

Automatic Parameter Selection Based on Residual Whiteness for Convex Non-convex Variational Restoration



Alessandro Lanza, Serena Morigi, and Fiorella Sgallari

Abstract Image restoration is a well-known ill-posed inverse problem whose aim is to recover a sharp clean image from the corresponding blur- and noise-corrupted observation. Variational methods penalize solutions deemed undesirable by incorporating regularization techniques. A popular strategy relies on using sparsity promoting regularizers; it is well known that, in general, nonconvex regularizers hold the potential for promoting sparsity more effectively than convex regularizers. Recently a new class of convex non-convex (CNC) variational models has been proposed which includes a general parametric nonconvex nonseparable regularizer. However, the performance of this approach depends critically on the regularization parameter. In this paper we propose to use a parametric CNC variational restoration model within a bilevel framework, where the parameter is tuned by minimizing a measure of the restoration residual whiteness. Some preliminary numerical experiments are shown which indicate the effectiveness of the proposal.

Keywords Sparsity-inducing regularization · Variational methods · Ill-posed problems · Non-convex non-smooth regularization · Optimization · Additive white gaussian noise.

Dedicated to Raymond H. Chan on the occasion of his 60th birthday.

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

1 Introduction

In this paper, we consider the problem of restoring 2-D gray-scale images corrupted by blur and additive white Gaussian noise (AWGN).

These images can be represented by the discretization of a real valued function defined on a 2-D compact rectangular domain. Let $x \in \mathbb{R}^n$, with $n = n_1 n_2$, be the unknown $n_1 \times n_2$ clean image concatenated into an n -vector, $A \in \mathbb{R}^{n \times n}$ be a known blurring operator and $\epsilon \in \mathbb{R}^n$ be an unknown realization of the noise process, which we assume white Gaussian with zero-mean and standard deviation σ . The discrete imaging model of the degradation process which relates the observed degraded image $b \in \mathbb{R}^n$ with x , can be expressed as follows:

$$b = Ax + \epsilon. \tag{1}$$

Given A and b , our goal is to solve the inverse problem of recovering an accurate estimate of x , which is known as deconvolution or deblurring. When A is the identity operator, recovering x is referred as denoising.

Image deblurring is a discrete ill-posed problem, as such further a priori assumptions on the solution can help to determine a meaningful approximation of x . Assuming the image is corrupted by AWGN, then an estimate x_λ^* of x can be obtained as a solution—i.e., a global minimizer—of the following variational model which is the sum of a convex smooth (quadratic) fidelity term and a regularization term:

$$x_\lambda^* \in \arg \min_{x \in \mathbb{R}^n} \mathcal{J}(x; \lambda), \quad \mathcal{J}(x; \lambda) := \frac{1}{2} \|Ax - b\|_2^2 + \lambda \mathcal{R}(x), \tag{2}$$

where $\|v\|_2$ denotes the ℓ_2 norm of vector v and λ represents the classical regularization parameter which controls the trade-off between data-fidelity and regularization.

The regularizer $\mathcal{R}(x)$ encodes a priori knowledge on the solution. Focusing on the recovery of images characterized by some sparsity property, we consider the general class of sparsity-inducing variational models described in [15] to determine solutions x_λ^* which are close to the data b according to the observation model and, at the same time, for which the transformed solution vector $y_\lambda^* = G(Lx_\lambda^*)$ is sparse with $L \in \mathbb{R}^{r \times n}$ a linear operator and $G: \mathbb{R}^r \rightarrow \mathbb{R}^s$ a possibly nonlinear vector-valued function—see [15].

The drawback of the proposal in [15], which will be briefly illustrated in Sect. 3, is that it requires a trial-and-error procedure for tuning the regularization parameter λ and a manual stopping. This represents a crucial aspect in variational restoration methods and has been subject of several research works.

We propose an automatic criterion for adjusting the regularization parameter λ . More precisely, our proposal is based on the key idea that if the restored image is a good estimate of the target clean image, then the residual image must resemble the realization of the noise process, thus being spectrally white.

Hence, starting from a sufficiently small λ value, we iteratively increase λ until a suitable whiteness maximality criterion is satisfied.

In Sect. 2 we review some related works on the choice of the regularization parameter. The class of CNC variational models introduced in [15] is briefly illustrated in Sect. 3. In Sect. 4 we define the residual whiteness strategy, and Sect. 5 is devoted to the description of the proposed algorithmic framework. Numerical results are presented in Sect. 6. Conclusions are drawn in Sect. 7.

2 Related Work

A crucial issue in the regularization of ill-posed inverse problems is the choice of the regularization parameter. The quality of the solution is affected by the value of λ : a too large value of λ gives an over-smoothed solution that lacks details that the desired original solution may have, while a too small value of λ yields a computed solution that is unnecessarily, and possibly severely, contaminated by propagated error that stems from the error ϵ in b .

The discrepancy principle (DP) [22] chooses the regularization parameter so that the variance of the residual equals that of the noise; the DP thus requires an accurate estimate of the noise variance and is known to yield overregularized estimates [7]. The sensitivity of λ and of the computed solution to the inaccuracies in an available estimate of $\|\epsilon\|$ has been investigated by Hamarik et al. [6], who proposed alternatives to the discrepancy principle when only a poor estimate of $\|\epsilon\|$ is known. Automatic procedures for selecting the λ parameter based on the DP has been proposed in literature, see e.g. [9].

Parameter choice methods when no estimate of $\|\epsilon\|$ is available are commonly referred to as “heuristic”, because they may fail in certain situations; see [5].

