Electronic Transactions on Numerical Analysis.
Volume 59, pp. 202–229, 2023.

ISSN 1068–9613.
DOI: 10.1553/etna_vol59s202

ETNA
Kent State University and
Johann Radon Institute (RICAM)

PARAMETER-FREE RESTORATION OF PIECEWISE SMOOTH IMAGES*

ALESSANDRO LANZA†, MONICA PRAGLIOLA‡, AND FIORELLA SGALLARI†

Abstract. We propose a novel strategy for the automatic estimation of the two regularization parameters arising in the image decomposition variational model employed for the restoration task when the underlying corrupting noise is known to be additive white Gaussian. In the model of interest, the target image is decomposed in its piecewise constant and smooth components, with a total variation term penalizing the former and a Tikhonov term acting on the latter. The proposed criterion, which relies on the whiteness property of the noise, extends the residual whiteness principle, originally introduced in the case of a single regularization parameter. The structure of the considered decomposition model allows for an efficient estimation of the pair of unknown parameters, that can be automatically adjusted along the iterations with the alternating direction method of multipliers employed for the numerical solution. The proposed multi-parameter residual whiteness principle is tested on different images with different levels of corruption. The performed tests highlight that the whiteness criterion is particularly effective and robust when moving from a single-parameter to a multi-parameter scenario.

Key words. image restoration, image decomposition, whiteness principle, ADMM

AMS subject classifications. 68U10, 94A08, 65K10.

1. Introduction.

In this paper, we are interested in the restoration of piecewise smooth images corrupted by blur and additive white Gaussian (AWG) noise. The image formation (or degradation) model considered can be formally written in vectorized form as follows

$$\boldsymbol{y} = \mathbf{K}\bar{\boldsymbol{x}} + \bar{\boldsymbol{\epsilon}} \qquad \text{with } \bar{\boldsymbol{\epsilon}} \text{ realisation of } \boldsymbol{\mathcal{E}} \sim \boldsymbol{\mathcal{G}}\left(\mathbf{0}_n, \sigma^2\mathbf{I}_n\right), \tag{1.1}$$

where $\bar{\boldsymbol{x}}, \bar{\boldsymbol{\epsilon}}, \boldsymbol{y} \in \mathbb{R}^n$ denote the vectorized (column-major) forms of the $n_1 \times n_2$ unknown target image, unknown AWG noise realisation and observed blur- and noise-corrupted image, respectively, with $n = n_1 n_2$ the total number of pixels. The n-dimensional random vector $\boldsymbol{\mathcal{E}}$ is Gaussian-distributed with null mean-vector and scalar covariance matrix ($\mathbf{0}_n$ and $\mathbf{I}_n$ denote the n-dimensional null vector and the $n \times n$ identity matrix, respectively), with σ indicating the noise standard deviation. Finally, $\mathbf{K} \in \mathbb{R}^{n\times n}$ is the coefficient matrix of a linear blurring operator, which we assume to be known and space-invariant, hence $\mathbf{K}$ is the coefficient matrix of a discrete 2-dimensional convolution operator.

Recovering $\bar{\boldsymbol{x}}$ starting from the knowledge of the observation $\boldsymbol{y}$, the blurring matrix $\mathbf{K}$ and, possibly, the noise standard deviation σ is typically an ill-conditioned (if not ill-posed) inverse problem, so that one rather seeks for an estimate $\widehat{\boldsymbol{x}}$ of the target image $\bar{\boldsymbol{x}}$ which solves a regularized, well-posed problem as close as possible to the original one. In the variational framework, the estimate $\widehat{\boldsymbol{x}}$ of $\bar{\boldsymbol{x}}$ in (1.1) is obtained as the solution of a general minimization problem which we refer to as the $\mathcal{R}$-L_2 class of variational models and typically takes the form

$$\begin{aligned} \widehat{\boldsymbol{x}}(\boldsymbol{\mu}) &\in \operatorname*{argmin}_{\boldsymbol{x}\in\mathbb{R}_n} \left\{\mathcal{J}(\boldsymbol{x};\boldsymbol{\mu}) := \mathcal{R}(\boldsymbol{x};\boldsymbol{\mu}) + \mathrm{L}_2(\boldsymbol{x};\boldsymbol{y},\mathbf{K})\right\}, \\ \mathrm{L}_2(\boldsymbol{x};\boldsymbol{y},\mathbf{K}) &:= \frac{1}{2}\left\|\mathbf{K}\boldsymbol{x} - \boldsymbol{y}\right\|_2^2, \end{aligned} \tag{$\mathcal{R}$-L_2}$$

where the cost (or energy) functional $\mathcal{J} : \mathbb{R}^n \to \mathbb{R}$ is made by the sum of two terms. The quadratic L_2 *data fidelity term* comes deductively from applying the maximum likelihood

*Received December 19, 2022. Accepted July 20, 2023. Published online on October 30, 2023. Recommended by A. Buccini.

†Department of Mathematics, University of Bologna, Italy
({alessandro.lanza2, fiorella.sgallari}@unibo.it).

‡Department of Mathematics and Applications, University of Naples Federico II, Italy
(monica.pragliola@unina.it).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

approach to the estimation of $\bar{\boldsymbol{x}}$ in (1.1), after explicitly considering the Gaussian distribution of the additive noise corruption. On the other hand, the *regularization term* $\mathcal{R}(\boldsymbol{x};\boldsymbol{\mu})$ encodes information or beliefs that may be available a priori on the target image $\bar{\boldsymbol{x}}$, such as smoothness or sparsity properties, and in general depends on a vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p) \in \mathbb{R}^p_{++}$ of free parameters called the *regularization parameters* (in the apaper, we denote by $\mathbb{R}_+$ and $\mathbb{R}_{++}$ the sets of non-negative and positive real numbers, respectively).

One of the most popular regularization terms in image processing is the Total Variation (TV) semi-norm [15], which corresponds to

$$\mathcal{R}(\boldsymbol{x};\mu) = \mu \mathrm{TV}(\boldsymbol{x}) := \mu \sum_{i=1}^{n} \|(\boldsymbol{\nabla x})_i\|_2, \tag{TV}$$

where $(\boldsymbol{\nabla x})_i \in \mathbb{R}^2$ represents the discrete gradient of image $\boldsymbol{x}$ computed at pixel i. The TV term, which induces sparsity of gradient magnitudes, is known to be particularly effective for the restoration of piecewise constant images; however, it is also well-established that the TV regularizer tends to promote edges thus producing the so-called *staircasing effect* on the smooth parts of the image.

As a way to partially overcome the classical drawbacks of TV, one can employ the TV_2 regularizer (see, e.g., [5]) defined by

$$\mathcal{R}(\boldsymbol{x};\mu) = \mu \mathrm{TV}_2(\boldsymbol{x}) := \mu \sum_{i=1}^{n} \|(\boldsymbol{\nabla}^2 \boldsymbol{x})_i\|_F, \tag{TV_2}$$

with $(\boldsymbol{\nabla}^2 \boldsymbol{x})_i \in \mathbb{R}^{2\times 2}$ indicating the discrete Hessian of image $\boldsymbol{x}$ at pixel i and $\|\cdot\|_F$ denoting the Frobenius norm. The TV_2 regularizer promotes piecewise-affine structures in the image, however its ability to recover sharp edges is less than TV.

A widely used regularizer aimed to promote smoothness of the restored image is the TV_2^2 regularizer defined by

$$\mathcal{R}(\boldsymbol{x};\mu) = \mu \mathrm{TV}_2^2(\boldsymbol{x}) := \mu \sum_{i=1}^{n} \|(\boldsymbol{\nabla}^2 \boldsymbol{x})_i\|_F^2. \tag{TV_2^2}$$

All the (TV), (TV_2) and (TV_2^2) regularization terms depend on one scalar regularization parameter $\mu \in \mathbb{R}_{++}$ only and are inherently unable to deal effectively with the restoration of piecewise smooth images.

An effective regularizer for this class of images has been proposed in [7] and relies on the explicit assumption that the target unknown image $\bar{x}$ in (1.1) is given by the sum of a piecewise constant (or cartoon) component $\bar{\boldsymbol{c}}$ and a smooth component $\bar{\boldsymbol{s}}$; in formulas,

$$\bar{\boldsymbol{x}} = \bar{\boldsymbol{c}} + \bar{\boldsymbol{s}}. \tag{1.2}$$

Based on (1.2), the authors in [7] propose a "composite" regularizer - that we refer to as TV-TV_2^2 - given by the weighted sum of a TV regularizer for the cartoon component and a TV_2^2 regularizer for the smooth component, with the weights $\mu_1, \mu_2 \in \mathbb{R}_{++}$ regarded as two free regularization parameters:

$$\mathcal{R}(\boldsymbol{x};\mu_1,\mu_2) = \mu_1 \mathrm{TV}(\boldsymbol{c}) + \mu_2 \mathrm{TV}_2^2(\boldsymbol{s}). \tag{TV-TV_2^2}$$

After introducing the two first- and second-order differential matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ defined by

$$\mathbf{D}_1 = (\mathbf{D}_h; \mathbf{D}_v) \in \mathbb{R}^{2n\times n}, \quad \mathbf{D}_2 = (\mathbf{D}_{hh}; \mathbf{D}_{vv}; \mathbf{D}_{hv}; \mathbf{D}_{vh}) \in \mathbb{R}^{4n\times n}, \tag{1.3}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with $\mathbf{D}_h, \mathbf{D}_v, \mathbf{D}_{hh}, \mathbf{D}_{vv}, \mathbf{D}_{hv}, \mathbf{D}_{vh} \in \mathbb{R}^{n\times n}$ finite difference matrices discretizing the first-order partial derivatives of the vectorized $n_1 \times n_2$ image $\boldsymbol{c} \in \mathbb{R}^n$ in the horizontal and vertical direction and the second-order partial derivatives of the vectorized $n_1 \times n_2$ image $\boldsymbol{s} \in \mathbb{R}^n$ in the horizontal, vertical, mixed horizontal-vertical and mixed vertical-horizontal directions (with $\mathbf{D}_{vh} = \mathbf{D}_{hv}$), respectively, the restoration-by-decomposition variational model proposed in [7], which we will henceforth refer to RBD-L_2 model, reads

$$\widehat{\boldsymbol{x}}\,(\mu_1,\mu_2) = \widehat{\boldsymbol{c}}\,(\mu_1,\mu_2) + \widehat{\boldsymbol{s}}\,(\mu_1,\mu_2)\,, \tag{1.4}$$

$$\{\widehat{\boldsymbol{c}}\,(\mu_1,\mu_2)\,,\widehat{\boldsymbol{s}}\,(\mu_1,\mu_2)\} \in \underset{\boldsymbol{c},\boldsymbol{s}\in\mathbb{R}^n}{\operatorname{argmin}} \mathcal{J}\,(\boldsymbol{c},\boldsymbol{s};\mu_1,\mu_2)\,, \tag{1.5}$$

$$\mathcal{J}\,(\boldsymbol{c},\boldsymbol{s};\mu_1,\mu_2) = \sum_{i=1}^{n} \|(\mathbf{D}_1\boldsymbol{c})_i\|_2 + \frac{\mu_1}{2}\,\|\mathbf{D}_2\boldsymbol{s}\|_2^2 + \frac{\mu_2}{2}\,\|\mathbf{K}\,(\boldsymbol{c}+\boldsymbol{s}) - \boldsymbol{y}\|_2^2\,, \tag{1.6}$$

where, with a little abuse of notation, we indicate by $(\mathbf{D}_1\boldsymbol{c})_i := \big((\mathbf{D}_h\boldsymbol{c})_i;(\mathbf{D}_v\boldsymbol{c})_i\big) \in \mathbb{R}^2$ the discrete gradient of image $\boldsymbol{c}$ at pixel i and where the dependence of the solution image $\widehat{\boldsymbol{x}}$, the solution components $\widehat{\boldsymbol{c}}$, $\widehat{\boldsymbol{s}}$ and the cost function $\mathcal{J}$ on the two regularization parameters μ_1, μ_2 has been made explicit. We notice that in (1.6) the parameters μ_1 and μ_2 multiply the TV_2^2 regularizer and the fidelity term, respectively, rather than the TV and TV_2^2 regularizers as would follow from the definition of the (TV-TV_2^2) regularizer. This choice, although it clearly leads to a completely equivalent RBD-L_2 variational model, is crucial for the automatic parameter selection strategy proposed in this paper.

Besides the choice of a suitable regularization term $\mathcal{R}(\boldsymbol{x};\boldsymbol{\mu})$, the selection of the regularization parameter vector $\boldsymbol{\mu}$ can strongly influence the quality of the output restoration provided by the ($\mathcal{R}$-L_2) class of variational models, hence also by the RBD-L_2 model (1.4)–(1.6). For the case of a single scalar regularization parameter μ - that is, when the (TV), (TV_2) or (TV_2^2) regularizers are used - some effective selection criteria have been proposed in literature; cf. [4, 9]. The most popular one is the *discrepancy principle* (DP) (cf. [8, 13]) which sets μ so that the value of the quadratic data fidelity term in the ($\mathcal{R}$-L_2) model is equal to a prescribed value depending on the number of pixels and on the noise standard deviation. In formula:

$$\text{(DP)} \qquad \text{Select } \ \mu = \widehat{\mu} \ \text{ such that } \ \|\widehat{\boldsymbol{r}}(\mu)\|_2 = \sqrt{n}\sigma, \ \text{ with } \ \widehat{\boldsymbol{r}}(\mu) := \mathbf{K}\widehat{\boldsymbol{x}}(\mu) - \boldsymbol{b}$$

indicating the μ-dependent residual image. Recently another selection criterion, called the *residual whiteness principle* (RWP), has been proposed first for the case of images corrupted by AWG noise [12, 14], then also for Poisson noise-corrupted images [1, 2]. The idea is to select the μ-value yielding a residual image $\widehat{\boldsymbol{r}}(\mu)$ (or a standardized version of it for the Poisson noise case) which most resembles the realization of a white noise random process. The RWP not only outperforms the DP but also does not require the knowledge of AWG noise standard deviation σ. According to the definition in [12], the RWP applied to the selection of a single regularization parameter μ in the ($\mathcal{R}$-L_2) class of restoration models - which, we recall, is suitable for AWG noise corruptions - can be formulated as follows:

$$\text{(RWP)} \qquad \text{Select } \ \mu = \widehat{\mu} \ \text{ such that } \ \widehat{\mu} \in \underset{\mu\in\mathbb{R}_{++}}{\operatorname{argmin}} W(\mu), \quad W(\mu) = \frac{\|\widehat{\boldsymbol{r}}(\mu) * \widehat{\boldsymbol{r}}(\mu)\|_2^2}{\|\widehat{\boldsymbol{r}}(\mu)\|_2^4},$$

where $*$ denotes the discrete 2-dimensional correlation operator and the scalar function W measures whiteness of the μ-dependent residual image $\widehat{\boldsymbol{r}}(\mu)$ by means of its normalized auto-correlation—see [12] for details. However, the RWP in its existing form does not apply to the case of more than one regularization parameter and thus can not be used for the joint

ETNA
Kent State University and
Johann Radon Institute (RICAM)

automatic selection of μ_1, μ_2 in the RBD-L_2 model (1.4)–(1.6) for the restoration of piecewise smooth images, which is the focus of this work.

The contribution of this paper can be summarized as follows:

(i) we analyze the original RBD-L_2 model (1.4)–(1.6) proposed in [7] and prove that the associated minimization problem (1.5)–(1.6) is not well-posed, as it admits an infinity of solutions. Then, we demonstrate that (1.4)–(1.6) is equivalent to a linearly constrained model of reduced dimensionality, whose associated minimization problem is well-posed and can be solved numerically in a faster way;

(ii) we extend for the first time the RWP proposed in [12] to the case of more than one regularization parameter. In particular, we formalize a multi-parameter RWP (MRWP) for the selection of the pair (μ_1, μ_2) of regularization parameters in the reduced RBD-L_2 model;

(ii) we propose a numerical optimization approach for the solution of the reduced RBD-L_2 model based on the Alternating Direction Method of Multipliers (ADMM) which allows to apply the MRWP along the ADMM iterations. This makes the proposed MRWP fully automatic;

(iv) from an applicative point of view, we propose and experimentally validate a novel, fully automatic and efficient approach for the restoration of piecewise smooth images corrupted by AWG noise, which does not require to know the noise standard deviation.

The paper is organized as follows. In Section 2 we analyze the original RBD-L_2 model and then introduce its equivalent constrained version of reduced dimensionality. In Section 3 we show how the multi-parameter RWP admits an explicit formulation for models in which a quadratic regularizer is coupled with an L_2 fidelity term and, then, in Section 4 we exploit this result to illustrate how the application of the MRWP to the reduced RBD-L_2 model can be embedded along the iterations of a suitable ADMM-based optimization scheme. In Section 5 we experimentally test the overall proposed fully automatic approach for the restoration of piecewise smooth images and, finally, in Section 6 we draw conclusions.

Notations. Throughout the paper, we denote by $\mathbb{R}_+$ the set of nonnegative real numbers, by $\mathbb{R}_{++} = \mathbb{R}_+ \setminus \{0\}$ the set of positive real numbers, by $\mathbf{1}_n$ the n-dimensional column vector of all ones and by $\iota_{\mathcal{S}}$ the indicator function of set $\mathcal{S}$, with $\iota_{\mathcal{S}}(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \mathcal{S}$, $+\infty$ otherwise. We denote by $\mathbf{F}$, $\mathbf{F}^* \in \mathbb{C}^{n\times n}$ the unitary matrices representing the 2D discrete Fourier transform operator and its conjugate transpose (i.e., its inverse) applied to vectorized $n_1 \times n_2$ images. The action of $\mathbf{F}$ on a vectorized image $\boldsymbol{u}$ is denoted by $\widetilde{\boldsymbol{u}}$ (i.e., $\widetilde{\boldsymbol{u}} = \mathbf{F}\boldsymbol{u}$ is the Fourier-transformed version of $\boldsymbol{u}$), and for a complex number c we indicate by $|c|$ and $\overline{c}$ its modulus and conjugate, respectively.

