Hessian-Based Norm Regularization for Image Restoration With Biomedical Applications

Stamatios Lefkimmiatis, *Member, IEEE*, Aurélien Bourquard, *Member, IEEE*, and Michael Unser, *Fellow, IEEE*

***Abstract*—We present nonquadratic Hessian-based regularization methods that can be effectively used for image restoration problems in a variational framework. Motivated by the great success of the total-variation (TV) functional, we extend it to also include second-order differential operators. Specifically, we derive second-order regularizers that involve matrix norms of the Hessian operator. The definition of these functionals is based on an alternative interpretation of TV that relies on mixed norms of directional derivatives. We show that the resulting regularizers retain some of the most favorable properties of TV, i.e., convexity, homogeneity, rotation, and translation invariance, while dealing effectively with the staircase effect. We further develop an efficient minimization scheme for the corresponding objective functions. The proposed algorithm is of the iteratively reweighted least-square type and results from a majorization–minimization approach. It relies on a problem-specific preconditioned conjugate gradient method, which makes the overall minimization scheme very attractive since it can be applied effectively to large images in a reasonable computational time. We validate the overall proposed regularization framework through deblurring experiments under additive Gaussian noise on standard and biomedical images.**

***Index Terms*—Biomedical imaging, Frobenius norm, Hessian matrix, image deblurring, linear inverse problems, majorization–minimization (MM) algorithms, spectral norm.**

I. Introduction

DURING image acquisition, artifacts are mainly caused by blurring, which is linked to imaging processes such as diffraction or aberrations, and by noise, which is intrinsic to the detection process. To alleviate these effects, image restoration can serve as a desirable preprocessing technique. In this paper, we are dealing with regularization-based restoration methods where the regularizer consists of second-order derivative terms.

The choice of the regularizer significantly affects the quality of the restored image. Several regularization approaches have already been proposed for image restoration in a variational framework, e.g., the total-variation (TV) seminorm that currently provides state-of-the-art results. TV regularization was initially introduced for denoising [1] and has been effectively applied to other inverse problems as well, such as image restoration [2]–[4], inpainting [5], and zooming [6]. The success and widespread use of the TV functional in the past two decades can be mainly attributed to its ability to produce well-preserved and sharp edges. Moreover, its convexity permits the design of efficient algorithms [7].

While TV regularization is extremely popular in a variety of applications, it also gives rise to some undesired effects. In particular, it leads to the well-known staircase effect when applied to signals that are not necessarily piecewise constant [8]–[10]. Indeed, TV favors piecewise-vanishing first-order derivatives and thus yields solutions belonging to the class of piecewise-constant functions. This tendency can be highly undesirable, particularly in applications such as biomedical imaging where image interpretation can be severely obstructed. The main challenge is thus to deal with the staircase effect while preserving image sharpness and convexity.

To attenuate the staircase effect, there is a growing interest in the literature for replacing TV by the L_1 norm of some higher order differential operator. The motivation behind such attempts is to restore potentially a wider class of images, which comprise more than merely piecewise-constant regions. The majority of these regularizers involve second-order differential operators because piecewise-vanishing second-order derivatives lead to piecewise-linear solutions that better fit smooth intensity changes.

Second-order regularization schemes that have been considered so far in the literature were mainly targeted to the problem of image denoising. In this case, the proposed variational methods are either combining a second-order regularizer with the TV seminorm [8], [9], [11], [12] or employing a second-order regularizer in a standalone way [10], [13]–[15]. The regularizers that have already been used on this problem include the L_1 norm of the Laplacian operator [8], [13], [15]

$$R_L(f) = \int_{\mathbf{\Omega}} |\Delta f(\mathbf{x})|\, d\mathbf{x} \tag{1}$$

where $\mathbf{\Omega} \subset \mathbb{R}^2$, $\Delta f(\mathbf{x}) = f_{xx}(\mathbf{x}) + f_{yy}(\mathbf{x})$ with $f_{ij}(\mathbf{x}) = (\partial^2/\partial i \partial j) f(\mathbf{x})$, the modified Laplacian [14], [15]

$$R_M(f) = \int_{\mathbf{\Omega}} |f_{xx}(\mathbf{x})| + |f_{yy}(\mathbf{x})|\, d\mathbf{x} \tag{2}$$

the Frobenius norm of the Hessian [9], [14], [15]

$$R_F(f) = \int_{\mathbf{\Omega}} \sqrt{f_{xx}^2(\mathbf{x}) + 2f_{xy}^2(\mathbf{x}) + f_{yy}^2(\mathbf{x})}\, d\mathbf{x} \tag{3}$$

and the affine TV functional [10]

$$R_A(f) = \int_{\mathbf{\Omega}} \left(\sqrt{f_{xx}^2(\mathbf{x}) + f_{xy}^2(\mathbf{x})} + \sqrt{f_{yx}^2(\mathbf{x}) + f_{yy}^2(\mathbf{x})} \right) d\mathbf{x}. \tag{4}$$

Manuscript received April 07, 2011; revised July 16, 2011 and September 01, 2011; accepted September 01, 2011. Date of publication September 19, 2011; date of current version February 17, 2012. This work was supported in part by the Hasler Foundation and in part by the Indo–Swiss Joint Research Program. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Brendt Wohlberg.

The authors are with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: stamatis.lefkimmiatis@epfl.ch; aurelien.bourquard@epfl.ch; michael.unser@epfl.ch).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2011.2168232

For image denoising, these regularizers have been shown either to perform better than TV or to complement it favorably. This motivates us to investigate higher order regularization methods that constitute valid extensions of TV and that can be applied effectively to the more general problem of image restoration/deblurring.

In this paper, our main contributions are twofold. First, we introduce two second-order extensions of the TV functional. Their definition is based on an alternative interpretation of TV that we derive and that relies on mixed norms of the first directional derivative. The obtained second-order functionals turn out to be well-suited to the task of image restoration, particularly for biomedical images. In particular, we prove that they preserve some of the most favorable properties of TV, i.e., convexity, homogeneity, rotation, and translation invariance. Second, we propose a unifying computational framework for image restoration based on second-order regularization. More specifically, we develop an iteratively reweighted least-squares-type (IRLS) algorithm that is derived following a majorization–minimization (MM) approach and is applicable to all the second-order regularizers that we have referred to. To further speed up our minimization scheme, we also devise a problem-specific preconditioned conjugate gradient (CG) algorithm. We show experimentally that one can significantly accelerate the optimization task, as compared with the standard CG algorithm or with other commonly used preconditioning schemes. Finally, we provide systematic comparisons between the studied second-order schemes and TV, which currently provides state-of-the-art results.

The rest of this paper is organized as follows. In Section II, we present the mathematical formulation of the image-deblurring problem and discuss higher order regularization schemes. Considering first the 1-D setting in Section II-B, we then introduce second-order extensions of TV for the 2-D case in Section II-C. In Section III, we obtain tight quadratic upper bounds for the second-order regularizers and describe our minimization algorithm. Then, in Section IV, we assess the performance of our approach on two different data sets consisting of standard and biomedical images. We finally conclude this paper in Section V.

II. Higher Order Regularization for Inverse Problems

A. Problem Formulation

Image deblurring belongs to the general family of inverse problems and amounts to estimating an image f given measurements y. The most commonly used image observation model involves linear measurements and can be expressed as

$$y = \mathcal{A}f + w \tag{5}$$

where $\mathcal{A}$ is a linear blur operator and w is the unknown noise affecting the measurements. The recovery of f from measurements y is an ill-posed problem [16], due to the presence of the noise and to the fact that the blurring operator $\mathcal{A}$ is usually ill-conditioned or, even, noninvertible. In general, the ill-posed nature of the problem implies that there are too many ways one can obtain an approximate solution. A reasonable strategy is to reformulate the image-deblurring problem by taking into account the image formation and acquisition process as well as any available prior information about the properties of the image to be restored.

The most common approach for restoring signal f is to form an objective function, which quantifies the quality of a given estimate and has the following form:

$$J(f) = J_{\text{data}}(f) + \tau R(f). \tag{6}$$

The first term is known as data fidelity and measures the consistency between the estimation and the measurements, whereas the second one is the regularization term, designed to penalize solutions that deviate significantly from the expected properties of the image to be restored. The regularization parameter $\tau \geq 0$ balances the contribution of the two terms. Image restoration can then be cast as the minimization of (6).

B. One-Dimensional Regularization

Given 1-D signals f of finite spatial support $\Omega \subset \mathbb{R}$ with appropriate continuity properties, the general form of regularizers based on L_q norms of derivatives can be expressed as

$$R(f; p, q) = \int_{\Omega} |D^p f(x)|^q \, dx = \|D^p f\|_q^q \tag{7}$$

where $D^p = \partial^p / \partial x^p$ is the pth-order derivative and q determines how "irregularities" are penalized. It is common to choose $q \geq 1$ to ensure the convexity of the regularizer.

In this paper, we only deal with cases $q = 1, 2$, which are the most commonly used in practice and which correspond to the L_1 and L_2 norms of the pth-order derivative of f. For $q = 2$, the corresponding regularizer is known as Tikhonov regularization [16] and has been widely used in the past because of its simple closed-form solution. However, in the past decade, there has been an increasing interest in using the L_1 norm. The main reason is that L_1-norm regularization has been found to be more appropriate and to lead to improved results for the problem of image restoration [1]. Its use still ensures convexity of the problem, implying that the computation of an optimal solution can be efficiently performed [7], even if no closed-form solution exists. The underlying philosophy in prefering L_1-based over L_2-based regularization is that L_1 norms are less sensitive to outliers and are therefore associated with better edge reconstruction capabilities. Indeed, edges correspond to high local magnitude values of $|D^p f|$, which are less penalized *a priori* than with norms of higher order.