A large number of heuristic parameter choice methods have been proposed in the literature due to the importance of being able to determine a suitable value of the regularization parameter when the DP cannot be used; see, e.g., [3, 7, 22]. These methods include the L-curve criterion, and generalized cross validation [3].

These methods are outperformed by more recent criteria based on Steins unbiased risk estimate (SURE) [4, 18]. SURE provides an estimate of the mean squared error (MSE), assuming knowledge of the noise distribution and requiring an accurate estimate of its variance [23].

Recently, the fact that the additive noise is the realization of a white random process and, hence, that the restoration residual image must be uncorrelated, has been used not only as an a-posteriori performance evaluation criterion (see, e.g., [21]), but also as a key idea in the design of new fidelity terms [11, 12, 14]. 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, 8, 20, 21] 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].

3 The Class of CNC Variational Models

The considered class of CNC variational models proposed in [15] relies on a general strategy for constructing non-convex non-separable regularizers starting from any convex regularizer $\mathcal{R} : \mathbb{R}^n \rightarrow \mathbb{R}$ of the form

$$\mathcal{R}(x) := \Phi(y), \quad y := G(Lx), \quad (3)$$

with $L \in \mathbb{R}^{r \times n}$, $G: \mathbb{R}^r \rightarrow \mathbb{R}^s$ a possibly nonlinear vector-valued function with $g_i: \mathbb{R}^r \rightarrow \mathbb{R}, i = 1, \dots, s$, representing its scalar-valued components and $\Phi: \mathbb{R}^s \rightarrow \mathbb{R}$ a sparsity-promoting penalty function [10, 13, 17]. The CNC model associated with the regularizer \mathcal{R} is as follows

$$x_\lambda^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{J}_B(x; \lambda), \quad \mathcal{J}_B(x; \lambda) := \frac{1}{2} \|Ax - b\|_2^2 + \lambda \mathcal{R}_B(x), \quad (4)$$

with the parameterized non-convex non-separable regularizer \mathcal{R}_B defined by

$$\mathcal{R}_B(x) := \mathcal{R}(x) - (\mathcal{R} \square \frac{1}{2} \|B \cdot\|_2^2)(x), \quad (5)$$

where \square denotes the infimal convolution operator and $B \in \mathbb{R}^{q \times n}$ is a matrix of parameters.

According to Proposition 8 in [15], a sufficient condition for \mathcal{J}_B to be strongly convex—hence, for the variational model in (4) to admit a unique solution—is that the matrix B satisfies

$$B^\top B \prec (1/\lambda)A^\top A. \quad (6)$$

A simple yet effective strategy for constructing a matrix $B^\top B \in \mathbb{R}^{n \times n}$ satisfying the convexity condition in (6) has been presented in [15]. Since the matrix $A^\top A \in \mathbb{R}^{n \times n}$ is symmetric and positive semidefinite, it admits the eigenvalue decomposition

$$A^\top A = VEV^\top, \quad E, V \in \mathbb{R}^{n \times n}, \quad E = \operatorname{diag}(e_1, \dots, e_n), \quad V^\top V = VV^\top = I_n, \quad (7)$$

with $e_i, i = 1, \dots, n$, indicating the real non-negative eigenvalues of $A^\top A$. By setting

$$B^\top B = \frac{1}{\lambda} V \Gamma E V^\top, \quad \Gamma := \operatorname{diag}(\gamma_1, \dots, \gamma_n), \quad \gamma_i \in [0, 1) \quad \forall i \in \{1, 2, \dots, n\}, \quad (8)$$

then (6) is clearly satisfied. A special case is to set a unique parameter $\gamma = \gamma_i \in [0, 1) \forall i$, which corresponds to setting $B = \sqrt{\gamma/\lambda} A$.

In the present work, which addresses the specific problem of image restoration, we consider as a first regularization function \mathcal{R} in (3) the popular Total Variation (TV) semi-norm [19]. In this case Φ is the ℓ_1 norm function and $L = [D_h^\top, D_v^\top]^\top$, with $D_h, D_v \in \mathbb{R}^{n \times n}$ representing finite difference approximations of the first-order partial derivatives along the horizontal and vertical directions, respectively, then we have:

$$\mathcal{R}(x) = \text{TV}(x) = \|G(Lx)\|_1 = \sum_{i=1}^n |g_i(Lx)|, \quad g_i(Lx) = \sqrt{(D_h x)_i^2 + (D_v x)_i^2}. \quad (9)$$

It is well known that TV-based reconstructions favor piecewise-constant solutions, but present staircase effects in the restoration of smooth parts of the images. To avoid this artifact, in the reconstruction of piecewise-affine solutions, a second-order extensions of the TV regularizer can be considered which promotes sparsity of the Hessian Schatten norms instead of the gradient norms. That is, the sum of the Schatten p -norms of the Hessian matrices computed at every pixel of the image is minimized [16], where, we recall, the Schatten p -norm $\|M\|_{\mathcal{S}_p}$ of a matrix $M \in \mathbb{R}^{z \times z}$ is defined by

$$\|M\|_{\mathcal{S}_p} := \left(\sum_{i=1}^z \sigma_i^p(M) \right)^{\frac{1}{p}}, \quad p > 0, \quad (10)$$

with $\sigma_i(M)$ indicating the i -th singular value of matrix M . Let $L = [D_{hh}^\top, D_{vv}^\top, D_{hv}^\top]^\top$ with $D_{hh}, D_{vv}, D_{hv} \in \mathbb{R}^{n \times n}$ representing finite difference approximations of second-order derivatives along horizontal, vertical and mixed horizontal/vertical directions, respectively. Then the Hessian Schatten p -norm regularizer is defined by

$$\mathcal{R}(x) = \mathcal{S}_p H(x) = \|G(Lx)\|_1 = \sum_{i=1}^n |g_i(Lx)|, \quad g_i(Lx) = \left\| \begin{bmatrix} (D_{hh}x)_i (D_{hv}x)_i \\ (D_{vh}x)_i (D_{vv}x)_i \end{bmatrix} \right\|_{\mathcal{S}_p}. \quad (11)$$

We recall that the Schatten p -norm reduces to the nuclear norm when $p = 1$.