2. Analysis and reformulation of the RBD-L_2 model. In this section, also based on similar results in [10, 11], we first analyze the original RBD-L_2 model (1.4)–(1.6) proposed in [7] and prove that the associated minimization problem (1.5)–(1.6) is not well-posed, as it admits an infinity of solution components. Then, we demonstrate that (1.4)–(1.6) is equivalent to a linearly constrained model with associated well-posed minimization problem of reduced dimensionality (dimension n instead of $2n$).

Some of the reported results (or their proof) can depend on the discretization choices for the first- and second-order differential matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ in (1.3). However, analogous results could be obtained in a similar manner for other discretization schemes. Here, we adopt unscaled forward finite difference discretizations for the first-order horizontal and vertical partial derivatives (i.e., for matrices $\mathbf{D}_h$ and $\mathbf{D}_v$), unscaled centered finite difference discretizations for the second-order horizontal and vertical partial derivatives (i.e., for matrices $\mathbf{D}_{hh}$ and $\mathbf{D}_{vv}$) and unscaled forward finite difference discretizations for the second-order mixed horizontal-vertical partial derivatives (i.e., for matrices $\mathbf{D}_{hv} = \mathbf{D}_{vh}$). More formally,

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the matrices $\mathbf{D}_h, \mathbf{D}_v, \mathbf{D}_{hh}, \mathbf{D}_{vv}, \mathbf{D}_{hv} \in \mathbb{R}^{n\times n}$ in (1.3) are 2D discrete convolution matrices acting on vectorized $n_1 \times n_2$ images and defined by the following point spread functions (or convolution kernels):

$$\begin{gathered}\mathbf{D}_h\colon (+1, \boldsymbol{-1})\,, \quad \mathbf{D}_v\colon \begin{pmatrix} +1 \\ \boldsymbol{-1} \end{pmatrix}, \\ \mathbf{D}_{hh}\colon (+1, \boldsymbol{-2}, +1)\,, \quad \mathbf{D}_{vv}\colon \begin{pmatrix} +1 \\ \boldsymbol{-2} \\ +1 \end{pmatrix}, \quad \mathbf{D}_{hv}\colon \begin{pmatrix} +1 & -1 \\ -1 & \boldsymbol{+1} \end{pmatrix},\end{gathered} \tag{2.1}$$

with boldface cells indicating the center of application of the point spread functions.

The space-invariant blurring matrix $\mathbf{K}$ is also a convolution matrix and can be very ill-conditioned or even singular, hence in subsequent analyses we allow the null space of $\mathbf{K}$ to be non-trivial. However, all results reported in this section are valid under the following assumption, that is satisfied in most practical situations and, hence, we assume to hold true.

Assumption. The null spaces of the space-invariant blurring matrix $\mathbf{K} \in \mathbb{R}^{n\times n}$ in (1.1) and of the finite difference matrices $\mathbf{D}_1 \in \mathbb{R}^{2n\times n}$, $\mathbf{D}_2 \in \mathbb{R}^{4n\times n}$ in (1.3) satisfy

$$\operatorname{null}(\mathbf{K}) \cap \operatorname{null}(\mathbf{D}_1) \;=\; \operatorname{null}(\mathbf{K}) \cap \operatorname{null}(\mathbf{D}_2) \;=\; \{\mathbf{0}_n\}\,. \tag{2.2}$$

Moreover, we adopt periodic boundary conditions for all convolution matrices $\mathbf{K}$, $\mathbf{D}_h$, $\mathbf{D}_v$, $\mathbf{D}_{hh}$, $\mathbf{D}_{vv}$, and $\mathbf{D}_{hv}$ which, hence, are all block-circulant with circulant blocks (from now on, BCCB) matrices and can be diagonalized by the 2D discrete Fourier transform; in formula

$$\begin{aligned} \widetilde{\mathbf{K}} &= \mathbf{F}\mathbf{K}\mathbf{F}^*, & \widetilde{\mathbf{D}}_h &= \mathbf{F}\mathbf{D}_h\mathbf{F}^*, & \widetilde{\mathbf{D}}_v &= \mathbf{F}\mathbf{D}_v\mathbf{F}^*, \\ \widetilde{\mathbf{D}}_{hh} &= \mathbf{F}\mathbf{D}_{hh}\mathbf{F}^*, & \widetilde{\mathbf{D}}_{vv} &= \mathbf{F}\mathbf{D}_{vv}\mathbf{F}^*, & \widetilde{\mathbf{D}}_{hv} &= \mathbf{F}\mathbf{D}_{hv}\mathbf{F}^*, \end{aligned} \tag{2.3}$$

with the diagonal matrices $\widetilde{\mathbf{K}}$, $\widetilde{\mathbf{D}}_h$, $\widetilde{\mathbf{D}}_v$, $\widetilde{\mathbf{D}}_{hh}$, $\widetilde{\mathbf{D}}_{vv}$ and $\widetilde{\mathbf{D}}_{hv} \in \mathbb{C}^{n\times n}$ defined by

$$\begin{aligned} \widetilde{\mathbf{K}} &= \operatorname{diag}\big(\widetilde{k}_1, \ldots, \widetilde{k}_n\big), & \widetilde{\mathbf{D}}_h &= \operatorname{diag}\big(\widetilde{d}_{h,1}, \ldots, \widetilde{d}_{h,n}\big), \\ \widetilde{\mathbf{D}}_v &= \operatorname{diag}\big(\widetilde{d}_{v,1}, \ldots, \widetilde{d}_{v,n}\big), & \widetilde{\mathbf{D}}_{hh} &= \operatorname{diag}\big(\widetilde{d}_{hh,1}, \ldots, \widetilde{d}_{hh,n}\big), \\ \widetilde{\mathbf{D}}_{vv} &= \operatorname{diag}\big(\widetilde{d}_{vv,1}, \ldots, \widetilde{d}_{vv,n}\big), & \widetilde{\mathbf{D}}_{hv} &= \operatorname{diag}\big(\widetilde{d}_{hv,1}, \ldots, \widetilde{d}_{hv,n}\big). \end{aligned} \tag{2.4}$$

To simplify the notations, we introduce the total optimization variable $\boldsymbol{t} := (\boldsymbol{c}; \boldsymbol{s}) \in \mathbb{R}^{2n}$ of the RBD-L_2 model.

In Lemma 2.1 we analyze the two finite difference matrices $\mathbf{D}_1$, $\mathbf{D}_2$ defined in (1.3), (2.1) and other four newly introduced matrices that will be of interest later. In Lemma 2.2 we analyze the cost function $\mathcal{J}$ in (1.6) and, then, in Proposition 2.3 we outline some important properties of the original RBD-L_2 model (1.4)–(1.6). Finally, in Proposition 2.4 we state the equivalence of the original RBD-L_2 model with a linearly constrained version of it having reduced dimensionality. To improve readability, the proofs of Lemmas 2.1, 2.2, and of Proposition 2.4 are postponed to Appendix A.

LEMMA 2.1. *Let $\mathbf{D}_1 \in \mathbb{R}^{2n\times n}$ and $\mathbf{D}_2 \in \mathbb{R}^{4n\times n}$ be the finite difference matrices in (1.3) with convolution kernels in (2.1), let $\mathbf{K} \in \mathbb{R}^{n\times n}$ be a space-invariant blurring matrix satisfying (2.2) and let $\mathbf{M}_1(\mu_1, \mu_2), \mathbf{M}_2(\mu_1, \mu_2), \mathbf{A}(\mu_1, \mu_2), \mathbf{B}(\mu_1, \mu_2) \in \mathbb{R}^{n\times n}$ be the (μ_1, μ_2)-dependent matrices defined by*

$$\mathbf{M}_1(\mu_1, \mu_2) = (\mu_1/\mu_2)\mathbf{D}_2^{\mathrm{T}}\mathbf{D}_2 + \mathbf{K}^{\mathrm{T}}\mathbf{K}, \tag{2.5}$$

$$\mathbf{M}_2(\mu_1, \mu_2) = -\mathbf{M}_1^{-1}(\mu_1, \mu_2)\mathbf{K}^{\mathrm{T}}\mathbf{K}, \tag{2.6}$$

$$\mathbf{A}(\mu_1, \mu_2) = \mathbf{I}_n + \mathbf{M}_2(\mu_1, \mu_2), \tag{2.7}$$

$$\mathbf{B}(\mu_1, \mu_2) = \mu_2\mathbf{A}(\mu_1, \mu_2)\mathbf{K}^{\mathrm{T}}\mathbf{K}. \tag{2.8}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Then, under the assumption of periodic boundary conditions, for any $(\mu_1, \mu_2) \in \mathbb{R}^2_{++}$ *the above matrices* $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{A}$, $\mathbf{B}$ *are all symmetric, BCCB matrices, with* $\mathbf{M}_1$ *positive definite,* $\mathbf{M}_2$ *negative semidefinite,* $\mathbf{A}$ *and* $\mathbf{B}$ *positive semidefinite. Moreover, the null spaces of the matrices* $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{A}$, $\mathbf{B}$ *satisfy*

$$\text{null}\,(\mathbf{D}_1) = \text{null}\,(\mathbf{D}_2) = \text{null}\,(\mathbf{A}(\mu_1, \mu_2)) = \text{span}\,(\mathbf{1}_n) \qquad \forall (\mu_1, \mu_2) \in \mathbb{R}^2_{++}, \tag{2.9}$$

$$\text{null}\,(\mathbf{M}_1(\mu_1, \mu_2)) = \{\mathbf{0}_n\} \qquad \forall (\mu_1, \mu_2) \in \mathbb{R}^2_{++}, \tag{2.10}$$

$$\text{null}\,(\mathbf{M}_2(\mu_1, \mu_2)) = \text{null}\,(\mathbf{B}(\mu_1, \mu_2)) = \text{null}\,(\mathbf{K}) \qquad \forall (\mu_1, \mu_2) \in \mathbb{R}^2_{++}. \tag{2.11}$$

LEMMA 2.2. *For any data* $\boldsymbol{y} \in \mathbb{R}^n$ *and any* $(\mu_1, \mu_2) \in \mathbb{R}^2_{++}$, *the cost function* $\mathcal{J}$ *in* (1.6) *is proper, continuous, convex, bounded below by zero and non-coercive jointly in* $\boldsymbol{t} = (\boldsymbol{c}; \boldsymbol{s})$. *In fact,* $\mathcal{J}$ *is constant along straight lines in its domain* $\mathbb{R}^{2n}$ *of direction defined by the vector*

$$\boldsymbol{d} := (2n)^{-1/2}\,(\mathbf{1}_n; -\mathbf{1}_n)\,. \tag{2.12}$$

However, the restriction of $\mathcal{J}$ *to any hyperplane* $\mathcal{H}^{\theta}_{\boldsymbol{v}} \subset \mathbb{R}^{2n}$ *not parallel to* $\boldsymbol{d}$,

$$\mathcal{H}^{\theta}_{\boldsymbol{v}} := \left\{\boldsymbol{t} \in \mathbb{R}^{2n} : \boldsymbol{v}^{\mathrm{T}}\boldsymbol{t} = \theta\right\}, \quad \textit{with} \quad \theta \in \mathbb{R} \quad \textit{and} \quad \boldsymbol{v} \in \mathbb{R}^{2n} : \boldsymbol{v}^{\mathrm{T}}\boldsymbol{d} \neq 0, \tag{2.13}$$

is coercive, hence $\mathcal{J}$ *admits (an infinity of) global minimizers over* $\mathbb{R}^{2n}$.

PROPOSITION 2.3. *For any data* $\boldsymbol{y} \in \mathbb{R}^n$ *and any* $(\mu_1, \mu_2) \in \mathbb{R}^2_{++}$, *the RBD-L*$_2$ *model proposed in [7] and defined in* (1.4)–(1.6) *admits a unique solution image* $\widehat{\boldsymbol{x}}(\mu_1, \mu_2)$, *however the minimization problem* (1.5)–(1.6) *admits an infinity of pairs of solution components*

$$\widehat{\boldsymbol{c}}_{\alpha}(\mu_1, \mu_2) = \widehat{\boldsymbol{c}}(\mu_1, \mu_2) + \alpha\mathbf{1}_n, \quad \widehat{\boldsymbol{s}}_{\alpha}(\mu_1, \mu_2) = \widehat{\boldsymbol{s}}(\mu_1, \mu_2) - \alpha\mathbf{1}_n, \quad \alpha \in \mathbb{R}, \tag{2.14}$$

with $(\widehat{\boldsymbol{c}}\,(\mu_1, \mu_2)\,; \widehat{\boldsymbol{s}}\,(\mu_1, \mu_2))$ *denoting any particular global minimizer of* $\mathcal{J}$ *in* (1.6).

Proof. The infinitely many pairs of solution components as defined in (2.14) comes in a straightforward manner from statement of Lemma 2.2, in particular from $\mathcal{J}$ being constant along straight lines of direction $\boldsymbol{d}$ given in (2.12). Then, it is immediate to note that the infinitely many solution components in (2.14) leads to a unique solution image. In fact, according to (1.4), we have

$$\widehat{\boldsymbol{x}}(\mu_1, \mu_2) = \widehat{\boldsymbol{c}}_{\alpha}(\mu_1, \mu_2) + \widehat{\boldsymbol{s}}_{\alpha}(\mu_1, \mu_2) = \widehat{\boldsymbol{c}}(\mu_1, \mu_2) + \widehat{\boldsymbol{s}}(\mu_1, \mu_2)\ \forall \alpha \in \mathbb{R}. \qquad \square$$

PROPOSITION 2.4. *For any data* $\boldsymbol{y} \in \mathbb{R}^n$ *and any* $(\mu_1, \mu_2) \in \mathbb{R}^2_{++}$, *the RBD-L*$_2$ *model proposed in [7] and defined in* (1.4)–(1.6) *is equivalent—that is, it admits the same solution image* $\widehat{\boldsymbol{x}}\,(\mu_1, \mu_2)$—*to the following linearly constrained, reduced one:*

$$\widehat{\boldsymbol{x}}\,(\mu_1, \mu_2) = \mathbf{A}\,(\mu_1, \mu_2)\,\widehat{\boldsymbol{c}}\,(\mu_1, \mu_2) + \boldsymbol{a}\,(\mu_1, \mu_2)\,, \tag{2.15}$$

$$\widehat{\boldsymbol{c}}\,(\mu_1, \mu_2) = \underset{\boldsymbol{c} \in \mathbb{R}^n}{\text{argmin}}\, \mathcal{J}_{\boldsymbol{c}}\,(\boldsymbol{c}; \mu_1, \mu_2)\,, \tag{2.16}$$

$$\mathcal{J}_c\,(\boldsymbol{c}; \mu_1, \mu_2) = \sum_{i=1}^{n} \|(\mathbf{D}_1\boldsymbol{c})_i\|_2 + \boldsymbol{c}^{\mathrm{T}}\mathbf{B}\,(\mu_1, \mu_2)\,\boldsymbol{c} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{b}\,(\mu_1, \mu_2) + \iota_{\mathcal{C}}(\boldsymbol{c}), \tag{2.17}$$

with matrices $\mathbf{A}\,(\mu_1, \mu_2)\,, \mathbf{B}\,(\mu_1, \mu_2) \in \mathbb{R}^{n \times n}$ *defined in* (2.5)–(2.8), *with column vectors* $\boldsymbol{a}(\mu_1, \mu_2), \boldsymbol{b}(\mu_1, \mu_2) \in \mathbb{R}^n$ *given by*

$$\boldsymbol{a}(\mu_1, \mu_2) = \mathbf{M}_1^{-1}(\mu_1, \mu_2)\mathbf{K}^{\mathrm{T}}\boldsymbol{y}, \tag{2.18}$$

$$\boldsymbol{b}(\mu_1, \mu_2) = \mu_2\mathbf{A}(\mu_1, \mu_2)\mathbf{K}^{\mathrm{T}}\boldsymbol{y}, \tag{2.19}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and with $\mathcal{C} \subset \mathbb{R}^n$ the $(n-1)$-dimensional linear subspace defined by

$$\mathcal{C} = \left\{ \boldsymbol{c} \in \mathbb{R}^n : \sum_{i=1}^{n} c_i = 0 \right\}. \tag{2.20}$$

As $\mathcal{C}$ can be seen as a particular hyperplane of the class $\mathcal{H}^{\theta}_{\boldsymbol{v}}$ in (2.13), with $\theta = 0$ and $\boldsymbol{v} = (\mathbf{1}_n; \mathbf{0}_n)$, then the convex minimization problem (2.16)–(2.17) admits a unique solution.

REMARK 2.5. We remark that uniqueness of the solution of problem (2.16)–(2.17) holds true by constraining the optimization variable $\boldsymbol{c}$ to belong to any set in the class

$$\mathcal{C}^{\theta}_{\boldsymbol{v}} := \left\{ \boldsymbol{c} \in \mathbb{R}^n : \boldsymbol{v}^{\mathrm{T}} \boldsymbol{c} = \theta \right\}, \quad \text{with} \quad \theta \in \mathbb{R} \quad \text{and} \quad \boldsymbol{v} \in \mathbb{R}^n : \boldsymbol{v}^{\mathrm{T}} \mathbf{1}_n \neq 0. \tag{2.21}$$

However, our special choice in (2.20), which corresponds to $\theta = 0$ and, more important, to $\boldsymbol{v} = \mathbf{1}_n$, will be crucial for the automatic parameter selection strategy outlined in Section 3 and Section 4. More precisely, only taking $\boldsymbol{v} = \kappa \mathbf{1}_n$ with $\kappa \in \mathbb{R} \setminus \{0\}$ in (2.21) will allow us to solve very efficiently (by means of the 2D fast Fourier transform) the constrained linear system arising as an ADMM sub-problem and simultaneously set the pair of regularization parameters (μ_1, μ_2) satisfying the residual whiteness principle.

