

Variational Image Restoration with Constraints on Noise Whiteness

Alessandro Lanza · Serena Morigi · Fiorella Sgallari

Received: 24 February 2014 / Accepted: 17 November 2014 / Published online: 4 December 2014
© Springer Science+Business Media New York 2014

Abstract We propose a novel variational framework for image restoration based on the assumption that noise is additive and white. In particular, the proposed variational model uses total variation regularization and forces the resemblance of the residue image to a white noise realization by imposing constraints in the frequency domain. The *whiteness constraint* constitutes the key novelty behind our approach. The restored image is efficiently computed by the constrained minimization of an energy functional using an alternating directions methods of multipliers procedure. Numerical examples show that the novel approach is particularly suited for textured image restorations.

Keywords Image restoration · Whiteness · Image deconvolution/deblurring · Variational methods

1 Introduction

Image restoration refers to the recovery of a clean sharp image from a noisy, and potentially blurred, observation. In this paper, we consider the problem of restoring images corrupted by blur and additive white noise. Several image restoration algorithms such as nonlinear-diffusion partial differential equations (PDE)-based methods and TV regularized strategies, succeeded in obtaining good quality edge preserving restorations, especially for noise removal. However, they

modify images towards piecewise constant functions, in such a way that important information, encoded in image features like textures and details, is often compromised in the restoration process.

We propose a novel variational formulation which exploits the relevant information on the whiteness of the noise thus resulting in a method that is particularly suitable for the restoration of partly-textured perturbed images, since it preserves fine scale features in the restoration process. In particular, we take advantage of the fact that while textures are characterized as repeated and meaningful structure of small patterns, noise is characterized as uncorrelated random patterns.

One aspect generally missing from the state-of-the-art image restoration methods is a full exploitation of all the available information about the noise. More precisely, most algorithms recover an estimate of the clean image by using prior knowledge only about the variance of the noise, but totally neglect the significantly more information-rich property of the noise being the realization of a white random process. This is to some extent surprising provided that, on the other hand, the whiteness of the residue image is often used as an a-posteriori criterion for evaluating the performance of restoration algorithms (see, e.g., [18]). In particular, by evaluating the resemblance of the residue image to a white noise realization, one can check, to some extents, the quality of the restored image. In [1, 10, 17, 18] the measures of residual spectral whiteness have been exploited for adjusting the regularization parameter and/or the number of iterations of the algorithms for deconvolution problems. Comparisons among several state of the art methods have been documented in [2].

The *whiteness constraint* constitutes the key novelty behind our approach. A penalty term to favor whiteness in the residue image has been proposed in [14] for the image

A. Lanza · S. Morigi (✉) · F. Sgallari
Department of Mathematics, University of Bologna, Bologna, Italy
e-mail: serena.morigi@unibo.it

A. Lanza
e-mail: alessandro.lanza2@unibo.it

F. Sgallari
e-mail: fiorella.sgallari@unibo.it

denoising problem. One of the contribution of this paper is to incorporate such a relevant information into the restoration process when the image is both blurred and corrupted by additive white noise. The second one is the proposal of an efficient algorithm for the solution of the new model.

Without loss of generality, we consider grayscale images with a square $d \times d$ domain. Let $u \in \mathbb{R}^{d^2}$ be the unknown $d \times d$ clean image concatenated into an d^2 -vector, $K \in \mathbb{R}^{d^2 \times d^2}$ be a known blurring operator and $n \in \mathbb{R}^{d^2}$ be an unknown realization of the random noise process, which we assume white with zero-mean and standard deviation σ . The discrete imaging model of the degradation process which relates the observed degraded image $g \in \mathbb{R}^{d^2}$ with u , can be expressed as follows:

$$g = Ku + n. \quad (1.1)$$

Given K and g , our goal is to solve the inverse problem of recovering an estimate u^* of u from g , which is known as deconvolution or deblurring. When K is the identity operator, recovering u is referred as denoising.

It is well known that deblurring by direct inversion of K , i.e. $u^* = K^{-1}g$, can yield meaningless results due to the fact that the linear blur operator K is typically ill-conditioned or singular [11]. To obtain a meaningful restoration, some prior information are commonly included in the formulation. A classical regularization approach consists in solving the optimization problem

$$u^* \leftarrow \arg \min_{u \in \mathbb{R}^{d^2}} \{ \mathcal{R}(u) + \mu \mathcal{F}(u, g) \}, \quad (1.2)$$

where the regularization term $\mathcal{R}(u)$ enforces certain prior constraints on the clean image u , the fidelity term $\mathcal{F}(u, g)$ measures how u fits the observed image g , and μ is a positive regularization parameter balancing the two terms. In particular, the fidelity term is typically related to the type of noise corrupting the image, while the regularization term commonly enforces smoothness of the solution.

Many different regularization functionals have been proposed, ranging from the classical Tikhonov [3, 11], to total variation (TV) [16], Mumford-Shah [12, 13], and many other variants of these. In this paper, we consider the TV regularization simply for its popularity, any other regularizers could be substituted as well. A popular and effective choice for images corrupted by additive Gaussian noise, is the TV- L_2 functional formulation:

$$u^* \leftarrow \arg \min_{u \in \mathbb{R}^{d^2}} \left\{ TV(u) + \frac{\mu}{2} \|Ku - g\|_2^2 \right\}, \quad (1.3)$$

where $TV(u)$ denotes the TV semi-norm of u defined as

$$TV(u) = \sum_{i=1}^{d^2} \|(\nabla u)_i\|_2, \quad (1.4)$$

with $(\nabla u)_i := (D_{x,i}u, D_{y,i}u)$ denoting the discrete gradient of u at pixel i and $D_{x,i}, D_{y,i}$ representing the i -th rows of the x and y -directional first order finite difference operators $D_x, D_y \in \mathbb{R}^{d^2 \times d^2}$, respectively. The regularization parameter μ plays a critical role in the success of the restoration process [21].

When the standard deviation σ of noise is available, under the discrepancy principle, the unconstrained TV- L_2 model in (1.3) can be equivalently reformulated as the following constrained optimization problem:

$$u^* \leftarrow \arg \min_{u \in U} TV(u), \quad (1.5)$$

with the feasible set defined as

$$U = \left\{ u \in \mathbb{R}^{d^2} : \|Ku - g\|_2^2 \leq \tau^2 d^2 \sigma^2 \right\}, \quad (1.6)$$

where τ is a pre-determined scalar parameter controlling the variance of the residue image $Ku - g$.

In this paper, we propose to explicitly incorporate the constraints on the whiteness of the residue image by modifying the feasible set defined in (1.6). The proposed TV-W model is

$$u^* \leftarrow \arg \min_{u \in W} TV(u), \quad (1.7)$$

where the new feasible set $W \subset \mathbb{R}^{d^2}$ contains solutions u such that the corresponding residue image $Ku - g$ satisfies the discrepancy principle and resembles a white noise realization. We will propose a particular whiteness set W and, accordingly, we will present an efficient minimization method based on the ADMM strategy which was originally developed in the 1970s [6, 7], and recently applied to the image restoration problem [5].

More recently, a fast TV deconvolution algorithm called FTVd was proposed in [20] for image restoration with L_2 fidelity, which is a quadratic penalty method. In [8] an augmented Lagrangian method has been successfully applied to overcome the difficulties due to the penalty parameter. Alternating direction method (ADM), a variant of the classic augmented Lagrangian method for structured optimization, has been introduced in [19]. Finally, in [5] the ADM strategy has been extended to the solution of constrained TV problems and named ADMM.

The paper is organized as follows. In Sect. 2 we present and motivate our choice for the whiteness set W . In Sect. 3 we illustrate in detail the ADMM-based algorithm used to minimize the functional together with a discussion of the computational aspects. In Sect. 4 we present experimental results assessing the performance of the proposed model. In Sect. 5 we draw conclusions.

2 Imposing Whiteness Constraints

The statistical version of the degradation model in (1.1) with images in matrix-form is

$$G(i, j) = \bar{v}(i, j) + \bar{N}(i, j), \quad (i, j) \in \Omega = \{1, \dots, d\}^2, \tag{2.1}$$

with capital letters indicating random quantities and where $\bar{v}(i, j) = (K u)(i, j)$ denotes the value of the blurred image at pixel (i, j) . In particular, the additive noise is modeled as a $d \times d$ discrete random process $\tilde{N} := \{\tilde{N}(i, j) : (i, j) \in \Omega\}$ with $\tilde{N}(i, j)$ denoting the scalar random variable modeling noise at pixel (i, j) . This means that different images of the same subject under the same blurring operator will differ due to the inherently random nature of the additive noise. The ensemble auto-correlation of \tilde{N} is a function $\rho_{\tilde{N}}$ which maps pairs of pixel locations $(i_1, j_1), (i_2, j_2) \in \Omega$ into a scalar value given by

$$\rho_{\tilde{N}}(i_1, j_1, i_2, j_2) := E [\tilde{N}(i_1, j_1) \tilde{N}(i_2, j_2)], \tag{2.2}$$

where E is the expectation operator [9].

Since we assume that noise is zero-mean and white, i.e. wide-sense stationary and uncorrelated, the auto-correlation of \tilde{N} depends only on the lag between the two pixel locations $(l, m) = (i_2 - i_1, j_2 - j_1)$ and, with a little abuse of notation, (2.2) can be rewritten as follows

$$\begin{aligned} \rho_{\tilde{N}}(i_1, j_1, i_2, j_2) &= \rho_{\tilde{N}}(0, 0, i_2 - i_1, j_2 - j_1) \\ &= \rho_{\tilde{N}}(l, m) = E [\tilde{N}(i, j) \tilde{N}(i + l, j + m)] \end{aligned} \tag{2.3}$$

$$= \begin{cases} \sigma^2, & \text{if } (l, m) = (0, 0) \\ 0, & \text{otherwise} \end{cases}, \tag{2.4}$$

$$(l, m) \in \Theta = \{-(d - 1), \dots, d - 1\}^2.$$

Equation (2.3) expresses stationarity of the noise process and holds independently for every $(i, j) \in \Omega$ such that $(i + l, j + m) \in \Omega$, whereas (2.4) says that a zero-mean white noise process is characterized by zero values of the auto-correlation function at all non-vanishing lags.

Given a single realization $n := \{n(i, j) \in \mathbb{R} : (i, j) \in \Omega\}$ of the noise process \tilde{N} , that is the series of noise values corrupting the particular observed image according to the deterministic degradation model in (1.1), the sample auto-correlation of n is a function r_n mapping all the possible lags $(l, m) \in \Theta$ into a scalar value given by

$$\begin{aligned} r_n(l, m) &:= \frac{1}{d^2} (n \star n)_{l,m} = (n * n')_{l,m} \\ &= \frac{1}{d^2} \sum_{(i,j) \in \Omega} n(i, j) n(i + l, j + m), \quad (l, m) \in \Theta, \end{aligned} \tag{2.5}$$

where \star and $*$ denote the 2-D discrete correlation and convolution operators, respectively, and where $n'(i, j)$

$= n(-i, -j)$. Clearly, for (2.5) being defined for all lags $(l, m) \in \Theta$, the noise realization n must be padded with at least $d - 1$ samples in all directions.