4 Residual Whiteness

Given a realization $\epsilon := \{\epsilon(i, j) \in \mathbb{R} : (i, j) \in \Omega\}$, $\Omega = \{1, 2, \dots, n_1\} \times \{1, 2, \dots, n_2\}$ of a 2D $n_1 \times n_2$ random noise process, that is the series of noise values corrupting the particular observed image according to the deterministic degradation model in (1), the *sample* auto-correlation of ϵ is a function a_ϵ mapping all the possible lags $(l, m) \in \Theta = \{-(n_1 - 1), \dots, n_1 - 1\} \times \{-(n_2 - 1), \dots, n_2 - 1\}$ into a scalar value given by

$$\begin{aligned}
a_\epsilon(l, m) &:= \frac{1}{n} (\epsilon \star \epsilon)_{l, m} = (\epsilon * \epsilon')_{l, m} \\
&= \frac{1}{n} \sum_{(i, j) \in \Omega} \epsilon(i, j) \epsilon(i + l, j + m), \quad (l, m) \in \Theta, \quad n = n_1 n_2, \quad (12)
\end{aligned}$$

where \star and $*$ denote the 2-D discrete correlation and convolution operators, respectively, and where $\epsilon'(i, j) = \epsilon(-i, -j)$. Clearly, for (12) being defined for all lags $(l, m) \in \Theta$, the noise realization ϵ must be padded with at least n_1 samples in the vertical direction and n_2 samples in the horizontal direction. We assume here periodic boundary conditions for ϵ , such that \star and $*$ in (12) denote circular correlation and convolution, respectively. If the noise process ϵ is white, then it is well known that the auto-correlation a_ϵ satisfies the following asymptotic property:

$$\lim_{n \rightarrow +\infty} a_\epsilon(l, m) = 0 \quad \forall (l, m) \in \Theta_0 = \Theta \setminus \{(0, 0)\}. \quad (13)$$

For noise corruptions affecting images of finite dimensions—namely, $n < +\infty$ —we can say that the auto-correlation values for all non-zero lags are small. Some important examples of distributions of additive white noises are the uniform, the Gaussian, the Laplacian and the Cauchy [14].

Clearly the nearest to the uncorrupted image is the restored image x_λ^* , the closer the residual image $r_\lambda^* = b - A x_\lambda^*$ is to the realization ϵ in (1) of a white noise process.

Our proposal is to seek for the regularization parameter value λ^* yielding the whitest restoration residual, which can be formally defined as follows:

$$\lambda^* \in \arg \min_{\lambda \in \mathbb{R}_+} \{W(\lambda) := \mathcal{W}(r_\lambda^*)\}, \quad (14)$$

with $\mathcal{W} : \mathbb{R}^n \rightarrow \mathbb{R}$ one of the two following residual whiteness measures:

$$\mathcal{W}_1(r_\lambda^*) = \frac{\sqrt{\sum_{(l, m) \in \Theta_0} (a_{r_\lambda^*}(l, m))^2}}{a_{r_\lambda^*}(0, 0)}, \quad \mathcal{W}_2(r_\lambda^*) = \frac{\max_{(l, m) \in \Theta_0} |a_{r_\lambda^*}(l, m)|}{a_{r_\lambda^*}(0, 0)}. \quad (15)$$

We notice that, according to definition (12), the term $a_{r_\lambda^*}(0, 0)$ represents nothing else than the sample variance of the residual image r_λ^* .

5 The Proposed Algorithmic Framework

The proposed bilevel framework consists of an iterative procedure for computing an approximate solution x_λ^* of the class of CNC models proposed in [15] and defined in (4), or also of the associated purely convex models in (2), with \mathcal{R} any sparsity-promoting convex regularizer of the form in (3) and λ^* satisfying the whiteness maximality criterion in (14).

In Algorithm 1 we report the main computational steps of the overall proposed bilevel framework for image restoration.

The algorithm starts with a sufficiently small value of the parameter λ yielding a large value of the residual whiteness measure W in (14)–(15); at each iteration λ is increased by a multiplicative factor $\theta > 1$ to strengthen the effect of the regularization term. The iterative procedure is terminated as soon as the residual whiteness measure stops decreasing, for a certain λ^* value. Such a scheme relies on the assumption that the residual whiteness function $W(\lambda)$ in (14) is monotonically decreasing on the λ interval between 0 and the function minimizer λ^* . This property of the whiteness function $W(\lambda)$ is very hard to be proved theoretically but we verified it empirically and the evidence of such behavior is reported in Sect. 6.

At each (outer) iteration h of Algorithm 1, the restored image $x^{(h)}$ is computed—that is, the corresponding optimization problem is solved—by using the Primal-Dual Forward-Backward (PDFB) algorithm described in [15] for the CNC models and the Alternating Direction Method of Multipliers (ADMM) for the associated purely convex models. We remark that, for any given λ value, the considered variational models are strongly convex—hence they admit a unique global minimizer—and the PDFB and ADMM minimization algorithms are guaranteed to converge towards such minimizer.