3. Multi-parameter RWP for the reduced RBD-L$_2$ model. Starting from the definition of the single-parameter (RWP), the formulation of the multi-parameter RWP (MRWP) for the reduced, linearly constrained RBD-L$_2$ model in (2.15)–(2.17) is quite straightforward. In fact, based on the idea to select (μ_1, μ_2) so that the residual image $\widehat{\boldsymbol{r}}(\mu_1, \mu_2) = \mathbf{K}\widehat{\boldsymbol{x}}(\mu_1, \mu_2) - \boldsymbol{y}$ of the reduced RBD-L$_2$ model is maximally white, the MRWP reads

$$\text{Select} \quad (\hat{\mu}_1, \hat{\mu}_2) = (\mu_1, \mu_2) \quad \text{such that} \quad \{\hat{\mu}_1, \hat{\mu}_2\} \in \underset{\mu_1, \mu_2 \in \mathbb{R}_{++}}{\operatorname{argmin}} \; W(\mu_1, \mu_2), \tag{3.1}$$

with the multi-parameter *whiteness measure function* $W : \mathbb{R}^2_{++} \to \mathbb{R}_+$ defined by

$$W(\mu_1, \mu_2) := \frac{\| \widehat{\boldsymbol{r}}(\mu_1, \mu_2) * \widehat{\boldsymbol{r}}(\mu_1, \mu_2) \|_2^2}{\| \widehat{\boldsymbol{r}}(\mu_1, \mu_2) \|_2^2}, \tag{3.2}$$

and where the residual image, according to definition (2.15), can be written as

$$\widehat{\boldsymbol{r}}(\mu_1, \mu_2) = \mathbf{K}\widehat{\boldsymbol{x}}(\mu_1, \mu_2) - \boldsymbol{y} = \mathbf{K}\left(\mathbf{A}\left(\mu_1, \mu_2\right)\widehat{\boldsymbol{c}}\left(\mu_1, \mu_2\right) + \boldsymbol{a}\left(\mu_1, \mu_2\right)\right) - \boldsymbol{y}, \tag{3.3}$$

with matrix $\mathbf{A}\left(\mu_1, \mu_2\right)$ and vector $\boldsymbol{a}\left(\mu_1, \mu_2\right)$ defined in (2.7) and (2.18), respectively, and with $\widehat{\boldsymbol{c}}\left(\mu_1, \mu_2\right)$ the (unique) solution component of the reduced minimization problem (2.16)–(2.17). We note that, upon the assumption of periodic boundary conditions for the three convolution matrices $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{K}$, analogously to the single-parameter whiteness measure function $W(\mu)$ in (RWP) (see [12, Proposition 3.1]) also $W(\mu_1, \mu_2)$ in (3.2) can be equivalently written in terms of the 2D Fourier transformed version of the residual image $\widetilde{\widehat{\boldsymbol{r}}}(\mu_1, \mu_2) = \mathbf{F}\widehat{\boldsymbol{r}}(\mu_1, \mu_2)$:

$$W(\mu_1, \mu_2) = \frac{\displaystyle\sum_{i=1}^{n} w_i^4(\mu_1, \mu_2)}{\left(\displaystyle\sum_{i=1}^{n} w_i^2(\mu_1, \mu_2)\right)^2}, \quad \text{with} \quad w_i(\mu_1, \mu_2) = \left|\widetilde{\widehat{r}}_i(\mu_1, \mu_2)\right|, \tag{3.4}$$

where $\widetilde{\widehat{r}}_i(\mu_1, \mu_2) \in \mathbb{C}$ is the i-th component of $\widetilde{\widehat{\boldsymbol{r}}}(\mu_1, \mu_2)$ and $|\cdot|$ indicates the modulus.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Due to non-smoothness of the cost function $\mathcal{J}_c$ in (2.17), the solution $\widehat{\boldsymbol{c}}(\mu_1,\mu_2)$ of the minimization problem (2.16)–(2.17) and, hence, the solution image $\widehat{\boldsymbol{x}}(\mu_1,\mu_2)$ in (2.17) and the associated residual image $\widehat{\boldsymbol{r}}(\mu_1,\mu_2)$ in (3.3) can not be written as explicit functions of the pair of parameters (μ_1,μ_2). Therefore, in principle, the MRWP outlined above can only be applied *a posteriori*: one can compute the solution $\widehat{\boldsymbol{c}}(\mu_1,\mu_2)$ - and, then, the residual $\widehat{\boldsymbol{r}}(\mu_1,\mu_2)$ - of the reduced RBD-L$_2$ model (2.15)–(2.17) for different values of (μ_1,μ_2) selected on a pre-defined or dynamically defined 2D grid and then calculate the corresponding whiteness measures $W(\mu_1,\mu_2)$ by (3.2) or, equivalently, (3.4). The optimal pair $(\hat{\mu}_1,\hat{\mu}_2)$ can thus be selected as the one minimizing $W(\mu_1,\mu_2)$ on the chosen discrete grid. Nonetheless, such approach is particularly demanding in terms of computing time and requires choosing a priori a grid search strategy.

As a way to drastically reduce the computational cost, inspired by [12], we propose to update the parameters μ_1,μ_2 along the iterations of the algorithm employed for the solution of (2.15)–(2.17) so as to satisfy the MRWP at convergence. We remark that this approach requires to solve the optimization problem (2.15)–(2.17) only once.

4. Iterated multi-parameter RWP for the reduced RBD-L$_2$ model. In the previous section we formalized the MRWP for the reduced RBD-L$_2$ model in (2.15)–(2.17), and discussed how it can be applied *a posteriori*. In what follows, we are going to show how the proposed parameter selection criterion can be rather adopted along the iterations of the optimization algorithm employed for the numerical solution of the model. In particular, first we introduce a standard two-blocks Alternating Direction Method of Multipliers (ADMM) iterative approach [3] for the solution of our minimization problem (2.15)–(2.17), then we illustrate how this approach can be equipped with an interlaced updating of the regularization parameters μ_1,μ_2 aimed to satisfy the proposed MRWP in (3.1)–(3.3) at convergence.

We resort to the variable splitting strategy and introduce the new variable $\boldsymbol{g}=\mathbf{D}_1\boldsymbol{c}\in\mathbb{R}^{2n}$, so that minimization problem (2.16)–(2.17) can be written in the following equivalent form:

$$\{\widehat{\boldsymbol{c}}\,(\mu_1,\mu_2)\,,\widehat{\boldsymbol{g}}\,(\mu_1,\mu_2)\}=\underset{\boldsymbol{c}\in\mathbb{R}^n,\,\boldsymbol{g}\in\mathbb{R}^{2n}}{\operatorname{argmin}}\ \{F_1(\boldsymbol{g})+F_2(\boldsymbol{c};\mu_1,\mu_2)\}\ \text{s.t.}\ \ \boldsymbol{g}=\mathbf{D}_1\boldsymbol{c}, \tag{4.1}$$

with functions $F_1:\mathbb{R}^n\to\mathbb{R}_+$ and $F_2:\mathbb{R}^{2n}\to\mathbb{R}_+$ defined by

$$F_1(\boldsymbol{g})=\sum_{i=1}^{n}\|\boldsymbol{g}_i\|_2\,,\quad F_2(\boldsymbol{c};\mu_1,\mu_2)=\frac{1}{2}\boldsymbol{c}^{\mathrm{T}}\mathbf{B}\,(\mu_1,\mu_2)\,\boldsymbol{c}-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b}\,(\mu_1,\mu_2)+\iota_{\mathcal{C}}(\boldsymbol{c}), \tag{4.2}$$

where, with a little abuse of notation, we indicate by $\boldsymbol{g}_i=(\mathbf{D}_1\boldsymbol{c})_i=\big((\mathbf{D}_h\boldsymbol{c})_i;(\mathbf{D}_v\boldsymbol{c})_i\big)\in\mathbb{R}^2$ the discrete gradient of image $\boldsymbol{c}$ at pixel i and $\iota_{\mathcal{C}}(\boldsymbol{c})$ denotes the indicator function of the hyperplane constraint $\mathcal{C}$ defined in (2.20).

The Lagrangian function $\mathcal{L}$ and the β-augmented Lagrangian function $\mathcal{L}_\beta$ associated to the linearly constrained problem (4.1)–(4.2) read

$$\mathcal{L}(\boldsymbol{c},\boldsymbol{g},\boldsymbol{\rho};\mu_1,\mu_2)=F_1(\boldsymbol{g})+F_2(\boldsymbol{c};\mu_1,\mu_2)+\boldsymbol{\rho}^{\mathrm{T}}(\mathbf{D}_1\boldsymbol{c}-\boldsymbol{g})\,, \tag{4.3}$$

$$\mathcal{L}_\beta(\boldsymbol{c},\boldsymbol{g},\boldsymbol{\rho};\mu_1,\mu_2)=\mathcal{L}(\boldsymbol{c},\boldsymbol{g},\boldsymbol{\rho};\mu_1,\mu_2)+\frac{\beta}{2}\|\mathbf{D}_1\boldsymbol{c}-\boldsymbol{g}\|_2^2, \tag{4.4}$$

where $\boldsymbol{\rho}\in\mathbb{R}^{2n}$ is the vector of Lagrange multipliers associated to the system of $2n$ linear constraints in (4.1) and $\beta\in\mathbb{R}_{++}$ is a penalty parameter.

Solving (4.1)–(4.2) amounts to seek for the saddle point(s) of the augmented Lagrangian function $\mathcal{L}_\beta$ in (4.4). The saddle-point problem takes the form:

$$\{\widehat{\boldsymbol{c}}(\mu_1,\mu_2),\widehat{\boldsymbol{g}}(\mu_1,\mu_2),\widehat{\boldsymbol{\rho}}(\mu_1,\mu_2)\}\in\underset{\boldsymbol{c}\in\mathbb{R}^n,\,\boldsymbol{g}\in\mathbb{R}^{2n}}{\operatorname{argmin}}\left\{\max_{\boldsymbol{\rho}\in\mathbb{R}^{2n}}\mathcal{L}_\beta(\boldsymbol{c},\boldsymbol{g},\boldsymbol{\rho};\mu_1,\mu_2)\right\}. \tag{4.5}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Upon suitable initialization, and for any $k \geq 0$, the k-th iteration of the standard two-blocks ADMM [3] applied to solving the saddle-point problem (4.5) reads as follows:

$$\boldsymbol{g}^{(k+1)} = \underset{\boldsymbol{g}\in\mathbb{R}^{2n}}{\operatorname{argmin}}\, \mathcal{L}_\beta\left(\boldsymbol{c}^{(k)}, \boldsymbol{g}, \boldsymbol{\rho}^{(k)}; \mu_1, \mu_2\right), \tag{4.6}$$

$$\boldsymbol{c}^{(k+1)} = \underset{\boldsymbol{c}\in\mathbb{R}^{n}}{\operatorname{argmin}}\, \mathcal{L}_\beta\left(\boldsymbol{c}, \boldsymbol{g}^{(k+1)}, \boldsymbol{\rho}^{(k)}; \mu_1, \mu_2\right), \tag{4.7}$$

$$\boldsymbol{\rho}^{(k+1)} = \boldsymbol{\rho}^{(k)} + \beta\left(\mathbf{D}_1\boldsymbol{c}^{(k+1)} - \boldsymbol{g}^{(k+1)}\right). \tag{4.8}$$

We remark that, since the two cost functions F_1 and F_2 in the (standard) two-blocks separable minimization problem (4.1), defined in (4.2), are both clearly proper, lower semicontinuous and convex, it is a standard result that, in case that the two regularization parameters μ_1, μ_2 (and the penalty parameter β) are kept fixed along the iterations, the ADMM scheme in (4.6)–(4.8) converges if the Lagrangian function $\mathcal{L}$ in (4.3) admits a saddle point; see [6]. Nonetheless, the proof of convergence of the interlaced scheme we are going to detail remains an open issue.

We note in advance that the parameters μ_1, μ_2 are only involved in sub-problem (4.7); in Section 4.1 we thus analyze the $\boldsymbol{g}$ sub-problem, while in Section 4.2 we are going to discuss the $\boldsymbol{c}$-update as well as the adjustment of the parameters according to the MRWP.

4.1. Solving the subproblem for $\boldsymbol{g}$. After removing in the expression (4.4) of the augmented Lagrangian function $\mathcal{L}_\beta$ all the terms that do not depend on $\boldsymbol{g}$, we have that the explicit expression of the $\boldsymbol{g}$-update problem in (4.6) reads

$$\begin{aligned}\boldsymbol{g}^{(k+1)} &= \underset{\boldsymbol{g}\in\mathbb{R}^{2n}}{\operatorname{argmin}} \left\{ F_2(\boldsymbol{g}) + \left(\boldsymbol{\rho}^{(k)}\right)^{\mathrm{T}}\left(\mathbf{D}_1\boldsymbol{c}^{(k)} - \boldsymbol{g}\right) + \frac{\beta}{2}\|\mathbf{D}_1\boldsymbol{c}^{(k)} - \boldsymbol{g}\|_2^2 \right\} \\ &= \underset{\boldsymbol{g}\in\mathbb{R}^{2n}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} \|\boldsymbol{g}_i\|_2 + \frac{\beta}{2}\|\boldsymbol{g} - \boldsymbol{u}^{(k)}\|_2^2 \right\}, \quad \boldsymbol{u}^{(k)} = \mathbf{D}_1\boldsymbol{c}^{(k)} + \frac{1}{\beta}\boldsymbol{\rho}^{(k)}.\end{aligned}$$

The above $2n$-variate problem can be split into n independent bivariate problems of the form

$$\boldsymbol{g}_i^{(k+1)} \in \underset{\boldsymbol{g}_i\in\mathbb{R}^{2}}{\operatorname{argmin}} \left\{ \|\boldsymbol{g}_i\|_2 + \frac{\beta}{2}\|\boldsymbol{g}_i - \boldsymbol{u}_i^{(k+1)}\|_2^2 \right\} = \operatorname{prox}_{\frac{1}{\beta}\|\cdot\|_2}\left(\boldsymbol{u}_i^{(k+1)}\right), \quad i = 1, \ldots, n,$$

with prox_f indicating the proximity operator of function f that, when applied to the Euclidean norm, has the following closed form expression (soft-thresholding shrinkage operator)

$$\boldsymbol{g}_i^{(k+1)} = \max\left(1 - \frac{1}{\beta\|\boldsymbol{u}_i^{(k)}\|_2}, 0\right)\boldsymbol{u}_i^{(k)}, \quad i = 1, \ldots, n, \tag{4.9}$$

where $\max\{1 - 1/0, 0\} = 0$ is assumed.

4.2. Solving the subproblem for $\boldsymbol{c}$ and updating μ_1, μ_2. After dropping out all terms not depending on $\boldsymbol{c}$ in the augmented Lagrangian function $\mathcal{L}_\beta$, introducing the vector

$$\boldsymbol{q}^{(k)} = \boldsymbol{g}^{(k+1)} - \frac{1}{\beta}\boldsymbol{\rho}^{(k)},$$

and writing the indicator function $\iota_{\mathcal{C}}(\boldsymbol{c})$ as a hard constraint, the $\boldsymbol{c}$-update problem (4.7) reads

$$\boldsymbol{c}^{(k+1)}(\mu_1, \mu_2) = \underset{\boldsymbol{c}\in\mathcal{C}}{\operatorname{argmin}}\, Q^{(k)}(\boldsymbol{c}; \mu_1, \mu_2), \tag{4.10}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with the quadratic cost function, which depends on (μ_1, μ_2) and also on $\boldsymbol{q}^{(k)}$, given by

$$Q^{(k)}(\boldsymbol{c};\mu_1,\mu_2)=\frac{1}{2}\boldsymbol{c}^{\mathrm{T}}\overbrace{\left(\mathbf{B}\left(\mu_1,\mu_2\right)+\beta\mathbf{D}_1^{\mathrm{T}}\mathbf{D}_1\right)}^{\mathbf{S}(\mu_1,\mu_2)}\boldsymbol{c}-\boldsymbol{c}^{\mathrm{T}}\overbrace{\left(\boldsymbol{b}\left(\mu_1,\mu_2\right)+\beta\mathbf{D}_1^{\mathrm{T}}\boldsymbol{q}^{(k)}\right)}^{\boldsymbol{z}(\mu_1,\mu_2)}. \tag{4.11}$$

The $\boldsymbol{c}$-update problem (4.10)–(4.11) is a quadratic program which admits a unique solution, as stated (and proved) in more general settings in Lemma 4.1, whose proof is postponed to the Appendix A.

LEMMA 4.1. *Let $\mathbf{S} \in \mathbb{R}^{n\times n}$ be a symmetric, positive definite, BCCB matrix. Then, for any vector $\boldsymbol{z} \in \mathbb{R}^n$ the quadratic program*

$$\widehat{\boldsymbol{p}} = \operatorname*{argmin}_{\boldsymbol{p}\in\mathbb{R}^n}\left\{Q(\boldsymbol{p}) := \frac{1}{2}\boldsymbol{p}^{\mathrm{T}}\mathbf{S}\boldsymbol{p} - \boldsymbol{p}^{\mathrm{T}}\boldsymbol{z}\right\} \quad \textit{subject to} \quad \sum_{i=1}^{n} p_i=0, \tag{4.12}$$

admits a unique solution given by

$$\widehat{\boldsymbol{p}} = \mathbf{F}^*\widetilde{\widehat{\boldsymbol{p}}}, \quad \textit{with} \quad \widetilde{\widehat{\boldsymbol{p}}}_i = \begin{cases} 0 & \text{for } i = 1, \\ \widetilde{\widehat{\boldsymbol{p}}}_i^{(U)} & \text{for } i = 2,\dots,n, \end{cases}$$

where $\widetilde{\widehat{\boldsymbol{p}}}^{(U)} \in \mathbb{C}^n$ denotes the (unique) Fourier-transformed solution of the unconstrained version of problem (4.12), which reads

$$\widetilde{\widehat{\boldsymbol{p}}}^{(U)} = \widetilde{\mathbf{S}}^{-1}\widetilde{\boldsymbol{z}}, \quad \text{with} \quad \widetilde{\mathbf{S}} = \mathbf{F}\mathbf{S}\mathbf{F}^* = \operatorname{diag}(\widetilde{s}_1,\dots,\widetilde{s}_n) \in \mathbb{C}^n. \tag{4.13}$$

The $\boldsymbol{c}$-update problem (4.10)–(4.11) has exactly the same form as (4.12). Moreover, it is easy to prove that, for any $(\mu_1, \mu_2) \in \mathbb{R}^2_{++}$, the Hessian matrix of the quadratic cost function $Q^{(k)}$ in (4.11), namely $\mathbf{S}(\mu_1, \mu_2) = \mathbf{B}\left(\mu_1, \mu_2\right) + \beta\mathbf{D}_1^{\mathrm{T}}\mathbf{D}_1$, is symmetric, positive definite and BCCB under our assumption of periodic boundary conditions for $\mathbf{K}$, $\mathbf{D}_1$ and $\mathbf{D}_2$. In fact, according to Lemma 2.1, matrix $\mathbf{B}\left(\mu_1, \mu_2\right)$ is a symmetric, positive semidefinite BCCB matrix with null space equal to the null space of the blurring matrix $\mathbf{K}$, whereas $\mathbf{D}_1^{\mathrm{T}}\mathbf{D}_1$ is clearly a symmetric positive semidefinite BCCB matrix with null space equal to the null space of $\mathbf{D}_1$. Then, due to assumption (2.2), the matrix $\mathbf{S}(\mu_1, \mu_2)$ in (4.11) is a symmetric positive definite BCCB matrix.