It can be demonstrated that, with the further assumption that the noise process is also ergodic [9], the sample auto-correlation is a good estimate of the ensemble auto-correlation. In particular, we have:

$$\lim_{d \rightarrow \infty} r_n(l, m) = \rho_{\tilde{N}}(l, m). \tag{2.6}$$

A set of white noise realizations could thus be defined in the spatial domain by constraining the values of the sample auto-correlation function r_n in (2.5) to lie within a band around the theoretical limit in (2.6). In [14] the authors proposed a denoising method based on a novel fidelity term penalizing the autocorrelation function of the residual. The numerical solution of the obtained integro-differential equation was computed by an explicit finite difference method. The extension of such strategy to our setting, due to the presence of the blur operator, would lead to very strict stability conditions and consequently to a very slow procedure. In this paper we followed a new strategy considering a whiteness definition in the frequency domain and, accordingly, we propose a novel whiteness set.

The periodogram $p_n \in \mathbb{R}^{d^2}$ of the noise realization n is defined as [9]

$$p_n = |Tn|, \tag{2.7}$$

where $|\cdot|$ denotes the component-wise modulus of the complex vector Tn , where $T \in \mathbb{C}^{d^2 \times d^2}$ is obtained using the unitary matrix representing the 1D discrete Fourier transform (DFT) $T^{(1)} \in \mathbb{C}^{d \times d}$

$$T = T^{(1)} \otimes T^{(1)}, \tag{2.8}$$

and where \otimes denotes the Kronecker product operator.

It can be demonstrated that the elements of the vector p_n are independent random variables. Moreover, the first element of p_n has a χ distribution with 1 degree of freedom and all the other elements are distributed as a χ distribution with 2 degrees of freedom [9]. We recall that the probability density function $p_k(x)$ and the mean value μ_k of a χ -distributed random variable with k degrees of freedom are, respectively:

$$p_k(x) = \frac{2^{1-k/2} x^{k-1} e^{-x^2/2}}{\Gamma(k/2)} \quad \text{and} \quad \mu_k = \sqrt{2} \frac{\Gamma((k + 1)/2)}{\Gamma(k/2)}, \tag{2.9}$$

where Γ is the Gamma function. To obtain a vector of random variables having all mean equal to one, and following from the properties of the χ distribution, we scale the vector p_n by means of a diagonal $d^2 \times d^2$ normalization matrix M , i.e. $\hat{p}_n = Mp_n$, with M defined as

$$M = \frac{1}{\tau \sigma d^2} \text{diag} \left(\sqrt{2} \Gamma(1.5), \frac{1}{\Gamma(1.5)}, \dots, \frac{1}{\Gamma(1.5)} \right). \quad (2.10)$$

Next we sort the elements of the normalized periodogram \hat{p}_n in increasing order of spatial frequency. At this aim, we introduce the $d \times d$ matrix Ψ with elements $\Psi_{i,j} = i^2 + j^2$ and the $d^2 \times d^2$ permutation matrix Π such that the elements of $\Pi \xi$ appear in nondecreasing order, with $\xi \in \mathbb{R}^{d^2}$ denoting the column-vector form of Ψ . Then the vector $\Pi \hat{p}_n$ holds all the elements of \hat{p}_n in order of increasing spatial frequency. We define the *normalized cumulative periodogram* (NCP) of n as the vector $c_n \in \mathbb{R}^{d^2}$ given by

$$c_n := S \Pi \hat{p}_n, \quad (2.11)$$

where S is the $d^2 \times d^2$ partial sums matrix, i.e. a lower triangular matrix with all nonzero entries equal to one.

In Sect. 4 we test experimentally by a Montecarlo simulation that in case that n is the realization of a zero-mean white noise process, as the image dimension d increases the associated NCP vector c_n tends to a limit vector c_w , that we call theoretical whiteness NCP vector, given by the discretization of the straight line between $(1, d^{-2})$ and $(d^2, 1)$, that is:

$$c_w = d^{-2} (1, 2, \dots, d^2)^T. \quad (2.12)$$

In this paper, we will exploit such a property of white noise to constrain the NCP vector c_n of the restored residue image $n = Ku - g$ to lie within a narrow band around the theoretical whiteness NCP vector c_w . In particular, we define the whiteness set W to be used in our model (1.7) as follows:

$$W = \left\{ u \in \mathbb{R}^{d^2} : b^- \leq S \Pi M |T(Ku - g)| \leq b^+ \right\}, \quad (2.13)$$

where the inequalities must be intended component-wise and $b^+, b^- \in \mathbb{R}^{d^2}$ are vectors containing the upper and lower limits of the whiteness set, respectively. In Sect. 4, we will illustrate how the values of b^+ and b^- can be selected according to probabilistic arguments. By constraining the restored image u to belong to the set W in (2.13) with appropriate set limits b^- and b^+ , we are thus implicitly forcing the restoration residual to resemble the realization of a white noise process.

However, from the definitions of the whiteness set in (2.13), we notice that the complicated constraint on u can be turned into a simple box-constraint on the NCP of the residue image c_n , that is c_n must belong to the following whiteness set:

$$B = \left\{ c_n \in \mathbb{R}^{d^2} : b^- \leq c_n \leq b^+ \right\}, \quad (2.14)$$

where b^- and b^+ are the pre-computed limits of the box-constraint.

In Fig. 1 we illustrate the capability of the set W in discriminating white signals, by showing the NCP vectors for two different 1D signals: the first signal in Fig. 1a is the realization of a white gaussian process with zero-mean, standard deviation $\sigma = 4$, and number of samples $d = 1,000$, the second signal in Fig. 1b is obtained from the first signal by setting one-fifth of the samples to a constant value equal to 4. In Fig. 1c, d we show the NCP vectors c of the 1D sampled signals depicted in Fig. 1a, b, respectively. We note that the NCP vector c of the white signal depicted in Fig. 1c is very similar to the theoretical whiteness NCP vector c_w , while a significant deviation occurs for the partially modified signal, as visible in Fig. 1d.

3 Applying ADMM to the Proposed Model

Constrained problems are in general much more difficult to solve than the unconstrained ones. Moreover, for constrained TV problems, the singularity of the TV functional prohibits the application of Newton-like methods [15]. Recently Chan et al. [5] successfully adapted the ADMM algorithm for solving both the constrained TV-L2 and TV-L1 problems. Thus in the following we introduce a suitable variant of the basic ADMM approach to solve the proposed constrained minimization problem in (1.7) with the whiteness set W defined in (2.13).

At this aim, we first introduce two auxiliary variables t and z to reformulate (1.7) into the following equivalent form:

$$\begin{aligned} u^* &\leftarrow \arg \min_{u,t,z} \left\{ \sum_{i=1}^{d^2} \|t_i\|_2 + \iota_W(z) \right\} \\ \text{s.t. : } &t = Du, z = u, \end{aligned} \quad (3.1)$$

where $D = (D_x; D_y) \in \mathbb{R}^{2d^2 \times d^2}$, ι_W is the indicator function of the whiteness set W , with the convention that $\iota_W(z)$ takes the value 0 for $z \in W$ and ∞ otherwise. The auxiliary variable t is introduced to transfer the discrete gradient operator $(\nabla u)_i$ out of the non-differentiable term $\|\cdot\|_2$. The variable z plays the role of u within the whiteness constraint so that the constraint is now imposed on z instead of u .

As we have already noticed the whiteness set W on u , i.e. on z , can be reformulated into a simple box-constraint set B on c_n , where B is defined in (2.14). To exploit this new simple set we introduce three new auxiliary variables $f = F(Ku - g)$, $p = \Pi M |f|$ and $c = Sp$, where $F = (T_{real}; T_{imag}) \in \mathbb{R}^{2d^2 \times d^2}$.

Then the minimization problem replacing (3.1) is written as:

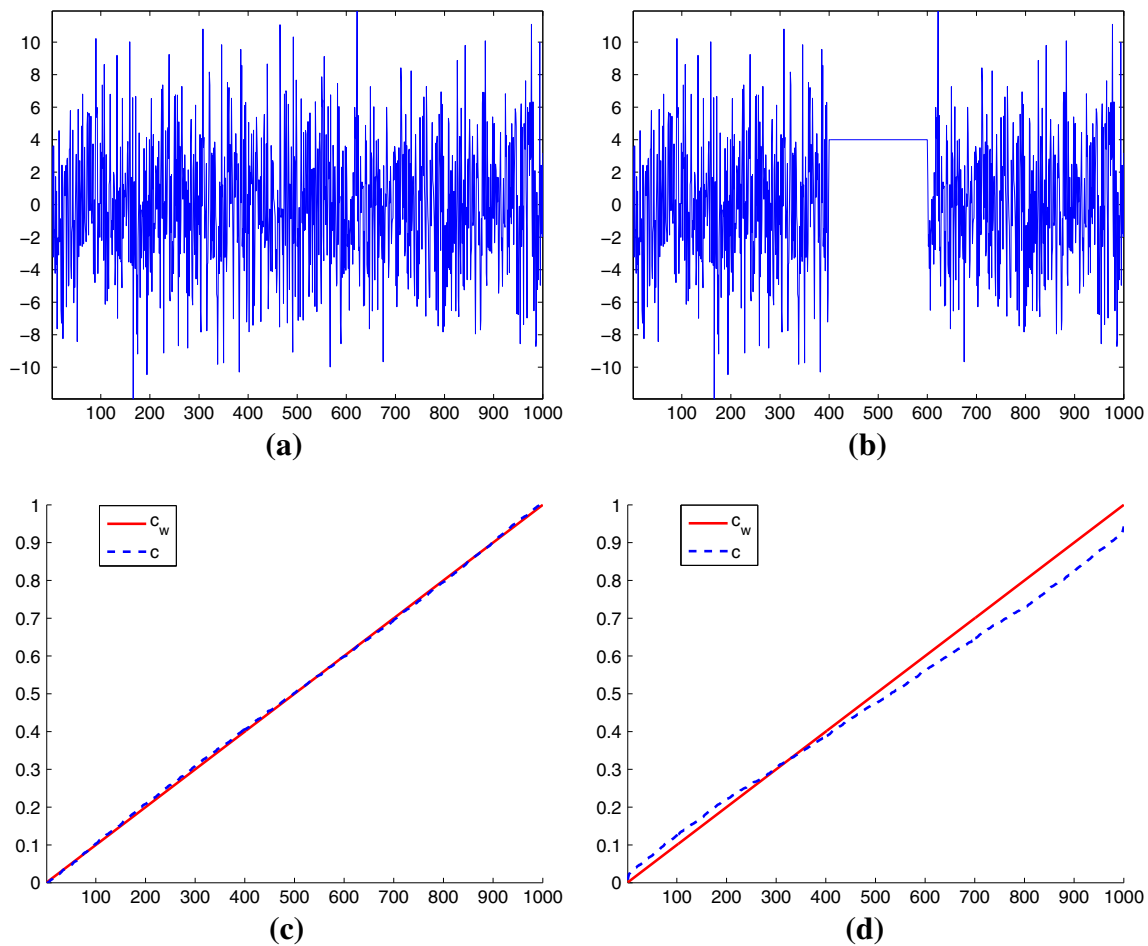


Fig. 1 Sample 1D signals (*top*) and associated NCP vectors (*bottom*)

$$\begin{aligned}
 u^* &\leftarrow \arg \min_{u,t,f,p,c} \left\{ \sum_{i=1}^{d^2} \|t_i\|_2 + \iota_B(c) \right\} \\
 \text{s.t. : } &t = Du, \quad f = F(Ku - g), \quad p = \Pi M|f|, \quad c = Sp,
 \end{aligned} \tag{3.2}$$

where ι_B is the indicator function of the whiteness set B .