To illustrate this issue, we consider a denoising problem (i.e., $\mathcal{A}$ matches the identity operator $\mathcal{I}$) for a synthetic signal consisting of piecewise-constant, piecewise-linear, and piecewise-parabolic regions. The results are shown in Fig. 1 and are obtained using regularizers of norm $q = 1, 2$ for differential operators of order $p = 1, 2$. Irrespectively of order p, the results clearly show that the use of the L_1 norm is more appropriate since $q = 1$ better restores the abrupt transitions of the signal by creating less oscillations. Meanwhile, we observe that the first-order solution (TV) restores the piecewise-constant regions of the signal properly but creates staircase artifacts elsewhere. However, for $q = 1$ and $p = 2$, smoother transitions are more accurately restored. These results motivate the use of

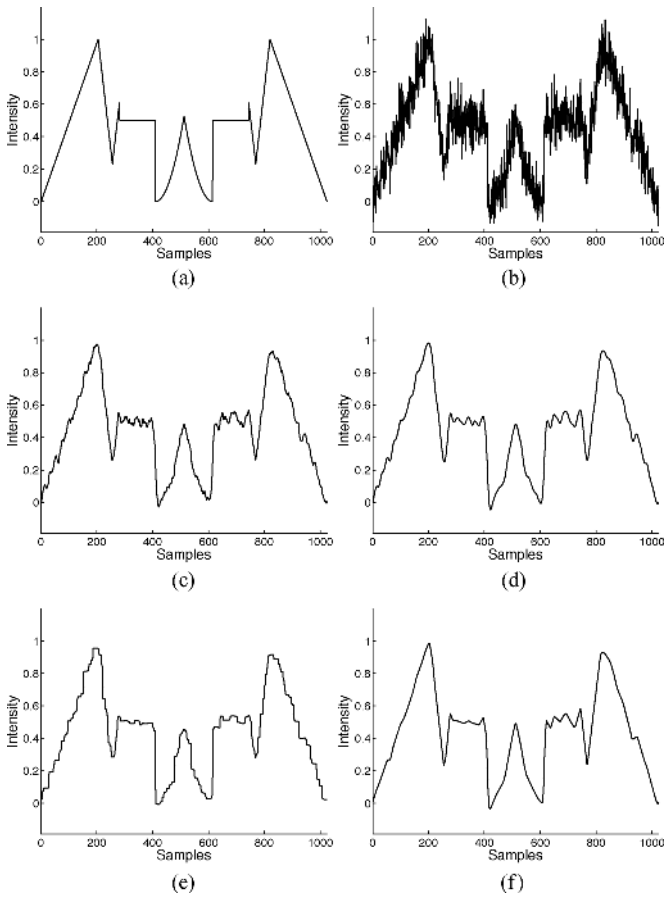


Fig. 1. Denoising a synthetic signal consisting of piecewise-constant, piecewise-linear, and piecewise-parabolic regions. Intensity profile of (a) original signal, (b) signal degraded by Gaussian noise of standard deviation $\sigma = 0.1$, (c) first-order L_2-norm solution, (d) second-order L_2-norm solution, (e) first-order L_1-norm solution, and (f) second-order L_1-norm solution.

higher order operators in two dimensions because, as in the 1-D case, natural images contain regions that deviate from a piecewise-constant model.

C. Two-Dimensional Regularization

The standard extension of TV to two dimensions, assuming that image f is continuously differentiable, is

$$\mathrm{TV}(f) = \int_{\mathbf{\Omega}} \|\nabla f(\mathbf{x})\|_2 \, d\mathbf{x}. \tag{8}$$

It involves the image gradient ∇f, where $\|\nabla f\|_2$ is its Euclidean norm and where $\mathbf{\Omega} \subset \mathbb{R}^2$. This regularizer is closely linked to its 1-D counterpart $R(f;1,1)$. Indeed, the 2-D TV functional can be equivalently written as

$$\mathrm{TV}(f) = \frac{1}{h(q)} \int_{\mathbf{\Omega}} \left\| D^1_{\theta} f(\mathbf{x}) \right\|_{L_q[0,2\pi]} d\mathbf{x} \tag{9}$$

where $h(q) = \|\cos(\theta)\|_{L_q[0,2\pi]}$ and D^1_{θ} is the first directional derivative along angle θ, defined as $D^1_{\theta} f = \langle \nabla f, \mathbf{u}_{\theta} \rangle = \nabla f \cdot \mathbf{u}_{\theta}$ with $\mathbf{u}_{\theta} = (\cos\theta, \sin\theta)$. Therefore, according to (9), TV can be interpreted as a mixed $L_1 - L_q$ norm where the L_1 norm acts on the image domain, whereas the L_q norm acts along the angular orientation of the directional derivative. This particular expression also highlights the properties of convexity, homogeneity[1], rotation, and translation invariance of the TV functional.

In two dimensions, a natural way to compute second derivatives is to use the Hessian operator

$$\mathcal{H}_f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}. \tag{10}$$

While the use of second derivatives for constructing a regularizer is straightforward in the 1-D setting, there are several possible choices in 2-D. In the literature, several attempts at the definition of second-order regularizers have been reported, as mentioned in Section I. They typically lead to (1) –(4), where it has been required upfront that the regularizers satisfy certain properties such as rotation invariance, convexity, or both.

In this paper, we want to derive second-order regularizers that promote invariances and, at the same time, constitute valid extensions of TV, in the sense that they extend definition (9) to the second-order case. To do so, we increase the order of differentiation using second-order bidirectional derivatives and define our second-order regularizers accordingly. In this case, however, the resulting functional is not the same for different choices of L_q norms. In this paper, we are thus considering the mixed norms $L_1 - L_\infty$ and $L_1 - L_2$, which result in the definitions

$$R_S(f) = \int_{\mathbf{\Omega}} \left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_\infty[0,2\pi]^2} d\mathbf{x} \tag{11}$$

$$R_F(f) = \frac{1}{\pi} \int_{\mathbf{\Omega}} \left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_2[0,2\pi]^2} d\mathbf{x} \tag{12}$$

where $D^2_{\theta,\phi}$ is the second directional derivative along the directions given by angles θ and ϕ, defined as $D^2_{\theta,\phi} f(\mathbf{x}) = D^1_{\theta}(D^1_{\phi} f) = \mathbf{u}_{\theta}^T \mathcal{H}_f(\mathbf{x}) \mathbf{v}_{\phi}$, with $\mathbf{v}_{\phi} = (\cos\phi, \sin\phi)$. For alternative higher order extensions of TV based on tensor algebra, we refer the interested reader to two recent works in [17] and [18].

In the following two lemmas, we prove that the integrands of (11) and (12) correspond to the spectral and Frobenius norms of the Hessian, respectively.

Lemma 1: The L_∞ norm of the second directional derivative of f at coordinates $\mathbf{x}$ is equal to the spectral norm of the Hessian matrix $\|\mathcal{H}_f(\mathbf{x})\|_2$.

Proof: The spectral norm of a matrix $\mathbf{B}$ is defined as [19, Ch. 5]

$$\|\mathbf{B}\|_2 = \max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^H \mathbf{B} \mathbf{x}|$$

where ^{H} denotes the Hermitian transpose. Based on this definition, we have

$$\begin{aligned} \left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_\infty[0,2\pi]^2} &= \max_{\theta,\phi} \left| D^2_{\theta,\phi} f(\mathbf{x}) \right| \\ &= \max_{\theta,\phi} \left| \mathbf{u}_{\theta}^T \mathcal{H}_f(\mathbf{x}) \mathbf{v}_{\phi} \right| \\ &= \max_{\|\mathbf{u}\|_2=1} \max_{\|\mathbf{v}\|_2=1} \left| \mathbf{u}^T \mathcal{H}_f(\mathbf{x}) \mathbf{v} \right| \\ &= \|\mathcal{H}_f(\mathbf{x})\|_2 . \end{aligned} \tag{13}$$

■

[1]Homogeneity of function f is considered in the sense that f presents a multiplicative scaling behavior, i.e., $f(\alpha \mathbf{x}) = \alpha^k f(\mathbf{x})$.

TABLE I
DEFINITION OF DIFFERENTIAL OPERATORS

Scalar Operators	Vectorial Operators
$\Delta = \partial_{xx} + \partial_{yy}$	$\nabla = (\partial_x,\, \partial_y)$
$\bar{\Delta} = \partial_{xx} - \partial_{yy}$	$\mathcal{U} = (\partial_{xx} - \partial_{yy},\, 2\partial_{xy})$
$\Gamma = 2\partial_{xy}$	$\mathcal{V} = \left(\partial_{xx}, \sqrt{2}\partial_{xy},\, \partial_{yy}\right)$

The Hessian spectral norm can be alternatively defined as

$$\|\mathcal{H}_f(\mathbf{x})\|_2 = \max_{i=1,2} \left(|\lambda_i f(\mathbf{x})|\right) \tag{14}$$

where $\lambda_i f(\mathbf{x})$ are the eigenvalues of the Hessian matrix of f at coordinates $\mathbf{x}$. These two eigenvalues are expressed as

$$\lambda_{1,2} f(\mathbf{x}) = \frac{\Delta f(\mathbf{x}) \pm \sqrt{\left(\bar{\Delta} f(\mathbf{x})\right)^2 + \left(\Gamma f(\mathbf{x})\right)^2}}{2} \tag{15}$$

where the associated differential operators are defined in Table I. The eigenvalues of the Hessian operator are called principal directions. They are the directions of pure curvature, and are invariant to the rotation of the system of coordinates.

Lemma 2: The L_2 norm of the second directional derivative of f at coordinates $\mathbf{x}$ is proportional to the Frobenius norm of the Hessian matrix $\|\mathcal{H}_f(\mathbf{x})\|_F$.