Hence, according to Lemma 4.1, the unique solution of the $\boldsymbol{c}$-update problem (4.10)–(4.11) reads

$$\boldsymbol{c}^{(k+1)}(\mu_1,\mu_2) = \mathbf{F}^*\widetilde{\boldsymbol{c}}^{(k+1)}(\mu_1,\mu_2), \tag{4.14}$$

with

$$\widetilde{\boldsymbol{c}}_i^{(k+1)}(\mu_1,\mu_2) = \begin{cases} 0 & \text{for } i = 1, \\ \dfrac{\widetilde{z}_i(\mu_1,\mu_2)}{\widetilde{s}_i(\mu_1,\mu_2)} & \text{for } i = 2,\dots,n, \end{cases} \tag{4.15}$$

where $\widetilde{\boldsymbol{z}}(\mu_1, \mu_2) = \mathbf{F}\boldsymbol{z}(\mu_1, \mu_2)$ with $\boldsymbol{z}(\mu_1, \mu_2)$ defined in (4.11) and where $\widetilde{s}_i(\mu_1, \mu_2)$ is the i-th diagonal element of the Fourier diagonalization of matrix $\mathbf{S}(\mu_1, \mu_2)$, again defined in (4.11).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

It follows from the definition in (3.3) that, at any ADMM iteration k, the updated residual image $\boldsymbol{r}^{(k+1)}(\mu_1,\mu_2)$ can be written as an affine function of $\boldsymbol{c}^{(k+1)}(\mu_1,\mu_2)$,

$$\boldsymbol{r}^{(k+1)}(\mu_1,\mu_2) = \mathbf{K}\left(\mathbf{A}\left(\mu_1,\mu_2\right)\boldsymbol{c}^{(k+1)}\left(\mu_1,\mu_2\right) + \boldsymbol{a}\left(\mu_1,\mu_2\right)\right) - \boldsymbol{y}, \tag{4.16}$$

hence also - by exploiting (4.14) - as an explicit function of the pair of regularization parameters (μ_1,μ_2). In the following Proposition 4.2 we show how, relying on (4.14)–(4.16), at any ADMM iteration it is possible to derive the explicit expression of the whiteness measure W as a function of μ_1,μ_2, and to select efficiently the two regularization parameters satisfying the MRWP.

PROPOSITION 4.2. *Let $\boldsymbol{c}^{(k+1)}(\mu_1,\mu_2)$ be the updated cartoon component in (4.14)–(4.15) and $\boldsymbol{r}^{(k+1)}(\mu_1,\mu_2)$ the associated updated residual image defined according to (3.3), with $\widetilde{\boldsymbol{r}}^{(k+1)}(\mu_1,\mu_2)$ its Fourier transform. Then, the parameters $\mu_1^{(k+1)},\mu_2^{(k+1)}$ yielding the whitest residual image according to the scalar whiteness metric defined in (3.2) or, equivalently, in (3.4), are given by*

$$\left\{\mu_1^{(k+1)},\mu_2^{(k+1)}\right\} \in \underset{\mu_1,\mu_2\in\mathbb{R}_{++}}{\operatorname{argmin}} \left\{ W^{(k+1)}(\mu_1,\mu_2) = \frac{\displaystyle\sum_{i=1}^{n}\left(w_i^{(k+1)}(\mu_1,\mu_2)\right)^4}{\left(\displaystyle\sum_{i=1}^{n}\left(w_i^{(k+1)}(\mu_1,\mu_2)\right)^2\right)^2} \right\}, \tag{4.17}$$

with

$$w_i^{(k+1)}(\mu_1,\mu_2) = |\widetilde{r}_i^{(k+1)}(\mu_1,\mu_2)| = \begin{cases} \dfrac{\eta_i|\widetilde{y}_i|}{\eta_i+\epsilon_i(\mu_2/\mu_1)} & i=1, \\ \dfrac{\eta_i\phi_i^{(k)}}{\zeta_i(\eta_i+\epsilon_i(\mu_2/\mu_1))+\eta_i\epsilon_i(\mu_2/\beta)} & i=2,\dots,n, \end{cases} \tag{4.18}$$

where

$$\begin{aligned} \epsilon_i &= |\widetilde{k}_i|^2, \quad \zeta_i = |\widetilde{d}_{h,i}|^2 + |\widetilde{d}_{v,i}|^2, \quad \eta_i = |\widetilde{d}_{hh,i}|^2 + 2|\widetilde{d}_{hv,i}|^2 + |\widetilde{d}_{vv,i}|^2, \\ \phi_i^{(k)} &= |\widetilde{k}_i\widetilde{z}_i^{(k)} - \zeta_i\widetilde{y}_i|, \quad \widetilde{z}_i^{(k)} = \overline{\widetilde{d}}_{h,i}\widetilde{q}_{1,i}^{(k)} + \overline{\widetilde{d}}_{v,i}\widetilde{q}_{2,i}^{(k)}, \end{aligned} \tag{4.19}$$

and $\widetilde{k}_i$, $\widetilde{d}_{h,i}$, $\widetilde{d}_{v,i}$, $\widetilde{d}_{hh,i}$, $\widetilde{d}_{vv,i}$, $\widetilde{d}_{hv,i}$ are defined in (2.3) and (2.4).

Based on the above proposition, $\mu_1^{(k+1)},\mu_2^{(k+1)}$ can be computed by minimizing the whiteness function $W^{(k+1)}$ expressed in (4.17), (4.18), and (4.19) by means of the Newton algorithm; that is, tantamount to find the zero of the gradient $\nabla W^{(k+1)}$. Notice that all the quantities in (4.19), except $\widetilde{z}_i^{(k)},\phi_i^{(k)}$, can be computed once for all at the beginning of the ADMM iterations. The updated parameters $\mu_1^{(k+1)},\mu_2^{(k+1)}$ can then be plugged into (4.14)–(4.15) so as to get the solution of (4.10)–(4.11).

The main steps of the overall proposed approach, to which we refer as Iterated MRWP-ADMM (IMRWP-ADMM) are summarized in Algorithm 1.

5. Numerical results. In this section, we evaluate experimentally the performance of the proposed MRWP-based automatic procedure for selecting the pair of regularization parameters (μ_1,μ_2) inthe reduced RBD-L_2 image restoration variational model in (1.6). The aim of our analysis is twofold:

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 1: Proposed parameter-free IMRWP-ADMM approach for the solution of the reduced RBD-L_2 variational model (2.15)–(2.17).

inputs: observed degraded image $\boldsymbol{y} \in \mathbb{R}^n$,
blur matrix, difference matrices $\mathbf{K} \in \mathbb{R}^{n\times n}$, $\mathbf{D}_1 \in \mathbb{R}^{2n\times n}$, $\mathbf{D}_2 \in \mathbb{R}^{4n\times n}$

output:restored image $\widehat{\boldsymbol{x}} \in \mathbb{R}^n$, piecewise constant component $\widehat{\boldsymbol{c}} \in \mathbb{R}^n$
smooth component $\widehat{\boldsymbol{s}} \in \mathbb{R}^n$

1. **initialize:** set $\boldsymbol{c}^{(0)} = \boldsymbol{y}$
2. **for** *k = 0, 1, 2, . . . until convergence* **do**:
3. · compute $\mu_1^{(k+1)}, \mu_2^{(k+1)}$ by solving (4.17), (4.18), and (4.19)
4. · compute $\boldsymbol{c}^{(k+1)}$ by (4.14), (4.15)
5. · compute $\boldsymbol{g}^{(k+1)}$ by (4.9)
6. · compute $\boldsymbol{\rho}^{(k+1)}$ by (4.8)
7. **end for**
8. $\widehat{\boldsymbol{c}} = \boldsymbol{c}^{(k+1)}$, $\widehat{\boldsymbol{s}} = \mathbf{M}_2\left(\mu_1^{(k+1)}, \mu_2^{(k+1)}\right)\widehat{\boldsymbol{c}} + \boldsymbol{a}\left(\mu_1^{(k+1)}, \mu_2^{(k+1)}\right)$, $\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{c}} + \widehat{\boldsymbol{s}}$

(i) first, we want to assess the robustness of the whiteness principle when transferred from a single parameter to a multiple parameters scenario;
(ii) then, we want to assess that the proposed IMRWP-ADMM is capable of selecting $(\hat{\mu}_1, \hat{\mu}_2)$, i.e., the parameters minimizing the whiteness function, in a very robust and efficient way.

For what concerns point (i), the performance of the MRWP for the selection of (μ_1, μ_2) in the RBD-L_2 variational model will be compared with the performance of the RWP for the selection of μ in the TIK-L_2, TV-L_2 variational models.

The quality of the restorations $\widehat{\boldsymbol{x}}$ for different values of $\hat{\mu}$ in the TIK-L_2, TV-L_2 models, and for different $(\hat{\mu_1}, \hat{\mu_2})$ in the RBD-L_2 variational model, with respect to the original uncorrupted image $\overline{\boldsymbol{x}}$, will be assessed by means of two scalar measures, namely the Improved Signal-to-Noise Ratio (ISNR),

$$\mathrm{ISNR}(\boldsymbol{y}, \overline{\boldsymbol{x}}, \widehat{\boldsymbol{x}}) := 10\log_{10}\frac{\|\boldsymbol{y} - \overline{\boldsymbol{x}}\|_2^2}{\|\widehat{\boldsymbol{x}} - \overline{\boldsymbol{x}}\|_2^2},$$

and the Structural Similarity Index (SSIM) [16]. The larger the ISNR and SSIM values, the higher the quality of the restoration.

For all tests, iterations of the IRWP-ADMM approach in Algorithm 1 are stopped as soon as

$$\delta_{\boldsymbol{x}}^{(k)} := \frac{\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right\|_2}{\left\|\boldsymbol{x}^{(k-1)}\right\|_2} < 10^{-6}, \quad k \in \mathbb{N}\setminus\{0\},$$

and the ADMM penalty parameter β has been set manually. It is also worth remarking that, as far as the minimization of function $W^{(k+1)}$ is concerned, we observed that at each iteration of Algorithm 1 a global minimum exists and it is always attained, i.e., Newton algorithm for finding a zero of $\nabla W^{(k+1)}$ converges in few iterations.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

(a) `phantom` (b) `peppers`

FIG. 5.1. *Original test image* `phantom` *(200×200) (a) and* `peppers` *(256×256) (b).*

We consider the `phantom` (200×200) and `peppers` (256×256) test images, with pixel values normalized in the range [0,1], shown in Figure 5.1. Each image is corrupted by a space-invariant Gaussian blur defined by a convolution kernel generated using the Matlab routine `fspecial` with parameters `band`$= 5$ and `sigma`$= 1$. The `band` parameter represents the side length (in pixels) of the square support of the kernel, whereas `sigma` is the standard deviation (in pixels) of the isotropic bivariate Gaussian distribution defining the kernel in the continuous setting. The blurred images are further corrupted by AWG noise with standard deviation $\sigma = 0.05$.

First, we employ the TIK-L_2 and the TV-L_2 variational models for the restoration of the test image `phantom`. The two models are applied for different values of μ selected on fine one-dimensional grids. Then, for what concern the reduced RBD-L_2 model, we consider the reparametrization introduced in the proof of Proposition 4.2 and recalled below

$$\lambda := \frac{\mu_1}{\mu_2}, \quad \gamma := \frac{\mu_2}{\beta}.$$

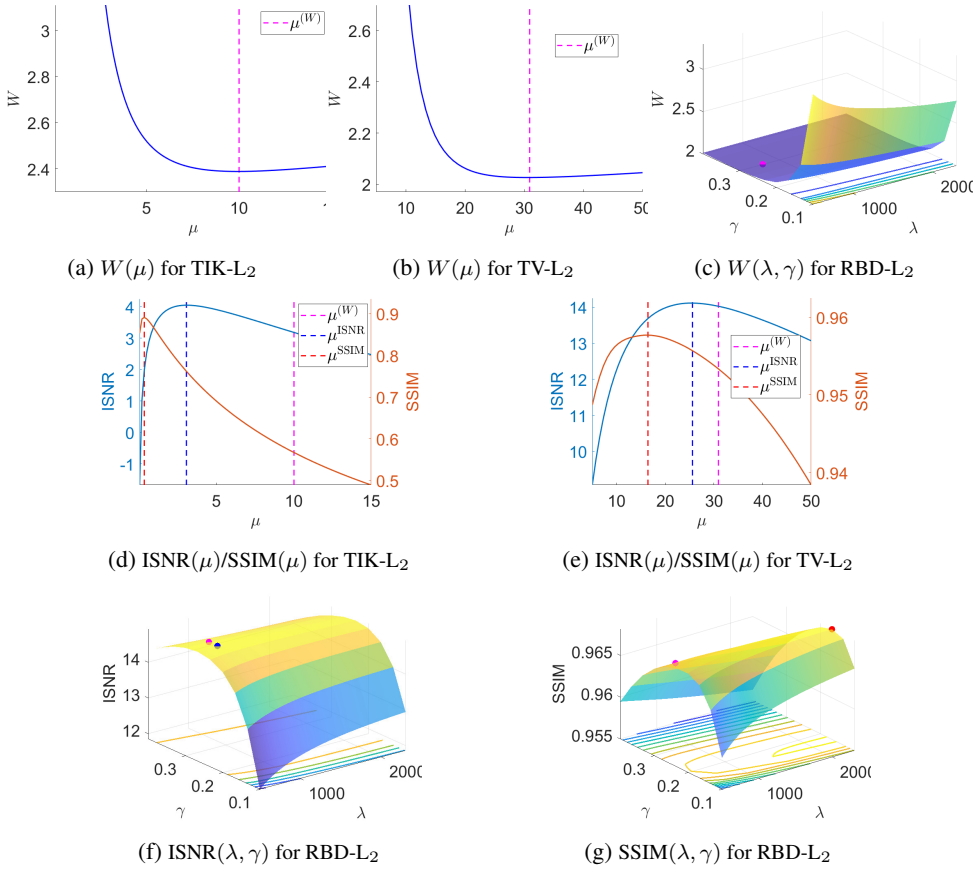
Hence, we let (λ, γ) vary on a fine two-dimensional grid.

For each output restoration we record the ISNR and SSIM achieved, and the corresponding whiteness measure. In the first line of Figure 5.2, we show the behaviour of the one-dimensional whiteness curves $W(\mu)$ for the TIK-L_2 and the TV-L_2 models, as well as the two-dimensional whiteness surface $W(\lambda, \gamma)$ for the RBD-L_2 model. The vertical magenta lines in the left panels represent the μ-value detected by the RWP, and denoted by $\mu^{(W)}$, while the magenta circle in the third panel detects the couple (λ, γ) estimated by the MRWP for the RBD-L_2 model. In Figures 5.2d and 5.2e we plot the ISNR and SSIM values achieved by the TIK-L_2 and TV-L_2 models, respectively; the vertical blue and red lines represent the μ-values corresponding to the maximum ISNR and SSIM obtained, and denoted by μ^{ISNR}, μ^{SSIM}, respectively. Finally, on the bottom row of Figure 5.2 we show the behaviour of the quality metrics for the RBD-L_2 on the selected grid for (λ, γ). The blue circle in the Figure 5.2f and the red circle in Figure 5.2g represent the highest ISNR and SSIM achieved, respectively.

From Figures 5.2e and 5.2f it seems clear that the RWP for the TV-L_2 and the MRWP for the RBD-L_2 return output restorations which are very close to the one achieving the maximum ISNR. Notice also that, although apparently far, the magenta and the red circles in Figure 5.2g belong to very close level curves.

In order to assess the robustness of the whiteness criterion when employed for a multi-parameter selection, in the left part of Table 5.1 we report the ISNR and SSIM values achieved by the the RWP for the TIK-L_2 and TV-L_2 models and by the MRWP for the RBD-L_2, together

ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) $W(\mu)$ for TIK-L$_2$ (b) $W(\mu)$ for TV-L$_2$ (c) $W(\lambda,\gamma)$ for RBD-L$_2$

(d) ISNR(μ)/SSIM(μ) for TIK-L$_2$ (e) ISNR(μ)/SSIM(μ) for TV-L$_2$

(f) ISNR(λ,γ) for RBD-L$_2$ (g) SSIM(λ,γ) for RBD-L$_2$

FIG. 5.2. *Whiteness measure functions for the three variational models performed (top row) and ISNR/SSIM values (middle and bottom rows) for the test image* `phantom` *corrupted by Gaussian blur with* `band` *= 5,* `sigma` *= 1 and AWGN with* $\sigma = 0.05$.

TABLE 5.1
ISNR and SSIM values achieved by the RWP for the TIK-L$_2$, TV-L$_2$ models, and by the MRWP for the RBD-L$_2$ model, together with the corresponding percentage differences for the test images `phantom` *and* `peppers` *corrupted by Gaussian blur with* `band` *= 5,* `sigma` *= 1, and AWGN with* $\sigma = 0.05$.