The augmented Lagrangian functional associated with (3.2) is

$$\begin{aligned}
 \mathcal{L}(u, t, f, p, c; \lambda_t, \lambda_f, \lambda_p, \lambda_c) &= \sum_{i=1}^{d^2} \|t_i\|_2 + \iota_B(c) \\
 &- \langle \lambda_t, t - Du \rangle + \frac{\beta_t}{2} \|t - Du\|_2^2 \\
 &- \langle \lambda_f, f - F(Ku - g) \rangle + \frac{\beta_f}{2} \|f - F(Ku - g)\|_2^2 \\
 &- \langle \lambda_p, p - \Pi M|f| \rangle + \frac{\beta_p}{2} \|p - \Pi M|f|\|_2^2 \\
 &- \langle \lambda_c, c - Sp \rangle + \frac{\beta_c}{2} \|c - Sp\|_2^2,
 \end{aligned} \tag{3.3}$$

where $\beta_t, \beta_f, \beta_p, \beta_c > 0$ are scalar penalty parameters and $\lambda_t, \lambda_f \in Q, \lambda_p, \lambda_c \in V$ are the vectors of Lagrangian multipliers, with $V = \mathbb{R}^{d^2}, Q = \mathbb{R}^{2d^2}$.

Solving (3.2) is thus equivalent to search for the solutions of the following saddle point problem:

$$\begin{aligned}
 \text{Find } &(x^*; \lambda^*) \in X \times \Lambda \\
 \text{s.t. } &\mathcal{L}(x^*; \lambda) \leq \mathcal{L}(x^*; \lambda^*) \leq \mathcal{L}(x; \lambda^*) \\
 &\forall (x; \lambda) \in X \times \Lambda,
 \end{aligned} \tag{3.4}$$

with \mathcal{L} defined in (3.3) and where, for simplicity of notations, we set $x = (u, t, f, p, c), \lambda = (\lambda_t, \lambda_f, \lambda_p, \lambda_c), X = V \times Q \times Q \times V \times V$ and $\Lambda = Q \times Q \times V \times V$.

Starting at $u = u^k, \lambda_t = \lambda_t^k, \lambda_f = \lambda_f^k, \lambda_p = \lambda_p^k, \lambda_c = \lambda_c^k$, the ADMM iterative scheme applied to the solution of (3.2) reads as follows:

$$t^{k+1} \leftarrow \arg \min_{t \in Q} \mathcal{L}(u^k, t, f^k, p^k, c^k; \lambda_t^k, \lambda_f^k, \lambda_p^k, \lambda_c^k) \tag{3.5}$$

$$f^{k+1} \leftarrow \arg \min_{f \in Q} \mathcal{L}(u^k, t^{k+1}, f, p^k, c^k; \lambda_t^k, \lambda_f^k, \lambda_p^k, \lambda_c^k) \tag{3.6}$$

$$p^{k+1} \leftarrow \arg \min_{p \in V} \mathcal{L}(u^k, t^{k+1}, f^{k+1}, p, c^k; \lambda_t^k, \lambda_f^k, \lambda_p^k, \lambda_c^k) \tag{3.7}$$

$$c^{k+1} \leftarrow \arg \min_{c \in V} \mathcal{L}(u^k, t^{k+1}, f^{k+1}, p^{k+1}, c; \lambda_t^k, \lambda_f^k, \lambda_p^k, \lambda_c^k) \tag{3.8}$$

$$u^{k+1} \leftarrow \arg \min_{u \in V} \mathcal{L}(u, t^{k+1}, f^{k+1}, p^{k+1}, c^{k+1}; \lambda_t^k, \lambda_f^k, \lambda_p^k, \lambda_c^k) \tag{3.9}$$

$$\begin{pmatrix} \lambda_t^{k+1} \\ \lambda_f^{k+1} \\ \lambda_p^{k+1} \\ \lambda_c^{k+1} \end{pmatrix} \leftarrow \begin{pmatrix} \lambda_t^k - \gamma \beta_t (t^{k+1} - Du^{k+1}) \\ \lambda_f^k - \gamma \beta_f (f^{k+1} - F(Ku^{k+1} - g)) \\ \lambda_p^k - \gamma \beta_p (p^{k+1} - \Pi M|f^{k+1}|) \\ \lambda_c^k - \gamma \beta_c (c^{k+1} - Sp^{k+1}) \end{pmatrix}, \tag{3.10}$$

where γ is a relaxation parameter chosen in the interval $(0, (\sqrt{5} + 1)/2)$, as analyzed in [5].

In the following subsections we show in detail how to solve the five minimization sub-problems (3.5)–(3.9), then we present the overall iterative ADMM-based minimization algorithm. As we will see, the first two sub-problems admit a very efficient closed-form solution based on the following proposition, whose proof is reported in the Appendix.

Proposition 3.1 *Let $\alpha \in \mathbb{R}, \beta \in \mathbb{R}^+$ and $\lambda, v \in \mathbb{R}^n$ with $n \geq 1$ be given constants. Then, the solution x^* of the n -dimensional minimization problem*

$$\arg \min_{x \in \mathbb{R}^n} \left\{ \alpha \|x\|_2 - \langle \lambda, x \rangle + \frac{\beta}{2} \|x - v\|_2^2 \right\} \tag{3.11}$$

is given by:

$$\begin{aligned} \text{a) } x^* &= \max \left\{ 1 - \frac{\alpha}{\beta \|q\|_2}, 0 \right\} q \quad \text{if } \|q\|_2 > 0, \forall \alpha \\ \text{b) } x^* &= 0 \quad \text{if } \|q\|_2 = 0 \text{ and } \alpha \geq 0 \\ \text{c) } x^* &\in \left\{ x \in \mathbb{R}^n : \|x\|_2 = -\frac{\alpha}{\beta} \right\} \quad \text{if } \|q\|_2 = 0 \text{ and } \alpha < 0 \end{aligned} \tag{3.12}$$

where we set

$$q = v + \frac{1}{\beta} \lambda. \tag{3.13}$$

3.1 Solving the Sub-problem for t

Given the definition of the augmented Lagrangian functional in (3.3), the minimization sub-problem for t in (3.5) can be written as follows:

$$\begin{aligned} t^* &\leftarrow \arg \min_{t \in Q} \left\{ \sum_{i=1}^{d^2} \|t_i\|_2 - \langle \lambda_t, t - Du \rangle + \frac{\beta_t}{2} \|t - Du\|_2^2 \right\} \\ &\leftarrow \arg \min_{t \in Q} \left\{ \sum_{i=1}^{d^2} \|t_i\|_2 - \langle \lambda_t, t \rangle + \frac{\beta_t}{2} \|t - Du\|_2^2 \right\} \end{aligned} \tag{3.14}$$

$$\leftarrow \arg \min_{t \in Q} \left\{ \sum_{i=1}^{d^2} \left(\|t_i\|_2 - \langle \lambda_{t,i}, t_i \rangle + \frac{\beta_t}{2} \|t_i - D_i u\|_2^2 \right) \right\}. \tag{3.15}$$

Note that in (3.14) we omitted the constant terms while in (3.15) the functional is written in an explicit component-wise form. The minimization in (3.15) is equivalent to the following d^2 problems:

$$\begin{aligned} t_i^* &\leftarrow \arg \min_{t_i \in \mathbb{R}^2} \left\{ \|t_i\|_2 - \langle \lambda_{t,i}, t_i \rangle + \frac{\beta_t}{2} \|t_i - D_i u\|_2^2 \right\}, \\ &i = 1, \dots, d^2. \end{aligned} \tag{3.16}$$

Based on Proposition (3.1) and setting

$$q_i := D_i u + \frac{1}{\beta_t} \lambda_{t,i}, \quad i = 1, \dots, d^2, \tag{3.17}$$

the solution of (3.16) is given explicitly by the following d^2 operations:

$$t_i^* = \max \left\{ \|q_i\|_2 - \frac{1}{\beta_t}, 0 \right\} \frac{q_i}{\|q_i\|_2}, \quad i = 1, \dots, d^2. \tag{3.18}$$

where $0 \cdot (0/0) = 0$ is assumed. We notice that the computational cost of (3.17)–(3.18) is linear with respect to the number of pixels d^2 .

3.2 Solving the Sub-problem for f

The minimization sub-problem for f in (3.6) is as follows:

$$\begin{aligned} f^* &\leftarrow \arg \min_{f \in Q} \left\{ -\langle \lambda_f, f - F(Ku - g) \rangle \right. \\ &\quad \left. - \langle \lambda_p, p - \Pi M|f| \rangle + \frac{\beta_f}{2} \|f - F(Ku - g)\|_2^2 \right. \\ &\quad \left. + \frac{\beta_p}{2} \|p - \Pi M|f|\|_2^2 \right\} \\ &\leftarrow \arg \min_{f \in Q} \left\{ -\langle \lambda_f, f \rangle + \langle \lambda_p, \Pi M|f| \rangle \right. \\ &\quad \left. + \frac{\beta_f}{2} \|f - F(Ku - g)\|_2^2 + \frac{\beta_p}{2} \|p - \Pi M|f|\|_2^2 \right\} \\ &\leftarrow \arg \min_{f \in Q} \left\{ -\langle \lambda_f, f \rangle + \langle \Pi^T \lambda_p, M|f| \rangle \right. \\ &\quad \left. + \frac{\beta_f}{2} \|f - F(Ku - g)\|_2^2 + \frac{\beta_p}{2} \|\Pi^T p - M|f|\|_2^2 \right\}. \end{aligned} \tag{3.19}$$

In (3.19) we omitted the constant terms while in (3.20) we rearranged the second and fourth term of the functional by exploiting the orthogonality of Π .