Proof: The second directional derivative of f can be written as a function of the Hessian eigenvalues. Specifically, since the Hessian matrix is symmetric, we can use the spectral decomposition theorem and express the second directional derivative as

$$\begin{aligned} D^2_{\theta,\phi} f(\mathbf{x}) &= \mathbf{u}_\theta^T \mathbf{Q} \mathbf{\Lambda}_f(\mathbf{x}) \mathbf{Q}^T \mathbf{v}_\phi \\ &= (\mathbf{Q}^T \mathbf{u}_\theta)^T \mathbf{\Lambda}_f(\mathbf{x}) (\mathbf{Q}^T \mathbf{v}_\phi) \end{aligned} \tag{16}$$

where $\mathbf{\Lambda}_f$ is a diagonal matrix with the Hessian eigenvalues and $\mathbf{Q}$ is a 2×2 orthogonal matrix with the eigenvectors in its columns. We can always choose $\mathbf{Q}^T$ to be of the form

$$\mathbf{Q}^T = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

where $a^2 + b^2 = 1$. It is then a rotation matrix by definition. We can therefore write (16) as

$$D^2_{\theta,\phi} f(\mathbf{x}) = \lambda_1 f(\mathbf{x}) \sin(\theta') \sin(\phi') + \lambda_2 f(\mathbf{x}) \cos(\theta') \cos(\phi') \tag{17}$$

where θ' and ϕ' correspond to the rotated versions of θ and ϕ, respectively. Based on (17), we then obtain

$$\begin{aligned} \left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_2[0,2\pi]^2} &= \left(\int_0^{2\pi} \int_0^{2\pi} \left| D^2_{\theta,\phi} f(\mathbf{x}) \right|^2 d\theta d\phi \right)^{1/2} \\ &= \sqrt{\pi^2 \left[\lambda_1^2 f(\mathbf{x}) + \lambda_2^2 f(\mathbf{x})\right]} \\ &= \pi \left\| \mathcal{H}_f(\mathbf{x}) \right\|_F . \end{aligned} \tag{18}$$

■

Definitions (11) and (12) lead to convex regularizers since they arise from the integration of the norm of a linear operator [7, Ch. 3]. In addition, they are also rotation invariant, homogeneous, and translation invariant. The rotation-invariance property becomes explicit when considering that the Hessian matrix norms are fully determined by the eigenvalues. The last two properties are induced by the connection between the regularizers and the second directional derivative. The resulting regularizers are therefore suitable extensions of TV for the second-order case as they retain invariances while following the same principles. It is worthwhile to note that the Hessian Frobenius norm has been introduced as a regularization model in [9], [14], and [15] for image denoising. Meanwhile, the connection we establish between the Frobenius norm and the L_2 norm of $D^2_{\theta,\phi} f$ justifies this regularizer as a valid extension of TV.

We now focus on the interpretation of the spectral-norm regularizer, which does not appear to have been used before for image-processing applications. Using the following identity:

$$\max\left(|\alpha + \beta|, |\alpha - \beta|\right) = |\alpha| + \beta, \quad \forall \beta \geq 0$$

and combining it with (11), (13), and (15), we further expand the spectral-norm regularizer in terms of 2-D second-order differential operators, which yields

$$R_S(f) = \frac{1}{2} \int_{\mathbf{\Omega}} \left(|\Delta f(\mathbf{x})| + \|\mathcal{U} f(\mathbf{x})\|_2 \right) d\mathbf{x} \tag{19}$$

where the vectorial operator $\mathcal{U}$ is defined in Table I. This regularization term can thus be seen as an equally weighted compound regularizer, which is a linear combination of two distinct regularization terms with equal weights, whose first term corresponds to the L_1-norm Laplacian.

To demonstrate the performance of the spectral-norm regularizer, we present a restoration experiment on a synthetic image consisting of both piecewise-constant and piecewise-linear regions. In this example, the image is degraded by Gaussian noise while the blurring operator is, similar to the 1-D case, equal to the identity, which allows us to better evaluate the regularization performance. The results for both the TV and the spectral-norm regularizer are depicted in Fig. 2. As expected, TV restoration preserves sharp edges well at the cost of introducing pronounced staircase artifacts in smooth regions. On the other hand, the Hessian spectral norm keeps a better balance between restoration of piecewise-smooth regions and edge preservation.

Among all possible second-order regularizers, those that seem to be of more interest are those that also satisfy the properties of convexity, homogeneity, translation, and rotation invariance. The L_1-norm Laplacian (1) can be easily verified to satisfy these properties. In particular, convexity ensures that one can find an efficient solution of the problem, whereas the other three properties imply that the solution will remain stable under intensity transformations as well as shifts and rotations of the observed image. On the other hand, the modified Laplacian (2) and affine TV (4) are not rotation invariant. This stems from the fact that the latter two regularizers cannot be fully determined by the Hessian eigenvalues. Since a preferential choice of a coordinate system on the model is not desirable for the task of image restoration, we only consider rotation-invariant

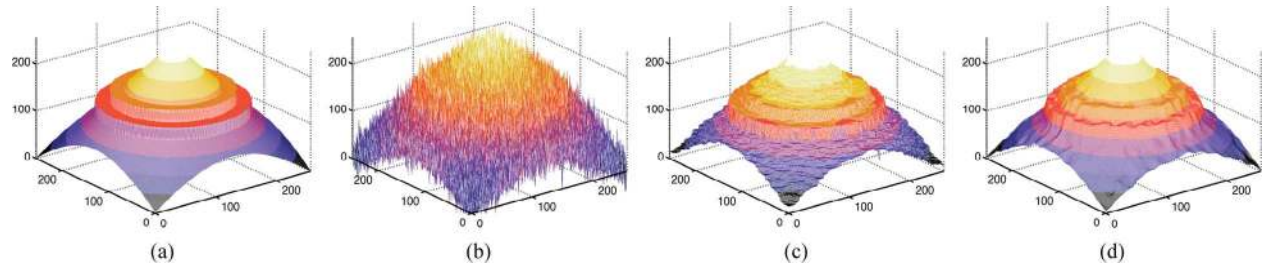


Fig. 2. Denoising of a synthetic image. Intensity profile of (a) original image, (b) degraded image by a Gaussian noise of standard deviation $\sigma = 20$, (c) TV solution, and (d) Spectral-norm solution.

regularizers in the rest of this paper. Notwithstanding that, the minimization method we are proposing in the following section is well suited to the rotation-variant regularizers as well.

III. Minimization of the Objective Function

A. Discrete Problem Formulation

Let us now consider the discrete formulation[2] of (6). In our analysis, we assume that Gaussian noise affects the measurements. Thus, a quadratic data term is appropriate, and the objective function is

$$J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau R(\mathbf{f}) \tag{20}$$

where $\mathbf{A} \in \mathbb{R}^{N\times N}$ is the convolution matrix describing the blurring operation, and $\mathbf{y}$ and $\mathbf{f} \in \mathbb{R}^N$ are N-dimensional vectors that contain the rasterized, observed, and unknown images, respectively, with $N = m \times n$. Depending on the assumed boundary conditions [16], matrix $\mathbf{A}$ can be block circulant with circulant blocks (periodic boundary conditions), block Toeplitz with Toeplitz blocks (zero boundary conditions), or a sum of four different kinds of block matrices (reflexive boundary conditions). To discretize the differential operators described in Section II, we assume periodic boundary conditions for the images and use forward finite differences for approximating second-order derivatives [20, Chap. 4] so that

$$\begin{aligned}
\mathbf{f}_{xx}[i,j] &= \mathbf{f}[i,j] - 2\mathbf{f}[i+1,j] + \mathbf{f}[i+2,j],\\
\mathbf{f}_{yy}[i,j] &= \mathbf{f}[i,j] - 2\mathbf{f}[i,j+1] + \mathbf{f}[i,j+2],\\
\mathbf{f}_{xy}[i,j] &= \mathbf{f}[i,j] - \mathbf{f}[i+1,j] - \mathbf{f}[i,j+1] + \mathbf{f}[i+1,j+1].
\end{aligned} \tag{21}$$

Based on this convention, all discretized operators are realized as block-circulant convolution matrices.

B. MM via Quadratic Upper Bounds

In the MM framework [21], an iterative algorithm for solving the minimization problem

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} J(\mathbf{f}) \tag{22}$$

takes the form

$$\mathbf{f}^{(t+1)} = \arg\min_{\mathbf{f}} Q\left(\mathbf{f}; \mathbf{f}^{(t)}\right) \tag{23}$$

[2]In the sequel, boldface symbols indicate matrices and vectors, in reference to the discretization of the problem.

where $Q(\mathbf{f}; \mathbf{f}^{(t)})$ is the majorizer of function $J(\mathbf{f})$ at a fixed point $\mathbf{f}^{(t)}$ satisfying the following two conditions:

$$Q\left(\mathbf{f}; \mathbf{f}^{(t)}\right) \geq J(\mathbf{f}), \quad \forall \mathbf{f} \tag{24a}$$

$$Q\left(\mathbf{f}^{(t)}; \mathbf{f}^{(t)}\right) = J\left(\mathbf{f}^{(t)}\right). \tag{24b}$$

Instead of minimizing the actual function $J(\mathbf{f})$, we first upper-bound it by a suitable majorizer $Q(\mathbf{f}; \mathbf{f}^{(t)})$ and then minimize this majorizing function to produce the next iterate $\mathbf{f}^{(t+1)}$. Given the properties of the majorizer in (24), iteratively minimizing $Q(\cdot; \mathbf{f}^{(t)})$ also decreases the objective function $J(\cdot)$ [21].

To develop an MM-based algorithm, we derive appropriate quadratic majorizers $Q_R(\mathbf{f}; \mathbf{f}')$ of the studied second-order penalty functions $R(\mathbf{f})$. Since the data fidelity term $J_{\text{data}}(\mathbf{f})$ in (20) is quadratic, expression $Q(\mathbf{f}; \mathbf{f}') = J_{\text{data}}(\mathbf{f}) + Q_R(\mathbf{f}; \mathbf{f}')$ is itself a quadratic majorizer of the complete objective function $J(\mathbf{f})$. This implies that the minimization of the resulting majorizer amounts to solving a system of linear equations, a task for which there are excellent methods available in the literature.