	`phantom`			`peppers`		
	TIK-L$_2$	TV-L$_2$	RBD-L$_2$	TIK-L$_2$	TV-L$_2$	RBD-L$_2$
ISNR	3.1795	14.0253	14.8631	2.2198	3.2020	3.3745
SSIM	0.5674	0.9533	0.9653	0.7613	0.7984	0.8047
Δ_{ISNR}	21.4228	0.6221	0.5148	12.1158	9.3110	6.7113
Δ_{SSIM}	36.3393	0.4552	0.2946	0.4374	0.8346	0.6636

with the percentage differences

$$\Delta_{\mathrm{ISNR}} = 100 \times \frac{\mathrm{ISNR}_{\max} - \mathrm{ISNR}(\boldsymbol{y}, \overline{\boldsymbol{x}}, \widehat{\boldsymbol{x}})}{\mathrm{ISNR}_{\max}}, \quad \Delta_{\mathrm{SSIM}} = 100 \times \frac{\mathrm{SSIM}_{\max} - \mathrm{SSIM}(\overline{\boldsymbol{x}}, \widehat{\boldsymbol{x}})}{\mathrm{SSIM}_{\max}}.$$

One can easily observe that, when comparing the TV-L$_2$ and the RBD-L$_2$, the whiteness criterion presents a very stable behaviour, as the percentage differences are particulalry close.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

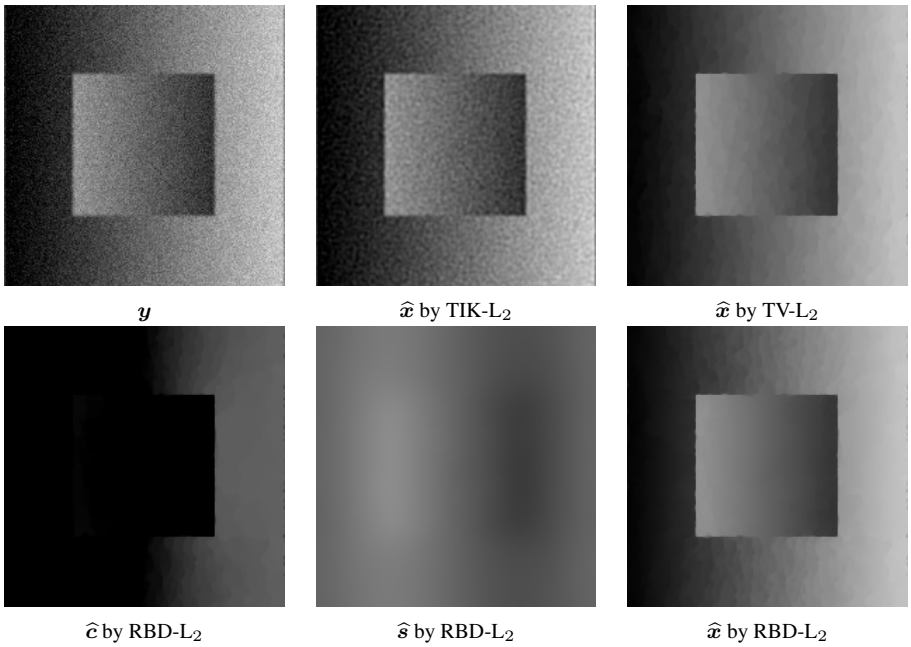


FIG. 5.3. *Output restorations for the test image* `phantom` *corrupted by Gaussian blur with* `band` *= 5,* `sigma` *= 1, and AWGN with* $\sigma = 0.05$.

Finally, in Figure 5.3 we show the output restorations for the three models considered combined with the RWP/MRWP. Notice that for the RBD-L$_2$, in the bottom row, we also show the piecewise constant component $\widehat{\boldsymbol{c}}$ and the smooth component $\widehat{\boldsymbol{s}}$.

For the second test image `peppers`, we show the monitored ISNR and SSIM values, as well as the whiteness measure function, for different selection of parameters in the three models considered. Also in this case, we notice that the output restoration obtained by the MRWP for the RBD-L$_2$ model is characterized by quality metrics that are very close to the highest achievable.

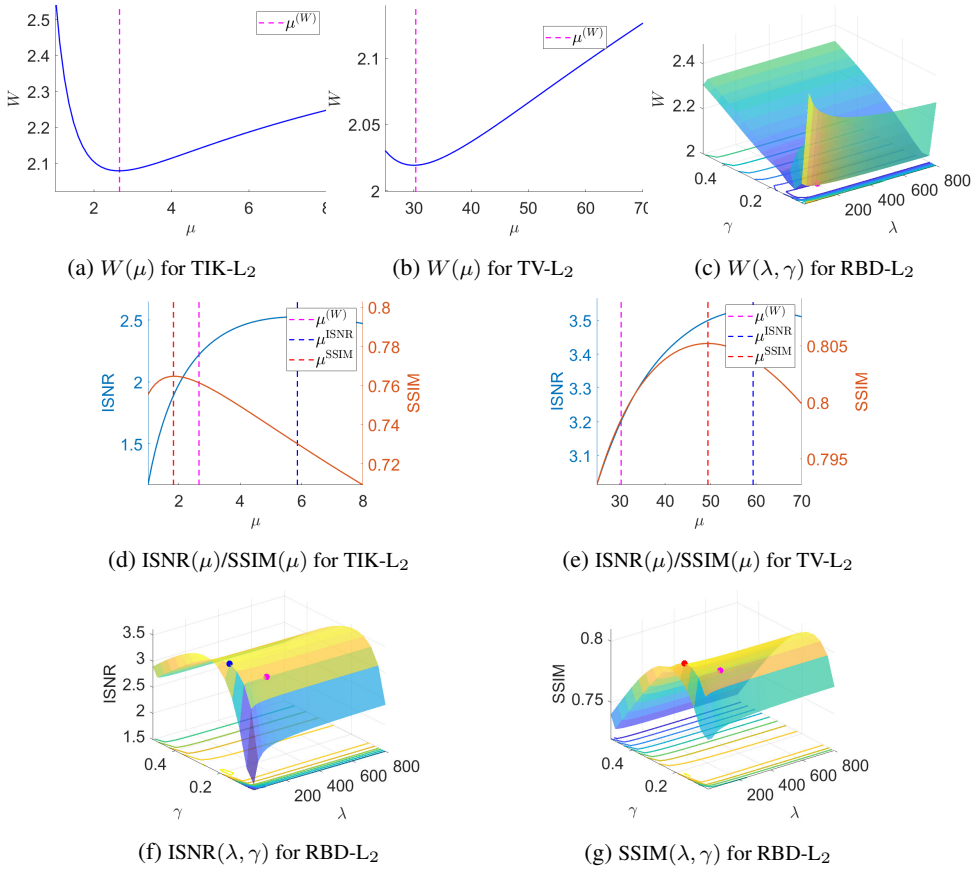
In the right part of Table 5.1, we report the ISNR and SSIM values, as well as the percentage differences, obtained for the test image `peppers`. Notice that Δ_{ISNR}, Δ_{SSIM} present a very stable behaviour for the three models considered, thus reflecting the robustness of the whiteness criterion along the different scenarios in which it is employed.

The output restorations obtained by the RWP/MRWP for the three models are shown in Figure 5.5.

Then, we considered a more severe degradation for the two test images, which have been corrupted by a space-invariant Gaussian blur with `band` = 9 and `sigma` = 2, and by AWGN with $\sigma = 0.1$. In Figures 5.6 and 5.8 we show the behaviour of the quality metrics and of the whiteness measure for the TIK-L$_2$, TV-L$_2$ and RBD-L2 model for the test image `phantom` and `peppers`, respectively. Also in this case, for the `phantom` image, the whiteness criterion returns an output restoration which is very close to the one with the highest ISNR for the TV-L$_2$ and the RBD-L$_2$ models. On the other hand, for the test image `peppers` the restorations obtained by means of the whiteness principle are either very close to the one with highest SSIM - in the case of TIK-L$_2$ - or to the one with highest ISNR - in the case of TV-L$_2$ - or to both - in the case of the RBD-L$_2$.

The values reported in Table 5.2 confirm again the robustness of the whiteness principle also in higher degradation scenarios, as the percentage differences are stable when moving

ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) $W(\mu)$ for TIK-L$_2$ (b) $W(\mu)$ for TV-L$_2$ (c) $W(\lambda,\gamma)$ for RBD-L$_2$

(d) ISNR(μ)/SSIM(μ) for TIK-L$_2$ (e) ISNR(μ)/SSIM(μ) for TV-L$_2$

(f) ISNR(λ,γ) for RBD-L$_2$ (g) SSIM(λ,γ) for RBD-L$_2$

FIG. 5.4. *Whiteness measure functions for the three variational models performed (top row) and ISNR/SSIM values (middle and bottom rows) for the test image* `peppers` *corrupted by Gaussian blur with* `band` *= 5,* `sigma` *= 1, and AWGN with* $\sigma = 0.05$.

TABLE 5.2
ISNR and SSIM values achieved by the RWP for the TIK-L$_2$, TV-L$_2$ models, and by the MRWP for the RBD-L$_2$ model, together with the corresponding percentage differences for the test images `phantom` *and* `peppers` *corrupted by Gaussian blur with* `band` *= 9,* `sigma` *= 2, and AWGN with* $\sigma = 1$.

	`phantom`			`peppers`		
	TIK-L$_2$	TV-L$_2$	RBD-L$_2$	TIK-L$_2$	TV-L$_2$	RBD-L$_2$
ISNR	6.5437	14.4841	15.0326	4.6037	5.0583	5.3749
SSIM	0.6028	0.9229	0.9376	0.6502	0.6804	0.6956
Δ_{ISNR}	6.2763	0.1724	0.8331	6.5130	6.0155	3.2684
Δ_{SSIM}	30.0668	0.7075	0.5521	0.1375	1.5632	0.8294

from a single to a coupled estimate.

Finally, the ouput restorations for the `phantom` and `peppers` are shown in Figures 5.7 and 5.9, respectively.

After assessing the robustness of the MRWP when employed *a posteriori*, that is the RBD-L$_2$ is solved over a two-dimensional grid of parameters and the optimal couple of parameters is selected according to the output values of the whiteness measure, it is also worth evaluating

ETNA
Kent State University and
Johann Radon Institute (RICAM)

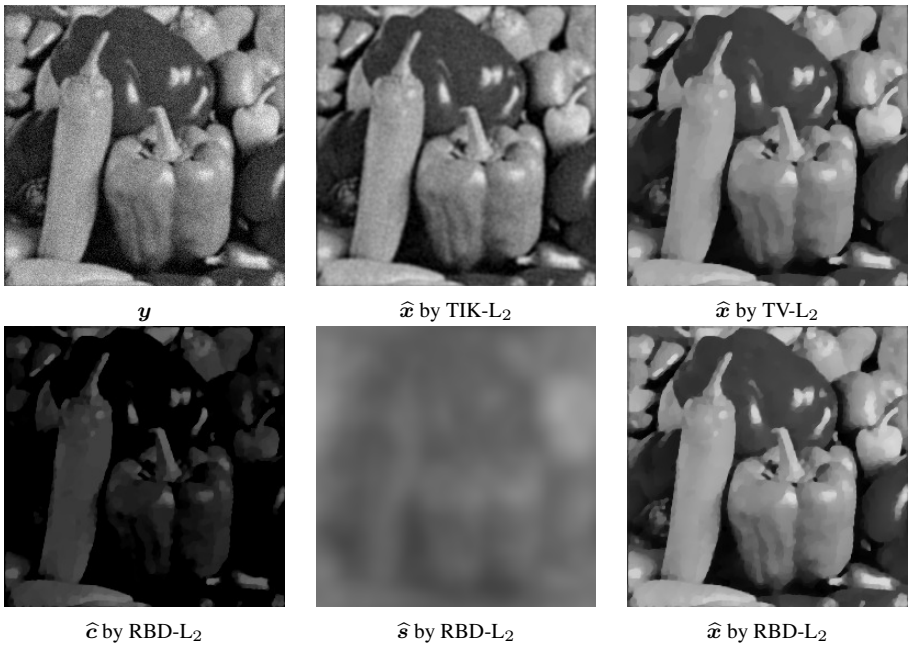


FIG. 5.5. *Output results for the test image* `peppers` *corrupted by Gaussian blur with* `band` *= 5,* `sigma` *= 1, and AWGN with* $\sigma = 0.05$.

TABLE 5.3
Output ISNR/SSIM values and μ_1, μ_2 *obtained by employing the MRWP and the IMRWP for the solution of the RBD-L*$_2$ *model for images corrupted by Gaussian blur with* `band` *= 5,* `sigma` *= 1, and AWGN with* $\sigma = 0.05$.

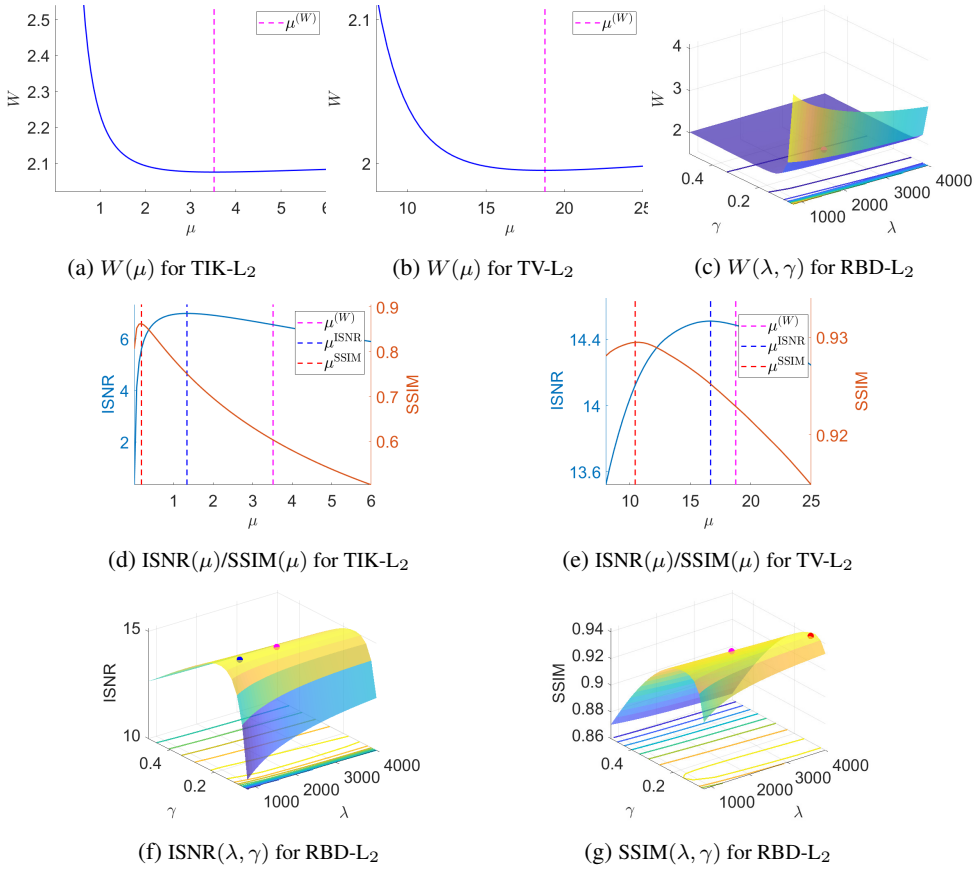
	`phantom`		`peppers`	
	MRWP	IMRWP-ADMM	MRWP	IMRWP-ADMM
μ_1	23698.31	21090.56	6082.34	6835.42
μ_2	30.08	27.98	36.57	32.29
ISNR	14.8631	14.7092	3.3745	3.3168
SSIM	0.9653	0.9648	0.8047	0.8013

the robustness of the IMRWP-ADMM. To this purpose, we run the IMRWP-ADMM outlined in Algorithm 1 for the restoration of the test images `phantom` and `peppers` in the lower degradation scenario considered. In Table 5.3 we report the output ISNR/SSIM and the estimated μ_1, μ_2 obtained by solving the RBD-L$_2$ model with the MRWP in its *a posteriori* and iterated version. We remark that μ_1, μ_2 can be recovered from the output parameters λ, γ - that are actually estimated - via the transformation in (A.16). One can observe that the results achieved by the IMRWP-ADMM are very close to the one obtained by employing the MRWP *a posteriori*.

The output restorations by the IRWP-ADMM , together with the piecewise constant and smooth components, for the test images `phantom` and `peppers` are displayed in Figure 5.10.

The same analysis is carried out for the larger degradation case; the output quality metrics and parameters are reported in Table 5.4, while the ouptut restorations are shown in Figure 5.11. Again, we observe that the IMRWP-ADMM is capable of obtaining results which are very close to the ones achieved *a posteriori*.

ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) $W(\mu)$ for TIK-L_2 (b) $W(\mu)$ for TV-L_2 (c) $W(\lambda,\gamma)$ for RBD-L_2

(d) ISNR(μ)/SSIM(μ) for TIK-L_2 (e) ISNR(μ)/SSIM(μ) for TV-L_2

(f) ISNR(λ,γ) for RBD-L_2 (g) SSIM(λ,γ) for RBD-L_2

FIG. 5.6. *Whiteness measure functions for the three variational models performed (top row) and ISNR/SSIM values (middle and bottom rows) for the test image* `phantom` *corrupted by Gaussian blur with* `band`*=9,* `sigma`*=2 and AWGN with* $\sigma = 0.1$.

6. Conclusions. We proposed an automatic strategy for the selection of the regularization parameters μ_1, μ_2 in the RBD-L_2 variational model for restoration of piecewise smooth images. The proposed strategy relies on the extension of the well-known residual whiteness principle to the multi-parameter scenario, to which we referred as Multi-parameter residual whiteness principle, in short MRWP. The MRWP has been applied for the solution of the RBD-L_2 model when the only information available on the corrupting additive Gaussian noise is its whiteness. The MRWP has been also embedded along the iterations of a suitable ADMM-based optimization scheme, thus yielding the IMRWP-ADMM approach, which allows for an efficient and automatic selection of the target parameters based on the whiteness criterion. The proposed approach has been tested on the restoration of different test images for different degradation levels, whence we concluded that the MRWP inherits the effectiveness of the RWP, originally designed for single-parameter estimates. As a future research direction, we plan to extend the MRWP in its *a posteriori* version to different variational models involving in their expressions multiple unknown parameters, possibly related to the employment of more articulated regularizers, or fidelity terms yielded by mixed noise corruption.