To simplify notations, we denote by $\tilde{\lambda}_p = \Pi^T \lambda_p$ and $\tilde{p} = \Pi^T p$ the inversely-permuted versions of vectors λ_p

and p , respectively, by M_i the i -th element on the main diagonal of M in (2.10) and by $w_i = F_i(Ku - g)$ the two-dimensional vector containing the real and imaginary parts of the i th component of the 2D-DFT of the residue image. The component-wise form of (3.20) reads thus as follows:

$$f^* \leftarrow \arg \min_{f \in Q} \sum_{i=1}^{d^2} \left\{ -\langle \lambda_{f,i}, f_i \rangle + \tilde{\lambda}_{p,i} M_i \|f_i\|_2 + \frac{\beta_f}{2} \|f_i - w_i\|_2^2 + \frac{\beta_p}{2} (\tilde{p}_i - M_i \|f_i\|_2)^2 \right\} \quad (3.21)$$

$$\leftarrow \arg \min_{f \in Q} \sum_{i=1}^{d^2} \left\{ -\langle \lambda_{f,i}, f_i \rangle + \tilde{\lambda}_{p,i} M_i \|f_i\|_2 + \frac{\beta_f}{2} (\|f_i\|_2^2 + \|w_i\|_2^2 - 2\langle f_i, w_i \rangle) + \frac{\beta_p}{2} (\tilde{p}_i^2 + M_i^2 \|f_i\|_2^2 - 2\tilde{p}_i M_i \|f_i\|_2) \right\} \quad (3.22)$$

$$\leftarrow \arg \min_{f \in Q} \sum_{i=1}^{d^2} \left\{ M_i (\tilde{\lambda}_{p,i} - \beta_p \tilde{p}_i) \|f_i\|_2 - \langle \lambda_{f,i} + \beta_f w_i, f_i \rangle + \frac{\beta_f + M_i^2 \beta_p}{2} \|f_i\|_2^2 \right\}, \quad (3.23)$$

where the constant terms have been omitted in (3.23). The minimization in (3.23) is equivalent to the following d^2 problems:

$$f_i^* \leftarrow \arg \min_{f_i \in \mathbb{R}^2} \left\{ \alpha_i \|f_i\|_2 - \langle \rho_i, f_i \rangle + \frac{\kappa_i}{2} \|f_i\|_2^2 \right\}, \quad i = 1, \dots, d^2, \quad (3.24)$$

where the constants $\alpha_i \in \mathbb{R}$, $\rho_i \in \mathbb{R}^2$ and $\kappa_i \in \mathbb{R}$ in (3.24) are defined as

$$\begin{cases} \alpha_i = M_i (\tilde{\lambda}_{p,i} - \beta_p \tilde{p}_i) \\ \rho_i = \lambda_{f,i} + \beta_f w_i \\ \kappa_i = \beta_f + \beta_p M_i^2 \end{cases}, \quad i = 1, \dots, d^2. \quad (3.25)$$

Similar to the solution of (3.16), based on Proposition (3.1) and setting

$$q_i := \frac{1}{\kappa_i} \rho_i, \quad i = 1, \dots, d^2, \quad (3.26)$$

the solution of (3.24) is given explicitly by the following d^2 operations:

$$f_i^* = \max \left\{ \|q_i\|_2 - \frac{\alpha_i}{\kappa_i}, 0 \right\} \frac{q_i}{\|q_i\|_2}, \quad i = 1, \dots, d^2. \quad (3.27)$$

where $0 \cdot (0/0) = 0$ is assumed. The computational cost of (3.25)–(3.27) is dominated by the 2D-DFT of the residue image $F(Ku - g)$ required to compute w_i .

3.3 Solving the Sub-problem for p

Given f and c , the minimization sub-problem for p in (3.7) is as follows:

$$p^* \leftarrow \arg \min_{p \in V} \left\{ -\langle \lambda_p, p - \Pi M|f| \rangle - \langle \lambda_c, c - Sp \rangle + \frac{\beta_p}{2} \|p - \Pi M|f|\|_2^2 + \frac{\beta_c}{2} \|c - Sp\|_2^2 \right\} \leftarrow \arg \min_{p \in V} \left\{ -\langle \lambda_p, p \rangle + \langle \lambda_c, Sp \rangle + \frac{\beta_p}{2} \|p - \Pi M|f|\|_2^2 + \frac{\beta_c}{2} \|c - Sp\|_2^2 \right\}. \quad (3.28)$$

The problem (3.28) is a quadratic optimization problem whose optimality conditions read as follows:

$$-\lambda_p + S^T \lambda_c + \beta_p (p - \Pi M|f|) - \beta_c S^T (c - Sp) = 0 \quad (3.29)$$

that is

$$\left(S^T S + \frac{\beta_p}{\beta_c} I \right) p = \frac{1}{\beta_c} (\lambda_p + \beta_p \Pi M|f| + S^T (\beta_c c - \lambda_c)). \quad (3.30)$$

Recalling that S is a lower triangular matrix with all nonzero entries equal to one, the expressions for the $d^2 \times d^2$ matrix $S^T S$ and its inverse $(S^T S)^{-1}$ is quite straightforward to derive:

$$S^T S = \begin{bmatrix} d^2 & \dots & 4 & 3 & 2 & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 4 & \dots & 4 & 3 & 2 & 1 \\ 3 & \dots & 3 & 3 & 2 & 1 \\ 2 & \dots & 2 & 2 & 2 & 1 \\ 1 & \dots & 1 & 1 & 1 & 1 \end{bmatrix} (S^T S)^{-1} = \begin{bmatrix} 1 & -1 & 0 & \dots & \dots & \dots & 0 \\ -1 & 2 & -1 & \ddots & & & \vdots \\ 0 & -1 & 2 & \ddots & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & \ddots & 2 & -1 & 0 \\ \vdots & & & \ddots & -1 & 2 & -1 \\ 0 & \dots & \dots & \dots & 0 & -1 & 2 \end{bmatrix} \quad (3.31)$$

Since $S^T S$ is symmetric positive definite and $\frac{\beta_p}{\beta_c} > 0$, the coefficient matrix of the linear system in (3.30) is symmetric positive definite. Moreover, to exploit the sparsity of the matrix $(S^T S)^{-1}$ both sides of (3.30) are multiplied by $\frac{\beta_c}{\beta_p} (S^T S)^{-1}$ thus obtaining the following equivalent linear

system:

$$\begin{aligned} & \left((S^T S)^{-1} + \frac{\beta_c}{\beta_p} I \right) p \\ &= \frac{1}{\beta_p} S^{-1} \left(S^{-T} (\lambda_p + \beta_p \Pi M |f|) + \beta_c c - \lambda_c \right). \end{aligned} \quad (3.32)$$

The coefficient matrix in (3.32) is a tridiagonal symmetric positive definite matrix and is constant with the iterations, hence its bidiagonal Cholesky factorization can be computed as a preliminary step. At each iteration, the computational cost for solving (3.32) is twice the cost for solving a bidiagonal system.

3.4 Solving the Sub-problem for c

Given p , the minimization sub-problem for c in (3.8) is as follows:

$$\begin{aligned} c^* &\leftarrow \arg \min_{c \in V} \left\{ \iota_B(c) - \langle \lambda_c, c - Sp \rangle + \frac{\beta_c}{2} \|c - Sp\|_2^2 \right\} \\ &\leftarrow \arg \min_{c \in B} \left\{ \frac{\beta_c}{2} \left\| c - \left(Sp + \frac{\lambda_c}{\beta_c} \right) \right\|_2^2 \right\}. \end{aligned} \quad (3.33)$$

The solution of (3.33) is thus given by a simple Euclidean projection of the vector $Sp + \lambda_c/\beta_c$ onto the box-constraints defined by the whiteness set B in (2.14):

$$c^* = \mathcal{P}_B \left(Sp + \frac{\lambda_c}{\beta_c} \right). \quad (3.34)$$

Such a projection can be obtained straightforwardly by computing the following d^2 component-wise projections:

$$c_i^* = \max \left\{ \min \left\{ \left(S_i p + \frac{\lambda_{c,i}}{\beta_c} \right), b_i^+ \right\}, b_i^- \right\}, \quad i = 1, \dots, d^2, \quad (3.35)$$

where $S_i p$ is the i -th component of the vector of partial sums of p , and where $b_i^+, b_i^- \in \mathbb{R}$ represent the selected box-constraint limits of the i -th component of the NCP vector.

The computational complexity of this sub-problem is linear in the number of pixels d^2 . In fact, the matrix-vector product Sp in (3.34) can be efficiently obtained by computing the partial sums of the vector p incrementally.

3.5 Solving the Sub-problem for u

The minimization sub-problem for u in (3.9) can be re-written as follows:

$$\begin{aligned} u^* &\leftarrow \arg \min_{u \in V} \left\{ -\langle \lambda_t, t - Du \rangle - \langle \lambda_f, f - F(Ku - g) \rangle \right. \\ &\quad \left. + \frac{\beta_t}{2} \|t - Du\|_2^2 + \frac{\beta_f}{2} \|f - F(Ku - g)\|_2^2 \right\} \end{aligned}$$

$$\begin{aligned} &\leftarrow \arg \min_{u \in V} \left\{ \langle \lambda_t, Du \rangle + \langle \lambda_f, F(Ku) \rangle \right. \\ &\quad \left. + \frac{\beta_t}{2} \|t - Du\|_2^2 + \frac{\beta_f}{2} \|f - F(Ku - g)\|_2^2 \right\}. \end{aligned} \quad (3.36)$$

The problem (3.36) is a quadratic optimization problem whose optimality conditions read as follows:

$$\begin{aligned} D^T \lambda_t + K^T F^T \lambda_f - \beta_t D^T (t - Du) \\ - \beta_f K^T F^T (f - F(Ku - g)) = 0 \end{aligned} \quad (3.37)$$

that is

$$\begin{aligned} & \left(\beta_t D^T D + \beta_f K^T F^T F K \right) u \\ &= -D^T \lambda_t - K^T F^T \lambda_f + \beta_t D^T t + \beta_f K^T F^T (f + Fg) \end{aligned} \quad (3.38)$$

by dividing by β_t and considering that $F^T F = I$, we have

$$\begin{aligned} & \left(D^T D + \frac{\beta_f}{\beta_t} K^T K \right) u \\ &= D^T \left(t - \frac{1}{\beta_t} \lambda_t \right) + \frac{\beta_f}{\beta_t} K^T \left(g + F^T \left(f - \frac{1}{\beta_f} \lambda_f \right) \right). \end{aligned} \quad (3.39)$$

The coefficient matrix of the linear system (3.39) does not change with iterations. Moreover, since under periodic boundary conditions for u both $D^T D$ and $K^T K$ are block circulant matrices with circulant blocks, the coefficient matrix can be diagonalized once for all by the 2D-DFT using the FFT implementation. Therefore, at each iteration the linear system (3.39) can be solved by one forward FFT and one inverse FFT, each at a cost of $O(d^2 \log d)$. For the computation of the right-hand side, we point out that the matrix-vector product $F^T x$ can be obtained by taking the real part of the inverse FFT of the vector x .

3.6 ADMM Iterative Scheme

To solve the proposed whiteness constrained minimization problem (3.2), we use the ADMM iterative scheme reported in Algorithm 1.

The overall computational cost of the algorithm is dominated by step 5 which computes the u update by solving the linear system (3.39) of dimension d^2 . However, we should notice that the cost of the steps 2, 3 and 4 is strongly reduced in our implementation thanks to the symmetry of the 2D-DFT. In particular, the dimension of the variables f, p, c is reduced from d^2 to $s \simeq \frac{d^2}{2}$. Clearly, in step 5, we then need to expand the f variable in order to cover the entire frequency range according to the symmetry of the DFT of real 2D images.