Defining the ℓ_1 norm of a vector field $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2 \ldots, \mathbf{u}_N) \in \mathbb{R}^{N\times k}$ as $\|\mathbf{u}\|_1 = \sum_{1\leq i\leq N} |\mathbf{u}_i|$, where $|\mathbf{u}_i| = (\sum_{1\leq j\leq k} \mathbf{u}_{i,j}^2)^{1/2}$, we rewrite the three regularizers of interest defined in (1), (3), and (19) in their discrete form as

$$R_L(\mathbf{f}) = \|\boldsymbol{\Delta}\mathbf{f}\|_1 = \sum_{i=1}^{N} |(\boldsymbol{\Delta}\mathbf{f})_i| \tag{25}$$

$$\begin{aligned}
R_F(\mathbf{f}) &= \|\mathbf{V}\mathbf{f}\|_1\\
&= \frac{\sqrt{2}}{2}\sum_{i=1}^{N} \sqrt{(\boldsymbol{\Delta}\mathbf{f})_i^2 + (\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2}
\end{aligned} \tag{26}$$

$$\begin{aligned}
R_S(\mathbf{f}) &= \frac{1}{2}\left(\|\boldsymbol{\Delta}\mathbf{f}\|_1 + \|\mathbf{U}\mathbf{f}\|_1\right)\\
&= \frac{1}{2}\left(\sum_{i=1}^{N} |(\boldsymbol{\Delta}\mathbf{f})_i| + \sqrt{(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2}\right)
\end{aligned} \tag{27}$$

where $(\cdot)_i$ corresponds to the ith element of the corresponding matrix–vector product. Next, we consider the two inequalities holding for a general function $g(\cdot) : \mathbb{R} \mapsto \mathbb{R}$

$$|g(x)| \leq \frac{|g(y)|}{2} + \frac{g(x)^2}{2\,|g(y)|}, \quad \forall x, \forall y : g(y) \neq 0 \tag{28a}$$

$$\sqrt{g(x)} \leq \frac{\sqrt{g(y)}}{2} + \frac{g(x)}{2\sqrt{g(x)}}, \quad \forall x : g(x) > 0, \forall y : g(y) \geq 0 \tag{28b}$$

with equality if and only if $g(x) = g(y)$. Using (28a) and the fact that the majorization relation is closed under the formation of sums and nonnegative products [22], we show that the function $Q_L(\mathbf{f}; \mathbf{f}^{(t)})$ defined as

$$Q_L\left(\mathbf{f}; \mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\boldsymbol{\Delta}\mathbf{f}^{(t)}\right\|_1 + \frac{1}{2}\sum_{i=1}^{N}\frac{(\boldsymbol{\Delta}\mathbf{f})_i^2}{\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|} \tag{29}$$

is a majorizer of $R_L(\mathbf{f})$ at the fixed point $\mathbf{f}^{(t)}$. Similarly, (28b) implies that functions $Q_U(\mathbf{f}; \mathbf{f}^{(t)})$ and $Q_F(\mathbf{f}; \mathbf{f}^{(t)})$ defined as

$$Q_U\left(\mathbf{f}; \mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\mathbf{U}\mathbf{f}^{(t)}\right\|_1 + \frac{1}{2}\sum_{i=1}^{N}\frac{\left[(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}} \tag{30}$$

$$Q_F\left(\mathbf{f}; \mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\mathbf{V}\mathbf{f}^{(t)}\right\|_1 + \frac{\sqrt{2}}{4} \times \sum_{i=1}^{N}\frac{\left[(\boldsymbol{\Delta}\mathbf{f})_i^2 + (\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}} \tag{31}$$

are majorizers of $\|\mathbf{U}\mathbf{f}\|_1$ and $R_F(\mathbf{f})$, respectively, at the same fixed point $\mathbf{f}^{(t)}$. By combining (29) and (30), we also obtain a valid majorizer for the spectral-norm regularizer $R_S(\mathbf{f})$ that is expressed as

$$\begin{aligned} Q_S\left(\mathbf{f}; \mathbf{f}^{(t)}\right) &= \frac{1}{2}\left(Q_L\left(\mathbf{f}; \mathbf{f}^{(t)}\right) + Q_U\left(\mathbf{f}; \mathbf{f}^{(t)}\right)\right) \\ &= \frac{1}{4}\sum_{i=1}^{N}\frac{\left[(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}} \\ &\quad + \frac{1}{4}\sum_{i=1}^{N}\frac{(\boldsymbol{\Delta}\mathbf{f})_i^2}{\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|} + \text{const.} \end{aligned} \tag{32}$$

where the constant in (32) is independent of $\mathbf{f}$ but depends only upon the fixed point $\mathbf{f}^{(t)}$. We can now rewrite the majorizers (29), (31), and (32) in the more compact form[3] as follows:

$$Q_R\left(\mathbf{f}; \mathbf{f}^{(t)}\right) = \frac{1}{2}\mathbf{f}^T\left(\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)\mathbf{f} + \text{const.} \tag{33}$$

where for the case of the Laplacian regularizer $\mathbf{M} = \boldsymbol{\Delta}$ and $\mathbf{W}^{(t)} = \mathbf{W}_L^{(t)}$ is a diagonal matrix whose diagonal elements are defined in (35a). For the spectral- and Frobenius-norm regularizers, $\mathbf{M} = [\boldsymbol{\Delta}^T, \bar{\boldsymbol{\Delta}}^T, \boldsymbol{\Gamma}^T]^T$ is a $3N \times N$ matrix, and $\mathbf{W}^{(t)}$ is the block-diagonal $3N \times 3N$ matrix

$$\mathbf{W}^{(t)} = \begin{bmatrix} \mathbf{W}_1^{(t)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2^{(t)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W}_2^{(t)} \end{bmatrix} \tag{34}$$

with $N \times N$ diagonal blocks $\mathbf{W}_1^{(t)}$ and $\mathbf{W}_2^{(t)}$. For the spectral-norm regularizer, it holds that $\mathbf{W}_1^{(t)} = \mathbf{W}_L^{(t)}$ and $\mathbf{W}_2^{(t)} = \mathbf{W}_U^{(t)}$, whereas for the Frobenius one, $\mathbf{W}_1^{(t)} = \mathbf{W}_2^{(t)} = \mathbf{W}_F^{(t)}$. Their diagonal elements are defined accordingly as

$$\left(\mathbf{W}_L^{(t)}\right)_{ii} = \frac{1}{2\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|}, \quad i = 1 \ldots N \tag{35a}$$

$$\left(\mathbf{W}_U^{(t)}\right)_{ii} = \frac{1}{2\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}} \tag{35b}$$

$$\left(\mathbf{W}_F^{(t)}\right)_{ii} = \frac{\sqrt{2}}{2\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}}. \tag{35c}$$

Using (33), we can now write a quadratic majorizer for the complete objective function as

$$\begin{aligned} Q\left(\mathbf{f}; \mathbf{f}^{(t)}\right) &= \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau Q_R\left(\mathbf{f}; \mathbf{f}^{(t)}\right) \\ &= \frac{1}{2}\mathbf{f}^T\left(\mathbf{A}^T\mathbf{A} + \tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)\mathbf{f} \\ &\quad - 2\mathbf{f}^T\mathbf{A}^T\mathbf{y} + \text{const.} \end{aligned} \tag{36}$$

where the constant in (36) does not depend on $\mathbf{f}$ and is thus irrelevant to the optimization task.

Having the total majorizer at hand, we readily proceed with the MM scheme. Since $Q(\mathbf{f}; \mathbf{f}^{(t)})$ is a quadratic function of $\mathbf{f}$, its minimum with respect to $\mathbf{f}$ corresponds to the solution of the linear system of equations

$$\underbrace{\left(\mathbf{A}^T\mathbf{A} + \tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)}_{\mathbf{S}^{(t)}}\mathbf{f}^{(t+1)} = \mathbf{A}^T\mathbf{y}. \tag{37}$$

Obtaining $\mathbf{f}^{(t+1)}$ directly from (37) can be computationally expensive since it demands the inversion of a very large matrix $\mathbf{S}^{(t)}$. For this reason, a common practice is to solve iteratively the problem by employing the CG algorithm [23]. The CG algorithm is a very powerful method in terms of computational cost and memory management since it relies on simple matrix–vector products and does not require the inversion or storage of the matrix $\mathbf{S}^{(t)}$. In fact, the exact solution of (37) is not required since we can still solve (22) by using the weaker condition of decreasing majorizer $Q(\mathbf{f}; \mathbf{f}^{(t)})$ in each iteration, instead of minimizing it. Therefore, in each external iteration, it suffices to run only a few CG iterations. The resulting method then falls into the category of generalized MM algorithms.

C. PCG Method

We use a preconditioned CG (PCG) method to improve the condition number of $\mathbf{S}^{(t)}$ and to achieve a significant decrease in the majorization function within fewer internal iterations. We devise our preconditioner $\mathbf{P}$ as the best circulant approximation of system $\mathbf{S}^{(t)}$, in the sense of [24] and [25]. Specifically, we define

$$\mathbf{P} = \arg\min_{\mathbf{X}\in\mathcal{C}}\left\|\mathbf{S}^{(t)} - \mathbf{X}\right\|_F \tag{38}$$

where $\|\cdot\|_F$ is the Frobenius norm and $\mathcal{C}$ denotes the class of circulant matrices. Since the Frobenius norm of a matrix is unitarily invariant [19], we equivalently have

$$\mathbf{P} = \arg\min_{\mathbf{X}\in\mathcal{C}}\left\|\mathbf{F}\mathbf{S}^{(t)}\mathbf{F}^H - \mathbf{F}\mathbf{X}\mathbf{F}^H\right\|_F \tag{39}$$

[3]For the modified Laplacian and the affine TV regularizer, we can obtain a quadratic majorizer in exactly the same way, where we use inequality (28a) for the former and inequality (28b) for the latter.

where $\mathbf{F}$ denotes the normalized forward discrete Fourier transform (DFT) matrix. Given that $\mathbf{X}$ is diagonalized by the DFT, let us consider the diagonal matrix $\mathbf{X}' = \mathbf{F}\mathbf{X}\mathbf{F}^H$ that is associated to the solution $\mathbf{X}$. Then, we write the best $\mathbf{X}'$ corresponding to $\mathbf{P}$ as

$$\mathbf{P}' = \arg\min_{\mathbf{X}' \in \mathcal{D}} \left\| \mathbf{F}\mathbf{S}^{(t)}\mathbf{F}^H - \mathbf{X}' \right\|_F \tag{40}$$

where $\mathcal{D}$ denotes the class of diagonal matrices. The relation between $\mathbf{X}$ and $\mathbf{X}'$ implies that the solution can be written as

$$\mathbf{P} = \mathbf{F}^H \mathbf{P}' \mathbf{F}. \tag{41}$$

The solution of the constrained minimization problem in (40) in the DFT domain is simply the projection $\mathcal{P}_{\mathcal{D}}$ of $\mathbf{F}\mathbf{S}^{(t)}\mathbf{F}^H$ onto the space of diagonal matrices, which, when combined with (41), yields

$$\mathbf{P} = \mathbf{F}^H \mathcal{P}_{\mathcal{D}} \left(\mathbf{F}\mathbf{S}^{(t)}\mathbf{F}^H \right) \mathbf{F}. \tag{42}$$