We adopt for efficiency purposes the so called warm-starting strategy to initialize the algorithm at the next inner optimization step using the estimated values at the previous step.

The considered CNC variational approach requires the design of a matrix B satisfying the convexity condition (6) for the functional \mathcal{J}_B . Many such matrices B exist, see [15]. In the following experiments, we set Γ to be a two-dimensional dc-notch filter defined by $\Gamma = I - H$ where H is a two-dimensional low-pass filter with a dc-gain of unity and $H \leq I$. In our experiments, we set $\gamma = 0.98$ and $H = H_0^T H_0$ where H_0 is the most basic two-dimensional low-pass filter: the moving-average filter with square support.

6 Numerical Examples

In this section, we report some experimental results aimed at assessing the effectiveness of the automatic parameter selection procedure illustrated in Sects. 4 and 5 for image restoration by using the CNC variational models recently proposed in [15] and briefly outlined in Sect. 3.

We consider the three test images shown in the first row of Fig. 1: `qrcode` which belongs to the class of piecewise constant images, `roof` which is a piecewise affine image, and the popular photographic `cameraman` image. The test images have been synthetically corrupted by space-invariant Gaussian blur generated by the Matlab command `fspecial('gaussian',band,sigma)` with parameters $(\text{band}, \text{sigma}) = (5, 1.5)$, and AWGN of standard deviation $\sigma = 40$, so as to obtain the three degraded images shown in the second row of Fig. 1. The `qrcode` and `roof` images are

Algorithm 1 Bilevel Framework based on Residual Whiteness

inputs: degraded image $b \in \mathbb{R}^n$, blur operator $A \in \mathbb{R}^{n \times n}$
outputs: regularization parameter $\lambda^* > 0$, restored image $x_{\lambda^*}^* \in \mathbb{R}^n$
parameters: $\lambda_{min} > 0, \theta > 1$

initialization: $h = 0, \lambda^{(h)} = \lambda_{min}, W^{(h)} = +\infty$

repeat

- **update iteration counter and regularization parameter:**
 - $h = h + 1, \lambda^{(h)} = \theta \lambda^{(h-1)}$
- **compute restored image by solving the optimization problem:**
 - $x^{(h)} = \arg \min_{x \in \mathbb{R}^n} \mathcal{J}_B(x; \lambda^{(h)})$ (or $\mathcal{J}(x; \lambda^{(h)})$)
- **compute residual whiteness:**
 - $r^{(h)} = b - Ax^{(h)}$, and $W^{(h)} = \mathcal{W}(r^{(h)})$, by (15)

until $W^{(h)} > W^{(h-1)}$

$\lambda^* = \lambda^{(h-1)}, x_{\lambda^*}^* = x^{(h-1)}$



Fig. 1 Original (first row) and degraded (second row) test images `qrCode` (left column), `roof` (center column) and `cameraman` (right column)

characterized by very sparse first- and second-order derivatives, respectively, hence the convex TV and Schatten 1-norm regularization terms in (9)–(11) and their non-convex non-separable counterparts defined according to (5) are suitable to get good restorations.

Hence, in the experiments we perform restoration by using the proposed bilevel framework outlined in Algorithm 1 applied to the two purely convex variational models

$$\text{TV-}\ell_2: x_\lambda^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \lambda \text{TV}(x) \right\}, \quad (16)$$

$$\mathcal{S}_1H\text{-}\ell_2: x_\lambda^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \lambda \mathcal{S}_1H(x) \right\}, \quad (17)$$

and the two associated CNC counterparts

$$\text{CNC-TV-}\ell_2: x_\lambda^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \lambda \left(\text{TV} - \text{TV} \square \frac{1}{2} \|B \cdot\|_2^2 \right)(x) \right\}, \quad (18)$$

$$\text{CNC-}\mathcal{S}_1H\text{-}\ell_2: x_\lambda^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \lambda \left(\mathcal{S}_1H - \mathcal{S}_1H \square \frac{1}{2} \|B \cdot\|_2^2 \right)(x) \right\}. \quad (19)$$

For all the tests, i.e. for all images and all restoration models, the bilevel framework outlined in Algorithm 1 has been used in order to automatically select the regularization parameter λ^* yielding the whitest restoration residual according to both the whiteness measures $\mathcal{W}_1, \mathcal{W}_2$ defined in (15); we denote by λ_1^*, λ_2^* such two optimal values and by $W_1^* := \mathcal{W}_1(\lambda_1^*), W_2^* := \mathcal{W}_2(\lambda_2^*)$ the associated (minimum) whiteness measure values.

The (inner) iterations of the minimization algorithms used to determine the restored image for any given λ value—namely, ADMM for the $\text{TV-}\ell_2$ and $\mathcal{S}_1H\text{-}\ell_2$ models, PDFB for the $\text{CNC-TV-}\ell_2$ and $\text{CNC-}\mathcal{S}_1H\text{-}\ell_2$ models—are terminated as soon as two successive iterates satisfy

$$\frac{\|x^{(h)} - x^{(h-1)}\|_2}{\|x^{(h-1)}\|_2} < 10^{-5}. \quad (20)$$

The quality of the obtained restorations is evaluated by means of both the Signal-to-Noise Ratio (SNR) and the Structural Similarity Index (SSIM). We indicate by $\text{SNR}_1^*, \text{SNR}_2^*$ and $\text{SSIM}_1^*, \text{SSIM}_2^*$ the SNR and SSIM values of the restored images associated with the optimal values λ_1^*, λ_2^* . In order to quantitatively evaluate the ability of the proposed approach in automatically selecting λ values yielding restored images of good quality, we also introduce—and compute for each test—the following quantities:

$$\text{LQ}_j^* = 100 \frac{\bar{Q} - Q_j^*}{\bar{Q}}, \quad Q \in \{\text{SNR}, \text{SSIM}\}, \quad j \in \{1, 2\}, \quad (21)$$

where \bar{Q} denotes the maximum value of the quality measure - SNR or SSIM—achievable by letting λ vary in its domain. These quantities represent the loss of

restoration quality (in percentage) yielded by the proposed automatic selection procedure with respect to the maximum achievable.