Algorithm 1 ADMM for the whiteness constrained problem (3.2)

Input: $g, K, \beta_t > 0, \beta_f > 0, \beta_p > 0, \beta_c > 0$
Output: approximate solution u^* of (1.7)
 1. *Initialize:* $u^0 = g, \lambda_t^0 = 0, \lambda_f^0 = 0, \lambda_p^0 = 0, \lambda_c^0 = 0$;
 2. **For** $k = 0, 1, 2, \dots$ *until convergence:*
 1) given u^k, λ_t^k , compute t^{k+1} according to (3.17)–(3.18)
 2) given $u^k, p^k, \lambda_f^k, \lambda_p^k$, compute f^{k+1} according to (3.25)–(3.27)
 3) given $f^{k+1}, c^k, \lambda_p^k, \lambda_c^k$, compute p^{k+1} according to (3.32)
 4) given p^{k+1}, λ_c^k , compute c^{k+1} according to (3.35)
 5) given $t^{k+1}, f^{k+1}, \lambda_t^k, \lambda_f^k$, compute u^{k+1} by solving (3.39)
 6) given $t^{k+1}, f^{k+1}, p^{k+1}, c^{k+1}, u^{k+1}$, compute $\lambda_t^{k+1}, \lambda_f^{k+1}, \lambda_p^{k+1}, \lambda_c^{k+1}$ by (3.10)

End For

4 Numerical Examples

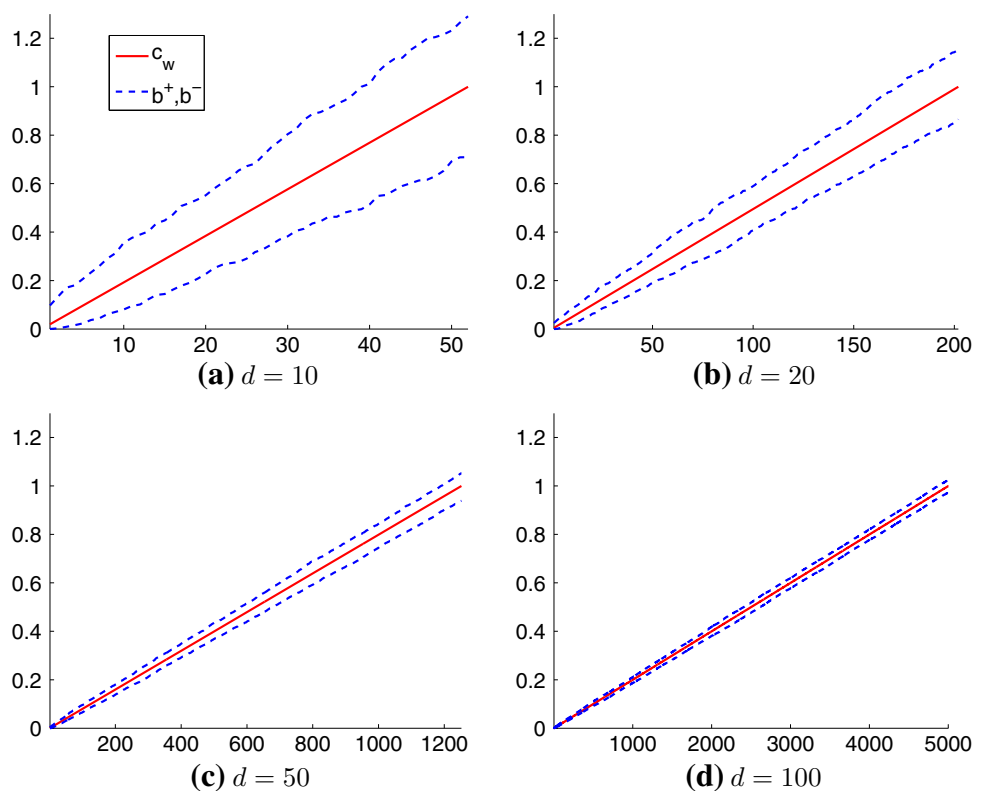
In this section we first discuss the approach we followed to select the limits b^+, b^- for the box-constraint in (2.14), then we present experimental results to test the proposed algorithm.

For the selection of the whiteness set limits $b^+, b^- \in \mathbb{R}^s$, we propose the following simple Montecarlo approach. Given a standard deviation σ for the noise and a selected image dimension d , we generate a large number of $d \times d$ images containing different realizations of a zero-mean white Gaussian noise with standard deviation σ . For each realiza-

tion, we compute the NCP vector c according to the definition in (2.11). Finally, we compute the minimum (b^-) and the maximum (b^+) values within all the realizations for all the components of the NCP vector. In Fig. 2 we show the results obtained by using four different image dimensions $d = 10, 20, 50, 100$. It is worth also pointing out that, thanks to the normalization adopted for the computation of the NCP vector, see (2.10), the upper and lower limits b^+, b^- do not depend on the standard deviation σ of the noise.

We notice how the size of the band gets smaller as the image dimension d increases. In other words, the bigger the

Fig. 2 Some results of Montecarlo simulations: in *dashed blue* we depict the computed upper and lower limits $b^+, b^- \in \mathbb{R}^s$ of the whiteness box-constraints on the NCP vector $c \in \mathbb{R}^s$ for zero-mean white Gaussian noise with standard deviation $\sigma = 4$, obtained for $d \times d$ images with dimension **a** $d = 10$, **b** $d = 20$, **c** $d = 50$, **d** $d = 100$. In *solid red* the theoretical whiteness NCP vector c_w defined in (2.12) is shown (Color figure online)



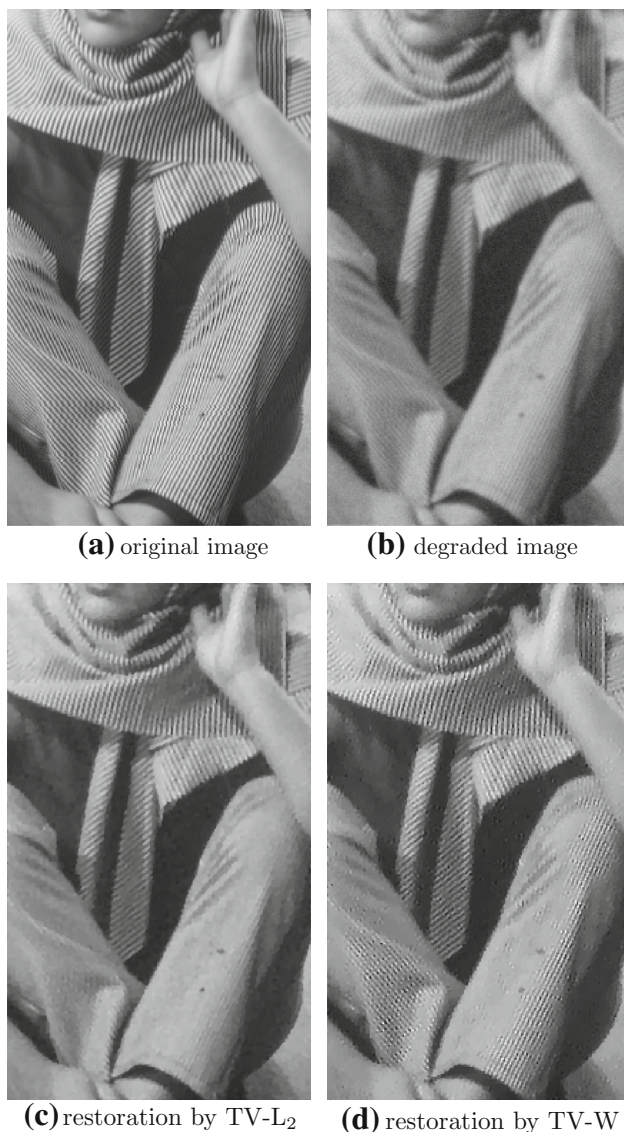


Fig. 3 Example 1. Restoration results obtained by TV-L₂ and TV-W models on the original barbara image degraded by Gaussian blur with $\text{band} = 5$, $\text{sigma} = 1.0$ and additive zero-mean white Gaussian noise with standard deviation $\sigma = 5$

image dimension d is, the narrower the band of the whiteness constraint will be. In the limit, as d tends to infinity, the NCP vector c converges to the theoretical whiteness NCP vector c_w defined in (2.12).

In the rest of this section we evaluate the performance of the proposed restoration algorithm when applied to images synthetically corrupted by blur and white Gaussian noise.

We compare the proposed algorithm, referred to as TV-W, with the well-known Rudin-Osher-Fatemi (ROF) model [16], based on the minimization of the TV-L₂ functional (1.3). The TV-L₂ approach is implemented by the ADM, a variant of the classic augmented Lagrangian method for structured optimization which reformulates a TV problem

as a linear equality constrained problem. The ADM algorithm is stable, efficient and in particular faster than most of the state-of-the-art restoration algorithms [19]. The software package is freely available at <http://www.caam.rice.edu/optimization/L1/ftvd/v4.0/> and described in detail in [19].

For the proposed TV-W model, we used the ADMM minimization procedure illustrated in Algorithm 1) which requires a number of tuning parameters to be estimated. We used the following parameters setting: $\beta_t = \beta_c = \beta_p = \beta_f = 10$, $\gamma = 1.618$, which is a pretty standard choice. In fact we observed that the restoration results are relatively insensitive to the choice of these parameters. However, for a faster convergence the β_t parameter should be tuned towards higher values.

In the extensive experiments we run to test the iterative algorithms for TV-L₂ and TV-W, we noticed that after a few iterations they both approach toward a given stable solution, thus we decided to report in this section the results of the two algorithms obtained as soon as the relative difference between two successive iterates satisfies $e^k := \|u^k - u^{k-1}\|_2 / \|u^{k-1}\|_2 < 10^{-4}$ or after a maximum of 100 iterations.

For a fair comparison between the TV-L₂ and TV-W algorithms which are implemented as constrained models, we set $\tau = 1$ both in (1.6) and in (2.10), where σ is the available standard deviation of the noise.

The accuracy of the methods is evaluated by the improved signal to noise ratio (ISNR) and blurred signal to noise ratio (BSNR), defined by

