Ternary image decomposition with automatic parameter selection via auto- and cross-correlation

Ternary Image Decomposition with Automatic Parameter Selection via Auto- and Cross-Correlation

Laura Girometti[1], Alessandro Lanza[1], Serena Morigi[1⋆]

[1] Department of Mathematics, University of Bologna, Piazza di Porta S. Donato 5, 40126 Bologna (BO), Italy

laura.girometti2@unibo.it, alessandro.lanza2@unibo.it, serena.morigi@unibo.it

Abstract

This paper is devoted to the decomposition of a images into cartoon, texture and noise components. A two-stage variational model is proposed which is parameter-free and both context- and noise-unaware. In the first stage the additive white noise component is separated and then the denoised image is further split into cartoon and texture, in the second stage. Auto-correlation and cross-correlation principles represent the key aspects of the two variational stages. The solutions of the two optimization problems are efficiently obtained by the Alternating Directions Method of Multipliers (ADMM). Numerical results show the potentiality of the proposed approach for decomposing images corrupted by different kinds of additive white noises.

Keywords: Variational Image Decomposition, Whiteness, Cross-correlation, Automatic Parameter Selection.

1. Introduction

Decomposing an image into meaningful components is an important and challenging inverse problem in image processing. This task requires the separation of an image into semantically different contents which are assumed to be meaningful parts. A classic example is the image denoising problem in which the goal is to decompose a given degraded image b into a component u representing the *true* (noise-free) image and a component n containing the noise realization. In case of additive noise degradation, these quantities are usually related by $b = u + n$. A well-assessed variational approach for image denoising relies on the search for a solution u in Bounded Variation (BV) space, the set of functions of bounded variation, or with limited total variation (TV). This is the natural space for modeling 'cartoon' type images, since elements of BV consist of homogeneous regions with sharp object boundaries. TV-based variational approaches successfully extract the piecewise-constant (cartoon) structure of the image, while removing noise as well as, eventually present, small scale repeated patterns (texture) from the image. This is due to the fact that both types of patterns (additive noise and texture) can be modelled by oscillatory functions of zero mean taking both positive and negative values. Y. Meyer in [17] introduced the G space, where zero-mean oscillatory functions show a small G-norm, and a BV-G-based variational model able to decompose an image into a cartoon part c containing only geometric objects, and an oscillatory part n containing noise and textures. Inspired by this idea, several variational models have been proposed, some of them addressing numerical issues on the decomposition approach [1, 7], and others introducing the use of a negative Sobolev norm as numerical treatable approximations of the G-norm [19, 20, 18, 9, 24]. Other different approaches have been proposed to compute such a cartoon + texture decomposition: from statistical methods or wavelet techniques, e.g. [16], to the so-called morphological component analysis, which introduces appropriate dictionaries to play the role of discriminants between the different content types [22]. Finally, a recently proposed three-component non-convex formulation is presented in [9], but devoted to simultaneously separate smooth trend, cartoon and oscillatory parts.

However, in the presence of additive noise and fine granularity textures, all the described proposals would incorporate textures and noise into the same component. The critical problem therefore remains to distinguish between noise and texture parts: both texture and noise are indeed characterized by oscillatory small-scale details. Nevertheless, noise is random and is typically characterized by spatially uncorrelated values. This will be the underlying motivation of the proposed decomposition framework which aims to

split a degraded $h \times w$ vectorized image $b \in \mathbb{R}^N$ - with $N = h \times w$ - into three components:

$$b = c + t + n, \tag{1}$$

where the first component c is well-structured, the so called 'cartoon' part, and has a simple geometric description: it models the homogeneous objects which are present in the image. The second and third components t and n contain the oscillating patterns which represent textures and noise, respectively.

Given the desired properties of c, t and n, a natural decomposition model has been first formulated in [2] as an optimization problem which minimizes the sum of three different norms: total variation, G norm for the textured part and E-norm (where E is the generalized Besov space) for the noise component. The intrinsic difficulty with the proposed minimization problem comes from the numerical intractability of the considered norms. For the E-norm, wavelets have been applied to define the equivalent norm. In the last two decades many papers were published along this direction, addressing the modeling aspects and some image applications [3, 5, 6]. However, a very small number of publications proposed to characterize the solutions of such models with respect to the choice of the parameters [1, 7, 8]. Indeed, selecting the regularization parameters in any image variational framework is, in general, of crucial importance, since it can highly influence the quality of the final result. For image decomposition, in the past, this has been done more by trial and error rather than by any well understood theory. An interesting approach to select the regularization parameter is proposed in [3] and applied in a two-term TV-Gabor decomposition model, by relying on the assumption that texture and structured components of an image are poorly correlated.

Furthermore, most of the variational image decomposition approaches rely on more than two energy terms, and inherently on more than one parameter. This clearly allows for high flexibility of the model, but, on the other hand, the parameter selection issue becomes even more critical.

In this work, we propose an image decomposition approach whose novelty relies on two statistical characterizations. From one side, the noise component is extracted based on the expected properties of its auto-correlation, on the other side the texture component is separated based on its cross-correlation with the cartoon component. More precisely, we propose a two stage variational decomposition model which, in the first stage, extracts the noise and, in the second stage, separates the texture and cartoon components. Summarizing, the main contribution of this work is twofold:

- an effective texture-preserving variational model for extracting the noise component from images corrupted by additive white noise, which does not require any a priori knowledge on the noise distribution. It relies on a novel, normalized version of the so-called *residual whiteness set* proposed in [11], which, unlike [11], does not require to know/estimate the noise standard deviation.
- a fully automatic variational model for extracting the texture component, where the regularization parameter is selected based on a novel *cross-correlation principle*. The parameter yielding the maximum dissimilarity between the texture and cartoon components is chosen.

From a numerical point of view, for both the two novel decomposition stages, we propose an efficient numerical solution approach based on the popular Alternating Direction Method of Multipliers (ADMM).

Overall, the proposed two stage decomposition model is parameter-free. In particular, stage I model contains a fidelity term in the form of a whiteness hard constraint without any free regularization parameter; instead, the regularization parameter of stage II model is automatically updated along the ADMM iterations based on the cross-correlation principle.

The work is organised as follows. In Sec. 2 the two-stage variational decomposition model is formulated. In Sec. 3, we begin by recalling some basic definitions on 2D normalized cross-correlation and auto-correlation. Then the whiteness principle and statistical characterization of the sample normalized auto-correlation are introduced in Sec. 4. The proposed cartoon-texture separation approach via normalized cross-correlation minimization is discussed in Sec. 5, while in Sec. 6 the computational ADMM framework is detailed. We report some experimental results in Sec. 7, and conclusions in Sec. 8.

2. The proposed two-stage variational decomposition model

The proposed model for the decomposition of images into cartoon, texture and noise components takes the following form.

Starting from an observed image b, additively composed as in (1), compute estimates c^*, t^*, n^* of components c, t, n by the following two-stage variational framework

- **Stage I:**

$$u^* \in \arg\min_{u\in\mathbb{R}^N} \left\{ \mathscr{J}_1(u;\alpha,a_1) = \sum_{i=1}^{N} \phi\left(\|(\mathsf{D}u)_i\|_2;a_1\right) + \iota_{\mathscr{W}_\alpha}(b-u) \right\}, \tag{2}$$

$$n^* = b-u^*, \tag{3}$$

- **Stage II:**

$$c^* \in \arg\min_{c\in\mathbb{R}^N} \left\{ \mathscr{J}_2(c;\mu,a_2) = \sum_{i=1}^{N} \phi\left(\|(\mathsf{D}c)_i\|_2;a_2\right) + \frac{\mu}{2}\|c-u^*\|_2^2 \right\}, \tag{4}$$

$$t^* = u^*-c^*, \tag{5}$$

where $\mathsf{D} := (\mathsf{D}_h;\mathsf{D}_v) \in \mathbb{R}^{2N\times N}$, with $\mathsf{D}_h, \mathsf{D}_v \in \mathbb{R}^{N\times N}$ finite difference operators discretizing the first-order horizontal and vertical partial derivatives of an image, respectively, and where, with a little abuse of notation, $(\mathsf{D}x)_i := ((\mathsf{D}_h x)_i; (\mathsf{D}_v x)_i) \in \mathbb{R}^2$ indicates the discrete gradient of image x at pixel i. The function $\iota_{\mathscr{W}_\alpha} : \mathbb{R}^N \to \overline{\mathbb{R}} := \mathbb{R}\cup\{+\infty\}$ in (2) is the indicator function of the set $\mathscr{W}_\alpha \subset \mathbb{R}^N$, namely $\iota_{\mathscr{W}_\alpha} = 0$ for $x \in \mathscr{W}_\alpha$, $\iota_{\mathscr{W}_\alpha} = +\infty$ for $x \notin \mathscr{W}_\alpha$. The parametric set $\mathscr{W}_\alpha$, referred to as the *normalized whiteness set* with $\alpha \in \mathbb{R}_{++}$ called the *whiteness parameter*, will be formally defined in Section 4.

The parametric function $\phi(\cdot;a) : \mathbb{R}_+ \to \mathbb{R}_+$ in (2) and (4) is a re-parameterized and re-scaled version of the minimax concave (MC) penalty [25], namely a simple piece-wise quadratic function defined by:

$$\phi(t;a) = \begin{cases} -\dfrac{a}{2}t^2 + \sqrt{2a}\,t & \text{for } t \in \left[0, \sqrt{2/a}\right), \\ 1 & \text{for } t \in \left[\sqrt{2/a}, +\infty\right), \end{cases} \tag{6}$$

with $a \in \mathbb{R}_+$ a free parameter called the *concavity parameter* of penalty ϕ. In fact, since $a = -\min_t \phi''(t;a)$, it represents a measure of the degree of non-convexity of ϕ.

The common regularization function ϕ used in both the Stage I model (2) and the Stage II model (4) represents a well-known, tunable, non-convex generalization of the popular TV regularizer [21]. Thanks to its non-convexity, depending on the value of the concavity parameter a, the regularizer in (2), (4) holds the potential for inducing sparsity of the image gradient magnitudes more strongly than TV. We refer to [8] for details on the sparsity-inducing properties of the regularizer.

Finally, the scalar $\mu \in \mathbb{R}_{++}$ in (4) represents the regularization parameter of the stage II variational model.

The Stage I variational model (2)-(3), which will be discussed in detail in Section 4, is aimed at denoising the observed image b or, equivalently, extracting the noise component n from b, with n any realization from the wide class of additive white noise (AWN) random processes. This class includes important noises such as those characterized by Gaussian (AWGN), uniform (AWUN), Laplacian (AWLN), Cauchy (AWCN) or even mixed distributions, which can be found in many applications [11]. The task is not easy, as the image to denoise is assumed to contain a texture component. Decomposing noise from texture is in fact hard due to both components sharing the property of being highly oscillatory. An effective approach for achieving a good deblur/denoise of textured images consists in enforcing whiteness of the residue image, that is of the noise component, in the data fidelity term [10, 11]. In fact, whiteness is a property typically exhibited by noise but not by texture. In particular, in [11] the authors proposed a variational image restoration model containing a TV regularizer and a fidelity term in the form of a hard constraint on the so-called residual whiteness set. However, the definition of this set requires that the noise standard deviation is known. Here, we build on the idea in [11] but we propose a normalized version of the whiteness set which does not require any knowledge on the noise distribution and standard deviation. In particular, in Section 4 we introduce a normalized whiteness set $\mathscr{W}_\alpha$ to which the residue image $b-u$ of stage I denoising model (2) must belong. The idea is that a good-quality estimate u^* yields a residual $b-u^*$ which resembles the realization n of a white noise process, hence by constraining $b-u^*$ to be white, an accurate u^* and, hence, $n^* = b-u^*$ are achieved.

In the Stage II variational model (4)-(5), devoted to decomposing the result u^* of Stage I into a cartoon and a texture component, the regularization parameter μ is automatically selected by imposing the minimal cross-correlation between the two parts. In particular, as it will be illustrated in Section 5, model (4)-(5)

is recasted within a bilevel optimization framework. The use of the non-convex regularizer instead of the standard TV for extracting the piecewise constant part of the image has been very well assessed in recent works [8].

3. Preliminaries on 2D normalized cross- and auto-correlation

Let us consider two non-zero images in matrix form $x, y \in \mathbb{R}^{h\times w}$

$$x = \left\{x_{i,j}\right\}_{(i,j)\in\Omega},\ y = \left\{y_{i,j}\right\}_{(i,j)\in\Omega},\quad \Omega := \{0,\dots,h-1\}\times\{0,\dots,w-1\}.$$

The *sample normalized cross-correlation* of the two images x and y and the *sample normalized auto-correlation* of image x are the two matrix-valued functions $\rho : \mathbb{R}^{h\times w}\times\mathbb{R}^{h\times w} \to \mathbb{R}^{(2h-1)\times(2w-1)}$ and $\varphi : \mathbb{R}^{h\times w} \to \mathbb{R}^{(2h-1)\times(2w-1)}$ defined by

$$\rho(x,y) = \left\{\rho_{l,m}(x,y)\right\}_{(l,m)\in\Theta},\quad \varphi(x) = \left\{\varphi_{l,m}(x)\right\}_{(l,m)\in\Theta}, \tag{7}$$

$$\Theta := \{-(h-1),\dots,0,\dots,h-1\}\times\{-(w-1),\dots,0,\dots,w-1\}, \tag{8}$$

with scalar components $\rho_{l,m}(x,y) : \mathbb{R}^{h\times w}\times\mathbb{R}^{h\times w} \to \mathbb{R}$ and $\varphi_{l,m}(x) : \mathbb{R}^{h\times w} \to \mathbb{R}$ given by

$$\begin{aligned}\rho_{l,m}(x,y) &= \frac{1}{\|x\|_2\|y\|_2}\left(x\star y\right)_{l,m} = \frac{1}{\|x\|_2\|y\|_2}\left(x'*y\right)_{l,m}\\ &= \frac{1}{\|x\|_2\|y\|_2}\sum_{(i,j)\in\Omega} x_{i,j}\,y_{i+l,j+m},\quad (l,m)\in\Theta,\end{aligned} \tag{9}$$

$$\varphi_{l,m}(x) = \rho_{l,m}(x,x) = \frac{1}{\|x\|_2^2}\sum_{(i,j)\in\Omega} x_{i,j}\,x_{i+l,j+m},\quad (l,m)\in\Theta, \tag{10}$$

respectively. In (9)-(10), $\|\cdot\|_2$ denotes the Frobenius matrix norm, index pairs (l,m) are commonly called *lags*, and the symbols $\star$ and $*$ denote the 2-D discrete correlation and convolution operators, respectively, with $x'(i,j) = x(-i,-j)$.