Acknowledgements. This work was supported in part by MIUR RFO projects, in part by the National Group for Scientific Computation (INDAM-GNCS), Research Projects 2022

ETNA
Kent State University and
Johann Radon Institute (RICAM)

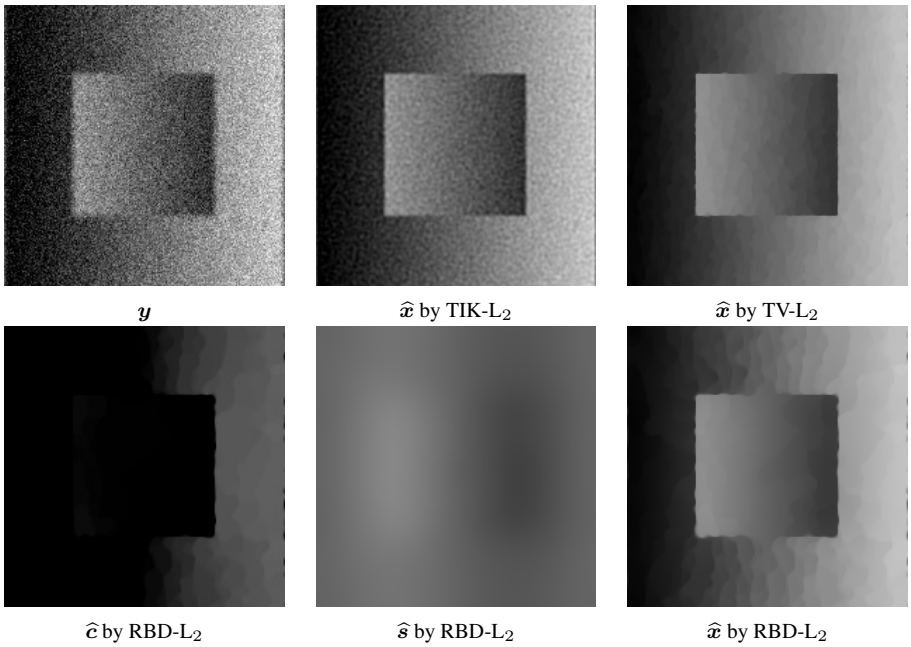


FIG. 5.7. *Output results for the test image* `phantom` *corrupted by Gaussian blur with* `band` *= 9,* `sigma` *= 2, and AWGN with* $\sigma = 0.1$.

TABLE 5.4
Output ISNR/SSIM values and μ_1, μ_2 *obtained by employing the MRWP and the IMRWP for the solution of the RBD-L*$_2$ *model for images corrupted by Gaussian blur with* `band` *= 9,* `sigma` *= 2, and AWGN with* $\sigma = 0.1$.

	`phantom`		`peppers`	
	MRWP	IMRWP-ADMM	MRWP	IMRWP-ADMM
μ_1	44223.67	34206.12	6720.54	7056.82
μ_2	19.21	17.10	24.53	27.14
ISNR	15.0326	15.0122	5.3749	5.3612
SSIM	0.9376	0.9369	0.6956	0.6945

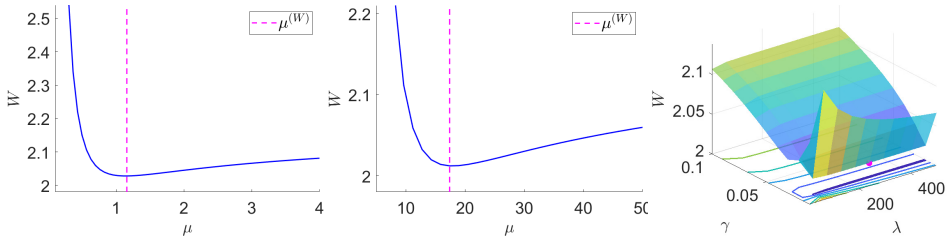
"Metodi numerici per l'imaging: dal 2D al 3D" (code CUP_E55F22000270001) and "Modelli e metodi avanzati in Computer Vision" (code CUP_E53C22001930001).

Appendix A. In what follows, we prove the results stated in Sections 2 and 4.

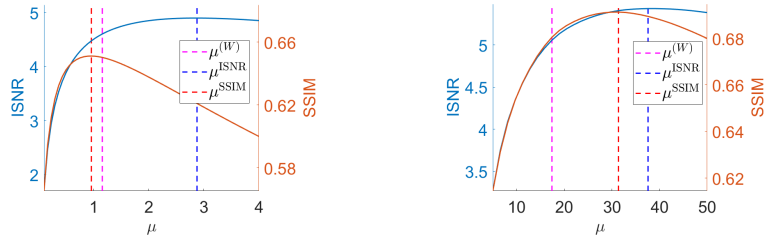
Proof of Lemma 2.1. Under the assumption of periodic boundary conditions, it is quite well-known that the finite difference matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ in (1.3) with convolution kernels in (2.1) have the same null space containing constant images, as stated in (2.9). Then, it follows easily from assumption (2.2) and from definition (2.5) that matrix $\mathbf{M}_1$ is a symmetric, BCCB, positive definite matrix and, hence, admits symmetric, BCCB positive definite inverse $\mathbf{M}_1^{-1}$, for any $(\mu_1, \mu_2) \in \mathbb{R}^2_{++}$. Matrix $\mathbf{M}_2$ in (2.6) is given by the negative product of the two square symmetric matrices $\mathbf{M}_1^{-1}$ and $\mathbf{K}^{\mathrm{T}}\mathbf{K}$, with $\mathbf{M}_1^{-1}$ positive definite and $\mathbf{K}^{\mathrm{T}}\mathbf{K}$ positive semidefinite. Since both matrices are BCCB matrices, they commute, i.e.,

$$\mathbf{M}_1^{-1}(\mathbf{K}^{\mathrm{T}}\mathbf{K}) = (\mathbf{K}^{\mathrm{T}}\mathbf{K})\mathbf{M}_1^{-1}.$$

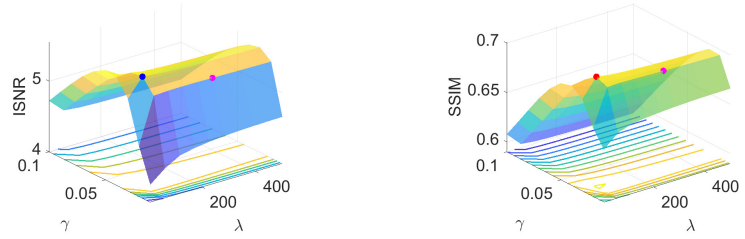
ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) $W(\mu)$ for TIK-L$_2$ (b) $W(\mu)$ for TV-L$_2$ (c) $W(\lambda,\gamma)$ for RBD-L$_2$



(d) ISNR(μ)/SSIM(μ) for TIK-L$_2$ (e) ISNR(μ)/SSIM(μ) for TV-L$_2$



(f) ISNR(λ,γ) for RBD-L$_2$ (g) SSIM(λ,γ) for RBD-L$_2$

FIG. 5.8. *Whiteness measure functions for the three variational models performed (top row) and ISNR/SSIM values (middle and bottom rows) for the test image* `peppers` *corrupted by Gaussian blur with* `band` *= 9,* `sigma` *= 2, and AWGN with* $\sigma = 0.1$.

Hence, their product is symmetric and positive semidefinite with null space equal to the null space of $\mathbf{K}$. Matrix $\mathbf{M}_2$ in (2.6) is thus symmetric negative semidefinite with the same null space. Matrix $\mathbf{A}$ in (2.7) is clearly a symmetric, BCCB matrix and, by resorting to its Fourier diagonalization, it is easy to prove that it is positive semidefinite with null space equal to the null space of $\mathbf{D}_1$. Finally, matrix $\mathbf{B}$ in (2.8) is the product of two symmetric, BCCB matrices, hence it is a symmetric BCCB matrix with null space equal to the null space of $\mathbf{K}$. □

Proof of Lemma 2.2. Function $\mathcal{J}$ in (1.6) is given by the sum of three terms which are all clearly proper, continuous, convex and bounded below by zero independently of the observed data $\boldsymbol{y} \in \mathbb{R}^n$ and of the parameter pair $(\mu_1, \mu_2) \in \mathbb{R}^2_{++}$, so $\mathcal{J}$ also has these properties. Then, replacing in the expression (1.6) of $\mathcal{J}$ all points belonging to straight lines $\ell_{\bar{\boldsymbol{t}}}(t) \subset \mathbb{R}^{2n}$ parallel to the versor $\boldsymbol{d}$ in (2.12) and passing through a generic point $\bar{\boldsymbol{t}} = (\bar{\boldsymbol{c}}; \bar{\boldsymbol{s}})$, in formula

$$\ell_{\bar{\boldsymbol{t}}}(t) := \bar{\boldsymbol{t}} + t\boldsymbol{d} = (\bar{\boldsymbol{c}} + \alpha \mathbf{1}_n, \bar{\boldsymbol{s}} - \alpha \mathbf{1}_n), \qquad t \in \mathbb{R},\ \alpha = t(2n)^{-1/2} \in \mathbb{R},$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

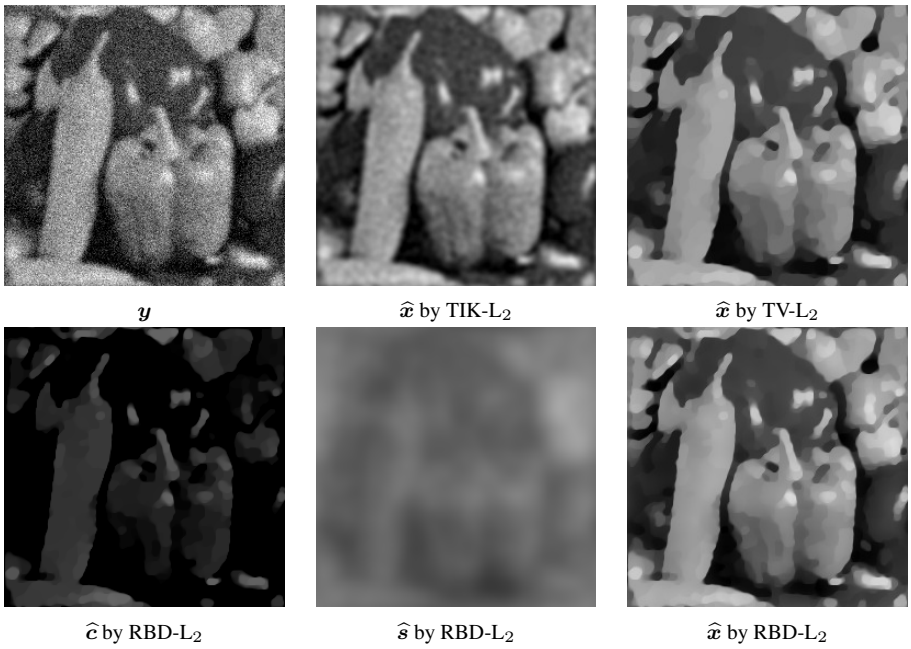


FIG. 5.9. *Output results for the test image* `peppers` *corrupted by Gaussian blur with* `band` *= 9,* `sigma` *= 2, and AWGN with* $\sigma = 0.1$.

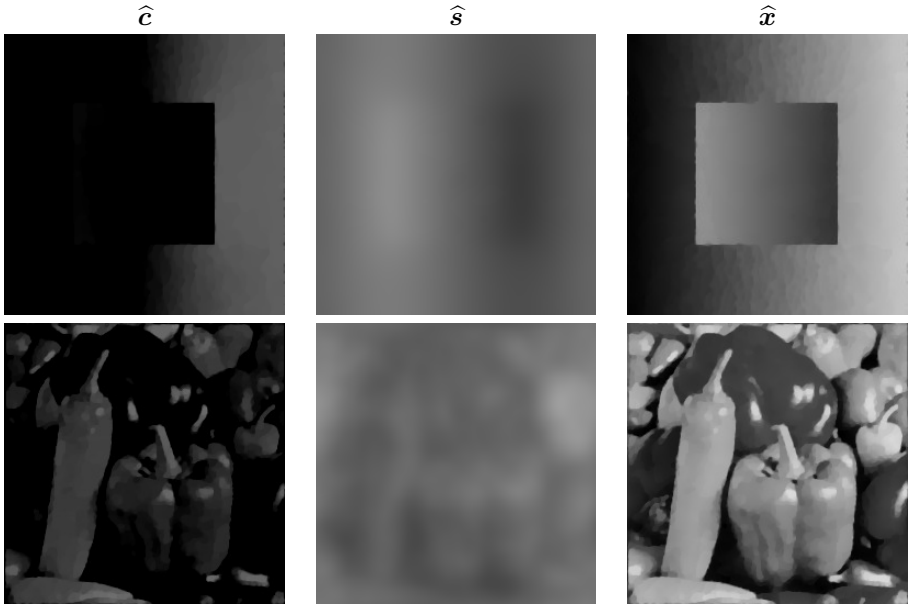


FIG. 5.10. *Output restorations by the IRWP-ADMM employed for the solution of the RBD-*L_2 *model for the restoration of* `phantom` *(top) and* `peppers` *(bottom) corrupted by Gaussian blur with* `band` *= 5,* `sigma` *= 1, and AWGN with* $\sigma = 0.05$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

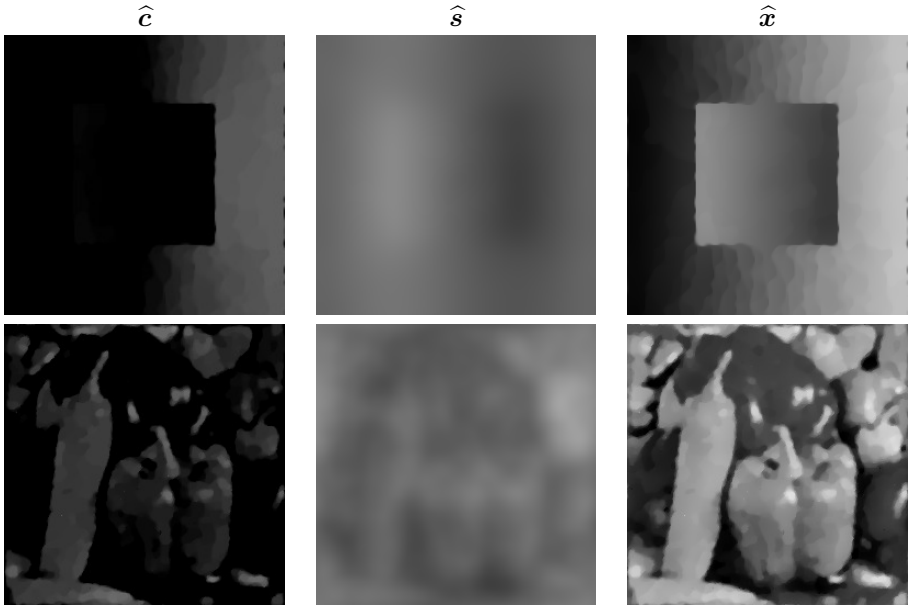


FIG. 5.11. *Output restorations by the IRWP-ADMM employed for the solution of the RBD-L_2 model for the restoration of* `phantom` *(top) and* `peppers` *(bottom) corrupted by Gaussian blur with* `band` *= 9,* `sigma` *= 2, and AWGN with* $\sigma = 0.1$.

one gets

$$\begin{aligned}\mathcal{J}\left(\ell_{\bar{\boldsymbol{t}}}(t);\mu_1,\mu_2\right) &= \sum_{i=1}^{n}\left\|\left(\mathbf{D}_1\left(\bar{\boldsymbol{c}}+\alpha\mathbf{1}_n\right)\right)_i\right\|_2+\frac{\mu_1}{2}\left\|\mathbf{D}_2\left(\bar{\boldsymbol{s}}-\alpha\mathbf{1}_n\right)\right\|_2^2\\ &\quad+\frac{\mu_2}{2}\left\|\mathbf{K}\left(\bar{\boldsymbol{c}}+\alpha\mathbf{1}_n+\bar{\boldsymbol{s}}-\alpha\mathbf{1}_n\right)-\boldsymbol{y}\right\|_2^2\\ &= \sum_{i=1}^{n}\left\|\left(\mathbf{D}_1\bar{\boldsymbol{c}}\right)_i\right\|_2+\frac{\mu_1}{2}\left\|\mathbf{D}_2\bar{\boldsymbol{s}}\right\|_2^2+\frac{\mu_2}{2}\left\|\mathbf{K}\left(\bar{\boldsymbol{c}}+\bar{\boldsymbol{s}}\right)-\boldsymbol{y}\right\|_2^2 \qquad \text{(A.1)}\\ &= \mathcal{J}\left(\bar{\boldsymbol{t}};\mu_1,\mu_2\right)\quad \forall\bar{\boldsymbol{t}}\in\mathbb{R}^{2n},\qquad \forall t\in\mathbb{R},\end{aligned}$$

where in (A.1) we applied the properties $\mathbf{D}_1\mathbf{1}_n = \mathbf{0}_{2n}$ and $\mathbf{D}_2\mathbf{1}_n = \mathbf{0}_{4n}$, which come from (2.9). Function $\mathcal{J}$ is thus constant along straight lines of direction $\boldsymbol{d}$, therefore it is not coercive. Finally, (2.13) can be proved after noting that by restricting $\mathcal{J}$ to any hyperplane which intersects in a single point any straight line of direction $\boldsymbol{d}$ yields coercivity of $\mathcal{J}$. □

Proof of Proposition 2.4. First, we notice that, for any fixed value of the first optimization variable (component) $\boldsymbol{c}$, the function $\mathcal{J}(\boldsymbol{c},\boldsymbol{s};\mu_1,\mu_2)$ in (1.6) is quadratic and strongly convex in the second optimization variable $\boldsymbol{s}$, independently of the observed data $\boldsymbol{y}\in\mathbb{R}^n$ and of the parameter pair $(\mu_1,\mu_2)\in\mathbb{R}^2_{++}$. In fact, it is easy to verify that the gradient $\boldsymbol{\nabla}_{\boldsymbol{s}}\mathcal{J}\in\mathbb{R}^n$ and Hessian matrix $\boldsymbol{\nabla}^2_{\boldsymbol{s}}\mathcal{J}\in\mathbb{R}^{n\times n}$ of function $\mathcal{J}$ with respect to $\boldsymbol{s}$ are given by

$$\begin{aligned}\boldsymbol{\nabla}_{\boldsymbol{s}}\mathcal{J}\left(\boldsymbol{c},\boldsymbol{s};\mu_1,\mu_2\right) &= \mu_2\mathbf{M}_1(\mu_1,\mu_2)\boldsymbol{s}+\mu_2\mathbf{K}^{\mathrm{T}}\left(\mathbf{K}\boldsymbol{c}-\boldsymbol{y}\right),\\ \boldsymbol{\nabla}^2_{\boldsymbol{s}}\mathcal{J}\left(\boldsymbol{c},\boldsymbol{s};\mu_1,\mu_2\right) &= \mu_2\mathbf{M}_1(\mu_1,\mu_2),\end{aligned}$$

where matrix $\mathbf{M}_1(\mu_1,\mu_2)$, according to its definition in (2.5) and to the statement of Lemma 2.1, is symmetric and positive definite, hence non-singular. It follows that, for any fixed $\boldsymbol{c}$, the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

global minimiser of function $\mathcal{J}$ with respect to $\boldsymbol{s}$ is unique and coincides with the zero of the (linear) gradient function $\boldsymbol{\nabla}_{\boldsymbol{s}}\mathcal{J}$ in (A). After some simple algebraic manipulations, this implies that the global minimizers of function $\mathcal{J}$ in (1.6) take the following form

$$\{\widehat{\boldsymbol{c}}\,(\mu_1,\mu_2)\,,\widehat{\boldsymbol{s}}\,(\mu_1,\mu_2)\}\,,\quad \text{with}\ \ \widehat{\boldsymbol{s}}\,(\mu_1,\mu_2)=\mathbf{M}_2(\mu_1,\mu_2)\widehat{\boldsymbol{c}}\,(\mu_1,\mu_2)+\boldsymbol{a}(\mu_1,\mu_2), \tag{A.2}$$

with the matrix $\mathbf{M}_2(\mu_1,\mu_2)$ defined in (2.6) and the vector $\boldsymbol{a}(\mu_1,\mu_2)$ defined in (2.18). Hence, the solution image $\widehat{\boldsymbol{x}}(\mu_1,\mu_2)$ in the original RBD-L$_2$ model (1.4)–(1.6) clearly reads

$$\widehat{\boldsymbol{x}}(\mu_1,\mu_2)=\widehat{\boldsymbol{c}}(\mu_1,\mu_2)+\mathbf{M}_2\widehat{\boldsymbol{c}}\,(\mu_1,\mu_2)+\boldsymbol{a}(\mu_1,\mu_2)=\mathbf{A}(\mu_1,\mu_2)\widehat{\boldsymbol{c}}(\mu_1,\mu_2)+\boldsymbol{a}(\mu_1,\mu_2),$$

with the matrix $\mathbf{A}(\mu_1,\mu_2)$ defined in (2.7).

