 1, \ldots$

1. compute $\mathbf{r}_n = \mathbf{y}^\delta - T\mathbf{x_n}$,

2. solve $\min_{\mathbf{h}} \left\| T\mathbf{h} - \mathbf{r}_n \right\|^2 + \alpha_n \left\| L\mathbf{h} \right\|^2$,

3. update $\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{h}_n$.

Under the previous condition (5), solving the point 2. is equivalent to

$$\mathbf{x}_\alpha = (T^*T + \alpha_n L^*L)^{-1}T^*\mathbf{y}^\delta,$$

which can be rewritten as

$$\mathbf{x}_\alpha = T^*(TT^* + \alpha_n LL^*)^{-1}\mathbf{y}^\delta,$$

if L commutes with T. Hence the previous iterated generalized Tikhonov algorithm can be formulated as the iteration

$$\mathbf{x}_{n+1} = \mathbf{x}_n + (T^*T + \alpha_n L^*L)^{-1}T^*(\mathbf{y}^\delta - T\mathbf{x_n}), \qquad n = 0, 1, \ldots. \tag{9}$$

3. Approximated Iterated Tikhonov

We now describe the preconditioned iteration proposed in [8]. We need the following assumption that it will be necessary also for our algorithms in the next sections.

Assumption 1. *Let C be a linear operator such that*

$$\|(C-T)\mathbf{z}\| \le \rho \|T\mathbf{z}\|, \qquad \forall \mathbf{z} \in \mathcal{X}, \tag{10}$$

for some $0 < \rho < \frac{1}{2}$. We say that C is spectrally equivalent to T.

Under this assumption it holds a preliminary result useful for the convergence analysis.
First we define the residual at the n-th step

$$\mathbf{r}_n = \mathbf{y}^\delta - T\mathbf{x}_n.$$

Lemma 2 ([8]). *Assume that (2) and Assumption 1 hold, if $\tau_n = \|\mathbf{r}_n\|/\delta > \tau_* = (1+\rho)/(1-2\rho)$, then it follows that*

$$\|\mathbf{r}_n - C\mathbf{e}_n\| \le \left(\rho + \frac{1+\rho}{\tau_n}\right)\|\mathbf{r}_n\| < (1-\rho)\|\mathbf{r}_n\|.$$

Algorithm 1 (AIT). *Let $\mathbf{x}_0 \in \mathcal{X}$ be fixed and set $n = 0$. Choose $\tau = \frac{1+2\rho}{1-2\rho}$ with ρ as in (10), and fix $q \in [2\rho, 1]$.*

While $\|\mathbf{r}_n\| > \tau\delta$, let $\tau_n = \|\mathbf{r}_n\|/\delta$ and $q_n = \max\left\{q, 2\rho + \frac{1+\rho}{\tau_n}\right\}$, compute

$$\mathbf{h}_n = C^*(CC^* + \alpha_n I)^{-1}\mathbf{r}_n,$$

where α_n is such that

$$\|\mathbf{r}_n - C\mathbf{h}_n\| = q_n\|\mathbf{r}_n\|,$$

and update

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{h}_n,$$
$$\mathbf{r}_{n+1} = \mathbf{y}^\delta - T\mathbf{x}_{n+1}.$$

We summarize the main theoretical results proved in [8] about the convergence and the monotonic decrease of the norm of the error for AIT. We denote the iteration error $\mathbf{e}_n$ with

$$\mathbf{e}_n = \mathbf{x}^\dagger - \mathbf{x}_n.$$

Proposition 3 ([8]). *Under Assumption 1, while $\|\mathbf{r}_n\| > \tau\delta$, with $\tau = (1+2\rho)/(1-2\rho)$ the norm of the reconstruction error $\mathbf{e}_n$ decreases monotonically, namely $\|\mathbf{e}_{n+1}\| \le \|\mathbf{e}_n\|$, for $n = 0, 1, \ldots$.*

Theorem 4 ([8]). *Assume that the data are exact, i.e., $\delta = 0$, and that $\mathbf{x}_0$ is not a solution of problem (1). Then the sequence $(\mathbf{x}_n)_n$ converges as $n \to \infty$ to the solution of (1) which is nearest to $\mathbf{x}_0$.*

Theorem 5 ([8]). *Let $\delta \mapsto \mathbf{y}^\delta$ be a function from $\mathbb{R}^+$ to $\mathcal{Y}$ such that (2) holds true for all $\delta > 0$. Under Assumption 1, for two fixed parameters τ and q, denote with $\mathbf{x}^\delta$ the resulting approximation obtained with AIT. Then for $\delta \to 0$ we have that $\mathbf{x}^\delta \to \mathbf{x}_0^\dagger$ which is the nearest solution of (1) to $\mathbf{x}_0$.*

In [8] the choice of $\mathbf{x}_0$ was not deeply investigated. For many iterative regularization methods, like Krylov methods, the null vector is usually a good choice for $\mathbf{x}_0$. Nevertheless, for AIT this is not a good choice and setting $\mathbf{x}_0 = T^*\mathbf{y}^\delta$, which is the initial solution subspace vector for LSQR, usually provides better results. This is confirmed by several numerical experiments with image deblurring problems and follows from the next observation. The approximation of T by C is motivated by the fact that the error equation, used for the iterative refinement, allows a slight misfit due to the noise already present in the problem. If we choose $\mathbf{x}_0 = \mathbf{0}$ then $\mathbf{r}_0 = \mathbf{y}^\delta$ and $\mathbf{x}_1 = \mathbf{x}_0 + C^*(CC^* + \alpha_n I)^{-1}\mathbf{y}^\delta$ which is exactly the Tikhonov solution for the operator C instead of T. Although, from a theoretical point of view, this should not be a problem, numerically this can lead to some issues. For example if C do not approximate well T then $\mathbf{x}_1$ could contain large error components. This components may be hard to reduce in the following iterations to the point of producing a slightly worse solution than the one obtained with $\mathbf{x}_0 = T^*\mathbf{y}^\delta$.

4. Approximated Iterated Tikhonov with general penalty term (AIT-GP)

In this section we combine the idea of AIT with the generalized iterated Tikhonov method (9), i.e., we introduce the regularization operator L in Algorithm 1.

We need a couple of assumptions that link the matrix L with T and C similarly to the basic assumption (5).

Assumption 6. *Let L and C be two linear operators such that*

(i) $C|_{\mathcal{N}(L)} = T|_{\mathcal{N}(L)}$;

(ii) L and C are diagonalized by the same unitary transformation.

Assumption 6(ii) is restrictive, but it is needed for the proofs that follows. This kind of requirements can be satisfied for particular choices of C and L and in particular for certain classes of structured matrices. In Section 6 we show an example.

Note that thanks to (10) $\mathcal{N}(T) = \mathcal{N}(C)$ and hence (5) implies that $\mathcal{N}(L) \cap \mathcal{N}(C) = \{0\}$.

Remark 7. *Under the assumption* (ii) *on C and L we have that*

$$C : \mathcal{X} \to \mathcal{X},$$
$$L : \mathcal{X} \to \mathcal{X}.$$

It is indeed possible to choose an $L : \mathcal{X} \to \mathcal{Z}$ and then transform it into an operator to $\mathcal{X}$ either via an appropriate zero padding *or using its QR factorization. However, it can be challenging proving Assumption 6*(ii) *after the transformation.*

We define the orthogonal projection over $\mathcal{N}(L)$

$$P_{\mathcal{N}(L)} = I - L^\dagger L$$

and the orthogonal projection over $\mathcal{N}(L)^\perp$

$$P_{\mathcal{N}(L)^\perp} = L^\dagger L.$$

From Remark 7 and Assumption 6(ii) we have the following

Lemma 8. *Let L and C be operator from $\mathcal{X}$ to $\mathcal{X}$ that commute and let $L_C^\dagger$ be the operator defined in (7), then it holds*

(i) $C^\dagger C$ commutes with $L^\dagger L$;

(ii) $(I - L^\dagger L)C = C(I - L^\dagger L)$;

(iii) $(C(I - L^\dagger L))^\dagger = (I - L^\dagger L)C^\dagger$.

Proof. From Assumption 6(ii) there exists Q orthogonal such that

$$C = Q^{-1}\Gamma Q,$$
$$L = Q^{-1}\Lambda Q,$$

with $Q^{-1} = Q^*$.

Consider now $C^\dagger C$ and $L^\dagger L$, these are the projection onto the orthogonal of the null spaces of C and L respectively, it is trivial to see then that

$$C^\dagger C = Q^{-1}\Gamma^\dagger \Gamma Q,$$
$$L^\dagger L = Q^{-1}\Lambda^\dagger \Lambda Q. \tag{11}$$

This is because this operator just need to switch of the components lying into the null spaces of C or L and leave unchanged the others.

From (11) the proof ot point (i) comes immediately. In fact

$$C^\dagger C L^\dagger L = Q^{-1}\Gamma^\dagger \Gamma Q Q^{-1}\Lambda^\dagger \Lambda Q$$
$$= Q^{-1}\Gamma^\dagger \Gamma \Lambda^\dagger \Lambda Q = Q^{-1}\Lambda^\dagger \Lambda \Gamma^\dagger \Gamma Q$$
$$= Q^{-1}\Lambda^\dagger \Lambda Q Q^{-1}\Gamma^\dagger \Gamma Q = L^\dagger L C^\dagger C,$$

where we have used the fact that diagonal operators commute with each other.