Clearly, for (9)-(10) being defined for all lags $(l,m)\in\Theta$, images x, y must be (implicitly) padded with at least $h-1$ samples in the vertical direction (up and down) and $w-1$ samples in the horizontal direction (left and right). Upon the assumption of periodic boundary conditions for x, y, the cross- and auto-correlation functions ρ, φ present symmetries that allow to consider only lags

$$(l,m)\in\overline{\Theta} := \{0,\dots,h-1\}\times\{0,\dots,w-1\}.$$

Moreover, it is well known that

$$\rho_{l,m}(x,y),\ \varphi_{l,m}(x,y)\ \in [-1,1]\ \ \forall (l,m)\in\overline{\Theta}.$$

The normalized auto-correlation φ in (9) - see, e.g., [14] - is the key aspect of the first stage of our decomposition model. In fact, the normalized whiteness set $\mathscr{W}_\alpha$ in (2) will be defined in Section 4 by inferring bounds on the normalized auto-correlation values of an AWN realization.

The normalized cross-correlation ρ in (10) is used in the second stage of our decomposition model. In particular, as it will be detailed in Section 5, the regularization parameter μ of stage II variational model (4) is automatically selected based on mimizing the normalized cross-correlation between the decomposed cartoon and texture components. To this aim, we introduce the following non-negative scale-independent scalar measure of correlation $\mathscr{C} : \mathbb{R}^{h\times w}\times\mathbb{R}^{h\times w} \to \mathbb{R}_+$ between the images x and y:

$$\mathscr{C}(x,y) := \frac{1}{N}\,\|\rho(x,y)\|_2^2 = \frac{1}{N}\,\frac{1}{\|x\|_2^2\,\|y\|_2^2}\,\|x\star y\|_2^2, \tag{11}$$

where the second equality in (11) follows from the definition of the normalized cross-correlation ρ in (9). Note that the cross-correlation measure $\mathscr{C}(\cdot,\cdot)$ defined in (11) is scale-independent in the sense that $\mathscr{C}(x,y) = \mathscr{C}(\alpha x,\beta y)$, for any $\alpha,\beta\in\mathbb{R}\setminus\{0\}$.

In Prop. 3.1 below, we give the expression of $\mathscr{C}$ in the Fourier domain.

Proposition 3.1. Let $x, y \in \mathbb{R}^{h\times w}$ and let $\widetilde{x}, \widetilde{y} \in \mathbb{C}^{h\times w}$ be their 2-D discrete Fourier transforms. Then, upon the assumption of periodic boundary conditions for x, y, the function $\mathscr{C}$ defined in (11) satisfies:

$$\mathscr{C}(x,y) = \widetilde{\mathscr{C}}(\tilde{x},\tilde{y}) := \frac{\sum\limits_{(l,m)\in\overline{\Theta}} \left|\widetilde{x}_{l,m}\right|^2 \left|\widetilde{y}_{l,m}\right|^2}{\sum\limits_{(l,m)\in\overline{\Theta}} \left|\widetilde{x}_{l,m}\right|^2 \cdot \sum\limits_{(l,m)\in\overline{\Theta}} \left|\widetilde{y}_{l,m}\right|^2}. \tag{12}$$

Proof. First, the convolution - or, better, correlation - theorem allows to write

$$\widetilde{x \star y} \;=\; \sqrt{N}\left(\overline{\widetilde{x}} \odot \widetilde{y}\right) = \sqrt{N}\left\{\overline{\widetilde{x}}_{l,m}\;\widetilde{y}_{l,m}\right\}_{(l,m)\in\overline{\Theta}}, \tag{13}$$

where $\odot$ denotes the Hadamard matrix product operator and $\bar{z}$, $|z|$ indicate the conjugate and the modulus of complex number z, respectively. Then, by applying the Parseval's theorem we have

$$\|x \star y\|_2^2 \;=\; \|\widetilde{x \star y}\|_2^2 = N \sum_{(l,m)\in\overline{\Theta}} \left|\overline{\widetilde{x}}_{l,m}\,\widetilde{y}_{l,m}\right|^2 = N \sum_{(l,m)\in\overline{\Theta}} \left|\widetilde{x}_{l,m}\right|^2 \left|\widetilde{y}_{l,m}\right|^2, \tag{14}$$

$$\left\|x\right\|_2^2 \left\|y\right\|_2^2 \;=\; \|\widetilde{x}\|_2^2\,\|\widetilde{y}\|_2^2 \;=\; \sum_{(l,m)\in\overline{\Theta}} \left|\widetilde{x}_{l,m}\right|^2 \cdot \sum_{(l,m)\in\overline{\Theta}} \left|\widetilde{y}_{l,m}\right|^2, \tag{15}$$

where the second equality in (14) comes from (13). The expression of $\widetilde{\mathscr{C}}$ in (12) comes straightforwardly by replacing (14) and (15) into the expression of $\mathscr{C}$ in (11). □

4. Stage I: automatic noise separation via normalized auto-correlation

Stage I is aimed at computing an estimate u^* of the cartoon+texture component, that is at separating the noise component $n^* = b - u^*$, under the general assumption that noise is additive white. The key novelty characterizing Stage I variational model in (2)-(3) is the normalized whiteness set $\mathscr{W}_\alpha$, which will be motivated, formally defined and validated by simulations in the next subsection. Then, in subsection 4.2 we outline the ADMM-based numerical solution of Stage I model.

4.1. The normalized whiteness set

If the component $n \in \mathbb{R}^{h\times w}$ in (1) represents the realization of a 2-D discrete white noise random process $\mathscr{N}$, then it is well known that the sample normalized auto-correlation $\varphi(n)$ defined in (10) satisfies the following asymptotic property:

$$\lim_{N\to+\infty} \varphi_{l,m}(n) = \begin{cases} 1 & \text{if } (l,m) = (0,0)\,, \\ 0 & \text{if } (l,m) \in \overline{\Theta}_0 := \overline{\Theta} \setminus \{(0,0)\}\,. \end{cases}$$

For white noise corruptions of realistic (finite) size N, one can roughly say that the normalized auto-correlation values at non-zero lags $(l,m) \in \overline{\Theta}_0$ have all small magnitude, depending on the total number of pixels N. As a general rule, the larger is the image size N, the smaller will be (on average) the sample normalized auto-correlation values of the white noise realization n.

More precisely, regarding all quantities $\varphi_{l,m}(n)$, $(l,m) \in \overline{\Theta}_0$ as realizations of random variables $\Phi_{l,m}(\mathscr{N})$, it can be proved the following result.

Proposition 4.1. Let us consider a $h \times w$ white random process with distribution having (finite) variance σ^2, and let the process be stationary to the fourth-order and with finite fourth-order moments. Then, as the dimension $N = hw$ tends to $+\infty$, the random variables $\Phi_{l,m}(\mathscr{N})$ representing the sample normalized auto-correlation at any nonzero lag $(l,m) \in \overline{\Theta}_0$ are asymptotically uncorrelated and their limiting distribution is a Gaussian distribution defined as follows:

$$\Phi_{l,m}(\mathscr{N}) \sim \mathscr{G}(0,\sigma_\Phi) \quad \forall\,(l,m) \in \overline{\Theta}_0, \quad \sigma_\Phi = N^{-1/2}. \tag{16}$$

The proof of Proposition 4.1 follows quite easily from the one of Proposition 1 in [11], where the distribution of the *unnormalized* auto-correlation is considered and the standard deviation of the zero-mean Gaussian

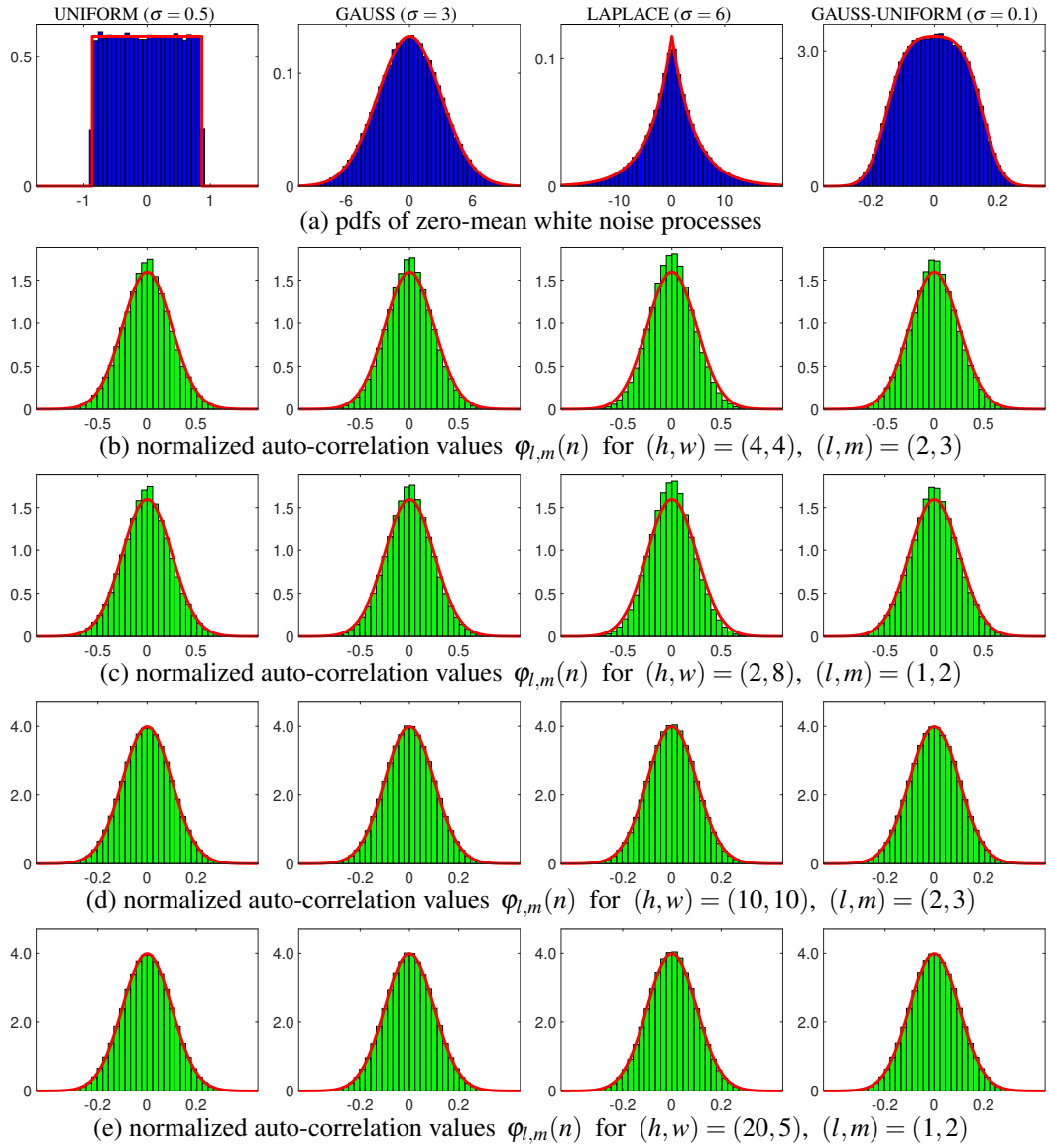

Figure 1: Results of the Montecarlo simulation carried out for validating the theoretical property (16) of normalized auto-correlation values $\varphi_{l,m}(n)$ at non-zero lags (l,m) on realizations $n \in \mathbb{R}^{h \times w}$ of white noise processes with different pdfs and standard deviations.

pdf is $\sigma^2 N^{-1/2}$. Intuitively, statement 16 comes from considering that the normalizing factor (denominator) in the definition (10) of the *normalized* auto-correlation is proportional to σ^2.

In order to provide clear experimental evidence for property (16), in Figure 1 we report the results of a Montecarlo simulation, that we detail in the following.

For some different image sizes (h,w), we generated pseudo-randomly a large number – namely, 10^6 – of realizations $n \in \mathbb{R}^{h \times w}$ from each of four different zero-mean white noise processes, characterized by uniform, Gaussian, Laplace and mixed Gauss-uniform distributions with standard deviations $\sigma = 0.5$, $\sigma = 3$, $\sigma = 6$ and $\sigma = 0.1$, respectively. In the top row of Figure 1 we report (solid red line) the four theoretical noise distributions considered and also the normalized histograms of the generated random samples (in blue) to validate the employed pseudo-random sampling procedure. Then, for each realization n we computed the sample normalized auto-correlation $\varphi(n)$ according to definition (10) and, finally, we constructed, for each non-zero lag $(l,m) \in \overline{\Theta}_0$, the normalized histogram of the obtained 10^6 sample auto-correlation values $\varphi_{l,m}(n)$. In the second to fifth rows of Figure 1 we show the obtained histograms (in green) and the theoretical pdf (solid red line) defined in (16) for four different small image sizes $(h,w) \in \{(4,4),(2,8),(10,10),(20,5)\}$ and different auto-correlation lags $(l,m) \in \{(2,3),(1,2),(2,3),(1,2)\}$. The results in the second and third row correspond to images having $N = 16$ pixels, in the fourth and fifth row

to images with $N = 100$ pixels. Hence, the solid red lines in Figure 1 indicate a zero-mean Gaussian pdf with standard deviation $\sigma_\Phi = 16^{-1/2} = 0.25$ in (b)-(c) and $\sigma_\Phi = 100^{-1/2} = 0.1$ in (d)-(e), respectively.

First, one can notice that, for a given number of pixels N, the histograms are practically independent of the noise pdf and standard deviation, of the image height and width and of the auto-correlation lag. Moreover, for $N = 16$ the histograms are only approximately similar to the theoretical pdf in (16), but already for images made of $N = 100$ pixels, the theoretical pdf almost coincides with the histograms.