Given the structure of $\mathbf{S}^{(t)}$, our preconditioner satisfies the properties of symmetry and positive definiteness, as required for the CG method [23]. We then use left–right preconditioning, resulting in the modified system as follows:

$$\mathbf{S}'^{(t)}\mathbf{f}'^{(t+1)} = \mathbf{y}' \tag{43}$$

where

$$\mathbf{S}'^{(t)} = \mathbf{P}^{-\frac{1}{2}}\mathbf{S}^{(t)}\mathbf{P}^{-\frac{1}{2}} \tag{44}$$

$$\mathbf{y}' = \mathbf{P}^{-\frac{1}{2}}\mathbf{A}^T\mathbf{y}. \tag{45}$$

Accordingly, solution $\mathbf{f}^{(t+1)}$ is obtained as

$$\mathbf{f}^{(t+1)} = \mathbf{P}^{-\frac{1}{2}}\mathbf{f}'^{(t+1)}. \tag{46}$$

The specified form of (42) can be seen as a DFT counterpart of Jacobi preconditioning[4] . By linearity, and from (37) and (42), the circulant preconditioner $\mathbf{P}$ is finally determined as

$$\mathbf{P} = \mathbf{A}^T\mathbf{A} + \tau \sum_k \bar{w}_k \mathbf{M}_k^T \mathbf{M}_k \tag{47}$$

where each scalar $\bar{w}_k$ is the average of the diagonal terms of the corresponding matrix $\mathbf{W}_k^{(t)}$ and $\mathbf{M}_k$ refers to the kth matrix component of the block matrix $\mathbf{M}$.

We compare the performance of our circulant method to that of diagonal preconditioning and standard CG. To do so, we consider averaged results on 1000 deblurring experiments involving 64×64 image patches, which are small enough to obtain closed-form solutions by direct matrix inversion. Specifically, we are minimizing the first quadratic upper bounds of the associated convex objective functions by solving the corresponding linear problems. The performance of the employed methods is then measured in terms of MSE with respect to the closed-form solutions, as a function of the number of iterations.

[4]In case that we are assuming reflexive boundary conditions and that the blurring matrix $\mathbf{A}$ is symmetric, we can still use the same preconditioner with the only difference being that $\mathbf{F}$ now corresponds to the normalized forward Discrete Cosine Transform matrix.

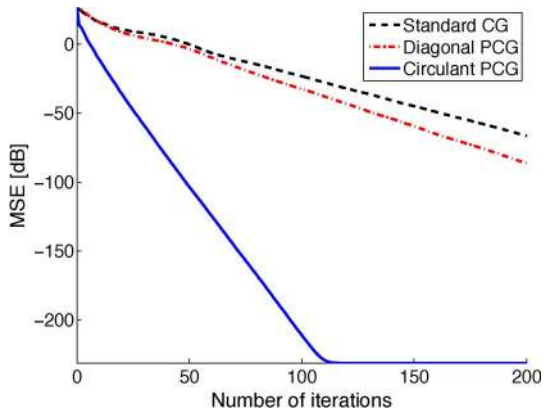


Fig. 3. Comparisons of the convergence rates for the standard CG method versus its diagonal and circulant preconditioned versions. The presented results refer to the average MSE evolution (per iteration) obtained for several instances of a large-scale linear problem.

TABLE II
COMPUTATIONAL PERFORMANCE OF CG VERSUS DIAGONAL AND CIRCULANT PCG VARIANTS

	Time to Convergence	Iterations
CG	25.35 s	211
dPCG	15.92 s	128
cPCG	7.62 s	51

As presented in Fig. 3, the proposed preconditioner shows a significant improvement regarding the convergence of the algorithm. We observe that our PCG method finds the solution with an accuracy value of -50 dB within less than 30 iterations. On the other hand, the diagonal PCG and the standard CG need about 135 and 165 iterations, respectively, to reach the same level of precision. Note also that the circulant PCG practically converges to the exact solution after only 100 iterations; the floor behavior indicates that the solution has reached the maximum available precision. Rounding errors explain why we cannot do much better than -200 dB.

To also provide results of more practical interest, we run the same experiment for an image of size 512×512. In this case, the stopping criterion for the different variants of the CG algorithm is set to reaching a relative normed difference of 10^{-7}. The corresponding results are reported in Table II and clearly demonstrate the superiority of the circulant PCG approach over the other schemes.

D. IRLS Minimization Algorithm

We summarize our proposed algorithm in Table III. Note that the definition of the weighting matrix $\mathbf{W}^{(t)}$ in (34) and (35) requires proper care. Specifically, if, for some pixel i of the image, $(\boldsymbol{\Delta}\mathbf{f}^{(t)})_i$, $|(\mathbf{U}\mathbf{f}^{(t)})_i|$, or $|(\mathbf{V}\mathbf{f}^{(t)})_i|$ are zero, depending on the utilized regularizer, then the corresponding diagonal elements of the matrices $\mathbf{W}_1^{(t)}$ and $\mathbf{W}_2^{(t)}$ take an infinite value, which can prevent convergence. This is known in the MM literature as the singularity issue [22]. It appears because the majorizer cannot be defined when image $\mathbf{f}^{(t)}$ belongs to the null space of the corresponding operator.

This question arises in all the MM-based methods involving quadratic majorizers for nondifferentiable functions, such as

TABLE III
SUMMARY OF THE PROPOSED IRLS ALGORITHM

Input:
$\mathbf{y}$: observed degraded image.
λ: regularization parameter.
N: number of MM iterations.
Output:
$\mathbf{f}^*$: An approximation of the optimal solution of (22).
Initialization:
Set $\mathbf{f}^{(0)} = \mathbf{y}$.
for $t = 0 : N - 1$ **do**
  Update matrix $\mathbf{W}^{(t)}$ according to (34)–(35).
  Set $\mathbf{f}^{(t+1)} = \mathbf{f}^{(t)}$.
  Compute $\mathbf{S}'^{(t)}$, $\mathbf{f}'^{(t+1)}$ and $\mathbf{y}'$ according to (43)–(45).
  while *CG stopping criterion is not satisfied* **do**
    Update $\mathbf{f}'^{(t+1)}$ with one CG iteration for:
    $\mathbf{S}'^{(t)}\mathbf{f}'^{(t+1)} = \mathbf{y}'$.
  end
  Set $\mathbf{f}^{(t+1)} = \mathbf{P}^{-\frac{1}{2}}\mathbf{f}'^{(t+1)}$.
end

the ones developed for TV regularization [3], [26]. A common workaround is to redefine the diagonal matrices as

$$\left(\mathbf{W}_L^{(t)}\right)_{ii} = \frac{1}{2\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \varepsilon}}, \qquad i = 1 \ldots N \tag{48a}$$

$$\left(\mathbf{W}_U^{(t)}\right)_{ii} = \frac{1}{2\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2 + \varepsilon}} \tag{48b}$$

$$\left(\mathbf{W}_F^{(t)}\right)_{ii} = \frac{\sqrt{2}}{2\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2 + \varepsilon}}. \tag{48c}$$

The presence of the small constant $\varepsilon > 0$ bounds the diagonal elements of $\mathbf{W}$ and prevents them to take an infinite value, irrespectively of the values that $(\boldsymbol{\Delta}\mathbf{f}^{(t)})_i$, $|(\mathbf{U}\mathbf{f}^{(t)})_i|$ or $|(\mathbf{V}\mathbf{f}^{(t)})_i|$ might take.

Meanwhile, Figueiredo *et al.* have recently presented some interesting results in [22], which show that the singularity issue can be handled without recourse to ε. They proved that every component remains nonsingular after a finite number of iterations, provided that the MM algorithms of this class are properly initialized. For our regularization scheme, this amounts to choosing an initial $\mathbf{f}^{(0)}$ such that no element $(\boldsymbol{\Delta}\mathbf{f}^{(0)})_i$, $|(\mathbf{V}\mathbf{f}^{(0)})_i|$, or $|(\mathbf{U}\mathbf{f}^{(0)})_i|$ is zero-valued.

The convergence of the proposed algorithm can be established by noting its direct relation to a variant of the lagged diffusivity method that was first proposed in [26] for TV minimization. This method is based on the Euler–Lagrange equation and, for TV, leads to an elliptic partial differential equation (PDE). To solve this PDE, Vogel and Oman pursued a fixed-point iterative approach [26], [27, Ch. 8], and the global convergence of its discrete approximations were proven in [28]. For our problem, we can follow a similar approach and form the corresponding fourth-order PDEs. Then, we can show that the discrete version of the fixed-point iterative approach for solving these PDEs coincides with the MM method we are proposing. Using similar arguments, as in [28], we can then prove the global convergence of our method.

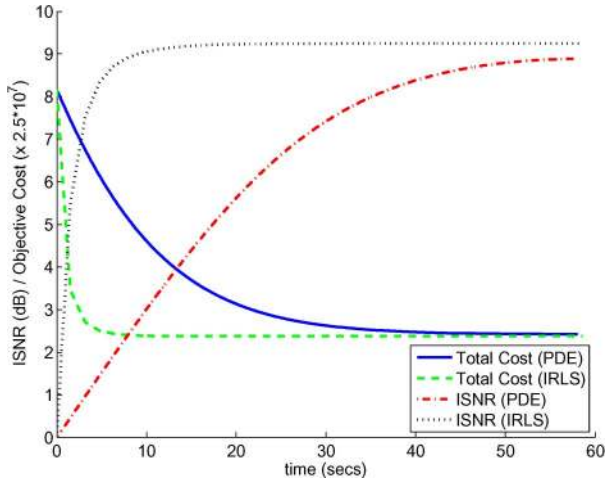


Fig. 4. Comparison of the proposed IRLS algorithm with the PDE-based method of [9] for the task of image denoising using the Frobenius-norm regularization. The presented results refer to the ISNR and objective cost evolution along time.