We move now to point (ii). From (11) we have that

$$(I - L^\dagger L) = (Q^{-1}Q - Q^{-1}\Lambda^\dagger \Lambda Q) = Q^{-1}(I - \Lambda^\dagger \Lambda)Q.$$

We can then write

$$(I - L^\dagger L)C = Q^{-1}(I - \Lambda^\dagger \Lambda)QQ^{-1}\Gamma Q$$
$$= Q^{-1}(I - \Lambda^\dagger \Lambda)\Gamma Q = Q^{-1}\Gamma Q Q^{-1}(I - \Lambda^\dagger \Lambda)Q$$
$$= C(I - L^\dagger L),$$

proving point (ii).

Finally we can prove point (iii). In order to do that we show the four properties that characterize the Moore-Penrose pseudo-inverse. In particular X is the Moore-Penrose pseudo-inverse of A if and only if

$$AXA = A, \quad XAX = X, \quad (AX)^* = AX, \quad (XA)^* = XA.$$

We now prove the above four properties with $A = C(I - L^\dagger L)$ and $X = (I - L^\dagger L)C^\dagger$.

$AXA = A$. Using the fact that $(I - L^\dagger L)(I - L^\dagger L) = (I - L^\dagger L)$, point (ii) and the properties of the pseudo-inverse $C^\dagger$ we have

$$\left[C(I - L^\dagger L)\right]\left[(I - L^\dagger L)C^\dagger\right]\left[C(I - L^\dagger L)\right] = C(I - L^\dagger L)C^\dagger C(I - L^\dagger L) = (I - L^\dagger L)CC^\dagger C(I - L^\dagger L)$$
$$= (I - L^\dagger L)C(I - L^\dagger L) = C(I - L^\dagger L)(I - L^\dagger L) = C(I - L^\dagger L).$$

$XAX = X$. Considering also point (i) we have

$$\left[(I - L^\dagger L)C^\dagger\right]\left[C(I - L^\dagger L)\right]\left[(I - L^\dagger L)C^\dagger\right] = (I - L^\dagger L)C^\dagger C(I - L^\dagger L)C^\dagger$$
$$= (I - L^\dagger L)(I - L^\dagger L)C^\dagger CC^\dagger = (I - L^\dagger L)C^\dagger$$

$(AX)^* = AX$. Noting that $(C^\dagger C)^* = C^\dagger C$ and $(L^\dagger L)^* = L^\dagger L$ by the properties of the pseudo-inverse we get

$$\left(\left[C(I - L^\dagger L)\right]\left[(I - L^\dagger L)C^\dagger\right]\right)^* = \left(C(I - L^\dagger L)C^\dagger\right)^* = \left((I - L^\dagger L)CC^\dagger\right)^*$$
$$= CC^\dagger(I - L^\dagger L) = (I - L^\dagger L)CC^\dagger = C(I - L^\dagger L)C^\dagger$$
$$= \left[C(I - L^\dagger L)\right]\left[(I - L^\dagger L)C^\dagger\right].$$

$(XA)^* = XA$. Analogously we get

$$\left(\left[(I - L^\dagger L)C^\dagger\right]\left[C(I - L^\dagger L)\right]\right)^* = \left((I - L^\dagger L)(I - L^\dagger L)C^\dagger C\right)^*$$
$$= \left((I - L^\dagger L)C^\dagger C\right)^* = C^\dagger C(I - L^\dagger L)$$
$$= C^\dagger C(I - L^\dagger L)(I - L^\dagger L) = \left[(I - L^\dagger L)C^\dagger\right]\left[C(I - L^\dagger L)\right].$$

Which conclude the proof of point (iii). $\square$

We are now in the position of proving

Lemma 9. *With the same assumptions and notations of Lemma 8 it holds*

$$L_C^\dagger = L^\dagger$$

Proof. Let us write the expression for $L_C^\dagger$ and use the results shown in Lemma 8

$$
\begin{aligned}
L_C^\dagger &= (I - (C(I - L^\dagger L))^\dagger C)L^\dagger \\
&= (I - (I - L^\dagger L)C^\dagger C)L^\dagger \\
&= (I - C^\dagger C + L^\dagger L C^\dagger C)L^\dagger \\
&= L^\dagger - C^\dagger C L^\dagger + L^\dagger L C^\dagger C L^\dagger \\
&= L^\dagger - C^\dagger C L^\dagger + C^\dagger C L^\dagger L L^\dagger \\
&= L^\dagger - C^\dagger C L^\dagger + C^\dagger C L^\dagger = L^\dagger
\end{aligned}
$$

$\square$

We define

$$
\overline{C} = C L_C^\dagger = C L^\dagger. \tag{12}
$$

Algorithm 2 (AIT-GP). *Let L and C be linear operators that fulfills Assumptions 1 and 6 for a fixed $0 < \rho \leq \frac{1}{2}$. Let $\mathbf{x}_0 \in \mathcal{X}$ be fixed and set $n = 0$. Choose $\tau = \frac{1+2\rho}{1-2\rho}$ with ρ from (10), and fix $q \in [2\rho, 1]$.*

While $\|\mathbf{r}_n\| > \tau\delta$, let $\tau_n = \|\mathbf{r}_n\|/\delta$ and $q_n = \max\left\{q, 2\rho + \frac{1+\rho}{\tau_n}\right\}$, compute

$$
\mathbf{h}_n = C^*(CC^* + \alpha_n L L^*)^{-1}\mathbf{r}_n,
$$

where α_n is such that

$$
\|\mathbf{r}_n - C\mathbf{h}_n\| = q_n \|\mathbf{r}_n\|, \tag{13}
$$

and update

$$
\begin{aligned}
\mathbf{x}_{n+1} &= \mathbf{x}_n + \mathbf{h}_n, \\
\mathbf{r}_{n+1} &= \mathbf{y}^\delta - T\mathbf{x}_{n+1}.
\end{aligned}
$$

We refer to Algorithm 2 as *Approximated Iterated Tikhonov with General Penalty term* since this method can be seen as a preconditioned iterative method whose preconditioner is obtained by approximated Tikhonov with a general regularization operator L.

We define

$$
\begin{cases}
\mathbf{h}_n^{(0)} = (C(I - L^\dagger L))^\dagger \mathbf{r}_n \\
\overline{\mathbf{r}}_n = \mathbf{r}_n - C\mathbf{h}_n^{(0)}
\end{cases}. \tag{14}
$$

Note that if L is invertible then $\overline{\mathbf{r}}_n = \mathbf{r}_n$.

Lemma 10. *Let $\overline{\mathbf{r}}_n$ be defined in (14), then it holds*

$$
\left\| L^\dagger \mathbf{r}_n \right\| = \left\| L^\dagger \overline{\mathbf{r}}_n \right\|.
$$

Proof. From the definition of $\overline{\mathbf{r}}_n$ and $\mathbf{h}_n^{(0)}$ in (14) it follows that

$$
\left\| L^\dagger \overline{\mathbf{r}}_n \right\| = \left\| L^\dagger (\mathbf{r}_n - C\mathbf{h}_n^{(0)}) \right\| = \left\| L^\dagger \mathbf{r}_n - L^\dagger C(C(I - L^\dagger L))^\dagger \mathbf{r}_n \right\|
$$

proving that $L^\dagger C(C(I - L^\dagger L))^\dagger = 0$ will conclude the proof. Consider the results in Lemma 8

$$
L^\dagger C(C(I - L^\dagger L))^\dagger = L^\dagger C(I - L^\dagger L)C^\dagger = L^\dagger(I - L^\dagger L)CC^\dagger.
$$

Note that, since $L : \mathcal{X} \to \mathcal{X}$ it holds that $\mathcal{N}(L^\dagger) = \mathcal{N}(L^\dagger)$ thus, being $(I - L^\dagger L) = P_{\mathcal{N}(L)}$, we have that $L^\dagger(I - L^\dagger L) = 0$. Using this last equality we have that

$$
L^\dagger C(C(I - L^\dagger L))^\dagger = L^\dagger(I - L^\dagger L)CC^\dagger = 0,
$$

which concludes the proof.

In particular we also obtained that $\mathbf{h}_n^{(0)} \in \mathcal{N}(L)$.

$\square$

Lemma 11. *Let $\overline{\mathbf{r}}_n$ and $\overline{C}$ be defined in (14) and (12), respectively, and define*

$$\overline{\mathbf{h}}_n := L\mathbf{h}_n.$$

Then it holds

$$\left\|\overline{\mathbf{r}}_n - \overline{C}\overline{\mathbf{h}}_n\right\| = \|\mathbf{r}_n - C\mathbf{h}_n\|.$$

Proof. This results has been shown in [10], we give here a proof with our notation for completeness.

