Generalized Structure Preserving Preconditioners for Frame-Based Image Deblurring

Davide Bianchi[1,†] and Alessandro Buccini[2,†,*]

[1] Department of Science and High Technology, University of Insubria, 22100 Como, Italy; email: d.bianchi9@uninsubria.it

[2] Department of Mathematics and Computer Science, University of Cagliari, 09124 Cagliari, Italy; email: alessandro.buccini@unica.it

[*] Correspondence: alessandro.buccini@unica.it

[†] These authors contributed equally to this work.

Abstract: We are interested in fast and stable iterative regularization methods for image deblurring problems with space invariant blur. The associated coefficient matrix has a Block Toeplitz Toeplitz Blocks (BTTB) like structure plus a small rank correction depending on the boundary conditions imposed on the imaging model. In the literature, several strategies have been proposed in the attempt to define proper preconditioner for iterative regularization methods that involve such linear systems. Usually, the preconditioner is chosen to be a Block Circulant with Circulant Blocks (BCCB) matrix because it can efficiently exploit Fast Fourier Transform (FFT) for any computation, including the (pseudo-)inversion. Nevertheless, for ill-conditioned problems, it is well known that BCCB preconditioners cannot provide a strong clustering of the eigenvalues. Moreover, in order to get an effective preconditioner, it is crucial to preserve the structure of the coefficient matrix. On the other hand, thresholding iterative methods have been recently successfully applied to image deblurring problems, exploiting the sparsity of the image in a proper wavelet domain. Motivated by the results of recent papers, the main novelty of this work is combining nonstationary structure preserving preconditioners with general regularizing operators which hold in their kernel the key features of the true solution that we wish to preserve. Several numerical experiments shows the performances of our methods in terms of quality of the restorations.

---

1. Introduction

In image deblurring we are concerned in reconstructing an approximation of an image from blurred and noisy measurements. This process can be modeled by an integral equation of the form

$$g(x,y) = (\kappa * f)(x,y) = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \kappa(x,y,x',y')f(x',y')\,dx'\,dy' + \eta(x,y), \qquad (x,y) \in \Omega \subset \mathbb{R}^2, \quad (1)$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is the original image and $g : \mathbb{R}^2 \to \mathbb{R}$ is the observed imaged which is obtained from a combination of an (Hilbert-Schmidt) integral operator, represented by $\kappa : \mathbb{R}^4 \to \mathbb{R}$, and the add of some (unavoidable) noise $\eta : \mathbb{R}^2 \to \mathbb{R}$ coming from, e.g., perturbations on the observed data, measurement errors, and approximation errors. By assuming the kernel κ to be compactly supported and considering the ideal case $\eta = 0$ equation (1) becomes

$$g = K \cdot f,$$

where K is a compact linear operator. In this contest, κ is generally called *point spread function* (PSF).

Figure 1. Field of view. We see what is inside the square box.

Considering an uniform grid, images are represented by their color intensities measured on the grid (pixels). In this paper, for the sake of simplicity, we will deal only with square and gray-scale images, even if all the techniques presented here carry over to images of different sizes and colors as well.

Collected images are available only in a finite region, the field of view (FOV), and the measured intensities near the boundary are affected by data which lie outside the FOV; see Figure 1 for an illustration.

Denoting by **g** and **f** the stack ordered vectors corresponding to the observed image and the true image, respectively, the discretization of (1) becomes

$$\mathbf{g} = K\mathbf{f} + \boldsymbol{\eta}, \tag{2}$$

where the matrix K is of size $m^2 \times k^2$, being m and k the dimensions (in pixels) of the original picture. The matrix K is often called the *blurring matrix*. When imposing proper Boundary Conditions (BCs), the matrix K becomes square $m^2 \times m^2$ and in some cases, depending on the BCs and the symmetry of the PSF, it can be diagonalized by discrete trigonometric transforms. Indeed, specific BCs induce specific matrix structures that can be exploited to lessen the computational costs using fast algorithms. Of course, since BCs are artificially introduced, their advantages could come with drawbacks in terms of reconstruction accuracy, depending on the type of problem. The BCs approach forces a functional dependency between the elements of **f** external to the FOV and those internal to this area. If the BC model is not a good approximation of the real world outside the FOV, the reconstructed image can be severely affected by some unwanted artifacts near the boundary, called ringing effects; see, e.g., [1].

The choice of the different BCs can be driven by some additional knowledge on the true image and/or from the availability of fast transforms to diagonalize the matrix K within $O(m^2 \log(m))$ arithmetic operations. Indeed, the matrix-vector product can be always computed by the 2D FFT, after a proper padding of the image to convolve, (see, e.g., [2]), while the availability of fast transforms to diagonalize the matrix K depends on the BCs. Among the BCs present in the literature, we consider: Zero (Dirichlet), Periodic, Reflective and Anti-Reflective, but our approach can be extended to other BCs like, e.g., synthetic BCs [3] or high order BCs [4,5]. See Figure 2 for an illustration of the described BCs.

On the other hand, since equation (2) is the product of the discretization of a compact operator, K is severely ill-conditioned and may be singular. Such linear systems are commonly referred to as linear discrete ill-posed problems; see, e.g., [6] for a discussion. Therefore a good approximation of **f** cannot be obtained from the algebraic solution (e.g., the least-square solution) of (2), but regularization methods are required. The basic idea of regularization is to replace the original ill-conditioned problem with a nearby well-conditioned problem, whose solution approximates the
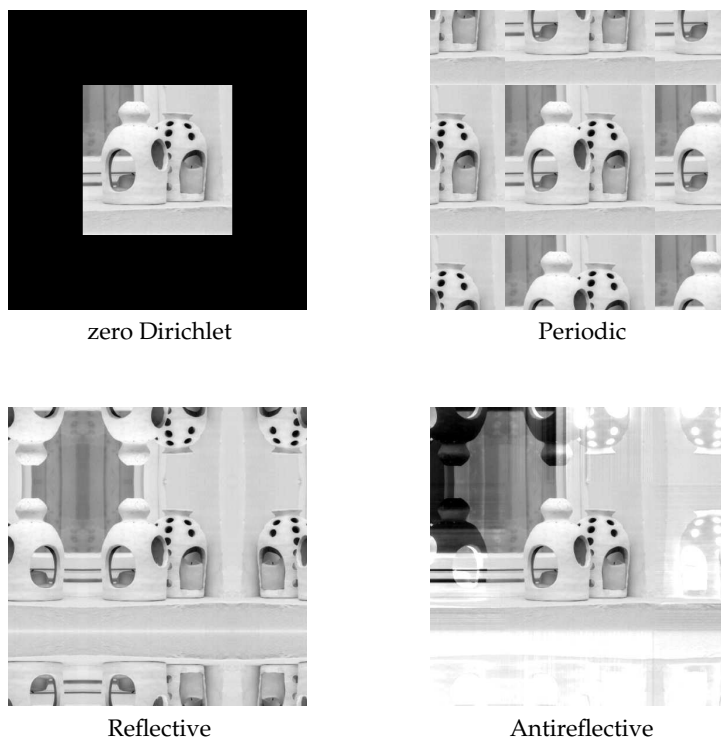
Figure 2. Examples of boundary conditions.

true solution. One of the popular regularization techniques is Tikhonov regularization and it amounts in solving

$$\min_{\mathbf{f}}\{\|K\mathbf{f} - \mathbf{g}\|_2^2 + \mu\|\mathbf{f}\|_2^2\}, \tag{3}$$

where $\|\cdot\|_2$ denotes the vector 2-norm and $\mu > 0$ is a regularization parameter to be chosen. The first term in (3) is usually refereed to as fidelity term and the second as regularization term. This translates into solving a linear problem for which many efficient methods have been developed for computing its solution and for estimating the regularizing parameter μ, see [6]. This approach unfortunately comes with a drawback: the edges of restored images are usually over-smoothed. Therefore, nonlinear strategies have been employed, in order to overcome this unpleasant property, like total variation (TV) [7] and thresholding iterative methods [8,9]. That said, typically many nonlinear regularization methods have an inner step that apply a least-square regularization and therefore can benefit from strategies previously developed for such simpler model.

In the present paper, both the regularization strategies that we propose share two common ingredients: wavelet decomposition and ℓ_1-norm minimization on the regularization term. This is motivated by the fact that the wavelet coefficients (under some basis) of most real images are usually very sparse. In particular, here we consider the tight frame systems used in [10–12], but the result can be easily extended to any framelet/wavelet system. Let W^* be a wavelet or tight-frame synthesis operator ($W^*W = I$), the wavelets or tight-frame coefficients of the original image $\mathbf{f}$ are $\mathbf{x}$ such that

$$\mathbf{f} = W^*\mathbf{x}, \qquad \text{and the blurring operator becomes } A = KW^*.$$

Within this frame set, the model equation (2) translates into

$$\mathbf{g} = A\mathbf{x}. \tag{4}$$

Recently, in [13], a new technique was proposed which directly applies a single preconditioning operator P to 4. The new preconditioned system becomes

$$P\mathbf{g} = PA\mathbf{x}.$$

Combining this approach with a soft-thresholding technique such as the modified linearized Bregman splitting algorithm [14], in order to mimic the ℓ_1-norm minimization, leads to the following preconditioned iterative scheme

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \tau P(\mathbf{g} - A\S_\mu(\mathbf{x}_n)),$$

where τ is a positive relaxation parameter and $\S_\mu(\cdot)$ is the soft-thresholding function as defined in 6. Hereafter we will put $\tau = 1$: this is justified by applying an implicit rescaling of the preconditioned system matrix PA.