In Table 1 we report all the obtained quantitative results, whereas in Figs. 2, 3, 4 and 5 we show some visual and graphical results related to the restoration of the cameraman test image by the four variational models considered. In particular, the results obtained by using the whiteness measure \mathcal{W}_1 are in Figs. 2 and 3, by whiteness measure \mathcal{W}_2 in Figs. 4 and 5. Each column in these figures corresponds to a different restoration model. In the third, fourth and fifth row we report the plots of the SNR and SSIM values of the restored image and the plots of the whiteness measure of the restoration residual as functions of the regularization parameter λ , respectively. The dashed vertical red lines indicate the “optimal” regularization parameter values, namely those yielding the smallest residual whiteness measures. It is worth noticing that for all reported tests the residual whiteness measure function $W(\lambda)$ with both the choices of \mathcal{W} introduced in (15)—shown in the last row of Figs. 2, 3, 4 and 5—exhibit a monotonically decreasing behavior on the λ interval between 0 and the functions minimizer λ^* .

In the first and second row of Figs. 2, 3, 4 and 5 we show the restored images obtained by using such optimal λ values and the associated absolute error images, respectively.

In Table 1 the best results are marked in boldface. The results obtained by hand-tuning λ (labeled as \bar{Q}) indicate that, as expected, TV-based models perform better on the piecewise constant images `qrcode` and `cameraman` whereas \mathcal{S}_1H -based models outperform TV-based models on the piecewise affine image `roof`. More precisely, the CNC models perform better than their associated purely convex counterparts. This is due to the stronger sparsity-promoting effect produced by non-convex regularization.

For what regards the optimal residual whiteness measures W_1^* and W_2^* reported in the last two columns of Table 1, it is worth observing that the lowest results (in boldface) are obtained in correspondence of the best performing models for each restoration test. This in principle should allow to use the proposed automatic parameter selection strategy in order to automatically select the best regularization term for each problem.

Finally, for any given model, the proposed automatic parameter selection strategy seems to perform very well as indicated by the small values of the quality losses LQ_j^* , $j = 1, 2$, reported in Table 1 and visually supported by the plots in the figures.

Visual inspection and comparison of the restored images are consistent with the results in Table 1.

Table 1 SNR/SSIM results obtained by restoring the test images `qr`code, `roof` and `cameraman`

qr code		\bar{Q}	Q_1^*	Q_2^*	LQ_1^*	LQ_2^*	W_1^*	W_2^*
Q = SNR	TV- ℓ_2	12.857	12.778	12.682	0.6	1.4	0.3982	1.7928
	CNC-TV- ℓ_2	14.262	13.708	13.631	3.9	4.4	0.3947	1.6594
	S_1H - ℓ_2	9.052	8.836	8.520	2.4	5.9	0.4242	4.0721
	CNC- S_1H - ℓ_2	8.994	8.847	8.514	1.6	5.3	0.4245	4.2746
Q = SSIM	TV- ℓ_2	0.803	0.803	0.801	0.0	0.3	0.3982	1.7928
	CNC-TV- ℓ_2	0.869	0.850	0.849	2.2	2.3	0.3947	1.6594
	S_1H - ℓ_2	0.565	0.565	0.561	0.0	0.6	0.4242	4.0721
	CNC- S_1H - ℓ_2	0.558	0.557	0.556	0.0	0.4	0.4245	4.2746
roof		\bar{Q}	Q_1^*	Q_2^*	LQ_1^*	LQ_2^*	W_1^*	W_2^*
Q = SNR	TV- ℓ_2	22.927	22.599	22.697	1.4	1.0	0.5096	2.2322
	CNC-TV- ℓ_2	22.658	22.658	22.526	0.0	0.6	0.5100	2.2410
	S_1H - ℓ_2	39.567	39.199	39.514	0.9	0.1	0.5025	2.1634
	CNC- S_1H - ℓ_2	42.171	41.720	41.906	1.1	0.6	0.5002	2.1481
Q = SSIM	TV- ℓ_2	0.919	0.896	0.916	2.5	0.3	0.5096	2.2322
	CNC-TV- ℓ_2	0.913	0.902	0.910	1.2	0.3	0.5100	2.2410
	S_1H - ℓ_2	0.999	0.999	0.999	0.0	0.0	0.5025	2.1634
	CNC- S_1H - ℓ_2	0.999	0.999	0.999	0.0	0.0	0.5002	2.1481
cameraman		\bar{Q}	Q_1^*	Q_2^*	LQ_1^*	LQ_2^*	W_1^*	W_2^*
Q = SNR	TV- ℓ_2	11.009	10.712	10.808	2.7	1.8	0.3950	1.6809
	CNC-TV- ℓ_2	11.125	10.780	10.974	3.1	1.4	0.3946	1.6775
	S_1H - ℓ_2	10.445	10.006	9.938	4.2	4.9	0.3994	1.9923
	CNC- S_1H - ℓ_2	10.211	9.951	9.821	2.5	3.8	0.3992	2.0405
Q = SSIM	TV- ℓ_2	0.701	0.699	0.701	0.2	0.0	0.3950	1.6809
	CNC-TV- ℓ_2	0.707	0.699	0.707	1.1	0.0	0.3946	1.6775
	S_1H - ℓ_2	0.663	0.663	0.663	0.0	0.0	0.3994	1.9923
	CNC- S_1H - ℓ_2	0.650	0.648	0.650	0.3	0.0	0.3992	2.0405