Form [10] we know that

$$\mathbf{h}_n = L^\dagger \overline{C}^*(\overline{C}\overline{C}^* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n + \mathbf{h}_n^{(0)},$$

and so, since $\overline{C}^*(\overline{C}\overline{C}^* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n \in \mathcal{N}(L)^\perp$ for construction of $\overline{\mathbf{r}}_n$ and $\overline{C}$ and $\mathbf{h}_n^{(0)} \in \mathcal{N}(L)$ (see Lemma 10), we get

$$\overline{\mathbf{h}}_n = L\mathbf{h}_n = LL^\dagger\overline{C}^*(\overline{C}\overline{C}^* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n + L\mathbf{h}_n^{(0)} = \overline{C}^*(\overline{C}\overline{C}^* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n,$$

and so it holds

$$\overline{\mathbf{h}}_n = \overline{C}^*(\overline{C}\overline{C}^* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n. \tag{15}$$

Moreover

$$\left\|\overline{\mathbf{r}}_n - \overline{C}\overline{\mathbf{h}}_n\right\| = \|\mathbf{r}_n - C\mathbf{h}_n\|, \tag{16}$$

in fact

$$\left\|\overline{\mathbf{r}}_n - \overline{C}\overline{\mathbf{h}}_n\right\| = \left\|\mathbf{r}_n - C\mathbf{h}_n^{(0)} - \overline{C}\overline{\mathbf{h}}_n\right\| = \left\|\mathbf{r}_n - C\left(L^\dagger\overline{\mathbf{h}}_n + \mathbf{h}_n^{(0)}\right)\right\| = \|\mathbf{r}_n - C\mathbf{h}_n\|,$$

where, in the last step, we have used the definition of $\overline{C} = CL^\dagger$. $\qquad\square$

Now we divide the space $\mathcal{X} = \mathcal{N}(L) \oplus \mathcal{N}(L)^\perp$ and we analyze the behavior of Algorithm 2 on each subspace. We call

$$\mathbf{e}_n^\perp = P_{\mathcal{N}(L)^\perp}(\mathbf{e}_n) = P_{\mathcal{N}(L)^\perp}(\mathbf{x}^\dagger) - P_{\mathcal{N}(L)^\perp}(\mathbf{x}_n)$$

and

$$\mathbf{e}_n^{(0)} = P_{\mathcal{N}(L)}(\mathbf{e}_n) = P_{\mathcal{N}(L)}(\mathbf{x}^\dagger) - P_{\mathcal{N}(L)}(\mathbf{x}_n).$$

On the two subspaces the Algorithm 2 has different behaviors. First, in Remark 12 we concentrate on the space $\mathcal{N}(L)$.

Remark 12. *Let us consider the projection onto $\mathcal{N}(L)$ of the very first iteration*

$$P_{\mathcal{N}(L)}(\mathbf{x}_1) = P_{\mathcal{N}(L)}(\mathbf{x}_0 + \mathbf{h}_0) = P_{\mathcal{N}(L)}(\mathbf{x}_0) + P_{\mathcal{N}(L)}(\mathbf{h}_0),$$

since $\mathbf{h}_n = L^\dagger\overline{\mathbf{h}}_n + \mathbf{h}_n^{(0)}$, we get

$$P_{\mathcal{N}(L)}(\mathbf{h}_0) = P_{\mathcal{N}(L)}(\mathbf{h}_0^{(0)}) = \mathbf{h}_0^{(0)} = \left(C\big|_{\mathcal{N}(L)}\right)^\dagger(\mathbf{y}^\delta - T\mathbf{x}_0).$$

In force of Assumption 6 we have

$$P_{\mathcal{N}(L)}(\mathbf{h}_0) = \left(T\big|_{\mathcal{N}(L)}\right)^\dagger(\mathbf{y}^\delta - T\mathbf{x}_0) = P_{\mathcal{N}}(T^\dagger\mathbf{y}^\delta) - P_{\mathcal{N}(L)}(\mathbf{x}_0).$$

And thus

$$P_{\mathcal{N}(L)}(\mathbf{x}_1) = P_{\mathcal{N}(L)}(\mathbf{x}_0) + P_{\mathcal{N}}(T^\dagger\mathbf{y}^\delta) - P_{\mathcal{N}(L)}(\mathbf{x}_0) = P_{\mathcal{N}}(T^\dagger\mathbf{y}^\delta),$$

so in the null space of L we directly invert the operator T at the very first step.

Proposition 13. *Let $\overline{\mathbf{e}}_n = L\mathbf{e}_n^\perp$, under Assumption 6 the norm of $\overline{\mathbf{e}}_n$ of Algorithm 2 decreases monotonically.*

$$\|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_{n+1}\|^2 \geq 2\rho\left\|(\overline{C}\overline{C}^* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n\right\|\|\mathbf{r}_n\|$$

Proof. This proof is in the spirit of the original result showed in [8, Proposition 2]. Let us consider $\|\overline{\mathbf{e}}_n\| = \|L\mathbf{e}_{n+1}^{\perp}\| = \|L\mathbf{e}_{n+1}\|$ so we have that

$$\|\overline{\mathbf{e}}_{n+1}\|^2 = \langle L\mathbf{e}_{n+1}, L\mathbf{e}_{n+1}\rangle = \langle L\mathbf{e}_n - L\mathbf{h}_n, L\mathbf{e}_n - L\mathbf{h}_n\rangle$$
$$= \|L\mathbf{e}_n\|^2 - 2\langle L\mathbf{e}_n, L\mathbf{h}_n\rangle + \|L\mathbf{h}_n\|^2.$$

Using the definition of $\mathbf{h}_n$ we have that, denoting with $Q_n = (\overline{CC}^* + \alpha_n I)$, it holds

$$\|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_{n+1}\|^2 = 2\langle L\mathbf{e}_n, L\mathbf{h}_n\rangle - \|L\mathbf{h}_n\|^2 \geq 2\langle L\mathbf{e}_n, L\mathbf{h}_n\rangle - 2\|L\mathbf{h}_n\|^2$$
$$\geq 2\left\langle L\mathbf{e}_n, \overline{C}^* Q_n^{-1}\overline{\mathbf{r}}_n\right\rangle - 2\left\langle \overline{\mathbf{r}}_n, \overline{CC}^* Q_n^{-2}\overline{\mathbf{r}}_n\right\rangle.$$

since $L\mathbf{h}_n = \overline{\mathbf{h}}_n = \overline{C}^* Q_n^{-1}\overline{\mathbf{r}}_n$ thanks to (15). Therefore

$$\|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_{n+1}\|^2 = 2\langle \overline{\mathbf{r}}_n, Q_n^{-1}\overline{\mathbf{r}}_n\rangle - 2\left\langle \overline{\mathbf{r}}_n, \overline{CC}^* Q_n^{-2}\overline{\mathbf{r}}_n\right\rangle$$
$$- 2\left\langle \overline{\mathbf{r}}_n - \overline{C}L\mathbf{e}_n, Q_n^{-1}\overline{\mathbf{r}}_n\right\rangle$$
$$= 2\langle \overline{\mathbf{r}_n}, Q_n^{-1}\overline{\mathbf{r}}_n\rangle - 2\left\langle \overline{\mathbf{r}}_n, \overline{CC}^* Q_n^{-2}\overline{\mathbf{r}}_n\right\rangle$$
$$- 2\left\langle \mathbf{r}_n - C\mathbf{e}_n^{\perp}, Q_n^{-1}\overline{\mathbf{r}}_n\right\rangle$$
$$= 2\left\langle \overline{\mathbf{r}}_n, \left[Q_n^{-1} - \overline{C}C^* Q_n^{-2}\right]\overline{\mathbf{r}}_n\right\rangle - 2\left\langle \overline{\mathbf{r}}_n - C\mathbf{e}_n^{\perp}, Q_n^{-1}\overline{\mathbf{r}}_n\right\rangle$$
$$\geq 2\alpha_n\langle \overline{\mathbf{r}}_n, Q_n^{-2}\overline{\mathbf{r}}_n\rangle - 2\left\|\overline{\mathbf{r}}_n - C\mathbf{e}_n^{\perp}\right\|\left\|Q_n^{-1}\overline{\mathbf{r}}_n\right\|$$
$$= 2\alpha_n\left\|Q_n^{-1}\overline{\mathbf{r}}_n\right\|^2 - 2\left\|\overline{\mathbf{r}}_n - C\mathbf{e}_n^{\perp}\right\|\left\|Q_n^{-1}\overline{\mathbf{r}}_n\right\|$$
$$= 2\left\|Q_n^{-1}\overline{\mathbf{r}}_n\right\|\left[\left\|\alpha_n Q_n^{-1}\overline{\mathbf{r}}_n\right\| - \left\|\overline{\mathbf{r}}_n - C\mathbf{e}_n^{\perp}\right\|\right]$$
$$\geq 2\left\|Q_n^{-1}\overline{\mathbf{r}}_n\right\|\left[\|\mathbf{r}_n - C\mathbf{h}_n\| - \|\mathbf{r}_n - C\mathbf{e}_n\|\right],$$

where the last step is obtained by considering (16)

$$\|\mathbf{r}_n - C\mathbf{h}_n\| = \left\|\overline{\mathbf{r}}_n - \overline{C\mathbf{h}}_n\right\| = \left\|\overline{\mathbf{r}}_n - \overline{CC}^*(\overline{CC}* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n\right\|$$
$$= \left\|\left[I - \overline{CC}^*(\overline{CC}* + \alpha_n I)^{-1}\right]\overline{\mathbf{r}}_n\right\| = \left\|\alpha_n Q_n^{-1}\overline{\mathbf{r}}_n\right\|,$$

and by

$$\left\|\overline{\mathbf{r}}_n - C\mathbf{e}_n^{\perp}\right\| = \left\|P_{\mathcal{N}(L)^{\perp}}(\mathbf{r}_n - C\mathbf{e}_n)\right\| \leq \|\mathbf{r}_n - C\mathbf{e}_n\|,$$

since $\left\|P_{\mathcal{N}(L)^{\perp}}\right\| = \|L^{\dagger}L\| = 1$.