The paper is organized as follows: in Section 2 we propose a generalization of an approximated iterative Tikhonov scheme that was firstly introduced in [15] and then developed and adapted into different settings in [16,17]. Here the preconditioner P takes the form

$$P = B^* \left(BB^* + \alpha_n \Lambda\Lambda^* \right)^{-1},$$

where B is an approximation of A, in the sense that $B = CW^*$ with C the discretization of the same problem (1) as the original blurring matrix K but imposing Periodic BCs. The operator $\Lambda\Lambda^*$ can be a function of CC^* or the discretization of a differential operator. The method is nonstationary and the parameter α_n is computed by solving a nonlinear problem with a computational cost of $O(m^2)$. Related work on this kind of preconditioner can be found in [18–20]. In Section 3 we define a class of preconditioners P endowed with the same structure of the system matrix A, as initially proposed in [21] and then further developed in [22]. It is called structure preserving reblurring preconditioning strategy and we combine it with the generalized regularization filtering approach of the preceding Section 2. The idea is to preserve both the informations carried over by the spectra of the operator A and the structure itself of the operator induced by the best fitting BCs. Section 4 contains a selection of significant numerical examples which confirm the robustness and quality of the proposed regularization schemes. Section 5 provides a resume of the techniques presented in this work and draws some conclusions. Finally, in Appendix A are provided proofs of convergence and regularization properties of the proposed algorithms.

2. Preconditioned Iterated Soft-thresholding Tikhonov with general regularizing operator

2.1. Preliminary definitions

Before proceeding further, let us introduce here some definitions and notations that will be used in the forthcoming sections. We consider

$$K : \left(\mathbb{R}^{m^2}, \| \cdot \| \right) \to \left(\mathbb{R}^{m^2}, \| \cdot \| \right)$$

to be the discretization of a compact linear operator

$$\mathbf{g} = K\mathbf{f},$$

where the Euclidean 2-norm $\| \cdot \|$ is induced by the standard Euclidean inner product

$$\langle \mathbf{f}^{(1)}, \mathbf{f}^{(2)} \rangle_{\mathbb{R}^{m^2}} = \sum_{j=1}^{m^2} f_j^{(1)} f_j^{(2)}.$$

Hereafter, we will specify the vector space where the inner product $\langle \cdot, \cdot \rangle$ acts only whenever it is necessary for disambiguation. The analysis that will follow in the next sections will be performed generally on a perturbed data $\mathbf{g}^\delta$, namely

$$\mathbf{g}^\delta = K\mathbf{f},$$

with $\mathbf{g}^\delta = \mathbf{g} + \boldsymbol{\eta}$, and where $\boldsymbol{\eta}$ is a noise vector such that $\|\boldsymbol{\eta}\| = \delta$, δ is called the noise level.

Let

$$C : \left(\mathbb{R}^{m^2}, \| \cdot \| \right) \to \left(\mathbb{R}^{m^2}, \| \cdot \| \right)$$

be the discretization of a compact linear operator that approximates K, in a sense that will be specified later. Let

$$W : \left(\mathbb{R}^{m^2}, \| \cdot \| \right) \to (\mathbb{R}^s, \| \cdot \|)$$

be such that

$$W^*W = I,$$

where $W^* : \mathbb{R}^s \to \mathbb{R}^{m^2}$ indicates the adjoint operator of W, i.e., $\langle W\mathbf{f}, \mathbf{u} \rangle_{\mathbb{R}^s} = \langle \mathbf{f}, W^*\mathbf{u} \rangle_{\mathbb{R}^{m^2}}$ for each pair $\mathbf{f} \in \mathbb{R}^{m^2}, \mathbf{u} \in \mathbb{R}^s$. We define

$$\mathbf{x} = W\mathbf{f}, \qquad A = KW^*, \qquad B = CW^*.$$

Let us introduce the following matrix norm. Given a generic linear operator

$$L : (\mathbb{R}^s, \| \cdot \|_\infty) \to \left(\mathbb{R}^{m^2}, \| \cdot \| \right),$$

where $\| \cdot \|_\infty$ is the sup norm, let us define the matrix norm $\|\|\cdot\|\|$ as

$$\|\|L\|\| := \sup_{\|\mathbf{x}\|_\infty \equiv 1} \|L\mathbf{x}\|_2. \tag{5}$$

Finally, let $\mu \geq 0$ and let $\S_\mu : \mathbb{R}^s \to \mathbb{R}^s$ be such that

$$[\S_\mu(\mathbf{u})]_i = S_\mu(u_i), \tag{6}$$

with S_μ the soft-thresholding function

$$S_\mu(u_i) = \mathrm{sgn}(u_i) \max \left\{ |u_i| - \mu, 0 \right\}.$$

2.2. General regularization operator as $h(CC^*)$

Let $h : [0, \|CC^*\|^2] \to \mathbb{R}$ be a continuous function such that

$$0 < c_1 \leq h(\sigma^2) \leq c_2,$$

where c_1, c_2 are positive constants, and define $c := c_1/c_2$. By the continuity of h and by well-known facts from functional analysis [23] we can write $h(CC^*)$ as the operator defined by

$$h(CC^*)[\mathbf{f}] := \int h(\sigma^2) dE_{\sigma^2}(\mathbf{f}) = \sum_{k=1}^\infty h(\sigma_k^2) \langle \mathbf{f}, \mathbf{u}_k \rangle \mathbf{u}_k,$$

where $\{E_{\sigma^2}\}_{\sigma^2 \in \sigma(CC^*)}$ is the spectral decomposition of a (generic) self-adjoint operator CC^* and $(\sigma_k, \mathbf{v}_k, \mathbf{u}_k)_{k \in \mathbb{N}}$ is the singular value expansion of C. We summarize the computations in Algorithm 1.

A rigorous and full detailed analysis of the preceding algorithm will be performed in Appendix A. In order to prove all the desired properties we will need a couple of assumptions on the operators K, C, and on the parameter μ, that we present here below.

Algorithm 1 PISTA_h

Fix $\mathbf{z}^0 \in \mathbb{R}^s$, $\delta > 0$ and set $\mathbf{x}^0 = \S_\mu(\mathbf{z}^0)$, $\mathbf{r}^0 = \mathbf{g} - A\mathbf{x}^0$.

Set $\rho \in (0, c/2)$ and $q \in (2\rho, c)$.

Compute $\tau = \frac{1+2\rho}{c-2\rho}$ and $\mathbf{r}^n = \mathbf{g} - A\mathbf{x}^n$.

while $\|\mathbf{r}^n\| > \tau\delta$ **do**

  Compute $\tau_n := \|\mathbf{r}^n\|/\delta$.

  Compute $q_n := \max\{q, 2\rho + (1+\rho)/\tau_n\}$.

  Compute α_n such that

$$\alpha_n \|(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}^n\| = \frac{q_n}{c_1}\|\mathbf{r}^n\|. \tag{7}$$

Compute

$$\mathbf{h}^n = WC^*(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}^n. \tag{8}$$

Compute

$$\begin{cases} \mathbf{z}^{n+1} = \mathbf{z}^n + \mathbf{h}^n, \\ \mathbf{x}^{n+1} = \S_\mu(\mathbf{z}^{n+1}). \end{cases}$$

Assumption 1.

$$\|(C - K)\mathbf{f}\| \leq \rho\|K\mathbf{f}\|, \qquad \forall\, \mathbf{f} \in \mathbb{R}^{m^2}, \tag{9a}$$

and

$$\mu \leq \frac{\rho\delta}{\||B\||}, \tag{9b}$$

with a fixed $0 < \rho < c/2$, where $\delta = \|\boldsymbol{\eta}\|$ is the noise level and where $\||\cdot\||$ is the operator norm defined in (5).

Let us observe that Equation (9a) translates into

$$\|(B - A)\mathbf{u}\| \leq \rho\|A\mathbf{u}\|, \qquad \forall\, \mathbf{u} \in \mathbb{R}^s. \tag{10}$$

Let us spread some light on the preceding conditions. Assumption (9a), or equivalently (10), is a strong assumption. It may be hard to satisfy it for every specific problem, as it implies

$$(1 - \rho)\|K\mathbf{v}\| \leq \|C\mathbf{v}\| \leq (1 + \rho)\|K\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{R}^{m^2},$$

or equivalently

$$(1 - \rho)\|A\mathbf{u}\| \leq \|B\mathbf{u}\| \leq (1 + \rho)\|A\mathbf{u}\| \qquad \text{for all } \mathbf{u} \in \mathbb{R}^s, \tag{11}$$

that is, K and C are spectrally equivalent. Nevertheless, in image deblurring the boundary conditions have a very local effect, i.e., the approximation error $C - K$ can be decomposed as

$$C - K = E + R,$$

where E is a matrix of small norm (and the zero matrix if the PSF is compactly supported), and R is a matrix of small rank, compared to the dimension of the problem. This suggests that Assumption (9a) needs to be satisfied only in a relatively small subspace, supposedly being a zero measure subspace. In particular only for every $\mathbf{e}_\delta^n$, with $n \geq N$ and N fixed, such that Proposition A1 in Appendix A could hold. All the numerical experiments are consistent with this observation but for a deeper understanding and a full treatment of this aspect we refer the reader to [15, Section 4].

On the other hand instead, Assumption (9b) is quite natural. It is indeed equivalent to require that

$$\| B \left(\mathbf{u} - \S_\mu(\mathbf{u}) \right) \| \leq \rho\delta$$

that is, the soft-thresholding parameter $\mu = \mu(\delta)$ is continuously noise-dependent and it holds that $\mu(\delta) \to 0$ as $\delta \to 0$.

2.3. General regularization operator as $\Lambda\Lambda^*$

In image deblurring, in order to better preserve the edges of the reconstructed solution, it is usually introduced a differential operator $\Lambda\Lambda^*$, where $\Lambda : \mathcal{X} \to \mathcal{Y}$ is chosen as a first or second order differential operator which holds in its kernel all these functions which posses the key features of the true solution that we wish to preserve. In particular, since we are interested to recover the edges and curves of discontinuities of the true image, it is a common choice to rely on the Laplace operator with Neumann BCs, see [24]. In these recent papers [25,26], observing the spectral distribution of the Laplacian, it was proposed to substitute $\Lambda\Lambda^*$ with

$$h(CC^*) = \left(I - \frac{CC^*}{\|CC^*\|} \right)^j,$$

with $j \in \mathbb{N}$.

Adding some new assumptions, we propose a modified version of the preceding Algorithm 1 that can take into account directly the operator Λ.