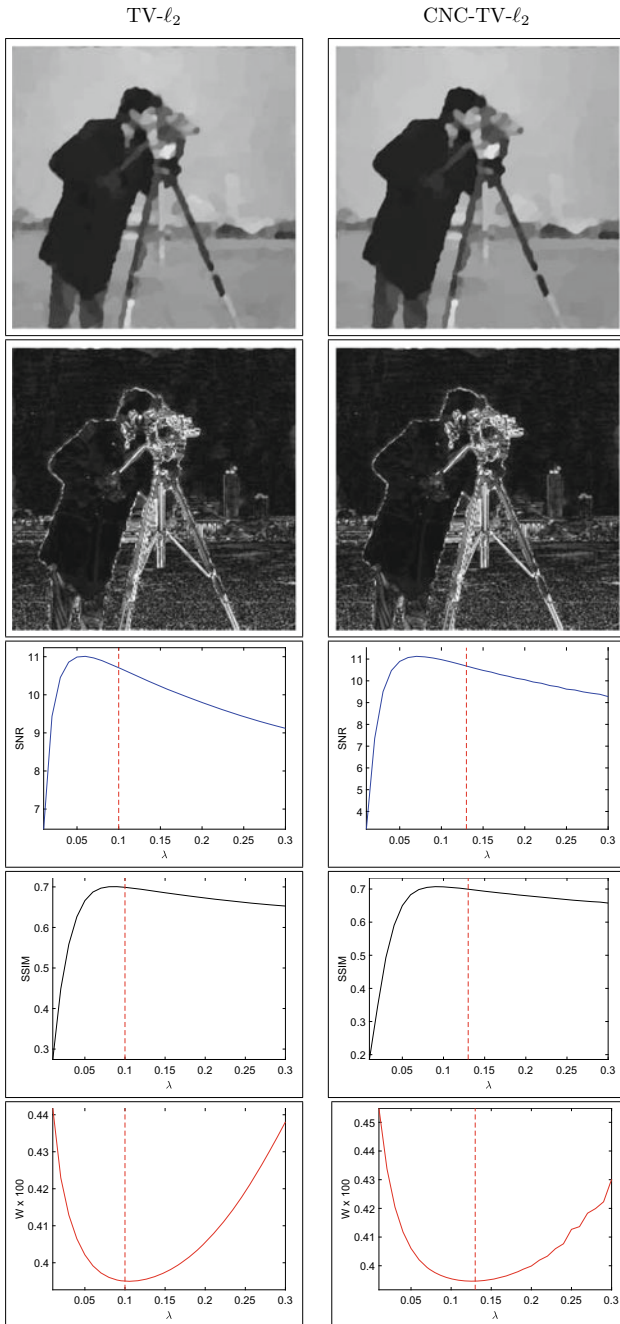


Fig. 2 Visual/graphical results obtained by using the \mathcal{W}_1 residual whiteness measure

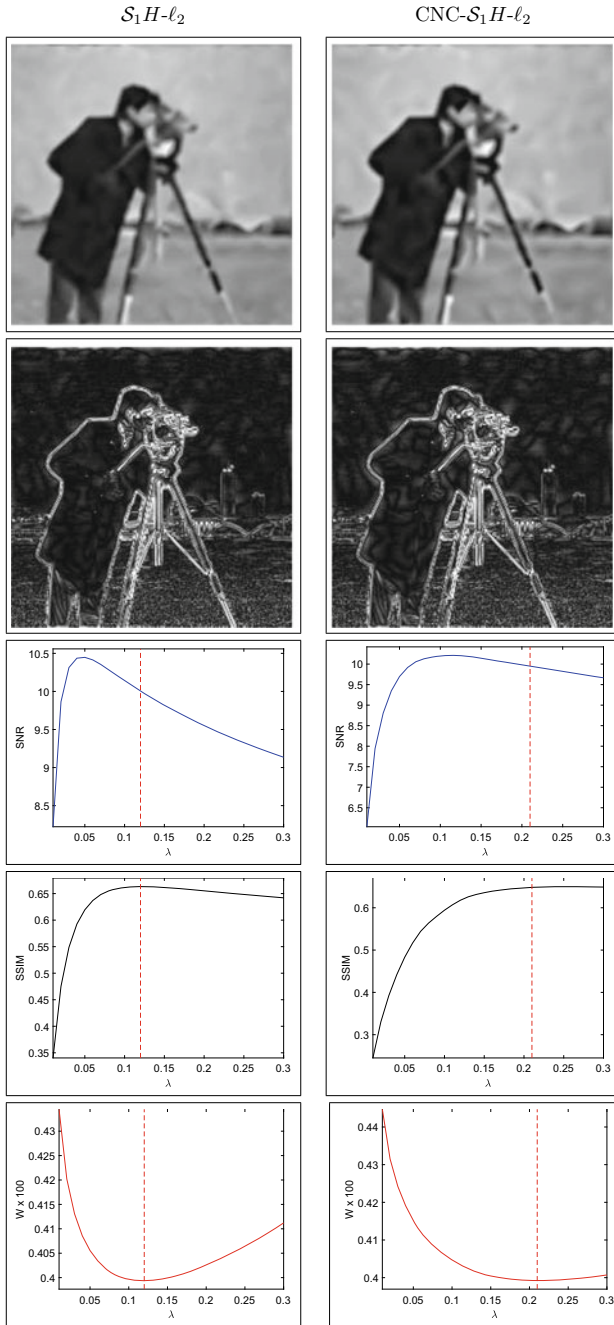


Fig. 3 Visual/graphical results obtained by using the \mathcal{W}_1 residual whiteness measure