In virtue of Proposition 2 and using equation (13) we have that

$$\|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_{n+1}\|^2 \geq 2\left\|Q_n^{-1}\overline{\mathbf{r}}_n\right\|\left[q_n\|\mathbf{r}_n\| - \|\mathbf{r}_n - C\mathbf{e}_n\|\right]$$
$$\geq 2\rho\left\|Q_n^{-1}\overline{\mathbf{r}}_n\right\|\|\mathbf{r}_n\| = 2\rho\left\|(\overline{CC}^* + \alpha_n I)^{-1}\overline{\mathbf{r}}_n\right\|\|\mathbf{r}_n\|.$$

$\square$

We call n^{δ} the iteration at which Algorithm 2 stops. From Corollary 14 we are going to be able to deduce that n^{δ} is finite if and only if $\delta > 0$.

Repeating the same steps that in [8] led to derive Corollary 3 from Proposition 2, the following result can be derived from Proposition 13.

Corollary 14. *With the notation and assumptions of Proposition 13, it holds*

$$\|\overline{\mathbf{e}}_0\|^2 \geq 2\rho \sum_{n=0}^{n^{\delta}-1}\left\|(\overline{CC}^* + \alpha_n I)^{-1}\mathbf{r}_n\right\|\|\mathbf{r}_n\| \geq c \sum_{n=0}^{n^{\delta}-1}\|\mathbf{r}_n\|^2.$$

Form the outer inequality in Corollary 14 we obtain that the sum of the squares of the norm of the residual (in $\mathcal{N}(L)^{\perp}$) is bounded and hence, if $\delta > 0$, there must be a first integer $\mathbb{N} \ni n^{\delta} < \infty$ that fulfills the stopping criterion. In fact suppose that the algorithm does not stop after finitely many iterations, we get that $\lim_{n\to\infty} \|\mathbf{r}_n\|^2 = 0$. Thus there exists $\overline{n}$ such that $\|\mathbf{r}_{\overline{n}}\| < \tau\delta$ which is absurd. In other words, if $\delta > 0$ Algorithm 2 terminates after a finite number of iterations. Conversely in Theorem 15 we show that, if $\delta = 0$ then the algorithms, even though it converges to a solution of the system, does not stop.

Theorem 15. *Assume that the data are exact, i.e., $\delta = 0$, and that $\mathbf{x}_0$ is not a solution of the problem. Then, the sequence $(\mathbf{x}_n)_n$ converges as $n \to \infty$ to a solution $\mathbf{x}_0^{\dagger}$ such that:*

(i) $T\mathbf{x}_0^{\dagger} = \mathbf{y}$;

(ii) $P_{\mathcal{N}(L)}(\mathbf{x}_0^{\dagger}) = P_{\mathcal{N}(L)}(\mathbf{x}^{\dagger})$;

(iii) the distance between $\mathbf{x}_0^{\dagger}$ and $\mathbf{x}_0$ is minimal with respect to the set of all the solutions.

Proof. The proof follows the same strategy of the analogous result in [8, Theorem 4]. Let us call $\mathbf{x}_{n\delta} = \mathbf{x}_0^{\dagger}$, since $\delta = 0$ the stopping criterion can only be fulfilled for $n = n^{\delta}$ and with $\|\mathbf{r}_n\| = 0$.

We now show that an infinite number of iterations is needed. If $n > 0$, then $\mathbf{h}_{n-1}$ must coincide with $\mathbf{e}_{n-1}$ up to an element in the null space of T, that is (thanks to Assumption 6) the null space of C, and so, using (13) and Proposition 2, we get

$$q_{n-1} \|\mathbf{r}_{n-1}\| = \|\mathbf{r}_{n-1} - C\mathbf{h}_{n-1}\| = \|\mathbf{r}_{n-1} - C\mathbf{e}_{n-1}\| \leq \left(\rho + \frac{1+\rho}{\tau_{n-1}}\right) \|\mathbf{r}_{n-1}\|.$$

This contradicts the definition of q_{n-1} and so the iteration does not terminate after finitely manly iterations for exact data if $\mathbf{x}_0$ is not a solution of the system.

Using Remark 12 the proof of point (ii) is immediate. It is left for us to show points (iii) and (i). In order to do that we first show that the sequence $(L\mathbf{x}_n)_n = (\overline{\mathbf{x}}_n)_n$ is a Cauchy sequence.

Let $m > l$ and let us consider $\|L\mathbf{x}_m - L\mathbf{x}_l\|^2$

$$\|L\mathbf{x}_m - L\mathbf{x}_l\|^2 = \|L\mathbf{e}_m - L\mathbf{e}_l\|^2 = \|\overline{\mathbf{e}}_m\|^2 - \|\overline{\mathbf{e}}_l\|^2 - 2\langle \overline{\mathbf{e}}_l, \overline{\mathbf{e}}_m - \overline{\mathbf{e}}_l\rangle$$
$$= \|\overline{\mathbf{e}}_m\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2\langle \overline{\mathbf{e}}_l, \overline{\mathbf{x}}_m - \overline{\mathbf{x}}_l\rangle.$$

Inserting the definition of $\mathbf{x}_n$ and of $\mathbf{h}_n$ we get

$$\|\overline{\mathbf{x}}_m - \overline{\mathbf{x}}_l\|^2 = \|\overline{\mathbf{e}}_m\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2\sum_{k=l}^{m-1} \langle \overline{\mathbf{e}}_l, \overline{\mathbf{h}}_k\rangle$$
$$= \|\overline{\mathbf{e}}_m\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2\sum_{k=l}^{m-1} \left\langle L\mathbf{e}_l, \overline{C}^*(\overline{CC}^* + \alpha_k)^{-1}\overline{\mathbf{r}}_k\right\rangle$$
$$= \|\overline{\mathbf{e}}_m\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2\sum_{k=l}^{m-1} \left\langle C\mathbf{e}_l^{\perp}, (\overline{CC}^* + \alpha_k)^{-1}\overline{\mathbf{r}}_k\right\rangle$$
$$\leq \|\overline{\mathbf{e}}_m\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2\sum_{k=l}^{m-1} \left\|C\mathbf{e}_l^{\perp}\right\| \left\|(\overline{CC}^* + \alpha_k)^{-1}\overline{\mathbf{r}}_k\right\|$$
$$\leq \|\overline{\mathbf{e}}_m\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2\sum_{k=l}^{m-1} \|C\mathbf{e}_l\| \left\|(\overline{CC}^* + \alpha_k)^{-1}\overline{\mathbf{r}}_k\right\|,$$

where in the last step we have used the fact that $\left\|C\mathbf{e}_n^{\perp}\right\| \leq \|C\mathbf{e}_n\|$.

Let us suppose now that $l \geq n$ and so

$$
\begin{aligned}
\|\overline{\mathbf{x}}_l - \overline{\mathbf{x}}_n\|^2 &= \|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2 \sum_{k=n}^{l-1} \left\langle \overline{\mathbf{e}}_l, \overline{\mathbf{h}}_k \right\rangle \\
&= \|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2 \sum_{k=n}^{l-1} \left\langle C\mathbf{e}_l^\perp, (\overline{CC}^* + \alpha_k)^{-1} \overline{\mathbf{r}}_k \right\rangle \\
&\leq \|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2 \sum_{k=n}^{l-1} \left\| C\mathbf{e}_l^\perp \right\| \left\| (\overline{CC}^* + \alpha_k)^{-1} \overline{\mathbf{r}}_k \right\| \\
&\leq \|\overline{\mathbf{e}}_n\|^2 - \|\overline{\mathbf{e}}_l\|^2 + 2 \sum_{k=n}^{l-1} \|C\mathbf{e}_l\| \left\| (\overline{CC}^* + \alpha_k)^{-1} \overline{\mathbf{r}}_k \right\|.
\end{aligned}
$$

Using the two inequalities together and Assumption 1 we get for general $m > n$ and any $l \in \{n, \dots, m-1\}$

$$
\begin{aligned}
\|L\mathbf{x}_m - L\mathbf{x}_n\|^2 &\leq 2 \|L\mathbf{x}_m - L\mathbf{x}_l\|^2 - 2 \|L\mathbf{x}_l - L\mathbf{x}_n\|^2 \\
&\leq 2 \|\overline{\mathbf{e}}_m\|^2 + 2 \|\overline{\mathbf{e}}_n\|^2 + \\
&\quad - 4 \|\overline{\mathbf{e}}_l\|^2 + 4 \sum_{k=n}^{m-1} \|C\mathbf{e}_l\| \left\| (\overline{CC}^* + \alpha_k)^{-1} \overline{\mathbf{r}}_k \right\| \\
&\leq 2 \|\overline{\mathbf{e}}_m\|^2 + 2 \|\overline{\mathbf{e}}_n\|^2 - 4 \|\overline{\mathbf{e}}_l\|^2 + \\
&\quad + 4(1 + \rho) \sum_{k=n}^{m-1} \|\mathbf{r}_l\| \left\| (\overline{CC}^* + \alpha_k)^{-1} \overline{\mathbf{r}}_k \right\|
\end{aligned}
$$