It is thus evident that, independently of the noise distribution type and standard deviation, for sufficiently large images the normalized auto-correlation values at all non-zero lags distribute according to the zero-mean Gaussian pdf in (16), where the standard deviation σ_Φ only depends on the image size N.

The validated property (16) motivates the introduction of the following *normalized whiteness set* to which the target white noise component n or, equivalently, the residue image $b-u$ of Stage I model (2), must belong:

$$\begin{aligned} \mathscr{W}_\alpha &:= \left\{ n \in \mathbb{R}^{h\times w} :\ -w_\alpha \le \varphi_{l,m}(n) \le w_\alpha \quad \forall\,(l,m) \in \overline{\Theta}_0 \right\} \\ &= \left\{ n \in \mathbb{R}^{h\times w} :\ -w_\alpha \le \frac{1}{\|n\|_2^2}\big(n\star n\big)(l,m) \le w_\alpha \quad \forall\,(l,m) \in \overline{\Theta}_0 \right\} \\ &= \left\{ n \in \mathbb{R}^{h\times w} :\ -w_\alpha\, n^T n \le \big(n\star n\big)(l,m) \le w_\alpha\, n^T n \quad \forall\,(l,m) \in \overline{\Theta}_0 \right\}, \end{aligned} \tag{17}$$

where the non-negative scalar w_α, referred to as the *whiteness bound*, represents a degree of freedom which allows to set the actual size of the whiteness set or, equivalently, the probability that the sample normalized auto-correlation $\varphi(n)$ of a white noise realization n belongs to the whiteness set. In particular, based on (16), it results natural both to use a unique symmetric bound w_α for any lag (l,m) in definition (17) and to select w_α as a non-negative multiple α of the standard deviation σ_φ , namely

$$w_\alpha = \alpha\,\sigma_\varphi = \alpha N^{-1/2},$$

such that the non-negative *whiteness coefficient* α allows to directly set the probability that the sample normalized auto-correlation of a white noise realization at any given non-zero lag falls inside the whiteness set. In fact, due to the pdf in (16) being Gaussian with zero mean, there is a well known one-to-one relationship between α and the probability that the normalized auto-correlation values belong to the interval $[-w_\alpha, +w_\alpha]$. For instance, setting $\alpha = 2$ yields a whiteness set $\mathscr{W}_{\alpha=2}$ which represents the set of smallest size containing, for any non-zero lag, about 95% of all the possible realizations of white noise processes.

We finally point out that the introduced normalized whiteness set $\mathscr{W}_\alpha$ in (17) is non-convex, similarly to the unnormalized one in [11]. This implies that Stage I represents a challenging nonconvex optimization problem.

4.2. Applying ADMM to the Stage I variational model (2)-(3)

By introducing the three auxiliary variables $g \in \mathbb{R}^{2N}$, $r \in \mathbb{R}^N$, $s \in \mathbb{R}^N$, model (2) can be rewritten in the following equivalent form:

$$\begin{aligned} \{u^*, g^*, r^*, s^*\} \ \in \ & \underset{u,r,s\in\mathbb{R}^N,\, g\in\mathbb{R}^{2N}}{\arg\min} \left\{ \sum_{i=1}^{N} \phi\left(\|g_i\|_2 ; a_1\right) + \iota_{\mathscr{W}_\alpha}(r,s) \right\} \\ & \text{subject to:} \qquad g = \mathsf{D}u,\ r = b-u,\ s = b-u, \end{aligned} \tag{18}$$

where, with a little abuse of notation, $g_i := (\mathsf{D}u)_i = \big((\mathsf{D}_h u)_i\,;(\mathsf{D}_v u)_i\big) \in \mathbb{R}^2$ represents the discrete gradient of image u at pixel i, and $\iota_{\mathscr{W}_\alpha}$ is the indicator function of the feasible set $\mathscr{W}_\alpha$ for variables r and s defined by

$$\mathscr{W}_\alpha := \left\{ (r,s) \in \mathbb{R}^N \times \mathbb{R}^N :\ -w_\alpha r^T s \le (r\star s)(l,m) \le w_\alpha r^T s \ \ \forall\,(l,m) \in \overline{\Theta}_0 \right\}. \tag{19}$$

The auxiliary variable g is introduced to transfer the discrete gradient operator $(\mathsf{D}u)_i$ out of the non-differentiable term $\phi(\|\cdot\|_2; a_1)$. The variables r and s both play the role of the noise component $n = b-u$ within the whiteness constraint (17) so that constraint (19) is now imposed jointly on r and s. To solve (18) we apply the ADMM-based iterative approach proposed in [11], where we simply replace the feasible set with the one defined in (19) and the convex TV regularizer with the non-convex one in (2), based on the penalty function ϕ in (6).

5. Stage II: automatic cartoon-texture decomposition via normalized cross-correlation

Stage II variational model in (4) is aimed at decomposing the output u^* of Stage I into a cartoon component c^* and a texture component $t^* = u^* - c^*$. As previously outlined, the non-convex regularization term in (4) allows to extract the piece-wise constant component c^* more effectively than using the standard TV regularizer. In particular, the larger the concavity parameter a_2 of the penalty function ϕ in (4), the stronger the gradient-sparsity promoting effect of the regularizer. However, by increasing a_2, the cost function $\mathscr{J}_2$ in (4) becomes more and more non-convex, with all the implied numerical difficulties (existence of local minimizers, convergence of minimization algorithms). Since the fidelity term in (4) is quadratic and strongly convex, then the so-called Convex Non-Convex (CNC) strategy can be used for achieving good sparsity-inducing effects while keeping the cost function $\mathscr{J}_2$ convex [4, 13].

Sufficient conditions to apply the CNC strategy to the cost function $\mathscr{J}_2$ in (4), are reported in Prop. 5.1 below; we refer to [15] for proof details.

Proposition 5.1. For any $a_2, \mu \in \mathbb{R}_{++}$, the cost function $\mathscr{J}_2 : \mathbb{R}^N \to \mathbb{R}$ in (4) is continuous, bounded below by zero and coercive. Then, a sufficient condition for $\mathscr{J}_2$ to be convex (strongly convex) is that the parameters a_2, μ satisfy

$$a_2 \mathsf{D}^T\mathsf{D} - \mu\, \mathsf{I}_N \succeq 0 \ (\succ 0) \iff a_2 = \tau \frac{\mu}{8}, \quad \tau \in [0,1] \ (\tau \in [0,1)) \,. \tag{20}$$

It follows that by choosing a_2 according to (20) with *convexity coefficient* $\tau \in [0,1)$, $\mathscr{J}_2$ is strongly convex and, hence, Stage II model (4) admits a unique solution for any $\mu \in \mathbb{R}_{++}$. Hence, we can use $\tau = 0.99$, so that the gradient-sparsity promoting effect of the regularizer is almost maximized in the convex regime, or $\tau > 1$ can be used when even stronger sparsifying effects are desired, but at the cost of dealing with a non-convex model.

The only free parameter in Stage II model is thus the regularization parameter μ. In the next subsection we propose a novel cross-correlation principle which allows to select automatically also the value of μ.

5.1. The cross-correlation principle for selecting μ

The idea at the basis of the proposed cross-correlation principle is to select the μ value which minimizes the normalized cross-correlation measure in (11) between the cartoon component c^* and the texture component $t^* = u^* - c^*$ obtained by solving the Stage II model (4)-(5). In fact, one can expect that the two components are poorly cross-correlated.

In order to formalize the application of the cross-correlation principle, we rewrite Stage II decomposition model (4)-(5) in the form of the following bi-level optimization problem:

$$\mu^* \in \underset{\mu \in \mathbb{R}_{++}}{\arg\min} \left\{ C(\mu) := \mathscr{C}\left(\hat{c}(\mu), \hat{t}(\mu)\right) \right\}, \tag{21}$$

subject to :

$$\begin{cases} \hat{c}(\mu) \ \in \ \underset{c \in \mathbb{R}^N}{\arg\min} \left\{ \displaystyle\sum_{i=1}^{N} \phi\left(\|(\mathsf{D}c)_i\|_2 ; a_2 \right) + \frac{\mu}{2} \|c - u^*\|_2^2 \right\}, \quad a_2 = \tau\, \frac{\mu}{8}, \\ \\ \hat{t}(\mu) \ = \ u^* - \hat{c}(\mu), \end{cases} \tag{22}$$

The bi-level problem consists in a lower-level minimization (22) whose solutions $\hat{c}(\mu), \hat{t}(\mu)$ are arguments of the higher-level minimization (21). The aim of the bi-level problem is then to find an "optimal" regularization parameter μ^* such that $C(\mu^*)$ attains a minimum value, to finally obtain $c^* = \hat{c}(\mu^*)$, $t^* = \hat{t}(\mu^*)$.

Note that in (21) we apply the proposed normalized cross-correlation principle, with the scalar cross-correlation measure $\mathscr{C}$ defined in (11), and that the bi-level problem (21)-(22) does not admit a closed-form solution due to the fact that the solution function $\hat{c}(\mu)$ does not admit an explicit form.

In Section 6 we will illustrate in detail the iterative approach proposed for computing estimates of the solutions of the bi-level problem (21)-(22). The lower-level minimization problem (22) is solved by means of a two-blocks ADMM algorithm, with the value of the regularization parameter μ updated along the ADMM iterations so as to solve the higher-level minimization problem (21) at convergence. In particular, the update of μ can be performed efficiently within one of the two ADMM primal sub-problems thanks to the fact that this sub-problem takes the form of a quadratic Tikhonov-regularized least-square problem.

In fact, as described in the next subsection 5.2, this kind of quadratic problem allows for a direct (and efficient) application of the proposed cross-correlation principle. Therefore, the results in subsection 5.2 will constitute a building-block for the ADMM-based iterative approach proposed in Section 6 for the solution of bi-level problem (21)-(22).

5.2. The cross-correlation principle applied to Tikhonov-regularized least-square problems

We consider the following Tikhonov-regularized least-square problem:

$$\hat{c}(\mu) = \underset{c\in\mathbb{R}^N}{\arg\min} \left\{ Q(c;\mu) := \frac{1}{2}\left\|\mathsf{D}\,c - q\right\|_2^2 + \frac{\mu}{2}\left\|c-u\right\|_2^2 \right\}, \tag{23}$$

with $q \in \mathbb{R}^{2N}$, $u \in \mathbb{R}^N$ given constant vectors (we drop the superscript * for u to simplify notation).

The quadratic cost function $Q(c;\mu) : \mathbb{R}^N \to \mathbb{R}$ in (23) is clearly continuous, bounded below by zero and strongly convex, hence it admits a unique global minimizer, solution of the first-order optimality condition

$$\left(\mathsf{D}^T\mathsf{D} + \mu\ \mathsf{I}_N\right) c \;=\; \mathsf{D}^T q + \mu\ u\,. \tag{24}$$

Recalling that $\mathsf{D} = (\mathsf{D}_h;\mathsf{D}_v)$ and defining accordingly $q = (q_h;q_v),\ q_h,q_v \in \mathbb{R}^N$, the linear system in (24) can be written as

$$\left(\mathsf{D}_h^T\mathsf{D}_h + \mathsf{D}_v^T\mathsf{D}_v + \mu\ \mathsf{I}_N\right) c = \mathsf{D}_h^T q_h + \mathsf{D}_v^T q_v + \mu\ u\,. \tag{25}$$

In order to write the linear system in the Fourier domain, first we note that the two matrices $\mathsf{D}_h, \mathsf{D}_v \in \mathbb{R}^{N\times N}$ are 2D convolution matrices that, under the assumption of periodic boundary conditions, can be diagonalized in $\mathbb{C}$ by using the 2D discrete Fourier transform. Formally, we can write:

$$\mathsf{D}_h = \mathsf{F}^*\widetilde{\mathsf{D}}_h\mathsf{F}, \qquad \mathsf{D}_v = \mathsf{F}^*\widetilde{\mathsf{D}}_v\mathsf{F}\,, \tag{26}$$

with $\mathsf{F}, \mathsf{F}^* \in \mathbb{C}^{N\times N}$ unitary matrices representing the 2D discrete Fourier transform operator and its conjugate transpose (i.e., its inverse), respectively, and with $\widetilde{\mathsf{D}}_h, \widetilde{\mathsf{D}}_v \in \mathbb{C}^{N\times N}$ diagonal matrices defined by

$$\widetilde{\mathsf{D}}_h = \mathrm{diag}\left(\tilde{d}_{h,1},\ldots,\tilde{d}_{h,N}\right), \qquad \widetilde{\mathsf{D}}_v = \mathrm{diag}\left(\tilde{d}_{v,1},\ldots,\tilde{d}_{v,N}\right).$$

In particular, it follows from properties of differential matrices $\mathsf{D}_h, \mathsf{D}_v$ that

$$\widetilde{d}_{h,1} = \widetilde{d}_{v,1} = 0\,, \qquad \left|\widetilde{d}_{h,i}\right|^2 + \left|\widetilde{d}_{v,i}\right|^2 \in (0,8] \ \ \forall\, i \in \{2,\ldots,N\}\,. \tag{27}$$

Replacing (26) into the linear system (25), after simple manipulations we obtain

$$\mathsf{F}^*\left(\left|\widetilde{\mathsf{D}}_h\right|^2 + \left|\widetilde{\mathsf{D}}_v\right|^2 + \mu\ \mathsf{I}_N\right)\mathsf{F}\,c \;=\; \mathsf{F}^*\left(\overline{\widetilde{\mathsf{D}}}_h\mathsf{F}\,q_h + \overline{\widetilde{\mathsf{D}}}_v\mathsf{F}\,q_v + \mu\ \mathsf{F}\,u\right). \tag{28}$$

Then, multiplying both sides of (28) by the unitary matrix F and denoting by $\tilde{x}$ the 2D discrete Fourier transform of image x, i.e. $\tilde{x} = \mathsf{F}x$, we get

$$\left(\left|\widetilde{\mathsf{D}}_h\right|^2 + \left|\widetilde{\mathsf{D}}_v\right|^2 + \mu\ \mathsf{I}_N\right)\widetilde{c} \;=\; \overline{\widetilde{\mathsf{D}}}_h\,\widetilde{q}_h + \overline{\widetilde{\mathsf{D}}}_v\,\widetilde{q}_v + \mu\ \widetilde{u}\,. \tag{29}$$

The coefficient matrix of the linear system (29) is clearly real and diagonal. Moreover, it follows from $\mu \in \mathbb{R}_{++}$ that the matrix is positive definite, hence non-singular.