IV. EXPERIMENTS

To validate the effectiveness of our second-order regularization scheme for the task of image restoration/deblurring, we provide experimental comparisons with TV, which currently produces state-of-the-art results. In addition, to point out the efficiency of the proposed IRLS algorithm for the simpler image denoising case, we provide an indicative denoising example.

A. Algorithmic Performance

In this section, we compare the computational performance of the proposed IRLS algorithm for image denoising employing the Frobenius-norm regularization. The denoising task is performed for the Lena image under Gaussian noise of standard deviation $\sigma = 20$. The motivation for this comparison is to evaluate the efficiency of our scheme for the simpler problem of image denoising, compared with the PDE-based noise removal method proposed in [9]. The latter work was the first to introduce the Frobenius-norm regularization for image denoising.

In the reported experiment, each external iteration of our method involved ten PCG iterations. The step size of the PDE-based method was chosen so as to achieve the best possible decrease in the objective function leading also to the largest increase in SNR (ISNR). In Fig. 4, we plot the evolution of the ISNR and the decrease in the objective cost for both methods. As we can clearly see, our proposed method converges much faster, leading to a smaller value of the objective and a better SNR.

B. Restoration Settings

We have used two distinct sets of images. The first set is composed of five 8-bit grayscale standard test images of size 512×512 pixels, shown in Fig. 5. Given that our main interest leans toward biomedical imaging, we further provide comparisons on a second set composed of eight cell images. These images, shown in Fig. 6, are part of the biomedical image database [29]. They were converted to grayscale, and their intensities were adjusted to lie in the range [0, 255]. Since their size was relatively

Fig. 5. Set of standard test images. From left to right: Barbara, Boat, Fingerprint, Fluorescent Cells, and Lena.

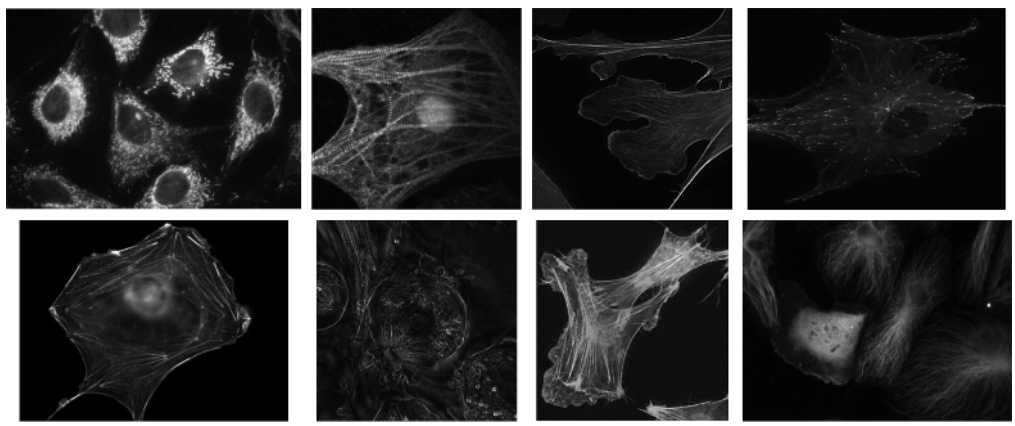


Fig. 6. Set of biomedical cell images. From left to right: CIL 248, CIL 1585, CIL 7053, CIL 7438, CIL 7762, CIL 9061, CIL 10016, and CIL 7437. CIL stands for the Cell Image Library identifier. Every image in the database [29] is assigned a unique identifier number.

large, they were also resized, with the largest resulting image having a dimension of 364×487 pixels.

The performance of all methods was examined for two blurring kernels and three different noise levels. For the blurring operation, we used a Gaussian point-spread function (PSF) of standard deviation $\sigma_b = 4$ and a uniform one, both of them having a support of 9×9 pixels. As an additional degradation factor, we considered the Gaussian noise of three distinct levels corresponding to a blurred SNR (BSNR) of 20, 30, and 40 dB, respectively. The BSNR is defined as $\text{BSNR} = \text{var}[\mathbf{Af}]/\sigma^2$, where σ is the standard deviation of the noise. The quality of the resulting images is measured in terms of an SNR increase (ISNR) measured in decibels. The ISNR is defined as

$$\text{ISNR} = 10 \log_{10} \left(\frac{\text{MSE}_{\text{in}}}{\text{MSE}_{\text{out}}} \right) \tag{49}$$

where MSE_{in} and MSE_{out} are the MSE between the degraded image and the original image, and the restored image and the original image, respectively.

Regarding the comparisons with TV regularization, we used the available code of the MM-based algorithm proposed in [3]. For the L_1-norm Laplacian regularization, instead of using forward finite differences to approximate derivatives, as we do for the rest of the second-order regularizers, we used the stencil

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

which arises from an approximation with central finite differences. The rationale behind this choice is that it yields better restoration results in terms of ISNR while the Laplacian operator remains self-adjoint, as in the continuous case. In addition, for each method under comparison, we used ten external iterations (i.e., ten successive quadratic-bound minimizations) to minimize the corresponding objective function. For the quadratic minimizations, we used the PCG algorithm with a stopping criterion set to either reaching a relative normed difference of 10^{-5} between two successive estimates or a maximum of 200 iterations[5] .

C. Standard Test Images

In Table IV, we provide the comparative results of our proposed second-order regularization scheme with TV, for the set of the standard test images and an extensive set of degradation conditions. For the sake of consistency, the reported results for each regularizer were obtained using the individualized regularization parameter τ that gave the best ISNR performance.

From this table, we see that, among the second-order regularizers, the Frobenius norm consistently gives the best results, with the spectral-norm regularizer achieving a slightly inferior ISNR. On the other hand, the L_1-norm Laplacian regularizer fails to be competitive irrespectively of the considered images or degradation conditions. In some cases, its performance can even fall behind by more than 1.5 dB. Meanwhile, the results obtained using the proposed regularization approach are similar to those of TV, irrespective of the choice of the Hessian-matrix-norm regularizer. It even turns out that, particularly when the Frobenius-norm is used, we manage to obtain better ISNR

[5]For the PCG method proposed in Section III-C, we never reached the maximum number of iterations, and the algorithm converged after a significantly smaller number of iterations.

TABLE IV
PERFORMANCE OF THE FOUR CONSIDERED REGULARIZERS IN TERMS OF ISNR ON THE SET OF STANDARD TEST IMAGES FOR TWO BLURRING KERNELS AND THREE DIFFERENT NOISE LEVELS

Image / BSNR		Uniform blur: 9×9				Gaussian blur: 9×9			
		TV	Laplacian	Spectral	Frobenius	TV	Laplacian	Spectral	Frobenius
Barbara	20 dB	**1.48**	1.10	1.41	1.44	**1.18**	0.87	1.13	1.17
	30 dB	2.07	1.62	2.06	**2.11**	1.68	1.27	1.65	**1.70**
	40 dB	3.67	2.98	3.68	**3.80**	2.53	2.05	2.54	**2.58**
Boat	20 dB	**3.95**	2.89	3.82	3.94	3.10	2.24	3.02	**3.11**
	30 dB	6.10	4.69	6.00	**6.18**	5.08	3.77	4.97	**5.14**
	40 dB	8.56	7.03	8.49	**8.70**	7.39	5.84	7.29	**7.48**
Fingerprint	20 dB	5.75	4.93	6.38	**6.48**	5.45	4.23	5.83	**5.93**
	30 dB	8.06	7.44	8.90	**9.04**	7.37	6.31	7.91	**8.02**
	40 dB	10.83	10.08	11.59	**11.79**	9.65	8.62	10.19	**10.34**
Fluor. Cells	20 dB	3.86	3.22	4.08	**4.16**	3.28	2.73	3.49	**3.57**
	30 dB	5.74	4.99	6.06	**6.18**	4.91	4.17	5.22	**5.32**
	40 dB	8.19	7.38	8.62	**8.79**	7.11	6.27	7.52	**7.67**
Lena	20 dB	**3.95**	2.83	3.71	3.82	**3.47**	2.37	3.20	3.29
	30 dB	5.37	4.27	5.36	**5.50**	4.69	3.52	4.56	**4.70**
	40 dB	7.60	6.49	7.71	**7.90**	6.51	5.35	6.50	**6.65**

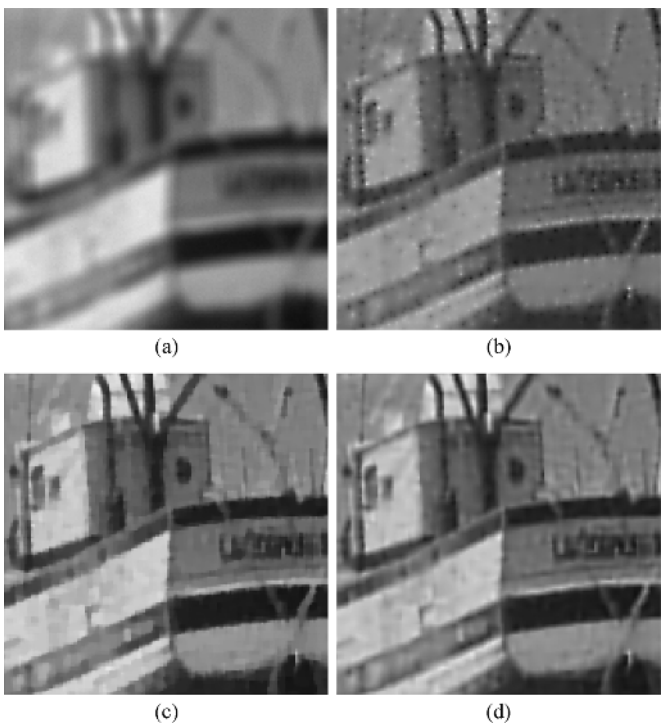


Fig. 7. Results on the Boat image degraded by a Gaussian PSF and noise corresponding to a BSNR level of 30 dB. Close-up of (a) Degraded image, (b) L_1-norm Laplacian solution (ISNR $= 3.77$), (c) TV solution (ISNR $= 5.08$), and (d) Frobenius-norm solution (ISNR $= 5.14$).