Assumption 2.

$$\mathrm{Ker}(K) \cap \mathrm{Ker}(\Lambda) = \{0\};$$

$$C_{|\mathrm{Ker}(\Lambda)} = K_{|\mathrm{Ker}(\Lambda)};$$

C and Λ are diagonalized by the same unitary transform.

Algorithm 2 PISTA$_\Lambda$

Fix $\mathbf{z}^0 \in \mathbb{R}^s$, $\delta > 0$ and set $\mathbf{x}^0 = \S_\mu(\mathbf{z}^0)$, $\mathbf{r}^0 = \mathbf{g} - A\mathbf{x}^0$.
Set $\rho \in (0, 1/2)$ and $q \in (2\rho, 1)$.
Compute $\tau = \frac{1+2\rho}{1-2\rho}$ and $\mathbf{r}^n = \mathbf{g} - A\mathbf{x}^n$.
while $\|\mathbf{r}^n\| > \tau\delta$ **do**
    Compute $\tau_n := \|\mathbf{r}^n\|/\delta$.
    Compute $q_n := \max\{q, 2\rho + (1+\rho)/\tau_n\}$.
    Compute α_n such that

$$\alpha_n \|(CC^* + \alpha_n\Lambda\Lambda^*)^{-1}\mathbf{r}^n\| = q_n\|\mathbf{r}^n\|.$$

Compute

$$\mathbf{h}^n = WC^*(CC^* + \alpha_n\Lambda\Lambda^*)^{-1}\mathbf{r}^n. \tag{12}$$

Compute

$$\begin{cases} \mathbf{z}^{n+1} = \mathbf{z}^n + \alpha_n\mathbf{h}^n, \\ \mathbf{x}^{n+1} = \S_\mu(\mathbf{z}^{n+1}). \end{cases}$$

We skip all the proofs of convergence since they can be recovered easily adapting the proofs in Section A with [16, Section 4].

3. Structured PISTA with general regularizing operator

The structured case is a generalization of what was developed in [21,22], merging these ideas with the general approach described in Section 2. We skip some details since they can be easily recovered from the aforementioned papers.

The blurring matrix K is made by two parts: the PSF and the BCs inherited by the discretization. Different types of structured matrices are given rise by this latter choice. Without loosing generality and for the sake of simplicity, we consider a square PSF $H_\kappa \in \mathbb{R}^{k \times k}$ and we suppose that the center of the PSF is known.

Consider the pixels $\kappa_{i,j}$ of the PSF, we can define the following generating function $\kappa : \mathbb{R}^2 \to \mathbb{C}$

$$\kappa(x_1, x_2) = \sum_{i,j=-m+1}^{m-1} \kappa_{i,j} e^{\hat{\imath}(ix_1 + jx_2)}, \tag{13}$$

where $\hat{\imath}^2 = -1$ and we assumed that $\kappa_{i,j} = 0$ if the entry $(\kappa_{i,j})$ does not belong to H_κ [5]. Observe that $\kappa_{j,j}$ can be seen as the Fourier coefficients of $\kappa \in \text{span}\{e^{\hat{\imath}(ix_1 + jx_2)}, i, j = -k, \ldots, k\}$, so that the same information is contained in the generating function κ and in H_κ.

Summarizing the notation that we set in the Introduction about the BCs, we have

Zero BCs:	$K = \mathcal{T}_m(\kappa)$,
Periodic BCs:	$K = \mathcal{C}_m(\kappa) = \mathcal{T}_m(\kappa) + \mathcal{B}_m^{\mathcal{C}}(\kappa)$,
Reflective BCs:	$K = \mathcal{R}_m(\kappa) = \mathcal{T}_m(\kappa) + \mathcal{B}_m^{\mathcal{R}}(\kappa)$,
Anti-Reflective BCs:	$K = \mathcal{AR}_m(\kappa) = \mathcal{T}_m(\kappa) + \mathcal{B}_m^{\mathcal{AR}}(\kappa)$.

We notice that, since the continuous operator is shift-invariant, in all these four cases K has a Toeplitz structure $\mathcal{T}_m(\kappa)$ which depends on κ plus a correction term $\mathcal{B}_m^{\mathcal{X}}(\kappa)$, $\mathcal{X} = \mathcal{C}, \mathcal{R}, \mathcal{AR}$ which depends on the chosen BCs.

In conclusion, we employ the unified notation $K = \mathcal{M}_m(\kappa)$, where $\mathcal{M}(\cdot)$ can be any of the classes of matrices just introduced (i.e. $\mathcal{T}, \mathcal{C}, \mathcal{R}, \mathcal{AR}$). With this notation we wish to highlight the two crucial elements that determine K: the blurring phenomenon associated with the PSF described by κ and the chosen BCs represented by $\mathcal{M}$.

Given the generating function κ (13) associated to the PSF H_κ, let us compute the eigenvalues $u_{i,j}$ of the corresponding BCCB matrix $\mathcal{C}_m(\kappa) := C$ by means of the 2D-FFT, where $i, j = 0, \cdots, m-1$. Fix a regularizing (differential) operator $\Lambda\Lambda^*$ as in Section 2, and suppose that the Assumptions 1 and 2 hold. The differential operator can be of the form $\Lambda\Lambda^* = h(CC^*)$, as in Algorithm 1 as well. Let now

$$v_{i,j} = \frac{\overline{u}_{i,j}}{|u_{i,j}|^2 + \alpha_n |\sigma_{i,j}|^2}, \quad (\overline{u}_{i,j} \text{ the complex conjugate of } u_{i,j})$$

be the new eigenvalues after the application of the Tikhonov filter to $u_{i,j}$, where $\sigma_{i,j}$ are the eigenvalues (singular values) of Λ and α_n is computed as in Algorithm 1-2. Let us compute now the coefficients $\hat{\kappa}_{i,j}$ of

$$\hat{\kappa}(x_1, x_2) = \sum_{i,j=-m+1}^{m-1} \hat{\kappa}_{i,j} e^{\hat{\imath}(ix_1 + jx_2)} \tag{14}$$

by the means of a 2D-iFFT and, finally, let us define

$$P = \mathcal{M}_m(\hat{\kappa}),$$

where $\mathcal{M}(\cdot)$ corresponds to the most fitting BCs for the model problem (1).

We are now ready to formulate the last method whose computation are reported in Algorithm 3.

Algorithm 3 Struct-PISTA$_\Lambda$

Fix H_κ, BCs, Λ.

Set $C = \mathcal{C}_m(\kappa)$.

Get $\{u_{i,j}\}_{i,j=0}^{n-1}$ by computing an FFT of H_κ.

Fix $\mathbf{z}^0 \in \mathbb{R}^s$, $\delta > 0$ and set $\mathbf{x}^0 = \S_\mu(\mathbf{z}^0)$, $\mathbf{r}^0 = \mathbf{g} - K\mathbf{x}^0$.

Set $\rho \in (0, 1/2)$ and $q \in (2\rho, 1)$.

Compute $\tau = \frac{1+2\rho}{1-2\rho}$ and $\mathbf{r}^n = \mathbf{g} - K\mathbf{x}^n$.

while $\|\mathbf{r}^n\| > \tau\delta$ **do**

    Compute $\tau_n := \|\mathbf{r}^n\|/\delta$.

    Compute $q_n := \max\{q, 2\rho + (1+\rho)/\tau_n\}$.

    Compute α_n such that

$$\alpha_n \|(CC^* + \alpha_n \Lambda\Lambda^*)^{-1}\mathbf{r}^n\| = q_n\|\mathbf{r}^n\|.$$

Compute $v_{i,j} = \dfrac{\overline{u}_{i,j}}{|u_{i,j}|^2 + \alpha_n|\sigma_{i,j}|^2}$.

Get the mask $\widetilde{H}$ of the coefficients $\hat{\kappa}_{i,j}$ of $\hat{\kappa}$ of (14) by computing an IFFT of $\{v_{i,j}\}_{i,j=0}^{m-1}$.

Generate the matrix $P := \mathcal{M}_m(\hat{\kappa})$ from the coefficient mask $\widetilde{H}$ and BCs.

Compute

$$\mathbf{h}^n = P\mathbf{r}^n. \tag{15}$$

Compute

$$\begin{cases} \mathbf{z}^{n+1} = \mathbf{z}^n + \mathbf{h}^n, \\ \mathbf{x}^{n+1} = \S_\mu(\mathbf{z}^{n+1}). \end{cases}$$

In the case that $\Lambda\Lambda^* = h(CC^*)$, then the algorithm is modified in the following way:

$$\rho \in (0, c/2), \quad q \in (2\rho, c), \quad \alpha_n \|(CC^* + \alpha_n\Lambda\Lambda^*)^{-1}\mathbf{r}^n\| = \frac{q_n}{c_1}\|\mathbf{r}^n\|$$

where $0 < c_1 \leq h(\sigma^2) \leq c_2$, $c := c_1/c_2$. We will denote this version by Struct-PISTA$_h$. We will not provide a direct proof of convergence for this last algorithm. Let us just observe that the difference between (15) and (12)-(8) is just a correction of small rank and small norm.

4. Numerical experiments

We now compare the proposed algorithms with some methods from the literature. In particular, we consider the AIT-GP algorithm described in [16] and the ISTA algorithm described in [8]. The AIT-GP method can be seen as Algorithm 2 with $\mu = 0$, while the ISTA algorithm is equivalent to iterations of Algorithm 2 without the preconditioner. These comparisons allow us to show how the quality of the reconstructed solution is improved by the presence of both the soft-thresholding and the preconditioner.

The ISTA method and our proposals require the selection of a regularization parameter. For all these methods we select the parameter that minimizes the relative restoration error defined by

$$\mathrm{RRE}(\mathbf{f}) = \frac{\|\mathbf{f} - \mathbf{f}_{\mathrm{true}}\|}{\|\mathbf{f}_{\mathrm{true}}\|}.$$

For the comparison of the algorithms we consider the Peak Signal to Noise Ratio (PSNR) defined by

$$\mathrm{PSNR}(\mathbf{f}) = 20\log_{10}\left(\frac{mM}{\|\mathbf{f} - \mathbf{f}_{\mathrm{true}\|}}\right),$$

where m^2 is the the number of elements of $\mathbf{f}$ and M denotes the maximum value of $\mathbf{f}_{\mathrm{true}}$. Moreover, we consider the Structure SIMilarity index (SSIM), the definition of the SSIM is involved, here we recall that this index measures how accurately the computed approximation is able to reconstruct the overall structure of the image. The higher the value of the SSIM the better the reconstruction is, and the maximum value achievable is 1; see [27] for a precise definition of the SSIM.