Thus, the quadratic optimization problem in (23) admits a unique solution given by the following explicit expression:

$$\hat{c}(\mu) \;=\; \mathsf{F}^*\,\widetilde{\hat{c}}(\mu)\,, \quad \text{with} \quad \widetilde{\hat{c}}_i(\mu) \;=\; \frac{\overline{\widetilde{d}}_{h,i}\,\widetilde{q}_{h,i} + \overline{\widetilde{d}}_{v,i}\,\widetilde{q}_{v,i} + \mu\ \widetilde{u}_i}{\left|\widetilde{d}_{h,i}\right|^2 + \left|\widetilde{d}_{v,i}\right|^2 + \mu}\,, \quad i = 1,\ldots,N\,. \tag{30}$$

In particular, we note that, based on the first property in (27), we have

$$\widetilde{\hat{c}}_1(\mu) = \widetilde{u}_1 \ \ \forall\, \mu \in \mathbb{R}_{++}\,.$$

This means that the solution $\hat{c}(\mu)$ of the quadratic minimization problem (23) has the same mean value of the vector u, independently of the regularization parameter μ, and that the associated residual $\hat{t}(\mu) = u - \hat{c}(\mu)$ has mean zero.

The 2D discrete Fourier transformed cartoon component $\widetilde{\hat{c}}(\mu)$ in (30) can be rewritten in more compact form as follows

$$\widetilde{\hat{c}}_i(\mu) = \frac{\varepsilon_i + \mu\ \widetilde{u}_i}{\omega_i + \mu}, \quad i = 1,\ldots,N, \tag{31}$$

where

$$\omega_i := \left|\widetilde{d}_{h,i}\right|^2 + \left|\widetilde{d}_{v,i}\right|^2 \in \mathbb{R}_+, \quad \varepsilon_i := \overline{\widetilde{d}}_{h,i}\,\widetilde{q}_{h,i} + \overline{\widetilde{d}}_{v,i}\,\widetilde{q}_{v,i} \in \mathbb{C}. \tag{32}$$

The associated texture component $\widetilde{\hat{t}}(\mu) = \widetilde{u} - \widetilde{\hat{c}}(\mu)$ thus reads

$$\widetilde{\hat{t}}_i(\mu) = \widetilde{u}_i - \widetilde{\hat{c}}_i(\mu) = \frac{\omega_i \widetilde{u}_i - \varepsilon_i}{\omega_i + \mu}, \quad i = 1,\ldots,N. \tag{33}$$

Based on (31)-(33) and on statement (12) of Prop. 3.1, the cross-correlation function $C(\mu)$ in (21) can be written in explicit form. In fact, we have

$$C(\mu) = \mathscr{C}\left(\hat{c}(\mu),\hat{t}(\mu)\right) = \mathscr{C}\left(\widetilde{\hat{c}}(\mu),\widetilde{\hat{t}}(\mu)\right) = \frac{\displaystyle\sum_{i=1}^{N}\left(f_{c,i}(\mu)\,f_{t,i}(\mu)\right)}{\left(\displaystyle\sum_{i=1}^{N} f_{c,i}(\mu)\right)\left(\displaystyle\sum_{i=1}^{N} f_{t,i}(\mu)\right)} \tag{34}$$

with functions $f_{c,i}, f_{t,i} : \mathbb{R}_{++} \to \mathbb{R}_+$ defined by

$$\begin{aligned} f_{c,i}(\mu) &= \left|\widetilde{\hat{c}}_i(\mu)\right|^2 = \frac{\left|\varepsilon_i + \mu\ \widetilde{u}_i\right|^2}{(\omega_i+\mu)^2} = \frac{\left|\varepsilon_i\right|^2 + \mu^2\left|\widetilde{u}_i\right|^2 + 2\,\mu\,\mathrm{Re}\left(\varepsilon_i\overline{\widetilde{u}}_i\right)}{(\omega_i+\mu)^2} \\ f_{t,i}(\mu) &= \left|\widetilde{\hat{t}}_i(\mu)\right|^2 = \frac{\left|\omega_i\widetilde{u}_i - \varepsilon_i\right|^2}{(\omega_i+\mu)^2} = \frac{\left|\varepsilon_i\right|^2 + \omega_i^2\left|\widetilde{u}_i\right|^2 - 2\,\omega_i\mathrm{Re}\left(\varepsilon_i\overline{\widetilde{u}}_i\right)}{(\omega_i+\mu)^2} \end{aligned}, \quad i = 1,\ldots,N, \tag{35}$$

where we applied in (35) the relation $\left|z_1+z_2\right|^2 = \left|z_1\right|^2 + \left|z_2\right|^2 + 2\,\mathrm{Re}\,(z_1\overline{z}_2)$, which holds for any two complex numbers $z_1, z_2 \in \mathbb{C}$.

Remark 5.2. We observe that a criterion for the selection of the regularization parameter based on the cross-correlation between the cartoon and texture components has been previously proposed in [3]. However, in [3] the regularization parameter of a TV-Gabor decomposition model is selected a-posteriori (i.e., unlike our proposal, the model is numerically solved for a grid of different μ values) as the minimizer of a different and less informative cross-correlation measure. More precisely, our proposal aims to minimize the cross-correlation measure in (11), which is defined in terms of the cross-correlation function values $\rho_{l,m}$ in (9) at all lags (l,m). Instead, the proposal in [3] only takes into consideration the cross-correlation function value at lag $(l,m) = (0,0)$ and it is explicitly stated that the criterion is expected to work well for simple decomposition problems. In order to highlight the benefit of using our more informative criterion, in Figure 2 we compare in terms of Signal-to-Noise Ratio (SNR) the quality of the decomposition results obtained on a simple example by using Stage II model (4)-(5) with μ selected according to both our criterion and the criterion in [3]. The SNR values of cartoon and texture components provided by our approach are 40.46 and 29.11, respectively, those by the method in [3] are 30.03 and 18.69.

6. Applying ADMM to the Stage II variational model (4)-(5)

In this section, we illustrate in detail the ADMM-based iterative algorithm used to numerically solve the Stage II decomposition model (4)-(5), recast as the bi-level optimization problem (21)-(22), with the penalty function ϕ defined in (6) and with the regularization parameter μ automatically updated along the ADMM iterations to satisfy the proposed normalized cross-correlation principle.

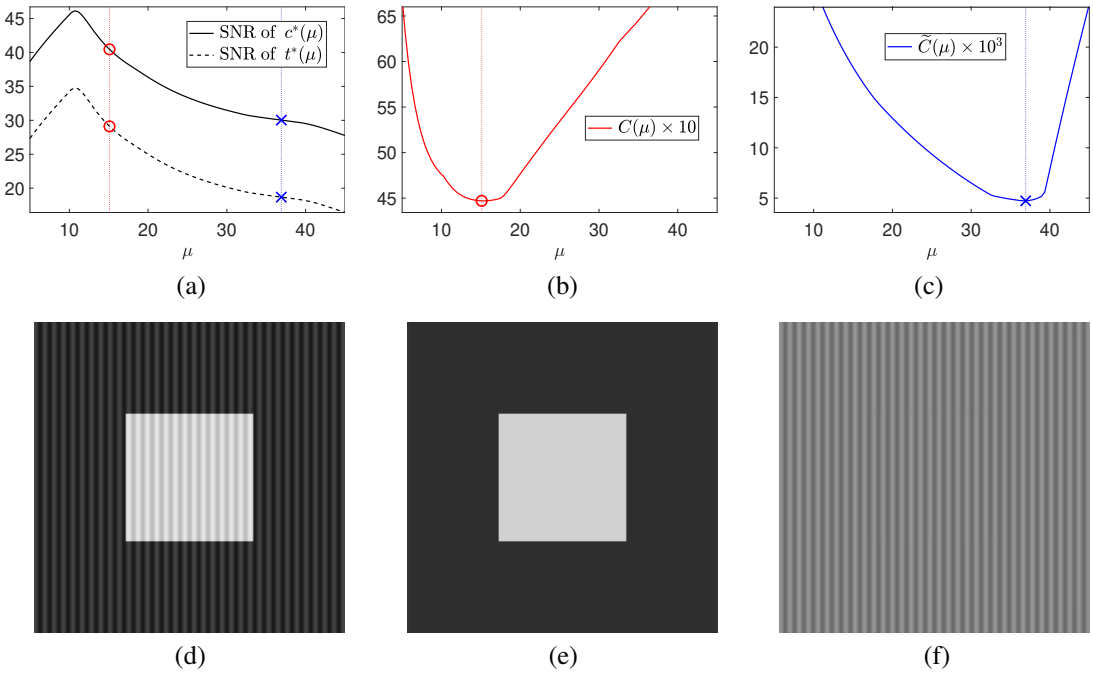

Figure 2: (a) SNR values of the cartoon and texture components obtained by applying the Stage II decomposition model (4)-(5) for a fine grid of different μ values, with circles and crosses indicating the SNR values achieved by using the proposed μ-selection criterion and the one in [3], respectively; (b) proposed cross-correlation measure $C(\mu)$ defined in (21), with the selected minimum marked by a circle; (c) cross-correlation measure proposed in [3], denoted by $\widetilde{C}(\mu)$, with the selected minimum marked by a cross; (d) test image input to decomposition; (e)-(f) cartoon and texture components obtained by using the proposed μ-selection criterion.

We recall that, unlike Stage I variational model (2) which is non-convex due to the feasible whiteness set $\mathscr{W}_\alpha$, Stage II model (4) contains a non-convex regularization term but it can be convexified by applying the previously illustrated CNC strategy [12], see Prop. 5.1.

Similarly to what we did for solving Stage I model, we introduce the auxiliary variable $g := \mathrm{D}c \in \mathbb{R}^{2N}$, which represents the discrete gradient of image c. Then, Stage II model (4) can be rewritten in the following equivalent form:

$$\begin{aligned} \{c^*, g^*\} \;\in\; & \underset{c\in\mathbb{R}^N, g\in\mathbb{R}^{2N}}{\arg\min} \left\{ \sum_{i=1}^{N} \phi\left(\|g_i\|_2; a_2\right) + \frac{\mu}{2}\|c-u^*\|_2^2 \right\}, \quad a_2 = \tau\,\frac{\mu}{8}, \\ & \text{subject to:} \quad g = \mathrm{D}c\,. \end{aligned} \tag{36}$$

To solve (36), we define the augmented Lagrangian function

$$\mathscr{L}(c,g,\lambda;\mu) = \sum_{i=1}^{N} \phi\left(\|g_i\|_2; a_2\right) + \frac{\mu}{2}\|c-u^*\|_2^2 - \langle \lambda, g-\mathrm{D}c\rangle + \frac{\beta}{2}\|g-\mathrm{D}c\|_2^2, \tag{37}$$

where $\lambda \in \mathbb{R}^{2N}$ is the vector of Lagrange multipliers associated with the linear constraints $g = \mathrm{D}c$ in (36) and $\beta \in \mathbb{R}_{++}$ is the ADMM penalty parameter.

Solving (36) amounts to seek for solutions of the saddle-point problem:

$$\begin{gathered} \text{Find } \{c^*, g^*, \lambda^*\} \in \mathbb{R}^N \times \mathbb{R}^{2N} \times \mathbb{R}^{2N} \text{ such that} \\ \mathscr{L}(c^*, g^*, \lambda; \mu) \le \mathscr{L}(c^*, g^*, \lambda^*; \mu) \le \mathscr{L}(c, g, \lambda^*; \mu) \;\; \forall \{c, g, \lambda\} \in \mathbb{R}^N \times \mathbb{R}^{2N} \times \mathbb{R}^{2N}. \end{gathered} \tag{38}$$

Given the previously computed (or initialized for $k = 0$) vectors $c^{(k)}$, $g^{(k)}$, $\lambda^{(k)}$, and the regularization parameter $\mu^{(k)}$, the k-th iteration of the proposed ADMM-based iterative scheme applied to the solution of the saddle-point problem (38) reads as follows:

$$c^{(k+1)} \in \underset{c\in\mathbb{R}^N}{\arg\min}\ \mathscr{L}\left(c, g^{(k)}, \lambda^{(k)}; \mu^{(k)}\right), \tag{39}$$

$$g^{(k+1)} \in \underset{g\in\mathbb{R}^{2N}}{\arg\min}\ \mathscr{L}\left(c^{(k+1)}, g, \lambda^{(k)}; \mu^{(k)}\right), \tag{40}$$

$$\lambda^{(k+1)} = \lambda^{(k)} - \beta\left(g^{(k+1)} - \mathsf{D}c^{(k+1)}\right). \tag{41}$$

In the following two subsections we show how to solve the two minimization sub-problems (39)-(40) for the primal variables g and c, then in subsection 6.0.3 we outline and analyze the main computational steps of the overall algorithm.