Let $l \in \{n, \dots, m-1\}$ be that particular index for which $\|\mathbf{r}_l\|$ is minimal, so that

$$
\begin{aligned}
\|L\mathbf{x}_m - L\mathbf{x}_n\|^2 \leq &\, 2 \|\overline{\mathbf{e}}_m\|^2 + 2 \|\overline{\mathbf{e}}_n\|^2 - 4 \|\overline{\mathbf{e}}_l\|^2 + \\
&+ 4(1 + \rho) \sum_{k=n}^{m-1} \|\mathbf{r}_k\| \left\| (\overline{CC}^* + \alpha_k)^{-1} \overline{\mathbf{r}}_k \right\|.
\end{aligned}
$$

The right-hand side of the inequality above becomes arbitrarly small, because the sequence $(\|\overline{\mathbf{e}}_k\|)_k$ is monotonically decreasing, in force of Proposition 13, and so converges to some limit $\epsilon \geq 0$ and the summation is the partial sum of a converging series (see Corollary 14). We have proved that the sequence $(\overline{\mathbf{x}}_n)_n$ is a Cauchy sequence and so converges to a certain limit $\overline{\mathbf{x}} \in \mathcal{X}$, and by continuity of $L^\dagger$ we get that

$$
P_{\mathcal{N}(L)^\perp}(\mathbf{x}_n) = L^\dagger L \mathbf{x}_n \to L^\dagger \overline{\mathbf{x}} = L^\dagger L \mathbf{x} = P_{\mathcal{N}(L)^\perp}(\mathbf{x}),
$$

for some $\mathbf{x} \in \mathcal{X}$. Accordingly the norm of the residual $P_{\mathcal{N}(L)^\perp}(\mathbf{r}_n) = P_{\mathcal{N}(L)^\perp}(\mathbf{y} - T\mathbf{x}_n)$ goes to $P_{\mathcal{N}(L)^\perp}(\mathbf{y} - T\mathbf{x})$, while in force of Corollary 14 the norm of this residual converges to zero and so $P_{\mathcal{N}(L)^\perp}(\mathbf{x})$ is the projection of a solution of the system, this with Remark 12 prooves point (i) of the theorem.

By construction, every iterate $\mathbf{x}_n$ satisfies

$$
\mathbf{x}_n - \mathbf{x}_0 = \sum_{k=0}^{n-1} \mathbf{h}_k \in \mathcal{R}(C*) = \mathcal{N}(C)^\perp,
$$

Therefore $\mathbf{x} - \mathbf{x}_0 \in \mathcal{N}(T)^\perp$, thanks to Assumption 6 (i) and so $\mathbf{x}$ is the particular solution of the system which is closest to $\mathbf{x}_0$ in the norm of $\mathcal{X}$ thus proving point (iii). □

Remark 16. *If $\mathbf{x}_0$ is a solution of the system then we have that $\|\mathbf{r}_0\| = \|T\mathbf{x}_0 - \mathbf{y}\| = 0$ and thus the algorithm does not start. In particular only a finite number of iteration is needed.*

Let us consider the inexact data case, in this circumstances Algorithm 2 is a regularization method, in fact we have the following

Theorem 17. *Assume that Assumption 6 holds for some $0 < \rho \leq \frac{1}{2}$ and let $\delta \mapsto \mathbf{y}^\delta$ be a function from $\mathbb{R}$ to $\mathcal{X}$ such that for all δ it holds $\left\| \mathbf{y} - \mathbf{y}^\delta \right\| \leq \delta$. For fixed τ and q denote with $\mathbf{x}^\delta$ the approximation of $\mathbf{x}^\dagger$ obtained with Algorithm 2. Then, as $\delta \to 0$, $\mathbf{x}^\delta$ goes to the solution of the system which is closest to $\mathbf{x}_0$.*

We omit the proof since it can be copied from [12][Theorem 2.3]; for further reference see also [1][Theorem 11.5]. Its essentials ingredients are the monotonicity proved in Proposition 13, the convergence to the exact solution in the exact data case proved in Theorem 15 and the continuity of the map $\delta \mapsto \mathbf{y}^\delta$.

5. Approximated Projected Iterated Tikhonov (APIT)

Let $\Omega \subset \mathcal{X}$ be closed and convex and such that $\mathbf{x}^\dagger \in \Omega$, let P_Ω be the metric projection of $\mathcal{X}$ on Ω and $T_\Omega = T\big|_\Omega$, $C_\Omega = C\big|_\Omega$. We want to constrain our problem so that $\forall n, \mathbf{x}_n \in \Omega$.

Definition 18. *We define the* metric projection *of $\mathbf{x} \in \mathcal{X}$ onto Ω as*

$$P_\Omega(\mathbf{x}) = \arg\min_{\mathbf{y} \in \Omega} \frac{1}{2} \left\| \mathbf{x} - \mathbf{y} \right\|^2.$$

Lemma 19. *Let Ω be a closed and convex subset of a Hilbert space $\mathcal{X}$, then P_Ω, the metric projection of $\mathcal{X}$ over n, is such that:*

(i) $\left\| P_\Omega(\mathbf{x}) - P_\Omega(\mathbf{y}) \right\|^2 \leq \left\| \mathbf{x} - \mathbf{y} \right\|^2 - \left\| (I - P_\Omega)(\mathbf{x}) - (I - P_\Omega)(\mathbf{y}) \right\|^2$;

(ii) $\left\| P_\Omega(\mathbf{x}) - P_\Omega(y) \right\|^2 \leq \langle \mathbf{x} - \mathbf{y}, P_\Omega(\mathbf{x}) - P_\Omega(\mathbf{y}) \rangle.$

Proof. The proof of the first can be found in [13]. The second is just a reformulation. $\square$

Remark 20. *Lemma 19 implies that the map P_Ω is* non-expansive.

In order to constrain Algorithm 1 we simply project at each iteration, obtaining the following

Algorithm 3 (APIT). *Let $\mathbf{x}_0 \in \mathcal{X}$ be fixed and set $n = 0$. Choose $\tau = \frac{1+2\rho}{1-2\rho}$ with ρ as in (10), and fix $q \in [2\rho, 1]$.*

While $\left\| \mathbf{r}_n \right\| > \tau\delta$, let $\tau_n = \left\| \mathbf{r}_n \right\| / \delta$ and $q_n = \max\left\{ q, 2\rho + \frac{1+\rho}{\tau_n} \right\}$, compute

$$\mathbf{h}_n = C^*(CC^* + \alpha_n I)^{-1} \mathbf{r}_n,$$

where α_n is such that

$$\left\| \mathbf{r}_n - C\mathbf{h}_n \right\| = q_n \left\| \mathbf{r}_n \right\|,$$

and update

$$\mathbf{x}_{n+1} = P_\Omega(\mathbf{x}_n + \mathbf{h}_n),$$
$$\mathbf{r}_{n+1} = \mathbf{y}^\delta - T\mathbf{x}_{n+1}.$$

We refer to Algorithm 3 as *Approximated Projected iterated Tikhonov* since this method can be seen as a preconditioned iterative method whose preconditioner is obtained by approximated Tikhonov and is projected at each iteration.

Remark 21. *Since $\mathbf{x}^\dagger \in \Omega$, we have $\left\| \mathbf{e}_n \right\| \leq \left\| \tilde{\mathbf{e}}_n \right\|$, where $\tilde{\mathbf{e}}_n$ is the error at the n-th iteration before of the projection into Ω, namely $\tilde{\mathbf{e}}_n = \mathbf{x}^\dagger - (\mathbf{x}_{n-1} + \mathbf{h}_{n-1})$.*

Using Lemma 19, the theoretical results reported in Section 3 for AIT can be easily extended to APIT.