We now describe how we construct the operator W. We use the tight frames determined by linear B-splines; see, e.g., [28]. For one-dimensional problems they are composed by a low-pass filter $W_0 \in \mathbb{R}^{m \times m}$ and two high-pass filters $W_1 \in \mathbb{R}^{m \times m}$ and $W_2 \in \mathbb{R}^{m \times m}$. These filters are determined by the masks are given by

$$u^{(0)} = \frac{1}{4}[1, 2, 1], \quad u^{(1)} = \frac{\sqrt{2}}{4}[1, 0, -1], \quad u^{(2)} = \frac{1}{4}[-1, 2, -1].$$

Imposing reflexive boundary conditions we determine the analysis operator W so that $W^*W = I$. Define the matrices

$$W_0 = \frac{1}{4}\begin{pmatrix} 3 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 2 & 1 \\ 0 & \dots & 0 & 1 & 3 \end{pmatrix}, \quad W_1 = \frac{\sqrt{2}}{4}\begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix},$$

and

$$W_2 = \frac{1}{4} \begin{pmatrix} 1 & -1 & 0 & \cdots & & 0 \\ -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ 0 & \cdots & 0 & & -1 & 1 \end{pmatrix}.$$

Then the operator W is defined by

$$W = \begin{pmatrix} W_0 \\ W_1 \\ W_2 \end{pmatrix}.$$

To construct the two-dimensional framelet analysis operator we use the tensor products

$$W_{i,j} = W_i \otimes W_j, \quad i, j = 0, 1, 2.$$

The matrix W_{00} is a low-pass filter; all the other matrices W_{ij} contain at least one high-pass filter. The analysis operator is given by

$$W = \begin{bmatrix} W_{00} \\ W_{01} \\ \vdots \\ W_{22} \end{bmatrix}.$$

In PISTA$_h$, following [26], we set

$$h(x) = \left(1 - \frac{x}{\|A\|^2}\right)^4 + 10^{-15}.$$

All the computations are performed on MATLAB R2018b running on a laptop with an Intel i7-8750H @2.20 GHz CPU and 16GB of RAM.

Cameraman

We first consider the cameraman image in Figure 3(a) and we blur it with the non-symmetric PSF in Figure 3(b). We then add 2% white Gaussian noise obtaining the blurred and noisy image in Figure 3(c). Note that we crop the boundaries of the image to simulate real data; see [1] for more details. Since the image is generic we impose reflexive BCs.

In Table 1 we report the results obtained with the different methods. We can observe that Struct-PISTA$_h$ provides the best reconstruction of all considered algorithms. Moreover, we can observe that, in general, the introduction of the structured preconditioner improves the quality of the reconstructed solutions, especially in term of SSIM. From the visual inspection of the reconstructions in Figure 4 we can observe that the introduction of the structured preconditioner allows us to evidently reduce the boundary artifacts as well as avoid the amplification of the noise.

Grain

We now consider the grain image in Figure 5(a) and blur it with the PSF, obtained by the superposition of two motions PSF, in Figure 5(b). After adding 3% of white Gaussian noise and cropping the boundaries we obtain the blurred and noisy image in Figure 5(c). According to the nature of the image we use reflexive bc's.

Again in Table 1 we report all the results obtained with the considered methods. In this case ISTA provides the best reconstruction in terms of RRE and PSNR. However, Struct-PISTA$_h$ provides the best reconstruction terms of SSIM and very similar results in term of PSNR and RRE. In Figure 6 we report some of the reconstructed solution. From the visual inspection of these reconstruction
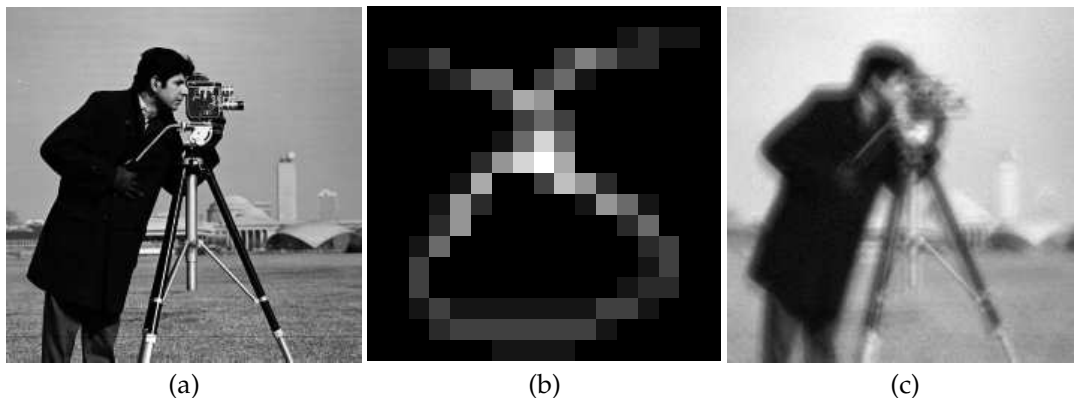
Figure 3. Cameraman test problem: (a) True image (238 × 238 pixels), (b) PSF (17 × 17 pixels), (c) Blurred and noisy image with 2% of white Gaussian noise (238 × 238 pixels).
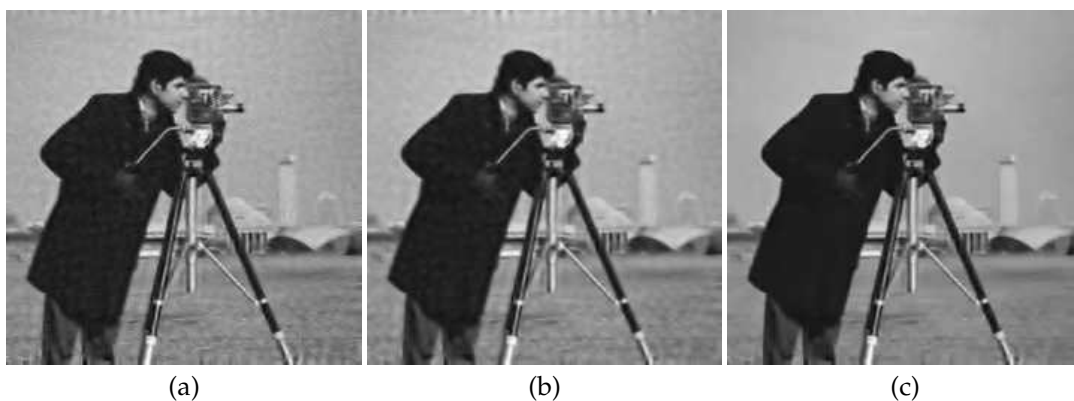

Figure 4. Cameraman test problem reconstructions: (a) ISTA, (b) PISTA$_h$, (c) Struct-PISTA$_h$.

we can see that the introduction of the structured preconditioner reduces the ringing and boundary effects in the computed solutions.

Satellite

Our final example is the `atmosphericBlur30` from the MATLAB toolbox RestoreTools [2]. The true image, PSF, and blurred and noisy image are reported in Figures 7(a), (b), and (c), respectively. Since we know the true image we can estimate the noise level in the image, which is approximately 1%. Since this is an astronomical image we impose zero bc's.

From the comparison of the computed results in Table 1 we can see that the Struct-PISTA$_h$ method provides the best reconstruction among all considered methods. We can observe that, in this particular example, ISTA provides a very low quality reconstruction both in term of RRE and SSIM. We report in Figure 8 some reconstructions. From the visual inspection of the computed solutions we can observe that both the approximations obtained with PISTA$_h$ and Struct-PISTA$_h$ do not present heavy ringing effects, while the reconstruction obtained by AIT-GP presents very heavy ringing around the "arms" of the satellite. This allows us to show the benefits of introducing the soft-thresholding into the AIT-GP method.

5. Conclusions

This work develops further and brings together all the techniques studied in [16,17,21,22,29]. The idea is to combine thresholding iterative methods, an approximate Tikhonov regularization scheme depending on a general (differential) operator and a structure preserving approach, with the main goal in mind to reduce the boundary artifacts which appear in the resulting de-blurred image when imposing artificial boundary conditions. The numerical results are promising and show
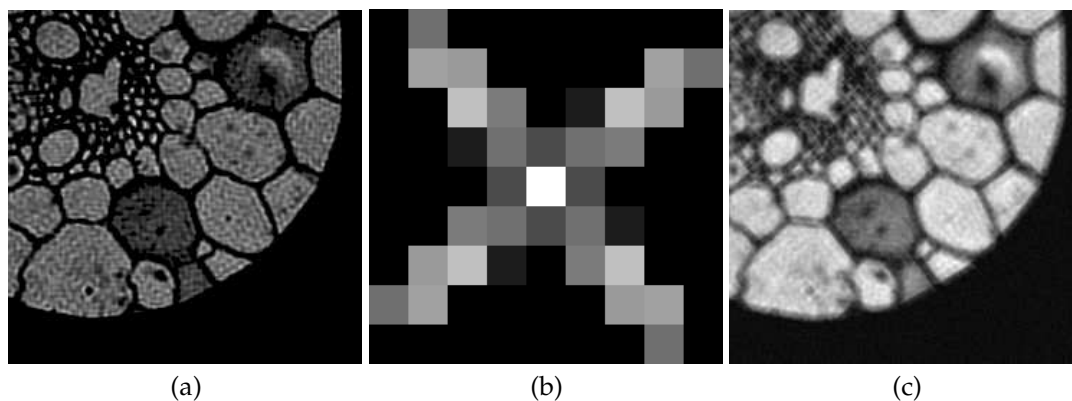
Figure 5. Grain test problem: (a) True image (246 × 246 pixels), (b) PSF (9 × 9 pixels), (c) Blurred and noisy image with 3% of white Gaussian noise (246 × 246 pixels).
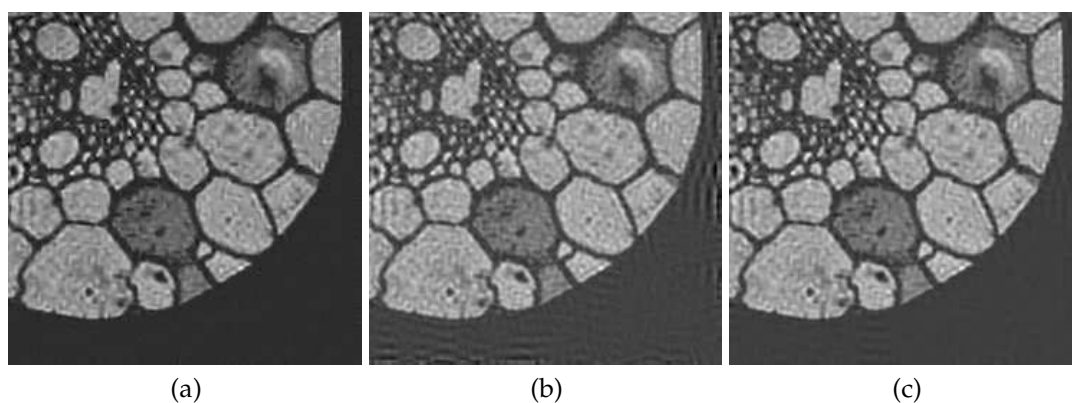