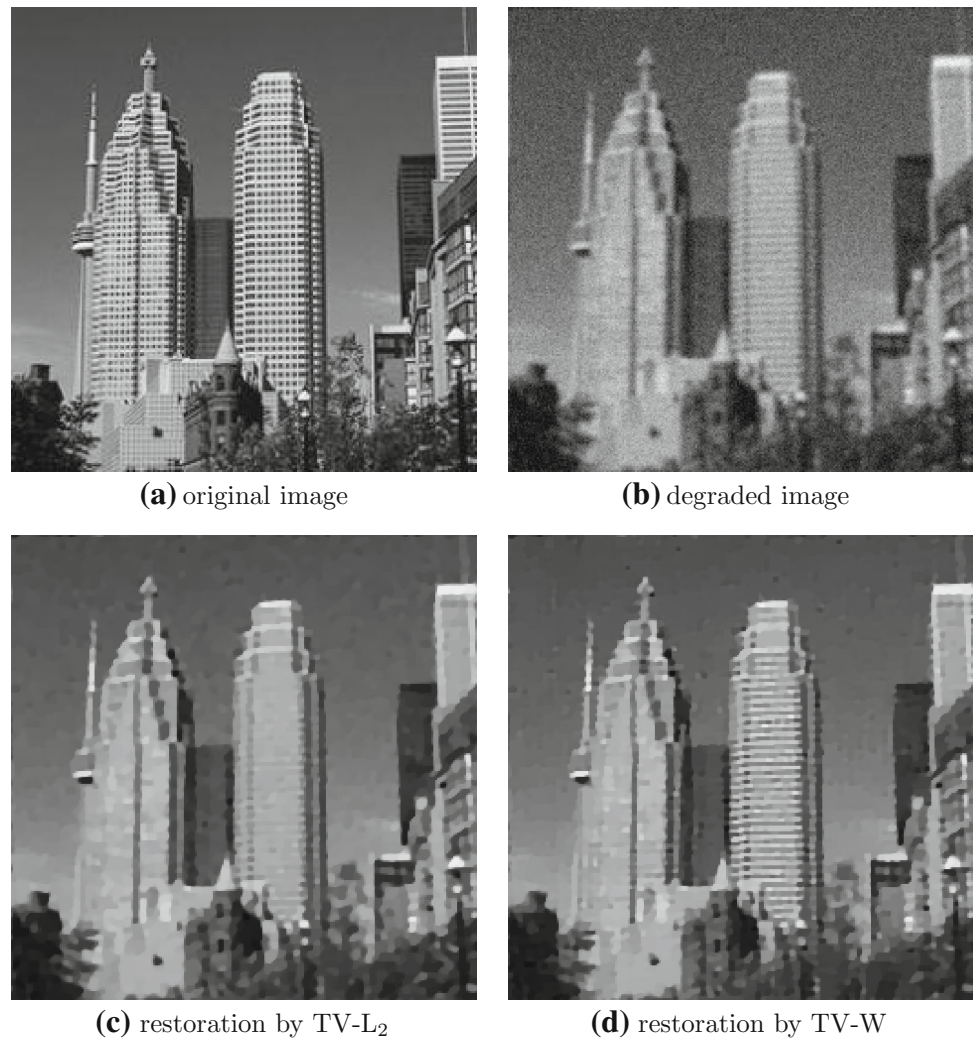
$$\begin{aligned} \text{ISNR}(u^*, u) &= 10 \log_{10} \frac{\|g - u\|_2^2}{\|u^* - u\|_2^2} \text{ dB}, \\ \text{BSNR}(u^*, u) &= 10 \log_{10} \frac{\|g\|_2^2}{\|n\|_2^2} \text{ dB}, \end{aligned} \quad (4.1)$$

where $u^* \in \mathbb{R}^{d^2}$ is the computed estimate of the uncorrupted image $u \in \mathbb{R}^{d^2}$. The ISNR quantity provides a quantitative measure of the improvement in the quality of the restored image: a high ISNR value indicates that u^* is an accurate approximation of u .

The corrupted images taken as input by the algorithms have been obtained as follows. First, the original $d \times d$ image u is blurred by a Gaussian kernel characterized by two parameters band and sigma . The former specifies the half-bandwidth of the Toeplitz blocks and the latter the variance of the Gaussian point spread function. The larger the sigma is, the more the blurring will be. The kernel is generated through the MATLAB command `fspecial('Gaussian', band, sigma)`. Then, the blurred image Ku is corrupted by additive zero-mean white Gaussian noise with standard deviation σ .

Example 1. We consider the restoration of three different images: *barbara* ($d = 512$), *skyscraper* ($d = 256$) and

Fig. 4 Example 1. Restoration results obtained by TV- L_2 and TV-W models on the original skyscraper image degraded by Gaussian blur with $\text{band} = 5$, $\text{sigma} = 1.5$ and additive zero-mean white Gaussian noise with standard deviation $\sigma = 10$



geometry ($d = 256$), which present interesting mixtures of textures, flat regions and shaded areas; the noise-free versions of the images are depicted in Figs. 3a, 4a, and 5a, respectively.

In the first part of this example we investigate the comparison between the restoration by TV- L_2 and by TV-W algorithms for a fixed additive Gaussian noise with standard deviation $\sigma = 5$ and variable Gaussian blur degradations, characterized by band and sigma , which significantly worsen the ill-conditioning of the image restoration problems. Table 1 reports in the first and second columns the parameters band and sigma of the Gaussian blur kernel, and the third column, labeled BSNR, reports the BSNR-values of the available corrupted image g . The columns labeled TV- L_2 and TV-W display ISNR-values for the restored images obtained by the two related algorithms. For increasing band and sigma parameters, the textured parts of the images tend to disappear, consequently the benefits of the whiteness constraint decrease until the worst case with $\text{band} = 7$ and $\text{sigma} = 2.5$ where the ISNR-values for the two algorithms are almost the same.

In the second part of this example, we evaluate quantitatively the robustness of the restoration performance for a fixed Gaussian blur and increasing standard deviation σ of the additive noise when the restoration methods are applied to the same images. Table 2 shows the resulting ISNR values for σ ranging from 5 to 40. For each image we chose different blur intensities in such a way that textured parts are not completely filtered out, and noise effect can be better highlighted.

In Figs. 3, 4, 5 we display the degraded images and the results by applying the two different image restoration methods.

A visual inspection of the figures of this example together with the results in the Tables 1, 2 allow us to conclude that by constraining the whiteness of the residual, we can go beyond the use of just the residual norm or the variance of the noise model, thus providing better restorations. In all the tables the highest ISNR values are in boldface.

A proof of theoretical convergence for the ADMM algorithm proposed for the TV-W model is beyond the scope

Fig. 5 Example 1. Restoration results obtained by TV- L_2 and TV-W models on the original geometry image degraded by Gaussian blur with $\text{band} = 7$, $\text{sigma} = 1.5$ and additive zero-mean white Gaussian noise with standard deviation $\sigma = 10$

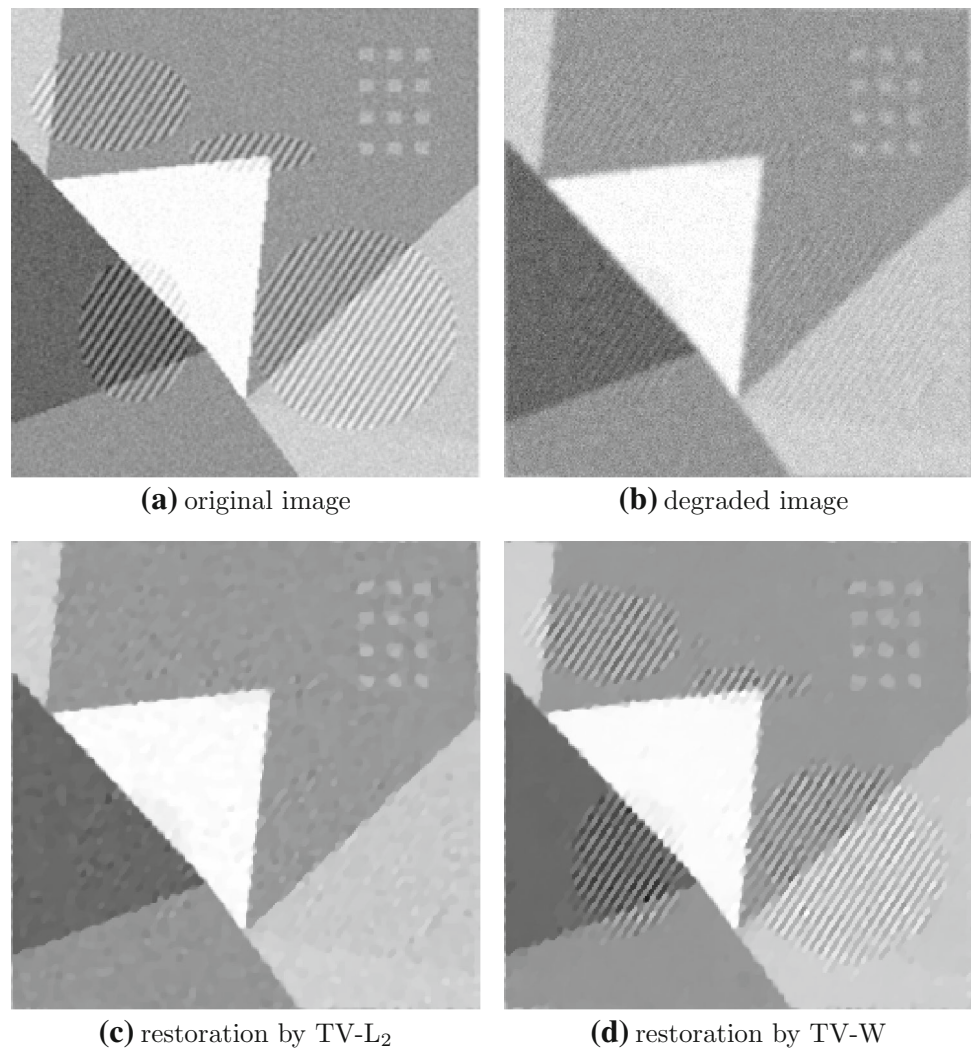


Table 1 Example 1. ISNR values obtained by restoring the images barbara, skyscraper and geometry corrupted by additive zero-mean white Gaussian noise with fixed standard deviation $\sigma = 5$ and Gaussian blur with increasing strength

band	sigma	barbara			skyscraper			geometry		
		BSNR	TV- L_2	TV-W	BSNR	TV- L_2	TV-W	BSNR	TV- L_2	TV-W
7	2.5	19.09	0.39	0.40	18.01	1.65	2.18	19.47	0.93	0.93
7	1.2	19.29	-0.04	0.55	18.42	1.88	2.83	19.63	3.29	3.96
5	1.0	19.37	0.31	1.22	18.64	2.43	3.06	19.73	2.84	3.81
5	0.7	19.57	1.55	3.06	19.04	2.80	2.98	19.88	1.53	2.58
3	1.0	19.48	0.69	1.66	18.91	2.36	2.76	19.84	1.94	2.69
3	0.5	19.95	1.46	2.59	19.51	1.81	2.06	20.04	0.32	0.92

Table 2 Example 1. ISNR values obtained by restoring the images barbara, skyscraper and geometry corrupted by fixed Gaussian blur and additive zero-mean white Gaussian noise with increasing standard deviation σ

barbara (band = 3, sigma = 0.5)				skyscraper (band = 5, sigma = 1.5)				geometry (band = 7, sigma = 1.5)			
σ	BSNR	TV- L_2	TV-W	σ	BSNR	TV- L_2	TV-W	σ	BSNR	TV- L_2	TV-W
5	19.95	1.46	2.59	5	18.37	1.14	1.73	5	19.56	2.4	4.15
10	13.93	-0.28	0.99	10	12.35	0.63	1.35	10	13.54	1.03	4.24
20	7.91	0.74	1.38	20	6.33	1.66	1.96	20	7.52	3.36	5.13
40	1.88	4.36	4.47	40	0.31	4.56	4.58	40	1.50	7.71	7.74

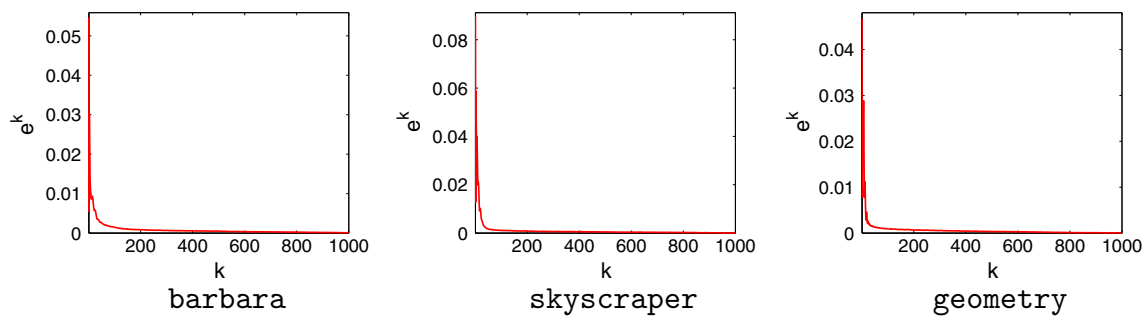


Fig. 6 Plot of the relative change of the approximate solution e^k computed by the proposed TV-W algorithm as a function of iterations count k for the example reported in the third row of Table 1 ($\text{band} = 5, \text{sigma} = 1.0$)

of this paper, and will be investigated in future work, since the convergence of a several block variables ADMM is still an open problem to the best of our knowledge. However, in Fig. 6 we provide empirical evidence of the numerical convergence of the proposed algorithm for the example reported in the third row of Table 1 ($\text{band} = 5, \text{sigma} = 1.0$). In particular, in Fig. 6 the relative change of the approximate solution computed as e_k is shown as a function of the iteration count k . The same convergent behavior can be observed for all the other tested examples.