Proposition 22. *With the same notations and assumptions of Proposition 3, the norm of the iteration error $\mathbf{e}_n$ decreases monotonically, namelly*

$$\left\| \mathbf{e}_n \right\|^2 - \left\| \mathbf{e}_{n+1} \right\|^2 \geq 2\rho \left\| (CC^* + \alpha_n I)^{-1} \mathbf{r}_n \right\| \left\| \mathbf{r}_n \right\|.$$

Proof. Using Lemma 19:

$$\|\mathbf{e}_n\|^2 - \|\mathbf{e}_{n+1}\|^2 = \|\mathbf{e}_n\|^2 - \left\|\mathbf{x}^\dagger - P_\Omega(\mathbf{x}_n + \mathbf{h}_n)\right\|^2 \geq$$

$$\geq \|\mathbf{e}_n\|^2 - \left\|\mathbf{x}^\dagger - (\mathbf{x}_n + \mathbf{h}_n)\right\|^2 = \|\mathbf{e}_n\|^2 - \|\mathbf{e}_n - \mathbf{h}_n\|^2.$$

Then, proceeding like in [8, Proposition 2], we have the thesis. □

Using the same approach of [8] to prove the results in Section 4, it can be shown that

Theorem 23. *Assume that the data are correct, i.e., $\delta = 0$, and that $\mathbf{x}_0$ is not a solution of the problem* (1). *Then, the sequence $\mathbf{x}_n$ converges as $n \to \infty$ to the solution on* (1) *which is closest to $\mathbf{x}_0$.*

Using this result and coping the proof of Theorem 2.3 in [12] we obtain

Theorem 24. *Let $\delta \mapsto \mathbf{y}^\delta$ be a function from $\mathbb{R}^+$ to $\mathcal{Y}$ such that* (2) *holds true for all $\delta > 0$. Under Assumption 1, for fixed parameters τ and q, denote by n_δ the corresponding stopping indices, and by $\mathbf{x}^\delta$ the resulting approximations. Then, as $\delta \to 0$, $\mathbf{x}^\delta$ converges to the solution of* (1) *that is closest to $\mathbf{x}_0$ in the norm of $\mathcal{X}$.*

5.1. Approximated Projected Iterated Tikhonov with General Penalty term (APIT-GP)

We now combine the previous two algorithms into a third one.

Algorithm 4 (APIT-GP)**.** *Let $\mathbf{x}_0 \in \mathcal{X}$ be fixed, set $n = 0$. Choose $\tau = \frac{1+2\rho}{1-2\rho}$ with ρ as as in* (10)*, and fix $q \in [2\rho, 1]$.*

While $\|\mathbf{r}_n\| > \tau\delta$, let $\tau_n = \|\mathbf{r}_n\|/\delta$ and $q_n = \max\left\{q, 2\rho + \frac{1+\rho}{\tau_n}\right\}$, compute

$$\mathbf{h}_n = C^*(CC^* + \alpha_n LL^*)^{-1}\mathbf{r}_n,$$

where α_n is such that

$$\|\mathbf{r}_n - C\mathbf{h}_n\| = q_n \|\mathbf{r}_n\|,$$

and update

$$\mathbf{x}_{n+1} = P_\Omega(\mathbf{x}_n + \mathbf{h}_n),$$
$$\mathbf{r}_{n+1} = \mathbf{y}^\delta - T\mathbf{x}_{n+1}.$$

We refer to Algorithm 4 as *Approximated Projected Iterated Tikhonov with General Penalty term* since this method can be seen as the combination of Algorithms 6 and 3.

From the numerical experiments in Section 6 we can see that this algorithm has good performances, however an eventual proof of its convergence meets some problems.

6. Numerical Examples

We apply our methods to the image deblurring problem. In this examples we use as T the blurring matrix with boundary conditions that respect the nature of the image and C the blurring matrix that has the same point spread function (PSF) of T with periodic boundary conditions.

As L we choose the discretization of the first derivative with periodic boundary conditions. This operator is obtained as the Kroneker product

$$L = L_1 \otimes I + L_1 \otimes I, \tag{17}$$

where L_1 is defined as

$$L_1 = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ -1 & & & & 1 \end{pmatrix},$$

and $\mathcal{N}(L_1) = \text{span}\{\mathbf{1}\}$, where $\mathbf{1}$ denotes the vector with all entries equal to 1. Therefore, L defined in (17) is singular and its null space is

$$\mathcal{N}(L) = \text{span}\{\mathbf{1}\}, \tag{18}$$

in other words the space of the constant vectors.

Since we are using the periodic boundary conditions the matrices C and L are block circulant matrices with circulant blocks (BCCB) and satisfy the Assumption 6 (ii), with Q being the bidimensional discrete Fourier transform [14].

Note that for L defined in (17), we have that $\mathcal{N}(L) \cap \mathcal{N}(C) = \{0\}$ thanks to (18). Indeed,

$$C\mathbf{1} = \mathbf{1}$$

because the sum of every row of C is equal to the sum of all entries of the PSF, which is equal to 1 to preserve the total light intensity.

Images can be seen as the measurement of the quantity of light received from a source and so they should not have negative values. Therefore we choose Ω to be the nonnegative cone

$$\Omega = \{\mathbf{x} \in \mathbb{R}^M : \forall i = 1, \ldots, M, \ x_i \geq 0\}.$$

Moreover, according to several numerical tests with different problems and the suggestions in [8], we fix

$$\rho = 10^{-3} \quad \text{and} \quad q = 0.7$$

in all our examples. To compare the quality of the restorations, we define the relative restoration error (RRE) as

$$\text{RRE} = \frac{\|\mathbf{x} - \mathbf{x}^\dagger\|}{\|\mathbf{x}\|},$$

where $\mathbf{x}$ is the computed solution. The minimum RRE in all table will be marked in bold.

For the construction of the examples we proceed in the following way. We first start with an image of $n_1 \times n_2$ pixels and blur it using any boundary conditions, e.g., the periodic one, using a PSF with of $m_1 \times m_2$ with $m_j < n_j$, $j = 1, 2$. Then, in order to simulate a real situation we cut out the the boundary from the blurred image of half the size of the PSF, i.e., of $\left\lceil \frac{m_j}{2} \right\rceil$. We then add some white Gaussian noise, we refer to

$$\nu = \frac{\delta}{\|\mathbf{y}\|} \tag{19}$$

as noise level.

We compare the restoration obtained with our methods with the original method AIT and with some other methods already present in the literature. In particular we consider the following methods:

- Hybrid [15];

- Two step iterative shrinkage/thresholding (Twist) [16];

- Range Restricted Arnoldi–Tikhonov (RRAT) [17];

- Flexible Arnoldi Tikhonov(FlexiAT) [18];

- Nonnegative Restarted Generalized Arnoldi Tikhonov (NN-ReStart-GAT) [18].

The Hybrid method is a Krylov method in which on the each Krylov space a Tikhonov regularization is implemented so to obtain a regularized solution, the regularization parameter is chosen with a particular modification of the generalized cross validation. In RRAT the Arnoldi Tikhonov decomposition is used to consider a certain Krylov space and then on this space the regularized solution is obtained using Tikhonov regularization, the regularization parameter is chosen solving the discrepancy principle equation. Twist is a method that combines regularization of the iterative shrinkage/thresholding methods and the splitting of the Iterative Re-Weighted Shrinkage methods. FlexiAT is a method that enables to introduce a regularization term into the equation and to adapt the Krylov subspace using the intermediate solutions in order to achieve better approximates the optimal regularization matrix. NN-ReStart-GAT is a projected version of ReStart-GAT, this method uses a restarted strategy, the inner iteration solves a Tikhonov regularized version of the problem exploiting Arnoldi
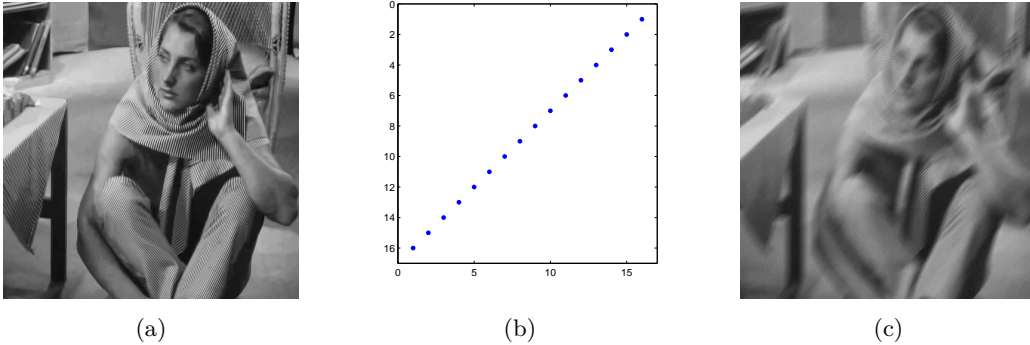
(a)   (b)   (c)

Figure 1: Barbara test case: (a) Test image (496×496 pixels), (b) Diagonal motion PSF (16×16 pixels), (c) Blurred image (496×496 pixels), $RRE = 0.16145$.

Tikhonov decomposition with a regularization term L such that $\|Lx\| \approx \|x\|_1$, then the outer iteration updates L so that the approximation of the $1-$norm gets better and better with the iterations.

In the following figures, the restored images are shown after a projection into Ω also for the methods that do not impose the nonnegative constraint. This allows a better visualization of the images in particular when they are affected by large ringing effects.

All the tests were performed using Matlab 9.0.0.341360 (R2016a) 64bit running on a laptop with an Intel core i7-6700HQ @ 2.60 GHz CPU and 8 GB of RAM.

Example 1 – Barbara. In this example we use the image Barbara (see Figure 1(a)), we blur the image with a diagonal motion PSF of 16 pixel (see Figure 1(b)) and add 3% of white Gaussian noise, i.e., $\nu = 0.03$ in (19) (for the result see Figure 1(c)).