6.0.1. Solving the subproblem for c and updating μ based on the cross-correlation principle

After dropping the constant terms (i.e., the terms not depending on the optimization variable c) of the augmented Lagrangian function $\mathscr{L}$ in (37) , the minimization sub-problem for c in (39) reads

$$c^{(k+1)} = \underset{c\in\mathbb{R}^N}{\arg\min}\left\{\frac{\mu}{2}\|c-u^*\|_2^2 - \langle\lambda^{(k)}, g^{(k)} - \mathsf{D}c\rangle + \frac{\beta}{2}\|g^{(k)} - \mathsf{D}c\|_2^2\right\}. \tag{42}$$

After simple algebraic manipulations, (42) can be equivalently rewritten as

$$c^{(k+1)} = \underset{c\in\mathbb{R}^N}{\arg\min}\left\{Q(c;\gamma) := \frac{1}{2}\|\mathsf{D}c - q^{(k)}\|_2^2 + \frac{\gamma}{2}\|c-u^*\|_2^2\right\} \tag{43}$$

with

$$q^{(k)} = g^{(k)} - \frac{1}{\beta}\lambda^{(k)} \in \mathbb{R}^{2N}, \quad \gamma = \frac{\mu}{\beta} \in \mathbb{R}_{++}. \tag{44}$$

The unconstrained quadratic optimization problem in (43)-(44) has exactly the same form of problem (23), with $q^{(k)}$ in place of q and γ in place of μ. Hence, according to the cross-correlation principle outlined in Section 5.2, we can preliminarily determine the value $\gamma^{(k+1)}$ – hence, the value $\mu^{(k+1)} = \beta\,\gamma^{(k+1)}$ – which minimizes the normalized cross-correlation between the obtained cartoon component $c^{(k+1)}$ and the associated texture component $t^{(k+1)} = u^* - c^{(k+1)}$. More precisely, $\gamma^{(k+1)}$ is selected by minimizing the cross-correlation measure $C(\gamma)$ defined as in (34), namely

$$\gamma^{(k+1)} \in \underset{\gamma\in\mathbb{R}_{++}}{\arg\min}\left\{C^{(k)}(\gamma) := \frac{\displaystyle\sum_{i=1}^{N}\left(f_{c,i}^{(k)}(\gamma)\, f_{t,i}^{(k)}(\gamma)\right)}{\left(\displaystyle\sum_{i=1}^{N} f_{c,i}^{(k)}(\gamma)\right)\left(\displaystyle\sum_{i=1}^{N} f_{t,i}^{(k)}(\gamma)\right)}\right\}. \tag{45}$$

Notice that, unlike in Section 5.2, here we added the superscript $^{(k)}$ to the functions $f_{c,i}$, $f_{t,i}$ and, hence, to the function C to highlight that they can change along the ADMM iterations due to vector $q^{(k)}$ in (44) changing. More precisely, based on definitions in (35), the functions $f_{c,i}^{(k)}, f_{t,i}^{(k)} : \mathbb{R}_{++} \to \mathbb{R}_+$ in (45) read

$$\begin{aligned} f_{c,i}^{(k)}(\gamma) &= \frac{\left|\varepsilon_i^{(k)}\right|^2 + \gamma^2\left|\widetilde{u}_i\right|^2 + 2\,\gamma\,\mathrm{Re}\left(\varepsilon_i^{(k)}\overline{\widetilde{u}}_i\right)}{(\omega_i+\gamma)^2} \\ f_{t,i}^{(k)}(\gamma) &= \frac{\left|\varepsilon_i^{(k)}\right|^2 + \omega_i^2\left|\widetilde{u}_i\right|^2 - 2\,\omega_i\mathrm{Re}\left(\varepsilon_i^{(k)}\overline{\widetilde{u}}_i\right)}{(\omega_i+\gamma)^2} \end{aligned}, \quad i = 1,\ldots,N, \tag{46}$$

where constants ω_i, $\varepsilon_i^{(k)}$, $i = 1,\ldots,N$, are defined as in (32), with $\varepsilon_i^{(k)}$ depending on vector $q^{(k)}$ in (44). To solve the 1-dimensional minimization problem in (45), we use a grid search procedure. We choose a discrete set of candidate γ values

$$\Gamma := \{\gamma_1, \gamma_2, \ldots, \gamma_M\}, \quad \gamma_j \in \mathbb{R}_{++},\ j = 1,\ldots,M, \tag{47}$$

then we select as $\gamma^{(k+1)}$ the γ_j yielding the minimum value of function $C^{(k)}(\gamma)$; in formula,

$$\gamma^{(k+1)} = \underset{\gamma\in\Gamma}{\arg\min}\; C^{(k)}(\gamma)\,. \tag{48}$$

Once $\gamma^{(k+1)}$ has been selected, we can compute $c^{(k+1)}$ as defined in (30), (31), namely

$$c^{(k+1)} = \mathsf{F}^*\,\widetilde{c}^{(k+1)}, \quad \text{with } \widetilde{c}_i^{(k+1)} = \frac{\varepsilon_i^{(k)} + \gamma^{(k+1)}\,\widetilde{u}_i}{\omega_i + \gamma^{(k+1)}}, \quad i = 1,\ldots,N\,. \tag{49}$$

.

6.0.2. Solving the sub-problem for g

Before solving the subproblem for g, we update the value of the concavity parameter a_2 of the penalty function ϕ based on the new value of the regularization parameter $\mu^{(k)}$ selected in the previous subproblem:

$$a_2^{(k)} = \tau\,\frac{\mu^{(k)}}{8}\,. \tag{50}$$

The minimization sub-problem for g in (40) can thus be equivalently written as follows:

$$g^{(k+1)} \in \underset{g\in\mathbb{R}^{2N}}{\arg\min}\; \left\{ \sum_{i=1}^{N} \phi\left(\|g_i\|_2\,;a_2^{(k)}\right) + \frac{\beta}{2}\left\|g - v^{(k)}\right\|_2^2 \right\}, \tag{51}$$

with vector $v^{(k)} \in \mathbb{R}^{2N}$ defined by

$$v^{(k)} = \mathsf{D}c^{(k+1)} + \frac{1}{\beta}\lambda^{(k)}\,.$$

The $2N$-dimensional problem (51) is equivalent to the following N independent 2-dimensional problems:

$$\begin{aligned} g_i^{(k+1)} &\in \underset{g_i\in\mathbb{R}^2}{\arg\min}\; \left\{ \phi\left(\|g_i\|_2\,;a_2^{(k)}\right) + \frac{\beta}{2}\left\|g_i - v_i^{(k)}\right\|_2^2 \right\} \\ &= \operatorname{prox}^{\beta}_{\phi(\|\cdot\|_2;a_2^{(k)})}\left(v_i^{(k)}\right), \quad i = 1,\ldots,N, \end{aligned} \tag{52}$$

with the constant vectors $v_i^{(k)} \in \mathbb{R}^2$ defined by

$$v_i^{(k)} = \left(\mathsf{D}c^{(k+1)}\right)_i + \frac{1}{\beta}\left(\lambda^{(k)}\right)_i, \quad i = 1,\ldots,N, \tag{53}$$

and where the (possibly set-valued) function $\operatorname{prox}^{\beta}_{\phi(\|\cdot\|_2;a_2^{(k)})} : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ denotes the proximal operator of function $\phi(\|\cdot\|_2;a_2^{(k)})$ with proximity parameter β.

A necessary and sufficient condition for all the cost functions in (52) to be strongly convex, such that the proximal maps in (52) are single-valued functions and, hence, the (global) minimizers $g_i^{(k+1)}$ exist and are unique, is that $\beta > a^{(k)}$, see [8]. In particular, we set

$$\beta^{(k)} = \max\left\{\beta\,,\,1.05\,a^{(k)}\right\}. \tag{54}$$

The proximal map in (52) admits a closed-form expression [8], such that the N problems in (52) admit solutions $g_i^{(k+1)}$, $i = 1,\ldots,N$, which read

$$g_i^{(k+1)} = \xi_i^{(k)} v_i^{(k)}, \quad \xi_i^{(k)} = \begin{cases} 0 \;\text{ if }\; \left\|v_i^{(k)}\right\|_2 = 0\,, \\ \max\left\{ \min\left\{ \dfrac{\beta^{(k)} - \sqrt{2a_2^{(k)}}\,/\left\|v_i^{(k)}\right\|_2}{\beta^{(k)} - a_2^{(k)}}\,,\,1 \right\},0 \right\} \;\text{ if }\; \left\|v_i^{(k)}\right\|_2 > 0\,. \end{cases} \tag{55}$$

Algorithm 1 ADMM-based method for the numerical solution of Stage II decomposition model (4)-(5)

inputs: image $u^* \in \mathbb{R}^N$ output of Stage I decomposition model (2)

output: estimated cartoon and texture components $c^*, t^* \in \mathbb{R}^N$, with $c^* + t^* = u^*$

1. **initialise:** set $g^{(0)} = \mathrm{D}\,u^*$, $\lambda^{(0)} = 0$
2. **for** *k = 0, 1, 2, ... until convergence* **do**:
3. · compute $\gamma^{(k+1)}$ by (44)-(48)
4. · compute $c^{(k+1)}$ by (49)
5. · compute $g^{(k+1)}$ by (50), (53)-(55)
6. · compute $\lambda^{(k+1)}$ by (41)
7. **end for**
8. $c^* = c^{(k+1)}$, $t^* = u^* - c^*$

6.0.3. Computational details and efficiency analysis

The main computational steps of the overall proposed ADMM-based approach for the numerical solution of Stage II decomposition model (4)-(5) with automatic selection of both the two model parameters μ, a_2 and (eventual) adjustment of the ADMM penalty parameter β are outlined in Alg. 1.

For what regards the computational cost per-iteration, first we note that solving the subproblem for g by (50), (53)-(55) and updating λ by (41) - lines 5-6 of Alg. 1 - both require $O(N)$ operations, whereas updating μ by (44)-(48) and solving the subproblem for c by (49) - lines 3-4 of Alg. 1 - require at least $O(N\log N)$ operations. In fact, a direct 2D discrete Fourier transform of $q^{(k)}$ in (44) is necessary to determine the complex constants $\varepsilon_i^{(k)}$ in (46) - see definition (32) - and an inverse 2D discrete Fourier transform of $\widetilde{c}^{(k+1)}$ is required in (49), both performed by using a 2D fast Fourier transform implementation.

Actually, the CPU time profiling of Alg. 1 shows that the real computational bottleneck resides in the μ-updating step in line 3, but not due to the Fourier transform. In fact, once the constants $\varepsilon_i^{(k)}$ has been calculated by Fourier transforming $q^{(k)}$, the computation of $\gamma^{(k+1)}$ by (45)-(48) requires $O(NM)$ operations, with M the number of different γ values tested in the grid-search minimization procedure (47)-(48). This step is the more costly since in order to achieve a sufficiently accurate estimate of $\gamma^{(k+1)}$, M can not be too small and, in general, must be taken larger than $\log N$. For this reason, in the following we outline the procedure used to evaluate M times - i.e., for $\gamma \in \{\gamma_1, \ldots, \gamma_M\}$ - the function $C^{(k)}(\gamma)$ in (45)-(46) as efficiently as possible, based on precomputing constants.

At the beginning of the ADMM algorithm, we compute once for all the constants entering the definition of $C^{(k)}(\gamma)$ which do not change along the iterations (this avoids recomputing the same quantities at each iteration). In particular, after calculating (by 2D discrete Fourier transform of u and by Fourier diagonalization of convolution matrices $\mathrm{D}_h, \mathrm{D}_v$) the complex numbers $\widetilde{u}_i, \widetilde{d}_{h,i}, \widetilde{d}_{h,i}$ and their conjugates $\overline{\widetilde{u}}_i, \overline{\widetilde{d}}_{h,i}, \overline{\widetilde{d}}_{h,i}$, $i = 1, \ldots, N$, we compute, in the following order, the constants

$$\left\{ \begin{array}{lcll} \omega_i & = & \overline{\widetilde{d}}_{h,i}\,\widetilde{d}_{h,i} + \overline{\widetilde{d}}_{v,i}\,\widetilde{d}_{v,i} & \in \mathbb{R}_+ \\ \eta_{1,i} & = & \overline{\widetilde{u}}_i\,\widetilde{u}_i & \in \mathbb{R}_+ \\ \eta_{2,i} & = & \omega_i^2\,\eta_{1,i} & \in \mathbb{R}_+ \end{array} \right. , \quad i = 1, \ldots, N\,.$$

Then, at each ADMM iteration (with iteration index k), after calculating the vector $q^{(k)}$ as defined in (44) and its 2D discrete Fourier transform (with complex components $\widetilde{q}_{h,i}^{(k)}, \widetilde{q}_{v,i}^{(k)}$, $i = 1, \ldots, N$), we compute,

Table 1: Example 1: decomposition of the `geometricR` image, shown in Fig.3, composed by a `radial` texture t and a `geometric` cartoon component c; SNR and SSIM values of c^* and t^* for different f factors.

f	c^*: SNR	t^*: SNR	c^*: SSIM	t^*: SSIM
0.55	16.20	14.34	0.930	0.881
0.60	17.38	13.74	0.934	0.889
0.65	18.73	13.23	0.939	0.899
0.70	20.22	12.74	0.946	0.909
0.75	21.88	12.22	0.956	0.920
0.80	23.88	11.71	0.967	0.930
0.85	26.39	11.20	0.978	0.941
0.90	29.94	10.73	0.988	0.955
0.95	34.97	9.27	0.998	0.975

in the following order, the constants:

$$\begin{cases} \varepsilon_i^{(k)} &= \overline{\widetilde{d}}_{h,i}\,\widetilde{q}_{h,i}^{(k)} + \overline{\widetilde{d}}_{v,i}\,\widetilde{q}_{v,i}^{(k)} & \in \mathbb{C} \\ \eta_{3,i}^{(k)} &= \overline{\varepsilon}_i^{(k)}\,\varepsilon_i^{(k)} & \in \mathbb{R}_+ \\ \eta_{4,i}^{(k)} &= 2\,\mathrm{Re}\left(\varepsilon_i^{(k)}\,\overline{\widetilde{u}}_i\right) & \in \mathbb{R} \\ \eta_{5,i}^{(k)} &= \eta_{2,i} + \eta_{3,i}^{(k)} - \omega_i\,\eta_{4,i}^{(k)} & \in \mathbb{R} \end{cases}, \quad i = 1,\ldots,N.$$

Hence, at each ADMM iteration, for any of the M different γ_j values belonging to the discrete set Γ defined in (47), we calculate

$$\begin{cases} f_{c,i}^{(k)}(\gamma_j) &= \dfrac{\eta_{3,i}^{(k)} + \gamma_j\left(\eta_{1,i}\,\gamma_j + \eta_{4,i}^{(k)}\right)}{\left(\omega_i + \gamma_j\right)^2} & \in \mathbb{R} \\ f_{t,i}^{(k)}(\gamma_j) &= \dfrac{\eta_{5,i}^{(k)}}{\left(\omega_i + \gamma_j\right)^2} & \in \mathbb{R} \end{cases}, \quad i = 1,\ldots,N,$$

and then compute the value $C^{(k)}(\gamma_j)$ by using the definition in (45).