Example 2. In this example we show the good performance of Algorithm 1) with whiteness constraints in recovering perturbed textured images with increasing size of the texture pattern. In [4] a strategy for the restoration of textured images has been proposed, where a texture detection preliminary procedure was required to set the fractional order of the regularization. Unlike, in TV-W method, the textured parts are automatically handled by the constraint in the variational formulation.

Synthetic 200×200 checkboard images are created with increasing pattern size as depicted in Fig. 7 (first column). The checkboard characterized by the finest pattern size can be considered as a classical textured image, while for increasing pattern size any automatic texture detection algorithm could even not succeed in classifying them as textured images. Therefore, all those algorithms that take advantage of a preliminary detection of textured regions to improve the restoration would fail in these cases. Moreover, even for the checkboard image characterized by the largest pattern size which is notoriously well recovered by TV- L_2 algorithm, TV-W algorithm provides significantly better results.

The checkboard images corrupted by white Gaussian noise with $\sigma = 20$ and Gaussian blur with $\text{band} = 7$ and $\text{sigma} = 1.5$ are displayed in Fig. 7 (second column), the restored images obtained by the TV- L_2 and TV-W algorithms are depicted in Fig. 7 third and fourth columns, respectively.

In Table 3 we report the results of the restorations for images depicted in Fig. 7 (second column). The initial BSNR values (labeled as $BSNR$) and those obtained by each

algorithm for different noise levels ($\sigma = 10, 20$), are reported. The ISNR values reported in Table 3 show that TV-W algorithm significantly improves the restoration quality compared to TV- L_2 , and the improvement increases for large pattern size.

As previously stated, in all the experiments reported, we used the parameter value $\tau = 1$ for both approaches, that is we imposed that the variance of the residue image is equal to the true noise variance.

However, it is known that using the discrepancy principle with $\tau = 1$ in the TV- L_2 model tends to lead to over-regularized solutions. Therefore the value of the τ parameter is usually adjusted for a better performance.

In Table 4 the examples shown in Table 3 has been rerun, where the τ parameters of the both approaches have been tuned. The corresponding τ values are reported in columns fourth and sixth for the TV- L_2 and TV-W, respectively.

We observe that the proposed model leads to better results than the TV- L_2 model after parameter tuning of both approaches, moreover it's worth noticing that the optimal τ values are less than 1 for TV- L_2 as expected, while they are greater than 1 in our model.

5 Conclusions

We have presented a new variational model for the restoration of images corrupted by blur and zero-mean white Gaussian noise. The proposed model takes advantage of residual components, rather than only the norm, to determine when the approximate solution is adequately separated from noise. In particular, the whiteness of the residual image has been integrated in the model as a new constraint on the solution. The constrained TV-W model has been efficiently solved by combining the effective ADMM method with an accurately chosen variable splitting strategy. Experimental comparison with the classical TV- L_2 model for image restoration demonstrates the effectiveness of the proposed model in particular for textured images. However, we notice that the present implementation is not optimized as the downloaded code for

Fig. 7 Restoration results obtained by TV-L₂ (third column) and TV-W (fourth column) models on the degraded images shown in the second column, obtained by corrupting the original images in the first column by a Gaussian blur with $\text{band} = 7$, $\text{sigma} = 1.5$, and additive white Gaussian noise with zero-mean and $\sigma = 20$

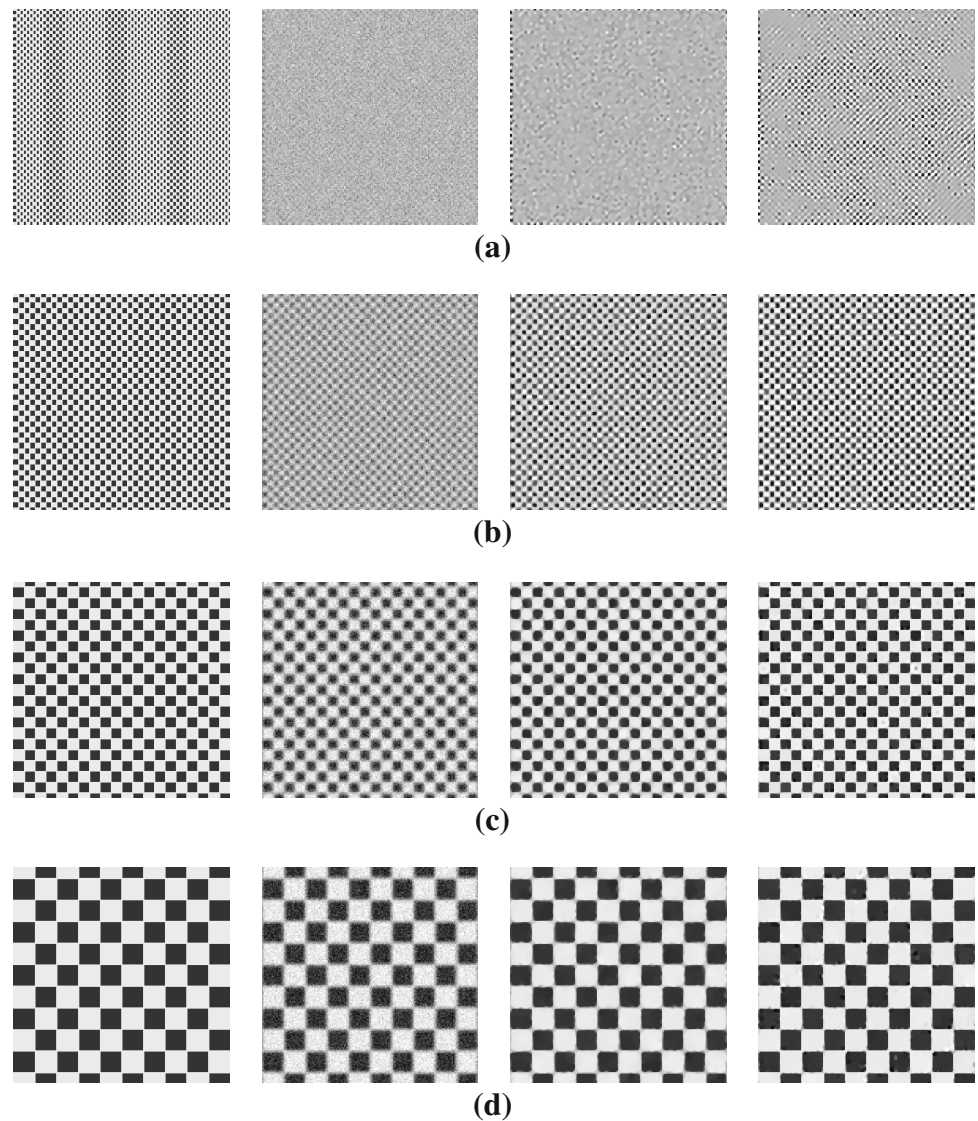


Table 3 Example 2. ISNR values obtained by restoring the synthetic checkerboard image corrupted by additive zero-mean white Gaussian noise of standard deviation σ for different checkerboard-sized images

	σ	BSNR	TV-L ₂	TV-W
Fig.7a	10	-1.88	1.56	3.27
	20	-7.90	-0.03	2.28
Fig.7b	10	11.09	3.87	4.98
	20	5.07	2.84	4.46
Fig.7c	10	16.59	6.75	13.11
	20	10.57	3.73	10.34
Fig.7d	10	18.46	7.66	13.54
	20	12.44	5.31	9.21

Table 4 Example 2. ISNR values obtained by restoring the synthetic checkerboard image corrupted by additive zero-mean white Gaussian noise of standard deviation σ for different checkerboard-sized images, using properly tuned parameter values τ

	σ	BSNR	τ	TV-L ₂	τ	TV-W
Fig.7a	10	-1.88	0.92	2.30	1.03	4.84
	20	-7.90	0.92	0.89	1.04	3.38
Fig.7b	10	11.09	0.96	4.22	1.01	5.78
	20	5.07	0.96	3.05	1.03	5.34
Fig.7c	10	16.59	0.95	7.46	1.02	13.79
	20	10.57	0.96	5.12	1.01	10.78
Fig.7d	10	18.46	0.97	9.36	1.00	13.54
	20	12.44	0.98	7.16	1.02	10.48

TV-L₂ is, therefore the TV-W presents an overhead with respect to the TV-L₂ between 30% and 50%. The method can also be used in conjunction with other fidelity terms or differ-

ent regularization operators provided that the noise is not correlated. Future work will investigate how this could improve

the overall performance. Moreover, the present approach could be easily extended to non-white noise perturbations in which the periodogram curve is known or can be inferred, without any particular modifications of the algorithm.

Appendix

Proof of Proposition (3.1). First, the problem in (3.11) can be equivalently rewritten as follows:

$$x^* \leftarrow \arg \min_{x \in \mathbb{R}^n} \left\{ \varphi(x) = \alpha \|x\|_2 + \frac{\beta}{2} \|x - q\|_2^2 \right\}, \quad (5.1)$$

with $q \in \mathbb{R}^n$ defined as in (3.13). In Fig. 8a we give a geometric representation of the problem in the 2-dimensional case. To prove the proposition statement in (3.12) we consider separately the two cases $\|q\|_2 > 0$ and $\|q\|_2 = 0$.