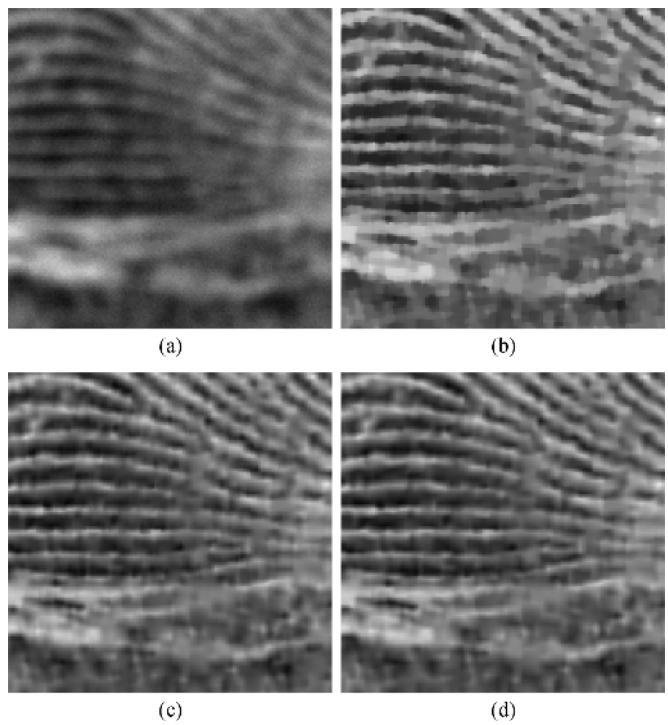


Fig. 8. Results on the Fingerprint image degraded by a uniform PSF and noise corresponding to a BSNR level of 20 dB. Close-up of (a) degraded image, (b) TV solution (ISNR $= 5.75$), (c) spectral-norm solution (ISNR $= 6.38$), and (d) Frobenius-norm solution (ISNR $= 6.48$).

scores than TV on all images for many blurring and noise conditions. For the remaining few cases, the difference in terms of ISNR performance is relatively small.

The success of the proposed method in terms of visual quality can be appreciated by inspecting the representative Boat and Fingerprint deblurring examples of Figs. 7 and 8. The Boat image has been degraded by Gaussian blurring and Gaussian noise corresponding to a BSNR level of 30 dB, whereas the Fingerprint image has been degraded by uniform blurring and Gaussian noise corresponding to a BSNR level of 20 dB. In the first example, for the second-order regularization case, we provide the results obtained for the L_1-norm Laplacian and the Frobenius-norm regularizers.

We clearly see that the L_1-norm Laplacian solution suffers from the presence of intense white-point artifacts. As reported in [13] and [15], this behavior seems to be also a standard for the case of image denoising. To deal with this problem, the authors in [13] proposed to apply a median-like filtering on the

TABLE V
PERFORMANCE OF THE FOUR CONSIDERED REGULARIZERS IN TERMS OF ISNR ON THE SET OF CELL IMAGES FOR TWO BLURRING KERNELS AND THREE DIFFERENT NOISE LEVELS

Image	BSNR	Uniform blur: 9×9				Gaussian blur: 9×9			
		TV	Laplacian	Spectral	Frobenius	TV	Laplacian	Spectral	Frobenius
CIL 248	20 dB	3.50	2.90	3.61	**3.65**	2.93	2.43	3.05	**3.08**
	30 dB	5.31	4.56	5.60	**5.66**	4.45	3.76	4.68	**4.72**
	40 dB	7.92	7.02	8.31	**8.42**	6.56	5.72	6.91	**7.00**
CIL 1585	20 dB	3.31	2.84	3.49	**3.54**	2.83	2.38	2.98	**3.02**
	30 dB	4.98	4.57	5.43	**5.46**	4.20	3.83	4.62	**4.63**
	40 dB	7.54	7.13	8.23	**8.31**	6.28	5.94	6.96	**7.01**
CIL 7053	20 dB	4.33	3.41	4.42	**4.59**	3.52	3.00	3.87	**4.00**
	30 dB	7.32	6.07	7.56	**7.87**	5.53	4.72	5.96	**6.18**
	40 dB	10.97	9.61	11.20	**11.53**	8.67	7.56	9.05	**9.26**
CIL 7437	20 dB	2.86	2.40	2.92	**2.98**	2.50	2.06	2.55	**2.60**
	30 dB	3.76	3.24	3.99	**4.07**	3.04	2.58	3.24	**3.32**
	40 dB	5.59	5.05	5.94	**6.06**	4.62	4.15	4.97	**5.06**
CIL 7438	20 dB	4.77	4.35	4.90	**4.95**	3.77	3.51	3.95	**3.98**
	30 dB	7.42	6.65	7.65	**7.71**	6.12	5.54	6.39	**6.54**
	40 dB	10.50	9.59	10.76	**10.99**	8.83	8.12	9.10	**9.33**
CIL 7762	20 dB	3.54	3.34	4.03	**4.14**	2.99	2.77	3.47	**3.59**
	30 dB	5.71	5.26	6.22	**6.37**	4.57	4.46	5.33	**5.47**
	40 dB	9.06	8.38	9.51	**9.71**	7.06	6.77	7.81	**7.98**
CIL 9061	20 dB	3.67	3.18	3.87	**3.94**	2.88	2.50	3.12	**3.17**
	30 dB	5.76	5.10	6.09	**6.21**	4.71	4.14	**5.02**	4.88
	40 dB	8.49	7.73	8.70	**8.76**	6.99	6.17	7.03	**7.10**
CIL 10016	20 dB	**3.20**	2.34	3.05	3.13	**2.57**	1.82	2.43	2.49
	30 dB	4.99	3.92	4.92	**5.07**	4.01	3.10	4.02	**4.13**
	40 dB	7.47	6.30	7.45	**7.64**	6.05	5.04	6.11	**6.28**

denoised images. We also tried to use this postprocessing approach on our restored results, but although we observed a small increase in terms of ISNR, the majority of the artifacts remained present. Meanwhile, the TV solution is characterized by sharper edges, as expected, but also results in blocky artifacts appearing in the smoother regions of the image. These artifacts are much less pronounced with the Frobenius-norm solution, which retains the quality of TV at the abrupt transitions of image intensities. While apparent on the Boat image in Fig. 7, this benefit becomes even more dramatic in Fig. 8, where the Fingerprint image mostly consists in smooth transitions and ridge features. From this example, we can also see that, although the Frobenius norm achieves a better ISNR than the spectral norm, the results look very similar, and it is difficult to observe any significant visual differences.

Regarding TV, heavy block artifacts are produced, which shuffle the details of the image and broaden its fine structures. It is also worthwhile to note that the sharp intensity transitions produced by TV do not necessarily correlate with high-quality restoration as they can be attributed to its inherent tendency to exaggerate contrast along the image contours. This exaggeration is linked to the staircasing effect itself [30].

D. Biomedical Cell Images

In this section, our main interest is to evaluate the performance of Hessian-based regularization on biomedical data, which is potentially more relevant for practical applications. To this end, we provide comparative deblurring results for the set of cell images shown in Fig. 6. In these comparisons, we again consider the same Gaussian and uniform PSFs for the blurring operations along with the same three levels of Gaussian noise.

By examining Table V, we see that the second-order regularization approach, employing either of the Hessian-based norm regularizers, almost always performs better than TV. The superiority of our approach for this class of images can be attributed to its ability to preserve the ridges and filament-like structures, which are commonly met in biomedical imaging. This property stems from the fact that, similar to simple ridge detectors [31], the second-order regularizers we are using are based on the eigendecomposition of the Hessian matrix. However, regarding the evaluation of the Laplacian regularizer, we are lead to the same conclusions as before. Indeed, the L_1-norm Laplacian cannot compete with any of the other methods and consistently gives the worst results.

Beyond ISNR considerations, we present qualitative comparative deblurring results on the cell images CIL 7762 and CIL 7437 in Figs. 9 and 10, respectively. In these examples, the former image has been blurred by a Gaussian PSF and further degraded by Gaussian noise corresponding to a BSNR of 30 dB, whereas the latter has been degraded by uniform blurring and noise resulting in a BSNR of 20 dB. The behavior of the L_1-norm Laplacian regularizer is consistent with the findings of the previous sections, i.e., that the obtained solutions suffer from the presence of white-point artifacts. This regularization approach therefore cannot be considered a reliable solution in biomedical applications as it may bias the interpretation of the

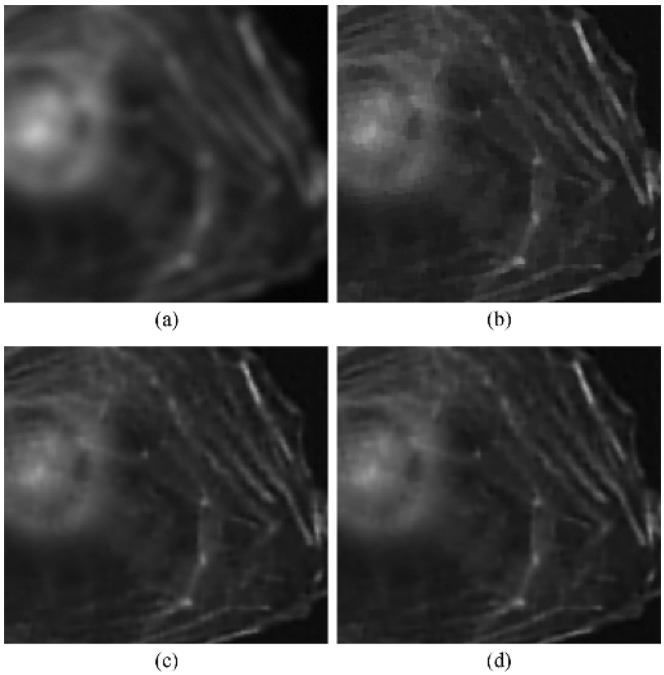


Fig. 9. Results on the CIL 7762 image degraded by a Gaussian PSF and noise corresponding to a BSNR level of 30 dB. Close-up of (a) degraded image, (b) TV solution (ISNR $= 4.57$), (c) spectral-norm solution (ISNR $= 5.33$), and (d) Frobenius-norm solution (ISNR $= 5.47$).