Finally, $\gamma^{(k+1)}$ is selected as the $\gamma_j \in \Gamma$ yielding the minimum value of $C^{(k)}(\gamma_j)$ - see the minimization problem in (48).

7. Numerical Results

In this section we assess the effectiveness of the proposed two-stages image decomposition method and the performance of the automatic selection of the regularization parameter μ under the cross-correlation minimization discussed in Section 5. More specifically, in the first example in Section 7.1 we focus only on the cartoon-texture decomposition according to the variational model (4)-(5) in Stage II where u^* is taken as a noise-free synthetically constructed image, to better highlight the benefits of the parameter μ selection. In the second example, discussed in Section 7.2, we present the results of the overall two-stages decomposition framework applied to photographic images corrupted by different kinds of additive white noise, where the textured parts are limited to small regions of the image and represent natural as well as geometric patterns. A comparison with other state-of-the-art variational models for the "c+t+n" decomposition is finally presented and commented in the third example in Section 7.3.

We tested the decomposition framework in MATLAB on both synthetic and photographic images, and we assessed the quality of the computed image decompositions by two quantitative measures: the Signal-to-Noise Ratio (SNR) and the Structural Similarity Index (SSIM). SNR is defined by the following formula

$$\mathrm{SNR}(x, x_{\mathrm{GT}}) := 10\log_{10}\frac{\|x_{\mathrm{GT}} - \mathbb{E}[x_{\mathrm{GT}}]\|_2^2}{\|x - x_{\mathrm{GT}}\|_2^2},$$

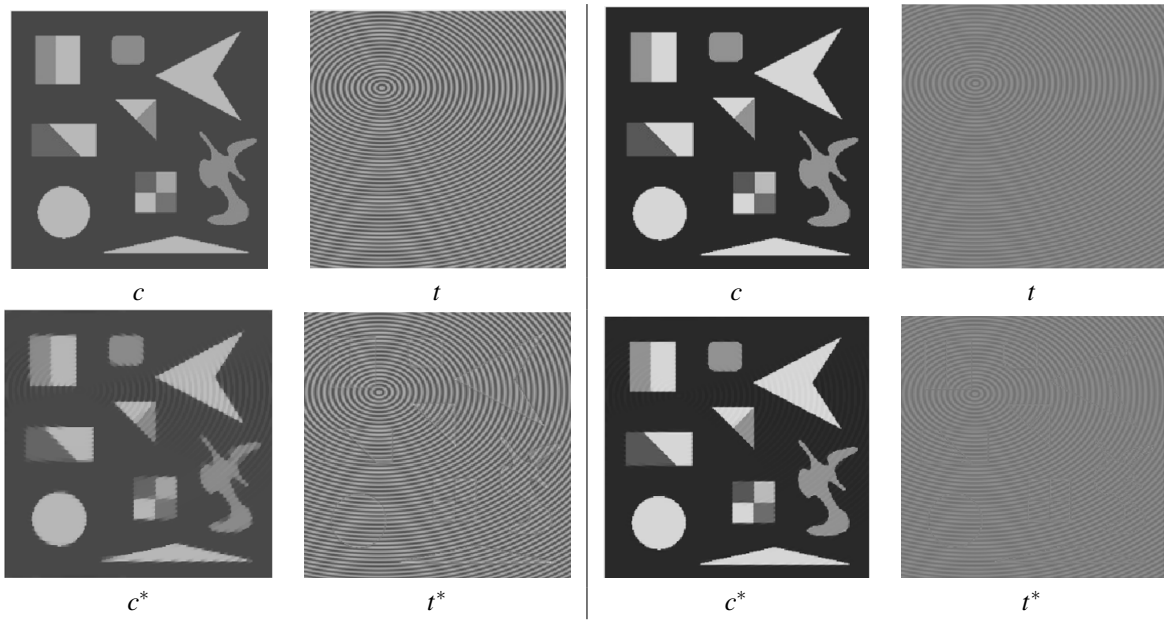

Figure 3: Example 1: decomposition of the `geometricR` image into cartoon c^* and textured t^* images for different blending factors $f = 0.55$ (left) and $f = 0.85$ (right).

Table 2: Example 1: decomposition of the `starR` image, shown in Fig.4, composed by blending a `radial` texture t and a `star` cartoon component c with factor $f = 0.85$; SNR and SSIM values of c^* and t^* for different texture scales T.

T	c^*: SNR	t^*: SNR	c^*: SSIM	t^*: SSIM
2	27.36	10.54	0.989	0.950
4	26.33	9.70	0.979	0.916
6	24.91	8.29	0.976	0.880

where x_{GT}, x are the ground truth image component and its reconstruction, respectively, and $\mathbb{E}[x_{\mathrm{GT}}]$ is the mean of x_{GT}. SSIM is used as a metric to measure the similarity between two given images x and y, and it is defined as

$$\mathrm{SSIM}(x,y) := \frac{(2\mathbb{E}[x]\mathbb{E}[y]+C_1)(2\sigma_{xy}+C_2)}{(\mathbb{E}[x]^2+\mathbb{E}[y]^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)},$$

where σ_{xy} represents the covariance between x and y, σ_x and σ_y the standard deviations of x and y, and C_1, C_2 constants with standard values as defined in the `ssim` Matlab command routine.

The parameters in the ADMM-based methods in Algorithm 1 are set as follows: τ in (20) is set to $\tau = 0.99$ in Examples 1 and 3, and, consequently, the concavity parameter for Stage II model is set as $a_2 = \tau\frac{\mu}{8}$ in order to guarantee to be in a CNC regime. In Example 2, the challenging decomposition of corrupted textured photographic images required a non-convex regime in Stage II with $\tau_c = 5$. Finally, the penalty parameter β of ADMM for Stage II satisfies (54), such that the ADMM sub-problem for g is strongly convex. Moreover, iterations of the ADMM-based approaches used for solving Stage I and II models are both stopped as soon as the iteration index k or the iterate $x^{(k)}$ satisfy

$$k > 2000, \qquad \delta^{(k)} := \frac{\left\|x^{(k)}-x^{(k-1)}\right\|_2}{\left\|x^{(k-1)}\right\|_2} \leq 10^{-5}, \tag{56}$$

where $x^{(k)} = u^{(k)}$ for Stage I, $x^{(k)} = c^{(k)}$ for Stage II.

7.1. Example 1: performance of stage II

In this section, we evaluate the performance of the proposed Stage II model (4)-(5) in separating textured and cartoon image components from a noise-free image b (corresponding to u^* in (4)). The success of this stage strongly depends on the cross-correlation principle applied to automatically set the regularization parameter μ in (4).

In our examples, the image b is synthetically constructed as a linear combination of an original cartoon image c with values normalized in the range $[0,1]$, and a zero-mean texture component t with values in the range $[-0.5, 0.5]$. Then b is a combination obtained by a blending scalar factor $f \in [0,1]$ as follows

$$b = fc + (1-f)t. \tag{57}$$

We will analyze how the blending factor f, the kind of texture and the geometric structure can affect the texture-cartoon image decomposition process.

First of all, we have experimentally inferred that textures with a more modest contribution to the overall intensity of image b (i.e., f large) result in higher quality decompositions. In Figure 3 we illustrate the decomposition results of images b synthetically constructed by applying (57) with two different blending factors: $f = 0.55$ (Figure 3, left panel), and $f = 0.85$ (Figure 3, right panel). In the second row of Figure 3 the illustrated images are the computed estimates c^*, t^* of the components c, t, where here $b = c + t$ (first row). In Table 1, quantitative evaluations in terms of SNR and SSIM values are reported for increasing f values.

We observe that, by increasing the blending factor the quality of the piece-wise component c^* improves, and corresponding SNR and SSIM values increase. Differently for the component t^*, where the SNR slightly decreases, remaining approximately constant, while the SSIM increases. This can be visually justified by the incorrect presence in t^* of edges of the geometric structures which is noticeable for low factors of f, and tends to disappear for higher values of f. We noted this trend for many different kinds of cartoon and texture components.

Another important characteristic of our decomposition framework that we have deduced experimentally is that textures with finer granularity are more easily separable from the cartoon part. From a mathematical point of view, while the definition of noise or piecewise-constant components of an image are pretty clear, the definition of texture is rather vague and highly dependent on the level of detail (granularity) of the textured pattern. To better understand which level of detail is better handled by our framework, we constructed test images with different granularity depending on a parameter T which controls the scale of the details

$$t = \sin\left(2\pi\frac{\mathrm{S}}{T}\right), \tag{58}$$

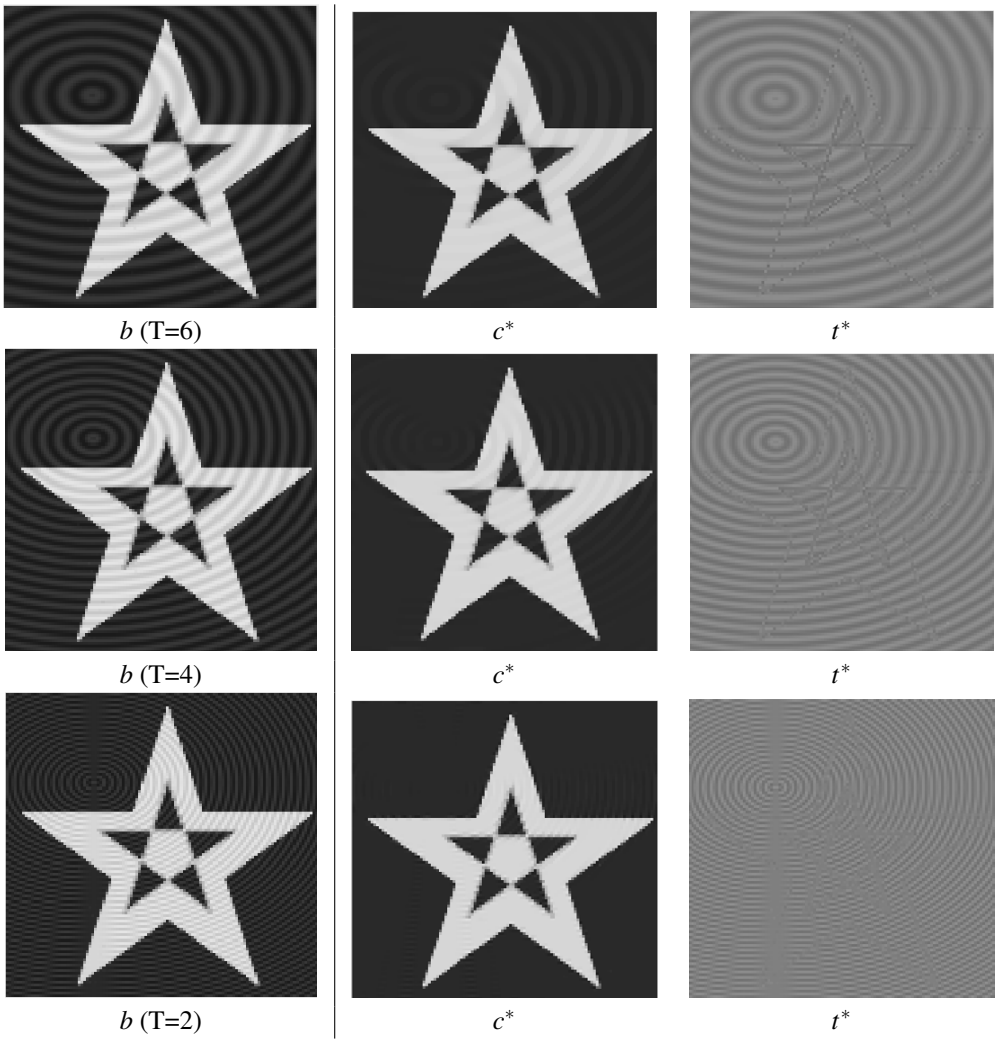

Figure 4: Example 1: decomposition of the `starR` images (first column) composed by blending with factor $f = 0.85$ a cartoon component c `star` with different textured images t `radial` generated using periods $T = 6, 4, 2$ in (58).

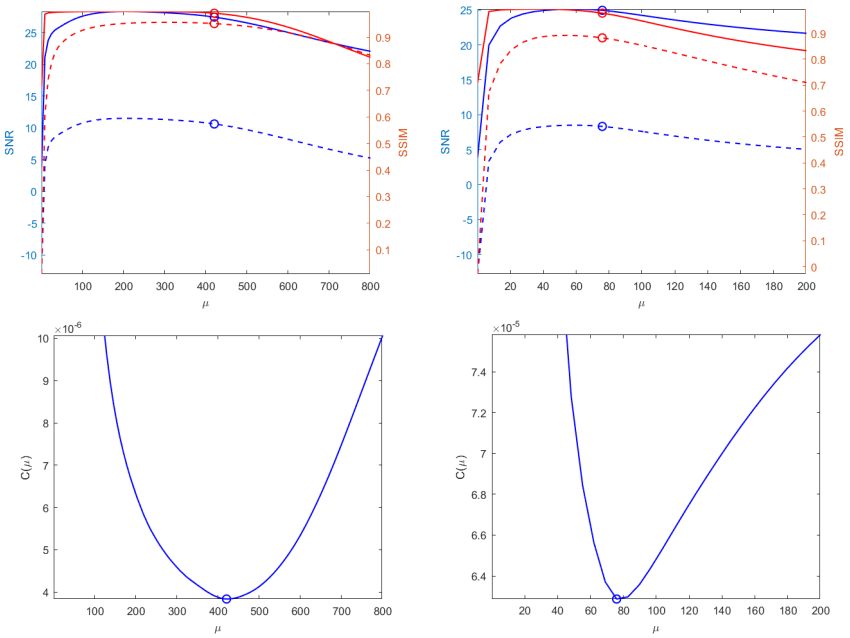

Figure 5: Example 1: first column, plots referred to the decompositions in Fig. 4(T=2); second column, plots referred to the decompositions in Fig. 4(T=6). For the two cases we have that the minumum $C(\mu)$ is 3.84×10^{-6} (left panel) and 6.29×10^{-5} (right panel)

where T represents the period of the texture and S is a matrix which characterizes the texture. For example, for radial sinusoidal textures, each entry of S represents the distance between the coordinates of a reference point and the coordinates of each pixel image, eventually rescaled by a factor.