Case $\|q\|_2 > 0$. First, we prove that the solution x^* of (5.1) lies on the half-line Oq with origin at the n -dimensional null vector O and passing through q , represented in solid red in Fig. 8a. To this purpose, we demonstrate that for every point p not lying on Oq there always exists a point p^* on Oq providing a lower value of the objective function in (5.1), that is a point p^* such that $\varphi(p) - \varphi(p^*) > 0$. In particular, we define p^* as the intersection point between the half-line Oq and the n -dimensional sphere with center O and passing through p , depicted in solid blue in Fig. 8a. Noting that $\|p^*\|_2 = \|p\|_2$ by construction, we can thus write:

$$\begin{aligned} \varphi(p) - \varphi(p^*) &= \alpha (\|p\|_2 - \|p^*\|_2) \\ &+ \frac{\beta}{2} (\|p - q\|_2^2 - \|p^* - q\|_2^2) \\ &= \frac{\beta}{2} (\|p\|_2^2 + \|q\|_2^2 - 2 \langle p, q \rangle - \|p^*\|_2^2 - \|q\|_2^2 \\ &\quad + 2 \langle p^*, q \rangle) \\ &= \beta \langle p^* - p, q \rangle \\ &= \beta \|p^* - p\|_2 \|q\|_2 \cos(\widehat{Op^*p}). \end{aligned} \quad (5.2)$$

Since $\beta > 0$, $\|q\|_2 > 0$ for hypothesis, $p \neq p^*$ by construction, and noting that the angle $\widehat{Op^*p}$ is always acute, we can conclude that the expression in (5.2) is positive. Hence, the solution x^* of (5.1) lies on the half-line Oq , i.e. $x^* = \xi^*q$, $\xi^* \geq 0$.

By setting $x = \xi q$, $\xi \geq 0$, the n -dimensional minimization problem in (5.1) can be transformed into the following 1-dimensional problem:

$$\begin{aligned} \xi^* &\leftarrow \arg \min_{\xi \geq 0} \left\{ \alpha \|\xi q\|_2 + \frac{\beta}{2} \|\xi q - q\|_2^2 \right\} \\ &\leftarrow \arg \min_{\xi \geq 0} \left\{ \alpha \|q\|_2 \xi + \frac{\beta}{2} \|q\|_2^2 (\xi - 1)^2 \right\} \end{aligned}$$

$$\begin{aligned} &\leftarrow \arg \min_{\xi \geq 0} \left\{ \alpha \xi + \frac{\beta}{2} \|q\|_2 (\xi - 1)^2 \right\} \\ &\leftarrow \arg \min_{\xi \geq 0} \left\{ g(\xi) = \left(\frac{\beta}{2} \|q\|_2 \right) \xi^2 + (\alpha - \beta \|q\|_2) \xi \right\}. \end{aligned} \quad (5.3)$$

Since the coefficient $\frac{\beta}{2} \|q\|_2$ is strictly positive (in fact, we are assuming that $\beta > 0$ and $\|q\|_2 > 0$), the function $g(\xi)$ in (5.3) represents a strictly convex parabola passing through the origin and having its unconstrained minimum at the abscissa of its vertex:

$$\xi_v = 1 - \frac{\alpha}{\beta \|q\|_2}. \quad (5.4)$$

Hence, for what concern the constrained minimization in (5.3) we have the two cases illustrated in Fig. 8b. In formulas:

$$\xi^* = \begin{cases} \xi_v & \text{if } \xi_v \geq 0 \\ 0 & \text{if } \xi_v < 0 \end{cases}, \quad \text{that is: } \xi^* = \max\{\xi_v, 0\}. \quad (5.5)$$

Therefore, from (5.4) and (5.5) the solution $x^* = \xi^*q$ of the n -dimensional problem in (5.1) is:

$$x^* = \xi^*q = \max \left\{ 1 - \frac{\alpha}{\beta \|q\|_2}, 0 \right\} q, \quad (5.6)$$

thus proving case a) of the proposition statement in (3.12).

Case $\|q\|_2 = 0$. Since q is the n -dimensional null vector, the objective function minimized in (5.1) depends on x only through its norm $\|x\|_2$. Hence, the minimizers will be all the vectors x^* belonging to the n -dimensional sphere $\|x\|_2 = r^*$ with radius $r^* \geq 0$ given by the solution of the constrained 1-dimensional minimization problem obtained from (5.1) by setting $q = 0$ and $r = \|x\|_2$, that is

$$r^* = \arg \min_{r \geq 0} \left\{ \alpha r + \frac{\beta}{2} r^2 \right\}. \quad (5.7)$$

The minimization problem in (5.7) is very similar to problem (5.3) and the solution can be analogously computed as follows:

$$r^* = \max \left\{ -\frac{\alpha}{\beta}, 0 \right\}. \quad (5.8)$$

Therefore, if $\|q\|_2 = 0$ the solution of (5.1) is given by:

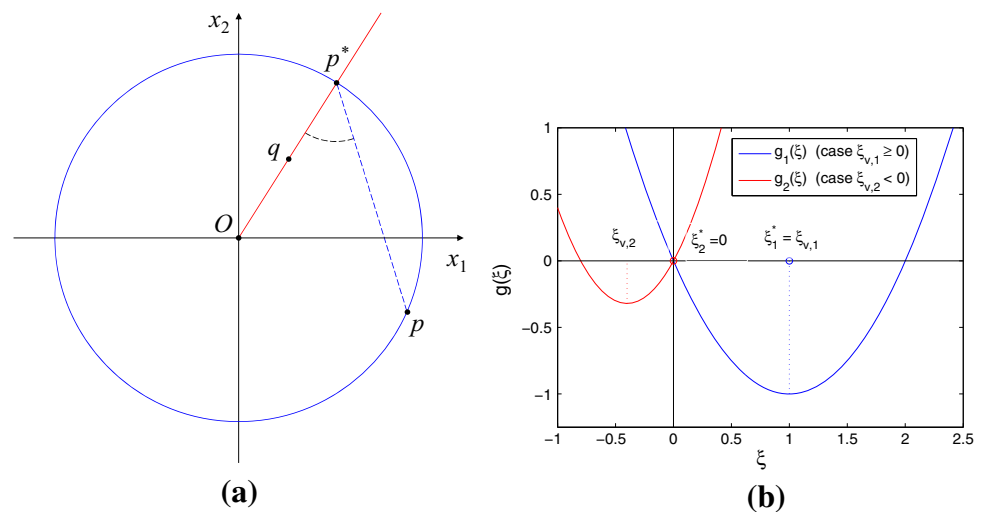
$$x^* = 0 \quad \text{if} \quad \alpha \geq 0, \quad (5.9)$$

while we get an infinite number of solutions

$$x^* \in \{x \in \mathbb{R}^n : \|x\|_2 = r^*\} \quad \text{if} \quad \alpha < 0, \quad (5.10)$$

with r^* defined in (5.8). This proves cases b) and c) of the proposition statement in (3.12). \square

Fig. 8 Geometric representation of the problem in (5.1)



References

- Almeida, M., Figueiredo, M.: Parameter estimation for blind and non-blind deblurring using residual whiteness measures. *IEEE Trans. Image Process.* **22**(7), 2751–2763 (2013)
- Bauer, F., Lukas, M.A.: Comparing parameter choice methods for regularization of ill-posed problems. *Math. Comput. Simul.* **81**(9), 1795–1841 (2011)
- Calvetti, D., Reichel, L.: Tikhonov regularization of large linear problems. *BIT Numer. Math.* **43**(2), 263–283 (2003)
- Chan, R.H., Lanza, A., Morigi, S., Sgallari, F.: An adaptive strategy for the restoration of textured images using fractional order regularization. *Numer. Math.* **6**(1), 276–296 (2013)
- Chan, R.H., Tao, M., Yuan, X.M.: Constrained total variational deblurring models and fast algorithms based on alternating direction method of multipliers. *SIAM J. Image. Sci.* **6**, 680–697 (2013)
- Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. *Comput. Math. Appl.* **2**, 17–40 (1976)
- Glowinski, R., Marrocco, A.: Sur l’approximation par éléments finis d’ordre un, et la résolution par pénalisation-dualité d’une classe de problèmes de Dirichlet non linéaires. *RAIRO Anal. Numér.* **9**, 41–76 (1975)
- Goldstein, T., Osher, S.: The split Bregman method for L1-Regularized Problems. *SIAM J. Image Sci.* **2**, 323–343 (2009)
- Gray, R.M., Davisson, L.D.: *An Introduction to Statistical Signal Processing*. Cambridge University Press, Cambridge, UK (2010)
- Hansen, P.C., Kilmer, M.E., Kjellden, R.H.: Exploiting residual information in the parameter choice for discrete ill-posed problems. *BIT Numer. Math.* **46**(1), 41–59 (2006)
- Hansen, P.C.: *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. SIAM, Philadelphia, PA, USA (1998)
- Jung, M., Bresson, X., Chan, T.F., Vese, L.A.: Nonlocal Mumford-Shah regularizers for color image restoration. *IEEE Trans. Image Process. (TIP)* **20**(6), 1583–1598 (2011)
- Jung, M., Vese, L.A.: *Image Restoration via Nonlocal Mumford-Shah Regularizers*, UCLA CAM report 09–09, (2009)
- Lanza, A., Morigi, S., Sgallari, F., Yezzi, A.J.: Variational image denoising based on autocorrelation whiteness. *SIAM J. Imaging Sci.* **6**(4), 1931–1955 (2013)
- Morini, B., Porcelli, M., Chan, R.H.: A reduced Newton method for constrained linear least-squares problems. *J. Comput. Appl. Math.* **233**, 2200–2212 (2010)
- Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D* **60**(1–4), 259–268 (1992)
- Rust, B.W., O’Leary, D. P.: Residual periodograms for choosing regularization parameters for ill-posed problems. *Inverse Probl.* **4**(3), 034005 (30pp) (2008)
- Rust, B.W.: Parameter selection for constrained solutions to ill-posed problems. *Comput. Sci. Stat.* **32**, 333–347 (2000)
- Tao, M., Yang, J.: *Alternating direction algorithms for total variation deconvolution in image reconstruction*, Tech. Rep. TR0918, Department of Mathematics, Nanjing University, China (2009)
- Wang, Y., Yang, J., Yin, W., Zhang, Y.: A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.* **1**, 948–951 (2008)
- Wen, Y., Chan, R.H.: Parameter selection for total variation based image restoration using discrepancy principle. *IEEE Trans. Image Process.* **21**(4), 1770–1781 (2012)



Alessandro Lanza received the MS degree in civil engineering from the University of Pavia, Italy, in 2000 and the European PhD degree in information technology from the Advanced Research Centre on Electronic Systems for Information and Communication Technologies (ARCES), University of Bologna, in 2007. In 2006, he spent six months at the Ecole Polytechnique Federale de Lausanne (EPFL) as a doctoral fellow. He is currently a postdoctoral fellow in the Department of Mathematics, University of Bologna.

His research interests include numerical methods for inverse problems and image processing, computer vision, pattern recognition.



Fiorella Sgallari is full professor of Numerical Analysis, Univ. Bologna. She is author of more than 100 publications. Her research concerns PDE models and numerical methods for image processing and solution of very large linear systems arising from discrete ill posed problems. She coordinated several national, international and industrial research projects. She is Associate Editor of several international journals. From 2006 to 2013 she was Chair of C. I. R. A.

M.- Research Centre of Applied Mathematics, Univ.Bologna.



Serena Morigi is a Professor of numerical analysis at the Department of Mathematics, University of Bologna, Italy. In 1996 she received the Ph.D. degree in Applied Mathematics from the University of Padova, Italy. In 1997 she did a postdoctoral fellowship at the Vanderbilt University, Tennessee, USA. Her research interests include discrete ill-posed problems and numerical PDE models for image processing. She is also interested in geometric modelling, surface

reconstruction, image processing and computer graphics. She is Associate Editor for two international journals in numerical analysis and computer graphics. She has participated in several national and international research projects.