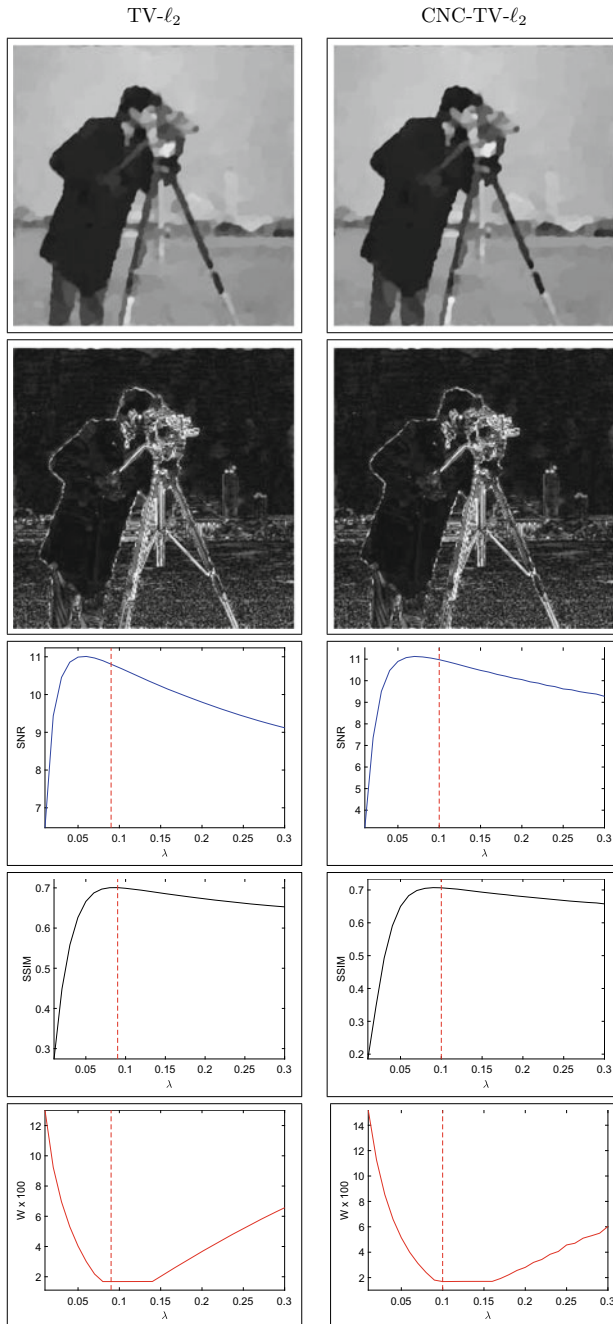


Fig. 4 Visual/graphical results obtained by using the \mathcal{W}_2 residual whiteness measure

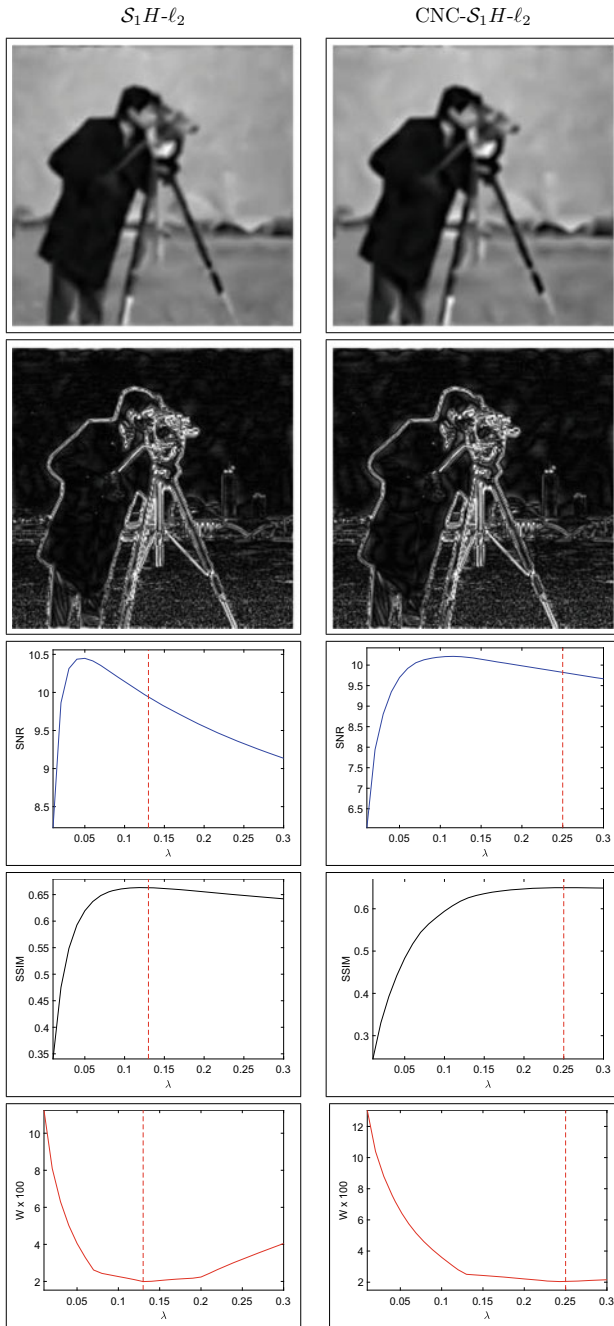


Fig. 5 Visual/graphical results obtained by using the \mathcal{W}_2 residual whiteness measure