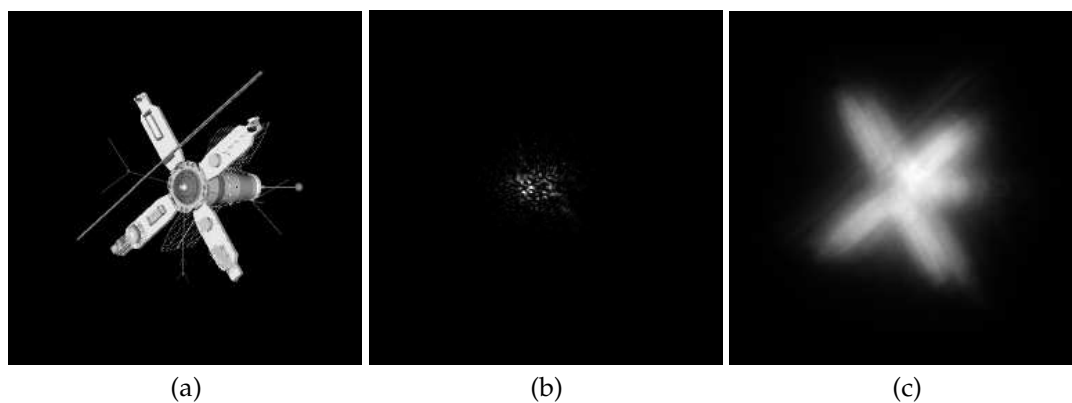**Figure 6.** Grain test problem reconstructions: (a) ISTA, (b) PISTA$_\Lambda$, (c) Struct-PISTA$_\Lambda$.

Figure 7. Satellite test problem: (a) True image (256 × 256 pixels), (b) PSF (256 × 256 pixels), (c) Blurred and noisy image with ≈ 1% of white Gaussian noise (256 × 256 pixels).
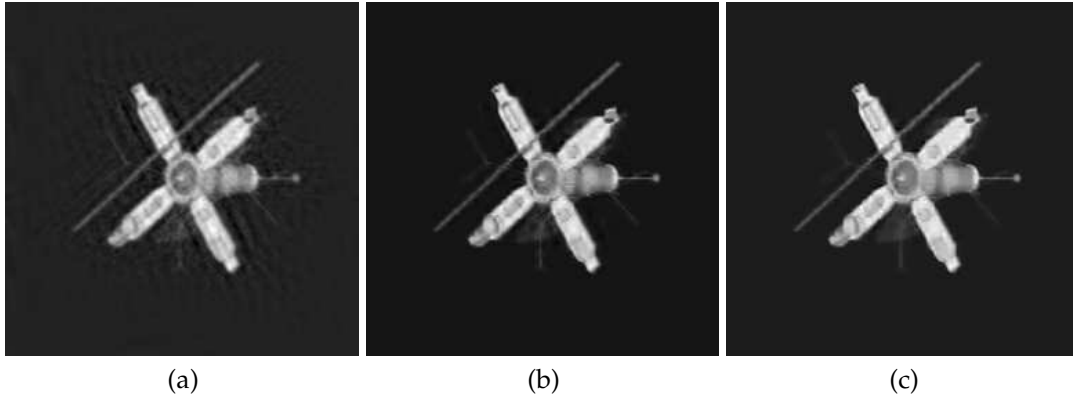
(a)   (b)   (c)

Figure 8. Satellite test problem reconstructions: (a) AIT-GP, (b) PISTA$_h$, (c) Struct-PISTA$_h$.

Table 1. Comparison of the quality of the reconstructions for all considered examples. We highlight in boldface the best result.

Example	Method	RRE	PSNR	SSIM
Cameraman	AIT-GP	0.111024	24.7798	0.729095
	ISTA	0.090921	26.5149	0.763217
	PISTA$_h$	0.096558	25.9924	0.790363
	PISTA$_\Lambda$	0.094853	26.1471	0.795061
	Struct-PISTA$_h$	**0.088796**	**26.7203**	**0.840145**
	Struct-PISTA$_\Lambda$	0.090182	26.5857	0.834532
Grain	AIT-GP	0.183796	25.9571	0.731407
	ISTA	**0.160655**	**27.1259**	0.845816
	PISTA$_h$	0.195516	25.4202	0.737254
	PISTA$_\Lambda$	0.181727	26.0554	0.748582
	Struct-PISTA$_h$	0.161715	27.0688	**0.859284**
	Struct-PISTA$_\Lambda$	0.168472	26.7133	0.830990
Satellite	AIT-GP	0.222783	26.6708	0.742416
	ISTA	0.286179	24.4956	0.657111
	PISTA$_h$	0.192146	27.9558	0.928584
	PISTA$_\Lambda$	0.193730	27.8844	0.916993
	Struct-PISTA$_h$	**0.187970**	**28.1466**	**0.934876**
	Struct-PISTA$_\Lambda$	0.189147	28.0924	0.924931

improvements with respect to known state-of-the-art deblurring algorithms. There are still open problems, mainly concerning the theoretical assumptions and convergence proofs which will be furtherly investigated in future works.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Proofs

Hereafter we analyze Algorithm 1, aiming to prove its convergence. The techniques carried out in the proofs of most of the following results can be tracked down to the papers [15–17], therefore bringing not much mathematical novelties other than the results themselves. Nevertheless, since the proofs are very technical such that even a slight change can produce non trivial difficulties, we will present a full treatment leaving no details, in order to make this paper self-contained and easily readable.

Following up Section 2.1, we need to set some more notations. Let us consider the singular value decomposition (SVD) of C as the triple (U, V, Σ) such that

$$C = U\Sigma V^*,$$

$$U, V \in \mathcal{O}(m^2, \mathbb{R}), \qquad \Sigma = \text{diag}_{j=1,\cdots,m^2}(\sigma_j) \text{ with } 0 \leq \sigma_{m^2} \leq \cdots \leq \sigma_1,$$

where $\mathcal{O}(m^2, \mathbb{R})$ is the orthonormal group and V^* is the adjoint of the operator V, i.e., $\langle V\mathbf{f}^1, \mathbf{f}^2 \rangle = \langle \mathbf{f}^1, V^*\mathbf{f}^2 \rangle$ for every pair $\mathbf{f}^1, \mathbf{f}^2 \in \mathbb{R}^{m^2}$. We will denote the spectrum of CC^* by

$$\sigma(CC^*) = \{0\} \cup \bigcup_{j=1}^{m^2} \{\sigma_j^2\}.$$

Hereafter, without loss of generality we will assume that

$$\|C\| = 1 \qquad \text{and} \qquad \|h(CC^*)\| = \max_{\sigma^2 \in [0,1]} h(\sigma^2) = 1.$$

The first issue we have to consider is the existence of the sequence $\{\alpha_n\}$.

Lemma A1. *Let $\|\mathbf{r}^n\| > \tau\delta$. Then for every fixed n there exists α_n that satisfies (7). It can be computed by the following iteration*

$$\alpha_n^{k+1} := \frac{\left(\alpha_n^k\right)^2 \Phi'(\alpha_n^k)}{\alpha_n^k \Phi'(\alpha_n^k) + \Phi(\alpha_n^k) - q_n^2 \|\mathbf{r}^n\|}, \tag{A1}$$

where

$$\Phi(\alpha) := \|\alpha(CC^* + \alpha h(CC^*))^{-1}\mathbf{r}^n\|^2,$$

$$\Phi'(\alpha) := \|\sqrt{2\alpha}CC^*(CC^* + \alpha h(CC^*))^{-3/2}\mathbf{r}^n\|^2.$$

The convergence is locally quadratic. The existence of the regularization parameter α_n and the locally quadratic convergence of the algorithm above are independent and uniform with respect to the dimension m^2.