In (58), as T decreases, the function increases its oscillatory behaviour (frequency) or, equivalently, the granularity of the texture component becomes finer.

In Table 2, SNR and SSIM values of c^* and t^* are reported for some values of T on the test image b, composed by a `radial` texture component and a `star` cartoon component, illustrated in Figure 4. As T decreases, the performance of the cross-correlation principle applied improves, in the sense that the resulting estimated components improve both qualitatively and in terms of SNR and SSIM values.

To strengthen this fact, we illustrate in Figure 5 the plots of the cross-correlation measure $C(\mu^*)$ and of the SNR and SSIM curves as function of increasing μ values for two different texture granularities $T = 2$ (left) and $T = 6$ (right). The regularization parameter μ, selected by our automatic algorithm as the global minimizer of the function $C(\mu)$ in (11), is depicted by a circle. One can notice that the selected μ value in both cases is near to the optimal one, but for finer granularity (left panel) the minimum cross-correlation measure $C(\mu^*)$ is one order of magnitude smaller. This reflects the fact that the achieved decomposition is of higher quality, as the two obtained components c^* and t^* are less correlated.

Finally, Stage II of the proposed decomposition framework has been validated against different geometric texture models without finding significant variation on its performance. In the examples illustrated in Figure 6 we show the cartoon-texture components of the `geometricR` images (first column) composed by blending with factor $f = 0.90$ a geometric cartoon component c with textured images t `linear`, `radial` and `chess` (row-wise); with coarse (top block) and fine (bottom block) textured patterns. A more complete overview is offered by the results in Table 3 for varying f values (first column). The quantitative performance associated with the images shown in Figure 6 are reported in the last block.

In general we found out that the radial textures get slightly worse results in terms of SNR and SSIM values and this is much more visible in coarse granularity textured images t. We will expand our insights in Example 2 where we will validate our framework on photographic images containing natural textures.

Remark 7.1 (**Note on convergence of the algorithm for Stage II**). All results reported in this section are obtained by using the ADMM-based approach outlined in Section 6, which computes estimates of both the cartoon/texture components and the regularization parameter by solving the bi-level optimization problem

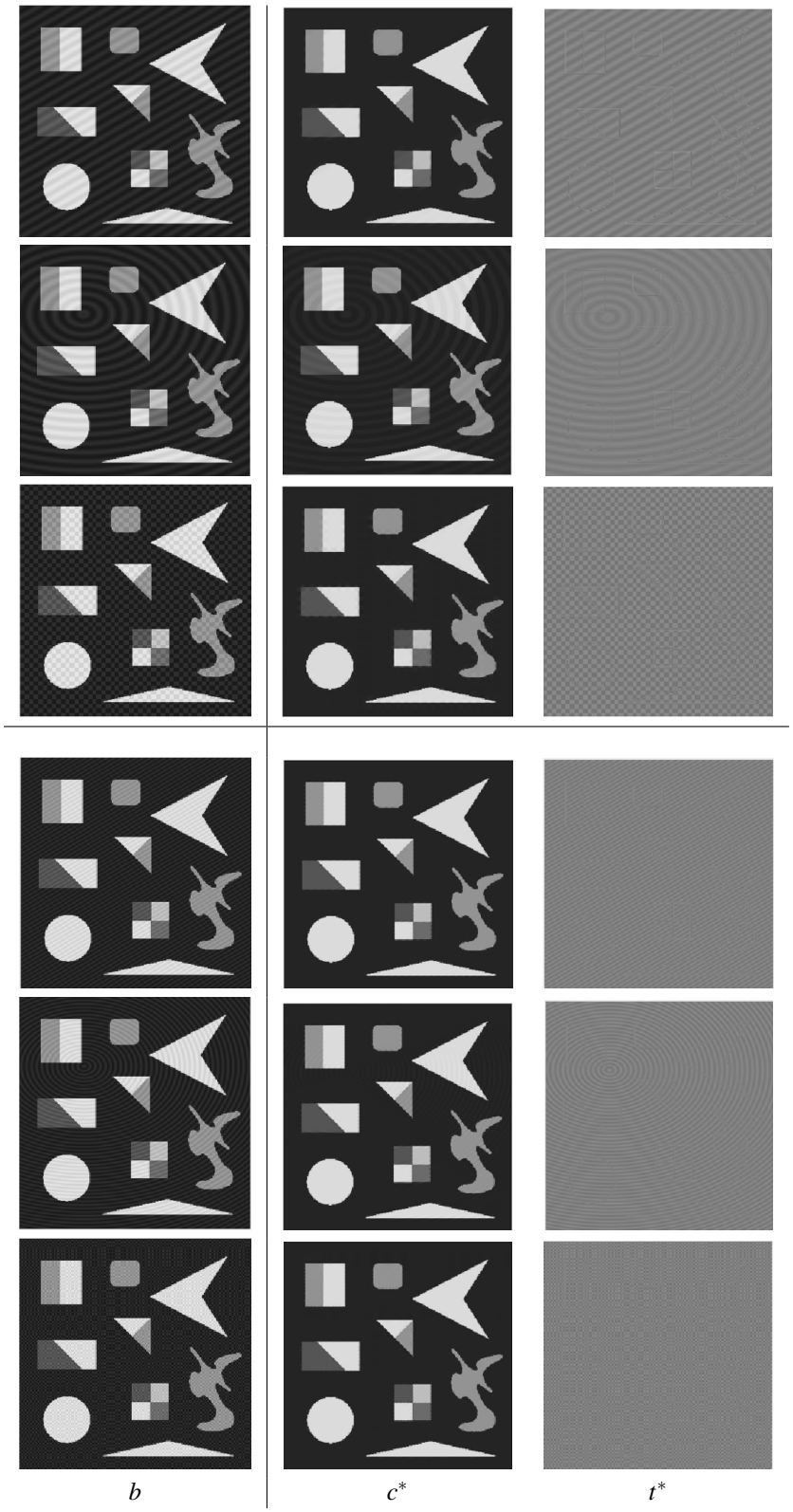

Figure 6: Example 1: decomposition of the `geometricR` images (first column) composed by blending with factor $f = 0.90$ a `geometric` cartoon component c with textured images t `linear`, `radial` and `chess`; with coarse (top block) and fine (bottom block) textured patterns.

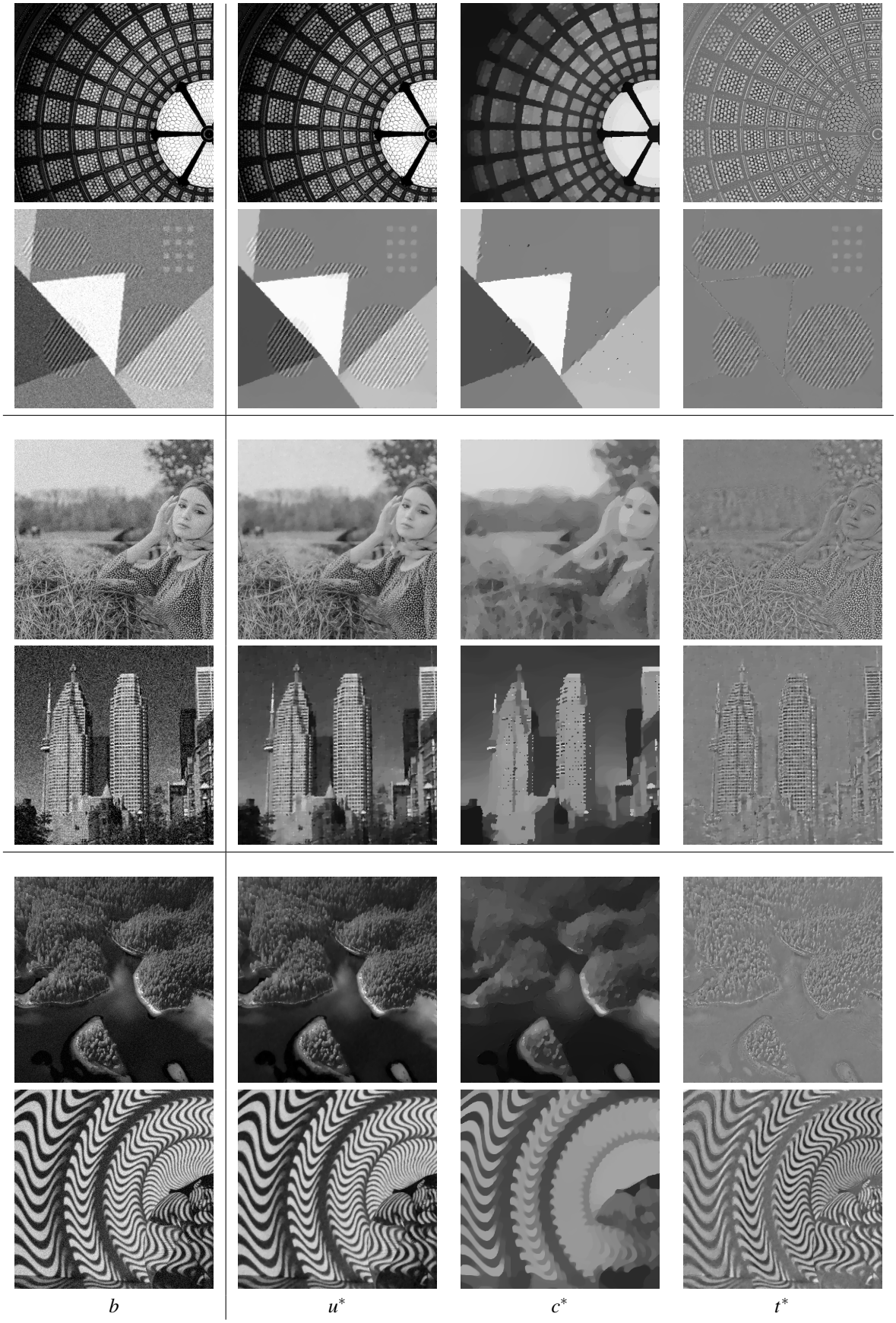

Figure 7: Example 2 - decomposition results for different types of additive white noise, namely AWL (first block), AWGN (second block) and AWU (third block), on test images - from top to bottom - `deco`, `geom`, `woman`, `sky`, `tree` and `girl`.

Table 3: Example 1: SNR and SSIM values of components c^* and t^* for different factors f and pattern grains - Fine (F) Coarse (C) - with `linear`, `radial` and `chess` textured images, given `geometric` as cartoon.

f	T	texture	c^*: SNR	t^*: SNR	c^*: SSIM	t^*: SSIM
0.6	C	lin	13.35	9.71	0.945	0.822
		rad	13.51	9.84	0.898	0.840
		chess	13.80	13.17	0.923	0.838
	F	lin	18.39	14.75	0.977	0.794
		rad	17.38	13.74	0.934	0.889
		chess	17.69	17.06	0.972	0.823
0.7	C	lin	17.39	9.91	0.968	0.860
		rad	16.79	9.28	0.925	0.842
		chess	16.99	12.52	0.948	0.853
	F	lin	21.54	14.06	0.986	0.891
		rad	20.22	12.74	0.946	0.909
		chess	20.26	15.79	0.982	0.881
0.8	C	lin	22.07	9.91	0.985	0.891
		rad	20.86	8.68	0.948	0.852
		chess	21.23	12.08	0.974	0.881
	F	lin	25.05	12.89	0.993	0.938
		rad	23.88	11.71	0.967	0.930
		chess	23.45	14.30	0.990	0.930
0.9	C	lin	27.40	8.20	0.995	0.923
		rad	24.72	5.49	0.940	0.800
		chess	26.94	10.75	0.991	0.928
	F	lin	30.99	11.79	0.997	0.967
		rad	29.92	10.73	0.988	0.955
		chess	28.95	12.75	0.996	0.961

(21)-(22). A theoretical proof of convergence of the algorithm is very challenging to be given. However, evidence of numerical convergence has emerged from all experiments in this section and, in particular, can be observed in Figure 5 for two specific examples. In fact, the blue curves shown in the bottom row represent the values of the normalized cross-correlation measure $C(\mu)$ in the higher-level minimization problem (21) and have been obtained by solving the lower-level problem (22) for a fine grid of different μ values. Hence, the global minimizer of these curves represents (approximately, due the discrete set of μ values considered) the theoretical solution μ^* of the higher-level problem (21). Since the blue circles in the same figures indicate the μ values selected by using our ADMM approach, this confirms the good convergence behaviour of our algorithm.

7.2. Example 2: performance of the two-stages decomposition framework

In what follows we focus on the joint decomposition model (Stage I + Stage II) solving the minimisation problems (2)-(3) and (4)-(5) on noise-corrupted photographic images. Following the degradation model (1) the images $u = c + t$ have been synthetically corrupted by additive white noises n of different types among uniform (AWUN), Gaussian (AWGN), Laplacian (AWLN). The corrupted images $b = u + n$, are shown in Fig. 7 (first column), and the resulting image components u^* (the denoised component), c^* (cartoon component), and t^* (textured parts), are reported in the remaining columns.

The images are characterized by a mixture of flat regions, neat edges and textured parts and they have been corrupted by adding a realization of white noise, from top to bottom: AWLN with standard deviation $\sigma \in \{2, 10\}$, AWGN with $\sigma \in \{15, 25\}$, AWUN with $\sigma \in \{10, 20\}$, respectively. All the experiments have been carried out by applying the proposed two-stages algorithm with automatic parameter selection, in pure non-convex regime ($\tau = 5.0$). The denoiser in Stage I solves the variational problem (2), without any a priori estimate on the noise distributions or specific knowledge on the content of the images. This model turned out to be a great denoiser for images with textured parts (see images u^* in Fig. 7, second column).