We can thus plug the expression of $\boldsymbol{s}$ as a function of $\boldsymbol{c}$ given in (A.2) into the original cost function $\mathcal{J}$ in (1.6), so that the minimization problem (1.5)–(1.6) turns into

$$\widehat{\boldsymbol{c}}\,(\mu_1,\mu_2)\in\underset{\boldsymbol{c}\in\mathbb{R}^n}{\operatorname{argmin}}\left\{\widetilde{\mathcal{J}}_{\boldsymbol{c}}\,(\boldsymbol{c};\mu_1,\mu_2)=\sum_{i=1}^{n}\|\mathbf{D}_1\boldsymbol{c}\|_2+\mathcal{Q}_{\boldsymbol{c}}\,(\boldsymbol{c};\mu_1,\mu_2)\right\}, \tag{A.3}$$

with the quadratic function $\mathcal{Q}_{\boldsymbol{c}}:\mathbb{R}^n\to\mathbb{R}$ defined by

$$\mathcal{Q}_{\boldsymbol{c}}\,(\boldsymbol{c};\mu_1,\mu_2)=\overbrace{\frac{\mu_1}{2}\left\|\mathbf{D}_2\left(\mathbf{M}_2\boldsymbol{c}+\boldsymbol{a}\right)\right\|_2^2}^{U(\boldsymbol{c};\mu_1,\mu_2)}+\overbrace{\frac{\mu_2}{2}\left\|\mathbf{K}\left(\boldsymbol{c}+\mathbf{M}_2\boldsymbol{c}+\boldsymbol{a}\right)-\boldsymbol{y}\right\|_2^2}^{V(\boldsymbol{c};\mu_1,\mu_2)}, \tag{A.4}$$

where for shortness we drop the dependencies of $\mathbf{M}_2$ and $\boldsymbol{a}$ on (μ_1,μ_2). We observe that

$$\begin{aligned}U(\boldsymbol{c};\mu_1,\mu_2)&=\frac{\mu_1}{2}(\mathbf{D}_2\mathbf{M}_2\boldsymbol{c}+\mathbf{D}_2\boldsymbol{a})^{\mathrm{T}}(\mathbf{D}_2\mathbf{M}_2\boldsymbol{c}+\mathbf{D}_2\boldsymbol{a})\\&=\frac{\mu_1}{2}\boldsymbol{c}^{\mathrm{T}}\mathbf{M}_2^{\mathrm{T}}\mathbf{D}_2^{\mathrm{T}}\mathbf{D}_2\mathbf{M}_2\boldsymbol{c}+\mu_1\boldsymbol{c}^{\mathrm{T}}\mathbf{M}_2^{\mathrm{T}}\mathbf{D}_2^{\mathrm{T}}\mathbf{D}_2\boldsymbol{a}+\kappa_U,\end{aligned} \tag{A.5}$$

and that

$$\begin{aligned}V(\boldsymbol{c};\mu_1,\mu_2)&=\frac{\mu_2}{2}\left\|\mathbf{K}(\mathbf{I}_n+\mathbf{M}_2)\boldsymbol{c}+\mathbf{K}\boldsymbol{a}-\boldsymbol{y}\right\|_2^2\\&=\frac{\mu_2}{2}\left\|\mathbf{K}\mathbf{A}\boldsymbol{c}+\mathbf{K}\boldsymbol{a}-\boldsymbol{y}\right\|_2^2\\&=\frac{\mu_2}{2}(\mathbf{K}\mathbf{A}\boldsymbol{c}+\mathbf{K}\boldsymbol{a}-\boldsymbol{y})^{\mathrm{T}}(\mathbf{K}\mathbf{A}\boldsymbol{c}+\mathbf{K}\boldsymbol{a}-\boldsymbol{y})\\&=\frac{\mu_2}{2}\boldsymbol{c}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{K}\mathbf{A}\boldsymbol{c}+\mu_2\boldsymbol{c}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}(\mathbf{K}\boldsymbol{a}-\boldsymbol{y})+\kappa_V,\end{aligned} \tag{A.6}$$

with the matrix $\mathbf{A}$ defined in (2.7) and where the terms κ_U in (A.5) and κ_V in (A.6) both depend on (μ_1,μ_2) but not on the optimization variable $\boldsymbol{c}$.

It follows from (A.4)–(A.6) that

$$\begin{aligned}\mathcal{Q}_{\boldsymbol{c}}\,(\boldsymbol{c};\mu_1,\mu_2)&=\frac{1}{2}\boldsymbol{c}^{\mathrm{T}}\left(\mu_1\mathbf{M}_2^{\mathrm{T}}\mathbf{D}_2^{\mathrm{T}}\mathbf{D}_2\mathbf{M}_2+\mu_2\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{K}\mathbf{A}\right)\boldsymbol{c}\\&\quad+\boldsymbol{c}^{\mathrm{T}}\left(\mu_1\mathbf{M}_2^{\mathrm{T}}\mathbf{D}_2^{\mathrm{T}}\mathbf{D}_2\boldsymbol{a}+\mu_2\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}(\mathbf{K}\boldsymbol{a}-\boldsymbol{y})\right)+\kappa_U+\kappa_V\qquad\text{(A.7)}\\&=\frac{1}{2}\boldsymbol{c}^{\mathrm{T}}\mathbf{B}\boldsymbol{c}-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b}+\kappa_U+\kappa_V,\qquad\text{(A.8)}\end{aligned}$$

with the matrix $\mathbf{B}$ defined in (2.8) and the vector $\boldsymbol{b}$ in (2.19), and where (A.8) comes from (A.7) by replacing $\mathbf{A}=\mathbf{I}_n+\mathbf{M}_2$ and then carrying out some algebraic manipulations. In fact, we

ETNA
Kent State University and
Johann Radon Institute (RICAM)

have

$$\begin{aligned}
\mathbf{B} &= \mu_1 \mathbf{M}_2^{\mathrm{T}} \mathbf{D}_2^{\mathrm{T}} \mathbf{D}_2 \mathbf{M}_2 + \mu_2 (\mathbf{I}_n + \mathbf{M}_2^{\mathrm{T}}) \mathbf{K}^{\mathrm{T}} \mathbf{K} (\mathbf{I}_n + \mathbf{M}_2) \\
&= \mu_1 \mathbf{M}_2^{\mathrm{T}} \mathbf{D}_2^{\mathrm{T}} \mathbf{D}_2 \mathbf{M}_2 + \mu_2 (\mathbf{K}^{\mathrm{T}} \mathbf{K} + \mathbf{M}_2^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} \mathbf{K} + \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_2 + \mathbf{M}_2^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_2) \\
&= \mathbf{M}_2^{\mathrm{T}} (\mu_1 \mathbf{D}_2^{\mathrm{T}} \mathbf{D}_2 + \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K}) \mathbf{M}_2 + \mu_2 (\mathbf{K}^{\mathrm{T}} \mathbf{K} + \mathbf{M}_2^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} \mathbf{K} + \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_2) \\
&= \mu_2 \mathbf{M}_2^{\mathrm{T}} \mathbf{M}_1 \mathbf{M}_2 + \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_2 + \mu_2 (\mathbf{K}^{\mathrm{T}} \mathbf{K} + \mathbf{M}_2^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} \mathbf{K}) \\
&= \mu_2 \big(- \mathbf{M}_1^{-1} \mathbf{K}^{\mathrm{T}} \mathbf{K} \big)^{\mathrm{T}} \mathbf{M}_1 \mathbf{M}_2 + \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_2 + \mu_2 (\mathbf{I}_n + \mathbf{M}_2^{\mathrm{T}}) \mathbf{K}^{\mathrm{T}} \mathbf{K} \qquad \text{(A.9)} \\
&= -\mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_2 + \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_2 + \mu_2 (\mathbf{I}_n + \mathbf{M}_2) \mathbf{K}^{\mathrm{T}} \mathbf{K} \qquad \text{(A.10)} \\
&= \mu_2 \mathbf{A} \mathbf{K}^{\mathrm{T}} \mathbf{K}, \qquad \text{(A.11)}
\end{aligned}$$

where in (A.9) we replaced the expression (2.6) of $\mathbf{M}_2$ and, then, in (A.10) and (A.11) we exploited the symmetry of matrices $\mathbf{M}_1, \mathbf{M}_2$ - as proved in Lemma 2.1 - and the definition (2.7) of matrix $\mathbf{A}$, respectively. Relying on the same properties and definitions, we have

$$\begin{aligned}
\boldsymbol{b} &= -\mu_1 \mathbf{M}_2^{\mathrm{T}} \mathbf{D}_2^{\mathrm{T}} \mathbf{D}_2 \boldsymbol{a} - \mu_2 \mathbf{A}^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} (\mathbf{K} \boldsymbol{a} - \boldsymbol{y}) \\
&= \mu_1 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{D}_2^{\mathrm{T}} \mathbf{D}_2 \boldsymbol{a} - \mu_2 (\mathbf{I}_n - \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1}) \mathbf{K}^{\mathrm{T}} \mathbf{K} \boldsymbol{a} + \mu_2 (\mathbf{I}_n - \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1}) \mathbf{K}^{\mathrm{T}} \boldsymbol{y} \\
&= \mu_1 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{D}_2^{\mathrm{T}} \mathbf{D}_2 \boldsymbol{a} - \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \boldsymbol{a} + \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{K}^{\mathrm{T}} \mathbf{K} \boldsymbol{a} \\
&\quad + \mu_2 \mathbf{K}^{\mathrm{T}} \boldsymbol{y} - \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{K}^{\mathrm{T}} \boldsymbol{y} \\
&= \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} (\mu_1 \mathbf{D}_2^{\mathrm{T}} \mathbf{D}_2 + \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K}) \boldsymbol{a} - \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \boldsymbol{a} + \mu_2 \mathbf{K}^{\mathrm{T}} \boldsymbol{y} - \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{K}^{\mathrm{T}} \boldsymbol{y} \\
&= \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{M}_1 \boldsymbol{a} - \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \boldsymbol{a} + \mu_2 \mathbf{K}^{\mathrm{T}} \boldsymbol{y} - \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{K}^{\mathrm{T}} \boldsymbol{y} \\
&= \mu_2 \mathbf{K}^{\mathrm{T}} \boldsymbol{y} - \mu_2 \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1} \mathbf{K}^{\mathrm{T}} \boldsymbol{y} = \mu_2 (\mathbf{I}_n - \mathbf{K}^{\mathrm{T}} \mathbf{K} \mathbf{M}_1^{-1}) \mathbf{K}^{\mathrm{T}} \boldsymbol{y} \\
&= \mu_2 (\mathbf{I}_n + \mathbf{M}_2^{\mathrm{T}}) \mathbf{K}^{\mathrm{T}} \boldsymbol{y} = \mu_2 \mathbf{A} \mathbf{K}^{\mathrm{T}} \boldsymbol{y}.
\end{aligned}$$

By replacing the obtained expression (A.8) of $\mathcal{Q}_{\boldsymbol{c}}$ into (A.3), dropping the constant (with respect to the optimization variable $\boldsymbol{c}$) term $\kappa_U + \kappa_V$ and including the indicator function of set $\mathcal{C}$ defined in (2.20), we get the optimization problem (2.16)–(2.17).

Then, based on definition (2.13), we have

$$\mathcal{H}^{\theta=0}_{\boldsymbol{v}=(\mathbf{1}_n;\mathbf{0}_n)} = \left\{ \boldsymbol{t} = (\boldsymbol{c}; \boldsymbol{s}) \in \mathbb{R}^{2n} : \mathbf{1}_n^{\mathrm{T}} \boldsymbol{c} = 0 \Longleftrightarrow \sum_{i=1}^{N} c_i = 0 \right\} = \mathcal{C} \times \mathbb{R}^n, \qquad \text{(A.12)}$$

with $\mathcal{C} \subset \mathbb{R}^n$ the set in (2.20). Since clearly

$$\boldsymbol{v}^{\mathrm{T}} \boldsymbol{d} = (\mathbf{1}_n, \mathbf{0}_n)(2n)^{-1/2} (\mathbf{1}_n; -\mathbf{1}_n) = (2n)^{-1/2} n = \sqrt{n/2} \neq 0,$$

then the set in (A.12) belongs to the class of hyperplane constraints introduced in (2.13) which, according to Lemma 2.2, guarantee coercivity of function $\mathcal{J}$ in (1.6) and, as a consequence of Proposition 2.3, also guarantee uniqueness of the pairs of solution components $\{\widehat{\boldsymbol{c}}(\mu_1, \mu_2), \widehat{\boldsymbol{s}}(\mu_1, \mu_2)\}$ of the minimization problem (1.5)–(1.6). The constraint set in (A.12) for the original RBD-L$_2$ model thus corresponds to the constraint set $\mathcal{C}$ in (2.20) for the reduced RBD-L$_2$ model. It follows that the reduced, linearly constrained minimization problem in (2.16)–(2.17) admits a unique solution. □

Proof of Lemma 4.1. Since by hypothesis matrix $\mathbf{S}$ is BCCB, it can be diagonalized by the 2D discrete Fourier transform as formalized in (4.13). Moreover, as $\mathbf{S}$ is assumed to be positive definite, hence nonsingular, it admits inverse and the quadratic program (4.12) admits a unique solution (minimization of a strongly convex quadratic function over an hyperplane

ETNA
Kent State University and
Johann Radon Institute (RICAM)

constraint). The Lagrangian function $L(\boldsymbol{p},\nu)$ associated to the quadratic program (4.12) reads

$$L(\boldsymbol{p},\nu) = Q(\boldsymbol{p}) + \nu\sum_{i=1}^{n} p_i = \frac{1}{2}\boldsymbol{p}^{\mathrm{T}}\mathbf{S}\boldsymbol{p} + \boldsymbol{p}^{\mathrm{T}}(\nu\mathbf{1}_n - \boldsymbol{z}).$$

The solution of the system of Lagrangian conditions $\boldsymbol{\nabla}_{\boldsymbol{p}} L(\boldsymbol{p},\nu) = 0 \wedge \boldsymbol{\nabla}_{\nu} L(\boldsymbol{p},\nu) = 0$ is

$$\widehat{\nu} = \frac{\mathbf{1}_n^{\mathrm{T}}\mathbf{S}^{-1}\boldsymbol{z}}{\mathbf{1}_n^{\mathrm{T}}\mathbf{S}^{-1}\mathbf{1}_n}, \qquad \widehat{\boldsymbol{p}} = \mathbf{S}^{-1}\boldsymbol{z} - \widehat{\nu}\,\mathbf{S}^{-1}\mathbf{1}_n = \widehat{\boldsymbol{p}}^{(U)} - \widehat{\nu}\,\mathbf{S}^{-1}\mathbf{1}_n,$$

where $\widehat{\boldsymbol{p}}^{(U)} = \mathbf{S}^{-1}\boldsymbol{z}$ denotes the solution of the unconstrained version of (4.12). Multiplying (on the left) by the Fourier transform matrix $\mathbf{F}$ both sides of the second equality above, and noting that, clearly, the inverse of $\mathbf{S}$ admits the Fourier diagonalizaton $\mathbf{S}^{-1} = \mathbf{F}^{*}\widetilde{\mathbf{S}}^{-1}\mathbf{F}$, we get

$$\mathbf{F}\widehat{\boldsymbol{p}} = \mathbf{F}\widehat{\boldsymbol{p}}^{(U)} - \widehat{\nu}\mathbf{F}\left(\mathbf{F}^{*}\widetilde{\mathbf{S}}^{-1}\mathbf{F}\right)\mathbf{1}_n \iff \widetilde{\widehat{\boldsymbol{p}}} = \widetilde{\widehat{\boldsymbol{p}}}^{(U)} - \widehat{\nu}\,\widetilde{\mathbf{S}}^{-1}(\mathbf{F}\mathbf{1}_n)\,, \tag{A.13}$$

with the Lagrange multiplier $\widehat{\nu}$ given by

$$\widehat{\nu} = \frac{\mathbf{1}_n^{\mathrm{T}}\left(\mathbf{F}^{*}\widetilde{\mathbf{S}}^{-1}\mathbf{F}\right)\boldsymbol{z}}{\mathbf{1}_n^{\mathrm{T}}\left(\mathbf{F}^{*}\widetilde{\mathbf{S}}^{-1}\mathbf{F}\right)\mathbf{1}_n} = \frac{(\mathbf{F}\mathbf{1}_n)^{*}\,\widetilde{\mathbf{S}}^{-1}\widetilde{\boldsymbol{z}}}{(\mathbf{F}\mathbf{1}_n)^{*}\,\widetilde{\mathbf{S}}^{-1}(\mathbf{F}\mathbf{1}_n)}. \tag{A.14}$$