Proof. The existence of α is an easy consequence of the monotonicity of

$$\phi_\alpha(\sigma^2) = \alpha \left(\sigma^2 + \alpha h(\sigma^2)\right)^{-1}$$

with respect to α. Indeed, let us rewrite (7) as follows

$$
\begin{aligned}
\alpha^2 \| (CC^* + \alpha h\,(CC^*))^{-1} \mathbf{r}^n \|^2 &= \| \phi_\alpha (CC^*) \mathbf{r}^n \|^2 \\
&= \int_{[0,1]} \phi_\alpha^2(\sigma^2) d\|\mathbf{r}^n\|(\sigma) \\
&= \int_{[0,1]} \frac{\alpha^2}{(\sigma^2 + \alpha h(\sigma^2))^2} d\|\mathbf{r}^n\|(\sigma) \\
&= \frac{q_n^2}{c_1^2} \int_{[0,1]} d\|\mathbf{r}^n\|(\sigma),
\end{aligned}
\tag{A2}
$$

where $d\|\mathbf{r}^n\|(\cdot)$ is the discrete spectral measure associated to $\mathbf{r}^n$ with respect to the SVD of C and $\sigma \in \sigma(C)$ are the singular values of the spectrum of C. Since $\frac{d\phi_\alpha}{d\alpha} > 0$ for every $\alpha \geq 0$, then by monotone convergence it holds that

$$
\begin{aligned}
\lim_{\alpha \to \infty} \int_{[0,1]} \frac{\alpha^2}{(\sigma^2 + \alpha h(\sigma^2))^2} d\|\mathbf{r}^n\|(\sigma) &= \int_{[0,1]} \lim_{\alpha \to \infty} \frac{\alpha^2}{(\sigma^2 + \alpha h(\sigma^2))^2} d\|\mathbf{r}^n\|(\sigma) \\
&\geq \|\mathbf{r}^n\|^2 > \frac{q_n^2}{c_1^2} \|\mathbf{r}^n\|^2.
\end{aligned}
$$

Indeed, it is not difficult to prove that $q_n/c_1 < 1$ whenever $\rho \in (0, c_1/2)$ and $\|\mathbf{r}^n\| > \tau\delta$, as assumed in the hypothesis. Since for $\alpha = 0$ the left-hand-side of (A2) is zero, then we conclude that there exists an unique $\alpha_n > 0$ such that equality holds in (7). Due to the generality of our proof and the fact that we could pass the limit under the sign of integral, the existence of such an α_n is granted uniformly with respect to the dimension m^2.

Since

$$
\phi_\alpha(\sigma^2) = \alpha \left(\sigma^2 + \alpha h(\sigma^2) \right)^{-1} = \left(\alpha^{-1}\sigma^2 + h(\sigma^2) \right)^{-1},
$$

fixing $\gamma = \alpha^{-1}$, let us now define the following function

$$
\psi_\gamma(\sigma^2) = \left(\gamma\sigma^2 + h(\sigma^2) \right)^{-1}.
$$

Since

$$
\frac{\partial \psi_\gamma^2(\sigma^2)}{\partial \gamma} = -2\sigma^2 \left(\gamma\sigma^2 + h(\sigma^2) \right)^{-3}, \qquad \frac{\partial^2 \psi_\gamma^2(\sigma^2)}{\partial \gamma^2} = 6\sigma^4 \left(\gamma\sigma^2 + h(\sigma^2) \right)^{-4},
\tag{A3}
$$

then there exists two constants d_1, d_2 independents of γ such that

$$
\left| \frac{\partial \psi_\gamma^2(\sigma^2)}{\partial \gamma} \right| \leq d_1, \qquad \left| \frac{\partial^2 \psi_\gamma^2(\sigma^2)}{\partial \gamma^2} \right| \leq d_2,
$$

and in particular $d_1, d_2 \in L^1([0,1], d\|\mathbf{r}^n\|)$ for every n and m. Therefore, if we define

$$
\Psi(\gamma) := \|\psi_\gamma(CC^*)\mathbf{r}^n\|^2,
$$

it holds that

$$
\Psi'(\gamma) = \frac{\partial}{\partial \gamma} \int_{[0,1]} \psi_\gamma^2(\sigma^2) d\|\mathbf{r}^n\|(\sigma) = \int_{[0,1]} \frac{\partial \psi_\gamma^2(\sigma^2)}{\partial \gamma} d\|\mathbf{r}^n\|(\sigma),
\tag{A4}
$$

$$
\Psi''(\gamma) = \frac{\partial}{\partial \gamma} \int_{[0,1]} \frac{\partial \psi_\gamma^2(\sigma^2)}{\partial \gamma} d\|\mathbf{r}^n\|(\sigma) = \int_{[0,1]} \frac{\partial^2 \psi_\gamma^2(\sigma^2)}{\partial \gamma^2} d\|\mathbf{r}^n\|(\sigma).
\tag{A5}
$$

Then the Newton iteration applied to $\Psi(\gamma) = q_n^2 \|\mathbf{r}^n\|^2$ yields the iteration

$$\gamma_n^{k+1} = \gamma_n^k + \frac{q_n^2 \|\mathbf{r}^n\|^2 - \Psi(\gamma_n^k)}{\Psi'(\gamma_n^k)}, \qquad k \geq 0.$$

By (A3), $\Psi(\gamma)$ is a decreasing convex function in γ. Since $\gamma_n = \lim_{k \to \infty} \gamma_n^{k+1} = 1/\alpha_n$, obviously we have that

$$\Psi(\gamma_n) = \frac{q_n^2}{c_1^2} \|\mathbf{r}^n\|^2. \tag{A6}$$

If

$$\Psi'(\gamma_n) = -\|\sqrt{2}(CC^*)(\gamma_n CC^* + h(CC^*))^{-3/2} \mathbf{r}^n\|^2 = 0,$$

then necessarily we would have that $CC^* \mathbf{r}^n = 0$. Hence, $(\gamma_n CC^* + h(CC^*))^{-1} \mathbf{r}^n = h(CC^*) \mathbf{r}^n$, and consequently

$$\Psi(\gamma_n) = \|h(CC^*) \mathbf{r}^n\|^2.$$

From (A6) we would deduce that $q_n \geq c_1$, but this is absurd since as already observed above, $q_n < c_1$ if $\|\mathbf{r}^n\| > \tau\delta$. Therefore $\Psi'(\gamma_n) \neq 0$ and by standard properties of the Newton iteration, γ_n^k converges to the minimizer γ_n from below and the convergence is locally quadratic. Finally, defining

$$\Phi(\alpha) = \Psi(1/\alpha),$$

then we get (A1), α_n^k converges monotonically from above to α_n and the convergence is locally quadratic. Again, thanks to (A4) and (A5), the rate of convergence is uniform with respect to the dimension of $\mathcal{Y}$. $\square$

From now on, instead of working with the error $\mathbf{e}_\delta^n = \mathbf{x} - \mathbf{x}_\delta^n$, in order to simplify the following proofs and notations, it is useful to consider the partial error with respect to $\mathbf{z}_\delta^n$, namely

$$\tilde{\mathbf{e}}_\delta^n = \mathbf{x} - \mathbf{z}_\delta^n. \tag{A7}$$

This will not affect the generality of our proofs, thanks to the continuity of $\S_\mu(\cdot)$ with respect to the noise level δ.

Proposition A1. *Under the assumptions (9), if $\|\mathbf{r}_\delta^n\| > \tau\delta$ and we define $\tau_n = \|\mathbf{r}_\delta^n\|/\delta$, then it follows that*

$$\|\mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n\| \leq \left(\rho + \frac{1 + 2\rho}{\tau_n} \right) \|\mathbf{r}_\delta^n\| < (1 - \rho)\|\mathbf{r}_\delta^n\|, \tag{A8}$$

where $\tilde{\mathbf{e}}^n$ is defined in (A7).

Proof. In the free noise case we have $\mathbf{g} = K\mathbf{x}$. As a consequence

$$\mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n = \mathbf{g}^\delta - K\mathbf{x}_\delta^n - B(\mathbf{x} - \mathbf{z}_\delta^n) + B\mathbf{x}_\delta^n - BS_\mu(\mathbf{z}_\delta^n)$$
$$= \mathbf{g}^\delta - \mathbf{g} + (K - B)\mathbf{e}_\delta^n + B(\mathbf{z}_\delta^n - S_\mu(\mathbf{z}_\delta^n)).$$

Using now assumptions (9), in particular (10), and $\|\mathbf{g}^\delta - \mathbf{g}\| \leq \delta$, we derive the following estimate

$$\begin{aligned}
\|\mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n\| &\leq \|\mathbf{g}^\delta - \mathbf{g}\| + \|(K - B)\mathbf{e}_\delta^n\| + \|B(\mathbf{z}_\delta^n - S_\mu(\mathbf{z}_\delta^n))\| \\
&\leq \|\mathbf{g}^\delta - \mathbf{g}\| + \rho\|K\mathbf{e}_\delta^n\| + \rho\delta \\
&\leq \|\mathbf{g}^\delta - \mathbf{g}\| + \rho(\|\mathbf{r}_\delta^n\| + \|\mathbf{g}^\delta - \mathbf{g}\| + \delta) \\
&\leq (1 + 2\rho)\delta + \rho\|\mathbf{r}_\delta^n\|.
\end{aligned}$$

The first inequality in (A8) now follows from the hypothesis $\delta = \|\mathbf{r}_\delta^n\|/\tau_n$. The second inequality follows from $\rho + \frac{1+2\rho}{\tau_n} < \rho + \frac{1+2\rho}{\tau}$. $\quad\square$

Combining the preceding proposition with (7), we are going to show that the sequence $\|\tilde{\mathbf{e}}_\delta^n\|$ is monotonically decreasing. We have the following result.