Since the image is generic we use the *antireflexive* boundary conditions for the operator T, see [19]. From the comparison of the RRE history in Figure 8(a) we can see that, since there are no important black parts in the image, the introduction of the projection does not give any relevant improvement, in fact the graphs of APIT and AIT are overlayed and the same happens for AIT-GP and APIT-GP. The introduction of L is able to make the method faster and more accurate. In Figure 2 we can see the reconstructions with AIT-GP, AIT, Twist and the optimal reconstruction obtained with Hybrid. We can see from those reconstructions that the introduction of the regularization operator L let us have a better reconstruction of edges and details even though there is some ringing effect. In Table 1 we can find the comparison of the RRE and computational times with some other method from the literature, we can see that usually the method proposed are able to get better reconstructions in a smaller amount of time. We want to stress the fact that the stopping criterion of Hybrid was not able to effectively stop the method, so we printed also the optimal error; from this we can see that, even though AIT and APIT are outperformed by Hybrid, the introduction of the regularization operator gives better reconstructions.

FlexiAT, RRAT and NN-ReStart-GAT do not seems to perform well, moreover they also reach the maximum number of iterations without converging. This effect might be due to the fact that this methods are constructed for images that are mostly black, like astronomical or biological images, and not for photographic images like the one we are using in this example.

Example 2 – Grain. For this example we use the image Grain (see Figure 3(a)), we blur the image with a non symmetric Gaussian PSF (see figure 3(b)) from the toolbox Restore Tools by S. Berisha and J. Nagy, see [20], and add 5% of white Gaussian noise, i.e., $\nu = 0.05$ in (19) (for the result see Figure 3(c)). This image having a very huge black area is very useful to see the improvements introduced by the non-negative constraint. Again, since the image is generic at the boundary, we use the antireflexive boundary conditions. From the RRE history in Figure 8(b) we can see that the projection in the nonnegative cone gives great improvements in the quality of the reconstructions. In Figure 4 we can find the reconstruction with the AIT-GP, APIT, APIT-GP and Hybrid methods, from these we can see that L helps reconstructing the edges and that the projection let us have a more homogeneous result in the black areas. In order to better notice that we show in Figure 5 a detail of $|\mathbf{e}_{n^\delta}|$ in color map jet. In fact the reconstructions in Figure 4 are visualized so that no negative values are introduced, if the negative values were permitted we would get high oscillations in the black areas for the non-projected algorithms. In Table 2 we can find the RRE and computational times of our algorithms compared with some other method from the literature.

15

Figure 2: Barbara test case, reconstructions obtained with different methods: (a) AIT-GP, (b) AIT (c) Hybrid at the optimal iteration (the method does not stop properly), (d) TwIST.

Method	RRE	Iterations	Computational Time (sec.)
AIT	0.13489	3	0.60364
AIT-GP	**0.13132**	3	0.75130
APIT	0.13489	3	0.57012
APIT-GP	**0.13132**	3	0.73076
Hybrid	0.15919 (Opt.: 0.13337)	33 (Opt.: 5)	9.4639 (Opt. 1.4339)
TwIST	0.13906	6	4.5313
FlexiAT	0.16613	50	6.6665
RRAT	0.17308	50	13.782
NN-ReStart-GAT	0.16471	500	109.10

Table 1: Barbara test case: Comparison of the methods in term of relative restoration error (RRE), number of iterations and computational time.
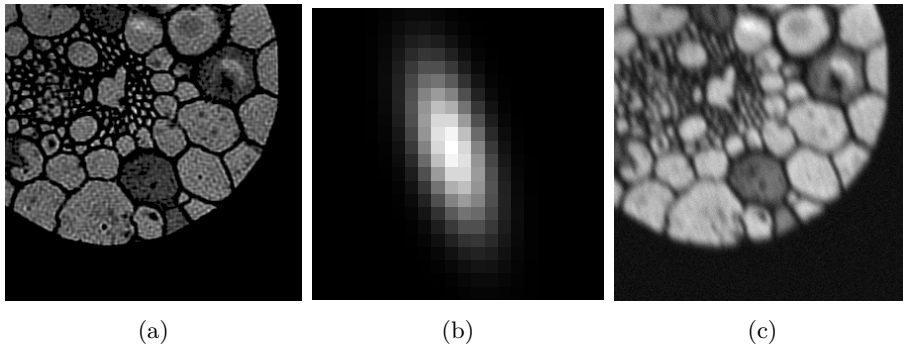
Figure 3: Grain test case: (a) Test image (300×300 pixels), (b) Non symmetric Gaussian PSF (22×22 px), (c) Blurred image (300×300 pixels), $RRE = 0.3680$.
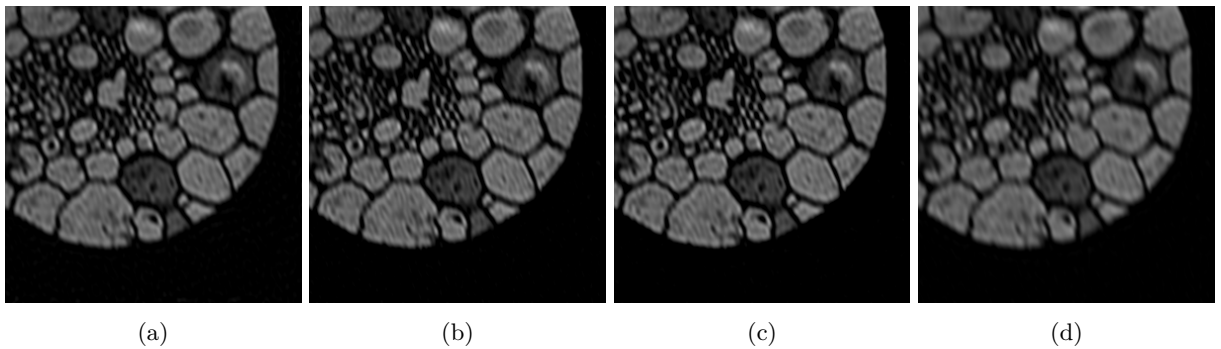
Figure 4: Grain test case, reconstructions obtained with different methods: (a) AIT-GP, (b) APIT, (c) APIT-GP, (d) Hybrid.
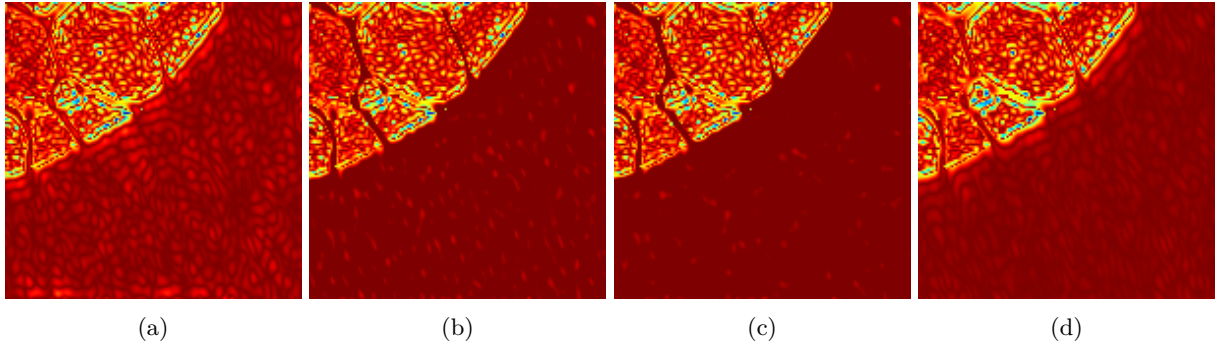
16

Figure 5: Grain test case, absolute value of the error in the south-east corner: (a) AIT-GP, (b) APIT, (c) APIT-GP, (d) Hybrid.

Method	RRE	Iterations	Computational Time (sec.)
AIT	0.28742	4	0.33530
AIT-GP	0.28485	4	0.33922
APIT	0.27393	30	1.4502
APIT-GP	**0.27063**	57	3.0685
Hybrid	0.32334	8	1.0395
TwIST	0.28743	16	5.3594
FlexiAT	0.35340	4	2.7535
RRAT	0.29767	9	0.084399
NN-ReStart-GAT	0.35044	52	4.7067

Table 2: Grain test case: Comparison of the methods in term of relative restoration error (RRE), number of iterations and computational time.