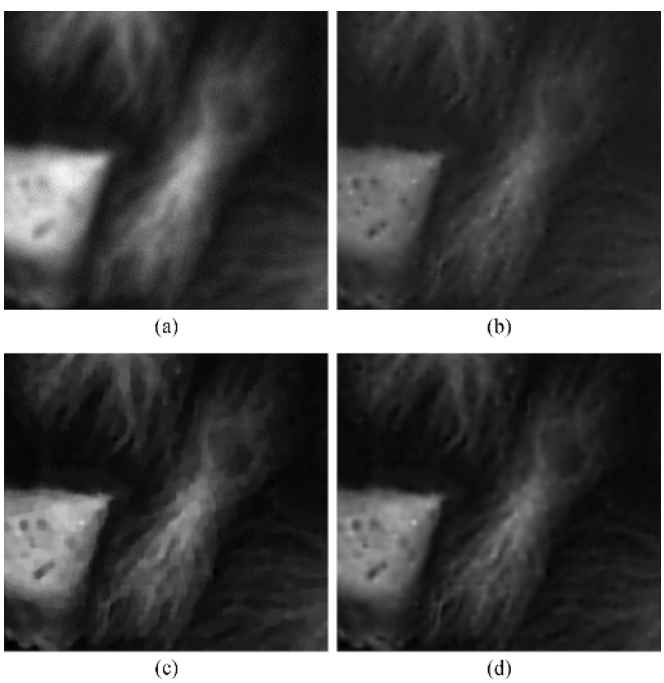


Fig. 10. Results on the CIL 7437 image degraded by a uniform PSF and noise corresponding to a BSNR level of 20 dB. Close-up of (a) degraded image, (b) L_1-norm Laplacian solution (ISNR $= 2.40$), (c) TV solution (ISNR $= 2.86$), and (d) spectral-norm solution (ISNR $= 2.92$).

structure of the image. As far as TV solutions are concerned, we observe that favoring piecewise-constant image patches tends to oversmooth certain features, which wipes out important details of the image structure. On the other hand, both Hessian-based regularizers manage to restore the fine details more accurately while maintaining their actual size even in cases where the level of noise is substantial. It is also worthwhile to mention that, once more, although the Frobenius regularizer yields a better ISNR than the spectral one, the visual quality of the corresponding results is similar.

V. Conclusion

In this paper, we have devised a novel second-order regularization scheme to solve linear inverse problems geared toward biomedical imaging applications. Specifically, we proposed a unifying framework for the minimization of the objective function employing several different second-order regularizers. The proposed optimization algorithm is of IRLS type and results from an MM approach. It relies on a problem-specific PCG method that we have also designed. Our PCG approach makes the overall minimization scheme quite attractive since it can be effectively applied to large images in a reasonable computational time. Furthermore, motivated by the great effectiveness of TV regularization, we have extended its definition to also include higher order differential operators. We have shown that the resulting Hessian-based second-order regularizers can be a good choice for deblurring images that are not dominated by piecewise-constant regions and sharp edges but, rather, by ridges or smoother transitions of intensity. Fluorescence microscopy images tend to contain many such features.

The performance and the practical relevance of the proposed regularization scheme for image restoration have been illustrated through comparisons with the state-of-the art TV method. Our regularization approach is very competitive, particularly in the case of biomedical images. The results we have obtained are promising, both in terms of restoration quality and computational performance.

Acknowledgment

The authors would like to thank the anonymous reviewers for their useful comments and P. Thévenaz for his careful proofreading of the manuscript.

References

[1] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992.

[2] P. L. Combettes and J. C. Pesquet, "Image restoration subject to a total variation constraint," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1213–1222, Sep. 2004.

[3] J. Bioucas-Dias, J. Oliveira, and M. Figueiredo, "Total variation-based image deconvolution: A majorization–minimization approach," in *Proc. ICASSP*, Toulouse, France, 2006, vol. 2, pp. 861–864.

[4] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[5] T. F. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM J. Appl. Math.*, vol. 62, no. 3, pp. 1019–1043, 2002.

[6] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imaging Vis.*, vol. 20, no. 1, pp. 89–97, Jan. 2004.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*. Norwell, MA: Kluwer, 2004.

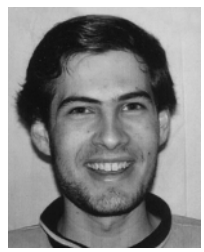
[8] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 22, no. 2, pp. 503–516, 2000.

[9] M. Lysaker and X.-C. Tai, "Iterative image restoration combining total variation minimization and a second-order functional," *Int. J. Comput. Vis.*, vol. 66, no. 1, pp. 5–18, Jan. 2006.

[10] J. Yuan, C. Schnörr, and G. Steidl, "Total-variation based piecewise affine regularization," in *Proc. 2nd Int. Conf. Scale Space Methods Variational Methods Comput. Vis.*, 2009, pp. 552–564.
[11] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.*, vol. 76, no. 2, pp. 167–188, Apr. 1997.
[12] T. F. Chan, S. Esedoglu, and F. E. Park, "Image decomposition combining staircase reduction and texture extraction," *J. Vis. Commun. Image Represent.*, vol. 18, no. 6, pp. 464–486, Dec. 2007.
[13] Y.-L. You and M. Kaveh, "Fourth-order partial differential equations for noise removal," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1723–1730, Oct. 2000.
[14] M. Lysaker, A. Lundervold, and X.-C. Tai, "Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1579–1590, Dec. 2003.
[15] G. Steidl, "A note on the dual treatment of higher-order regularization functionals," *Computing*, vol. 76, no. 1, pp. 135–148, Jan. 2006.
[16] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Spectra, and Filtering*. Philadelphia, PA: SIAM, 2006.
[17] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imaging Sci.*, vol. 3, no. 3, pp. 492–526, 2010.
[18] S. Setzer, G. Steidl, and T. Teuber, "Infimal convolution regularizations with discrete ℓ_1-type functionals," *Commun. Math. Sci.*, vol. 9, no. 3, pp. 797–827, 2011.
[19] C. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM, 2000.
[20] J. E. Denis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Philadelphia, PA: SIAM, 1983.
[21] D. Hunter and K. Lange, "A tutorial on MM algorithms," *Amer. Statistician*, vol. 58, no. 1, pp. 30–37, 2004.
[22] M. Figueiredo, J. Bioucas-Dias, and R. Nowak, "Majorization–minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2980–2991, Dec. 2007.
[23] J. R. Shewchuk, An introduction to the conjugate gradient method without the agonizing pain 1994 [Online]. Available: http://www.cs.cmu.edu/~jrs/jrspapers.html
[24] T. Chan, "An optimal circulant preconditioner for Toeplitz systems," *SIAM J. Sci. Stat. Comput.*, vol. 9, no. 4, pp. 766–771, 1988.
[25] R. Chan, J. Nagy, and R. Plemmons, "FFT-based preconditioners for Toeplitz-block least squares problems," *SIAM J. Numer. Anal.*, vol. 30, no. 6, pp. 1740–1768, Dec. 1993.
[26] C. R. Vogel and M. E. Oman, "Iterative methods for total variation denoising," *SIAM J. Sci. Comput.*, vol. 17, no. 1, pp. 227–238, 1996.
[27] C. R. Vogel, *Computational Methods for Inverse Problems*. Philadelphia, PA: SIAM, 2002.
[28] D. Dobson and C. R. Vogel, "Convergence of an iterative method for total variation denoising," *SIAM J. Numer. Anal.*, vol. 34, no. 5, pp. 1779–1791, 1997.
[29] D. Brandner and G. Withers, The Cell: An Image Library 2010 [Online]. Available: http://www.cellimagelibrary.org
[30] C. Louchet and L. Moisan, "Total variation denoising using posterior expectation," in *Proc. Eusipco*, Lausanne, Switzerland, 2008.
[31] M. Jacob and M. Unser, "Design of steerable filters for feature detection using Canny-like criteria," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1007–1019, Aug. 2004.

Stamatios Lefkimmiatis (S'08–M'11) received the M.Sc. degree in computer engineering and informatics (with highest honors) from the University of Patras, Patras, Greece, in 2004 and the Ph.D. degree from the National Technical University of Athens, Athens, Greece, in 2009.

Since 2010, he has been a Research Scientist with the Biomedical Imaging Group, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. His main research interests include image analysis, statistical modeling, and image reconstruction using convex optimization techniques.



Aurélien Bourquard (M'08) received the M.Sc. degree in microengineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 2008. He is currently working toward the Ph.D. degree with the Biomedical Imaging Group, EPFL.

His research interests include image reconstruction using convex optimization and multigrid techniques, as well as new acquisition methods in the framework of computational optics.

Michael Unser (M'89–SM'94–F'99) received the M.S. (*summa cum laude*) and Ph.D. degrees in electrical engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1981 and 1984, respectively.

From 1985 to 1997, he worked as a Scientist with the National Institutes of Health, Bethesda, MD. He is currently a Full Professor and a Director of the Biomedical Imaging Group, EPFL. He is the author of 200 journal papers. His main research interests include biomedical image processing, sampling theories, multiresolution algorithms, wavelets, and the use of splines for image processing.

Dr. Unser is a EURASIP Fellow and a member of the Swiss Academy of Engineering Sciences. He was the Associate Editor for the IEEE SIGNAL PROCESSING LETTERS in 1994–1998, for the IEEE TRANSACTIONS ON IMAGE PROCESSING in 1992–1995, and for the IEEE TRANSACTIONS ON MEDICAL IMAGING in 1999–2002 and 2006–2007. He has also held the position of associate Editor-in-Chief in 2003–2005 for the IEEE TRANSACTIONS ON MEDICAL IMAGING. He is currently a member of the editorial boards of *Foundations and Trends in Signal Processing* and *Sampling Theory in Signal and Image Processing*. He coorganized the first IEEE International Symposium on Biomedical Imaging and was the Founding Chair of the technical committee of the IEEE-SP Society on Bio Imaging and Signal Processing. He was the receipient of the 1995 and 2003 Best Paper Awards, the 2000 Magazine Award, and two IEEE Technical Achievement Awards (2008 SPS and 2010 EMBS). He is one of ISI's Highly Cited Authors in Engineering (http://isihighlycited.com).