Proposition A2. *Let $\tilde{\mathbf{e}}_\delta^n$ be defined in* (A7). *If the assumptions* (9) *are satisfied, then $\|\tilde{\mathbf{e}}_\delta^n\|$ of Algorithm 1 decreases monotonically for $n = 0, 1, \ldots, n_\delta - 1$. In particular, we deduce*

$$\|\tilde{\mathbf{e}}_\delta^n\|^2 - \|\tilde{\mathbf{e}}_\delta^{n+1}\|^2 \geq \frac{8\rho^2}{1+2\rho}\|(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n\|\|\mathbf{r}_\delta^n\| > 0. \tag{A9}$$

Proof. Recalling that $WC^* = B^*$ and that $BB^* = CC^*$, we have

$$
\begin{aligned}
\|\tilde{\mathbf{e}}_\delta^n\|^2 - \|\tilde{\mathbf{e}}_\delta^{n+1}\|^2 &= 2\langle \tilde{\mathbf{e}}_\delta^n, \mathbf{h}^n \rangle - \|\mathbf{h}^n\|^2 \\
&= 2\langle B\tilde{\mathbf{e}}_\delta^n, (CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n \rangle - \langle \mathbf{r}_\delta^n, CC^*(CC^* + \alpha_n h\,(CC^*))^{-2}\mathbf{r}_\delta^n \rangle \\
&= 2\langle \mathbf{r}_\delta^n, (CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n \rangle - \langle \mathbf{r}_\delta^n, CC^*(CC^* + \alpha_n h\,(CC^*))^{-2}\mathbf{r}_\delta^n \rangle \\
&\quad - 2\langle \mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n, (CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n \rangle \\
&\geq 2\langle \mathbf{r}_\delta^n, (CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n \rangle - 2\langle \mathbf{r}_\delta^n, CC^*(CC^* + \alpha_n h\,(CC^*))^{-2}\mathbf{r}_\delta^n \rangle \\
&\quad - 2\langle \mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n, (CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n \rangle \\
&= 2\alpha_n\langle \mathbf{r}_\delta^n, h\,(CC^*)\,(CC^* + \alpha_n h\,(CC^*))^{-2}\mathbf{r}_\delta^n \rangle \\
&\quad - 2\langle \mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n, (CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n \rangle \\
&\geq 2\alpha_n\langle \mathbf{r}_\delta^n, h\,(CC^*)\,(CC^* + \alpha_n h\,(CC^*))^{-2}\mathbf{r}_\delta^n \rangle \\
&\quad - 2\|\mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n\|\|(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n\| \\
&\geq 2\|(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n\|\left(\|c_1\alpha_n(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n\|\right. \\
&\quad \left. - \|\mathbf{r}_\delta^n - B\tilde{\mathbf{e}}_\delta^n\|\right) \\
&\geq 2\|(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n\| \cdot \left(q_n\|\mathbf{r}_\delta^n\| - \left(\rho + \frac{1+2\rho}{\tau_n}\right)\|\mathbf{r}_\delta^n\|\right) \\
&\geq \frac{8\rho^2}{1+2\rho}\|(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n\|\|\mathbf{r}_\delta^n\| > 0,
\end{aligned}
$$

where the relevant inequalities are a consequence of equation (7) and Proposition A1. The last inequality follows from (7) and $\tau_n > \tau = (1+2\rho)/(1-2\rho)$ for $\|\mathbf{r}_\delta^n\| > \tau\delta$. $\quad\square$

Corollary A1. *Under the assumptions* (9)*, there holds*

$$\|\tilde{\mathbf{e}}_\delta^0\| \geq \frac{8\rho^2}{1+2\rho}\sum_{n=0}^{n_\delta-1}\|(CC^* + \alpha_n h\,(CC^*))^{-1}\mathbf{r}_\delta^n\|\|\mathbf{r}_\delta^n\| \geq c\sum_{n=0}^{n_\delta-1}\|\mathbf{r}_\delta^n\|^2 \tag{A10}$$

for some constant $c > 0$, depending only on ρ and q in (7).

Proof. The first inequality follows by taking the sum of the quantities in (A9) from $n = 0$ up to $n = n_\delta - 1$.

For the second inequality, note that for every

$$\alpha > \frac{q_n}{c_1 - q_n}$$

and every $\sigma \in \sigma(C) \subset [0,1]$, we have

$$\frac{\alpha}{\sigma^2 + \alpha h(\sigma^2)} \geq \frac{\alpha}{1+\alpha} = (1+1/\alpha)^{-1} > \frac{q_n}{c_1},$$

and hence,

$$\alpha \|(CC^* + \alpha h(CC^*))^{-1} \mathbf{r}_\delta^n\| > \frac{q_n}{c_1} \|\mathbf{r}_\delta^n\|,$$

as $\|\mathbf{r}_\delta^n\| > 0$ for $n < n_\delta$. This implies that α_n in (7) satisfies $0 < \alpha_n \leq \frac{q_n}{c_1 - q_n}$, thus

$$\|(CC^* + \alpha_n h(CC^*))^{-1} \mathbf{r}_\delta^n\| = \frac{q_n}{c_1 \alpha_n} \|\mathbf{r}_\delta^n\| \geq (c_1 - q_n) \|\mathbf{r}_\delta^n\|.$$

According to the choice of parameters in Algorithm 1, we deduce

$$c_1 - q_n = \min\{c_1 - q, c_1 - 2\rho - (1+\rho)/\tau_n\},$$

and

$$c_1 - 2\rho - (1+\rho)/\tau_n = \frac{1+2\rho}{\tau} - \frac{1+\rho}{\tau_n} > \frac{1+2\rho}{\tau} - \frac{1+\rho}{\tau} = \frac{\rho}{\tau}.$$

Therefore, there exists $c > 0$, depending only on ρ and q such that

$$c_1 - q_n \geq c \left(\frac{8\rho^2}{1+2\rho} \right)^{-1},$$

and

$$\|(CC^* + \alpha_n h(CC^*))^{-1} \mathbf{r}_\delta^n\| \geq c \left(\frac{8\rho^2}{1+2\rho} \right)^{-1} \|\mathbf{r}_\delta^n\| \qquad \text{for } n = 0, 1, \cdots, n_\delta - 1.$$

Now the second inequality follows immediately. $\square$

From (A10) it can be seen that the sum of the squares of the residual norms is bounded, and hence, if $\delta > 0$, there must be a first integer $n_\delta < \infty$ such that (A10) is fulfilled, i.e., Algorithm 1 terminates after finitely many iterations.

Finally, we are ready to prove a convergence and regularity result.

Theorem A1. *Assume that $\mathbf{z}^0$ is not a solution of the linear system*

$$\mathbf{g} = AW^* \mathbf{x}, \tag{A11}$$

and that δ_m is a sequence of positive real numbers such that $\delta_k \to 0$ as $k \to \infty$. Then, if Assumption 1 is valid, the sequence $\{\mathbf{x}_{\delta_k}^{n(\delta_k)}\}_{k \in \mathbb{N}}$, generated by the discrepancy principle rule (A10), converges as $k \to \infty$ to the solution of (A11) which is closest to $\mathbf{z}^0$ in Euclidean norm.

Proof. We are going to show convergence for the sequence $\{\mathbf{z}_{\delta_k}^{n(\delta_k)}\}_{k \in \mathbb{N}}$ and then the thesis will follow easily from the continuity of $S_{\mu(\delta)}$, i.e.,

$$\lim_{k \to \infty} \mathbf{x}_{\delta_k}^{n(\delta_k)} = \lim_{k \to \infty} S_{\mu(\delta_k)}(\mathbf{z}_{\delta_k}^{n(\delta_k)}) = S_{\lim_{k \to \infty} \mu(\delta_k)}(\lim_{k \to \infty} \mathbf{z}_{\delta_k}^{n(\delta_k)}) = \lim_{k \to \infty} \mathbf{z}_{\delta_k}^{n(\delta_k)}.$$

The proof of the convergence for the sequence $\{\mathbf{z}_{\delta_k}^{n(\delta_k)}\}$ can be divided into two steps: at step one, we show the convergence in the free noise case $\delta = 0$. In particular, the sequence $\{\mathbf{z}^n\}$ converges to a solution of (A11) that is the closest to $\mathbf{z}^0$. At the second step, we show that given a sequence of

positive real numbers $\delta_k \to 0$ as $k \to \infty$, then we get a corresponding sequence $\{\mathbf{z}_{\delta_k}^{n(\delta_k)}\}$ converging as $k \to \infty$.

Step 1: Fix $\delta = 0$. It follows that $\mathbf{r}_\delta^n = \mathbf{r}^n$, and the sequence $\{\mathbf{z}^n\}$ will not stop, i.e., $n \to \infty$, since the discrepancy principle will not be satisfied by any n, in particular $n_\delta \to \infty$ for $\delta \to 0$. Set $n > l > j$, with $n, l, j \in \mathbb{N}$. It holds that

$$
\begin{aligned}
\|\mathbf{z}^n - \mathbf{z}^l\|^2 &= \|\tilde{\mathbf{e}}^n - \tilde{\mathbf{e}}^l\|^2 \\
&= \|\tilde{\mathbf{e}}^n\|^2 - \|\tilde{\mathbf{e}}^l\|^2 - 2\langle \tilde{\mathbf{e}}^l, \tilde{\mathbf{e}}^n - \tilde{\mathbf{e}}^l \rangle \\
&= \|\tilde{\mathbf{e}}^n\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2\langle \tilde{\mathbf{e}}^l, \mathbf{z}^n - \mathbf{z}^l \rangle \\
&= \|\tilde{\mathbf{e}}^n\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2\sum_{i=l}^{n-1}\langle \tilde{\mathbf{e}}^l, \mathbf{h}^i \rangle \\
&= \|\tilde{\mathbf{e}}^n\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2\sum_{i=l}^{n-1}\langle B\tilde{\mathbf{e}}^l, (CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i \rangle \\
&\leq \|\tilde{\mathbf{e}}^n\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2\sum_{i=l}^{n-1}\|B\tilde{\mathbf{e}}^l\|\|(CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i\| \\
&\leq \|\tilde{\mathbf{e}}^n\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2(1+\rho)\sum_{i=l}^{n-1}\|\mathbf{r}^l\|\|(CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i\|,
\end{aligned}
\tag{A12}
$$

where the last inequality comes from (11). At the same time, we have that

$$
\begin{aligned}
\|(\mathbf{z}^l - \mathbf{z}^k)\|^2 &= \|(\tilde{\mathbf{e}}^l - \tilde{\mathbf{e}}^k)\|^2 \\
&= \|\tilde{\mathbf{e}}^k\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2\langle \tilde{\mathbf{e}}^l, \tilde{\mathbf{e}}^k - \tilde{\mathbf{e}}^l \rangle \\
&= \|\tilde{\mathbf{e}}^k\|^2 - \|\tilde{\mathbf{e}}^l\|^2 - 2\langle \tilde{\mathbf{e}}^l, \mathbf{z}^l - \mathbf{z}^k \rangle \\
&= \|\tilde{\mathbf{e}}^k\|^2 - \|\tilde{\mathbf{e}}^l\|^2 - 2\sum_{i=k}^{l-1}\langle \tilde{\mathbf{e}}^l, \mathbf{h}^i \rangle \\
&\leq \|\tilde{\mathbf{e}}^k\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2\sum_{i=k}^{l-1}\|B\tilde{\mathbf{e}}^l\|\|(CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i\| \\
&\leq \|\tilde{\mathbf{e}}^k\|^2 - \|\tilde{\mathbf{e}}^l\|^2 + 2(1+\rho)\sum_{i=k}^{l-1}\|\mathbf{r}^l\|\|(CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i\|.
\end{aligned}
\tag{A13}
$$