Example 3 – Satellite. In this last example we use the dataset `satellite` from the toolbox Restore Tools[20]. In this case the image is blurred with an astronomical PSF. The noise level ν is approximately of the 4% and has been computed using the knowledge of the true image. See Figure 6 for the true image, the PSF and the blurred and noisy data. Like in the example before this image, having a very huge black area, is very useful to see the improvements introduced by the non-negative constraint. Since near the boundary the image is all black we use the *zero* boundary conditions. In Figure 8(c) we find the RRE history, we can see that all the three methods we introduced give better result than AIT, since the image is for the most part black the better result is achieved with APIT. In this case, however, the best reconstruction is not the one given by APIT-GP, this is due to the fact that the introduction of the regularization operator is able to enanche the edges and some small noise, in the black area is recognised as edge and preserved. We must notice, none the less, that the difference between APIT-GP and APIT is very small. Finally in Figure 7 we can see the reconstructions for APIT, APIT-GP, RRAT and NN-Restart-GAT. In Table 3 we can find the comparison of the RRE and computational times with some other method from the literature, we can see that all the method proposed are able to get better reconstructions, even though in some cases the computational time is higher.
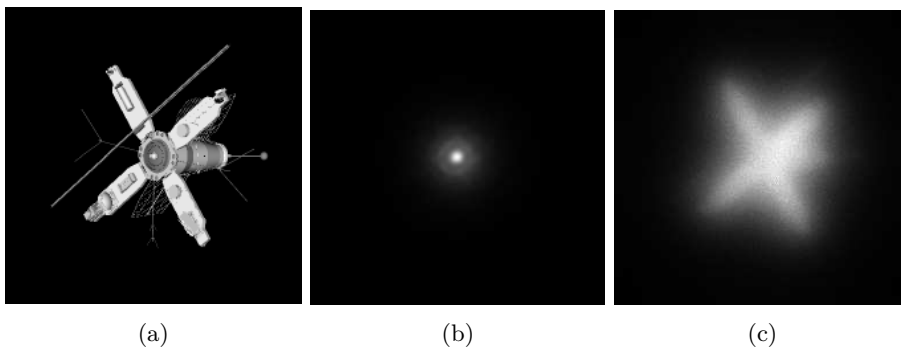

Figure 6: Satellite Test case: (a) Test image (256×256 pixels), (b) Astronomic PSF (256×256 pixels), (c) Blurred image (256×256 pixels), $RRE = 0.70464$.
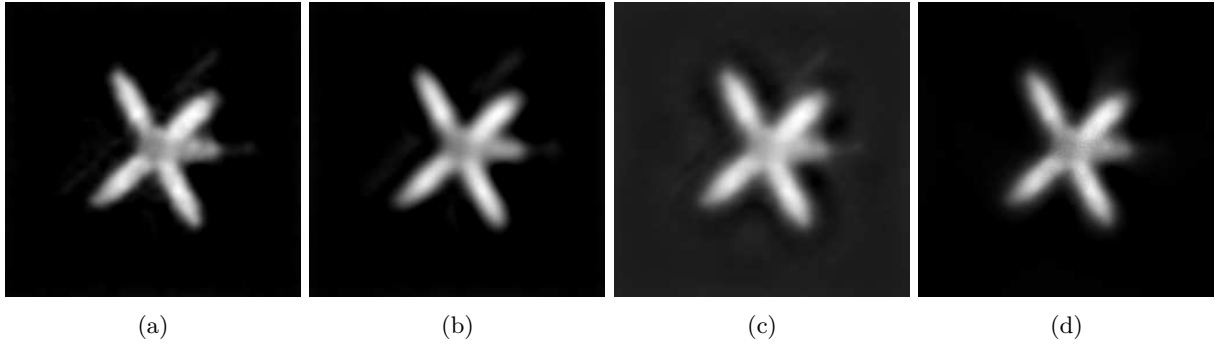
Figure 7: Satellite Reconstructions: (a) APIT, (b) APIT-GP, (c) RRAT, (d) FlexiAT.

Method	RRE	Iterations	Computational Time (sec.)
AIT	0.40996	7	0.48016
AIT-GP	0.42385	7	0.50797
APIT	**0.39801**	21	1.1806
APIT-GP	0.41129	32	1.7659
Hybrid	0.47663	50	4.5397
TwIST	0.47745	22	3.0313
FlexiAT	0.44875	8	0.18139
RRAT	0.45807	8	0.078776
NN-ReStart-GAT	0.83804	59	4.5397

Table 3: Satellite test case: Comparison of the methods in term of relative restoration error (RRE), number of iterations and computational time.
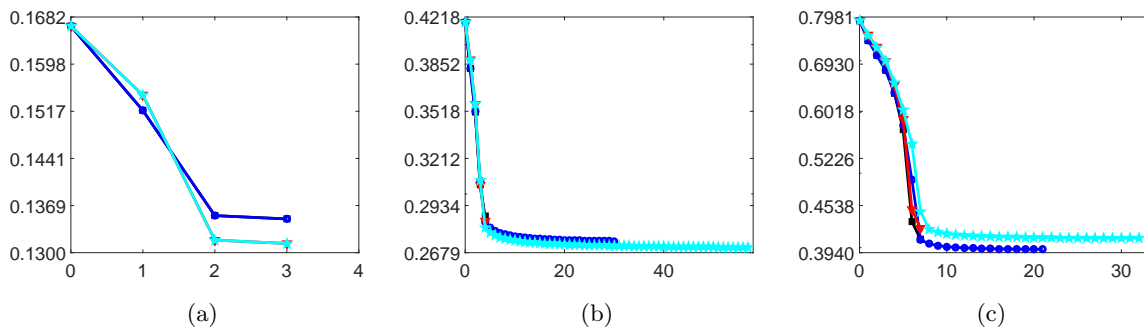

Figure 8: Evolution of the relative reconstruction error against the iterations for AIT, APIT, AIT-GP, and APIT-GP: (a) Barbara test case, (b) Cell test case, (c) Satellite test case. In black with stars AIT, in blue with circles APIT, in red with triangles AIT-GP, and in cyan with pentacles APIT-GP

7. Conclusions

In this paper, starting from the method proposed by Donatelli and Hanke in [8], we developed three algorithms which like the original method AIT, do not need any parameter estimate (a part from ρ which can be usually safely fixed equal to 10^{-3}). All three methods can be seen both as approximated version of iterated Tikhonov and as preconditioned algorithms in which the preconditioner is nonstationary meaning that, still preserving its structure, it changes at each iteration. The first developed method is ARIT and let us introduce a regularization term in order to achieve better reconstructions by adding features like edge sharpening. The second is APIT which introduces the projection in a closed and convex set. Finally we created APIT-GP which combines the two above. With all three methods we are able to embed the knowledge that we could have on the true solution (e.g. that has nonnegative entries) in the method and exploit it to get good reconstructions. Numerical experiments have proven that the introduction of the projection and the regularization is able to improve the quality of the restored images without changing to much the structure of the original method. The comparison with other methods in literature shows the robustness and accuracy of the algorithms developed.

[1] H. W. Engl, M. Hanke, A. Neubauer, Regularization of inverse problems, Vol. 375, Springer Science & Business Media, 1996.

[2] M. Hanke, P. C. Hansen, Regularization methods for large-scale problems, Survey on Mathematics for Industry 3 (4).

[3] M. Hanke, Conjugate gradient type methods for ill-posed problems, Vol. 327, CRC Press, 1995.

[4] C. W. Groetsch, The theory of Tikhonov regularization for Fredholm equations of the first kind, Boston MA:Pitman, 1984.

[5] M. Donatelli, A. Neuman, L. Reichel, Square regularization matrices for large linear discrete ill-posed problems, Numerical Linear Algebra with Applications 19 (6) (2012) 896–913.

[6] M. Hanke, C. W. Groetsch, Nonstationary iterated Tikhonov regularization, Journal of Optimization Theory and Applications 98 (1) (1998) 37–53.

[7] M. Donatelli, On nondecreasing sequences of regularization parameters for nonstationary iterated Tikhonov, Numerical Algorithms 60 (4) (2012) 651–668.

[8] M. Donatelli, M. Hanke, Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring, Inverse Problems 29 (9) (2013) 095008.

[9] P. Dell'Acqua, M. Donatelli, C. Estatico, Preconditioners for image restoration by reblurring techniques, Journal of Computational and Applied Mathematics 272 (2014) 313–333.

[10] L. Eldén, A weighted pseudoinverse, generalized singular values, and constrained least squares problems, BIT Numerical Mathematics 22 (4) (1982) 487–502.

[11] G. Huang, L. Reichel, F. Yin, Projected nonstationary iterated Tikhonov regularization, BIT Numerical Mathematics 56 (2) (2016) 467–487. `doi:10.1007/s10543-015-0568-7`.
URL `http://dx.doi.org/10.1007/s10543-015-0568-7`

[12] M. Hanke, A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems, Inverse problems 13 (1) (1997) 79.

[13] E. H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, University of Wisconsin, 1971.

[14] P. C. Hansen, J. G. Nagy, D. P. O'leary, Deblurring images: matrices, spectra, and filtering, Vol. 3, Siam, 2006.

[15] J. Chung, J. G. Nagy, D. P. O'Leary, A weighted GCV method for Lanczos hybrid regularization, Electronic Transactions on Numerical Analysis 28 (2008) 149–167.

[16] J. M. Bioucas-Dias, M. A. Figueiredo, A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration, Image Processing, IEEE Transactions on 16 (12) (2007) 2992–3004.

[17] A. Neuman, L. Reichel, H. Sadok, Algorithms for range restricted iterative methods for linear discrete ill-posed problems, Numerical Algorithms 59 (2) (2012) 325–331.

[18] S. Gazzola, J. G. Nagy, Generalized Arnoldi–Tikhonov method for sparse reconstruction, SIAM Journal on Scientific Computing 36 (2) (2014) B225–B247.

[19] S. Serra-Capizzano, A note on antireflective boundary conditions and fast deblurring models, SIAM Journal on Scientific Computing 25 (4) (2004) 1307–1325.

[20] S. Berisha, J. G. Nagy, Iterative methods for image restoration, Tech. rep., Department of Mahtematics and Computer Science, Emory University, `http://www.mathcs.emory.edu/~nagy/RestoreTools/IR.pdf` (2012).