Then, by recalling the well-know property that the (unitary) 2D discrete Fourier transform of a (vectorized) image of all ones satisfies

$$\mathbf{F}\mathbf{1}_n = \frac{1}{\sqrt{n}}(1;\mathbf{0}_{n-1}) \iff (\mathbf{F}\mathbf{1}_n)^{*} = \frac{1}{\sqrt{n}}(1;\mathbf{0}_{n-1})^{\mathrm{T}},$$

the complex vector $\widetilde{\widehat{\boldsymbol{p}}}$ in (A.13) with $\widetilde{z}$ in (A.14) reads

$$\begin{aligned}\widetilde{\widehat{\boldsymbol{p}}} &= \widetilde{\widehat{\boldsymbol{p}}}^{(U)} - \frac{\cancel{\frac{1}{\sqrt{n}}}(1;\mathbf{0}_{n-1})^{\mathrm{T}}\widetilde{\mathbf{S}}^{-1}\widetilde{\boldsymbol{z}}}{\cancel{\frac{1}{\sqrt{n}}}(1;\mathbf{0}_{n-1})^{\mathrm{T}}\widetilde{\mathbf{S}}^{-1}\cancel{\frac{1}{\sqrt{n}}}(1;\mathbf{0}_{n-1})}\widetilde{\mathbf{S}}^{-1}\cancel{\frac{1}{\sqrt{n}}}(1;\mathbf{0}_{n-1}) \\ &= \widetilde{\widehat{\boldsymbol{p}}}^{(U)} - \frac{\widetilde{z}_1/\widetilde{s}_1}{1/\widetilde{s}_1}\frac{1}{\widetilde{s}_1}(1;\mathbf{0}_{n-1}) = \widetilde{\widehat{\boldsymbol{p}}}^{(U)} - \frac{\widetilde{z}_1}{\widetilde{s}_1}(1;\mathbf{0}_{n-1}).\end{aligned} \tag{A.15}$$

The proof is completed by recalling that

$$\widetilde{\widehat{\boldsymbol{p}}}_i^{(U)} = \widetilde{z}_i/\widetilde{s}_i$$

and replacing this expression in (A.15). □

Proof of Proposition 4.2. In the following derivations, the iteration super-indices will be dropped out for the sake of better readability.

According to Lemma 4.1, the solution of the unconstrained version of problem (4.10) and (4.11) is given by

$$\boldsymbol{c}^{(U)}(\mu_1,\mu_2) = \mathbf{F}^{*}\widetilde{\mathbf{S}}^{-1}(\mu_1,\mu_2)\widetilde{\boldsymbol{z}}(\mu_1,\mu_2),$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with

$$\mathbf{S}(\mu_1,\mu_2)=\frac{1}{\beta}\mathbf{B}(\mu_1,\mu_2)+\mathbf{D}_1^{\mathrm{T}}\mathbf{D}_1,\qquad \boldsymbol{z}(\mu_1,\mu_2)=\frac{1}{\beta}\boldsymbol{b}(\mu_1,\mu_2)+\mathbf{D}_1^{\mathrm{T}}\boldsymbol{q}.$$

The above minimizer can be plugged into (3.3) so as to get the corresponding residual image

$$\begin{aligned}\boldsymbol{r}^{(U)}(\mu_1,\mu_2)&=\mathbf{K}(\mathbf{A}(\mu_1,\mu_2)\mathbf{F}^*\widetilde{\mathbf{S}}^{-1}(\mu_1,\mu_2)\widetilde{\boldsymbol{z}}(\mu_1,\mu_2)+\boldsymbol{a}(\mu_1,\mu_2))-\boldsymbol{y}\\&=\mathbf{K}(\mathbf{A}(\mu_1,\mu_2)\mathbf{F}^*\mathbf{F}\left(\frac{1}{\beta}\mathbf{B}(\mu_1,\mu_2)+\mathbf{D}_1^{\mathrm{T}}\mathbf{D}_1\right)^{-1}\mathbf{F}^*\mathbf{F}\left(\frac{1}{\beta}\boldsymbol{b}(\mu_1,\mu_2)+\mathbf{D}_1^{\mathrm{T}}\boldsymbol{q}\right)\\&\quad+\mathbf{K}\boldsymbol{a}(\mu_1,\mu_2)-\boldsymbol{y},\end{aligned}$$

with the index (U) indicating that in the above expression the constraint $\boldsymbol{c}\in\mathcal{C}$ has not been yet accounted. In order to simplify the following calculations, the parameters μ_1,μ_2 are mapped into the novel

$$\lambda:=\frac{\mu_1}{\mu_2},\quad \gamma:=\frac{\mu_2}{\beta}. \tag{A.16}$$

Notice that the matrices and the vectors defined in Lemma 2.1 and in Proposition 2.4 can be easily re-parametrized in terms of λ,γ as follows

$$\begin{aligned}\mathbf{M}_1(\lambda)&=\lambda\mathbf{D}_2^{\mathrm{T}}\mathbf{D}_2+\mathbf{K}^{\mathrm{T}}\mathbf{K},&\qquad \mathbf{M}_2(\lambda)&=-\mathbf{M}_1^{-1}(\lambda)\mathbf{K}^{\mathrm{T}}\mathbf{K},\\ \mathbf{A}(\lambda)&=\mathbf{I}_n+\mathbf{M}_2(\lambda),&\qquad \boldsymbol{a}(\lambda)&=\mathbf{M}_1^{-1}(\lambda)\mathbf{K}^{\mathrm{T}}\boldsymbol{y}\end{aligned}$$

while, with a little abuse of notations, in what follows we are going to denote

$$\mathbf{B}(\lambda,\gamma)=\gamma\mathbf{A}(\lambda)\mathbf{K}^{\mathrm{T}}\mathbf{K},\qquad \boldsymbol{b}(\lambda,\gamma)=\gamma\mathbf{A}(\lambda)\mathbf{K}^{\mathrm{T}}\boldsymbol{y}.$$

Hence, the $\boldsymbol{r}^{(U)}(\mu_1,\mu_2)$ can be equivalently re-written as

$$\boldsymbol{r}^{(U)}(\lambda,\gamma)=\mathbf{K}\mathbf{A}(\lambda)\mathbf{F}^*\mathbf{F}\left(\mathbf{B}(\lambda,\gamma)+\mathbf{D}_1^{\mathrm{T}}\mathbf{D}_1\right)^{-1}\mathbf{F}^*\mathbf{F}\left(\boldsymbol{b}(\lambda,\gamma)+\mathbf{D}_1^{\mathrm{T}}\boldsymbol{q}\right)+\mathbf{K}\boldsymbol{a}(\lambda)-\boldsymbol{y}. \tag{A.17}$$

Relying on the diagonalizations provided in (2.3) and (2.4), we have that the expression of the residual image in (A.17) can be manipulated by suitably multiplying by $\mathbf{F},\mathbf{F}^*$ and exploiting the property $\mathbf{F}^*\mathbf{F}=\mathbf{I}_n$, thus yielding

$$\boldsymbol{r}^{(U)}(\lambda,\gamma)=\mathbf{F}^*\widetilde{\mathbf{K}}(\mathbf{I}+\widetilde{\mathbf{M}}_2(\lambda))(\widetilde{\mathbf{B}}(\lambda,\gamma)+|\widetilde{\mathbf{D}}|_1^2)^{-1}(\widetilde{\boldsymbol{b}}(\lambda,\gamma)+\overline{\widetilde{\mathbf{D}}}_1\boldsymbol{q})+\mathbf{F}^*\widetilde{\mathbf{K}}\widetilde{\boldsymbol{a}}(\lambda)-\mathbf{F}^*\widetilde{\boldsymbol{y}},$$

whence

$$\widetilde{\boldsymbol{r}}^{(U)}(\lambda,\gamma)=\widetilde{\mathbf{K}}(\mathbf{I}+\widetilde{\mathbf{M}}_2(\lambda))(\widetilde{\mathbf{B}}(\lambda,\gamma)+|\widetilde{\mathbf{D}}|_1^2)^{-1}(\widetilde{\boldsymbol{b}}(\lambda,\gamma)+\overline{\widetilde{\mathbf{D}}}_1\boldsymbol{q})+\widetilde{\mathbf{K}}\widetilde{\boldsymbol{a}}(\lambda)-\widetilde{\boldsymbol{y}}.$$

We introduce

$$\begin{aligned}\epsilon_i&=|\widetilde{k}_i|^2,&\qquad \eta_i&=|\widetilde{d}_{hh,i}|^2+2|\widetilde{d}_{hv,i}|^2+|\widetilde{d}_{vv,i}|^2,\\ \zeta_i&=|\widetilde{d}_{h,i}|^2+|\widetilde{d}_{v,i}|^2,&\qquad \widetilde{z}_i&=\overline{\widetilde{d}}_{h,i}\widetilde{q}_{1,i}+\overline{\widetilde{d}}_{v,i}\widetilde{q}_{2,i},\end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

so that the i-th component of $\widetilde{\boldsymbol{r}}^{(U)}$ can be written as

$$\begin{aligned}
\widetilde{\widetilde{r}}_i^{(U)}(\lambda,\gamma) &= \tilde{k}_i\left(1-\frac{\epsilon_i}{\lambda\eta_i+\epsilon_i}\right)\left(\gamma\epsilon_i\left(1-\frac{\epsilon_i}{\lambda\eta_i+\epsilon_i}\right)+\zeta_i\right)^{-1}\times \\
&\quad\times\left[\tilde{z}_i+\gamma\left(1-\frac{\epsilon_i}{\lambda\eta_i+\epsilon_i}\right)\bar{\tilde{k}}_i\tilde{y}_i-\gamma\left(1-\frac{\epsilon_i}{\lambda\eta_i+\epsilon_i}\right)\frac{\epsilon_i}{\lambda\eta_i+\epsilon_i}\bar{\tilde{k}}_i\tilde{y}_i\right. \\
&\quad\left.+\gamma\lambda\frac{\epsilon_i\eta_i}{(\lambda\eta_i+\epsilon_i)^2}\bar{\tilde{k}}_i\tilde{y}_i\right]-\left(1-\frac{\epsilon_i}{\lambda\eta_i+\epsilon_i}\right)\tilde{y}_i \\
&= \tilde{k}_i\left(\frac{\lambda\eta_i}{\lambda\eta_i+\epsilon_i}\right)\left(\frac{\gamma\lambda\eta_i\epsilon_i+\zeta_i(\lambda\eta_i+\epsilon_i)}{\lambda\eta_i+\epsilon_i}\right)^{-1}\left[\tilde{z}_i+\gamma\frac{\lambda\eta_i}{\lambda\eta_i+\epsilon_i}\bar{\tilde{k}}_i\tilde{y}_i\right. \\
&\quad\left.\cancel{-\gamma\lambda\frac{\epsilon_i\eta_i}{(\lambda\eta_i+\epsilon_i)^2}\bar{\tilde{k}}_i\tilde{y}_i+\gamma\lambda\frac{\epsilon_i\eta_i}{(\lambda\eta_i+\epsilon_i)^2}\bar{\tilde{k}}\tilde{y}_i}\right]-\frac{\lambda\eta_i}{\lambda\eta_i+\epsilon_i}\tilde{y}_i \\
&= \frac{\lambda\eta_i}{\gamma\lambda\eta_i\epsilon_i+\zeta_i(\lambda\eta_i+\epsilon_i)}\left[\tilde{k}_i\tilde{z}_i+\gamma\lambda\frac{\eta_i\epsilon_i}{\lambda\eta_i+\epsilon_i}\tilde{y}_i\right. \\
&\quad\left.-\frac{\gamma\lambda\eta_i\epsilon_i+\zeta_i(\lambda\eta_i+\epsilon_i)}{\lambda\eta_i}\frac{\lambda\eta_i}{\lambda\eta_i+\epsilon_i}\tilde{y}_i\right] \\
&= \frac{\lambda\eta_i}{\gamma\lambda\eta_i\epsilon_i+\zeta_i(\lambda\eta_i+\epsilon_i)}\left[\tilde{k}_i\tilde{z}_i+\gamma\lambda\frac{\eta_i\epsilon_i}{\lambda\eta_i+\epsilon_i}\tilde{y}_i-\frac{\gamma\lambda\eta_i\epsilon_i}{\lambda\eta_i+\epsilon_i}\tilde{y}_i-\zeta_i\tilde{y}_i\right] \\
&= \frac{\lambda\eta_i}{\gamma\lambda\eta_i\epsilon_i+\zeta_i(\lambda\eta_i+\epsilon_i)}\left[\tilde{k}_i\tilde{z}_i-\zeta_i\tilde{y}_i\right].
\end{aligned}$$

As highlighted in Lemma 4.1, accounting for the constraint $\boldsymbol{c}\in\mathcal{C}$ amounts to require $\widetilde{c}_1=0$, so that the Fourier transform of the residual image corresponding to the solution of the original constrained problem (4.10) and (4.11) is given, componentwise, by

$$\widetilde{r}_i(\lambda,\gamma)=\begin{cases}-\left(1-\dfrac{\epsilon_i}{\lambda\eta_i+\epsilon_i}\right)\widetilde{y}_i & i=1,\\[2ex] \widetilde{r}_i^{(U)} & i=2,\dots,n.\end{cases}$$

Hence, after introducing

$$\phi_i=|\widetilde{k}_i\widetilde{z}_i-\zeta_i\widetilde{y}_i|,$$

we have that $w_i(\lambda,\gamma)$ takes the form

$$w_i(\lambda,\gamma)=|\widetilde{r}_i(\lambda,\gamma)|=\begin{cases}\dfrac{\lambda\eta_i}{\lambda\eta_i+\epsilon_i}|\widetilde{y}_i| & i=1,\\[2ex] \dfrac{\lambda\eta_i\phi_i}{\gamma\lambda\eta_i\epsilon_i+\zeta_i(\lambda\eta_i+\epsilon_i)} & i=2,\dots,n.\end{cases}$$

The above functions $w_i(\lambda,\gamma)$ can be plugged into (3.4) so as to get the final expression of the whiteness function W in terms of λ,γ, i.e., μ_1,μ_2. □

REFERENCES

[1] F. BEVILACQUA, A. LANZA, M. PRAGLIOLA, AND F. SGALLARI, *Masked unbiased principles for parameter selection in variational image restoration under Poisson noise*, Inverse Problems, 39 (2023), Art. 034002 (33 pages).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[2] ———, *Whiteness-based parameter selection for Poisson data in variational image processing*, Appl. Math. Model., 117 (2023), pp. 197–218.
[3] S. BOYD, N. PARIKH, E. CHU, B. PELEATO, AND J. ECKSTEIN, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learn., 3 (2011), pp. 1–122.
[4] A. BUCCINI, L. REICHEL, M. PRAGLIOLA, AND F. SGALLARI, *A comparison of parameter choice rules for ℓ_p-ℓ_q minimization*, Ann. dell'Università di Ferrara, 68 (2022), pp. 441—463.
[5] T. F. CHAN, S. ESEDOGLU, AND F. PARK, *A fourth order dual method for staircase reduction in texture extraction and image restoration problems*, in 2010 IEEE International Conference on Image Processing, IEEE Conference Proceedings, Los Alamitos, 2010, pp. 4137–4140.
[6] J. ECKSTEIN AND D. P. BERTSEKAS, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Mathematical Programming, 55 (1992), pp. 293–318.
[7] A. GHOLAMI AND S. HOSSEINI, *A balanced combination of Tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals*, Signal Process., 93 (2013), pp. 1945–1960.
[8] P. HANSEN, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
[9] M. HOCHSTENBACH, L. REICHEL, AND G. RODRIGUEZ, *Regularization parameter determination for discrete ill-posed problems*, J. Comput. Appl. Math., 273 (2015), pp. 132–149.
[10] M. HUSKA, S.-H. KANG, A. LANZA, AND S. MORIGI, *A variational approach to additive image decomposition into structure, harmonic, and oscillatory components*, SIAM J. Imaging Sci., 14 (2021), pp. 1749–1789.
[11] M. HUSKA, A. LANZA, S. MORIGI, AND I. SELESNICK, *A convex-nonconvex variational method for the additive decomposition of functions on surfaces*, Inverse Problems, 35 (2019), Art. 124008 (34 pages).
[12] A. LANZA, M. PRAGLIOLA, AND F. SGALLARI, *Residual whiteness principle for parameter-free image restoration*, Electron. Trans. Numer. Anal., 53 (2020), pp. 329–351.
https://etna.ricam.oeaw.ac.at/vol.53.2020/pp329-351.dir/pp329-351.pdf
[13] V. MOROZOV, *On the solution of functional equations by the method of regularization*, Dokl. Akad. Nauk SSSR, 167 (1966), pp. 510–512.
[14] M. PRAGLIOLA, L. CALATRONI, A. LANZA, AND F. SGALLARI, *ADMM-based residual whiteness principle for automatic parameter selection in single image super-resolution problems*, J. Math. Imaging Vision, 65 (2023), pp. 99–123.
[15] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[16] Z. WANG, A. BOVIK, H. SHEIKH, AND E. SIMONCELLI, *Image quality assessment: From error visibility to structural similarity*, IEEE Trans. Image Process., 4 (2004), pp. 600–612.