Combining together (A12) and (A13), we obtain that

$$
\begin{aligned}
\|\mathbf{z}^n - \mathbf{z}^k\|^2 &\leq 2\|\mathbf{z}^n - \mathbf{z}^l\|^2 + 2\|\mathbf{z}^l - \mathbf{z}^k\|^2 \\
&\leq 2\|\tilde{\mathbf{e}}^n\|^2 + 2\|\tilde{\mathbf{e}}^k\|^2 - 4\|\tilde{\mathbf{e}}^l\|^2 + 4(1+\rho)\sum_{i=k}^{n-1}\|\mathbf{r}^l\|\|(CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i\|.
\end{aligned}
$$

This is valid for every $l \in \{k+1, \cdots, n-1\}$. Choosing l such that $\|\mathbf{r}^l\| = \min_{i=k+1,\cdots,n-1}\|\mathbf{r}^i\|$, it follows that

$$
\|\mathbf{z}^n - \mathbf{z}^k\|^2 \leq 2\|\tilde{\mathbf{e}}^n\|^2 + 2\|\tilde{\mathbf{e}}^k\|^2 - 4\|\tilde{\mathbf{e}}^l\|^2 + 4(1+\rho)\sum_{i=k}^{n-1}\|\mathbf{r}^i\|\|(CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i\|.
$$

From Proposition A2, $\{\|\tilde{\mathbf{e}}^j\|^2\}_{j \in \mathbb{N}}$ is a converging sequence, and from Corollary A1

$$
\sum_{i=k}^{n-1}\|\mathbf{r}^l\|\|(CC^* + \alpha_i h\,(CC^*))^{-1}\mathbf{r}^i\| \to 0 \qquad \text{as } k, n \to \infty
$$

since it is the tail of a converging series. Therefore,

$$\|\mathbf{z}^n - \mathbf{z}^k\|^2 \to 0 \qquad \text{as } k, n \to \infty$$

and $\{\mathbf{z}^n\}_{n \in \mathbb{N}}$ is a Cauchy sequence, and then convergent.

Step 2: Let $\mathbf{x}$ be the converging point of the sequence $\{\mathbf{z}^n\}_{n \in \mathbb{N}}$ and let $\delta_k > 0$ be a sequence of positive real numbers converging to 0. For every δ_k, let $n = n(\delta_k)$ be the first positive integer such that (A10) is satisfied, whose existence is granted by Corollary A1, and let $\{\mathbf{z}_{\delta_k}^{n(\delta_k)}\}$ be the corresponding sequence. For every fixed $\epsilon > 0$, there exists $\overline{n} = \overline{n}(\epsilon)$ such that

$$\|\mathbf{x} - \mathbf{z}^n\| \le \epsilon/2 \qquad \text{for every } n > \overline{n}(\epsilon), \tag{A14}$$

and there exists $\overline{\delta} = \overline{\delta}(\epsilon)$ for which

$$\|\mathbf{z}^{\overline{n}} - \mathbf{z}_\delta^{\overline{n}}\| \le \epsilon/2 \qquad \text{for every } 0 < \delta < \overline{\delta}, \tag{A15}$$

due to the continuity of the operator $\mathbf{g} \mapsto \mathbf{z}^n$ for every fixed n. Therefore, let us choose $\overline{k} = \overline{k}(\epsilon)$ large enough such that $\delta_k < \overline{\delta}$ and such that $n(\delta_k) > \overline{n}$ for every $k > \overline{k}$. Such $\overline{k}$ does exists since $\delta_k \to 0$ and $n_\delta \to \infty$ for $\delta \to 0$. Hence, for every $k > \overline{k}$, we have

$$\begin{aligned} \|\mathbf{x} - \mathbf{z}_{\delta_k}^{n(\delta_k)}\| &= \|\tilde{\mathbf{e}}_{\delta_k}^{n(\delta_k)}\| \\ &\le \|\tilde{\mathbf{e}}_{\delta_k}^{\overline{n}}\| \\ &= \|\mathbf{x} - \mathbf{z}_{\delta_k}^{\overline{n}}\| \\ &\le \|\mathbf{x} - \mathbf{z}^{\overline{n}}\| + \|\mathbf{z}^{\overline{n}} - \mathbf{z}_{\delta_k}^{\overline{n}}\| \le \epsilon, \end{aligned}$$

where the first inequality comes from Proposition A2 and the last one from (A14) and (A15). $\square$

References

1. Hansen, P.C.; Nagy, J.G.; O'leary, D.P. *Deblurring images: matrices, spectra, and filtering*; Vol. 3, Siam, 2006.
2. Nagy, J.G.; Palmer, K.; Perrone, L. Iterative methods for image deblurring: a Matlab object-oriented approach. *Numerical Algorithms* **2004**, *36*, 73–93. See also: http://www.mathcs.emory.edu/~nagy/RestoreTools.
3. Almeida, M.S.; Figueiredo, M. Deconvolving images with unknown boundaries using the alternating direction method of multipliers. *IEEE Transactions on Image processing* **2013**, *22*, 3074–3086.
4. Dell'Acqua, P. A note on Taylor boundary conditions for accurate image restoration. *Advances in Computational Mathematics* **2017**, *43*, 1283–1304.
5. Donatelli, M. Fast transforms for high order boundary conditions in deconvolution problems. *BIT Numerical Mathematics* **2010**, *50*, 559–576.
6. Hanke, M.; Hansen, P.C. Regularization methods for large-scale problems. *Surveys on Mathematics for Industry* **1993**, *3*, 253–315.
7. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena* **1992**, *60*, 259–268.
8. Daubechies, I.; Defrise, M.; De Mol, C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences* **2004**, *57*, 1413–1457.
9. Figueiredo, M.A.; Nowak, R.D. An EM algorithm for wavelet-based image restoration. *IEEE Transactions on Image Processing* **2003**, *12*, 906–916.
10. Chan, R.H.; Chan, T.F.; Shen, L.; Shen, Z. Wavelet algorithms for high-resolution image reconstruction. *SIAM Journal on Scientific Computing* **2003**, *24*, 1408–1432.
11. Cai, J.F.; Osher, S.; Shen, Z. Linearized Bregman iterations for frame-based image deblurring. *SIAM Journal on Imaging Sciences* **2009**, *2*, 226–252.

12. Cai, J.F.; Osher, S.; Shen, Z. Split Bregman methods and frame based image restoration. *Multiscale modeling & simulation* **2010**, *8*, 337–369.

13. Dell'Acqua, P.; Donatelli, M.; Estatico, C. Preconditioners for image restoration by reblurring techniques. *Journal of Computational and Applied Mathematics* **2014**, *272*, 313–333.

14. Yin, W.; Osher, S.; Goldfarb, D.; Darbon, J. Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. *SIAM Journal on Imaging sciences* **2008**, *1*, 143–168.

15. Donatelli, M.; Hanke, M. Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring. *Inverse Problems* **2013**, *29*, 095008.

16. Buccini, A. Regularizing preconditioners by non-stationary iterated Tikhonov with general penalty term. *Applied Numerical Mathematics* **2017**, *116*, 64–81.

17. Cai, Y.; Donatelli, M.; Bianchi, D.; Huang, T.Z. Regularization preconditioners for frame-based image deblurring with reduced boundary artifacts. *SIAM Journal on Scientific Computing* **2016**, *38*, B164–B189.

18. Buccini, A.; Donatelli, M.; Reichel, L. Iterated Tikhonov regularization with a general penalty term. *Numerical Linear Algebra with Applications* **2017**, *24*, e2089.

19. Buccini, A.; Pasha, M.; Reichel, L. Generalized singular value decomposition with iterated Tikhonov regularization. *Journal of Computational and Applied Mathematics* **2020**, *373*, 112276.

20. Huang, G.; Reichel, L.; Yin, F. Projected nonstationary iterated Tikhonov regularization. *BIT Numerical Mathematics* **2016**, *56*, 467–487.

21. Dell'Acqua, P.; Donatelli, M.; Estatico, C.; Mazza, M. Structure preserving preconditioners for image deblurring. *Journal of Scientific Computing* **2017**, *72*, 147–171.

22. Bianchi, D.; Buccini, A.; Donatelli, M. Structure Preserving Preconditioning for Frame-Based Image Deblurring. In *Computational Methods for Inverse Problems in Imaging*; Springer, 2019; pp. 33–49.

23. Rudin, W. Functional analysis. 1991. *International Series in Pure and Applied Mathematics* **1991**.

24. Engl, H.W.; Hanke, M.; Neubauer, A. *Regularization of inverse problems*; Vol. 375, Springer Science & Business Media, 1996.

25. Bianchi, D.; Donatelli, M. On generalized iterated Tikhonov regularization with operator-dependent seminorms. *Electronic Transactions in Numerical Analysis* **2017**, *47*, 73–99.

26. Huckle, T.K.; Sedlacek, M. Tikhonov–Phillips regularization with operator dependent seminorms. *Numerical Algorithms* **2012**, *60*, 339–353.

27. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* **2004**, *13*, 600–612.

28. Buccini, A.; Reichel, L. An $\ell^2 - \ell^q$ Regularization Method for Large Discrete Ill-Posed Problems. *Journal of Scientific Computing* **2019**, *78*, 1526–1549.

29. Huang, J.; Donatelli, M.; Chan, R.H. Nonstationary iterated thresholding algorithms for image deblurring. *Inverse Problems & Imaging* **2013**, *7*, 717–736.