7 Conclusions

We presented a bilevel framework aimed at equipping the class of CNC variational models for image restoration proposed in [15] with an effective strategy for automatically selecting the regularization parameter based on maximizing the residual whiteness. The idea behind our proposal is that if the recovered image is well estimated, the residual image is spectrally white; on the contrary a poorly restored image exhibits structured artifacts which yield spectrally colored residual images. Numerical results for restoring images characterized by some sparsity properties strongly indicate that the considered class of CNC models with the proposed automatic parameter selection strategy outperforms classical convex models with non-smooth but convex regularizers. The proposed parameter selection strategy makes the considered class of CNC models automatic, in the sense that the regularization parameter is set without requiring any knowledge about the noise variance.

Acknowledgements We would like to thank the referees for comments that lead to improvements of the presentation. This research was supported in part by the National Group for Scientific Computation (GNCS-INDAM), Research Projects 2019.

References

1. M. Almeida, M. Figueiredo, Parameter estimation for blind and non-blind deblurring using residual whiteness measures. *IEEE Transactions on Image Processing* **22**(7), 2751–2763 (2013)
2. F. Bauer, M.A. Lukas, Comparing parameter choice methods for regularization of ill-posed problems. *Mathematics and Computers in Simulation* **81**(9), 1795–1841 (2011)
3. D. Calvetti, S. Morigi, L. Reichel, F. Sgallari, Tikhonov regularization and the L-curve for large discrete ill-posed problems. *Journal of Computational and Applied Mathematics* **123**(1–2), 423–446 (2000)
4. Y.C. Eldar, Generalized SURE for exponential families: Applications to regularization. *IEEE Transactions on Signal Processing* **57**(2), 471–481 (2009)
5. H.W. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems* (Kluwer, Dordrecht, 1996)
6. U. Haamarik, R. Palm, T. Raus, A family of rules for parameter choice in Tikhonov regularization of ill-posed problems with inexact noise level. *Journal of Computational and Applied Mathematics* **236**(8), 2146–2157 (2012)
7. P.C. Hansen, D.P. O’Leary, The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing* **14**(6), 1487–1503 (1993)
8. P.C. Hansen, M.E. Kilmer, R.H. Kjeldsen, Exploiting residual information in the parameter choice for discrete ill-posed problems. *BIT Numerical Mathematics* **46**(1), 41–59 (2006)
9. C. He, C. Hu, W. Zhang, B. Shi, A fast adaptive parameter estimation for total variation image restoration. *IEEE Transactions on Image Processing* **23**(21), 4954–4967 (2014)
10. A. Lanza, S. Morigi, F. Sgallari, Convex image denoising via non-convex regularization. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **9087**, 666–677 (2015)
11. A. Lanza, S. Morigi, F. Sgallari, A.J. Yezzi, Variational image denoising based on autocorrelation whiteness. *SIAM Journal on Imaging Sciences* **6**(4), 1931–1955 (2013)
12. A. Lanza, S. Morigi, F. Sgallari, Variational image restoration with constraints on noise whiteness. *Journal of Mathematical Imaging and Vision* **53**(1), 61–77 (2015)

13. A. Lanza, S. Morigi, F. Sgallari, Convex image denoising via non-convex regularization with parameter selection. *Journal of Mathematical Imaging and Vision* **56**(2), 195–220 (2016)
14. A. Lanza, S. Morigi, F. Sciacchitano, F. Sgallari, Whiteness constraints in a unified variational framework for image restoration. *Journal of Mathematical Imaging and Vision* **60**(9), 1503–1526 (2018)
15. A. Lanza, S. Morigi, I.W. Selesnick, F. Sgallari, Sparsity-inducing nonconvex nonseparable regularization for convex image processing. *SIAM Journal on Imaging Sciences* **12**(2), 1099–1134 (2019)
16. S. Lefkimmatis, J.P. Ward, M. Unser, Hessian Schatten-norm regularization for linear inverse problems. *IEEE Transactions on Image Processing* **22**(5), 1873–1888 (2013)
17. A. Parekh, I.W. Selesnick, Convex denoising using non-convex tight frame regularization. *IEEE Signal Processing Letters* **22**(10), 1786–1790 (2015)
18. S. Ramani, Z. Liu, J. Rosen, J. Nielsen, J.A. Fessler, Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods. *IEEE Transactions on Image Processing* **21**(8), 3659–3672 (2012)
19. L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physics D*, vol. 60(1–4), pp. 259–268, 1992
20. B.W. Rust, D.P. O’Leary, Residual periodograms for choosing regularization parameters for ill-posed problems. *Inverse Probl.* **24**(3) (2008)
21. B.W. Rust, Parameter selection for constrained solutions to ill-posed problems. *Computing Science and Statistics* **32**, 333–347 (2000)
22. A.M. Thompson, J.C. Brown, J.W. Kay, D.M. Titterton, A study of methods of choosing the smoothing parameter in image restoration by regularization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(4), 326–339 (1991)
23. X. Zhu, P. Milanfar, Automatic parameter selection for denoising algorithms using a no-reference measure of image content. *IEEE Transactions on Image Processing* **19**(12), 3116–3132 (2010)