As concerns the Stage II, the cross-correlation principle works very well in simple cases as shown in synthetic images in Example 1, where a very clear distinction between texture and structure has been

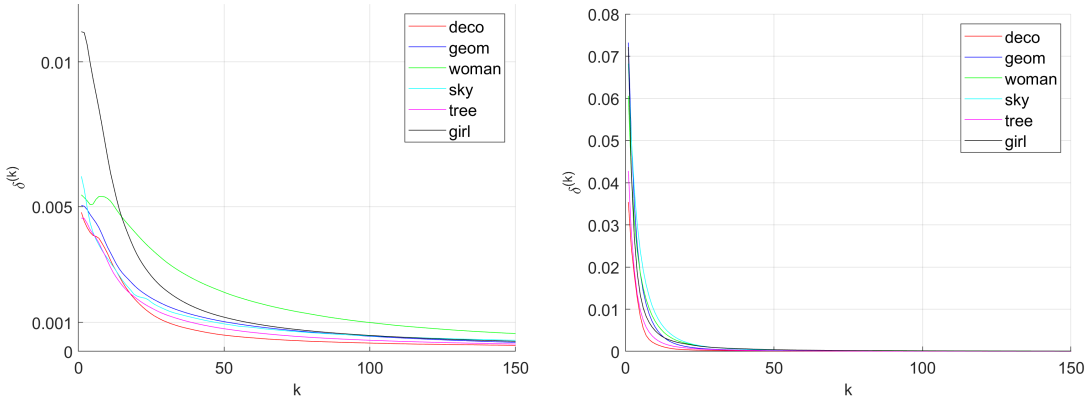

Figure 8: Example 2: plots of the relative change values $\delta^{(k)}$ of Stage I (left panel) and Stage II (right panel) ADMM iterates for the six numerical tests reported in Fig. 7.

obtained (see Fig. 5). In these cases $C(\mu)$ reaches a minimum approximately at the point where the texture is completely smoothed out from b, and then increases, when more of the structure gets into the c part. However, for photographic images, like those in Fig.7 where there are natural textures that present structures of different scales, the decomposition is a more challenging issue. In terms of correlation function, in fact, there is no more a single minimum and the function may oscillate. We adopted here the selection of μ which corresponds to the first local minimum reached by the correlation function.

From a visual inspection we can appreciate the clean decomposition obtained even in presence of severe corruptions and diversified texture patterns. The image in the last row of Fig. 7 contains textures of different granularity (waved strips). The textured part t^* naturally captured by our decomposition framework is the one with finer granularity, as already observed in Example 1.

In order to provide evidence for the good convergence behaviour of the ADMM-based iterative approaches proposed for the numerical solution of Stage I and II variational models, in Fig. 8 we report the relative change values $\delta^{(k)}$ - see definition in (56) - of iterates $u^{(k)}$ (ADMM for Stage I, left panel) and iterates $c^{(k)}$ (ADMM for Stage II, right panel) for all the six different numerical tests shown in Fig. 7.

7.3. Example 3: comparison with other variational decomposition models

In this last example we present the qualitative performance of the proposed decomposition approach with respect to other variational models which aim to separate an image into three components (structure, texture, and noise). Specifically, the original model in [2], based on total variation, G-norm, and Besov-norm (which amounts to a wavelet shrinkage), the adaptive model in [5], where differently weighted G-norms are used to separate noise from texture using a space-variant regularisation parameter, and, finally, the ternary variational decomposition introduced in [23], where a non-convex penalty function replaces the TV term, while G-norm, and Besov-norm are used to model the texture and noise components, respectively. To avoid introducing artifacts due to naive reproduction of the various algorithms and associated parameter setting within, the images compared and here reported in Fig. 9 and Fig. 10 were directly captured by the original papers.

In Fig. 9 the decomposition results are illustrated for a simple image b corrupted by AWGN with standard deviation $\sigma = 20$. The corresponding image components c^*, t^* and n^*, presented in [2] and [5] are illustrated in the second and third row, respectively. The latter model performs very well in texture detection, while suffering of a manual setting of the space-variant parameter map. Results obtained by our Stage I + Stage II algorithm are shown in Fig. 9 (first row). Stage I shows its effectiveness in removing noise without incorporating edges, and the CNC approach applied in solving Stage II seems to be particularly well suited to capture the piece-wise constant structure c^*.

A similar effect can be observed in Fig. 10 where our decomposition approach is compared with the results shown in [23]. The image b has been corrupted by AWGN with $\sigma = 20$. The difficulty in separating texture from noise by [23] is evident in the image n^* of the second line. Difficulty overcome by the proposed Stage I, as shown in image n^* in Fig. 10, first row.

Compared with these state-of-the-art image variational decomposition models, the proposed two-stage model yields the best qualitative performance, without neither requiring any critical parameter tuning, nor

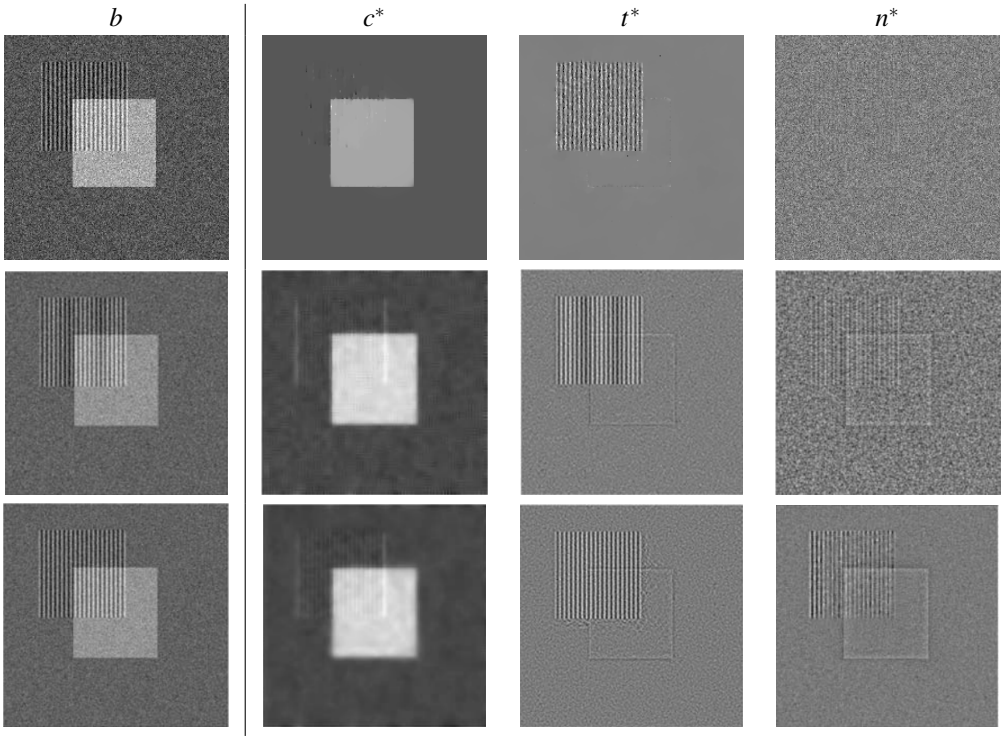

Figure 9: Example 3 - decompositions of images b (first column) into cartoon, texture and noise obtained by applying our method (first row), model in [2] (second row), model in [5] (third row)

assuming any knowledge of noise kind and variance, but taking advantage, instead, of an automatic parameter selection to achieve a sub-optimal texture - cartoon separation.

8. Conclusions

In this work we have presented a parameter-free two-stage variational model for the decomposition of images into cartoon, texture and noise components. A first peculiarity of the model is the capability to efficiently separate texture from noise, thanks to statistics based on whiteness and cross-correlation. The decomposition model does not require any knowledge on the noise distribution or standard deviation, which is a fundamental second distinguishing quality useful in practical contexts. And last, but not least, the parameter tuning is automatically performed on the basis of the morphology of the image, along the iterations of the numerical optimization procedure.

The trivial ℓ_2-norm texture characterization considered as penalty in the model (4) allows for a direct and explicit tuning of the regularization parameter μ and for a convex optimization due to the CNC numerical strategy. On the other hand, it could represent a weakness of the decomposition model in the sense that the underlying function space is very large, and the reconstructed t^* images could contain spurious fine-scale details. Future work will investigate other more suitable normed spaces that still allow to maintain an explicit relationship with the regularization parameter. Finally, to solve the two optimization problems in Stage I and Stage II, we designed ADMM-based algorithms which integrate the automatic computation of the involved parameters.

The results obtained are very promising, and motivated us to experiment and compare with other variational decomposition methods.

9. Acknowledgments

This work was supported in part by the National Group for Scientific Computation (GNCS-INDAM), Research Projects 2021, and in part by MIUR RFO projects.

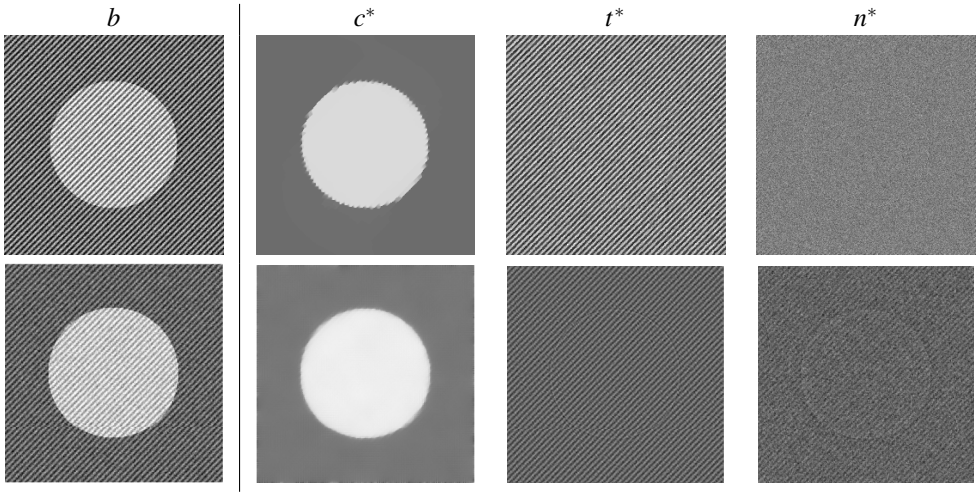

Figure 10: Example 3 - decompositions of images b (first column) into cartoon, texture and noise by applying our method (first row), and obtained by model in [23] (second row)

10. Conflict Of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] Aujol, J.F, Aubert, G., Blanc-Feraud, L., Chambolle,A.: Image Decomposition into a Bounded Variation Component and an Oscillating Component. J. Math. Imaging Vis. **22**, 71–88 (2005)

[2] Aujol, J.F, Chambolle,A.: Dual Norms and Image Decomposition Models. J. Math. Imaging Vis. **63**, 84–104 (2005)

[3] Aujol, J.F, Gilboa, G., Chan, T., Osher, S.: Structure-Texture Image Decomposition—Modeling, Algorithms, and Parameter Selection. Int J Comput Vision **67**, 111–136 (2006)

[4] Chan, R., Lanza, A., Morigi, S., Sgallari, F., Convex non-convex image segmentation, Numerische Mathematik, 2018, 138(3), 635–680

[5] Gilles, J.: Noisy Image Decomposition: A New Structure, Texture and Noise Model Based on Local Adaptivity. J. Math. Imaging Vis. **28**, 285–295 (2007)

[6] Gilles, J.: Noisy Multiscale Texture Separation. Multiscale Model. Simul. **10**, 1409–1427(2012)

[7] Gilles, J. and Meyer, Y.: Properties of BV-G structures + textures decomposition models. Application to road detection in satellite images. IEEE Trans. Image Process. **19**, 2793–2800 (2010)

[8] Huska, M., Lanza, A., Morigi, S., and Selesnick, I. A convex-nonconvex variational method for the additive decomposition of functions on surfaces. Inverse Problems, **35**, 124008 (2019).

[9] Huska, M., Kang, S. H., Lanza, A. and Morigi, S., A Variational Approach to Additive Image Decomposition into Structure, Harmonic, and Oscillatory Components, SIAM Journal on Imaging Sciences, 2021, **14**:4, 1749-1789

[10] Lanza, A., Morigi, and Sgallari, F., Variational Image Restoration with Constraints on Noise Whiteness. J. Math Imaging Vis **53**, 61–77 (2015).

[11] Lanza, A., Morigi, S., Sciacchitano, F. and Sgallari, F.. Whiteness Constraints in a Unified Variational Framework for Image Restoration. J. Math Imaging Vis **60**, 1503–1526 (2018).

[12] Lanza, A., Morigi, S. and Sgallari, F.: Convex image denoising via non-convex regularization with parameter selection. J. Math Imaging Vis **56**, 195—220 (2016).

[13] Lanza, A., Morigi, S., Selesnick, I.W., Sgallari, F., Sparsity-inducing nonconvex nonseparable regularization for convex image processing, SIAM Journal on Imaging Sciences, 2019, 12(2), 1099–1134.

[14] Lanza, A., Pragliola, M. and Sgallari, F. : Residual whiteness principle for parameter-free image restoration. Electron.Trans. Numer. Anal.,**53**, 329–351 (2020)

[15] A. Lanza, S. Morigi, I. W. Selesnick, F. Sgallari (2022). Convex non-Convex Variational Models, Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging, Springer.

[16] Malgouyres, F.,Minimizing the total variation under a general convex constraint for image restoration. IEEE Trans. Image Process., 11(12), 1450–1456, (2002).

[17] Meyer, Y. Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures. American Mathematical Society, USA, 2001.

[18] Osher, S.J., Sole, A., and Vese, L.A. 2003. Image decomposition and restoration using total variation minimization and the H^{-1} norm. SIAM Journal on Multiscale Modeling and Simulation, 1(3):349–370.

[19] T. M. Le and L. Vese. Image decomposition using total variation and div(bmo). CAM04-76, UCLA, 2004.

[20] L. A. Vese and S. J. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. J. Sci. Comput., 19(1-3):553–572, (2003).

[21] Rudin, L. I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D **60**, 259—268 (1992).

[22] J. -. Starck, M. Elad and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," in IEEE Transactions on Image Processing, vol. 14, no. 10, pp. 1570-1582, (2005)

[23] Tang, L., Wu, L., Fang, Z., Li, C. (2021). A non-convex ternary variational decomposition and its application for image denoising, IET Signal Processing.

[24] Vese, L and Osher, S.: Modeling textures with total variation minimization and oscillating patterns in image processing. J. Sci. Comput. **19**, 553–572 (2003)

[25] Zhang, C.H., Nearly unbiased variable selection under minimax concave penalty, The Annals of Statistics, Ann. Statist.,38, The Institute of Mathematical Statistics, 894–942 (2010)